Journal of Visualized Experiments

Substructure Analyzer, a user-friendly workflow for rapid exploration and accurate analysis of cellular bodies in fluorescence microscopy images.
--Manuscript Draft--

1 **TITLE:**
2 Substructure Analyzer: A User-Friendly Workflow for Rapid Exploration and Accurate Analysis of
3 Cellular Bodies in Fluorescence Microscopy Images
4
5 **AUTHORS AND AFFILIATIONS:**
6 Géraud Heckler[1], Christelle Aigueperse[1], Liza Hettal[1], Quentin Thuillier[1], Fabrice de Chaumont[2],
7 Stéphane Dallongeville[2], Isabelle Behm-Ansmant[1]
8
9 [1]Université de Lorraine, CNRS, IMoPA F54000 Nancy France
10 [2]Institut Pasteur, BioImage Analysis Unit, CNRS UMR 3691, Paris, France
11
12 Corresponding Author:
13 Isabelle Behm-Ansmant
14 Isabelle.behm@univ-lorraine.fr
15
16 Email Addresses of Co-authors**:**
17 Géraud Heckler   (geraud.heckler@univ-lorraine.fr)
18 Christelle Aigueperse   (christelle.aigueperse@univ-lorraine.fr)
19 Liza Hettal   (liza.hettal@univ-lorraine.fr)
20 Quentin Thuillier   (quentin.thuillier@univ-lorraine.fr)
21 Fabrice de Chaumont   (fabrice.de-chaumont@pasteur.fr)
22 Stéphane Dallongeville   (stephane.dallongeville@pasteur.fr)
23
24 **KEYWORDS:**
25 Fluorescence microscopy; Bio-image analysis; Cell segmentation; Foci counting; Cellular bodies;
26 Oxidative stress; Automatization; Translocation analysis; Substructure Analyzer; Icy; Image J
27
28 **SUMMARY:**
29 We present a freely available workflow built for rapid exploration and accurate analysis of cellular
30 bodies in specific cell compartments in fluorescence microscopy images. This user-friendly
31 workflow is designed on the open-source software Icy and also uses ImageJ functionalities. The
32 pipeline is affordable without knowledge in image analysis.
33
34 **ABSTRACT:**
35 The last decade has been characterized by breakthroughs in fluorescence microscopy techniques
36 illustrated by spatial resolution improvement but also in live-cell imaging and high-throughput
37 microscopy techniques. This led to a constant increase in the amount and complexity of the
38 microscopy data for a single experiment. Because manual analysis of microscopy data is very time
39 consuming, subjective, and prohibits quantitative analyses, automation of bioimage analysis is
40 becoming almost unavoidable. We built an informatics workflow called *Substructure Analyzer* to
41 fully automate signal analysis in bioimages from fluorescent microscopy. This workflow is
42 developed on the user-friendly open-source platform Icy and is completed by functionalities from
43 ImageJ. It includes the pre-processing of images to improve the signal to noise ratio, the
44 individual segmentation of cells (detection of cell boundaries) and the detection/quantification

45 of cell bodies enriched in specific cell compartments. The main advantage of this workflow is to
46 propose complex bio-imaging functionalities to users without image analysis expertise through a
47 user-friendly interface. Moreover, it is highly modular and adapted to several issues from the
48 characterization of nuclear/cytoplasmic translocation to the comparative analysis of different cell
49 bodies in different cellular sub-structures. The functionality of this workflow is illustrated through
50 the study of the Cajal (coiled) Bodies under oxidative stress (OS) conditions. Data from
51 fluorescence microscopy show that their integrity in human cells is impacted a few hours after
52 the induction of OS. This effect is characterized by a decrease of coilin nucleation into
53 characteristic Cajal Bodies, associated with a nucleoplasmic redistribution of coilin into an
54 increased number of smaller foci. The central role of coilin in the exchange between CB
55 components and the surrounding nucleoplasm suggests that OS induced redistribution of coilin
56 could affect the composition and the functionality of Cajal Bodies.
57
58 **INTRODUCTION:**
59 Light microscopy and, more particularly, fluorescence microscopy are robust and versatile
60 techniques commonly used in biological sciences. They give access to the precise localization of
61 various biomolecules like proteins or RNA through their specific fluorescent labeling. The last
62 decade has been characterized by rapid advances in microscopy and imaging technologies as
63 evidenced by the 2014 Nobel Prize in Chemistry awarding Eric Betzig, Stefan W. Hell and William
64 E. Moerner for the development of super-resolved fluorescence microscopy (SRFM)[1]. SFRM
65 bypasses the diffraction limit of traditional optical microscopy to bring it into the nanodimension.
66 Improvement in techniques like live-imaging or high throughput screening approaches also
67 increases the amount and the complexity of the data to treat for each experiment. Most of the
68 time, researchers are faced with high heterogeneous populations of cells and want to analyze
69 phenotypes at the single-cell level.
70
71 Initially, analyses such as foci counting were performed by eye, which is preferred by some
72 researchers since it provides full visual control over the counting process. However, manual
73 analysis of such data is too much time consuming, leads to variability between observers, and
74 does not give access to more complex features so that computer-assisted approaches are
75 becoming widely used and almost unavoidable[2]. Bioimage informatics methods substantially
76 increase the efficiency of data analysis and are free of the unavoidable operator subjectivity and
77 potential bias of the manual counting analysis. The increased demand in this field and the
78 improvement of computer power led to the development of a large number of image analysis
79 platforms. Some of them are freely available and give access to various tools to perform analysis
80 with personal computers. A classification of open access tools has been recently established[3] and
81 presents Icy[4] as a powerful software combining usability and functionality. Moreover, Icy has the
82 advantage of communicating with ImageJ.
83
84 For users without image analysis expertise, the main obstacle is to choose the appropriate tool
85 according to the problematic and correctly tuned parameters that are often not well understood.
86 Moreover, setup times are often long. Icy proposes a user-friendly point-and-click interface
87 named "Protocols" to develop workflow by combining some plugins found within an exhaustive
88 collection[4]. The flexible modular design and the point-and-click interface make setting up an

89 analysis feasible for non-programmers. Here we present a workflow called *Substructure Analyzer*,
90 developed in Icy's interface, whose function is to analyze fluorescent signals in specific cellular
91 compartments and measure different features brightness, foci number, foci size, spatial
92 distribution. This workflow addresses several issues such as quantification of signal translocation,
93 analysis of transfected cells expressing a fluorescent reporter, or analysis of foci from different
94 cellular substructures in individual cells. It allows the simultaneous processing of multiple images,
95 and output results are exported to a tab-delimited worksheet that can be opened in commonly
96 used spreadsheet programs.
97
98 The *Substructure Analyzer* pipeline is presented in **Figure 1**. First, all the images contained in a
99 specified folder are pre-processed to improve their signal to noise ratio. This step increases the
100 efficiency of the following steps and decreases the running time. Then, the Regions of Interest
101 (ROIs), corresponding to the image area where the fluorescent signal should be detected, are
102 identified and segmented. Finally, the fluorescent signal is analyzed, and results are exported into
103 a tab-delimited worksheet.
104
105 Object segmentation (detection of boundaries) is the most challenging step in image analysis,
106 and its efficiency determines the accuracy of the resulting cell measurements. The first objects
107 identified in an image (called primary objects) are often nuclei from DNA-stained images (DAPI
108 or Hoechst staining), although primary objects can also be whole cells, beads, speckles, tumors,
109 or whatever stained objects. In most biological images, cells or nuclei touch each other or overlap
110 causing the simple and fast algorithms to fail. To date, no universal algorithm can perform perfect
111 segmentation of all objects, mostly because their characteristics (size, shape, or texture)
112 modulate the efficiency of segmentation[5]. The segmentation tools commonly distributed with
113 microscopy software (such as the MetaMorph Imaging Software by Molecular Devices[6], or the
114 NIS-Elements Advances Research software by Nikon[7]) are generally based on standard
115 techniques such as correlation matching, thresholding, or morphological operations. Although
116 efficient in basic systems, these overgeneralized methods rapidly present limitations when used
117 in more challenging and specific contexts. Indeed, segmentation is highly sensitive to
118 experimental parameters such as cell type, cell density, or biomarkers, and frequently requires
119 repeated adjustment for a large data set. The *Substructure Analyzer* workflow integrates both
120 simple and more sophisticated algorithms to propose different alternatives adapted to image
121 complexity and user needs. It notably proposes the marker-based watershed algorithm[8] for
122 highly clustered objects. The efficiency of this segmentation method relies on the selection of
123 individual markers on each object. These markers are manually chosen most of the time to get
124 correct parameters for full segmentation, which is highly time-consuming when users face a high
125 number of objects. *Substructure Analyzer* proposes an automatic detection of these markers,
126 providing a highly efficient segmentation process. Segmentation is, most of the time, the limiting
127 step of image analysis and can considerably modify the processing time depending on the
128 resolution of the image, the number of objects per image, and the level of clustering of objects.
129 Typical pipelines require a few seconds to 5 minutes per image on a standard desktop computer.
130 Analysis of more complex images can require a more powerful computer and some basic
131 knowledge in image analysis.
132

