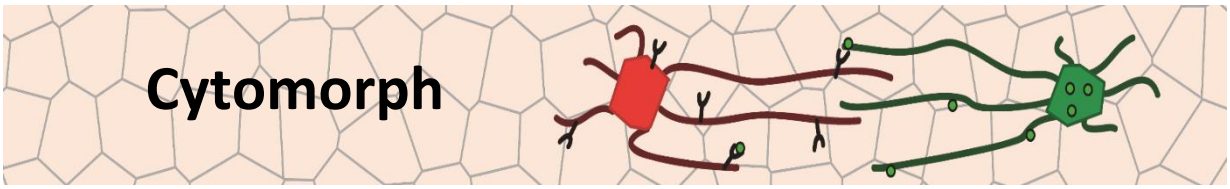


User manual

Introduction

Simulation guide (User instructions)

Results interpretation warning

References

Introduction

0.1 Cytomorph tool

Cytomorph software was developed in 2020 by Adrián Aguirre-Tamaral in the laboratory of Isabel Guerrero at CBMSO (CSIC-UAM).

It is an open source software under a FOSS license 3-clause BSD and implemented in Matlab for simulations of cytoneme-mediated gradient formations. Cytoneme signaling has been reported for signaling of several morphogen gradient formation in different animal developmental systems. Therefore, in order to apply Cytomorph to those systems the software was developed based on a general mathematical model and has been computationally divided in modules that can be updated to incorporate new data; those modules can also be remodeled to simulate specific requirements for the system under study.

0.2 Installation

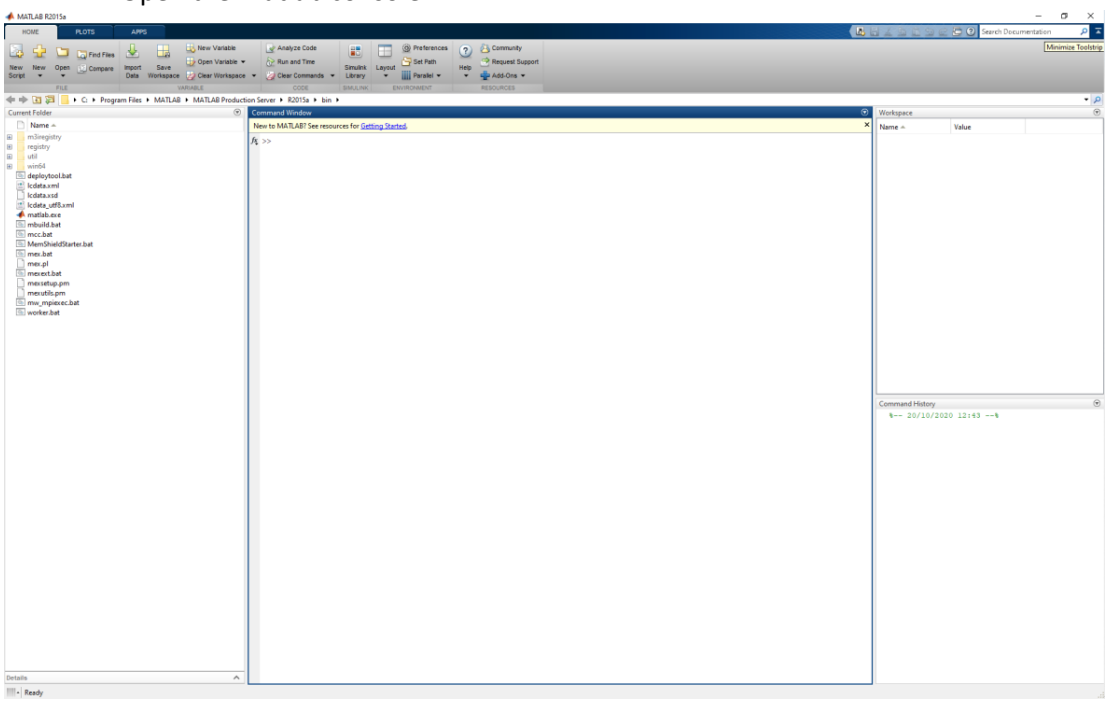
Warning: In order to use of Cytomorph it is first necessary to have Matlab program (Version R2015a or superior) installed in the computer.

1. Download the Cytomorph package from the repository: https://github.com/AdrianA-T/cytomorph

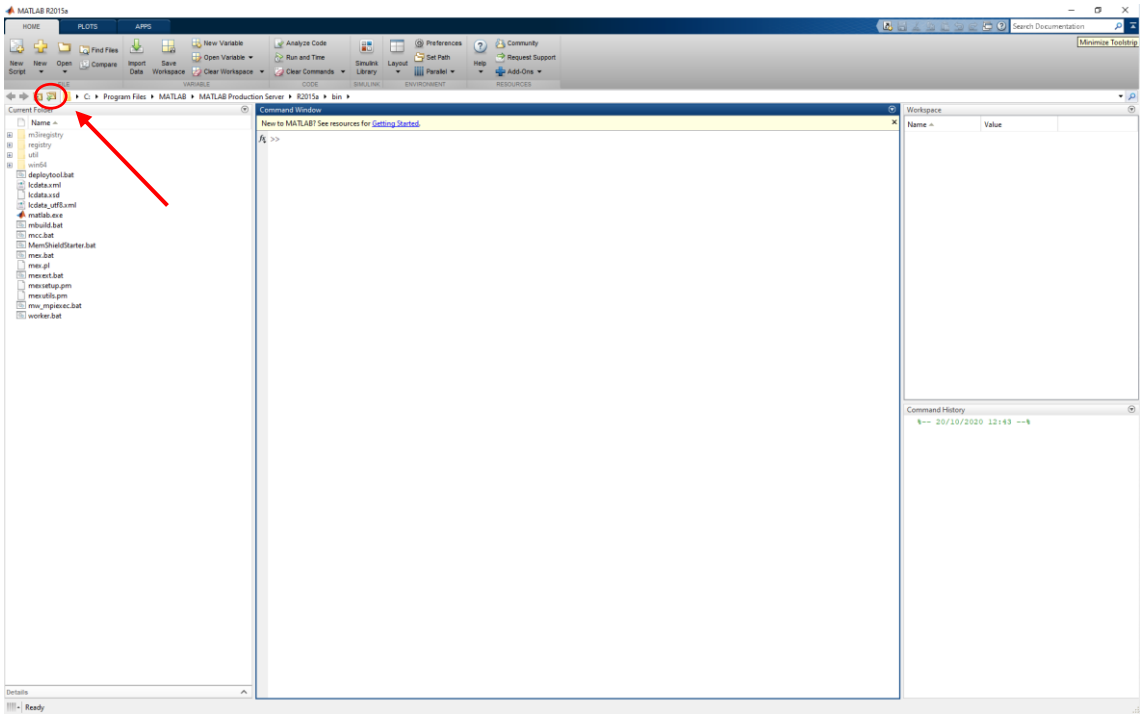
2. Extract the compressed package in the folder where you want to keep Cytomorph.

3. How to open Cytomorph?

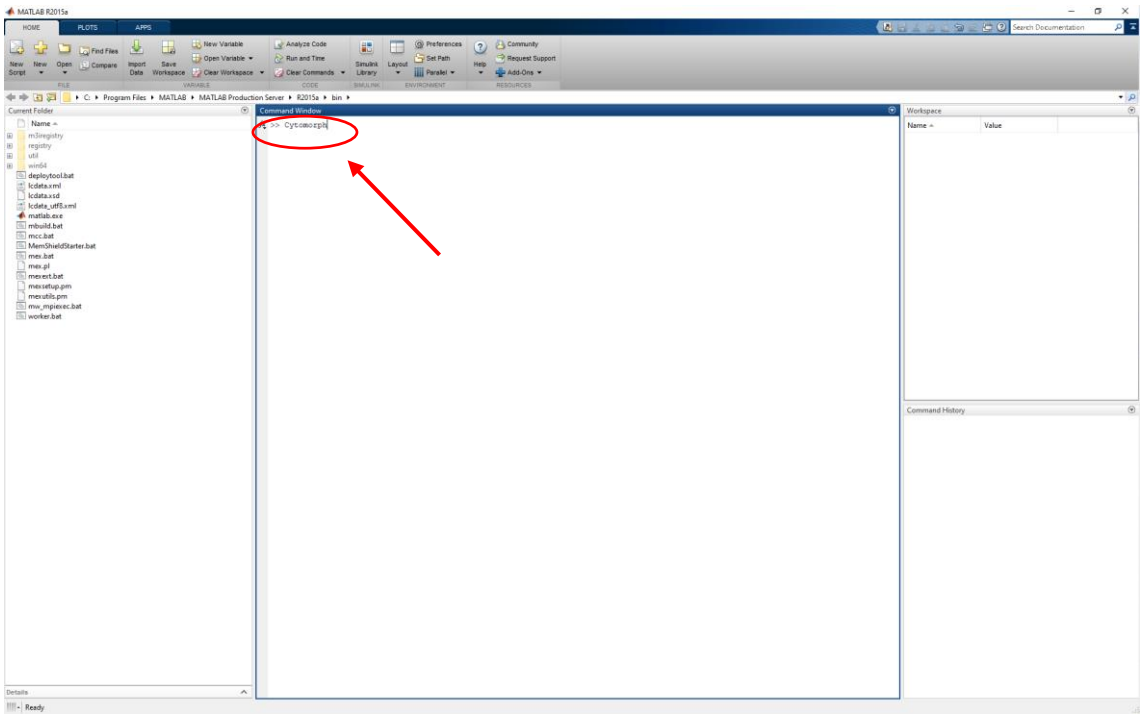
   i. Open the Matlab console



ii. Select the chosen Cytomorph folder path.

