Inferring Genetic Architecture from "Systems Genetics" Studies

Norbert Perrimon

Harvard Medical School

Q.1 How do we systematically analyze the function of genes?

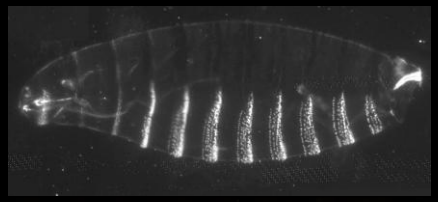
Q. 2 How do we assemble gene products into networks?

Drosophila genome: 15,000 protein coding genes

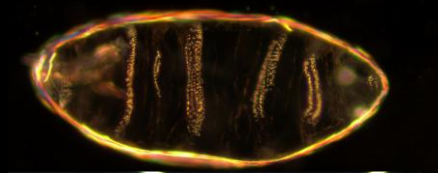
Drosophila cells: 7,500 genes

Forward Genetic Screens to identify pathway components

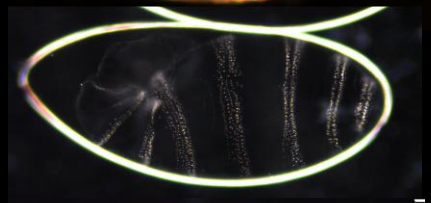
Specific and Penetrant Phenotypes



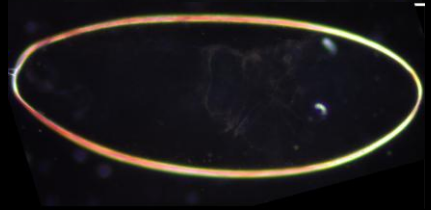
Wild type



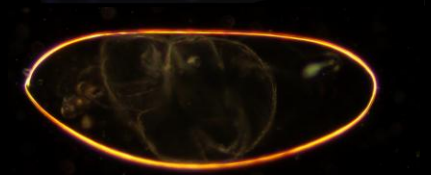
Unpaired



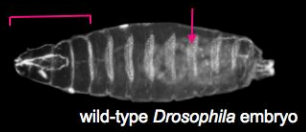
Decapentaplegic

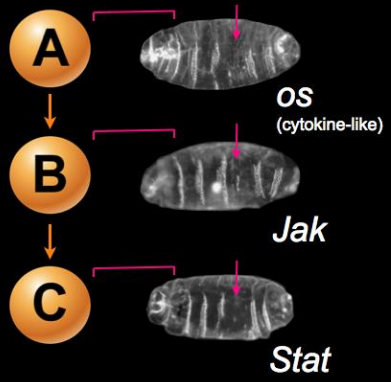


Notch

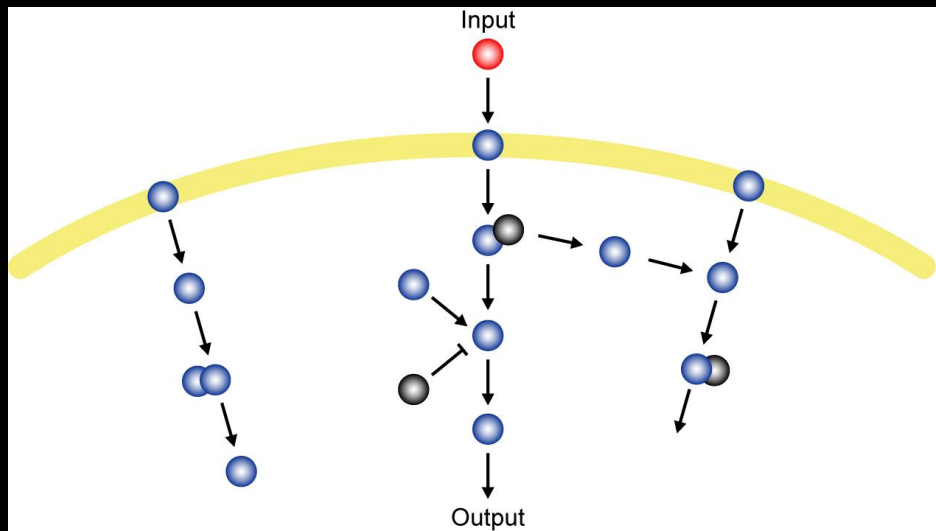


dorsal





Hypothesis: Genes with similar specific mutant phenotypes encode components of the same biochemical pathway.

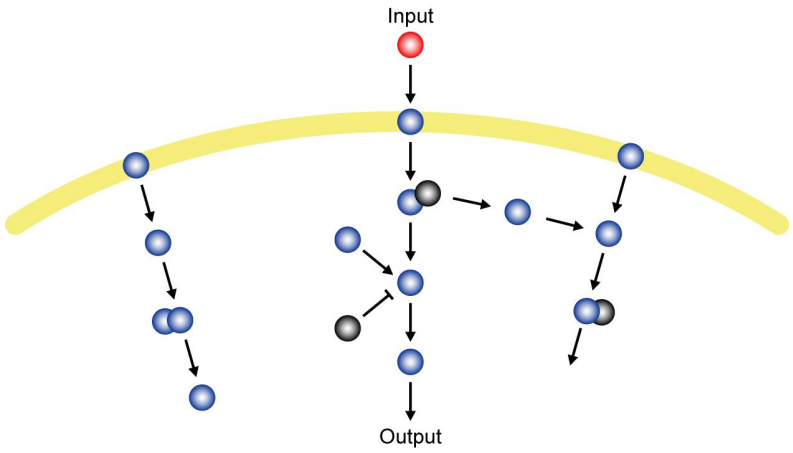


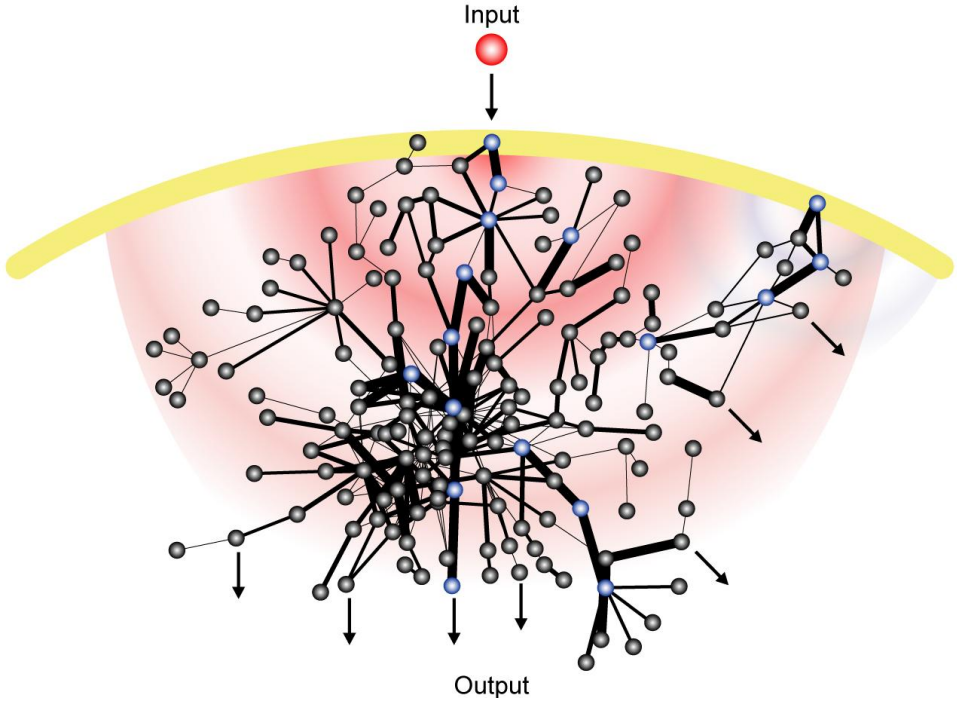
After 2000: Analyses of signaling networks

Classic Genetics
One gene/protein at a time
"Canonical Pathways"

RNAi
Genome Sequences

Omics
Global Network Analysis
"Signaling Networks"



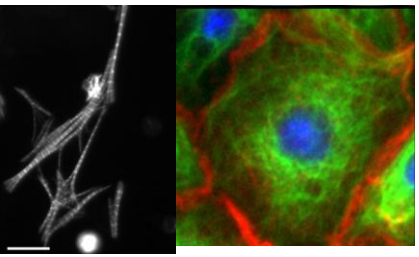


Overall Strategy and Approaches

From "Omics" in tissue culture cells

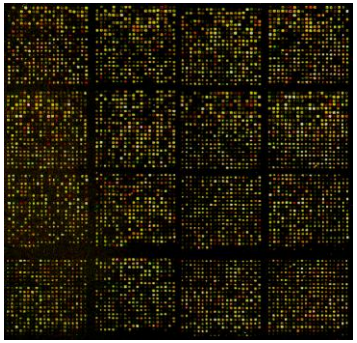
To Biology

Cell-based RNAi screens

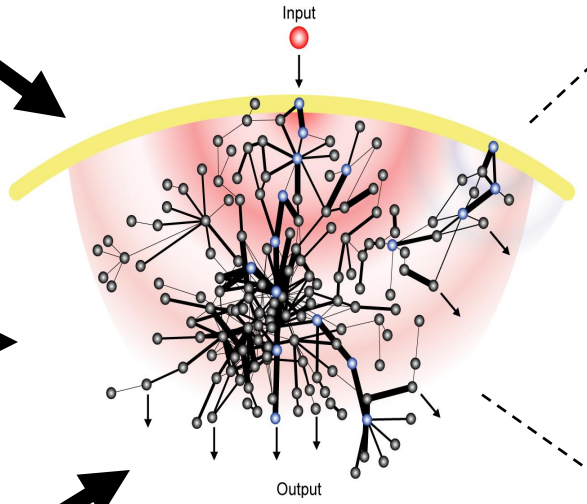


Mass Spec

Transcriptome
(Array, RNAseq, Nanostring)

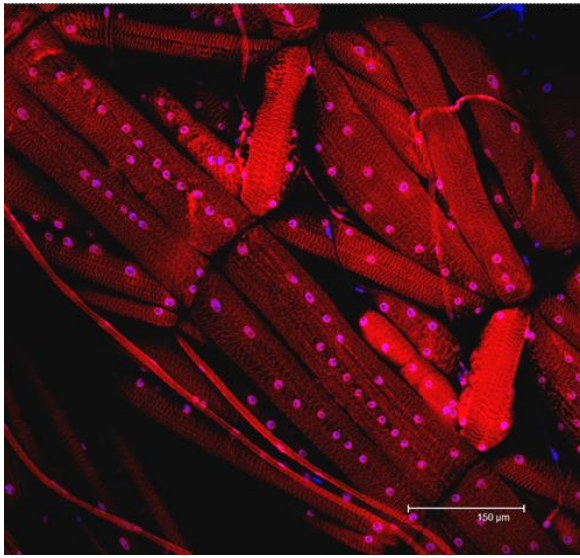


Network Topology



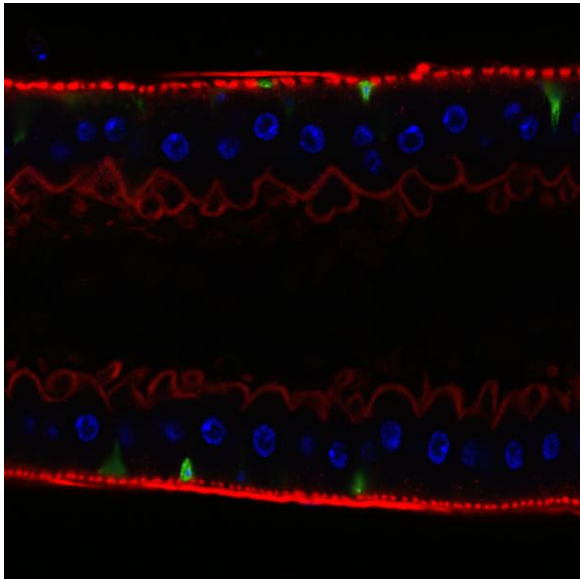
Computational tools

Muscle Growth



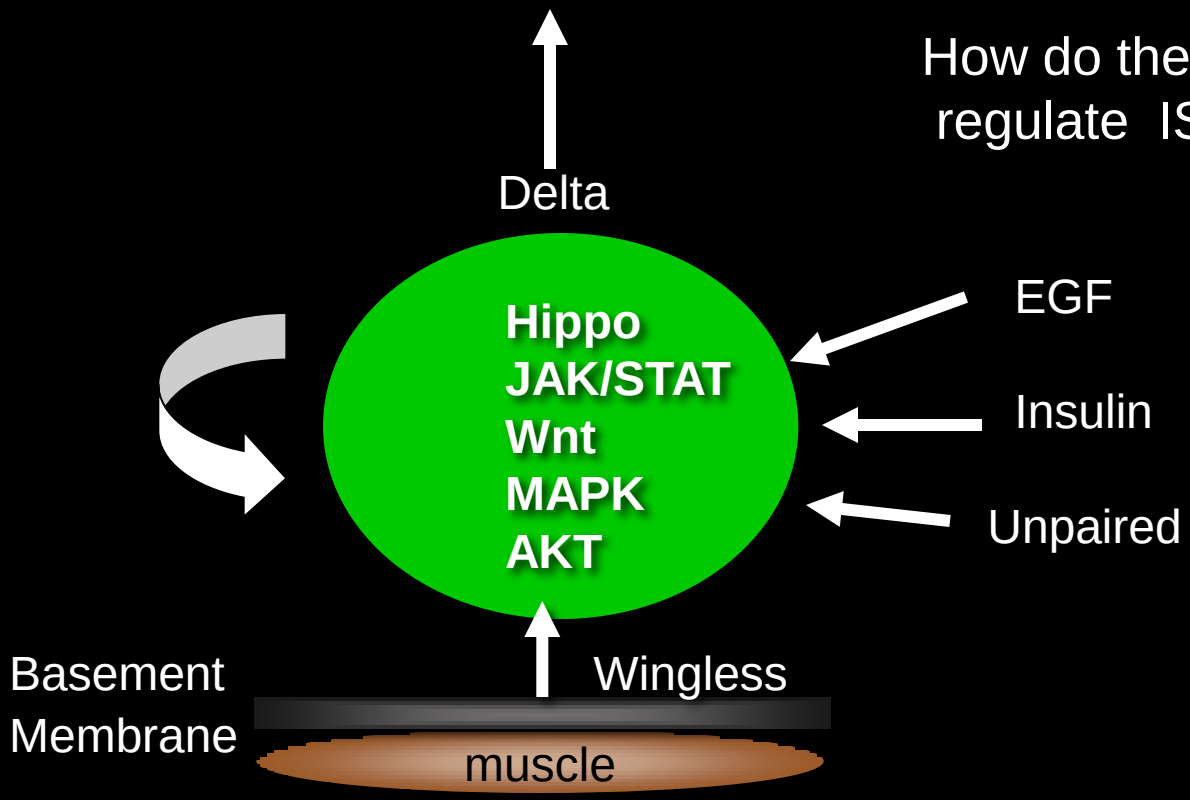
Validation by transgenic RNAi

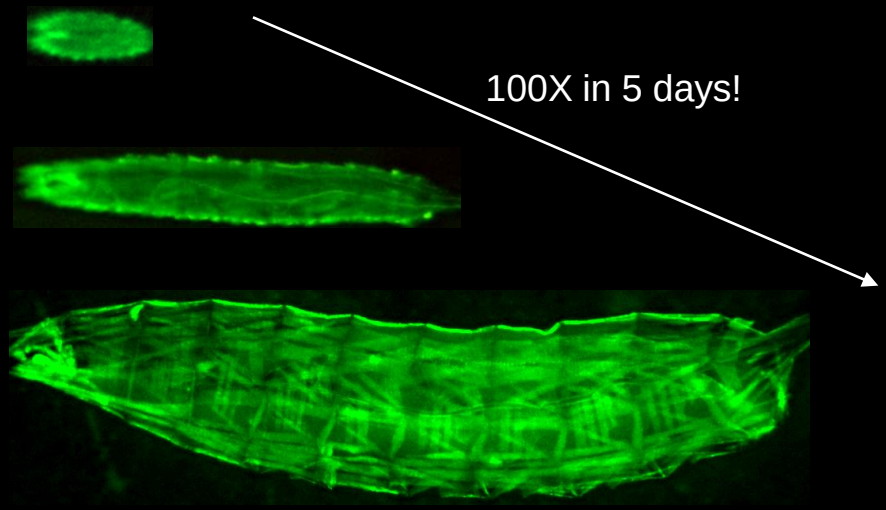
Gut Stem Cells



Signaling pathways integration and cross-talk

How do the various pathways that regulate ISCs proliferation interact ?