133 The flexibility and functionality of this workflow are illustrated with various examples in the
134 representative results. The advantages of this workflow are notably displayed through the study
135 of nuclear substructures under oxidative stress (OS) conditions. OS corresponds to an imbalance
136 of the redox homeostasis in favor of oxidants and is associated with high levels of reactive oxygen
137 species (ROS). Since ROS act as signaling molecules, changes in their concentration and
138 subcellular localization affect positively or negatively a myriad of pathways and networks that
139 regulate physiological functions, including signal transduction, repair mechanisms, gene
140 expression, cell death, and proliferation[9,10]. OS is thus directly involved in various pathologies
141 (neurodegenerative and cardiovascular diseases, cancers, diabetes, etc.), but also cellular aging.
142 Therefore, deciphering the consequences of OS on the human cell's organization and function
143 constitutes a crucial step in the understanding of the roles of OS in the onset and development
144 of human pathologies. It has been established that OS regulates gene expression by modulating
145 transcription through several transcription factors (p53, Nrf2, FOXO3A)[11], but also by affecting
146 the regulation of several co- and post-transcriptional processes such as alternative splicing (AS)
147 of pre-RNAs[12-14]. Alternative splicing of primary coding and non-coding transcripts is an essential
148 mechanism that increases the encoding capacity of the genomes by producing transcript
149 isoforms. AS is performed by a huge ribonucleoprotein complex called spliceosome containing
150 almost 300 proteins and 5 U-rich small nuclear RNAs (UsnRNAs)[15]. Spliceosome assembly and AS
151 are tightly controlled in cells and some steps of the spliceosome maturation occur within
152 membrane-less nuclear compartments named Cajal Bodies. These nuclear substructures are
153 characterized by the dynamic nature of their structure and their composition, which are mainly
154 conducted by multivalent interactions of their RNA and protein components with the coilin
155 protein. Analysis of thousands of cells with the *Substructure Analyzer* workflow allowed
156 characterization of never described effects of OS on Cajal Bodies. Indeed, obtained data suggest
157 that OS modifies the nucleation of Cajal Bodies, inducing a nucleoplasmic redistribution of the
158 coilin protein into numerous smaller nuclear foci. Such a change of the structure of Cajal Bodies
159 might affect the maturation of the spliceosome and participate in AS modulation by OS.

161 **PROTOCOL:**

163 NOTE: User-friendly tutorials are available on Icy's website **http://icy.bioimageanalysis.org**.

165 **1. Download Icy and the *Substructure Analyzer* protocol**

167 1.1. Download Icy from the Icy website (**http://icy.bioimageanalysis.org/download)** and
168 download the **Substructure Analyzer** protocol:
169 **http://icy.bioimageanalysis.org/protocols?sort=latest.**

171 NOTE: If using a 64-bit OS, be sure to use the 64-bit version of Java. This version allows for
172 increasing the memory allocated to Icy (**Preferences | General | Max memory**).

174 **2. Opening the protocol**

176 2.1. Open Icy and click on **Tools** in the **Ribbon** menu.

177

178   2.2.   Click on **Protocols** to open the **Protocols Editor** interface.

179

180   2.3.   Click on **Load** and open the protocol **Substructure Analyzer**. Protocol loading can take a
181   few seconds. Be sure that the opening of the protocol is complete before using it.

182

183   NOTE: The workflow is composed of **13 general blocks** presented in **Figure 2a**. Each block works
184   as a pipeline composed of several boxes performing specific subtasks.

185

186   3.   **Interacting with the workflow on Icy**

187

188   NOTE: Each block or box is numbered and has a specific rank within the workflow (**Figure 2b**). By
189   clicking on this number, the closest possible position to the first is assigned to the selected
190   block/box then the position of the other blocks/boxes is re-organized. Respect the right order of
191   the blocks when preparing the workflow. For example, Spot Detector block needs pre-defined
192   ROIs so that Segmentation blocks have to run before Spot Detector blocks. Do not modify the
193   position of boxes. Do not use "." in the image's name.

194

195   3.1.   By clicking on the upper left corner icon, collapse, expand, enlarge, narrow, or remove
196   the block (**Figure 2b**).

197

198   3.2.   Each pipeline of the workflow is characterized by a network of boxes connected via their
199   input and output (**Figure 2b**). To create a connection, click on **Output** and maintain until the
200   cursor attains an input. Connections can be removed by clicking on the **Output** tag.

201

202   **4.   Merging of the channels of an image**

203

204   4.1.   Use the block **Merge Channels** to generate merged images. If necessary, rename the files
205   so that sequences to be merged have the same name's prefix followed by a distinct separator.
206   For example, sequences of individual channels from an Image A are named: ImageA_red,
207   ImageA_blue.

208

209   NOTE: For the separator, do not use characters already present in the image's name.

210

211   4.2.   In the same folder, create one new folder per channel to merge. For example, to merge
212   red, green, and blue channels, create 3 folders, and store the corresponding sequences in these
213   folders.

214

215   4.3.   Only use the block **Merge Channels**, remove the other blocks, and save the protocol as
216   **Merge Channels**.

217

218   4.3.1.   Access the boxes to set parameters. For each channel, fill the boxes **Channel number X**
219   (boxes 1, 5 or 9), **Folder channel number X** (boxes 2, 6 or 10), **Separator channel number X** (boxes
220   3, 7 or 11) and **Colormap channel nb X** (boxes 4, 8 and 12) respectively.

221

222 NOTE: These boxes are horizontally grouped by four, each line corresponding to the same
223 channel. In each line, a display is also available (boxes 23, 24, or 25) to directly visualize the
224 sequence of the corresponding channel.

225

226 4.3.1.1.   In the box **Channel number X**, choose which channel to extract (in classical RGB
227 images, 0=Red, 1=Green, 2=Blue). The user quickly accesses the different channels of an image
228 within the **Inspector** window of Icy, in the **Sequence** tab. Write the smallest channel's value in
229 the upper line and the highest one in the bottom line.

230

231 4.3.1.2. In the box **Folder channel number X**, write the \Name of the folder containing images of
232 channel X.

233

234 4.3.1.3. In the box **Separator channel number X**, write the separator used for the image's name
235 (in the previous example: "_red", "_green" and "_blue").

236

237 4.3.1.4. In the box **Colormap channel nb X**, indicate with a number which colormap model to use
238 to visualize the corresponding channel in Icy. The available colormaps are visible in the **Sequence**
239 tab of the **Inspector** window.

240

241 4.3.1.5. In the box **Format of merged images** (box 28), write the extension to save merged
242 images: .tif, .gif, .jpg, .bmp or .png.

243

244 NOTE: To merge only 2 channels, do not fill the four boxes corresponding to the third channel.

245

246 4.3.2.   On the upper left corner of the **Merge Channels** block, click on the link directly to the right
247 of **Folder**. In the **Open** dialog box which appears, double-click on the folder containing sequences
248 of the first channel that has been defined in the box **Folder channel number 1** (box 2). Then, click
249 on **Open**.

250

251 4.3.3.   Run the protocol by clicking on the black arrow in the upper left corner of the **Merge**
252 **Channels** block (see part 7 for more details). Merged images are saved in a **Merge** folder in the
253 same directory as the folders of individual channels.

254

255 **5.   Segmentation of the regions of interest**

256

257 NOTE: Substructure Analyzer integrates both simple and more sophisticated algorithms to
258 propose different alternatives adapted to image complexity and user needs.

259

260 5.1. Select the adapted block.

261

262 5.1.1. If objects do not touch each other, or the user does not need to differentiate clustered
263 objects individually, use the block **Segmentation A: Non clustered objects**.

