V1.0

Test instructions

OpenCASA

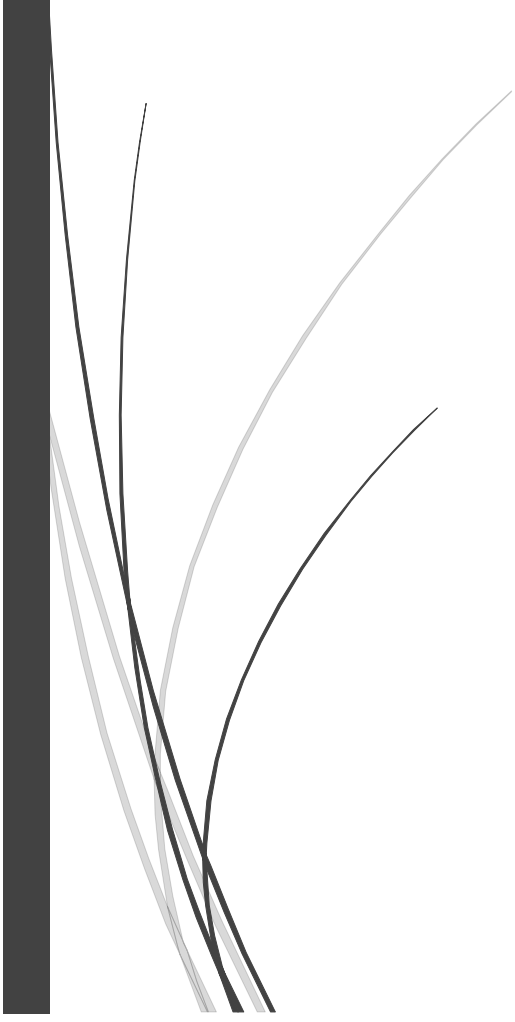


Table of Contents

Introduction 1

Simulation 1

Chemotaxis 2

- Dataset 2
- Analyze file 2
- Simulated experiment 3
  - Ch-Index 4
  - Bootstrapping 5
- Experiment analysis 6

Motility 7

- Analyze file 7
- Analyze directory 8
- Analyze multiple directories 8

Viability 9

Morphometry 11

Validation data 12

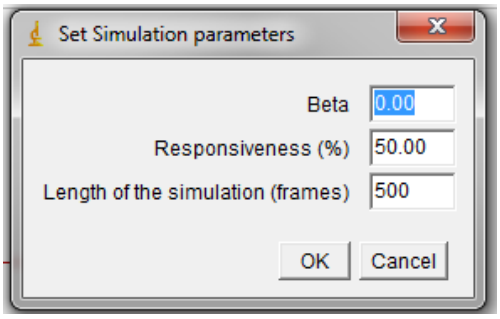
- Motility 12
- Viability 13
- Morphometry 14

Introduction

In this document are presented the instructions to test the software. Except for the simulation module, all configuration parameters of OpenCASA that were set to carry out each analysis, are available at the corresponding directory of the test dataset that goes with this document.

Simulation

This module generates a simulation of 100 cells in movement. There are three parameters needed to generate a simulation:

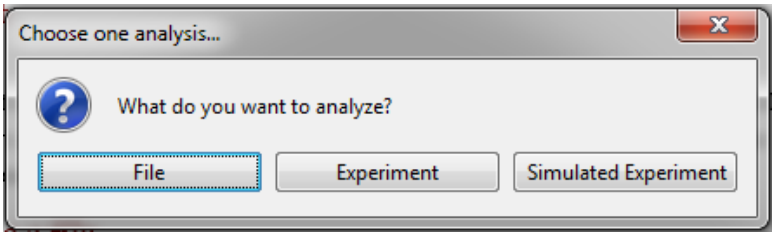


Beta equals 0 means the population will behave without any bias in the directionality of the swimming path. Values greater/lower than 0 will force the cells to move to the right/left respectively. Responsiveness is the percentage of cells that will be forced to some direction when beta is not zero. The length of the simulation is the duration (in frames) of the simulated video. In order to test visually the effect and aspect of the simulations, check the following situations:

Condition	Parameters	Expected response
Control	Beta=0; Responsiveness = indifferent	Random swimming paths in all directions
Attraction to the right (moderate)	Beta = 2 Responsiveness = 50	The half of the population will eventually move to the right side of the screen.
Attraction to the right (High level)	Beta = 2 Responsiveness = 100	All the population will eventually move to the right side of the screen.
Attraction to the left (moderate)	Beta = -2 Responsiveness = 50	The half of the population will eventually move to the left side of the screen.

Chemotaxis

The module allows the user to carry out three types of analysis: analyze a file, an experiment, or simulate an experiment. Each analysis has different purposes: analyze a single file allows to check the distributions of the instantaneous displacements of the whole population, and also to take a look at the relative trajectories; Experiment analysis tries to characterized the bias in the swimming direction of a particular set of sperm populations, comparing it to a control set; and Simulated Experiment helps the user to validate and learn the basis of an Experiment analysis.



Dataset

In order to test the module, various videos are provided. In the chemotaxis folder of the validation data, it is possible to find the "Experiment" folder, with 4 simulations: two about control conditions, and two of biased populations.

Analyze file

Click on ***Chemotaxis Module => Analyze File***. After selecting a video, the program will show two pictures: one about the relative trajectories detected by the module, and the other about the distribution of the instantaneous directionality angles of the whole population. In color green is drawn the chemotactic cone, and in color red the chemotaxis direction set by the user.

