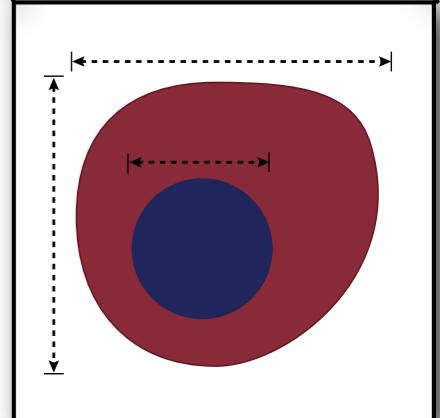


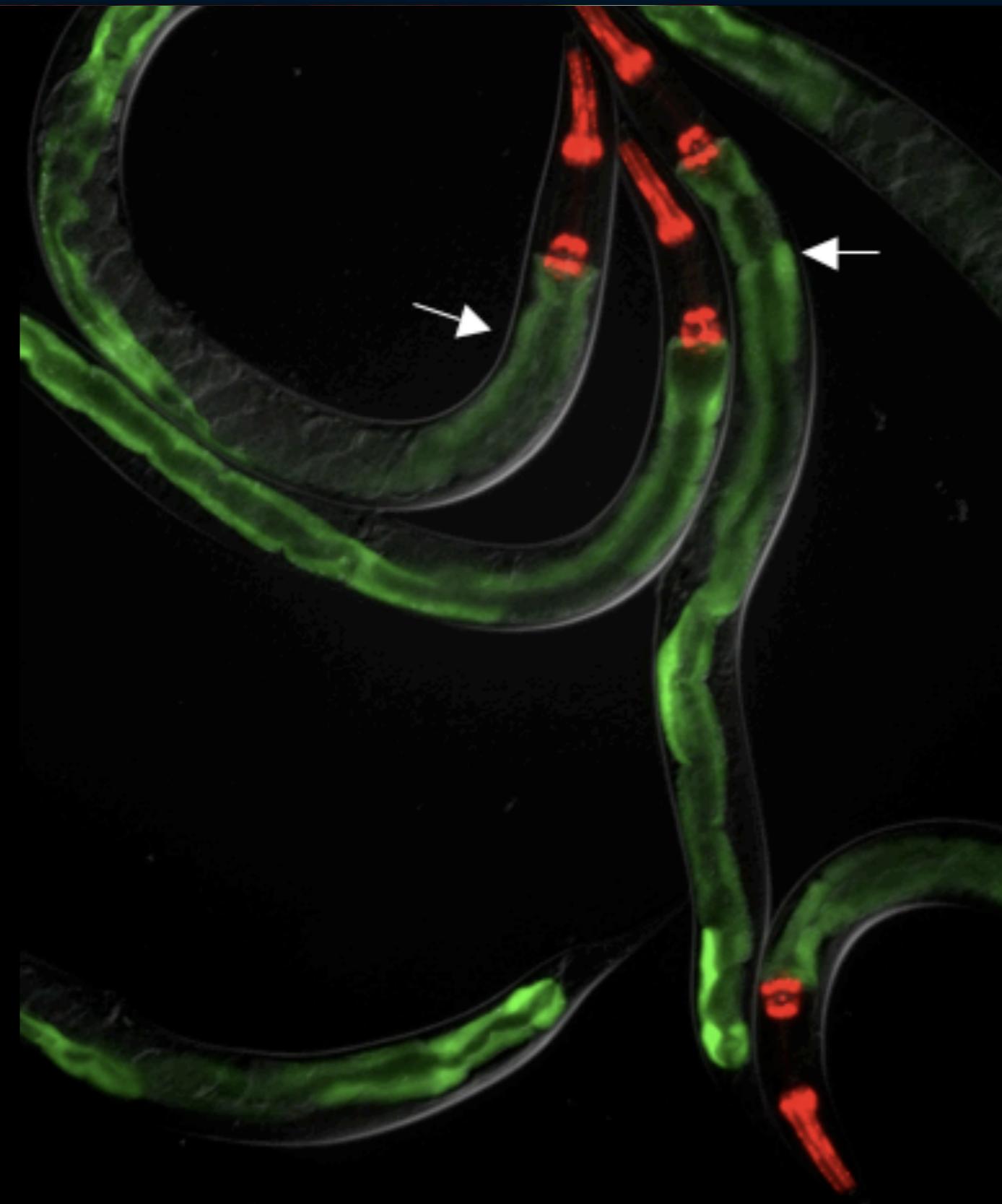
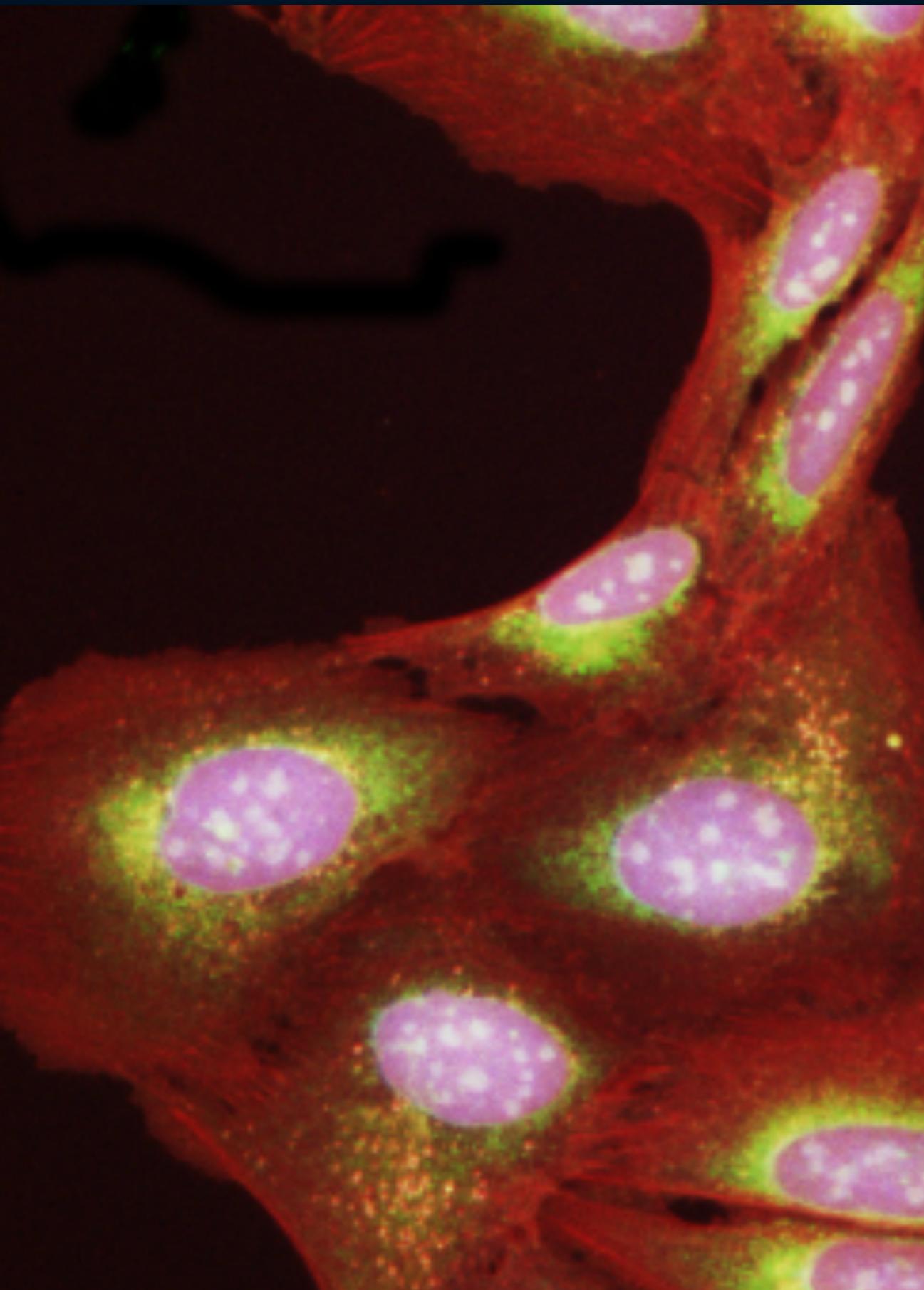
# Extracting rich information from biological images to tackle world health problems



0.4233  
54,454  
45.777  
0.6886  
0.0055  
6.9994  
83.333  
14.113  
1.5567  
0.0954  
0.5553  
...

Anne E. Carpenter, Ph.D.

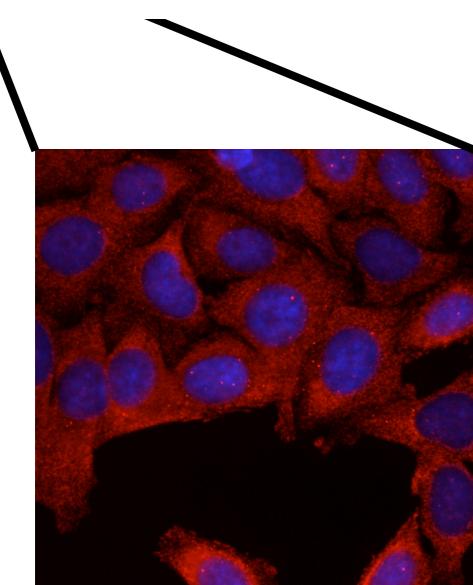
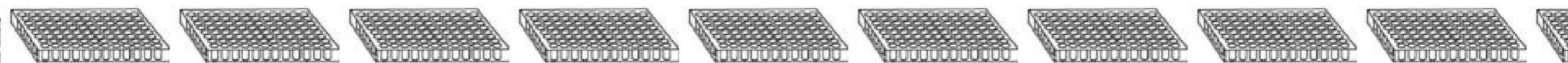
# Images contain a wealth of information



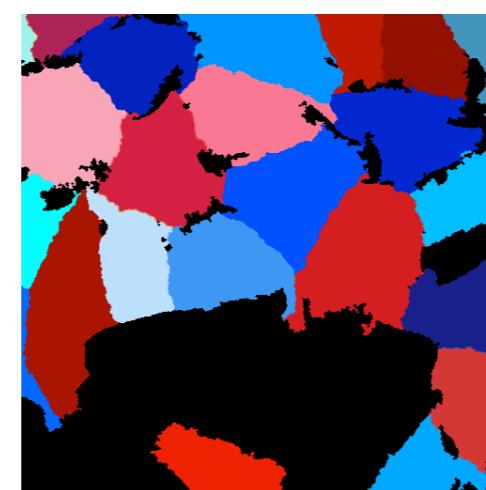
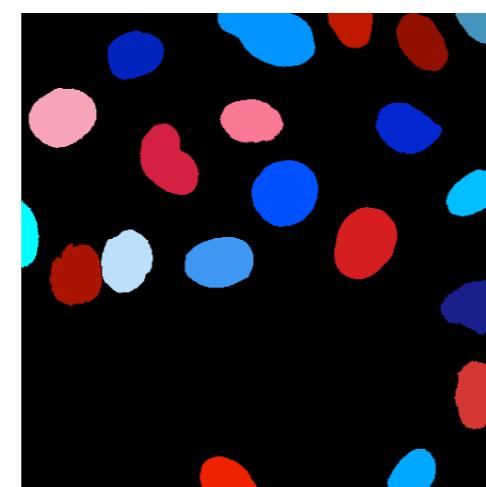
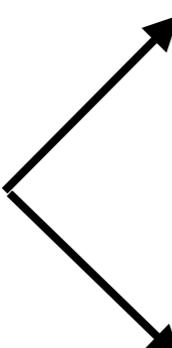
Images: Sigrun Gustafsdottir and Javier Irazoqui

# Screening to find genes and chemicals of interest

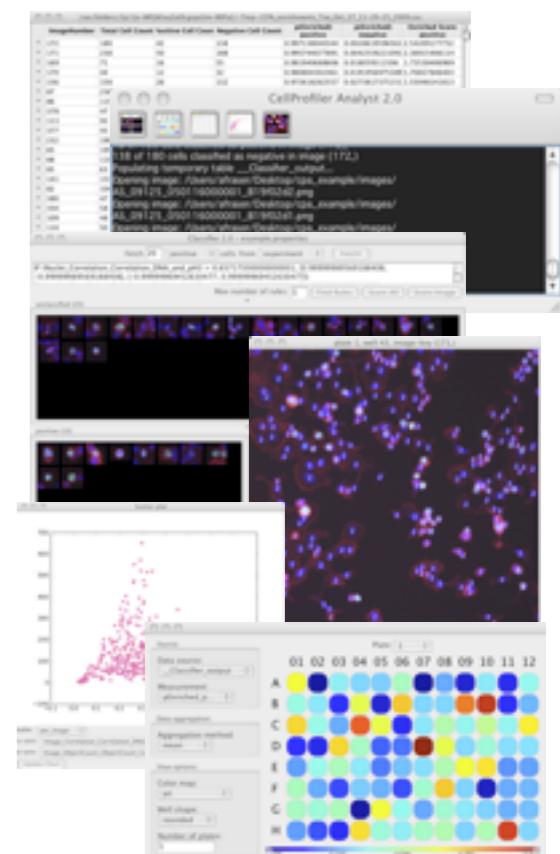
Cells or organisms in multiwell plates, each well treated with a gene or chemical perturbant



automated  
microscopy  
(any manufacturer)



- Cell measurements
- (size, shape, intensity, texture, etc.)



Data exploration  
& machine learning



**CellProfiler™**  
cell image analysis software



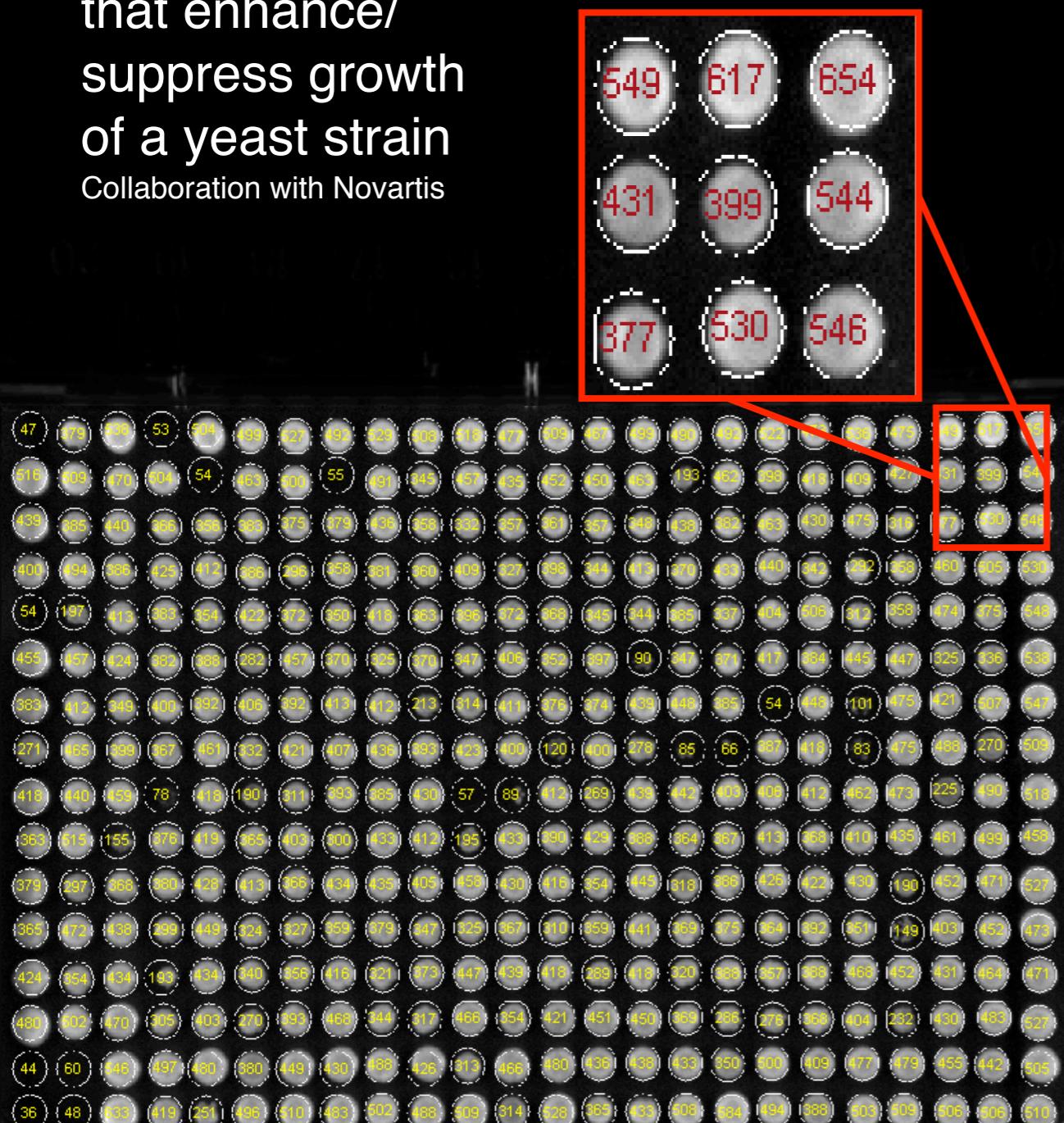
**CellProfiler™  
Analyst**  
data exploration software

# Yeast patch growth:

Goal: identify chemicals or genetic knockouts  
that enhance/  
suppress growth  
of a yeast strain



