Check for updates

SOFTWARE TOOL ARTICLE

REVISED SAPHIR: a Shiny application to analyze tissue section images [version 2; peer review: 2 approved]

Elodie Germani 📵 1, Hugues Lelouard2, Mathieu Fallet 📵 2

1Université de Rennes 1, Rennes, 35000, France
2Aix Marseille Univ, CNRS, INSERM, CIML, Marseille, France

Abstract
Study of cell populations in tissues using immunofluorescence is a powerful method for both basic and medical research. Image acquisitions performed by confocal microscopy notably allow excellent lateral resolution and more than 10 parameter measurements when using spectral or multiplex imaging. Analysis of such complex images can be very challenging and easily lead to bias and misinterpretation. Here, we have developed the Shiny Analytical Plot of Histological Image Results (SAPHIR), an R shiny application for histo-cytometry using scatterplot representation of data extracted by segmentation. It offers many features, such as filtering of spurious data points, selection of cell subsets on scatterplot, visualization of scatterplot selections back into the image, statistics of selected data and data annotation. Our application allows to characterize labeled cells, from their phenotype to their number and location in the tissue, as well as their interaction with other cells. SAPHIR is available from: https://github.com/elodiegermani/SAPHIR

Keywords
Tissue cellular quantification, spatial cellular profiling, cell-cell interactions, scatterplot, histo-cytometry, image cytometry

This article is included in the NEUBIAS - the Bioimage Analysts Network gateway.

Open Peer Review

Reviewer Status ✔ ✔

	Invited Reviewers	
	1	**2**
version 2 (revision) 08 Apr 2021	✔ report	✔ report
version 1 27 Oct 2020	? report	? report

1. **Mahmoud Ahmed** 📵, Gyeongsang National University School of Medicine, Jinju, South Korea

   Trang Huyen Lai 📵, Gyeongsang National University School of Medicine, Jinju, South Korea

2. **Seth Winfree**, Indiana University School of Medicine, Indianapolis, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Elodie Germani (elodie.germani@hotmail.fr)

Author roles: Germani E: Software, Writing – Original Draft Preparation; **Lelouard H**: Conceptualization, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing; **Fallet M**: Conceptualization, Project Administration, Resources, Supervision, Writing – Original Draft Preparation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

How to cite this article: Germani E, Lelouard H and Fallet M. **SAPHIR: a Shiny application to analyze tissue section images [version 2; peer review: 2 approved]** F1000Research 2021, **9**:1276 https://doi.org/10.12688/f1000research.27062.2

First published: 27 Oct 2020, **9**:1276 https://doi.org/10.12688/f1000research.27062.1

> **REVISED** **Amendments from Version 1**
>
> According to reviewers' comments, we updated our paper with details added on the Data requirement section, on the development of the application and on the Use case section. We also provided new Figures with a better resolution and new options implemented in SAPHIR to increase the visualization properties of the app.
>
> **Any further responses from the reviewers can be found at the end of the article**

Introduction

The identification, localization and quantification of cell subsets in tissues is a difficult but essential task for biologists to understand spatial cellular organization in different settings (e.g. homeostasis vs inflammatory diseases or cancer). Advances in optical microscopy allow image acquisitions with more than 10 channel measurements using spectral fluorescence imaging or multiplex imaging combined with z-axis optical slices of tissue sections ranging from 10 μm to more than 200 μm when clearing methods are used[1–4]. Analysis of such complex images is very challenging due to the size and complexity of data. It requires image segmentation in 3D that can be further improved using deep learning-based segmentation[5–7]. Then, like flow cytometry, complex image analysis can benefit from scatterplot representations that allow to gate cells of interest[2,8]. Surprisingly, in existing software, this scatterplot representation is rarely interactive with the image itself, although this would allow to locate selection results back into the image but also to filter or correct results and fine-tune the gates defining cell populations to obtain in return a better visualization of them into the image.

To this end, we developed the Shiny Analytical Plot of Histological Image Results (SAPHIR), an R/Shiny application for the quantitative analysis of tissue section images. Since image segmentation is a complex task in continuous development and highly dependent on image quality and information, the integration of a single type of segmentation method in an image analysis application is not necessarily recommended. Therefore, we decided to separate segmentation from SAPHIR, but we provided in a segmentation menu the link to Fiji and the possibility to run Fiji segmentation macros on user image. Two Fiji macro examples with associated images, which can be used to perform this task before running SAPHIR, are also provided with the application. SAPHIR offers many features such as interactive scatterplots with the image and data filtering and correction as described below.

Methods

Data requirement

The segmentation process that is required to use the application SAPHIR was carried out under Fiji[9], with homemade macros. These are available on GitHub in the Demonstration Files and can be loaded from the segmentation menu of the application, too. Cell nucleus detection was used for 3D segmentation and the optical section containing the central region of each cell (largest nucleus area) was used to evaluate the fluorescent intensity signal of each channel within a circular ring containing the cell membrane. Finally, the external ring was converted to simple ROI (polygon) and save in .roi for each Cell Of Interest (COI) to be easily readable in R. Each COI was associated with its main optical section (slice). In addition, regions of interest were defined using DBSCAN (optional)[10]. The segmentation process of the image to be analyzed should create a csv result file containing the intensity values of each channel (ranging from 0 to 255) and, if required, the positioning (x, y, z) linked to each COI. Importantly, many other morphological measurements like area, roundness, or solidity as well as other information (COI belonging to a region of interest, COI interaction with other cells) can be added to the segmentation result file. The result file should contain at least two columns that provide intensities or other parameters necessary to create a scatterplot.

Finally, the minimal requirements to run the application are the image in TIFF with multiple channels and optionally slices, a legend file with all channel information written in column with name headers (Ch1, Ch2,..) and the corresponding labelling (e.g., LT or LB,..), a segmentation result file with intensity values of each COI in csv and the roi.zip file containing contours of each COI. For the TIFF image, since large images may alter the fluidity of the application, we decided that images with a X or Y dimension higher than 1024 would be resized proportionaly so the highest dimension is 1024 (e.g. a 2048 * 1536 image will become a 1024 * 768 image). SAPHIR is able to load standard 2D ROI from Fiji. The segmentation process can be done in either 2D or 3D, but in the latter case only the largest ROI associated with the central slice of each cell is kept. It is possible to convert labelled images (or mask image) in .roi by applying the "analyze particles" function of Fiji or by loading polygon ROI in Fiji and saving them in .roi.

Operation

SAPHIR was built in the R programming language (version 4.0.2), which has very powerful stastistical packages and excellent interactive dashboard programing package, Shiny and its derivatives like ShinyDashboard or Shinyfiles. Several packages available on CRAN (e.g. ijtiff, magick, ggplot2, plotly) and one package (EBImage) available on Bioconductor were used.

The development of the app was made as follows. First, an interactive Shiny app was created, using the "shiny" library in R, to link a tissue section image and a scatterplot representing for each cell its fluorescence intensity on different channels. To do so, the libraries "ijtiff", "RImageROI" and "ggplot2" were used. In addition, the statistical part of the app and the "Image to Plot" section were added. Second, the interface of the app was improved with "shinydashboard", "shinyWidgets" and "shinycssloaders". Third, a menu "Select your result" was created to load needed files (see Data requirement section). Fourth, the interactivity between the scatter plot and the image was improved using "plotly. A "filtering" option was also added to the scatterplot. Fifth, the visualization options of the scatterplot selections in the image, such as cell ID display, channel overlays and thickness of cells contours were added to the app. This was done using "EBImage" and "magick" packages. Finally, the "Annotation" menu was added to allow result

corrections based on scatterplot to image analysis and "Download" buttons were created to allow the download of SAPHIR results.

SAPHIR has been tested on macOS and Windows 10. A user flowchart from image acquisition to obtention of quantitative data of COI with SAPHIR is shown in Figure 1.

