

Cell Images in Drug Discovery



Anne Carpenter



Runxi Shen

PURDUE UNIVERSITY

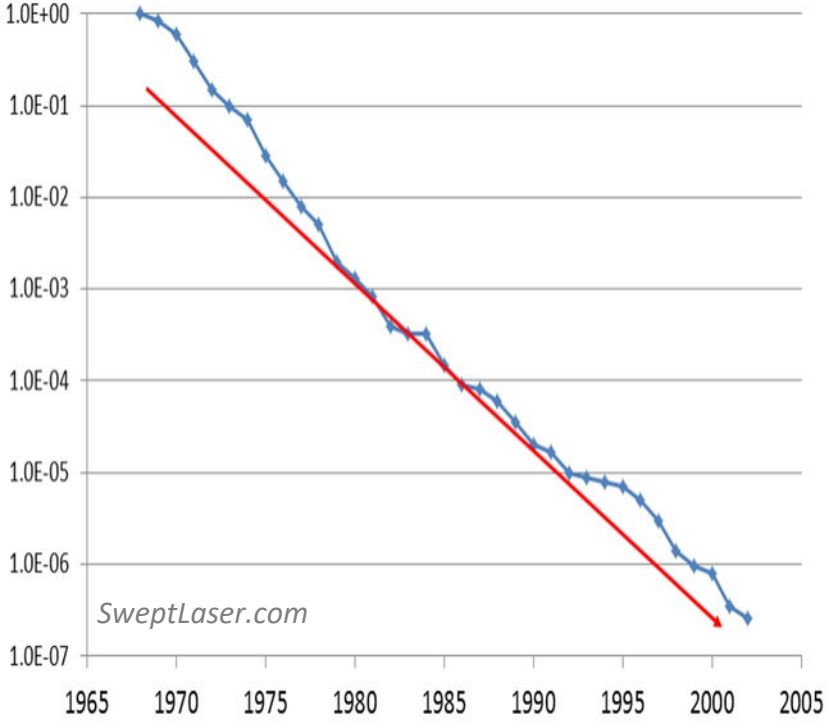
Download PDF of slides here: http://broad.io/CarpenterSlides

(talk content is non-confidential)

Tale of two industries

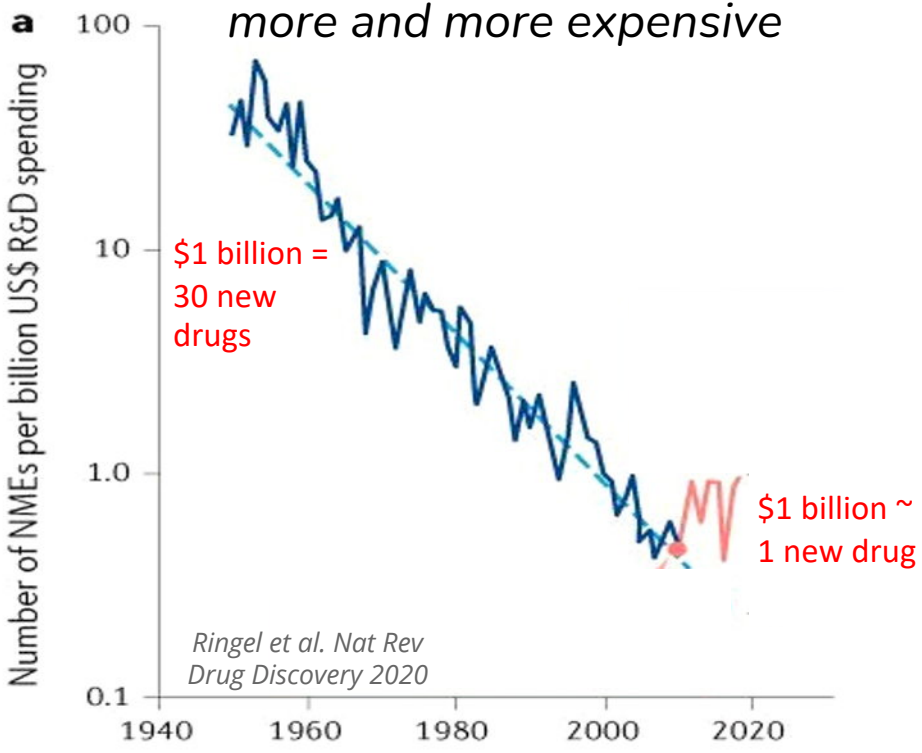
Moore's Law

Compute gets cheaper and cheaper

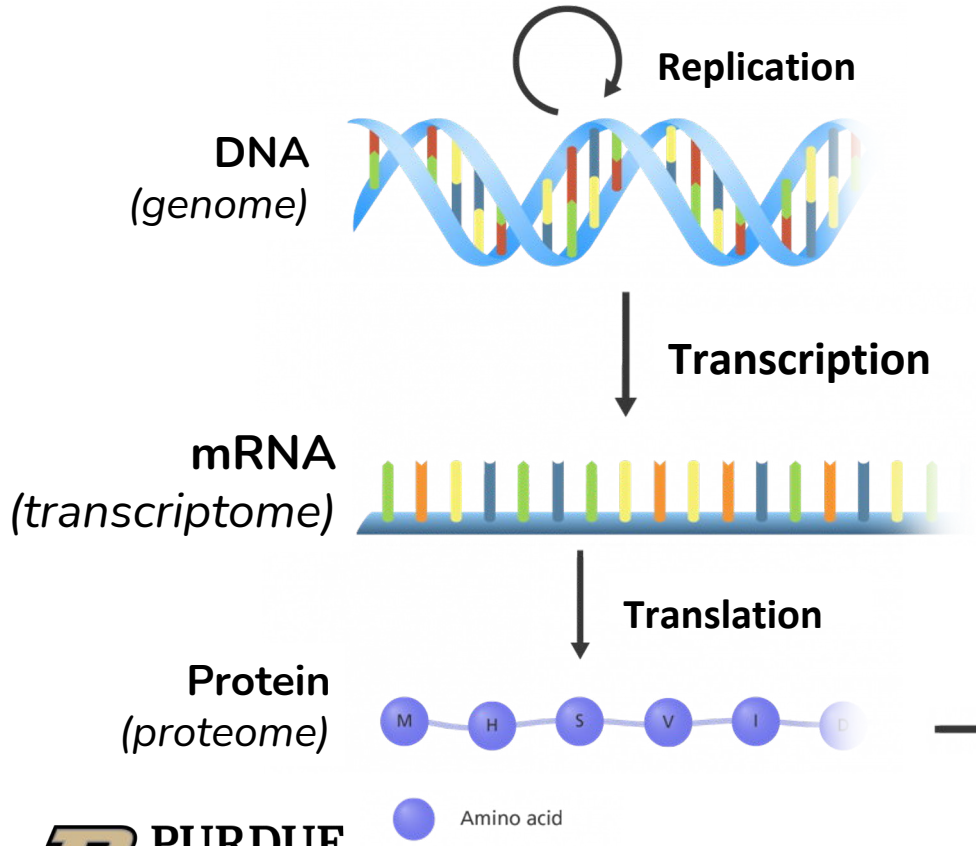


Eroom's Law

Discovering new medicines gets more and more expensive



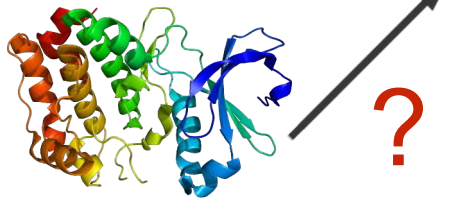
AI is getting better at some tasks faster than others



It's… messy

Connecting proteins to diseases:

- ✕ There are not simple rules we can follow.
- ✕ There's not enough direct data to train on.
- ✕ We cannot experiment on humans



Protein structure

?



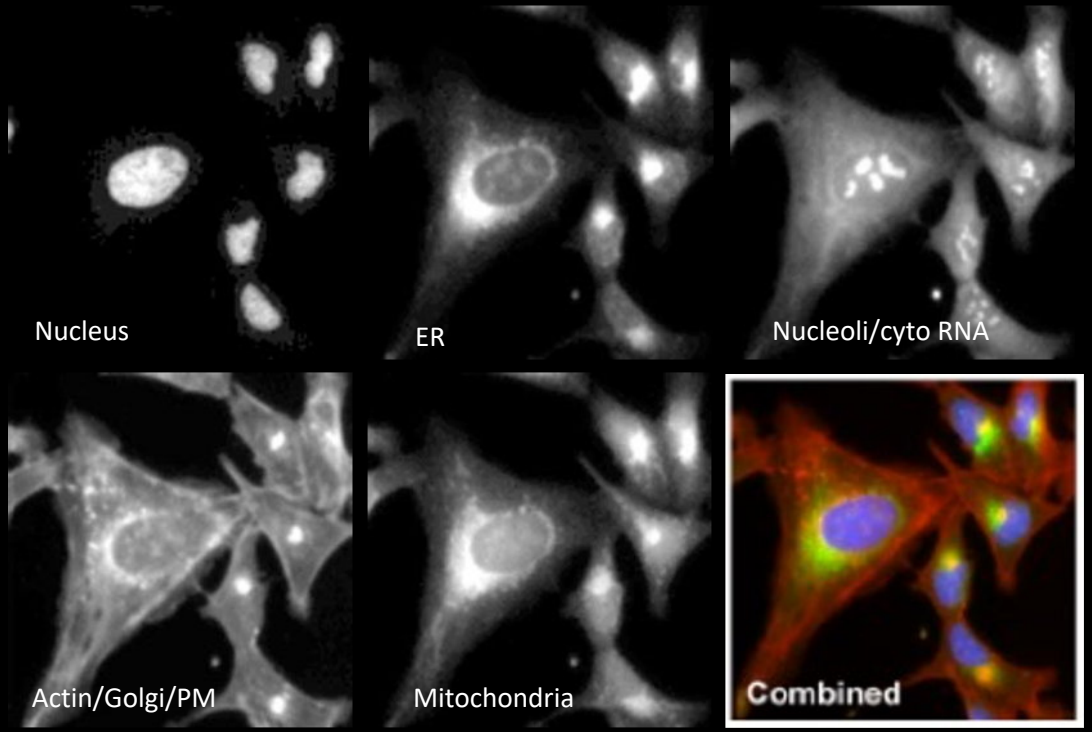
Cell phenotypes

Disease phenotypes



PURDUE UNIVERSITY

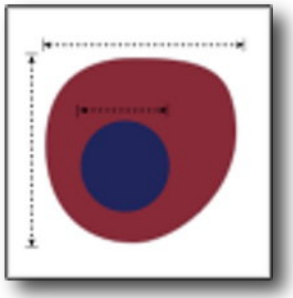
Images contain a wealth of information



1. single-cell resolution
2. complex model systems
3. multiplexed
4. quantitative

A revolution in how we use images

Measure known phenotypes



Profile to characterize samples







High-content screening

Image-based profiling
(using the Cell Painting assay)