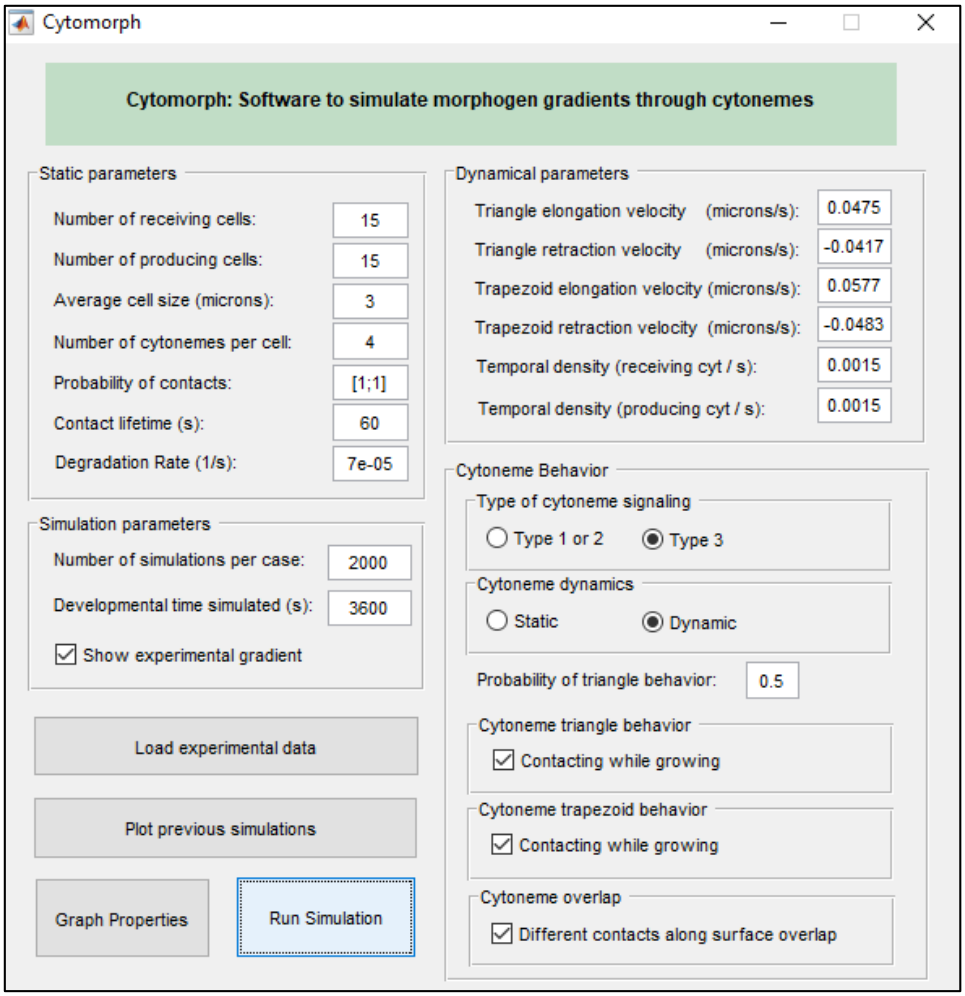


iii. Run the software introducing “Cytomorph” in the command window of Matlab



Simulation guide (User instructions)

If the installation was successful, once the software runs in Matlab, the following window should appear:



This GUI interface is where the initial conditions for simulations should be uploaded.

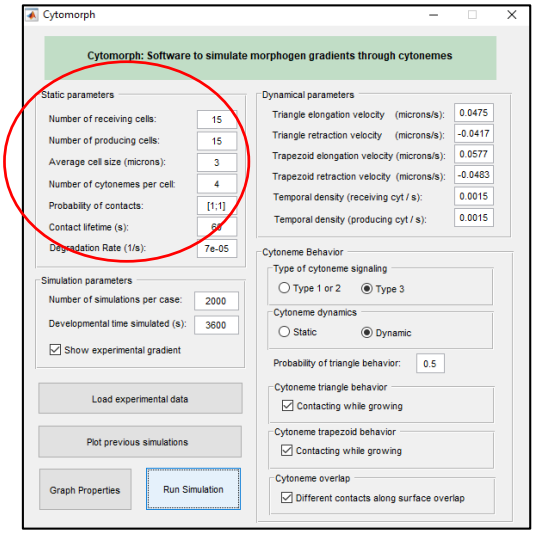
1. Inputs

To run the simulations, the model needs a set-up of experimental values and initial conditions, those inputs are described below:

1.1 GUI parameters

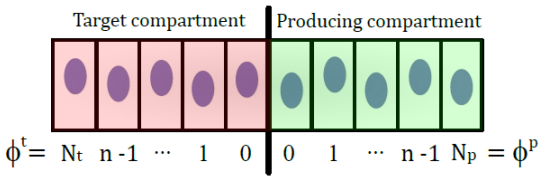
These are the parameters that can be entered in the initial Cytomporh console, in what follows we will see a full description **of each variable.**

1.1.1 Static parameters



Number of receiving cells (N_r):

Since we are working in 1D, the number of receptor cells is represented by the number of receptor rows that we want to model in the software (red cells in the figure).



Number of producing cells (N_p):

Since we are working in 1D, the number of producing cells is the number of producing rows that want to be model in the software (green cells in the figure).

Average cell size:

Cytomorph works internally in term of cell diameters, therefore the average size of cells in the tissue under simulation must be introduced in order to normalized the experimental data measured in microns.

Number of cytonemes per cell:

This is the number of cytoneme per cell that signal at the same time.

Since cytoneme are considered independent signaling units, the software simulates the dynamics of a single cytoneme per cell as many times as the number selected of cytonemes per cell and then uses the addition of all contributions.

Probability of contact:

If cytonemes satisfy the spatial condition of being closer than 1 cell diameter, then the model computes the possibility of contact under the value introduced in this box.

This value can be numerical:

> **[p1; p2]**
>
> Where the p1 is the probability of contact for Type 1 and type 2 signaling and p2 is the probability of contact in type 3 signaling for the region where only cytoneme to cytoneme interaction can occur.

Or no numerical:

> **'ON'**
>
> In this case the probability of contact is a function that depends on the position of the receiving cell.

This last case considers different biological mechanisms that can affect the contact between cytonemes; the implemented default software function is a simple linear case where the probability depends on the equation:

$$p(\phi^t) = \frac{2}{\lambda_t^{max} + \lambda_p^{max})}\left[\frac{-\phi^t}{\lambda_t^{max} + \lambda_p^{max})} + 1\right]$$

Where ϕ^r is the position of the receiving cell and λ_t^{max} and λ_p^{max} are the maximum lengths measured for receiving and producing cytonemes, respectively.

Notes: A different spatial dependence can be implemented if required by the system under study.

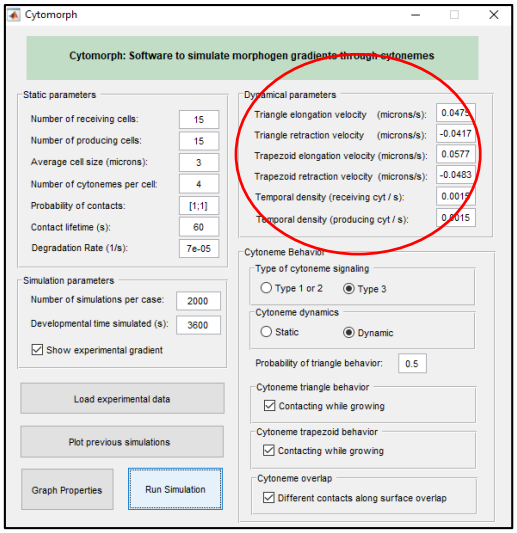
Contact life time:

This value is the average time that a contact can be transmitting morphogen proteins once the contact is stablished.

Degradation time:

It is the degradation rate of the morphogen protein in the receiving cells after reception (expressed in 1/s).

1.1.2 Dynamic parameters



Two different cytoneme dynamics have been described experimentally, therefore both (triangle and trapezoid) are consider for the cytoneme dynamic:

Triangle elongation velocity:

The average elongation velocity of the cytonemes following the triangle dynamics of elongation, expressed in microns/s.

Triangle retraction velocity:

The average retraction velocity of the cytonemes following the triangle dynamics of retraction, expressed in microns/s.

Trapezoid elongation velocity:

The average elongation velocity of the cytonemes following the trapezoid dynamics of elongation, expressed in microns/s.

Trapezoid retraction velocity:

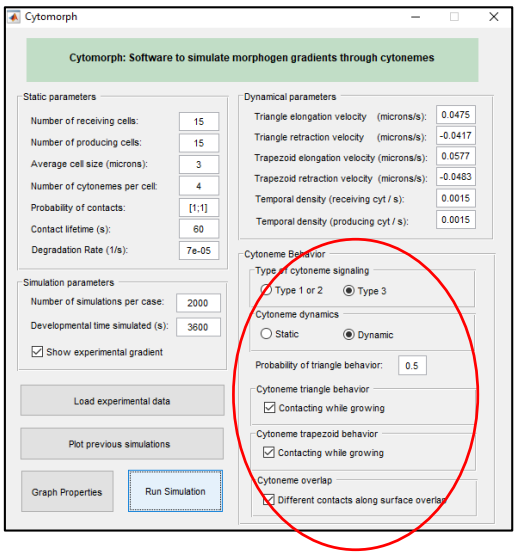
The average retraction velocity of the cytonemes following the trapezoid dynamics of retraction, expressed in microns/s.

Temporal density (receiving) and (producing):

The number of cytonemes that a receiving or producing cell can emit per unit of time. In other words, how often a cell emits a new cytoneme; the units can be calculated in number of cytonemes per second*;* for example, the default value 0.0015 is selected for 1 cytoneme each 11 min since the average life time of cytonemes in Hh singnaling in Drosophila tissues is in this order of magnitude.

- **Warning:** the use of the correct units described in the Cytomorph window is imperative.