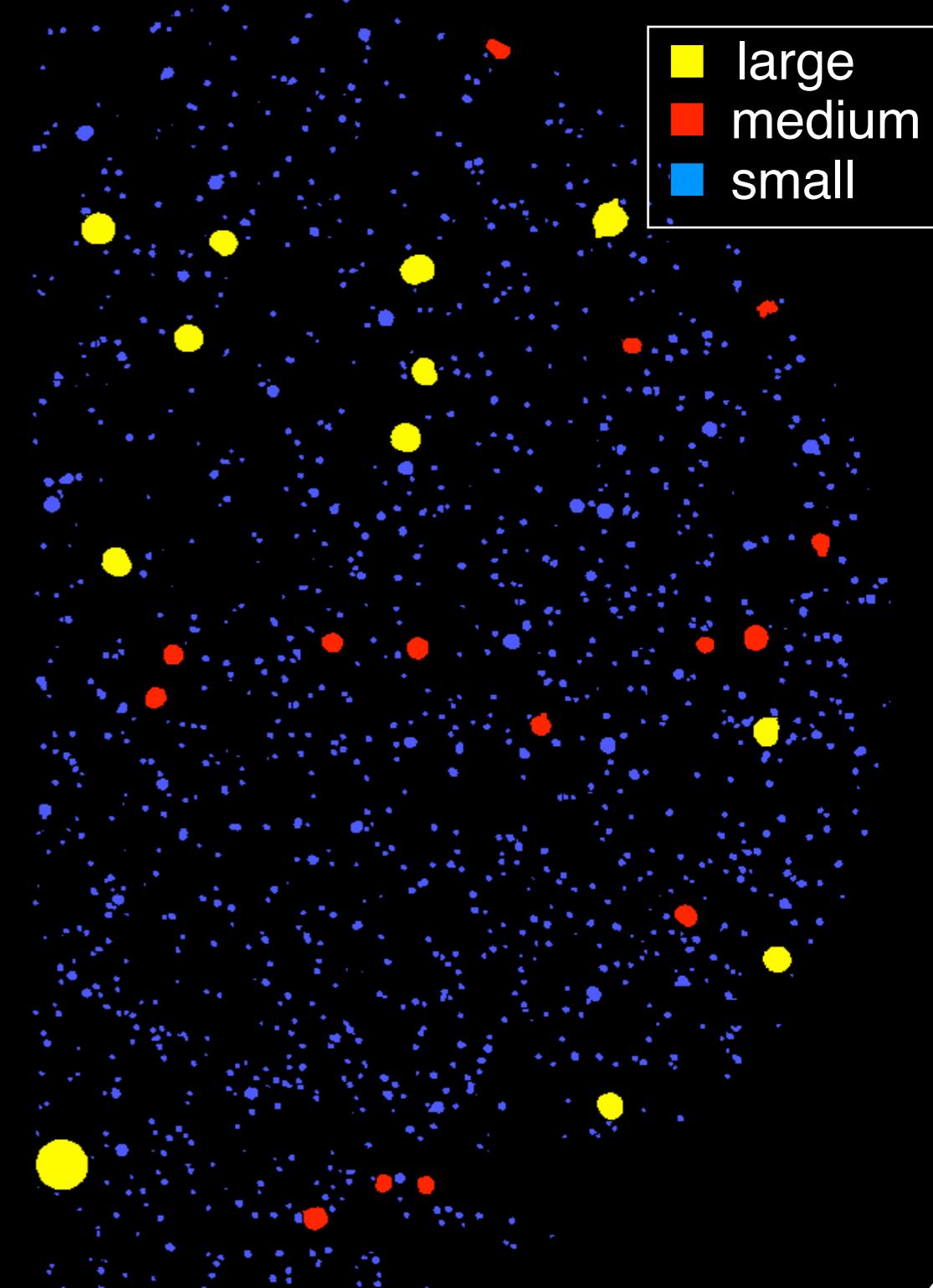
#### Collaboration with Novartis



# Yeast colony size:

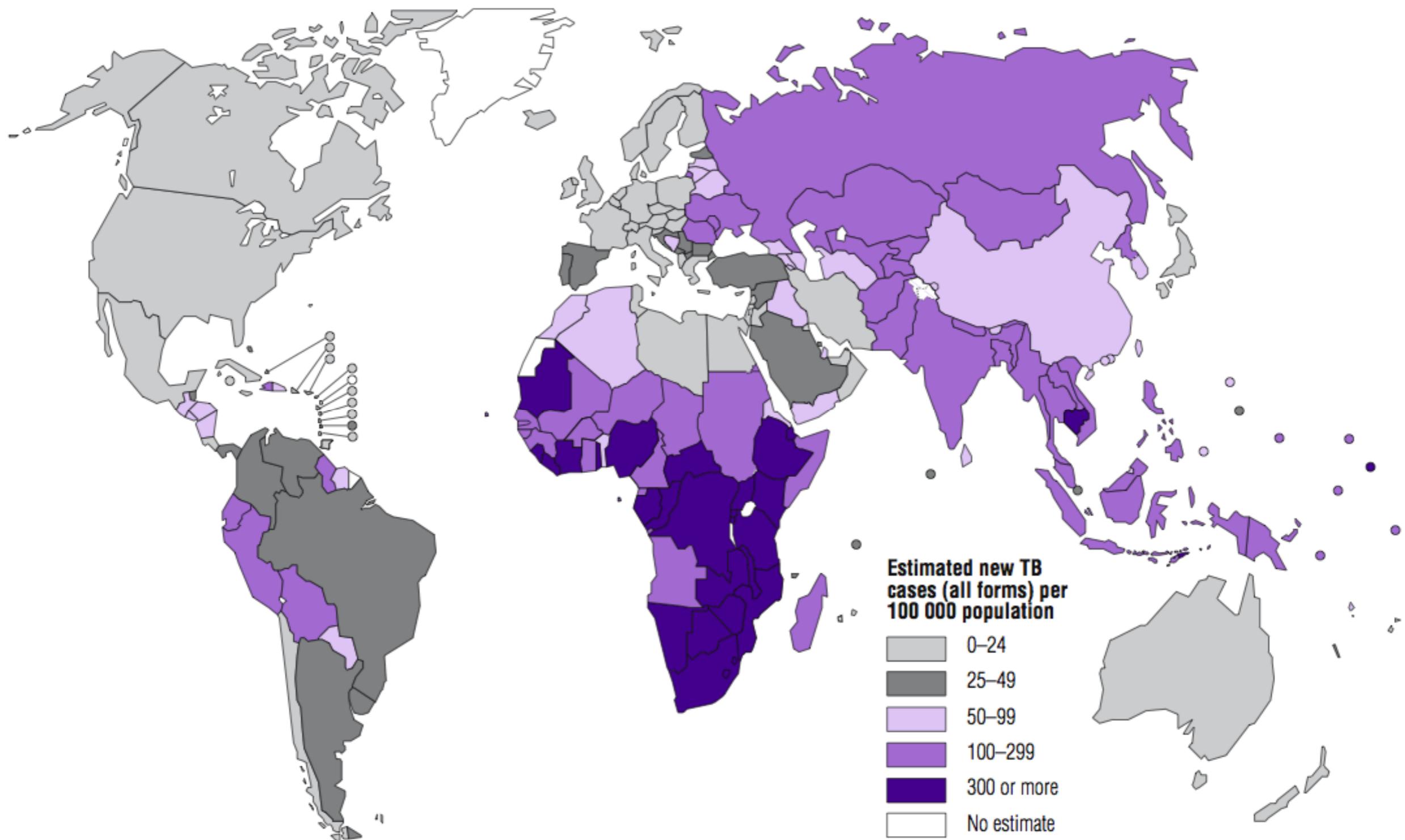
Goal: to understand pathways leading to drug-resistant yeast

Cowen, et al., *Eukaryotic Cell*, 2006



# Case study: Tuberculosis

Estimated TB incidence rates, by country, 2006

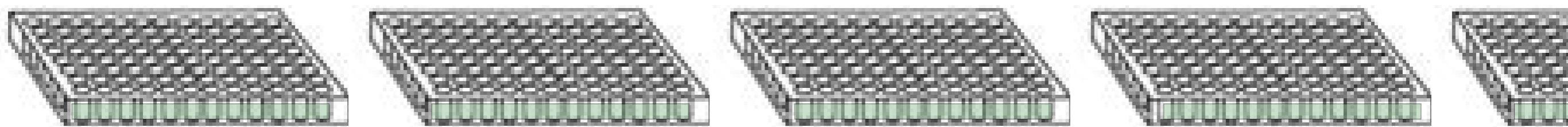


9.2 million new cases of tuberculosis in 2006  
1.7 million deaths in 2006

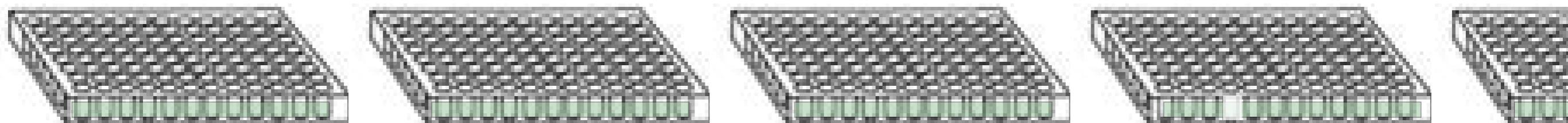
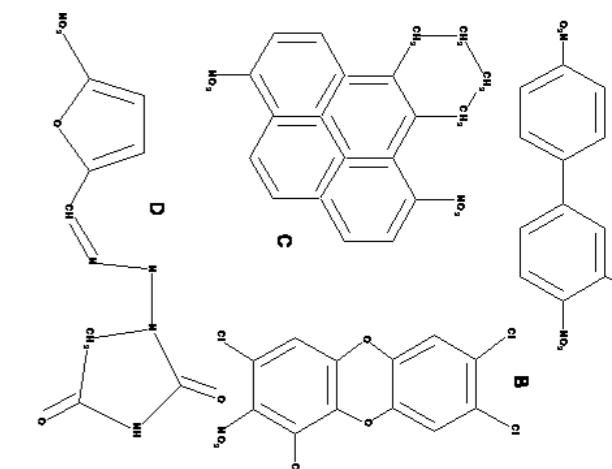
WHO Report, Global Tuberculosis Control 2008

# Traditional approach to find antibiotics

Try to kill **bacteria** in individual wells of multi-well plates

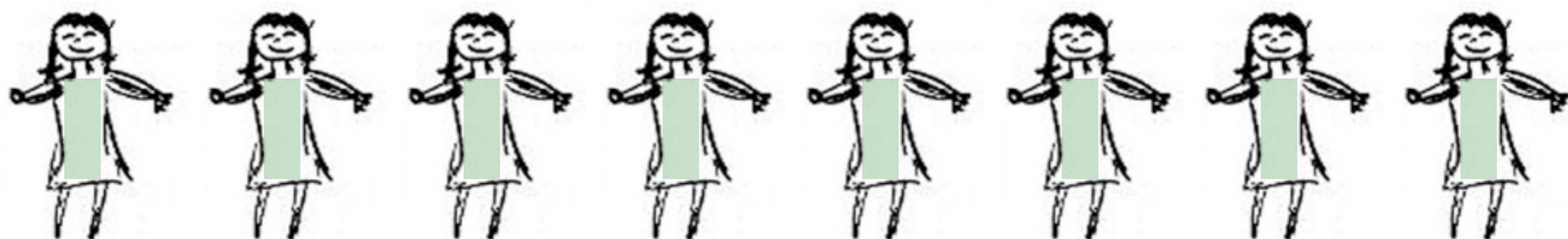


Add 1,000,000 test chemicals,  
each chemical in a different well

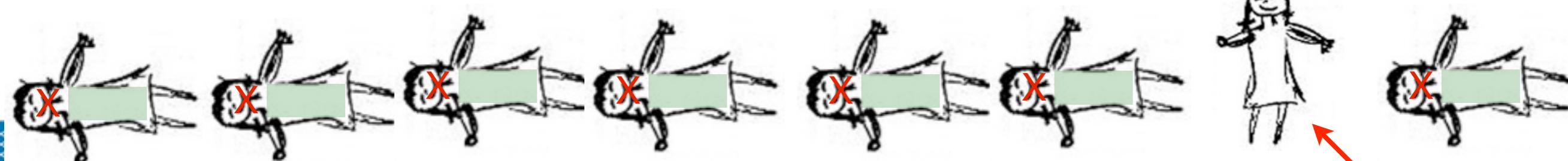
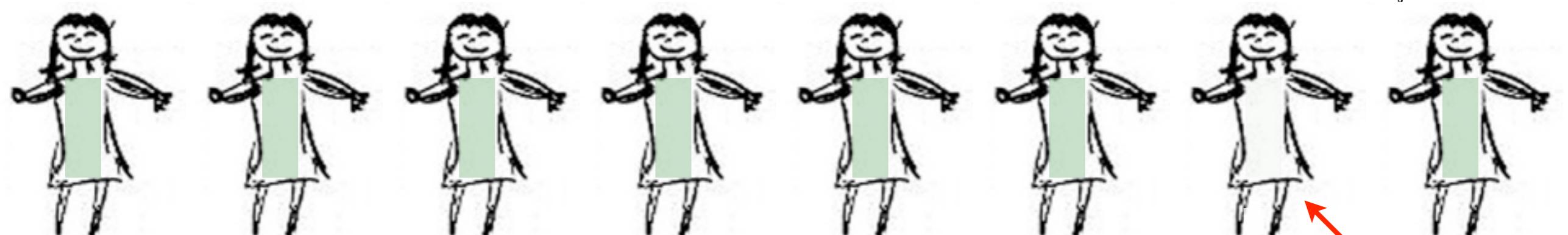
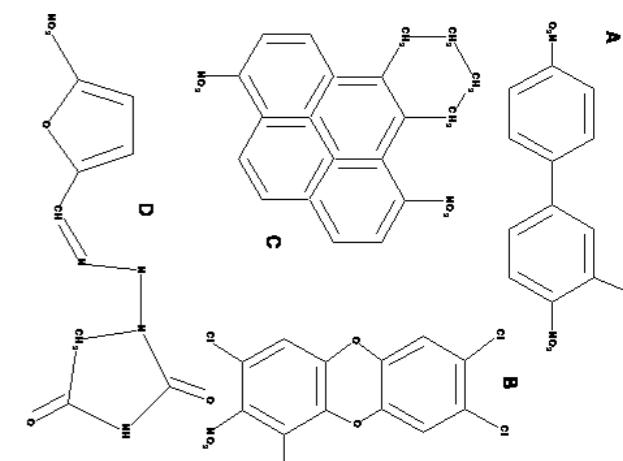


Measure amount of bacteria (fluorescence plate reader)  
and look for wells where bacteria died

# Alternate approach to find antibiotics (effective but non-ideal)

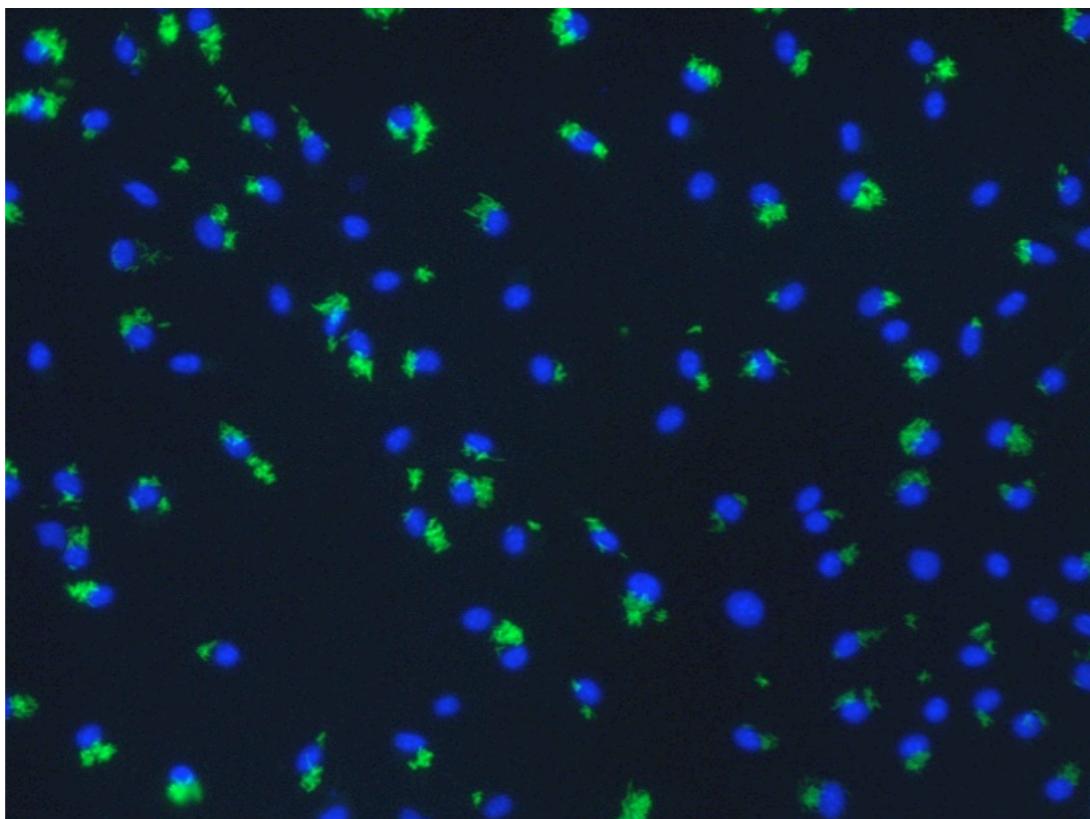


Add 1,000,000 test chemicals,  
each chemical in a different person



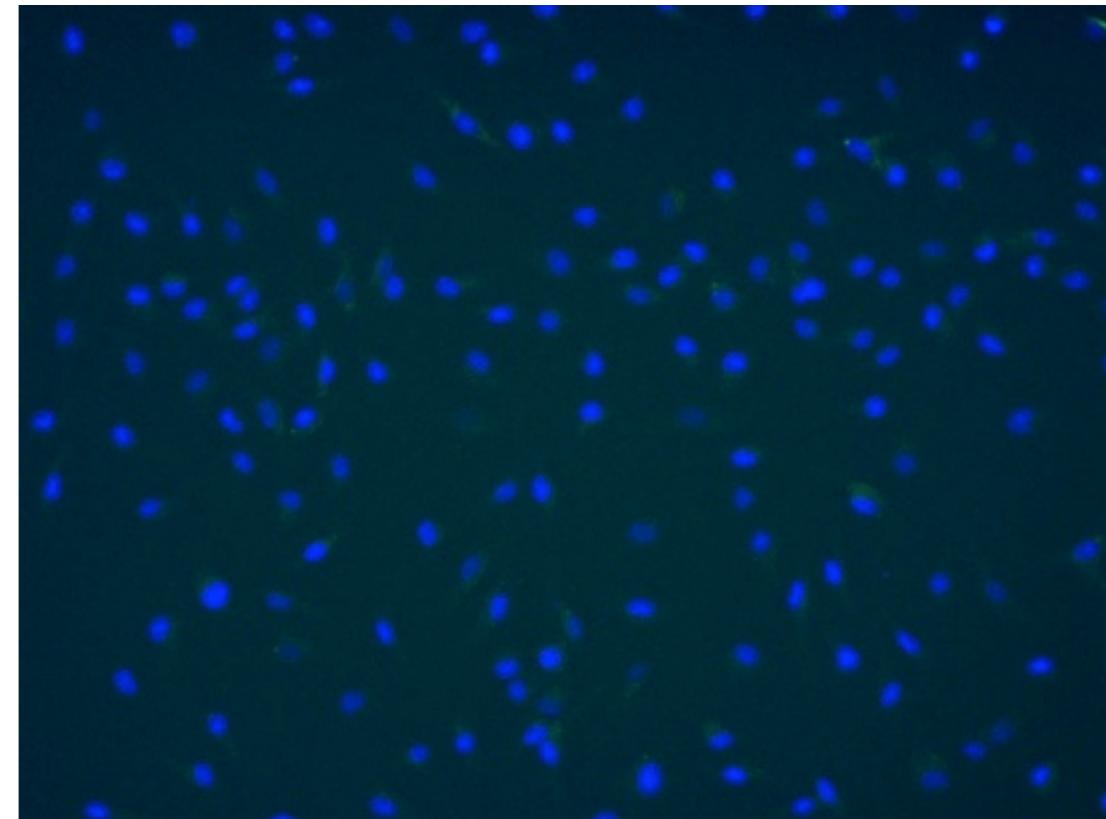
# Search for tuberculosis treatments

Without drug



mouse  
nuclei

With drug



tuberculosis  
bacteria



Martha  
Vokes



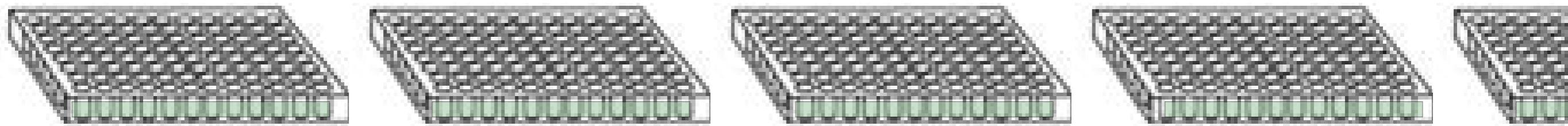
Mark  
Bray



Deb Hung,  
Broad/MGH    Sarah  
Stanley,  
postdoc    Amy  
Barczak,  
postdoc 8

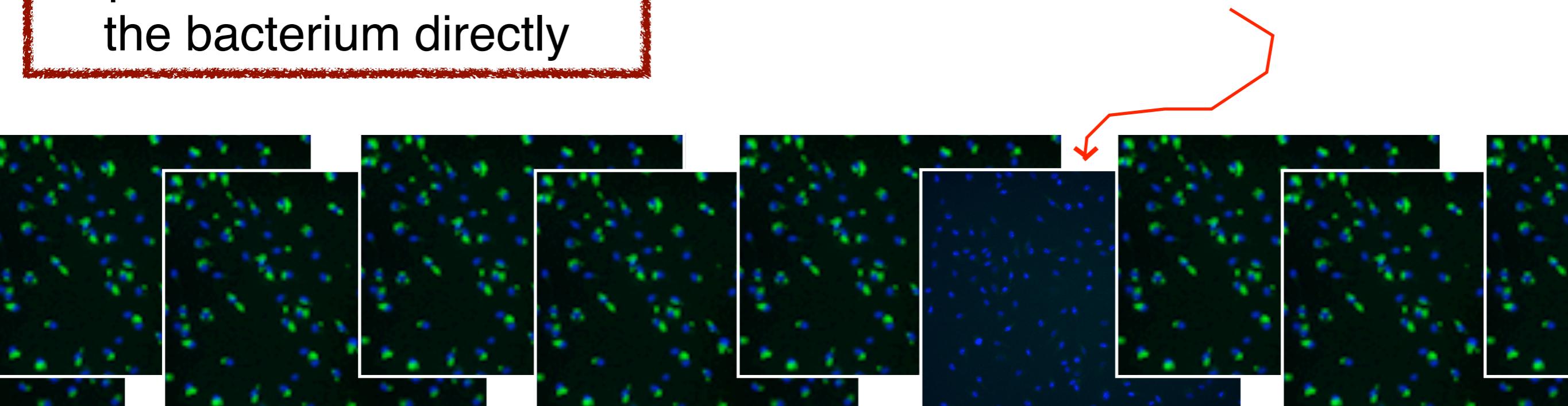
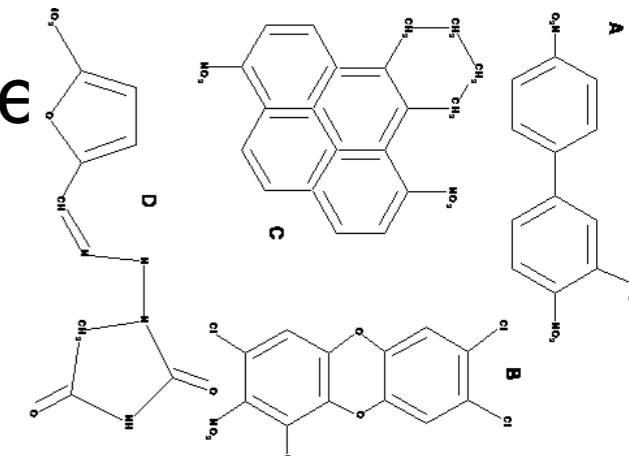
# Search for tuberculosis treatments

