

Robust open source informatics workflow for high-throughput microscopy

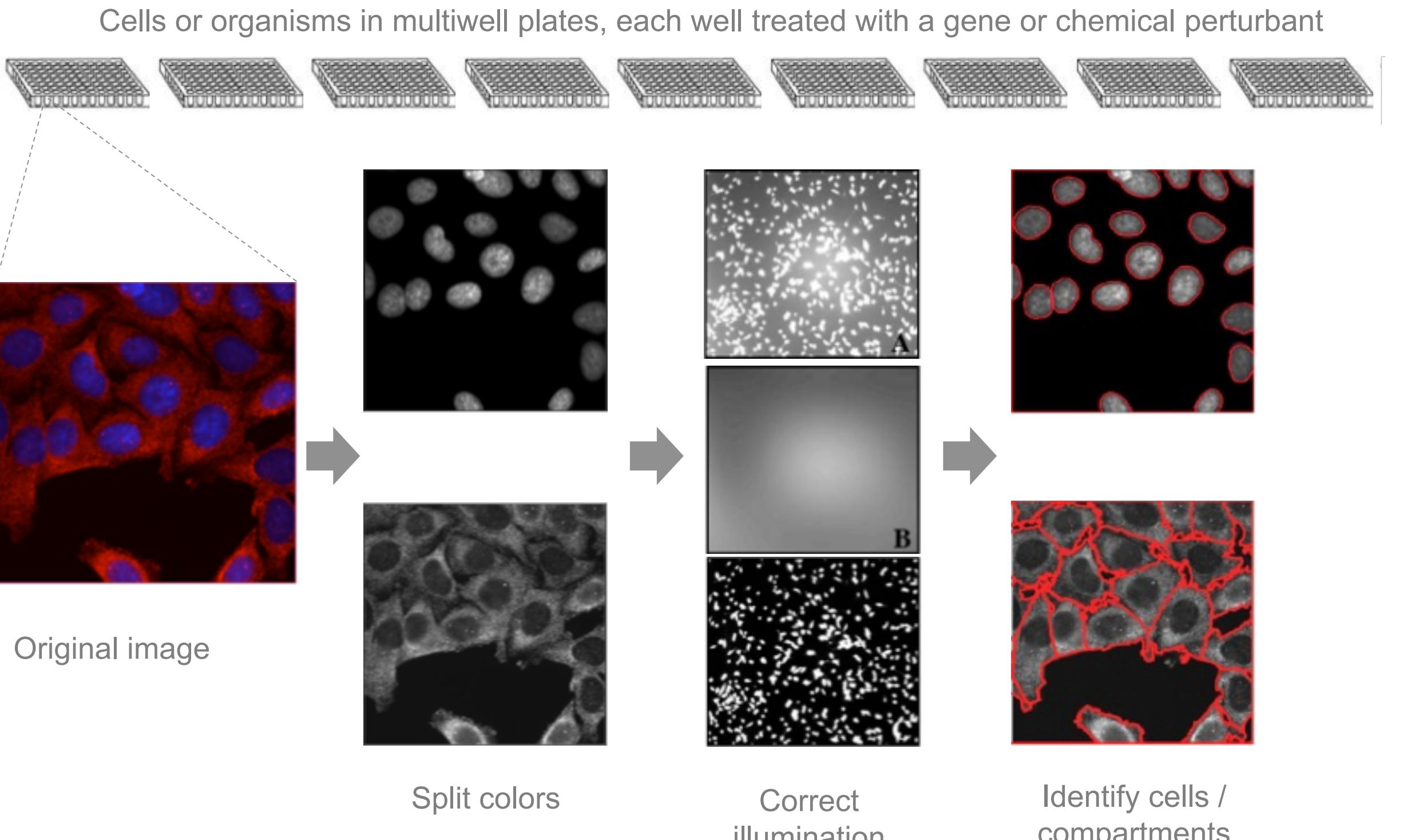
SHANTANU SINGH



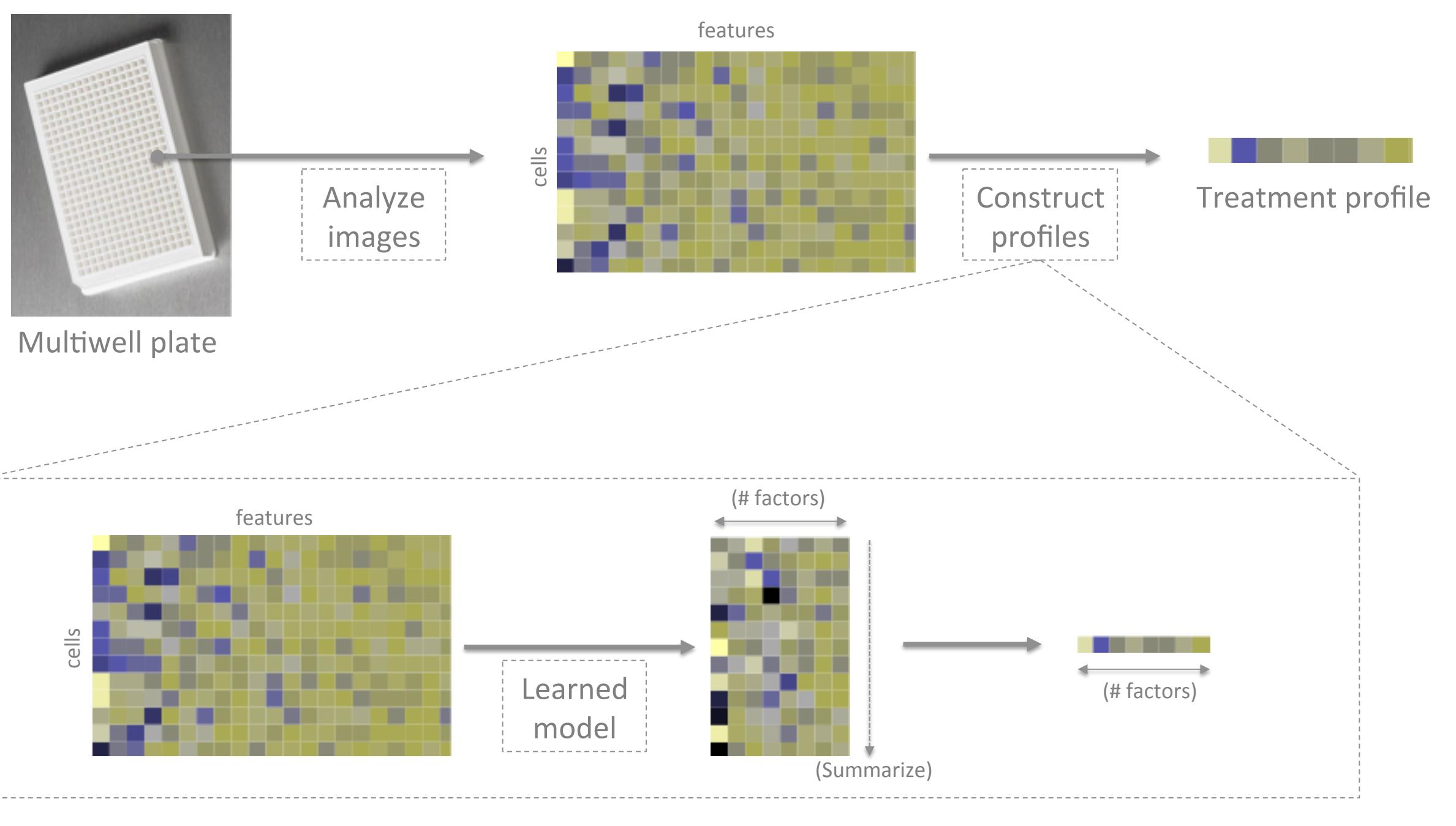
Imaging Platform, Broad Institute of MIT and Harvard

Large-scale imaging experiments

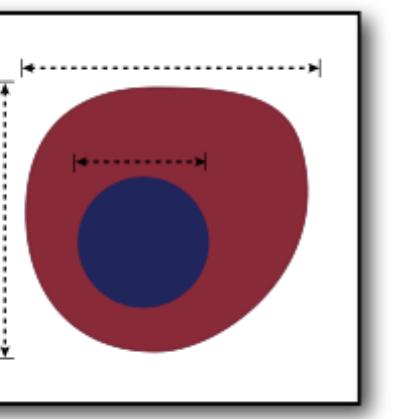
Image analysis pipelines can quantify images



