

EpiGraph: an open-source platform to quantify epithelial organization.

Pablo Vicente-Munuera^{1,†}, Pedro Gómez-Gálvez^{1,†}, Robert J. Tetley², Cristina Forja¹, Antonio Tagua¹, Marta Letrán¹, Melda Tozluoglu², Yanlan Mao^{2,3,*} and Luis M. Escudero^{1,*}.

1: Instituto de Biomedicina de Sevilla (IBiS), Hospital Universitario Virgen del Rocío/CSIC/Universidad de Sevilla and Departamento de Biología Celular, Universidad de Sevilla. 41013 Seville, Spain.

2: MRC Laboratory for Molecular Cell Biology, University College London, Gower Street, London WC1E 6BT, UK.

3: College of Information and Control, Nanjing University of Information Science and Technology, Nanjing, Jiangsu 210044, China

†: These authors contributed equally to this work.

*: Corresponding authors: Yanlan Mao (y.mao@ucl.ac.uk) and L.M. Escudero (lmescudero-ibis@us.es).

1 **SUMMARY**

2

3 During development, cells must coordinate their differentiation with their

4 growth and organization to form complex multicellular structures such as

5 tissues and organs. Healthy tissues must maintain these structures during

6 homeostasis. Epithelia are packed ensembles of cells from which the different

7 tissues of the organism will originate during embryogenesis. A large barrier to

8 the analysis of the morphogenetic changes in epithelia is the lack of simple

9 tools that enable the quantification of cell arrangements. Here we present

10 EpiGraph, an image analysis tool that quantifies epithelial organization. Our

11 method combines computational geometry and graph theory to measure the

12 degree of order of any packed tissue. EpiGraph goes beyond the traditional

13 polygon distribution analysis, capturing other organizational traits that

14 improve the characterization of epithelia. EpiGraph can objectively compare

15 the rearrangements of epithelial cells during development and homeostasis

16 to quantify how the global ensemble is affected. Importantly, it has been

17 implemented in the open-access platform FIJI. This makes EpiGraph very

18 user friendly, with no programming skills required.

19

1 INTRODUCTION

2 The development of any multicellular organism is based on coordinated
3 changes that transform the embryo into the adult individual. During
4 morphogenesis and growth, patterning, cell divisions and architectural
5 changes must perfectly fit together for the correct development of the body
6 plan. Any morphogenetic movement such as migration, extension or
7 invagination of epithelial cells is coupled with dramatic changes in the
8 organization of cells (Bertet et al., 2004; Blankenship et al., 2006; Escudero
9 et al., 2007; Farhadifar et al., 2007; Girdler and Roper, 2014; Lecuit and
10 Lenne, 2007; Pilot and Lecuit, 2005). After development, homeostatic tissues
11 must maintain their complex organization of cells in order to function correctly.

12 How tissues modulate and maintain their organization during development
13 and homeostasis is an important question that remains unsolved. This is
14 mainly due to the lack of simple and general methods that can capture and
15 quantify the arrangement of cells. It has been known for almost a hundred
16 years that epithelial tissues exhibit a degree of order. The analysis of epithelial
17 organization has been mainly based on the number of neighbours of the
18 epithelial cells, considering the apical surface of these cells as convex
19 polygons with the same number of sides as neighbours. In previous works,
20 we have investigated several aspects of the organization of packed tissues
21 using Voronoi tessellations to compare the polygon distributions of natural
22 and mathematical tessellations (Sanchez-Gutierrez et al., 2016). We have
23 described that the polygon distribution of natural tessellations is restricted to
24 a series of frequencies of polygons that match the Voronoi diagrams that
25 conform to the Centroidal Voronoi tessellation (CVT). This is what we call a
26 “CVT path” and was used as a scale to compare the organization of different
27 packed tissues. However, polygon distribution is not sufficient to completely
28 characterize tissue organization. Tissues with clearly different appearance
29 can present very similar polygon distribution (Sanchez-Gutierrez et al., 2016).

30 As an alternative approach, we have proposed that Graph Theory could
31 capture differences in the topology of tissues (Escudero et al., 2011;
32 Sanchez-Gutierrez et al., 2013; Sánchez-Gutiérrez et al., 2017). This is based
33 on the idea of converting the epithelium into a network of cell-to-cell contacts

(Escudero et al., 2011). The resulting “epithelial graph” can be analysed by combining the tools of network theory and multivariable statistical analysis (Escudero et al., 2011; Kursawe et al., 2016; Sanchez-Gutierrez et al., 2013; Yamashita and Michiue, 2014). This approach has been adapted to analyze biomedical tissue samples, useful in clinical research and the development of diagnostic tools (Csikász-Nagy et al., 2013; Guillaud et al., 2010; Sáez et al., 2013; Sánchez-Gutiérrez et al., 2017). Finding features and patterns that can describe the graphs is key in many diverse fields, including biology (Benson et al., 2016; Costa et al., 2007; Hayes et al., 2013). A network can be split up into different subgraphs named graphlets. The graphlet composition of a network has been used to quantify differences between complex systems (Hayes et al., 2013; Ho et al., 2010; Kuchaiev et al., 2011; Pržulj et al., 2004). These measurements are based on the comparison of the quantity of each subgraph in different networks, providing an index of distance between them. This feature has the advantage of integrating the differences between diverse networks into a single value, simplifying the analyses and allowing multiple comparisons.

In summary, there is a clear need for a method to specifically quantify tissue organization and aid the interpretation of biophysical and mechanical aspects of morphogenesis and tissue homeostasis. The advances in imaging techniques, together with the appearance of powerful methods for automated image analysis (Heller et al., 2016; Khan et al., 2014; Kursawe et al., 2016; Schindelin et al., 2012; Weigert et al., 2018) and new simulation resources (Bi et al., 2016, 2015; Blanchard et al., 2009; Etournay et al., 2016; Fletcher et al., 2014; Guirao et al., 2015; Mirams et al., 2013; Tanaka et al., 2015) provide a large amount of good quality source data that can now be analysed in terms of organization. Here we present an open source platform, EpiGraph, a new image analysis method that uses segmented images from real epithelia or simulations, to easily quantify and compare the organization of packed tissues.

RESULTS

Graphlet measurements as an approach to capture organization of packed tissues.

1 In previous studies, a set of 29 graphlets was used to distinguish between
2 different types of networks (Pržulj et al., 2004) (**Fig. S1**). This method
3 calculated the Graphlet degree Distribution agreement Distance (GDD)
4 between two networks (Pržulj, 2007). Therefore, the “GDD value”, that in
5 theory can range from 0 to 1, weighs the differences among the two
6 distributions of graphlets; the higher the value, the more different the
7 arrangements (**Fig. S1** and **Methods**). Epithelial images can be considered
8 as natural tessellations and converted into networks of cell-to-cell contacts
9 (Escudero et al., 2011). We have used the “graphlet” approach to capture the
10 topology of epithelial tissues, making a correlation between graphlets and
11 cellular motifs (compare **Fig. 1A** and **Fig. S1**). Tessellations give rise to
12 “geographic networks” (Albert and Barabasi, 2002) that only make sense in a
13 planar surface. For this reason, when we translated the set of graphlets to
14 cellular patterns, some of them were redundant or not possible (see
15 methods). Therefore, in this study we have used a total of 26 graphlets
16 corresponding to 29 different cellular motifs that account for the organization
17 of groups of up to 5 cells (**Fig. 1A, Fig. S1**). Most of the analyses performed
18 in this work were completed with only 17 motifs (17-motifs, **Fig. 1A**, mauve).
19 We found that, although all the motifs could be present in an actual tissue,
20 17-motifs minimized the redundancy of the information provided by the
21 graphlets. In addition, this set downplays the importance of rare cellular
22 geometries that could excessively weight GDD calculations (for example, a
23 high difference in GDD could appear when comparing an image with one or
24 two quadrilateral cells versus another image with no four-sided cells; this
25 effect is minimized using 17-motifs). However, it would be possible to use
26 other combinations such as all the motifs (29-motifs) or cellular motifs that
27 account for the organization of groups of up to 4 cells (10-motifs) (**Fig. 1A**).

