Image analysis software for high throughput cell-based screening of living cell microarrays

Anne E. Carpenter, Ph.D.
Whitehead Institute for Biomedical Research

New opportunities to determine gene function

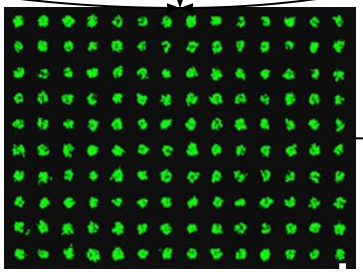
Set of RNA interference reagents, one for each gene

Living cell microarray technology

Automated image collection

Automated image analysis

High throughput biological assay

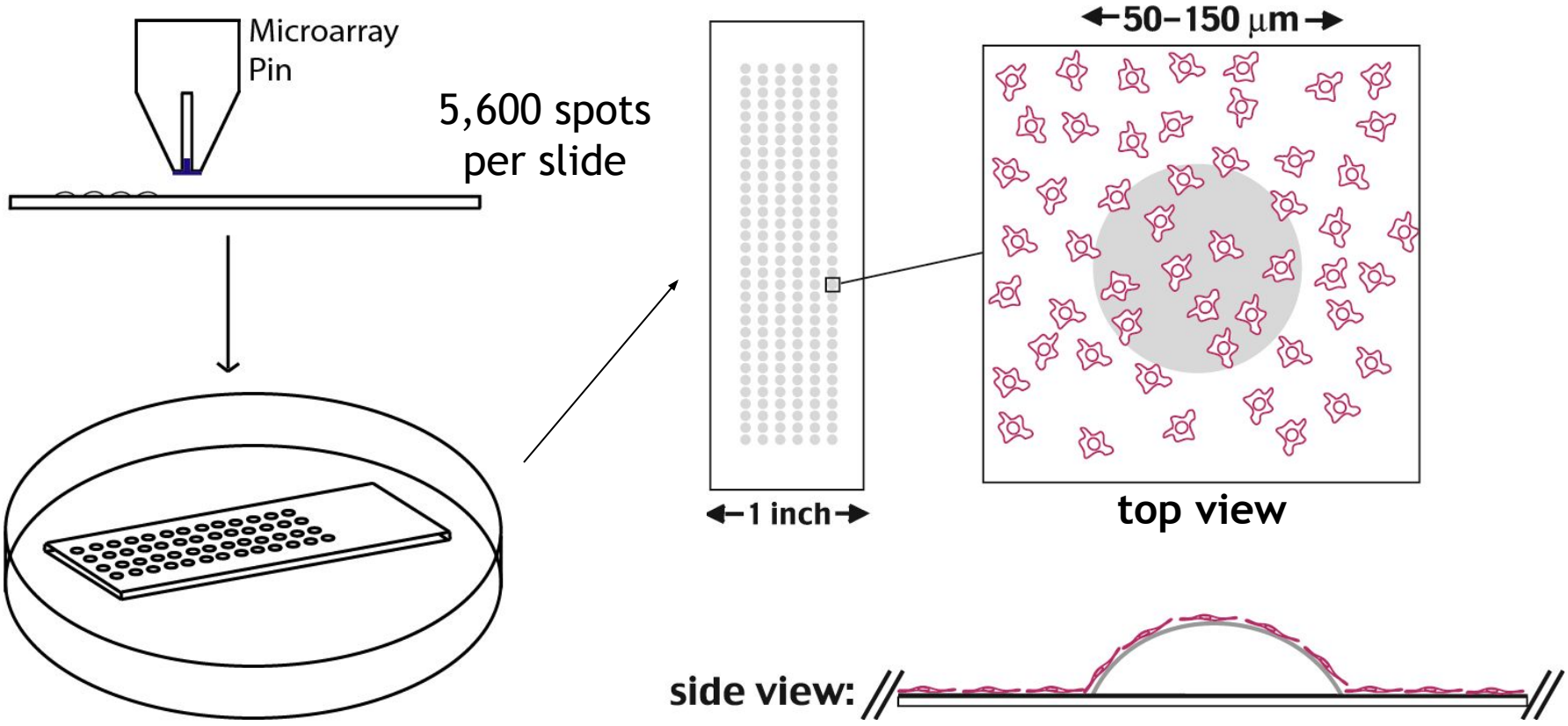


Determine the phenotypic effects of knocking down every gene in the genome

SYSTEMATIC GENOME-WIDE SCREENS OF GENE FUNCTION

Anne E. Carpenter and David M. Sabatini Nature Reviews Genetics 5:11-22

Living cell microarray technology

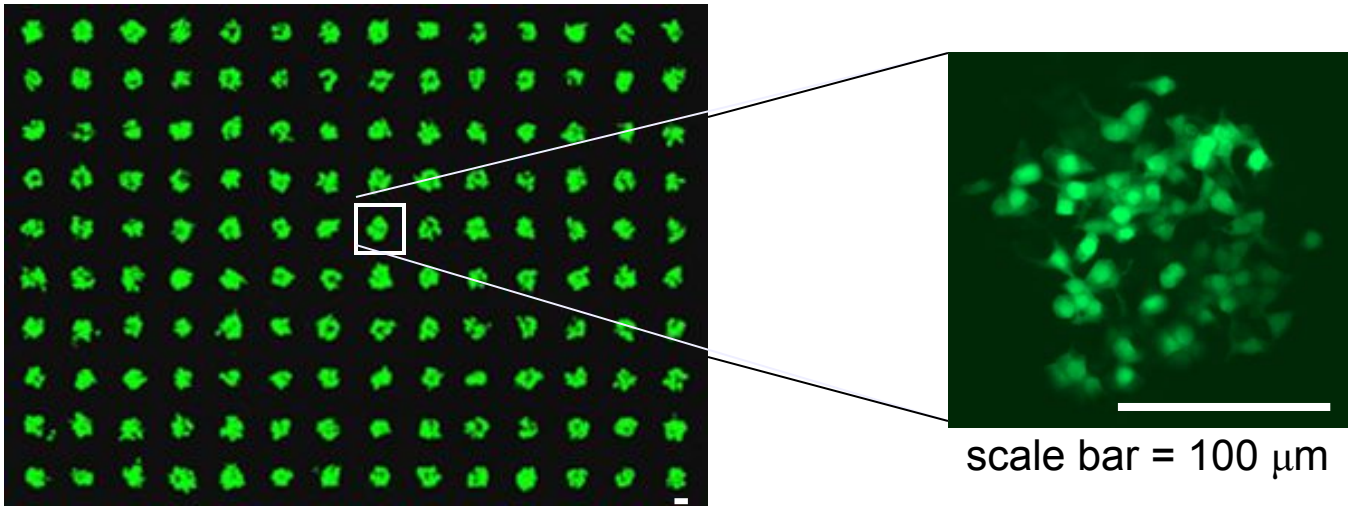


cDNA expression	→	gain of function
RNA interference	→	loss of function
small molecules	→	chemical genetics/ drug discovery

Ziauddin & Sabatini: Nature, 411:107-110 (2001)

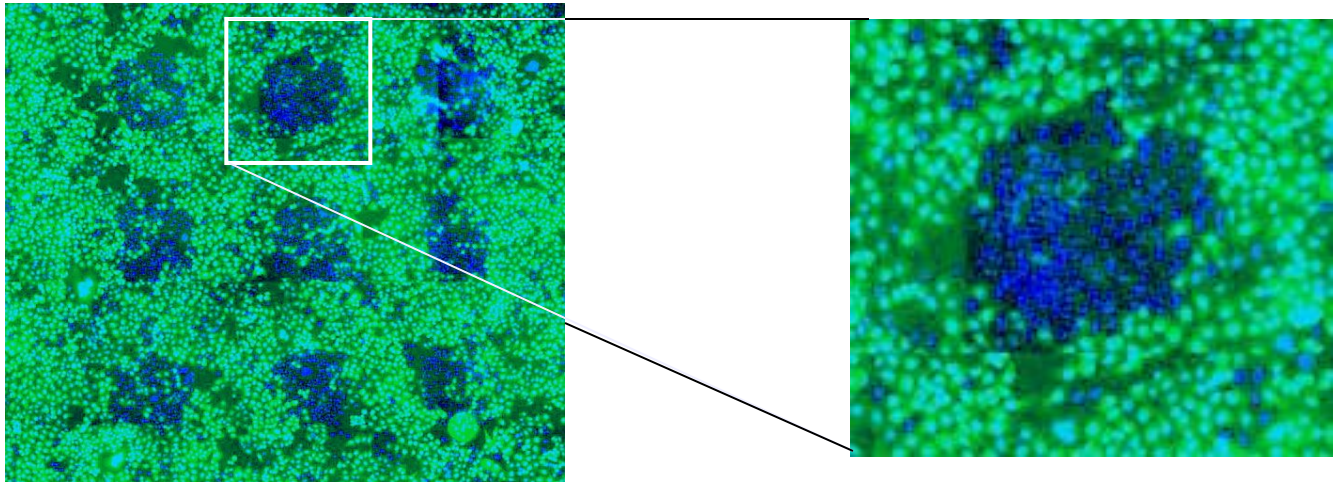
- Fast
- Cheap
- Requires little reagent/ few cells
- Uniform - can see subtle effects
- Synthetic genetic interactions easy
- High content screening/western blots

