Title

Rapid analysis and exploration of fluorescence microscopy images.

Permalink

Authors

Pavie, Benjamin
Rajaram, Satwik
Ouyang, Austin
et al.

Publication Date

2014

DOI

10.3791/51280

Peer reviewed

Video Article
Rapid Analysis and Exploration of Fluorescence Microscopy Images

Benjamin Pavie*[1], Satwik Rajaram*[1], Austin Ouyang[2], Jason M. Altschuler[1,3], Robert J. Steininger III[1], Lani F. Wu[1], Steven J. Altschuler[1]

[1]Green Center for Systems Biology, UT Southwestern Medical Center

[2]Advanced Imaging Research Center, UT Southwestern Medical Center

[3]Princeton University

*These authors contributed equally

Correspondence to: Lani F. Wu at Lani.Wu@utsouthwestern.edu, Steven J. Altschuler at Steven.Altschuler@utsouthwestern.edu

Abstract

Despite rapid advances in high-throughput microscopy, quantitative image-based assays still pose significant challenges. While a variety of specialized image analysis tools are available, most traditional image-analysis-based workflows have steep learning curves (for fine tuning of analysis parameters) and result in long turnaround times between imaging and analysis. In particular, cell segmentation, the process of identifying individual cells in an image, is a major bottleneck in this regard.

Here we present an alternate, cell-segmentation-free workflow based on PhenoRipper, an open-source software platform designed for the rapid analysis and exploration of microscopy images. The pipeline presented here is optimized for immunofluorescence microscopy images of cell cultures and requires minimal user intervention. Within half an hour, PhenoRipper can analyze data from a typical 96-well experiment and generate image profiles. Users can then visually explore their data, perform quality control on their experiment, ensure response to perturbations and check reproducibility of replicates. This facilitates a rapid feedback cycle between analysis and experiment, which is crucial during assay optimization. This protocol is useful not just as a first pass analysis for quality control, but also may be used as an end-to-end solution, especially for screening. The workflow described here scales to large data sets such as those generated by high-throughput screens, and has been shown to group experimental conditions by phenotype accurately over a wide range of biological systems. The PhenoBrowser interface provides an intuitive framework to explore the phenotypic space and relate image properties to biological annotations. Taken together, the protocol described here will lower the barriers to adopting quantitative analysis of image based screens.

Video Link

The video component of this article can be found at http://www.jove.com/video/51280/

Introduction

Over the past decade, rapid advances in imaging technology have given many labs the ability to perform high-throughput microscopy. The challenge has now shifted from one of imaging to one of analysis. How can we characterize, compare, and explore the torrent of complex yet subtle phenotypes generated by a typical high-throughput imaging screen?

A number of image-analysis and informatics platforms have given users sophisticated toolboxes[1-3] for extracting biological information from large collections of images. Yet many significant obstacles remain in analyzing data from high-throughput image-based screens. Difficulties involved with choosing appropriate characterizations of the images and fine tuning of algorithmic parameters (to the system of interest) often result in long setup times, particularly for users without image-analysis expertise. In particular, cell segmentation, the process of identifying individual cells in an image, is a major bottleneck in this regard. Segmentation can be highly sensitive to experimental parameters such as cell type, cell density and bio-markers, and frequently requires repeated adjustment for a large data set. For these reasons, image analysis is often an independent, terminal step performed by specialists rather than an integrated part of the experimental workflow. This precludes the rapid feedback cycle between analysis and experiment, a crucial part of assay development.

Here we describe an alternative workflow that is optimized for high-throughput fluorescence microscopy and avoids many of the aforementioned difficulties. To this end, we make use of PhenoRipper[4], an open-source software platform for rapid exploration and analysis of microscopy images. Significantly, PhenoRipper avoids many of the challenges involved with cell segmentation by not attempting to identify single cells. Instead, PhenoRipper divides images into pixel blocks of subcellular size and characterizes images in terms of the blocks they contain. Specifically, PhenoRipper profiles blocks in terms of marker-colocalization-based features and automatically identifies[4] the most representative block types in the training images. PhenoRipper then assigns to each block the most similar representative block type, and finally characterizes images based on the types of the blocks they contain.

