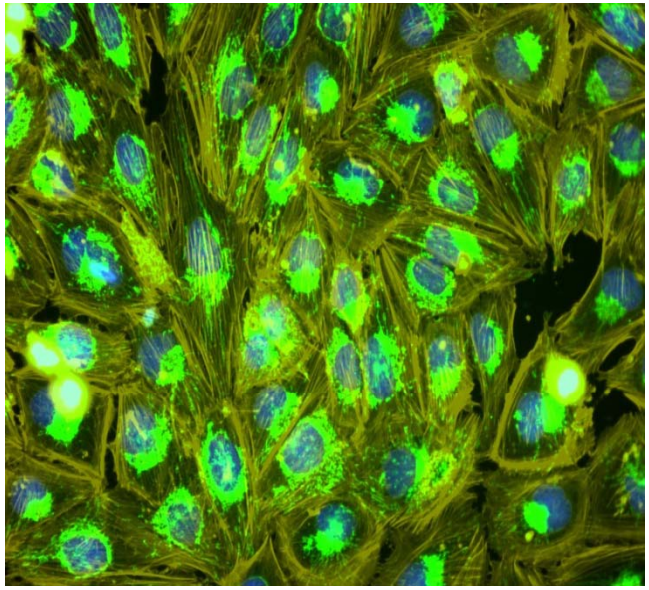
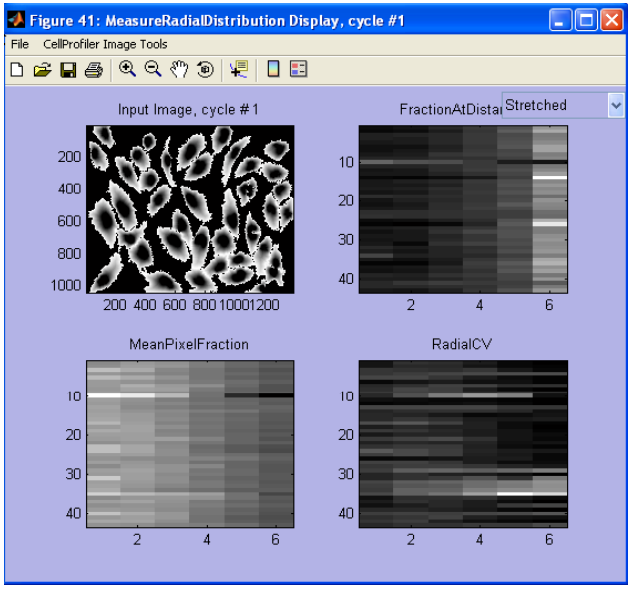
# Fingerprinting Cellular Physiology
# Using Multiparametric Immunofluorescent Cytometry





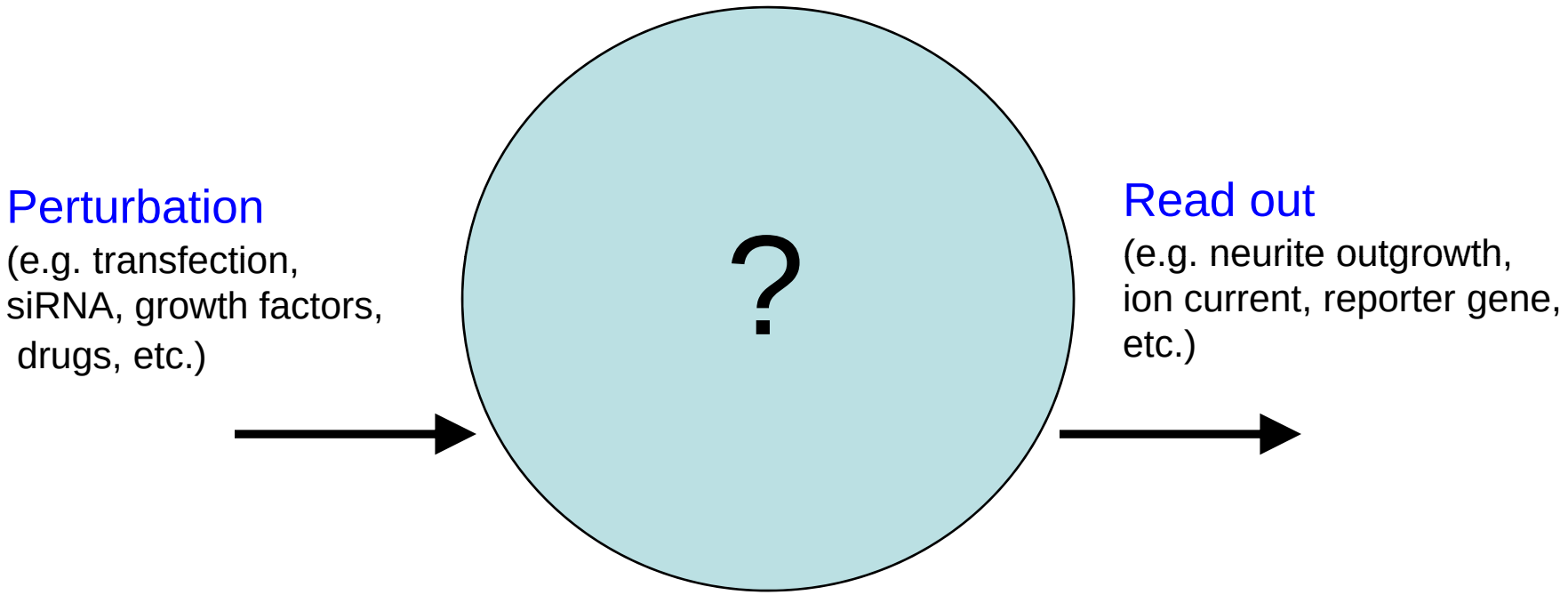

**Erik Gunther & Steve Strittmatter**

11-26-08

# Outline

- High dimensionality data in physiology fingerpringting
- Multiparametric immunocytometry as a source of high dimensionality data
- Multiparametric immunocytometry as a means of drug action analysis

# Experimental Strategy of the Cell Biologist

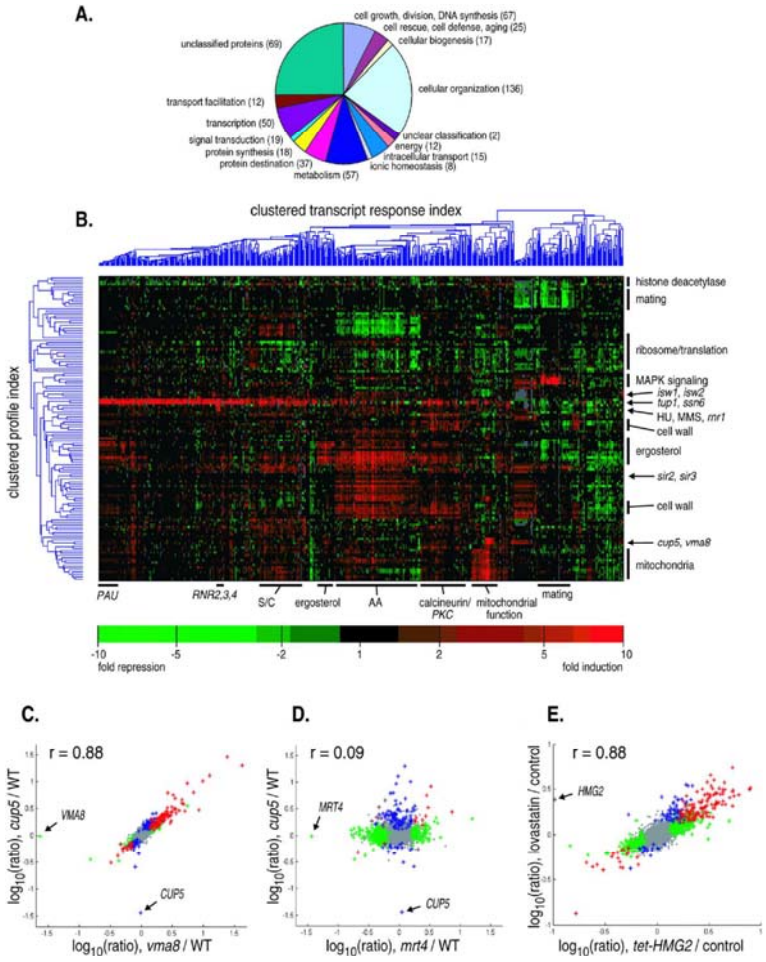



Problem: limited scope/understanding

# Functional discovery by statistical classification of gene expression profiles ("fingerprinting")

e.g.