28 **Graphlet measurements capture differences beyond polygon** 29 **distributions.**

30 We tested the power of graphlet-based measurements in quantifying
31 differences between sets of images with very similar polygon distributions
32 (**Fig. 1B**). In third instar larvae of *Drosophila*, the photoreceptors are
33 specified, giving rise to a particular repetitive arrangement of the presumptive

1 eye cells (Eye, **Fig. 1C**). This arrangement is very different to the irregular
2 distribution in a Voronoi tessellation where the initial seeds were placed in a
3 random way (Sanchez-Gutierrez et al., 2016), (Diagram 1, **Fig. 1C**). We
4 previously showed that it was not possible to discriminate between the
5 polygon distributions of these two tessellations (Sanchez-Gutierrez et al.,
6 2016). Using the graphlets approach, we obtained a GDD value of 0.086
7 when comparing these two sets of images (17-motifs, **Table S1, Fig. S2**). In
8 order to know if this difference was biologically relevant, we tried to set a
9 baseline, by comparing other images with very similar polygon distribution
10 that also presented an apparently similar arrangement. This was the case for
11 Diagram 4 of the CVT vs. the *Drosophila* wing imaginal disc in larvae (dWL)
12 and Diagram 5 of the CVT vs. the *Drosophila* wing imaginal disc in prepupae
13 (dWP) (**Fig. 1B-C**). Both results were in the same range, with a GDD value of
14 0.042 for Diagram 4 vs. dWL and 0.049 for Diagram 5 vs. dWP (**Fig. 1C**).
15 Similar results were obtained when comparing Diagram 4 vs. Diagram 5 and
16 dWL vs. dWP (**Fig. 1C**). These results suggested the existence of a baseline
17 in the range of 0.04-0.05 values that correspond to similar cellular
18 arrangements that cannot be well distinguished using the graphlets
19 distribution. Therefore, we interpreted the value of 0.086 obtained in the Eye
20 vs. Diagram 1 comparison as the reflection of actual differences between
21 these two sets. In all the mentioned cases, the results obtained using 17-
22 motifs and 29-motifs were equivalent (**Table S1**).

23 **EpiGraph quantitatively compares the organization of multiple sets of** 24 **images.**

25 The GDD had the limitation of comparing only 2 samples each time. Here
26 we have tried to overcome this limitation evaluating different images
27 simultaneously using a reference. Therefore, we designed EpiGraph, a
28 method that calculates the GDD of any epithelial tissue with another
29 tessellation that serves as a reference. We used three different references: i)
30 a tessellation formed by regular hexagons, representing the most ordered
31 way to pave the space (**Fig. 2A**, Epi-Hexagons). ii) the network motifs
32 emerging from a random Voronoi tessellation (**Fig. 2B**, Epi-Random). iii) a
33 Voronoi Diagram 5 from the CVT path (**Fig. 2C**, Epi-Voronoi5) that presents

1 a polygon distribution similar to the one from multiple examples in nature
2 (Gibson et al., 2006; Sanchez-Gutierrez et al., 2016).

3 We tested the method with epithelial images that have been previously
4 compared with the CVT path in terms of polygon distribution: chicken neural
5 tube (cNT), dWL, dWP, reduction of myosin II in the *Drosophila* prepupa wing
6 disc epithelium (dMWP) and Eye (**Fig. 2D**)(Sanchez-Gutierrez et al., 2016).
7 To have a scale and facilitate fast comparisons, we used the concept of the
8 CVT path (Sanchez-Gutierrez et al., 2016). We calculated the GDD values
9 for Epi-Hexagons, Epi-Random and Epi-Voronoi5 for all the Voronoi diagrams
10 and visualized these results with respect to the percentage of hexagons of
11 the corresponding diagram (the percentage of hexagons is indicative of the
12 proportions of the different types of polygons along the CVT, **Table S2**).
13 However, the CVT does not progress beyond the 70% of hexagons limiting
14 the possibilities of analysis. Therefore, we extended the Voronoi scale
15 spanning a wider range of polygon distributions. The algorithm that devises
16 the CVT was modified to introduce “noise” in the positioning of the seed that
17 produces the subsequent diagram. In this way, we obtained a “CVT noise”
18 (CVTn) whose last diagrams reached 90% of hexagons (**Fig. 2E-G, Fig. S3**
19 **and Material and methods**). Interestingly, the plot obtained using CVT and
20 CVTn diagrams was an optimum way to easily visualize these geometric
21 scales as a continuous “CVT path” and a “CVTn path”. Therefore, we used
22 this framework to analyse the values of Epi-Hexagons, Epi-Random and Epi-
23 Voronoi5 for each diagram in the scale. As expected, the Epi-Hexagons
24 values were higher in the initial diagrams and progressively decreased with
25 the increase in the percentage of hexagons of the Voronoi diagrams (**Fig. 2E,**
26 **left panel**). The opposite happened in the case of the Epi-Random values
27 (**Fig. 2E, central panel**). In the plot of percentage of hexagons vs Epi-
28 Voronoi5, the CVTn path presented the shape of a walking stick (**Fig. 2E,**
29 **right panel**). The Epi-Voronoi5 values of Voronoi Diagrams 1, 2, 3, and 4
30 were decreasing progressively, with Diagram 5 the closest to the zero value.
31 The values for the rest of the diagrams gradually increased, as in the case of
32 the Epi-Random.

1 We then plotted the values for the actual epithelia. We found that for cNT,
 2 dWL and dWP the Epi-Hexagons, Epi-Random and Epi-Voronoi5 values were
 3 similar to the CVTn at the same percentage of hexagons of the polygon
 4 distribution (**Fig. 2E**). In agreement with our previous results using the GDD,
 5 the Eye images presented a higher Epi-Random and Epi-Voronoi5 values
 6 than the expected for a 30% of hexagons (**Fig. 2E**). The differences with
 7 respect to the CVTn were even more clear when plotting Epi-Hexagons vs
 8 Epi-Random and Epi-Random vs Epi-Voronoi5 (**Fig. 2F-G**). We obtained
 9 similar results when analysed the dMWP set of images. In this case, our
 10 previous work showed a small deviation of the dMWP polygon distribution
 11 with respect the CVT (Sanchez-Gutierrez et al., 2016). However, using
 12 Epigraph, we observed that Epi-Random and Epi-Voronoi5 captured the clear
 13 differences in organization between these images and the CVTn (**Fig. 2E-G**,
 14 **Fig. S4**). These results suggested that EpiGraph is able to distinguish
 15 between different tessellations with a similar polygon distribution. In this
 16 regard, we have developed a statistical output using an outlier detection
 17 approach whose quantitative results represent how similar the organization
 18 of a tissue is when compared with the CVTn scale (**Fig. S3** and **Material and**
 19 **methods**). The test confirmed that cNT, dWL, and dWP were close to the
 20 CVTn and similar to the Voronoi diagrams 1, 4, and 6 respectively. In contrast,
 21 the Eye and dMWP samples were labelled as different (**Table S3**). In this way,
 22 EpiGraph provides a quantitative description of tissue organization.

23 **Epigraph can capture different organization traits.**

24 We further investigated the possible applications of EpiGraph and
 25 performed a series of experiments aimed at understanding what traits of
 26 tissue organization are being captured and quantified by the graphlet
 27 measurements. To this end, we have used images of different vertex model
 28 simulations that alter tissue organization by changing the biophysical
 29 properties of the cells (images taken from (Sanchez-Gutierrez et al., 2016)
 30 (**Material and methods** and **Fig. 3A-F**).

31 First, we analysed samples with 10% of the cells having increased effective
 32 cell-cell adhesion (**Material and methods, Fig. 3B**). This feature induced the
 33 formation of cells with a “quadrilateral shape” that often organized in motifs

1 presenting four-way vertex configurations. These images were compared with
2 simulations in which “elongated” cells appear (by simultaneously increasing
3 cell-cell adhesion and reducing ideal area, **Material and methods** and **Fig.**
4 **3C**). Epigraph analysis indicated that while control simulations gave similar
5 values to the CVTn, the “squared” and “elongated” sets of images were
6 different to the control and well separated from the CVTn. However,
7 EpiGraph failed to find clear differences between the “squared” and
8 “elongated” images (**Fig. 3G** and **Fig. S5**).

9 Second, we used a set of conditions to mimic the effect of a reduction of
10 myosin II in the *Drosophila* prepupa wing disc epithelium (dMWP, **Fig. 2D**). In
11 the control simulation (**Fig. 3D**), cells grow to double the original area and
12 then divide into two cells. In case III and case IV simulations there was a
13 random reduction of the tension parameter together with a requirement of a
14 minimum tension threshold to be able to divide (**Fig. 3E-F**). If the cells do not
15 reach this threshold, they continue to grow without dividing the cell body.
16 When this happens, the cells will be stuck in mitotic phase and will not start a
17 second round of cell division (Sanchez-Gutierrez et al., 2016) (**Material and**
18 **methods**). The control simulation gave similar values to the CVTn, while case
19 III, case IV and dMWP images presented a clear deviation in the Epi-Random
20 vs Epi-Voronoi5 graph (**Fig. 3G**). All these data-points distributed in the same
21 zone of the graph. Interestingly, we found that both sets of simulations
22 (squared and elongated vs Case III and Case IV) appeared in two
23 complementary regions, suggesting that the regions in the graph can reflect
24 the existence of different traits of organization in each condition (**Fig. 3G**).