264

265 5.1.2. When objects do not touch each other, but some of them are close, use the block
266 **Segmentation B: Poorly clustered objects**.
267
268 5.1.3. For objects with a high clustering level and a convex shape, use the block **Segmentation C:**
269 **Clustered objects with convex shapes**.
270
271 5.1.4. If objects present a high clustering level and have irregular shapes, use the block
272 **Segmentation D: Clustered objects with irregular shapes**.
273
274 5.1.5. Use the block **Segmentation E: Clustered cytoplasm** to segment touching cytoplasms
275 individually using segmented nuclei as markers. This block imperatively needs segmented nuclei
276 to process.
277
278 NOTE: Blocks adapted for primary object segmentation process independently so that several
279 blocks can be used in the same run to compare their efficiency for a particular substructure or to
280 segment different types of substructures. If the clustering level is heterogeneous within the same
281 set of images, then process small and highly clustered objects separately in the adapted blocks.
282
283 5.2. Link the **output0** (File) of the block **Select Folder** to the **folder** input of the chosen
284 segmentation block.
285
286 5.3. Set parameters of the chosen segmentation block.
287
288 5.3.1. Segmentation A: Non clustered objects and Segmentation C: Clustered objects with convex
289 shapes
290
291 5.3.1.1. In box **Channel signal** (box 1), set the channel of the objects to segment.
292
293 5.3.1.2. As an option, in box **Gaussian filter** (box 2), increase the X and Y sigma values if the signal
294 inside objects is heterogeneous. The Gaussian filter smooths out textures to obtain more uniform
295 regions and increases the speed and efficiency of nuclei segmentation. The smaller the objects,
296 the lower the sigma value is. Avoid high sigma values. Set default values to 0.
297
298 5.3.1.3. In box **HK-Means** (box 3), set the **Intensity classes** parameter and the approximate
299 minimum and maximum sizes (in pixels) of objects to be detected.
300
301 NOTE: For intensity classes, a value of 2 classifies pixels in 2 classes: background and foreground.
302 It is thus adapted when the contrast between the objects and the background is high. If
303 foreground objects have different intensities or if the contrast with the background is low,
304 increase the number of classes. The default setting is 2. Object size can be quickly evaluated by
305 drawing a ROI manually around the object of interest. Size of the ROI (Interior in pixels) appears
306 directly on the image when pointing it with the cursor or can be accessed in the **ROI statistics**
307 window (open it from the search bar). Optimal parameters detect each foreground object in a
308 single ROI**.** They can be manually defined in Icy (**Detection and Tracking | HK-Means**).

309

310 5.3.1.4. In the box **Active Contours** (box 4), optimize the detection of object borders. Exhaustive
311 documentation for this plugin is available online:
312 http://icy.bioimageanalysis.org/plugin/Active_Contours. Correct parameters can also be
313 manually defined in Icy (**Detection and Tracking | Active Contours**).

314

315 5.3.1.5. During the process, a folder is automatically created to save images of segmented
316 objects. In the box **Text** (box 6), name this folder (ex: Segmented nuclei). To set the format for
317 saving images of segmented objects (Tiff, Gif, Jpeg, BMP, PNG), fill the box **format of images of**
318 **segmented objects**. The folder is created in the folder containing merged images.

319

320 5.3.1.6. Run the workflow (for details, see part 7).

321

322 5.3.2. Segmentation B: Poorly clustered objects

323

324 5.3.2.1. Follow the same steps as in 5.3.1 to set parameters of boxes **Channel signal**, **HK-Means**,
325 **Active Contours**, **Extension to save segmented objects** and **Text** (Boxes ranks are not the same
326 as in step 5.3.1).

327

328 5.3.2.2. In the box **Call IJ plugin** (box 4), set the **Rolling** parameter to control background
329 subtraction. Set this parameter to at least the size of the largest object that is not part of the
330 background. Decreasing this value increases background removal, but can also induce loss of
331 foreground signal.

332

333 5.3.2.3. In the box **Adaptive histogram equalization** (box 6), improve the contrasts between the
334 foreground objects and the background. Increasing the slope gives more contrasted sequences.

335

336 5.3.2.4. Run the workflow (for details, see part 7).

337

338 5.3.3. Segmentation D: Clustered objects with irregular shapes

339

340 NOTE: Three different methods of segmentation apply to each image: firstly, **HK-means**
341 **clustering** combined with the **Active Contours method** is applied. Then, the **classical watershed**
342 **algorithm** (using the Euclidian distance map) is applied to previously mis-segmented objects.
343 Finally, a **marker-based watershed algorithm** is used. Only HK-means and marker-based
344 watershed methods need user intervention. For both methods, the same parameters can be
345 applied for all images (fully automated version) or be changed for each image (semi-automated
346 version). If the user is not trained in these segmentation methods, the semi-automated
347 processing is highly recommended. During the processing of this block, manual intervention is
348 needed. When a segmentation method is finished, the user must manually remove mis-
349 segmented objects before the beginning of the next segmentation method. Successfully
350 segmented objects are saved and not taken into account in the next step(s). This block has to be
351 connected with the block **Clustered/heterogeneous shapes primary objects segmentation**
352 **Dialog Box** to work correctly.

353
354 5.3.3.1. Download the ImageJ collection MorphoLibJ on
355 **https://github.com/ijpb/MorphoLibJ/releases**. The MorphoLibJ 1.4.0 version is used in this
356 protocol. Place the file **MorphoLibJ_-1.4.0.jar** in the folder icy/ij/plugins. More information about
357 the content of this collection is available on **https://imagej.net/MorphoLibJ.**

358
359 5.3.3.2. Follow the same steps as in step 5.3.1 to set parameters of boxes **Channel signal**,
360 **Gaussian filter**, **Active Contours**, **Extension to save segmented objects** and **Text**. Boxes ranks are
361 not the same as in step 5.3.1.

362
363 5.3.3.3. Set parameters of the box **Adaptive histogram equalization** (see step 5.3.2.3).

364
365 5.3.3.4. To activate **Subtract Background**, write yes in **Apply Subtract Background?** (box 5). Else,
366 write **No**. If the plugin is activated, set the rolling parameter (see step 5.3.2), in the box **Subtract**
367 **Background parameter** (box 7).

368
369 5.3.3.5. **Automatization of HK-means:** To apply the same parameters for all images (fully
370 automated processing), set the **Nb of classes** (box 11), the **Minimum size** (box 12), and the
371 **Maximum size** (box13) (see step 5.3.1). These parameters must be set to select a maximum of
372 foreground pixels and to optimize the individualization of foreground objects. For the semi-
373 automated processing version, no intervention is needed.

374
375 5.3.3.6. **Automatization of marker extractions:** for the fully-automated version, expand the box
376 **Internal Markers extraction** (box 27) and set the value of the "dynamic" parameter in line 13 of
377 the script. For the semi-automated processing version, no intervention is needed.

378
379 NOTE: Markers are extracted by applying an extended-minima transformation on an input image
380 controlled by a "dynamic" parameter. In the marker-based watershed algorithm, flooding from
381 these markers is simulated to perform object segmentation. For the successful segmentation of
382 foreground objects, one single marker per foreground object should be extracted. The setting of
383 the "dynamic" parameter for optimal markers extraction depends mostly on the resolution of
384 images. Thus, if being not familiar with this parameter, use the semi-automated version.

385
386 5.3.3.7. Run the workflow (for details, see part 7).

387
388 5.3.3.8. At the beginning of the processing, dialog boxes **HK-means parameters** and **Marker-**
389 **based watershed** successively open. To apply the same parameters for all images (fully-
390 automated version), click on **YES**. Otherwise, click on **NO**. An information box opens, asking to
391 "Determine optimal ROIs with HK-Means plugin and close image". Click on **OK** and manually apply
392 the HK-Means plugin (**Detection and Tracking|HK-Means**) on the image, which automatically
393 opens. Select the **Export ROIs** option in the HK-Means plugin box. Apply the best parameters to
394 have ROIs containing a maximum of foreground pixels and to optimize the individualization of
395 foreground objects. When optimal ROIs are found, directly close the image.

396

397 5.3.3.9. At the end of the first segmentation method, an information box opens and asks to
398 "Remove unwanted ROIs and close image". These ROIs correspond to the borders of the
399 segmented objects. Select **OK** and remove ROIs of mis-segmented objects in the image, which
400 automatically opens. An ROI can be easily removed by placing the cursor on its border and using
401 the "Delete" button of the keyboard. Close the image. Repeat the same procedure after
402 completion of the second segmentation step.
403
404 5.3.3.10. At this stage, if the YES button was selected for the full-automatization of the marker-
405 based watershed algorithm, the previously set parameters will be applied to all images.
406
407 5.3.3.11. If the NO button was selected, an information box opens, asking to "Determine and
408 adjust internal markers". Click on **OK** and within the ImageJ interface of Icy, go to **Plugins |**
409 **MorphoLibJ | Minima** and **Maxima|Extended Min & Max**. In Operation, select **Extended**
410 **Minima**.
411
412 5.3.3.12. Select **Preview** to pre-visualize on the automatically opened image the result of the
413 transformation. Move the dynamic until optimal markers are observed. Markers are groups of
414 pixels with a value of 255 (Not necessarily white pixels). Optimal parameters lead to one marker
415 per object. Focus on the remaining objects that have not been well segmented with the two
416 previous segmentation methods.
417
418 5.3.3.13. If necessary, improve the markers by applying additional morphological operations like
419 "Opening" or "Closing" (**Plugins | MorphoLibJ | Morphological Filters**). When getting the final
420 image of markers, keep it open and close all the other images ending with the image initially used
421 as an input for the **Extended Minima** operation. Click on **No** if an ImageJ box asks to save changes
422 on this image.
423
424 5.3.3.14. In the box **Nb of images with Information Box** (box 14), determine how many images
425 with Information Boxes should appear.
426
427 5.3.4. Segmentation E: Clustered cytoplasm
428
429 NOTE: This block uses previously segmented nuclei as individual markers to initiate cytoplasm
430 segmentation. Be sure that the block of nuclei segmentation has processed before using it.
431
432 5.3.4.1. In the box **Channel cytoplasm** (box 1), set the channel of the cytoplasmic signal.
433
434 5.3.4.2. In the box **Extension segmented nuclei** (box 2), write the format used to save images of
435 segmented nuclei (tif, jpeg, bmp, png). The default format is tif.
436
437 5.3.4.3. In the box **Text** (box3), write the \Name of the folder containing segmented nuclei.
438
439 5.3.4.4. In the box **Format of images of segmented cytoplasms** (box 4), set the format to use for
440 saving segmented objects images (Tiff, Gif, Jpeg, BMP, PNG).

