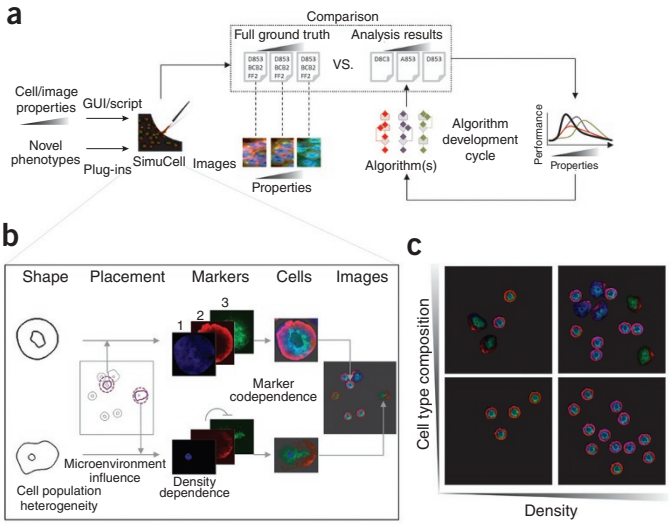

Figure 1 | Synthetic image generation using SimuCell. (**a**) Typical workflow of SimuCell use during algorithm development. (**b**) Steps involved in generation of synthetic images. In this example, microenvironment (local cell density) affects marker 1, and marker 2 influences marker 3. (**c**) SimuCell can be used to create images in which cell-population properties are varied independently.

range of phenotypes, encompassing nontrivial population-level effects such as cell-type heterogeneity or local cell-density effects (**Fig. 1c**). Although realistic synthetic data cannot replace true experimental data[6], SimuCell can be a useful part of the algorithm developer's toolbox by generating rich, flexible test image data sets containing specified, parameterized 'biological' effects.

Note: Supplementary information is available at http://www.nature.com/doifinder/10.1038/nmeth.2096.

ACKNOWLEDGMENTS

We thank R. Murphy and members of the Altschuler and Wu laboratories for helpful feedback and discussions. This research was supported by US National Institutes of Health R01 grants (GM085442 to S.J.A. and GM081549 to L.F.W.), the Welch Foundation (I-1619 to S.J.A. and I-1644 to L.F.W.), CPRIT RP10900 (L.F.W.) and the University of Texas Southwestern Medical Center (QP-SURF to N.E.F.H.).

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Satwik Rajaram[1,2,4], Benjamin Pavie[1,2,4], Nicholas E F Hac[3], Steven J Altschuler[1,2] & Lani F Wu[1,2]

[1]Green Center for Systems Biology, University of Texas Southwestern Medical Center, Dallas, Texas, USA. [2]Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas, USA. [3]University of Virginia School of Medicine, Charlottesville, Virginia, USA. [4]These authors contributed equally to the work.
e-mail: steven.altschuler@utsouthwestern.edu or lani.wu@utsouthwestern.edu

1. Snijder, B. *et al. Nature* **461**, 520–523 (2009).
2. Bray, M.A., Fraser, A.N., Hasaka, T.P. & Carpenter, A.E. *J. Biomol. Screen.* **17**, 266–274 (2012).
3. Zhao, T. & Murphy, R.F. *Cytometry A* **71A**, 978–990 (2007).
4. Svoboda, D., Kozubek, M. & Stejskal, S. *Cytometry A* **75A**, 494–509 (2009).
5. Lehmussola, A., Ruusuvuori, P., Selinummi, J., Huttunen, H. & Yli-Harja, O. *IEEE Trans. Med. Imaging* **26**, 1010–1016 (2007).
6. Anonymous. *Nat. Methods* **8**, 885 (2011).

PhenoRipper: software for rapidly profiling microscopy images

To the Editor: Recent advances in fluorescence microscopy have enabled unprecedented progress in many areas of biology. With the technology to perform high-content image-based screens now accessible to many labs, the analysis of the resulting large and complex data sets has become a bottleneck. Existing image analysis platforms[1–3] offer flexible and sophisticated toolboxes for extracting biological information from image data. However, they can require steep learning curves, tuning of many parameters and long computational runtimes. There is an unmet need for easy-to-use tools that enable bench scientists to rapidly interpret their image data sets. Here we describe PhenoRipper (**Supplementary Software**; updated versions available at http://www.phenoripper.org/), an open-source software tool designed for rapid exploration of high-content microscopy images (**Fig. 1a** and **Supplementary Fig. 1**). PhenoRipper permits rapid and intuitive comparison of images obtained under different experimental conditions based on image phenotype similarity.

