



# High-Content Screening - *Big image data meets big pharma*

Predictive drug discovery using quantitative image analysis, organotypic cell systems and phenotypic screening

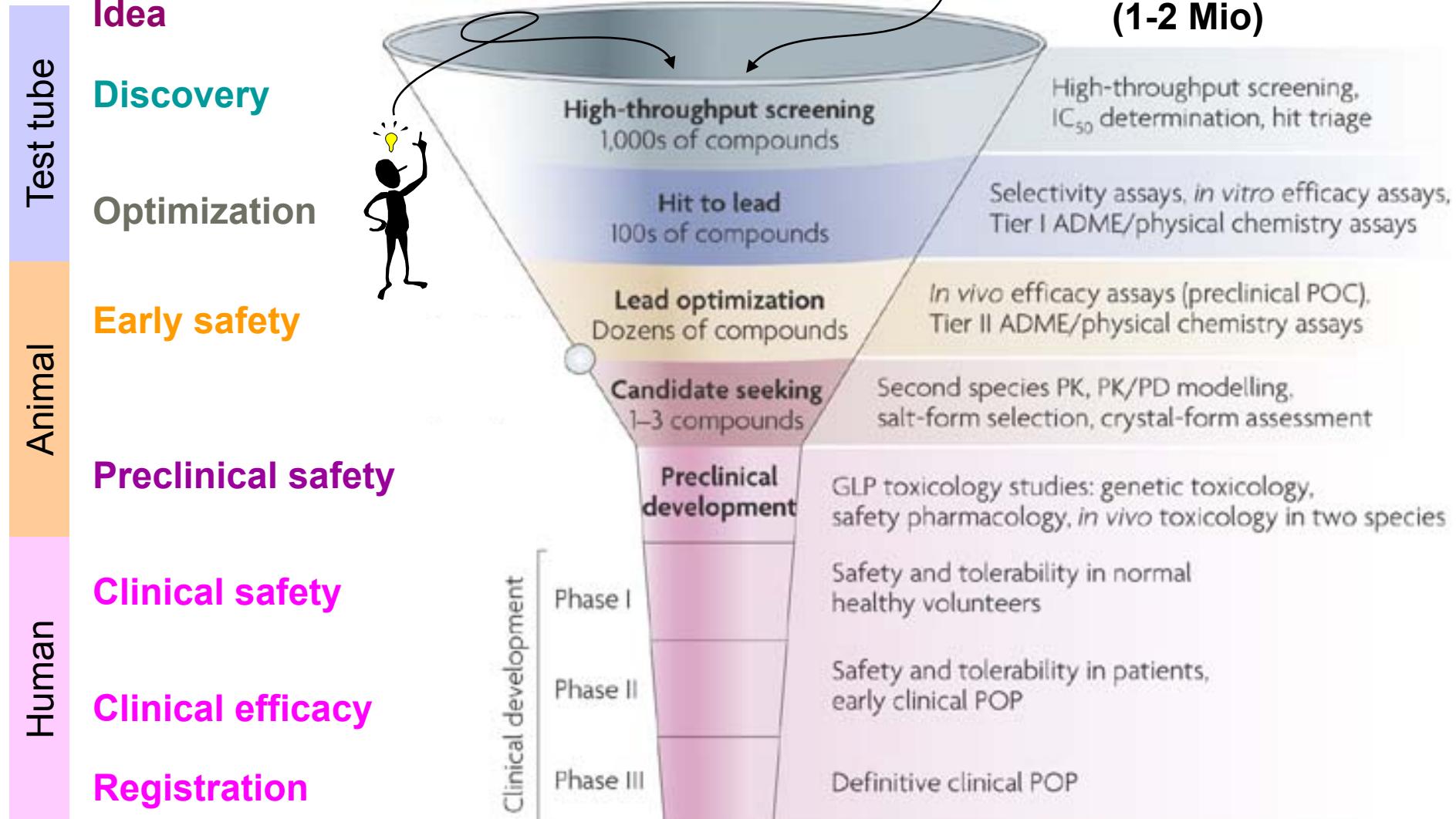
Michael Prummer, Biostatistician, NEXUS, ETH since April 2015  
2007-2015: Senior Scientist Discovery Technologies, F. Hoffmann-La Roche Ltd., Basel

# Disclaimer

This presentation may contain copyrighted material.

Do not distribute

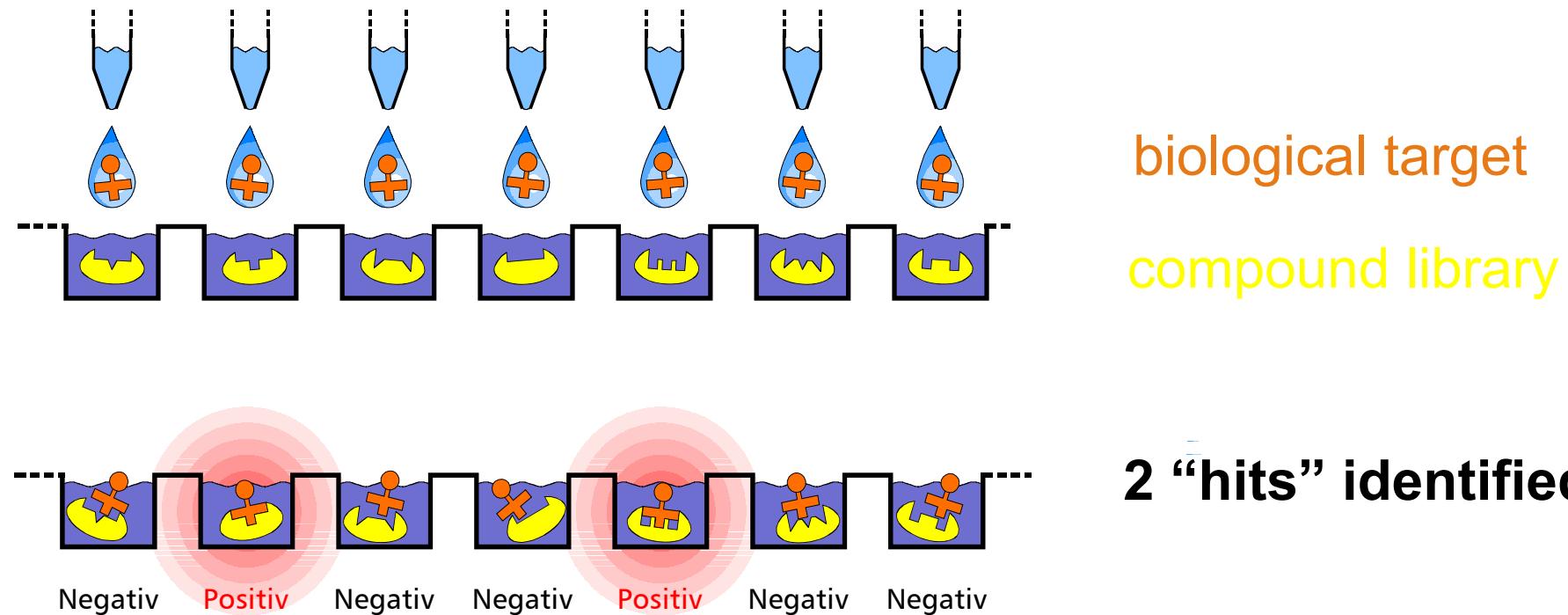
# The drug discovery pipeline



# What is High Throughput Screening (HTS)?

## Well-Well-Well-...

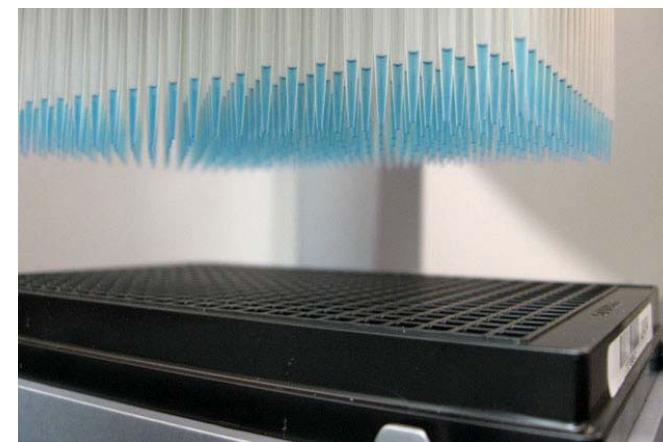
The systematic testing of a compound library against a biological target



