From heterogeneous morphogenetic fields to homogeneous regions as a step towards understanding complex tissue dynamics

Satoshi Yamashita, Boris Guirao and Francois Graner
DOI: 10.1242/dev.199034

Editor: Steve Wilson

Review timeline

Original submission:	26 November 2020
Editorial decision:	12 March 2021
First revision received:	30 June 2021
Editorial decision:	19 August 2021
Second revision received:	20 October 2021
Accepted:	25 October 2021

Original submission

First decision letter

MS ID#: DEVELOP/2020/199034

MS TITLE: From heterogenous morphogenetic fields to homogeneous regions as a step towards understanding complex tissue dynamics

AUTHORS: Satoshi Yamashita, Boris Guirao, and Francois Graner

I apologise hugely for theunacceptablylong period oftimethat it took to obtain referees' reports on the above manuscript. I have now, however received referees' comments. They are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

As you will see, one of the referees is largely happy with your revisions whereas the other two raise significant concerns that preclude publication of the study in its current form. It is unusual to allow manuscripts to go through a second round of major revisions at Development but I would like to give you the opportunity to consider whether you think it is feasible to address the issues of concern raised by the two more critical reviewers. Consequently, if you are able to revise the manuscript along the lines suggested,I will be happy receive a revised version of the manuscript. I do appreciate that having already revised the manuscript once, you may choose not to undertake further revisions and submit elsewhere. Please let us know if you do decide to do this.

We are aware that you may be experiencing disruption to the normal running of your lab that make experimental revisions challenging. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Please attend to all of the reviewers' comments and ensure that you clearly highlight all changes made in the revised manuscript. Please avoid using 'Tracked changes' in Word files as these are lost in PDF conversion. I should be grateful if you would also provide a point-by-point response detailing

how you have dealt with the points raised by the reviewers in the 'Response to Reviewers' box. If you do not agree with any of their criticisms or suggestions please explain clearly why this is so.

Reviewer 1

Advance summary and potential significance to field

The manuscript is much improved. Not entirely convinced of the generic applicability of the new methodology for the analysis of the many different tissue types developmental biologists might like to investigate. However the authors are now much clearer on the methods limitations, as well as benefits, and this is appreciated.

Comments for the author

The authors have satisfactorily addressed all my comments.

Reviewer 2

Advance summary and potential significance to field

Yamashita et al describe a method to partition a developing tissue into different regions, according to homogeneity in quantifications of cellular deformation. The authors start by testing different methods to identify regions which share a certain level of homogeneity in terms of cellular area change. They find that a combination of a label propagation algorithm and boundary smoothing returns relatively homogeneous and smooth regions. They then consider partitions based on alternative scalar quantifications of deformation such as the average rate of tissue elongation, parallel to the direction of elongation. The authors then compare the result of one of these segmentations to the anatomical regions of the notum, and apply the same analysis to wing blade morphogenesis.

Comments for the author

The analysis described by the authors is done with great care and follows a clear logical flow in trying and improving different methods. The authors also provide a clear and honest assessment of their results.

Despite these positive points, I am not convinced about the usefulness of the approach, for the following reasons:
1) The improvement in intra-regional homogeneity brought by the label propagation over region-growing algorithm is clearly exposed in the manuscript. But besides being better than a worse case scenario, it is not clear what makes the method proposed by the authors a good choice, more so than any other possible algorithm not tested in the manuscript. In other words, I am not sure how the authors decide that a given segmentation algorithm returns an acceptable result. For instance, I could imagine another type of method where local maxima and minima of the scalar field are detected, and used as seeds for a watershed algorithm.
Is it clear that this alternative method would be worse than what the authors propose?

2) It seems to me that the authors method offers little control over the properties of the segmented regions.
For instance I think it would be more interesting if the degree of homogeneity in the regions of a segmentation could be systematically varied, so that an optimal value of homogeneity that best matches gene expression patterns could be deducted.

3) The success of a major test of the method, Fig. 8, which compares regions detected by the algorithm against anatomical features of the adult not, is not clear. I do not know if the anatomical regions (“scutum scutellum, scutum-scutellum boundary”) are designated as such simply because of the order in which they appear in the segmentation along the anteroposterior axis, or if there is something more precise that allows to clearly distinguish which of these different regions give rise

to distinctive adult tissues (for instance, the size of a domain of gene expression).In addition, some of the detected subregions (e.g. pale blue and green in Fig. 8) are brought together as a single anatomic region; it is therefore not clear if there is any significance to the appearance of these subregions.
Overall, I think that the authors came up with an interesting idea to identify morphogenetic domains in a developing tissue, but a certain degree of arbitrariness and lack of control in the method proposed, and unclear validation with other anatomical features or gene expression pattern, reduce its potential utility.

Reviewer 3

Advance summary and potential significance to field

In this manuscript, Yamashita et al. propose an algorithmic pipeline by which one can coarse grain cellular-scale segmentations into larger, more homogeneous morphological regions. Such regions can be defined for any developmental feature of interest with a well-defined metric, e.g. cell area or time-averaged deformation rate. To my knowledge, no such algorithm currently exists and thus represents the main novelty of the manuscript. However, the algorithm depends upon many parameters. In its current form, the manuscript does not adequately quantify and discuss how the resulting partition depends upon variation of these parameters.