25 **EpiGraph: a method to capture epithelial organization implemented in** 26 **FIJI.**

27 Aiming to enhance the accessibility of the analysis of tissue organization to
28 the biology community, we have implemented EpiGraph as a plugin for FIJI
29 (Schindelin et al., 2012). EpiGraph consists of a pipeline of 5 very simple
30 steps. First, the skeleton of an epithelial image is uploaded and the individual
31 cells are identified. Second, the user selects the distance threshold to identify
32 two cells as neighbours. Here it is possible to select different thresholds and
33 to check the number of neighbours of every cell in each case. Third, a ROI is

1 selected. There are several possibilities such as a default ROI from the image
 2 or the selection of individual cells. Fourth, the graphlet information for the
 3 selected cells is calculated. These data are used to obtain the Epi-Hexagons,
 4 Epi-Random and Epi-Voronoi5. These values are incorporated into a table
 5 and serve as input data for a statistical analysis that indicates if a new image
 6 is inside or outside of the CVTn path and describes which Voronoi diagram
 7 presents the most similar organization to the sample (**Material and**
 8 **methods**). The fifth step includes the classification and labelling of different
 9 images in order to represent them in a new window. This final phase allows
 10 one to export the representation of the data in a three-dimensional graph.
 11 **Movie S1** shows an example of EpiGraph usage. A detailed description of
 12 EpiGraph can be found in the **Supplementary Material and methods**. A full
 13 set of tutorials explaining how to install and use EpiGraph is available at
 14 EpiGraph's wiki (<https://imagej.net/EpiGraph>).

15 **EpiGraph provides biological insights regarding homeostasis and** 16 **tissue fluidity transitions**

17 Epithelial tissues have the ability to behave as a fluid due to cellular
 18 rearrangements or to solidify as cellular rearrangements cease (Bi et al.,
 19 2016, 2015). The shape index is a characteristic of epithelial cells that has
 20 been shown, in vertex model simulations, to be able to capture the degree of
 21 rigidity, or fluidity, of a tissue (Bi et al., 2015). This study established the
 22 transition point between a soft (fluid) and a rigid (solid) tissue, described as a
 23 jamming transition, at the dimensionless shape index value of 3.81. We
 24 calculated the shape index for the CVTn path, finding that from Voronoi
 25 diagrams 1 to 20, the tessellations were behaving as a fluid (from diagram 21
 26 to 700 they behave as solid). Using this descriptor, all the images of biological
 27 tissues were placed in the fluid part as well as the four altered vertex model
 28 simulations shown in **Fig. 3 (Fig. S4, Fig. S5, Fig. S6 and Table S4)**.

29 We have investigated the dynamics of epithelial jamming in different
 30 conditions. First, to test the capabilities of EpiGraph in this regard, we
 31 analysed several snapshots from two simulations published by Bi and
 32 colleagues as supplementary movies (Bi et al., 2016). These videos show the
 33 movements of cells in two conditions: rigid state (shape index less than 3.81)

1 and soft state (shape index greater than 3.81) (**Fig. 4A**). As expected, the
2 snapshots of the soft tissue analysed appeared in different positions,
3 indicating that the simulated epithelia changed its organization during the
4 experiment. On the other hand, the different frames from the rigid simulation
5 were clustered (**Fig. 4B**), showing little cell rearrangements.

6 We next tested whether EpiGraph could detect changes in tissue fluidity in
7 real epithelia, which may be more ambiguous and noisier than simulations.
8 Real tissues also display fluid-to-solid jamming transitions which are
9 important for large scale tissue shape changes as well as for refining and
10 maintaining tissue shape (Curran et al., 2017; Mongera et al., 2018). In the
11 *Drosophila* pupal notum, the level of tissue fluidity is controlled by the global
12 level of myosin II activity (Curran et al., 2017). We wondered if the regulation
13 of myosin II could similarly impact on the fluidity state of the wing disc
14 epithelium and the cell rearrangements that have been described during the
15 late stages of normal wing disc development, where the overall tissue shape
16 does not dramatically change (Heller et al., 2016) (**Fig. 4C**). To this end, we
17 compared the WT organization with the effect of increasing myosin II activity
18 by knocking down Mbs (Myosin binding subunit of the myosin phosphatase,
19 which dephosphorylates myosin regulatory light chain, **Fig. 4D**) by RNAi
20 throughout the entire wing pouch. Based on work in the pupal notum (Curran
21 et al., 2017), we would predict that *Mbs-RNAi* discs behave as solids.
22 Interestingly, in the two cases, the shape index was greater than the
23 described shape index threshold of 3.81, suggesting that both tissues are in
24 a fluid state (**Fig. S6**).

25 We used EpiGraph to analyse the changes in organization of wing discs with
26 perturbed myosin II activity along time and compared them with a WT
27 condition. The snapshots for WT samples appeared clustered in the 3D
28 graph, indicating that the epithelia were not changing their organization during
29 the 30 minutes of analysis (**Fig. 4E**), despite previous work showing that cell
30 intercalations do occur (Heller et al., 2016). This suggests that during this
31 slow growing phase of wing disc development, any cell rearrangements that
32 occur do not drive large-scale morphogenesis, but act to maintain a
33 homeostatic tissue topology. The statistical analysis confirmed that all of the

1 WT wing discs were close to CVTn diagrams 3 and 4 (**Table S3**). In the case
 2 of the three samples from the *Mbs-RNAi* genotype, the data points presented
 3 different organizations (from similar to diagram 3 to close to diagram 13, see
 4 **Table S3**). In some cases, the dispersion was not only between samples, but
 5 occurred between images from each movie (**Fig. 4E** and **Table S3**). EpiGraph
 6 therefore predicted that these *Mbs-RNAi* wing discs are behaving very
 7 differently from WT wing discs, likely by changing their degree of fluidity.
 8 Accordingly, quantification of intercalation rates demonstrated that cell
 9 rearrangements happen significantly more frequently in WT than in *Mbs-RNAi*
 10 wing discs (**Fig. 4F**, 0.1281 ± 0.08 vs. 0.0076 ± 0.01 intercalations/cell/hour,
 11 Kolmogorov-Smirnov test, $p=0.0079$). As predicted, this resulted in more cells
 12 ‘jamming’ at 4-way vertex configurations as they fail to complete intercalations
 13 (**Fig. 4G**, 0.0065 ± 0.005 vs. 0.0116 ± 0.006 fourfold vertices/cell, Kolmogorov-
 14 Smirnov test, $p=0.029$). Interestingly, EpiGraph was able to detect this
 15 solidification of the tissue in the *Mbs-RNAi* discs, even though the shape index
 16 predicted a fluid tissue. Taken together, these results indicate that the
 17 quantification of tissue organization using EpiGraph can infer information
 18 about the fluidity of a tissue from several fixed snapshots, without the need to
 19 laboriously track individual frames of a time-lapse video.

20

21 **DISCUSSION**

22 Textbook definitions of morphogenesis include the term “organization” as
 23 key to explaining this fundamental developmental process (Dai and Gilbert,
 24 1991). The authors wondered, “How can matter organize itself so as to create
 25 a complex structure such as a limb or an eye?”. Later, changes in organization
 26 of adult tissues can reflect pathological traits due to defects in homeostasis
 27 (Csikász-Nagy et al., 2013; Soto and Sonnenschein, 2011). Here, we have
 28 provided a tool that can help to investigate these questions.

29 The analysis of the polygon sides of epithelial cells has been shown to be
 30 insufficient to completely understand tissue organization. Some tessellations
 31 can present very different arrangements yet have the same frequencies of
 32 number of neighbours. A second problem is the lack of a simple value as an

1 indicator of epithelial organization. This feature complicates the comparison
2 between morphogenesis of normal development and that of genetically
3 perturbed or diseased tissues. Our previous attempts to overcome this
4 caveat were based on multi-statistical analyses of graph features (Sanchez-
5 Gutierrez et al., 2013) and the creation of a Voronoi scale to statistically
6 compare groups of images with the CVT reference (Sanchez-Gutierrez et al.,
7 2016). Several recent works cover part of these integrative analyses
8 (Blanchard, 2017; Blanchard et al., 2009; Farrell et al., 2017; Guirao et al.,
9 2015; Jackson et al., 2017). However, we are aware that all these methods
10 are difficult to incorporate into the average biology or biomedicine lab.

11 We have developed EpiGraph, aiming to bring an easy way to quantify
12 tissue organization without the requirement for programming skills. EpiGraph
13 transforms the image into a graph of cell-to-cell contacts and extracts their
14 graphlet content to later compare with other images. These complex
15 algorithms are hidden behind the friendly user window of FIJI. This is the
16 most popular open-source biological image analysis platform. In addition, the
17 output data options of EpiGraph facilitate fast and clear representations and
18 interpretations of the results.

19 One of the strengths of EpiGraph is the comparison of any tessellation with
20 the hexagonal lattice, the “random” Voronoi tessellation and the Voronoi
21 tessellation that presents the “conserved polygon distribution” (Gibson et al.,
22 2006; Sanchez-Gutierrez et al., 2016) (**Fig. 2A-C**). We have tested EpiGraph
23 with different types of samples: as expected, the average of the natural
24 tessellations such as wing imaginal disc (dWL and dWP) or the chicken neural
25 tube (cNT) matched the CVTn path position (**Fig. 2D-G**). We interpret that
26 these three natural samples present similar polygon distributions and graphlet
27 compositions to some Voronoi Diagrams from the CVTn. On the other hand,
28 the average of the Eye samples appeared far from the CVTn when Epi-
29 Voronoi5 or Epi-Random values were plotted (**Fig. 2D-G**). These two
30 references were capturing differences in organization between the Eye and
31 any Voronoi Diagram (including Diagram 1, which presents a similar polygon
32 distribution to the Eye). This result supports the utility of EpiGraph to quantify
33 organizational traits that were not accessible until now. The same idea is

reinforced by the results obtained with the mutant samples for myosin II (dMWP, **Fig. 2D-G**). In previous work, we showed that this set of samples slightly deviated the CVT scale in terms of polygon distribution (Sanchez-Gutierrez et al., 2016). Here we show very clear differences in terms of the values of Epi-Voronoi5 and Epi-Random (**Fig. 2G**), suggesting a higher sensitivity of the new method when capturing differences in organization.