High Throughput Screening (HTS): 100'000 or more samples per day

# What is an assay?

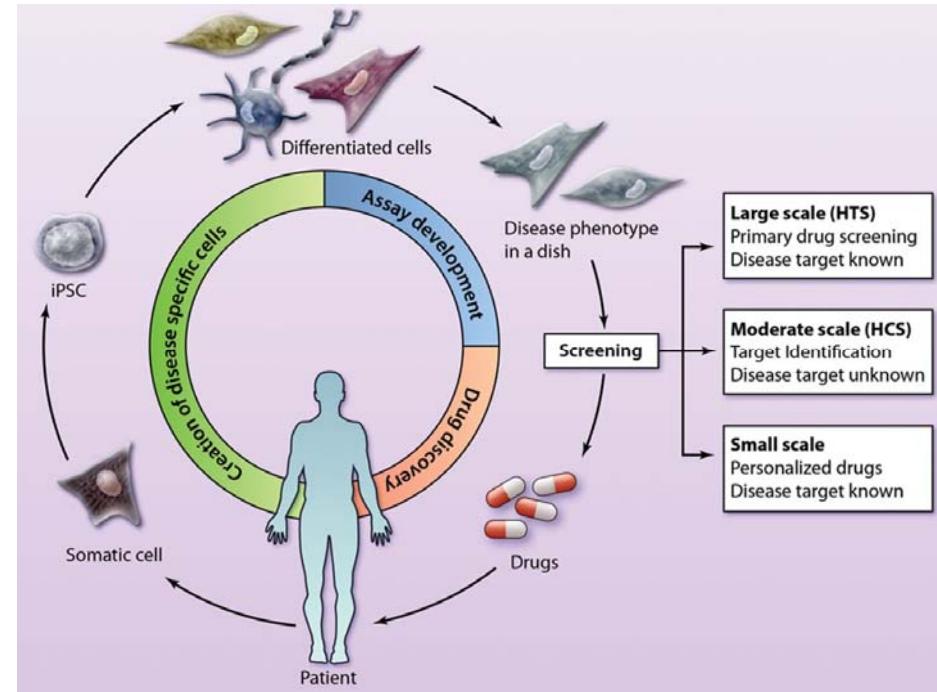
- An experiment run according to a protocol giving the same result every time and everywhere
  - ... provided the protocol is formulated to the required detail
  - ... provided the protocol is followed strictly



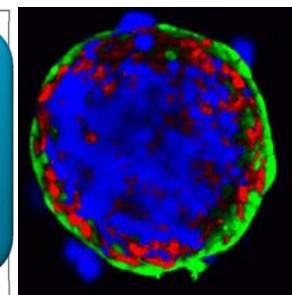
# Current bottlenecks in drug discovery

## *Possible improvements*

- **More predictive models**
  - Patient-derived or genome-edited reprogrammed cells instead of overexpressing immortalized cell lines
  - 3D co-culture organotypic systems instead of 2D mono culture
  
- **More predictive assays**
  - Phenotypic cell-based as opposed to target-based biochemical

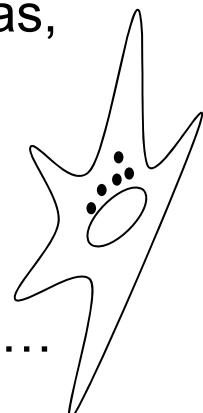


The modality of choice:  
High-Content Screening

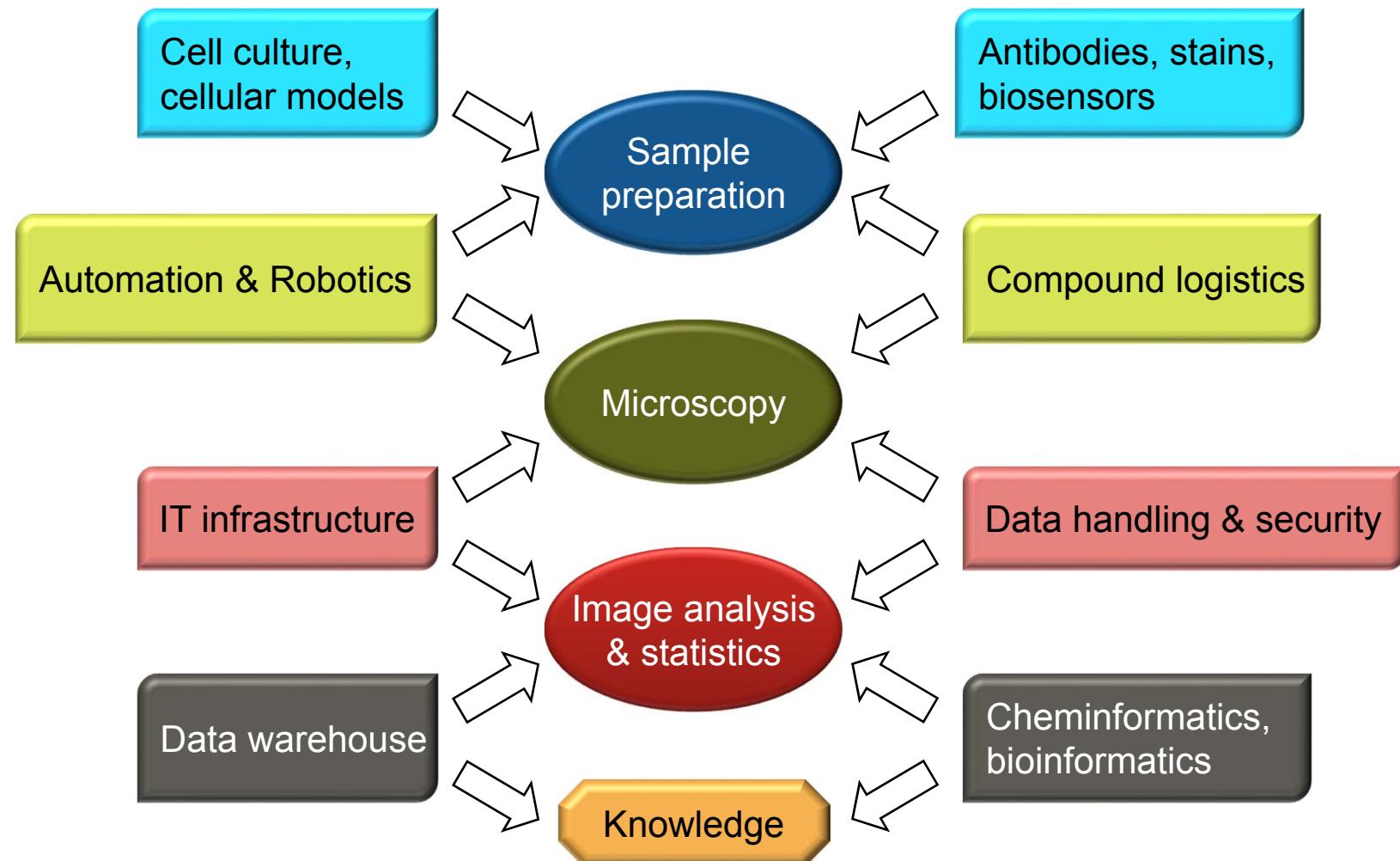


## Q: What is High-Content Screening

- A: automated (fluorescence) microscopy (of cultured cells *in vitro*) in combination with automated quantitative image analysis
- Q: what does “high-content” mean?
- A: (literally: result consists of more than 1 value.)
  - Many different readouts (features) derived from an image, such as,
    - \* nucleus area, circumference, roundness, length, width, ...
    - \* cytoplasm area, circumference, roundness, length, width, ...
    - \* intensity of marker 1 in nucleus, in cytoplasm, in spots, ...
    - \* intensity of marker 2 in nucleus, in cytoplasm, in spots, ...
    - \* number of nuclei per image, average number of spots per cell, ...
    - \* ...