Comments for the author

Major Comments:
1. I reiterate my original concern regarding the verification of the algorithm's accuracy against simulated data. I agree that simulating tissue dynamics of interacting regions presents a challenge that is outside the scope of the present work, however, it seems feasible to populate a vertex model with different regions of static geometric parameters, e.g. intrinsic cell area in the usual area elasticity, to see how accurately the boundaries of each region could be reconstructed from a known morphological quantity. Additionally, this would help clarify all subsequent questions below.

2. The authors have only partially addressed my original concern regarding the choice of the number of regions partitioned from an input tissue. I found the additions to the manuscript that explored this parameter incomplete.
(i) The authors state that an increase in the number of regions resulted in subdividing the regions already obtained? while referring to Figure S2. Upon inspection of Fig. S2 it appears this is true starting from 3 (top row) through 9 (bottom row) initial seeds. However, judging from Figures 7 and 9, 6 seeds were chosen for both the notum and 4 seeds for the pupal wing. Why were 6 and 4 chosen specifically? For example, 6 seeds on the notum appear to simply subdivide the regions obtained by 5? For example, the cellular dynamics between regions 2 and 3 of the notum look qualitatively similar, modulo a small 2nd bump in cell divisions. This also happened to be the partition added in going from 5 seeds to 6 seeds. How do we determine if this region is statistically significant?
(ii) The authors state that when the number is too large and a result of the initial label propagation included a too small region, the small region tended to disappear in the cellular Potts model smoothing, and thus the final label propagation resulted in undesired disconnected regions? This feels like an important part that requires more explanation. One should define too small? quantitatively for the reader. Does this imply the algorithm has a hard size cutoff, under which we should not expect reliable detection? Or is this a statement about relative sizes of regions, i.e. the algorithm will only detect regions of similar size?
(iii) How does the mean silhouette value depend upon the number of clusters? If you add too many do you begin to overfit Do different morphological features want to cluster under a different number of regions, allowing for an average silhouette value closer to 1? If it is strongly dependent, could one think of devising an information criterion to rigorously choose this number?

3. The manuscript's discussion of the Potts model is incomplete. As presented, the Potts model is run to both smooth the boundaries of the segmented regions obtained from the label propagation and to ensure all regions were simply connected. The manuscript would benefit greatly from an explicit discussion of a few features of this step of the algorithm:

(i) There are 50 metropolis steps, i.e. 50 pixels from the segmentation will be chosen at random to flip labels, for each Potts model run. Why is 50 hard-coded here? There is no reasoning given by the authors and no citation given that 50 is an adequate number of metropolis steps.
Intuitively, I would have assumed this number should depend extensively on the number of boundary pixels passed into the input, or at the very least the number of pixels in the original image.
Said another way, it would be useful to make this an explicit parameter of the algorithm and show in a supplement how the boundary area depends upon this number (and maybe how the average silhouette degrades).
(ii) How is the 4d parameter space screened?. Is it an exhaustive search on a fixed grid or is this explicitly optimized? Is it run on the label propagated result or on one instance of the Potts model minimization? How strongly do the results depend on our choice of parameters here? As written the manuscript doesn't give the reader a sense of this.
(iii) What is exactly gained by running the label propagation a second time, especially as it can spontaneously create disconnected regions after running the Potts model which is supposed to remove such regions from the first label propagation. Is it run simply because 50 metropolis steps poorly sample the Potts model energy landscape?

Minor Comments:
1. There are many typos and grammatical errors throughout the text that require attention. Examples (non-exhaustive) below:
* L8-9: missing articles, e.g. should be "Coordination between the invaginating mesoderm..."
* L26: "Recent formalisms have enabled [one] to measure..."
* L74: "in histology or of immuno-stained image..."
* L344: "followinig quantificationi"
2. I assume that plots of the dynamics of each region's cellular processes display the mean across the identified pixels. It would be useful to plot the standard deviation as an error bar so we can assess the variation across each region.

First revision

Author response to reviewers' comments

We thank the reviewers for the thoughtful and relevant comments which helped to improve significantly the manuscript. In below we answer point by point.

> Reviewer 2 Comments for the Author:
The analysis described by the authors is done with great care and follows a clear logical flow in trying and improving different methods. The authors also provide a clear and honest assessment of their results. Despite these positive points, I am not convinced about the usefulness of the approach, for the following reasons:

We warmly thank the referee for this positive opinion and for constructive comments. We answer them point by point, as detailed below. We believe these changes result in a significantly improved manuscript. We hope the revised manuscript now meets the standards for publication.

> 1) The improvement in intra-regional homogeneity brought by the label propagation over region-growing algorithm is clearly exposed in the manuscript. But besides being better than a worse case scenario, it is not clear what makes the method proposed by the authors a good choice, more so than any other possible algorithm not tested in the manuscript. In other words, I am not sure how the authors decide that a given segmentation algorithm returns an acceptable result. For instance, I could imagine another type of method where local maxima and minima of the scalar field are detected, and used as seeds for a watershed algorithm. Is it clear that this alternative method would be worse than what the authors propose?

We understand the point but we don’t have the alternative method to be compared. In future when other methods are proposed, this method can be used to test if they are better.