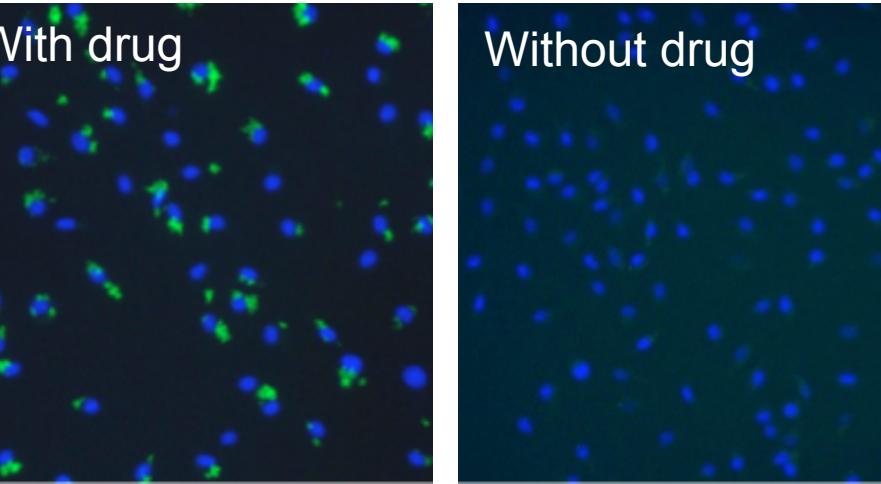
Quantified images can be used to create perturbation profiles



First Wave: Measure known phenotypes



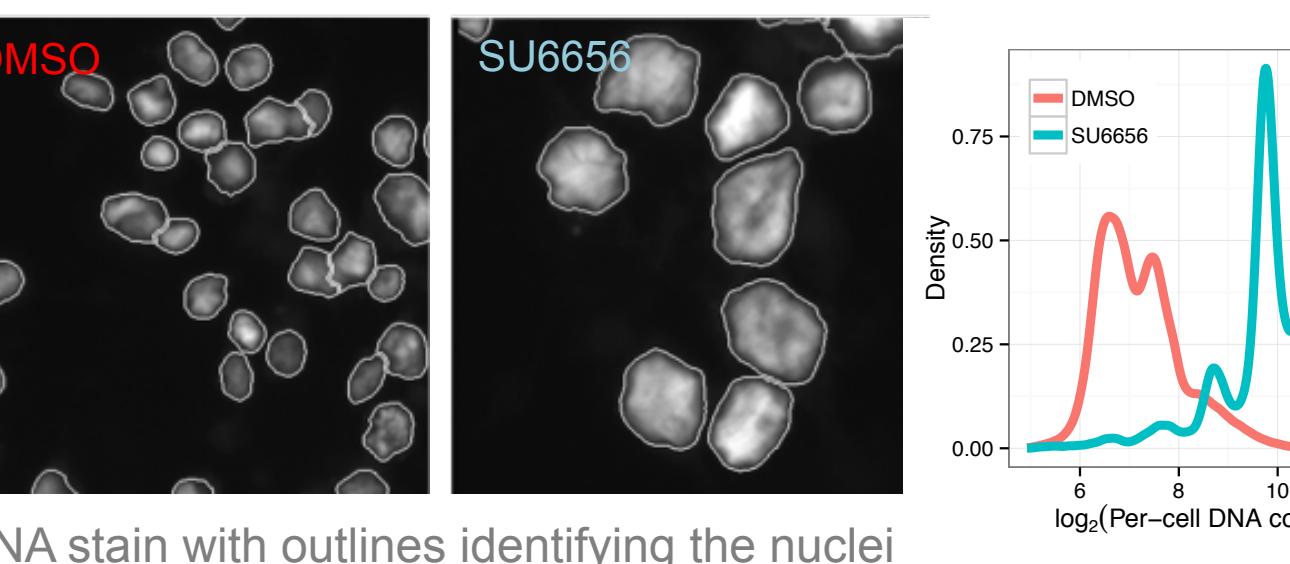
Search for Tuberculosis treatments



Identified hits that prevent bacterial infection/expansion but do NOT kill the bacterium directly