1.1.3 Cytoneme Behavior

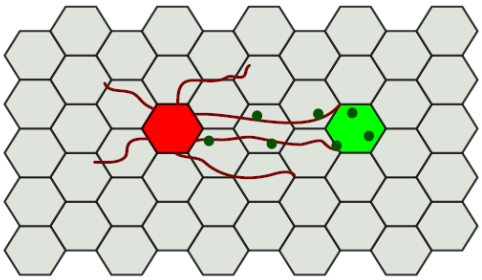


This option refers to *in silico* hypothesized cytoneme features that have not been experimentally described or are based on the system under study.:

Types of cytoneme behaviors:

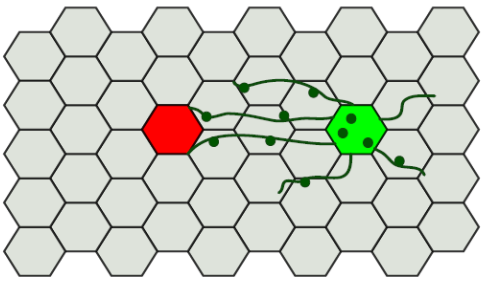
Different types of cell signaling have been described experimentally:

- Type 1:
  Cytonemes from receiving cells (red) to contact signal producing cell bodies (green):



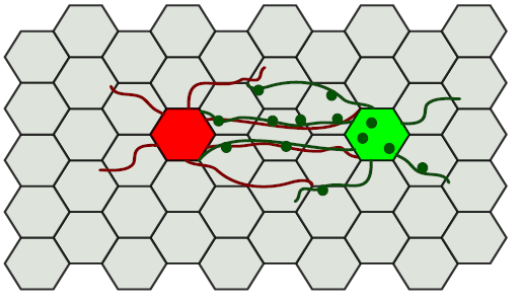
$$N(\phi^t, t) = \sum_{\phi^p=0}^{N_p} \left[C_{tc \to pc}(\phi^{t,p}, t) \right]$$

- Type 2:
  Cytonemes from signal producing cells (green) to contact receiving cell bodies (red):



$$N(\phi^t, t) = \sum_{\phi^p=0}^{N_p} \left[C_{pc \to tc}(\phi^{t,p}, t) \right]$$

- Type 3:
  Cytonemes from both signal producing (green) and receiving (red) cells:



$$N(\phi^t, t) = \sum_{\phi^p=0}^{N_p} \left[C_{tc\to pc}(\phi^{t,p}, t) + C_{pc\to tc}(\phi^{t,p}, t) + C_{cyt\to cyt}(\phi^{t,p}, t) \right]$$

According to the previous schemes, depending on the experimental system one equation will be chosen to simulate the gradient, and then one of the options below must be selected for the **Type of cytoneme signaling:**

a. In type 1 and type 2 cytonemes emanate either from producing or receiving cells only (mathematically both are equivalent and therefore, they appear together in the same option).

b. In type 3 cytonemes protrude from both, receiving and producing cells, contacting to deliver and receive the morphogen.

Cytoneme dynamics:

Experimentally, cytonemes have been reported as dynamic structures in some cases and static in others; both options have been implemented in the model.

Warning: The software is implemented mainly for dynamic cytonemes since in most of the experimental cases cytonemes are dynamic. Therefore, in case of simulating static cytonemes some hypothesis must be considered on how static behavior can affect other temporal parameters like, for example, the life time of contact.

Probability of triangles (p_t):

As mentioned, two types of behaviors have been reported experimentally for cytoneme dynamics; triangular and trapezoid. p_t is the probability to have a triangular behavior in the population of signaling cytonemes. In other words, it is the percentage of cytoneme population that follows the triangular dynamics (consequently, the percentage of cytoneme population with trapezoid dynamics will be $(1-p_t)\cdot 100$)).

Contacting while growing:

The molecular mechanisms underlying contact behavior are still unresolved and therefore, to simulate the gradient some hypotheses have to be proposed. In the current model 2 options are implemented:

The cytonemes can establish contact while growing and in the stationary phase (option selected clicking in the corresponding box) or cytonemes grow but only establish contacts under certain spatial conditions (in the current model for instance, these conditions are only satisfied when cytonemes arrive to a fixed maximum length; this option is simulated if the box is not selected).

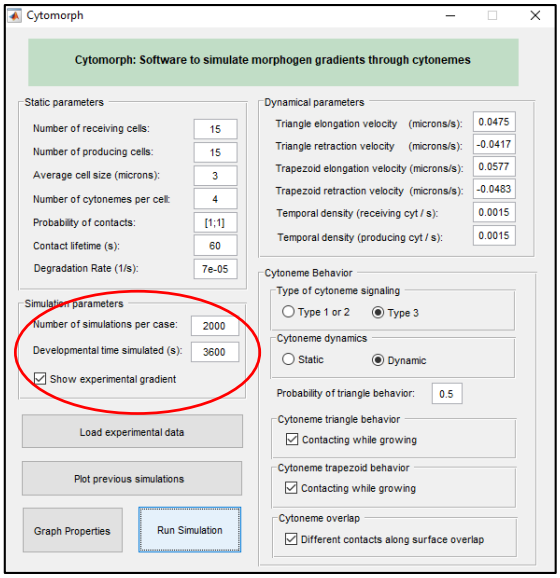
Note: The software accepts other hypothetical cases that could be implemented by the user.

Overlapping cytonemes:

Contacts between overlapping opposite cytonemes have been experimentally reported; thus, in the case of type 3 cytonemes, more than one contact can take place along their overlapping surface.

This option is mathematically implemented in Cytomorph and can be simulated selecting the box named *different contacts along overlapping surface.*

1.1.4 Simulation parameters



Number of simulations:

A subgroup of the experimental distribution data is selected stochastically in each simulation. This stochastic selection in combination with the variables that depend on probabilities requires a large number of simulations per case to obtain a meaningful prediction. This large number is referred as “number of simulations per case” and can be selected in the Cytomorph window.

Developmental time simulated:

Cytomorph allows to simulate the gradient formation along a given time; this time period is selected in seconds in this box.

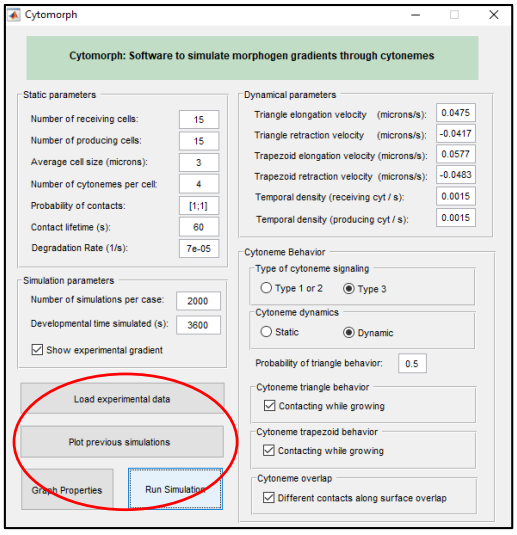
Note: This time begins at time zero when there is no gradient.

Warning: The software does not include cell division, therefore it is not recommended select times longer than the specific cell cycle lifetime of the tissue under study.

Show experimental gradient:

This box is selected by default in order to plot the predicted *in silico* gradient, together with the experimental gradient loaded; allowing comparison between real data and predictions from Cytomorph.

1.1.5 General keys:



Load experimental data:

Here the experimental data distribution is uploaded to the software to run the simulations with those parameters (for a detailed list of parameters in the Excel file see input 2.1.2).

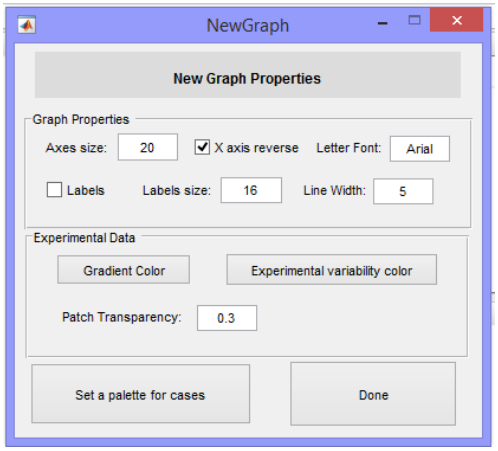
Warning: An Excel file containing those data is required to run Cytomorph.

Plot previous simulations:

Each time that Cytomorph is run, a Matlab file is generated containing the simulated results. These files can be uploaded again into Cytomorph to plot and analyze the same conditions without the computational cost of simulating again.

Graph properties:

This button will open a new window where the graph features of the plots can be chosen. A color palette can also be selected to compare different cases.



Run simulation:

Finally, this button will open a new dialog window to introduce new cases or to run the selected conditions only.

1.2 Experimental distributions

To simulate the biological system correctly, Cytomorph needs some experimental data to scale properly the predictions.

Since most of those parameters do not follow a normal distribution, instead of introducing an average value, a full raw distribution of a direct experimental quantification can be uploaded through an excel file in the order:

A) Whole data inputs

X gradient	Gradient concentration	Standard desviation of the Gradient distribution	Receiving cytoneme length	Producing cytoneme length	Elongation time of receiving triangles	Retraction time of receiving triangles	Elongation time of receiving trapezoids	Stationary time of receiving trapezoids	Retraction time of receiving trapezoids	Elongation time of producing triangles	Retraction time of producing triangles	Elongation time of producing trapezoids	Stationary time of producing trapezoids	Retraction time of producing trapezoids
x1	y1	std1	Lr1	Lp1	tetR1	trtR1	teTR1	ttR1	trTR1	tetP1	trtP1	teTP1	ttP1	trTP1
x2	y2	std2	Lr2	Lp2	tetR2	trtR2	teTR2	ttR2	trTR2	tetP2	trtP2	teTP2	ttP2	trTP2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
Xn	Yn	stdn	Lrn	Lpn	tetRn	trtRn	teTRn	ttRn	trTRn	tetPn	trtPn	teTPn	ttPn	trTPn

B) Zoom of the data

X gradient	Gradient concentration	Standard desviation of the Gradient distribution	Receiving cytoneme length	Producing cytoneme length	El
x1	y1	std1	Lr1	Lp1	
x2	y2	std2	Lr2	Lp2	
...	...	...	...	...	
...	...	...	...	...	
Xn	Yn	stdn	Lrn	Lpn	

th	Elongation time of receiving triangles	Retraction time of receiving triangles	Elongation time of receiving trapezoids	Stationary time of receiving trapezoids	Retraction time of receiving trapezoids	Elongation time of producing triangles	R
	tetR1	trtR1	teTR1	ttR1	trTR1	tetP1	
	tetR2	trtR2	teTR2	ttR2	trTR2	tetP2	
	...	...	...	...	...	...	
	...	...	...	...	...	...	
	tetRn	trtRn	teTRn	ttRn	trTRn	tetPn	

n time ucing les	Retraction time of producing triangles	Elongation time of producing trapezoids	Stationary time of producing trapezoids	Retraction time of producing trapezoids
1	trtP1	teTP1	ttP1	trTP1
2	trtP2	teTP2	ttP2	trTP2
	...	...	...	...
	...	...	...	...
n	trtPn	teTPn	ttPn	trTPn

Warnings and notes:

- The experimental gradient (first tree data columns) is just uploaded to compare it with the simulated gradient. If there are not experimental data for the gradient, write a zero in each column and select the option not to compare with the simulated gradient in the Cytomorph options window.
- The length of each column does not need to be the same.