PURDUE UNIVERSITY

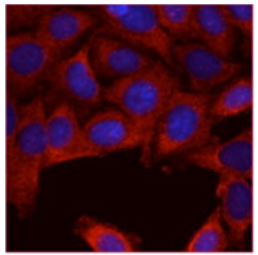
Extracting features from images

Sample prep



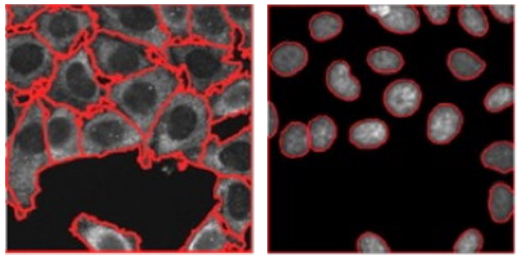
100's of 384-well plates

Microscopy



1000's of images/ plate; ~5 channels/image

Image analysis: Segmentation

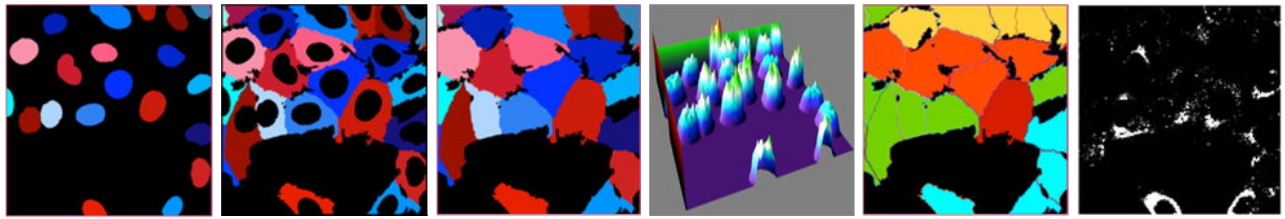


~500 cells/ image

CellProfiler™
cell image analysis software

Measure everything

Counts, Shapes, Sizes, Intensities, Textures, Correlations, Relationships



thousands of features / cell

60-80% of "high-content" studies use only 1 or 2 cellular features

Singh, et al. J Biomol Screen, 2014

CellProfiler: Open-source image analysis software

CellProfiler™
cell image analysis software

- Free and open-source; Windows, Mac, Linux
- Cited in **2,400+** papers per year, **20,000+** total
- Used in **most** pharma companies
- In the **Top 10** most popular papers in Genome Biology
- **Winner** of Bio-IT World's IT & Informatics Best Practices Award
- Ranked **most flexible** and **usable** in independent analysis (*Wiesmann et al.)*

Online Q&A! http://forum.image.sc

PURDUE UNIVERSITY

image.sc

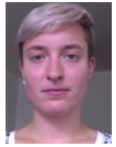

Anne Carpenter
2003-2006


Ray Jones
2007-2008


Lee Kamentsky
2009-2015


Allen Goodman
2015-2021


Beth Cimini
2016-now

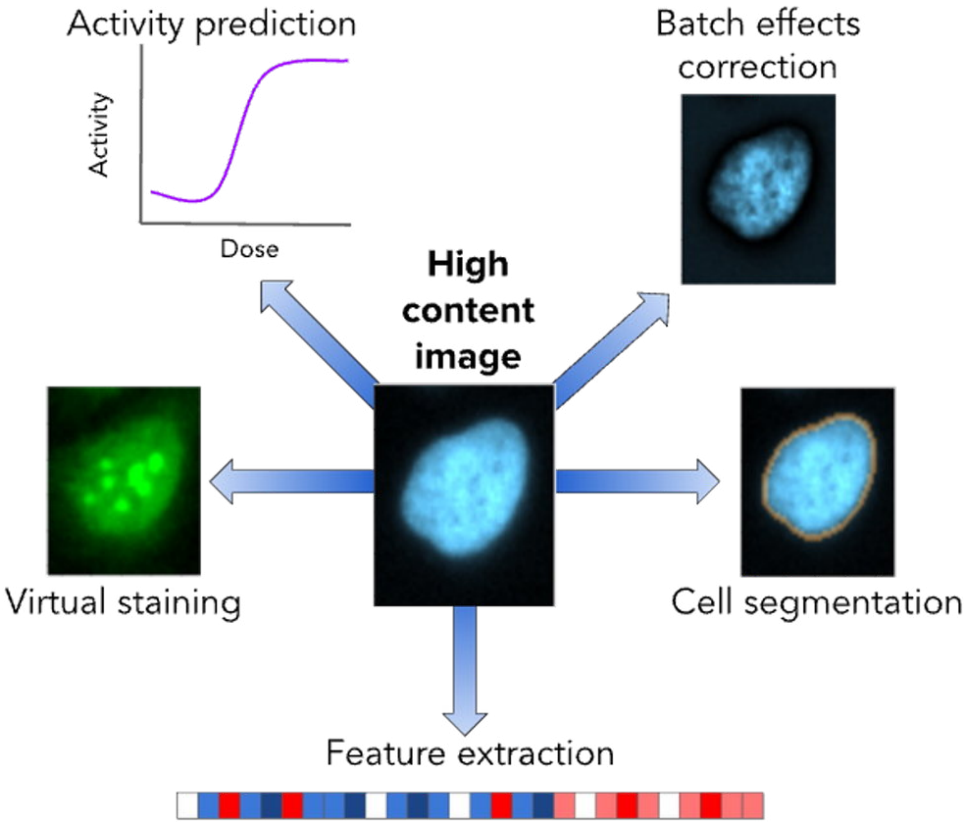

Claire McQuin
2017-2018


David Stirling
2019-2021


Alice Lucas
2020-now


Nodar Gogoberidze
2021-now

Impact of machine learning on image analysis



CellProfiler™
cell image analysis software

Cited in **2,400+** papers per year, **20,000+** total

OPIXIMI™

New! Deep learning for phenotype classification

image.sc

Online Q&A: **31,000+** posts/year



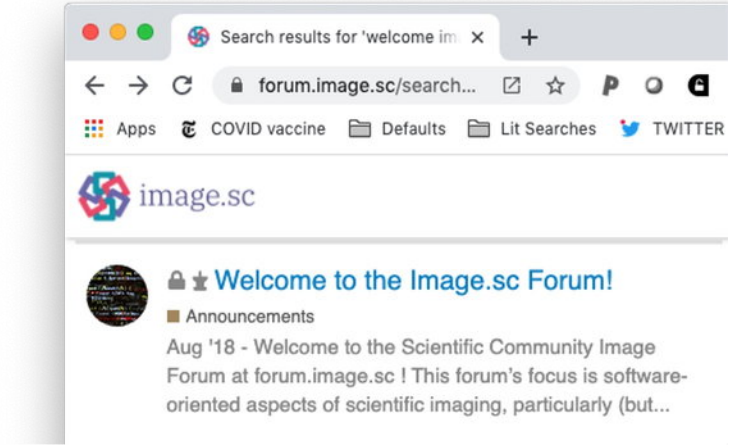
Beth Cimini's lab,
Broad Institute

PURDUE UNIVERSITY

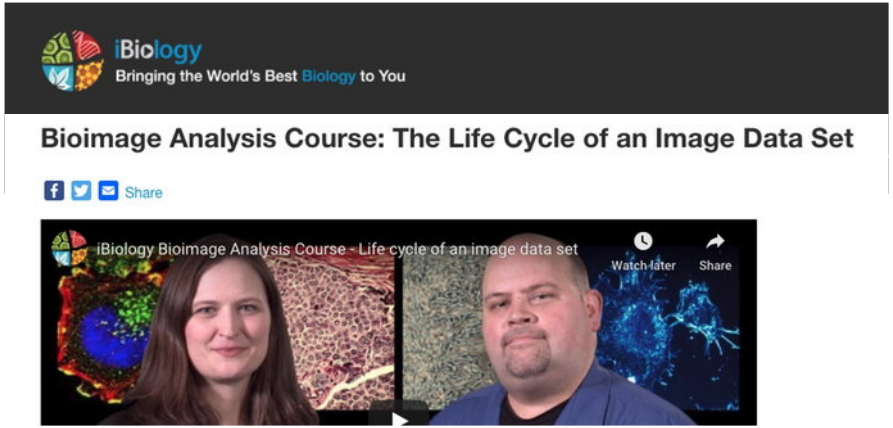
Caicedo and Cimini, Roadmap on Deep Learning for Microscopy arxiv 2023;
Lucas et al. Open-source deep-learning software for bioimage segmentation. Mol Bio Cell 2021

Just getting started in bioimage analysis?

The Scientific Community Image Forum:
https://forum.image.sc



iBiology video course:
https://www.ibiology.org/online-biology-courses/bioimage-analysis-course/



PURDUE UNIVERSITY

Download PDF of slides here: http://broad.io/CarpenterSlides

A revolution in how we use images

Measure known phenotypes

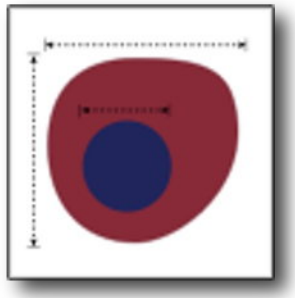


Image-based screening (High-content screening)

Beth Cimini lab

Profile to characterize samples

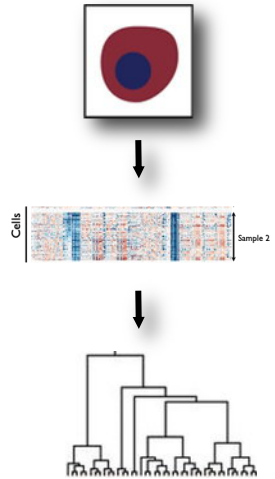


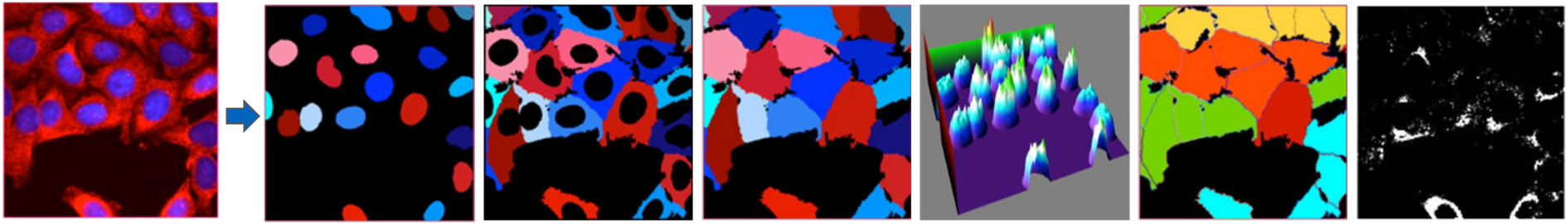
Image-based profiling (using the Cell Painting assay)

Carpenter–Shen lab

PURDUE UNIVERSITY

Measure everything

“Everything” = Counts, Shapes, Sizes, Intensities, Textures, Correlations, Relationships



~3,000 features / cell

CellProfiler™
cell image analysis software

Or, deep learning based feature extractors

Sample prep



hundreds of 384-well plates,
hundreds of thousands of samples

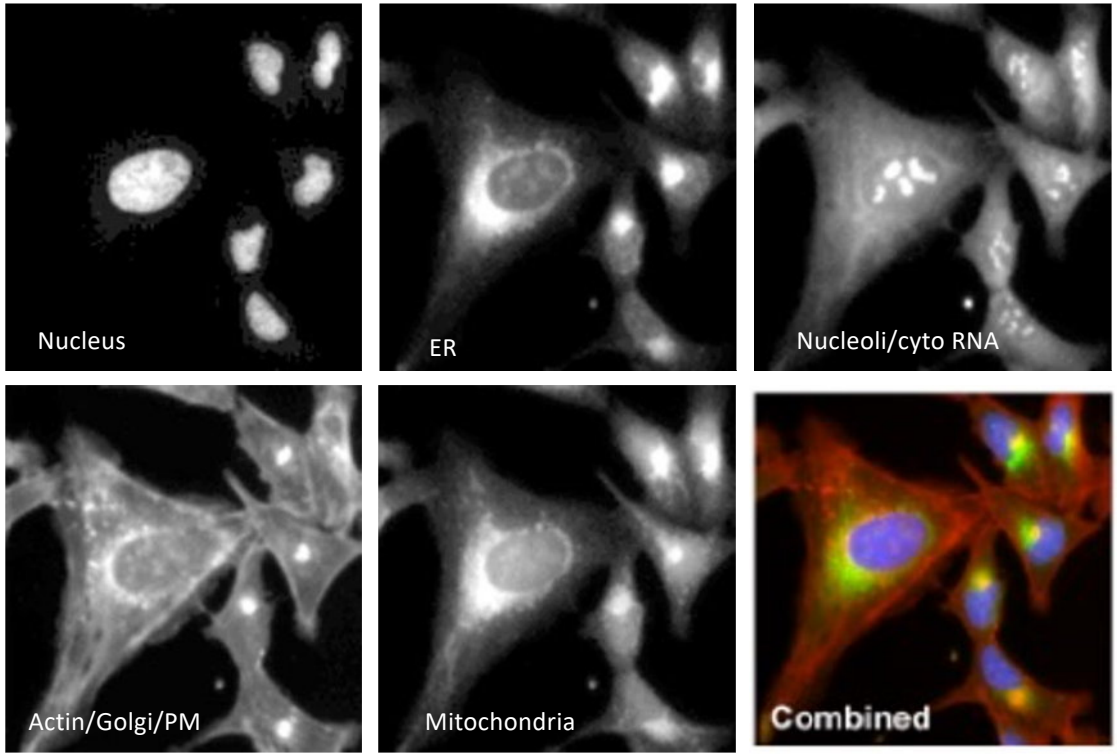
Why imaging?