To minimize user input, PhenoRipper automatically identifies features from the images; users may only be required to modify default values of a few, visually interpretable, parameters. To increase speed, we chose a segmentation-free approach[4,5]: the software breaks images down into a square grid of blocks[6–8] and performs analysis on these blocks rather than on individual cells. To capture heterogeneity, PhenoRipper identifies characteristic patterns of neighboring blocks and describes each image in terms of the occurrence frequencies of these patterns[6,8]. Finally, a simple graphical user interface, PhenoBrowser, is used to tie together images, features and profiles. Profiles can be annotated or combined (for example, by experimental or replicate conditions) to help interpret and explore their visual grouping. These design choices let users analyze their images an order of magnitude faster than existing unsupervised platforms (**Supplementary Fig. 2**). PhenoRipper does not replace traditional single cell–based analysis approaches[2,9,10] as it does not quantify properties such as area or average nuclear biomarker intensity. Nevertheless, the statistical properties of subcellular-scale phenotypes captured by PhenoRipper can be sufficient to accurately group cellular perturbations and identify outliers (**Supplementary Fig. 3a**).

PhenoRipper's engine performs four major steps (**Fig. 1a** and **Supplementary Fig. 1**). (i) PhenoRipper identifies foreground blocks. Images are gridded to a user-specified block size (20–30 blocks per cell works well), and blocks are selected when the intensities of >50% of their pixels exceed a foreground threshold. This threshold is precalculated based on a small subset of images (**Supplementary Methods**), but it can easily be changed by the user. (ii) PhenoRipper identifies the most common foreground block types. To do this, it characterizes blocks by their distributions of assigned pixel colors and applies cluster analysis to classify them into different block types. This measurement is not sensitive to cell orientation and captures more information than simple averages (for example, a block with 50% red and 50% blue pixels would be different from a block with 100% purple pixels). (iii) PhenoRipper uses cluster analysis to identify superblock types, which represent the most common block type co-occurrence patterns within 3×3 block neighborhoods. The use of blocks and superblocks helps

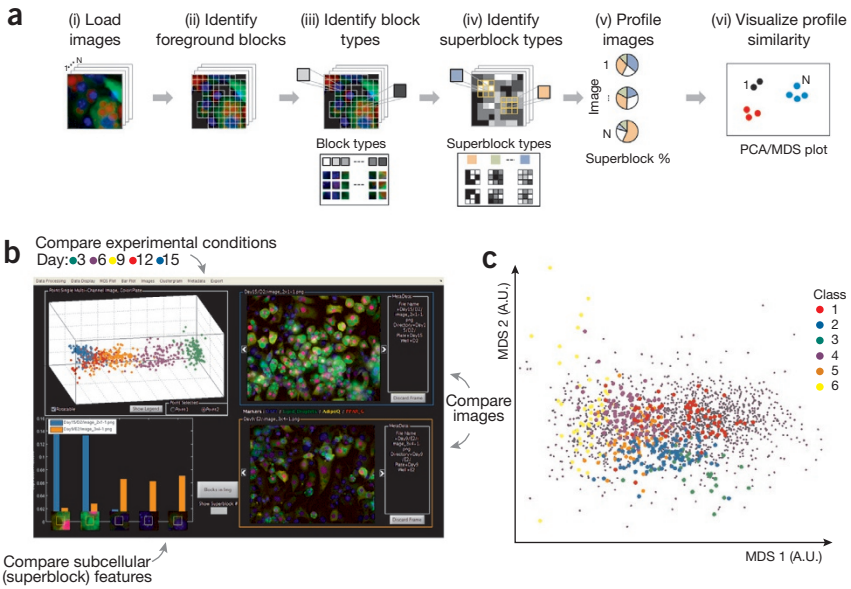

Figure 1 | Overview of the PhenoRipper platform. (**a**) Flow chart of analysis performed by PhenoRipper. (**b**) PhenoBrowser interface. Upper left: three-dimensional (3D) Multidimensional scaling (MDS) plot of profiles for images of 3T3-L1 cells on different days of differentiation to adipocytes. Right: two selected images from days 15 (top) and 9 (bottom) (blue, DNA; green, lipid droplets; yellow, AdipoQ; red, PPARγ). Lower left: superblock features that best distinguish the two selected images. (**c**) 2D MDS plot of PhenoRipper profiles for images of the 1,820 'hits' from the genome-wide siRNA screen described in ref. 10. Colors represent the six tightest phenotypic groups defined in ref. 10. Classes: 1, metaphase; 2, high-actin ratio; 3, lamellipodia + high-actin ratio; 4, proliferating cells; 5, small cells; 6, big cells. Gray dots, other/unclassified; A.U., arbitrary units.