Put **bacteria** and **mouse cells** in individual wells of multi-well plates



Status: pursuing hits that prevent bacterial infection/expansion but do NOT kill the bacterium directly

Add 10,000 bioactive chemicals, each chemical in a different well

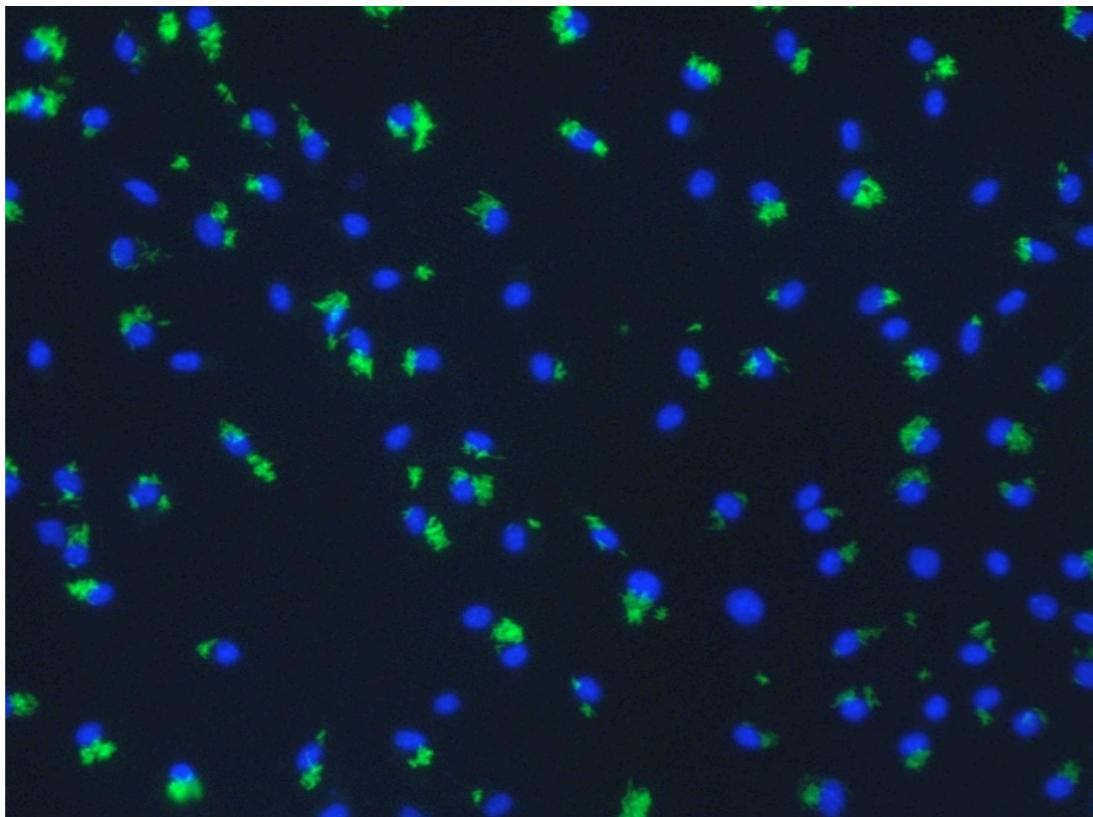


# Search for tuberculosis treatments

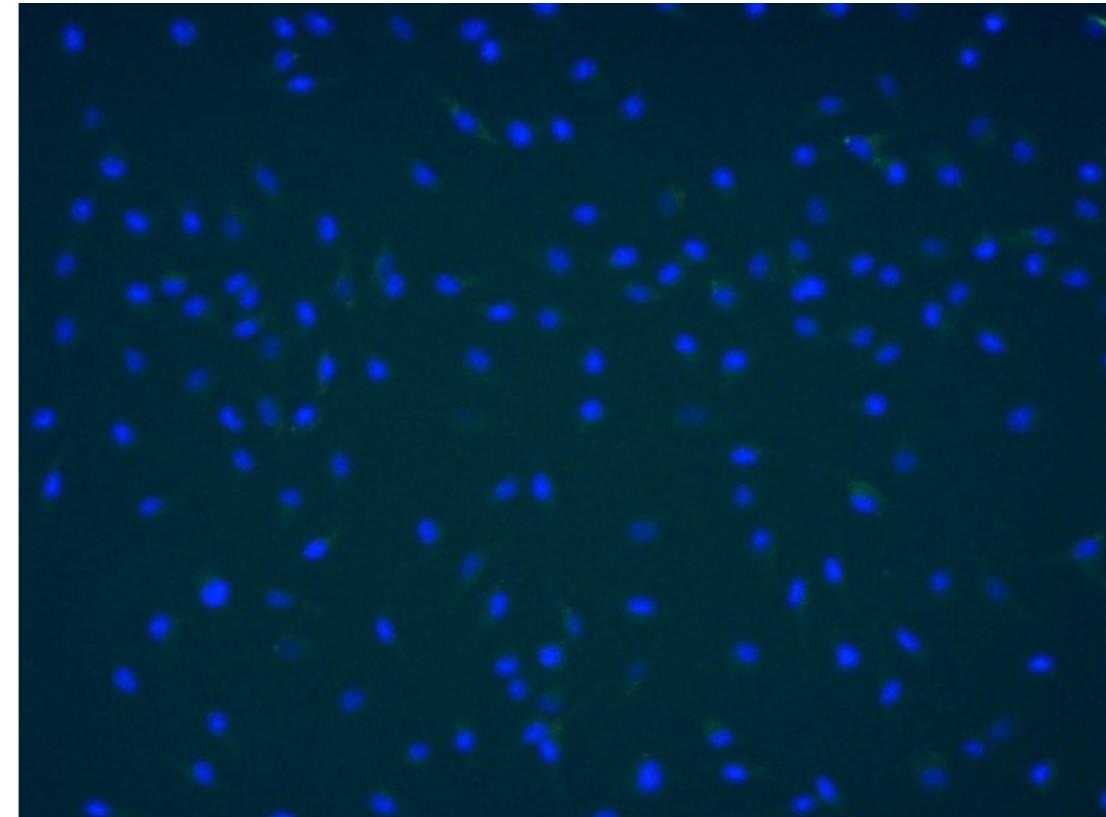
mouse  
nuclei

tuberculosis  
bacteria

Without drug



With drug



Status: pursuing hits that  
prevent bacterial infection/  
expansion but do NOT kill  
the bacterium directly



Martha  
Vokes



Mark  
Bray



Deb Hung,  
Broad/MGH



Sarah  
Stanley,  
postdoc

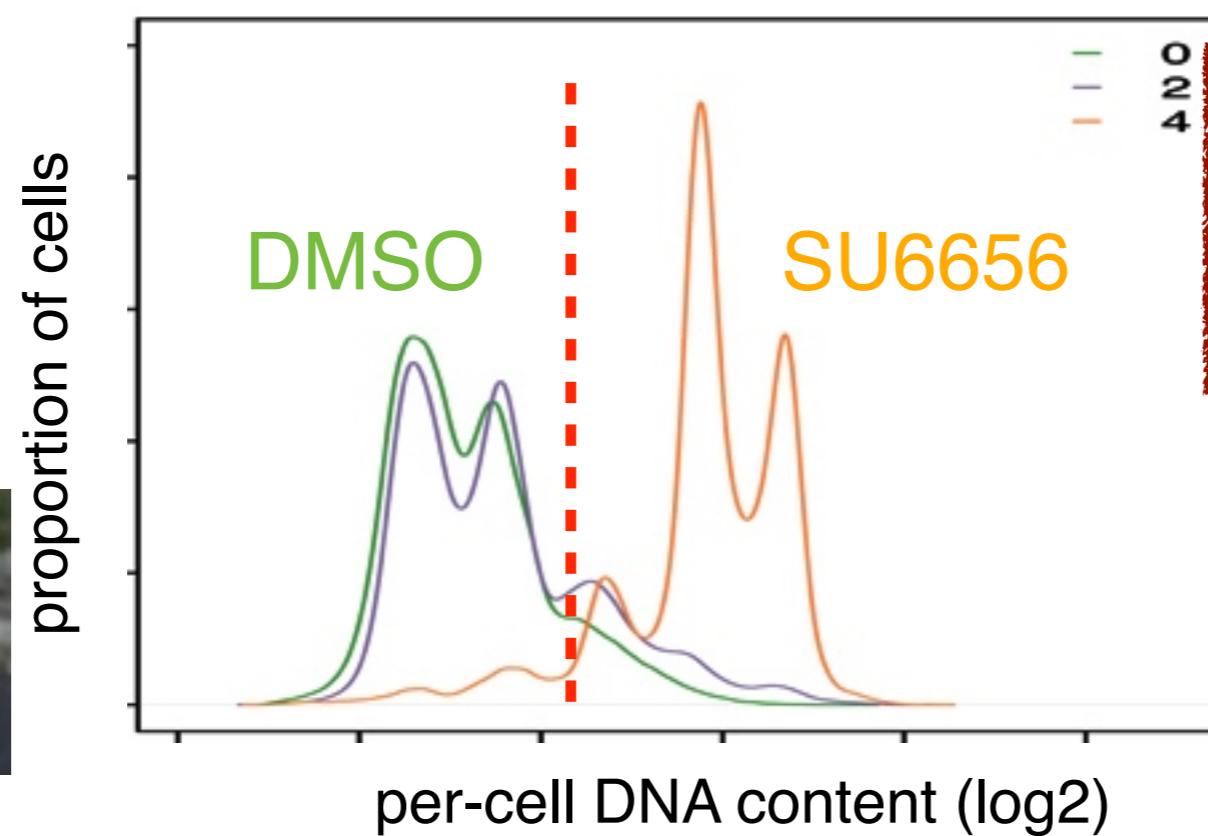
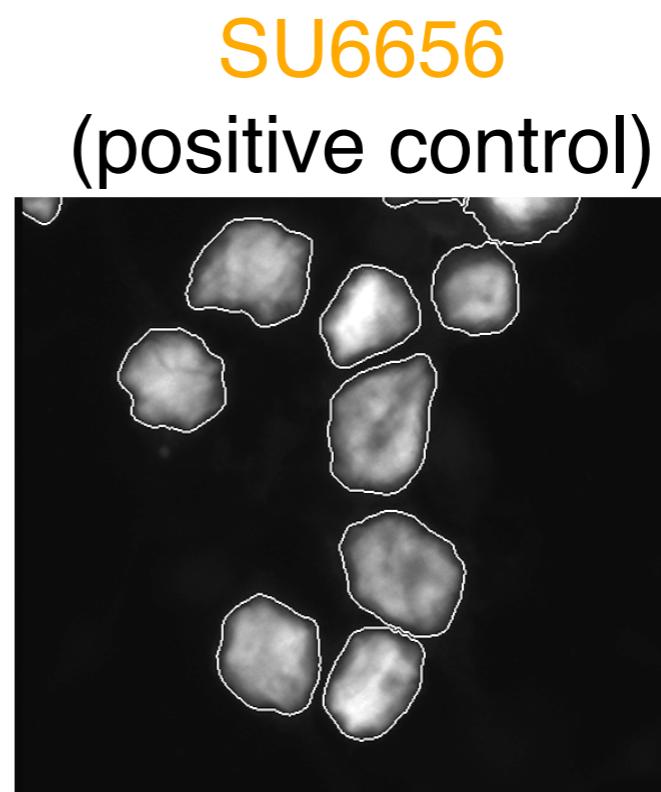
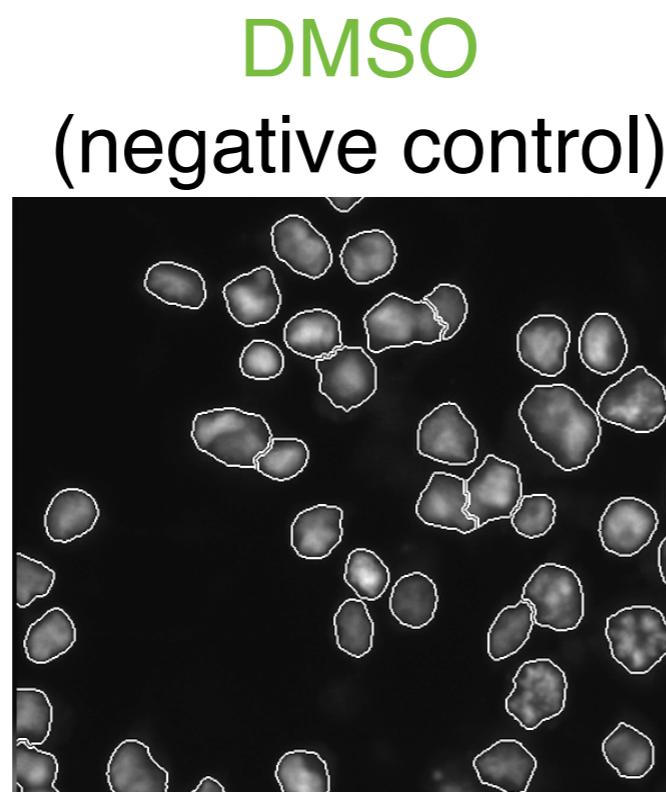


Amy  
Barczak,  
postdoc 10

*project in progress*

# Polyplloidization of megakaryocytes - AMKL (leukemia)

DNA stain,  
with  
outlines  
identifying  
the nuclei



Status: recommended  
clinical testing of AURKA  
inhibitor for AMKL



Martha  
Vokes



Mark  
Bray



John Crispino,  
Northwestern  
University



Jeremy  
Wen,  
postdoc

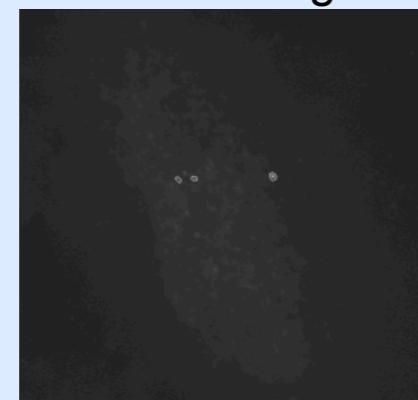
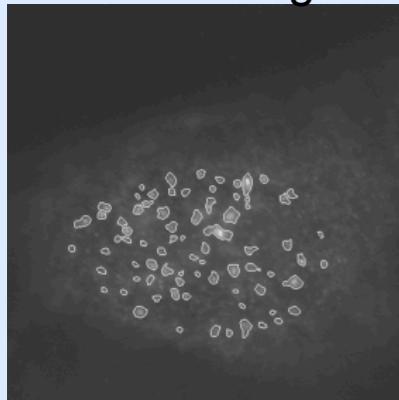
Wen, ..., Carpenter ... Crispino, et al. Cell 2012

## Cancer radiation treatment

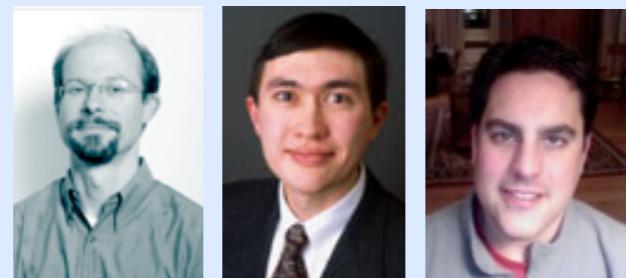
response to  
DNA damage

recovery from  
DNA damage

p-H2AX  
foci

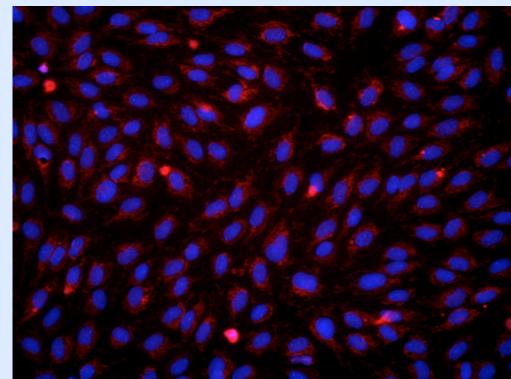


Mike Yaffe's lab at MIT's Center  
for Cancer Research (Scott Floyd  
& Michael Pacold, postdocs)  
& The RNAi platform/TRC at the  
Broad Institute

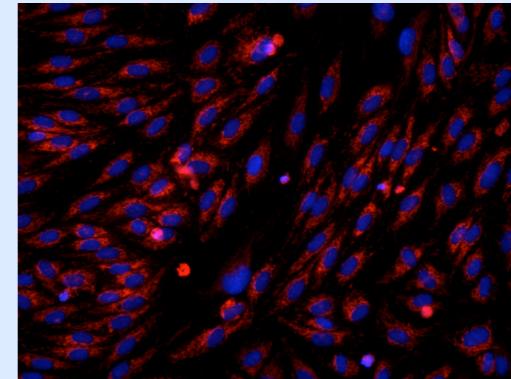


## Mitochondrial abundance

Negative control



Positive control



DNA

Mito-  
tracker

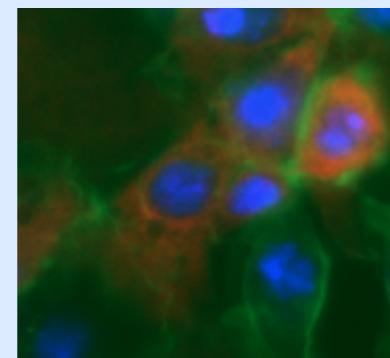
Vamsi Mootha's lab at Harvard  
Medical School/Mass General  
Hospital (Toshi Kitami, postdoc)



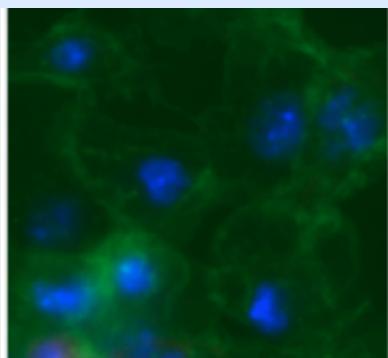
*Kitami ... Carpenter ... Mootha, PLoS ONE, 2012*

## S6K signalling

Control



dTOR RNAi



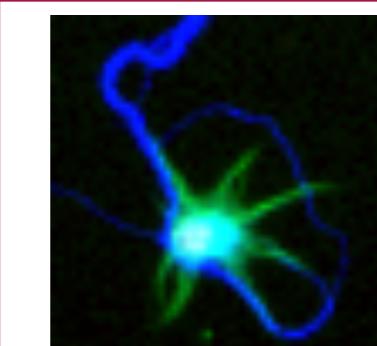
blue = DNA  
green = actin  
red = p-S6

Anne Carpenter & David Sabatini at  
the Whitehead Institute for  
Biomedical Research (Rob  
Lindquist, student)