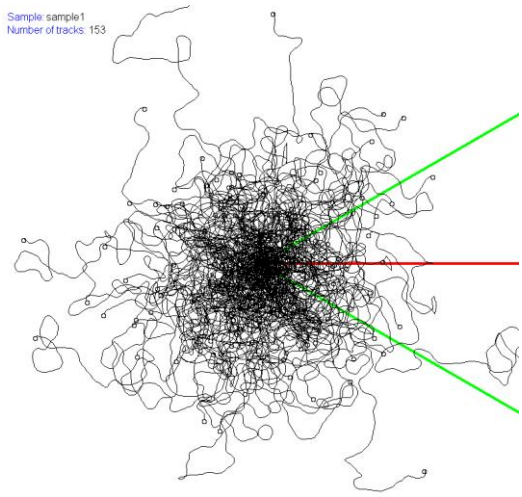


Figure 1 - Relative trajectories detected by the software. Each trajectory is drawn as if it started from the origin of coordinates.

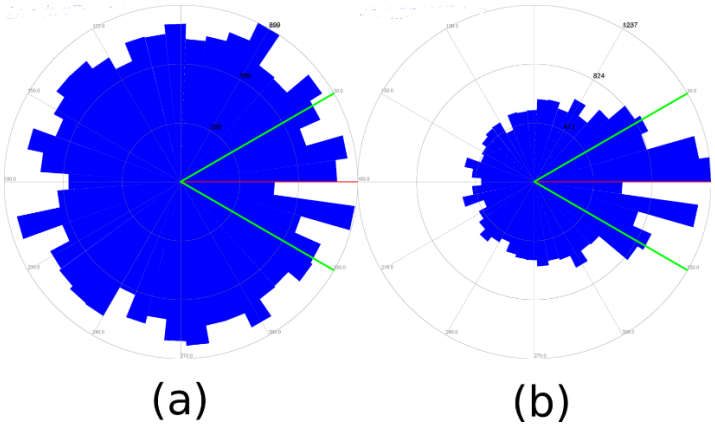
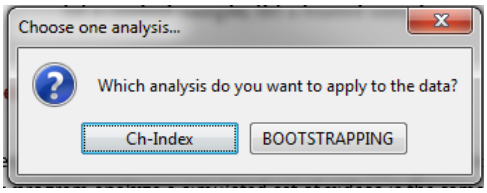


Figure 2- Differences on the distributions of the instantaneous directionality angles. (a) a control sample; (b) a biased sample

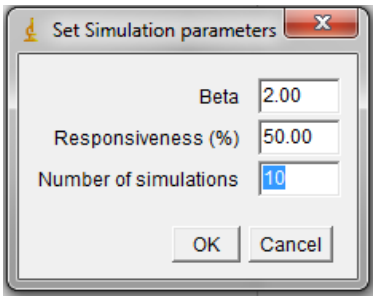
Simulated experiment

This analysis emerge from the need of validate the bootstrapping analysis for a large set of files. The way the program analyze a simulated set of videos is the same that is carried out with a real ones, but the process of creating the simulations is embedded in order to save time to the user. Other way would be to create manually a set of videos using the simulation module, place them in a folder (correctly structured as it is explained in the user's manual), and later carry out an "Experiment analysis".

To simulate an experiment: click on "**Chemotaxis --> Simulated Experiment**". There will appear two options: Ch-Index or Bootstrapping.

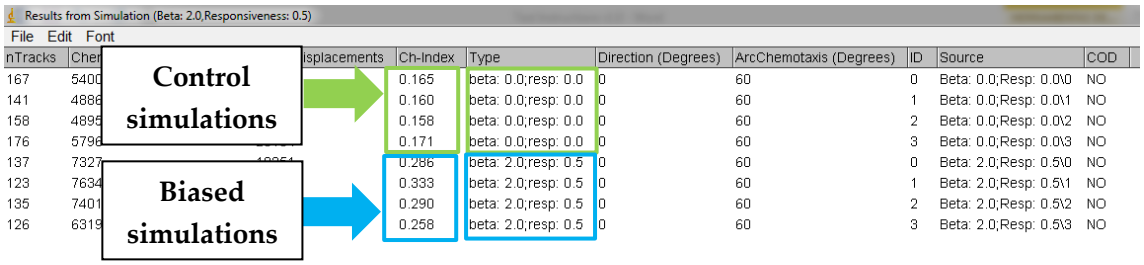


Each option is explained below but for both cases, the user will have to set the parameters of the simulations. As in the Simulation module, Beta equals 0 means the population will behave without any bias in the directionality of the swimming path. Values greater or lower than 0 will force the cells to move to the right or left respectively. Responsiveness is the percentage of cells that will be forced when beta is not zero. The length of all simulations in this module is fixed to 500 frames. The number of simulations means the number of pairs control-biased samples that will be simulated. For example, a value of 10 will create 10 simulations with Beta equals 0, paired up with another 10 simulations created with those parameters set by the user (in total, 20 simulations).



Ch-Index

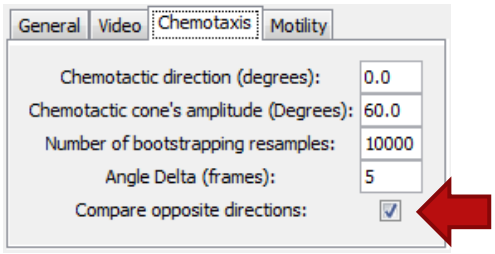
The Ch-Index analysis gives a measure of the percentage of the instantaneous directionality angles that point to the gradient direction respect to those not pointing to that gradient (an explanation in more detail can be found in the attached research article). In order to validate this analysis, it is necessary to check that, in control conditions (no bias in the directionality), the percentage of the instantaneous displacements pointing to any direction has to be proportional to the area taken into account. For example, if the user defines a chemotactic cone of 60 degrees, if there is no bias in the directionality, statistically the Ch-Index should be $\frac{60°}{360°} * 100 \approx 16{,}67\%$. Else, if there exists a bias in the chemotaxis direction, the percentage should be greater. This test can be seen in Figure 3, setting Beta equals 2 and Responsiveness equals 50%.