Functional discovery via a compendium of expression profiles.
Hughes et al, Cell. 2000 Jul 7;102(1):109-26.




## Method

1. Mutate 300 genes independently
2. Derive microarray expression profiles
3. Cluster per changes in overall expression
4. Infer function on basis of clustering

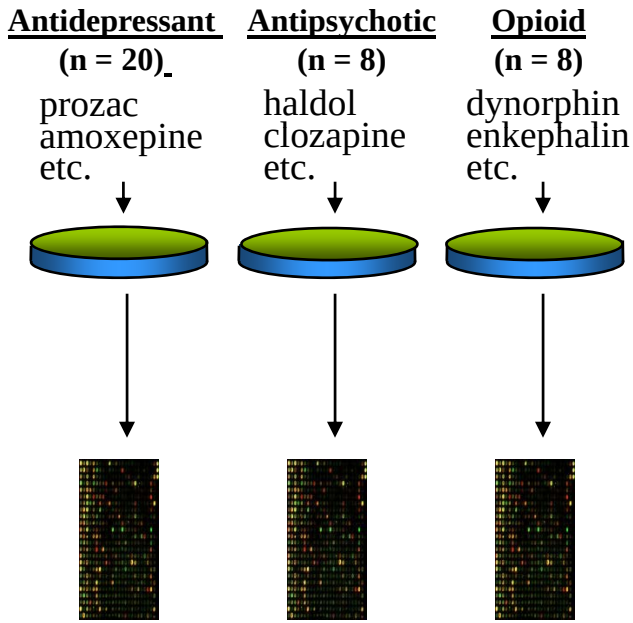
↓

Discovery of genes involved in cell wall synthesis, ergosterol synthesis & mitochondrial respiration.

# Listening to the cells with chips – a *lingua franca* of drug action

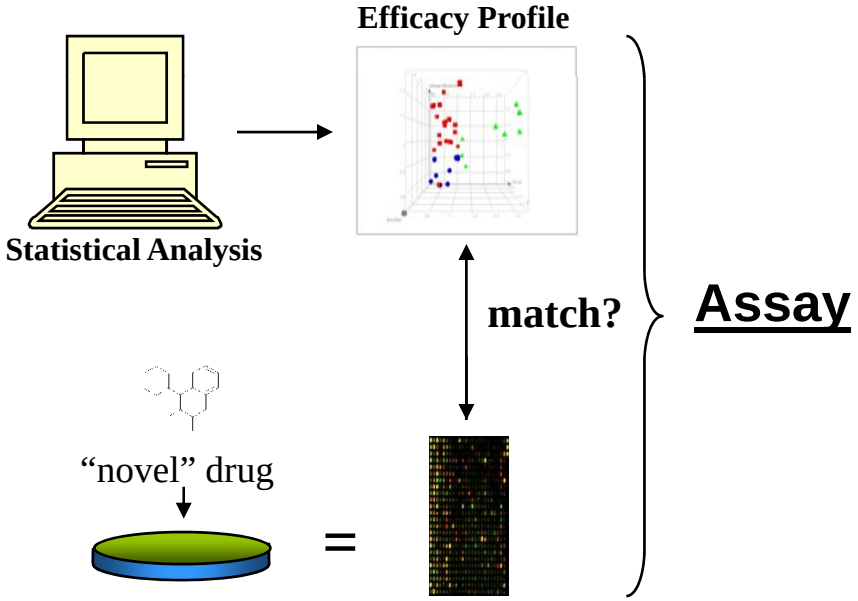
## Method

**1.** Treat neurons with multiple antidepressants, antipsychotics and opioid drugs




**2.** Generate microarray expression data for multiple members of each class

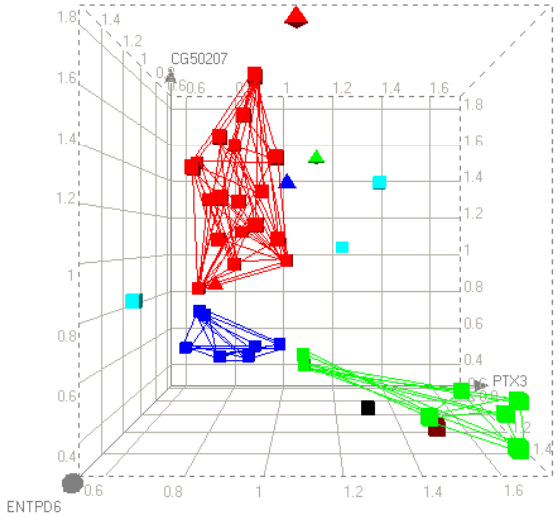
**3.** Statistically analyze biomarker expression: establish signature "efficacy profiles" of drug action




**4.** Novel therapeutic compounds are recognized by the signature expression profile they induce.

# Pharamcogenomic derivation of drug efficacy biomarker profiles

## Supervised Classification Picks Out Class-Distinguishing Features




Drug class separation provided by three example biomarkers

- **antidepressant**
- **antipsychotic**
- **opioid agonist**
- **control**

**Result:** “Classification Tree” statistical method identifies biomarkers that separate drug classes in n-dimensional gene space.

Gunther et al., PNAS, 2003 Aug 5; 100(16):9608-13.
**Prediction of clinical drug efficacy by classification of drug-induced genomic expression profiles *in vitro*.**

# Drug efficacy profiles predict drug class

| Compound | "True" Class | CT Prediction | RF Prediction | Functional Subclass |
|---|---|---|---|---|
| amoxapine | AD | AD | AD | tricyclic |
| clomipramine | AD | AD | AD | tricyclic |
| desipramine | AD | OP | AD | tricyclic |
| doxepin | AD | AD | AD | tricyclic |
| imipramine | AD | AD | AD | tricyclic |
| maprotiline | AD | AD | AD | tricyclic |
| nortryptyline | AD | AD | AD | tricyclic |
| protriptyline | AD | AD | AD | tricyclic |
| trimipramine | AD | AD | AD | tricyclic |
| amitriptyline | AD | AD | AD | tricyclic |
| citalopram | AD | AD | AD | SSRI |
| paroxetine | AD | AD | AD | SSRI |
| sertraline | AD | AD | AD | SSRI |
| fluoxetine | AD | AD | AD | SSRI |
| fluvoxamine | AD | AD | AD | SSRI |
| tranylcypromine | AD | AD | AD | MAOI |
| phenylzine | AD | AP | AP | MAOI |
| iproniazid | AD | AD | AP | MAOI |
| trazadone | AD | AD | AD | atypical |
| bupropion | AD | AD | AD | atypical |
| chlorpromazine | AP | AP | AP | classic |
| trifluperazine | AP | AP | AP | classic |
| triflupromazine | AP | AP | AP | classic |
| pimozide | AP | AP | AP | classic |
| clozapine | AP | AD | AD | atypical |
| haloperidol | AP | AP | AP | atypical |
| risperidone | AP | AP | AP | atypical |
| loxapine | AP | AP | AD | atypical |
| BW373U86 | OP | OP | AD | δ OPR |
| Enkephalin | OP | AD | AD | δ OPR |
| U50488 | OP | OP | OP | κ OPR |
| U62066 | OP | OP | OP | κ OPR |
| Endomorphin | OP | OP | OP | κ OPR |
| DALDA | OP | OP | OP | μ OPR |
| DAMGO | OP | OP | OP | μ OPR |
| Dynorphin A | OP | OP | AD | μ OPR |
| **% "correct"** | | **88.9** | **83.3** | |