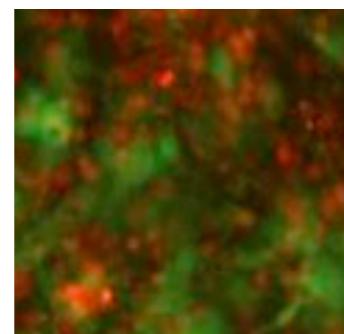


*Lindquist ... Carpenter, Genome Research, 2011*

# Screens to identify genes & chemicals: increasing physiological relevance



Cell-based assays  
with complex  
morphologies



Co-cultured  
mixtures of cell  
types



Whole organisms

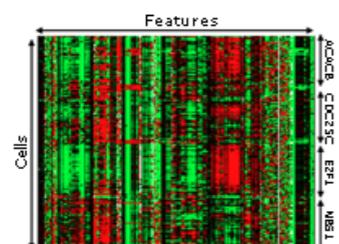
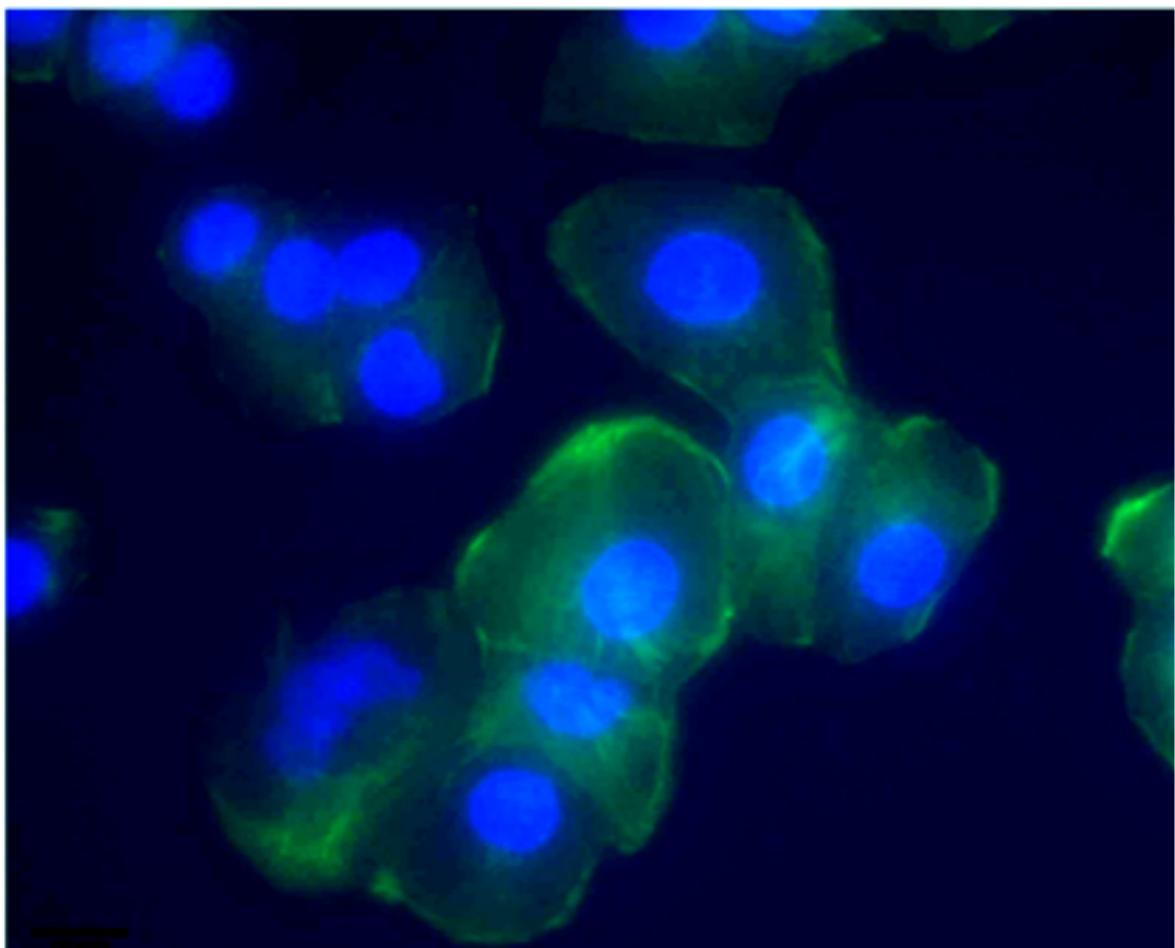


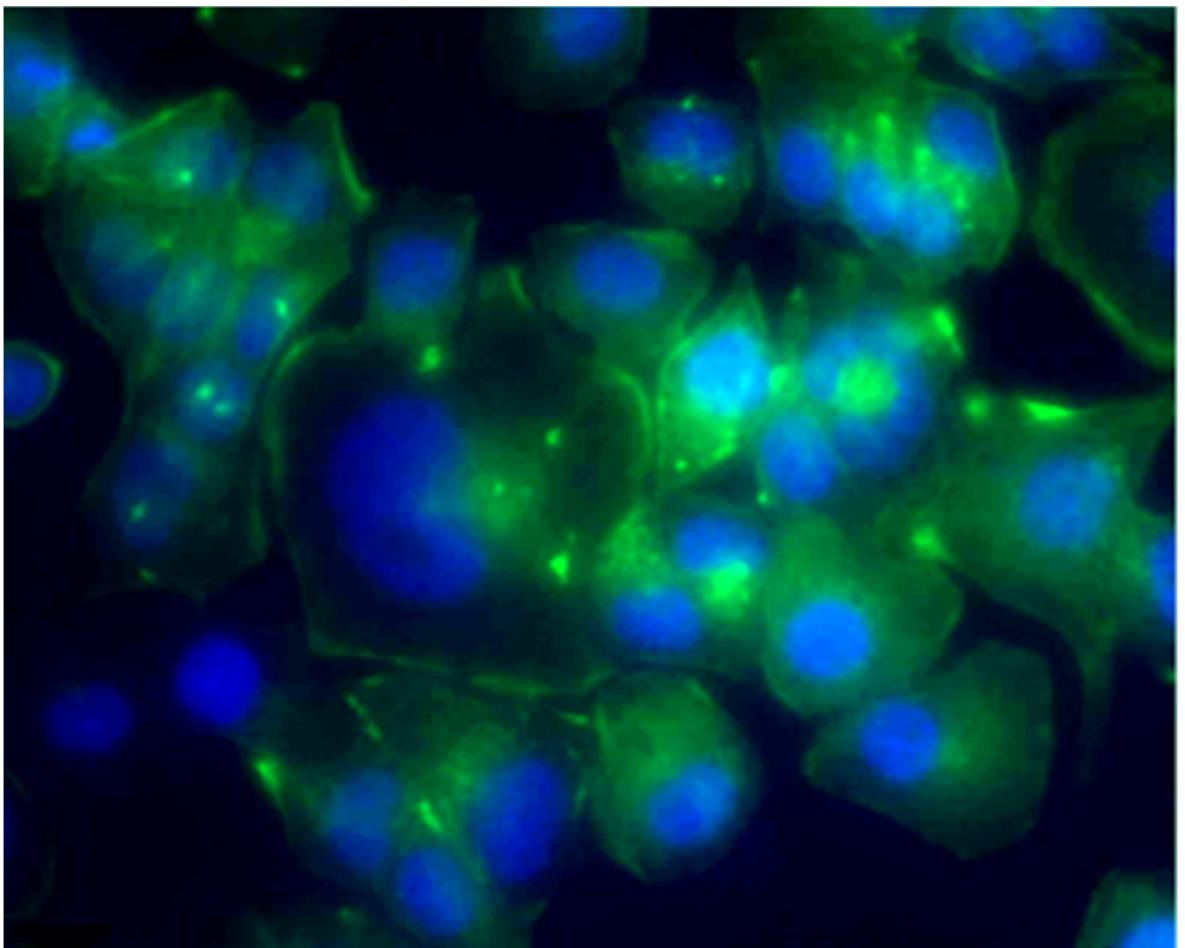
Image-based profiling: extracting ‘signatures’

# Breast cancer

Control



+ Growth factor



DNA

Actin



Ray  
Jones



Anne  
Carpenter



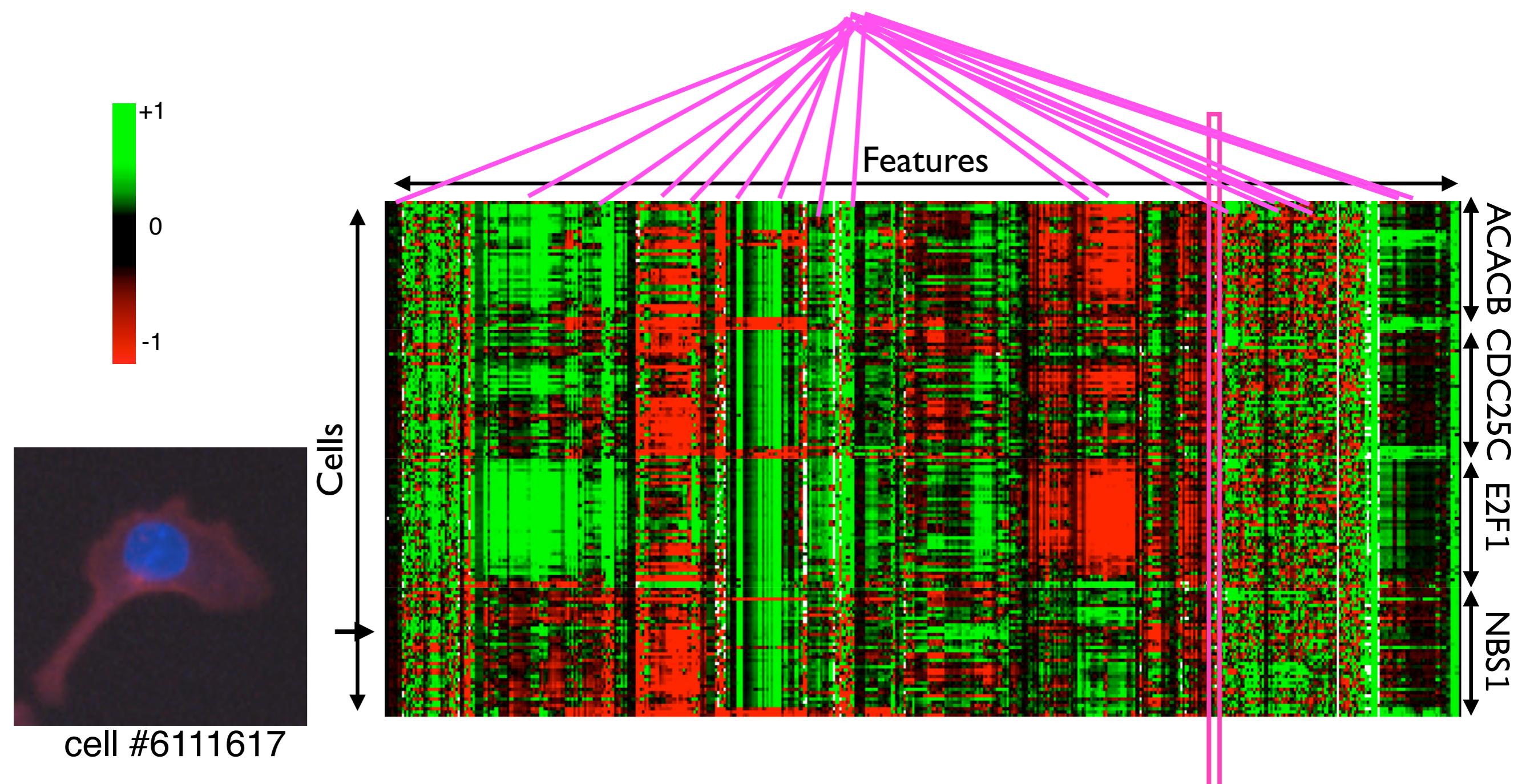
Eric Lander,  
Broad  
Institute



Piyush Gupta,  
postdoc

*project in progress*

# MEASURE EVERYTHING...ASK QUESTIONS LATER.



“Cytological profile”: collection of measurements describing the appearance of a cell

Perlman, et al. Science 2004



# User-friendly machine learning

Classifier 2.0 – /Users/afraser/CPA/properties/nirht\_area\_test.properties

Fetch cells

Fetch 25 mitotic cells from Gene gene: NME1 Fetch!

Train Classifier

```
IF(CYTOPLASM_AREASHAPE_SOLIDITY>0.5721039772033691, 0.12450471498804594, -0.5991868036084859)
IF(NUCLEI_TEXTURE_1_PH3_CORRELATION>0.7244470119476318, 0.2329807465866336, -0.3858751618560514)
IF(CELLS_CORRELATION_CORRELATION_DNA_AND_ACTIN>0.4226300120353699, -0.7973272282899164, 0.11330265089003502)
IF(NUCLEI_AREASHAPE_ECCENTRICITY>0.7200030088424683, 0.28075676668884736, -0.2779378462168646)
IF(NUCLEI_TEXTURE_3_PH3_INVERSEDIFFERENCEMOMENT>0.2568340003490448, -0.16926779984027135, 0.467381716198041)
```

unclassified (25)

positive      negative

fetching 25 mitotic cells from group Gene: gene=NME1



Ray  
Jones

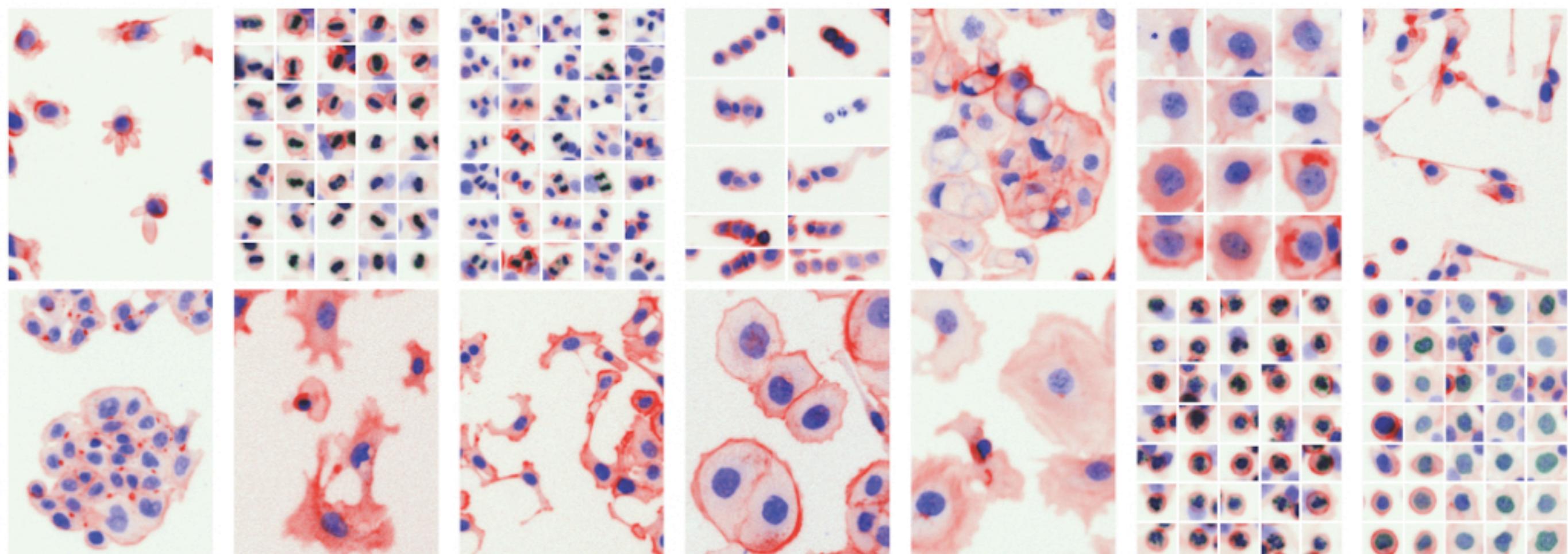


Adam  
Fraser

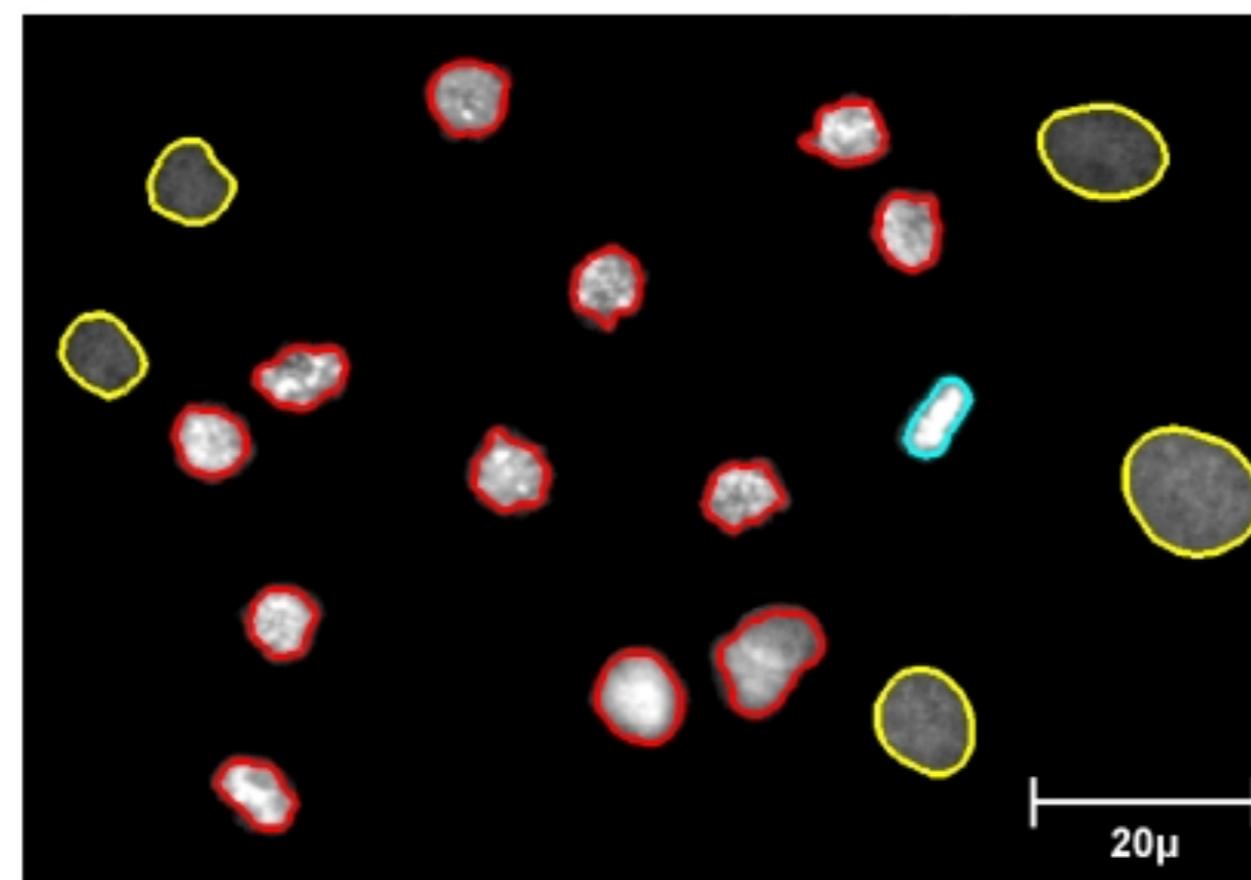
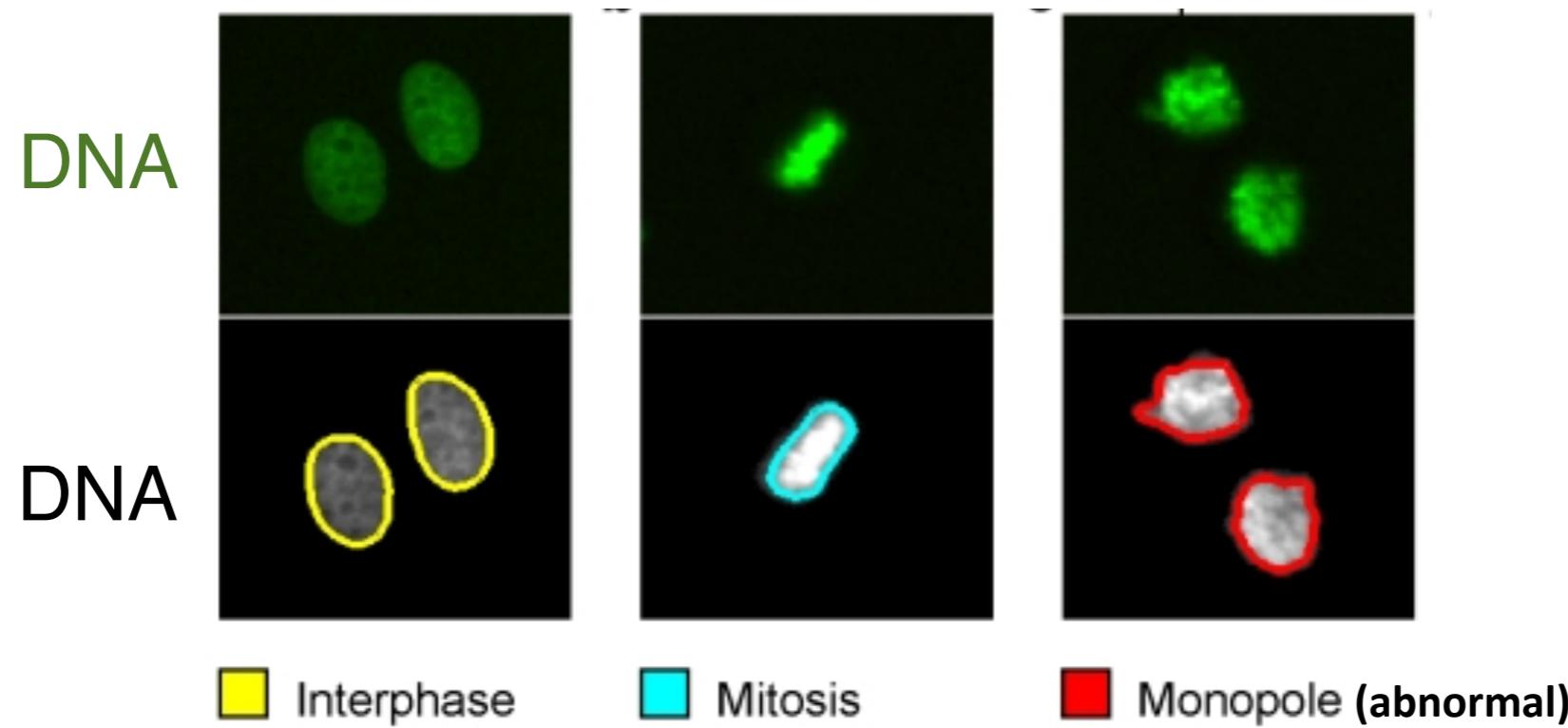


IMAGING  
PLATFORM

# Challenging cellular phenotypes



# Regulators of cell division



Tim Mitchison,  
Harvard Med.