10x cheaper than bulk mRNA and 1000-plex protein profiling

1000x cheaper than single-cell RNAseq

But similar predictive success!
And, complementary to mRNA/protein

The Cell Painting assay maximizes information for image-based profiling



6 stains imaged in 5 channels reveal 8 cellular components

Gustafsdottir, et al. PLOS ONE 2013
Bray, et al. Nature Protocols 2016
Cimini et al. Nature Protocols 2023


Kate Sokolnicki


Ray Jones


Maria Alimova


Kathy Hartland


Sigrun Gustafsdottir


Paul Clemons


Aly Shamji

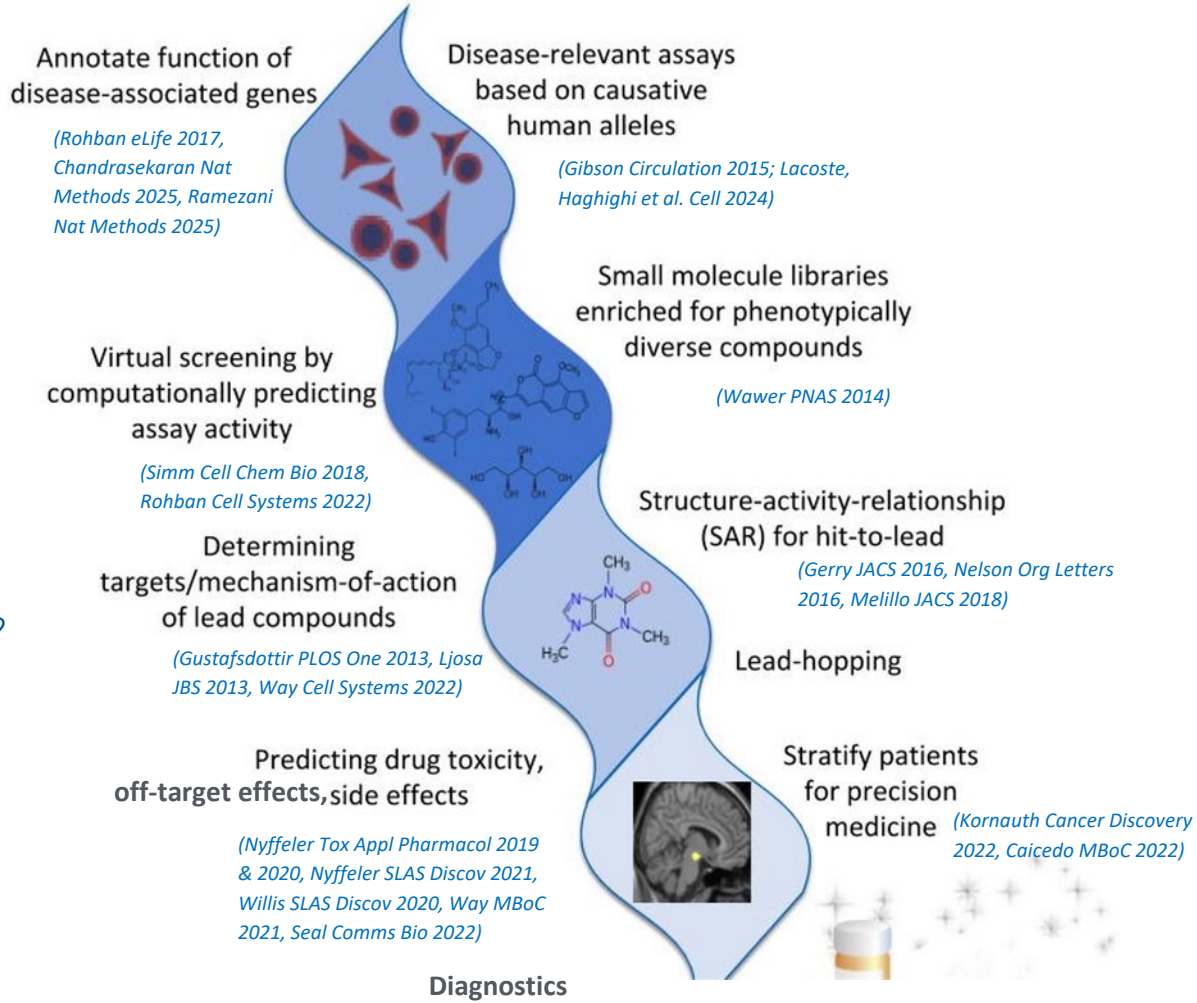

Stuart Schreiber, Broad Institute

PURDUE UNIVERSITY

Image-based profiling: accelerating many steps of drug discovery

Can we speed up individual steps using automation, clever biotech, AI?

Can we **surpass human capabilities**?



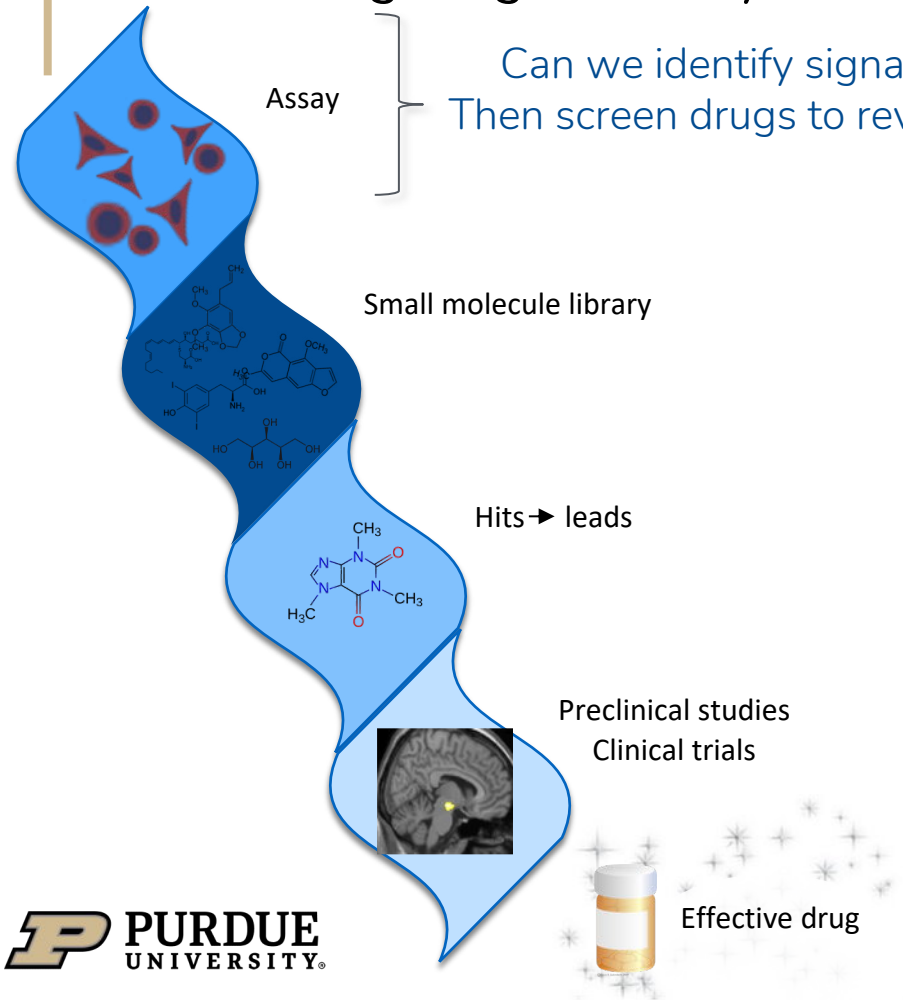
Applications reviewed in: Chandrasekaran, et al. Nat Reviews Drug Discovery (2020); Seal, et al. Nature Methods (2024)

Best practices reviewed in: Caicedo, et al. Nature Methods (2017)

Annual conferences: CytoData, SBI2

Download PDF of slides here: http://broad.io/CarpenterSlides

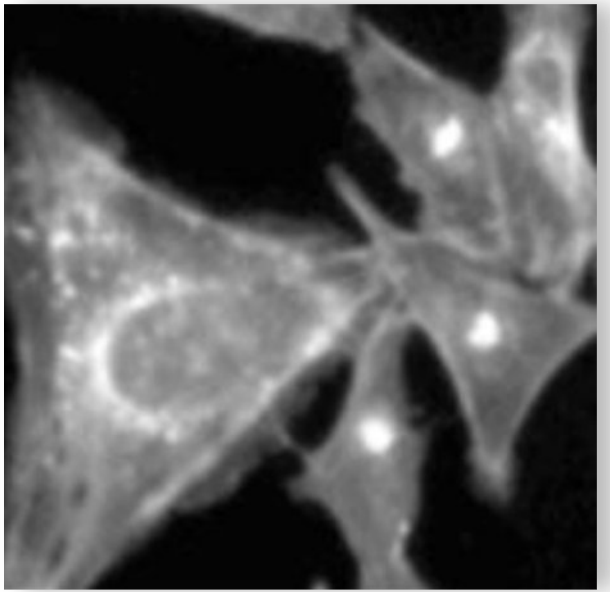
Transforming drug discovery via image-based profiling



Can we identify signatures of disease?
Then screen drugs to reverse the signature?