Implementation

The first step of SAPHIR is either to run a segmentation program to obtain appropriate files or to load the four required files (image.tif, roi.zip, results.csv, legend.csv) in the "select your results" menu of the application. Here, result, region of interest (ROI) and legend files were obtained with a homemade segmentation macro developed under Fiji as described in the Data requirement section. Fiji and the macro can be launched from the segmentation menu of the application.

The first tab of the "Plot to Image" menu is shown in Figure 2 and allows the users to filter their data depending on one or two parameters displayed on histogram or scatterplot, respectively. This can be used to remove cells that have been badly segmented. Indeed, such cells display an aberrant area or volume (doublets or aggregates in red in Figure 2). Here, only cells within the Gaussian curve were retained (in blue in Figure 2). This can also be used to work only on given subpopulations

filtered based on their location, interaction with other cells or other criteria (see Use case section below). Then, filtered cells can be separated into subsets thanks to a two-parameter (most often channel intensities) scatterplot (Figure 3A). In addition, users can add a third parameter that is displayed on the scatterplot through a change of symbol shape for each COI beyond a threshold defined by the users for a given parameter value (Figure 3A).

Gating can be made easily with quadrants, rectangles and lassos in single or multiple selection mode. Interactivity between scatterplot and image is optional to avoid slow interaction when the size of images exceeds 300 MB, especially using Docker. In the latter case, we recommend users to perform their gating before allowing their selection to be shown in the image by ticking the appropriate box. Users can easily change the slice and the displayed channels with buttons to monitor selected COI spatial distribution and fluorescence intensity in the image (Figure 3B). Moreover, several options are available such as cell identity display or channel overlay (Figure 3B). Finally, statistics of the gated COI are downloadable, and gates are saveable for further analyses.

The SAPHIR menu "Image to Plot" allows to select a region in the image and to display cells of this region on a two-parameter scatterplot (Figure 4).
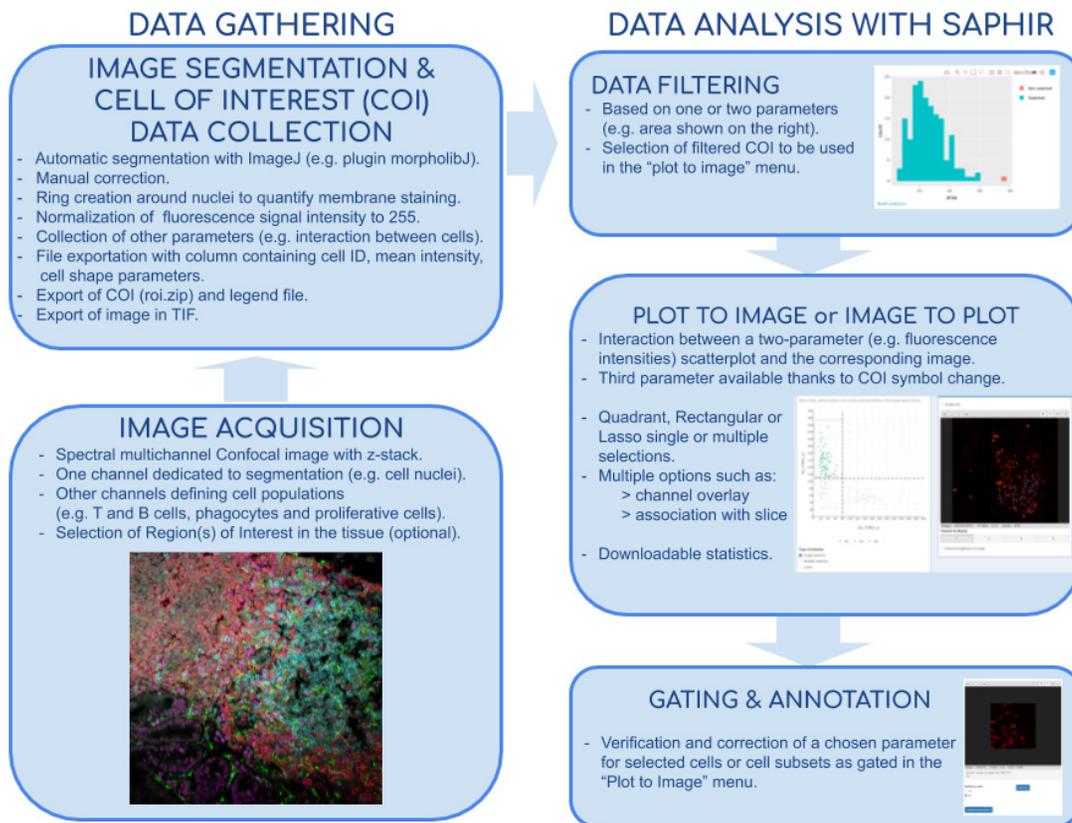
Figure 1. Flow chart of tissue image analysis from image acquisition and segmentation (left side) to extracted data analysis with SAPHIR (right side).
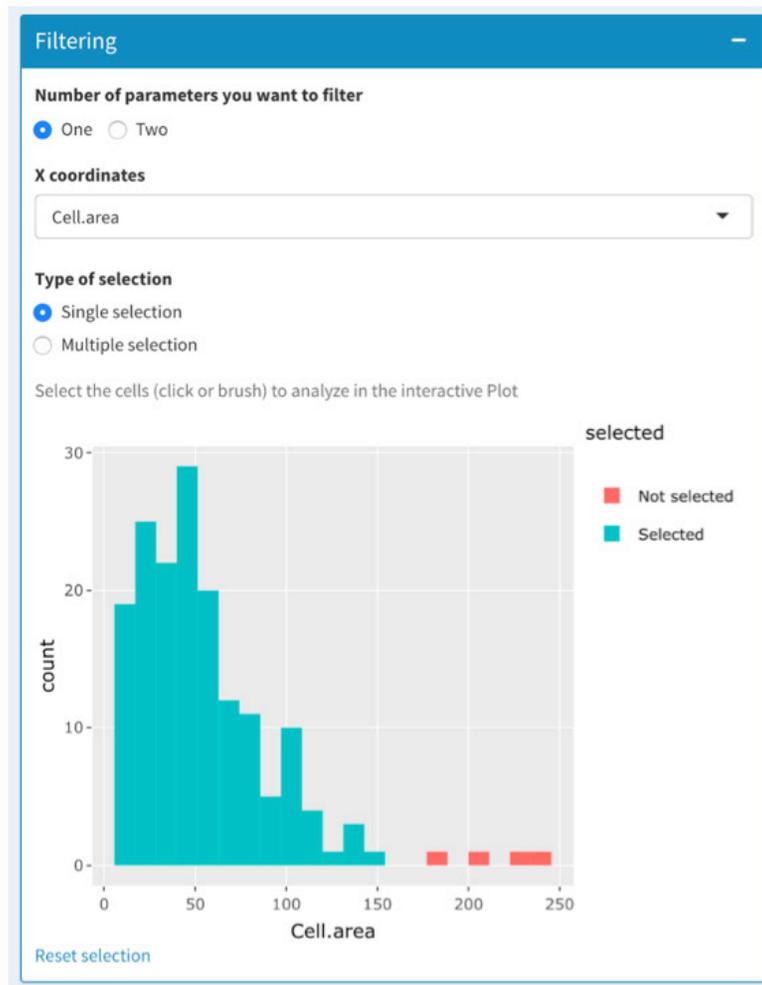
Figure 2. Menu "Plot to Image" of SAPHIR application. The data-filtering tab allows the selection of cells to be analyzed based on one or two parameters. Based on the cell area parameter, only cells that displayed a conventional cell area (in blue) were selected. Badly segmented cells (large area in red) were discarded.