Ray  
Jones



Martha  
Vokes



Tiao Xie

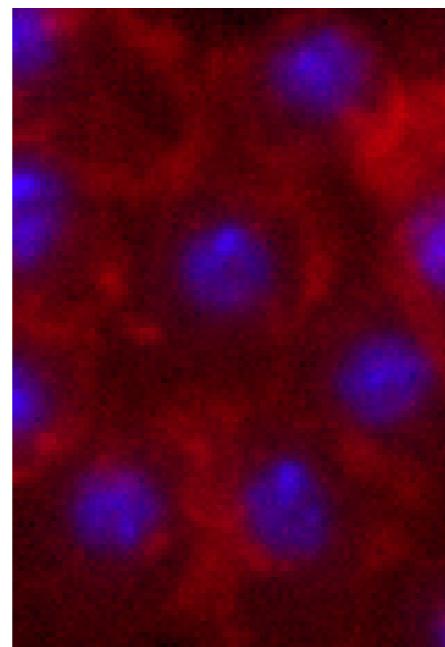


Melody  
Tsui

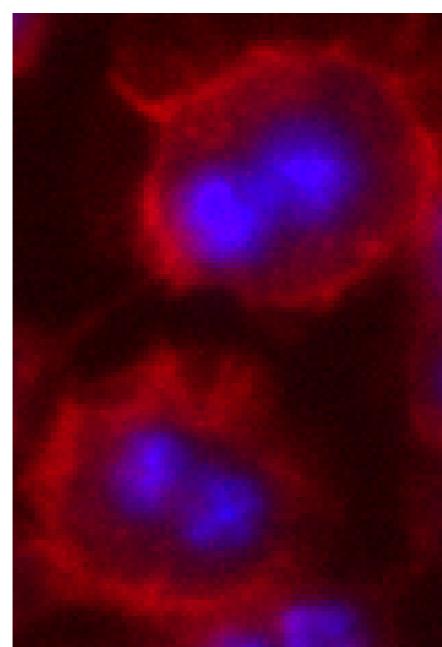
# Regulators of cell division

DNA

Actin



Normal:  
one **nucleus**  
per **cell**



Abnormal:  
two **nuclei**  
per **cell**



Ray  
Jones



Martha  
Vokes



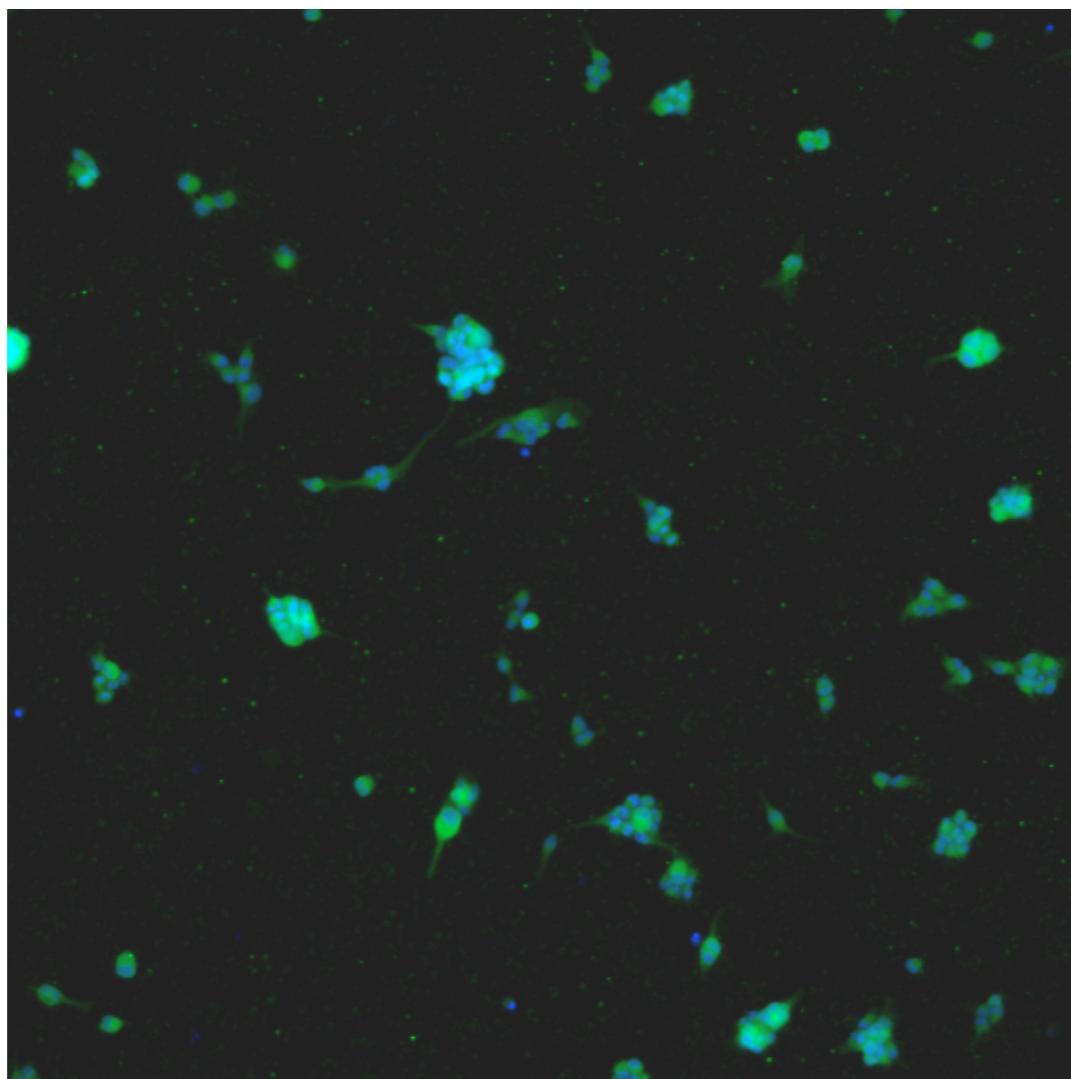
Riki Eggert,  
Harvard Med.



Adam  
Castoreno

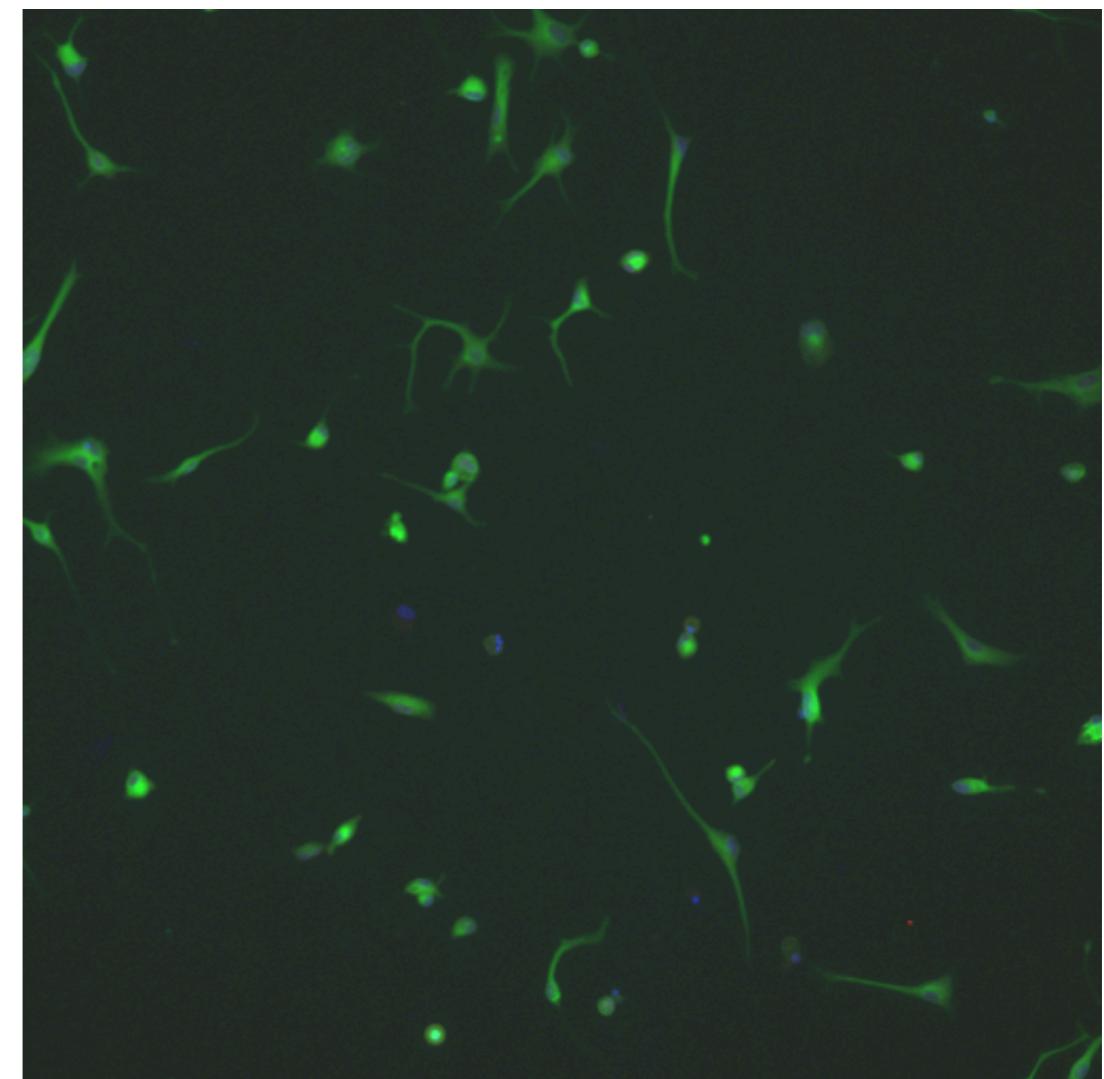
# RNAi screen: glioblastoma proliferation & differentiation

Neurosphere phenotype



DNA /  
Tubulin

Flat, elongated phenotype



Martha  
Vokes



Mark  
Bray



David  
Sabatini,  
Whitehead  
Institute



Yakov  
Chud-  
novsky,  
postdoc



William  
Hahn,  
Broad  
Institute

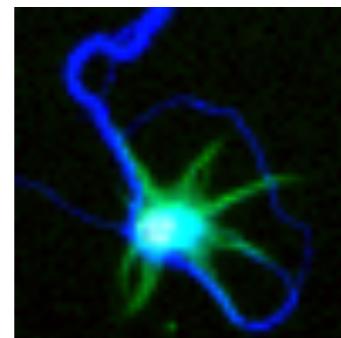


Milan  
Chheda,  
postdoc

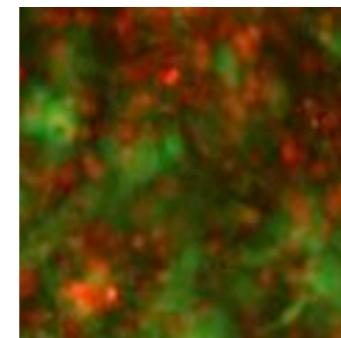


David  
Root,  
Broad Inst.

# Screens to identify genes & chemicals: increasing physiological relevance



Cell-based assays  
with complex  
morphologies



Co-cultured  
mixtures of cell  
types



Whole organisms

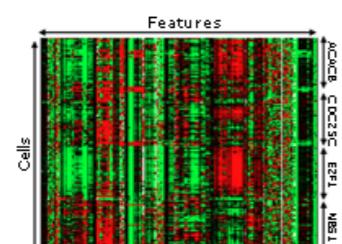
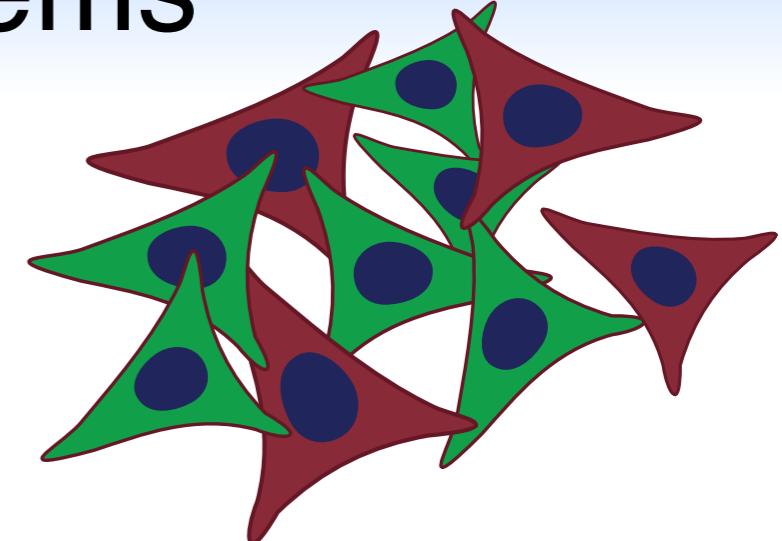


Image-based profiling: extracting ‘signatures’

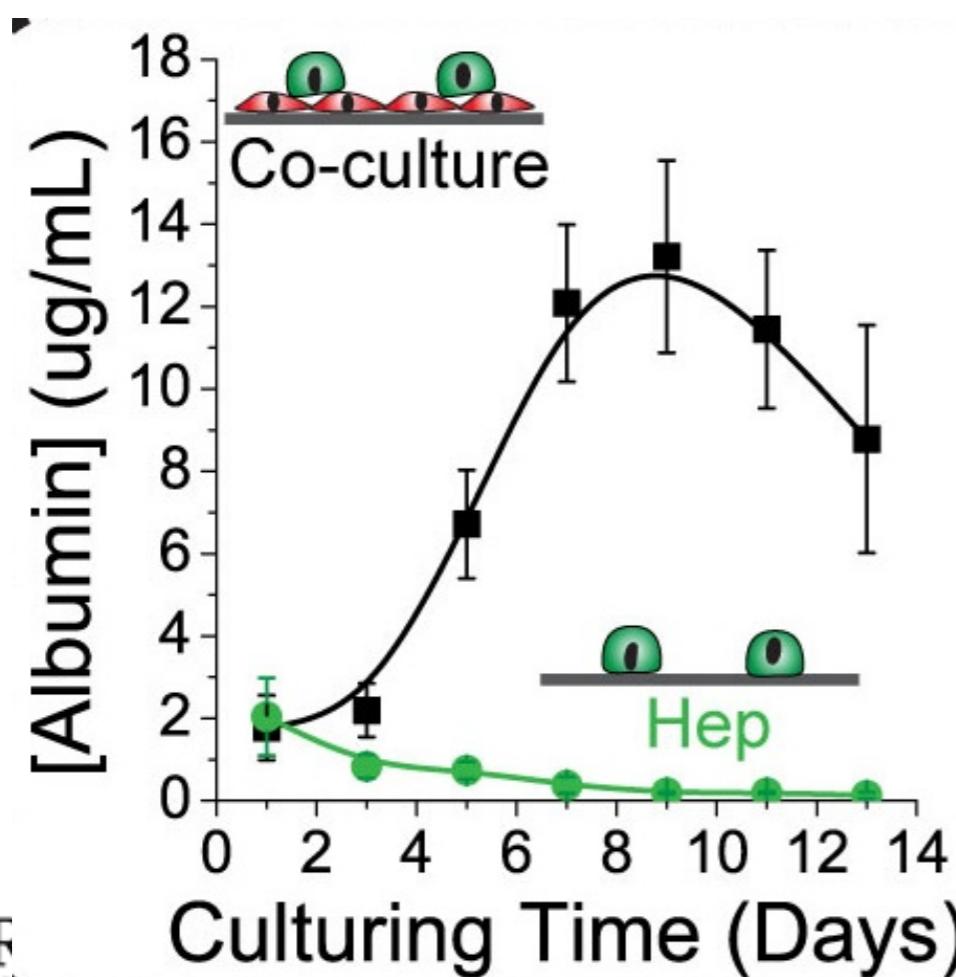
# Co-cultured cell systems

Two or more cell types cultured together in order to maintain physiological conditions



## Necessary

- Many primary cell types lose their physiological functions when grown in isolation



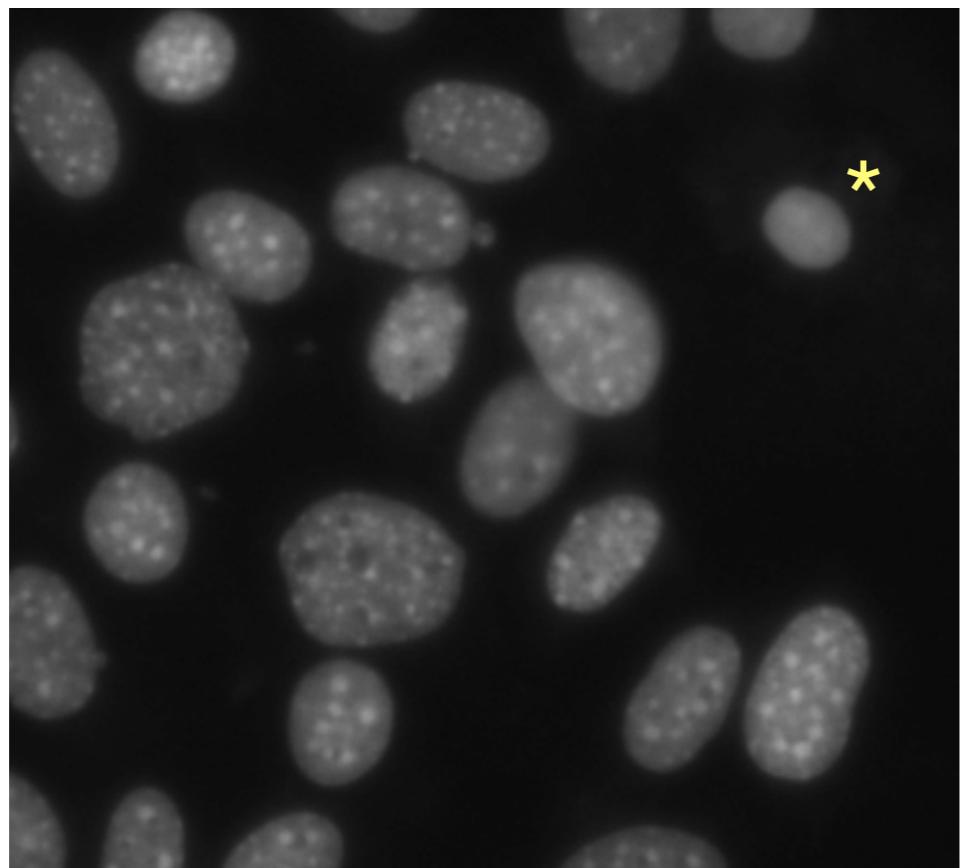
## Challenging

- Culture conditions are difficult to optimize and less robust
- Need to distinguish the cell type of interest from the co-cultured cells, ideally without using additional cellular stains

