Substructure Analyzer: A User-Friendly Workflow for Rapid Exploration and Accurate Analysis of Cellular Bodies in Fluorescence Microscopy Images

Géraud Heckler, Christelle Aigueperse, Liza Hettal, Quentin Thuillier, Fabrice de Chaumont, Stéphane Dallongeville, Isabelle Behm-Ansmant

jove

Substructure Analyzer: A User-Friendly Workflow for Rapid Exploration and Accurate Analysis of Cellular Bodies in Fluorescence Microscopy Images

Géraud Heckler[1], **Christelle Aigueperse**[1], **Liza Hettal**[1], **Quentin Thuillier**[1], **Fabrice de Chaumont**[2], **Stéphane Dallongeville**[2], **Isabelle Behm-Ansmant**[1]

[1] Université de Lorraine, CNRS, IMoPA F54000 Nancy France [2] Institut Pasteur, BioImage Analysis Unit, CNRS UMR 3691, Paris, France.

Corresponding Authors

Christelle Aigueperse
christelle.aigueperse@univ-lorraine.fr
Isabelle Behm-Ansmant
isabelle.behm@univ-lorraine.fr

Citation

Heckler, G., Aigueperse, C., Hettal, L., Thuillier, Q., de Chaumont, F., Dallongeville, S., Behm-Ansmant, I. Substructure Analyzer: A User-Friendly Workflow for Rapid Exploration and Accurate Analysis of Cellular Bodies in Fluorescence Microscopy Images. *J. Vis. Exp.* (), e60990, doi:10.3791/60990 (2020).

Date Published

June 16, 2020

DOI

10.3791/60990

URL

jove.com/video/60990

Abstract

The last decade has been characterized by breakthroughs in fluorescence microscopy techniques illustrated by spatial resolution improvement but also in live-cell imaging and high-throughput microscopy techniques. This led to a constant increase in the amount and complexity of the microscopy data for a single experiment. Because manual analysis of microscopy data is very time consuming, subjective, and prohibits quantitative analyses, automation of bioimage analysis is becoming almost unavoidable. We built an informatics workflow called *Substructure Analyzer* to fully automate signal analysis in bioimages from fluorescent microscopy. This workflow is developed on the user-friendly open-source platform Icy and is completed by functionalities from ImageJ. It includes the pre-processing of images to improve the signal to noise ratio, the individual segmentation of cells (detection of cell boundaries) and the detection/quantification of cell bodies enriched in specific cell compartments. The main advantage of this workflow is to propose complex bio-imaging functionalities to users without image analysis expertise through a user-friendly interface. Moreover, it is highly modular and adapted to several issues from the characterization of nuclear/cytoplasmic translocation to the comparative analysis of different cell bodies in different cellular sub-structures. The functionality of this workflow is illustrated through the study of the Cajal (coiled) Bodies under oxidative stress (OS) conditions. Data from fluorescence microscopy show that their integrity in human cells is impacted a few hours after the induction of OS. This effect is characterized by a decrease of coilin nucleation into characteristic Cajal Bodies, associated with a nucleoplasmic redistribution of coilin into an increased number of smaller foci. The central role of coilin in the exchange between CB components and the surrounding nucleoplasm

suggests that OS induced redistribution of coilin could affect the composition and the functionality of Cajal Bodies.

Introduction

Light microscopy and, more particularly, fluorescence microscopy are robust and versatile techniques commonly used in biological sciences. They give access to the precise localization of various biomolecules like proteins or RNA through their specific fluorescent labeling. The last decade has been characterized by rapid advances in microscopy and imaging technologies as evidenced by the 2014 Nobel Prize in Chemistry awarding Eric Betzig, Stefan W. Hell and William E. Moerner for the development of super-resolved fluorescence microscopy (SRFM)[1]. SFRM bypasses the diffraction limit of traditional optical microscopy to bring it into the nanodimension. Improvement in techniques like live-imaging or high throughput screening approaches also increases the amount and the complexity of the data to treat for each experiment. Most of the time, researchers are faced with high heterogeneous populations of cells and want to analyze phenotypes at the single-cell level.

Initially, analyses such as foci counting were performed by eye, which is preferred by some researchers since it provides full visual control over the counting process. However, manual analysis of such data is too time consuming, leads to variability between observers, and does not give access to more complex features so that computer-assisted approaches are becoming widely used and almost unavoidable[2]. Bioimage informatics methods substantially increase the efficiency of data analysis and are free of the unavoidable operator subjectivity and potential bias of the manual counting analysis. The increased demand in this field and the improvement of computer power led to the development of a large number of image analysis platforms. Some of them are freely available and give access to various tools to perform analysis with personal computers. A classification of open access tools has been recently established[3] and presents Icy[4] as a powerful software combining usability and functionality. Moreover, Icy has the advantage of communicating with ImageJ.

For users without image analysis expertise, the main obstacles are to choose the appropriate tool according to the problematic and correctly tune parameters that are often not well understood. Moreover, setup times are often long. Icy proposes a user-friendly point-and-click interface named “Protocols” to develop workflow by combining some plugins found within an exhaustive collection[4]. The flexible modular design and the point-and-click interface make setting up an analysis feasible for non-programmers. Here we present a workflow called *Substructure Analyzer*, developed in Icy’s interface, whose function is to analyze fluorescent signals in specific cellular compartments and measure different features as brightness, foci number, foci size, and spatial distribution. This workflow addresses several issues such as quantification of signal translocation, analysis of transfected cells expressing a fluorescent reporter, or analysis of foci from different cellular substructures in individual cells. It allows the simultaneous processing of multiple images, and output results are exported to a tab-delimited worksheet that can be opened in commonly used spreadsheet programs.

The *Substructure Analyzer* pipeline is presented in **Figure 1**. First, all the images contained in a specified folder are pre-

processed to improve their signal to noise ratio. This step increases the efficiency of the following steps and decreases the running time. Then, the Regions of Interest (ROIs), corresponding to the image areas where the fluorescent signal should be detected, are identified and segmented. Finally, the fluorescent signal is analyzed, and results are exported into a tab-delimited worksheet.

Object segmentation (detection of boundaries) is the most challenging step in image analysis, and its efficiency determines the accuracy of the resulting cell measurements. The first objects identified in an image (called primary objects) are often nuclei from DNA-stained images (DAPI or Hoechst staining), although primary objects can also be whole cells, beads, speckles, tumors, or whatever stained objects. In most biological images, cells or nuclei touch each other or overlap causing the simple and fast algorithms to fail. To date, no universal algorithm can perform perfect segmentation of all objects, mostly because their characteristics (size, shape, or texture) modulate the efficiency of segmentation[5]. The segmentation tools commonly distributed with microscopy software (such as the MetaMorph Imaging Software by Molecular Devices[6], or the NIS-Elements Advances Research software by Nikon[7]) are generally based on standard techniques such as correlation matching, thresholding, or morphological operations. Although efficient in basic systems, these overgeneralized methods rapidly present limitations when used in more challenging and specific contexts. Indeed, segmentation is highly sensitive to experimental parameters such as cell type, cell density, or biomarkers, and frequently requires repeated adjustment for a large data set. The *Substructure Analyzer* workflow integrates both simple and more sophisticated algorithms to propose different alternatives adapted to image complexity and user needs. It notably proposes the marker-based watershed algorithm[8] for highly clustered objects. The efficiency of this segmentation method relies on the selection of individual markers on each object. These markers are manually chosen most of the time to get correct parameters for full segmentation, which is highly time consuming when users face a high number of objects. *Substructure Analyzer* proposes an automatic detection of these markers, providing a highly efficient segmentation process. Segmentation is, most of the time, the limiting step of image analysis and can considerably modify the processing time depending on the resolution of the image, the number of objects per image, and the level of clustering of objects. Typical pipelines require a few seconds to 5 minutes per image on a standard desktop computer. Analysis of more complex images can require a more powerful computer and some basic knowledge in image analysis.