For the watershed algorithm, it can only divide a scalar field, but not a field of a quantity without order relation. In contrast to it, our method can accept any quantity of biological interest. We added this in the introduction P4L38:
“In contrast to other image segmenting methods like the watershed algorithm, this method is designed to accept any kind of quantity of biological interest, not only a scalar but also a vector, a tensor, and combination of them.”

> 2) It seems to me that the authors method offers little control over the properties of the segmented regions.
For instance I think it would be more interesting if the degree of homogeneity in the regions of a segmentation could be systematically varied, so that an optimal value of homogeneity that best matches gene expression patterns could be deducted.

We agree that the method should have a control over the resultant regions shape. Since the homogeneity was employed as a measurement in the smoothing, it is a variable quantity but not a tunable parameter. Instead, we tried changing the minimum circularity for the smoothing parameter screening, and could change the regions boundary complexity.
We included this new data in P10L211, Fig. S2.

> 3) The success of a major test of the method, Fig. 8, which compares regions detected by the algorithm against anatomical features of the adult not, is not clear. I do not know if the anatomical regions (“scutum, scutellum, scutum-scutellum boundary”) are designated as such simply because of the order in which they appear in the segmentation along the anteroposterior axis, or if there is something more precise that allows to clearly distinguish which of these different regions give rise to distinctive adult tissues (for instance, the size of a domain of gene expression).In addition, some of the detected subregions (e.g. pale blue and green in Fig. 8) are brought together as a single anatomic region; it is therefore not clear if there is any significance to the appearance of these subregions.

For the identification of scutum and scutellum and their boundary, we can refer to positions of macrochaetes (large bristles). We added this explanation in Fig. 9:
“The anatomical regions were identified according to positions of macrochaetae (orange circles in B).” For the green region, it might be close to a wing hinge which is attached to a lateral side of the notum. So it is reasonable to find a distinctive region there. However, the wing hinge was outside of plane taken by our microscopy, and we could not confirm it. Therefore in this paper we included the green region into the scutum.

> Overall, I think that the authors came up with an interesting idea to identify morphogenetic domains in a developing tissue, but a certain degree of arbitrariness and lack of control in the method proposed, and unclear validation with other anatomical features or gene expression pattern, reduce its potential utility.

Thanks to the suggestive comments, our manuscript is now revised and clarified as replied above.

Reviewer 3 Advance Summary and Potential Significance to Field:
> In this manuscript, Yamashita et al. propose an algorithmic pipeline by which one can coarse grain cellular-scale segmentations into larger, more homogeneous, morphological regions. Such regions can be defined for any developmental feature of interest with a well-defined metric, e.g. cell area or time-averaged deformation rate. To my knowledge, no such algorithm currently exists and thus represents the main novelty of the manuscript. However, the algorithm depends upon many parameters. In its current form, the manuscript does not adequately quantify and discuss how the resulting partition depends upon variation of these parameters.

We warmly thank the reviewer for insightful and constructive comments.
We answer them point by point, as detailed below. As pointed by the reviewer, data and explanations about the influential parameters were not sufficient in the previous manuscript. So we implemented new simulations, and now it shows how the results were affected by parameters, and clarified our methods efficiency and limitation.
We believe these changes result in a significantly improved manuscript. We hope the revised manuscript now meets the standards for publication.

> Reviewer 3 Comments for the Author:
Major Comments:

> 1. I reiterate my original concern regarding the verification of the algorithm's accuracy against simulated data. I agree that simulating tissue dynamics of interacting regions presents a challenge that is outside the scope of the present work, however, it seems feasible to populate a vertex model with different regions of static geometric parameters, e.g. intrinsic cell area in the usual area elasticity, to see how accurately the boundaries of each region could be reconstructed from a known morphological quantity.Additionally, this would help clarify all subsequent questions below.

We agree that simulations may show how our method would efficiently divide cell populations and when it fails. So we run a simulation of uniaxially compressed tissue. Cells are compressed in a direction, and elongated in the perpendicular direction. They rearrange and relax back towards a more rounded shape. As suggested by the Reviewer 3, there are two regions with different parameters. Rather than having different areas (which is readily seen by the human eye), we propose a more sensitive and challenging test by changing cell-cell junction tension (less detectable by the human eye), which yields a different rearrangement rate. We thus use a cellular Potts model, because it is more efficient than the vertex model to simulate cell rearrangements. Indeed, in our simulations, cell rearrangements are very similar in both regions, and difficult to distinguish by eye, but our method succeeds to distinguish them. Also, dividing into three regions showed an example of over-segmentation.
We add this new result in P7L141, Fig. 5.

> 2. The authors have only partially addressed my original concern regarding the choice of the number of regions partitioned from an input tissue. I found the additions to the manuscript that explored this parameter incomplete.

> (i) The authors state that an increase in the number of regions resulted in subdividing the regions already obtained? while referring to Figure S2. Upon inspection of Fig. S2 it appears this is true starting from 3 (top row) through 9 (bottom row) initial seeds. However, judging from Figures 7 and 9, 6 seeds were chosen for both the notum and 4 seeds for the pupal wing. Why were 6 and 4 chosen specifically? For example, 6 seeds on the notum appear to simply subdivide the regions obtained by 5? For example, the cellular dynamics between regions 2 and 3 of the notum look qualitatively similar, modulo a small 2nd bump in cell divisions.
This also happened to be the partition added in going from 5 seeds to 6 seeds. How do we determine if this region is statistically significant?