**Process:** Statistical models built with each individual drug left out, then predicted

**Result:** Biomarker efficacy profiles correctly identify clinical utility of "novel" psychiatric drugs with **88.9%** accuracy

**Significance:** Robust signatures of drug efficacy are recognizable

Gunther et al., PNAS, 2003 Aug 5; 100(16):9608-13.
**Prediction of clinical drug efficacy by classification of drug-induced genomic expression profiles *in vitro*.**

# Prediction of “novel” drug subclasses

| Compound | “True” Class | RF Prediction | Functional Subclass |
|---|---|---|---|
| amoxapine | AD | AD | tricyclic |
| clomipramine | AD | AD | tricyclic |
| desipramine | AD | AD | tricyclic |
| doxepin | AD | AD | tricyclic |
| imipramine | AD | AD | tricyclic |
| maprotiline | AD | AD | tricyclic |
| nortryptyline | AD | AD | tricyclic |
| protriptyline | AD | AD | tricyclic |
| trimipramine | AD | AD | tricyclic |
| amitriptyline | AD | AD | tricyclic |
| citalopram | AD | AD | SSRI |
| paroxetine | AD | AD | SSRI |
| sertraline | AD | AD | SSRI |
| fluoxetine | AD | AD | SSRI |
| fluvoxamine | AD | AD | SSRI |
| tranylcypromine | AD | | MAOI |
| phenylzine | AD | | MAOI |
| iproniazid | AD | | MAOI |
| trazadone | AD | | atypical |
| bupropion | AD | | atypical |
| chlorpromazine | AP | | classic |
| trifluperazine | AP | | classic |
| triflupromazine | AP | | classic |
| pimozide | AP | | classic |
| clozapine | AP | | atypical |
| haloperidol | AP | | atypical |
| risperidone | AP | | atypical |
| loxapine | AP | | atypical |
| BW373U86 | OP | | δ OPR |
| Enkephalin | OP | | δ OPR |
| U50488 | OP | | κ OPR |
| U62066 | OP | | κ OPR |
| Endomorphin | OP | | κ OPR |
| DALDA | OP | | μ OPR |
| DAMGO | OP | | μ OPR |
| Dynorphin A | OP | | μ OPR |

**Tricyclic correct: 100%**

**Process:** Statistical models built with entire Tricyclic or SSRI subclasses left out, then asked to predict

**Result:** Tricyclic-naïve model predicts antidepressant efficacy of tricyclics with 100% accuracy. Same with SSRIs.

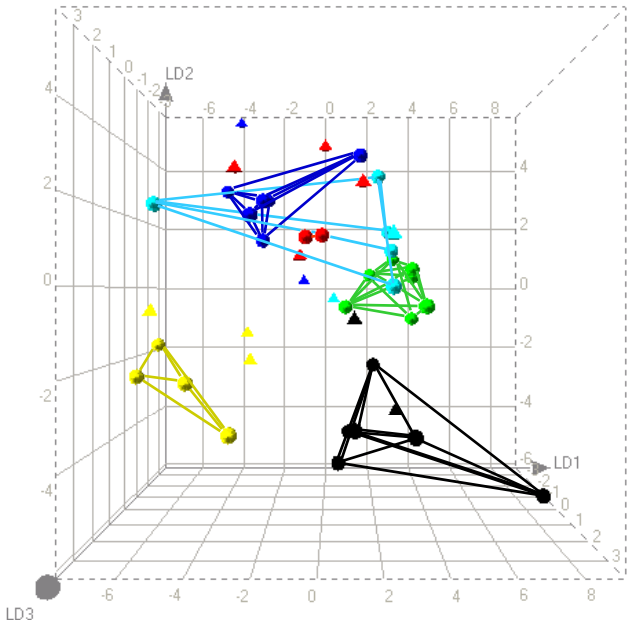
**Significance:** “Higher-order” signatures of drug efficacy make clinical utility predictable without knowledge of target

Gunther et al., PNAS, 2003 Aug 5; 100(16):9608-13.
**Prediction of clinical drug efficacy by classification of drug-induced genomic expression profiles *in vitro*.**

# Biomarkers define numerous drug efficacies

**Process:** Statistical models (linear discriminant analysis; LDA) built from drug classes active in HCN1A cells

**Result:** Multiple drug efficacies are separable in gene space

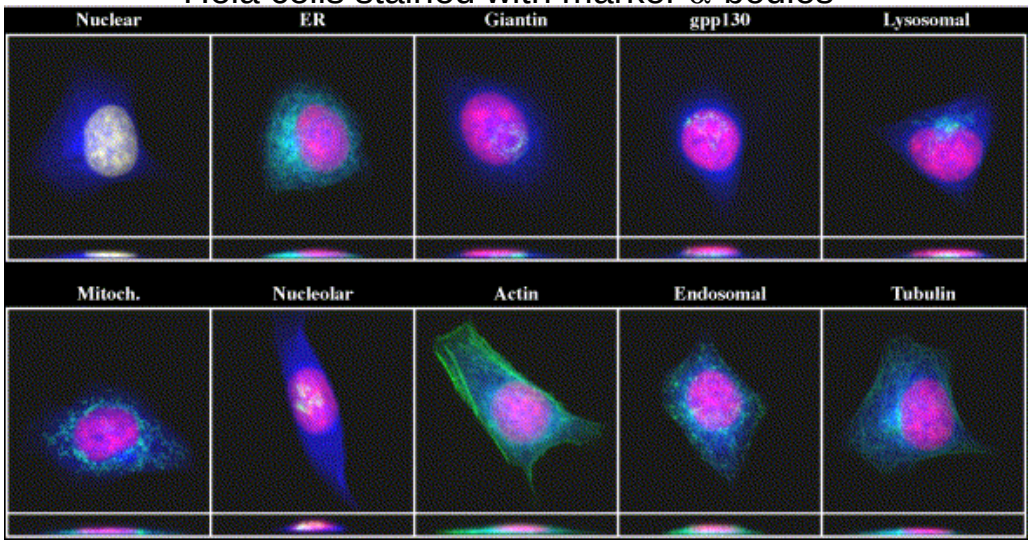



- **antianxiety**
- **anticonvulsant**
- **antihypertensive**
- **COX inhibitor**
- **opioid agonist**
- **DMSO control**

Axes represent 1st three linear discriminants of top 12 biomarkers identified by LDA

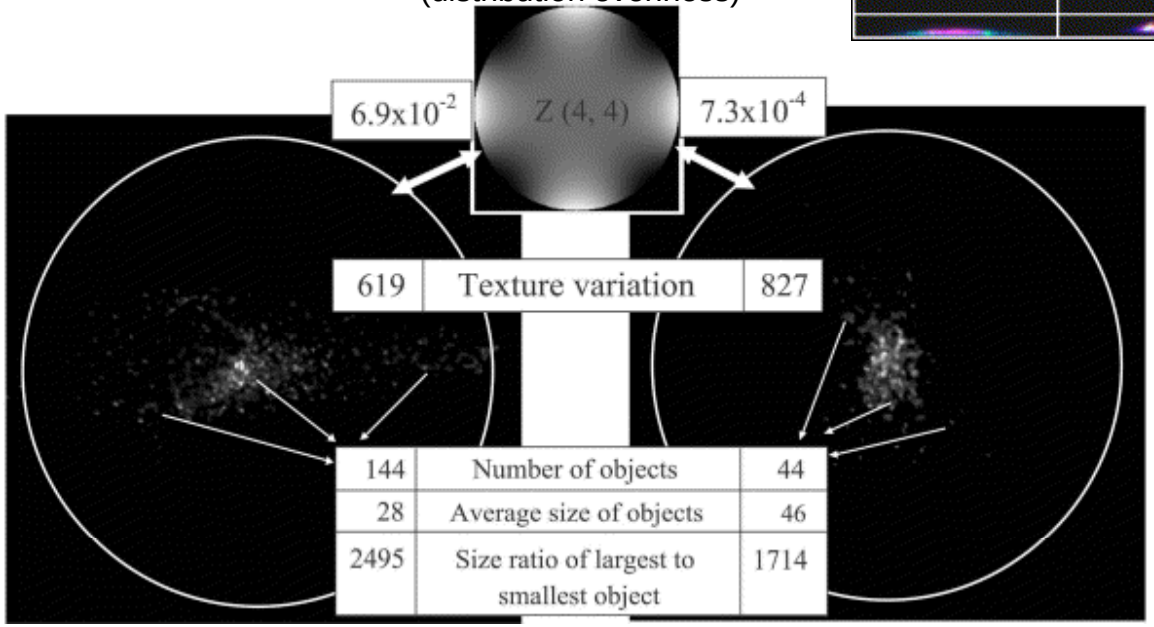
# Multiparametric Immunofluorescent Cytometry