Compared to more traditional single-cell-based approaches, this approach requires minimal fine tuning and expertise. As such, users only need to set two parameters: the block size and a threshold intensity (to discard noncellular portions of an image), both of which are set with graphical feedback. Importantly, the workflow described here has been shown[4] to scale easily to much larger data sets, where the speed and minimal fine tuning provides large time and reliability gains. In practice, we find that this approach reduces the time required both for setup and running by multiple orders of magnitude[4].

The primary output of PhenoRipper is a visual representation of image similarity, which enables the user to identify groups of experimental conditions that cause cells to display similar phenotypes. The groupings PhenoRipper finds typically have comparable "biological interpretability" to those generated by more time-consuming, cell-based methods[4]. In practice, this means that often, within half an hour of performing a typical 96-well fluorescence microscopy experiment, an experimenter without prior image-analysis experience can get a reliable readout about the performance of his or her experiment. The experimenter can then start exploring the relationships between different experimental conditions and relating these relationships to image phenotypes.

We demonstrate this workflow here by analyzing the effects of drugs in distinct mechanistic classes on HeLa cells stained for DNA and cytoskeletal markers. Images of cells treated with the same class of drugs were grouped together, and poor quality images (staining artifacts, out of focus cells and so on) appeared as outliers, making them easy to identify and potentially discard.

While this workflow is useful for a large number of image-based assays, there are clearly many situations where a different approach could potentially be more informative. For example, the output from PhenoRipper is primarily a relative comparison of fluorescence patterns across experimental conditions. If an absolute characterization of a specific single cell trait was required instead (for example the number of speckles in a cell), a single-cell-based analysis would be required. Nonetheless, even in this type of situation, PhenoRipper is likely to detect changes in these single-cell phenotypes and therefore be a useful tool for assay optimization.

Protocol

1. Preparing Samples and Imaging

1. Cell culture
    1. Split cultures routinely and maintain at 20-70% confluence. One day prior to the experiment, grow HeLa cells in DMEM (Dulbecco's Modified Eagle Medium, see Table of Materials) and supplemented with 10% FBS (Fetal Bovine Serum) and 1x penicillin/streptomycin.
    2. On day of experiment, split cells growing in culture (about 50% confluent).
    3. Wash cells with PBS (Phosphate Buffered Saline), expose to approximately 2 ml of trypsin/EDTA (Ethylenediaminetetraacetic acid) for a few seconds. Aspirate excess trypsin/EDTA and keep cells bathed in a thin film of trypsin/EDTA in the hood at RT for ~3 min.
    4. Mechanically shake flasks twice and detach cells with serum supplemented media. Count cells and dilute to 5,000 cells/50 µl of media and immediately plate 50 µl/well onto a 384-well glass bottom plate using an automated liquid handling robot (only 30 wells of the plate are used here).
    5. Incubate plate at RT for 30 min to minimize edge effects in wells[5]. After RT incubation, place plates in a 37 °C/5% CO_2 incubator overnight.
    6. Each drug will be assayed in duplicate. Add 25 µl of drug to the corresponding wells on the plate already containing media.
    7. Fix plate with 4% paraformaldehyde solution in PBS at 24 hr post drug treatment.

2. Staining
   The specific primary antibodies, fluorescently-labeled secondary antibodies and other fluorescent stains that were used here are listed in Table of Materials.
    1. Permeabilize fixed cells with 0.2% Triton X-100 solution in TBS (Tris-buffered saline) for 3 min, and wash with TBST (Tris-buffered saline + Tween 20) 3x. Add 5% BSA (Bovine Serum Albumin) solution in TBST for blocking.
    2. Incubate plate at RT for 2 hr and then wash with TBST 3x.
    3. Make primary antibody dilutions in 5% BSA in TBST (see Table of Materials) and add 10 µl/well. Incubate plate at RT for 2 hr, and then wash 3x with TBST.
    4. Make secondary antibody dilutionin 5% BSA in TBST (see Table of Materials) and add 10 µl/well.
    5. Incubate plate in the dark at RT for 2 hr and then wash twice with TBST.
    6. Prepare dilutions for Hoechst and phalloidin (see Table of Materials) in PBS and add 10 µl/well.
    7. Incubate the plate in the dark for 30 min at RT.
    8. After 2 TBST washes, cover plates in foil and store at 4 °C O/N.