# The elements of HCS

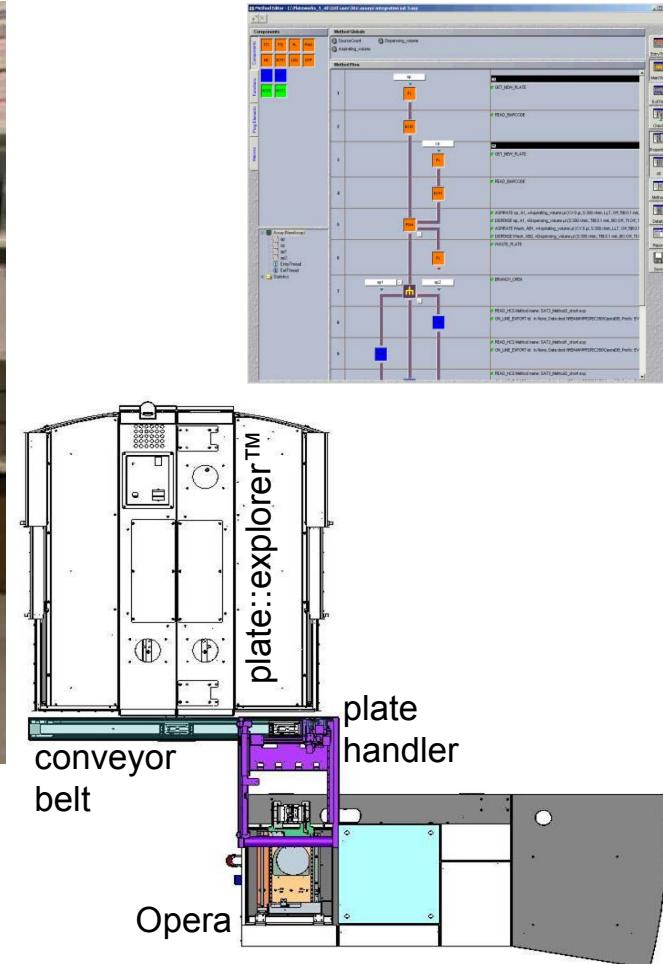


## Industrial HCS facility

F. Hoffmann-La Roche Ltd. 2014



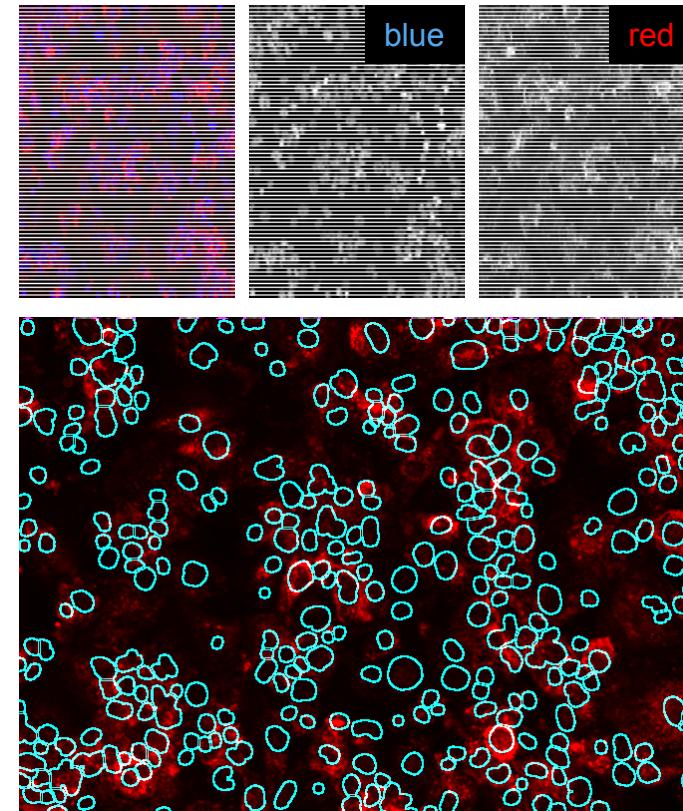
50.000 images in 24h



# Image analysis in HCS mode

## *Characteristics*

- Typical image type:
  - 2-5 channels, quasi monochrome
  - 2D (95%), 3D (1%),  
2D+time (4%), 3D+time (0%)
- Typical image objects:
  - Nucleus, Cytoplasm, Spots, Tubes
- Typical segmentation steps
  - Nucleus: good starting point
  - Cytoplasm: expand nucleus



# Image analysis in HCS mode

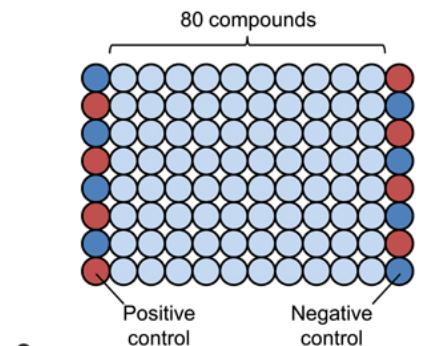
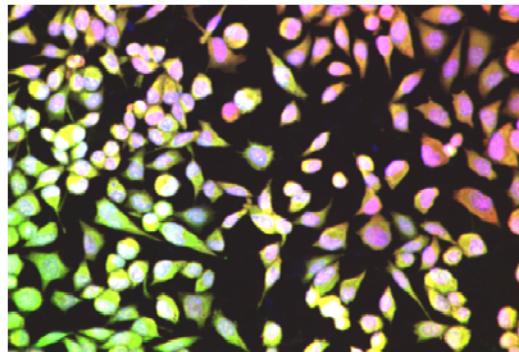
## *Requirements*

- Throughput per image field:
  - $1 \text{ s}^{-1}$  for screening (HT)
  - $0.01 \text{ s}^{-1}$  for profiling (MT)
- Robust without human intervention:
  - Heterogeneous cell population
  - Training data are never representative
    - Always expect the unexpected
  - Often low signal to noise, variable background

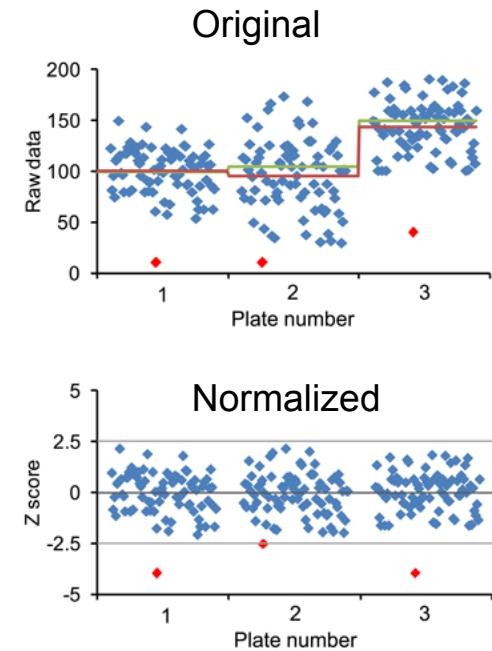
# HCS: all is not well

## *Correction of systematic errors*

- Plate to plate variation:
  - biological batch effect
- Well to well variation:
  - geometric patterns, pipetting errors
- Within field variation:
  - inhomogeneous illumination