The regions number 6 for notum was chosen so that it could be compared with the large grid segmentation. The large grid included not 4 but 6 regions so that it shows 1) some regions may differ in terms of homogeneity (Fig. 8C, G) even though the average cellular processes effective contributions look similar (Fig. S4C, D; regions 2-6).
For wing, segmentations into more than 4 regions included small regions, indicating over-segmentation. The silhouette analysis can show the significance of the obtained regions compared to control segmentations of the same number, but cannot compare segmentations of different numbers of regions.
For now, we can only use the over-segmented small regions to infer if the number of regions is too large or not.
We now mention it in P15L362:
“Also, it is hard to determine the number of regions.
We only know that a tissue might be over-segmented when it includes at least one small region.
In a practical application, the tissue shall be segmented into various number of regions, and one of them can be chosen by comparing it to gene expression patterns, analyzing cell behaviors inside each region, or other characteristics of interest.”

> (ii) The authors state that when the number is too large and a result of the initial label propagation included a too small region, the small region tended to disappear in the cellular Potts model smoothing, and thus the final label propagation resulted in undesired disconnected regions? This feels like an important part that requires more explanation.

Undesired small/disconnected regions for too large number of regions was not an expectation but an observation (Fig. S2). Disappearance of small regions in the cellular Potts model is due to the fact that the area constraint is small. Forcing the label propagation to return 6 regions from a set of similar 5-regions segmentations caused one of the five regions being over-segmented.
However, if the small region is included in a result of the first label propagation, it implies that the regions number is too large, and thus there is no need to smooth it but the tissue should be re-segmented into fewer regions.

> One should define too small? quantitatively for the reader. Does this imply the algorithm has a hard size cutoff, under which we should not expect reliable detection? Or is this a statement about relative sizes of regions, i.e. the algorithm will only detect regions of similar size?

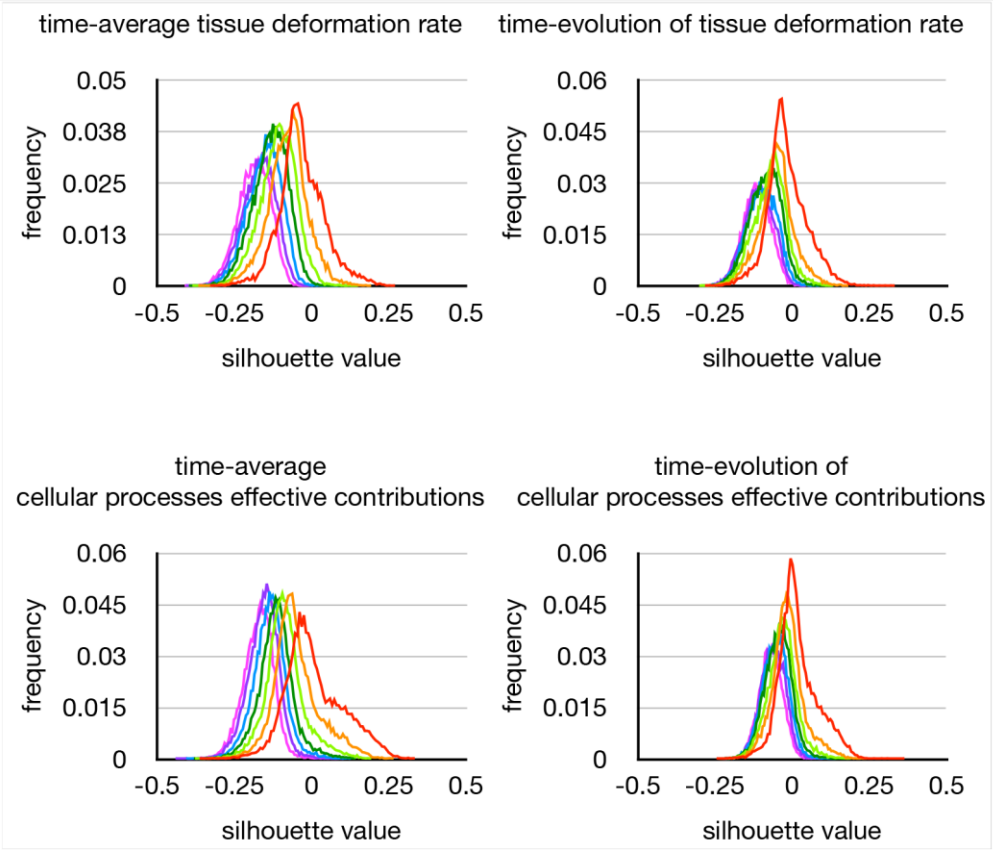
We observed regions smaller than 7 pixels when the tissue was divided into too many regions. However we are not sure if it is the smallest size of such over-segmented regions. Also, those small regions were totally removed during the boundary smoothing and made the following label propagation unstable as described above.
Theoretically, if the small region is enough distinctive and changing its shape increases the energy, then the region will not be absorbed into surrounding regions.
In other tissue, there are groups of few cells with distinctive behavior such as epithelial invagination.
Therefore it is hard to define the size of the over-segmented regions, but it might be assessed based on characteristics of the tissue, regions, and cell behaviors inside it. We now clarified that the small regions were due to the over-segmentation in P10L221: “Those small regions were absorbed into surrounding regions during the smoothing by cellular Potts model. Then the final label propagation tried to integrate regions smaller than the final segmentation, returning small and sometimes disconnected regions (Fig. S3 third column below third row and fourth column below sixth row), again like one of simulated data (Fig. 5G).
Thus the existence of small regions suggest that it is over-segmented.”