The flexibility and functionality of this workflow are illustrated with various examples in the representative results. The advantages of this workflow are notably displayed through the study of nuclear substructures under oxidative stress (OS) conditions. OS corresponds to an imbalance of the redox homeostasis in favor of oxidants and is associated with high levels of reactive oxygen species (ROS). Since ROS act as signaling molecules, changes in their concentration and subcellular localization affect positively or negatively a myriad of pathways and networks that regulate physiological functions, including signal transduction, repair mechanisms, gene expression, cell death, and proliferation[9,10]. OS is thus directly involved in various pathologies (neurodegenerative and cardiovascular diseases, cancers, diabetes, etc.), but also cellular aging. Therefore, deciphering the consequences of OS on the human cell's organization and function constitutes a crucial step in the understanding of the roles of OS in the onset and development of human pathologies. It

has been established that OS regulates gene expression by modulating transcription through several transcription factors (p53, Nrf2, FOXO3A)[11], but also by affecting the regulation of several co- and post-transcriptional processes such as alternative splicing (AS) of pre-RNAs[12,13,14]. Alternative splicing of primary coding and non-coding transcripts is an essential mechanism that increases the encoding capacity of the genomes by producing transcript isoforms. AS is performed by a huge ribonucleoprotein complex called spliceosome, containing almost 300 proteins and 5 U-rich small nuclear RNAs (UsnRNAs)[15]. Spliceosome assembly and AS are tightly controlled in cells and some steps of the spliceosome maturation occur within membrane-less nuclear compartments named Cajal Bodies. These nuclear substructures are characterized by the dynamic nature of their structure and their composition, which are mainly conducted by multivalent interactions of their RNA and protein components with the coilin protein. Analysis of thousands of cells with the *Substructure Analyzer* workflow allowed characterization of never described effects of OS on Cajal Bodies. Indeed, obtained data suggest that OS modifies the nucleation of Cajal Bodies, inducing a nucleoplasmic redistribution of the coilin protein into numerous smaller nuclear foci. Such a change of the structure of Cajal Bodies might affect the maturation of the spliceosome and participate in AS modulation by OS.

Protocol

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

1. Download Icy and the *Substructure Analyzer* protocol

1. Download Icy from the Icy website (**http://icy.bioimageanalysis.org/download**) and download the **Substructure Analyzer** protocol: **http://icy.bioimageanalysis.org/protocols?sort=latest**.
   NOTE: If using a 64-bit OS, be sure to use the 64-bit version of Java. This version allows for increasing the memory allocated to Icy (**Preferences | General | Max memory**).

2. Opening the protocol

1. Open Icy and click on **Tools** in the **Ribbon** menu.
2. Click on **Protocols** to open the **Protocols Editor** interface.
3. Click on **Load** and open the protocol **Substructure Analyzer**. Protocol loading can take a few seconds. Be sure that the opening of the protocol is complete before using it.
   NOTE: The workflow is composed of **13 general blocks** presented in **Figure 2a**. Each block works as a pipeline composed of several boxes performing specific subtasks.

3. Interacting with the workflow on Icy

NOTE: Each block or box is numbered and has a specific rank within the workflow (**Figure 2b**). By clicking on this number, the closest possible position to the first is assigned to the selected block/box then the position of the other blocks/boxes is re-organized. Respect the right order of the blocks when preparing the workflow. For example, Spot Detector block needs pre-defined ROIs so that Segmentation blocks have to

run before Spot Detector blocks. Do not modify the position of boxes. Do not use "." in the image's name.

1. By clicking on the upper left corner icon, collapse, expand, enlarge, narrow, or remove the block (**Figure 2b**).
2. Each pipeline of the workflow is characterized by a network of boxes connected via their input and output (**Figure 2b**). To create a connection, click on **Output** and maintain until the cursor attains an input. Connections can be removed by clicking on the **Output** tag.

4. Merging of the channels of an image

1. Use the block **Merge Channels** to generate merged images. If necessary, rename the files so that sequences to be merged have the same name's prefix followed by a distinct separator. For example, sequences of individual channels from an Image A are named: ImageA_red, ImageA_blue.
   NOTE: For the separator, do not use characters already present in the image's name.
2. In the same folder, create one new folder per channel to merge. For example, to merge red, green, and blue channels, create 3 folders, and store the corresponding sequences in these folders.
3. Only use the block **Merge Channels**, remove the other blocks, and save the protocol as **Merge Channels**.
   1. Access the boxes to set parameters. For each channel, fill the boxes **Channel number X** (boxes 1, 5 or 9), **Folder channel number X** (boxes 2, 6 or 10), **Separator channel number X** (boxes 3, 7 or 11) and **Colormap channel nb X** (boxes 4, 8 and 12) respectively.
      NOTE: These boxes are horizontally grouped by four, each line corresponding to the same channel. In each line, a display is also available (boxes 23, 24, or 25) to directly visualize the sequence of the corresponding channel.
      1. In the box **Channel number X**, choose which channel to extract (in classical RGB images, 0=Red, 1=Green, 2=Blue). The user quickly accesses the different channels of an image within the **Inspector** window of Icy, in the **Sequence** tab. Write the smallest channel's value in the upper line and the highest one in the bottom line.
      2. In the box **Folder channel number X**, write the \Name of the folder containing images of channel X.
      3. In the box **Separator channel number X**, write the separator used for the image's name (in the previous example: "_red", "_green" and "_blue").
      4. In the box **Colormap channel nb X**, indicate with a number which colormap model to use to visualize the corresponding channel in Icy. The available colormaps are visible in the **Sequence** tab of the **Inspector** window.
      5. In the box **Format of merged images** (box 28), write the extension to save merged images: .tif, .gif, .jpg, .bmp or .png.
         NOTE: To merge only 2 channels, do not fill the four boxes corresponding to the third channel.
   2. On the upper left corner of the **Merge Channels** block, click on the link directly to the right of **Folder**. In the **Open** dialog box which appears, double-click on the folder containing sequences of the first channel that has been defined in the box **Folder channel number 1** (box 2). Then, click on **Open**.

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

5. Segmentation of the regions of interest

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

1. Select the adapted block.
    1. If objects do not touch each other, or the user does not need to differentiate clustered objects individually, use the block **Segmentation A: Non clustered objects**.
    2. When objects do not touch each other, but some of them are close, use the block **Segmentation B: Poorly clustered objects**.
    3. For objects with a high clustering level and a convex shape, use the block **Segmentation C: Clustered objects with convex shapes**.
    4. If objects present a high clustering level and have irregular shapes, use the block **Segmentation D: Clustered objects with irregular shapes**.
    5. Use the block **Segmentation E: Clustered cytoplasm** to segment touching cytoplasms individually using segmented nuclei as markers. This block imperatively needs segmented nuclei to process.