Hela cells stained with marker $\alpha$-bodies




Geometric descriptors of texture can distinguish different antibody staining patterns

Zernike moment (4,4)
(distribution evenness)




transferrin receptor

lysosomal protein LAMP2

Procedure:

1. Find objects (the individual endosomes or lysosomes) by thresholding the image
2. Count and measure the objects.

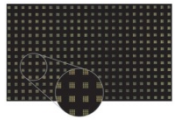
**Automated interpretation of subcellular patterns from immunofluorescence microscopy**
**Yanhua Hu and Robert F. Murphy**
**Journal of Immunological Methods Volume 290, Issues 1-2 , July 2004, Pages 93-105**

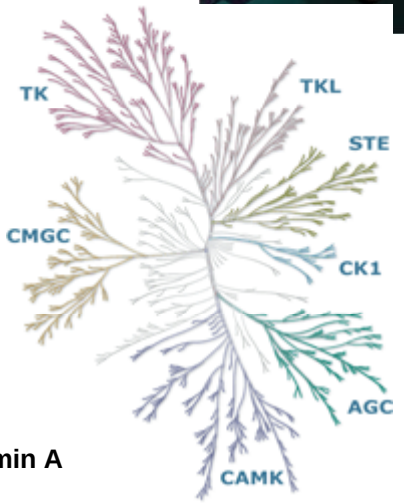
# High-throughput multidimensional cell imaging platform

## CNNR Cell Imaging Resource (BCMM 402)

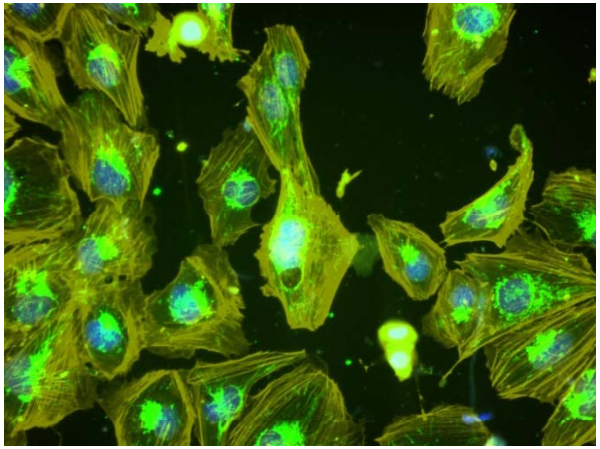
Instrument: Molecular Devices ImageXpress Micro (25,000 images/day)
Robotic plate handling/bar code reading
3 cytometric workstations (CellProfiler software)

Materials: Cell Signaling Technology antibodies (currently 320)
384-well plates





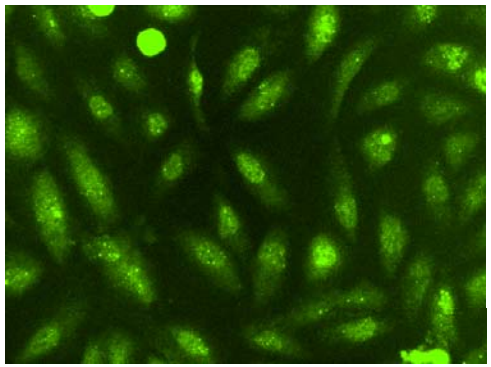

DAOY medulloblastoma cells



**Cleaved Lamin A**
**SRC-3**
**Phospho-EGF Receptor**
**HER2/ErbB2**
**c-Jun**
**S6 Ribosomal Protein**
**Chk1**
**CEA/CD66e**
**HSP27**
**PCNA**
**Alpha-Synuclein**
**Zap-70**
**Neurofilament-L**
**Etc….**

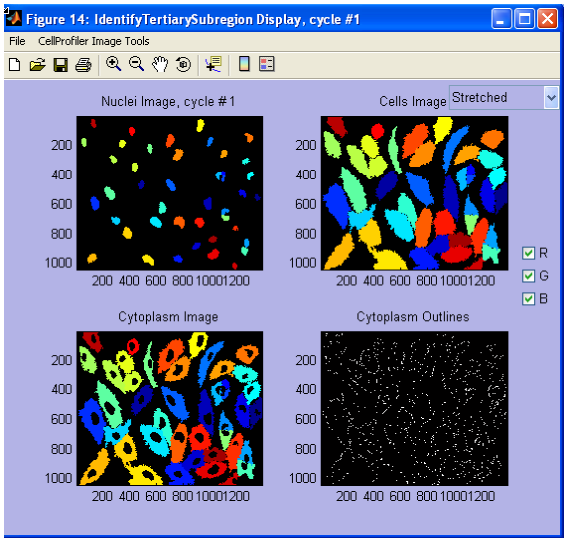
**Cyclin D1**
**Smac/Diablo**
**Akt**
**Ezh2**
**NQO1**
**CD56 (NCAM)**
**GFAP**
**Phospho-Myosin Light Chain**
**Tau**
**Cyclin B1**
**Etc….**

# Multiparametric Analysis with CellProfiler

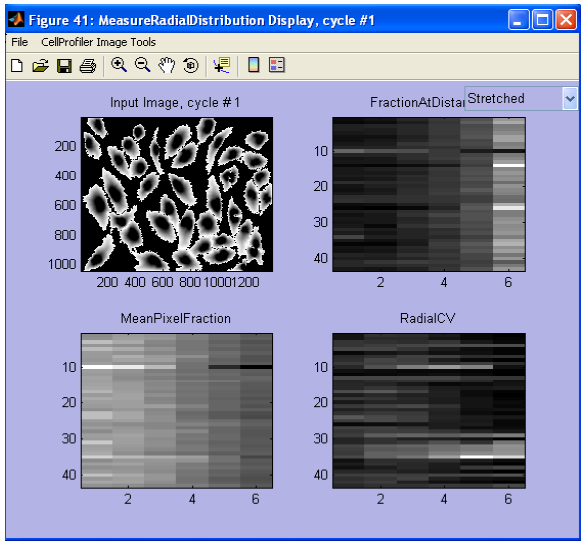


Performed on FITC channel

Step 1: cell segmentation




Step 2: feature extraction




For nucleus, cytoplasm & whole cell (2 wavelengths):

## Shape

**Area**
**Eccentricity**
**Solidity**
**Perimeter**
**Form factor**
**Major axis**
**Minor axis**

## Intensity

**Integrated**
**Mean**
**Std**
**Min**
**Max**
**Edge**
**Mass displacement**
**Ratios nuc/cyt**
**Correlation FITC,DAPI**

## Texture

**Angular second moment**
**Contrast**
**Correlation**
**Variance**
**Inverse difference moment**
**Sum entropy**
**Info measure 1**
**Info measure 2**
**Gabor X**
**Gabor Y**