Finally, the menu "Annotation" allows correction of the data from the result csv file but also from the scatterplot-gated cells when saved in the "Plot to Image" menu. For each selected cell, a cropped image of this cell is displayed, and based on the previous analyses users can change its parameters (e.g., cell identity) if necessary (Figure 5).

Use case

To show the usefulness of our application, we used it on a project that aims to characterize in murine Peyer's patches (PP) the interaction between phagocytes and proliferative immune effector cells, i.e. B and T cells. PP are immune inductive sites distributed along the small intestine in charge of sampling noxious antigens and mounting an immune response against them. In PP, antigens are taken up by phagocytes, which, upon stimulation, migrate in the T cell zone and its periphery to interact with prime naïve T cells. The periphery of the PP T cell zone is indeed an area of intense proliferation of immune effector cells after stimulation, suggesting that this region is a privileged site for their activation[11]. We therefore decided to use SAPHIR to examine the evolution of proliferative cell number, to determine their identity (B or T cells) and to analyze their interaction with migratory phagocytes in this region during the course of a stimulation.

We selected a spectral confocal image of a stained stimulated PP. It consisted of 12 optical slices with 6 channels distinguishing proliferative nuclei, B cells, T cells, monocyte-derived phagocytes, early-activated T cells and total nuclei. All optical slices but only the first four channels were necessary for the purpose of this study. First, we used two homemade Fiji macros as described in Data requirement. These macros are available on GitHub in the Demonstration Files and can be launched from the Segmentation menu of the application. The first macro allows the counting of proliferative cells through segmentation of Ki-67$^+$ nuclei, provides their identity through

A   B
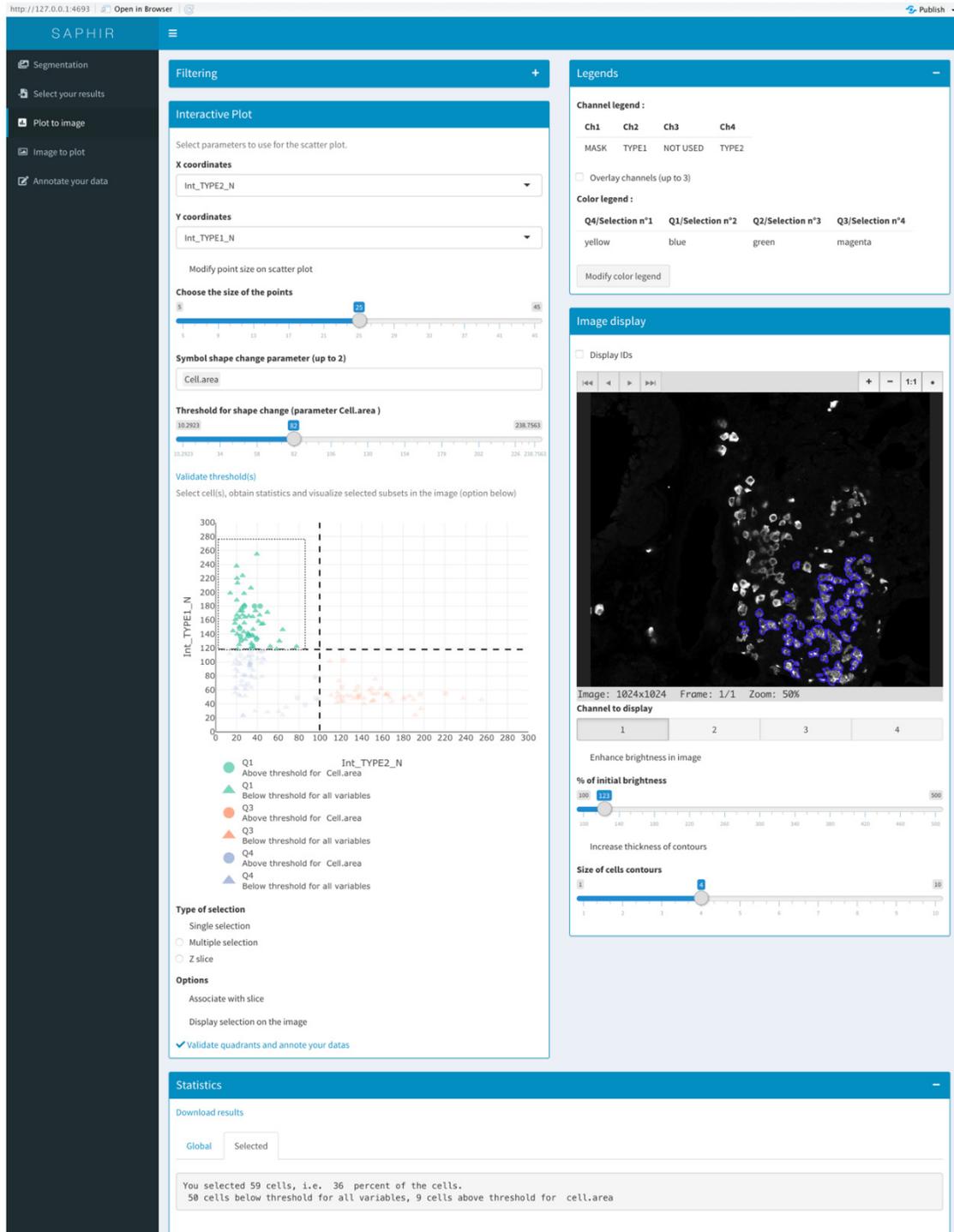
Figure 3. Menu "Plot to Image" of SAPHIR application. (**A**) Scatterplot of two channel intensity parameters used to gate COI. User can add a third parameter that is displayed on the scatterplot through a change of symbol shape for COI beyond a threshold defined by the user. (**B**) Visualization of selected cells in the image. Some of the available options are shown (displayed channel(s), channel overlay, cell identity number, brightness).
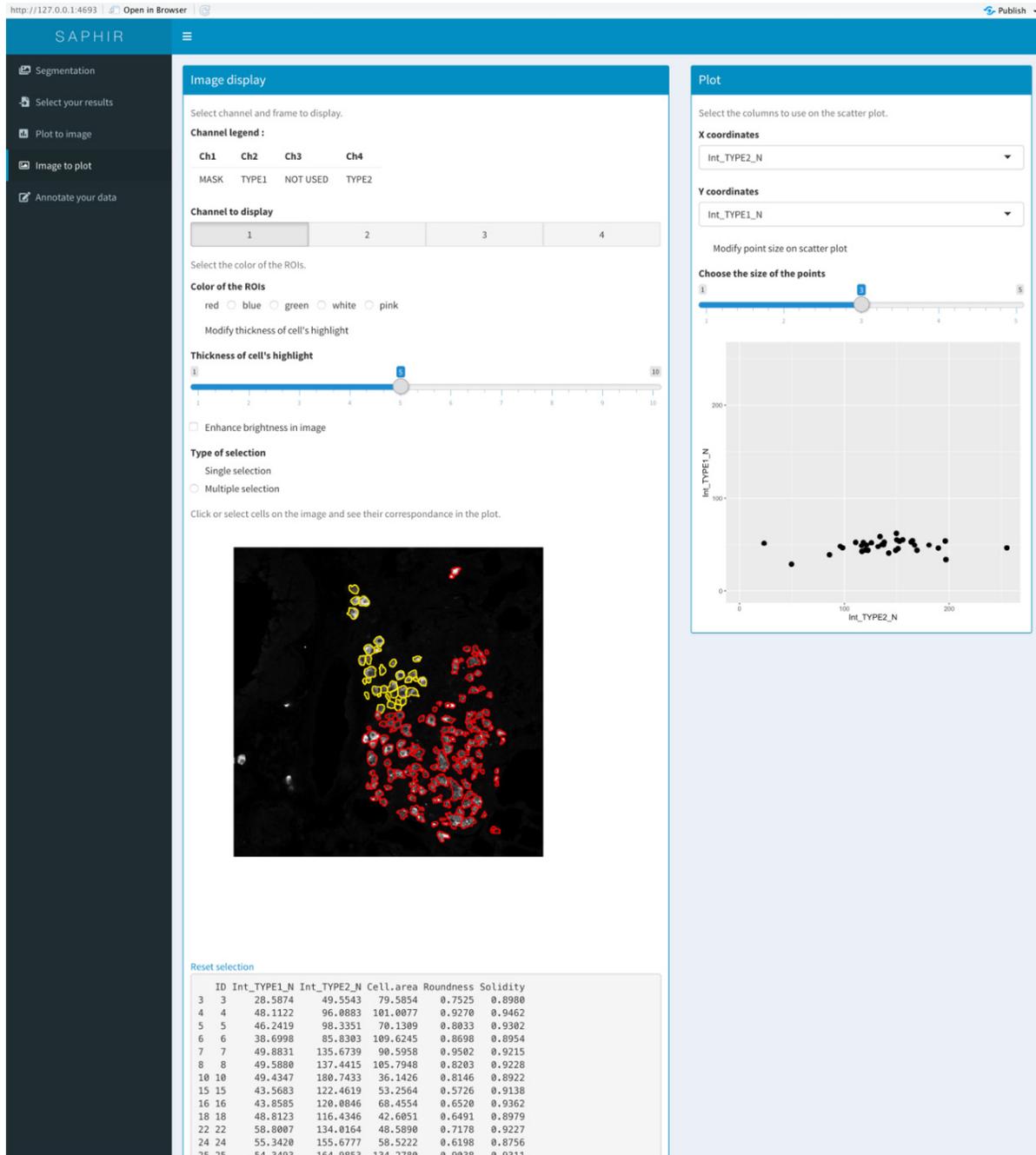
Figure 4. Menu "Image to Plot" of SAPHIR application. Selection of one region in the image (left) and visualization of COI of this region in a corresponding scatterplot where two parameters are displayed (right).