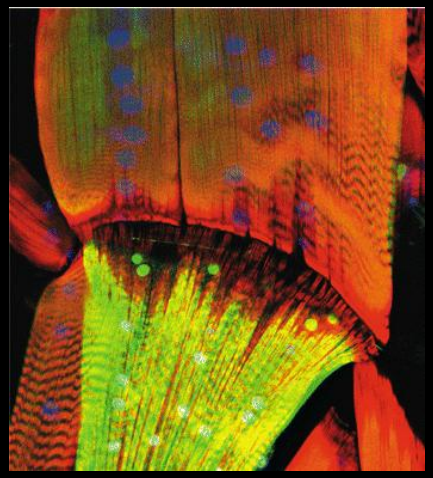




Drosophila larval muscle growth: A system to understand how the developmental programs of complex tissues are influenced by genetic and environmental factors

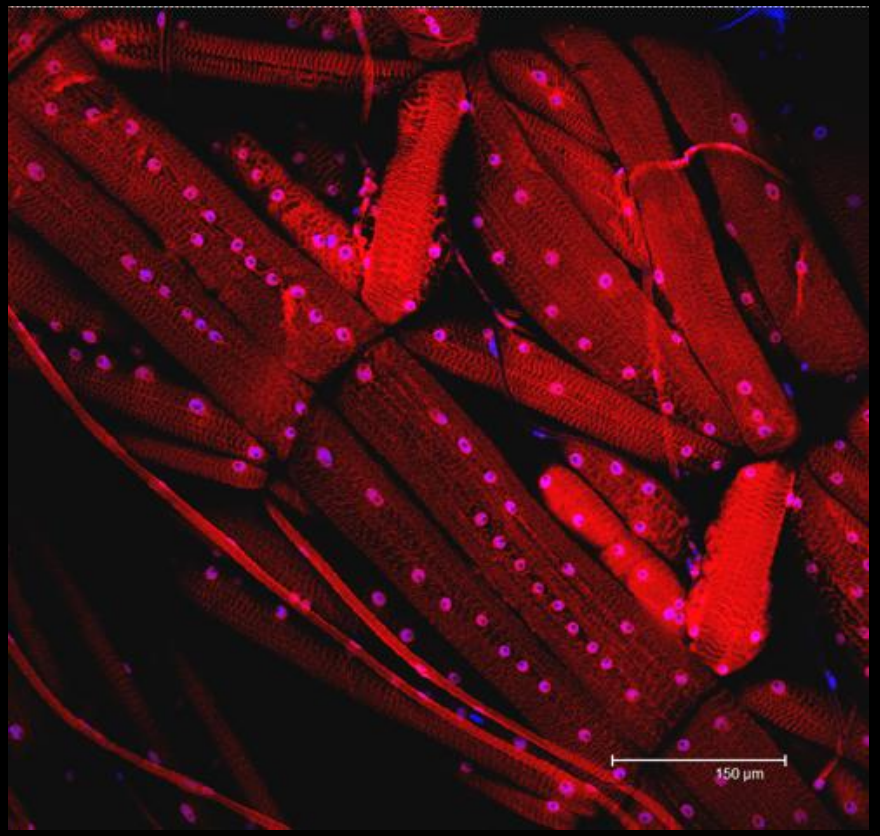
- **tremendous cytoplasmic growth**
- **stereotype pattern**
- **multinucleated cells**
- **abundant tissue**

P157-Gal4
UAS-PTEN
UAS-GFP



A decrease in Insulin signaling impairs growth (Demontis and Perrimon, 2009)

Insulin ⊣ Foxo ⊣ Dmyc → Growth



Inferring Genetic Architecture from "Systems Genetics" Studies

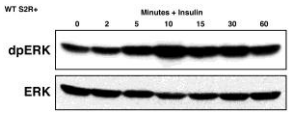
1.Building high confidence Networks

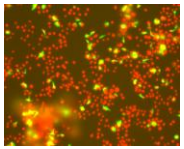
. RNAi
. Mass Spec
. Transcriptomics

2. Network analyses

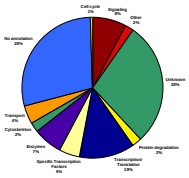
. Protein complexes
. Genetic interaction signatures "Epistasis-map"
. Flow
- transcriptome signatures
- phosphorylation signatures

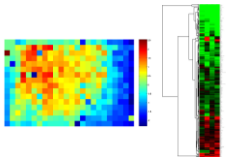
Genome-wide RNAi Screening Platform

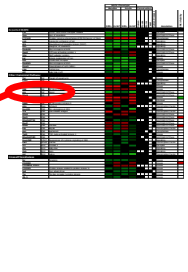




384 well plates







Assay development: optimization and validation

Quantitative screens: Plate Readers

Qualitative screens: Autoscope

Full genome screen (x 2) using *DRSC 2.0*

(17,000 dsRNAs)

Data normalization and analysis

Select top 1% for secondary tests

Confirmation with independent

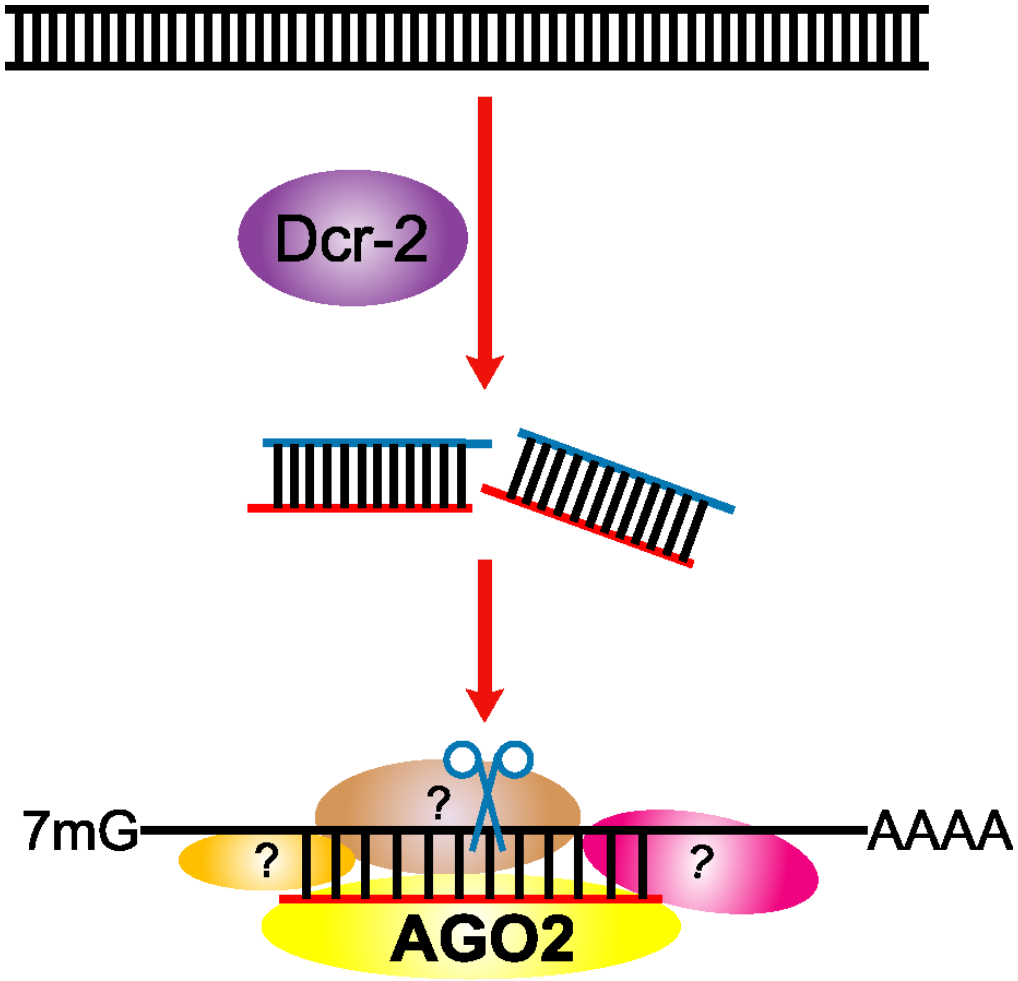
Additional assays

Final hit table

Database mining

Validation

Drosophila RNAi in tissue culture cells



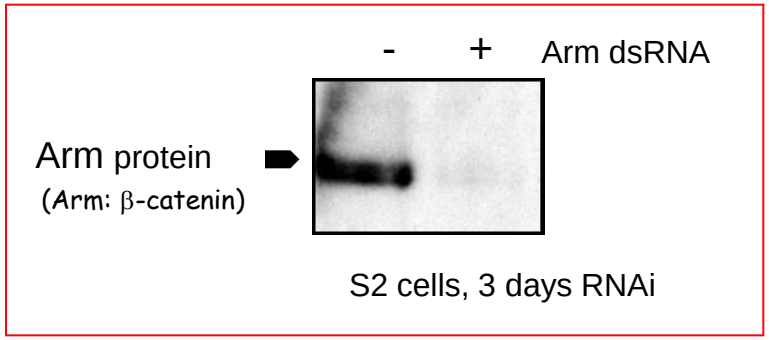
Add dsRNA to serum-free medium

(most cell lines do not require transfection)