> (iii) How does the mean silhouette value depend upon the number of clusters? If you add too many do you begin to overfit Do different morphological features want to cluster under a different number of regions, allowing for an average silhouette value closer to 1? If it is strongly dependent, could one think of devising an information criterion to rigorously choose this number?

The average silhouette value of control segmentations decreased little when the number of regions was increased (histograms in next page). However we could not find any significant dependency on the number of regions to determine the best number.

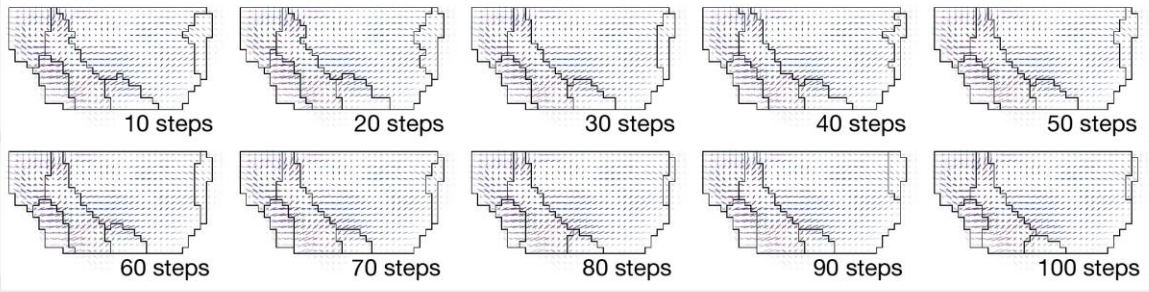
> 3. The manuscript's discussion of the Potts model is incomplete. As presented, the Potts model is run to both smooth the boundaries of the segmented regions obtained from the label propagation and to ensure all regions were simply connected. The manuscript would benefit greatly from an explicit discussion of a few features of this step of the algorithm:

(i) There are 50 metropolis steps, i.e. 50 pixels from the segmentation will be chosen at random to flip labels, for each Potts model run. Why is 50 hard-coded here?
There is no reasoning given by the authors and no citation given that 50 is an adequate number of metropolis steps.
Intuitively, I would have assumed this number should depend extensively on the number of boundary pixels passed into the input, or at the very least the number of pixels in the original image.
Said another way, it would be useful to make this an explicit parameter of the algorithm and show in a supplement how the boundary area depends upon this number (and maybe how the average silhouette degrades).



Histograms of average silhouette value of control segmentations. The dorsal thorax was divided randomly into 3-9 regions and silhouette value was assessed in the four properties. Red, orange, light green, dark green, light blue, dark blue, purple, and magenta lines show the histograms for 3, 4, 5, 6, 7, 8, and 9 regions respectively.

It is correct that the number of steps must correspond to the size of image, number of regions, and length of boundary between the regions. And it would be very informative if we could show how the number of steps is optimized to various image sizes, regions numbers, and boundary lengths. However we have only two tissues.
In the notum with 506 pixels, there was no large change in shape of regions when the number of steps was larger than 30 (figure below).
Also, as Reviewer 2 commented above, the optimal smoothness of the boundary might be determined by a purpose of application, whether to be matched with gene expression pattern or to study mechanical interactions between regions, and it can be controlled not by the number of steps but by the minimum circularity.
We added this new result in P10L211, Fig. S2.



The boundary smoothing by different number of steps. The segmentation of the dorsal thorax based on the time-evolution of cellular processes effective contributions was smoothed for different number of updates.

> (ii) How is the 4d parameter space screened?. Is it an exhaustive search on a fixed grid or is this explicitly optimized? Is it run on the label propagated result or on one instance of the Potts model minimization? How strongly do the results depend on our choice of parameters here? As written the manuscript doesn't give the reader a sense of this.

The parameters were screened in a pairwise testing manner for the cellular Potts model. We now mention it in P22L524:
“The screening was performed in a pairwise testing manner on a grid, and grid was converged to the highest homogeneity.”
Also, Fig. S2 shows that different parameters resulted in different boundary shapes.

> (iii) What is exactly gained by running the label propagation a second time, especially as it can spontaneously create disconnected regions after running the Potts model which is supposed to remove such regions from the first label propagation. Is it run simply because 50 metropolis steps poorly sample the Potts model energy landscape?

Since the cellular Potts model is a stochastic simulation and the image is small, it is hard to assess a convergence. Also it is not assured that a converged result is unique but there might be multiple minima.
If we were to search parameters with which the cellular Potts model reaches a convergence and regions circularity are higher than a threshold and homogeneity are as high as possible, it will be computationally impractical. A temperature tested in the screening might be too high for a convergence.
By limiting the number of steps we could reach practical and satisfying results. And the label propagation gives a unique result P7L131.