analysis of T and B cell staining intensity of Ki-67[+] cells, and determines their interaction with phagocytes. The second macro identifies the area of high density in proliferative cells as ROI, thanks to the DBSCAN algorithm[10] and determines whether previously analysed cells belong to this ROI. Finally, we obtained a result file that includes 176 COI (proliferative cells of the ROI) analyzed on 10 parameters (the intensity of LT and LB channels; the coordinates x, y, z; 3 physical parameters: area, circularity and solidity; interaction = 1 or not = 0 with a phagocyte; and the belonging = 1 or not = 0 to the region of strong proliferation).

Then, we integrated these data into the SAPHIR application and generated scatterplots and statistical analyses. Thus, with
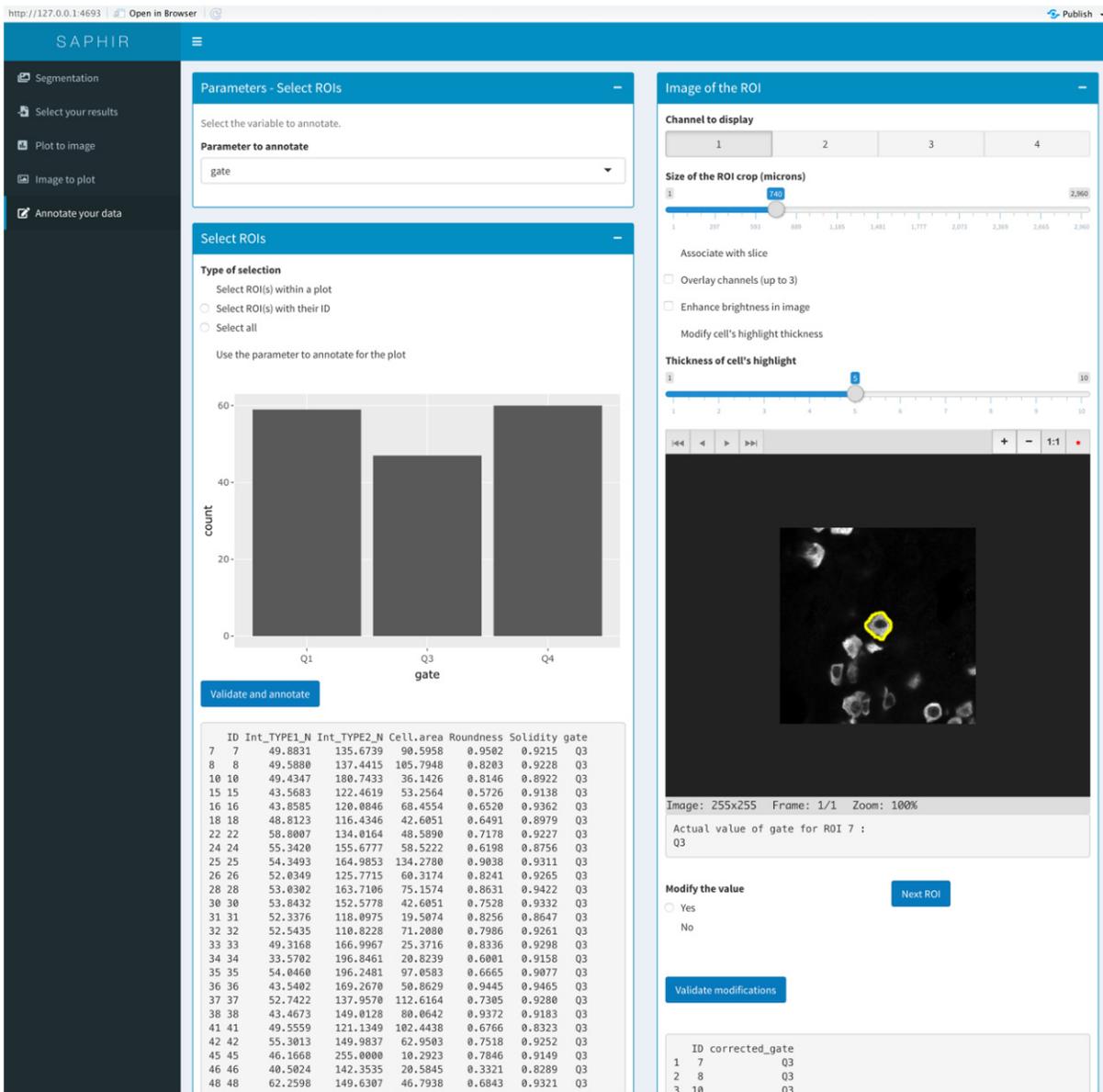
Figure 5. Menu "Annotate your data" of SAPHIR application. The results to annotate can be selected based on the Plot to Image scatterplot gated cells that have been previously saved (left panel). Each cell of the selection can be visualized in the image (right panel) with the possibility to change image size (magnification), the displayed channel(s), the slice (z optical section) and the brightness. Based on the different analysis, results for the selected COI can be modified and final results saved and exported.

the "Filtering" tab we selected only proliferative cells belonging to the area of intense proliferation at the periphery of the T cell zone from which we wanted to obtain statistical data (green rectangle in the upper left histogram of Figure 6). Then, these ROI-proliferative cells were split into B and T cells using the scatter plot (Figure 6, left mid-panel). We were thus able to identify proliferative T cells (lower right quadrant, red contour), proliferative B cells (upper left quadrant) and uncharacterized proliferative cells (lower left quadrant). Finally,

we used the "symbol shape change parameter" option to highlight ROI-proliferative cells that interacted with phagocytes (circles vs triangles). Selecting T cells, we could localize each of them on the optical slices and check their interaction or not with phagocytes directly into the image (Figure 6, right panel and higher magnification inserts).