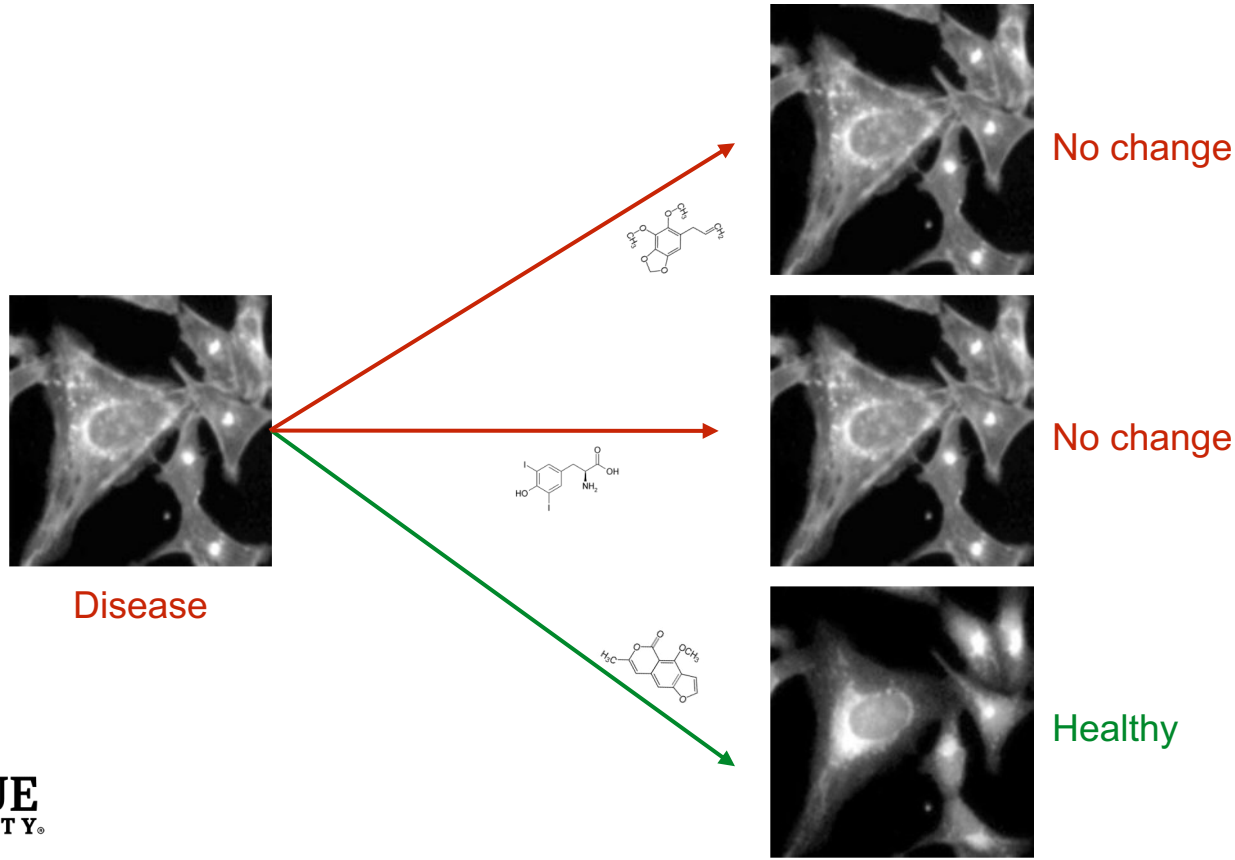
PURDUE UNIVERSITY

What if a diseased cell
could tell us
what's wrong with it?



PURDUE UNIVERSITY

If cell images can tell us they're diseased… then we could screen drugs at scale



PURDUE UNIVERSITY

Images do often faithfully reflect disease



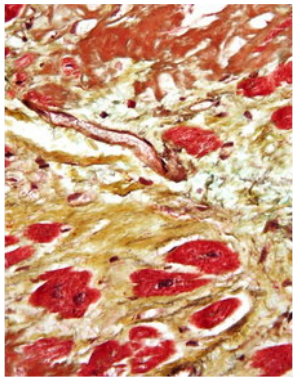
Prediction:Down's
Confidence:0.999



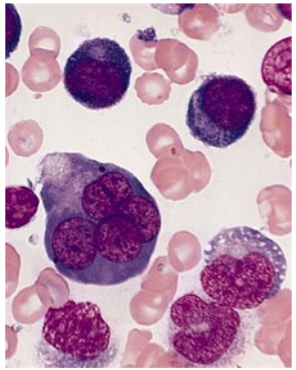
Prediction:Angelman
Confidence:0.997

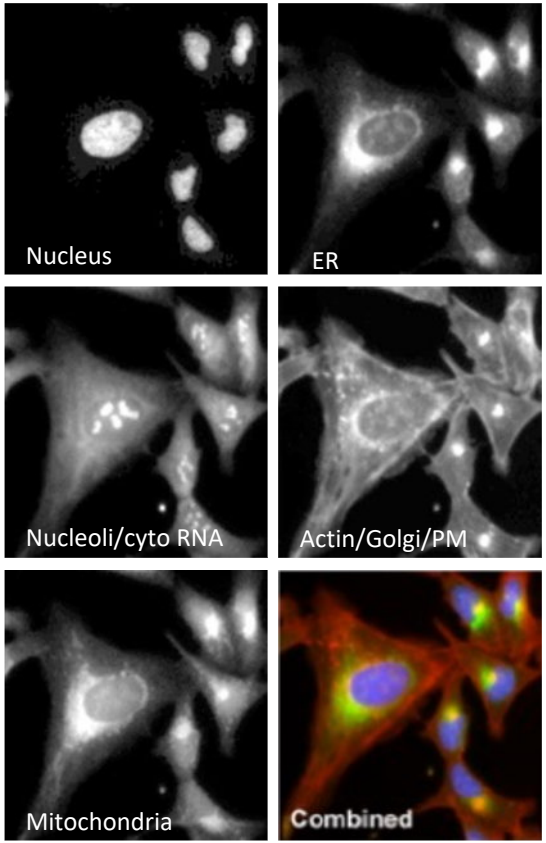
Singh & Kisku, IEEE ISED 2018 eLife
DB from Ferry, et al. 2014





Wikipedia



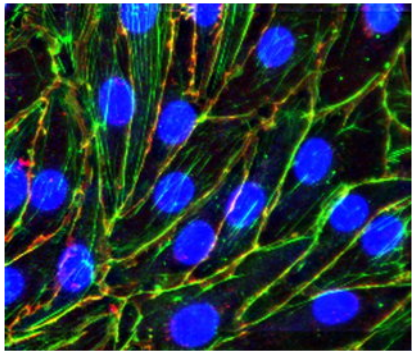


The Cell Painting assay

PURDUE UNIVERSITY

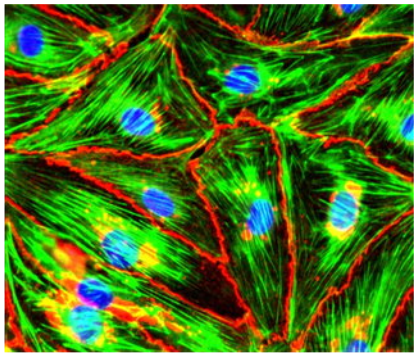
Can we find a phenotype for every disease?

Healthy



+ drug?

Disease (CCM knockdown)



Can we find drugs that reverse every disease phenotype?

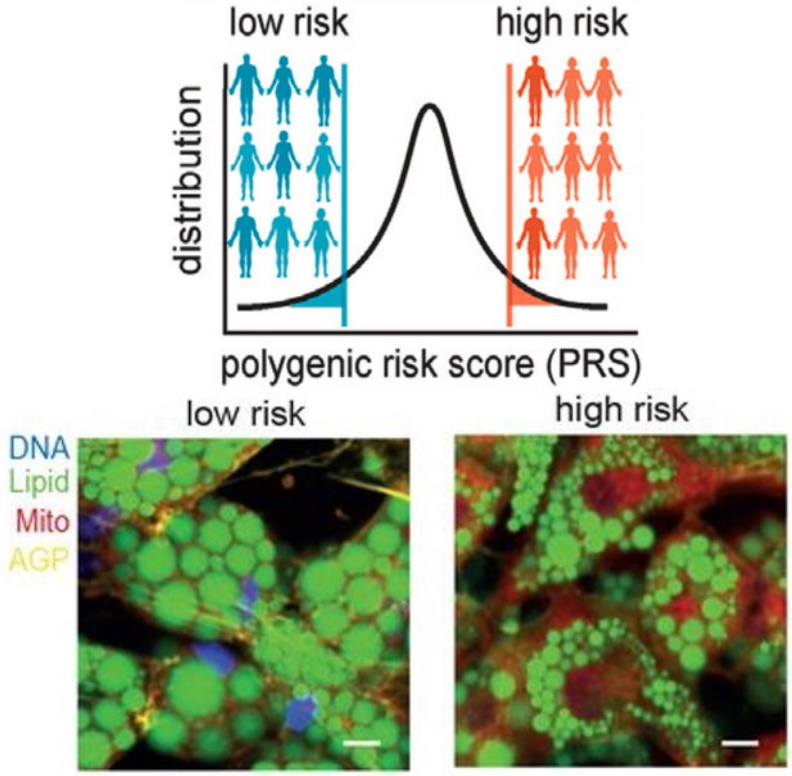
Disclosure: Anne Carpenter serves on Recursion's Scientific Advisory Board

Now: hundreds of disease models available for screening in parallel. And, five clinical trials launched!

PURDUE UNIVERSITY

Gibson, et al. Circulation 2015

Obesity example: patient cells at high vs. low risk
We can identify disease phenotypes using patient cell lines



LipocyteProfiler: Melina Claussnitzer lab

- ➔ Discovered cellular programs of intrinsic and extrinsic drivers of metabolic traits
- ➔ Found a non-coding variant associated with ‘Metabolically Obese Normal Weight’
- ➔ Found that pervasive pleiotropy mediates metabolic risk at a lipodystrophy-like metabolism locus

PURDUE UNIVERSITY

Melina Claussnitzer lab - Laber et al. Cell Genomics 2023 and Glunk, et al. Nat Metabolism 2023

Psychiatric example: patient cells with/without a condition


Marzieh Haghighi

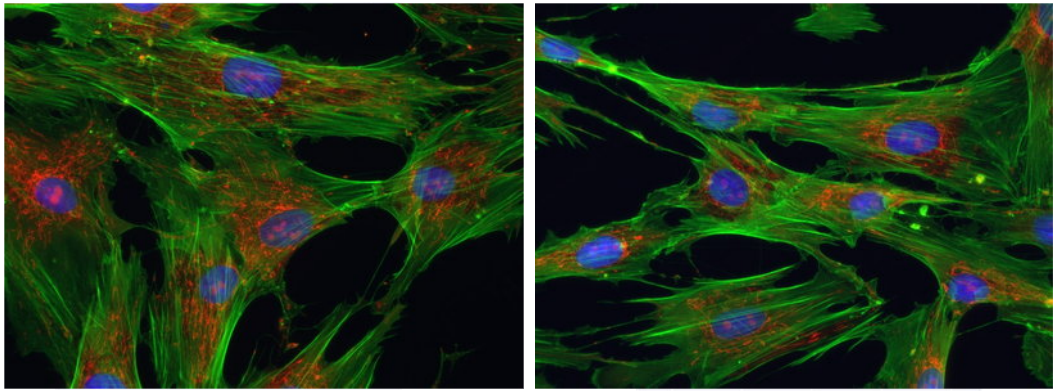

Mohammad Rohban


Kyle Karhohs

Step 1: Identify morphology change

(less mitochondrial dispersion)

Controls → Bipolar



Step 2: Test drugs to reverse the unhealthy morphology (NIH NCATS)


Bruce Cohen, McLean Hospital


Donna McPhie

PURDUE UNIVERSITY

Haghighi, McPhie, et al. Biorxiv 2025

Rare diseases present a major unmet need



That's 10%!
(330 million worldwide)



They cause 35% of deaths in the first year of life

10,000+
rare diseases



APPROVED TREATMENTS ARE AVAILABLE
FOR ONLY **5%** OF
ALL RARE DISEASES

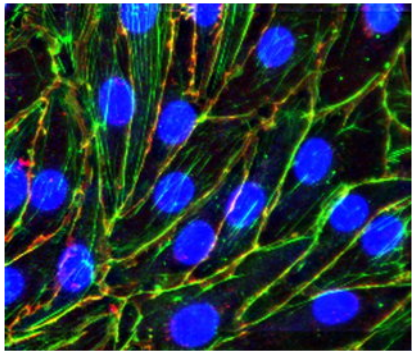
Current state of the field:

- Customize a protein or cell assay for the gene's function (~6 months) or do Cell Painting on patient cell lines, then screen novel chemicals to find a potential hit (~6 months)
- **Total cost $2 million per disease to get a hit** (estimate: Bay Bridge Bio)

Can we find a phenotype for every disease?