Stanley SA, et al. (2014). A chemical screen to identify host-targeting small molecules that restrict intracellular Mycobacterium tuberculosis growth. *PLoS Pathogens* 10(2):e1003946

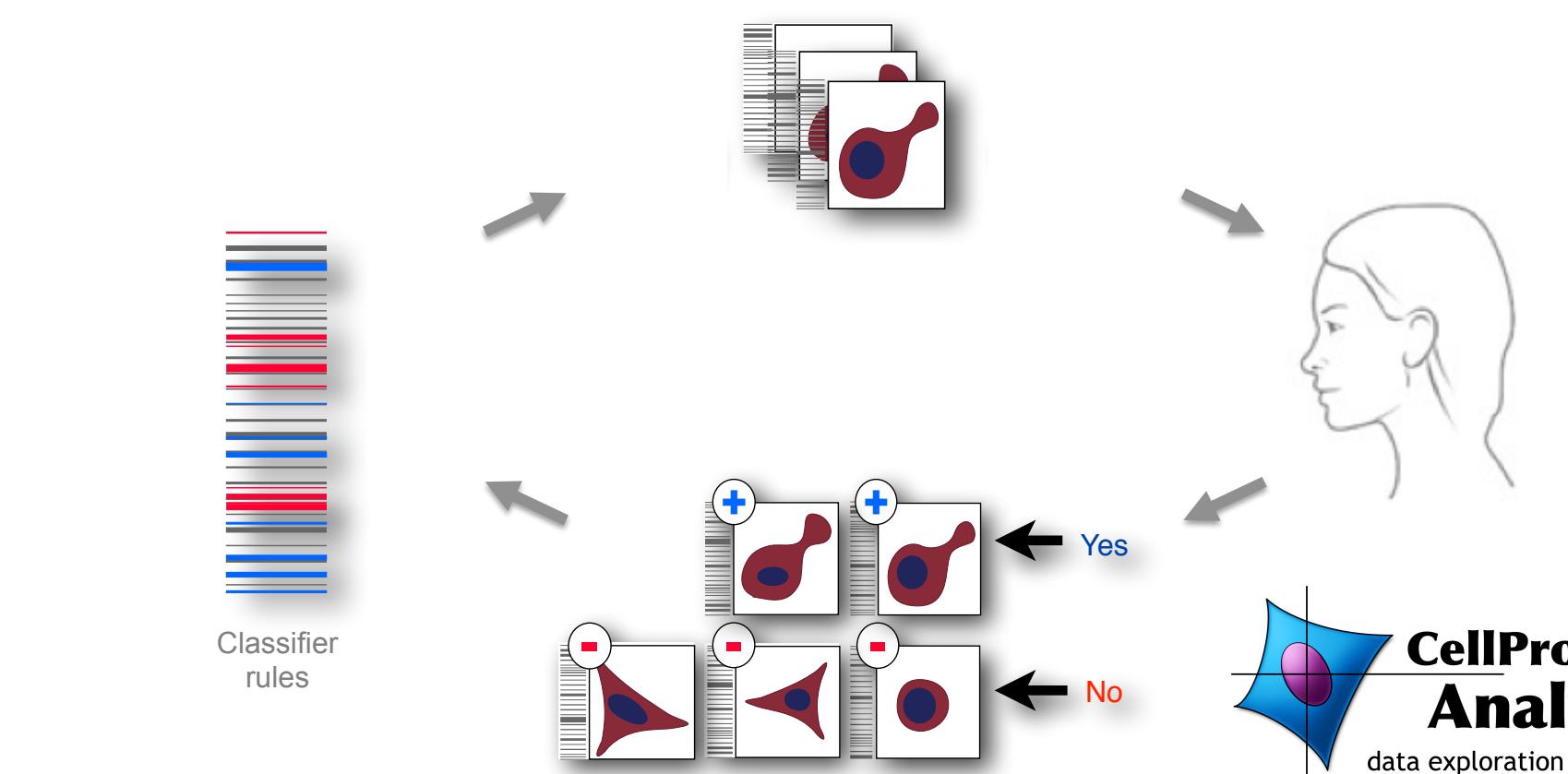
Polyplloidization of megakaryocytes - AMKL (leukemia)