Finally, we obtained exportable statistical data for all gated cells validated or not by direct visualization onto the image
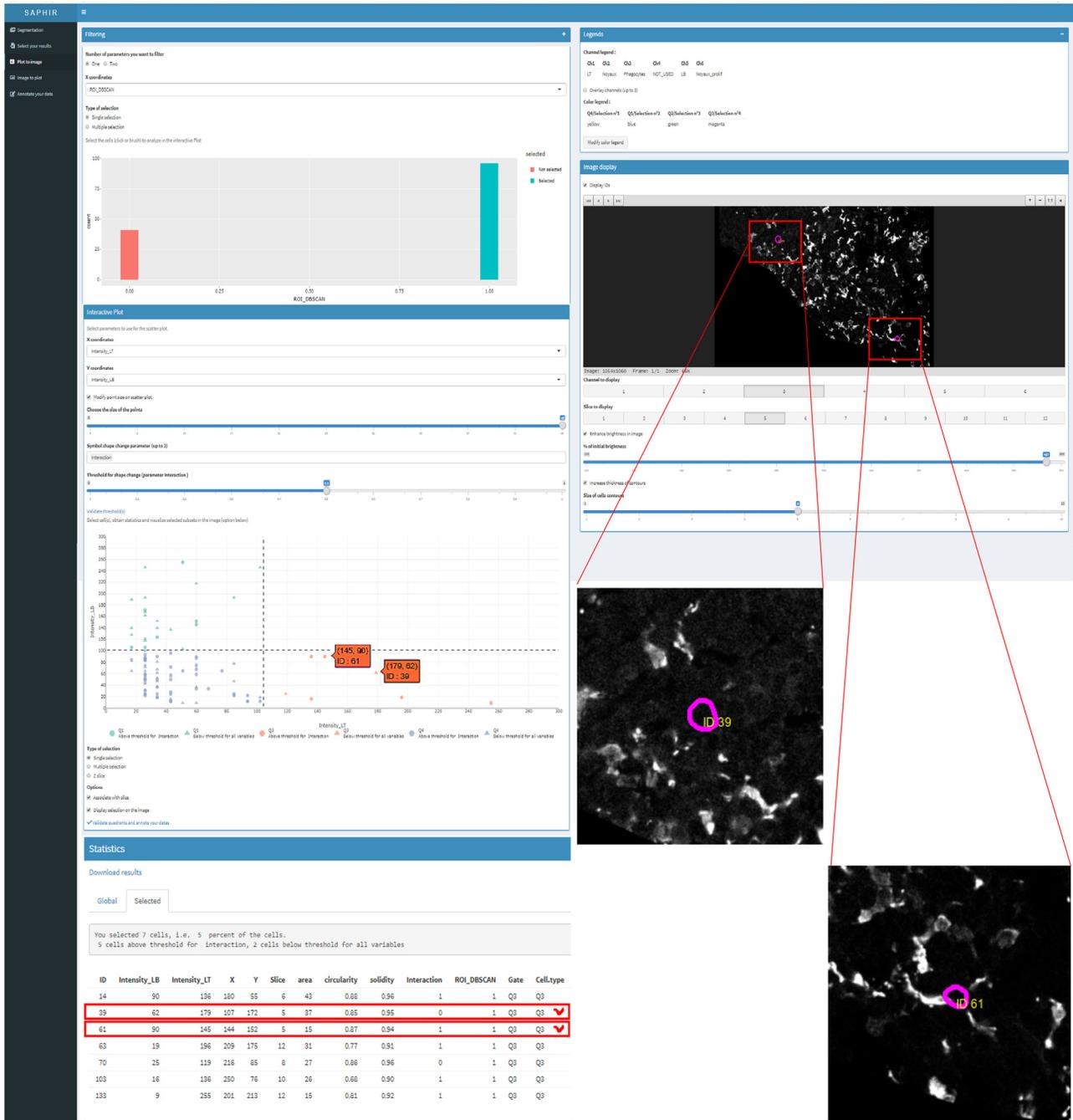
Figure 6. Use case: analysis of proliferative immune effector cells interacting with phagocytes in the T cell zone of murine Peyer's patches. Upper left: Proliferative cells belonging to the proliferative area of the T cell zone periphery (ROI_DBSCAN) were selected using the Filtering tab (green rectangle). Mid-left: ROI proliferative cells were split into T (orange symbols) and B cells (green symbols) on the scatterplot and phagocyte-interacting cells highlighted (circles) thanks to the symbol shape change parameter. Right: Proliferative T cells selected in the lower left scatter plot were visualized on each optical section (magenta contour). On the shown optical section two proliferative T cells were observed. Higher magnification of the boxed area showed phagocytes (grey cells) interacting with one proliferative T cell (ID 61) but not with the other (ID39). T cell ID numbers allow to locate them back in the scatterplot (pointed out in orange) and in the statistical table (boxed in red) and to check that one but not the other was indeed identified as interacting with phagocytes (circle vs triangle in the scatterplot and 1 vs 0 in the interaction column).

(Figure 6, lower left panel with selected cells boxed in red). In our case, no errors were detected but if they were, the annotation menu would have allowed to correct them. Finally, based on the analysis of this image and several other similar images, we could document the propensity of proliferating T cells to interact with phagocytes at the periphery of the T cell zone upon stimulation (manuscript in preparation).

Conclusion

The SAPHIR application provides a simple and user-friendly interface to obtain quantitative data from tissue images as well as COI positioning in the tissue. It is based on the interactivity between quantitative data extracted from the image and the image itself. This is especially useful to locate cells that have been selected based on the extracted data. It may thus reveal particular and unsuspected locations for the selected cells within the tissue. It can also be used to check some parameters of the selected cells (e.g., phenotype and interacting partners) on the image to correct any possible errors generated by the automatic analysis (e.g., segmentation and cell identity). Further developments of SAPHIR will include use of pyramidal images to minimize loading and analysis time and will implement R spatial analysis algorithm to analyze cell clusters and cell neighborhood[12].

Data and availability

SAPHIR is provided with segmentation data from two demo pictures (https://github.com/elodiegermani/SAPHIR/tree/master/Demonstration%20files): one with only two intensity channels and one optical slice to test the application in a very easy and quickly way, and a more complex one with 6 intensity channels, 12 optical slices and two additional parameters (belonging to a ROI and cell-cell interaction) for advanced testing of the application.

These data are parts of several projects conducted in the Center of Immunology of Marseille-Luminy (France). Acquisition of the images was made with spectral confocal microscopy and analyzed with Fiji with the provided macro.

Software availability

Source code available from: www.github.com/elodiegermani/SAPHIR

Archived source code as at time of publication: https://doi.org/10.5281/zenodo.4656582[13]

License: GNU General Public License v3.0

References

1. Dodt HU, Leischner U, Schierloh A, *et al.*: **Ultramicroscopy: three-dimensional visualization of neuronal networks in the whole mouse brain.** *Nat Methods.* 2007; **4**(4): 331–336.
   PubMed Abstract | **Publisher Full Text**

2. Gerner MY, Kastenmuller W, Ifrim I, *et al.*: **Histo-cytometry: a method for highly multiplex quantitative tissue imaging analysis applied to dendritic cell subset microanatomy in lymph nodes.** *Immunity.* 2012; **37**(2): 364–376.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Goltsev Y, Samusik N, Kennedy-Darling J, *et al.*: **Deep Profiling of Mouse Splenic Architecture with CODEX Multiplexed Imaging.** *Cell.* 2018; **174**(4): 968–981.e915.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. Lelouard H, Mailfert S, Fallet M: **A ten-color spectral imaging strategy to reveal localization of gut immune cell subsets.** *Zeiss Application note.* 2018.
   Reference Source

5. Legland D, Arganda-Carreras I, Andrey P: **MorphoLibJ: integrated library and plugins for mathematical morphology with ImageJ.** *Bioinformatics.* 2016; **32**(22): 3532–3534.
   PubMed Abstract | **Publisher Full Text**