- The model randomly selects values from the distribution to compute the correct distribution, but if for some of data there is not experimental distribution a single value can be updated. In this case, the software uses it as an average value for this parameter.
- It is important to introduce the correct units.

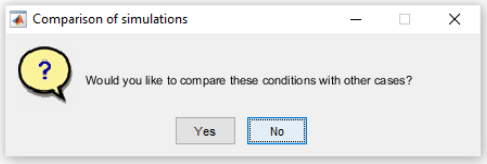
Units

X gradient	Gradient concentration	Standard desviation of the Gradient distribution	Receiving cytoneme length	Producing cytoneme length	E
Cell diameters	Dimensionless (normalized to maximum value of the gradient)	Dimensionless	Microns (μm)	Microns (μm)	...

th	Elongation time of receiving triangles	Retraction time of receiving triangles	Elongation time of receiving trapezoids	Stationary time of receiving trapezoids	Retraction time of receiving trapezoids	Elongation time of producing triangles	Retraction time of producing triangles	Elongation time of producing trapezoids	Stationary time of producing trapezoids	Retraction time of producing trapezoids
	seconds (s)	seconds (s)	seconds (s)	seconds (s)	seconds (s)	seconds (s)	seconds (s)	seconds (s)	seconds (s)	seconds (s)

2. *In silico* settings

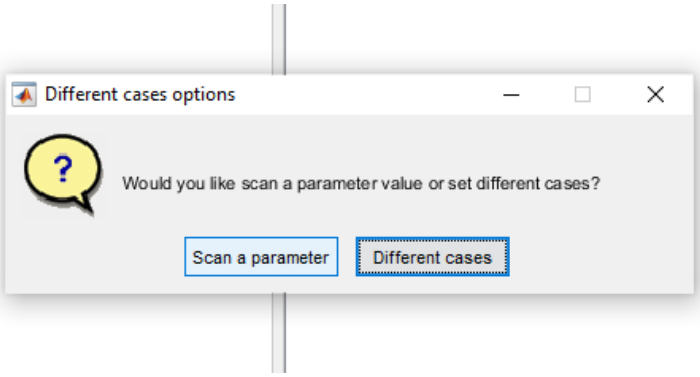
As mentioned, after introducing the previous inputs and clicking in running simulation, a message window will appear asking if you want to introduce other cases to compare predictions.



If "No" is selected, then Cytomorph simulates just one set of conditions and plots outputs in order to see the *in silico* gradient and the properties predicted for that set of conditions.

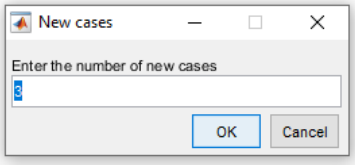
If "yes" is selected, then different sets of conditions are simulated and plotted in different outputs to study and compare cases.

For the user convenience, there are two ways to introduce new conditions: Scan variable and new cases. The software will ask in a dialog window which one to use to compare cases:



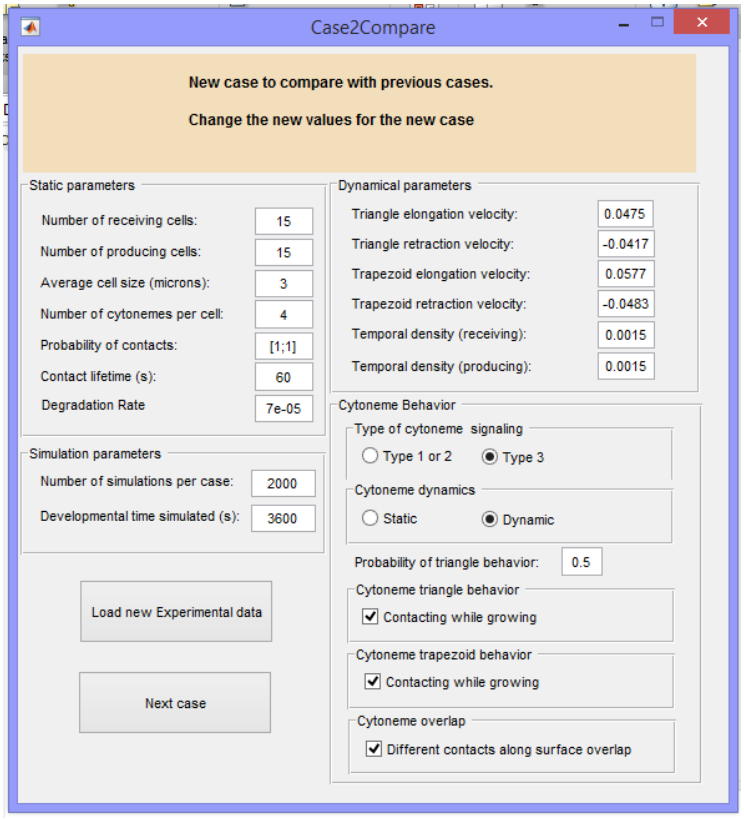
New cases window is designed to help the comparison of cases in which different conditions or combination of conditions are wanted to be tested.

If this option is selected, a window will appear asking how many new cases you want to compare:



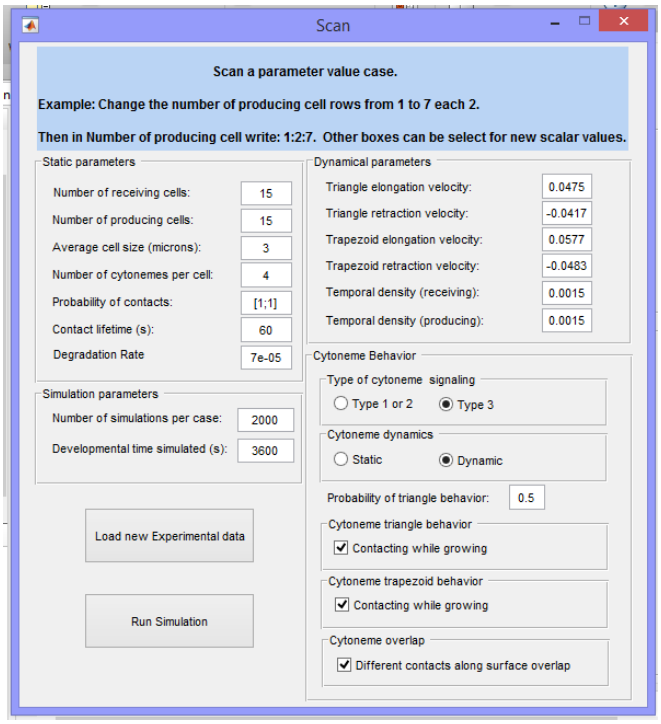
Note: (The number does not include the reference case) This number is just the new cases that you want to compare with the reference case.

The next window will appear as many cases you had selected to introduce all the new case conditions:



Each time you must to select the new conditions for each case and the click button to introduce the next case.

Scan variable window is designed to help the comparison of cases where only a single numerical value is wanted to be tested.



The variable to be scanned must be introduced in the scan window with the range and the coefficient of variation under study, writing that variable as:

minimum value: coefficient of variation: maximum value

To analyze the impact of how the number of producing cell rows could affect the gradient, a scan can be done for that variable to look at the changes of outputs (gradient properties); for example, having **1** cell row (**minimum value**), to **7** cell rows (**maximum value**) and comparing every **2** cell rows (**coefficient of variation**).

In this case, in the white box of producing cells ***1:2:7*** has to be written.

In other words, we are comparing 4 different cases where all the inputs are the same, except the number of producing cells involved, but instead of manually introducing 4 times the conditions but changing N_p=1, N_p=3, N_p=5, N_p=7, we use scan variable window that simplify the process just writing N_p=*1:2:7*

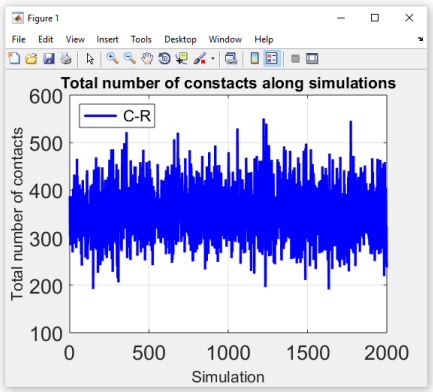
3. Outputs

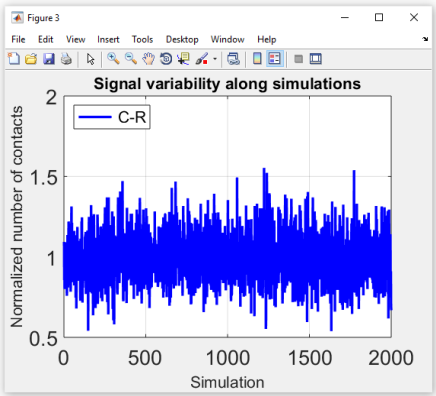
3.1 Single simulation

3.1.1 Contact distribution

The software computes the number of contacts for each cell diameter along the simulated time. This is repeated over a wide range of simulations per case and the resulting contact distribution array is the base for computing the rest of the parameters.