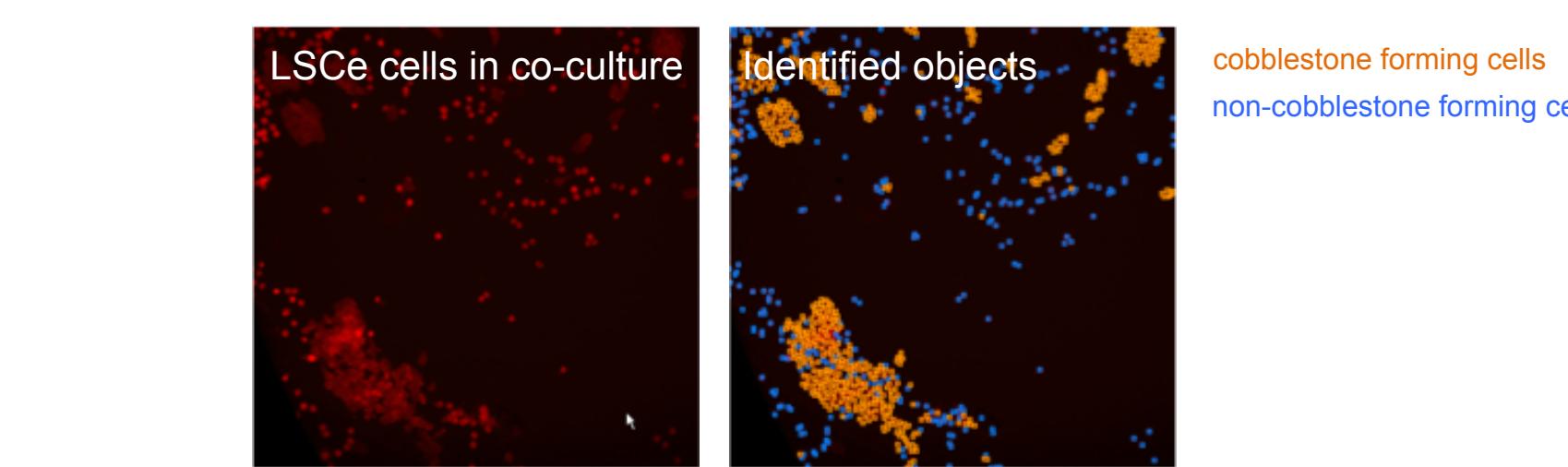
Recommended clinical testing of AURKA inhibitor for AMKL

Wen Q, et al. (2012). Identification of regulators of polyplloidization presents therapeutic targets for treatment of AMKL. *Cell* 150(3):575-89

