Scoring diverse cellular morphologies in image-based screens with iterative feedback and machine learning

Thouis R. Jones[a,b,c,1], Anne E. Carpenter[a,b,1,2], Michael R. Lamprecht[b], Jason Moffat[b,3], Serena J. Silver[a], Jennifer K. Grenier[a], Adam B. Castoreno[d], Ulrike S. Eggert[d], David E. Root[a], Polina Golland[c], and David M. Sabatini[a,b,e]

[a]The Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142; [b]Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142; [c]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA 02139; [d]Dana-Farber Cancer Institute and Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, 250 Longwood Avenue, Boston, MA 02115; and [e]Department of Biology, Massachusetts Institute of Technology, 31 Ames Street, Cambridge, MA 02139

Edited by Edward M. Scolnick, The Broad Institute, Cambridge, MA, and approved December 12, 2008 (received for review September 8, 2008)

Many biological pathways were first uncovered by identifying mutants with visible phenotypes and by scoring every sample in a screen via tedious and subjective visual inspection. Now, automated image analysis can effectively score many phenotypes. In practical application, customizing an image-analysis algorithm or finding a sufficient number of example cells to train a machine learning algorithm can be infeasible, particularly when positive control samples are not available and the phenotype of interest is rare. Here we present a supervised machine learning approach that uses iterative feedback to readily score multiple subtle and complex morphological phenotypes in high-throughput, image-based screens. First, automated cytological profiling extracts hundreds of numerical descriptors for every cell in every image. Next, the researcher generates a rule (i.e., classifier) to recognize cells with a phenotype of interest during a short, interactive training session using iterative feedback. Finally, all of the cells in the experiment are automatically classified and each sample is scored based on the presence of cells displaying the phenotype. By using this approach, we successfully scored images in RNA interference screens in 2 organisms for the prevalence of 15 diverse cellular morphologies, some of which were previously intractable.

high-content screening | high-throughput image analysis | phenotype

The history of biology has been dramatically shaped by classic visual screens in model organisms, including *Drosophila melanogaster* (1–3), *Saccharomyces cerevisiae* (4), *Caenorhabditis elegans* (5), and the zebrafish *Danio rerio* (6, 7). In each case, biological pathways were discovered because researchers were intrigued by groups of peculiar-looking mutants and identified the genes underlying their phenotypes. Because researchers have favored the extensive study of relatively few genes (8), classic, wide-net approaches like screening are as relevant as ever to probe known biological pathways and discover new ones. Modern technology now enables large-scale experiments in cultured cells to identify human genes that underlie biological processes via RNAi. Automation also allows the screening of chemical libraries to identify perturbants useful as research tools or drugs.

Despite these advances, scoring cells in images for rare and unusual morphologies has, in general, remained a significant bottleneck (9–12). Cell image analysis allows accurate identification and measurement of cells' features, enabling automated analysis of certain phenotypes that were previously intractable (13–26). However, many interesting phenotypes require the assessment of several measured features of cells. Machine learning methods that select and combine multiple features for automated cell classification have been used to score many phenotypes (15–26). These methods require the provision of example cells that do and do not display the morphology of interest (i.e., positive and negative cells). Finding positive cells is straightforward when positive control samples are available and most of the cells therein show the phenotype. However, when this is not the case, as in classic exploratory screens, finding a sufficient number of positive cells can be prohibitively difficult. Even when positive control samples are available, using positive example cells from only those samples can lead to inaccurate scoring because of overfitting of the machine learning algorithm.

Here we describe our approach to scoring multiple complex and subtle phenotypes in large-scale, image-based screens. It is particularly effective when positive control samples are not available or not highly penetrant, as is often the case in RNAi and chemical screens. Our approach uses: (*a*) a biologist's ability to identify an "interesting" phenotype, (*b*) automatic measurement of multiple features for each cell, (*c*) a computer's ability to rapidly test multiple combinations of features using machine learning algorithms, and (*d*) a computer's ability to quickly and objectively classify millions of individual cells based on their measured features. We used our approach to score 15 diverse cellular phenotypes in large-scale RNAi screens in human and *D. melanogaster* cells, demonstrating that automated scoring for image-based chemical and genetic screens for multiple complex, low-penetrance phenotypes is now feasible.