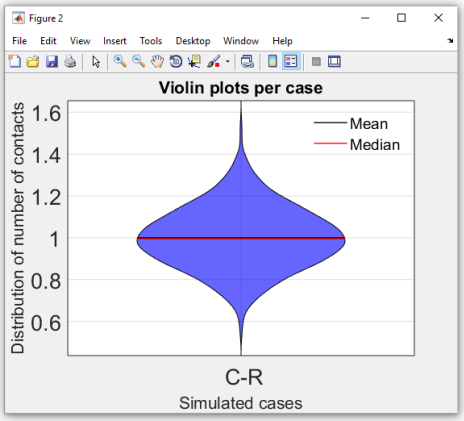
For example, using published experimental data, we can simulate the number of contacts for a specific receiving cell (ϕ^r) and look at the predicted value per simulation:





Those date are graphically plotted in raw numbers (left) and normalized to the average value (right).

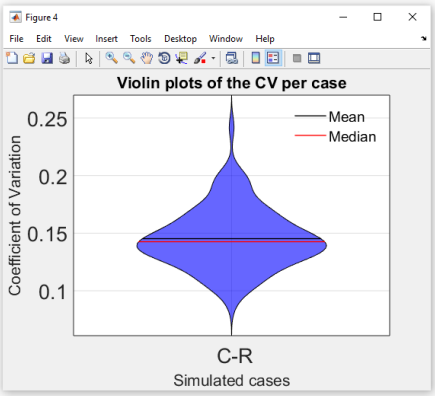
Since in each simulation the software collects a random subgroup from the experimental data, in order to better interpret these data, the previous graphs can be visualized using violin plots representing contact distributions per simulation:



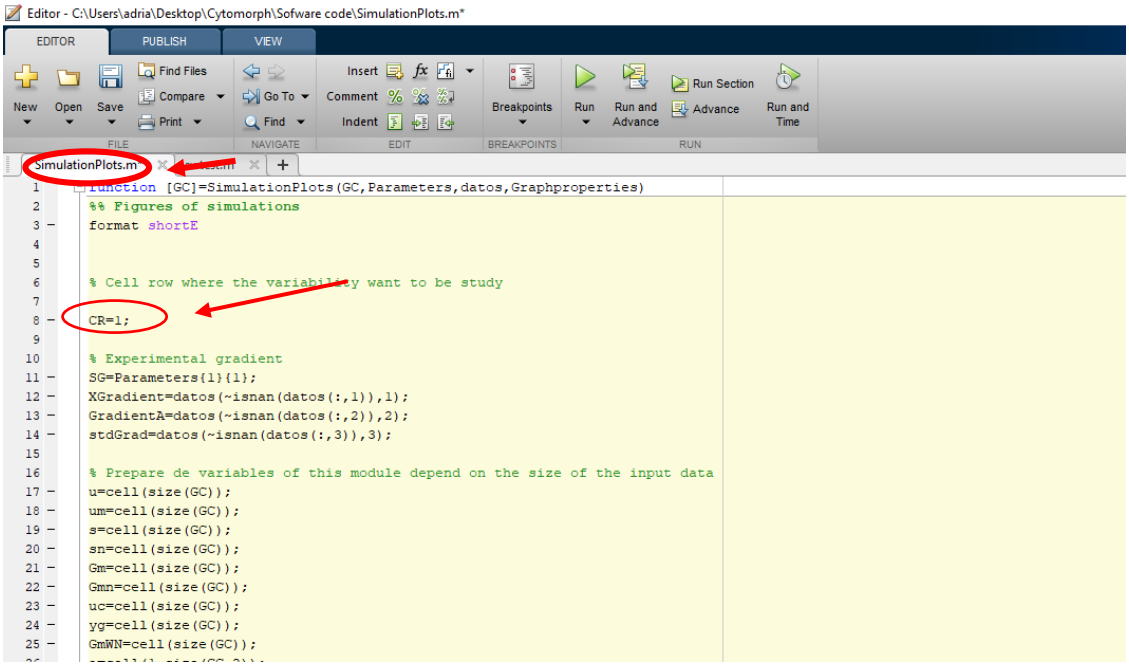
3.1.2 Signal variability

The predicted values for simulation variability can be determined from the different simulations. To better study this parameter and also have the opportunity to statistically compare different cases, we computed the distribution of coefficient of variations per case. To do that, we divided the N simulations in 100 subgroups of N/100 samples each. Then the coefficient of variation distribution per case was performed over those 100 subgroups.

The resulting data are presented as violin plots:

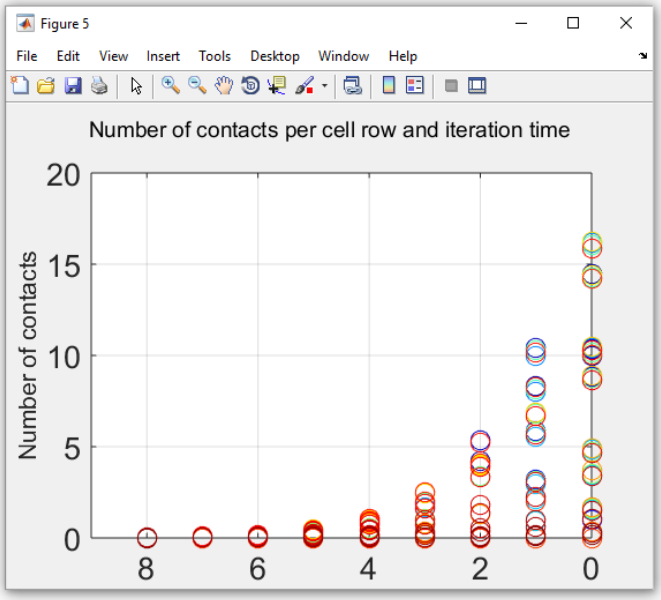


Note: The previous plots showed the data for a specific receiving cell, by default the software plots the first cell row (CR=1). This can be selected changing CR to other number in the SimulationPlots script:

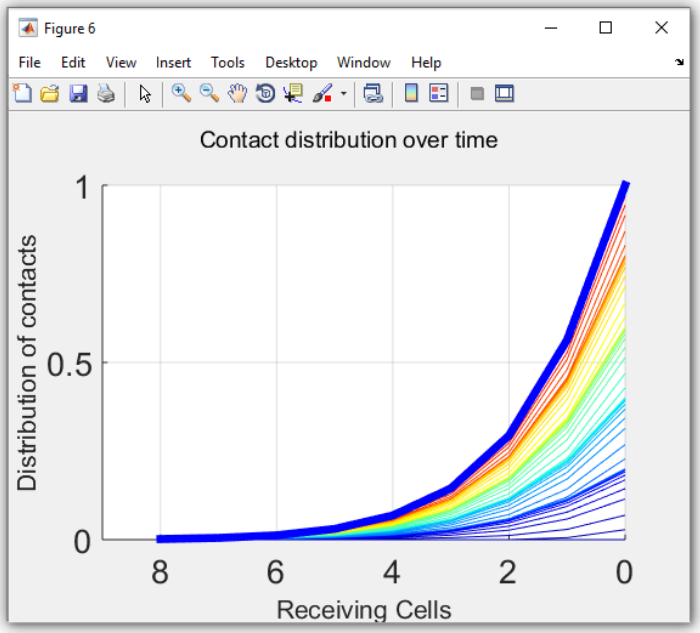


3.1.3 Temporal evolution

We can also observe the number of contacts in each receiving cell per time lapse:



Or the total evolution of the contact distribution and gradient shape over the simulated signaling time:



3.1.4 Gradient distribution

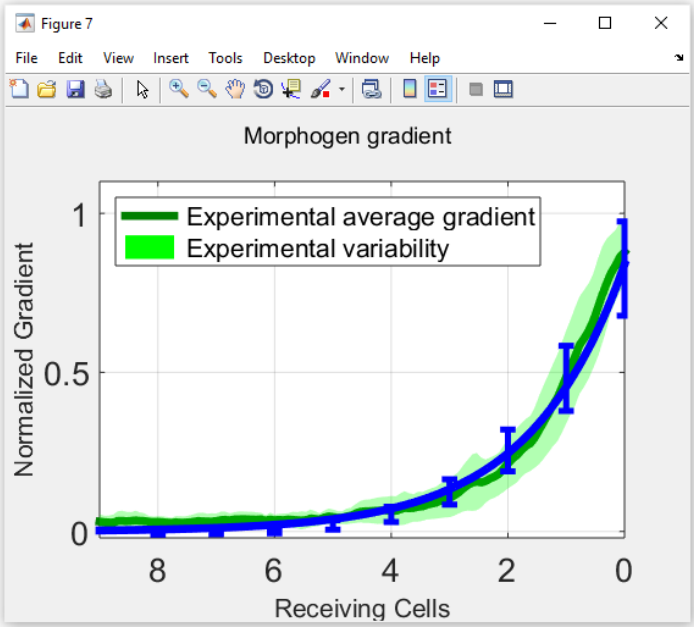
Assuming that each contact transmits a specific amount of morphogen, the distribution of the morphogen can be estimated for each interaction by the following proportional relation:

$$[protein] = D \cdot \#contact - \delta \cdot [protein]$$

Where δ is the degradation rate of the morphogen. The software normalizes and computes in each iteration the equation Subsequently, for a validation of the model, the software plots the simulation versus the experimental data of the morphogen gradient, visualizing the accuracy and the correlation of the *in silico* predictions with the experimental quantification of the morphogen gradient.

As mentioned before the software gives a numerical estimation for the number of contacts per receiving cell diameter along time. The plotted gradient is a normalized gradient calculated from the previous equation.

This gradient is an exponential fit of the numerical simulations and the error bars are the standard deviation of the simulations performed per case:



Note: There may be a small shift between the fit curve and the error bars, as the error bars are plotted at the position of the numerical data and not on the final fit.

3.2 Comparison between cases

It is interesting to study in silico different conditions and compare how they alter the gradient features. To facilitate this type of studies a set of plotting graphs was implemented in Cytomorph.