3-5 days

Partial to complete loss of protein

Uniform penetrance



Approaches to describe cellular phenotypes in tissue culture

Ligand

↓

Receptor

↓

Kinase

↓

Transcription factors

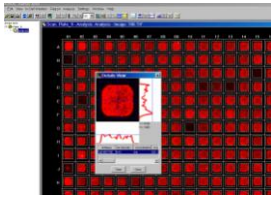
↓

Gene Expression

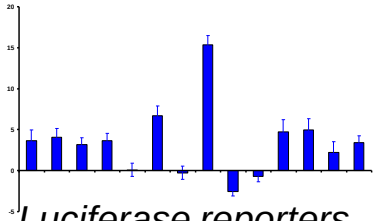
↓

Cell morphology

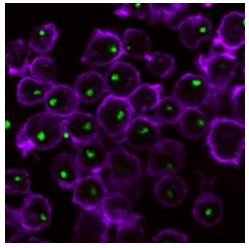
Assays



Phospho-Antibodies



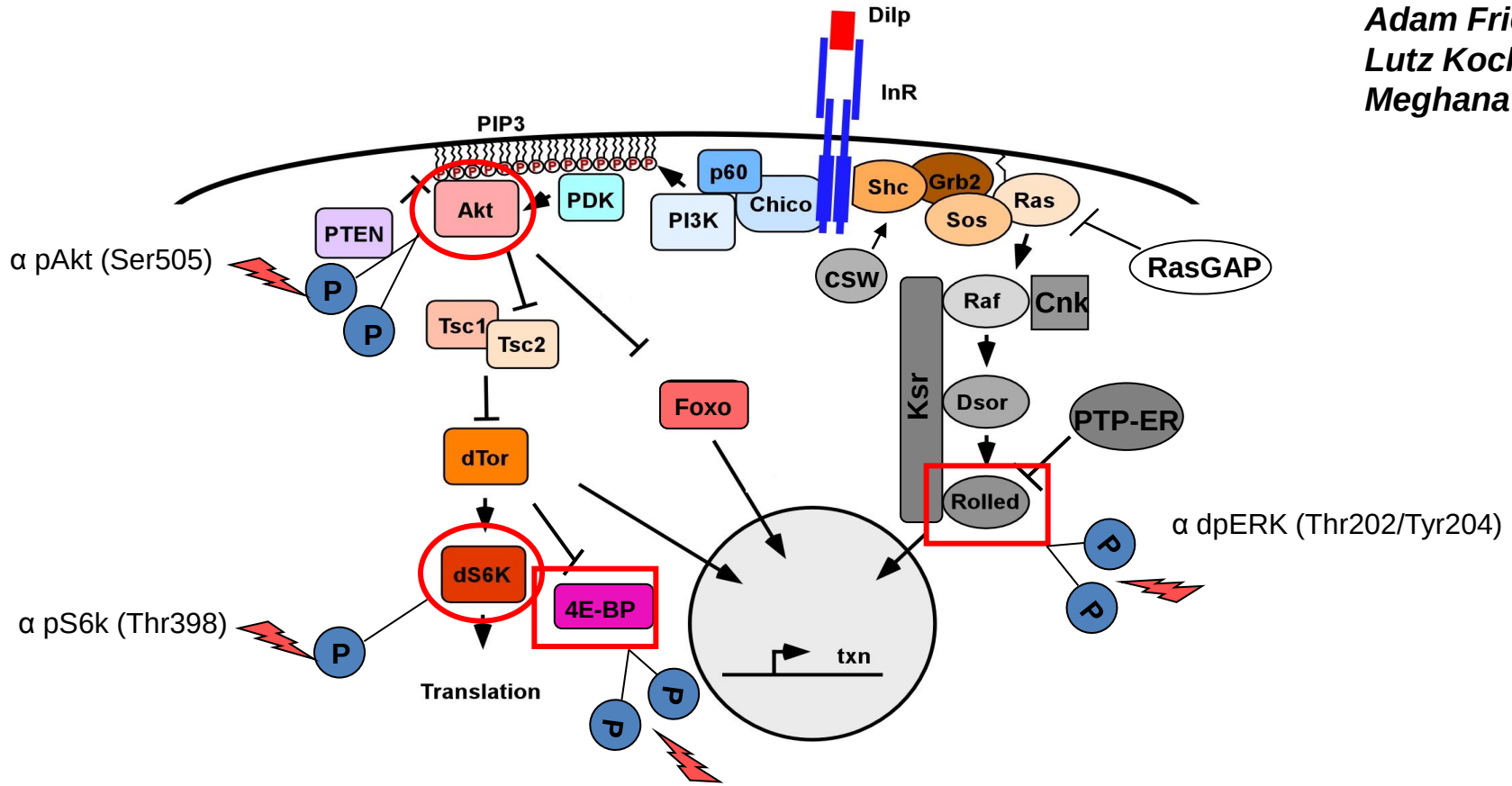
Luciferase reporters



Simple image-based screen

RTK signaling protein interaction network

Adam Friedman
Lutz Kockel
Meghana Kulkarni



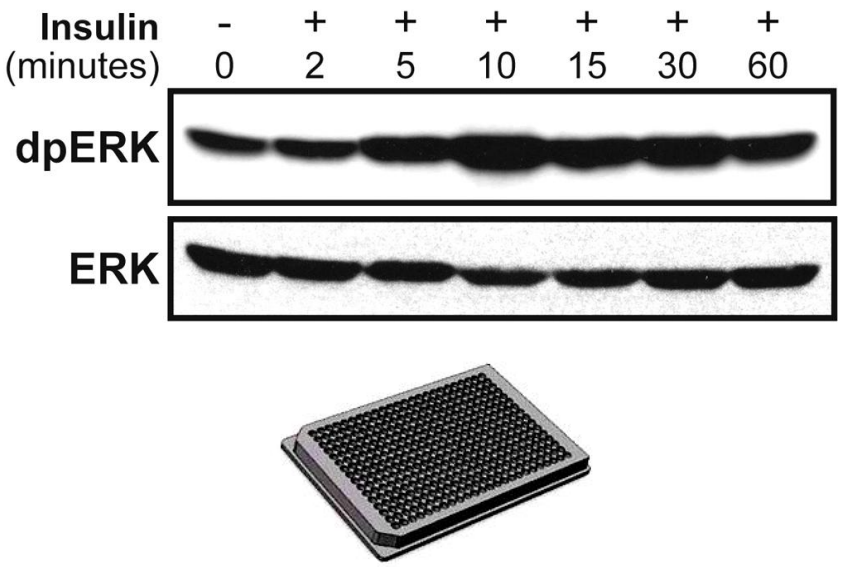
Glucose uptake and utilization

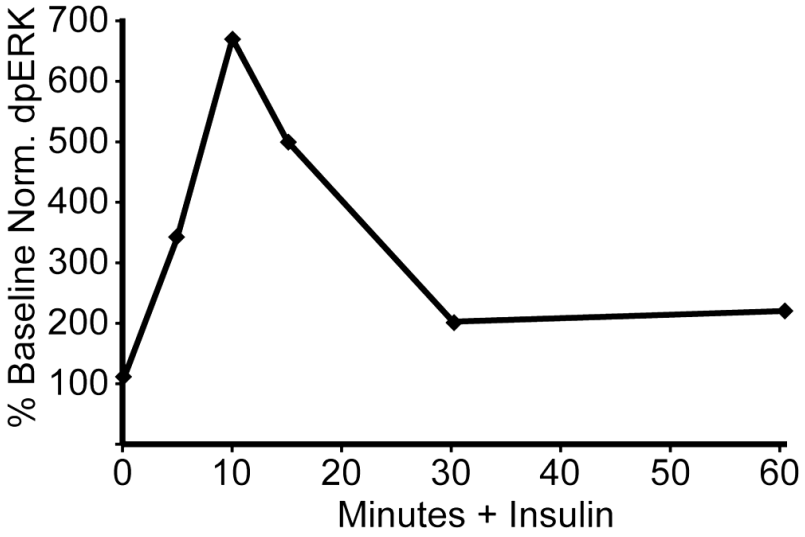
Lipid synthesis

Protein synthesis

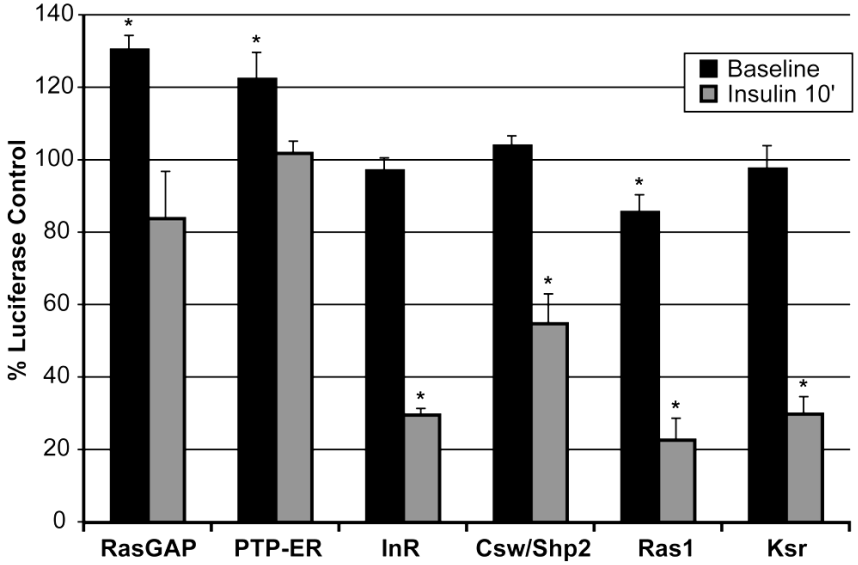
Growth and gene expression

MAPK/ERK RNAi screen: assay

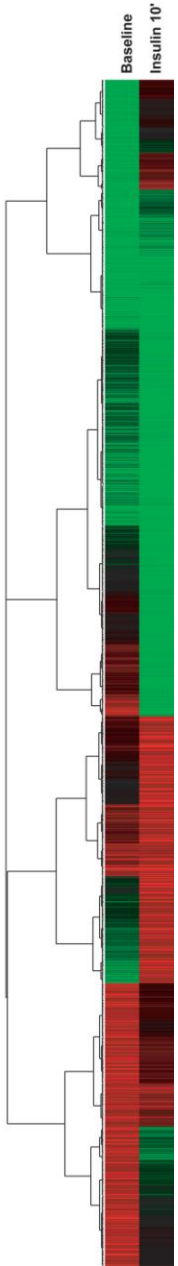


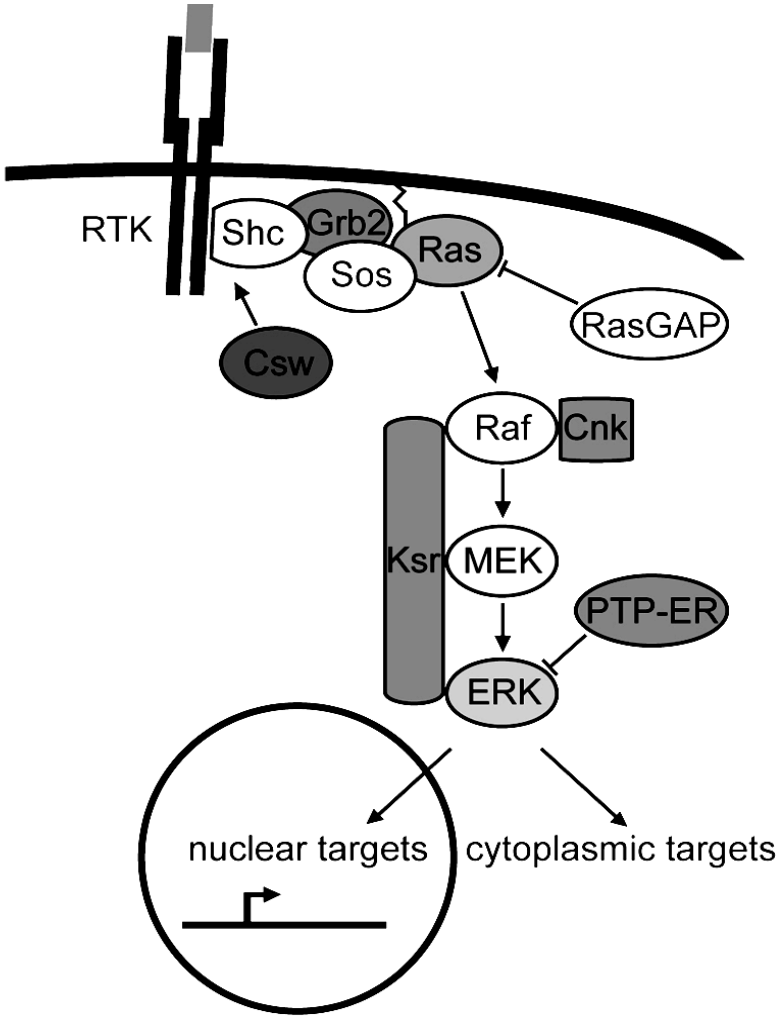


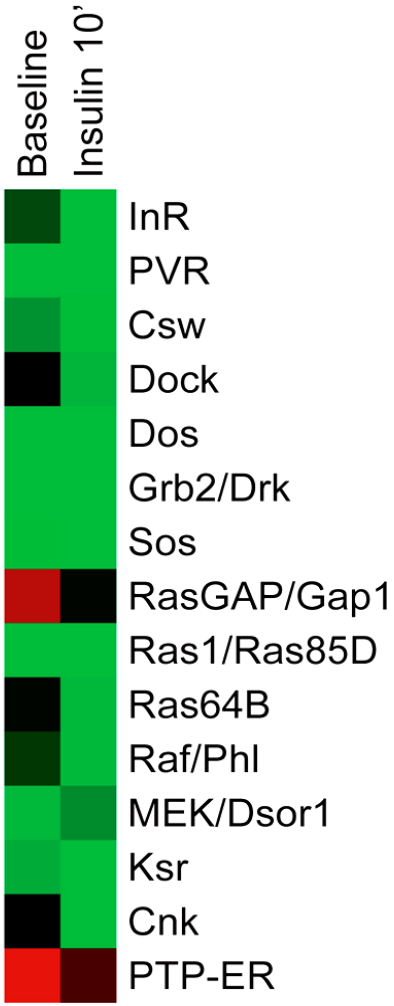
- Seed cells in 384-well plates with DRSC RNAi collection (available at http//flyrnai.org)
- 4 day incubation
- Stimulation with insulin, fixation, staining with fluorescently-conjugated dpERK antibody
- Normalization to total ERK, data analysis and filtering
- dsRNA scored as *Z* - score