Results

Overview of the Approach. We have developed and validated a method for researchers to rapidly train a computer to score unusual cell morphologies automatically (Fig. 1). First, we automatically identify and measure every cell in every image in the experiment by using the cell-image analysis software CellProfiler (13), which generates a cytological profile (27), or cytoprofile, for each cell. This cytoprofile consists of a set of numbers that describe the cell's characteristics, including size, shape, and the intensity and texture of various stains in various compartments (Fig. 1*A*). Next, the researcher initiates the training phase by identifying a few positive example cells that display a phenotype of interest and negative example cells without the phenotype (Fig. 1*B*). These cells can be from control samples if the screen has been designed to address a particular phenotype, or selected at random if the screen's goal is to uncover previously uncharacterized phenotypes in an exploratory screen. Most commonly, these example cells are taken from the full population without reference to the particular sample from

Author contributions: T.R.J., A.E.C., D.E.R., P.G., and D.M.S. designed research; T.R.J., A.E.C., M.R.L., J.M., S.J.S., J.K.G., A.B.C., and U.S.E. performed research; T.R.J., A.E.C., and P.G. analyzed data; and T.R.J. and A.E.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]T.R.J. and A.E.C. contributed equally to this work.

[2]To whom correspondence should be addressed. E-mail: anne@broad.mit.edu.

[3]Present address: Banting and Best Department of Medical Research, Donnelly Centre for Cellular and Biomolecular Research, University of Toronto, 160 College Street, Room 802, Toronto, Ontario, Canada M5S 3E1.

This article contains supporting information online at www.pnas.org/cgi/content/full/0808843106/DCSupplemental.

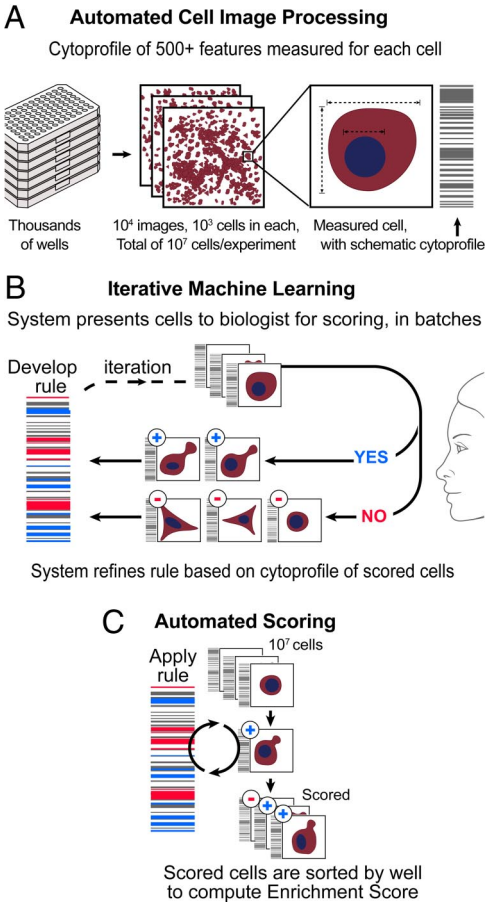

Fig. 1. Scoring cell morphologies via cytological profiling, iterative feedback, and machine learning. (*A*) Images of cell populations for each treatment condition (RNAi or chemical) are processed with cell-image analysis software (e.g., CellProfiler) to identify and measure individual cells, in order to generate a cytological profile, containing a collection of measurements of features of each cell, represented schematically here as a bar code. (*B*) The software system presents the researcher with individual cells for classification, sampled randomly from the screen-wide population. After a few dozen cells are classified, the researcher can begin the iterative machine learning phase, in which the computer generates a tentative rule based on the classified cells and presents the researcher with cells classified according to that rule. In general, larger training sets produce more accurate rules, and using too small a training set can result in the computer training to a too-narrow definition of the phenotype (Fig. S10). Generating a large training set without iterative feedback can be difficult when the phenotype is rare or no positive control samples are available; these are the cases where the iterative nature of our approach is most useful. The optimal initial training set size depends on the complexity of the phenotype and the scarcity of positive cells in the experiment. After the researcher corrects errors and retrains for several rounds, the rule becomes more accurate. (*C*) When the accuracy of the rule is sufficient, it is used to classify all cells in the experiment in order to calculate the number of positive and negative cells in each sample.

which they are derived. This action is taken to avoid overfitting the machine learning algorithm to a few particular samples.

Once a few dozen individual cells have been classified by the researcher, a machine learning algorithm is used to determine a tentative rule (i.e., a classifier) that distinguishes the cytoprofiles of the positive and negative example cells, using the GentleBoosting algorithm applied to regression stumps (28). Other machine learning methods are likely to be equally effective, based on their performance in previous work (15–24). The system then presents the researcher with a new batch of cells, which it has classified based on the tentative rule, and the researcher corrects errors. The corrections are used to refine the rule. After several rounds of error