        NOTE: Blocks adapted for primary object segmentation process independently so that several blocks can be used in the same run to compare their efficiency for a particular substructure or to segment different types of substructures. If the clustering level is heterogeneous within the same set of images, then process small and highly clustered objects separately in the adapted blocks.
2. Link the **output0** (File) of the block **Select Folder** to the **folder** input of the chosen segmentation block.
3. Set parameters of the chosen segmentation block.
    1. Segmentation A: Non clustered objects and Segmentation C: Clustered objects with convex shapes
        1. In box **Channel signal** (box 1), set the channel of the objects to segment.
        2. As an option, in box **Gaussian filter** (box 2), increase the X and Y sigma values if the signal inside objects is heterogeneous. The Gaussian filter smooths out textures to obtain more uniform regions and increases the speed and efficiency of nuclei segmentation. The smaller the objects, the lower the sigma value is. Avoid high sigma values. Set default values to 0.
        3. In box **HK-Means** (box 3), set the **Intensity classes** parameter and the approximate minimum and maximum sizes (in pixels) of objects to be detected.

            NOTE: For intensity classes, a value of 2 classifies pixels in 2 classes: background and foreground. It is thus adapted when the contrast between the objects and the background is high. If foreground objects have different intensities or if the contrast with the background is low, increase the number of classes. The default setting is 2. Object size can be quickly evaluated by drawing

a ROI manually around the object of interest. Size of the ROI (Interior in pixels) appears directly on the image when pointing it with the cursor or can be accessed in the **ROI statistics** window (open it from the search bar). Optimal parameters detect each foreground object in a single ROI**.** They can be manually defined in Icy (**Detection and Tracking | HK-Means**).

4. In the box **Active Contours** (box 4), optimize the detection of object borders. Exhaustive documentation for this plugin is available online: **http://icy.bioimageanalysis.org/plugin/Active_Contours**. Correct parameters can also be manually defined in Icy (**Detection and Tracking | Active Contours**).
5. During the process, a folder is automatically created to save images of segmented objects. In the box **Text** (box 6), name this folder (ex: Segmented nuclei). To set the format for saving images of segmented objects (Tiff, Gif, Jpeg, BMP, PNG), fill the box **format of images of segmented objects**. The folder is created in the folder containing merged images.
6. Run the workflow (for details, see part 7).

2. Segmentation B: Poorly clustered objects
   1. Follow the same steps as in 5.3.1 to set parameters of boxes **Channel signal**, **HK-Means**, **Active Contours**, **Extension to save segmented objects** and **Text** (Boxes ranks are not the same as in step 5.3.1).
   2. In the box **Call IJ plugin** (box 4), set the **Rolling** parameter to control background subtraction. Set this parameter to at least the size of the largest object that is not part of the background. Decreasing this value increases background removal but can also induce loss of foreground signal.
   3. In the box **Adaptive histogram equalization** (box 6), improve the contrasts between the foreground objects and the background. Increasing the slope gives more contrasted sequences.
   4. Run the workflow (for details, see part 7).
3. Segmentation D: Clustered objects with irregular shapes

   NOTE: Three different methods of segmentation apply to each image: firstly, **HK-means clustering** combined with the **Active Contours method** is applied. Then, the **classical watershed algorithm** (using the Euclidian distance map) is applied to previously mis-segmented objects. Finally, a **marker-based watershed algorithm** is used. Only HK-means and marker-based watershed methods need user intervention. For both methods, the same parameters can be applied for all images (fully automated version) or be changed for each image (semi-automated version). If the user is not trained in these segmentation methods, the semi-automated processing is highly recommended. During the processing of this block, manual intervention is needed. When a segmentation method is finished, the user must manually remove mis-segmented objects before the beginning of the next segmentation method. Successfully segmented objects are saved and not considered in the next step(s). This block has to be connected with the block

Clustered/heterogeneous shapes primary objects segmentation Dialog Box to work correctly.

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

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

3. Set parameters of the box **Adaptive histogram equalization** (see step 5.3.2.3).

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

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

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

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

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

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

9. At the end of the first segmentation method, an information box opens and asks to “Remove unwanted ROIs and close image”. These ROIs correspond to the borders of the segmented objects. Select **OK** and remove ROIs of mis-segmented objects in the image, which automatically opens. An ROI can be easily removed by placing the cursor on its border and using the “Delete” button of the keyboard. Close the image. Repeat the same procedure after completion of the second segmentation step.

10. At this stage, if the YES button was selected for the full-automatization of the marker-based watershed algorithm, the previously set parameters will be applied to all images.

11. If the NO button was selected, an information box opens, asking to “Determine and adjust internal markers”. Click on **OK** and within the ImageJ interface of Icy, go to **Plugins | MorphoLibJ | Minima** and **Maxima|Extended Min & Max**. In Operation, select **Extended Minima**.

12. Select **Preview** to pre-visualize on the automatically opened image the result of the transformation. Move the dynamic until optimal markers are observed. Markers are groups of pixels with a value of 255 (Not necessarily white pixels). Optimal parameters lead to one marker per object. Focus on the remaining objects that have not been well segmented with the two previous segmentation methods.

13. If necessary, improve the markers by applying additional morphological operations like “Opening” or “Closing” (**Plugins | MorphoLibJ | Morphological Filters**). When getting the final image of markers, keep it open and close all the other images ending with the image initially used as an input for the **Extended Minima** operation. Click on **No** if an ImageJ box asks to save changes on this image.

14. In the box **Nb of images with Information Box** (box 14), determine how many images with Information Boxes should appear.

4. Segmentation E: Clustered cytoplasm

NOTE: This block uses previously segmented nuclei as individual markers to initiate cytoplasm segmentation. Be sure that the block of nuclei segmentation has processed before using it.

1. In the box **Channel cytoplasm** (box 1), set the channel of the cytoplasmic signal.

2. In the box **Extension segmented nuclei** (box 2), write the format used to save images of segmented nuclei (tif, jpeg, bmp, png). The default format is tif.

3. In the box **Text** (box3), write the \Name of the folder containing segmented nuclei.

4. In the box **Format of images of segmented cytoplasms** (box 4), set the format to use for saving segmented objects images (Tiff, Gif, Jpeg, BMP, PNG).

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

6. Follow the same steps as in step 5.3.1 to set parameters of boxes **Gaussian filter** and **Active Contours** (Be careful, box ranks are not the same as in step 5.3.1).
7. Run the workflow (for details, see part 7).

6. Fluorescent signal detection and analysis

1. Select the adapted block.
    1. In the block **Fluorescence Analysis A: 1 Channel**, perform detection and analysis of foci in one channel inside one type of segmented object: detection of coilin foci (red channel) within the nucleus.
    2. In the block **Fluorescence Analysis B: 2 Channels in the same compartment**, perform detection and analysis of foci in two channels inside one type of segmented object: detection of coilin (red channel) and 53BP1 (green channel) foci within the nucleus.
    3. In the block **Fluorescence Analysis C: 2 Channels in two compartments**, perform detection and analysis of foci in one or two channels, specifically inside the nuclei and their corresponding cytoplasm: detection of Coilin foci (red channel) both within the nucleus and its corresponding cytoplasm or detection of Coilin foci (red channel) within the nucleus and G3BP foci (green channel) within the corresponding cytoplasm.
    4. In the block **Fluorescence Analysis D: Global Translocation**, calculate the percentage of signal from one channel in two cellular compartments (a and b). For example, in a cytoplasm/nucleus translocation assay, export calculated percentages of nuclear and cytoplasmic signals for each image in the final "Results" spreadsheet. The formula used to calculate the nuclear signal percentage is shown below. This block can be used for any subcellular compartment:

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

    5. In the block **Fluorescence Analysis E: Individual Cell Translocation**, calculate the percentage of signal from one channel in two cellular compartments for each cell. This block is specially optimized for nucleus/cytoplasm translocation assay at the single-cell level.
    NOTE: Because the block **Fluorescence Analysis E: Individual Cell Translocation** performs analysis at the single-cell level, efficient segmentation of nucleus and cytoplasm is needed.
2. Link the **output0** (File) of the block **Select Folder** (block 1) to the **folder** input (white arrows in black circles) of the chosen block.
3. Set the parameters of the chosen block.
    1. Fluorescence Analysis A: 1 Channel, Fluorescence Analysis B: 2 Channels in the same compartment and Fluorescence Analysis C: 2 Channels in two compartments
        1. In the box **Folder images ROI**, write the name of the folder containing images of segmented objects preceded by a backslash. (For example: \Segmented nuclei).
        2. In the box **Format of images of segmented objects** (box 2), write the format used to save images of segmented objects (tif, jpeg, bmp, png). The default format is tif.
        3. In the box **Kill Borders?**, write **Yes** to remove border objects. Otherwise, write **No**. The

installation of the MorphoLibJ collection of ImageJ is required to use this function (see step 5.3.3).

4. In the box(es) **Channel spots signal**, set the channel where spots have to be detected. In classical RGB images, 0=Red, 1=Green, and 2=Blue.

5. In the boxes **Name of localized molecule**, write the name of the molecule localizing into the spots. The number of fields to enter depends on the number of molecules.

6. In the box(es) **Wavelet Spot Detector Block**, set spot detection parameters for each channel. Set the scale(s) (referred to spot size), and the sensitivity of the detection (a smaller sensitivity decreases the number of detected spots, the default value being 100 and the minimum value being 0). Exhaustive documentation of this plugin is available online: **http://icy.bioimageanalysis.org/plugin/Spot_Detector**. Parameters can also be manually defined in Icy (**Detection and Tracking | Spot Detector**).

7. As an option, in the box **Filter ROI by size**, filter the segmented objects where spots are detected by setting an interval of size (in pixels). This step is especially useful to remove under- or over-segmented objects. To manually estimate the size of objects, see step 5.3.1. Default parameters do not include the filtering of ROIs by size. The block 2 Channels in two compartments contains two boxes: Filter nuclei by size (box 19) and Filter cytoplasm by size (box 46).

8. Optionally, in the box **Filter spots by size**, filter the detected spots according to their size (in pixels) to remove unwanted artifacts. To manually estimate spot size, click on **Detection and Tracking** and open the **Spot detector** plugin. In **Output** options, select **Export to ROI**. Be careful that default parameters do not include filtering of spots by size, and that filtered spots are not taken into account for the analysis. The number of fields to enter depends on the number of channels.

9. Optionally, in the box **Filter spots**, apply an additional filter (contrast, homogeneity, perimeter, roundness) on the detected spots. Be careful that default parameters do not include spot filtering and that filtered spots are not taken into account for the analysis. The number of fields to enter depends on the number of channels.

10. Optionally, in the boxes **Spot size threshold**, set a threshold for the area (in pixels) of analyzed spots. The number of counted spots below and above this threshold is exported in the final **Results** spreadsheet. The number of boxes to be informed depends on the number of channels.

11. Run the workflow (for details, see part 7). Data are exported in a **Results** spreadsheet saved in the folder containing merged images.

2. Fluorescence Analysis D: Global Translocation and Fluorescence Analysis E: Individual Cell Translocation:

    1. In the boxes **Folder images** (boxes 1 and 2), write the \Name of the folder containing images of segmented objects. In the block **Fluorescence**

Analysis D: Global Translocation, the two types of ROI are identified as ROI a and ROI b. For the block **Fluorescence Analysis E: Individual Cell Translocation**, in **Folder images segmented nuclei** and **Folder images segmented cytoplasms** boxes, write the name of the folder containing segmented nuclei and cytoplasms, respectively.

2. In the box **Channel signal** (box 3), enter the channel of the signal.
3. In the box **Format of images of segmented objects** (box 4), write the format used to save images of segmented objects (tif, jpeg, bmp, png). The default format is tif. The **Kill Borders** option is also available to remove border objects (see step 6.3.1).
4. Optionally, in boxes **Filter ROI by size**, filter the segmented objects by setting an interval of size (in pixels). This step could be useful to remove under- or over-segmented objects. To manually estimate object size, see step 5.3.1. There are two fields to enter, one per channel. Default parameters do not include ROI filtering by size.
5. Run the workflow (for details, see part 7). Export data in a spreadsheet **Results** saved in the folder containing merged images.

7. Run the protocol

1. To process one block in a run, remove the connection between the selected block and the block **Select Folder**. Place the wanted block at the 1st rank. On the upper left corner of the wanted block, click on the link directly to the right of **folder**. In the **Open** dialog box which appears, double-click on the folder containing the merged images. Then, click on **Open**. Click on **Run** to start the workflow. The processing can be stopped by clicking on the **Stop** button.
2. To process different blocks in a run, keep connections of chosen blocks with the block **Select Folder** (block 1). Make sure that their rank allows the good processing of the workflow. For example, if a specific block needs segmented objects to process, be sure that the segmentation block processes before. Before running the workflow, remove unused blocks and save the new protocol with another name.
3. Click on **Run** to start the workflow. When the open dialog box appears, double-click on the folder containing the merged images. Then, click on **Open**. The workflow automatically runs. If necessary, stop the processing by clicking on the **Stop** button.
4. At the end of the processing, check that the message **The workflow executed successfully** appeared in the lower right-hand corner and that all the blocks are flagged with a green sign (**Figure 2b**). If not, the block and the inside box presenting the error sign indicate the element to correct (**Figure 2b**).

   NOTE: After the workflow executed successfully, a new run cannot be directly started, and to process the workflow again, at least one block should be flagged with the sign “ready to process”. To change the state of a block, either delete and re-create a link between two boxes inside this block or simply close and re-open the protocol. If an error occurs during the processing, a new run can directly be started. During a new run, all the blocks of the pipeline are processed, even if some of them are flagged with the green sign.

Representative Results

All the described analyses have been performed on a standard laptop (64-bit, quad-core processor at 2.80 GHz with 16 GB random-access memory (RAM)) working with the 64-bit version of Java. Random-access memory is an important parameter to consider, depending on the amount and the resolution of images to analyze. Using the 32-bit version of Java limits the memory to about 1300 MB, which could be unsuitable for big data analysis, whereas the 64-bit version allows increasing the memory allocated to Icy. **Figure 3** reports the time needed for segmentation for different types of images and different resolutions. It confirms that high resolution substantially increases the time of primary object segmentation.

The data presented in the next paragraphs demonstrate that the *Substructure Analyzer* workflow can be used to solve most of the common problematics encountered in cellular and molecular biology (cell counting, foci counting and analysis, translocation analysis).