3. Imaging
   Acquire images using an epifluorescence microscope using a 20X objective with 1x1 camera binning. Acquire sixteen images for each well, for a total of 480 images (x3 channels).

2. Analyzing Images

Download and install PhenoRipper from phenoripper.org, and launch PhenoRipper to begin analysis of images. PhenoRipper currently supports TIFF, PNG, and JPEG and all other formats supported by MATLAB (microscope software typically support export into TIFF).

2.1 Loading Data

1. No standard naming/organization convention exists for microscopy images, which can make loading images and associating metadata (*i.e.* experimental annotation such as treatment type, well number and so on) the most challenging part of the workflow. Choose one of the presented methods (described below).

1. PhenoLoader: Choose this option to semi-automatically load data when the file name and/or directory structure contain enough information to group images as desired (*e.g.* B16\HeLa_Drug5-DAPI.png might indicate that the image is from well B16, cells are treated with the drug whose ID is 5, and the image captures the nuclear DAPI stain - with this naming convention, all images from well A1 or treated with Drug 5 could be easily grouped). The wizard will guide the user step by step to associate metadata with image groups and define the markers as follows:
    1. Click on "File Structure" to access to the PhenoLoader window.
    2. Click on "Define Root Directory" to select the directory containing the image dataset. Then filter the files that are not used (Unused Files) for the analysis (optional, can be left empty).
    3. Click on "Define Markers" to specify the biomarkers used in the experiment and associate them with the images. Specifically, click on an image file then select a portion of the filename that uniquely identifies the biomarkers. Finally, enter a name for each biomarker.
    4. Click on "Define Metadata" to associate information with each image. Specifically, click on an image file, and then select a portion of the filename identifying image information (*e.g.* "B16" for the well name). A window asking for a group name (*e.g.* "Well") will pop up, enter a name. Members of the group extracted from the filename will be shown on the group table (*e.g.* the full list of well names). To add additional information not present in the filename, click on the "(optional) Next" button. Click on "Add Metadata Group", enter the additional group name (*e.g.* "Drug") and for each value of the predefined group (*e.g.* "Well"), enter its value (*e.g.* "Taxol"). There is no limit to the number of additional groups that can be added.
    5. Finally, click on "Create Metadatafile" to generate the metadata file and continue the analysis. If a metadata file has been previously generated, the steps above may be avoided by directly loading the file using the option "Metadata File" (described in 3 below).

    2. Regular Expression Based Templates: Regular expressions are a standard approach for matching patterns in text. Choose this option when the file name and/or directory structure contain enough information to group images and the naming convention matches one of the pre-existing templates. Custom templates based on regular expressions may also be manually defined and reused. This advanced option is only recommended for users familiar with regular expressions.
    3. Metadata File: Choose this option for ultimate customizability. Define and load a comma separated values (CSV) file containing filenames and metadata associated with each image. An example of the file format is available on the download part of the web site (phenoripper.org). Simply load the file through the interface to display the information it contains.

2. Select the metadata field that defines replicates in the experiment to avoid redundant sampling. For data sets that contain replicates (*e.g.* all images in a well should be similar), grouping them together (*e.g.* by well) can improve performance. In most cases it is reasonable to choose the default option "Random (don't know)" which selects a random subset.

2.2 Setting Analysis Parameters

Click on "Set Parameters" to define the parameters required to process the dataset. The parameters to define are on the left panel. The right panel displays a merged sample of the dataset. Click on the left/right arrows to switch among different images.

1. Threshold Intensity: Choose an appropriate threshold value to roughly identify the cellular portions (rather than the background, noncellular regions) of images. A threshold is precalculated by PhenoRipper, but may be adjusted to improve highlighting of cellular regions using the scrollbar available on top of the image. If a chosen threshold does not work across all images, lower the threshold so as to include noncellular regions rather than dropping cellular regions.
2. Block Size: Choose an appropriate block size (in pixels) to grid the image. Blocks of specified size will be superimposed on the image. Adjust block size to obtain an average of 20-30 blocks/cell. To superimpose the grid over the image click on the check box in front of the "Block Size" field.
3. Optional Parameters: If the auto-scaled images do not display markers with desired relative intensities (*e.g.* one color is saturated) or if markers need to be dropped from the analysis, use the "Adjust Channel Intensity" option. To identify cellular regions on a subset of markers use the "Channels used for Threshold" option. Default choices of other parameters are usually adequate, but if required, adjust them using the "Advanced Analysis Options".
4. Run the analysis by pressing "PhenoRip" on the main application panel. PhenoRipper will automatically identify recurrent block types in the images. Phenotypic profiles will be generated for each image based on the fractions of the different block types in it.