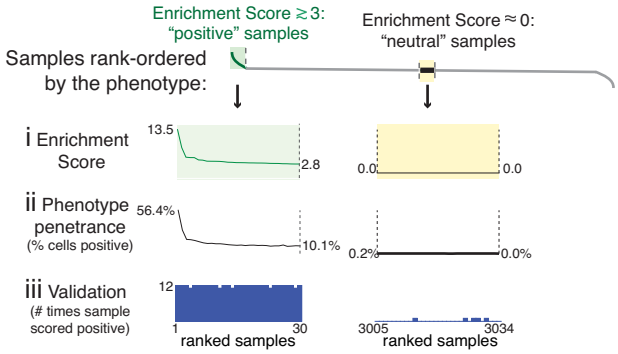

Fig. 2. Validation example of actin blebs phenotype. (*i*) The approach rank-orders samples (populations of cells under the same treatment condition) by their enrichment score (see *Methods*) and allows selection of positive and neutral samples based on this automated scoring. (*ii*) The corresponding phenotype penetrance is shown for the positive and neutral samples. Phenotype penetrance is typically correlated with enrichment score except that a low number of cells in a sample can decrease the score despite a high penetrance. (*iii*) The corresponding validation data are shown for the positive and neutral samples. The height of the bar for each sample indicates how many times a human observer chose that sample as a positive in a forced-choice comparison (see *Methods*). In this example, samples that were scored as positives (*Left*) were also chosen by the researchers as positives (11 or 12 times, of 12 comparisons per sample), and none of the neutral samples (*Right*) were routinely chosen as positive (0 or 1 time of 12 comparisons). Corresponding data for all phenotypes is shown in Figs. 3 and 4.

correction and rule refinement, the researcher has classified a few hundred cells, and these are used to produce a rule specific to the phenotype of interest. In the final step (Fig. 1*C*), the rule is applied to the cytoprofiles of every cell in the experiment, classifying each cell as positive or negative. Ultimately, the goal of the screen is to score each sample, which is a population of cells subjected to a particular RNAi or chemical treatment. Because simply ranking samples by the percentage of cells that are positive can be misleading for samples with few cells, we developed an "enrichment score" to rank each sample (see Fig. 2 and *Methods*). The researcher may continue to conduct further rounds of error correction and rule refinement based on images from samples with many positive cells, ultimately producing a rule with satisfactory accuracy. Although highly dependent on the complexity of the phenotype and the scarcity of positive example cells, the entire process of training for a phenotype typically takes a few hours.

Scoring RNAi Screens for Diverse Phenotypes in Human Cells. We used this iterative approach to recognize and score 14 diverse phenotypes (Figs. 3 and 4) based on measurements acquired from ≈8.3 million human cells contained within 40,000 previously acquired fluorescence images (14). The cytological profile for each cell contained 610 measurements (see *SI Text*), resulting in more than 5 billion measurements total. Some of the phenotypes we chose are well-known—cells in particular subphases of mitosis, for example. Others, such as crescent-shaped nuclei (Fig. 3*E*) and blebs of actin that sometimes formed tubular projections (Fig. 3*A*), have no clear biological interpretation.

Nearly every phenotype we attempted to score could be scored accurately without customization of the image processing. That is, the standard cytoprofiles were sufficient for accurate classification in all but the Peas in a Pod phenotype. We added one feature (angle between a nucleus' 2 nearest neighbors) to the image-analysis step to better identify this phenotype (Fig. 4*C*). Also, we abandoned attempts to train a rule to identify a "sparkly actin" phenotype (Fig. S1); few positive example cells could be found, and it is possible that our cytoprofiles did not contain appropriate texture measurements.

Features from the cytoprofiles that were used to classify cells for each phenotype usually included a mixture of measurements of