> Minor Comments:

> 1. There are many typos and grammatical errors throughout the text that require attention. Examples (non-exhaustive) below:

* L8-9: missing articles, e.g. should be "Coordination between the invaginating mesoderm..."
* L26: "Recent formalisms have enabled [one] to measure..."
* L74: "in histology or of immuno-stained image..."
* L344: "followinig quantificationi"

> 2. I assume that plots of the dynamics of each region's cellular processes display the mean across the identified pixels. It would be useful to plot the standard deviation as an error bar so we can assess the variation across each region.

We understand the point, but the plots included one tissue deformation rate and four cellular processes with similar values, and overlap between their error bars made the panel hard to see. So we omitted it for visibility.

Second decision letter

MS ID#: DEVELOP/2020/199034

MS TITLE: From heterogenous morphogenetic fields to homogeneous regions as a step towards understanding complex tissue dynamics

AUTHORS: Satoshi Yamashita, Boris Guirao, and Francois Graner

Many apologies for the delay in obtaining the referee reports on your manuscript. However I have now received two reports on the above manuscript, and have reached a decision. The referees' comments are appended below, or you can access them online: please go to BenchPress and click on the 'Manuscripts with Decisions' queue in the Author Area.

The reviewers come to somewhat different opinions with one happy to see the study published and the other less enthusiastic. This more negative reviewer still thinks that you should address some of the points raised in his/her initial review. I would like you to consider the points raised by this reviewer but fully addressing them is not a condition for publication of the manuscript in Development. If you do not agree with any of their criticisms or suggestions explain clearly why this is so.

We are aware that you may currently be unable to access the lab to undertake experimental revisions. If it would be helpful, we encourage you to contact us to discuss your revision in greater detail. Please send us a point-by-point response indicating where you are able to address concerns raised (either experimentally or by changes to the text) and where you will not be able to do so within the normal timeframe of a revision. We will then provide further guidance. Please also note that we are happy to extend revision timeframes as necessary.

Reviewer 2

Advance summary and potential significance to field

As mentioned in the first review.

Comments for the author

Yamashita et al have answered previous comments from the first review. As in the first review, I find that the study is well-done, careful, but the authors answers did not answer my comments regarding the relevance and applicability of the segmentation method.
Specifically:
1)the authors state that they do not have an alternative method to compare theirs to; I suggested one based on watershed which the authors did not use. I understand that the method I suggested only works for scalar fields, but since that is also what the authors are using in practice, a comparison could still be done.
2) I suggested to see if the method could offer some control on the degree of homogeneities of the region. I don't understand the authors' reply that " Since the homogeneity was employed as a measurement in the smoothing, it is a variable quantity but not a tunable parameter." I don't see what fundamental reason there would be against imposing a given degree of homogeneity in detected regions.
3) Regarding the biological relevance of the segmentation, the authors have added the position of macrochaetes.
This is helpful but I don't see a clear correspondence with detected regions, notably since these landmarks are single points as opposed to two-dimensional regions. Based on this it seems to me that the denomination of scutum-scutellum boundary is arbitrary and might not correspond to real biological regions.
Overall this is a valid and interesting article, I am just wondering if it is appropriate for the journal "development"
given that the article does not directly demonstrate the method biological relevance.

Reviewer 3

Advance summary and potential significance to field

The manuscript is much improved with the addition of verification against synthetic data. I find the choice of segmenting against cellular rearrangements convincing and bolsters the credibility of methodology significantly.
Furthermore, the added discussion of the algorithm's signatures of over-segmentation is clear and serves as a good description of both the scope and applicability of the algorithm.

Comments for the author

The authors have sufficiently answered my concerns.

Second revision

Author response to reviewers' comments

We thank the reviewers for the thoughtful and relevant comments which helped to improve significantly the manuscript. In below we answer point by point.

> Reviewer 2 Comments for the Author:
Yamashita et al have answered previous comments from the first review. As in the first review, I find that the study is well-done, careful, but the authors answers did not answer my comments regarding the relevance and applicability of the segmentation method.

We warmly thank the referee for this positive opinion and for constructive comments. We answer them point by point, as detailed below. We believe these changes result in a significantly improved manuscript. We hope the revised manuscript now meets the standards for publication.

> Specifically:
1)the authors state that they do not have an alternative method to compare theirs to; I suggested one based on watershed which the authors did not use. I understand that the method I suggested only works for scalar fields, but since that is also what the authors are using in practice, a comparison could still be done.

We tried segmentation by the watershed algorithm with the scalar field of tissue local expansion/contraction. Since a general watershed algorithm leaves 1-pixel width boundary between regions, we prepared a custom script to assign every pixel to one of the regions. To highlight boundary between expanding regions and contracting regions, we tried four operators,

Prewitt operator $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix}$,

Sobel operator $\begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}$,

edge detection kernel with 4 connects $\begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$, and

edge detection kernel with 8 connects $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$.

Gradient calculated by the Prewitt operator and the Sobel operator were then converted to magnitudes of the gradients.
By the watershed algorithm, the fields of highlighted boundary was segmented into many small regions (Fig. 1).