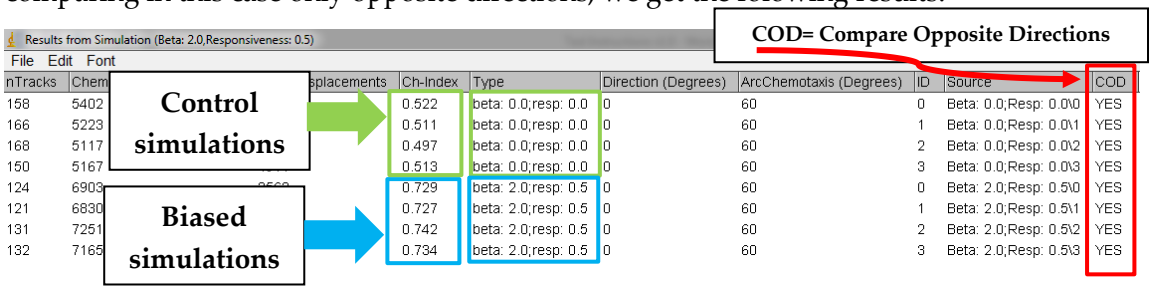
nTracks	Ch-Index	Type	Direction (Degrees)	ArcChemotaxis (Degrees)	ID	Source	COD
167	0.165	beta: 0.0;resp: 0.0	0	60	0	Beta: 0.0;Resp: 0.0\0	NO
141	0.160	beta: 0.0;resp: 0.0	0	60	1	Beta: 0.0;Resp: 0.0\1	NO
158	0.158	beta: 0.0;resp: 0.0	0	60	2	Beta: 0.0;Resp: 0.0\2	NO
176	0.171	beta: 0.0;resp: 0.0	0	60	3	Beta: 0.0;Resp: 0.0\3	NO
137	0.286	beta: 2.0;resp: 0.5	0	60	0	Beta: 2.0;Resp: 0.5\0	NO
123	0.333	beta: 2.0;resp: 0.5	0	60	1	Beta: 2.0;Resp: 0.5\1	NO
135	0.290	beta: 2.0;resp: 0.5	0	60	2	Beta: 2.0;Resp: 0.5\2	NO
126	0.258	beta: 2.0;resp: 0.5	0	60	3	Beta: 2.0;Resp: 0.5\3	NO

Figure 3 - Ch-Index analysis of 4 pairs of simulations. The biased ones were created with this parameters: Beta=2; Responsiveness=50%. The Ch-Index was calculated taking into account all directions on the instantaneous displacements (option 1, see paper).

However, there is an option to take into account only the displacements in the same and opposite direction of the chemotaxis gradient. In that case, if the user defines a chemotactic cone of 60 degrees, if there is no bias in the directionality, statistically the Ch-Index should be $\frac{60°}{60°+60°} * 100 = 50\%$. This option can be set in the chemotaxis tab of the settings menu:



Analysing a set of simulations with the same parameters as the previous example, but comparing in this case only opposite directions, we get the folowing results:



nTracks	Ch-Index	Type	Direction (Degrees)	ArcChemotaxis (Degrees)	ID	Source	COD
158	0.522	beta: 0.0;resp: 0.0	0	60	0	Beta: 0.0;Resp: 0.0\0	YES
166	0.511	beta: 0.0;resp: 0.0	0	60	1	Beta: 0.0;Resp: 0.0\1	YES
168	0.497	beta: 0.0;resp: 0.0	0	60	2	Beta: 0.0;Resp: 0.0\2	YES
150	0.513	beta: 0.0;resp: 0.0	0	60	3	Beta: 0.0;Resp: 0.0\3	YES
124	0.729	beta: 2.0;resp: 0.5	0	60	0	Beta: 2.0;Resp: 0.5\0	YES
121	0.727	beta: 2.0;resp: 0.5	0	60	1	Beta: 2.0;Resp: 0.5\1	YES
131	0.742	beta: 2.0;resp: 0.5	0	60	2	Beta: 2.0;Resp: 0.5\2	YES
132	0.734	beta: 2.0;resp: 0.5	0	60	3	Beta: 2.0;Resp: 0.5\3	YES

Bootstrapping

The Bootstrapping analysis act as a binary classifier (positive/negative) looking at the distribution of the instantaneous displacements of two populations: a control condition and a hypothesized biased one. The software will label a sample as positive if, statistically compared to a control condition, exists a bias in the direction of the chemotaxis gradient. By definition, this clasification has a 5% rate of false positives, so obtaining percentage of positives upper this value would indicate a chemotaxis phenomena in our samples. To test this analysis, we can carry out two different analysis. First, we can test if, analysing a set of NON-biased samples, we obtain only a 5% of false positives (take into account that this

percentage in the theoretical one obtained when the sample size is infinite, so a little variance around this value could be expected). To do it, we can simulate for example, 100 samples with Beta equals 0, and count the number of positives that we obtain. It is recommended to do it several times in order to be sure that this number is close to the theoretical 5%. Later, it is necessary to test that this percentage grows up when a bias is introduced. To do that, just simulate again setting the value of Beta above 0 (for example, Beta = 0.5). The value of Responsiveness is also important because is the number of cells that actually respond to the chemotactic effect. For the test below, Responsiveness has been set to 50%.