3. Exploring Results

1. Explore the relationships between the phenotypic profiles of the images using the PhenoBrowser (**Figure 1**). The PhenoBrowser interface consists of four panels: the top left panel is a graphical representation (scatter plot) of the relative similarity between different images/conditions, the two panels on the right display sample images of selected conditions, and the final panel (bar chart) compares their phenotypic profiles.
2. "Discard Empty Frames" using the "Data Processing" menu.
3. Remove poor quality images: First, to study individual images turn off grouping from the "Data Processing" menu. Next, click through the outlier points in the scatter plot and discard low quality images (*e.g.* out of focus, staining artifact).
4. Explore the data: Decide which metadata field defines a basic unit of comparison (*e.g.* to compare drugs, images with the same value of the metadata field "drug" need to be grouped together). Collapse all images with the same value of this metadata field into a single data point with "Group By". For example, grouping by drug will average the images within each drug to yield one point per drug. Nearby (or distant) points in the scatter plot reflect conditions that elicit similar (or dissimilar) cellular phenotypes. Color or label points (data display menu) based on available metadata to aid interpretability. The 3D plot is rotatable by clicking on the "rotatable" check box and by clicking on the left button and dragging the mouse over it.
5. Select two different points to compare them directly. The image panels will display sample images for each point, and the bar chart panel will show the block types whose occurrence frequencies are most different between the two points.

6. Launch the clustergram from the "Clustergram menu" to provide a more global view of the relationship between block types and experimental conditions. By grouping both block types and conditions, this representation highlights the cellular phenotypes that best distinguish groups of conditions (**Figure 5**).

Representative Results

Here we tested the ability of this workflow to group drugs based on their mechanism of action. HeLa cells were seeded in 30 wells of a 384-well plate and stained for DNA/actin/α-tubulin. The specific primary antibodies, fluorescently-labeled secondary antibodies, and other fluorescent stains that were used are listed in Table of Materials. Wells were treated with 15 drugs belonging to three mechanistic classes (histone deacetylase inhibiting, microtubule targeting and DNA damaging) for 24 hr and imaged using an epifluorescence microscope at 20X objective magnification. Sixteen images were acquired for each well, for a total of 480 images (x3 channels). The data described here can be downloaded from www.phenoripper.org.

As described above, PhenoRipper was used to analyze these images. The foreground threshold was set to 5% (*i.e.* 5% of the maximum possible intensity) and the block size to 20 pixels (**Figure 2**). Analysis of this data set took approximately 10 min on a standard desktop computer (a 64 bit, 8 core machine at 3.2 GHz, with 16 GB RAM). The PhenoBrowser interface (**Figure 3**) was used to explore the results. Imaging artifacts such as poor staining and empty frames were easy to identify since they stood out as clear outliers (**Figure 4**). The relative phenotypic similarities between images are visualized in the top left panel of **Figure 3** using multidimensional scaling. Each point represents a single image colored by its drug's mechanism of action. This plot shows that phenotypic profiles can be used to groups drugs by their mechanism of action. A clustergram of the data (**Figure 5**) relates this grouping to the phenotypic profiles.
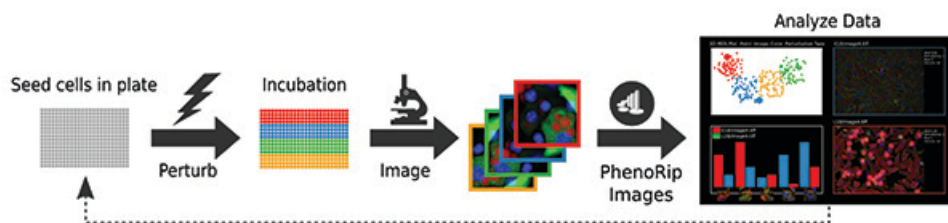