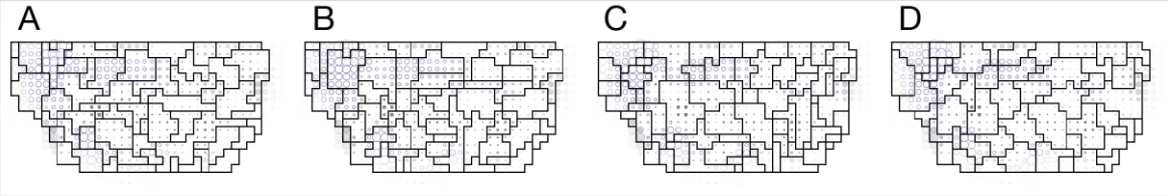


Figure 1. Segmentations by watershed algorithm. The field representing tissue local expansion/contraction was processed by Prewitt operator (A), Sobel operator (B), edge detection kernel with 4 connects (C), or edge detection kernel with 8 connects (D), and then segmented by the watershed algorithm.
Next we processed the fields of highlighted boundary with a disk blur filter or a Gaussian blur filter. It decreased a number of the regions (Fig. 2).

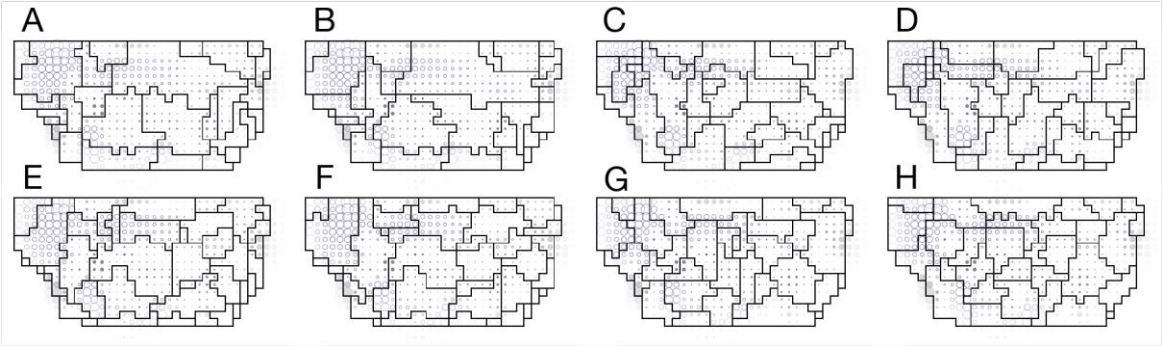


Figure 2. Segmentations of blurred images. (A-D) Processed by disk blur. (E-H) Processed by Gaussian blur. (A, E) Prewitt operator. (B, F) Sobel operator. (C, G) Edge detection kernel with 4 connects. (D, H) Edge detection kernel with 8 connects.

We also tried cutting off lower 30% pixels, and it further decreased the number of the regions (Fig. 3). In some cases it succeeded to enclose a posterior expanding region (Fig. 3E, F) and a lateral posterior expanding region (Fig. 3I, J). However, most of the regions included expanding and contracting points, failed in the segmentation.

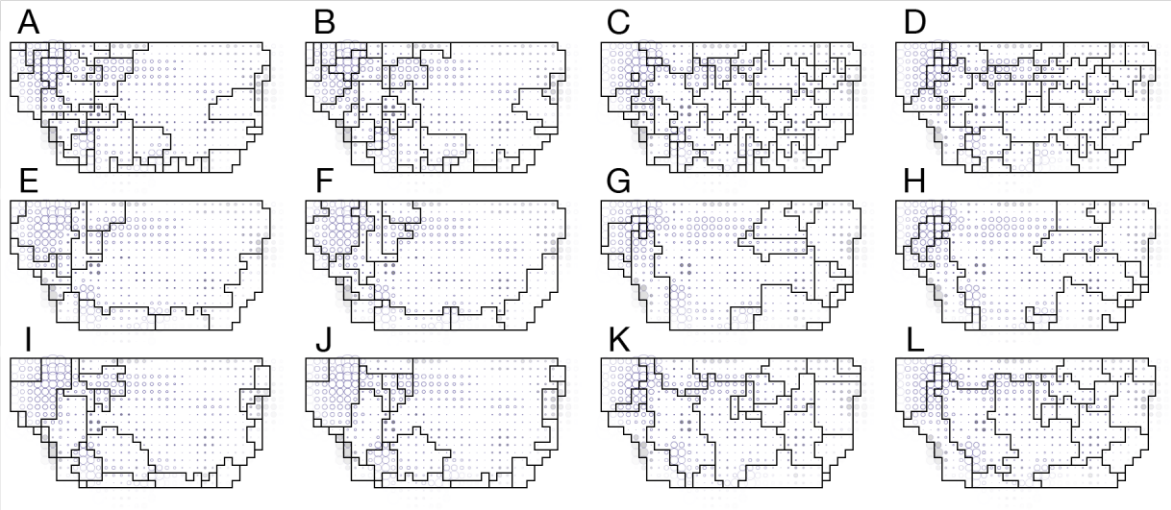


Figure 3. Segmentations with thresholded images. (A-D) No blur. (E-H) Disk blur. (I-L) Gaussian blur. (A, E, I) Prewitt operator. (B, F, J) Sobel operator. (C, G, K) Edge detection kernel with 4 connects. (D, H, L) Edge detection kernel with 8 connects.