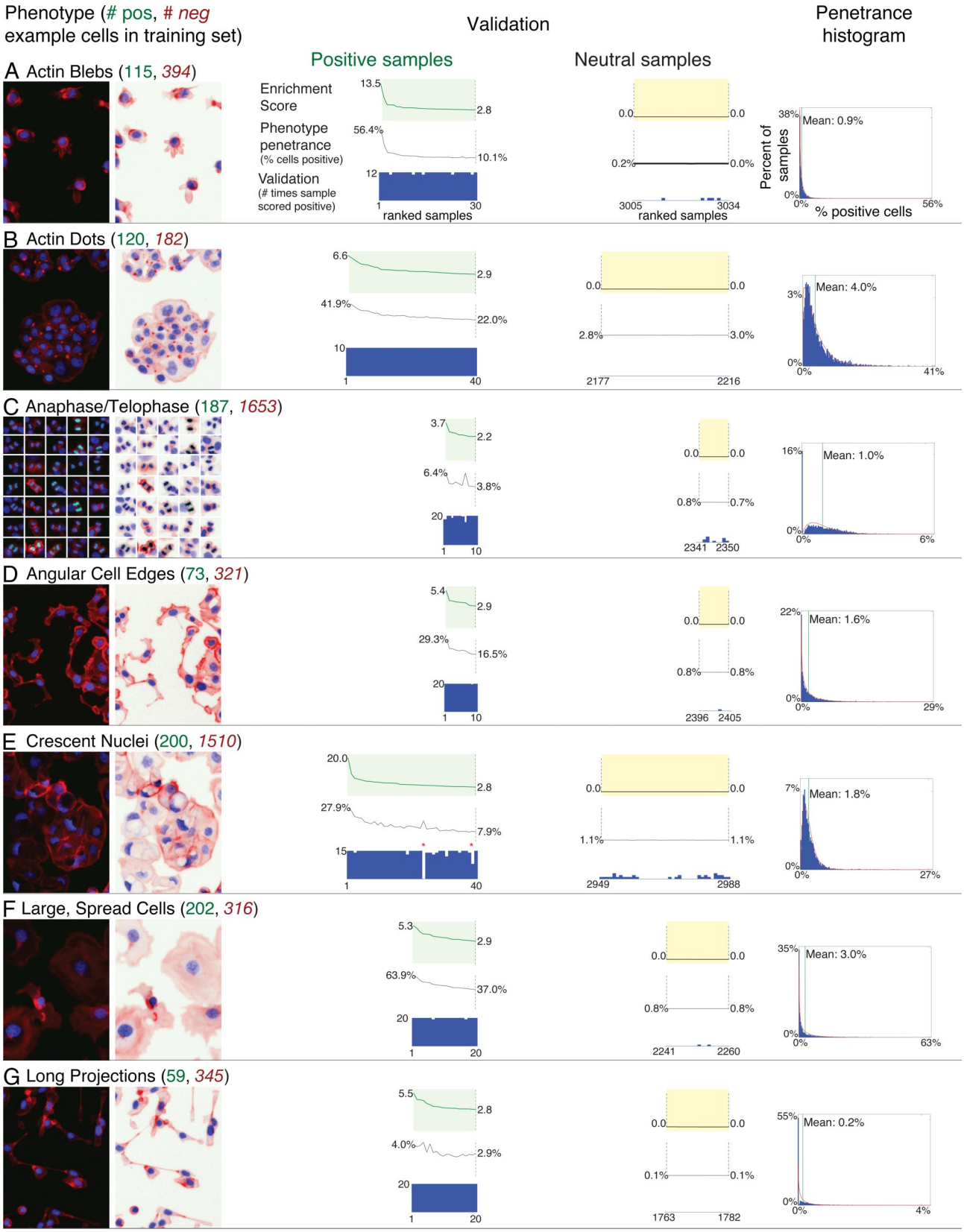

Fig. 3. Results of the phenotype-scoring system, for diverse cellular morphologies in human cells. Each row shows images and data for a different cellular morphology that the system was trained to recognize and score. The phenotype column shows the name of each phenotype along with the number of positive and negative example cells in the training set after all rounds of iteration were completed by the researcher. Images for each phenotype follow a color scheme: blue, DNA (contrast-stretched); red, actin (contrast-stretched); green, phospho-histone H3 (absolute scale). (*Left*) Traditional pseudocoloring of the fluorescence microscopy images. (*Right*) Color-adjustment using the ''Invert For Printing'' module of CellProfiler. The width of each image (or montage, for multiframe images) is 102 μm. For details on the validation column, see Fig. 2. The penetrance histogram column shows the distribution of per-sample penetrance for each phenotype, along with the mean (shown as text and with a green line) and the model fit to the data (red line).