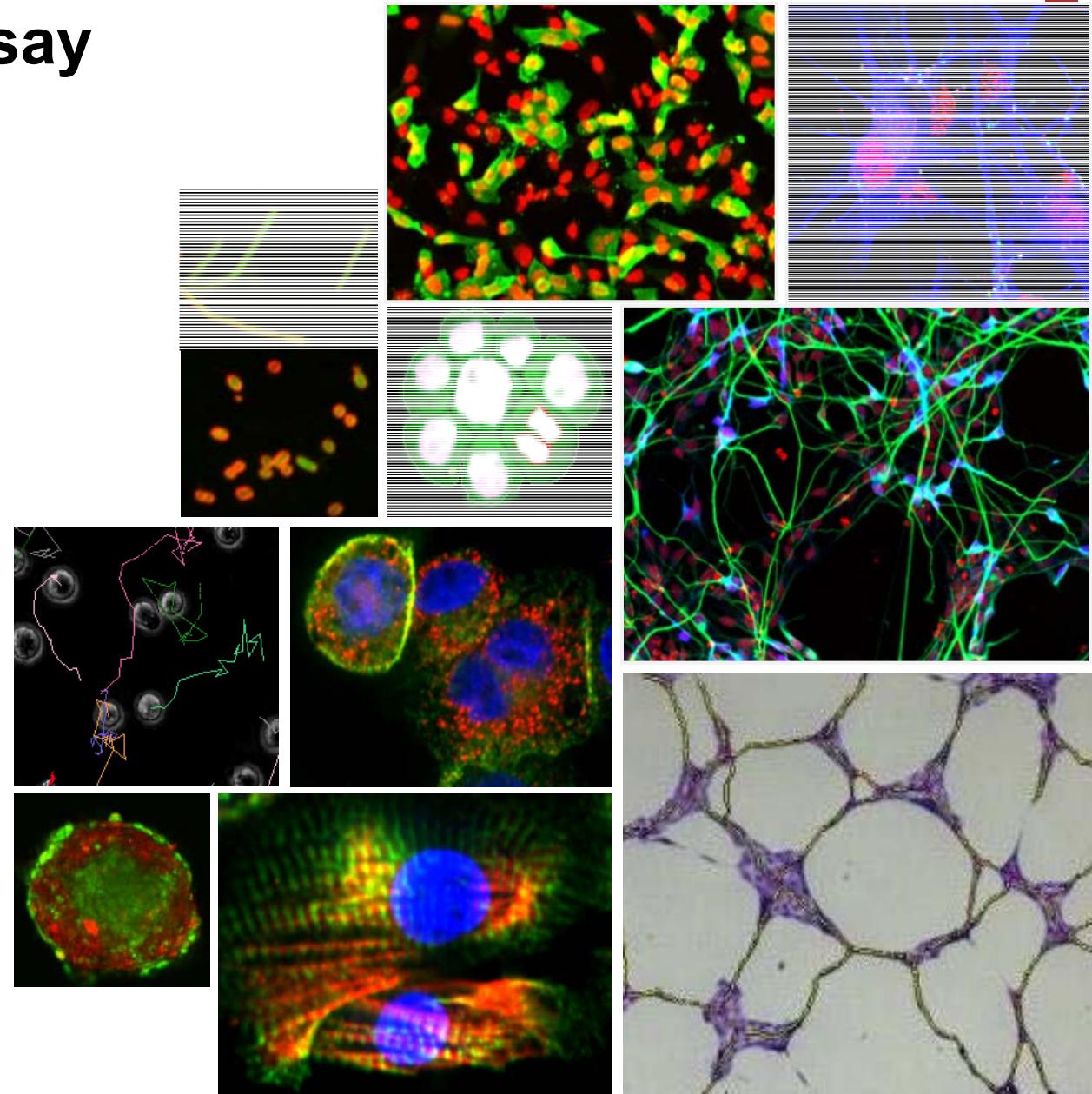


	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	-100	-66	-1	10	31	1	7	3	19	34	13	17	9	20	7	15	22	20	12	21	22	18	23	29
B	-97	-99	-32	5	31	-61	-47	-12	71	6	-10	42	61	-50	-14	42	15	9	-14	-48	2	0	-20	-15
C	101	-103	41	27	5	14	20	16	9	10	16	25	14	3	10	16	10	40	22	19	39	18	33	25
D	-90	-153	-47	-22	76	-32	-44	-18	-48	-53	-22	77	-10	-18	-8	-13	-24	-10	-9	-41	-5	-11	-45	-45
E	104	-108	13	20	10	13	11	15	2	14	14	5	14	16	16	15	1	23	30	39	44	22	-5	30
F	-99	-100	63	44	49	-77	-39	-11	-20	51	-9	13	-9	7	-12	-12	-50	-21	-2	-14	-52	-21	65	-50
G	101	-100	3	14	4	3	15	21	14	11	26	-16	23	16	23	12	14	18	-7	32	8	11	35	16
H	-99	-97	-42	-45	51	-18	-23	-43	-5	14	-69	63	-55	-49	-55	-33	-103	-148	-21	-20	-28	-111	79	-67
I	101	-102	-5	11	7	10	2	9	5	19	-13	16	15	29	7	15	21	14	20	20	7	26	21	14
J	104	97	-37	-42	-73	-53	-66	-53	-72	-8	-23	-81	-49	-6	-10	-101	-75	-79	-5	-28	-31	8	-38	-38
K	-97	-101	35	10	-11	9	11	11	11	11	12	8	11	20	5	25	5	15	15	21	15	28	33	-26
L	-104	-100	1	38	-18	-37	-51	-29	-68	-24	-12	-35	-54	-67	-5	-38	-48	-49	-44	-43	-23	-8	-42	-61
M	-98	-101	-5	2	-2	5	5	11	9	9	2	9	0	10	9	1	12	95	15	32	30	11	19	11
N	-106	99	-50	-63	54	-17	-25	-56	-31	-8	-21	-23	84	30	-48	-30	-36	-57	-95	-9	-75	91	-31	-11
O	-98	-103	5	37	3	-3	8	5	3	8	-5	10	6	-3	15	6	24	23	28	31	6	12	11	9
P	-102	-105	4	71	-19	-28	-37	-44	-30	-16	-47	-67	52	-15	-26	-69	4	-31	-28	-72	-23	-79	34	-59



# Typical HCS assay repertoire

- Micronucleus formation
- Internalization
- Trafficking
- Translocation
- Subcellular localization
- Calcium flux
- Apoptosis
- Proliferation
- Migration
- Sarcomere structure
- Neurite outgrowth
- Synapse formation
- Endothelial tube formation
- 3D spheroid microtissue



# Image analysis in HCS: a case study



Cell Reports 9, 810–820, November 6, 2014

Cell Reports  
Report

## Disease Modeling and Phenotypic Drug Screening for Diabetic Cardiomyopathy using Human Induced Pluripotent Stem Cells

Faye M. Drawnel,<sup>1</sup> Stefano Boccardo,<sup>1,2</sup> Michael Prummer,<sup>1</sup> Frédéric Delobel,<sup>1</sup> Alexandra Graff,<sup>3</sup> Michael Weber,<sup>1</sup> Régine Gérard,<sup>1</sup> Laura Badi,<sup>1</sup> Tony Kam-Thong,<sup>1</sup> Lei Bu,<sup>4</sup> Xin Jiang,<sup>4</sup> Jean-Christophe Hoflack,<sup>1</sup> Anna Kialainen,<sup>1</sup> Elena Jeworutzki,<sup>1</sup> Natsuyo Aoyama,<sup>5</sup> Coby Carlson,<sup>5</sup> Mark Burcin,<sup>1</sup> Gianni Gromo,<sup>1</sup> Markus Boehringer,<sup>1</sup> Henning Stahlberg,<sup>3</sup> Benjamin J. Hall,<sup>1</sup> Maria Chiara Magnone,<sup>1</sup> Kyle Kolaja,<sup>5</sup> Kenneth R. Chien,<sup>6,7</sup> Jacques Bailly,<sup>1</sup> and Roberto Iacone<sup>1,\*</sup>

<sup>1</sup>Roche Pharma Research & Early Development, Roche Innovation Center Basel, 4070 Basel, Switzerland

<sup>2</sup>Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA

<sup>3</sup>Center for Cell Imaging and Nano Analytics, Biozentrum, Department for Biosystems Science and Engineering, University of Basel, 4058 Basel, Switzerland

<sup>4</sup>The Cardiovascular Research Center, Richard B. Simches Research Center, Massachusetts General Hospital, Suite 3201, Boston, MA 02114, USA

<sup>5</sup>Cellular Dynamics International, Madison, WI 53711, USA

<sup>6</sup>Departments of Cell and Molecular Biology and of Medicine Huddinge, Karolinska Institutet, 171 77 Stockholm, Sweden