The output images from EpiGraph show the CVTn path as a clear reference for proliferative epithelia such the wing imaginal disc or the chicken neural tube and for vertex model control simulations. We have incorporated a statistical test into EpiGraph that indicates if a new tissue is within or outside of the CVTn path, and which is the Voronoi diagram with the closest organization. The different results comparing Epi-Hexagons, Epi-Random and Epi-Voronoi5 values also suggested that Epi-Hexagons had better resolution for images with a higher percentage of hexagons while Epi-Random and Epi-Voronoi5 were more sensitive to the differences between images with less than 40% of hexagons. For this reason, we have designed the visualization step of the program to easily change the three axes and check the different results using any combination of these GDD references and the “percentage of hexagons”.

Using different sets of simulations, we are able to distinguish two different types of organization: The cases where a subset of cells adopts a particular arrangement inside a mostly ordered tissue (**Fig 3B, C, G**) and the cases where the global topology of the tissue is altered and the cell sizes are very heterogeneous (**Fig. 3E-G** and **Fig. S5**). These two patterns create a “map” of arrangements that are out the CVTn, and they will help to other researchers to study the degree of order in their samples.

The dynamics of the transition between a tissue behaving as a fluid or a solid is an emerging problem in developmental biology and biomedicine (Curran et al., 2017; Firmino et al., 2016; Mongera et al., 2018; Park et al., 2015; Petridou et al., 2018; Tetley and Mao, 2018). We have used the capabilities of EpiGraph to study how the fluidity state can affect the organization of a tissue. The utility of EpiGraph in this regard is supported by its ability to quantify dynamic changes in organization due to cell

1 rearrangements in a vertex model simulation of a soft tissue (**Fig. 4A-B**).
2 Therefore, in these simulations, cell movements are captured as changes in
3 the organization of the tissue by EpiGraph. However, cell rearrangements do
4 not necessarily have to lead to changes in tissue organization, as is often the
5 case in more homeostatic tissues. Although, it has been shown that the late
6 third instar *Drosophila* imaginal disc can exchange neighbours and rearrange
7 during development (Heller et al., 2016), we were not able to see changes in
8 organization combining live imaging of the WT discs and EpiGraph analysis
9 (**Fig. 4E**). Therefore, we interpret that the multiple re-arrangements of the WT
10 disc conserve the organization of the tissue, at least in the time framework
11 analysed (30 min). On the contrary, the hyperactivation of myosin II (*Mbs*-
12 *RNAi*) produced a clear change in the organization of the tissue as detected
13 by EpiGraph. The decrease of intercalation rate and the increase of fourfold
14 vertices in the *Mbs-RNAi* discs suggest that EpiGraph is capturing a change
15 in tissue fluidity (**Fig. 4E-G**). In this respect, we think that EpiGraph analyses
16 provide information beyond previous parameters that have been used to
17 capture the fluidity in cell arrangements such as the shape index (Bi et al.,
18 2015). All the real images analysed in this work have a high shape index (**Fig.**
19 **S6**). These samples include the *Mbs-RNAi* mutant discs, that do not
20 intercalate. Altogether, our results suggest that the shape index is not a
21 sufficient parameter to define fluidity from a still image of a real sample.

22 In biomedicine, a robust and efficient analysis of histopathological images is
23 required. Computerized image tools have an enormous potential to improve
24 the quality of histological image interpretation, offering objective analyses that
25 can aid the pathologist's diagnoses. Changes in organization have proven to
26 be related to the onset of disease in very different contexts, being critical for
27 early detection (Emmanuele et al., 2015; Guillaud et al., 2010; Park et al.,
28 2015; Sáez et al., 2013; Tsuboi et al., 2018). We propose that EpiGraph is
29 able to efficiently detect mutant phenotypes related to changes in
30 organization and/or in tissue fluidity. Importantly, this can be done from a few
31 snapshots in time, without the need for sophisticated time-lapse imaging and
32 tracking. This may provide a simple detection tool for the early onset of

1 disease, where changes in organization can occur, and only limited tissue
2 samples are available from patients.

3 **EpiGraph limitations.**

4 Although EpiGraph accepts a wide range of images as inputs, we have
5 specified some minimum requirements. It is not recommended to use input
6 images bigger than 3000 pixels of width or 3000 pixels of height, since
7 processing them could be computationally intensive. In addition, EpiGraph
8 only accepts single images. Images from time series should be adapted to
9 single frames before uploading them to EpiGraph.

10 Computers with little RAM memory (less than 16gb) will work but with a
11 series of restrictions. To ensure usability, it is not recommended computing
12 images with a high number of cells (more than 1000) due to a possible lack
13 of memory. In the same way, we suggest skeletonizing the edges of the
14 images and using a small radius, i.e. 3 pixels of radius for skeletonized image
15 (we recommend don't overpass a radius value of 10 pixels to avoid
16 overloading the system) to calculate the cells neighbourhood. Choosing a
17 high radius value could slow down the work queue, increasing the use of RAM
18 memory.

19 If any of these requirements are not satisfied, the program alerts the user,
20 allowing him/her to change the image provided. Importantly, the images and
21 ROIs require a minimum number of valid cells (cells without touching the
22 borders or an invalid region of the image) in order to get coherent graphlets.
23 Therefore, to get any result, EpiGraph must detect at least a 3-distance valid
24 cell (see **Fig. 2**) in the case of 7-motifs or 10-motifs or a 4-distance valid cell
25 (see **Fig. 2**) in the case of 17-motifs and 29-motifs. In any case, we strongly
26 recommend having a greater number of 3-distance and 4-distance valid cells
27 to get results that can be trusted in terms of capturing the organization of a
28 tissue. Regarding the 3D visualization tool, it allows the user to see the
29 position of the samples from different angles. However, the resolution of the
30 exported file is only 72 pixels per inch (dpi). This could be too low for
31 publications and therefore EpiGraph provides an excel table with all the
32 information needed to represent it with other programs.

1 In summary, we have generated a very accessible, open source method to
2 produce a quantitative description of developmental events. This quantitative
3 aspect is reinforced by the statistical comparison with the CVT path that
4 serves as a scale for tissue organization. We anticipate that our tool will
5 improve the study of tissue dynamics and morphogenesis by permitting the
6 comparative analysis of epithelial organization in genetically mutated or
7 diseased tissues during time.
8

1 MATERIAL AND METHODS

2 Source images used in the study.

3 Centroidal Voronoi Tessellation (CVT) diagrams and variations

4 For the generation of this set of paths we have used the software Matlab
5 R2014b to iteratively apply Lloyd's algorithm to a random Voronoi tessellation
6 (Lloyd, 1957). This implies that the centroid of a cell in a Voronoi diagram is
7 the seed for the same cell in the next iteration.

8 - *Centroidal Voronoi Tessellation (CVT) diagrams*

9 Centroidal Voronoi Tessellation diagrams were obtained as described
10 previously by our group (Sanchez-Gutierrez et al., 2016). The 20 original
11 Voronoi diagrams were created placing 500 seeds randomly in an image of
12 1024x1024 pixels. A total of 700 iterations were generated for each initial
13 image.

14 - *Centroidal Voronoi Tessellation noise (CVTn) diagrams*

15 We have developed a variation of the CVT path, named the CVT noise
16 (CVTn) path (**Fig. S3**). We started from the same 20 initial random diagrams
17 described above. The development process of the CVTn path was modified
18 so that the new seeds were not strictly the centroid from the previous iteration.
19 In even iterations, we selected a region of 5 pixels of radius from the centroid
20 position, in which seeds could be placed randomly. In odd iterations, the
21 system was stabilized, applying the original Lloyd algorithm. A total of 700
22 iterations were generated for each initial image.

23 Natural packed tissues and vertex model simulations

24 The details of the obtaining and processing of the epithelial images were
25 described in (Escudero et al., 2011). Control vertex model simulations include
26 cell proliferation and are the basis for the other two cases. Case III
27 corresponds to a vertex model simulation with heterogeneous reduction of
28 line tension and an impairment of cell division when tension value is under 30
29 percent of the initial value. Case IV is a similar simulation to Case III with a
30 threshold of 40 percent. Regarding simulations with no cell proliferation, as a
31 baseline, the control had homogeneous parameters for contractility, line
32
33

1 tension and ideal area. ‘Elongated’ simulations were as the control, but with
2 ten percent of cells having a reduced line tension and ideal area, while
3 ‘squared’ ones had ten percent of cells with only line tension reduced. The
4 exact conditions for the vertex model simulations were described in
5 (Sanchez-Gutierrez et al., 2016).