441

442 5.3.4.5. During the process, a folder is automatically created to save images of segmented
443 cytoplasms. In the box **Text** (box 5), name this folder (ex: Segmented cytoplasms). The folder is
444 created in the folder containing merged images.

445

446 5.3.4.6. Follow the same steps as in step 5.3.1 to set parameters of boxes **Gaussian filter** and
447 **Active Contours** (Be careful, box ranks are not the same as in step 5.3.1).

448

449 5.3.4.7. Run the workflow (for details, see part 7).

450

451 **6. Fluorescent signal detection and analysis**

452

453 6.1. Select the adapted block.

454

455 6.1.1. In the block **Fluorescence Analysis A: 1 Channel**, perform detection and analysis of foci in
456 one channel inside one type of segmented object: detection of coilin foci (red channel) within
457 the nucleus.

458

459 6.1.2. In the block **Fluorescence Analysis B: 2 Channels in the same compartment**, perform
460 detection and analysis of foci in two channels inside one type of segmented object: detection of
461 coilin (red channel) and 53BP1 (green channel) foci within the nucleus.

462

463 6.1.3. In the block **Fluorescence Analysis C: 2 Channels in two compartments**, perform detection
464 and analysis of foci in one or two channels, specifically inside the nuclei and their corresponding
465 cytoplasm: detection of Coilin foci (red channel) both within the nucleus and its corresponding
466 cytoplasm or detection of Coilin foci (red channel) within the nucleus and G3BP foci (green
467 channel) within the corresponding cytoplasm.

468

469 6.1.4. In the block **Fluorescence Analysis D: Global Translocation**, calculate the percentage of
470 signal from one channel in two cellular compartments (a and b). For example, in a
471 cytoplasm/nucleus translocation assay, export calculated percentages of nuclear and cytoplasmic
472 signals for each image in the final "Results" spreadsheet. The formula used to calculate the
473 nuclear signal percentage is shown below. This block can be used for any subcellular
474 compartment:

475 $$\% \, nuclear \, signal = \frac{Sum \, of \, all \, nuclei \, signals \, X \, 100}{(Sum \, of \, all \, nuclei \, signals + Sum \, of \, all \, cytoplasmic \, signals \,)}$$

476

477 6.1.5. In the block **Fluorescence Analysis E: Individual Cell Translocation**, calculate the
478 percentage of signal from one channel in two cellular compartments for each cell. This block is
479 specially optimized for nucleus/cytoplasm translocation assay at the single-cell level.

480

481 NOTE: Because the block **Fluorescence Analysis E: Individual Cell Translocation** performs
482 analysis at the single-cell level, efficient segmentation of nucleus and cytoplasm is needed.

483

484   6.2. Link the **output0** (File) of the block **Select Folder** (block 1) to the **folder** input (white arrows
485   in black circles) of the chosen block.
486
487   6.3. Set the parameters of the chosen block.
488
489   6.3.1. Fluorescence Analysis A: 1 Channel, Fluorescence Analysis B: 2 Channels in the same
490   compartment and Fluorescence Analysis C: 2 Channels in two compartments
491
492   6.3.1.1. In the box **Folder images ROI**, write the name of the folder containing images of
493   segmented objects preceded by a backslash. (For example: \Segmented nuclei).
494
495   6.3.1.2. In the box **Format of images of segmented objects** (box 2), write the format used to save
496   images of segmented objects (tif, jpeg, bmp, png). The default format is tif.
497
498   6.3.1.3. In the box **Kill Borders?**, write **Yes** to remove border objects. Otherwise, write **No**. The
499   installation of the MorphoLibJ collection of ImageJ is required to use this function (see step 5.3.3).
500
501   6.3.1.4. In the box(es) **Channel spots signal**, set the channel where spots have to be detected. In
502   classical RGB images, 0=Red, 1=Green, and 2=Blue.
503
504   6.3.1.5. In the boxes **Name of localized molecule**, write the name of the molecule localizing into
505   the spots. The number of fields to enter depends on the number of molecules.
506
507   6.3.1.6. In the box(es) **Wavelet Spot Detector Block**, set spot detection parameters for each
508   channel. Set the scale(s) (referred to spot size), and the sensitivity of the detection (a smaller
509   sensitivity decreases the number of detected spots, the default value being 100 and the minimum
510   value being 0). Exhaustive documentation of this plugin is available online:
511   **http://icy.bioimageanalysis.org/plugin/Spot_Detector**. Parameters can also be manually
512   defined in Icy (**Detection and Tracking | Spot Detector**).
513
514   6.3.1.7. As an option, in the box **Filter ROI by size**, filter the segmented objects where spots are
515   detected by setting an interval of size (in pixels). This step is especially useful to remove under-
516   or over-segmented objects. To manually estimate the size of objects, see step 5.3.1. Default
517   parameters do not include the filtering of ROIs by size. The block 2 Channels in two compartments
518   contains two boxes: Filter nuclei by size (box 19) and Filter cytoplasm by size (box 46).
519
520   6.3.1.8. Optionally, in the box **Filter spots by size**, filter the detected spots according to their size
521   (in pixels) to remove unwanted artifacts. To manually estimate spot size, click on **Detection and**
522   **Tracking** and open the **Spot detector** plugin. In **Output** options, select **Export to ROI**. Be careful
523   that default parameters do not include filtering of spots by size, and that filtered spots are not
524   taken into account for the analysis. The number of fields to enter depends on the number of
525   channels.
526
527   6.3.1.9. Optionally, in the box **Filter spots**, apply an additional filter (contrast, homogeneity,

528 perimeter, roundness) on the detected spots. Be careful that default parameters do not include
529 spot filtering and that filtered spots are not taken into account for the analysis. The number of
530 fields to enter depends on the number of channels.
531
532 6.3.1.10. Optionally, in the boxes **Spot size threshold**, set a threshold for the area (in pixels) of
533 analyzed spots. The number of counted spots below and above this threshold is exported in the
534 final **Results** spreadsheet. The number of boxes to be informed depends on the number of
535 channels.
536
537 6.3.1.11. Run the workflow (for details, see part 7). Data are exported in a **Results** spreadsheet
538 saved in the folder containing merged images.
539
540 6.3.2. Fluorescence Analysis D: Global Translocation and Fluorescence Analysis E: Individual Cell
541 Translocation:
542
543 6.3.2.1. In the boxes **Folder images** (boxes 1 and 2), write the \Name of the folder containing
544 images of segmented objects. In the block **Fluorescence Analysis D: Global Translocation**, the
545 two types of ROI are identified as ROI a and ROI b. For the block **Fluorescence Analysis E:**
546 **Individual Cell Translocation**, in **Folder images segmented nuclei** and **Folder images segmented**
547 **cytoplasms** boxes, write the name of the folder containing segmented nuclei and cytoplasms,
548 respectively.
549
550 6.3.2.2. In the box **Channel signal** (box 3), enter the channel of the signal.
551
552 6.3.2.3. In the box **Format of images of segmented objects** (box 4), write the format used to save
553 images of segmented objects (tif, jpeg, bmp, png). The default format is tif. The **Kill Borders**
554 option is also available to remove border objects (see step 6.3.1).
555
556 6.3.2.4. Optionally, in boxes **Filter ROI by size**, filter the segmented objects by setting an interval
557 of size (in pixels). This step could be useful to remove under- or over-segmented objects. To
558 manually estimate object size, see step 5.3.1. There are two fields to enter, one per channel.
559 Default parameters do not include ROI filtering by size.
560
561 6.3.2.5. Run the workflow (for details, see part 7). Export data in a spreadsheet **Results** saved in
562 the folder containing merged images.
563
564 **7. Run the protocol**
565
566 7.1. To process one block in a run, remove the connection between the selected block and the
567 block **Select Folder**. Place the wanted block at the 1st rank. On the upper left corner of the wanted
568 block, click on the link directly to the right of **folder**. In the **Open** dialog box which appears,
569 double-click on the folder containing the merged images. Then, click on **Open**. Click on **Run** to
570 start the workflow. The processing can be stopped by clicking on the **Stop** button.
571