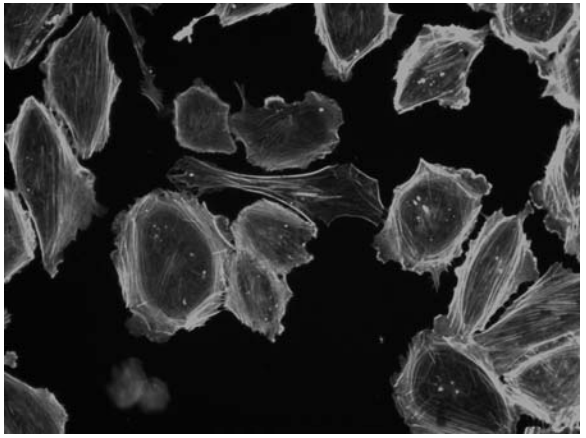
## Data Dimensionality

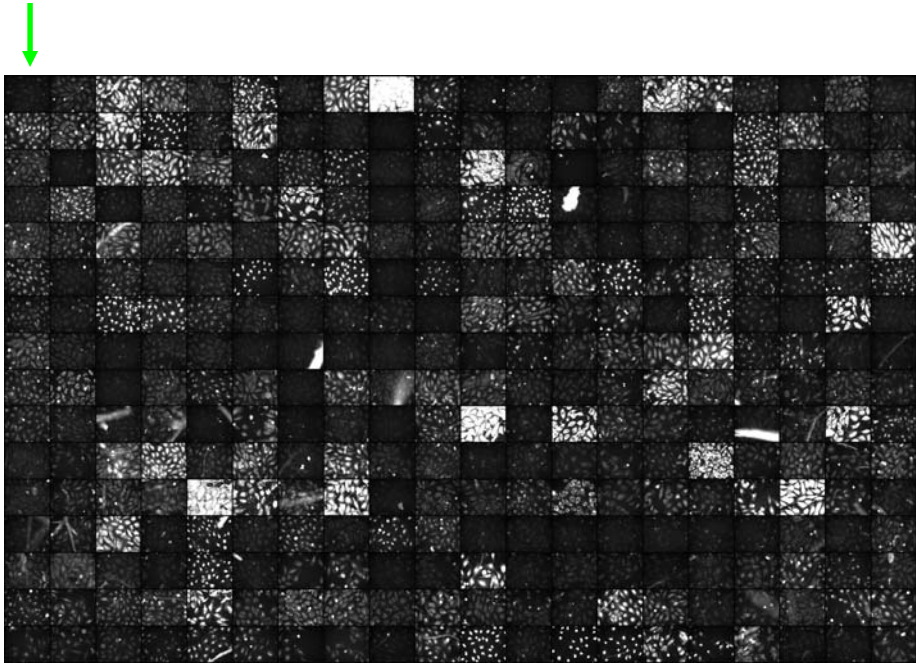
- 233 descriptors (parameters)
- 300 αbs
- 233 descriptors x 300 αbs = 70,000 data points/treatment (similar scale to microarray)

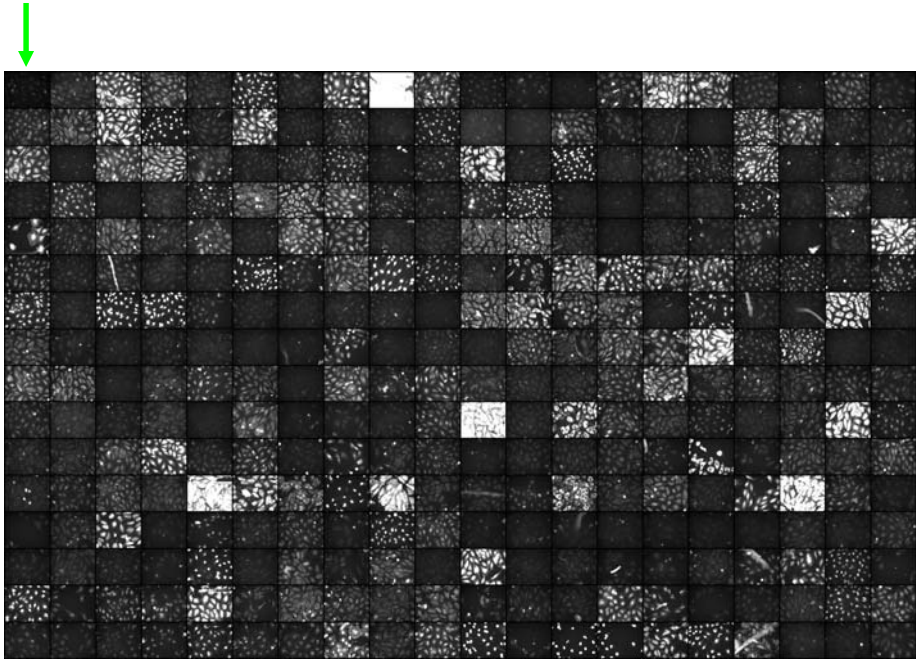
# Antibody Immunoreactivity/CAMP Responsiveness

Pilot experiment:

- two conditions: 8-Br-CAMP, DMSO control
- triplicate treatments, three images/well
- stain each treatment with all 320 antibodies
  (six 384-well plates; 5400 images total)
- image
- extract features (CellProfiler)
- quantify differences