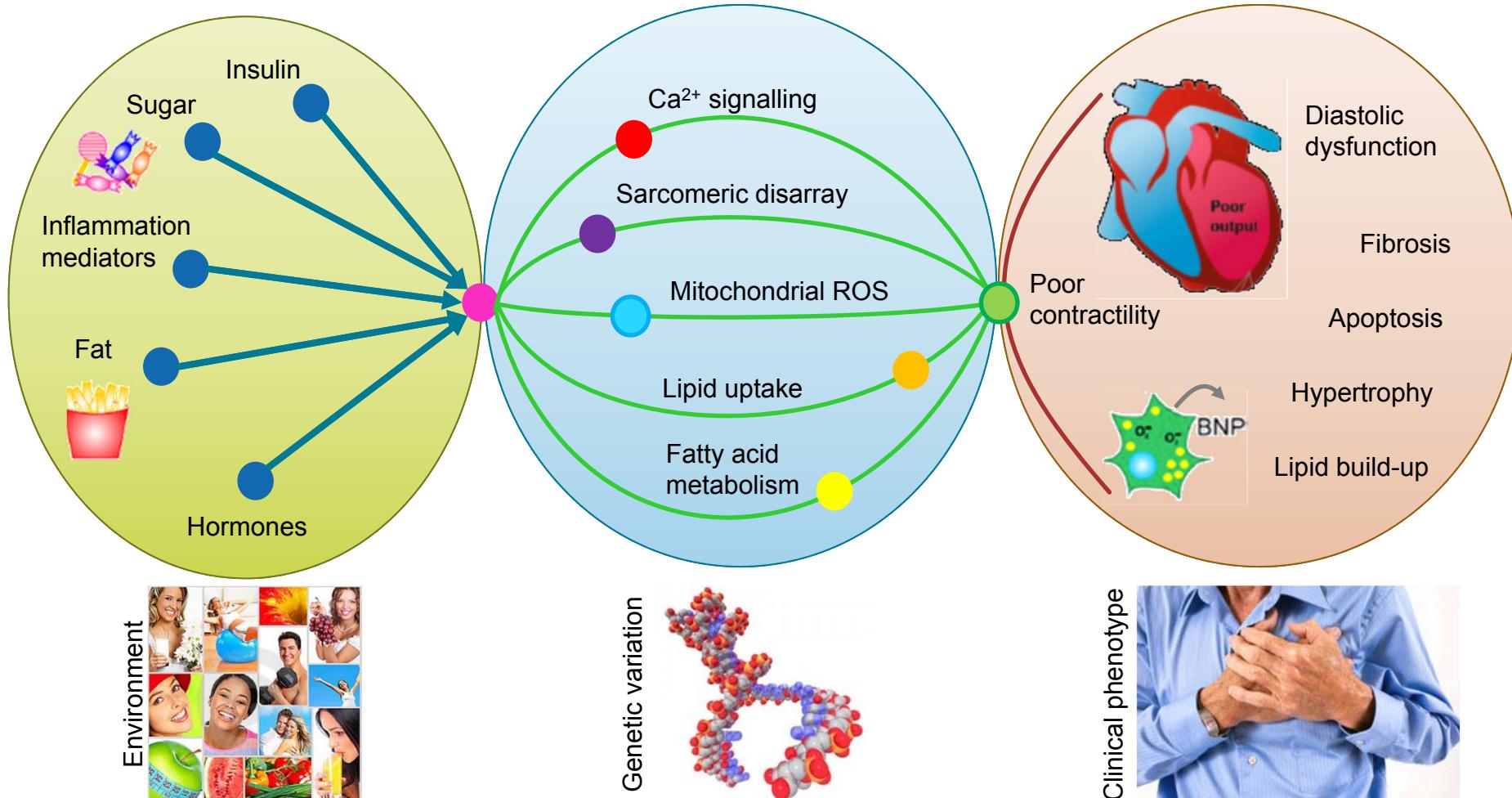
<sup>7</sup>Harvard Department of Stem Cell and Regenerative Biology, Harvard University, Cambridge, MA 02138, USA

\*Correspondence: [roberto.iacone@roche.com](mailto:roberto.iacone@roche.com)

<http://dx.doi.org/10.1016/j.celrep.2014.09.055>

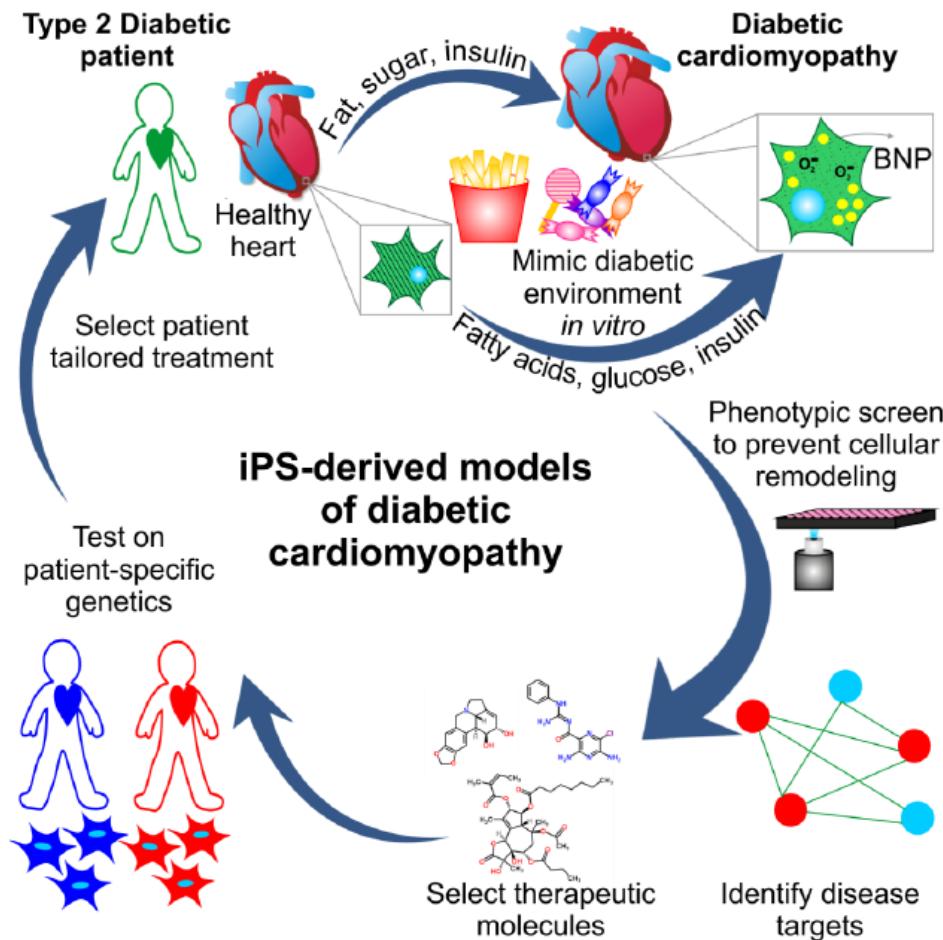
# Diabetic cardiomyopathy

## *Multiple inputs, multiple outputs*



# Diabetic cardiomyopathy modeling

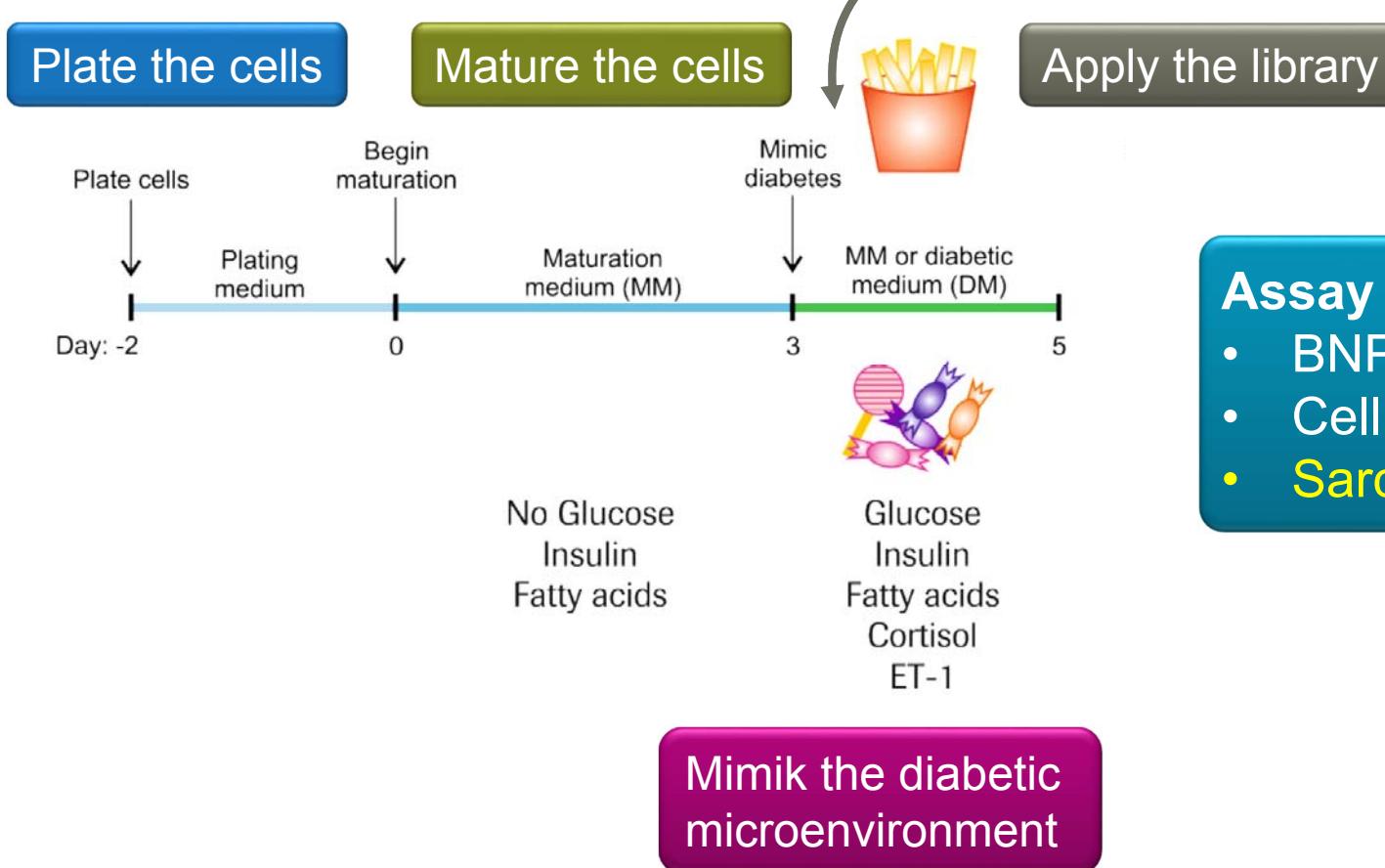
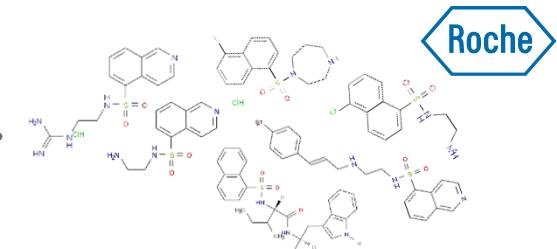
*«Patient in a dish»*