572   7.2. To process different blocks in a run, keep connections of chosen blocks with the block **Select**
573   **Folder** (block 1). Make sure that their rank allows the good processing of the workflow. For
574   example, if a specific block needs segmented objects to process, be sure that the segmentation
575   block processes before. Before running the workflow, remove unused blocks and save the new
576   protocol with another name.
577
578   7.3. Click on **Run** to start the workflow. When the open dialog box appears, double-click on the
579   folder containing the merged images. Then, click on **Open**. The workflow automatically runs. If
580   necessary, stop the processing by clicking on the **Stop** button.
581
582   7.4. At the end of the processing, check that the message **The workflow executed successfully**
583   appeared in the lower right-hand corner and that all the blocks are flagged with a green sign
584   (**Figure 2b**). If not, the block and the inside box presenting the error sign indicate the element to
585   correct (**Figure 2b**).
586
587   NOTE: After the workflow executed successfully, a new run cannot be directly started, and to
588   process the workflow again, at least one block should be flagged with the sign "ready to process".
589   To change the state of a block, either delete and re-create a link between two boxes inside this
590   block or simply close and re-open the protocol. If an error occurs during the processing, a new
591   run can directly be started. During a new run, all the blocks of the pipeline are processed, even if
592   some of them are flagged with the green sign.
593
594   **REPRESENTATIVE RESULTS:**
595   All the described analyses have been performed on a standard laptop (64-bit, quad-core
596   processor at 2.80 GHz with 16 GB RAM) working with the 64-bit version of Java. Random-access
597   memory is an important parameter to consider, depending on the amount and the resolution of
598   images to analyze. Using the 32-bit version of Java limits the memory to about 1300 MB, which
599   could be unsuitable for big data analysis, whereas the 64-bit version allows increasing the
600   memory allocated to Icy. **Figure 3** reports the time needed for segmentation for different types
601   of images and different resolutions. It confirms that high resolution substantially increases the
602   time of primary object segmentation.
603
604   The data presented in the next paragraphs demonstrate that the *Substructure Analyzer* workflow
605   can be used to solve most of the common problematics encounter in cellular and molecular
606   biology (cell counting, foci counting and analysis, translocation analysis).
607
608   Measurement in numerous cells of the precise proportion of molecules concentrated within a
609   given compartment or the change of localization across subcellular domains can be a very
610   cumbersome and error-prone task, especially when it is performed manually. **Figure 4** illustrates
611   the ability of the workflow to rapidly and precisely quantify the well-described nuclear
612   translocation of the transcription factor NFκB in response to TNFα stimulation. Images used for
613   the analysis were graciously compiled by Ilya Ravkin (http://www.ravkin.net/SBS/Image-
614   Library.htm) and are publicly available at the Broad Bioimage Benchmark Collection[16]. This set
615   contains images of MCF7 and A549 cell lines treated with increasing concentrations of TNFα. The

616   analysis has been performed on more than 40,000 cells and 96 images (48 images for each
617   channel) from the MCF7 cell line (**Figure 4a**). The whole analysis (from the import of images to
618   data saving) took 26 minutes. Each image corresponds to a specific TNFα concentration. The
619   nuclear/cytoplasmic proportions of NFκB were homogeneous across all cells for a given image.
620   For this reason, for each image, the values of all nuclear or cytoplasmic pixels were summed to
621   calculate the nuclear and cytoplasmic proportions of NFκB, and individualization of touching
622   nuclei/cytoplasms during their segmentation was not required. DAPI images were processed in
623   the **Segmentation C: Clustered objects with convex shapes** block to segment nuclei and the
624   green channel was used to delineate the cytoplasms with the **Segmentation E: Clustered**
625   **cytoplasm** block (**Figure 4b**). The **Fluorescence Analysis D: Global Translocation** block was used
626   to export the sum of nuclear and cytoplasmic pixel values. The obtained data are represented in
627   a dose-response curve (**Figure 4c**) and illustrate the increase of NFκB nuclear proportion after
628   stimulation with increasing concentrations of TNFα.

630   The workflow not only analyzes the global intensity signal in different sub-cellular compartments
631   but can also be used to detect foci and extract specific information about their features. **Figure**
632   **5a** illustrates the detection of P-bodies in individual cells by localizing the enhancer of mRNA
633   decapping 4 (EDC4) protein. The **Segmentation A: Non clustered objects** block was used to
634   segment nuclei that present a low level of clustering, the **Segmentation E: Clustered cytoplasm**
635   block to delineate the cytoplasms and the **Fluorescence Analysis C: 2 channels in two**
636   **compartments** to detect EDC4 foci. Segmented objects are identified (Segmented nuclei_0,
637   Segmented cytoplasms_0), and multiple information is collected into the corresponding
638   worksheet (Segmented nuclei Information and Segmented cytoplasms Information) like the EDC4
639   mean intensity or the number/size of detected bodies in each ROI (**Figure 5b**). Before being
640   analyzed, detected bodies can be preliminarily filtered according to their area, which can be
641   useful to exclude objects below the resolution power of the microscope. To estimate the size of
642   conserved objects, an additional area threshold can be introduced. Areas of all the detected
643   bodies are reported in dedicated worksheets (Segmented nuclei_Distribution foci size,
644   Segmented cytoplasms_Distribution foci size). In this example, we highlight the analysis of two
645   cells (Segmented nuclei_0 and _1 and Segmented cytoplasm_0 and _1), where cytoplasmic EDC4
646   foci are detected. Note that even though the signal of the EDC4 protein is detected in both the
647   nuclear and cytoplasmic compartments, only the cytoplasmic foci correspond to P-bodies. The
648   nuclear signal most likely results from the cross-reactivity of the antibody used to perform
649   immunofluorescence experiments. The size in pixels of each of the foci is given.

651   Then, as the effect of oxidative stress (OS) on Cajal Bodies was still unknown, we took advantage
652   of the versatility of the workflow to study it during time kinetics (2 to 20 hours after OS induction
653   with 500 µM H_2O_2) (**Figure 6**). Cajal Bodies are dynamic structures nucleating and dissolving
654   within the nucleoplasm under the control of specific parameters like the transcription rate or cell
655   cycle progression. Cajal Bodies were visualized by localizing their main structural and functional
656   component, the coilin protein. Information about the number and the size of Cajal bodies were
657   collected by studying 2300 individual cells from high-resolution images (3840X3072) containing
658   3 channels: DAPI (blue), 53BP1 (green) and coilin (red) (**Figure 6a**). Full processing took less than
659   1 h. Because the nuclei presented convex shape and some of them were clustered, the block

Segmentation C: Clustered objects with convex shapes was chosen to perform the segmentation. As OS is known to induce DNA double-strand breaks (DDSBs), the p53-binding protein 1 (53BP1), known to be an essential effector of DNA damage signaling pathways, was used as a marker to discriminate between stressed and non-stressed cells. Indeed, in the absence of DNA damages, 53BP1 is homogeneously distributed within the nucleoplasm, whereas following DNA damages, it concentrates on DDSBs, which form easily distinguishable foci in microscopy. The number and the size of both 53BP1 and coilin nuclear foci in each cell were analyzed by using the block **Fluorescence Analysis B: 2 Channels in the same compartment**. Firstly, the analysis of 53BP1 data helped to determine for each time point of the kinetics the best threshold to classify cells according to their stress status. Among the whole kinetics, OS induces a significant increase of the 53BP1 foci number with an 11-fold increase at 2, 4, and 8 h after the induction of OS (**Figure S1**). This effect is less pronounced at 6 h (6.5-fold) because of a higher initial level of DSB in non-stressed cells. Even after 20 h, a sustained increase remains (almost 6-fold). These data reflect the efficiency of the treatment to induce the OS and establish 53BP1 as an efficient stress-marker for both early and late observations in microscopy. On the other hand, a small proportion of stressed-cells with a low level of DDSBs was observed at each time point of the kinetics, suggesting that cells are not homogeneously stressed, enforcing the importance of using a stress marker in stress experiments. Based on the ROC analysis, an overall threshold of 17 53BP1 foci was selected to discriminate between stressed and unstressed cells throughout the kinetics (**Figure S1**).

Then the number and size of nuclear foci of coilin were analyzed according to the cell's stress status. A significant increase in the number of foci was observed 2, 4, and 6 h after stress induction (**Figure 6b**). The most sustained effect is observed at 2 h, the number of cells having more than 10 foci increasing from 0% in the unstressed cells to 75% in the stressed cells. Moreover, more than 25% of stressed cells of that time point had more than 20 coilin foci. This effect progressively decreases until 8 h. At 20 h, a small increase of the first and the third quartiles is observed, but data also show that 84% and 80% of cells present less than 8 foci in non-stressed and stressed cells, respectively. These data show that OS also has late effects on parameters controlling the nucleation of Cajal Bodies, which might be different from early effects. Interestingly, careful analysis of the features of the coilin nuclear foci revealed that the observed increase of the number of coilin foci associates with a decrease in their size (**Figure 6c**). For instance, 2 h after OS induction, the proportion of coilin foci with an area below 0.2 μm^2 increases from 26% to 64%. This effect decreases until 8 h, but, unlike Cajal Bodies' number, no significant change is observed at 20 h. This might reflect that early and late nucleoplasmic redistribution of Cajal Bodies are different events.