6. Weigert M, Schmidt U, Haase R, *et al.*: **Star-convex Polyhedra for 3D Object Detection and Segmentation in Microscopy.** *2020 IEEE Winter Conference on Applications of Computer Vision (WACV); 2020 1–5 March 2020.* 2020; 3655–3662.
   Reference Source

7. Stringer C, Michaelos M, Pachitariu M: **Cellpose: a generalist algorithm for cellular segmentation.** *bioRxiv.* 2020; 2020.2002.2002.931238.
   Publisher Full Text

8. Winfree S, Khan S, Micanovic R, *et al.*: **Quantitative Three-Dimensional Tissue Cytometry to Study Kidney Tissue and Resident Immune Cells.** *J Am Soc Nephrol.* 2017; **28**(7): 2108–2118.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. Schindelin J, Arganda-Carreras I, Frise E, *et al.*: **Fiji: an open-source platform for biological-image analysis.** *Nat Methods.* 2012; **9**(7): 676–682.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

10. Hahsler M, Piekenbrock M, Doran D: **dbscan: Fast Density-Based Clustering with R.** *J Stat Softw.* 2019; **91**(1): 30.
    Publisher Full Text

11. Wagner C, Bonnardel J, Da Silva C, *et al.*: **Differentiation Paths of Peyer's Patch LysoDCs Are Linked to Sampling Site Positioning, Migration, and T Cell Priming.** *Cell Rep.* 2020; **31**(1): 107479.
    PubMed Abstract | **Publisher Full Text**

12. Stoltzfus CR, Filipek J, Gern BH, *et al.*: **CytoMAP: A Spatial Analysis Toolbox Reveals Features of Myeloid Cell Organization in Lymphoid Tissues.** *Cell Rep.* 2020; **31**(3): 107523.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. elodiegermani, matfallet: **elodiegermani/SAPHIR: Second release: Minor modifications (Version v1.1.1).** *Zenodo.* 2021.
    http://www.doi.org/10.5281/zenodo.4656582

Open Peer Review

Current Peer Review Status: ✓ ✓

Version 2

Reviewer Report 26 April 2021

https://doi.org/10.5256/f1000research.55654.r82892

✓   **Seth Winfree**

Division of Nephrology, Department of Medicine, Indiana University School of Medicine, Indianapolis, IN, USA

Thank you for taking the time to respond and thoroughly and thoughtfully to my review. I found your responses to my concerns regarding "Rationale for tool development," "Interpretation of output, results tool performance", and figure presentation more than adequate.

With regard to my minor comments: I believe that addressing bias and inaccuracy is an important and worthwhile pursuit, however, it is unclear how begin able to reclassify cells would specifically address this. In fact, it could be argued that the included functionality could lead to more biased data by allowing for manipulation without clear mechanisms for tracking analysis and changelogs. This appears to be an opportunity for additional contribution in this space (image-based cytometry) and in all multi-step complicated image analysis protocols. How should we qualify segmentation and classification accuracy? However, this is obviously beyond the scope of your work and I would not consider it an impediment for approval.

Thank you for taking the time for putting together SAPHIR, I hope to incorporate it into my workflows that, at times, bounce between R and ImageJ/FIJI.

Competing Interests: Developer of a tool for use in tissue cytometry.

Reviewer Expertise: light microscopy, multispectral microscopy, image analysis, tissue cytometry, kidney immunology, neutrophil biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 23 April 2021

https://doi.org/10.5256/f1000research.55654.r82891

✔ **Mahmoud Ahmed** [iD]

Department of Biochemistry and Convergence Medical Sciences, Institute of Health Sciences, Gyeongsang National University School of Medicine, Jinju, South Korea

Based on their response, the modified manuscript, and the updated app, the authors addressed most of the issues raised in the review. These include a description of the app's development process and a detailed interpretation of the results in the use case. A few changes were made in the app itself, including adding a segmentation menu with links to the required macros, a docker file, and a modified script to address the issues with installing and running the app.

By contrast, the authors made a few arguments and committed to following up with changes to address other issues. Those are

- They explain that the goal of the example use case is to illustrate the functionality rather than provide a real example of the analysis workflow.

- No viable option was provided for users to try/test the app before investing in installing and running it locally. The authors explained that they are working on a strategy to address this issue.

- No substantial improvements to the code were made however the authors committed to cleaning up the code in the future.

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: image-analysis, genomics, adipocyte-differentiation, autophagy

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

- -

Version 1

Reviewer Report 30 November 2020

? **Seth Winfree**

Division of Nephrology, Department of Medicine, Indiana University School of Medicine,

Indianapolis, IN, USA

SAPHIR is an R based tool for image cytometry. SAPHIR provides a novel R based Shinny application for analyzing an existing segmentation, an associated image and measured features of the segmented objects. The tool is intended to meet an unmet need for a flow cytometry-like exploration tool that enables 1) the highlighting of cells-of-interest in the original image volume defined by cells gated on staining intensity or morphological features and 2) the plotting of cells to a scatterplot of staining intensity or morphological features based on an image region-of-interest (ROI).

Major criticisms:
1. Rationale for tool development: A number of tools previously described or referenced in this manuscript include functionality similar to or overlapping with SAPHIR albeit with different implementations. Although the rationale presented by the authors is clearly explained, it is unclear what the rationale is for a new implementation of these functionalities in R (versus CellAnalyst in Python and CodexMAV and Volumetric Tissue Exploration and Analysis in Java). A tool in R could prove very useful especially in light of major R-based projects in transcriptomic analyses and data visualization and the platform's unrivaled strength in statistical analysis, but this is not addressed. Furthermore, R's strengths and weaknesses in imaging (versus other platforms) are not considered or addressed. These points require further justification and discussion.

2. Rationale for tool development: It is unclear if this tool is designed as a generalizable solution or intended only for ImageJ/FIJI users as it is tied to ImageJ/FIJI functionality to generate a number of input files-especially the segmentation which is based on the ImageJ ROI class. Although there may be a path or utility for using ROIs from Icy, CellProfiler, Ilastik or other common segmentation tools, this is not discussed and needs to be clarified.

3. Interpretation of output results and tool functionality: The authors mention the importance of supporting multiplexed fluorescence images in the introduction, "more than 10 channel measurements", but in neither the results nor supporting material is it clear to what degree SAPHIR supports these kinds of datasets (e.g. Lelouard *et al.* the manuscript). Furthermore, the support for 3D datasets is unclear, as ImageJ ROIs are limited to 2D and the link referenced in the supporting material for the "Complex Image Analysis", https://zenodo.org/record/4056001#.X8HVRC2ZM_W, suggests the scripts used to generate the input files only consider a single optical slice. Please clarify these functionalities.

4. Tool performance: The claim of the application's quickness is hard to assess. Although, SAPHIR appears to function well with the demonstration datasets the authors also qualify their quickness claim: "Interactivity between scatterplot and image is optional to avoid slow interaction when high-resolution images are used. In the latter case, we recommend users to perform their gating before allowing their selection to be shown in the image by ticking the appropriate box." This apparent contradiction is unqualified, beyond "high-resolution", the claim of quickness is poorly clarified. Please elaborate, and demonstrate if necessary, this limitation-for instance with larger or more complex datasets.

5. Manuscript: Figures are difficult to interpret due to size and resolution, especially figure 6 in which the "orange" color could not be easily identified nor the "magenta contour". This is

critical to demonstrating the behaviors of SAPHIR. Please adjust to facilitate the interpretation of the results.

Minor criticisms:
1. <u>Tool performance:</u> One of the major goals appears to be addressing "bias and misinterpretation" but this is not directly discussed in the results. It is unclear in what way(s) SAPHIR directly addresses bias.