# Hepatocyte proliferation

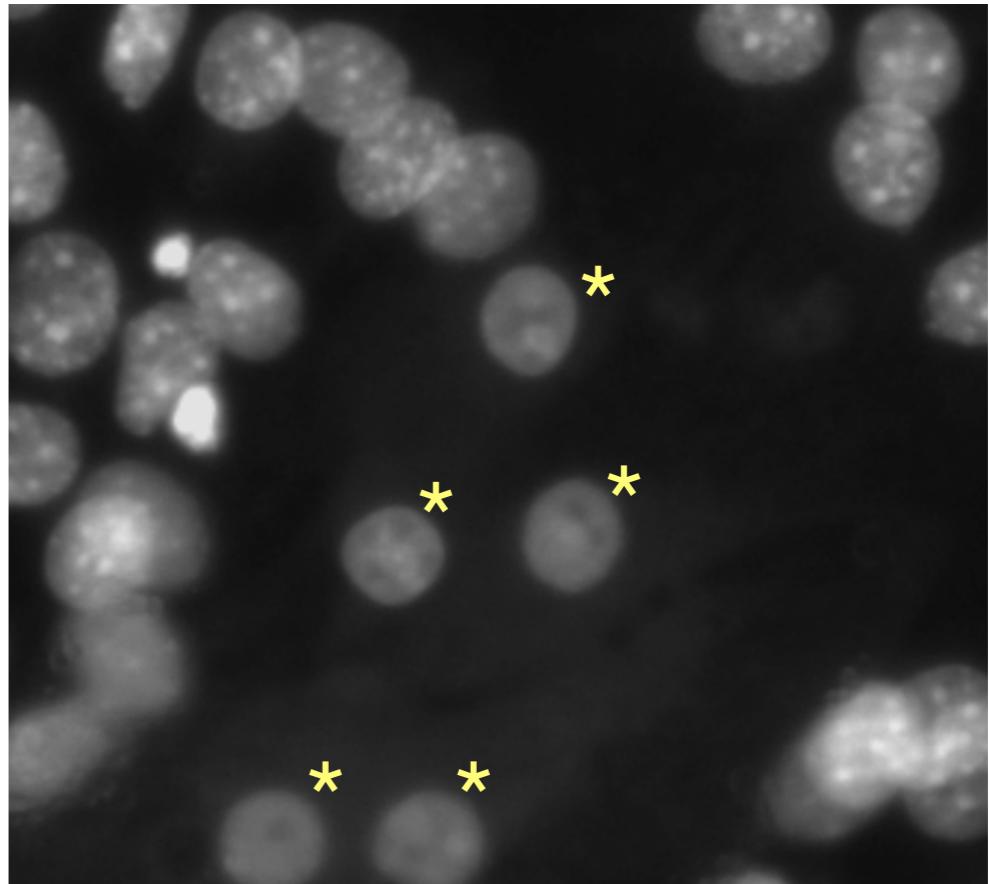
using human primary liver cells co-cultured with fibroblasts

Control



DNA

Hepatocyte-enriched



$Z'$  factor = 0.29 (*positive control = 2x hepatocytes*)



David  
Logan

Status: identified chemicals  
that stimulate primary human  
hepatocyte proliferation

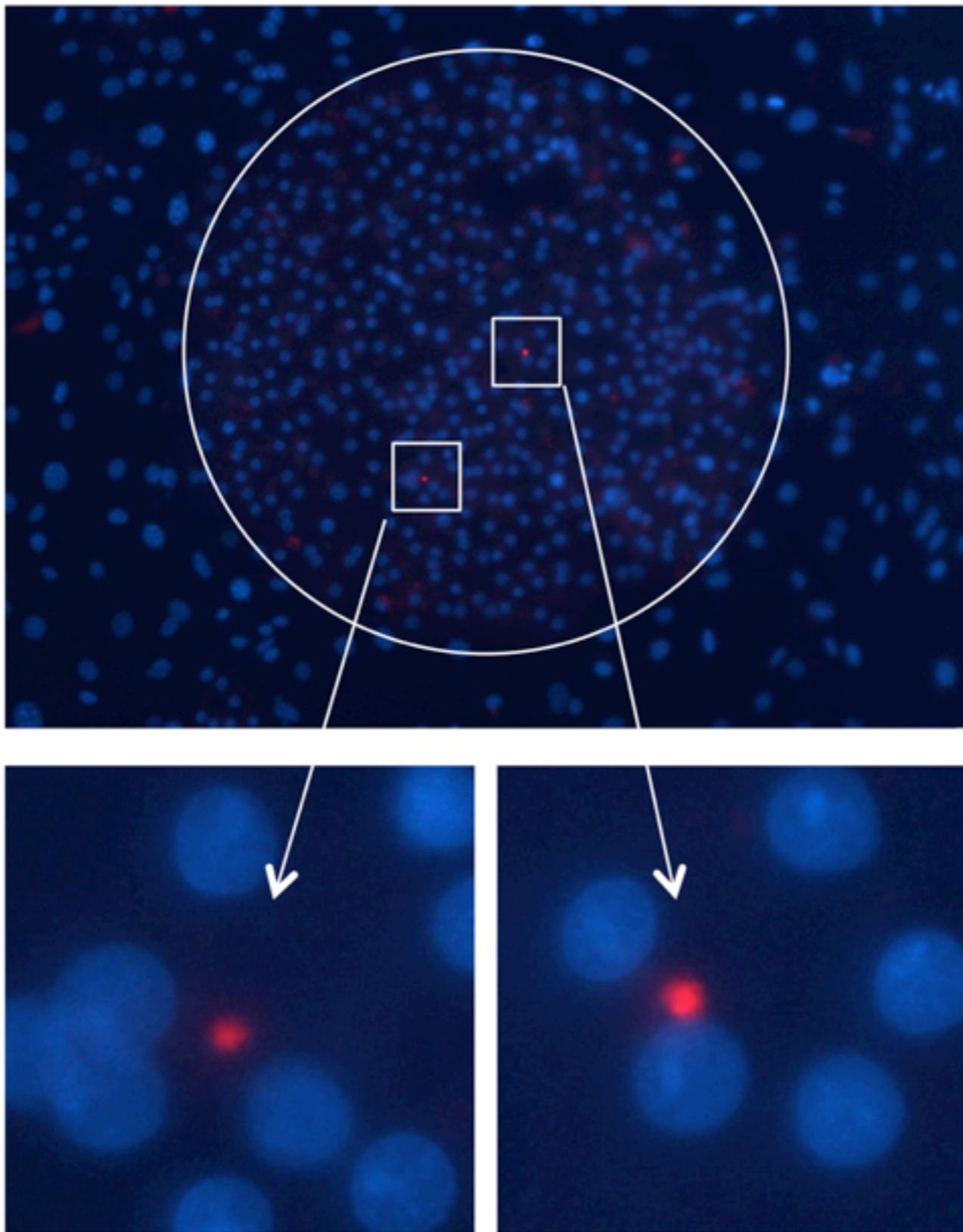


Sangeeta  
Bhatia,  
MIT



Meghan  
Shan,  
student

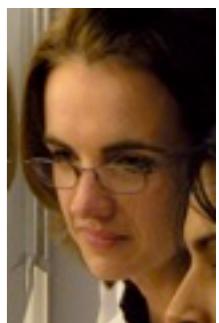
# Studying liver-stage malaria using human primary liver cells



David  
Logan



Sangeeta  
Bhatia,  
MIT

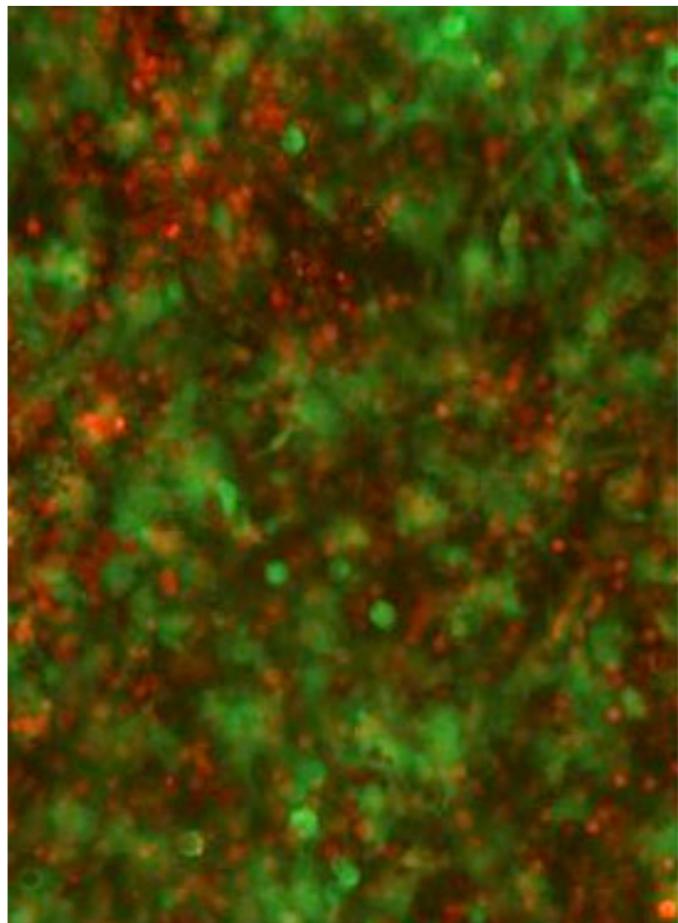


Sandra  
March

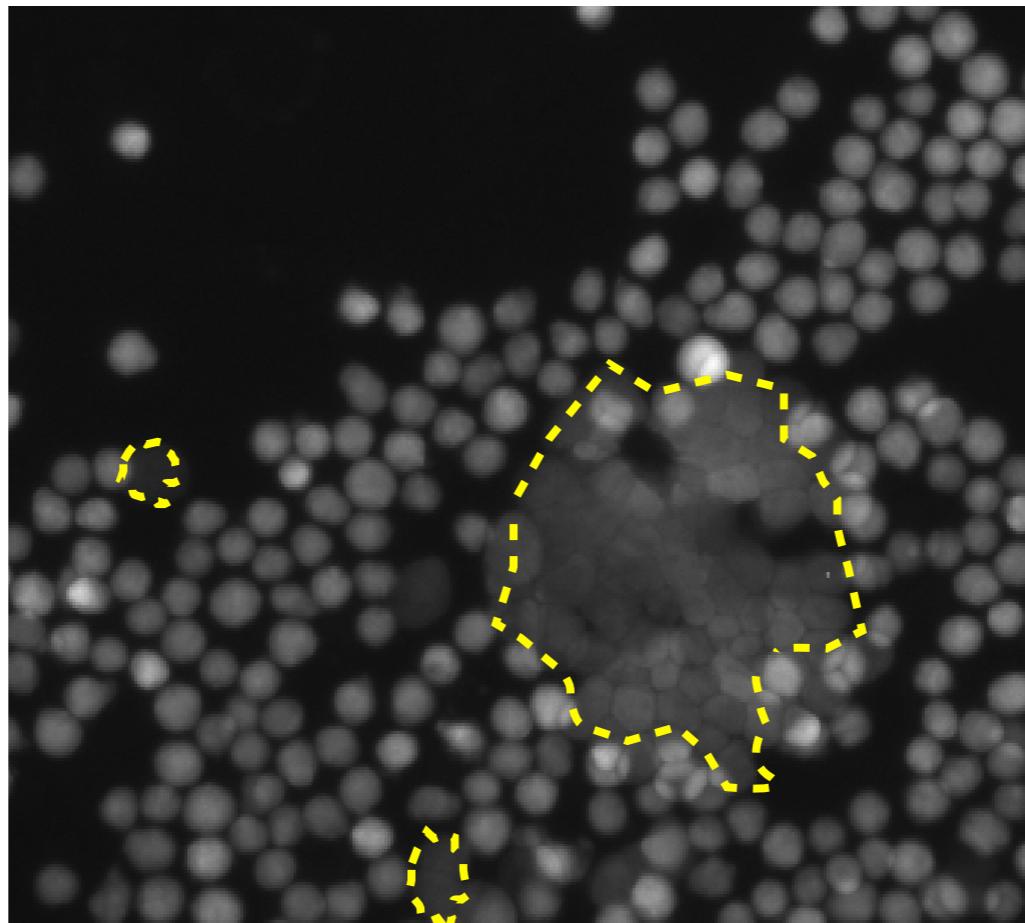
# Leukemic & hematopoietic stem cells (HSCs/LSCs)

using mouse primary HSCs or LSCs co-cultured with stromal cells

Co-cultured **LSCs**  
and **stroma**



LSC channel only:  
live, no DNA stain



**Cobblestones**



David  
Logan

Status: identified drugs  
that preferentially kill  
leukemic cells



Todd  
Golub,  
Broad  
Institute



Ben  
Ebert,  
HMS,  
Dana  
Farber



Gary  
Gilliland,  
Brigham &  
Women's  
Hospital



David  
Scadden,  
Mass.  
General  
Hospital

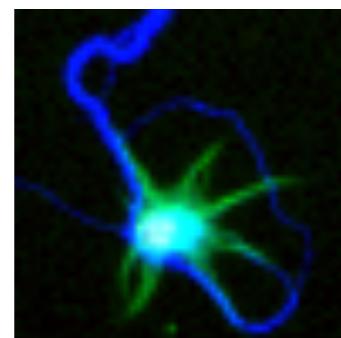


Stuart  
Schreiber,  
Broad  
Institute

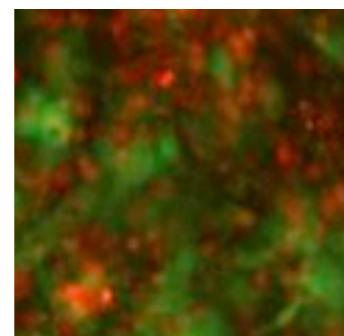
postdocs  
and  
students:  
Kimberly  
Hartwell,  
Alison  
Stewart,  
Peter  
Miller, et  
al.

*project in progress*

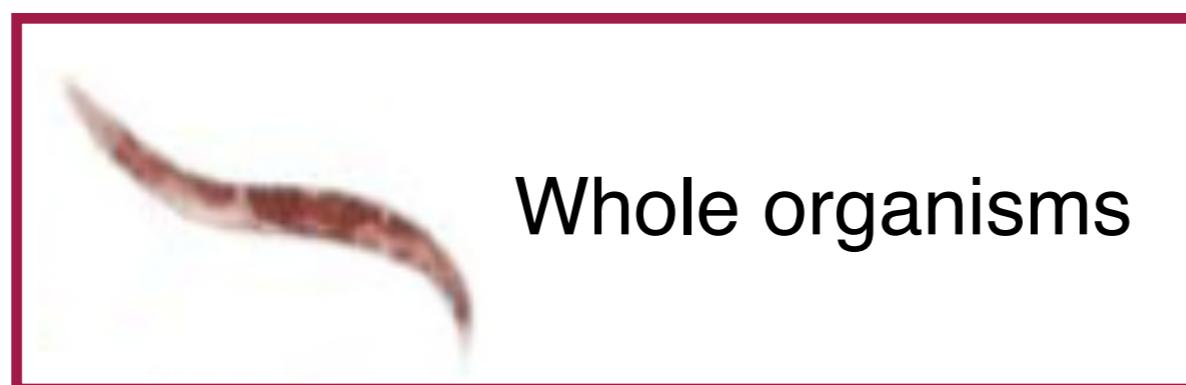
# Screens to identify genes & chemicals: increasing physiological relevance



Cell-based assays  
with complex  
morphologies



Co-cultured  
mixtures of cell  
types



Whole organisms

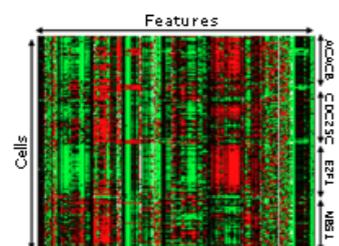


Image-based profiling: extracting ‘signatures’

# *C. elegans* High-Throughput Screening Facility

at Mass. General Hospital & Broad Institute Chem. Biol.

Consistent sample preparation & imaging at 4-6 plates per hour  
96-well agar or 384-well liquid culture



Media Dispensing



Worm sorting & dispensing  
(COPAS BioSort)



Image acquisition on  
automated microscope

## Ausubel & Ruvkun labs



Annie  
Lee-Conery



Gang  
Wu



Eyleen  
O'Rourke



Terry  
Moy

Jonah  
Larkins-Ford

# Novel anti-infectives against *E. faecalis*

*C. elegans* + *E. faecalis*  
= death

Brightfield

(SYTOX stain  
not shown)



*C. elegans* + *E. faecalis*  
+ rescuing drug = survival



37,214 compounds and extracts

80 known antibiotics and analogs

Six structural classes with “anti-infective” profile; i.e. cure infection but are not antibiotics