Over and under-expression by transfection



Spots = GFP expression plasmid

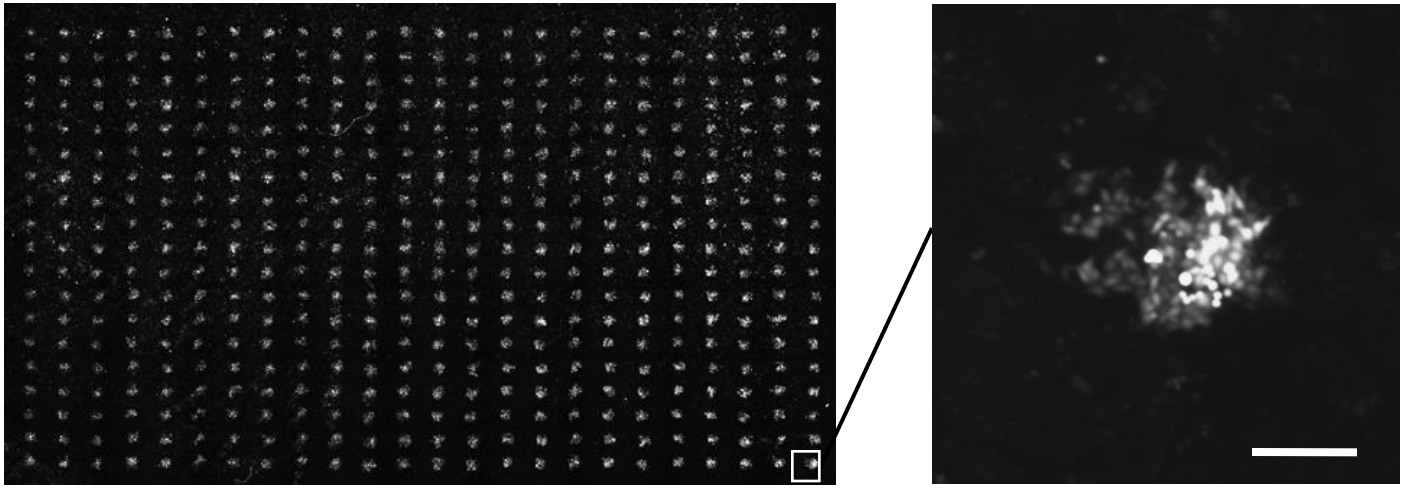
Cells = unmodified mammalian cells



Spots = synthetic siRNA against GFP

Cells = HeLa cells stably expressing GFP

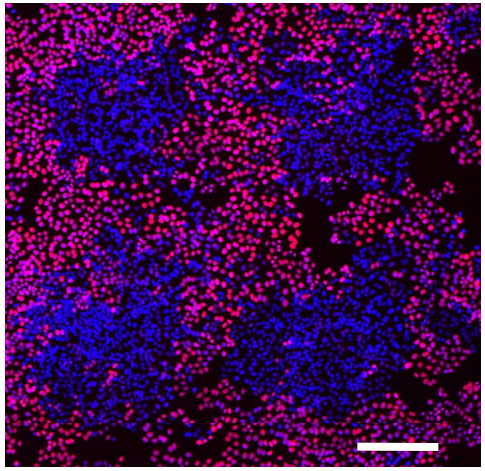
Over and under expression with lentivirus



Spots = GFP expressing lentivirus

Cells = unmodified mammalian cells

The RNAi Consortium (TRC)