Measurement in numerous cells of the precise proportion of molecules concentrated within a given compartment or the change of localization across subcellular domains can be a very cumbersome and error-prone task, especially when it is performed manually. **Figure 4** illustrates the ability of the workflow to rapidly and precisely quantify the well-described nuclear translocation of the transcription factor NFκB in response to TNFα stimulation. Images used for the analysis were graciously compiled by Ilya Ravkin (http://www.ravkin.net/SBS/Image-Library.htm) and are publicly available at the Broad Bioimage Benchmark Collection[16]. This set contains images of MCF7 and A549 cell lines treated with increasing concentrations of TNFα. The analysis has been performed on more than 40,000 cells and 96 images (48 images for each channel) from the MCF7 cell line (**Figure 4a**). The whole analysis (from the import of images to data saving) took 26 minutes. Each image corresponds to a specific TNFα concentration. The nuclear/cytoplasmic proportions of NFκB were homogeneous across all cells for a given image. For this reason, for each image, the values of all nuclear or cytoplasmic pixels were summed to calculate the nuclear and cytoplasmic proportions of NFκB, and individualization of touching nuclei/cytoplasms during their segmentation was not required. DAPI images were processed in the **Segmentation C: Clustered objects with convex shapes** block to segment nuclei and the green channel was used to delineate the cytoplasms with the **Segmentation E: Clustered cytoplasm** block (**Figure 4b**). The **Fluorescence Analysis D: Global Translocation** block was used to export the sum of nuclear and cytoplasmic pixel values. The obtained data are represented in a dose-response curve (**Figure 4c**) and illustrate the increase of NFκB nuclear proportion after stimulation with increasing concentrations of TNFα.

The workflow not only analyzes the global intensity signal in different sub-cellular compartments but can also be used to detect foci and extract specific information about their features. **Figure 5a** illustrates the detection of P-bodies in individual cells by localizing the enhancer of mRNA decapping 4 (EDC4) protein. The **Segmentation A: Non clustered objects** block was used to segment nuclei that present a low level of clustering, the **Segmentation E: Clustered cytoplasm** block to delineate the cytoplasms and the **Fluorescence Analysis C: 2 channels in two compartments** to detect EDC4 foci. Segmented objects are identified (Segmented nuclei_0, Segmented cytoplasms_0), and multiple information is collected into

the corresponding worksheet (Segmented nuclei Information and Segmented cytoplasms Information) like the EDC4 mean intensity or the number/size of detected bodies in each ROI (**Figure 5b**). Before being analyzed, detected bodies can be preliminarily filtered according to their area, which can be useful to exclude objects below the resolution power of the microscope. To estimate the size of conserved objects, an additional area threshold can be introduced. Areas of all the detected bodies are reported in dedicated worksheets (Segmented nuclei_Distribution foci size, Segmented cytoplasms_Distribution foci size). In this example, we highlight the analysis of two cells (Segmented nuclei_0 and_1 and Segmented cytoplasm_0 and_1), where cytoplasmic EDC4 foci are detected. Note that even though the signal of the EDC4 protein is detected in both the nuclear and cytoplasmic compartments, only the cytoplasmic foci correspond to P-bodies. The nuclear signal most likely results from the cross-reactivity of the antibody used to perform immunofluorescence experiments. The size in pixels of each of the foci is given.

Then, as the effect of oxidative stress (OS) on Cajal Bodies was still unknown, we took advantage of the versatility of the workflow to study it during time kinetics (2 to 20 hours after OS induction with 500 µM H_2O_2) (**Figure 6**). Cajal Bodies are dynamic structures nucleating and dissolving within the nucleoplasm under the control of specific parameters like the transcription rate or cell cycle progression. Cajal Bodies were visualized by localizing their main structural and functional component, the coilin protein. Information about the number and the size of Cajal bodies were collected by studying 2300 individual cells from high-resolution images (3840X3072) containing 3 channels: DAPI (blue), 53BP1 (green) and coilin (red) (**Figure 6a**). Full processing took less than 1 h. Because 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 a modification of the structure of Cajal Bodies could be linked to changes in the phosphorylation status of coilin. Moreover, recent works demonstrated that Cajal Bodies are not randomly localized within the nucleus, and can affect gene expression through proximal association with specific gene loci[17] . Given all these facts, it would not be surprising if an effect on Cajal bodies' functionality would accompany changes in their structure. From these results, we can also ask if such Cajal Bodies' remodeling is a passive consequence of OS or participates in the response by interacting with specific gene loci, for example.

To test if a change in coilin expression could alter its localization and change the nucleation of Cajal Bodies, we overexpressed an exogenous GFP-coilin fusion protein. Nuclei were not clustered, so segmentation was performed with the block **Segmentation A: Non clustered objects**. Then, to precisely quantify the number and size of the coilin foci (red signal, channel 1) according to the level of GFP-coilin (green signal, channel 2) overexpression, the block **Fluorescence Analysis B: 2 Channels in the same compartment** was used. The level of GFP-coilin was reflected by the mean intensity of the GFP signal in individual nucleus (**Figure 7a**). In cells with medium or high level of GFP-coilin overexpression (GFP intensity higher than 10), the number of Cajal Bodies per nucleus significantly increases with some nuclei displaying up to 21 foci (**Figure 7b**). We also noticed that the brightness (mean intensity) of Cajal Bodies considerably increases in parallel with the overexpression level leading to intensity saturation. Such saturated regions might thus mask the presence of smaller and fainter foci. In cells overexpressing a medium or high level of GFP-coilin, the size of Cajal Bodies also significantly increases, the median values rising from 1 to 2 μm^2 (**Figure 7c**). Moreover, more than 25% of cells with a high GFP-coilin expression level present foci with an area larger than 2.5 μm^2 , whereas it never exceeds 0.8 μm^2 in non-transfected cells. In conclusion, GFP-Coilin overexpression leads to a significant increase in both the number and size of Cajal

Bodies. Since OS increases the number of Cajal Bodies but reduces their size, these data might reflect that OS effect on the structure of Cajal Bodies is most probably induced by a change of their composition rather than by an effect on the cellular amount of coilin.

