Rapid, high-content genome-wide assays using cell microarrays

Anne E. Carpenter, Ph.D.
David M. Sabatini Lab
Whitehead Institute for Biomedical Research
What are all the genes doing?

- Cell size, count & morphology
- Nuclear size & morphology
- Nucleolar size, count & morphology
- Cell cycle distribution (DNA content)
- Apoptosis
- Protein content
- Actin content/morphology
- Phosphorylation of signaling proteins (S6, S6K, Akt)
- Localization of signaling proteins
- Mitochondrial size & morphology
- Upregulation of translation
- Amount & distribution of signaling lipids
- Upregulation of transcription

SYSTEMATIC GENOME-WIDE SCREENS OF GENE FUNCTION

Anne E. Carpenter and David M. Sabatini

Nature Reviews Genetics
Technologies to quickly determine gene function

Create a set of RNA interference reagents, one for each gene

Create spots of cells, each treated with a different RNAi reagent

Stain cells for a specific phenotype

Automated image collection

“Living cell microarray”

Automated image analysis with CellProfiler

Data analysis

Determine the phenotypic effects of knocking down every gene in the genome
Living cell microarray technology

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

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


5,600 spots per slide
Genome-wide screens in Drosophila

One glass slide: 5600 spots of dsRNA, each knocking down a different gene.

Actin
DNA: dark spot = lack of cells

DNA: described in Wheeler...Sabatini, Nature Methods 2004
reviewed in Wheeler, Carpenter, Sabatini, Nature Genetics suppl., June 2005
How can we measure cells automatically?

Result: hundreds of thousands of cell images

- GFP
- PPV
- dTOR
- PP2A CatSub
- CG9006
- String
- Cyclin A
- rpL12

...plus ~20,000 more images

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

- less tedious, less biased, quantitative
Sophisticated algorithms needed

Drosophila Kc167 cells

DNA (nuclei)

Actin (cell edges)

Jones, Carpenter & Golland (2005) ICCV Workshop on Computer Vision for Biomedical Image Applications
The CellProfiler project

CellProfiler™
cell image analysis software

Runs on Mac/PC/Unix, plugs into Matlab, can make use of cluster computing

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

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

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

Free!
Typical CellProfiler pipeline:

original images

Image processing modules

processed images

Illumination correction modules

illumination-corrected images

Object identification modules

identified objects (nuclei and cells)

Measurement modules

measurements for every cell in every image (number, location, size, shape, intensity, texture) can be analyzed by:

1. built-in CellProfiler data tools
2. exporting to spreadsheet
3. exporting to database
4. analyzing in MATLAB
Measurable cell features

- Location: X,Y
- Cell count
- Object count within cells (e.g. speckles within nucleus)
- Neighbors
- Size
- Shape

For all colors:
- Intensity (of entire object and of the edge of the object)
- Texture
- Correlation between different colors
Cell count validation

Human HT29

Drosophila
Cell area validation

![Graph showing cell area validation for different conditions and treatments. The x-axis represents different treatments (no dsRNA, Mad2, String, Anillin, Cyclin A), and the y-axis represents mean cell area in arbitrary units. Two datasets are compared: Coulter Counter and CellProfiler.]
Validation for DNA content (cell cycle)-Drosophila

Number of cells (thousands)

- no dsRNA
- Mad2
- String
- Anillin
- Cyclin A

DNA content, log scale (arbitrary units)
Slide scale normalization

Per-nucleus DNA content:

Normalized by local median:

Number of cells

DNA content

Number of cells
Field-of-view illumination correction

image from Steve Bailey, Sabatini lab
Field of view illumination correction

![Image showing field of view illumination correction](image)

The diagram illustrates the comparison between images with and without illumination correction. The graphs show the number of cells (in thousands) plotted against DNA content on a log scale (in arbitrary units). The corrections applied are:

- No illumination correction
- CellProfiler illumination correction
- White reference illumination correction
Validation for DNA content

Human HT29

Nuclei image  CellProfiler-outlined nuclei

mouse

wild-type  knockout

DNA content

Images from Andrew Baltus, Page lab, Whitehead Institute
Validation for DNA content
Time lapse movies of Drosophila embryos

Goal: identify nuclei & measure morphology & GFP content

movie from Victoria Foe, Univ. Washington
Time lapse movies of Drosophila embryos

Goal: identify nuclei & measure morphology & GFP content

movie from Victoria Foe, Univ. Washington

GFP content

Area

Shape
Time lapse movies of Drosophila embryos
Antibody staining intensity - Mouse tissue

Goal: score cells as positive or negative for the red-stained Mvh

Image from Andrew Baltus, Page lab, Whitehead Institute
Membrane localization

Goal: quantify the localization of proteins

RFP

GFP

Thy1

Cytoplasm-nucleus translocation assay

Goal: quantify the localization of proteins

CellProfiler results
Z-factor  Wortmannin  LY294002
0.91  0.86  0.83

V-factor

images from BioImage

Dose of Wortmannin (nM)

error bars = SEM
Speckle-counting assay

Goal: count and measure phospho-Histone2AX speckles

images from Scott Floyd, MIT
First genome-wide screen: in progress

**DNA staining:**
- cell count
- cell cycle distribution
- chromatin texture
- nuclear size
- nuclear morphology

**Actin staining:**
- cell size
- cell morphology
- actin content
- actin texture

**phospho-H3:**
- p-H3 amount
- p-H3 localization

Every gene can be screened in a single experiment using four microscope slides!
Data analysis: Population measures

- Cell count
- Average cell area
- Correlation between actin and phospho-Akt staining

black = low values, white = high values
Data analysis: Population measures

True high-content data set produced by multi-parameter phenotypic analysis
Discovering the function of undescribed genes

Phenotypes

<table>
<thead>
<tr>
<th>Gene</th>
<th>cell count</th>
<th>cell size</th>
<th>actin content</th>
<th>nucleus size</th>
<th>DNA content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gene #1</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene #2</td>
<td>267</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene #3</td>
<td>202</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In progress: data exploration with CellVisualizer
What are all the genes doing?

- Cell size, count & morphology
- Nuclear size & morphology
- Nucleolar size, count & morphology
- Cell cycle distribution (DNA content)
- Apoptosis
- Protein content
- Actin content/morphology
- Phosphorylation of signaling proteins (S6, S6K, Akt)
- Localization of signaling proteins
- Mitochondrial size & morphology
- Amount & distribution of signaling lipids
- Upregulation of translation
- Upregulation of transcription

SYSTEMATIC GENOME-WIDE SCREENS OF GENE FUNCTION

Nature Reviews Genetics

Anne E. Carpenter and David M. Sabatini
Thanks to...

David M. Sabatini Laboratory
Anne Carpenter
Colin Clarke
Xana (Maria) Frias
David Guertin
Kalyani Guntur
Thouis R. Jones
Mike Lamprecht
Wyman Li
Susan Ma
Jason Moffat
Kathleen Ottina
Tim Peterson
Dos Sarbassov
Yasemin Sancak
Shomit Sengupta
Joon-Ho Sheen
Carson Thoreen
Doug Wheeler

CellProfiler

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 help from:
Michael Lamprecht, Colin Clarke, In Han Kang, Ola Friman, Steve Lowe, Joo Han Chang, Susan Ma

CellProfiler has been supported by funding from:
• Novartis postdoc fellowship from the Life Sciences Research Foundation
• Merck/MIT Computational & Systems Biology postdoc fellowship
• Society for Biomolecular Screening Small Grant Award
• DOD Tuberous Sclerosis Complex Grant
• Sabatini Laboratory