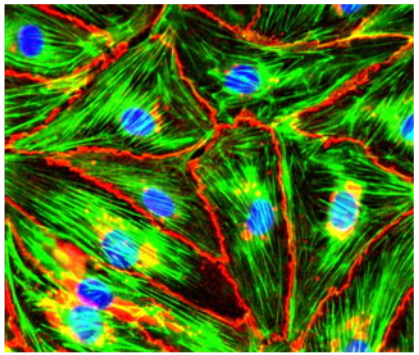
(more cheaply)

Healthy



existing + drug?

Disease



Can we find drugs that reverse every disease phenotype?

PURDUE UNIVERSITY

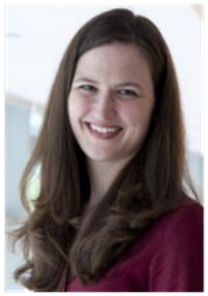
Project “VarChAMP”: Variant characterization across the Mendelian proteome



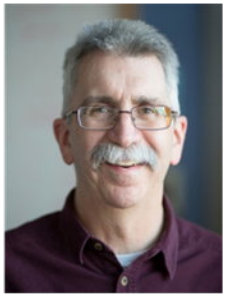
Marc Vidal



Mike Calderwood



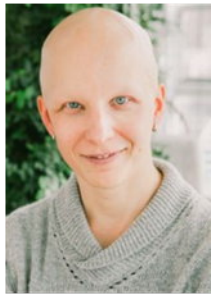
Anne Carpenter



Dave Hill



Fritz Roth



Mikko Taipale

Supported by a cast of dozens from these labs!

Current lead in Carpenter–Shen lab: Runxi Shen



BROAD INSTITUTE

Center for Cancer Systems Biology

PURDUE UNIVERSITY | Sinai Health | Lunenfeld-Tanenbaum Research Institute

Donnelly Centre Cellular & Biomolecular Research UNIVERSITY OF TORONTO

BLAVATNIK INSTITUTE HARVARD MEDICAL SCHOOL

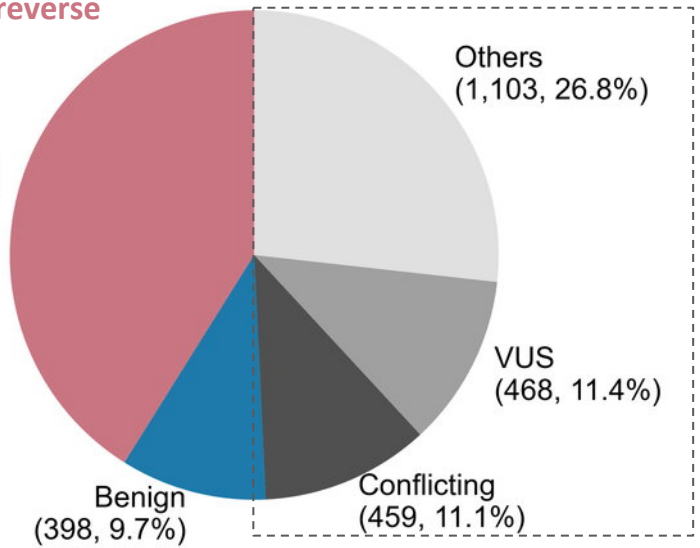
Dana-Farber Cancer Institute

VarChAMP aims to predict the impact of variants

- To guide patients and doctors
- To further research
- To identify phenotypes for drug screening

Why pathogenic?
What are their functional impacts?
Can we screen drugs to reverse these effects?

Are they pathogenic or not?



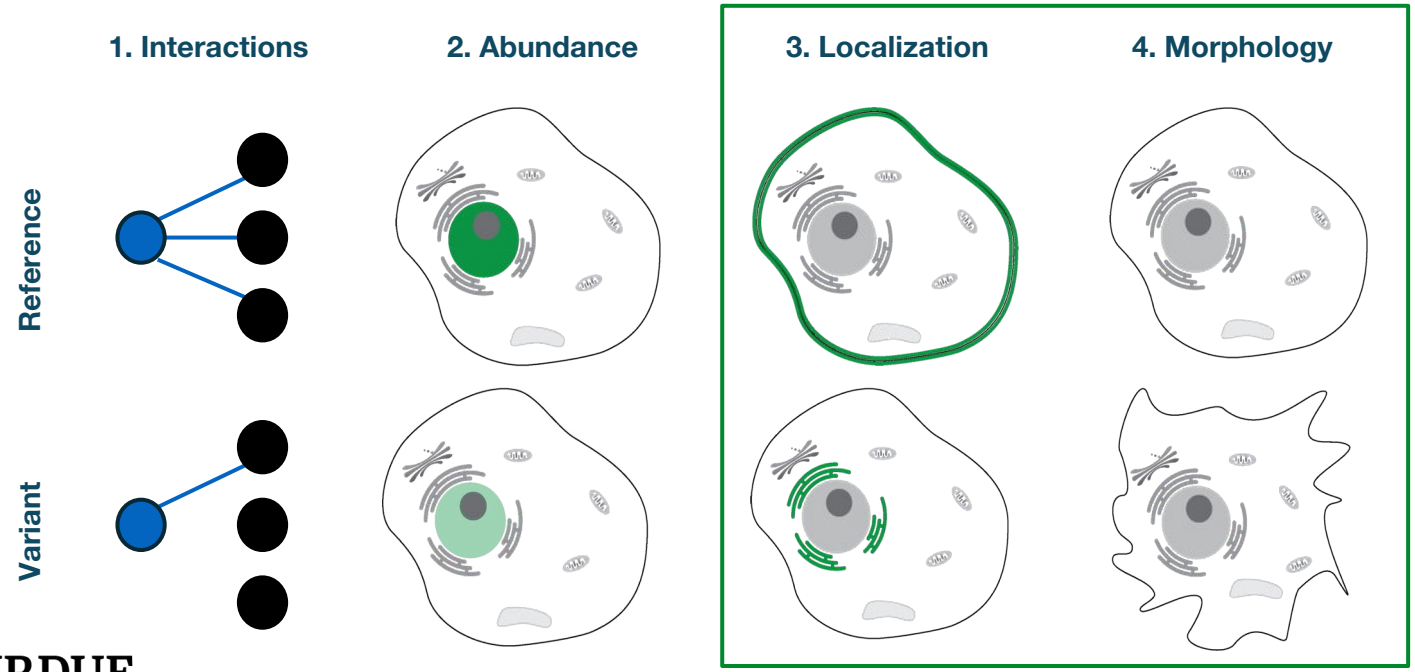
Project in progress

PURDUE UNIVERSITY

Capture the impact of coding variants at scale

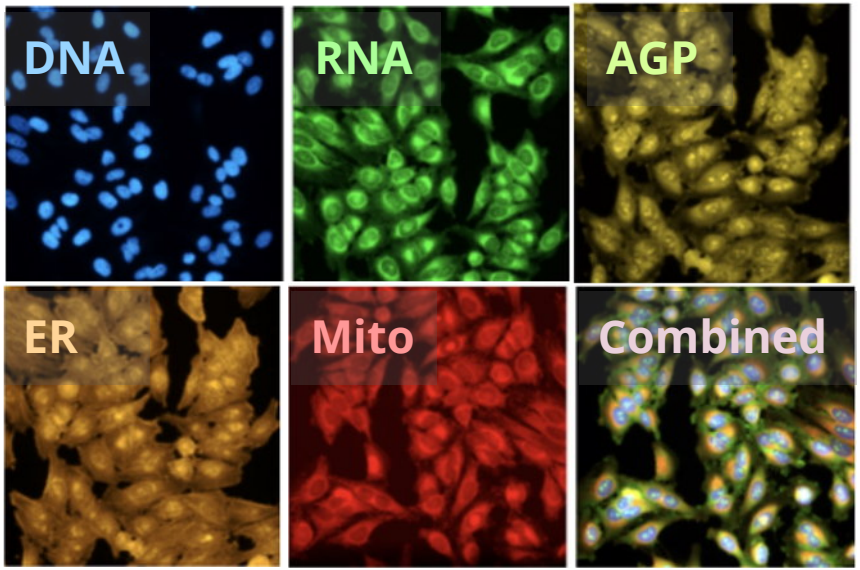
Goal: 30,000 coding variants in 3,000+ genes

Measure all the things!