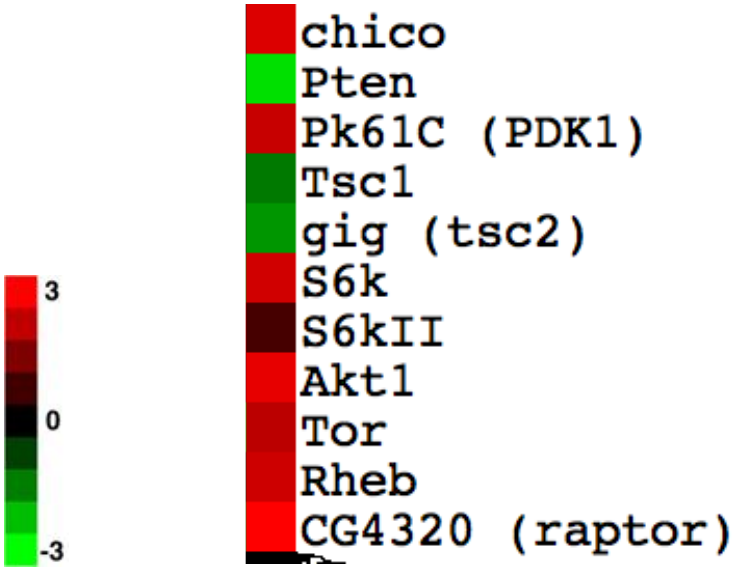
Functional RTK/ERK RNAi screen

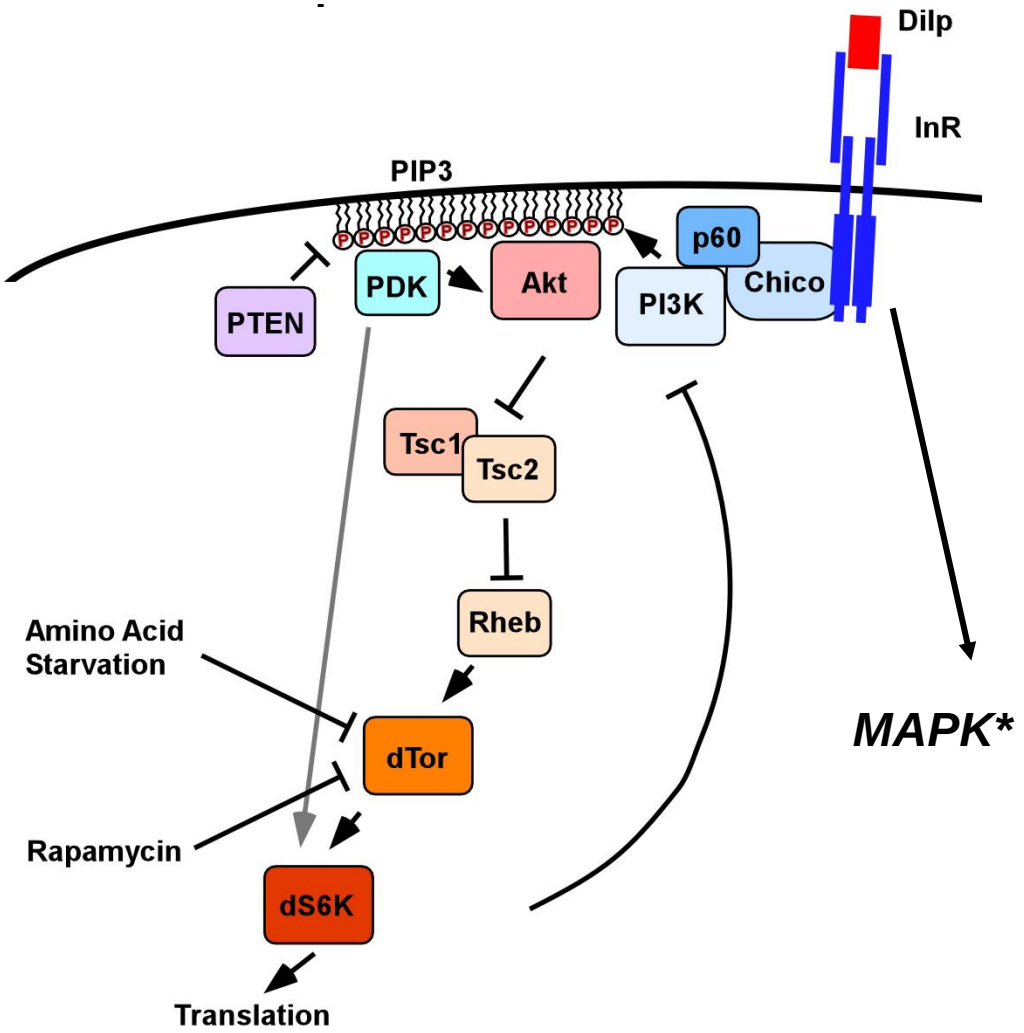


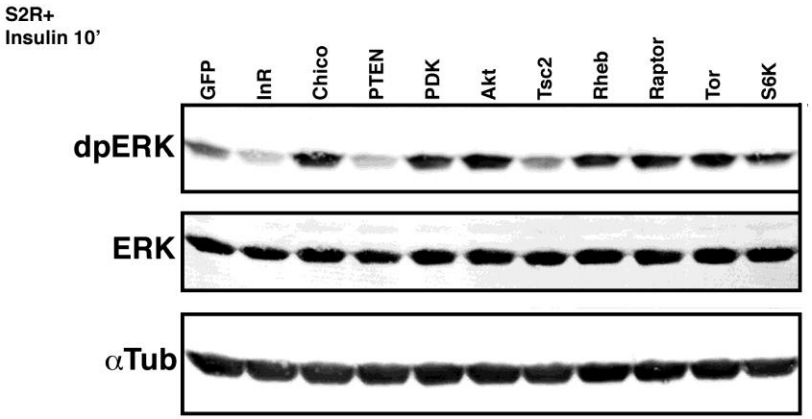




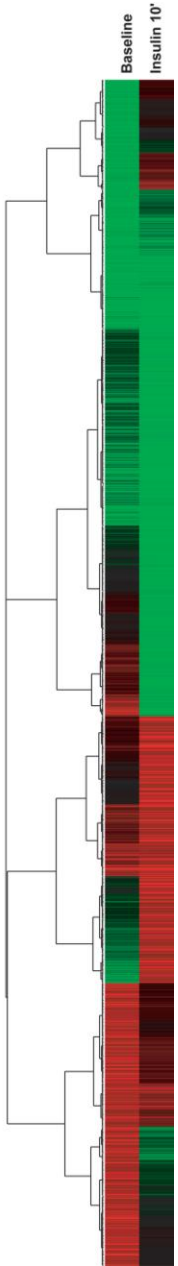
MAPK Phosphorylation captures both MAPK and AKT regulators because of the S6K feedback loop

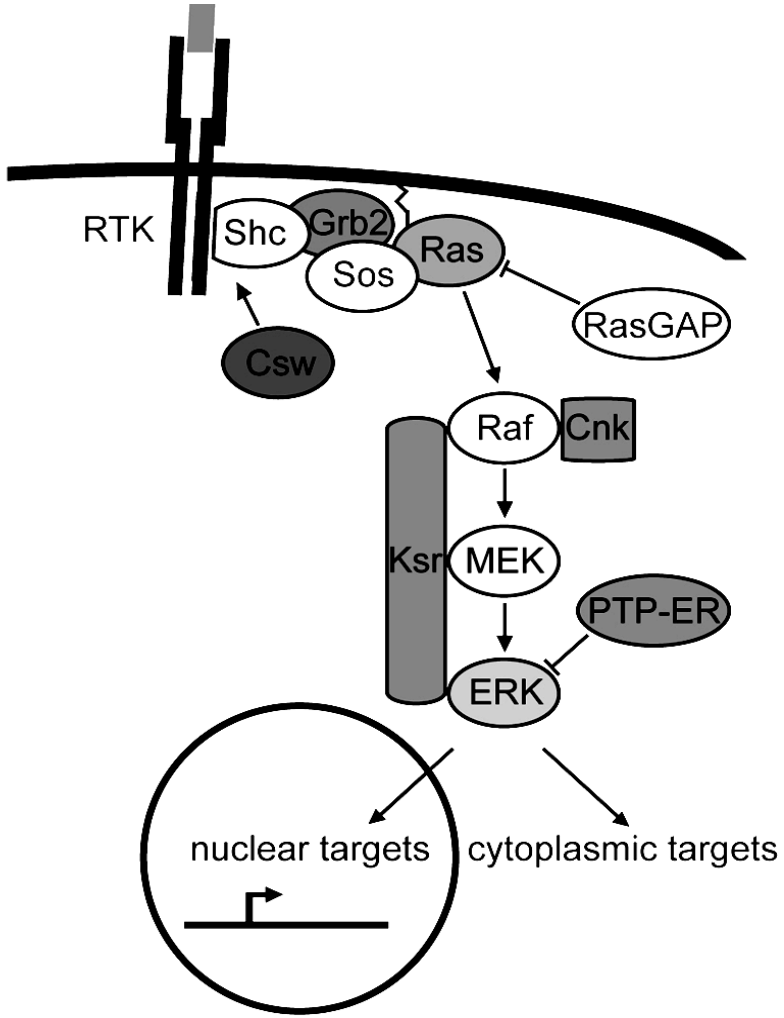


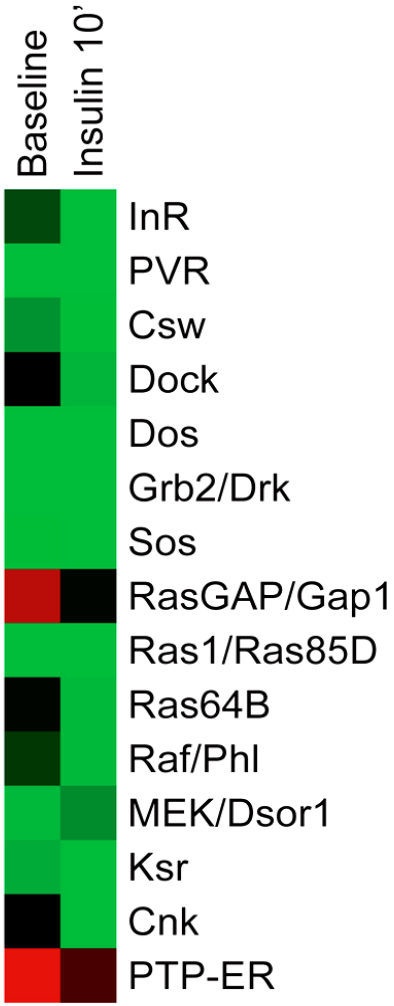




Functional RTK/ERK RNAi screen

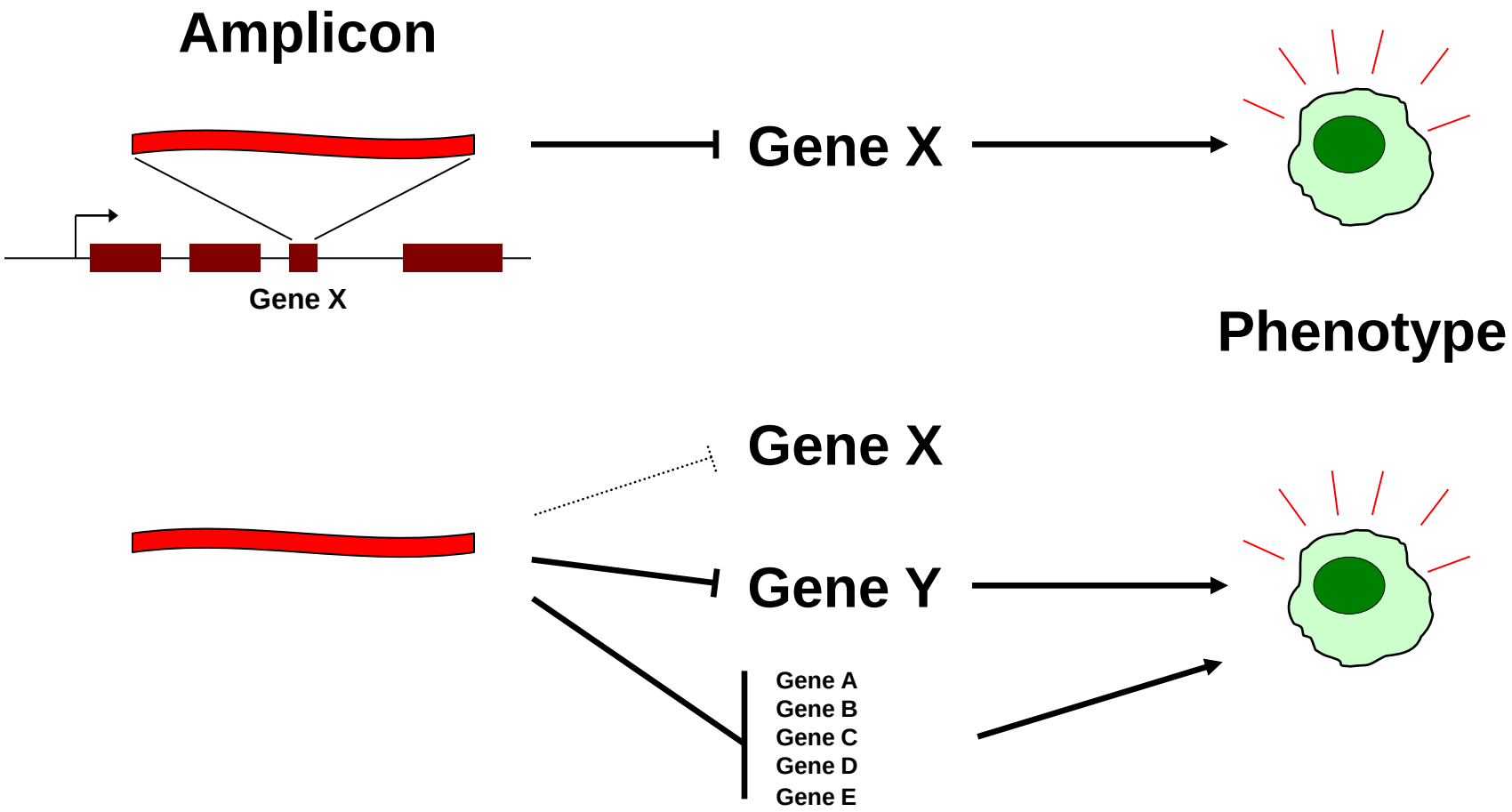






- > 350 putative regulators:
- Which hits are false + and which ones act indirectly?
- Which hits are InR-specific or cell type specific?