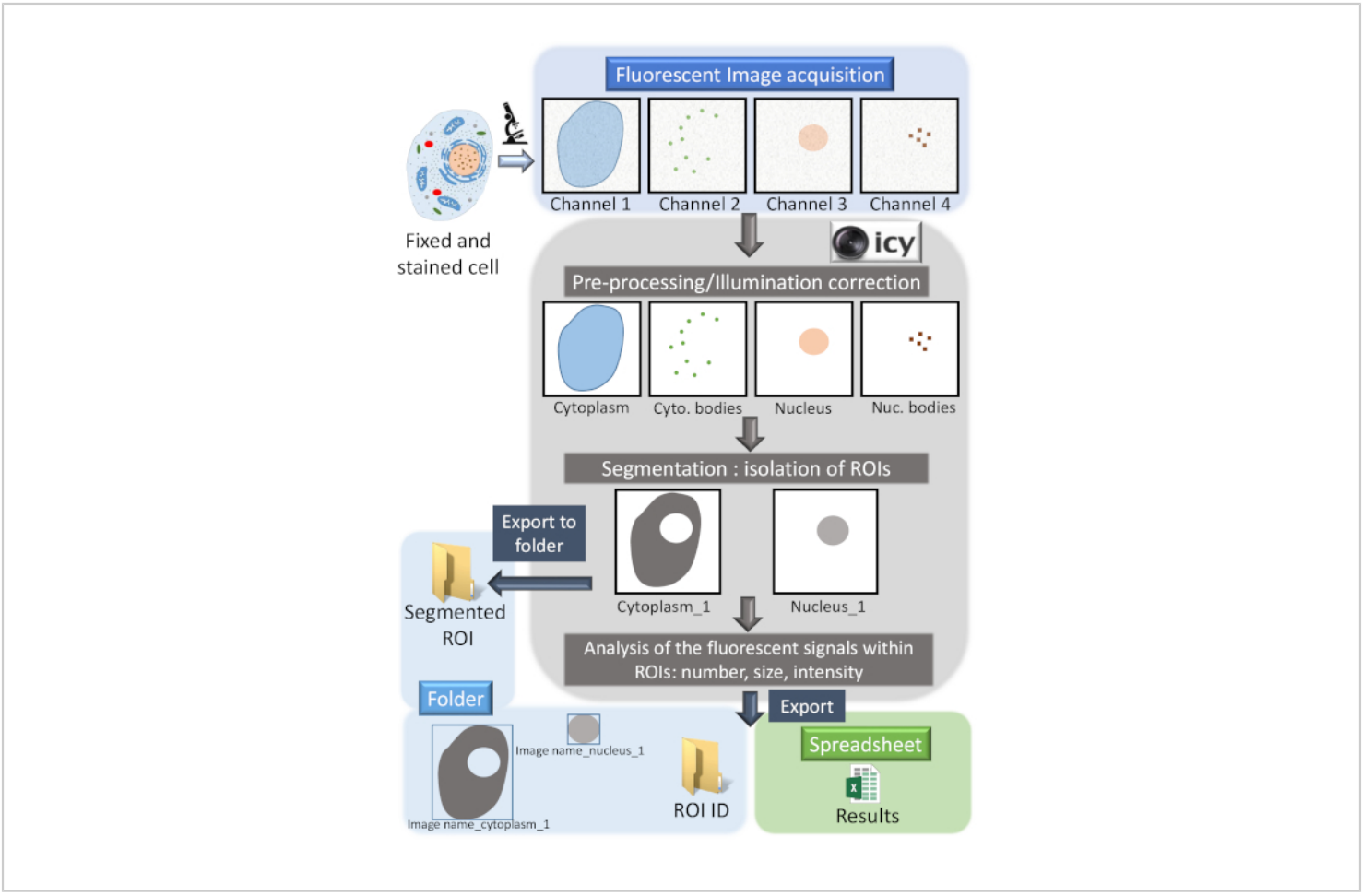

Figure 1: The *Substructure Analyzer* pipeline

The workflow is developed in Icy, an open community platform for bioimage informatics, and performs automatic analysis of multiple fluorescence microscopy images. Firstly, multi-channel images are automatically loaded within the protocol and pre-processed to improve, if necessary, the signal to noise ratio and to remove imaging artifacts. Then, image segmentation isolates regions of interest (ROIs) from the background. Several methods of segmentation are available depending on the level of clustering and the nature of the objects of interest. Segmented objects are saved with a specific descriptor (for example, Image name_Nucleus_1) in a specific folder and can be used/re-used for subsequent analysis. Fluorescent signals like spots are then analyzed within the ROIs and multiple features (location, size, shape, intensity, texture, spot number, and size) are exported into an automatically created spreadsheet. All the measured features, like the number of detected spots, are reported to the descriptor of the corresponding ROI. Please click here to view a larger version of this figure.

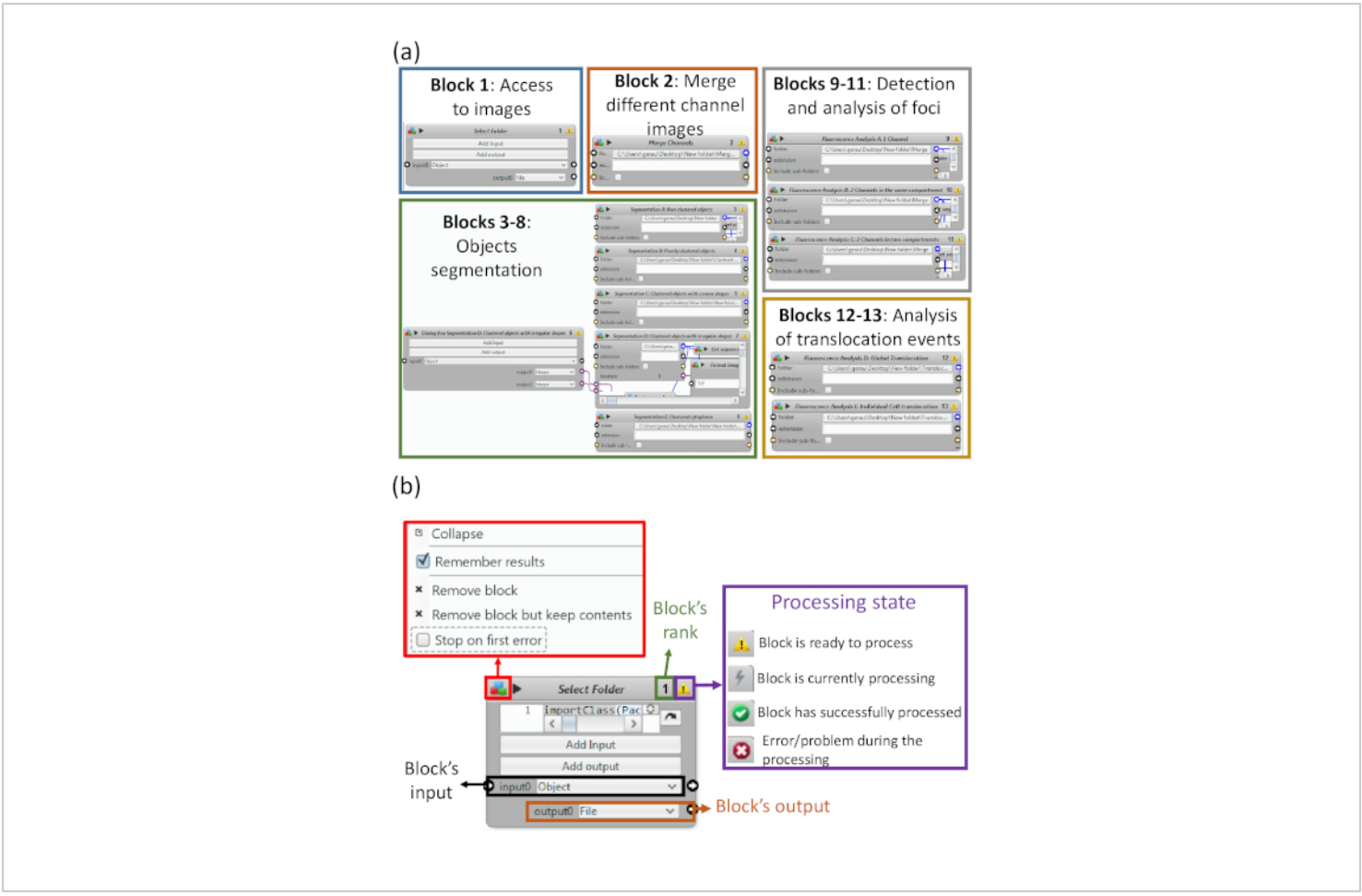

Figure 2: Graphical interface of the Icy workflow

(**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. Please click here to view a larger version of this figure.