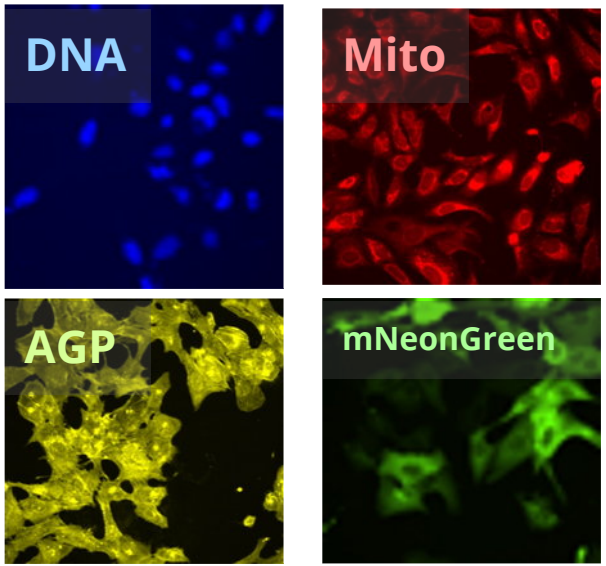


PURDUE UNIVERSITY

Variant Painting Assay can profile genetic variants



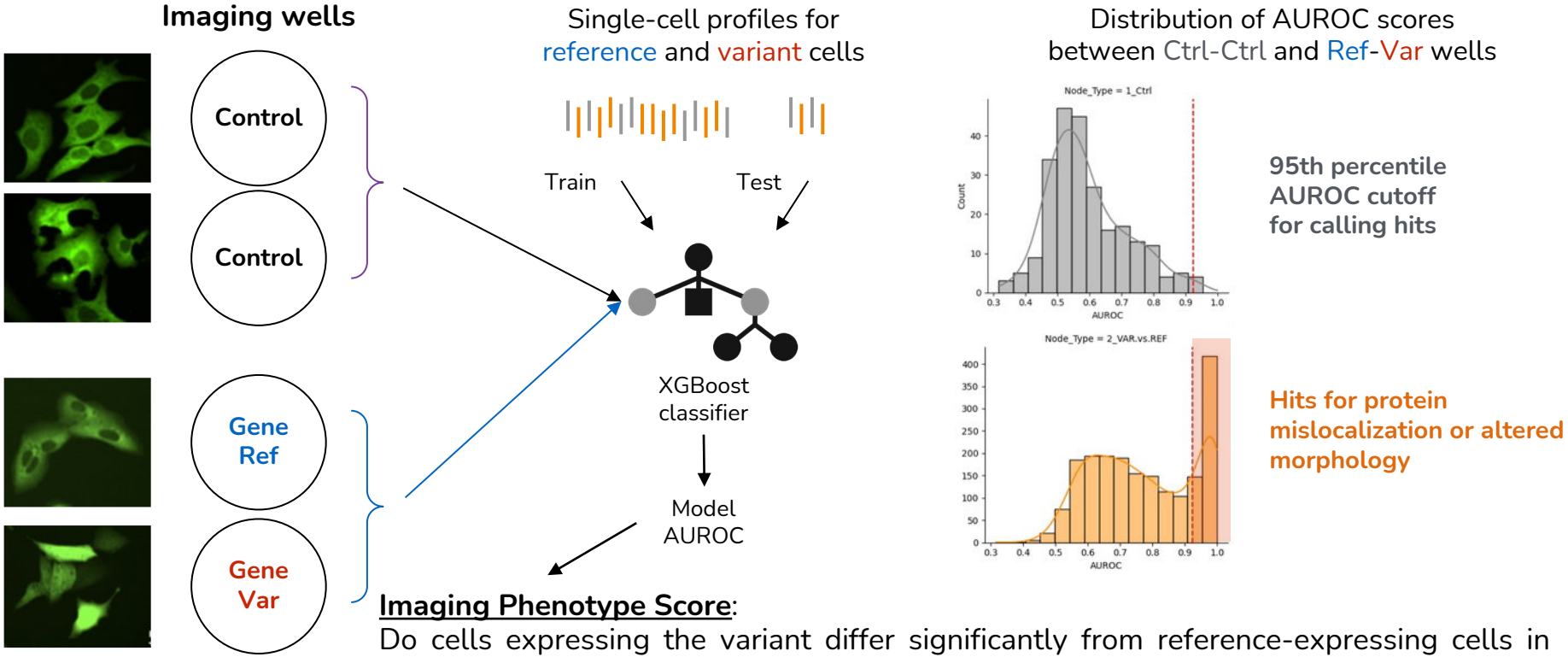
Standard Cell Painting
(general cell structures)



Variant Painting
(Morphology + Localization)

PURDUE UNIVERSITY

A machine learning workflow detects variant cells with imaging phenotypes

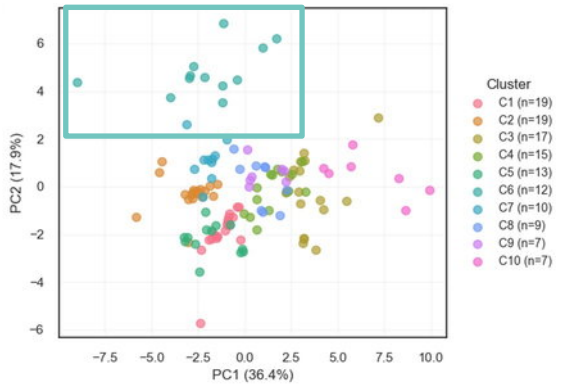


Imaging Phenotype Score:
Do cells expressing the variant differ significantly from reference-expressing cells in protein localization or morphology?

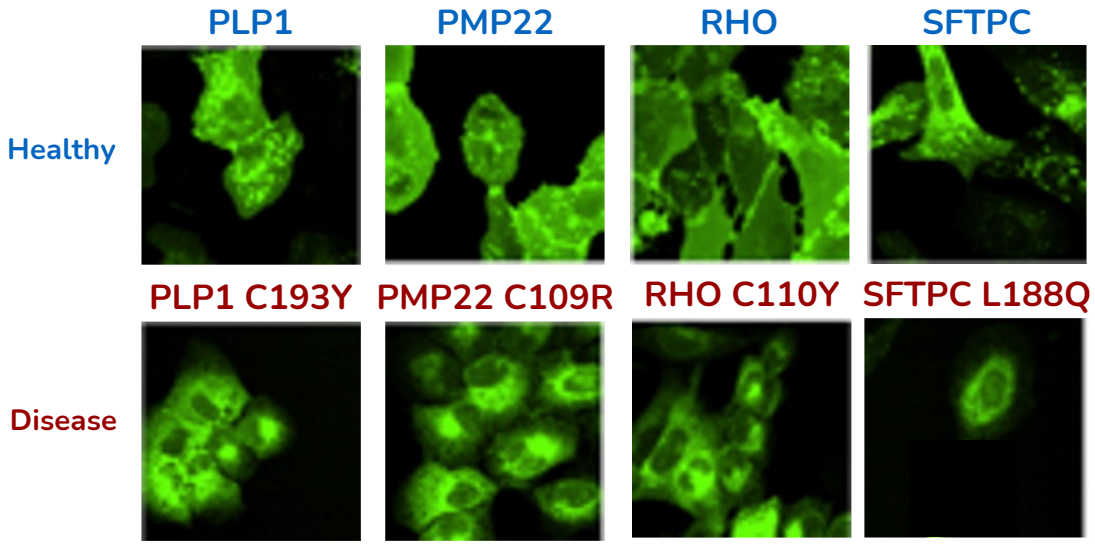
PURDUE UNIVERSITY

Jessica Ewald, Sam (Zitong) Chen

A shared protein mislocalization phenotype could suggest a unified therapeutic strategy



Shared protein mislocalization patterns were observed across different genes, cell types, and disease modules.



Could all these shared mislocalization patterns potentially be rescued with the same drug?

- ❑ **TMED9-targeted clearance (Dvela-Levitt et al., Cell 2019)**
- ❑ **A compound to release the misfolded proteins trapped in TMED9 cargo receptor-enriched compartments**

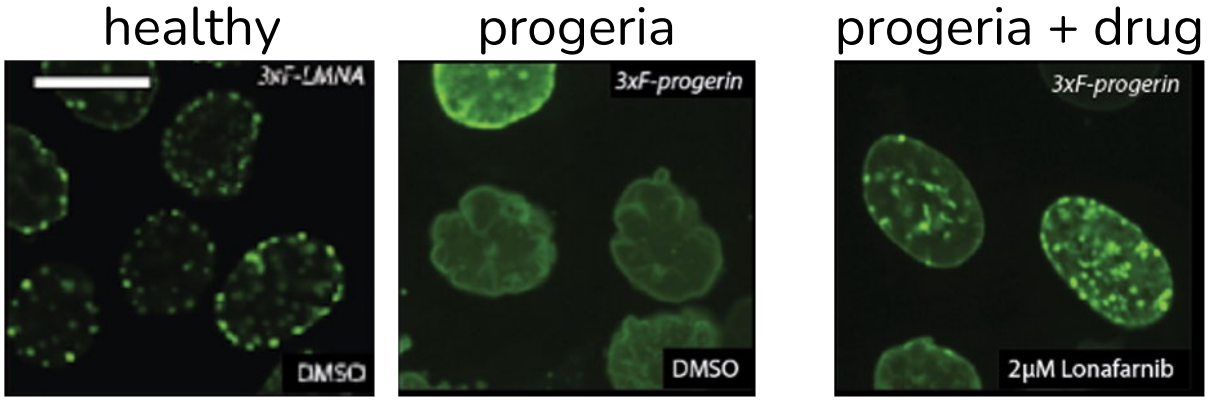
4,500 human genetic variants tested → 350+ disease phenotypes discovered

NIH IGVF Consortium: testing 30,000 new variants

VISTA Consortium: screen ~300 drugs x ~100 diseases

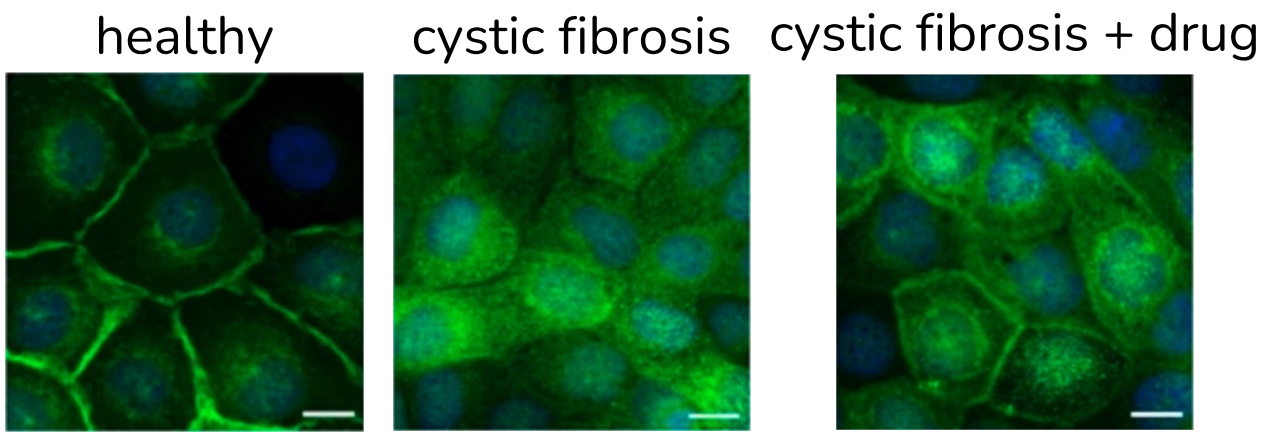
Some successful drugs reverse disease mislocalizations

Lonafarnib relocalizes LMNA in progeria syndrome (HGPS)



Lacoste, et al. unpublished

Lumacaftor relocalizes CFTR in cystic fibrosis



Capurro, et al. IJ Mol Sci 2021

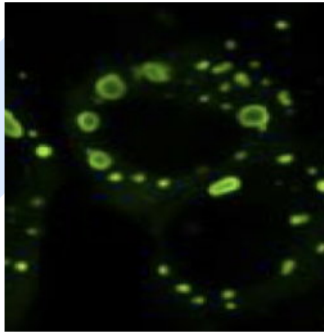
PURDUE UNIVERSITY

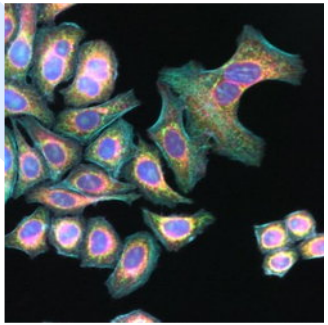
Scaling up drug discovery:
Optical pooled screening (via in situ sequencing)



Each well =
1 compound +
200 reference genes
200 disease variants

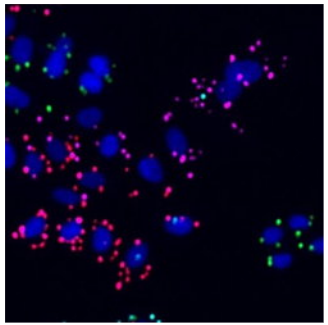
Localization & morphology





+

in situ barcoding




Shantanu Singh


Hamdah Abbasi


Beth Cimini


Jason Lim


Erin Weisbart


Marzieh Haghighi


JT Neal, Broad


David Feldman


Avtar Singh


Runxi Shen


Maria Lozada


Allie Brill


Julia Bauman


Celeste Diaz


Eddy Leardini


Jess Ewald


Paul Blainey, MIT


Sanam Kavari

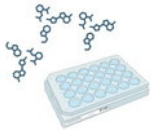

Varsha Prakash

PURDUE UNIVERSITY

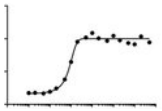
Project in progress, based on Pooled Optical Profiling: Feldman, et al. Cell 2019