Second Wave: Train for known phenotypes



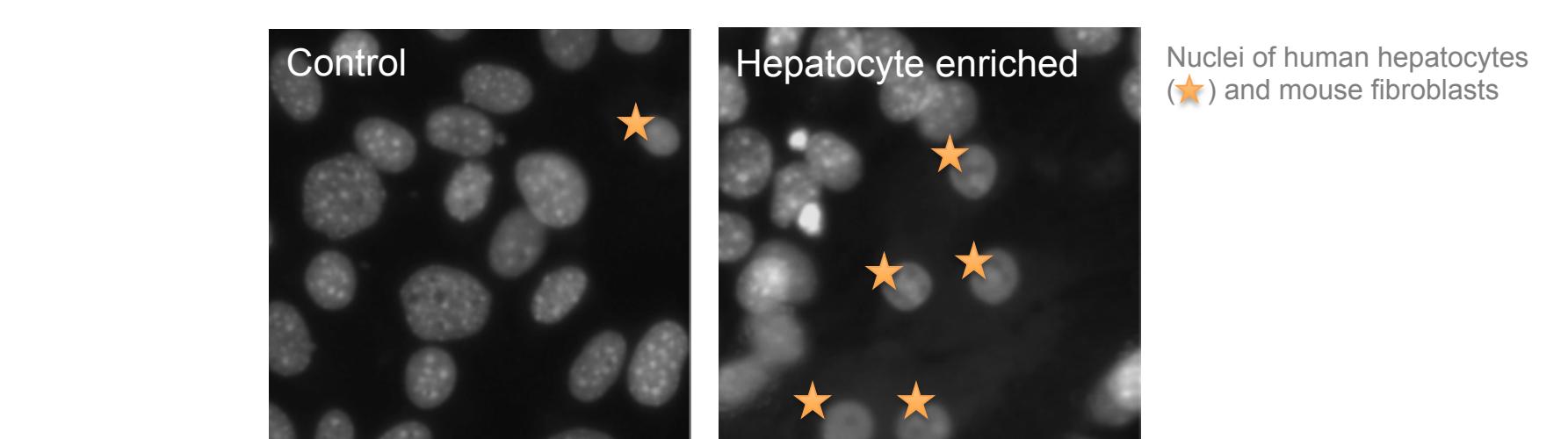
Leukemic & hematopoietic stem cells



Identified drugs that preferentially kill leukemic cells

Hartwell KA, et al. (2013). Niche-based screening identifies small-molecule inhibitors of leukemia stem cells. *Nature Chemical Biology* 12(8):840-848

Hepatocyte proliferation human primary liver cells co-cultured with fibroblasts



Identified chemicals that stimulate primary human hepatocyte proliferation

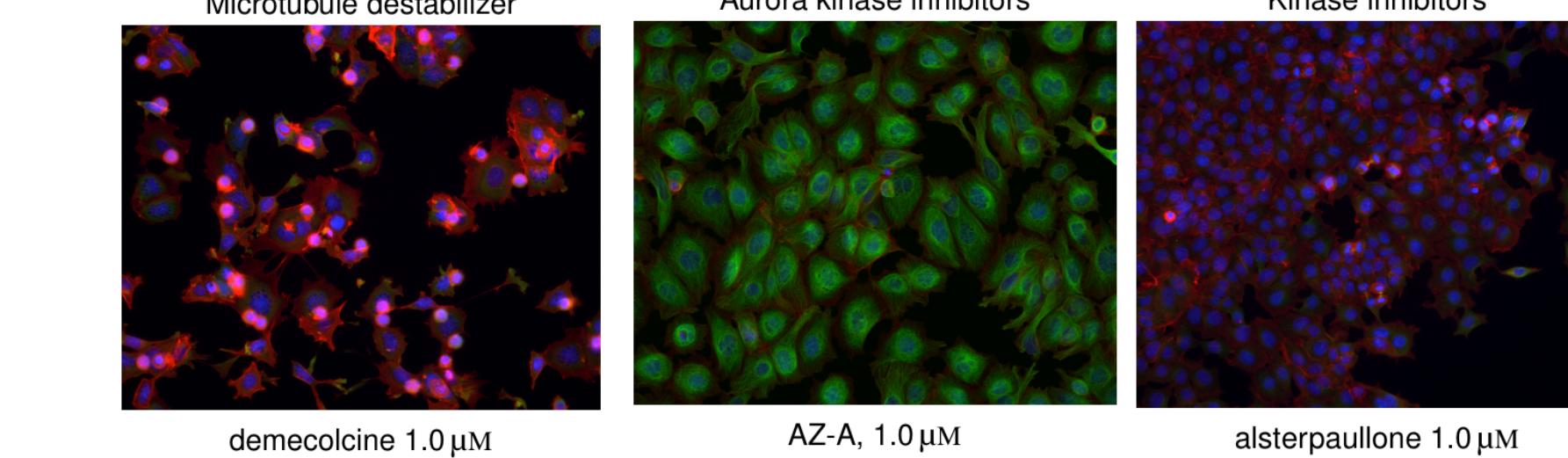
Shan J, et al. (2013). Identification of small molecules for human hepatocyte expansion and iPSC differentiation. *Nature Chemical Biology* 9(8):514-520