Figure 1. Workflow for fluorescence microscopy using PhenoRipper. Cell lines are seeded into a plate, then perturbations are added to each well and the plate is incubated. Images are acquired using an automated microscope and then rapidly analyzed using PhenoRipper. Thus, results of the current experiment are known shortly after imaging and can, if needed, be used to refine the setup for future iterations. Click here to view larger image.
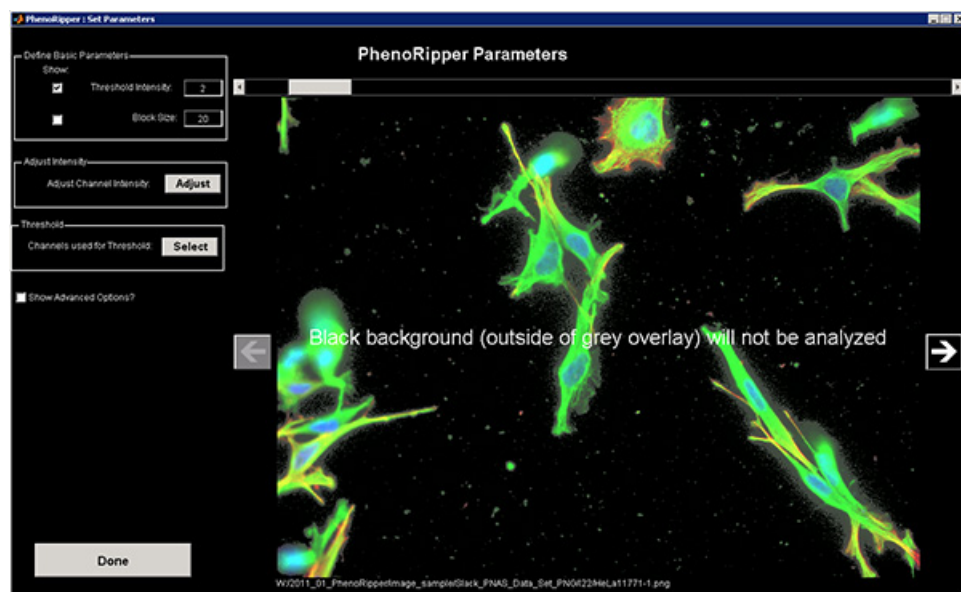