Workflow for finding drug-disease hypotheses…

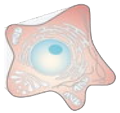
1) Primary screen (pooled)



2) Confirmation screen for hits (non-pooled, dose response, replicates)



3) Confirm in disease-relevant cell lines



4) Engage the rare disease community of patients, clinicians, and researchers



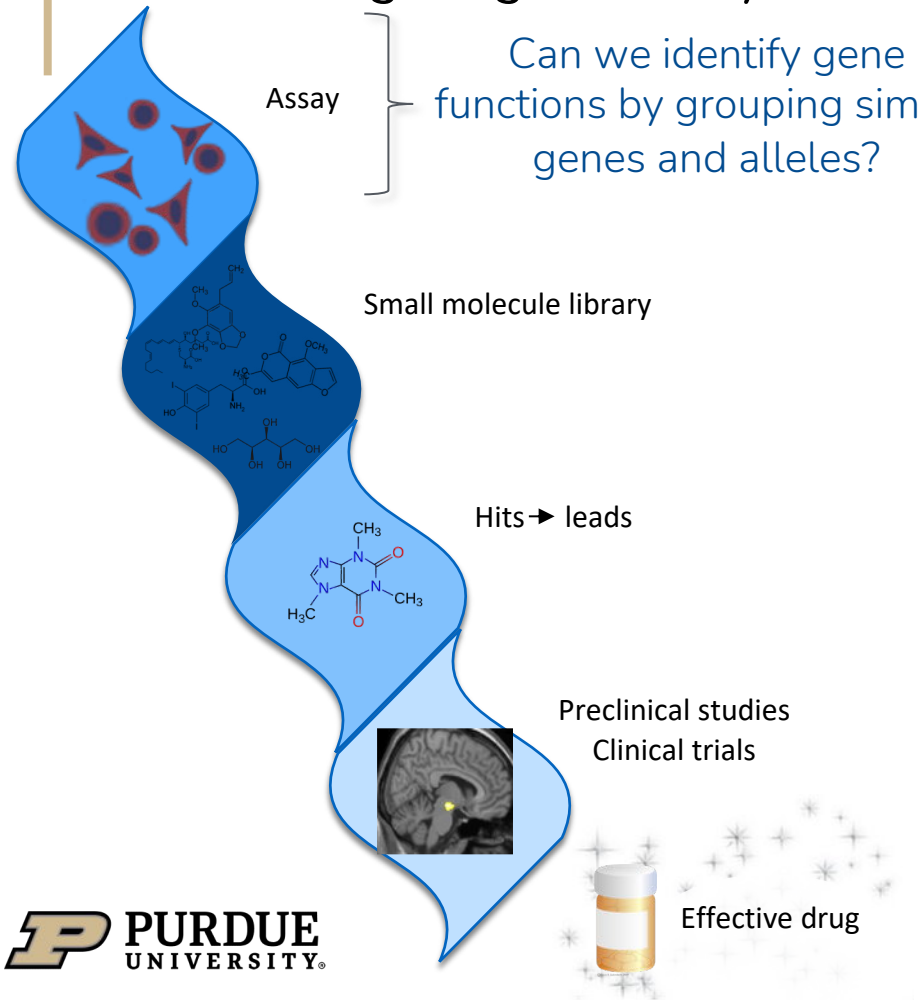




- Cell-based experiments (disease-relevant cell types or assay readouts)
- Animal experiments
- Human clinical trials/off-label use
- Encourage pharma investment for diseases with shared mislocalization

PURDUE UNIVERSITY

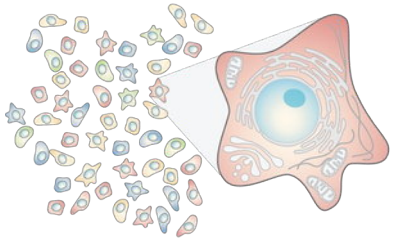
Transforming drug discovery via Cell Painting



PURDUE UNIVERSITY

JUMP-Cell Painting Consortium dataset is now available

(Joint Undertaking for Morphological Profiling)



- **13,000 overexpressed genes**
- **8,000 CRISPR knockouts**
- **116,000 (+ 30,000) small molecules**

5 replicates each in human U2OS cells

1.6 billion single cells *(mRNA: scBaseCount - 300 million cells)*

2,378 plates (384 well) = 913,152 wells

200 TB of images, similar amount of extracted features

Partners:
- Broad Institute
- Amgen
- AstraZeneca
- Bayer
- Biogen
- Eisai
- Janssen
- Ksilink
- Merck Healthcare KGaA (Germany)
- Pfizer
- Servier
- Takeda

Supporting Partners:
- Nomic bio
- PerkinElmer
- Verily
- Google Research
- Horizon Discovery
- Ardigen

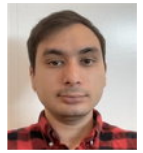
Collaborators:
- Pistoia Alliance
- Umea University
- Stanford University

JUMP Hub to access the data, exploration tools, and learn more:
http://broad.io/jump

Data Storage at the Cell Painting Gallery supported by the AWS Open Data Sponsorship Program

Chandrasekaran, et al. Biorxiv 2024

JUMP Cell Painting data and tools now online!



Alán Fernando Muñoz González

Try them yourself:
broad.io/jump → Interactive tools

For any gene or chemical perturbation:

- **How do the cells look?**
- **Which other profiles are the most similar?**
- **Which are its most distinctive morphological features?**

home / data

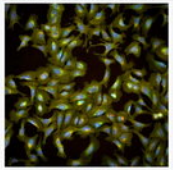
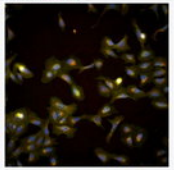
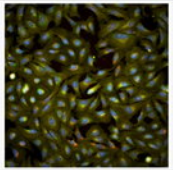
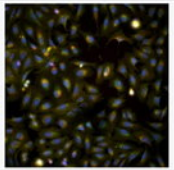
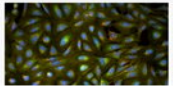
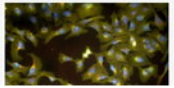
ORF Matches

756,550 rows

- column - = Apply

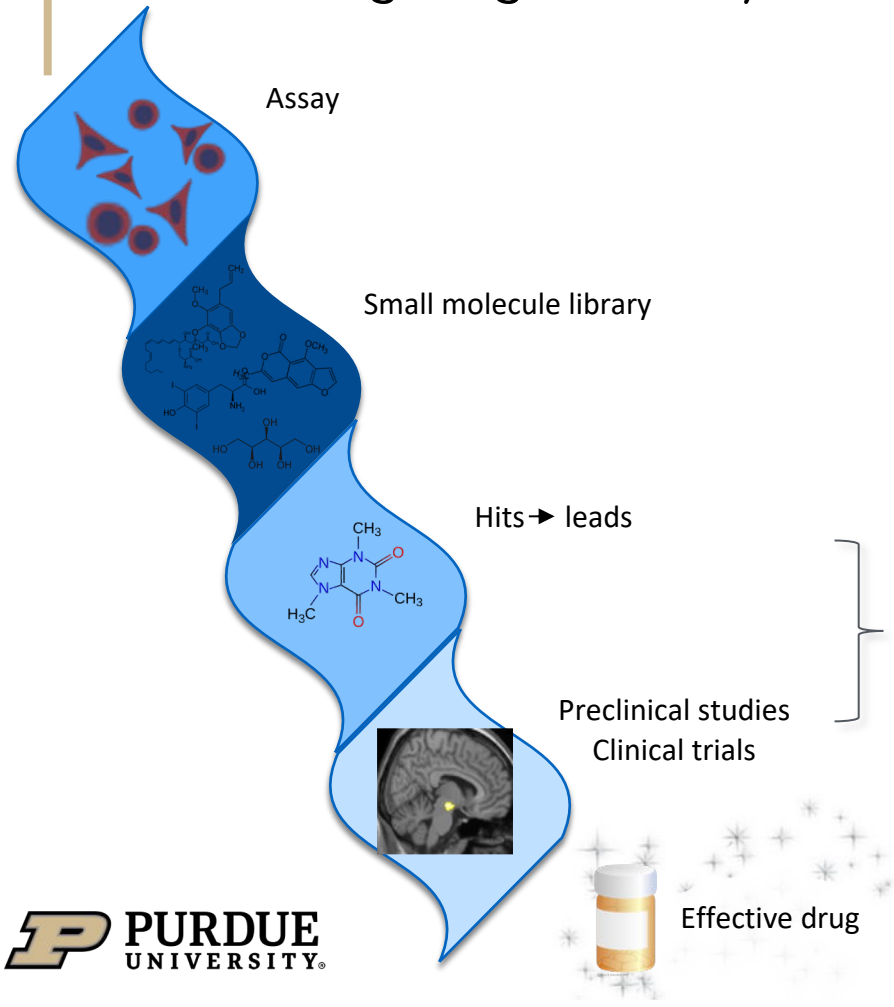
View and edit SQL

This data as json, CSV (advanced)

Link	rowid ▼	index	Gene/Compound	Match	Gene/Compound Example	Match Example	Similarity	Match resources	JCP2022
1	1	0	HLA-DRB1	TCP11L2			0.30702412128448486	External	JCP2022_910143
2	2	1	HLA-DRB1	TMEM208			0.3065337538719177	External	JCP2022_910143
3	3	2	HLA-DRB1	UBE2S			0.2936822175979614	External	JCP2022_910143

PURDUE UNIVERSITY

Transforming drug discovery via Cell Painting



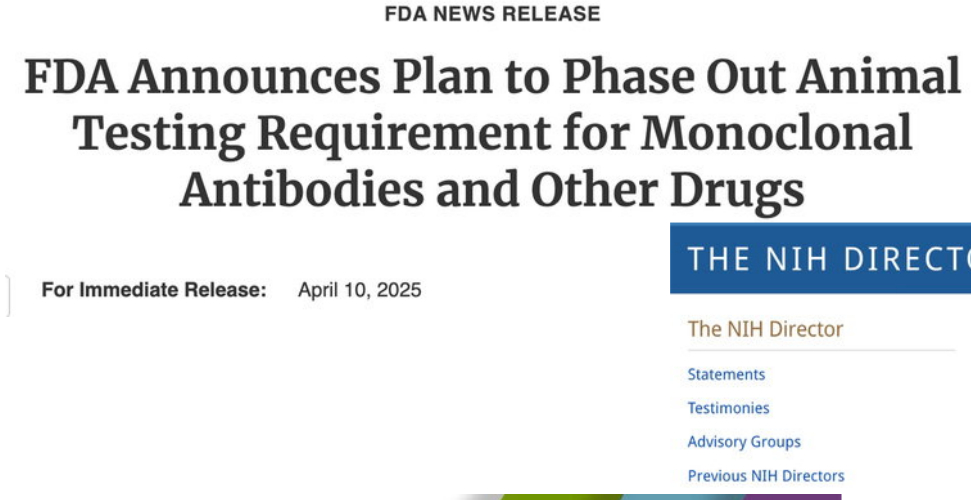
Can we identify various types of toxicity?

PURDUE UNIVERSITY

Can we predict chemical toxicity from images?

30% of Phase 1 and 25% of Phase 2 clinical trials fail due to tox missed by animal models

300,000 commercial compounds have never been assessed for safety!