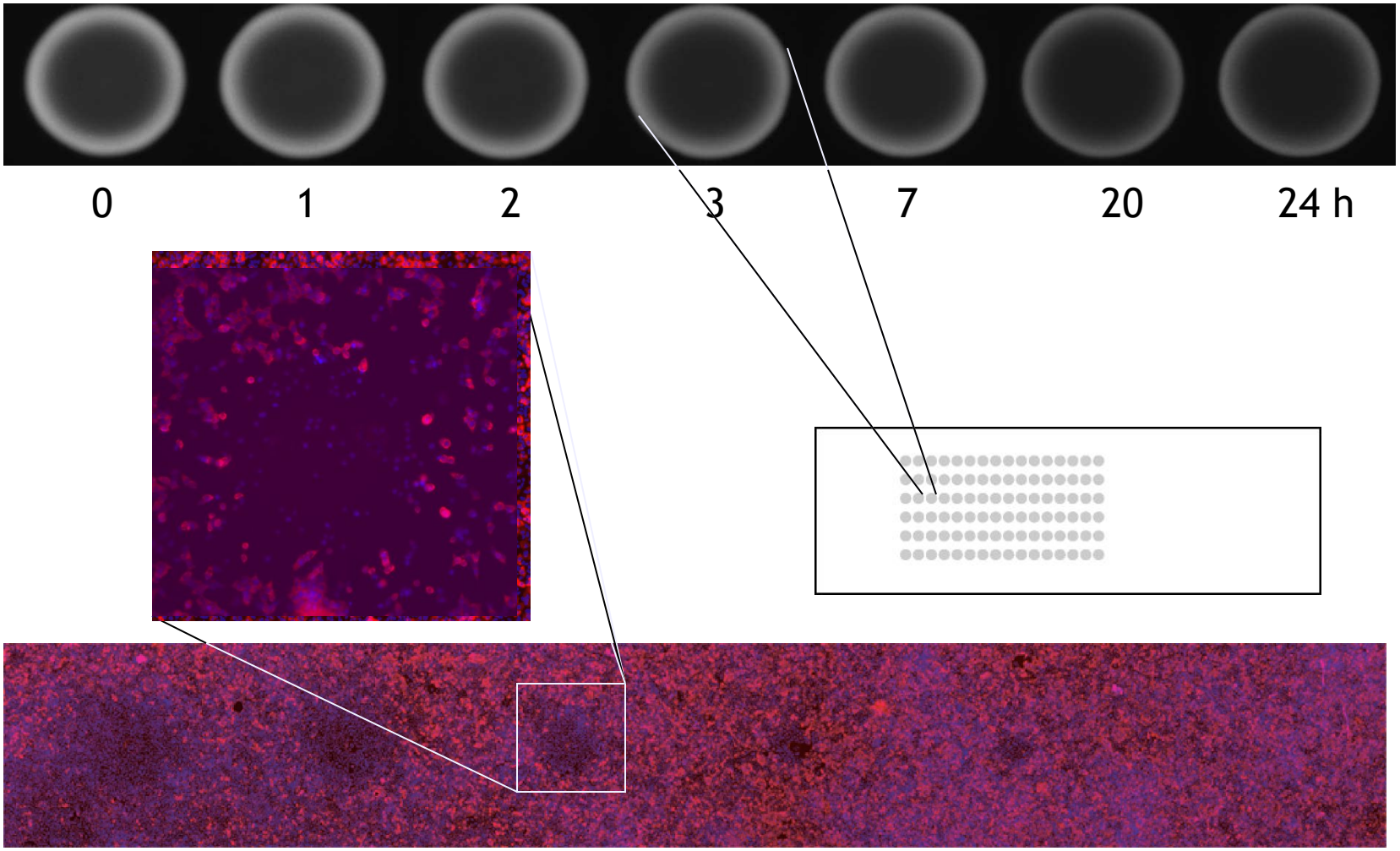


Spots = lamin A shRNA-lentiv irus

Cells = unmodified mammalian cells

scale bars = 100 μm

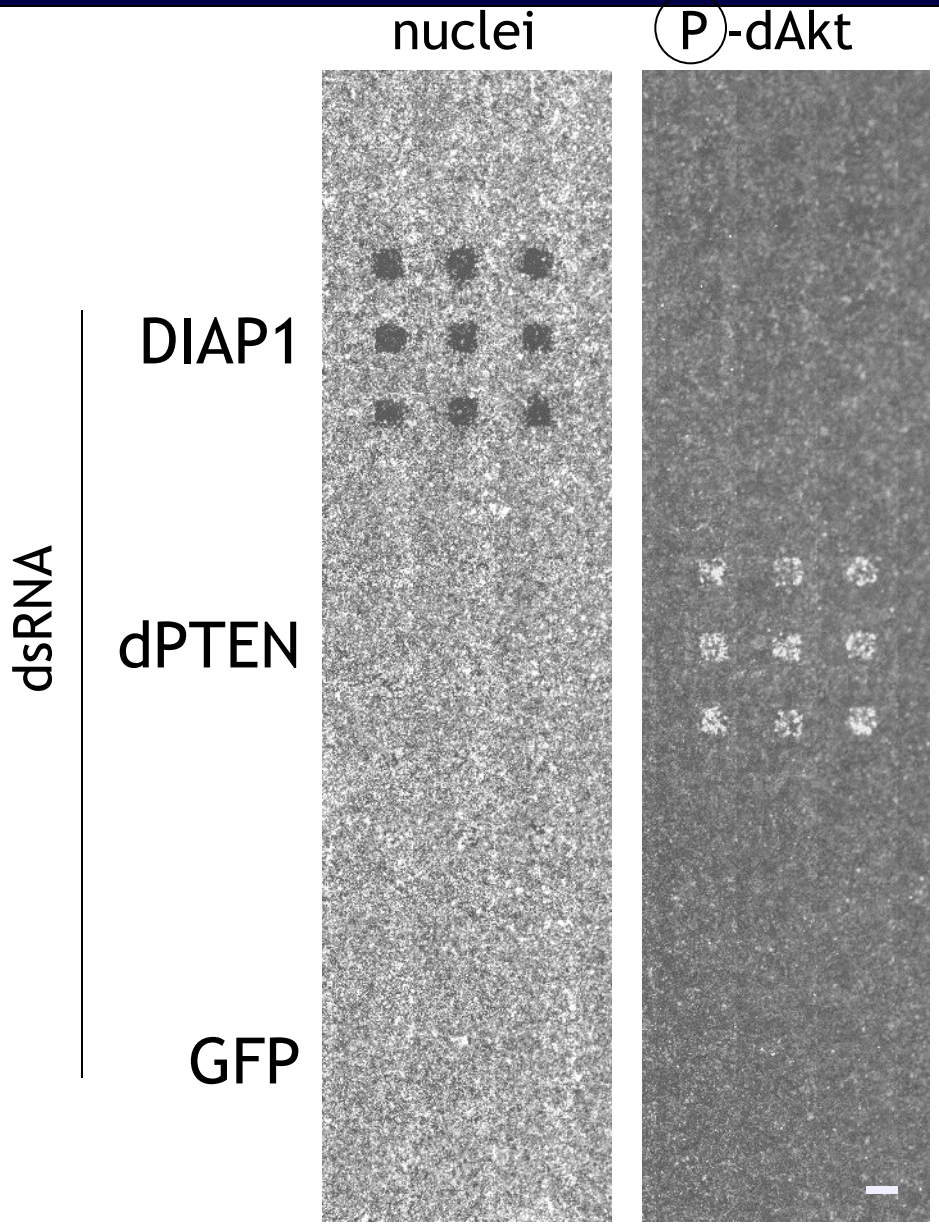
Small molecule living cell microarrays



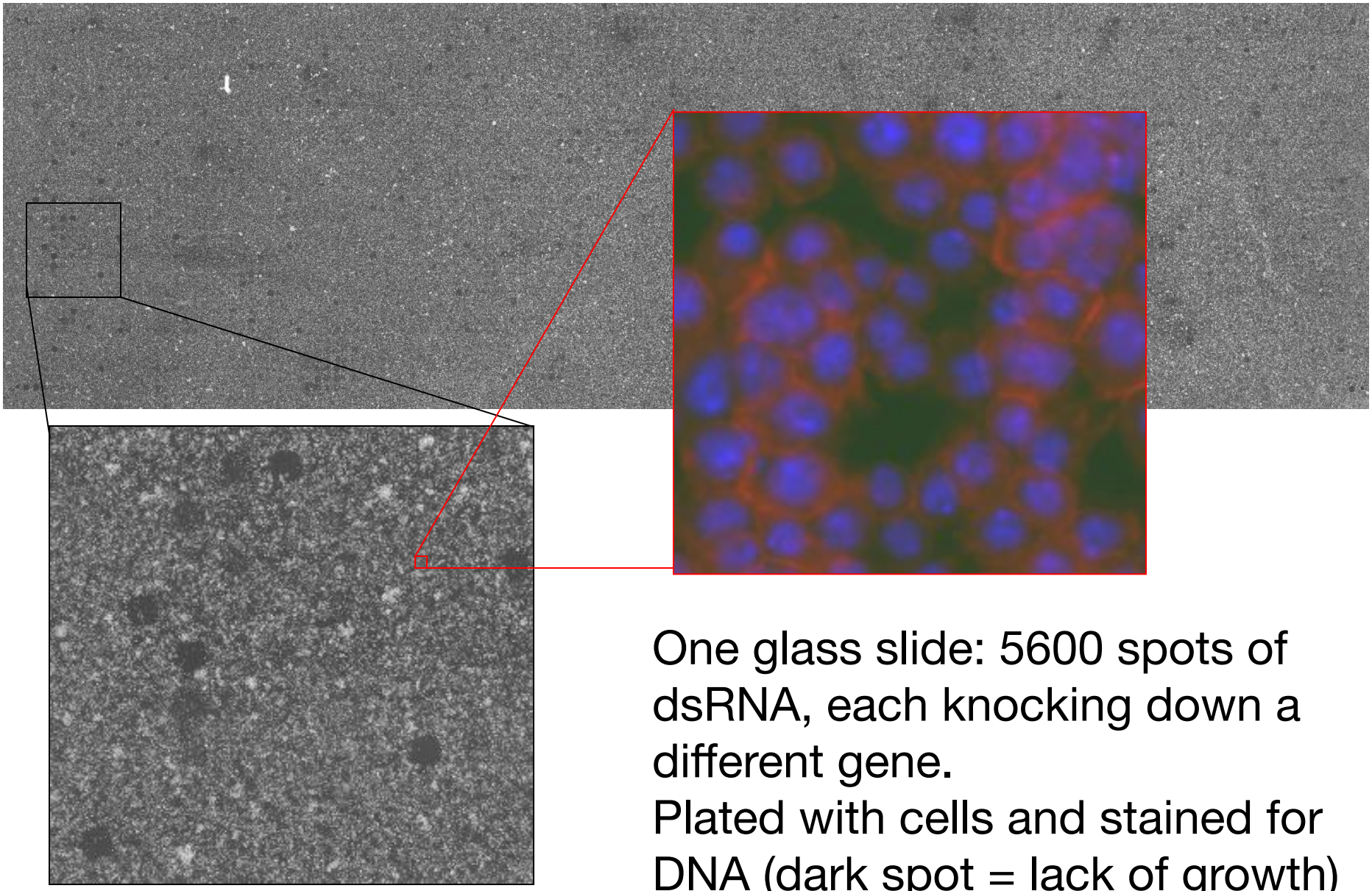
Application to *Drosophila*

Drosophila:

- RNAi effective, reliable, and convenient
- Genome-wide libraries available
- No interferon response
- Less redundancy vs. mammals



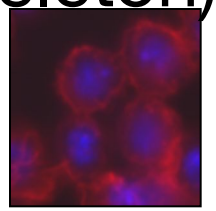
Genome-wide cell-based experiments

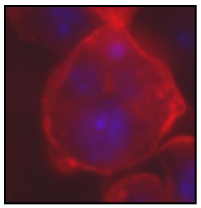


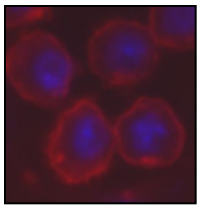
One glass slide: 5600 spots of dsRNA, each knocking down a different gene.
Plated with cells and stained for DNA (dark spot = lack of growth)

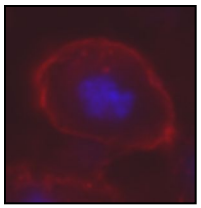
How can we identify unusual cells automatically?

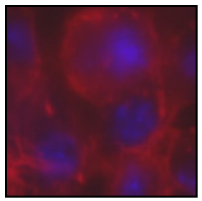
Zoomed in: blue = DNA, red = actin (cytoskeleton)

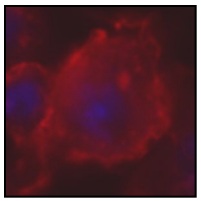

GFP

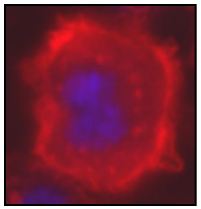

PPV

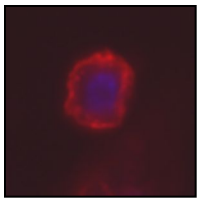

dTOR


PP2A Cat. Sub.


CG9006


String


Cyclin A

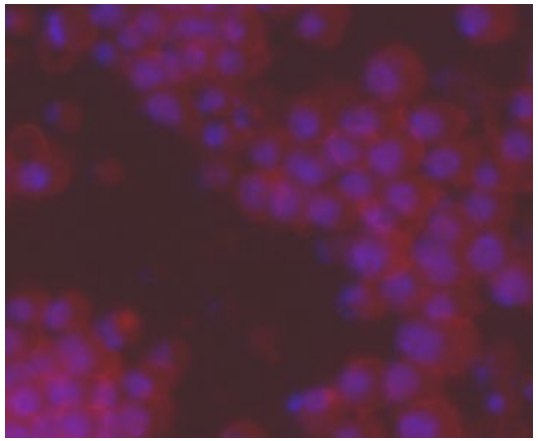

rpL12

…plus 19,000 more images

We want to know quantitatively: size, shape, intensity, texture, overlap of colors, etc for *every* cell in *every* image.

How can we identify unusual phenotypes?

Analyze >20 phenotypes by automated microscopy of genome-wide RNAi-treated living cell microarrays: identify required genes and decipher the signaling network



Cell size, count & morphology

Nuclear size & morphology

Nucleolar size, count & morphology

Cell cycle distribution (DNA content)

Upregulation of transcription

Apoptosis

Amount & distribution of signaling lipids

Upregulation of translation

Mitochondrial size & morphology

Localization of signaling proteins

Phosphorylation of signaling proteins (S6, S6K, Akt)

Actin content/ morphology

Protein content

Quantitative automatic image analysis is needed

Runs on Mac/PC/Unix, plugs into Matlab

Image file types: tif, jpg, bmp, gif, cur, hdf, ico, pbm, pcx, pgm, png, ppm, ras, xwd

Allows quantitative analysis of various cell phenotypes in thousands of images (high-throughput experiments, time lapse, etc.)