Altogether these data strongly suggest that OS changes the nucleation power of Cajal Bodies, inducing a nucleoplasmic redistribution into numerous smaller nuclear foci. Since nucleation of Cajal Bodies is driven by their protein and RNA components, this suggests that the OS effect might change the composition of Cajal Bodies. Within Cajal Bodies, the coilin protein interacts with several different protein partners such as components of the SMN complex and Sm proteins. These interactions often involved phosphodomains of coilin, and one can imagine that modification of the structure of Cajal Bodies could be linked to changes in the phosphorylation

704 status of coilin. Moreover, recent works demonstrated that Cajal Bodies are not randomly
705 localized within the nucleus, and can affect gene expression through proximal association with
706 specific gene loci[17]. Given all these facts, it would not be surprising if an effect on Cajal bodies'
707 functionality would accompany changes in their structure. From these results, we can also ask if
708 such Cajal Bodies' remodeling is a passive consequence of OS or participates in the response by
709 interacting with specific gene loci, for example.
710
711 To test if a change in coilin expression could alter its localization and change the nucleation of
712 Cajal Bodies, we overexpressed an exogenous GFP-coilin fusion protein. Nuclei were not
713 clustered, so segmentation was performed with the block **Segmentation A: Non clustered**
714 **objects**. Then, to precisely quantify the number and size of the coilin foci (red signal, channel 1)
715 according to the level of GFP-coilin (green signal, channel 2) overexpression, the block
716 **Fluorescence Analysis B: 2 Channels in the same compartment** was used. The level of GFP-coilin
717 was reflected by the mean intensity of the GFP signal in individual nucleus (**Figure 7a**). In cells
718 with medium or high level of GFP-coilin overexpression (GFP intensity higher than 10), the
719 number of Cajal Bodies per nucleus significantly increases with some nucleus displaying until 21
720 foci (**Figure 7b**). We also noticed that the brightness (mean intensity) of Cajal Bodies considerably
721 increases in parallel with the overexpression level leading to intensity saturation. Such saturated
722 regions might thus mask the presence of smaller and fainter foci. In cells overexpressing a
723 medium or high level of GFP-coilin, the size of Cajal Bodies also significantly increases, the median
724 values rising from 1 to 2 μm² (**Figure 7c**). Moreover, more than 25% of cells with a high GFP-coilin
725 expression level present foci with an area larger than 2.5 μm², whereas it never exceeds 0.8 μm²
726 in non-transfected cells. In conclusion, GFP-Coilin overexpression leads to a significant increase
727 in both the number and size of Cajal Bodies. Since OS increases the number of Cajal Bodies but
728 reduces their size, these data might reflect that OS effect on the structure of Cajal Bodies is most
729 probably induced by a change of their composition rather than by an effect on the cellular
730 amount of coilin.
731
732 **FIGURE AND TABLE LEGENDS:**
733 **Figure 1: The *Substructure Analyzer* pipeline**
734 The workflow is developed in Icy, an open community platform for bioimage informatics, and
735 performs automatic analysis of multiple fluorescence microscopy images. Firstly, multi-channel
736 images are automatically loaded within the protocol and pre-processed to improve, if necessary,
737 the signal to noise ratio and to remove imaging artifacts. Then, image segmentation isolates
738 regions of interest (ROIs) from the background. Several methods of segmentation are available
739 depending on the level of clustering and the nature of the objects of interest. Segmented objects
740 are saved with a specific descriptor (for example, Image name_Nucleus_1) in a specific folder and
741 can be used/re-used for subsequent analysis. Fluorescent signals like spots are then analyzed
742 within the ROIs and multiple features (location, size, shape, intensity, texture, spot number, and
743 size) are exported into an automatically created spreadsheet. All the measured features, like the
744 number of detected spots, are reported to the descriptor of the corresponding ROI.
745
746 **Figure 2: Graphical interface of the Icy workflow**

(**a**) The workflow is composed of 13 general blocks, which can be grouped according to their general function: **Select Folder** (block 1) allow access to images; **Merge Channels** (block 2) is used to merge different channels of images. Blocks 3 to 8 are dedicated to object segmentation, each one being adapted for a specific context: **Segmentation A: Non clustered objects**, **Segmentation B: Poorly clustered objects**, **Segmentation C: Clustered objects with convex shapes**, **Segmentation D: Clustered objects with irregular shapes** (block 7, works in association with block 6), **Segmentation E: Clustered cytoplasms**. Blocks 9 to 11 are used to count/analyze foci within subcellular compartments: **Fluorescence Analysis A: 1 Channel**, **Fluorescence Analysis B: 2 Channels in the same compartment** and **Fluorescence Analysis C: 2 Channels in two compartments**. Blocks 12 and 13 are adapted for the analysis of translocation events: **Fluorescence Analysis D: Global Translocation** and **Fluorescence Analysis E: Individual Cell Translocation**. (**b**) Global presentation of a block architecture.

Figure 3: Time needed for segmentation as a function of image resolution and object clustering level
(**a**) Graphical illustration of the time needed to perform segmentation depending on image resolution and clustering level. Two different segmentation blocks have been tested: segmentation A (for non-clustered objects) and segmentation D (for clustered objects with irregular shapes). The time of processing (in seconds per image) is plotted as a function of image resolution (number of pixels). (**b**) Example of images analyzed with either segmentation A or segmentation D blocks. DAPI staining and segmented nuclei are indicated.

Figure 4: Workflow analysis of an NFκB nuclear translocation assay in response to TNFα stimulation
(**a**) Example of images of human MCF7 cells (human breast adenocarcinoma cell line) treated with TNFα from the BBBC014 image set (Broad Bioimage Benchmark Collection)[12]. Localization of the transcription factor NFκB has been studied in cells treated with 12 increasing concentrations of TNFα (from 0 to 10^{-7}g/mL). For each concentration, 4 replicates were done. DAPI images (blue channel) were processed in the Segmentation C: Clustered objects with convex shapes block to segment nuclei, and then NFκB staining with FITC (green channel) was used to delineate the cytoplasms with the Segmentation E: Clustered cytoplasm block. The Fluorescence Analysis D: Global Translocation block was used to export the sum of nuclear and cytoplasmic pixel values. (**b**) Segmented nuclei (grey) and cytoplasms (white) are used to determine the total intensity of the NFκB fluorescent signal in each compartment. (**c**) The dose-response curve obtained from data analysis illustrates the increase of NFκB nuclear proportion as TNFα concentration increases. Images are in an 8-bit BMP format, and the image size is 1360 x 1024 pixels. For each concentration, the standard deviation between the 4 replicates is calculated and illustrated as error bars.

Figure 5: Quantification of EDC4 foci (P-bodies) within the cytoplasm in individual cells
(**a**) Co-staining with DAPI (blue channel) and Phalloidin (red channel) allows the segmentation and identification of cell nuclei and F-actin, respectively. The EDC4 protein (green channel) is used as a marker for the detection of P-bodies known to be exclusively cytoplasmic. Scale bar, 10 µm.

790   (**b**) Foci are counted, and several features (here their area in pixels) are exported to the *Results*
791   spreadsheet structured in several sheets. Two sheets are dedicated to the analysis of nuclei and
792   two to the analysis of cytoplasms. Each raw embed the information of a specific ROI. In
793   *Nucleus/Cytoplasm Information* sheets, the characteristics of segmented objects are exported
794   (ROI ID, size, and mean intensity), followed by the number of detected foci within each ROI. If
795   necessary, a size threshold (100 pixels in this example) can be added, and the number of foci
796   below and above this threshold is reported in the two last columns. In
797   *Nucleus/Cytoplasm_Distribution of foci size* sheets, the area (in pixels) of all the detected foci is
798   reported in the raw of the corresponding ROI. In this example, we highlight the analysis of two
799   cells (Segmented cytoplasm_0 and_1), where cytoplasmic EDC4 foci have been detected. The size
800   in pixels of each of the foci is also given.

801

802   **Figure 6: Kinetics of coilin and 53BP1 nucleoplasmic distribution after oxidative stress**
803   The analysis of the number and size of foci was based on more than 110 individual cells for each
804   time point from 3 independent kinetics. (**a**) Immunofluorescence detection of coilin and 53BP1
805   foci in HeLa S3 cells throughout treatment kinetics with 500 µM H_2O_2. 53BP1 concentrates on
806   DNA Double-Strand Break sites and is used to assess the efficiency of the treatment on each cell.
807   Coilin is enriched in Cajal Bodies. Cells were fixed 2, 4, 6, 8, and 20 h after oxidative stress
808   induction and co-stained with anti-coilin (red channel) and anti-53BP1 (green channel)
809   antibodies. Scale bar, 10 µm. (**b**) The number of coilin nuclear foci after oxidative stress. Results
810   are shown as box plots, where the central mark is the median, the lower and upper edge of the
811   box are the 25th and 75th percentiles. In cells treated with H_2O_2, a significant increase inthe
812   number of coilin foci is detected and is maximal after 2 and 4 h of treatment. Wilcoxon - Mann
813   Whitney test analysis demonstrates a significant difference between untreated and H_2O_2 treated
814   cells (*, $p<0.05$; **, $p<0.01$; ***, $p<0.001$). (**c**) The proportion of cells in function of the coilin foci
815   area (μm^2). The majority of foci have a smaller area at 2 h and 4 h time points as compared to
816   the control.