6 Perturbing myosin II activity in *Drosophila* wing discs and calculating 7 intercalation rates

8 *Drosophila* were raised in standard conditions. Wing discs were dissected
9 from third instar larvae and cultured under filters as described by (Zartman et
10 al., 2013). Discs were cultured in Shields and Sang M3 media supplemented
11 with 2% FBS, 1% pen/strep, 3ng/ml ecdysone and 2ng/ml insulin. The
12 following alleles and transgenes were used; *shg*-GFP (Ecad-GFP, Huang et
13 al., 2009), *UAS-Mbs-RNAi* (KK library, VDRC), *rn*-GAL4 (RMCE-MiMIC
14 Trojan-GAL4 collection). The following experimental genotypes were used;
15 Ecad-GFP (WT) and Ecad-GFP/*UAS-Mbs-RNAi*; *rn*-GAL4/+ (*Mbs-RNAi*). For
16 EpiGraph analysis, discs were imaged on a Zeiss LSM 880 microscope with
17 Airyscan at 512x512 resolution with a 63x objective (NA 1.4) at 1.4x zoom for
18 a total of 30 minutes with 1-minute time intervals and a z-step of 0.5µm. Time-
19 lapse image sequences were segmented using Epitools (Heller et al., 2016).

20 To quantify intercalation rates, 5 WT and 5 *Mbs-RNAi* wing discs were
21 imaged using the same methods as above, except using 5x zoom and 3
22 minutes intervals for a total of 2 hours. Intercalation rate data was exported
23 from the “EDGE_T1_TRANSITIONS” overlay in the “CellOverlay” plugin in
24 Epitools. To exclude mistakes generated when 4-way junctions were not
25 recognised, junctions less than 0.075µm in length were assigned a length of
26 0µm. A productive intercalation event was scored when a neighbour
27 exchange was stabilised for at least 2 time points (6 minutes). The total
28 number of tracked cells was also quantified, allowing the intercalation rate to
29 be expressed as the number of intercalations per cell per hour.

30 We also counted the number of fourfold vertices per cell in both WT and
31 *Mbs-RNAi* conditions. In particular, we quantified the number of vertices in
32 which four or more cells were touching each other, using Matlab R2014b. The

1 cells closest to the border of the image were excluded from the analysis. In
2 this way, we obtained the percentage of fourfold vertices per valid cell for
3 each image and calculated a Kolmogorov-Smirnov test to check if the
4 distributions of both conditions were different.

5 **Soft and Rigid tissue simulations**

6 We have extracted a set of screenshots from two videos that simulated
7 different dynamical behaviour of vertex model simulations. These videos are
8 presented as Supplemental Material in (Bi et al., 2016). The first video
9 represents a rigid behaviour in the simulation:
10 [https://journals.aps.org/prx/supplemental/10.1103/PhysRevX.6.021011/solid](https://journals.aps.org/prx/supplemental/10.1103/PhysRevX.6.021011/solid_tissue_v0_0.2_p0_3.5_Dr_0.1.mp4)
11 [_tissue_v0_0.2_p0_3.5_Dr_0.1.mp4](https://journals.aps.org/prx/supplemental/10.1103/PhysRevX.6.021011/solid_tissue_v0_0.2_p0_3.5_Dr_0.1.mp4); the second one represents a soft
12 behaviour: [https://journals.aps.org/prx/supplemental/10.1103/PhysRevX.6.02](https://journals.aps.org/prx/supplemental/10.1103/PhysRevX.6.021011/fluid_tissue_v0_0.2_p0_3.8_Dr_0.1.mp4)
13 [1011/fluid_tissue_v0_0.2_p0_3.8_Dr_0.1.mp4](https://journals.aps.org/prx/supplemental/10.1103/PhysRevX.6.021011/fluid_tissue_v0_0.2_p0_3.8_Dr_0.1.mp4). In both videos, we selected a
14 total of 13 frames with steps of 3.333 seconds (from $t = 0$ to 40 seconds).

15 **Graphlets and motifs selection.**

16 The different images from the previous section were used to create a graph
17 of cell-to-cell contacts ((Escudero et al., 2011) and **Supplementary Material**
18 **and methods**) that served as the source for the graphlet analysis (Pržulj,
19 2007; Pržulj et al., 2004). First, we adapted the graphlet analysis performed
20 by EpiGraph to the nature of our samples (tessellations). Three graphlets
21 were discarded since they were not possible in the context of an epithelial
22 tissue (**Fig. 1** and **Fig. S1**). Second, we used the computer program for
23 graphlet identification and calculation ORCA (Orbit Counting Algorithm)
24 (Hočevár and Demšar, 2014), to extract the different conformations of nodes
25 assembling the graphlets, called orbits (Pržulj, 2007). We computed the
26 Graphlet degree Distribution of the 73 given orbits from the 29 graphlets, and
27 then we removed the non-used ones. The reason to remove these graphlets
28 was that they were either redundant or not possible in a planar tissue. On the
29 first case, G5 and G27 were redundant since, in order to achieve G5 in a
30 plane, there must be a centre cell with 4 sides, the same centre cell captured
31 on G27 (**Fig 1** and **Fig. S1**). It may occur that more than one cell is inside G5,
32 which could not be captured by G27, still it would be captured by G5 and the
33 chances of encounter this setting would be very low. Regarding the second

1 case, G20, G22 and G25 were not possible to achieve in a planar tessellation
2 since it is assumed the convexity of the cells. Therefore, we removed them.

3 **Shape index calculation**

4 We have extracted the shape index, as an indicator of rigidity, from each
5 natural and simulated image, based on (Bi et al., 2015). The global shape
6 index in a tissue was measured as the median of the shape index of the
7 individual valid cells. We quantified the cell area and perimeter using Matlab
8 R2014b. We performed the following approach: We captured the vertex
9 coordinates for each valid cell. Then, we calculated the Euclidean distance
10 between each adjacent vertex, and adding all of them, we got the cell
11 perimeter. From these vertices, a polygon was inferred and we calculated its
12 contained area using the “polyarea” Matlab function.

13 **Statistical analysis.**

14 We have estimated the closest CVTn diagram of a given image in terms of
15 the three GDDs measured in EpiGraph (Epi-Hexagons, Epi-Random and Epi-
16 Voronoi5). We computed the centre of the point cloud formed by the 20
17 randomizations of a particular CVTn diagram as the mean of those twenty
18 images, obtaining a 3D point. Then, we calculated the Euclidean distance
19 between all the CVTn diagrams central points and the three calculated
20 parameters of the input image, obtaining its closest point, which corresponds
21 to its closest diagram. Furthermore, we checked if this image belonged to the
22 closest diagram point cloud using an outlier detection approach. In particular,
23 we tested if the inclusion of the image into a CVTn diagram point cloud would
24 increase or decrease the standard deviation of the original group. We
25 assigned the probability of being an inlier, which is defined as follows:

$$26 \quad confidence\ score = \frac{1}{n} \sum_{i=1}^n \frac{std(closest\ CVTn\ cloud)_i}{std(closest\ CVTn\ cloud + newImg)_i}$$

27 Where n is the total number of coordinates, which in our case is 3 due to
28 the three-dimensional space; the parameter stands for every different
29 coordinate (Epi-Hexagons, Epi-Random and Epi-Voronoi5); represents the
30 values of the 20 images from the closest CVTn diagram in a specific
31 coordinate and is the value of the input image for the same coordinate. The

1 values range from 0 (very far from point cloud) to $+\infty$ (inside point cloud). We
2 have estimated that with a confidence of > 0.95 the input image is considered
3 to be an inlier.