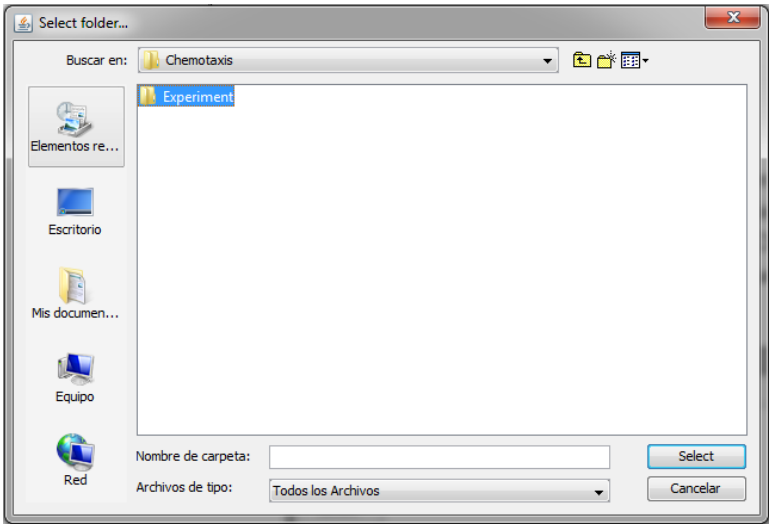
Results from Simulation (Beta: 0.5,Responsiveness: 0.5)

File Edit Font

ID	OR	Threshold	Result	Type	Source	COD
0	1.344	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\0	NO
1	1.366	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\1	NO
2	1.362	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\2	NO
3	1.293	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\3	NO
4	1.196	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\4	NO
5	1.395	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\5	NO
6	1.450	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\6	NO
7	1.527	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\7	NO
8	1.334	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\8	NO
9	1.374	1.116	POSITIVE	beta: 0.5;resp: 0.5	Beta: 0.5;Resp: 0.5\9	NO

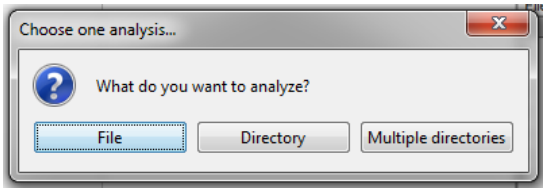
Experiment analysis

In order to test the third analysis, four videos are provided. After clicking on the button "Experiment Analysis" a dialog window is opened. After choosing between Ch-Index or Bootstrapping analysis, select the folder "Experiment" inside the Chemotaxis directory and the analysis will be carried out as in the previous section for simulated data.



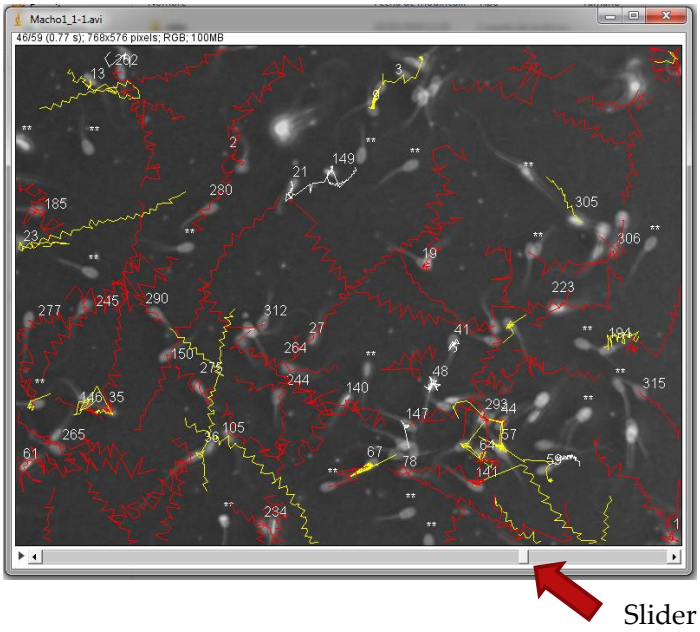
Motility

In order to test the motility module, a set of videos is provided. This module allows the user to analyze files at three levels: one single file, a directory with different videos, or a directory with multiple directorie.

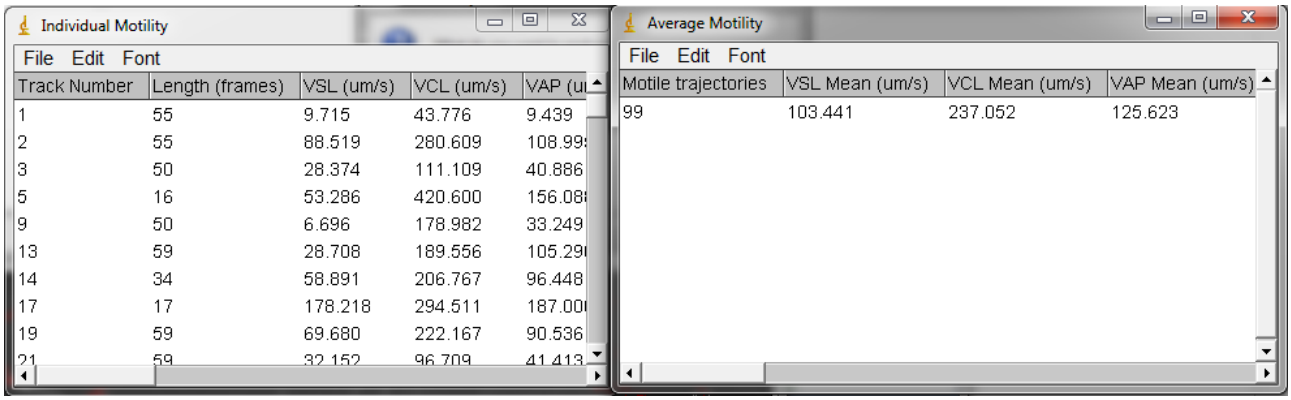


Analyze file

When a single file is analyzed, the original video is shown with the detected trajectories overwriting it. The user can check the correspondences between the trajectories and the cells, using the slider at the bottom of the window:



Also, the analysis return two reports: one for the individual analysis, and other with the average values for the whole sample.