817

818   **Figure 7: Study of Cajal Bodies' nucleation in cells overexpressing a GFP-coilin fusion protein**
819   (**a**) HeLa S3 cells were transiently transfected in biological triplicates with 500 ng of a plasmid
820   expressing a GFP-Coilin fusion protein using a standard procedure following the manufacturer's
821   recommendations. After 48 h, cells were fixed and stained with an antibody against coilin. DAPI
822   (blue channel) permits to segment nuclei. For these analyses, more than 100 cells were analyzed.
823   GFP signal reflects ectopic coilin-GFP protein (green channel) localization, whereas the coilin
824   signal (red channel) corresponds to both endogenous and exogenous protein localization. For
825   each nucleus, the mean intensity of the GFP signal reflects the level of coilin expression.
826   According to this parameter, the features of Coilin foci have been analyzed using the block
827   **Fluorescence Analysis B: 2 Channels in the same compartment**. Scale bar, 10 µm. (**b**)
828   Relationship between the number of coilin foci and mean GFP intensity per nucleus. Results are
829   shown as box plots, where the central mark is the median, the lower and upper edge of the box
830   are the 25th and 75th percentiles. Wilcoxon - Mann Whitney test analysis demonstrates a
831   significant increase of the number of foci for a GFP intensity >10 (*, $p<0.05$; **, $p<0.01$; ***,
832    $p<0.001$). (**c**) Relationship between the area of coilin foci and mean GFP intensity per nucleus.
833   Results are shown as box plots, where the central mark is the median, the lower and upper edge

834   of the box are the 25th and 75th percentiles. Wilcoxon - Mann Whitney test analysis demonstrates
835   a significant increase of the area of coilin foci for a GFP intensity >10 (*, p<0.05; **, p<0.01; ***,
836   p<0.001).

837

838   **Figure S1: Kinetics of 53BP1 nucleoplasmic distribution after oxidative stress**
839   (**a**) Quantification of 53BP1 nuclear foci in untreated or H_2O_2 treated cells after different
840   incubation times. Results are shown as box plots, where the central mark is the median, the lower
841   and upper edge of the box are the 25th and 75th percentiles. Wilcoxon - Mann Whitney test
842   analysis demonstrates a significant difference between untreated and H_2O_2 treated cells (*,
843   p<0.05; **, p<0.01; ***, p<0.001). (**b**) For each time point of the kinetics, a ROC (Receiver
844   Operating Characteristic) test has been done to determine the best threshold to discriminate
845   between stressed and unstressed cells. The parameters of each test are reported in the table
846   (Sensitivity, specificity, TP: true positive, TN: true negative, FP: false positive, FN: false negative).
847   From these data, a global threshold of 17 53BP1 foci was determined to discriminate stressed
848   from non-stressed cells through the kinetics.

849

850   **DISCUSSION:**
851   An increasing number of free software tools are available for the analysis of fluorescence cell
852   images. Users must correctly choose the adequate software according to the complexity of their
853   problematic, to their knowledge in image processing, and to the time they want to spend in their
854   analysis. Icy, CellProfiler, or ImageJ/Fiji are powerful tools combining both usability and
855   functionality[3]. Icy is a stand-alone tool that presents a clear graphical user interface (GUI), and
856   notably its "Protocols" point and click interface through which workflows can be easily designed
857   or manipulated[4]. The functionality of this software is enhanced by supporting and utilizing ImageJ
858   plug-ins, and a lot of documentation is available thanks to its large community of users. We used
859   this strong combination of usability and functionality of the Icy software to develop the
860   *Substructure Analyzer* workflow. Its main objective is to propose an automated solution to
861   perform a full analysis of cell fluorescent signal from multiple images. The analysis encompasses
862   image pre-processing, object segmentation, and fluorescent signal analysis.

863

864   Several protocols are gratefully shared on the Icy website and propose to use Icy functionalities
865   to detect and analyze fluorescent signals like cellular foci. However, some of these protocols only
866   process one image per run, do not export the results in a specific file or analyze a single channel.
867   Others do not contain preliminary segmentation to determine regions of interest (ROIs) or
868   propose simple algorithms like "HK-means" that are not suited for the segmentation of highly
869   clustered objects. *Substructure Analyzer* automates all the steps of the image analysis from image
870   pre-processing to the segmentation of objects and the detection/analysis of fluorescent signals.
871   The protocol also proposes simple or more complex methods for object segmentation and export
872   of the data (names of the segmented objects, foci analysis) is realized in optimized outputs
873   adapted to the problematic. CellProfiler also proposes powerful pipelines composed of modules
874   that encapsulate functionalities[18,19]. The available pipelines are well adapted to specific
875   problems. Icy workflows are a network of different boxes, each of them performing a specific
876   task. Through the "Protocols" interface, boxes composing the network are intuitive and easy to
877   manipulate. *Substructure Analyzer* is adapted to several contexts like the simple merging of

878   channels, the quantification of fluorescent signals distribution between two subcellular
879   compartments, the analysis of foci in one or several channels from one or two cellular
880   compartments in individual cells. Several displays allow visualizing the intermediate results
881   during each run to control the processing. Moreover, Icy presents icon bars that group the
882   methods by topic and from which functionalities found in the workflow can be separately
883   manipulated to manually test specific parameters on some images before using them on a whole
884   image set.

886   Like in the bioimage analysis field, the most critical step of this protocol is the object
887   segmentation. The protocol proposes the segmentation of primary objects, which are most of
888   the time cell nuclei, and also contains another block which task is to segment a second type of
889   objects within the cell. Nuclei segmentation can be difficult when the objects are highly clustered
890   so that they touch or overlap together. Different alternatives for the segmentation of primary
891   objects are contained in the protocol, according to the shape of the objects and the level of
892   clustering even if to date no universal algorithm success in the full segmentation of highly
893   clustered objects. The segmentation process is mainly limited by the determination of correct
894   parameters that could successfully segment a part of the objects but lead to under- or over-
895   segmentation of another part. The user would have to perform different runs with different
896   parameters to get a final correct segmentation, especially for clustered objects presenting
897   heterogeneous and non-convex shapes.

899   This protocol is mostly limited by the segmentation step, which could be highly time consuming
900   and need powerful computer configuration for the most complex cases. More specifically, the
901   more the images are complex (clustered objects to segment, a high number of pixels in the
902   images) or numerous, the more the user will require random-access memory (RAM). Globally,
903   for any problematic, the user will need to run the protocol on a 64 bits Java Runtime Environment
904   (freely available on Java website) and to use a 64 bits operating system (OS) to increase the
905   available memory used by Icy (Memory is limited to 1300 MB for 32 bits JRE). Because of the
906   script, the protocol does not work correctly on a Mac computer. Moreover, the number of
907   parameters increases with the complexity of the problem and will require more knowledge in
908   image analysis. Concerning the functionality criteria, this protocol has not yet been adapted for
909   the analysis of stacked images (time stacks, z stacks) even if Icy efficiently performs analysis on
910   this type of data.

912   Future updates will be realized notably to improve the segmentation of complex clustered
913   objects. Additional blocks will be developed to adapt the segmentation to clustered cellular
914   structures with irregular non-convex shapes. We will also adapt the workflow for the analysis of
915   time and z-stacks (for live imaging and 3D data).

DISCLOSURES:

The authors have nothing to disclose.

REFERENCES:

1. Möckl, L., Lamb, D.C., Bräuchle, C. Super-resolved fluorescence microscopy: Nobel Prize in Chemistry 2014 for Eric Betzig, Stefan Hell, and William E. Moerner. *Angewandte Chemie*. **53** (51), 13972-13977 (2014).

2. Meijering, E., Carpenter, A.E., Peng, H., Hamprecht, F.A., Olivo-Marin, J.-C. Imagining the future of bioimage analysis. *Nature Biotechnology*. **34** (12), 1250–1255 (2016).

3. Wiesmann, V., Franz, D., Held, C., Münzenmayer, C., Palmisano, R., Wittenberg, T. Review of free software tools for image analysis of fluorescence cell micrographs. *Journal of Microscopy*. **257** (1), 39–53 (2015).

4. de Chaumont, F. et al. Icy: an open bioimage informatics platform for extended reproducible research. *Nature Methods*. **9** (7), 690–696 (2012).