4

5

6

7

REFERENCES

- Albert R, Barabasi A-L. 2002. Statistical mechanics of complex networks. *Rev Mod Phys* **74**:47–97.
- Benson AR, Gleich DF, Leskovec J. 2016. Higher-order organization of complex networks. *Science (80-)* **353**:163–166. doi:10.1038/s41588-018-0165-1
- Bertet C, Sulak L, Lecuit T. 2004. Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation. *Nature* **429**:667–671.
- Bi D, Lopez JH, Schwarz JM, Manning ML. 2015. A density-independent rigidity transition in biological tissues. *Nat Phys* **11**:1074–1079. doi:10.1038/nphys3471
- Bi D, Yang X, Marchetti MC, Manning ML. 2016. Motility-driven glass and jamming transitions in biological tissues. *Phys Rev X* **6**. doi:10.1103/PhysRevX.6.021011
- Blanchard GB. 2017. Taking the strain: Quantifying the contributions of all cell behaviours to changes in epithelial shape. *Philos Trans R Soc B Biol Sci*. doi:10.1098/rstb.2015.0513
- Blanchard GB, Kabla AJ, Schultz NL, Butler LC, Sanson B, Gorfinkel N, Mahadevan L, Adams RJ. 2009. Tissue tectonics: morphogenetic strain rates, cell shape change and intercalation. *Nat Methods* **6**:458–464. doi:nmeth.1327 [pii]10.1038/nmeth.1327
- Blankenship JT, Backovic ST, Sanny JS, Weitz O, Zallen JA. 2006. Multicellular rosette formation links planar cell polarity to tissue morphogenesis. *Dev Cell* **11**:459–470.
- Costa LDAF, Rodrigues FA, Travieso G, Boas PR V. 2007. Characterization of complex networks: a survey of measurements. *Adv Phys* **56**:167–242.
- Csikász-Nagy A, Escudero LM, Guillaud M, Sedwards S, Baum B, Cavaliere M. 2013. Cooperation and competition in the dynamics of tissue architecture during homeostasis and tumorigenesis. *Semin Cancer Biol* **23**. doi:10.1016/j.semcancer.2013.05.009
- Curran S, Strandkvist C, Bathmann J, de Gennes M, Kabla A, Salbreux G, Baum B. 2017. Myosin II Controls Junction Fluctuations to Guide Epithelial Tissue Ordering. *Dev Cell* **43**:480–492.e6. doi:10.1016/j.devcel.2017.09.018
- Dai JD, Gilbert LI. 1991. Metamorphosis of the corpus allatum and degeneration of the prothoracic glands during the larval-pupal-adult transformation of *Drosophila melanogaster*: a cytophysiological analysis of the ring gland. *Dev Biol* **144**:309–326.
- Emmanuele V, Kubota A, Garcia-Diaz B, Garone C, Akman HO, Sánchez-Gutiérrez D, Escudero LM, Kariya S, Homma S, Tanji K, Quinzii CM, Hirano M. 2015. Fhl1 W122S causes loss of protein function and late-onset mild myopathy. *Hum Mol Genet* **24**. doi:10.1093/hmg/ddu490
- Escudero LM, Bischoff M, Freeman M. 2007. Myosin II regulates complex cellular arrangement and epithelial architecture in *Drosophila*. *Dev Cell* **13**:717–729.
- Escudero LM, Costa Lda F, Kicheva A, Briscoe J, Freeman M, Babu MM. 2011. Epithelial organisation revealed by a network of cellular contacts. *Nat*

1 *Commun* **2**:526. doi:10.1038/ncomms1536ncomms1536 [pii]

2 Etournay R, Merkel M, Popović M, Brandl H, Dye NA, Aigouy B, Salbreux G,
3 Eaton S, Jülicher F. 2016. TissueMiner: A multiscale analysis toolkit to
4 quantify how cellular processes create tissue dynamics. *Elife* **5**:773–786.
5 doi:10.7554/eLife.14334

6 Farhadifar R, Roper JC, Aigouy B, Eaton S, Jülicher F. 2007. The influence
7 of cell mechanics, cell-cell interactions, and proliferation on epithelial
8 packing. *Curr Biol* **17**:2095–2104.

9 Farrell DL, Weitz O, Magnasco MO, Zallen JA. 2017. SEGGA: a toolset for
10 rapid automated analysis of epithelial cell polarity and dynamics.
11 *Development* **144**:1725–1734. doi:10.1242/dev.146837

12 Firmino J, Rocancourt D, Saadaoui M, Moreau C, Gros J. 2016. Cell
13 Division Drives Epithelial Cell Rearrangements during Gastrulation in Chick.
14 *Dev Cell* **36**:249–261. doi:10.1016/j.devcel.2016.01.007

15 Fletcher AG, Osterfield M, Baker RE, Shvartsman SY. 2014. Vertex models
16 of epithelial morphogenesis. *Biophys J* **106**:2291–2304.
17 doi:10.1016/j.bpj.2013.11.4498

18 Gibson MC, Patel AB, Nagpal R, Perrimon N. 2006. The emergence of
19 geometric order in proliferating metazoan epithelia. *Nature* **442**:1038–1041.

20 Girdler GC, Roper K. 2014. Controlling cell shape changes during salivary
21 gland tube formation in *Drosophila*. *Semin Cell Dev Biol* **31**:74–81.
22 doi:10.1016/j.semcdb.2014.03.020

23 Guillaud M, Clem C, Macaulay C. 2010. An in silico platform for the study of
24 epithelial pre-invasive neoplastic development. *Biosystems* **102**:22–31.
25 doi:10.1016/j.biosystems.2010.07.008S0303-2647(10)00116-4 [pii]

26 Guirao B, Rigaud SU, Bosveld F, Bailles A, López-Gay J, Ishihara S,
27 Sugimura K, Graner F, Bellaïche Y. 2015. Unified quantitative
28 characterization of epithelial tissue development. *Elife* **4**:773–786.
29 doi:10.7554/eLife.08519

30 Hayes W, Sun K, Pržulj N. 2013. Graphlet-based measures are suitable for
31 biological network comparison. *Bioinformatics* **29**:483–491.
32 doi:10.1093/bioinformatics/bts729

33 Heller D, Hoppe A, Restrepo S, Gatti L, Tournier AL, Tapon N, Basler K,
34 Mao Y. 2016. EpiTools: An Open-Source Image Analysis Toolkit for
35 Quantifying Epithelial Growth Dynamics. *Dev Cell* **36**:103–116. doi:S1534-
36 5807(15)00797-2 [pii]10.1016/j.devcel.2015.12.012

37 Ho H, Milenković T, Memisević V, Aruri J, Pržulj N, Ganesan AK. 2010.
38 Protein interaction network topology uncovers melanogenesis regulatory
39 network components within functional genomics datasets. *BMC Syst Biol*
40 **4**:84. doi:10.1186/1752-0509-4-84

41 Hočevár T, Demšar J. 2014. A combinatorial approach to graphlet counting.
42 *Bioinformatics* **30**:559–565. doi:10.1093/bioinformatics/btt717

43 Jackson MD, Xu H, Duran-Nebreda S, Stamm P, Bassel GW. 2017.
44 Topological analysis of multicellular complexity in the plant hypocotyl. *Elife*
45 **6**. doi:10.7554/eLife.26023

46 Khan Z, Wang Y-C, Wieschaus EF, Kaschube M. 2014. Quantitative 4D
47 analyses of epithelial folding during *Drosophila* gastrulation. *Development*

1 **141**:2895–2900. doi:10.1242/dev.107730

2 Kuchaiev O, Stevanović A, Hayes W, Pržulj N. 2011. GraphCrunch 2:
3 Software tool for network modeling, alignment and clustering. *BMC*
4 *Bioinformatics* **12**. doi:10.1186/1471-2105-12-24

5 Kursawe J, Bardenet R, Zartman JJ, Baker RE, Fletcher AG. 2016. Robust
6 cell tracking in epithelial tissues through identification of maximum common
7 subgraphs. *J R Soc Interface* **13**. doi:10.1098/rsif.2016.0725

8 Lecuit T, Lenne PF. 2007. Cell surface mechanics and the control of cell
9 shape, tissue patterns and morphogenesis. *Nat Rev Mol Cell Biol* **8**:633–
10 644.

11 Lloyd S. 1957. Least square quantization in PCM's. *Bell Teleph Lab* **28**:129–
12 137.

13 Mirams GR, Arthurs CJ, Bernabeu MO, Bordas R, Cooper J, Corrias A,
14 Davit Y, Dunn SJ, Fletcher AG, Harvey DG, Marsh ME, Osborne JM,
15 Pathmanathan P, Pitt-Francis J, Southern J, Zemzemi N, Gavaghan DJ.
16 2013. Chaste: an open source C++ library for computational physiology and
17 biology. *PLoS Comput Biol* **9**:e1002970.
18 doi:10.1371/journal.pcbi.1002970PCOMPBIOL-D-12-01337 [pii]

19 Mongera A, Rowghanian P, Gustafson HJ, Shelton E, Kealhofer DA, Carn
20 EK, Serwane F, Lucio AA, Giammona J, Campàs O. 2018. A fluid-to-solid
21 jamming transition underlies vertebrate body axis elongation. *Nature*
22 **561**:401–405. doi:10.1038/s41586-018-0479-2

23 Park JA, Kim JH, Bi D, Mitchel JA, Qazvini NT, Tantisira K, Park CY, McGill
24 M, Kim SH, Gweon B, Notbohm J, Steward R, Burger S, Randell SH, Kho
25 AT, Tambe DT, Hardin C, Shore SA, Israel E, Weitz DA, Tschumperlin DJ,
26 Henske EP, Weiss ST, Manning ML, Butler JP, Drazen JM, Fredberg JJ.
27 2015. Unjamming and cell shape in the asthmatic airway epithelium. *Nat*
28 *Mater* **14**:1040–1048. doi:10.1038/nmat4357

29 Petridou NI, Grigolon S, Salbreux G, Hannezo E, Heisenberg C-P. 2018.
30 Fluidization-mediated tissue spreading by mitotic cell rounding and non-
31 canonical Wnt signalling. *Nat Cell Biol* **1**. doi:10.1038/s41556-018-0247-4

32 Pilot F, Lecuit T. 2005. Compartmentalized morphogenesis in epithelia: from
33 cell to tissue shape. *Dev Dyn* **232**:685–694.