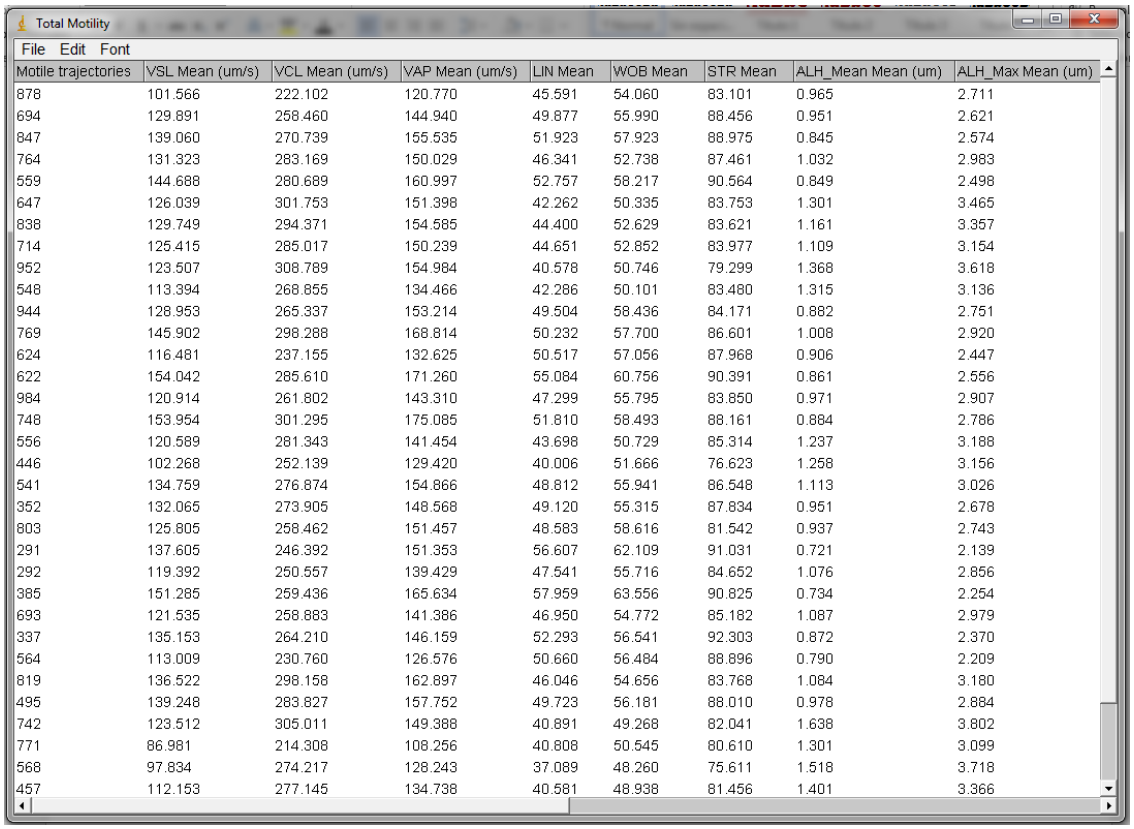


Analyze directory

The directory analysis is just the automation of the file analysis but for multiple files. The only difference is that in this analysis, the trajectories over the videos are not shown.

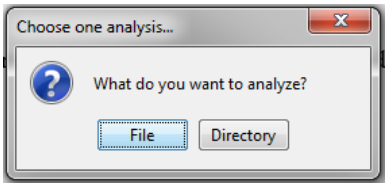
Analyze multiple directories

This analysis is the automation for the previous analysis. Selecting this option, multiple directories are analyzed, but in this case, the report given by the program shows only the average values calculated for each subdirectory, independently of how many samples has each directory. The report shows the average values (one directory for each row):



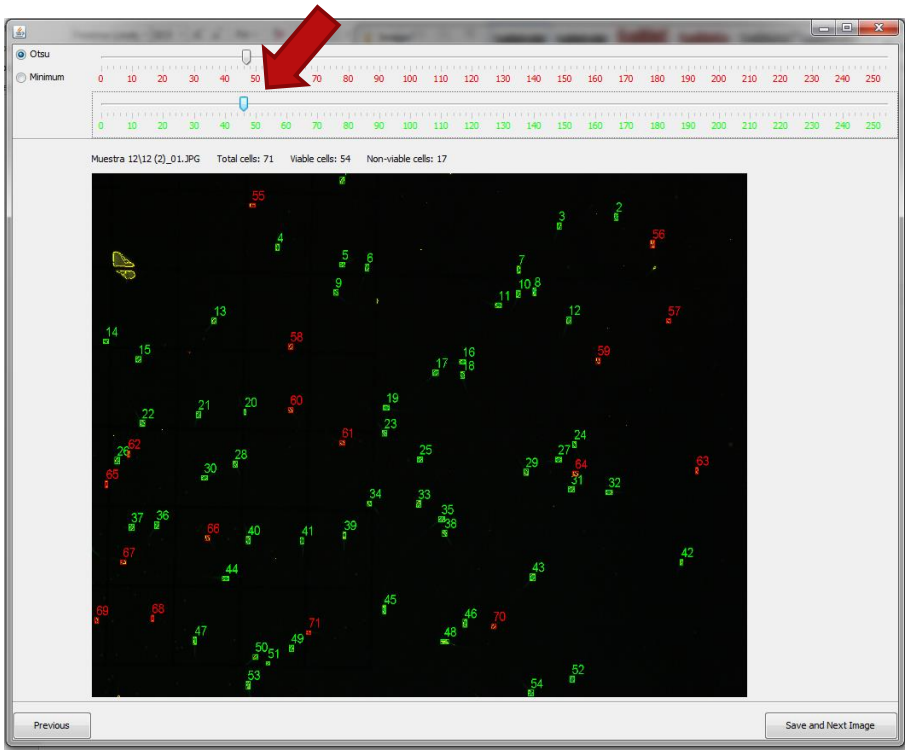
Viability

In order to test the viability module, a set of images is provided. This module allows the user to choose between a single file analysis (only one image), or a directory analysis (all images inside the selected directory). Both analysis are equivalent so here we will explain only how to carry out a directory analysis.