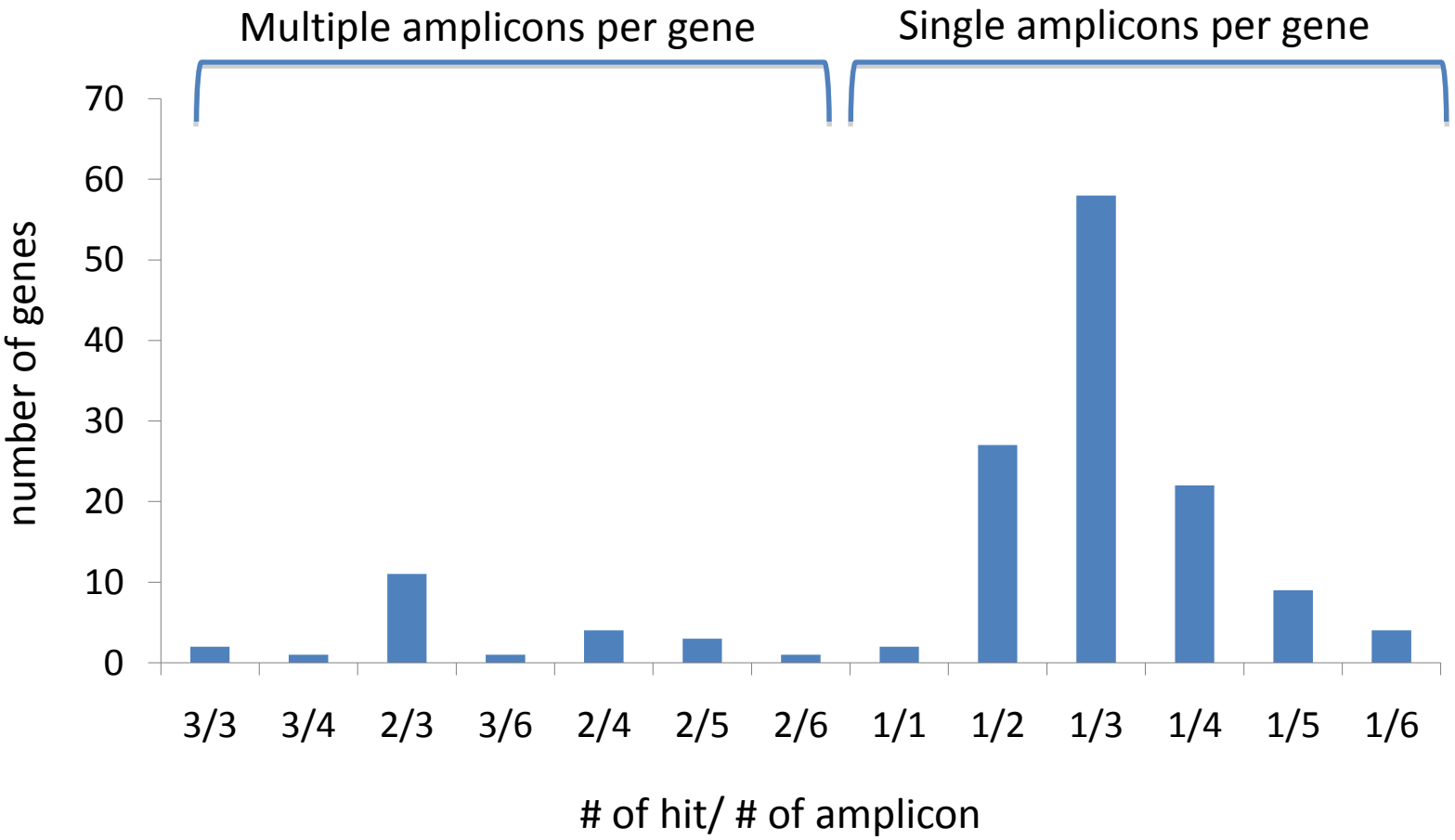
RNAi as a genetic reagent



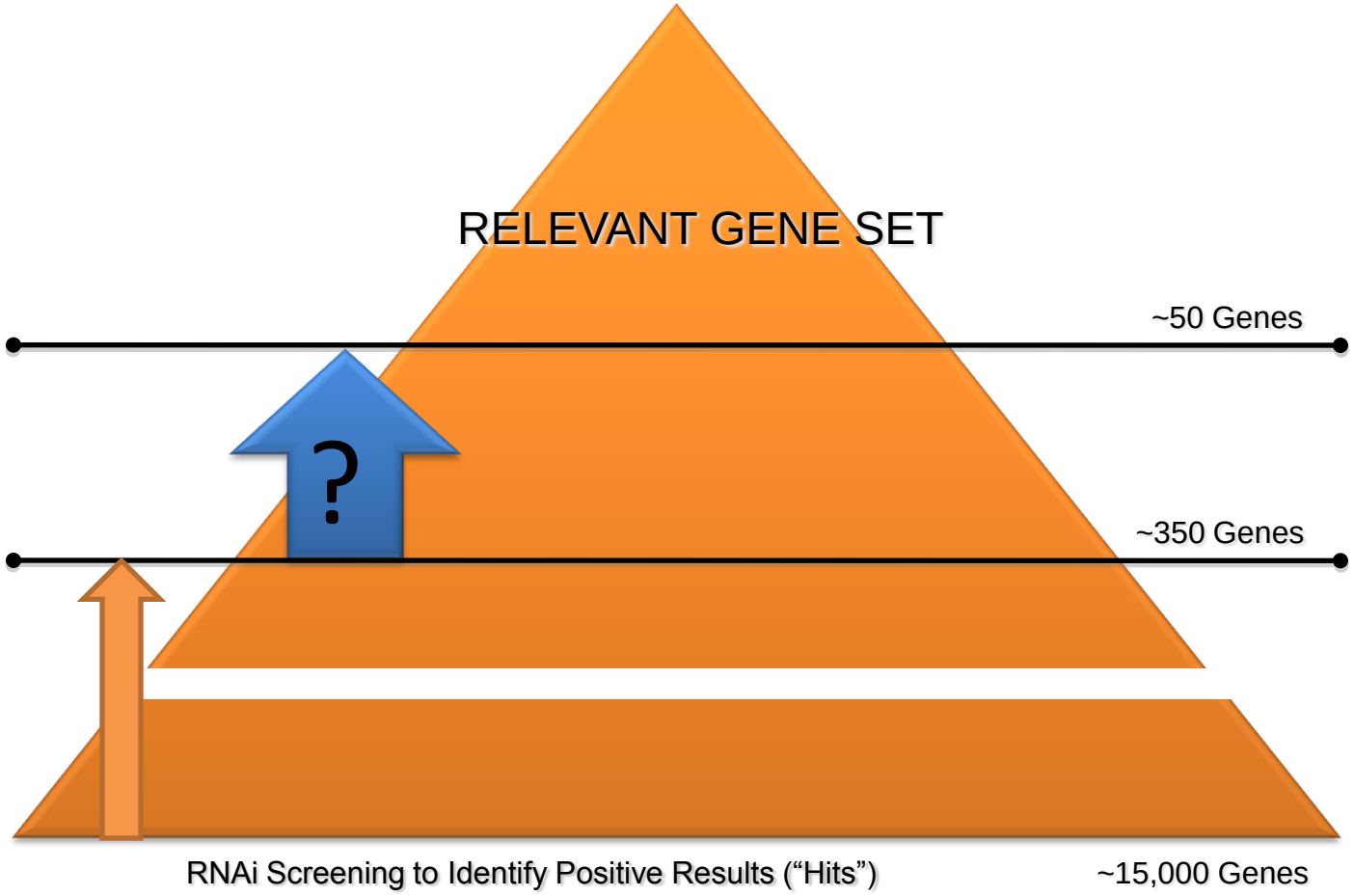
- "Off-target" effect (OTE) dissociates **Amplicon** from **Gene**, and thus restricts dsRNA as a genetic tool.

Distribution of hits

Kinase & Phosphatase set
- 468 genes
- average number of amplicons per gene: 3



Bottom-Up Genome-wide RNAi approach to Pathway analyses



. Multiple dsRNAs

. qPCR

. expression

. Internal validation (protein complex)

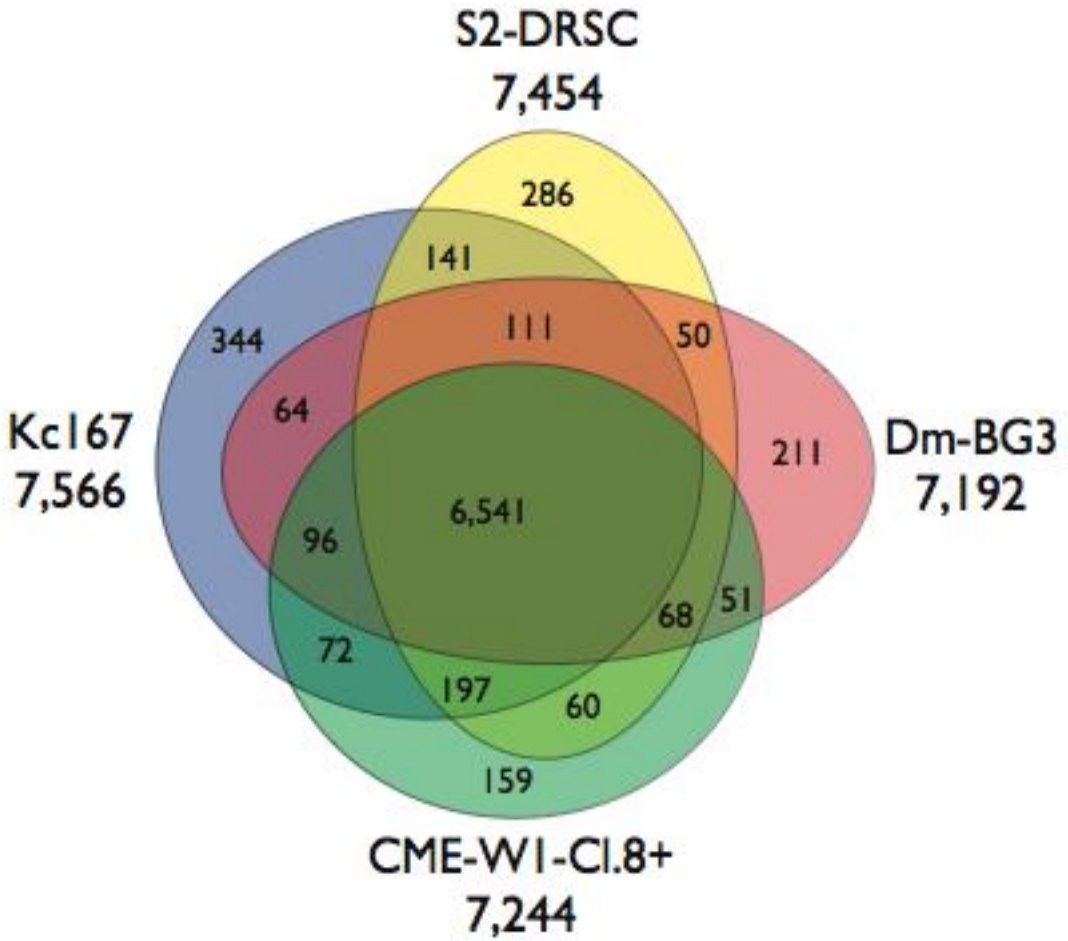
. Literature

. Mass Spec

. Rescue

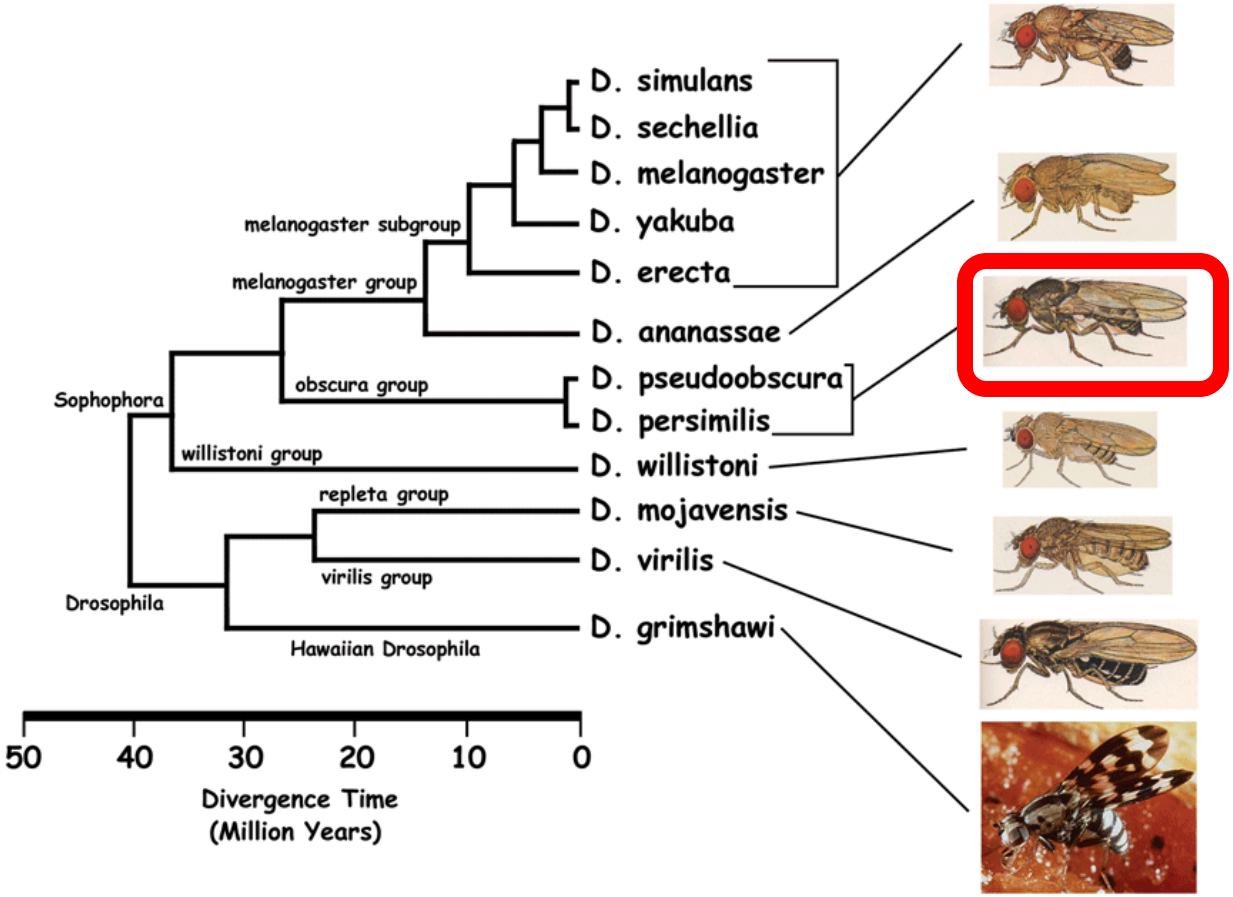
Gene expression in cell lines: Solexa sequencing

Anastasia Samsonova
in collaboration with modEncode



- *50-55% of the genes are expressed in one cell line*
- *85 -90% overlap*

Gold Standard in RNAi screens: *Rescuing RNAi-induced phenotypes in D. melanogaster cells by genomic DNA fragments of D. pseudoobscura*

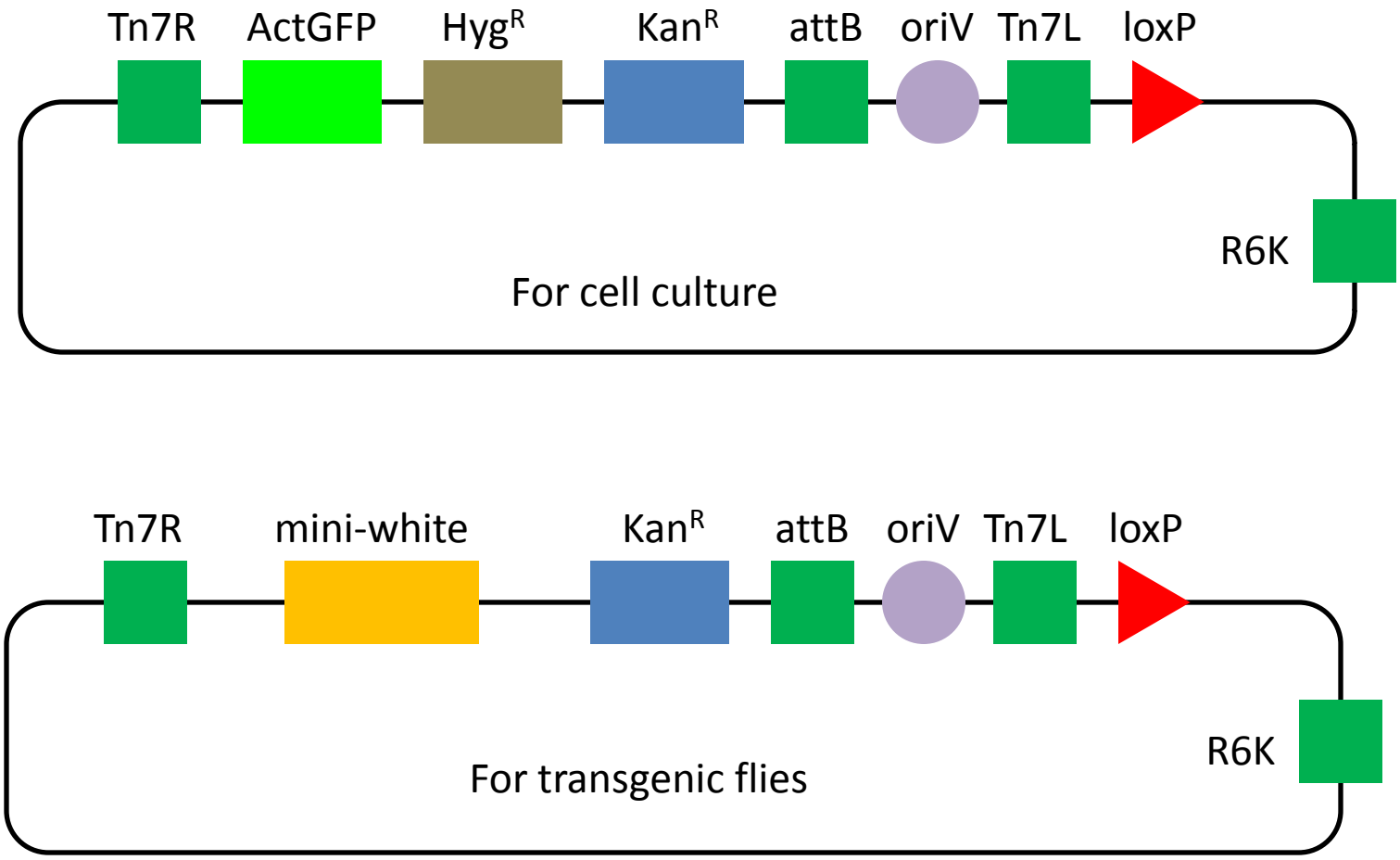


D. melanogaster and *D. pseudoobscura* diverged about 20 million years ago

More than 90% of *D. melanogaster* genes have an ortholog in *D. pseudoobscura*.