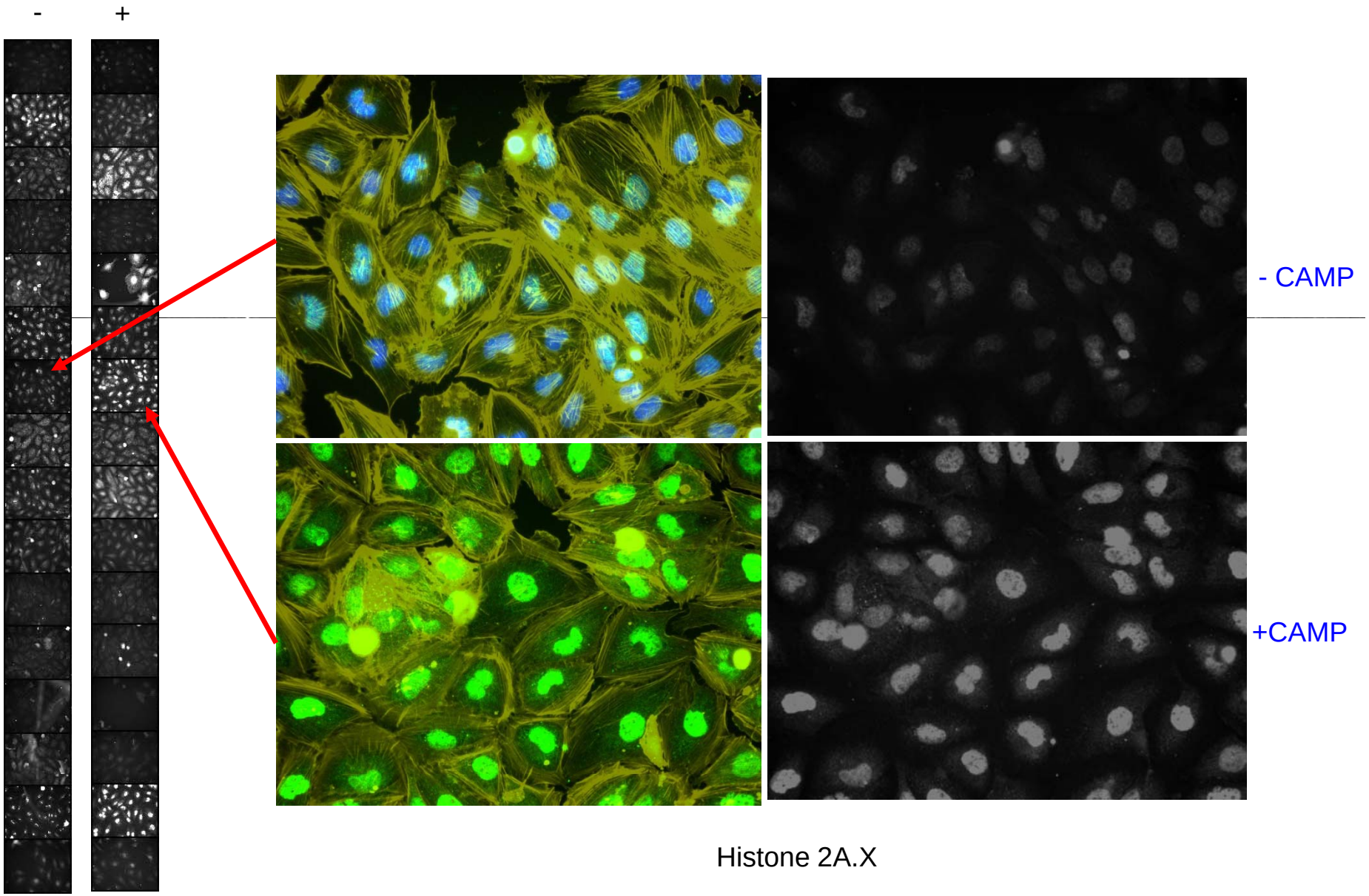


- CAMP

Daoy Mosaic
320 antibodies, 2 conditions

+CAMP

# Daoy CAMP Responsiveness



# Clustering of data from 8-Br-CAMP treatment

## Data Dimensionality

- 233 descriptors (parameters)
- 300 αbs
- 233 descriptors x 300 αbs = 70,000 data points/treatment (similar scale to microarray)

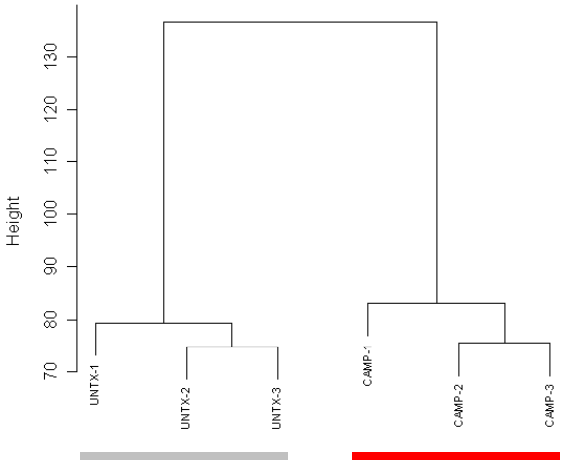
## Analysis

- Principle components analysis (PCA) used to identify the first 10 linear combinations of antibody characteristics that maximize variance explained.
- Cluster the data, see how samples segregate on basis of treatment.

## Result

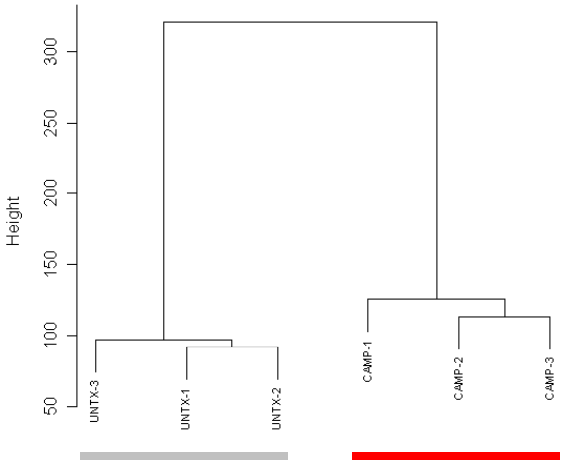
- 75% of the variance is explained by the first 10 components.
- All 10 components form discrete clusters when using unsupervised hierarchical clustering.

PC1



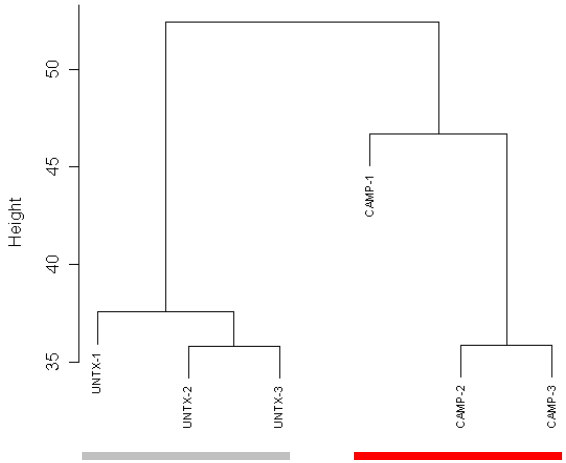

Untreated + CAMP

PC2…




Untreated + CAMP

PC10




Untreated + CAMP

# Generalization Across Multiple Distinct Drug Classes

## Experimental Design

**42 drugs in 9 drug classes, quadruplicate replicates**
**160 plates (61,440 wells), 300 αbs (Yale Center for Genomics & Proteomics)**
**368,640 images, 1.2 terabytes of data (CNNR Cell Imaging Resource)**
**70 hrs supercomputer time (90 nodes) (262 CPU days) (HPCC)**
**Random Forest analysis**

## Drug Classes

**adenosine agonist**
**aldose reductase inhibitor**
**COX inhibitor**
**dihydrofolate reductase inhibitor**
**DMSO control**
**glucocorticoid receptor agonist**
**phosphodiesterase inhibitor**
**protein kinase A activator**
**vitamin D receptor agonist**

Predicted Class

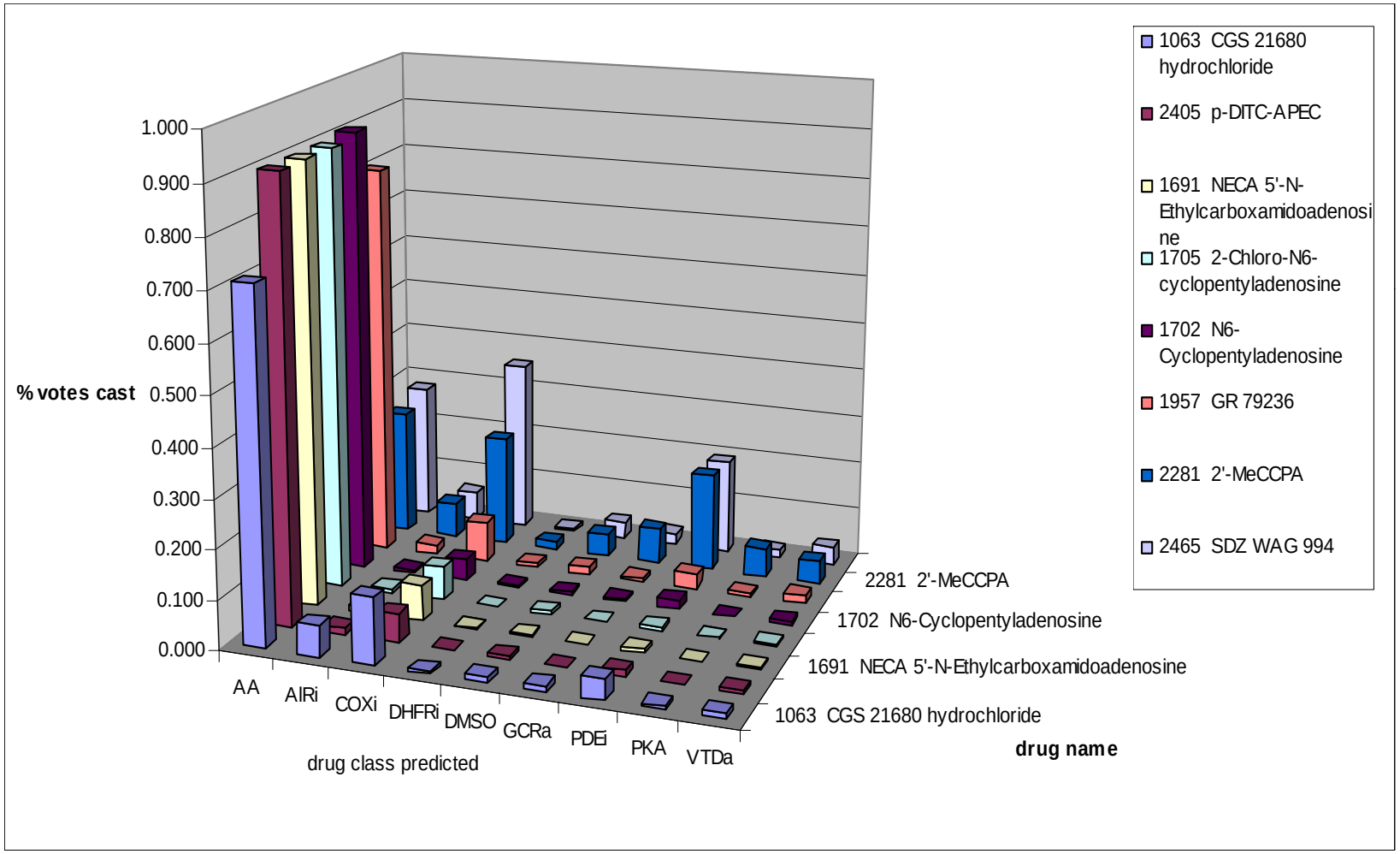
| True Class | AA | AlRi | COXi | DHFRi | DMSO | GCRa | PDEi | PKA | VTDa | class error |
|---|---|---|---|---|---|---|---|---|---|---|
| AA | 28 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 0 | 0.125 |
| AlRi | 0 | 10 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0.166 |
| COXi | 0 | 0 | 21 | 0 | 0 | 0 | 7 | 0 | 0 | 0.250 |
| DHFRi | 0 | 0 | 0 | 7 | 1 | 0 | 0 | 0 | 0 | 0.125 |
| DMSO | 0 | 0 | 0 | 0 | 15 | 0 | 0 | 0 | 1 | 0.062 |
| GCRa | 0 | 0 | 0 | 0 | 0 | 12 | 3 | 0 | 1 | 0.250 |
| PDEi | 0 | 0 | 2 | 0 | 0 | 2 | 19 | 0 | 1 | 0.208 |
| PKA | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16 | 0 | 0.000 |
| VTDa | 0 | 1 | 0 | 0 | 0 | 0 | 2 | 0 | 13 | 0.187 |