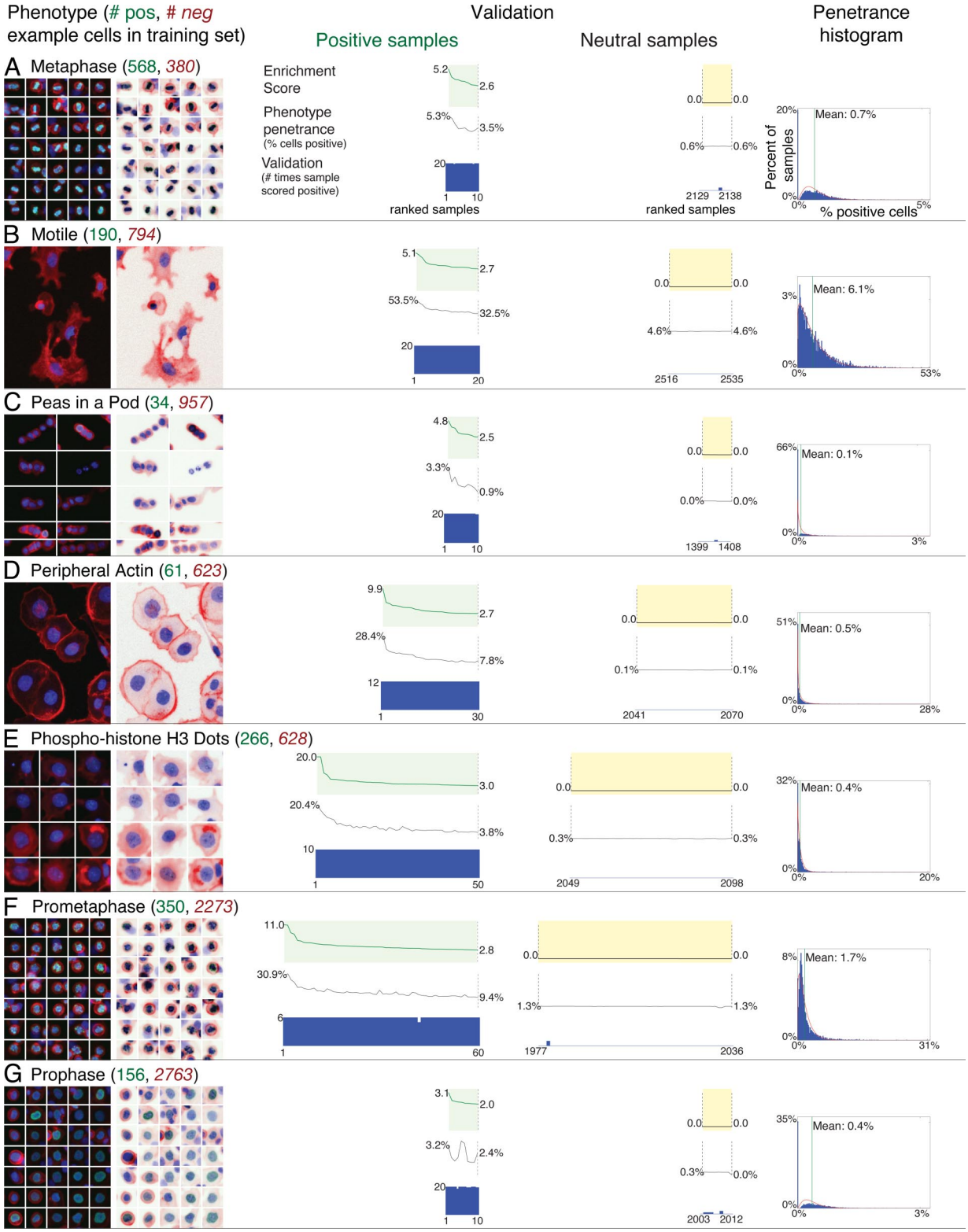

Fig. 4. More results of the phenotype-scoring system, for diverse cellular morphologies in human cells. See Fig. 3 for details.

intensity, texture, and area/shape (Fig. S2 and *SI Text*). Some features were unexpected, implying that choosing features manually by using biological or image-analysis expertise would have overlooked useful features. The features also served to generate hypotheses about phenotypes that were otherwise uncharacterized. For example, cells showing the actin blebs and peripheral actin phenotypes tend to have 4N DNA content, indicating an unexpected relationship to the cell cycle (Fig. S3).

For most phenotypes, we knew of no samples that could be considered positive controls, precluding our use of existing methods that require highly penetrant controls (15, 19, 20). Typically, our only exemplars were unusual phenotypes that we observed at a low

frequency in WT cells. Factors like cell cycle, local environment, stochastic noise, and epigenetics all play a role in generating nonuniform populations of cells (29, 30). We therefore wondered whether any samples would have an unusually high proportion of cells showing these naturally occurring rare morphologies. Interestingly, every phenotype we pursued yielded at least some RNAi samples in which the phenotype was significantly enriched. This is consistent with the possibility that the number of phenotypic states that are possible for a cell is fairly limited, and natural variation in mRNA expression levels can push cells into one of these states, even without the influence of RNAi. In any event, the system enabled us to indulge our curiosity by pursuing unusual and uncharacterized cellular morphologies, as in classic genetic screens.

Validation, Comparison to Previous Methods, and Flexibility. We tested our method's accuracy at ranking samples by having researchers score samples (that is, images showing a population of cells) by eye. The biologically relevant score for a sample is enrichment of cells that display the phenotype, rather than a hard "positive" or "negative" label, because samples in screens typically do not fall into clear positive and negative classes (particularly when judged by different researchers), but instead fall along a continuum (31). Our goal is to bring highly enriched samples to the attention of the researcher; therefore, our validation design (forced choice, described in *Methods*) (32) aimed to test whether top-ranking samples were indeed enriched relative to samples scored as neutral.