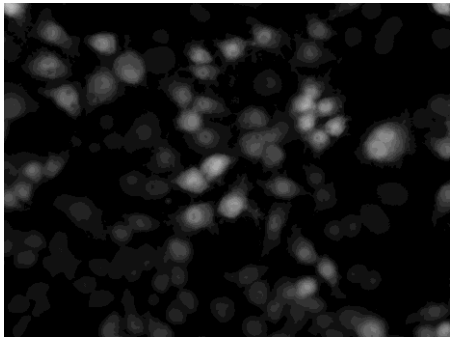
Even when the protein sequences are perfectly conserved, DNA sequences are less conserved due to accumulation of synonymous mutations.

Vectors for rescue of tissue culture and in vivo RNAi phenotypes

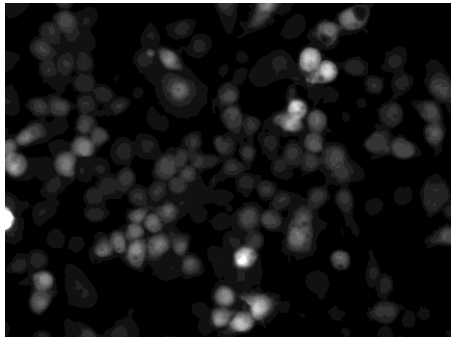


Shu Kondo

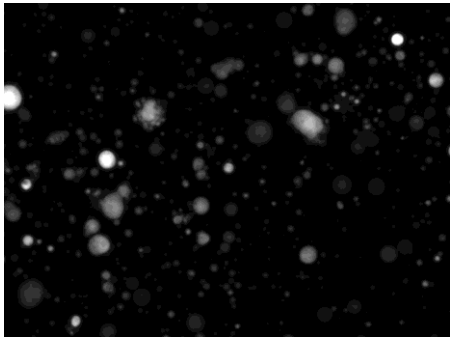
Rescuing th (DIAP1) dsRNA-induced phenotype



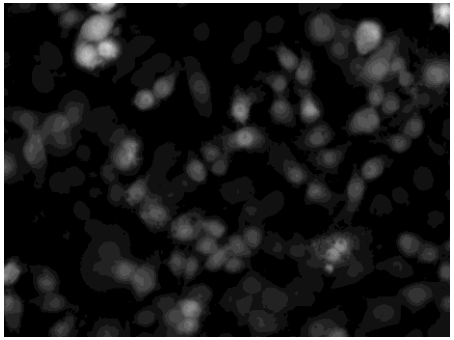
Control fosmid
+DW



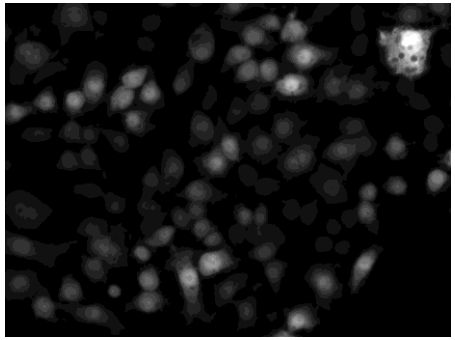
Control fosmid
+luc dsRNA



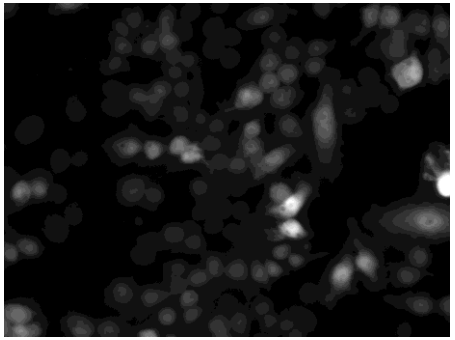
Control fosmid
+th dsRNA



th fosmid
+DW

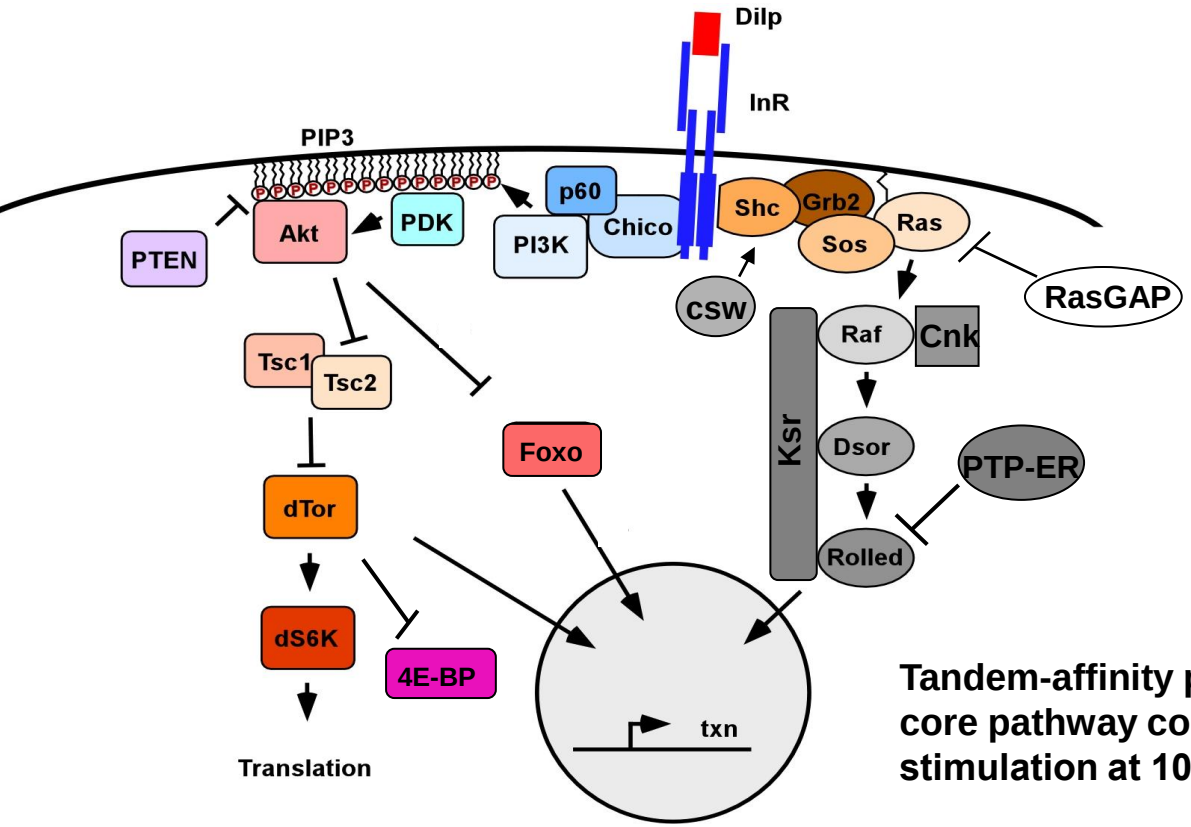


th fosmid
+luc dsRNA



th fosmid
+th dsRNA

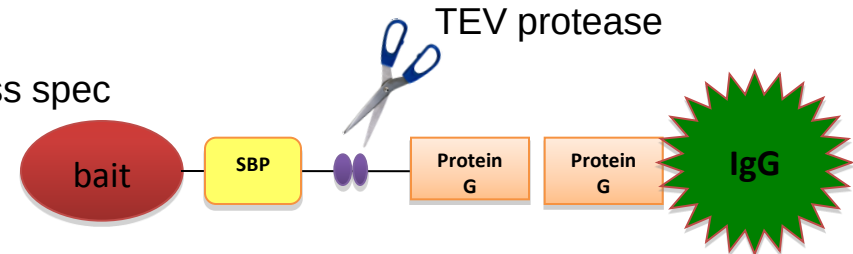
RTK signaling protein interaction network: *Mass Spec*



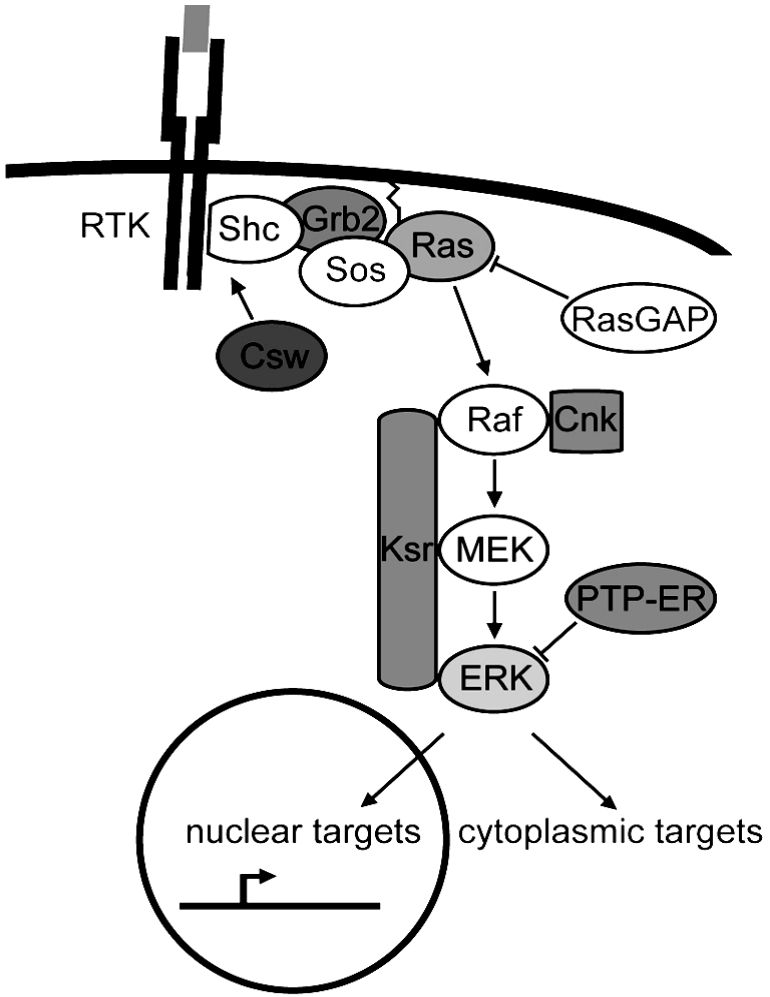
Adam Friedman
Meghana Kulkarni
John Asara (BI)
Pengyu Hong (Brandeis)
Bonnie Berger (MIT)

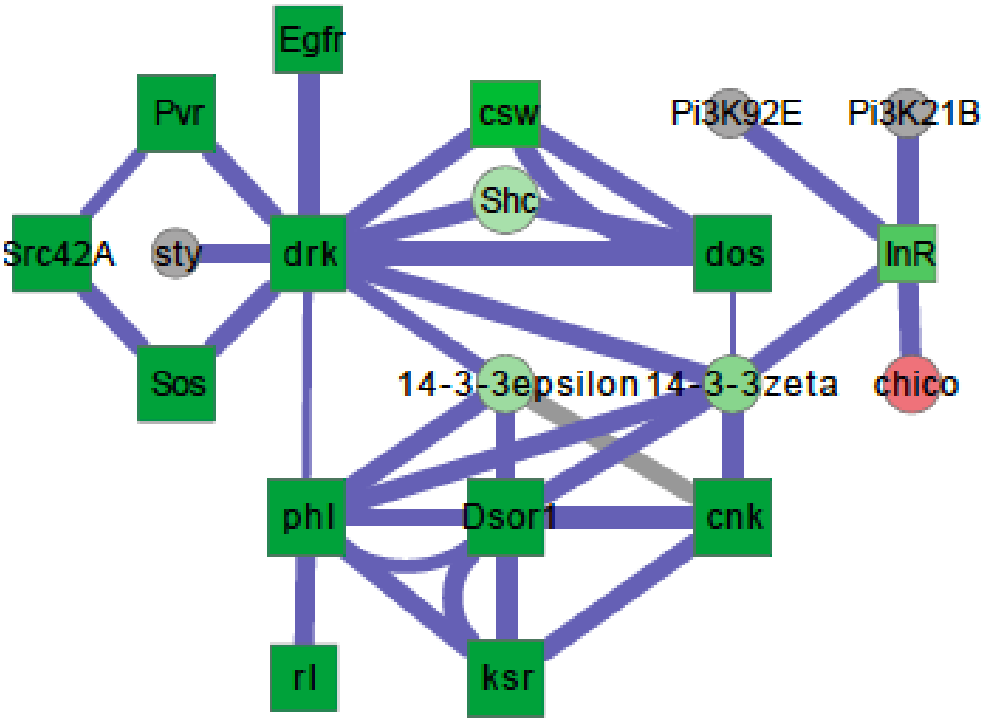
Tandem-affinity purification: LC/MS-MS (TAP/MS) of all 50 core pathway components at Baseline, Insulin, and EGF stimulation at 10' and 30' (400 Mass Spec runs)

. 400ml of culture – 1-2 x 10^9 S2 tissue culture cells
. Protein G and Streptavidin binding protein affinity purifications
. One dimensional LC-MS, reverse phase HPLC before the mass spec
. Significance Analysis of Interactome (SAINT) algorithm to filter background contaminants (Tyers et al. 2010)