2. <u>Tool performance:</u> The title and introduction suggest that large or mesoscale datasets that can be generated by modern imaging modalities are a unique challenge due to their size and complexity. It is unclear to what degree SAPHIR is able to analyze datasets with a large number of cells. This is an important qualification that is not directly addressed.

Is the rationale for developing the new software tool clearly explained?
Partly

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Yes

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Partly

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Partly

Competing Interests: I authored a article demonstrating a tool I developed with similar functionality.

Reviewer Expertise: light microscopy, multispectral microscopy, image analysis, tissue cytometry, kidney immunology, neutrophil biology

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

> Author Response 23 Mar 2021
> **Elodie GERMANI**, Université de Rennes 1, Rennes, France
>
> Dear reviewer, we would like to thank you for all your comments on our app and article that helped us improve them.

We studied every remark and wanted to give you an answer :

Major criticisms:
1. Rationale for tool development: A number of tools previously described or referenced in this manuscript include functionality similar to or overlapping with SAPHIR albeit with different implementations. Although the rationale presented by the authors is clearly explained, it is unclear what the rationale is for a new implementation of these functionalities in R (versus CellAnalyst in Python and CodexMAV and Volumetric Tissue Exploration and Analysis in Java). A tool in R could prove very useful especially in light of major R-based projects in transcriptomic analyses and data visualization and the platform's unrivaled strength in statistical analysis, but this is not addressed. Furthermore, R's strengths and weaknesses in imaging (versus other platforms) are not considered or addressed. These points require further justification and discussion.

Due to the short format chosen for our manuscript (allowed initial length of 1000 words), it was not possible to include a discussion on the advantage/disadvantage of each software and its environment. CodexMAV is a commercial software but available for free as an ImageJ plugin. It needs FCS files to run. The software is very powerful to do spatial analysis (t-SNE) of cell populations in tissues, but we didn't test it in our datasets, yet. CellProfiler (Analyst) is very interesting in term of segmentation and analysis but as far as we know, scatterplots and images are not interconnected. VTEA is also a great and easy application to use, as it is an ImageJ plugin. However, at the time we needed quantitative analysis to be done, segmentation was included in VTEA without all the flexibility that was required for the segmentation of our datasets. We therefore asked the VTEA developer to add the possibility to load ROIs from ImageJ, which is done now. However, the strong interconnection between the image and the scatterplot that we wanted to have to control and correct cells identity was lacking. We have added a sentence in section Operation to explain why we have chosen R for developing SAPHIR application.

2. Rationale for tool development: It is unclear if this tool is designed as a generalizable solution or intended only for ImageJ/FIJI users as it is tied to ImageJ/FIJI functionality to generate a number of input files-especially the segmentation which is based on the ImageJ ROI class. Although there may be a path or utility for using ROIs from Icy, CellProfiler, Ilastik or other common segmentation tools, this is not discussed and needs to be clarified.

Following reviewer recommendation, we have now added a sentence in the data requirement section to help users with the conversion of COI obtained from other software into the Fiji .roi format.

3. Interpretation of output results and tool functionality: The authors mention the importance of supporting multiplexed fluorescence images in the introduction, "more than 10 channel measurements", but in neither the results nor supporting material is it clear to what degree SAPHIR supports these kinds of datasets (e.g. Lelouard *et al.* the manuscript). Furthermore, the support for 3D datasets is unclear, as ImageJ ROIs

are limited to 2D and the link referenced in the supporting material for the "Complex Image Analysis", https://zenodo.org/record/4056001#.X8HVRC2ZM_W, suggests the scripts used to generate the input files only consider a single optical slice. Please clarify these functionalities.

SAPHIR is supporting multichannel images and Z stack image (ijtiff is used to load the image). We indeed used 3D images but following 3D segmentation only the main COI corresponding to the largest nucleus area was kept for calculating the intensity of other channels. In Fiji, the ROI are linked to specific slices allowing this operation. Thus, we used 2D ROI of 3D images for SAPHIR analysis. This is now explained in the data requirement section.
Of note, it is now possible to obtain 3D ROI in Fiji using a Plugin from Thomas Boudier (https://imagejdocu.tudor.lu/plugin/stacks/3d_roi_manager/start#d_roi_manager).

4. Tool performance: The claim of the application's quickness is hard to assess. Although, SAPHIR appears to function well with the demonstration datasets the authors also qualify their quickness claim: "Interactivity between scatterplot and image is optional to avoid slow interaction when high-resolution images are used. In the latter case, we recommend users to perform their gating before allowing their selection to be shown in the image by ticking the appropriate box." This apparent contradiction is unqualified, beyond "high-resolution", the claim of quickness is poorly clarified. Please elaborate, and demonstrate if necessary, this limitation-for instance with larger or more complex datasets.

As the application could be slowed down with images >300MB, we decided to remove the term "quickly" from the abstract of the manuscript. Moreover, the limitation of 300MB is now specified in the implementation section. Big images (with multiple slices and channels) have to be stored on SSD drive to limit slow interaction. Since our application was mainly used for tissue section < 20 μm, we did not test datasets > 300MB, yet.

5. Manuscript: Figures are difficult to interpret due to size and resolution, especially figure 6 in which the "orange" color could not be easily identified nor the "magenta contour". This is critical to demonstrating the behaviors of SAPHIR. Please adjust to facilitate the interpretation of the results.

The resolution of the figures has been improved and Figure 6 has been modified to allow better visualization of specific parts of the panels. Moreover, the app tools now include options to adjust the cell contour thickness in the image and the spot size in the scatterplot.

Minor criticisms:
1. Tool performance: One of the major goals appears to be addressing "bias and misinterpretation" but this is not directly discussed in the results. It is unclear in what way(s) SAPHIR directly addresses bias.

By allowing to analyze for each cell of the scatterplot its position and its characteristics in the image, SAPHIR allows to check that the parameters obtained automatically with the segmentation and other analysis algorithms fit with the observation of the selected cells in

the image. Then, the annotation menu allows to correct any bad analysis (e.g., cell identity). We have now extended the use case section to better describe SAPHIR use.

2. Tool performance: The title and introduction suggest that large or mesoscale datasets that can be generated by modern imaging modalities are a unique challenge due to their size and complexity. It is unclear to what degree SAPHIR is able to analyze datasets with a large number of cells. This is an important qualification that is not directly addressed.

Our aim was clearly not to analyze mesoscale datasets. We apologize if that is the impression given by our title or introduction, although we see no allusion to it in the title and only a small reference in the introduction (i.e., "10 μm to more than 200 μm when clearing methods are used") but not related to the description of the application. Actually, as mentioned in the introduction our main goal was to obtain an interconnection between the representation of image parameters in scatterplot and the image itself in order to analyze positioning of scatterplot-selected cells in the image and to check and correct any analysis errors by direct observation of the selected cells.

We hope these responses and the new version of our article will satisfy you.

Competing Interests: No competing interests were disclosed.

Reviewer Report 11 November 2020

https://doi.org/10.5256/f1000research.29894.r73741

? **Trang Huyen Lai** [iD]

Department of Biochemistry and Convergence Medical Sciences, Institute of Health Sciences, Gyeongsang National University School of Medicine, Jinju, South Korea

Mahmoud Ahmed [iD]

Department of Biochemistry and Convergence Medical Sciences, Institute of Health Sciences, Gyeongsang National University School of Medicine, Jinju, South Korea

In this article, the authors describe a shiny app to load and explore the segmentation output of fluorescence images from ImageJ/Fiji. The app provides an interactive interface to filter, select, and annotate the regions of interest as defined by the ImageJ/Fiji macros. Besides, the app allows for performing statistical testing on the selected regions. The manuscript describes the motivation for developing the app, the basic functionality, and a use case. In the use case, the app was used to filter proliferative T and B cells based on cell membranes staining and determine that more T cells interacted with phagocytes in Peyer's patches images.