The results for actin blebs are shown in detail in Fig. 2, and data for all human cell phenotypes are shown in the validation column in Figs. 3 and 4. For each phenotype, we rank-ordered the 5,000 puromycin-treated samples by enrichment score (Fig. 2*A*), as would be done in a typical screen. For validation, researchers were forced to choose between pairs of samples. One sample in each pair had been scored by the computer as highly enriched for the phenotype and the other as neutral. We recorded the number of times each sample was chosen as positive by the researchers (bar chart, Fig. 2*C*).

Among all 360 samples identified as "hits" across the different phenotypes (Figs. 3 and 4, positive samples column), there were 0 false negatives among the 360 samples identified as neutral and 2 potential false positives (red stars in Fig. 3*E*). Note that false positives can be readily weeded out by eye after analysis and that we cannot estimate the actual false-negative rate without knowing a priori the number of true positive samples, which is not possible in this screen. Agreement between humans was comparable with that between humans and automated scoring (Table S1), indicating sufficient accuracy to bring samples enriched for each phenotype to the attention of the researcher.

The phenotypes we chose were particularly challenging because their average penetrance was low (0.2–6.1%), and even the strongest hits for some phenotypes contained <5% positive cells. All phenotypes were, nonetheless, readily scored by our method. Previous approaches (15, 19, 20) have succeeded on highly penetrant phenotypes where positive control samples are known, but none of the phenotypes in our study had positive control samples available, and most were low-penetrance. We chose 4 of the phenotypes in this study and retrospectively tested a positive control-based method on them (Fig. S4). The method worked well on the most highly penetrant, straightforward phenotype, large spread cells (Fig. S4*A*), but was inferior on the other 3 phenotypes of greater morphological complexity and lower penetrance, in some cases even failing to highly rank the training samples (Fig. S4 *B–D*).

Overfitting is a concern when using machine learning algorithms, but boosting variants are fairly resistant to it (28). Cross-validation results (Fig. S5) show that this is also the case for our approach. The classification accuracy is typically not significantly reduced as the number of individual regression stumps forming a rule for a phenotype increases. To increase the coverage of the training set and guard against training to a too-narrow definition of a phenotype, it is useful to inspect images of the top-ranked samples (or positive control samples, if available), in which positively classified cells are marked. From these images, it is easy to identify false-negative cells and add them to the training set during the iterative training phase.

We considered whether a rule will generalize to new experiments. A rule trained on one experiment is unlikely to be applicable to experiments involving different assay protocols, cellular stains, or image acquisition instrumentation, although with our approach, the time required to generate a new rule for the new experiment is minimal. For replicate experiments, creating a training set from one replicate and applying the rule it generates to another replicate risks negatively impacting its accuracy because of undetected experimental variation (Fig. S6*B*). The more robust approach is to create a training set spanning all replicates (Fig. S6*A*).

Lastly, we tested our method's flexibility by applying it to another large-scale image set. Previously, 288 genes were screened for a metaphase phenotype by RNAi in *Drosophila* by using living-cell microarrays (33). In our previous work, we identified cells in metaphase by empirically applying sequential gates based on 4 measured features of the DNA stain of each cell. This process took more than a week. With our new approach, we identified metaphase nuclei and accurately scored the entire screen within 4 h, of which only 1 h was hands-on time (Fig. S7 and Fig. S8). The top of the rank-ordered list of genes from the screen (*SI Text*) contained widerborst (CG5643, the one hit in our original study), as well as other cell-cycle-related genes, e.g., polo (CG12306) and microtubule star (CG7109). The gene most deenriched for metaphase nuclei was Nima-related kinase 2 (Nek2, CG17256; "Nima" derives from "never in mitosis"). As was the case for complex human phenotypes (Fig. S4), providing the positive control sample images directly to the machine learning algorithm was unsuccessful (Fig. S9).

Discussion

Together, this work indicates that automated scoring of a wide variety of morphologies can be accomplished quickly and easily, even when a phenotype is rare in the WT population and positive control samples are not available. Specifically, the approach is scalable to large-scale, image-based screens (chemical or genetic) in which multiple complex phenotypes are examined. Whereas screening for perturbations of general cellular functions like cell division has yielded large networks of genes (14, 34), the ability to identify more subtle and rare cellular morphologies should yield more tightly focused families of genes worthy of study (35). In particular, morphologies of unknown biological significance are likely to lead to the study of entirely new pathways in the spirit of classic genetic screens.