These results show that the watershed algorithm is not suitable for the tissue segmentation based on morphogenesis. Since the morphogenesis is quantified in a cluster of cells, a spatial resolution cannot be high and it makes the edge detection hard. Also, the morphogenetic quantity changes in a gradient, further making it hard to detect edges.
We added this in P4L63:
“A watershed algorithm is a widely used tool in general image segmentation and biology such as cell segmentation and segmentation of CT data, but it requires the property space to be scalar and clear boundary between regions.”

> 2) I suggested to see if the method could offer some control on the degree of homogeneities of the region. I don't understand the authors' reply that " Since the homogeneity was employed as a measurement in the smoothing, it is a variable quantity but not a tunable parameter." I don't see what fundamental reason there would be against imposing a given degree of homogeneity in detected regions.

The boundary smoothness and regions homogeneity were balanced by the surface tension *J* and the homogeneity weight coefficient (P22 Eq. 12). When *J* is dominant, the boundary will be smoothed at the expense of the homogeneity, while dominant will keep the homogeneity high and the boundary long and zigzag. We chose *J* and values according to a circularity (P23L526) and the homogeneity of resultant regions. As suggested by the referee, it is possible to control the homogeneity by giving a minimum threshold value so that the resultant homogeneity is higher than the given value and the circularity is as high as possible. Figure 4 shows results of boundary smoothing with parameters chosen by various minimum homogeneities. Since a range of homogeneity is dependent on the field to be divided, we gave the minimum homogeneity of higher 5%, 0.5%, and 0.05% of control segmentations. As expected, when given lower threshold, it smoothed boundary more, while it smoothed less when given higher threshold.

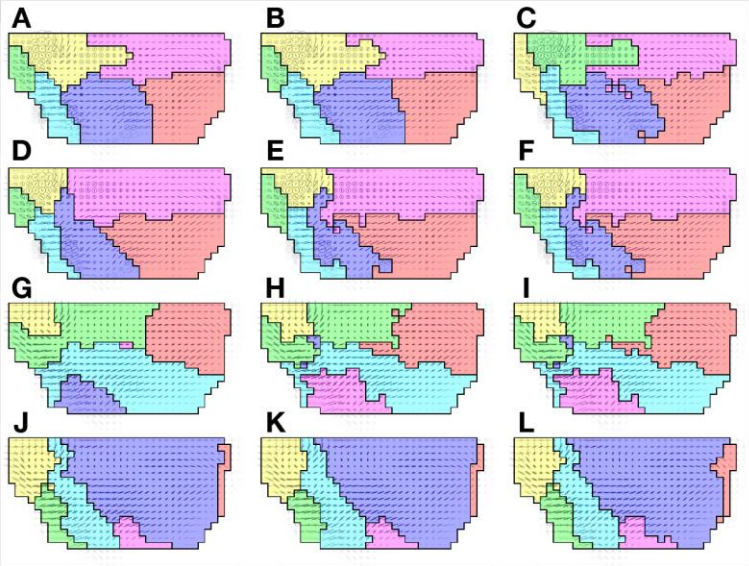


Figure 4. Boundary smoothing with various minimum homogeneity. The Drosophila notum was divided based on time-average tissue deformation rate (A-C), time-evolution of tissue deformation rate (D-F), time-average cellular processes effective contributions (G-I), and time-evolution of cellular processes effective contributions (J-L). They were smoothed with minimum homogeneity of higher 5% (A, D, G, J), 0.5% (B, E, H, K), and 0.05% (C, F, I, L) of control segmentations.

However, we cannot expect how much the boundary will be smoothed by the given threshold values.
Instead, we gave a threshold to the circularity (Fig. S2). It also succeeded in controlling the resultant homogeneity and boundary smoothness. We prefer this thresholding of circularity to the thresholding of homogeneity because it better corresponds to the resultant boundary shape.
We clarified how the parameters were chosen (P7L128):
“Resultant regions were evaluated by the homogeneity and a circularity (Bosveld et al., 2016) which represents a smoothness of the boundary. Parameters for the cellular Potts model were screened so that the circularity was higher than a given value and the homogeneity was as high as possible.”

> 3) Regarding the biological relevance of the segmentation, the authors have added the position of macrochaetes. This is helpful but I don't see a clear correspondence with detected regions, notably since these landmarks are single points as opposed to two- dimensional regions. Based on this it seems to me that the denomination of scutum- scutellum boundary is arbitrary and might not correspond to real biological regions.

We agree that the obtained regions and scutum-scutellum boundary correspondence is hard to confirm. We clarified that the obtained region was a “middle boundary region” and it overlapped with the scutum-scutellum boundary (P10L214, P13L285, and other corresponding sites).

Third decision letter

MS ID#: DEVELOP/2020/199034

MS TITLE: From heterogenous morphogenetic fields to homogeneous regions as a step towards understanding complex tissue dynamics

AUTHORS: Satoshi Yamashita, Boris Guirao, and Francois Graner
ARTICLE TYPE: Techniques and Resources Article

I am happy to tell you that your manuscript has been accepted for publication in Development, pending our standard ethics checks.