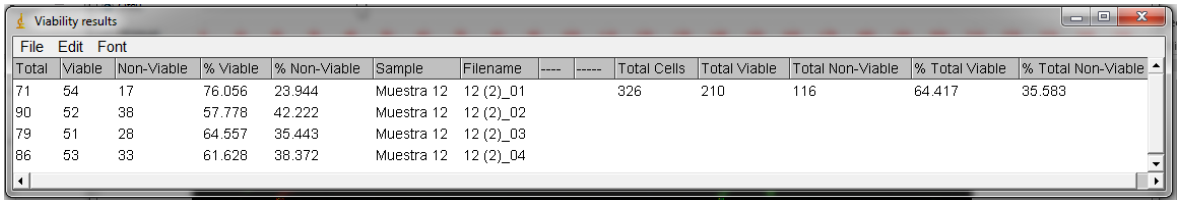


Once the directory has been selected, the program loads the first image and it tries to classify the cells inside the image between viable and non-viable cells. In case that some cells are labeled in a wrong way, the user can modify the thresholds used for red and green channels, just moving the corresponding slider:

Sliders

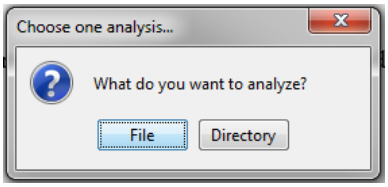


Once the user considers that the classification is correct, clicking on bottom-right button "Save and Next Image", the results will be added to the report window. In this window, on the left we will see the results for each analyzed image, and on the right, the absolute and relative number of analyzed cells. This second part is updated each time a new image is analyzed.

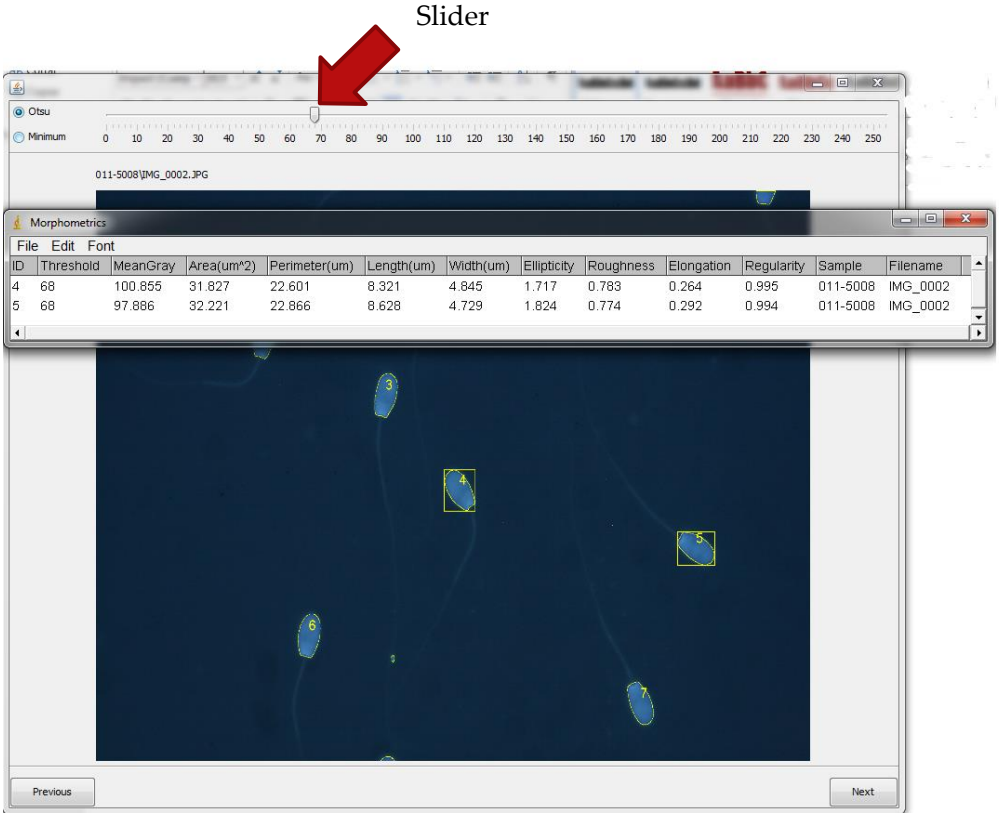


Morphometry

In order to test the morphometry module, a set of images is provided. This module allows the user to choose between a single file analysis (only one image), or a directory analysis (all images inside the directory). Both analysis are equivalent so here we will explain how to carry out a directory analysis.



Once the directory has been selected, the program loads the first image and it tries to determine the optimal threshold to detect the contour of the cell. In this module, the user has to select manually those cells that have been well outlined. When it occurs, the morphometric values of the cell are added to the report window. If a cell has not been well outlined, the user can adjust manually the threshold just moving the slider on the top of the window.