Most canonical pathway interactions are identified by Mass Spec

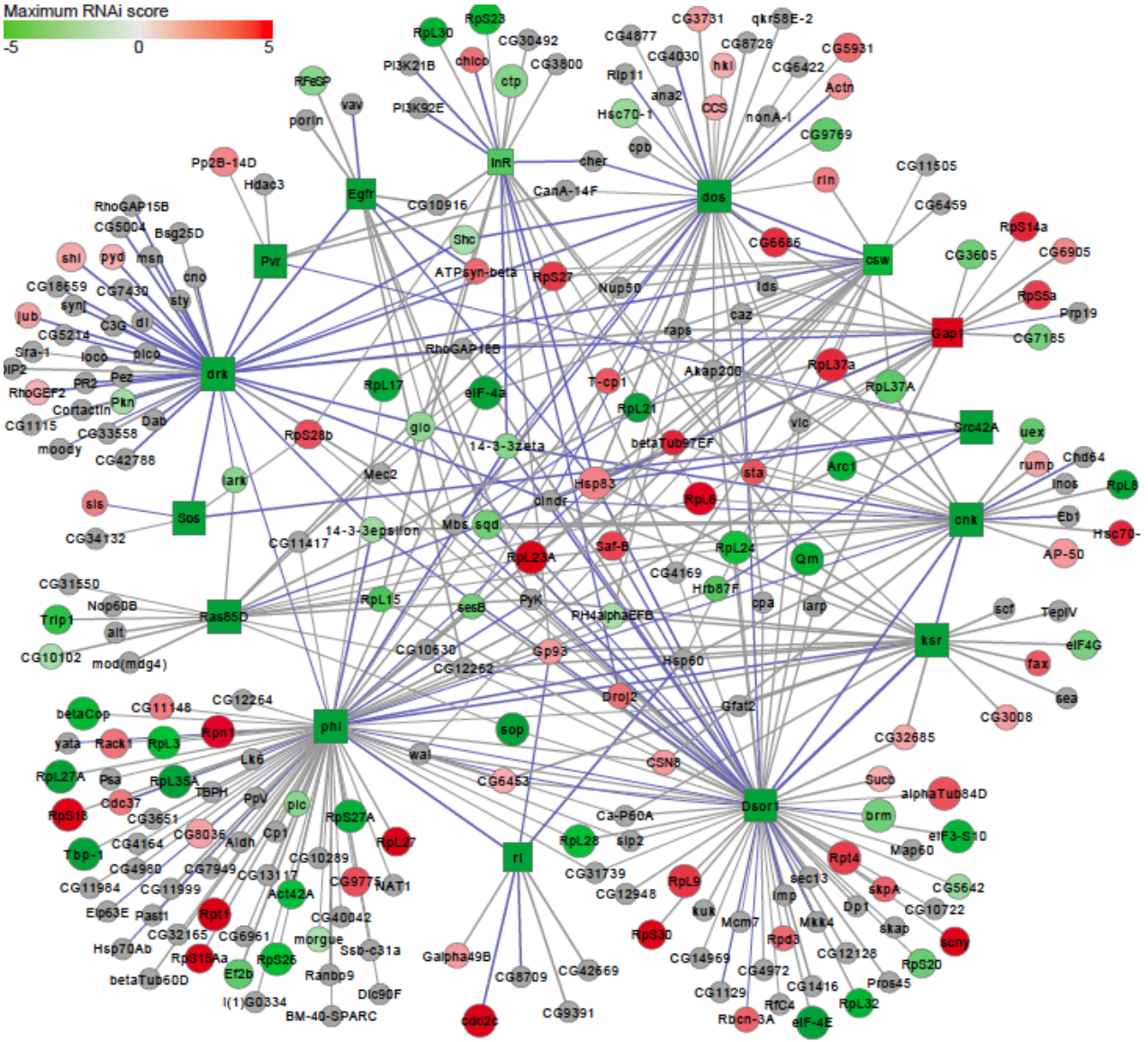




Filtered PPI network of the MAPK Pathway

. 386 interactions among 249 proteins surrounding the canonical components

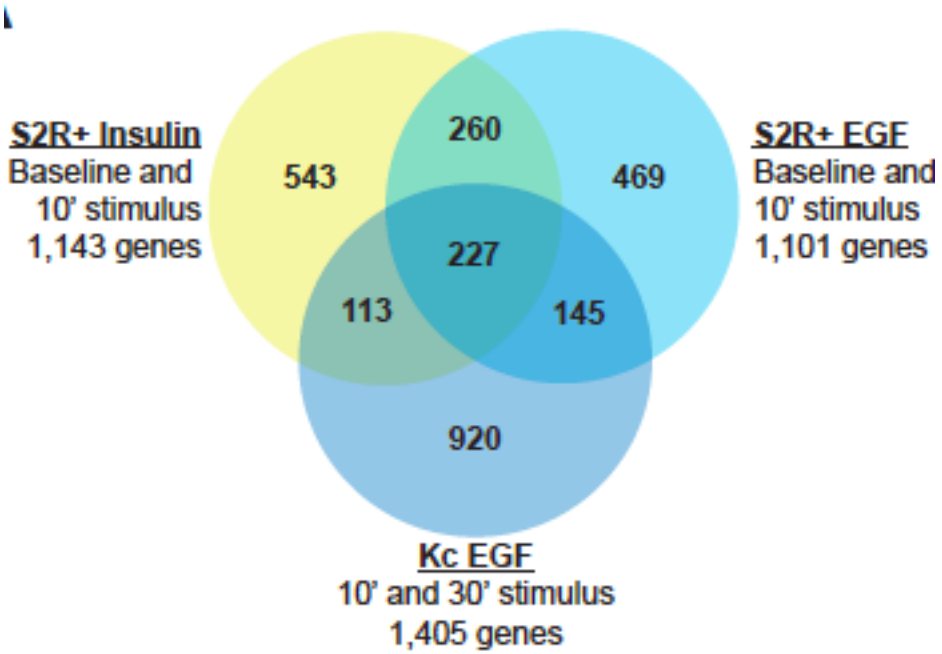
. 119 scored in the RNAi screens (48%)

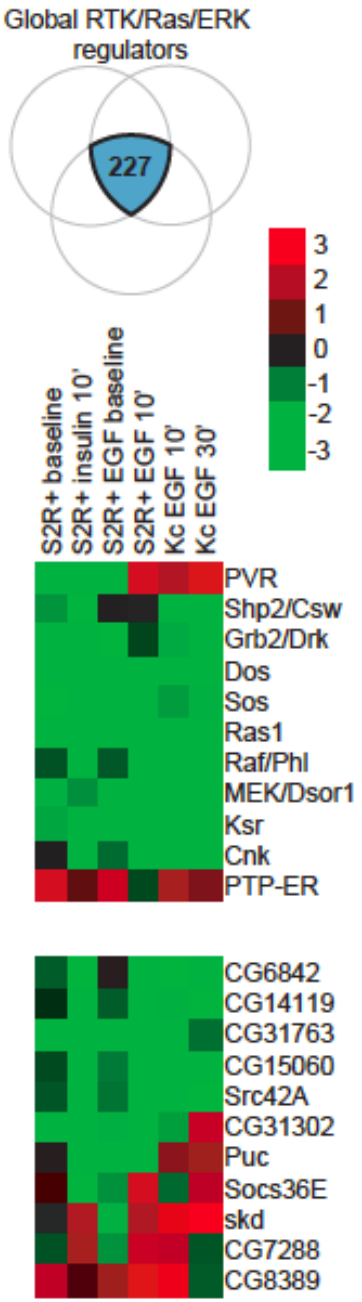


RTK conserved components

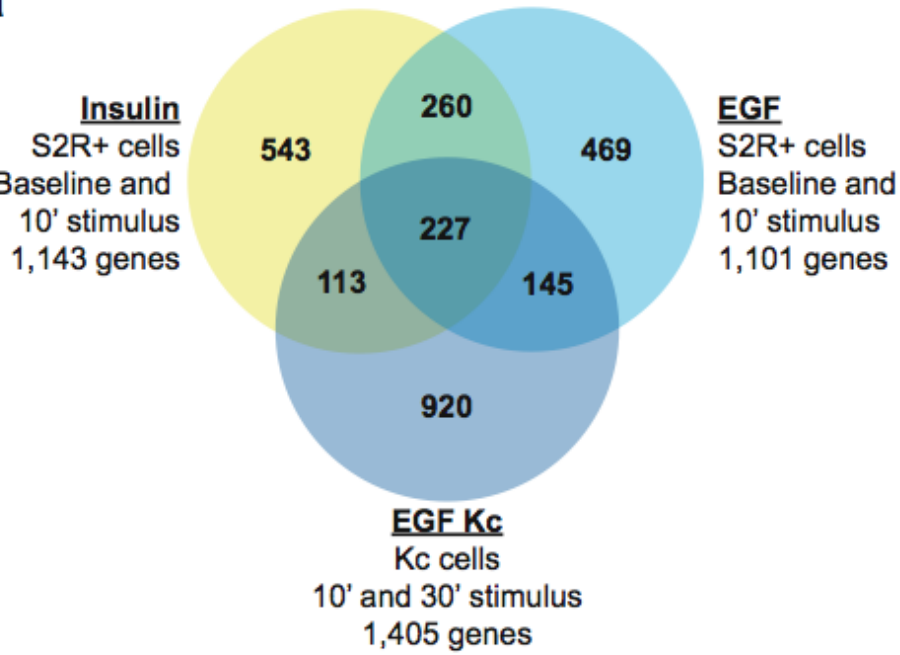
. different cell lines

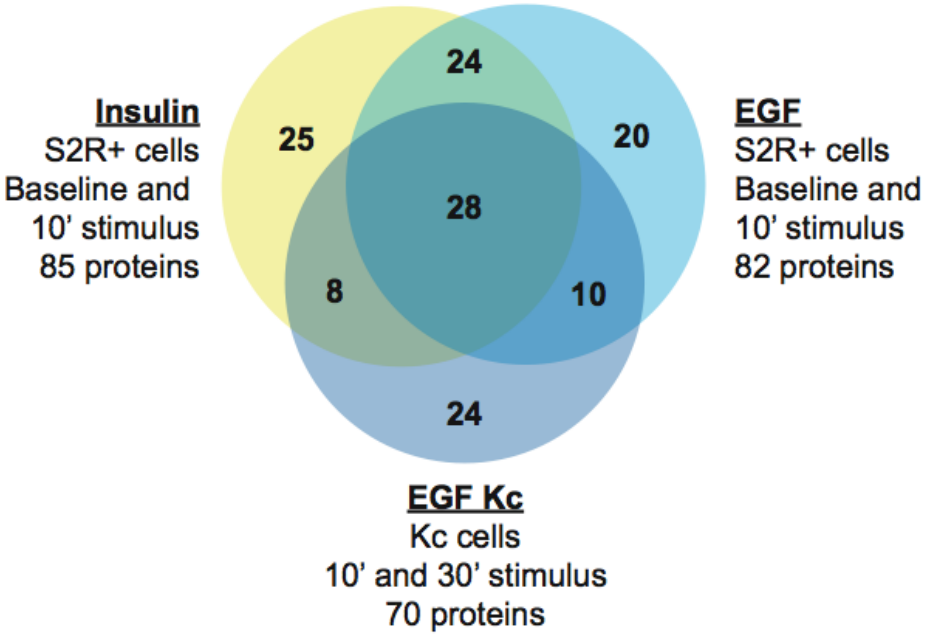
. different RTKs





RNAi + MS

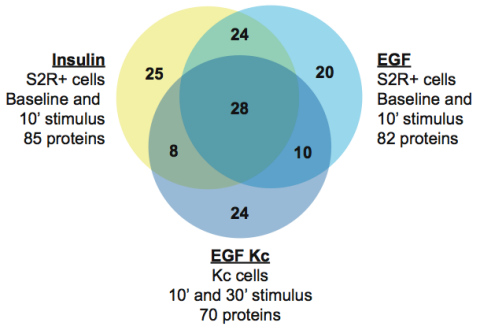




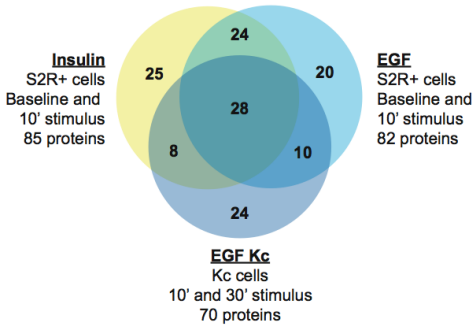
28 in common

149 total

The Canonical pathway is represented within the 28



Remaining 120 represent cell type and RTK specific regulators

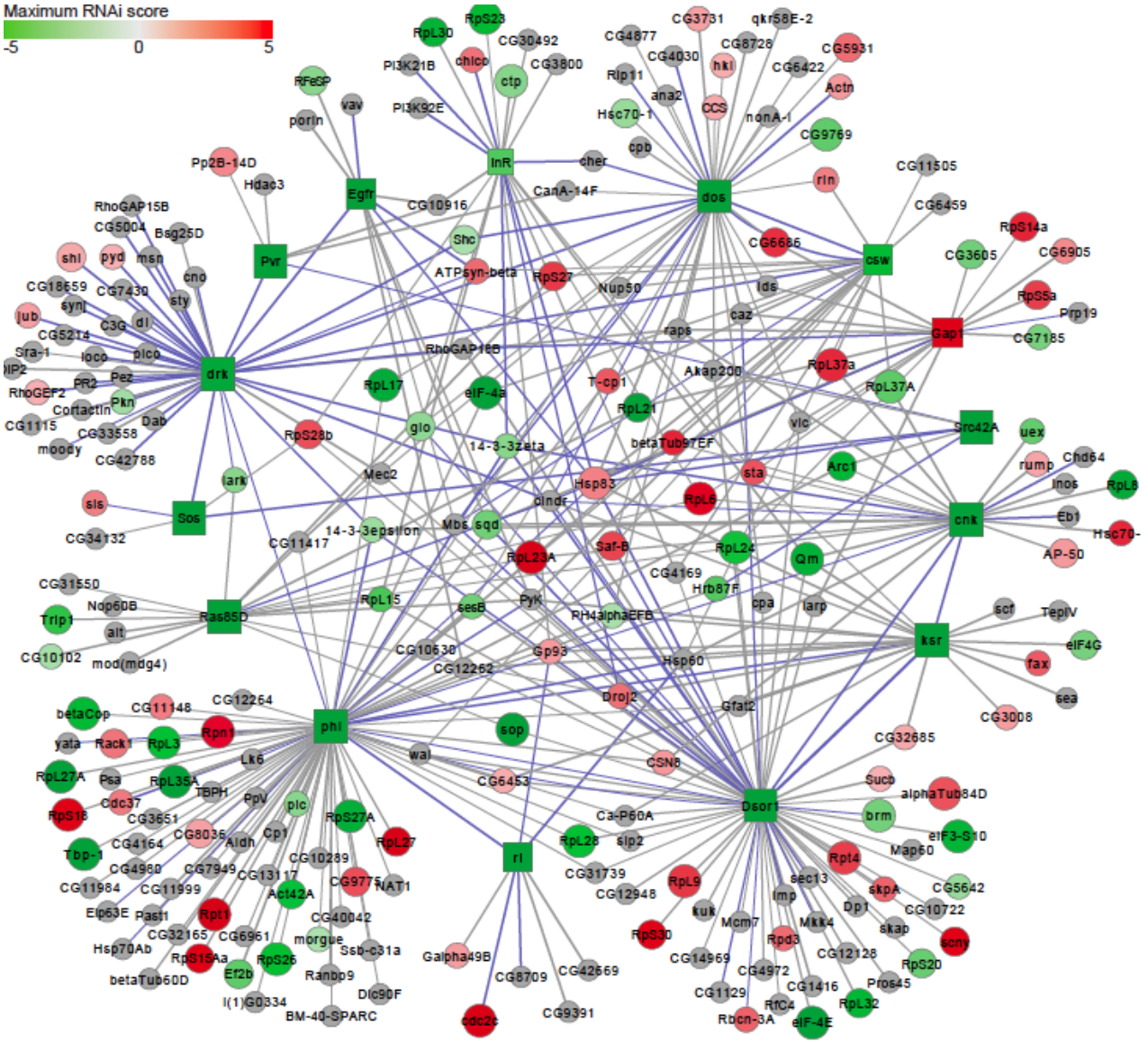


Insulin
S2R+ cells
Baseline and
10' stimulus
85 proteins

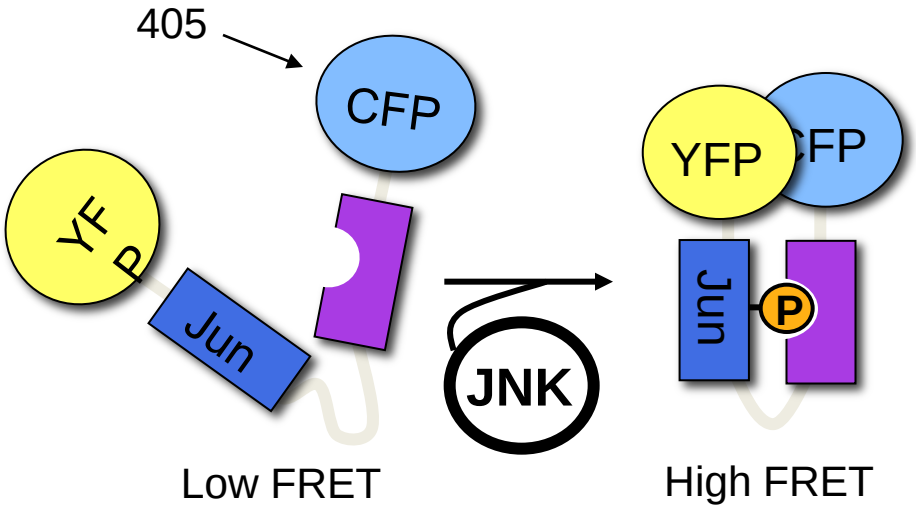
EGF
S2R+ cells
Baseline and
10' stimulus
82 proteins