**composite accuracy: 84.0%**

Conclusion: Multiple Drug Classes Can be Accurately Discriminated in a Single Experiment by a Single Objective Statistical Model

# Prediction Distributions

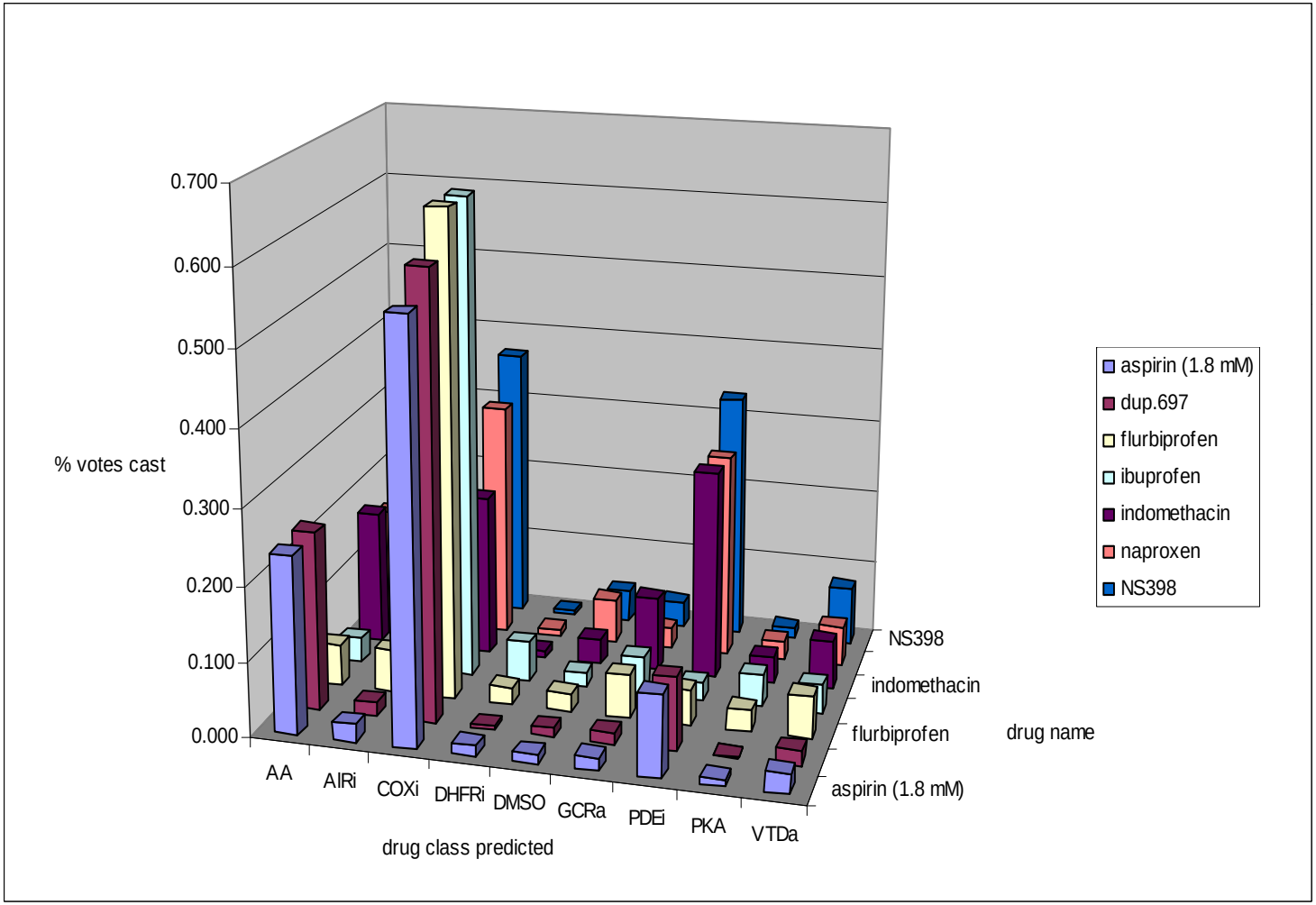
## ADENOSINE RECEPTOR AGONISTS




Six strong predictions, two weak ones (one misprediction)