Usable by cell biologists without programming knowledge

Modular design allows custom image analysis modules to be added

Whitehead Institute for Biomedical Research

Anne E. Carpenter
Whitehead Institute for Biomedical Research: Laboratory of David Sabatini

Thouis R. Jones
MIT Computer Sciences/ Artificial Intelligence Laboratory: Laboratory of Polina Golland

MIT

CellProfiler™ modules

Pre-process images:

Align
Apply Threshold
Crop
Correct Illumination
Invert Intensity
RGB Merge
RGB Split
Saturation Check

File Renumber
File Renamer
File Format Converter

CellProfiler™ modules

Object identification:

Primary objects (e.g. Nuclei):
AdaptiveThreshold
Distance
Intensity
Threshold

Secondary objects (e.g. Cells):
Distance
Watershed
Propagate

Tertiary objects (e.g. Cytoplasm):
Subregion

CellProfiler™ modules

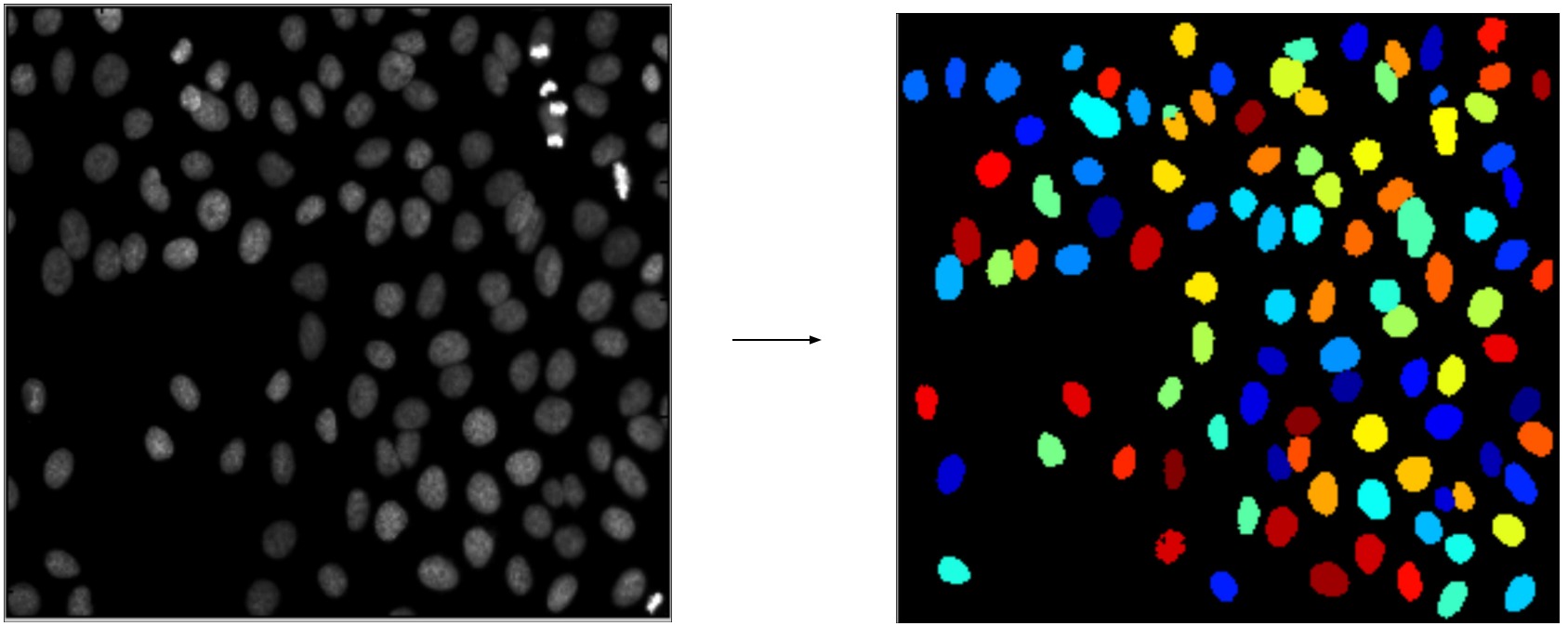
Measurement:

Each object:
Area
Shape
Intensity
Texture

Entire image:
Area Occupied
Total Intensity
Correlation

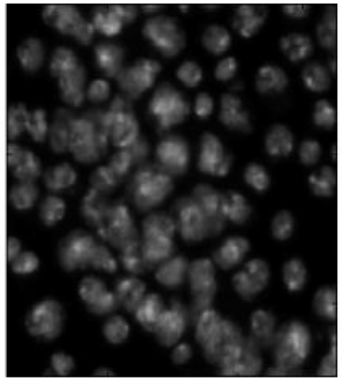
Improved identification of nuclei (Shape)

Identifies and separates clumps of objects when each object is fairly round. Ideal for HeLa nuclei.



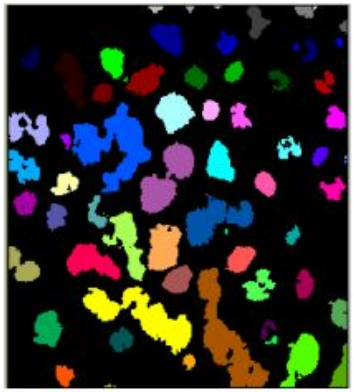
Improved identification of nuclei (Intensity)

Identifies and separates clumps of objects when each object is brighter in the center. Ideal for *Drosophila* nuclei.