The approach described here is compatible with automated image analysis systems and, importantly, is robust to the occasional segmentation errors produced by such systems. Previous work has demonstrated that machine learning algorithms can be successfully trained by using all cells from positive and negative control samples to create a training set, even for some phenotypes that cannot be visually distinguished by humans (25). Here we showed that, whereas this approach can be successful for highly penetrant phenotypes (Fig. S4*A*), it is not suitable when the phenotype is less penetrant (Fig. S4 *B–D* and Fig. S9). We have addressed these challenging situations, thus enabling screens for low-penetrance phenotypes that lack positive control samples. Even when positive control samples are available, leveraging the user's visual perception to select individual example cells helps prevent the machine learning algorithm from focusing on aspects of morphology that are irrelevant to the biological question at hand or from becoming tuned to cells that display some complex combination of phenotypes as the positive control samples (i.e., pleiotropic effects) rather than the specific phenotype of interest.

The machine learning approach presented here has been implemented and released as the "Classifier" feature in an open-source software package we developed previously for visualizing and exploring data from image-based screens, called CellProfiler Analyst (33).

Methods

Algorithms and Software. The software packages used in this work, CellProfiler and CellProfiler Analyst, are open-source (available from the Broad Institute at www.cellprofiler.org). The image-analysis pipeline, which can exactly recreate the analysis in CellProfiler, is provided along with a text description (*SI Text*). Based on code from Torralba *et al*. (36), the Classifier functionality was developed as a feature in CellProfiler Analyst for this study; its usage is described in a manual and an online demonstration video (available from the Broad Institute at www.cellprofiler.org/examples).

The time to compute a rule is on the order of a few seconds, and grows linearly with training set size and the number of features. Using the rule to classify 8 million cells in a database takes ≈2 min, with the same orders of growth, primarily limited by disk transfer speed, as the full dataset must be read to classify every cell. Image processing times to identify and measure cells using CellProfiler are currently on the order of 10 s to several minutes per image, depending on the particular experimental and image analysis used (for example, ≈2.5 min per 3-channel, 512- × 512-image on a 2.4-GHz Intel CPU with 8 gigabytes of RAM for the human cell images in this study). Cluster computing prevents this from becoming a bottleneck.

RNAi Screens, Images, and Cytological Profiles. Images used in the human screens presented here have been previously described (14). Cells were stained for DNA (Hoechst), actin (phalloidin), and phospho-histone-H3 serine 10 (antibody). Approximately 5 separate lentiviral-delivered shRNAs were tested for each of 1,028 genes, mostly kinases and phosphatases, with 2 samples for each shRNA (one with and one without the selection reagent for the shRNA, puromycin) and with 4 images captured per sample. We used the samples treated with puromycin (the selection agent for the shRNA vector) for the validation step shown in Figs. 3 and 4 because puromycin selection culls cells where the shRNA vector failed to infect, leading to more homogeneous populations in each sample and because puromycin affects phenotype penetrance in the WT population. Images (250 GB), the database of cytoprofiles (20 GB), and each phenotype's training set of positive and negative example cells are available on request. Images and data used in the *Drosophila* metaphase screen have also been previously described (33). Briefly, there were 5 replicates of a cell microarray, and each array had 3 replicate spots per gene, plus 256 negative control spots lacking an RNAi reagent.

ACKNOWLEDGMENTS. We thank InHan Kang (Massachusetts Institute of Technology) for creating the CellProfiler Analyst software infrastructure and engineering some of the machine learning functionality; the RNAi Consortium and the Broad Institute RNAi Platform for investment of time and resources in the project; the Broad Institute Imaging Platform members for image analysis, statistical analysis, and software engineering (especially Adam Fraser, Adam Papallo, and Martha Vokes); Shomit Sengupta for microscopy and assay guidance; Aviv Regev and Eric Lander for helpful comments on the manuscript; and David Bonnett, Renee Butterfield, Dan Card, Dianne Carpenter, Seth Carpenter, Christopher Lewis, and Themba Nyathi for scoring images for the project. This work was supported by the Broad Institute, the RNAi Consortium, a Novartis fellowship from the Life Sciences Research Foundation (to A.E.C.), a Society for Biomolecular Screening Academic grant (to A.E.C.), a L'Oreal for Women in Science fellowship (to A.E.C.), the Massachusetts Institute of Technology Department of Electrical Engineering and Computer Science/Whitehead/Broad Training Program in Computational Biology, National Institutes of Health Grant DK070069–01 (to T.R.J.), National Science Foundation CAREER Award 0642971 (to P.G.), National Institute of General Medical Sciences Grant R01 GM0725555 (to D.M.S.) and National Institute of Allergy and Infectious Diseases Grant RO1 AI047389 (to D.M.S.).