Figure 2. Setting PhenoRipper parameters: Block size and threshold intensity are set using a graphical interface with real time visual feedback. Click here to view larger image.
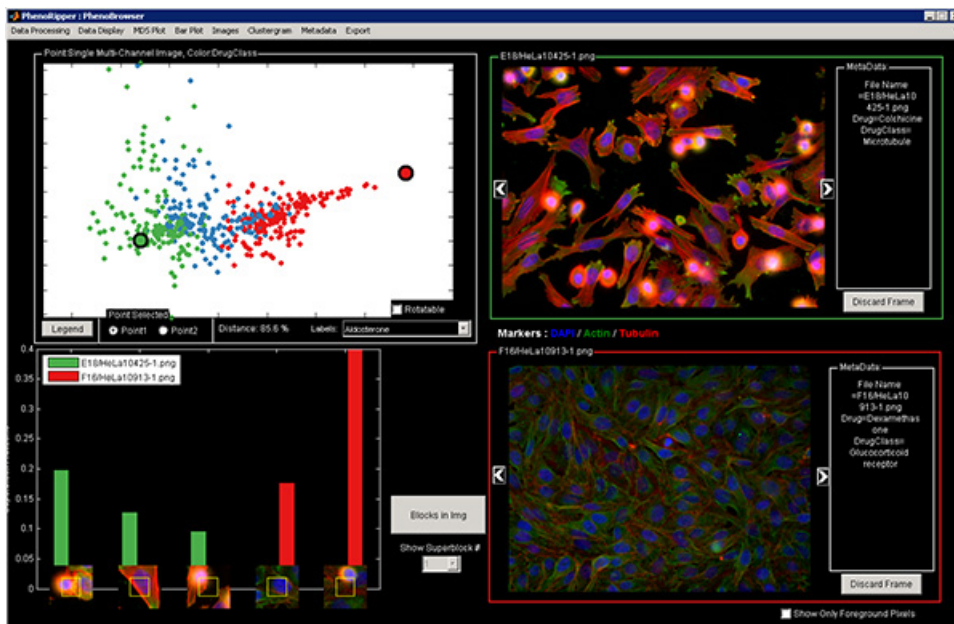
Figure 3. PhenoBrowser Interface for exploring PhenoRipper results: PhenoBrowser interface is divided into four panels: the top left panel is a scatter plot displaying the relative similarities between different images/conditions, the two panels on the right display sample images of selected conditions, and the lower left panel compares their phenotypic profiles. Click here to view larger image.
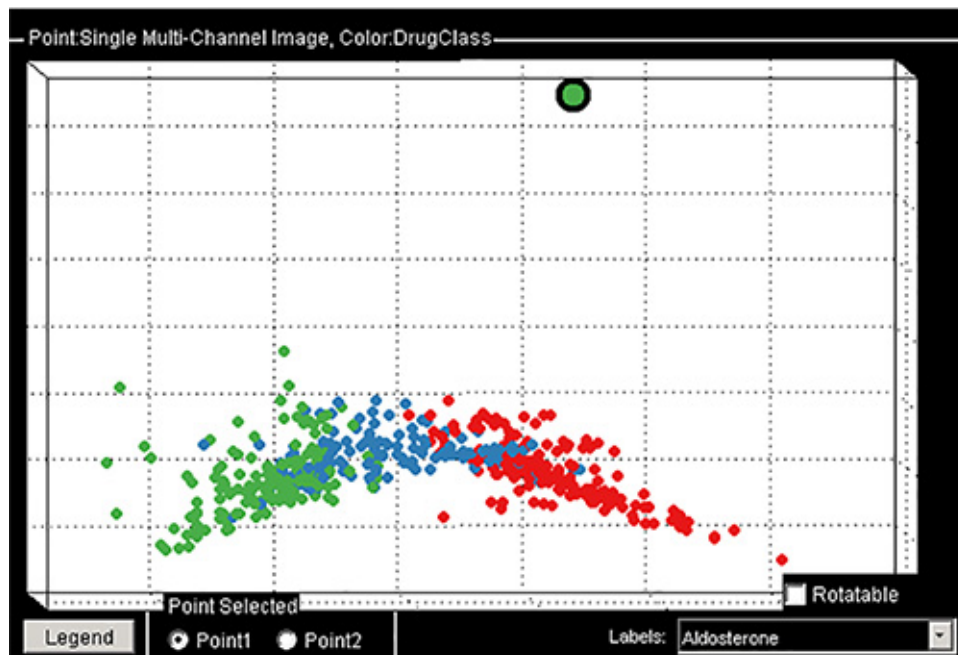
Figure 4. Example of an outlier in the scatter plot (highlighted) showing a staining artifact (see video). Click here to view larger image.
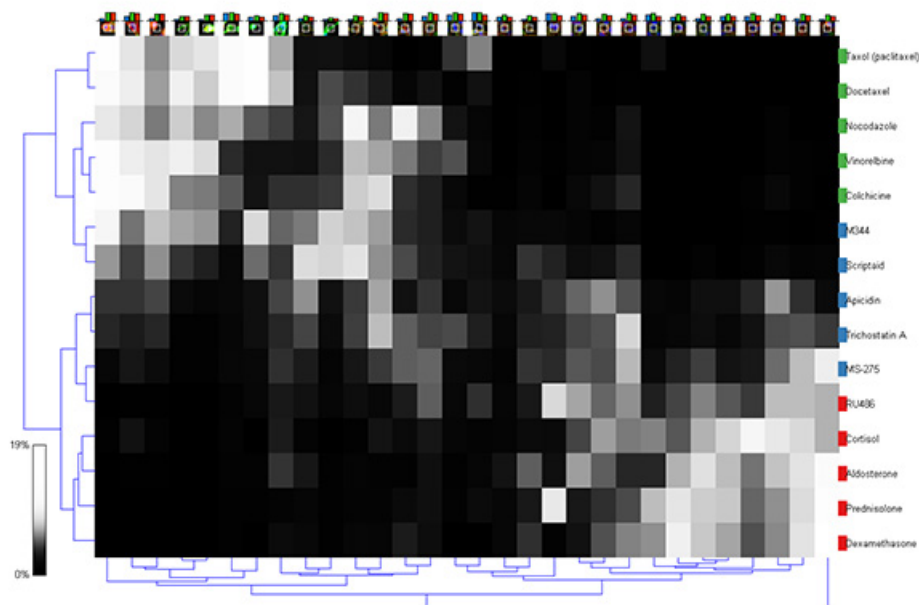
Figure 5. Clustergram representation of PhenoRipper profiles: Each row represents a single drug, and each column a block type. The gray-scale values reflect the phenotypic profiles, *i.e.* the fraction of the different block types for each drug. A lighter value indicates higher fraction of that block type. Hierarchical clustering largely separates the different drug by drug class (red/blue/green color on the right). Click here to view larger image.