Validation data

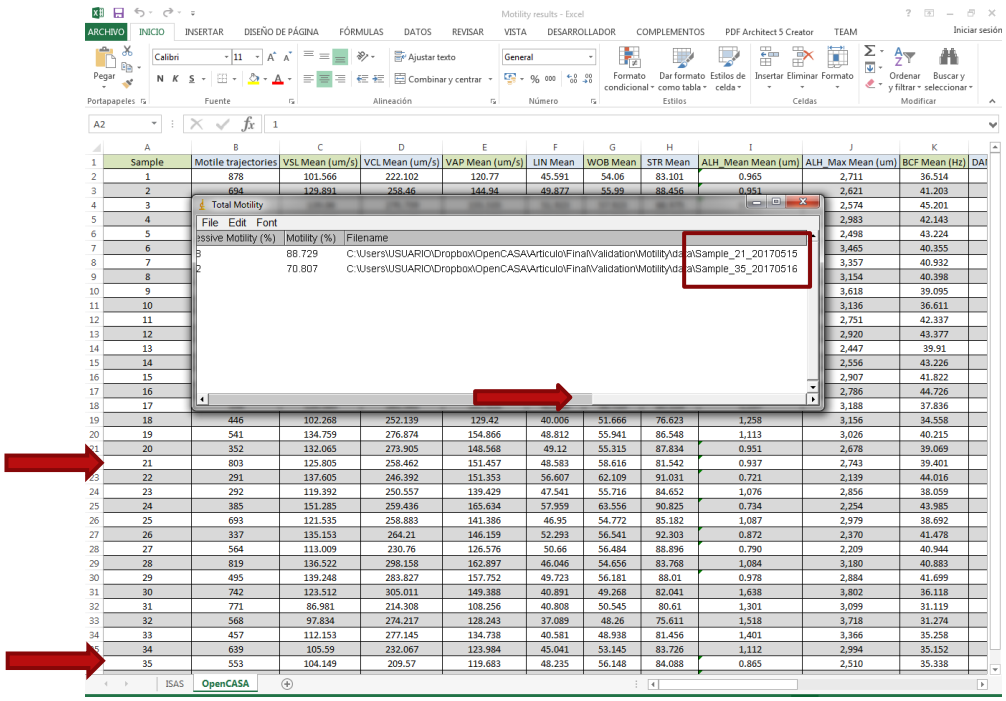
Considering that the simulation module can be tested visually, the chemotaxis module is perfectly validated with these simulations, and the rest of the modules have been validated using real data, the results of that validation are provided in order to allow the user to replicate and confirm the analysis. For that reason, a subset of the validation data is provided together with the results of the whole analysis. Taking into account the type of analysis, and the identification of each sample, the user has to be able to replicate the corresponding results.

Motility

The file "Motility results.xlsx" contains the results of the analysis carried out for each sample (both for the ISAS and OpenCASA software). Also, two samples with 8 videos each sample are provided. The steps that the user has to follow to replicate the results of the analysis of that samples are:

1. Set the OpenCASA parameters as said in the file "OpenCASA Parameters Motility.txt" placed in the motility folder.
2. Select the module Motility and click on "Multiple directories"
3. Select the folder "data" contained in the motility folder

After that, compare that the results obtained in the analysis are the same that those that figure on the excel file. To identify each sample, just move to the end of the report to check the last part of the filename and look for the sample ID on the first column of the spreadsheet:



Viability

The file "Viability results.xlsx" contains the results of the analysis carried out for each sample. Also, five samples with 52 images in total are provided. The steps that the user has to follow to replicate the results of the analysis of that samples are:

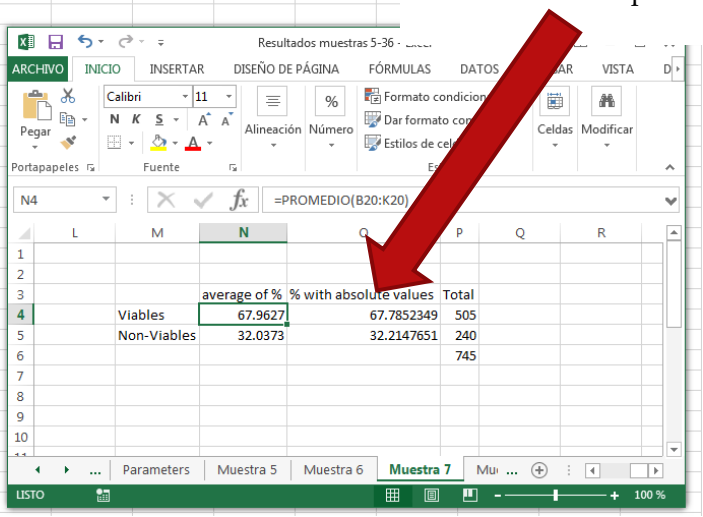
1. Set the OpenCASA parameters as said in the file "OpenCASA Parameters Viability.txt" placed in the viability folder.
2. Select the module Viability and click on "analyze directory"
3. Select the folder of one sample contained in the data folder.
4. Analyze the images of that sample and save the results
5. Repeat that process for each sample directory
6. After analyzing all samples, compare the values obtained with those of the same samples provided in the Viability results.xlsx file

The validation results

Sample	% viables	% non-viables	Total cells
Sample 5	65.43560606	34.56439394	1056
Sample 6	84.05797101	15.94202899	552
Sample 7	67.7852349	32.2147651	745
Sample 8	74.01247401	25.98752599	481
Sample 9	71.54255319	28.45744681	376
Sample 10	65.72769953	34.27230047	1065
Sample 11	61.10056926	38.89943074	1054
Sample 12	60.84099869	39.15900131	761
Sample 13	57.61589404	42.38410596	1057
Sample 14	82.47767857	17.52232143	896
Sample 15	73.24840764	26.75159236	942
Sample 16	59.82324949	40.17675051	1471
Sample 17	62.92613636	37.07386364	704
Sample 18	75.98499062	24.01500938	1066
Sample 19	68.41317365	31.58682635	668
Sample 20	64.70588235	35.29411765	391
Sample 21	86.35235732	13.64764268	806
Sample 22	89.03591682	10.96408318	529
Sample 23	71.56862745	28.43137255	612
Sample 24	64.98127341	35.01872659	534
Sample 25	73.26203209	26.73796791	748
Sample 26	46.42857143	53.57142857	392
Sample 27	88.69047619	11.30952381	168
Sample 28	90.14084507	9.85915493	284
Sample 29	75.8677686	24.1322314	605
Sample 30	79.57957958	20.42042042	333
Sample 31	76.96078431	23.03921569	408
Sample 35	45.85635359	54.14364641	362
Sample 36	43.42105263	56.57894737	380