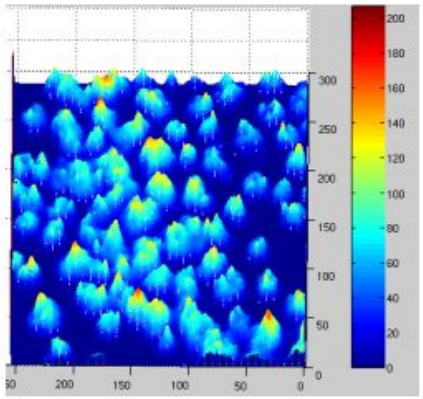
original image

OLD →

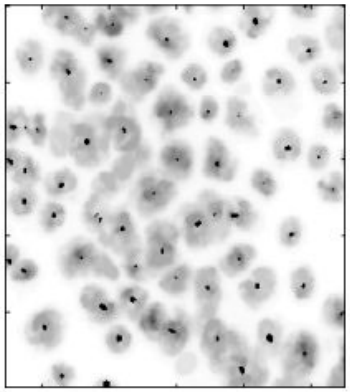


segmented nuclei

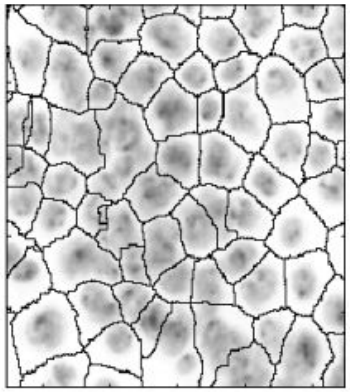
NEW →



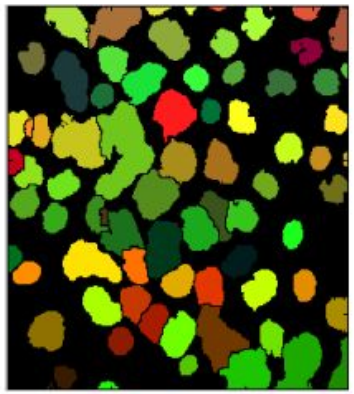
intensity represented in 3-D



intensity maxima



watershed lines



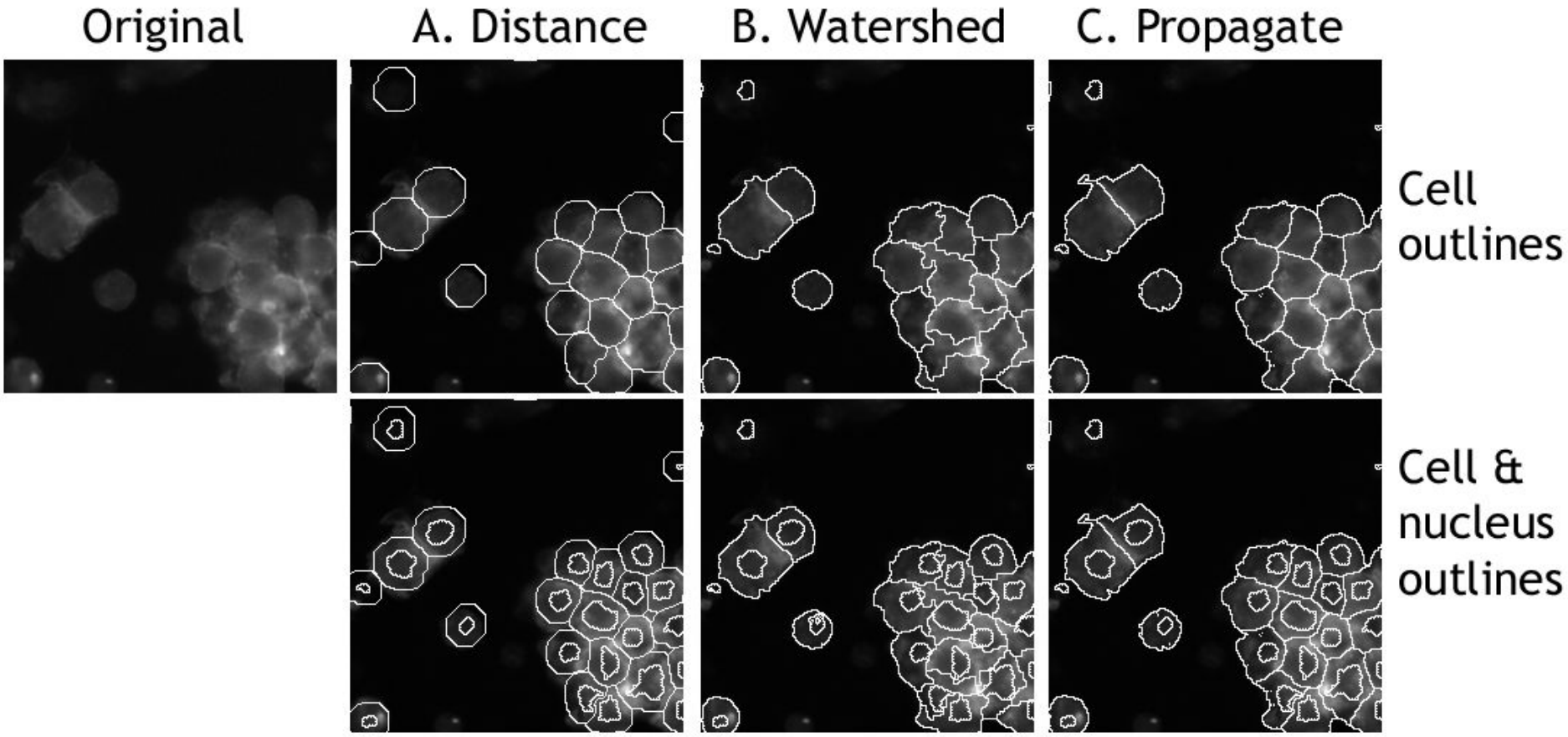
segmented nuclei

based on vision research by:

- Malpica & del Pozo
- Wahlby & Bengtsson
- Ortiz de Solorzano & Lockett

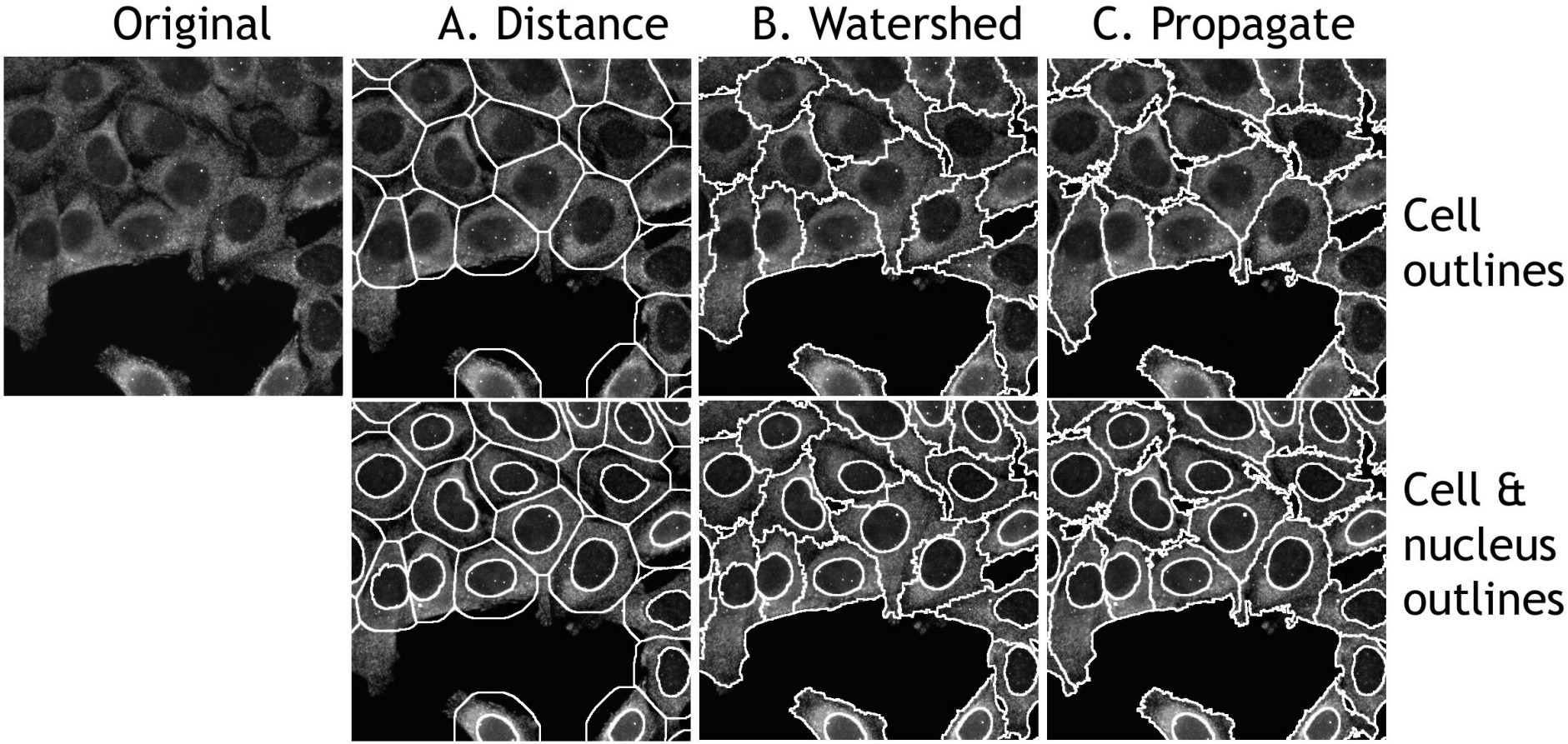
Cell edges can be identified using actin staining