Ray  
Jones



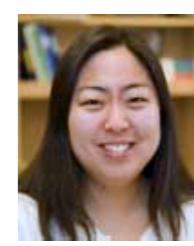
Anne  
Carpenter



Fred Ausubel,  
Harvard/  
Mass. General  
Hospital



Terry  
Moy



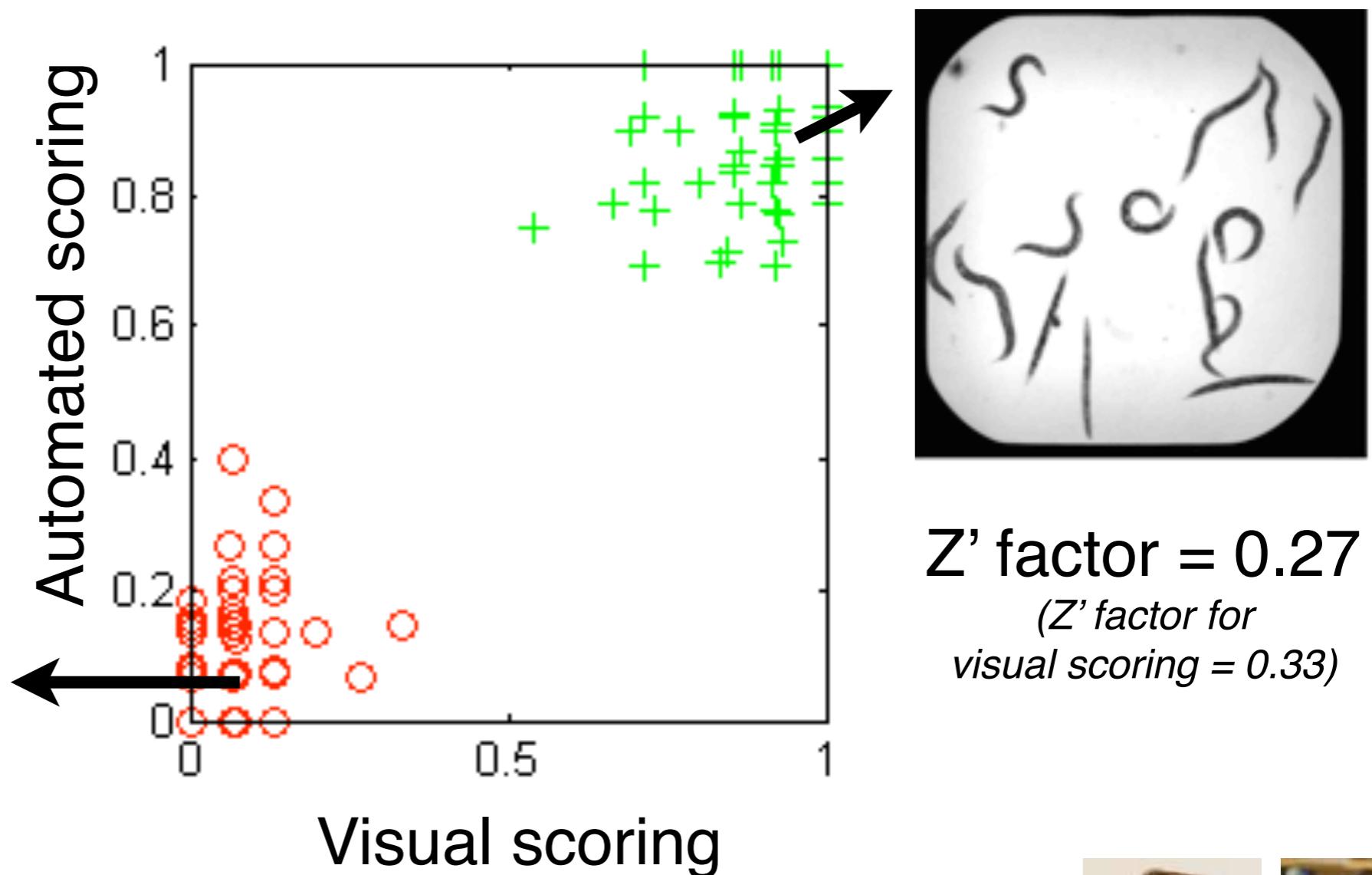
Annie  
Lee  
Conery



Gang  
Wu

# Screens for anti-infectives

Goal: score worm viability based on shape,  
after detangling the worm clusters



Visual scoring

Riklin-Raviv, Ljosa, Conery, Ausubel,  
Carpenter, Golland, Wählby: MICCAI 2010

Wählby, Riklin-Raviv, Ljosa, Conery,  
Golland, Ausubel, Carpenter: ISBI 2010

Wählby ...Carpenter, Nature Methods 2012



Carolina  
Wählby



Kate  
Sokolnicki

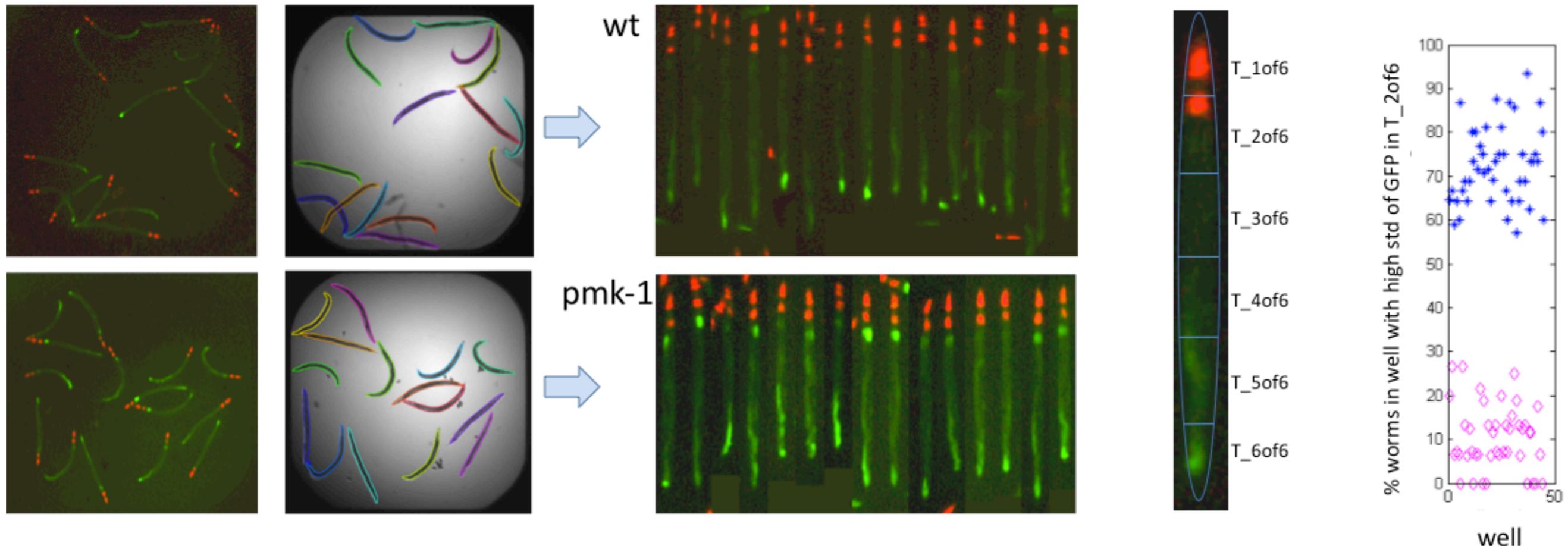


Fred  
Ausubel,  
HMS/MGH



Annie  
Lee-  
Conery

# Reporter expression in response to infection



Carolina  
Wählby



Kate  
Sokolnicki



Zihan  
Hans  
Liu



Lee  
Kamentsky



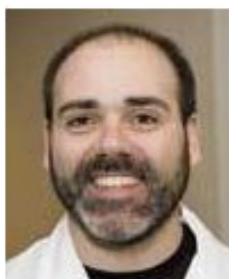
Polina  
Golland,  
MIT



Tammy  
Riklin-  
Raviv



Fred  
Ausubel



Javier  
Irazoqui

# Regulators of fat accumulation/metabolism

## Goal: Quantify Oil Red O staining



wild type



daf-2

Status: identified genes  
where RNAi alters fat  
metabolism



Carolina  
Wählby



Kate  
Sokolnicki



Gary Ruvkun  
HMS/MGH



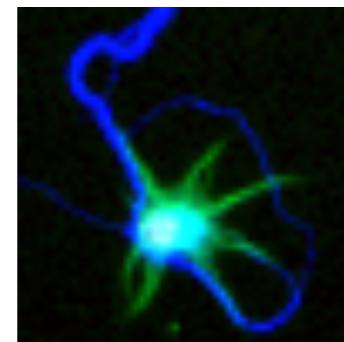
Eyleen  
O'Rourke



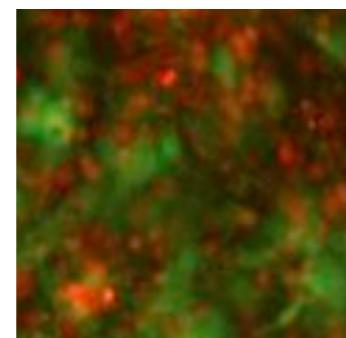
Annie  
Lee-  
Conery

*project in progress*

# Screens to identify genes & chemicals: increasing physiological relevance



Cell-based assays  
with complex  
morphologies



Co-cultured  
mixtures of cell  
types



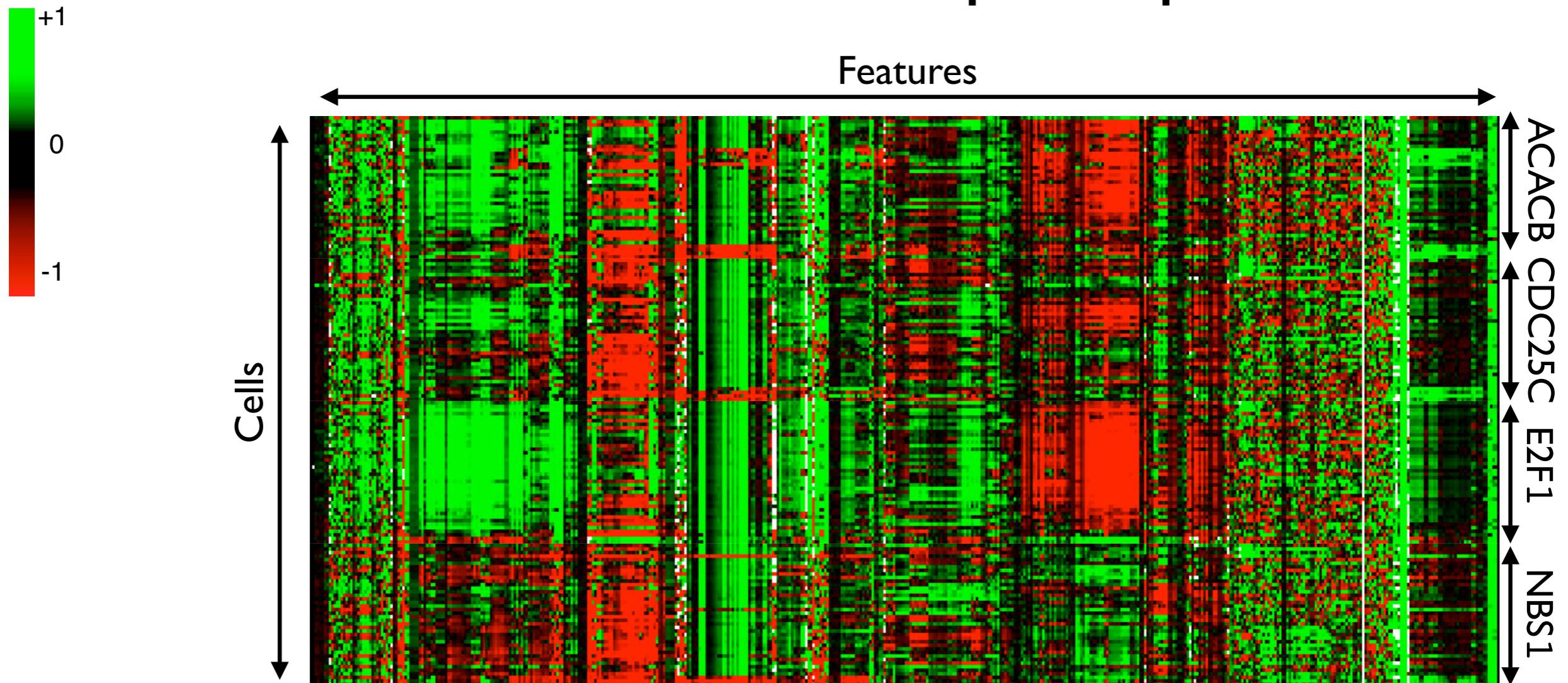
Whole organisms



Image-based profiling: extracting ‘signatures’

# Latent information in image-based experiments

- thousands of features per cell
- thousands of cells per sample

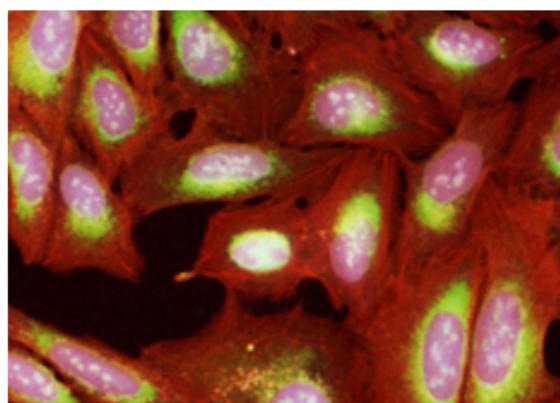


“Cytological profile”: collection of measurements describing the appearance of a cell

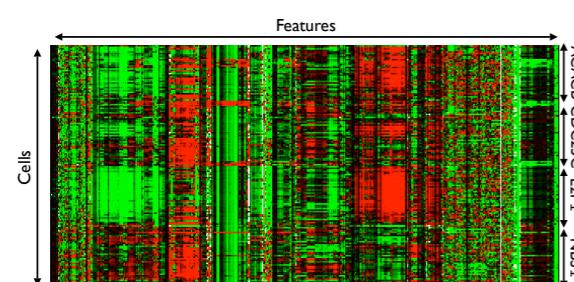
*Perlman, et al. Science 2004*

# Automatically extracting image-based phenotypes

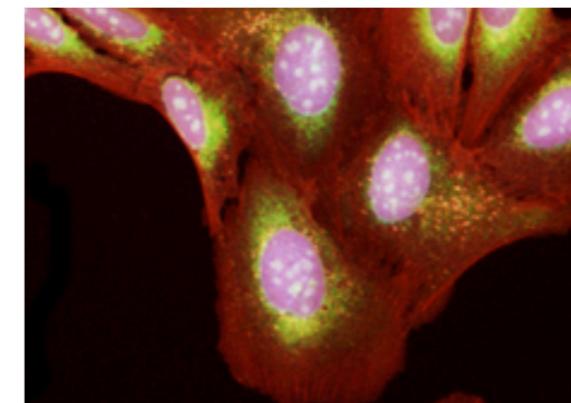
Wild-type cells



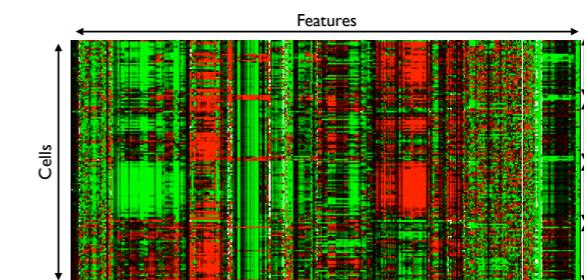
Extract  
features



Perturbed cells



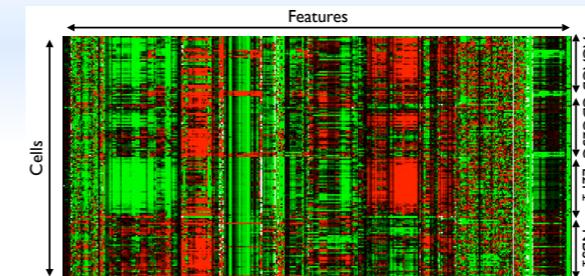
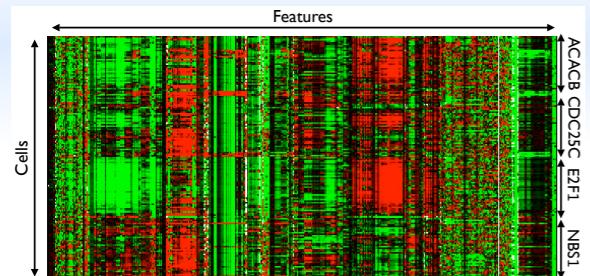
Extract  
features



Detectable  
phenotypic  
difference?

Identify difference in phenotype from image features  
*even if “invisible” to the human eye*

Screen for chemicals that can revert  
mutant phenotype -> wild type



Detectable phenotypic difference?

Cells from patient without mental illness



Cells from patient with bipolar disorder



Bruce Cohen,  
McLean Hosp.



Rakesh  
Karmacharya

Cells treated with hepatotoxic drug



Cells treated w/ non-hepatotoxic analog



Todd Golub,  
Broad Institute



David  
Thomas

Wild-type cells



Cells treated with RNAi against HDAC1, 2, 3...



Stuart Schreiber,  
Broad Institute



Angela  
Koehler



Ray  
Jones



Vebjorn  
Ljosa



Kate  
Sokolnicki



Auguste  
Genovesio



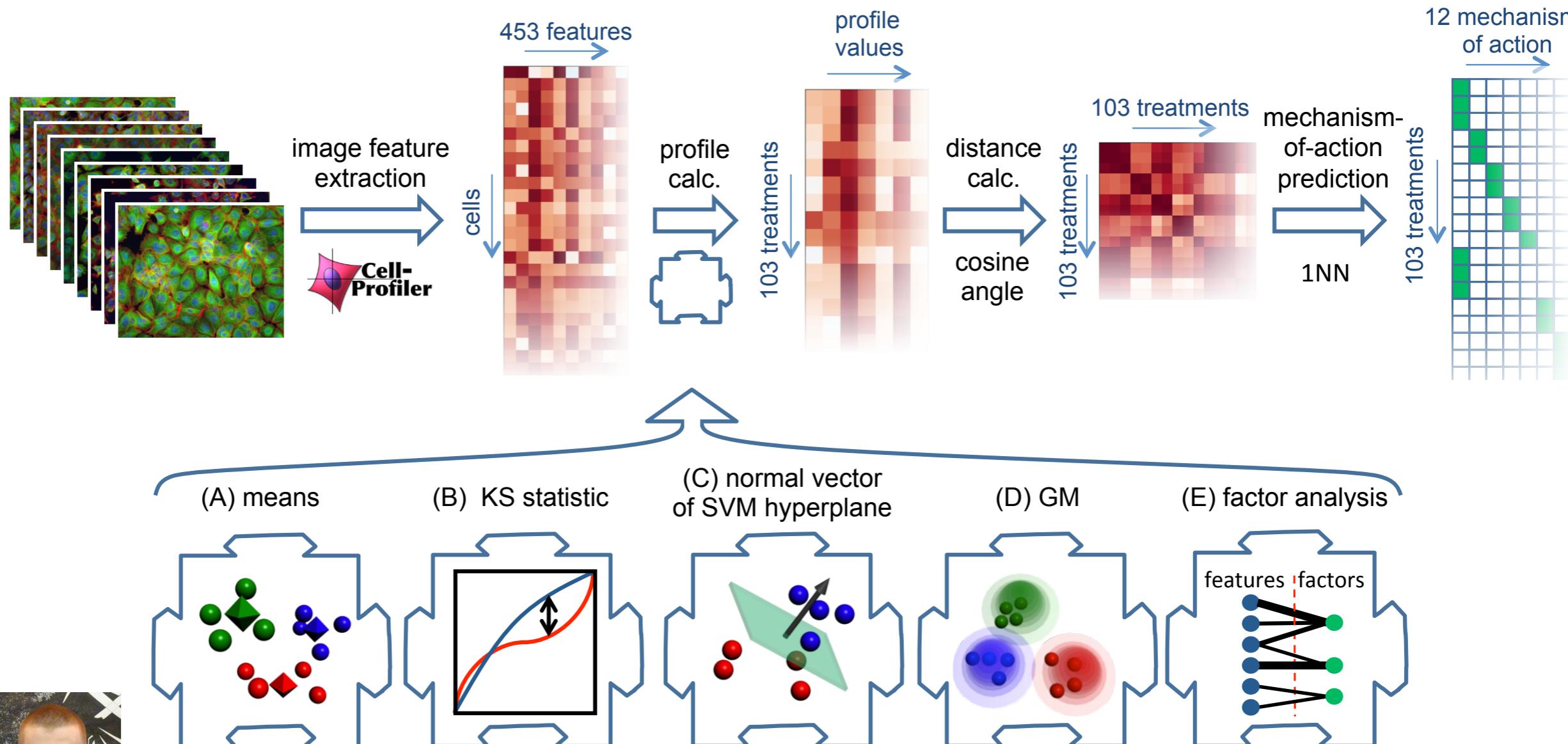
Shantanu  
Singh

Postdoctoral position available!