to capture information over different distance scales. To speed up the steps described above, this initial analysis randomly samples a subset of images (**Supplementary Methods**). (iv) PhenoRipper profiles each image by the frequency of occurrence of superblock types. Profiles of experimental conditions are computed by averaging the superblock fractions of their corresponding images. We have found that similarities between profiles are relatively insensitive to parameter variation (**Supplementary Figs. 3b** and **4**). These profiles provide compact, human- and machine-interpretable summaries of image phenotypes. Profile similarities can be used to infer relationships among experimental conditions and underlying mechanisms of perturbations.

We tested PhenoRipper on a data set (640 four-channel images) in which cells were difficult to segment and phenotypically heterogeneous[9] (**Fig. 1b** and **Supplementary Fig. 5**). This data set consists of images of 3T3-L1 preadipocytes that were monitored for multiple readouts of adipogenesis at different days after induction of differentiation. Our original study, in which image analysis was carried out by traditional single-cell analysis, required a tedious manual step of discarding poorly segmented cells. In contrast, PhenoRipper completed its analysis in ~6.5 minutes, selecting image features that could distinguish among images from different days of differentiation and identifying superblock types that corresponded roughly to subcellular features of previously identified subpopulations, at different stages of the differentiation process[8]. Thus PhenoRipper can reveal meaningful features of heterogeneous populations and images for which robust cell segmentation is not easily achieved.

Next we reanalyzed a data set (~10^5 three-channel images) whose scale and complexity is representative of high-throughput screens, which typically require dedicated image analysis platforms and analysis expertise[10] (**Fig. 1c**). This data set is from an experiment in which the effects of ~23,000 genome-wide RNAi-mediated knockdowns on HeLa cells were monitored using cytoskeletal markers. The previous analysis was reported to take over 300 CPU hours, which excludes the time required to optimize the analysis pathway. In comparison, PhenoRipper completed analysis of this data set in ~13 hours on a test desktop, without the need to tune any parameters other than threshold intensity and block size. To compare the profiling results, we focused on the 'hits' reported in the previous study (our analysis of these ~7,000 images took ~30 minutes). Visual grouping of PhenoRipper profiles, annotated by phenotypic classes defined in the previous study, suggested that similarities between knockdown profiles had been largely preserved between the two methods (**Fig. 1c**). Overall, similar profile pairs from PhenoRipper showed strong enrichment for similar biological function (**Supplementary Methods** and **Supplementary Fig. 6**). Thus PhenoRipper provides an approach for rapidly extracting biologically meaningful information from large, complex data sets.

PhenoRipper is designed to serve as an unsupervised exploratory tool for analysis of fluorescence microscopy images for both novices and experts. It may not always be the optimal tool—some applications may require quantification of specific features on single cells or may be more suitable for supervised classification. Nevertheless, the speed and simplicity of PhenoRipper make it a useful tool that enables bench scientists to perform rapid analysis of image data soon after acquisition.

Note: Supplementary information is available at http://www.nature.com/doifinder/10.1038/nmeth.2097.

ACKNOWLEDGMENTS

We thank R. Murphy, M. Slack, R.J. Steininger III, C.A. Thorne, P. Xie and members of the Altschuler and Wu laboratories for helpful feedback and discussions. This research was supported by National Institutes of Health grants R01 GM085442 (S.J.A.) and R01 GM081549 (L.F.W.), Welch Foundation grants I-1619 (S.J.A.) and I-1644 (L.F.W.) and CPRIT RP10900 (L.F.W).

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Satwik Rajaram[1,2], Benjamin Pavie[1,2], Lani F Wu[1,2] & Steven J Altschuler[1,2]

[1]Green Center for Systems Biology, University of Texas Southwestern Medical Center, Dallas, Texas, USA. [2]Department of Pharmacology, University of Texas Southwestern Medical Center, Dallas, Texas, USA.
e-mail: lani.wu@utsouthwestern.edu or steven.altschuler@utsouthwestern.edu