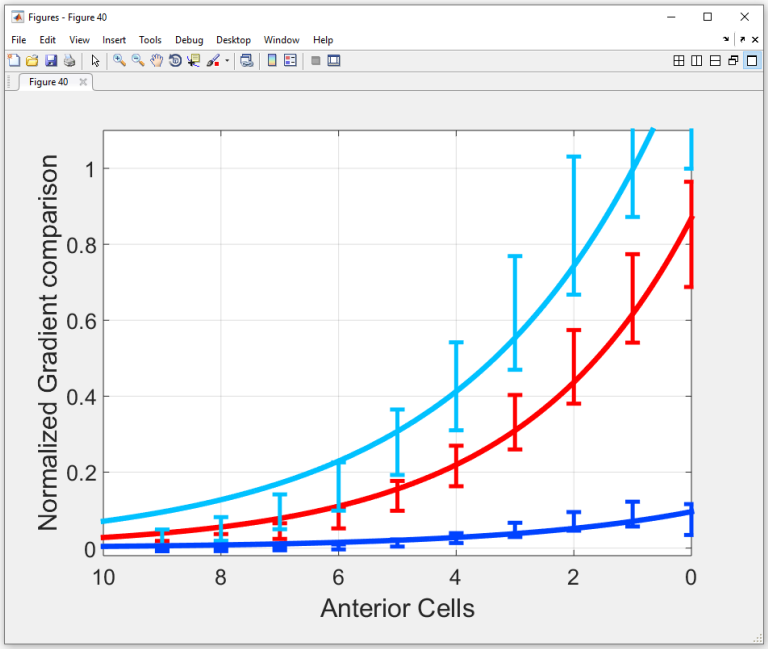
Note: Cytomorph predictions are normalized and, therefore, to compare the relative changes between different situations, the initial conditions (introduced in the initial window) are considered as the reference case.

3.2.1 Gradients distributions

Different sets of conditions might significantly change the number of contacts, and therefore. the gradient. Those changes can be visualized in the implemented plots below:

- **Final gradient**

In order to give a general view of how the gradient was affected, the different predicted gradients have been plotted together:

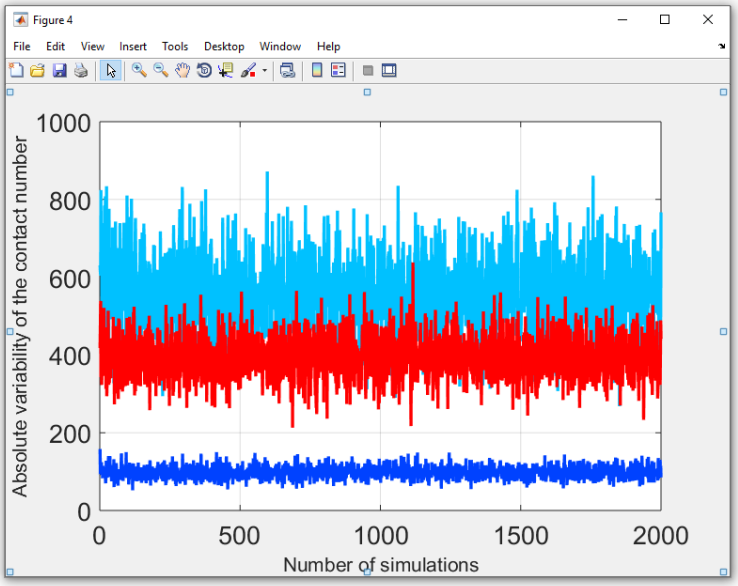


Note: Normalization equation

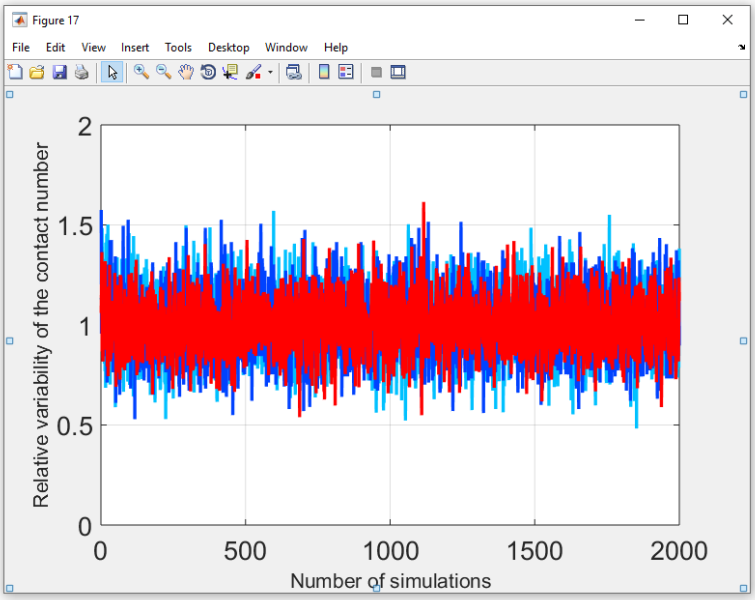
$$N_{c_i}(\phi)/max\left(N_{Ref}(\phi)\right)$$

- **Number of contacts and signal variability**

To compare those changes mathematically, Cytomorph plots the number of contacts per simulation:



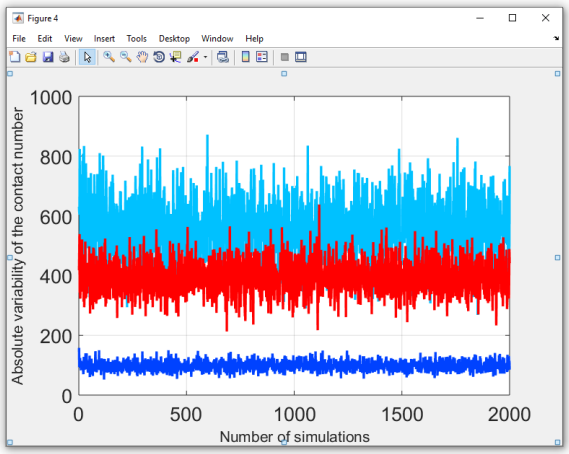
The normalized number of contact per simulation:

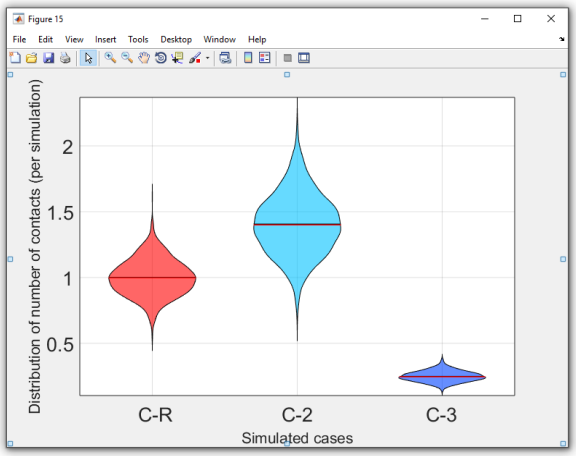


Note: Normalization equation

$$N_{s,c_i}(\phi^0)/\langle N_{Ref}(\phi^0)\rangle_s$$

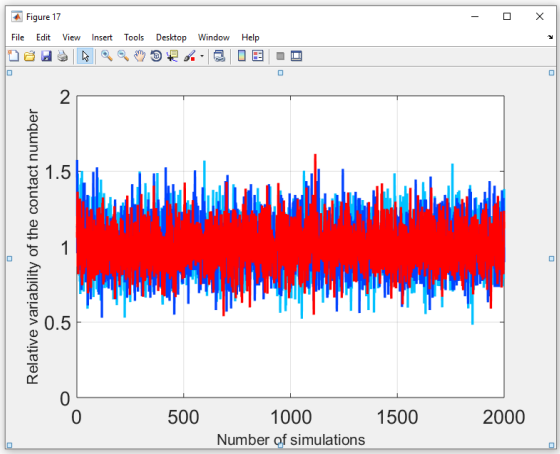
And the comparison between previous distributions in violin plots:





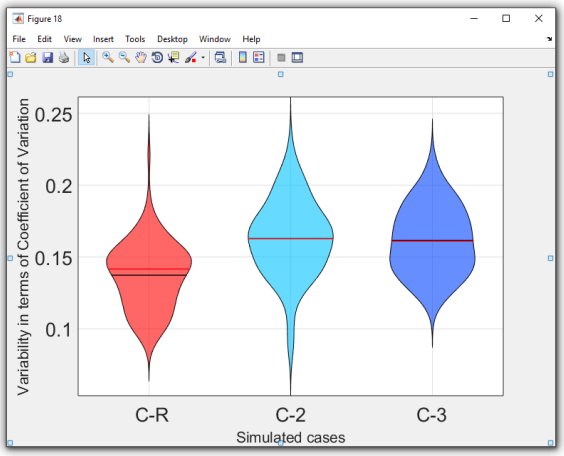
Note: Normalization equation

$$N_{s,c_i}(\phi^0)/\langle N_{Ref}(\phi^0)\rangle_s$$



Note: Normalization equation

$$N_{s,c_i}(\phi^0)/\langle N_{Ref}(\phi^0)\rangle_s$$



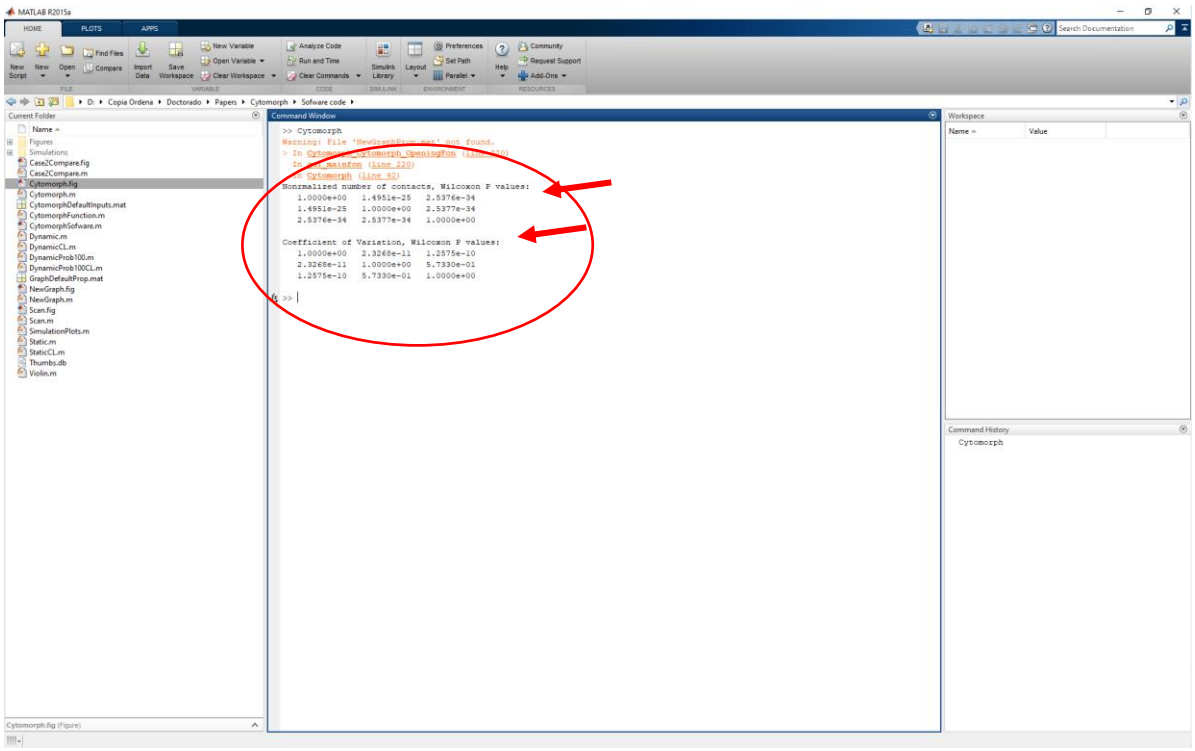
Note: Normalization equation

$$C.V = std\left(N_{s^*,c_i}(\phi^0)\right)/mean\left(N_{s^*,c_i}(\phi^0)\right)$$

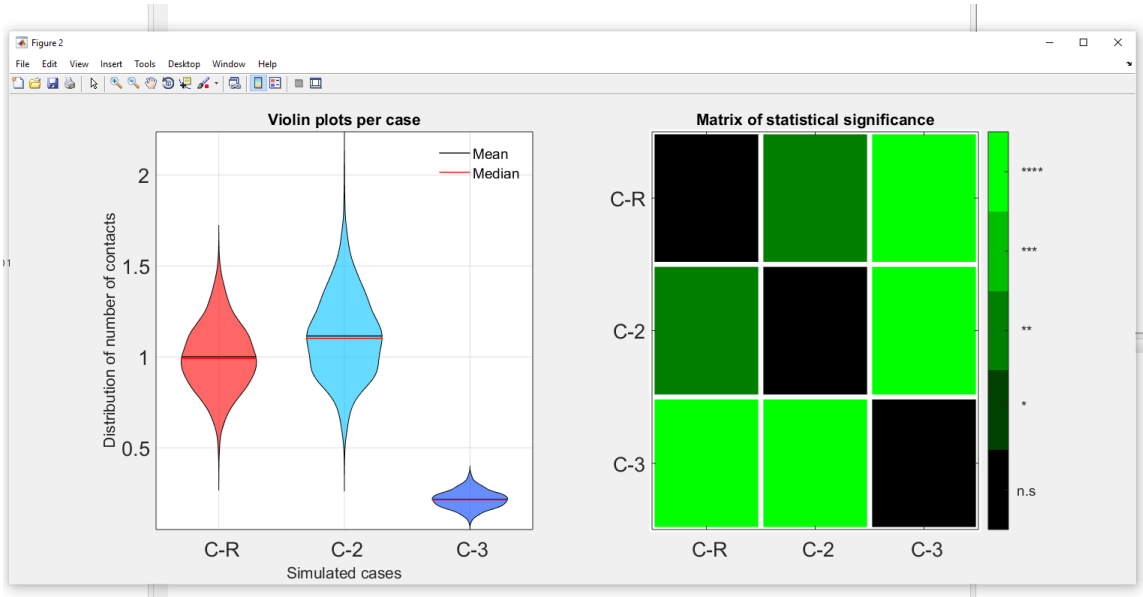
To have a better idea of previous changes, a statistical analysis was also implemented comparing the cases in pairs.

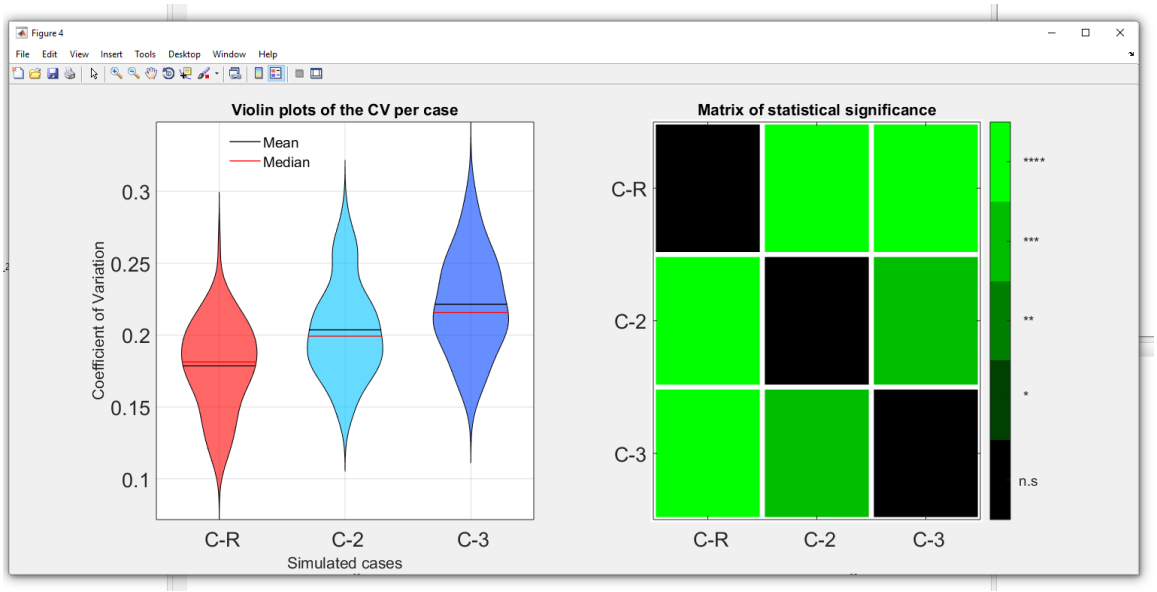
The normality of the data is tested by a Shapiro-Wilk test, if all the data sets follow a normal distributed population then the statistical analysis is performed using a T-test (Matlab function ttest2). If some of the data sets do not follow a normal distribution, then the software studies the statistical analysis using a Wilcoxon test used (Matlab function ranksum).

The statistical test used and the p-values are plotted in the command windows of Matlab:



For a quick visualization per case, the resulting p-values are also color-coded in a matrix

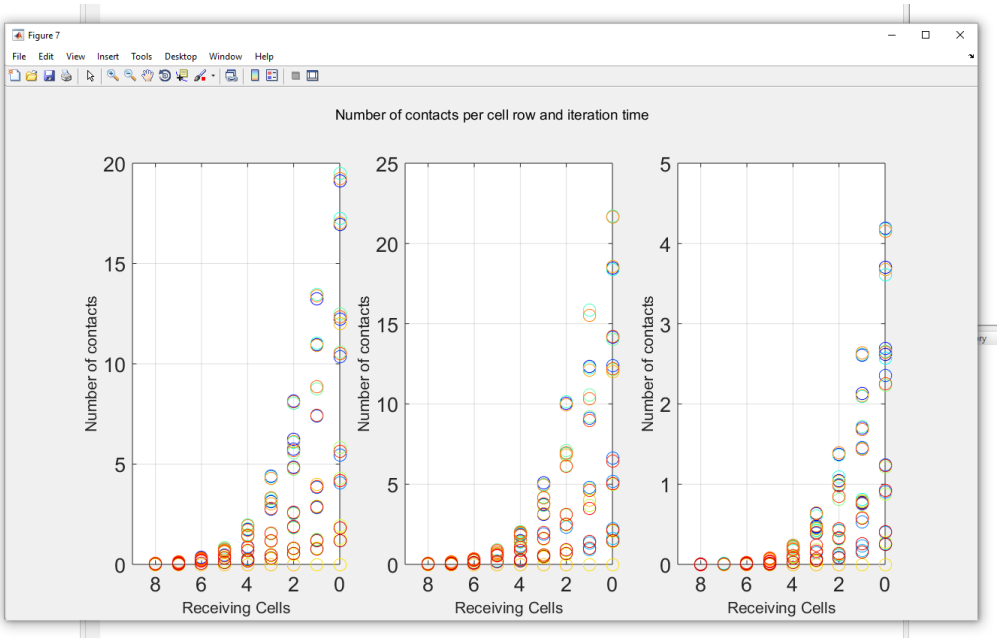




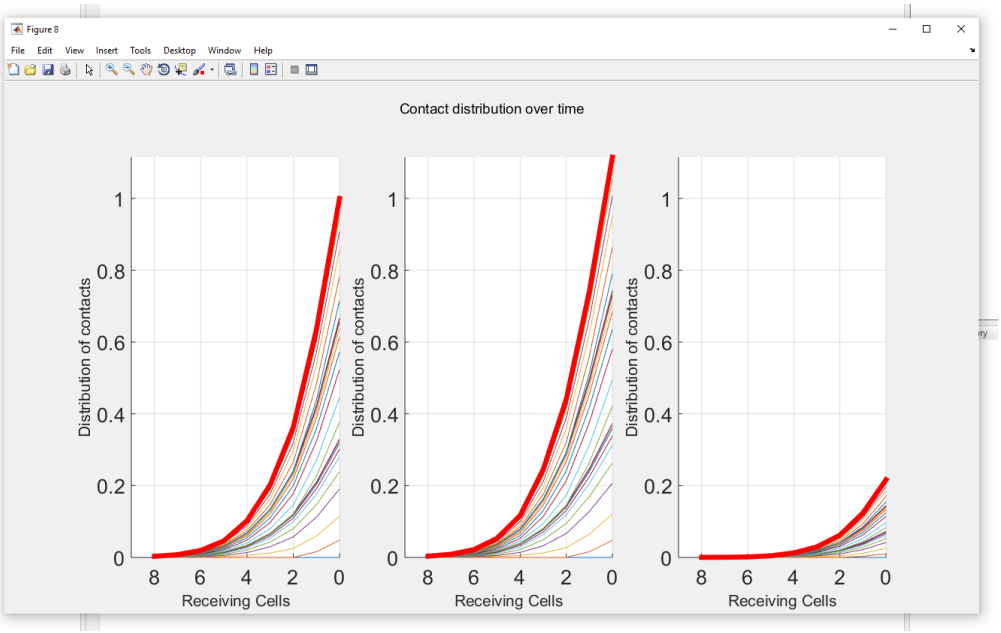
Where black is no significance (= n.s) and green means significance under the code: dark to light green for the p-values: * = p-value < 0.05, ** = p-value < 0.01, *** = p-value < 0.001, **** = p-value < 0.0001 respectively.