5. Girish, V., Vijayalakshmi, A. Affordable image analysis using NIH Image/ImageJ. *Indian J Cancer*. **41**(1):47 (2004).

5. Zaitoun, N. M., Aqel, M. J. Survey on image segmentation techniques. *Procedia Computer Science*. **65**, 797-806 (2015).

6. Devices, M. MetaMorph Microscopy Automation and Image Analysis Software. https://www.moleculardevices.com/products/cellular-imaging-systems/acquisition-and-analysis-software/metamorph-microscopy (2018).

7. Nikon. NIS-Elements Imaging Software. https://www.nikon.com/products/microscope-solutions/lineup/img_soft/nis-element (2014).

8. Meyer, F., Beucher, S. Morphological segmentation. *Journal of Visual Communication and Image Representation*. **1** (1), 21–46 (1990).

9. Schieber, M., Chandel, N.S. ROS Function in Redox Signaling and Oxidative Stress. *Current Biology*. **24** (10), R453–R462 (2014).

10. D'Autréaux, B., Toledano, M.B. ROS as signalling molecules: mechanisms that generate specificity in ROS homeostasis. Nature Reviews. *Molecular Cell Biology*. **8** (10), 813–824 (2007).

11. Davalli, P., Mitic, T., Caporali, A., Lauriola, A., D'Arca, D. ROS, Cell Senescence, and Novel Molecular Mechanisms in Aging and Age-Related Diseases. *Oxidative Medicine and Cellular Longevity*. **2016**, 3565127 (2016).

12. Disher, K., Skandalis, A. Evidence of the modulation of mRNA splicing fidelity in humans by oxidative stress and p53. *Genome*. **50** (10), 946–953 (2007).

13. Takeo, K. et al. Oxidative stress-induced alternative splicing of transformer 2β (SFRS10) and CD44 pre-mRNAs in gastric epithelial cells. *American Journal of Physiology - Cell Physiology*. **297** (2), C330–C338 (2009).

14. Seo, J. et al. Oxidative Stress Triggers Body-Wide Skipping of Multiple Exons of the Spinal Muscular Atrophy Gene. *PLOS ONE*. **11** (4), e0154390 (2016).

15. Will, C.L., Luhrmann, R. Spliceosome Structure and Function. *Cold Spring Harbor Perspectives in Biology*. **3** (7), a003707–a003707 (2011).

966   16. Ljosa, V., Sokolnicki, K.L., Carpenter, A.E. Annotated high-throughput microscopy image sets
967   for validation. *Nature Methods*. **9** (7), 637–637 (2012).
968   17. Wang, Q. et al. Cajal bodies are linked to genome conformation. *Nature Communications*. **7**
969   (2016).
970   18. Carpenter AE. et al. CellProfiler: image analysis software for identifying and quantifying cell
971   phenotypes. *Genome Biology.* **7**, R100 (2006).
972   19. McQuin C. et al. CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biology*.
973   **16** (7), e2005970 (2018).

Figure 1

FIGURE 1

Figure 2   Click here to access/download;Figure;Figure-2.psd ⬇

FIGURE 2

(a)

Block 1: Access to images

Block 2: Merge different channel images

Blocks 9-11: Detection and analysis of foci

Blocks 3-8: Objects segmentation

Blocks 12-13: Analysis of translocation events

(b)

- Collapse
- ✓ Remember results
- ✗ Remove block
- ✗ Remove block but keep contents
- ☐ Stop on first error

Block's rank

Select Folder

Processing state

- ⚠ Block is ready to process
- ⚡ Block is currently processing
- ✓ Block has successfully processed
- ✗ Error/problem during the processing

Block's input

Block's output

Figure 3

FIGURE 3

(a)

(b)

Figure 4

FIGURE 4

Figure 5

Click here to access/download;Figure;Figure-5-new(1).psd ⬇

FIGURE 5

(a)

(b)

Segmented ROIs

nuclei

cytoplasms

FIGURE 6

(a)

(b)

(c)

Figure 7

Click here to access/download;Figure;Figure-7.psd ⬇

FIGURE 7

Name of Material/ Equipment	Company	Catalog Number
16% Formaldehyde solution (w/v) methanol free	Thermo Fisher Scientific	28908
Alexa Fluor 488 of goat anti-rabbit	Thermo Fisher Scientific	A-11008
Alexa Fluor 555 of goat anti-mouse	Thermo Fisher Scientific	A-21425
Alexa Fluor 555 Phalloidin	Thermo Fisher Scientific	A34055
Bovine serum albumin standard (BSA)	euromedex	04-100-812-E
DMEM	Sigma-Aldrich	D5796-500ml
Duolink In Situ Mounting Medium with DAPI	Sigma-Aldrich	DUO82040-5ML
Human: HeLa S3 cells	IGBMC, Strasbourg, France	
Hydrogen peroxide solution 30% (H_2O_2)	Sigma-Aldrich	H1009-100ml
Lipofectamine 2000 Reagent	Thermo Fisher Scientific	11668-019
Mouse monoclonal anti-coilin	abcam	ab11822
Nikon Optiphot-2 fluorescence microscope	Nikon	
Opti-MEM I Reduced Serum Medium	Thermo Fisher Scientific	31985062
PBS pH 7.4 (10x)	gibco	70011-036
Rabbit polyclonal anti-53BP1	Thermo Fisher Scientific	PA1-16565
Rabbit polyclonal anti-EDC4	Sigma-Aldrich	SAB4200114
Triton X-100	Roth	6683

Comments/Description

to fix the cells
fluorescent secondary antibody
fluorescent secondary antibody
fluorescent secondary antibody

cell culture medium
mounting medium
cell line used to perform the experiments
used as a stressing agent
transfection reagent
Coilin-specific antibody
epifluoresecence microscope
transfection medium
to wash the cells
53BP1-specific antibody
EDC4-specific antibody
to permeabilize the cells

Behm-Ansmant Isabelle, PhD
Senior scientist (CRCN CNRS, HDR)
UMR7365 CNRS-Université de Lorraine
Vandoeuvre-les-Nancy
FRANCE
Tél : +33 3 72 74 66 50
Mail : isabelle.behm@univ-lorraine.fr

Friday, May 1st, 2020

Dear Editor in Chief of JoVE

Please find below the details of the revision of our video and manuscript (original manuscript number JoVE60990) according to editorial comments.

Please revise Figure 6B/6C and Supplemental Figure 1 to be units of concentration (uM) instead of distance (um): 0 uM H_2O_2 instead of 0 um H_2O_2. The M in the micromolar abbreviation should be capitalized. Figure 6 is used in the video as well at 12:38 and 13:39. These figure should be updated in the video as well.

Answer: All the required changes in Figure 6 and supplemental Figure 1 have been done. The video has also been updated accordingly.
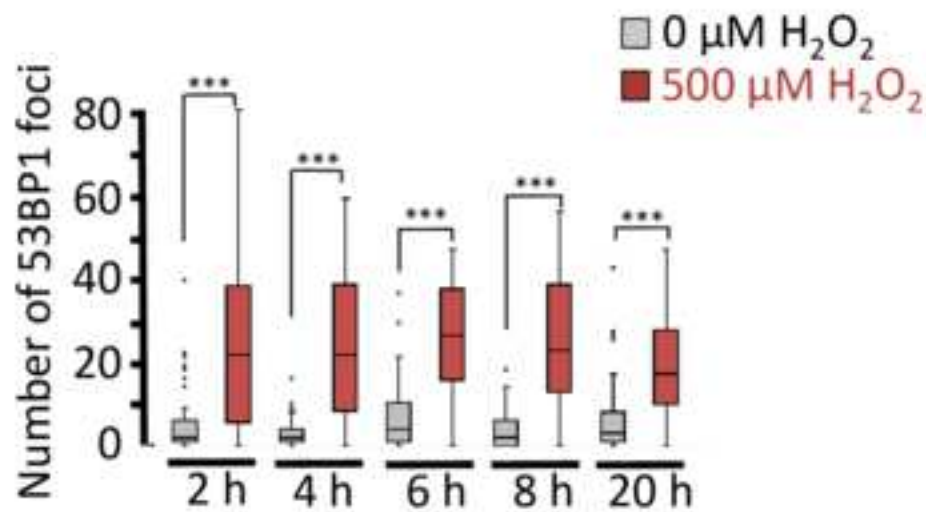
We hope that these revisions will improve the quality of our video and manuscript and make them suitable for publication in your journal,

Yours sincerely,

Isabelle BEHM-ANSMANT

Supplemental FIGURE S1

Time after stress induction	Test threshold	Sensitivity	Specificity	TP	TN	FP	FN
2h	16,000	0,579	0,893	33	50	6	24
4h	17,000	0,632	1,000	36	56	0	21
6h	17,000	0,754	0,929	43	52	4	14
8h	16,000	0,702	0,982	40	55	1	17
20h	15,000	0,632	0,911	36	51	5	21