Find compounds that protect cardiomyocytes against glucose-induced stress

# The sarcomere integrity assay

## Principle



- Assay readouts:**
- BNP secretion
  - Cell area
  - Sarcomere integrity

# Sarcomere structure in cardiomyocytes

## *Fluorescence staining*

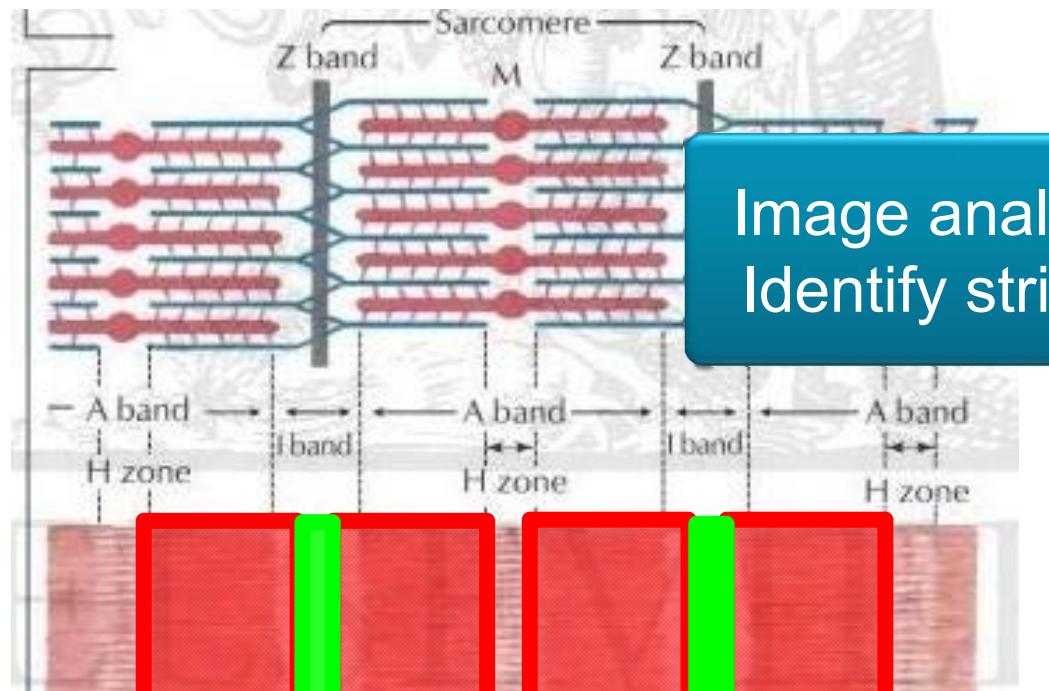
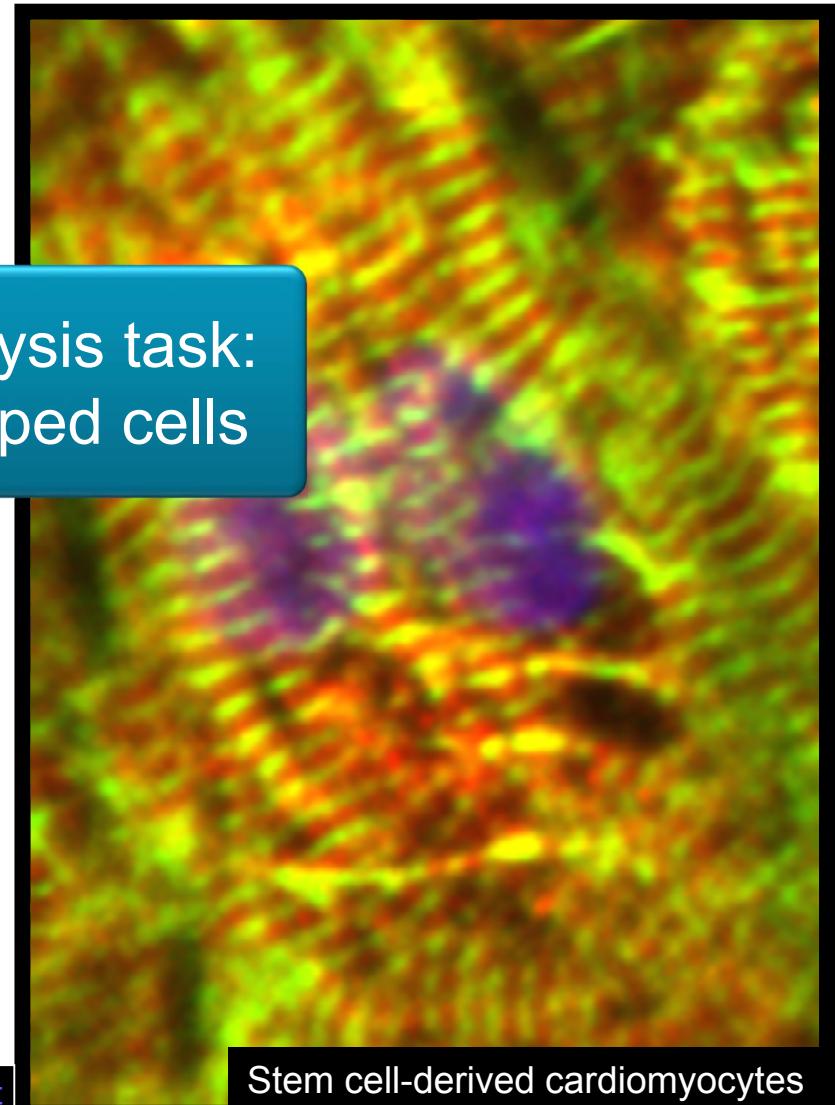