Total cells	
mean	std
670.551724	308.493411

The results obtained by the user for the same sample



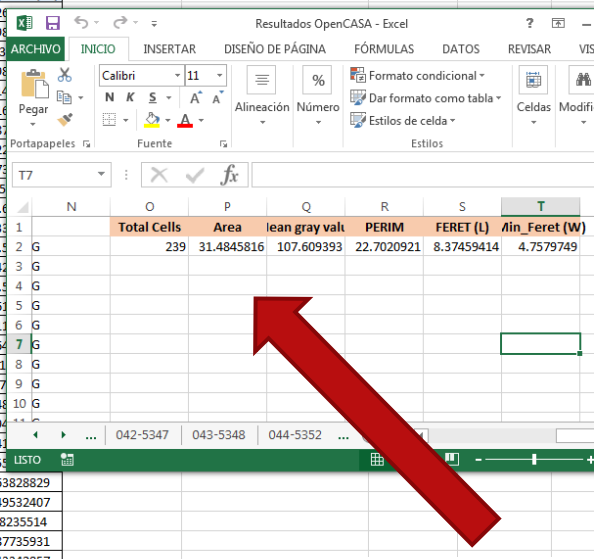
Morphometry

The file "Morphometry results.xlsx" contains the results of the analysis carried out for each sample (both for the ISAS and OpenCASA software). Also, three samples with 100 images in total are provided. The steps that the user has to follow to replicate the results of the analysis of that samples are:

1. Set the OpenCASA parameters as said in the file "OpenCASA Parameters Morphometry.txt" placed in the viability folder.
2. Select the module Morphometry and click on "analyze directory"
3. Select the folder of one sample contained in the data folder.
4. Analyze the images of that sample and save the results. Select around 200 cells/sample, equally distributed across all images.
5. Repeat that process for each sample directory
6. After analyzing all samples, compare the values obtained with those of the same samples provided in the Morphometry results.xlsx file.

The validation results

Sample ID	Total Cells	Area	Mean gray value	PERIM	FERET (L)	Min_Feret (W)
001-97690	202	30.52458911	112.0969851	22.63931683	8.287950495	4.626
002-81302	204	31.76346569	154.0151618	22.90215196	8.520553922	4.698
003-16384	204	30.96740686	156.232652	22.14803431	8.052196078	4.83
004-7170	204	30.65879902	126.9159755	22.69284804	8.448622549	4.598
005-97689	204	30.88681373	119.8051422	22.60832353	8.277617647	4.714
006-16375	204	31.93037255	111.265549	22.90313235	8.333818627	4.816
007-61871	204	31.9500049	139.5103431	23.01137745	8.498985294	4.687
008-16383	204	30.8322402	99.83638235	22.65676961	8.269102941	4.722
009-5347	205	31.13039512	128.1255707	22.58282927	8.31382439	4.773
010-753	200	29.201655	88.61499	21.9796	8.09697	4.5
011-5008	200	29.425945	110.025495	21.844905	8.054535	4.6
012-6085	208	30.27119231	93.92761058	22.35507692	8.267394231	4.633
013-6091	200	30.526355	104.0199	22.50462	8.402215	4.5
014-6125	205	31.55706341	93.25346341	22.84330732	8.414512195	4.742
015-7150	200	29.414505	87.20177	21.984995	8.127985	4.5
016-7151	202	31.35138119	91.36432673	22.71045545	8.328450495	4.761
017-7170	208	30.79721154	87.59507212	22.63678846	8.458913462	4.611
018-5008	204	28.81491667	122.6876765	21.57395098	7.970142157	4.564
020-5348	204	31.38675	111.287049	22.46971078	8.209176471	4.81
021-5350	200	31.268485	119.306	22.53635	8.276185	4.7
022-7161	204	31.071	135.4472353	22.51659314	8.291691176	4.748
023-9201	210	31.17112857	140.3046762	22.59303333	8.364585714	4.694
024-5347	208	30.71009615	102.6293077	22.44624519	8.252206731	4.741
025-5348	208	31.80598558	108.6089471	22.55971635	8.258072115	4.855
026-9201	222	29.66798649	101.7584505	22.07135586	8.18940991	4.563828829
027-97689	216	31.48284722	109.5584213	22.64615278	8.359361111	4.749532407
028-5008	214	29.06092991	121.9877009	21.6689486	8.003242991	4.58235514
029-5348	231	31.05004329	110.3488658	22.36737229	8.173909091	4.787735931
030-5350	210	30.09894286	120.5549714	22.15199048	8.161861905	4.643342857
031-7150	213	29.49221596	91.18520188	21.99075587	8.088267606	4.643685446
032-7151	201	31.25090547	96.65122388	22.53126368	8.278985075	4.784850746
033-9201	210	31.30850952	112.825019	22.60827619	8.362352381	4.714709524
034-80905	211	31.23971564	100.0855261	22.34725592	8.128838863	4.858075829
035-83992	209	31.37808612	104.1360957	22.740311	8.390253589	4.737708134
036-5349	216	30.75215278	98.29307407	22.24647222	8.1765	4.733416667



The results obtained by the user for the same sample