Drosophila

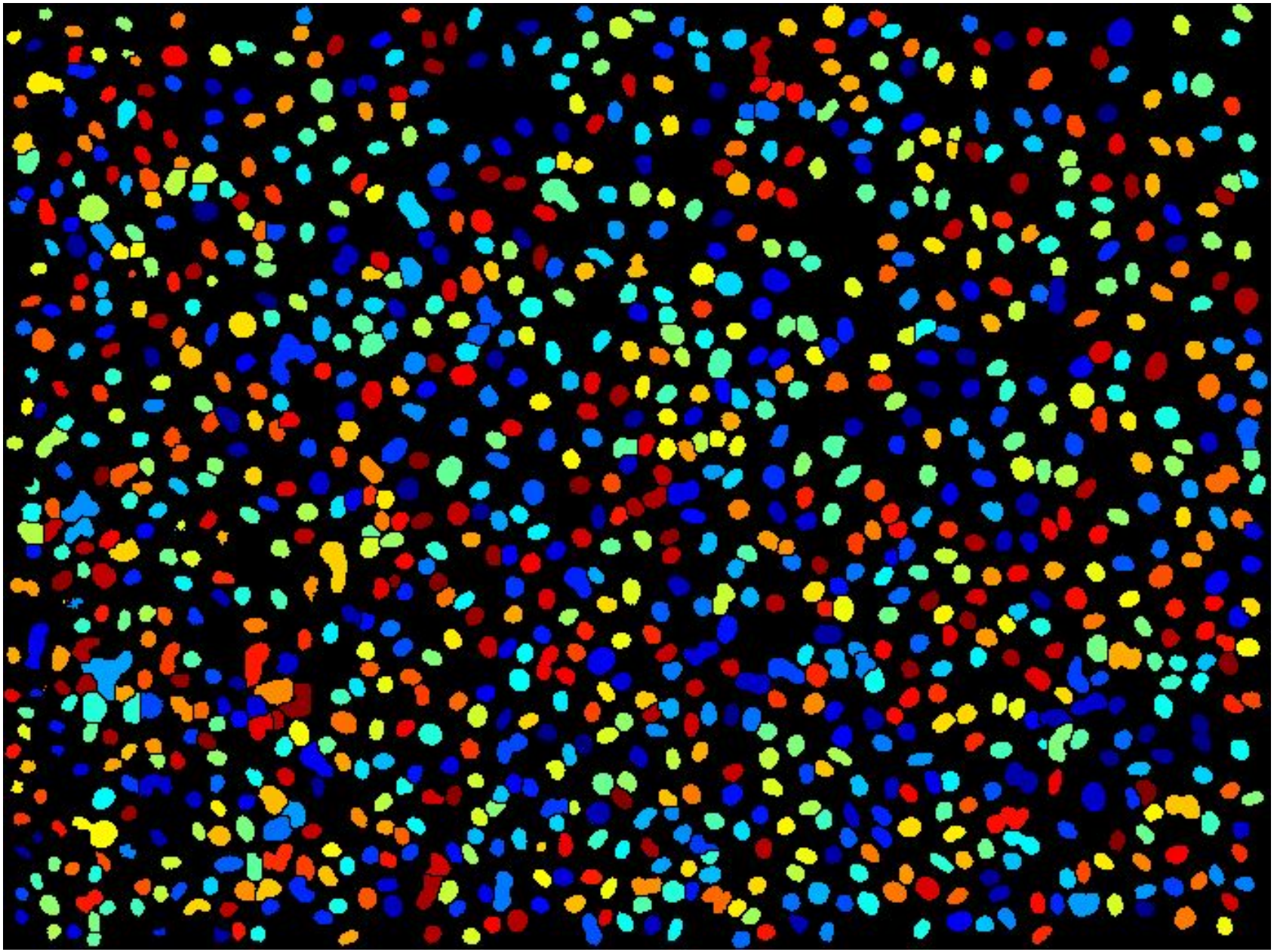


Cell edges can be identified using S6 staining

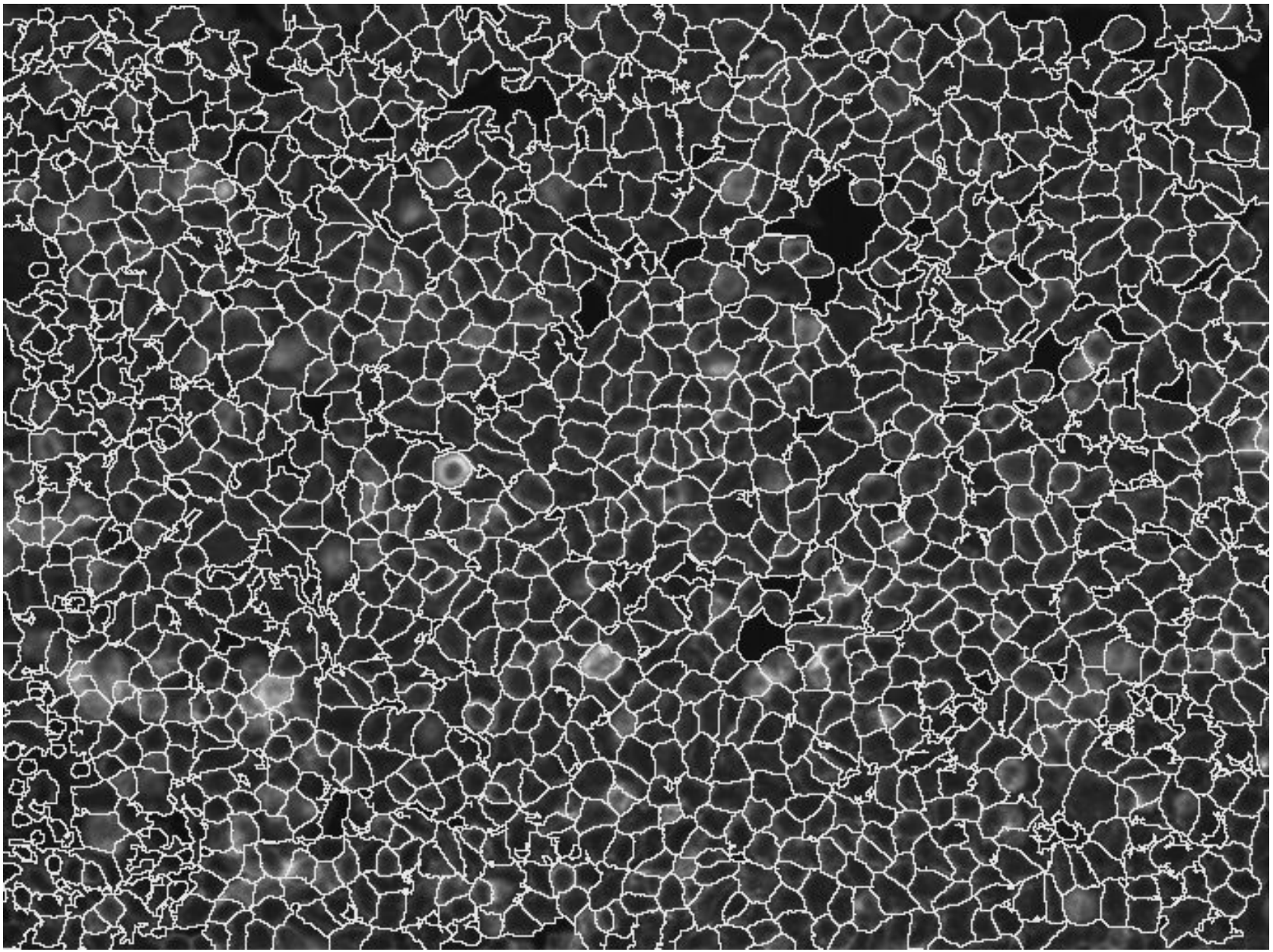
HeLa



Application to images from a chemical screen (ICCB/OME)

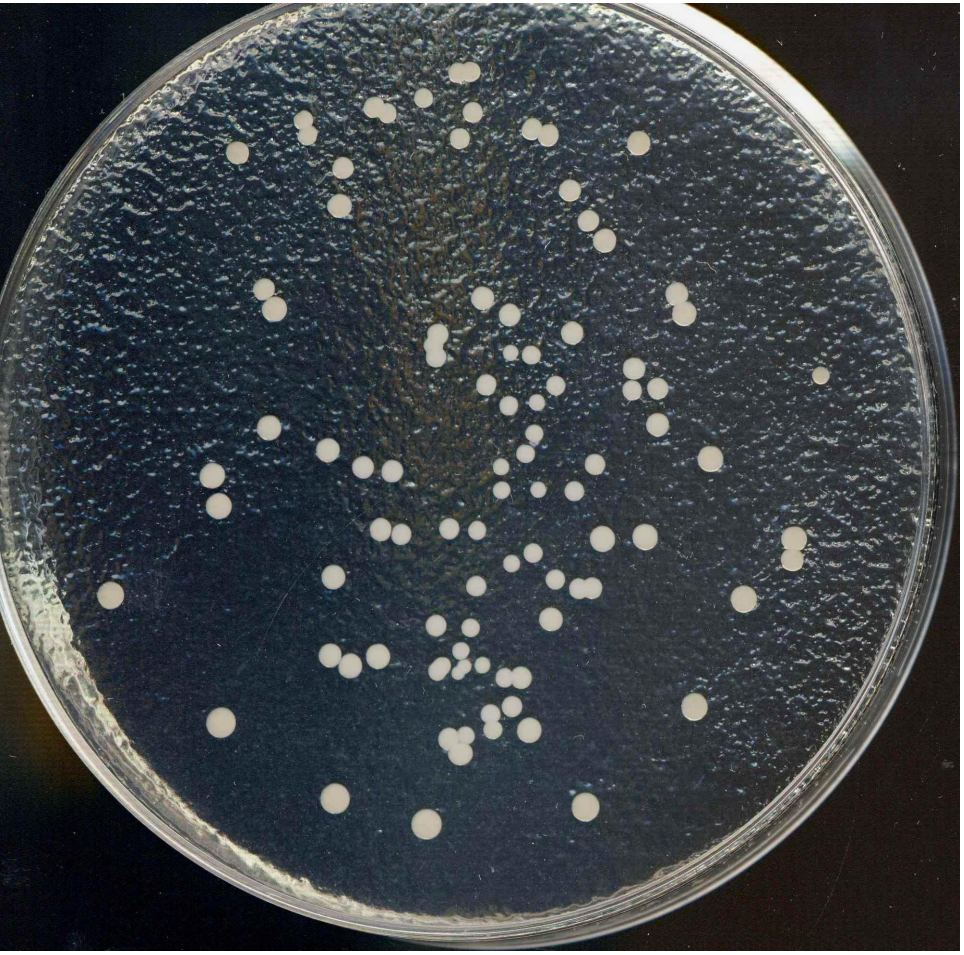


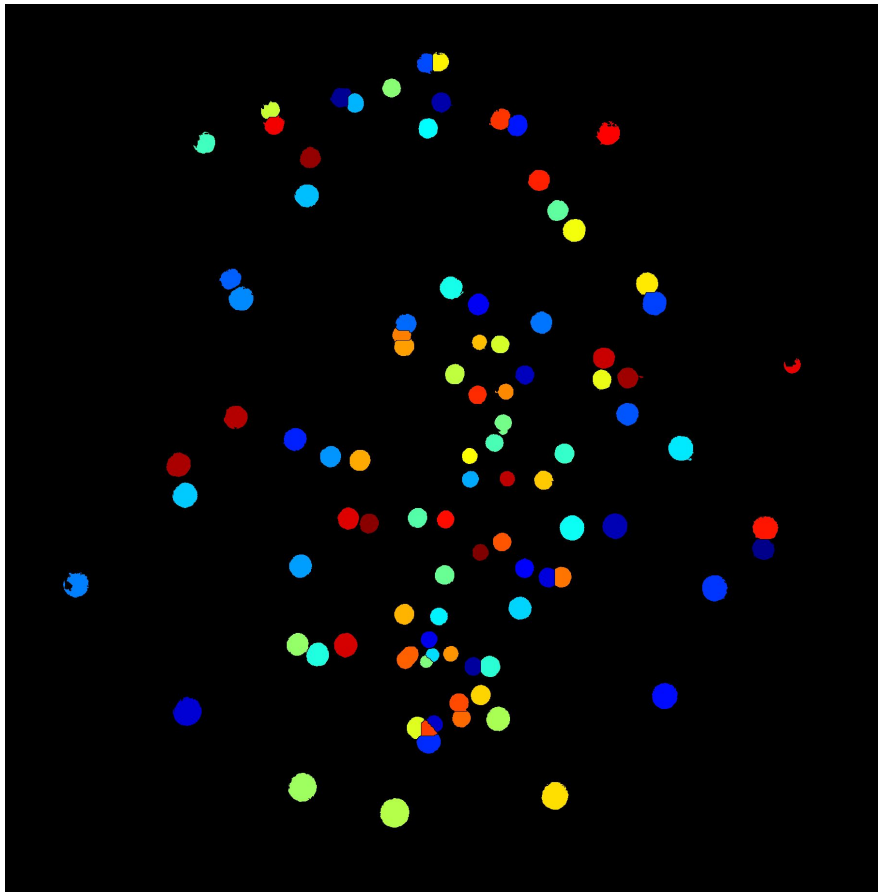
Application to images from a chemical screen (ICCB/OME)