Case Studies

Third Wave: Profile to identify phenotypes and characterize samples



Identifying mechanism of action of compounds using image-based profiling



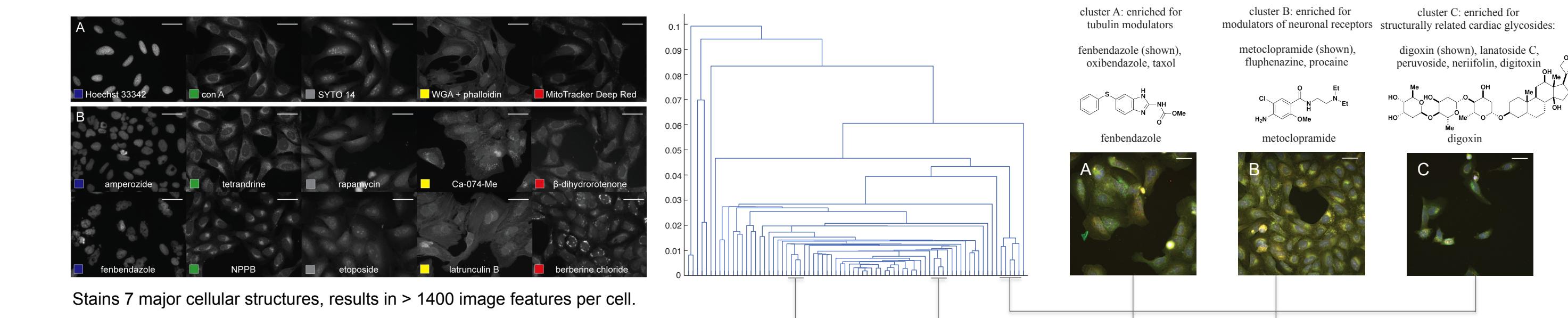
F: Factor analysis (mean of 20 models)	True mechanistic class	Predicted class	Acc.
Act	1	80 %	100 %
Actin disruptors	2	100 %	100 %
Aurora kinase inhibitors	3	94 %	94 %
Ch	4	82 %	82 %
Chk	5	77 %	77 %
Chk1	6	100 %	100 %
DNA damage	7	93 %	93 %
DD	8	95 %	95 %
DNA replication	9	100 %	100 %
DR	10	98 %	98 %
Eg5 inhibitors	11	100 %	100 %
Epi	12	100 %	100 %
Kinase inhibitors	13	100 %	100 %
Microtubule destabilizers	14	100 %	100 %
MD	15	100 %	100 %
Microtubule stabilizers	16	100 %	100 %
MS	17	100 %	100 %
Protein degradation	18	100 %	100 %
Protein synthesis	19	100 %	100 %
PS	20	100 %	100 %

Overall accuracy: 94%

A "generic" image-based assay can classify compounds into mechanisms with 94% accuracy

Ljosa V, et al. (2013). Comparison of Methods for Image-Based Profiling of Cellular Morphological Responses to Small-Molecule Treatment. *Journal of Biomolecular Screening* 18:1321-1329

Profiling compound libraries using "CellPainting"



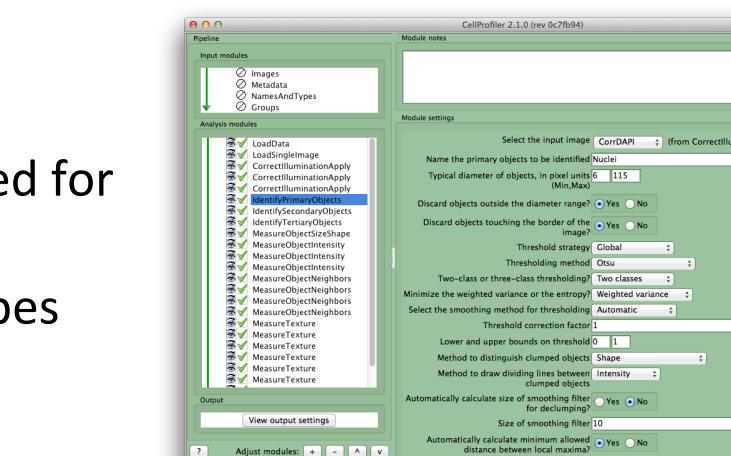
Designed assay to maximize potential insights gained from a single image-based profiling experiment. Clustering is able group compounds like modulators of neuronal receptors (B) and cardiac glycosides (C), despite not using directly relevant cell types or stains.