34 Pržulj N. 2007. Biological network comparison using graphlet degree
35 distribution. *Bioinformatics* **23**:853–854. doi:10.1093/bioinformatics/btl301

36 Pržulj N, Corneil DG, Jurisica I. 2004. Modeling interactome: Scale-free or
37 geometric? *Bioinformatics* **20**:3508–3515. doi:10.1093/bioinformatics/bth436

38 Sáez A, Rivas E, Montero-Sánchez A, Paradas C, Acha B, Pascual A,
39 Serrano C, Escudero LM. 2013. Quantifiable diagnosis of muscular
40 dystrophies and neurogenic atrophies through network analysis. *BMC Med*
41 **11**:77. doi:10.1186/1741-7015-11-77

42 Sánchez-Gutiérrez D, Sáez A, Gómez-Gálvez P, Paradas C, Escudero LM.
43 2017. Rules of tissue packing involving different cell types: Human muscle
44 organization. *Sci Rep* **7**:40444. doi:10.1038/srep40444

45 Sanchez-Gutierrez D, Saez A, Pascual A, Escudero LM. 2013. Topological
46 progression in proliferating epithelia is driven by a unique variation in
47 polygon distribution. *PLoS One* **8**:e79227.

1 doi:10.1371/journal.pone.0079227PONE-D-13-27658 [pii]
2 Sanchez-Gutierrez D, Tozluoglu M, Barry JD, Pascual A, Mao Y, Escudero
3 LM. 2016. Fundamental physical cellular constraints drive self-organization
4 of tissues. *EMBO J* **35**:77–88. doi:10.15252/embj.201592374
5 Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T,
6 Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ,
7 Hartenstein V, Eliceiri K, Tomancak P, Cardona A. 2012. Fiji: an open-
8 source platform for biological-image analysis. *Nat Methods* **9**:676–682.
9 doi:10.1038/nmeth.2019
10 Soto AM, Sonnenschein C. 2011. The tissue organization field theory of
11 cancer: A testable replacement for the somatic mutation theory. *BioEssays*
12 **33**:332–340. doi:10.1002/bies.201100025
13 Tanaka S, Sichau D, Iber D. 2015. LBIBCell: a cell-based simulation
14 environment for morphogenetic problems. *Bioinformatics* **31**:2340–7.
15 doi:10.1093/bioinformatics/btv147
16 Tetley RJ, Mao Y. 2018. The same but different: Cell intercalation as a
17 driver of tissue deformation and fluidity. *Philos Trans R Soc B Biol Sci*.
18 doi:10.1098/rstb.2017.0328
19 Tsuboi A, Ohsawa S, Umetsu D, Sando Y, Kuranaga E, Igaki T, Fujimoto K.
20 2018. Competition for Space Is Controlled by Apoptosis-Induced Change of
21 Local Epithelial Topology. *Curr Biol* **28**:2115–2128.
22 doi:10.1016/J.CUB.2018.05.029
23 Weigert M, Schmidt U, Boothe T, Müller A, Dibrov A, Jain A, Wilhelm B,
24 Schmidt D, Broaddus C, Culley S, Rocha-Martins M, Segovia-Miranda F,
25 Norden C, Henriques R, Zerial M, Solimena M, Rink J, Tomancak P, Royer
26 L, Jug F, Myers EW. 2018. Content-aware image restoration: pushing the
27 limits of fluorescence microscopy. *Nat Methods* **15**:1090–1097.
28 doi:10.1038/s41592-018-0216-7
29 Yamashita S, Michiue T. 2014. Quantitative analysis of cell arrangement
30 indicates early differentiation of the neural region during *Xenopus*
31 gastrulation. *J Theor Biol* **346**:1–7. doi:10.1016/j.jtbi.2013.12.029
32 Zartman J, Restrepo S, Basler K. 2013. A high-throughput template for
33 optimizing *Drosophila* organ culture with response-surface methods.
34 *Development* **140**:2848–2848. doi:10.1242/dev.098921
35
36
37

1 **Figure 1. Graphlets, cellular motifs and characterization of epithelial**
2 **organization. A)** A representation of the cellular motifs that correspond to
3 graphlets of up to five nodes. There are 29 motifs corresponding to 26
4 different graphlets (**Fig. S1**). Note that one graphlet can represent two cellular
5 motifs (G8, G23 and G26). Mauve motifs form the 17-motifs set. Prussian
6 Blue motifs stands for the set 29-motifs. In the first row are the motifs that
7 account for the organization of groups of up to 4 cells (10-motifs). Therefore,
8 7-motifs set is formed by the mauve coloured graphlets at the first row. **B)**
9 Polygon distribution comparison of images from: Voronoi diagram 1 (black
10 bar); Eye (orange), *Drosophila* eye disc: 3 samples; Voronoi diagram 4 (grey);
11 dWL (green), *Drosophila* larva wing disc: 15 samples; Voronoi diagram 5 (light
12 grey), dWP (red), *Drosophila* prepupal wing imaginal disc epithelium: 16
13 samples. Data shown refer to the mean \pm SEM. Diagram 1, 4 and 5: 20
14 replicates. **C)** GDD value calculation (17-motifs) between natural images and
15 Voronoi diagrams with similar polygon distribution. The data shown are the
16 mean of the GDD between each pair of images.

Figure 2. Epithelial organization of biological tissues with respect to the CVTn. **A-C)** Tessellations with the corresponding graph of cell-to-cell contacts for a perfect hexagonal arrangement (**A**) a Voronoi Diagram 1 (**B**) and a Voronoi Diagram 5 (**C**) from a CVTn. These tessellations represent the diagrams used as reference to calculate the Epi-Hexagons, Epi-Random and Epi-Voronoi5 respectively. The light blue edges in these panels represent the cellular connectivity network. The colourful nodes mark the valid cells that were involved in the cellular motifs to measure graphlets presence. The dark blue and green nodes are the 3-distance valid cells (cells connected exclusively to valid cells within a distance of 3 edges), which were used to calculate the graphlets for 10-motifs and 7-motifs. The green nodes are the 4-distance valid cells (cells connected exclusively to valid cells within a distance of 4 edges) that were used to quantify the graphlets for 29-motifs and 17-motifs. Cells without nodes were no valid cells for graphlet calculation. **D)** Representative images from the natural tessellations. **E)** Plots showing the different combinations of the values for 17-motifs of Epi-Hexagons, Epi-Random and Epi-Voronoi5 with the percentage of hexagons. The diagrams of the CVTn path from the iteration 1 until the iteration 700 are represented as a greyscale beginning in black and reducing its darkness with the increase of the iterations (from 1 to 20, from 30 to 100 by stepwise of 10 and from 100 to 700 by stepwise of 100). **F-G)** Charts representing the comparisons Epi-Hexagons against Epi-Random, and Epi-Random against Epi-Voronoi5, respectively. The CVTn path in both scatter plots, is formed by the diagrams with numbers between 1 and 100, in a greyscale as in (**E**). The natural tessellations are: dMWP (violet), *Drosophila* mutant wing disc: 3 samples; cNT (light blue), chicken neural tube: 16 samples; Eye, dWL and dWP are the same replicates than **Fig.1** and preserve their colour codes. Circumferences are individual values, circles are the average value obtained from the individual samples from each category.

Figure 3. Comparison of different simulations and mutants with the CVTn. A-C) Representative images for non-proliferative simulations. Control with homogeneous parameters (**A**). The ‘squared’ simulations are similar to control, but a ten percent of cells (randomly chosen) have a reduced line tension (**B**). The ‘elongated’ simulations have a ten percent of cells (randomly chosen) with its line tension and ideal area reduced, and the another 90% of cells have the same parameters than control simulations (**C**). **D)** Cell arrangement resulting from the control simulation that includes cell proliferation. **E-F)** Diagrams resulting from a vertex model simulation with an increase of the ideal area value, with respect the control, in some cells. Case III and Case IV slightly differ in the line-tension parameter conditions (see **Material and methods**). **G)** Plots showing the values of Epi-Random vs Epi-Voronoi5 and the percentage of hexagons vs Epi-Voronoi5 (17-motifs) for CVTn, dMWP, Eye, cNT, dWL, dWP; Proliferative Control (20 replicates, carnation pink), Case III (17 replicates, hot pink) and Case IV (15 replicates, purple); Non-proliferative control (20 replicates, blue bell), Squared (20 replicates, azure blue) and Elongated simulations (20 replicates, cornflower Blue). The diagrams of the CVTn path from the iteration 1 until the iteration 100 are represented as a greyscale beginning in black and reducing its darkness with the increase of the iterations; dMWP, Eye, cNT, dWL and dWP have the same replicates and colour codes than Fig.2; circumferences are individual values, circles are the average value obtained from the individual samples from each category.