Other applications:

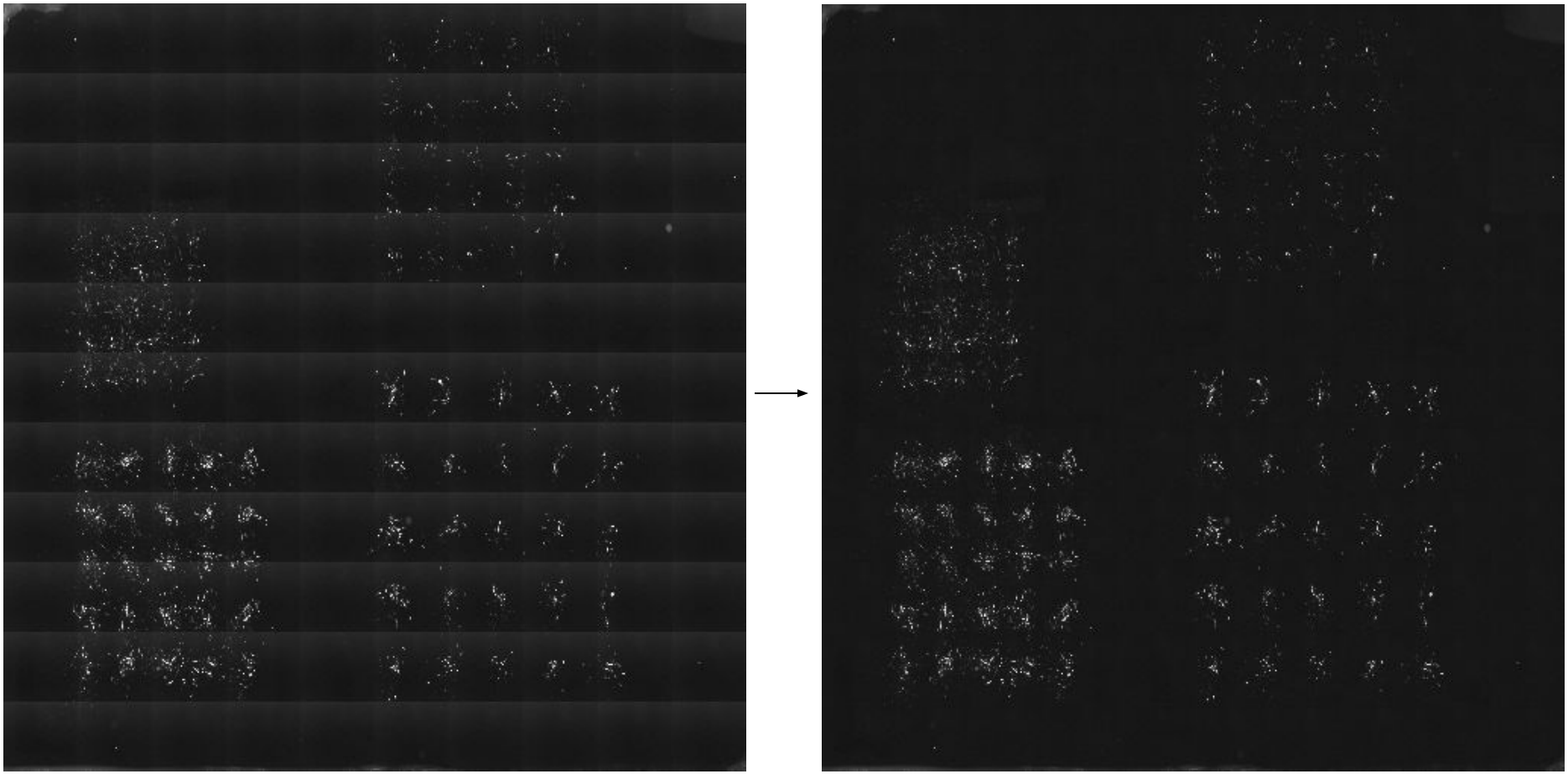
Yeast colony counting



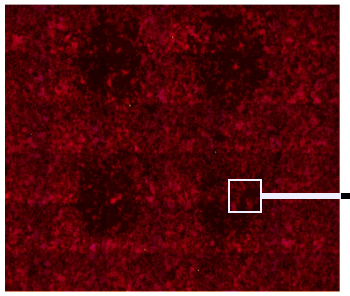


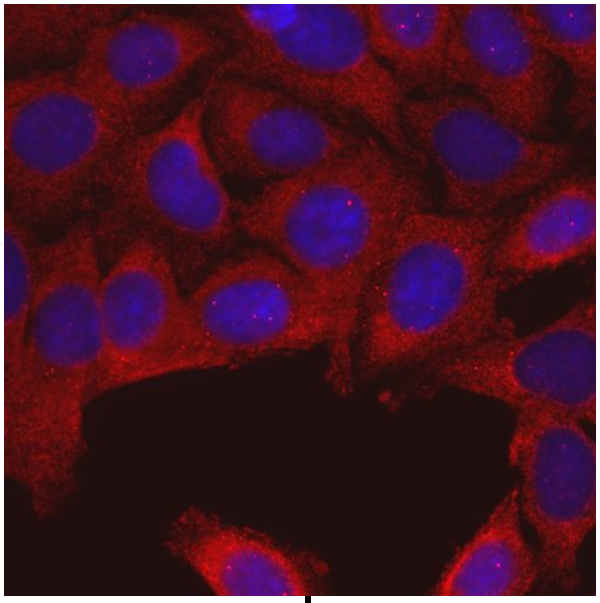
Other applications:

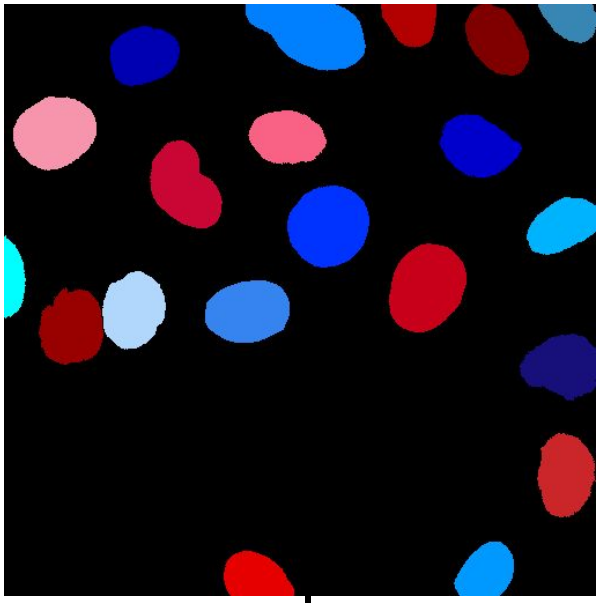
Illumination correction and image tiling

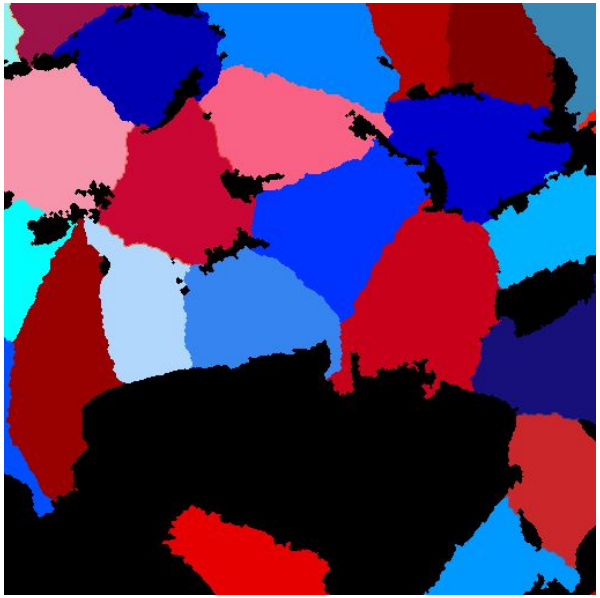


Multiple parameters are measured for every cell









cell #	1	2 …
Cell area	40.1	42.5
Cell perimeter	35.3	38.9
Cell aspect ratio	1.56	2.01
Cell staining intensity	4510	4939
Cell staining texture	16.8	17.2