Image analysis task:  
Identify striped cells

Troponin T,  $\alpha$  actinin, hoechst

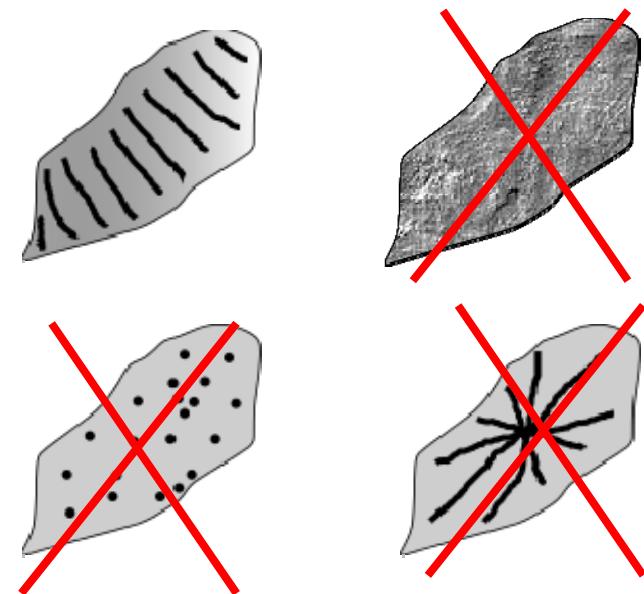


Stem cell-derived cardiomyocytes

# Image analysis: identify myocyte sarcomers

## *Concept*

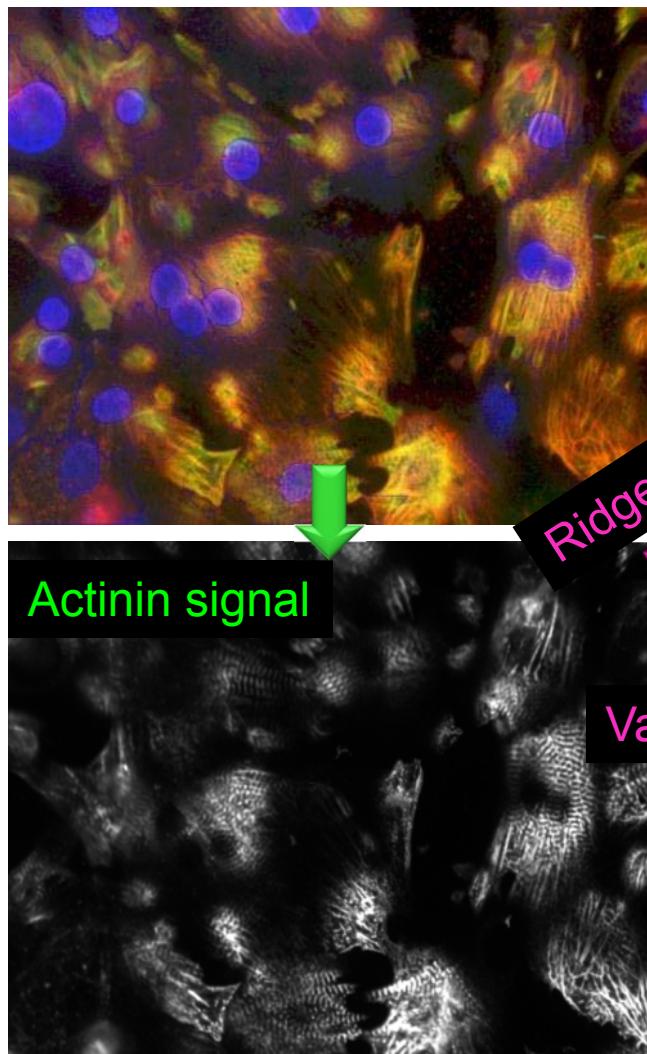
- Striped cells:
  - extended patches of parallel equidistant linear features
  - Not: homogeneous stain, spotted stain, star-shaped stain
- Object-based approach:
  - Stripe object lengths
  - Stripe object orientations
  - Stripe object distances
- Problem:
  - Very difficult and calculation intensive
  - Big bullet for «small» target



- Solution:
  - Linear features: texture filtering
  - Equidistant parallel features: Image auto-correlation / Fourier Transform

# Sarcomere integrity image analysis

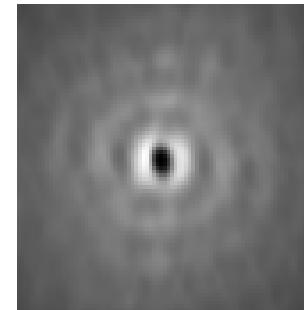
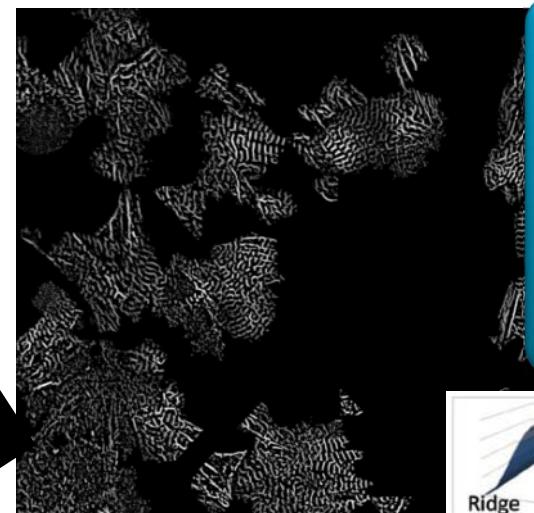
## *Ridge-Valley filtering*



Ridge filter

Actinin signal

Valley filter



Ridge-valley cross correlation features:

- Modulation amplitude: sarcomere strength
- Modulation frequency: sarcomere length