FDA NEWS RELEASE

FDA Announces Plan to Phase Out Animal Testing Requirement for Monoclonal Antibodies and Other Drugs

For Immediate Release: April 10, 2025



THE NIH DIRECTOR

The NIH Director

Statements

Testimonies

Advisory Groups

Previous NIH Directors

February 1, 2024

Statement on catalyzing the development of novel alternative methods



NIH National Institutes of Health
Office of Strategic Coordination–The Common Fund

Search Common Fund

Our Programs | Research Funding | News & Media | New Program Planning | Evaluation & Assessment

Complement Animal Research In Experimentation (Complement-ARIE) Program

OASIS Consortium for toxicity assessment

OASIS
Omics for Assessing Signatures for Integrated Safety Consortium

Public and private compounds with associated in vivo and/or human data

- 200 donated from industry partners
- 1400 from public domain

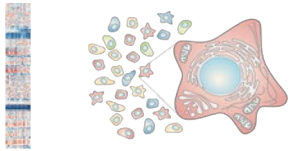
Cell model systems

(of increasing complexity, physiological relevance, and cost)

- U2OS human osteosarcoma immortal cell line
- HepaRG human hepatocyte-like cell line
- Primary rat hepatocytes
- Primary human hepatocytes
- Liver organoids, Liver-on-a-chip

Multi-omic profiling assays

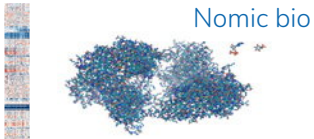
Cell Painting (images)



Transcriptomics (mRNA)



nELISA (protein)



Data analysis

- Translating *in vitro* to *in vivo* concentrations (qIVIVE)
- Distinguishing & interpreting mechanisms
- Integrating evidence across multiple data types/cell models

Publicly accessible data

Published methodologies

- Community of Practice



PURDUE UNIVERSITY

Rouquié, et al. Tox Sci 2025

Cell Painting detects bioactivity for more compounds and at lower concentrations than cytotoxicity assays

Axiom
www.axiombio.ai

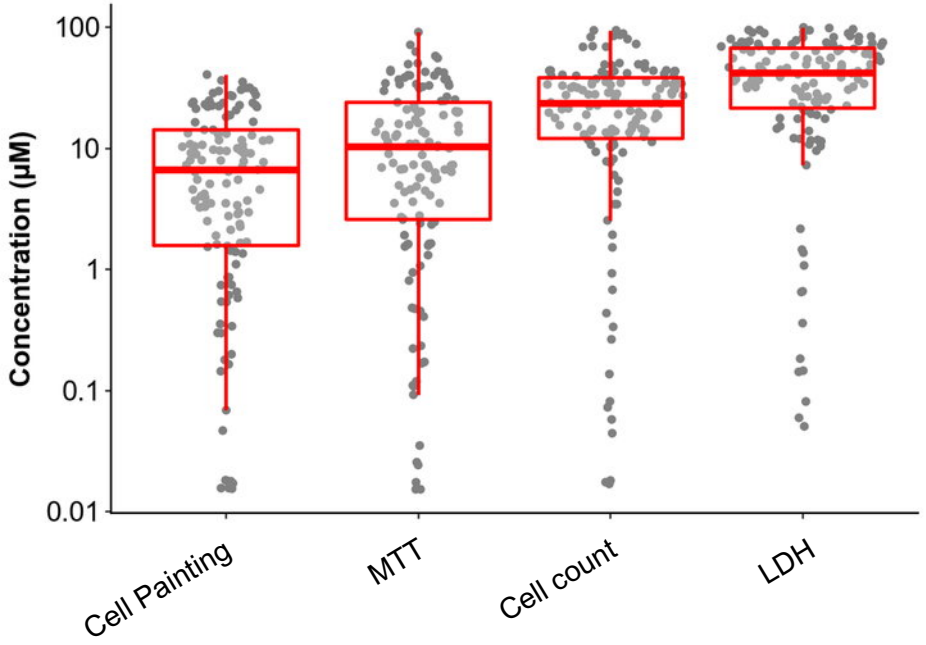


Jess Ewald

- Primary human hepatocytes
- 1085 compounds (8 doses, 2 replicates)
- Three assay readouts:
  - Biochemical assays (MTT and LDH)
  - Cell Painting assay (including cell count)

Assay	Number of bioactive compounds detected
Cell Painting (CellProfiler)	**607** (56%)
MTT	**431** (40%)
Cell count	**220** (20%)
LDH	**147** (14%)

Dose where chemical effect is seen
(lower is better)

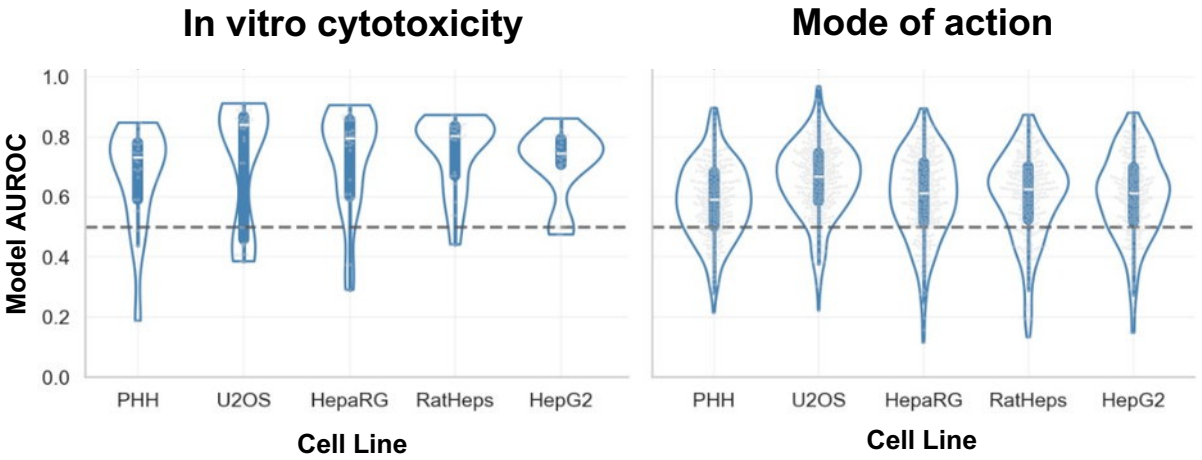


PURDUE UNIVERSITY

Ewald, et al. Cell Systems 2026

Morphology can predict cytotoxicity and drug modes of action across cell lines

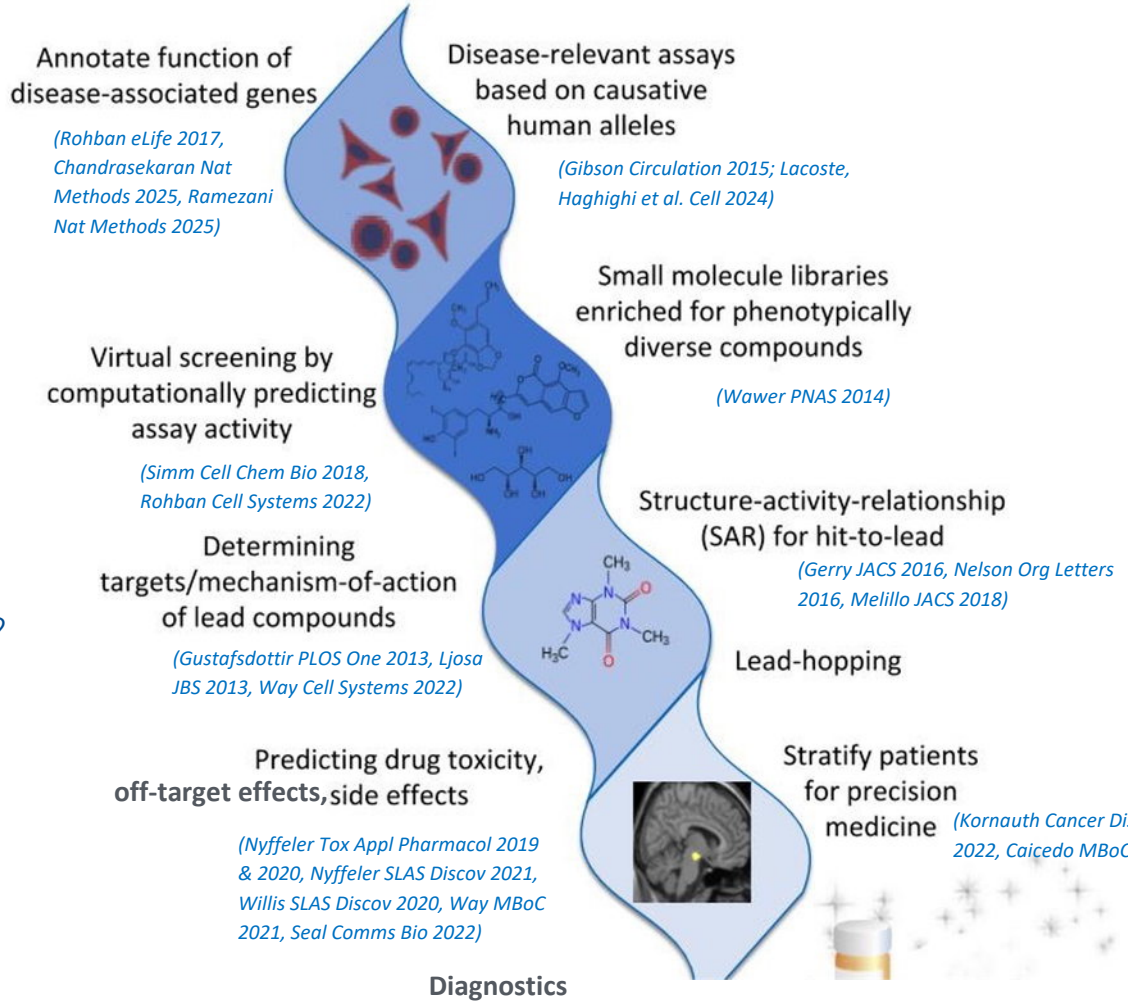
Cell Line	Number of bioactive compounds detected by morphology
PHH	607/1085 (56%)
U2OS	678/1453 (47%)
HepaRG	730/1453 (50%)
HepG2	145/360 (40%)
RatHeps	356/628 (57%)



PURDUE UNIVERSITY

Shen, et al. Manuscript in preparation

Image-based profiling: accelerating many steps of drug discovery



Can we speed up individual steps using automation, clever biotech, AI?

Can we **surpass human capabilities**?

Applications reviewed in: Chandrasekaran, et al. Nat Reviews Drug Discovery (2020); Seal, et al. Nature Methods (2024)

Best practices reviewed in: Caicedo, et al. Nature Methods (2017)

Annual conferences: CytoData, SBI2

Download PDF of slides here: http://broad.io/CarpenterSlides

Carpenter–Shen Lab


Anne E. Carpenter


Runxi Shen


Shantanu Singh


Alán Fernando Muñoz González


Jonne Rietdijk


Johan Fredin Haslum


Kazumasa Shimofuruta

Many thanks to two decades of Carpenter–Singh lab alumni and to our many collaborators, especially in Beth Cimini's lab!

Hiring at Purdue University!
Postdoc openings in data science/ML
https://broad.io/mlcbpostdoc

Recent major funding for this work provided by:
NIH NIGMS P41 GM135019
NIH NIGMS R35 GM122547 MIRA
Mass Life Sci Center

Disclosure: We receive honoraria for talks at various companies and serve as scientific advisors or founders for:
- Anne: Recursion, SyzOnc, Quiver Bioscience
- Shantanu: Dewpoint Therapeutics, Deepcell, Waypoint Bio

PURDUE UNIVERSITY

S.D.G.

Download PDF of slides here: http://broad.io/CarpenterSlides