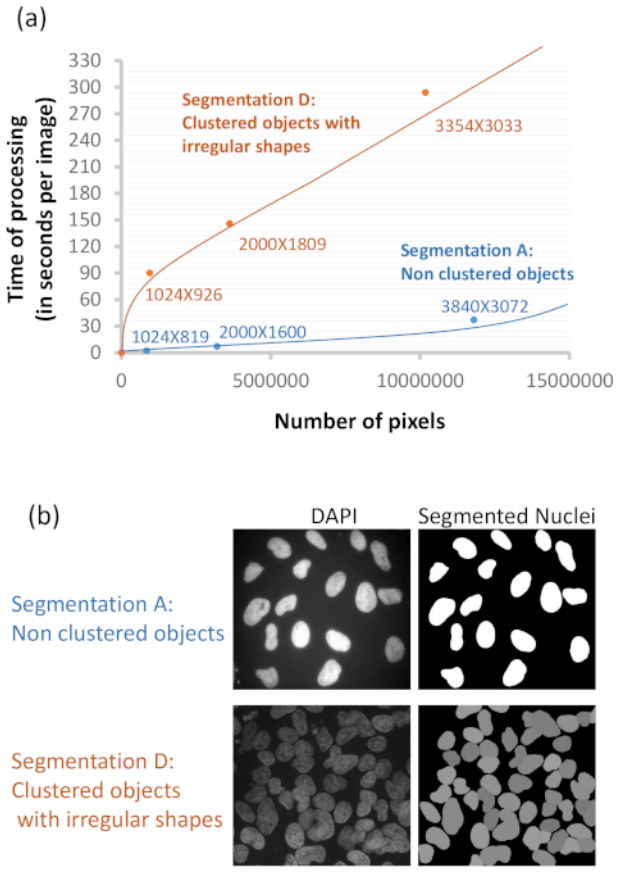

Figure 3: Time needed for segmentation as a function of image resolution and object clustering level

(**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** block. DAPI staining and segmented nuclei are indicated. Please click here to view a larger version of this figure.

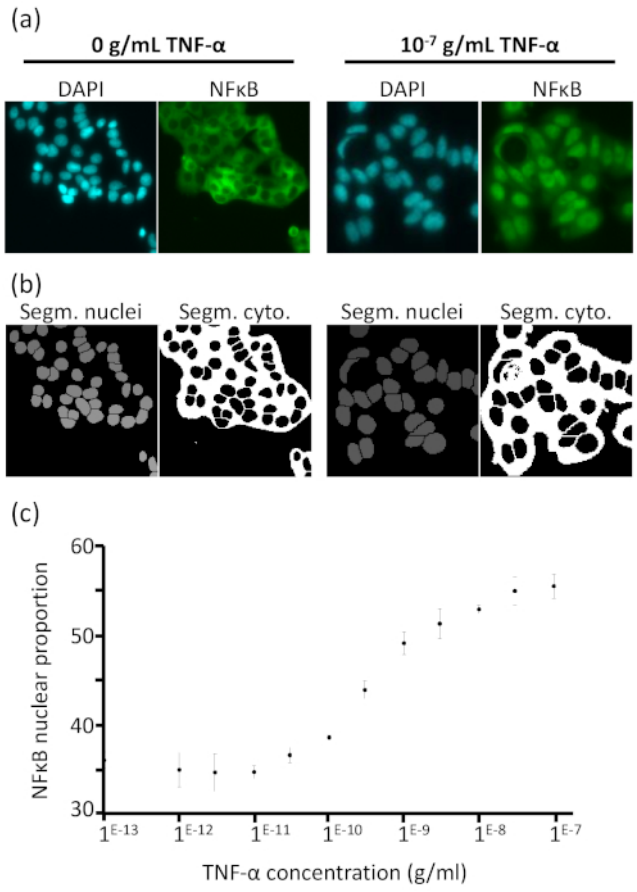

Figure 4: Workflow analysis of an NFκB nuclear translocation assay in response to TNFα stimulation

(**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. Please click here to view a larger version of this figure.

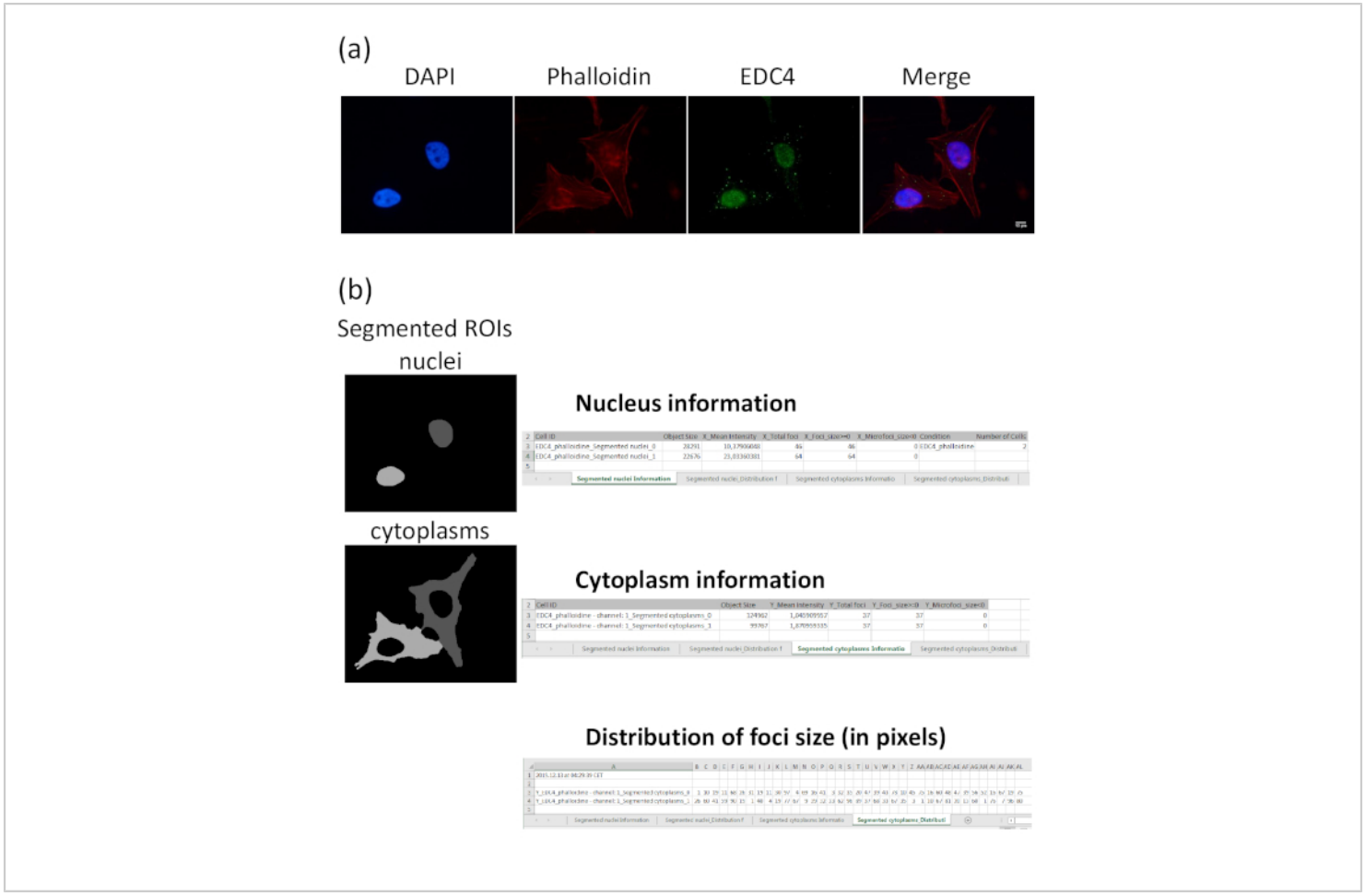

Figure 5: Quantification of EDC4 foci (P-bodies) within the cytoplasm in individual cells