…

\+ ~15 other measurements for cells

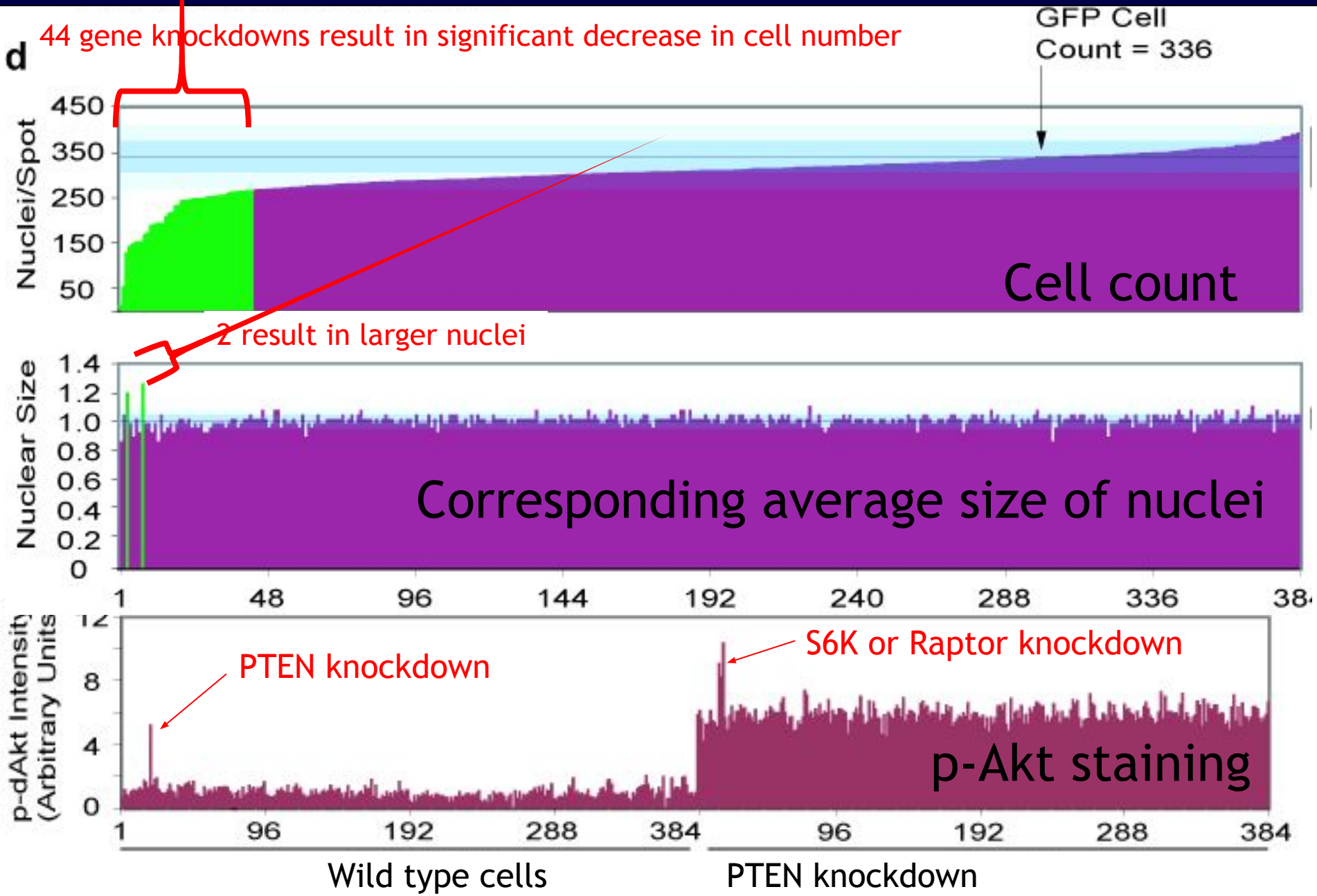
\+ ~20 measurements for nuclei

Phenotypes of interest in first screen

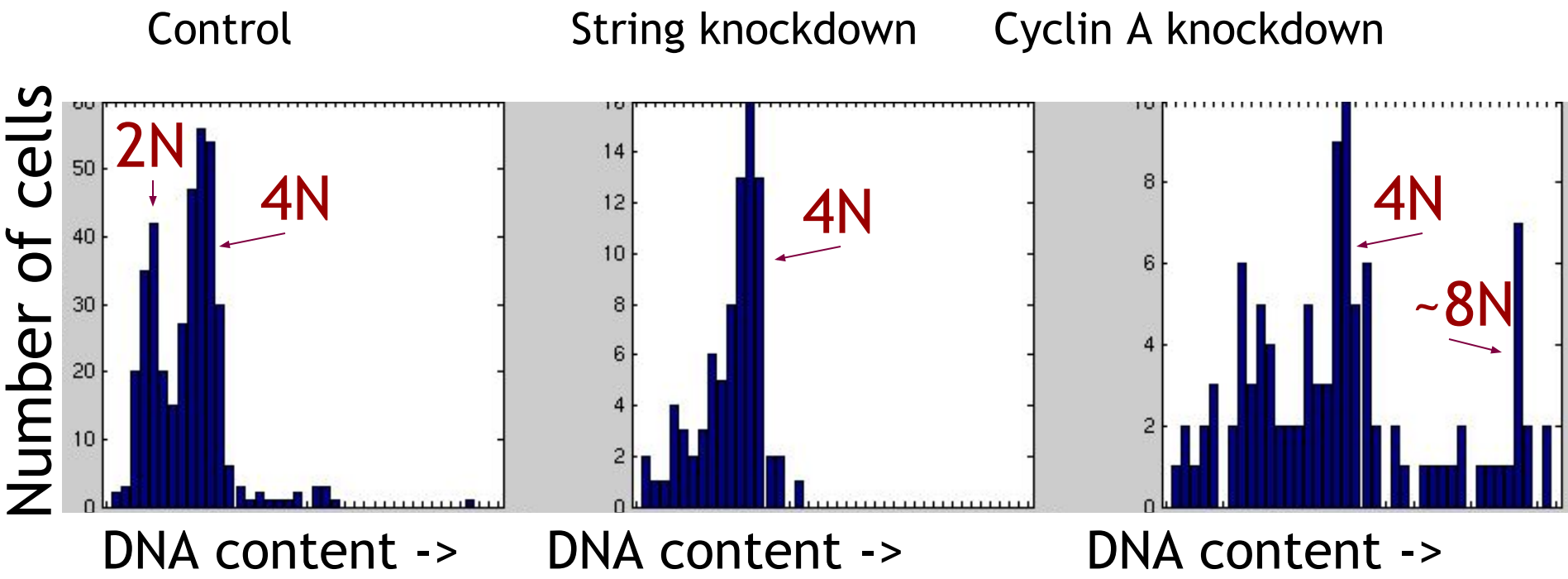
DNA staining: cell count
cell cycle distribution
chromatin texture
nuclear size
nuclear shape
Actin staining: morphology
actin content
phospho-Akt: p-Akt amount
p-Akt localization

We can screen every gene in one experiment using four slides!

Results: Cell number and nuclear size

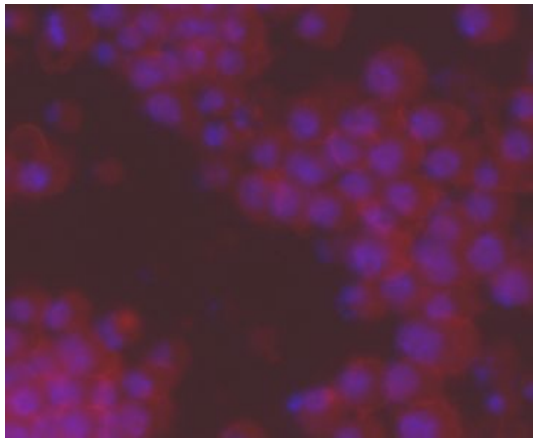


Results: Cell cycle



Future project design

Analyze >20 phenotypes by automated microscopy of genome-wide RNAi-treated living cell microarrays: identify required genes and decipher the signaling network



Cell size, count & morphology

Nuclear size & morphology

Nucleolar size, count & morphology

Cell cycle distribution (DNA content)

Upregulation of transcription

Apoptosis

Amount & distribution of signaling lipids

Upregulation of translation

Mitochondrial size & morphology

Localization of signaling proteins

Phosphorylation of signaling proteins (S6, S6K, Akt)

Actin content

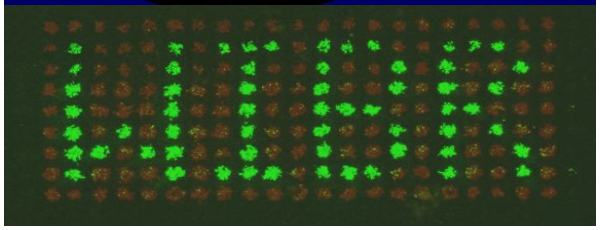
Protein content

Thanks to…

David M. Sabatini Laboratory

Siraj Ali
Alex Bagley
Steve Bailey
Anne Carpenter
Xana (Maria) Frias
Chris Gang
David Guertin
Kalyani Guntur
Thouis R. Jones
Jason Moffat
Kathleen Ottina
Dos Sarbassov
Yasemin Sancak
Shomit Sengupta
Joon-Ho Sheen
Carson Thoreen
* Doug Wheeler *

Whitehead Institute
for Biomedical Research



Lab Collaborators

Brent Stockwell - *Whitehead*

Lab Funding

Whitehead Institute, NIH, Mathers Foundation, DeCamp Foundation, Blum Foundation, Fidelity Foundation, Pew Charitable Trust

CellProfiler™
cell image analysis software
www.cellprofiler.org

Created by:
Anne E. Carpenter and Thouis R. Jones

In the laboratories of:
David M. Sabatini and Polina Golland

at:
the Whitehead Institute and MIT

with funding from:
•Sabatini Laboratory
•Merck/MIT Computational & Systems Biology Initiative postdoctoral fellowship
•Society for Biomolecular Screening Small Grant Award
•Novartis fellowship from the Life Sciences Research Foundation