# In order to detect phenotypic differences

- Biology
  - The relevant biological difference must be present (cell type, culture conditions, etc)
- Labels
  - We must have stained the samples so that the differences are visible
- Image features
  - Feature extraction algorithms must capture the differences
- Similarity measure
  - We must compare the features effectively

# Comparing profiling methods for predicting MOA (mechanism of action, for small molecules)



Vebjorn  
Ljosa



Rob ter  
Horst

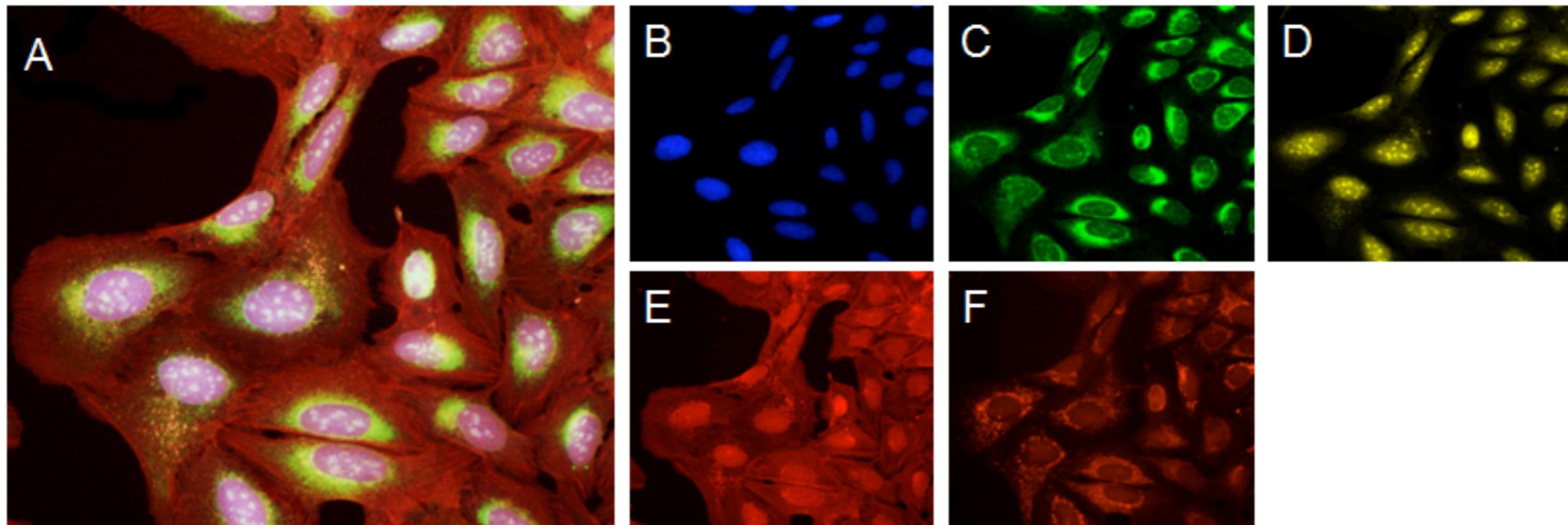
*project in progress*

# In order to detect phenotypic differences

- Biology
  - The relevant biological difference must be present (cell type, culture conditions, etc)
- Labels
  - We must have stained the samples so that the differences are visible
- Image features
  - Feature extraction algorithms must capture the differences
- Similarity measure
  - We must compare the features effectively

# Maximize information in profiling experiments

## Cell-painting assay



6 stains, 5 channels imaged, revealing:  
nucleus, nucleoli, actin, golgi, plasma membrane, mitochondria, ER



Ray  
Jones

Kate  
Sokolnicki



Stuart Schreiber,  
Broad Institute



Aly  
Shamji

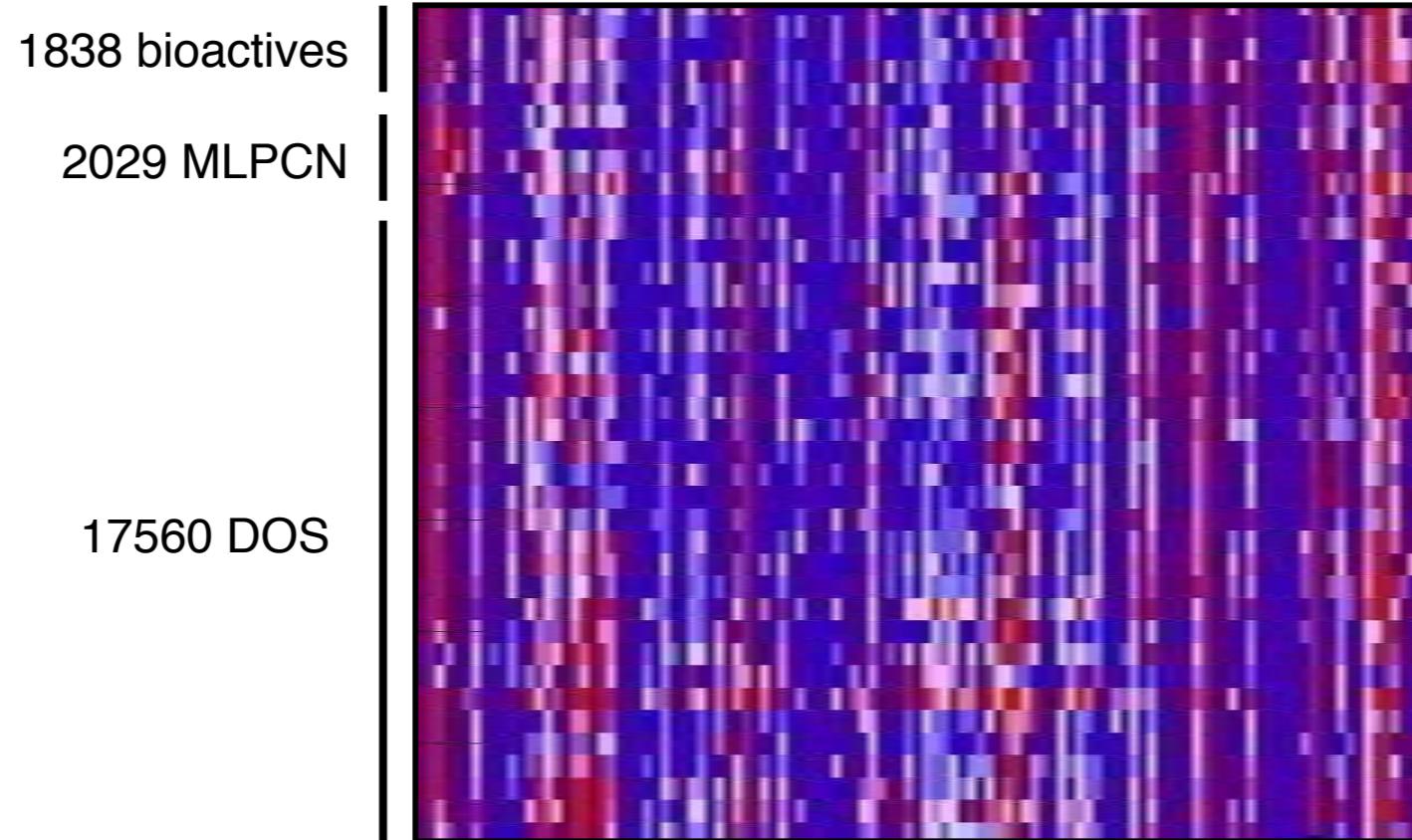


Paul  
Clemons



Sigrun  
Gustafsdottir

# Profiling more diverse compounds

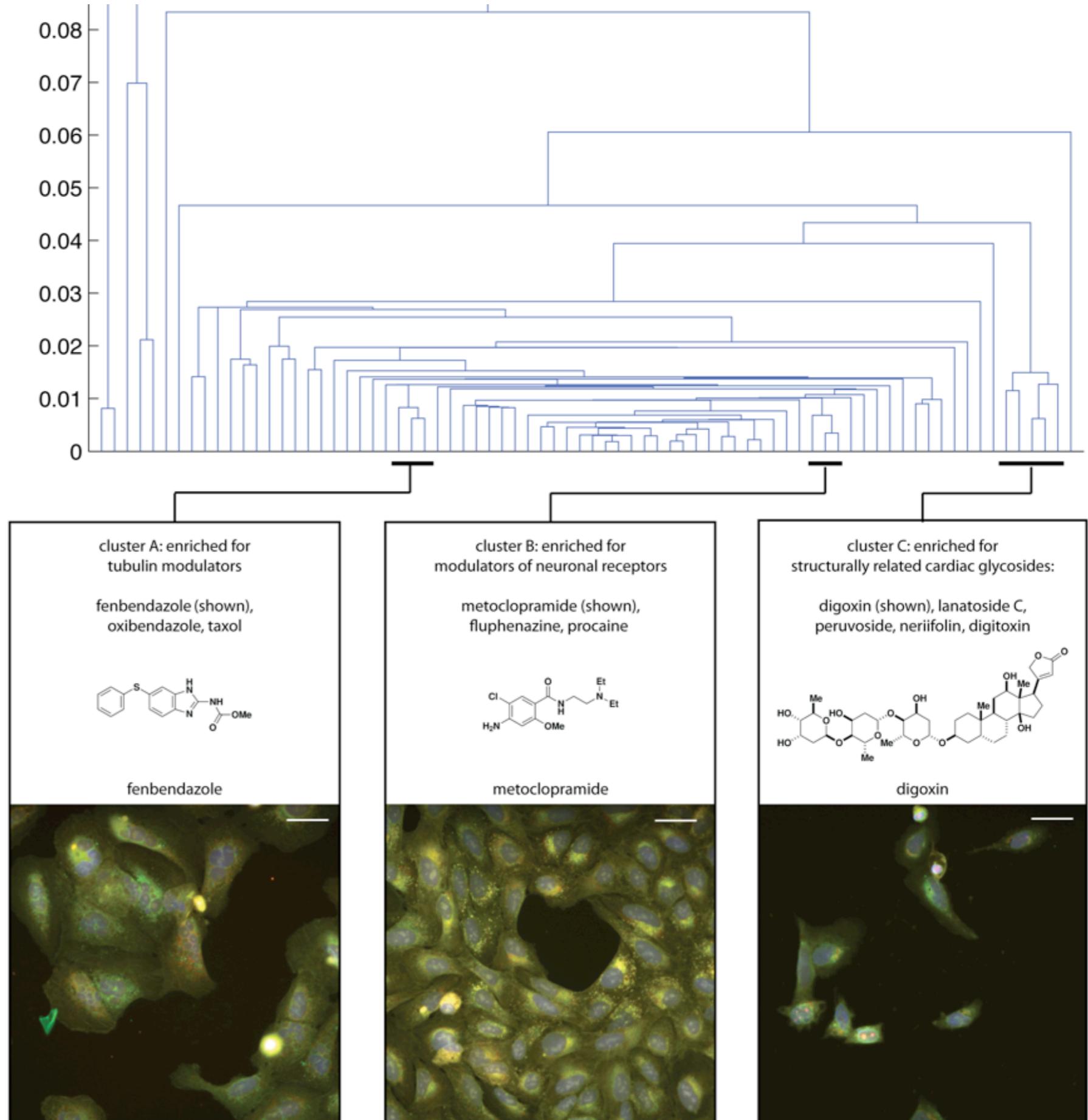


# Profiling more diverse compounds

203 out of  
1,600  
compounds  
yielded a  
phenotype

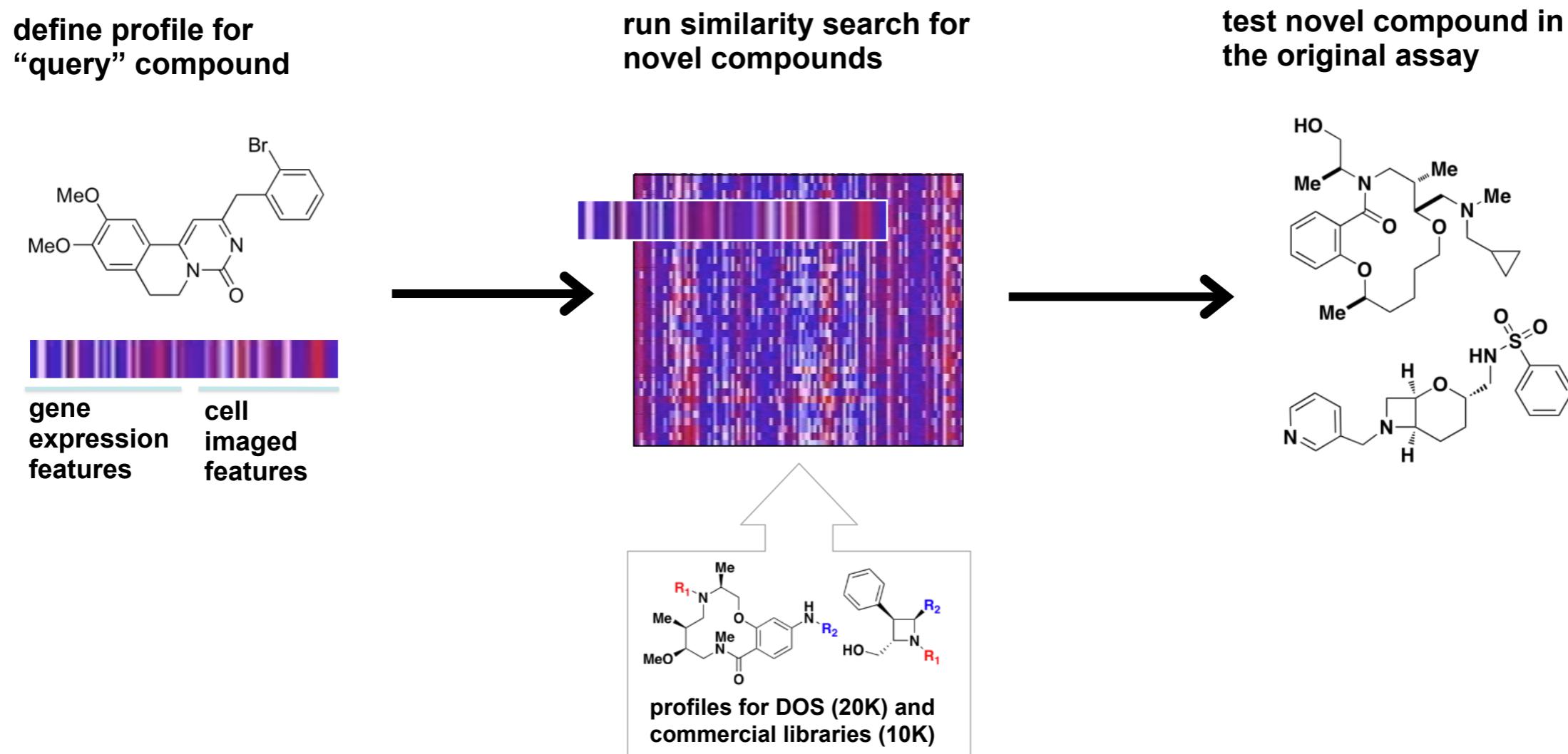


Vebjorn  
Ljosa



# Can profiling be used to identify similar compounds in a large-scale experiment?

- Profile ~200 “query” compounds from ~35 screens
- Identify similar-performing compounds from a library of 10,000 commercially available compounds + 20,000 novel compounds (DOS)
- Use imaging and gene expression for profiling





[www.cellprofiler.org](http://www.cellprofiler.org)

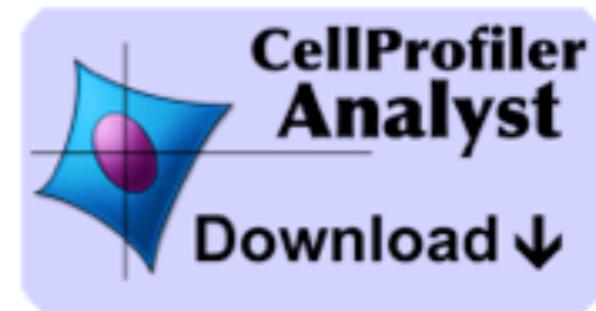
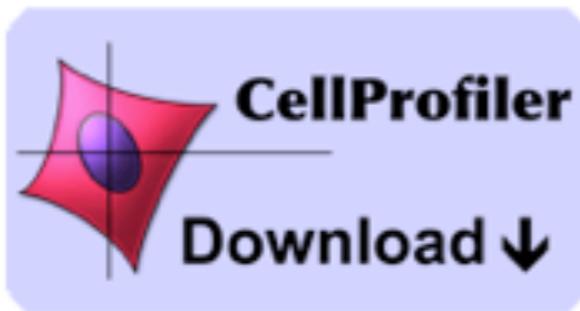
# CellProfiler

cell image analysis software



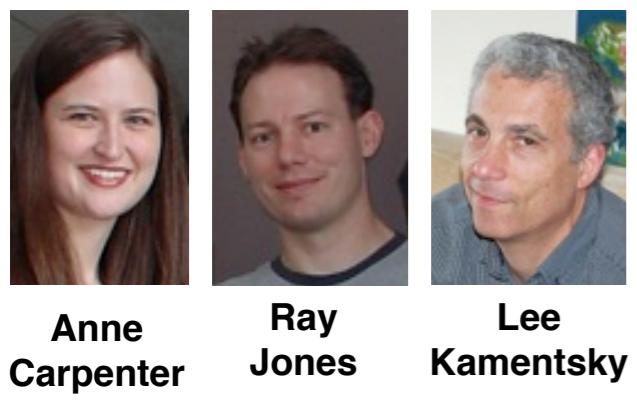
[home](#)   [getting started](#)   [download](#)   [for developers](#)   [forum](#)   [about us](#)

search



- Downloaded **>38,000** times
- Launched **>90,000** times/year
- Cited in **>600** papers
- One of the **Top 10** most-accessed papers of all time in Genome Biology
- Winner of Bio-IT World's IT & Informatics Best Practices Award in 2009

*As of December, 2012*



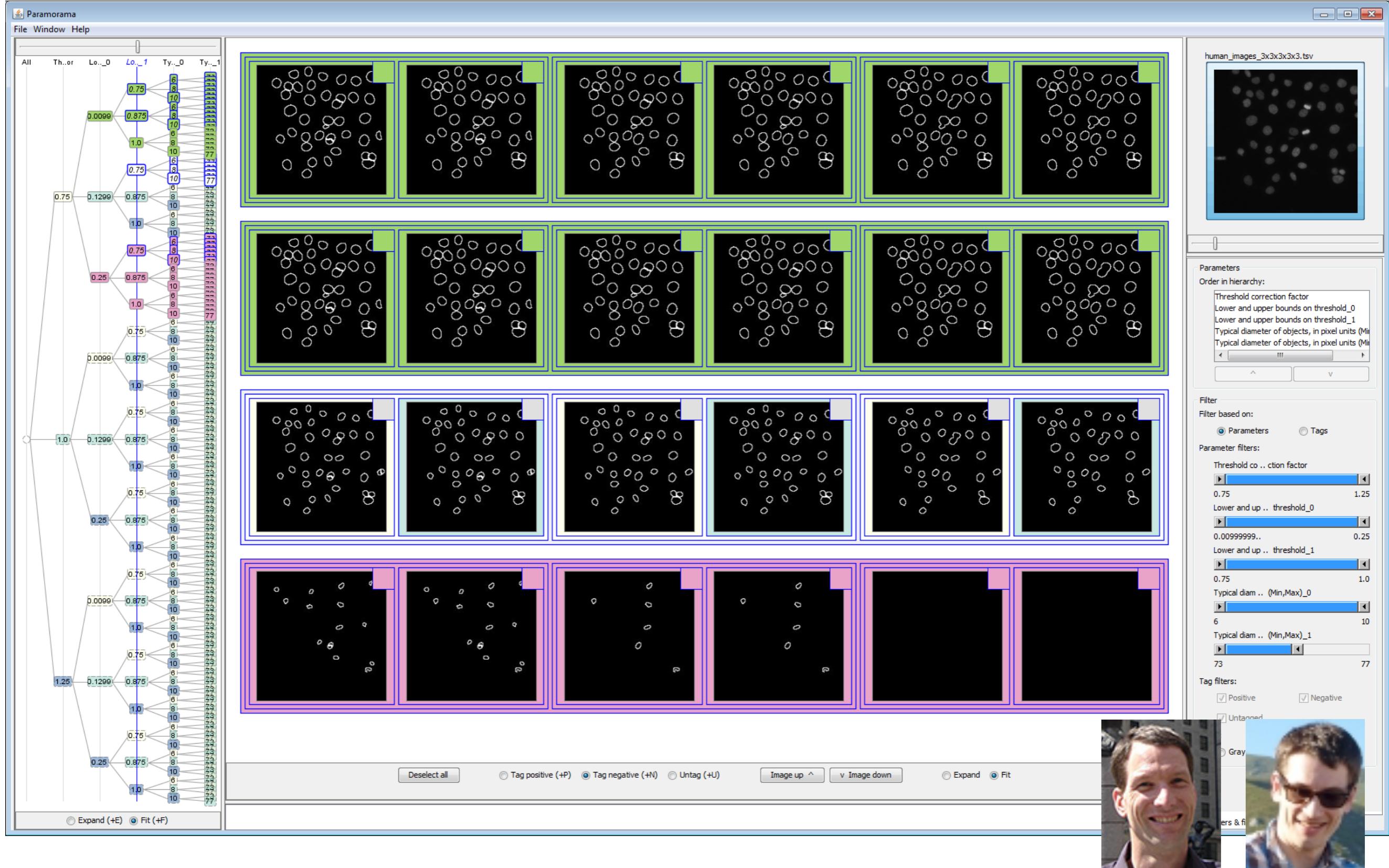
# ProtocolNavigator: reproducible research

The screenshot displays the ProtocolNavigator software interface. On the left, the 'Experiment Settings' window shows project details: Project Title (MIFlowCyt), Project Aim (Quantify Donor HSC contribution to lymphocytes and myeloid cells by flow cytometric analysis of peripheral blood cells stained with appropriate combinations of), and various experimental parameters like keywords, experiment number, start date, and status. Below this is the 'Bench' window, which tracks a timeline from 0:00 to 2689:37, listing actions such as incubation, spinning, washing, and harvesting. It also shows a plate layout for 'Plate3' with positions A, B, and C filled with yellow circles. On the right, the 'Experiment Lineage' window provides a visual representation of the experimental workflow, showing a grid of nodes connected by lines, with yellow nodes indicating active or completed steps.



Imtiaz  
Khan

# Paramorama: optimizing image processing



Roy Ruddle, Hannes  
Univ. Leeds Pretorius

# Helpful resources

[www.cellprofiler.org](http://www.cellprofiler.org) -> Tutorials

OPEN  ACCESS Freely available online

PLOS COMPUTATIONAL BIOLOGY  
2009, Vol. 5 No. 12

Education

## Introduction to the Quantitative Analysis of Two-Dimensional Fluorescence Microscopy Images for Cell-Based Screening

Vebjorn Ljosa, Anne E. Carpenter\*

Box 1. Resources for further exploration

Box 2. Practicalities

# Gratitude



free, at [www.cellprofiler.org](http://www.cellprofiler.org) :



## IMAGING PLATFORM

### Lab members

Mark Bray  
Lee Kamentsky  
Vebjørn Ljoså  
David Logan

Shantanu Singh  
Matthew Veneskey  
Carolina Wählby

Many thanks to our  
many biology collaborators

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)