Gustafsson SM, et al. (2013). Multiplex cytological profiling assay to measure diverse cellular states. *PLoS One* 8(12):e80999

Open source software



CellProfiler cell image analysis software is designed for biologists without training in computer vision or programming to quantitatively measure phenotypes from thousands of images automatically



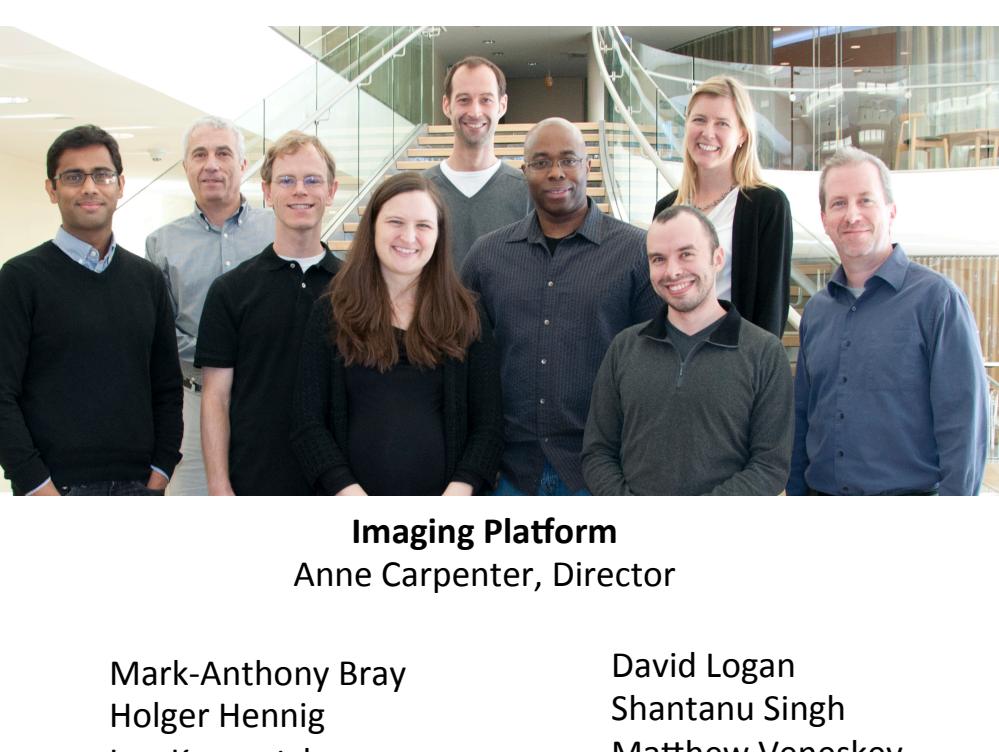
- Publications citing CellProfiler
-
- Launched >130,000 times/year
 - Downloaded >70,000 times
 - Winner of Bio-IT World's IT & Informatics Best Practices Award in 2009
 - One of the Top 10 most-accessed papers of all time in *Genome Biology* (> 940 citations)
- Getting Started
- Download CellProfiler at cellprofiler.org (Windows, Mac, Linux)
 - View CellProfiler demo movies
 - Try example pipelines – website has several biological assays, pick the ones that most resemble your phenotypes of interest
 - Go online for user support – www.cellprofiler.org/forum/

CellProfiler Analyst is an interactive data exploration tool that includes an iterative machine learning system which can be trained to recognize complicated and subtle phenotypes.

Support

Recent major funding for this work provided by:

NIH NIGMS R01 GM089652 (Carpenter)
NIH NIGMS R01 GM095672 (Wahlby)
NSF CAREER award (Carpenter)
Human Frontiers in Science Program (Carpenter)



Mark-Anthony Bray
Holger Hernig
Lee Kamensky
Vebjorn Ljosa
David Logan
Shantanu Singh
Matthew Veneskey
Carolina Wahlby