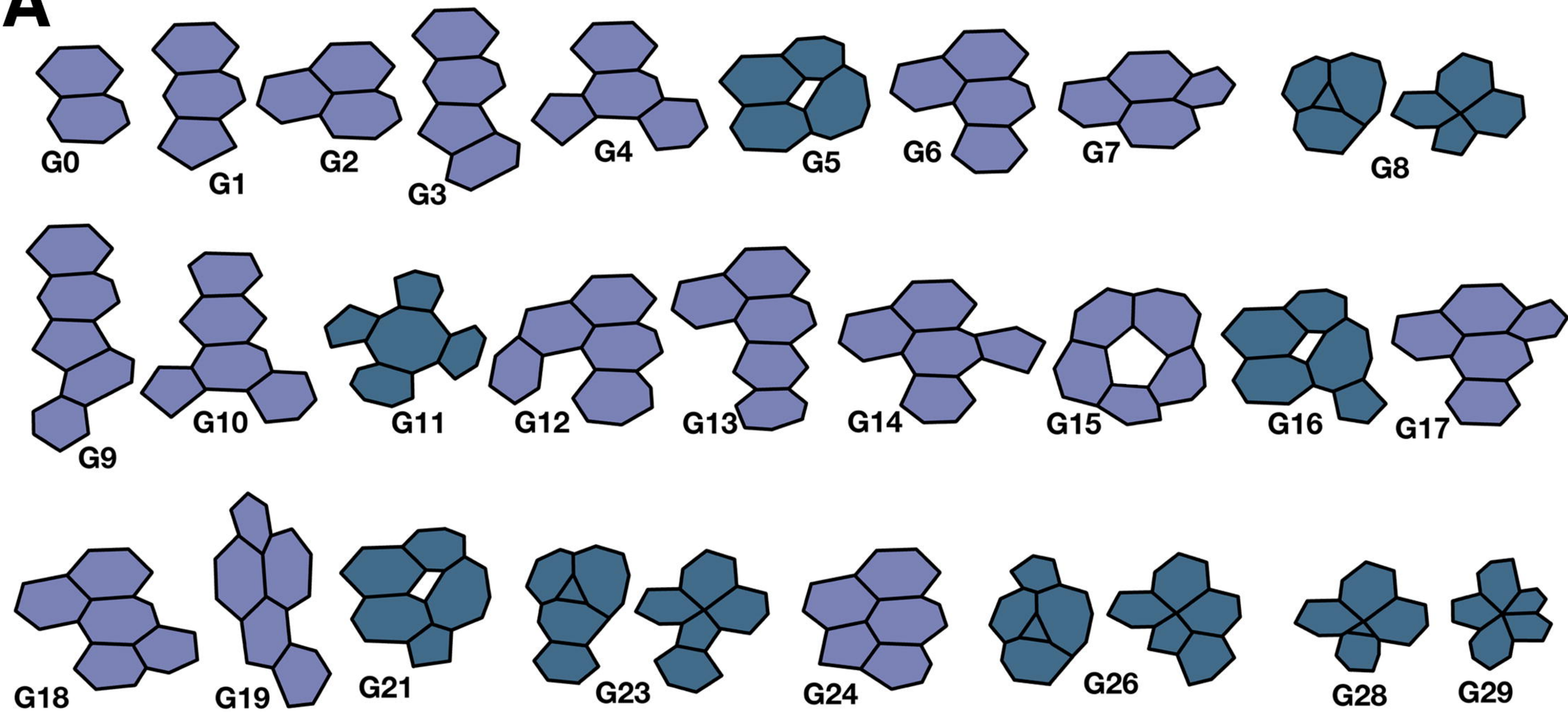
24
25

Figure 4. The rigidity/fluidity of a tissue can be assessed using EpiGraph. **A)** Initial and final frames of two simulations with different settings: a rigid and a soft tissue. **B)** EpiGraph's 3D plot with Epi-Random, Epi-Hexagons and Epi-Voronoi5 axes, showing the soft simulation tissue in green dots and the rigid simulation as orange dots. Each simulation is represented in 13 frames (see **Material and methods**). **C-D)** Representative examples of segmented images from the third instar *Drosophila* imaginal disc in different conditions: Wild Type (**C**); a solid mutant, *Mbs-RNAi* in (**D**). **E)** Plot comparing the fluidity and organization of the tissues in (**C-D**). CVTn (until diagram 30) displayed in greyscale. Dots in scales of blue represent the WT condition: wing disc 1, aquamarine; wing disc 2, light blue; wing disc 3, dark blue. Represented with points in tones of orange-red, *Mbs-RNAi*: sample 1, salmon colour; sample 2, orange; sample 3, red. **F-G)** Boxplot of the intercalation rate (**F**), which is the number of T1 transitions per cell per hour, and the fourfold vertices found per cell (**G**) (note that no fivefold vertices, or beyond, was found on any sample). Boxes stand for the data inside the upper and lower quartiles, while the vertical dashed lines (whiskers) indicate the variability outside them. Mean (dashed line) and median (thick line) of each condition is represented inside each box. The actual values are also presented as circles (and the outlier values as circumferences) with its correspondent colour. In addition, statistical significance, by means of a Kolmogorov-Smirnov test, is shown in the top of both panels (**F**: '***' $p < 0.01$, **G**: '**' $p < 0.05$). Each condition has 3 samples (different colour tone), and the numeric tag represents its frame. In WT have been taken 6 frames per sample in periods of 6 minutes. In the case of *Mbs-RNAi* were taken 3 frames per sample with time lapse of 15 minutes. All the conditions have been tracked for 30 minutes (see **Material and methods**).

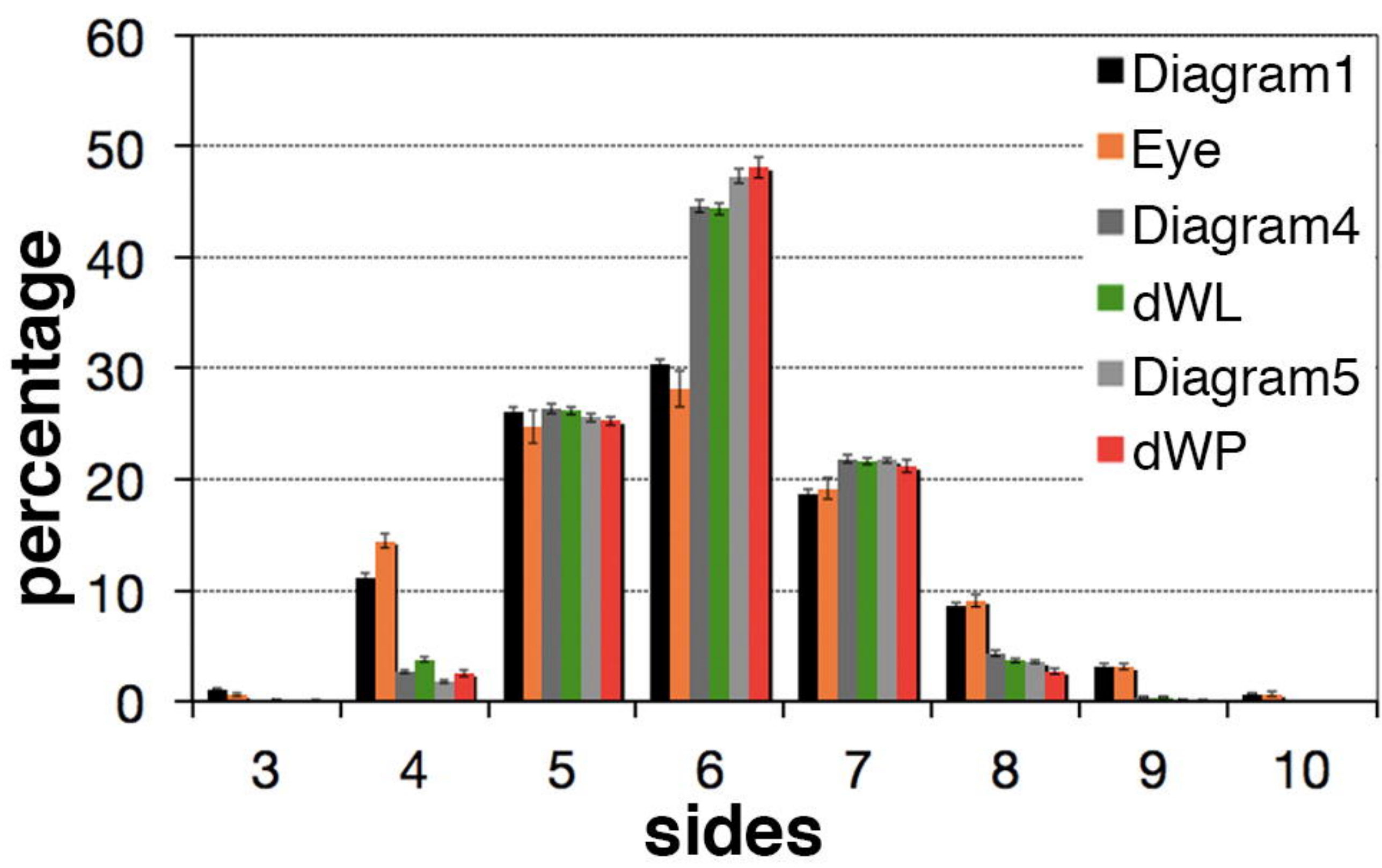
28

29

A



B



C