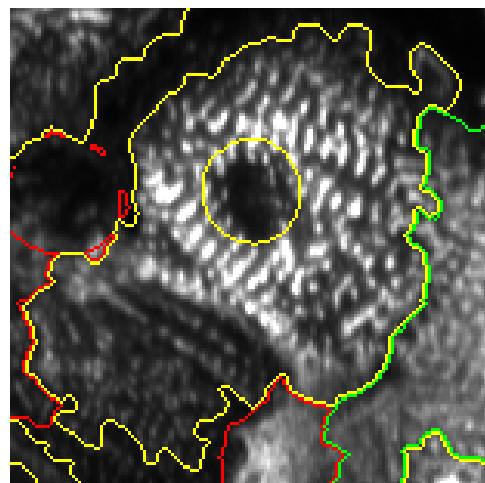
Whole image

- No cell classification
- No correlation with other cellular features

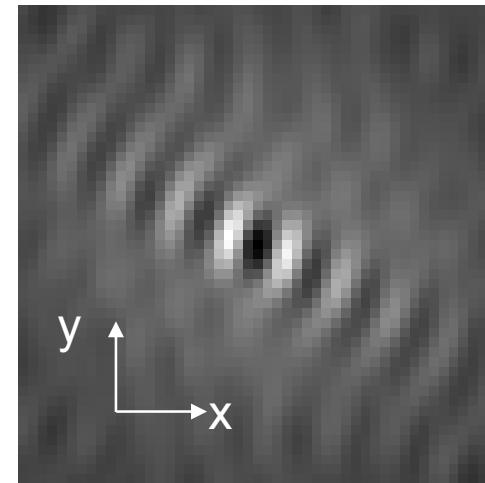


# Sarcomere ICCS analysis: per individual cell

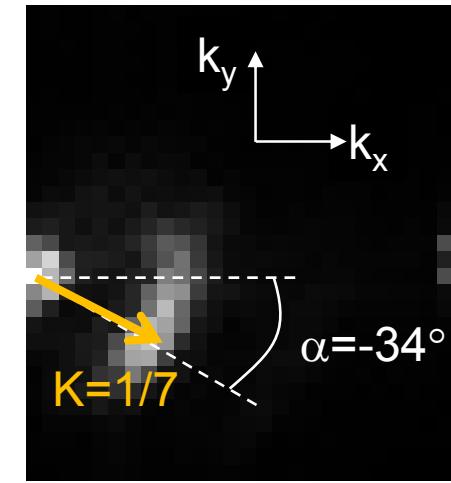
Actinin image



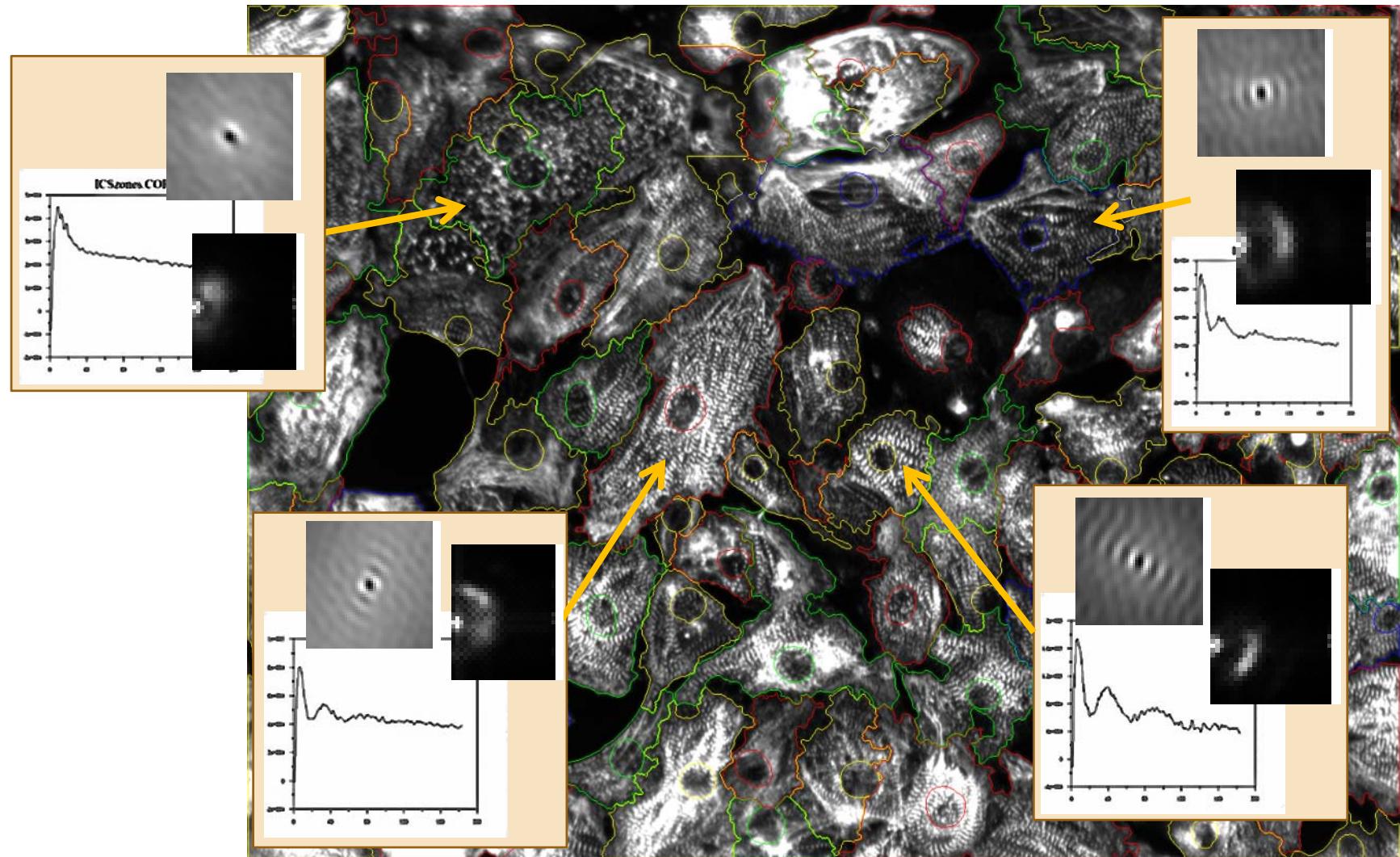
R-V cross correlogram



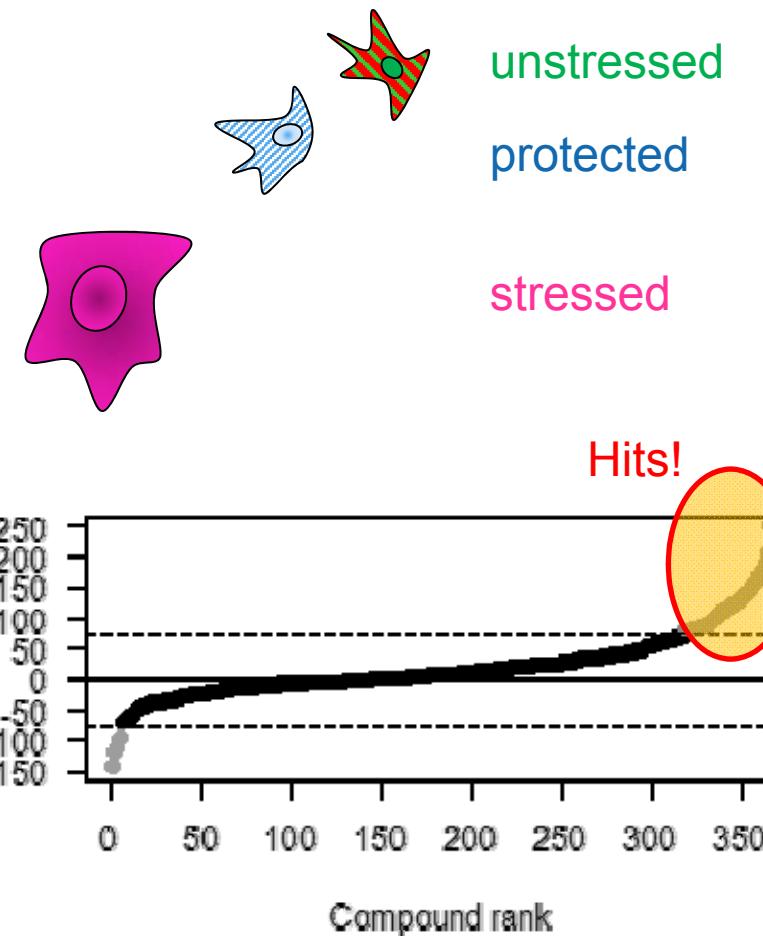
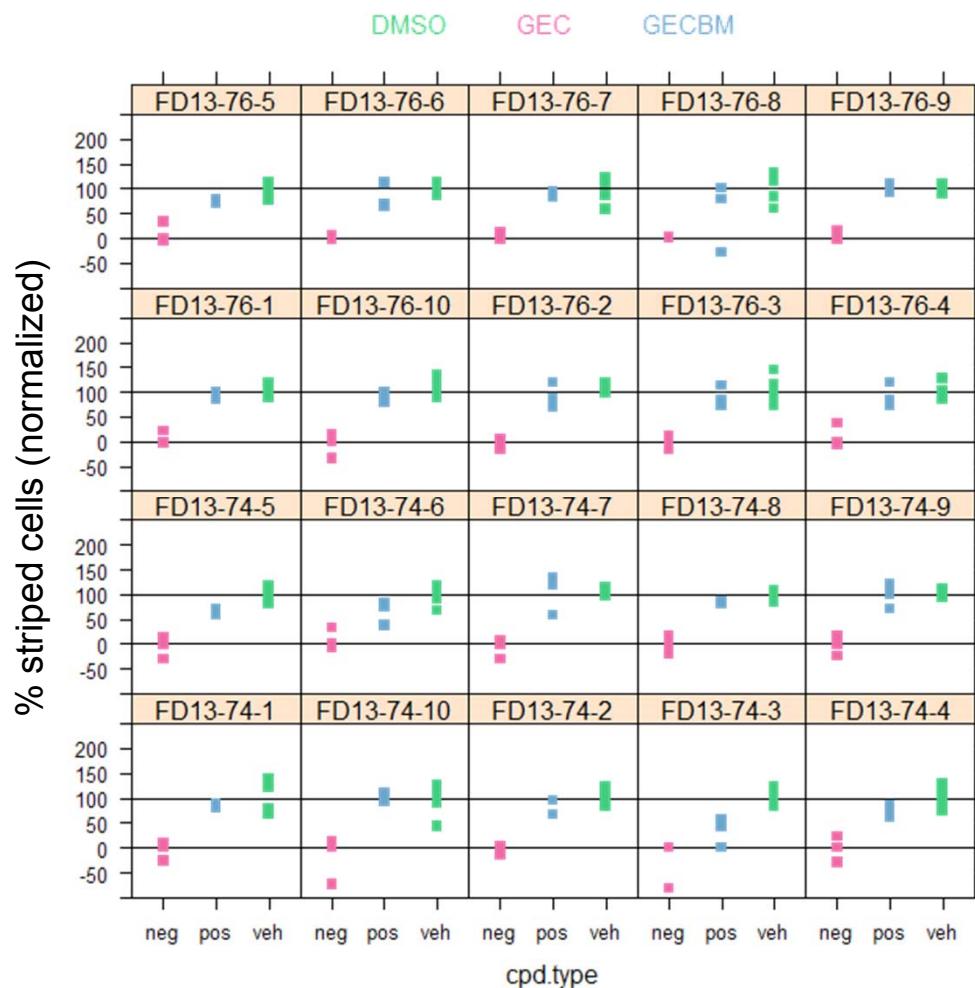
R-V cross spectrogram



# Sarcomere ICCS analysis: per individual cell



# Sarcomere integrity (SARITY) screen summary



# Summary

- Introduction to image-based screening (HCS) in drug discovery
- Requirements for image analysis in screening
- Prototypic HCS campaign:
  - Modern cellular disease model
  - Phenotypic assay development
  - Tailored image analysis for feature generation
  - Quality control of HCS results
- Next:
  - Using Computer Vision + Machine Learning for Aerial Image Segmentation