3.2.2 Gradient scaling

The gradient formation depends on the distribution of the contacts along the receiving cells; those values can be visually compared in Cytomorph:



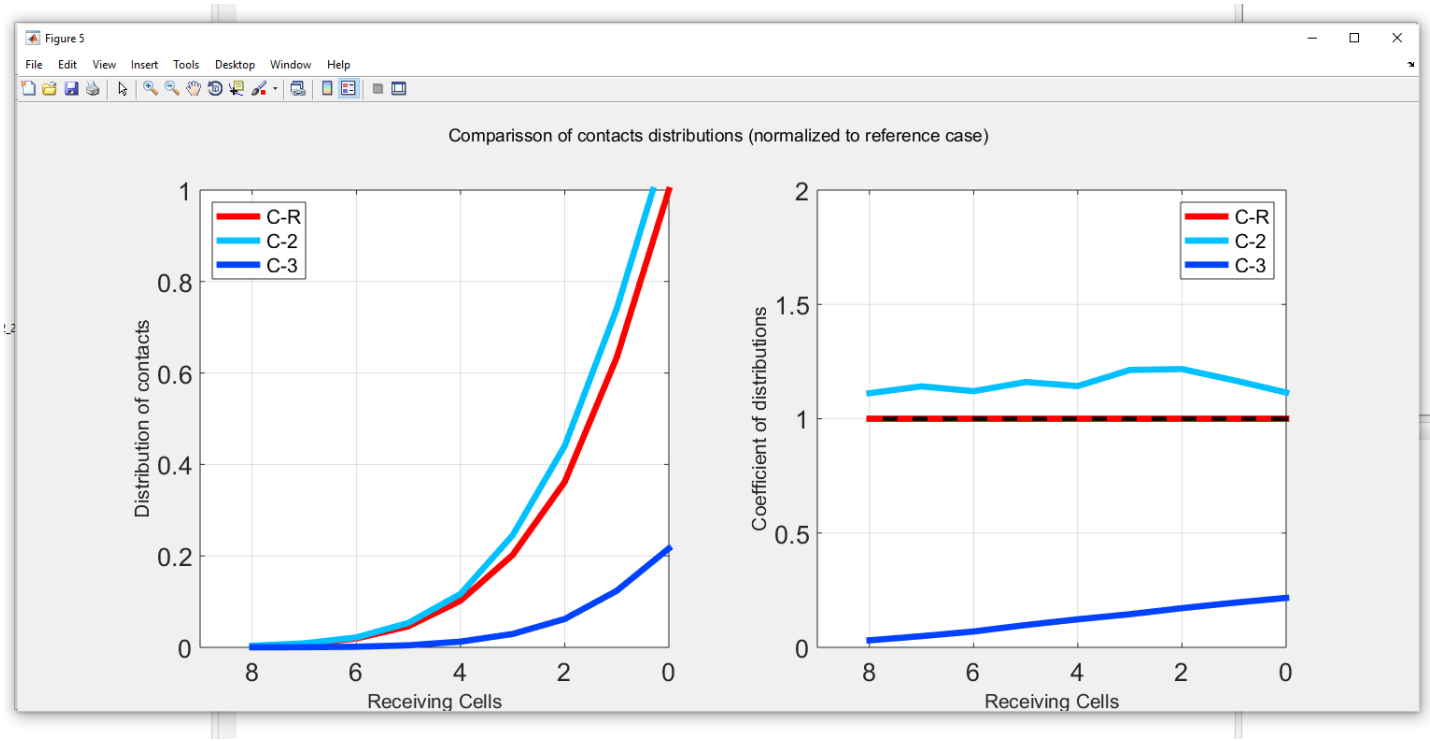
Together with the temporal evolution of the different cases:



In order to deeper analyze how the gradient is changing the next plots were implemented to visualize different aspect of the scaling between cases:

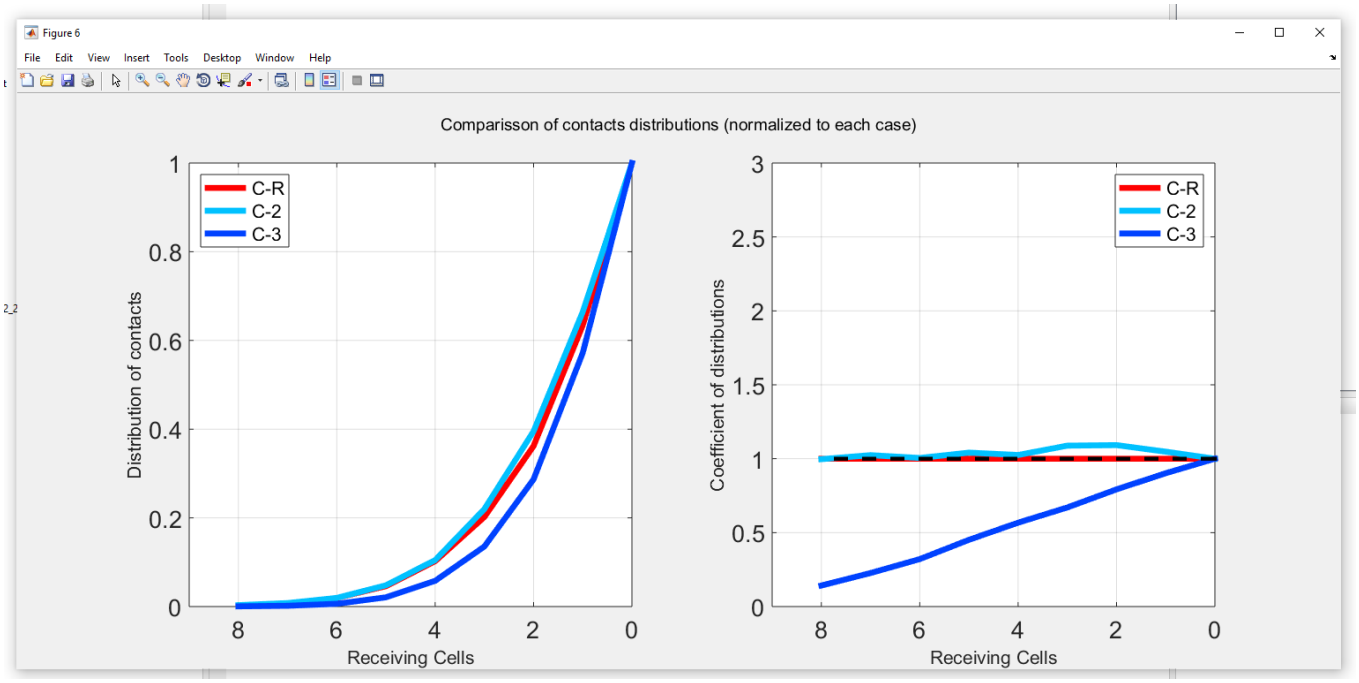
Exponential shapes and coefficient of change along receiving cells

In order to better see the changes, the different cases, the distribution of the contacts are plot together (left). The coefficient of the distribution of contact by the reference distribution is also plotted to better study the previous change in the shape distribution (right).



In order to see the distribution changes independently of the amount of morphogen, the distributions were normalized to their respective maximum (left). In this way we can see the changes in the exponential shape; in other words, the change between cases in the distribution of the contacts in receiving cells.

To better study the previous change in the shape distribution, the different cases where divided by the reference case (right) This plot gives the ratio between contact distributions; that is, what areas (along receiving cells) have bigger changes relative to the reference case and if those changes are increasing or decreasing the contacts.



Note: Normalization equation

$$N_{c_i}^N(\phi) = N_{c_i}(\phi)/max\left(N_{c_i}(\phi)\right)$$

Note: Normalization equation

$$N_{c_i}^N(\phi)/N_{Ref}^N(\phi)$$

Results interpretation warning

As a general warning, since some aspect of experimental mechanism of cytoneme signaling are still unresolved, all the data plotted are simulated using some working hypotheses. For that reason, most of the values and graphs are normalized and/or presented in relation to a reference case.

We strongly recommend not to assume the absolute values as unequivocal predictions since these magnitudes make sense better in comparisons. For example, instead of looking in a simulation at the exact number of contacts per cell row, what is more reliable is to see how this number compares with other cell rows in the same simulation or with the same cell row in other simulations.

References:

1. The experimental data to simulate the example plots were obtained from:
   González-Méndez, L., Seijo-Barandiarán, I. & Guerrero, I. Cytoneme-mediated cell-cell contacts for hedgehog reception. *Elife* **6**, (2017).

2. The Violin plots are not implemented in the base code of Matlab so in the software we externalized this plots using the script developed by Holger Hoffmann and available in MatlabWorks:

   Hoffmann H, 2015: violin.m - Simple violin plot using matlab default kernel density estimation. INRES (University of Bonn), Katzenburgweg 5, 53115 Germany.

   https://es.mathworks.com/matlabcentral/fileexchange/45134-violin-plot

3. Shapiro-Wilk test is also not implemented in Matlab code, for the statistical study of normality we used the scrip developed by Ahmed Ben Saïda and available in MatlabWorks:

   Copyright (c) 17 March 2009 by Ahmed Ben Saïda,

   Department of Finance, IHEC Sousse - Tunisia

   https://es.mathworks.com/matlabcentral/fileexchange/13964-shapiro-wilk-and-shapiro-francia-normality-tests