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

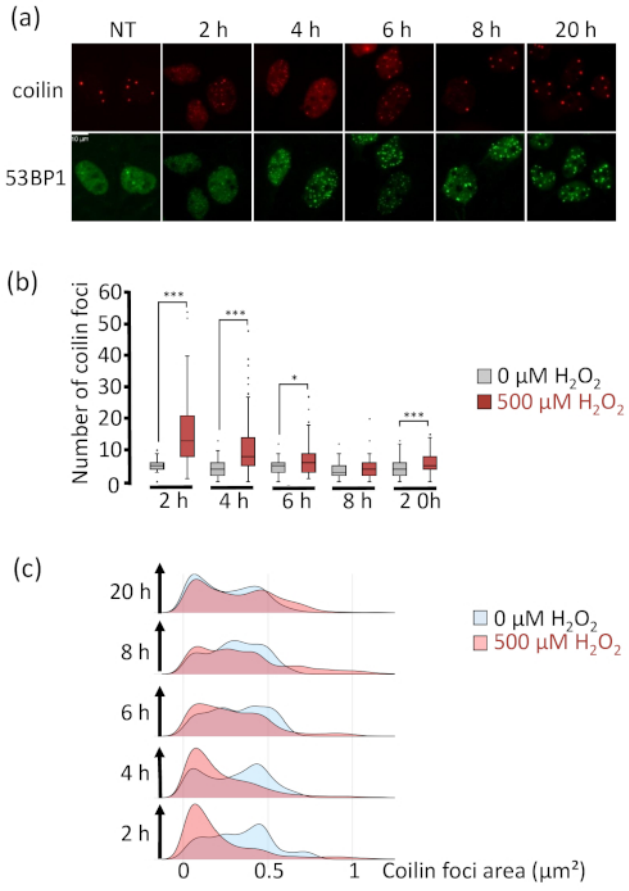

Figure 6: Kinetics of coilin and 53BP1 nucleoplasmic distribution after oxidative stress

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

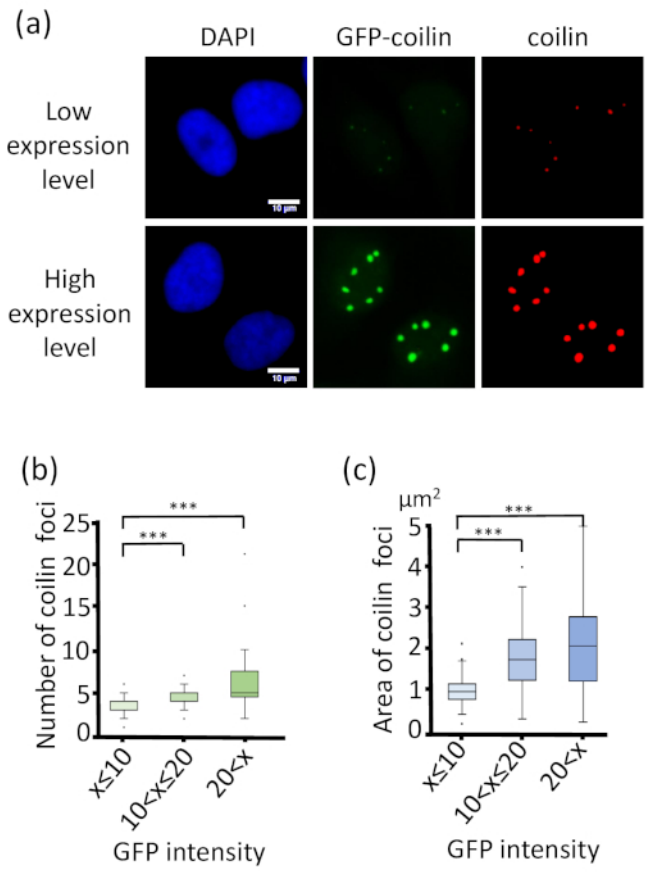

Figure 7: Study of Cajal Bodies' nucleation in cells overexpressing a GFP-coilin fusion protein

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

test analysis demonstrates a significant increase of the area of coilin foci for a GFP intensity >10 (*, p<0.05; **, p<0.01; ***, p<0.001). Please click here to view a larger version of this figure.

Figure S1: Kinetics of 53BP1 nucleoplasmic distribution after oxidative stress

(**a**) Quantification of 53BP1 nuclear foci in untreated or H_2O_2 treated cells after different incubation times. Results are shown as box plots, where the central mark is the median, the lower and upper edge of the box are the 25^{th} and 75^{th} percentiles. Wilcoxon - Mann Whitney test analysis demonstrates a significant difference between untreated and H_2O_2 treated cells (*, p<0.05; **, p<0.01; ***, p<0.001). (**b**) For each time point of the kinetics, a ROC (Receiver Operating Characteristic) curve has been established to determine the best threshold to discriminate between stressed and unstressed cells. The parameters of each test are reported in the table (Sensitivity, specificity, TP: true positive, TN: true negative, FP: false positive, FN: false negative). From these data, a global threshold of 17 53BP1 foci was determined to discriminate stressed from non-stressed cells through the kinetics. Please click here to download this file.

Discussion

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

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

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

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

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

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

Disclosures

The authors have nothing to disclose.

Acknowledgments

G.H. was supported by a graduate fellowship from the Ministère Délégué à la Recherche et aux Technologies. L.H. was supported by a graduate fellowship from the Institut de Cancérologie de Lorraine (ICL), whereas Q.T. was supported by a public grant overseen by the French National Research Agency (ANR) as part of the second “Investissements d’Avenir” program FIGHT-HF (reference: ANR-15-RHU4570004). This work was funded by CNRS and Université de Lorraine (UMR 7365).

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).
6. Zaitoun, N. M., Aqel, M. J. Survey on image segmentation techniques. *Procedia Computer Science.* 65, 797-806 (2015).
7. Devices, M. *MetaMorph Microscopy Automation and Image Analysis Software.* https://www.moleculardevices.com/products/cellular-imaging-systems/acquisition-and-analysis-software/metamorph-microscopy (2018).
8. Nikon. *NIS-Elements Imaging Software.* https://www.nikon.com/products/microscope-solutions/lineup/img_soft/nis-element (2014).
9. Meyer, F., Beucher, S. Morphological segmentation. *Journal of Visual Communication and Image Representation.* 1 (1), 21-46 (1990).
10. Schieber, M., Chandel, N.S. ROS Function in Redox Signaling and Oxidative Stress. *Current Biology.* 24 (10), R453-R462 (2014).
11. 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).
12. 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).
13. 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).
14. 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).
15. 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).
16. Will, C.L., Luhrmann, R. Spliceosome Structure and Function. *Cold Spring Harbor Perspectives in Biology.* 3 (7), a003707-a003707 (2011).
17. Ljosa, V., Sokolnicki, K.L., Carpenter, A.E. Annotated high-throughput microscopy image sets for validation. *Nature Methods.* 9 (7), 637-637 (2012).

18. Wang, Q. et al. Cajal bodies are linked to genome conformation. *Nature Communications.* 7 (2016).

19. Carpenter, A.E. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology.* 7, R100 (2006).

20. McQuin, C. et al. CellProfiler 3.0: Next-generation image processing for biology. *PLoS Biology.* 16 (7), e2005970 (2018).