The article is written in clear and concise language. The motivation for developing the app and the description of the operation is sound. A few issues remain:

Major Issues
1. The implementation section focuses on describing the interface but doesn't explain much about the inner working of the app or the development process. (related, question 3).

2. Four types of files are needed to use the app: Images, Data files, ROIs Files, and Legend files. Each file needs to be processed and prepared to match the formats and requirements of the app. The user has to perform many steps, including using ImageJ and R, to process the files to be able to use SAPHIR. Besides, if the user wants to combine 2 images, he or she has to prepare 4 files for each image. These complicated steps could limit the replication of the analysis by others. It is better if the providers could explain the steps and the tools that they have used to prepare their files. Or if there is an option in SAPHIR that we can use to directly prepare the files.

3. The article explains well using the app to perform different filtering and selection on an example image. However, it doesn't say much about the interpretation of the outputs, or the context in which these decisions could be made. (related, question 4)

4. The use case is very brief and it seems like many of the analysis steps were done outside of the app using custom macros. This makes it hard to evaluate the use case or determining how the app contributed to the final calculations. The authors could provide a more complete description of the use case starting from the raw images and describe how the macros work. It would be helpful to provide links to the data and the tools they used as well.

5. The two examples provided in the article use a single image. To be able to make any claims on the performance (related, question 5), the authors might need to perform testing with multiple images or provide a comparison with existing tools.

Minor Issues
1. Several screenshots from the app were included in the manuscript as figures. It might be helpful to change the size/resolution of the images for clarity. Also, highlighting the important parts of the image would make it easier to understand the interface.

2. The app depends on ImageJ so it should be bundled with the app or explain how to obtain and add to the app.

3. Related. Trying to run the app using rstudio server on rocker/verse failed. A dependency libpoppler-cpp-dev was required but not declared.

4. Provide a viable demo, for example using shinyapps.io or rstudio.cloud would make it easier to evaluate the tool and lowers the entry cost for users.

5. The app is written as one long R script. This makes it very difficult to examine, modify, or extend the code. Factoring the code or making it more modular would be helpful.

6. After installing the most recent versions of the required packages, trying to run the app in a fresh rstudio session through an error related to shinyjs::extendShinyjs. I was able to get it

to run by adding functions = 'winprint' as an argument. Source

Is the rationale for developing the new software tool clearly explained?
Yes

Is the description of the software tool technically sound?
Yes

Are sufficient details of the code, methods and analysis (if applicable) provided to allow replication of the software development and its use by others?
Partly

Is sufficient information provided to allow interpretation of the expected output datasets and any results generated using the tool?
Partly

Are the conclusions about the tool and its performance adequately supported by the findings presented in the article?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: image-analysis, genomics, adipocyte-differentiation, autophagy

We confirm that we have read this submission and believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however we have significant reservations, as outlined above.

Author Response 23 Mar 2021
Elodie GERMANI, Université de Rennes 1, Rennes, France

Dear reviewer, we would like to thank you for all your comments on our app and article that helped us improve them.

We studied every remark and wanted to give you an answer :

Major Issues

1. The implementation section focuses on describing the interface but doesn't explain much about the inner working of the app or the development process. (related, question 3).

Following the recommendation of the reviewer we have now explain the development process of the app in the operation section.

2. Four types of files are needed to use the app: Images, Data files, ROIs Files, and

Legend files. Each file needs to be processed and prepared to match the formats and requirements of the app. The user has to perform many steps, including using ImageJ and R, to process the files to be able to use SAPHIR. Besides, if the user wants to combine 2 images, he or she has to prepare 4 files for each image. These complicated steps could limit the replication of the analysis by others. It is better if the providers could explain the steps and the tools that they have used to prepare their files. Or if there is an option in SAPHIR that we can use to directly prepare the files.

We thank the reviewer for this suggestion. We have now added in the data requirement section a description of the segmentation process. A segmentation menu to is available in the application where Fiji and the two macros used in the manuscript can be launched.

3. The article explains well using the app to perform different filtering and selection on an example image. However, it doesn't say much about the interpretation of the outputs, or the context in which these decisions could be made. (related, question 4)

We agree with the reviewer and based on the use case we now describe the results we could obtain and the conclusions we were able to draw using SAPHIR on a specific application, which is the immune response initiation in murine Peyer's patches. We also show how useful the filtering tool is to focus on a given population of cells (i.e., in our case proliferative cells of the ROI) and how interaction between the image and the scatter plot can be used to focus on a given cell of interest and to control that analysis is correct.

4. The use case is very brief and it seems like many of the analysis steps were done outside of the app using custom macros. This makes it hard to evaluate the use case or determining how the app contributed to the final calculations. The authors could provide a more complete description of the use case starting from the raw images and describe how the macros work. It would be helpful to provide links to the data and the tools they used as well.

We followed the recommendation of the reviewer and have extended the section dedicated to the use case (see answer to point 3). However, since the segmentation process is now described in detail in the data requirement section, we only briefly described this step to rather focus on the contribution of SAPHIR to quantitative image analysis. As explained in the main text, all data and macros are available on GitHub in the Demonstration Files.

5. The two examples provided in the article use a single image. To be able to make any claims on the performance (related, question 5), the authors might need to perform testing with multiple images or provide a comparison with existing tools.

The main goal of this manuscript was to provide a brief description of the application and its usefulness using two types of image as demos. It is now applied in the lab to study biological processes using dozens of images to validate statistics and its use will be included in several studies, but it was not the aim of this manuscript to already provide this kind of analysis. We indeed tested other tools before developing SAPHIR, such as VTEA and FlowJo, but none was really adapted to our needs as we especially wanted to have a strong interconnection between images and scatterplots.

Minor Issues

1. Several screenshots from the app were included in the manuscript as figures. It might be helpful to change the size/resolution of the images for clarity. Also, highlighting the important parts of the image would make it easier to understand the interface.

The resolution of the figures has been improved and Figure 6 has been modified to allow better visualization of specific parts of the panels. Moreover, the app tools now include options to adjust the cell contour thickness in the image and the size of the dots in the scatterplots.

2. The app depends on ImageJ so it should be bundled with the app or explain how to obtain and add to the app.

Link to Fiji is now provided in the segmentation menu of the app, where Fiji and the two macros used in the manuscript can be launched and applied to any image.

3. Related. Trying to run the app using rstudio server on rocker/verse failed. A dependency libpoppler-cpp-dev was required but not declared.

The dependency to lipoppler-cpp-dev is due to the use of the magick package, which is required to use the application. We also had this problem when trying to deploy our app on shinyapps.io. This problem can be solved using a Unix operating system instead of Windows. We have created a Docker file to use the app without this kind of issues.

4. Provide a viable demo, for example using shinyapps.io or rstudio.cloud would make it easier to evaluate the tool and lowers the entry cost for users.

We are soon going to deploy a demo version on shinyapps.io. However, it will only be operable on our demo images and the user will not be able to change the images to use.

5. The app is written as one long R script. This makes it very difficult to examine, modify, or extend the code. Factoring the code or making it more modular would be helpful.

We agree with the reviewer and we are going to modularize the code in the following months.

6. After installing the most recent versions of the required packages, trying to run the app in a fresh rstudio session through an error related to shinyjs::extendShinyjs. I was able to get it to run by adding functions = 'winprint' as an argument. Source

We modified the script in order to correct this. Indeed, the package shinyjs was recently updated and a new argument in the function "extendShinyjs" was made mandatory. We added this modification to the latest version of our script.

We hope these responses and the new version of our article will satisfy you.

Competing Interests: No competing interests were disclosed.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias

- You can publish traditional articles, null/negative results, case reports, data notes and more

- The peer review process is transparent and collaborative

- Your article is indexed in PubMed after passing peer review

- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com   F1000Research