1. Nusslein-Volhard C, Wieschaus E (1980) Mutations affecting segment number and polarity in Drosophila. *Nature* 287:795–801.
2. Morgan TH (1911) The origin of five mutations in eye color in Drosophila and their modes of inheritance. *Science* 33:534–537.
3. Muller H (1927) Artificial Transmutation of the Gene. *Science* 66:84–87.
4. Hartwell LH, Culotti J, Reid B (1970) Genetic control of the cell-division cycle in yeast. I. Detection of mutants. *Proc Natl Acad Sci USA* 66:352–359.
5. Brenner S (1974) The genetics of Caenorhabditis elegans. *Genetics* 77:71–94.
6. Haffter P, *et al.* (1996) The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development* 123:1–36.
7. Driever W, *et al.* (1996) A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* 123:37–46.
8. Su AI, Hogenesch JB (2007) Power-law-like distributions in biomedical publications and research funding. *Genome Biol* 8:404.
9. Eggert US, Mitchison TJ (2006) Small molecule screening by imaging. *Curr Opin Chem Biol* 10:232–237.
10. Carpenter AE (2007) Image-based chemical screening. *Nat Chem Biol* 3:461–465.
11. Carpenter AE, Sabatini DM (2004) Systematic genome-wide screens of gene function. *Nat Rev Genet* 5:11–22.
12. Kiger A, *et al.* (2003) A functional genomic analysis of cell morphology using RNA interference. *J Biol* 2(4):27.
13. Carpenter AE, *et al.* (2006) CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* 7(10):R100.
14. Moffat J, *et al.* (2006) A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell* 124:1283–1298.
15. Bakal C, Aach J, Church G, Perrimon N (2007) Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science* 316:1753–1756.
16. Neumann B, *et al.* (2006) High-throughput RNAi screening by time-lapse imaging of live human cells. *Nat Methods* 3:385–390.
17. Orlov N, Johnston J, Macura T, Shamir L, Goldberg I (2007) Computer Vision for Microscopy Applications. *Vision Systems*: *Segmentation and Pattern Recognition*, eds Goro Obinata and Ashish Dutta (I-Tech, Vienna), pp 221–242.
18. Lin C, Mak W, Hong P, Sepp K, Perrimon N (2007) Intelligent Interfaces for Mining Large-Scale RNAi-HCS Image Databases. *IEEE 7th International Conference on Bioinformatics and Biomedical Engineering* (IEEE, Washington DC).
19. Chen X, Murphy RF (2006) Automated interpretation of protein subcellular location patterns. *Int Rev Cytol* 249:193–227.
20. Loo LH, Wu LF, Altschuler SJ (2007) Image-based multivariate profiling of drug responses from single cells. *Nat Methods* 4:445–453.
21. Adams CL, *et al.* (2006) Compound classification using image-based cellular phenotypes. *Methods Enzymol* 414:440–468.
22. Tanaka M, *et al.* (2005) An unbiased cell morphology-based screen for new, biologically active small molecules. *PLoS Biol* 3(5):e128.
23. Young DW, *et al.* (2008) Integrating high-content screening and ligand-target prediction to identify mechanism of action. *Nat Chem Biol* 4:59–68.
24. Wang J, *et al.* (2008) Cellular phenotype recognition for high-content RNA interference genome-wide screening. *J Biomol Screen* 13:29–39.
25. Boland MV, Murphy RF (2001) A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics* 17:1213–1223.
26. Boland MV, Markey MK, Murphy RF (1998) Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images. *Cytometry* 33:366–375.
27. Perlman ZE, *et al.* (2004) Multidimensional drug profiling by automated microscopy. *Science* 306:1194–1198.
28. Friedman JH, Hastie T, Tibshirani R (2000) Additive logistic regression: A statistical view of boosting. *Ann Stat* 28(2):337–407.
29. Sigal A, *et al.* (2006) Dynamic proteomics in individual human cells uncovers widespread cell-cycle dependence of nuclear proteins. *Nat Methods* 3:525–531.
30. Levsky JM, Singer RH (2003) Gene expression and the myth of the average cell. *Trends Cell Biol* 13:4–6.
31. Friedman A, Perrimon N (2007) Genetic screening for signal transduction in the era of network biology. *Cell* 128:225–231.
32. Wichmann FA, Graf ABA, Simoncelli EP, Bülthoff HH, Schölkopf B (2004) Machine learning applied to perception: Decision images for gender classification. *Adv Neural Info Processing Syst* 17:1489–1496.
33. Jones TR, *et al.* (2008) CellProfiler Analyst: Data exploration and analysis software for complex image-based screens. *BMC Bioinformatics* 9:482
34. Mukherji M, *et al.* (2006) Genome-wide functional analysis of human cell-cycle regulators. *Proc Natl Acad Sci USA* 103:14819–14824.
35. Echeverri CJ, *et al.* (2006) Minimizing the risk of reporting false positives in large-scale RNAi screens. *Nat Methods* 3:777–779.
36. Torralba A, Murphy KP, Freeman WT (2007) Sharing visual features for multiclass and multiview object detection. *IEEE Trans Pattern Anal Machine Intell* 29:854–869.