Discussion

The workflow described here allows fast and easy characterization and comparison of microscopy images. We first demonstrated how this workflow can help experiment optimization, for example by quickly performing quality control on microscopy images. Next, we demonstrated its potential for analyzing high-throughput screening data: we were able to group drugs based on their mechanism of action. The grouping of drugs was comparable to that found using more complex methods[6], even though PhenoRipper was orders of magnitude faster.

Although we have found this approach robust against sporadic experimental fluctuations, as with any image based approach, readouts will be affected by systematic nonbiological effects. For example, plate-to-plate variation in marker intensity may mask the biological effects of interest. The impact of such artifacts could be reduced by appropriate corrections[7,8] such as plate to plate normalization and background subtraction. The profiles generated by PhenoRipper could also be influenced by variations in cell density. While this behavior is usually desirable (reflecting survival or growth), its effect can be reduced by un-selecting the "Use Background Information" option in the "Advanced Parameters".

In choosing the appropriate method for analysis, users need to consider the tradeoff between speed, ease of use and level of analytical detail. Segmentation-free methods[9-11] such as PhenoRipper are typically faster and easier to use than single cell-based methods such as CellProfiler[2]. The disadvantage of segmentation-free methods is the inability to quantify specific single-cell properties. The user thus needs to decide on the appropriate approach based on his or her needs. When accurate comparison or grouping of experimental conditions is the main objective, we have found that the level of detail captured by PhenoRipper is more than adequate. In particular, researchers focused on high-throughput experiments will benefit greatly from the ability to quickly and easily explore the image data. In conclusion, the workflow presented here will allow bench scientists to easily explore their fluorescence microscopy images soon after acquisition and facilitate a quick turn-around between analyses and imaging. We believe this workflow will lower the barrier to performing quantitative image analysis.

Disclosures

The authors have nothing to disclose.

Acknowledgements

References

1. Collins, T. J. ImageJ for microscopy. *Biotechniques.* **43**, 25-30 (2007).
2. Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, R100 (2006).

3. Shamir, L., Delaney, J. D., Orlov, N., Eckley, D. M., & Goldberg, I. G. Pattern Recognition Software and Techniques for Biological Image Analysis. *PLoS Comput. Biol.* **6** (2010).
4. Rajaram, S., Pavie, B., Wu, L. F., & Altschuler, S. J. PhenoRipper: software for rapidly profiling microscopy images. *Nat. Methods.* **9**, 635-637 (2012).
5. Lundholt, B. K., Scudder, K. M., & Pagliaro, L. A simple technique for reducing edge effect in cell-based assays. *J. Biomol. Screen.* **8**, 566-570 (2003).
6. Slack, M. D., Martinez, E. D., Wu, L. F., & Altschuler, S. J. Characterizing heterogeneous cellular responses to perturbations. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 19306-19311 (2008).
7. Wolf, D. E., Samarasekera, C., & Swedlow, J. R. Quantitative analysis of digital microscope images. *Methods Cell. Biol.* **81**, 365-396 (2007).
8. Zwier, J. M., Van Rooij, G. J., Hofstraat, J. W., & Brakenhoff, G. J. Image calibration in fluorescence microscopy. *J Microsc.* **216**, 15-24, (2004).
9. Shamir, L. *et al.* Wndchrm - an open source utility for biological image analysis. *Source Code Biol. Med.* **3**, 13 (2008).
10. Hamilton, N. A., Wang, J. T., Kerr, M. C., & Teasdale, R. D. Statistical and visual differentiation of subcellular imaging. *BMC Bioinform.* **10**, 94 (2009).
11. Huang, K., & Murphy, R. F. Automated classification of subcellular patterns in multicell images without segmentation into single cells. *2004 2nd IEEE Int. Symp. Biomed. Imag. Macro Nano.* **2**, 1139-1142 (2004).