1. Collins, T.J. *Biotechniques* **43** Suppl, 25–30 (2007).
2. Carpenter, A.E. *et al. Genome Biol.* **7**, R100 (2006).
3. Shamir, L., Delaney, J.D., Orlov, N., Eckley, D.M. & Goldberg, I.G. *PLoS Comput. Biol.* **6**, e1000974 (2010).
4. Huang, K. & Murphy, R.F. in *Proc. of the 2004 2nd IEEE International Symp. on Biomedical Imaging*, 1139–1142 (2004).
5. Shamir, L. *et al. Source Code Biol. Med.* **3**, 13 (2008).
6. Bhattacharya, A. *et al. Proc. of the Fifth IEEE International Conf. on Data Mining*, 50–57 (2005).
7. Nanni, L. & Lumini, A. *Artif. Intell. Med.* **43**, 87–97 (2008).
8. Cruz-Roa, A., Caicedo, J.C. & González, F.A. *Artif. Intell. Med.* **52**, 91–106 (2011).
9. Loo, L.H. *et al. J. Cell Biol.* **187**, 375–384 (2009).
10. Fuchs, F. *et al. Mol. Syst. Biol.* **6**, 370 (2010).

Annotated high-throughput microscopy image sets for validation

To the Editor: Choosing among algorithms for analyzing biological images can be a daunting task, especially for nonexperts. Software toolboxes such as CellProfiler[1,2] and ImageJ[3] make it easy to try out algorithms on a researcher's own data, but it can still be difficult to assess whether an algorithm will be robust across an entire experiment based on the small subset of images that is practical to examine or annotate. Even if controls are available, a pilot high-throughput experiment may be insufficient to show that an algorithm will robustly identify rare phenotypes and handle the experimental artifacts that will invariably be present in a high-throughput experiment. It is therefore useful to know that a particular algorithm has proven superior on several similar image sets. The performance comparisons presented in papers that introduce new algorithms are often not very helpful for assessing this because each study typically relies on a different test image set (often to the advantage of the proposed algorithm), the algorithms compared may not be the ones the researcher is most interested in and the authors may not have implemented other algorithms as optimally as their own. Although biologists should always also validate algorithms on their own images, it would be useful if developers would quantitatively test new algorithms against a publicly available established collection of image sets. In this way, objective comparison can be made to other algorithms, as tested by the developers of those algorithms. We see a need for such a collection of image sets, together with ground truth and well-defined performance metrics.

Here we present the Broad Bioimage Benchmark Collection (BBBC), a publicly available collection of microscopy images intended as a resource for testing and validating automated image-analysis algorithms. The BBBC is particularly useful for high-throughput experiments and for providing biological ground truth for evaluating image-analysis algorithms. If an algorithm is sufficiently robust across samples to handle high-throughput experiments, low-thoughput applications also benefit because tolerance to variability in sample preparation and imaging makes the algorithm more likely to generalize to new image sets.

Each image set in the BBBC is accompanied by a brief description of its motivating biological application and a set of ground-truth data against which algorithms can be evaluated. The ground truth sets can consist of cell or nucleus counts, foreground and background pixels, outlines of individual objects, or biological labels based on treatment conditions or orthogonal assays (such as a dose-response curve or positive- and negative-control images). We describe canonical ways to measure an algorithm's performance so that algorithms can be compared against each other fairly, and we provide an optional framework to do so conveniently within CellProfiler. For each image set, we list any published results of which we are aware.

The BBBC is freely available from http://www.broadinstitute.org/bbbc/. The collection currently contains 18 image sets, including images of cells (*Homo sapiens* and *Drosophila melanogaster*) as well as of whole organisms (*Caenorhabditis elegans*) assayed in high throughput. We are continuing to extend the collection during the course of our research, and we encourage the submission of additional image sets, ground truth and published results of algorithms.

AUTHOR CONTRIBUTIONS

K.L.S. and V.L. curated image sets and oversaw collection of ground-truth annotations. K.L.S. developed benchmarking pipelines. V.L. defined benchmarking protocols. A.E.C. conceived the idea and guided the work. All authors wrote the manuscript.

ACKNOWLEDGMENTS

We thank the contributors of image sets as well as V. Uhlmann and the Carpenter laboratory members who have helped annotate them with ground truth.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Vebjorn Ljosa, Katherine L Sokolnicki & Anne E Carpenter

Broad Institute of Massachusetts Institute of Technology and Harvard, Cambridge, Massachusetts, USA.
e-mail: ljosa@broad.mit.edu or anne@broadinstitute.org

1. Carpenter, A.E. *et al. Genome Biol.* **7**, R100 (2006).
2. Kamentsky, L. *et al. Bioinformatics* **27**, 1179–1180 (2011).
3. Schneider, C.A., Rasband, W.S. & Eliceiri, K.W. *Nat. Methods* **9**, 671–675 (2012).