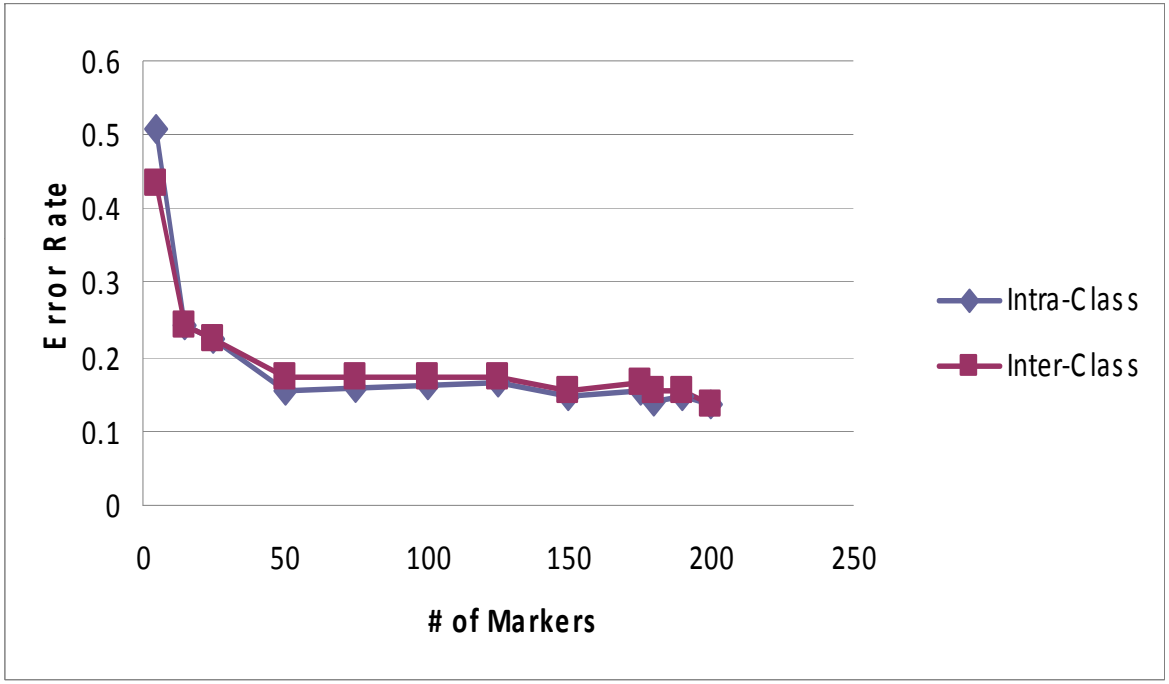
# Prediction Distributions

## CYCLOOXYGENASE INHIBITORS




three drugs with PDE inhibitor class votes

## Sufficient antibody number = 50




six-fold increase in throughput efficiency

## Scale-up:

A) create as many (therapeutically relevant) physiologies as possible

B) Screen unknowns against panel of physiology fingerprints

Easily manipulable targets expressed by DAOY cells

### 45 Receptors for:

| | |
|---|---|
| EGF | Interferon a2 |
| Eurythropoietin | IGF1, 2 |
| Oncostatin | Insulin |
| Activin A | Atrial naturetic peptide |
| Inhibin | TGF a |
| BMP2, 7 | TGF b |
| GDF9 | Hepatocyte GF |
| Cardiotrophin | Angiopoietin |
| LIF | IL6, 7 |
| TNF $\alpha$ | Semaphorin |
| TNF $\beta$ | CGRP |
| Pleiotrophin | SDF |
| PDGF | Adiponectin |
| Melanocortin | Ghrelin |
| Relaxin | Cholecystokinin |
| Nodal | Robo |
| FGF1 | VEGF |
| FGF2 | Netrin |
| RGMB | etc. |

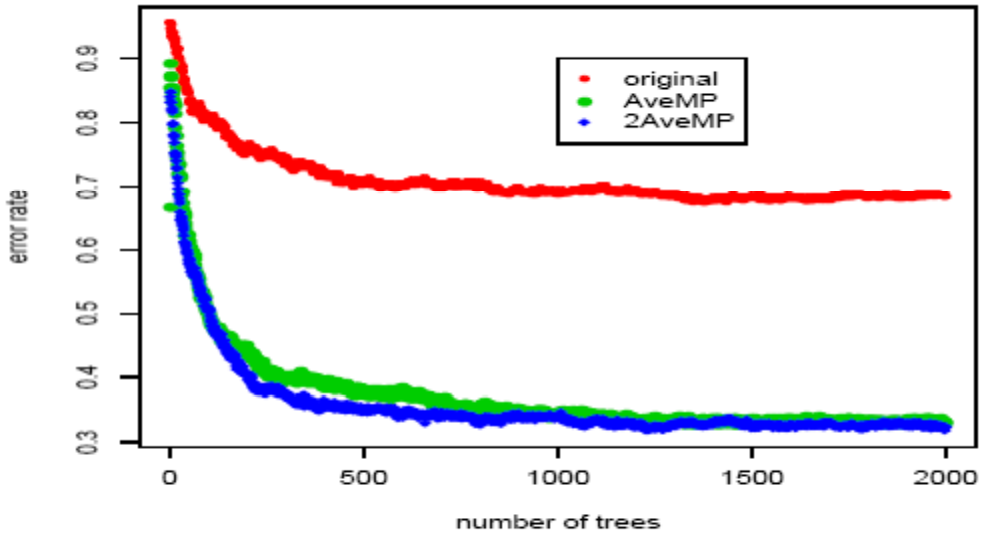
### 60 Targets with available drugs:

| | |
|---|---|
| ADRB2 beta 2 | GABBR1 |
| Butyrylcholinesterase | histone deacetylases |
| lipoprotein lipase | MAOA |
| carboxypeptidase A4 | cytochrome P450 |
| carbonic anhydrase | cysteinyl leukotriene receptor 2 |
| endothelin receptor B | glucocorticoid receptor alpha-2 |
| oxytocin receptor | ceramide glucosyltransferase |
| citrate transporter | ATPase, Na+/K+ trans, alpha 1 |
| adenosine receptors | dihydroorotate dehydrogenase |
| beta tubulin 1 | epoxide hydrolase |
| 5-alpha-reductase | guanylate cyclase |
| Tissue plasminogen activator | Vitamin K reductase |
| prostaglandin E receptor 2 | vacuolar ATP synthase |
| IkB kinase-a | PPAR alpha |
| NFKB | PPAR gamma |
| topoisomerase (DNA) II alpha | NMDAR 2A |
| retinoic acid receptor | vitamin D receptor |
| PDE3, 4 | Cycloxygenase 2 |
| PKA | etc. |

45 receptors + 60 targets = 105 physiologies

# 105 drug class experiment

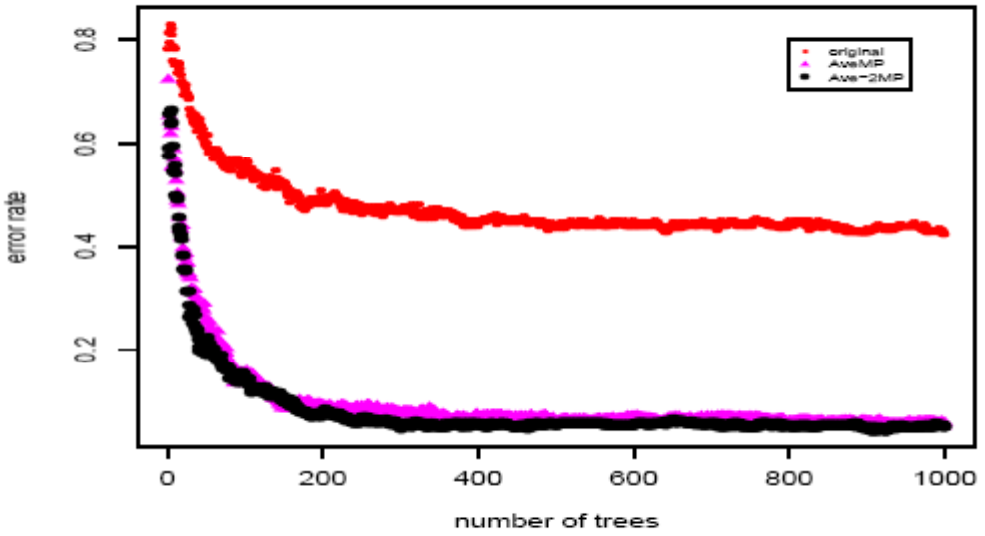
Error Rate Trace: All Drugs




un-normalized: 32% accuracy

normalized: 67% accuracy

Error Rate Trace: 24 Top Drugs




un-normalized: 55% accuracy

normalized: 95% accuracy

Ruiyan Luo

## SUMMARY

- Drug action can be precisely fingerprinted with high-dimensionality data (e.g. microarrays)
- Multiparametric cytometry provides suitably high-dimensionality proteomic data
- Single platform for recognizing hundreds of distinct drug physiologies in a single experiment

## FUTURE

- Screening novel compounds for therapeutic profiles
- Screening active compounds for toxic/undesirable profiles (e.g. addictive potential)
- Elucidating mechanisms of therapeutic/toxic/undesirable profiles

# Acknowledgements

Yale Center for Genomics & Proteomics

Janie Merkel & Mike Salcius

Center for High Performance Computing in Biology and Biomedicine

Rob Bjornson

Biostatististicians

Bob Gerwien, Hongyu Zhao & Ruiyan Luo