EGF Kc
Kc cells
10' and 30' stimulus
70 proteins

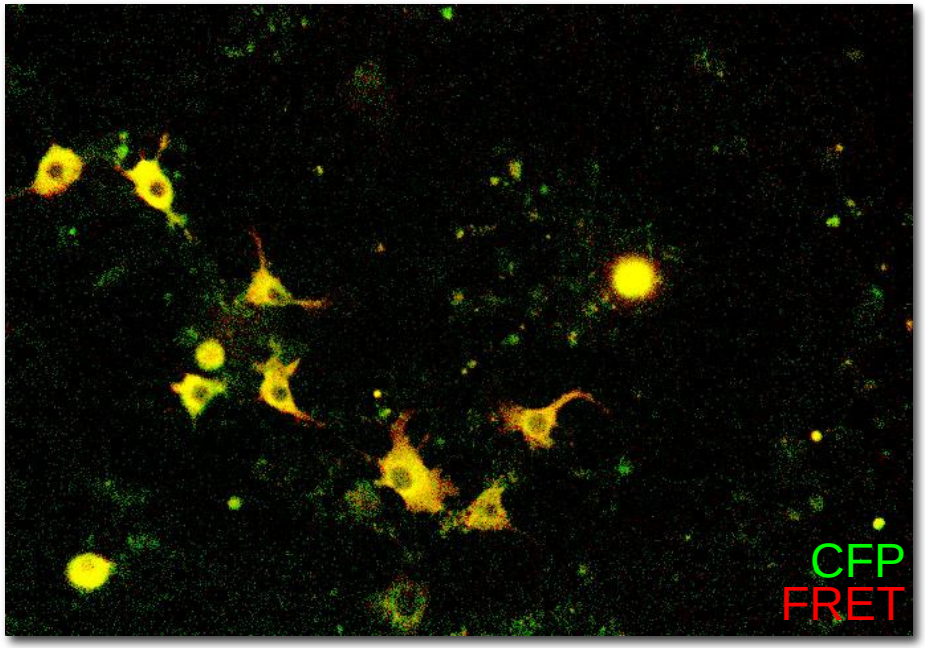
Issue of network redundancy: 50% of the components do not have phenotypes



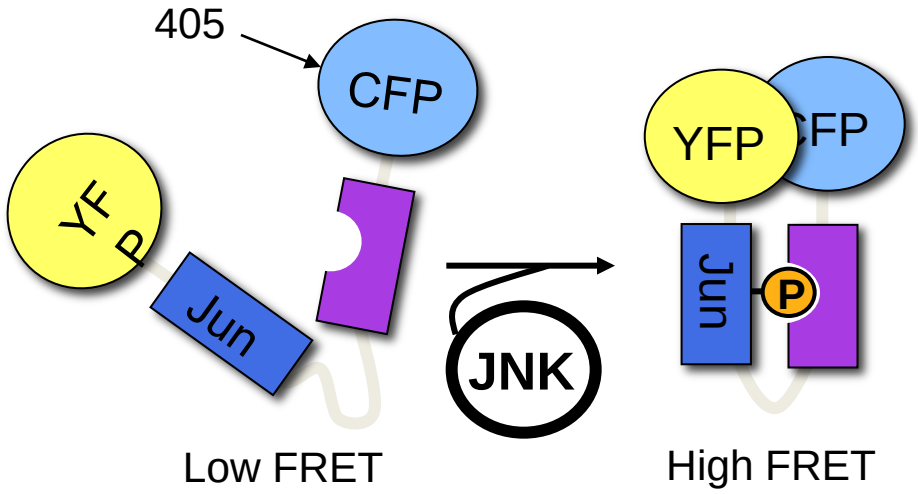
A Single-cell FRET-based JNK Assay

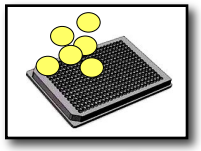


Chris Bakal

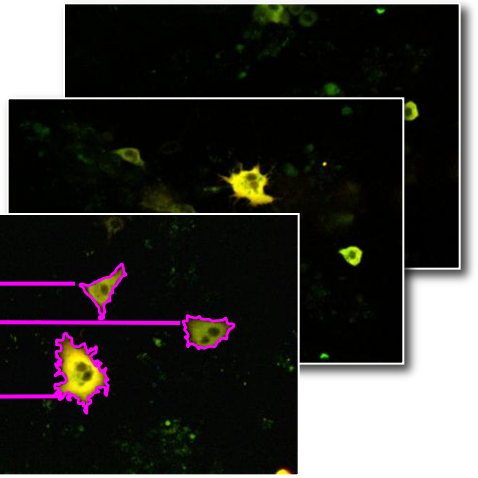


A Single-cell FRET-based JNK Assay





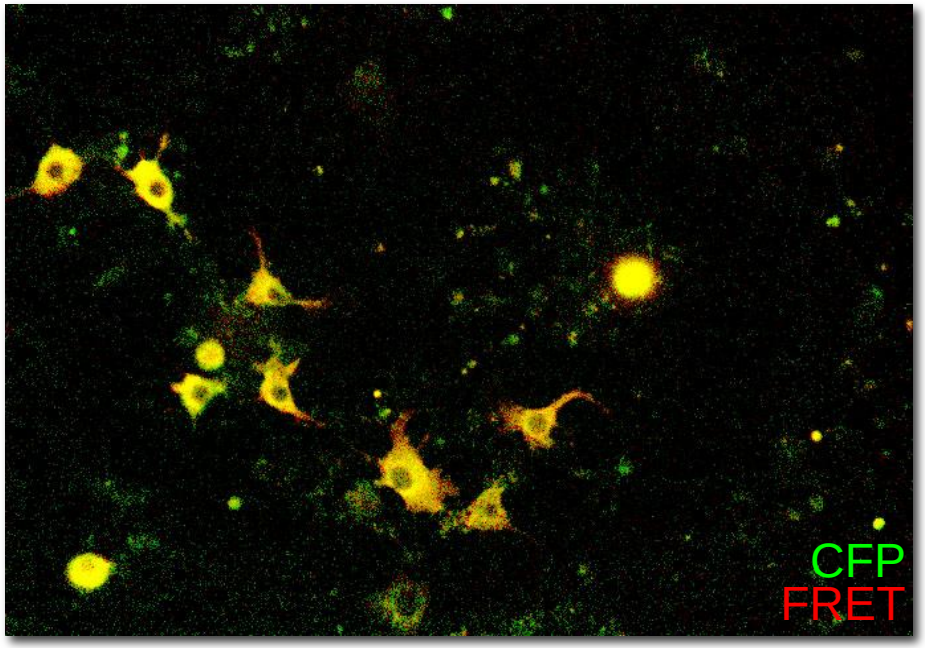
Drosophila
Kinase/Phosphatase
dsRNA library

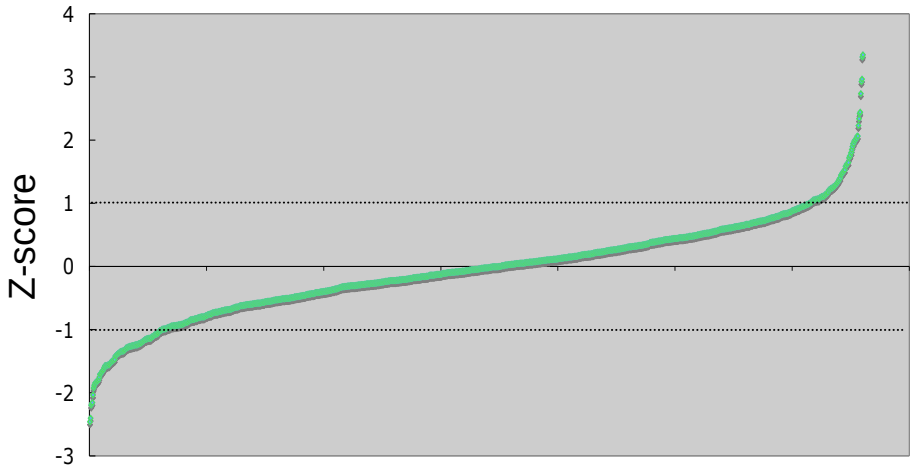


FRET/CFP Ratio

1.34
1.29
1.24
1.73

1.40 = Mean Ratio/Well

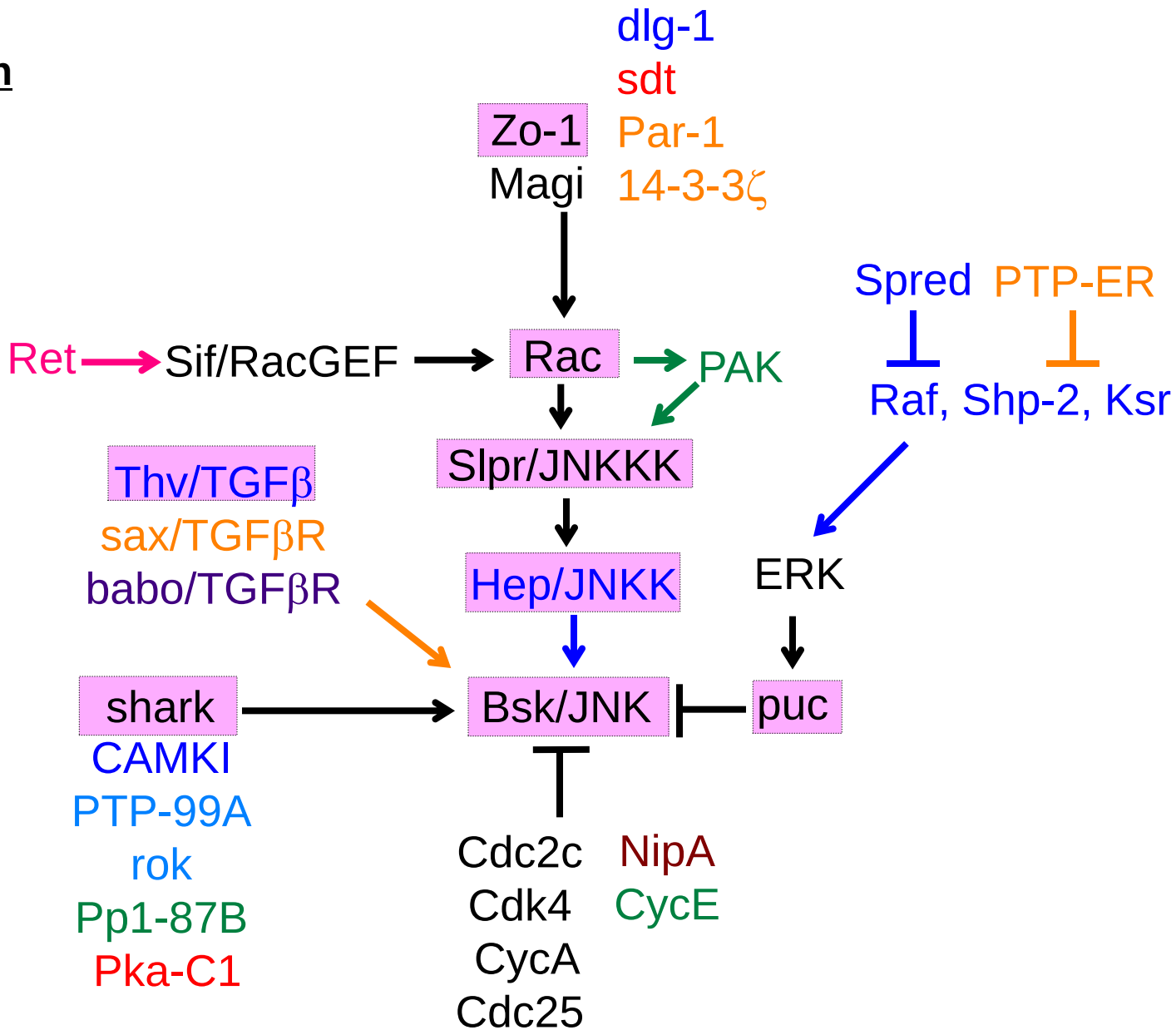




Decreasing False Negative Rates Through Sensitized Screens

Sensitized Screen

Rac1
Cdc42
p190RhoGAP
Akt
PTEN
ERK
Sif/Tiam1
puckered



Canonical JNK

Approaches to describe cellular phenotypes

Ligand

↓

Receptor

↓

Kinase

↓

Transcription factors

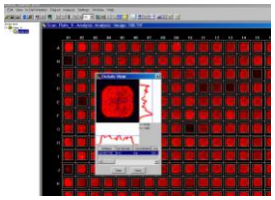
↓

Gene Expression

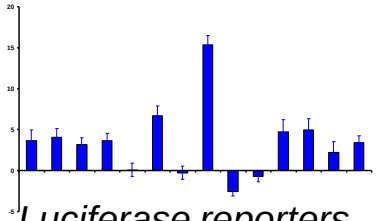
↓

Cell morphology

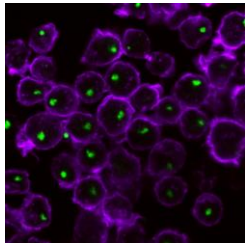
Assays



Phospho-Antibodies

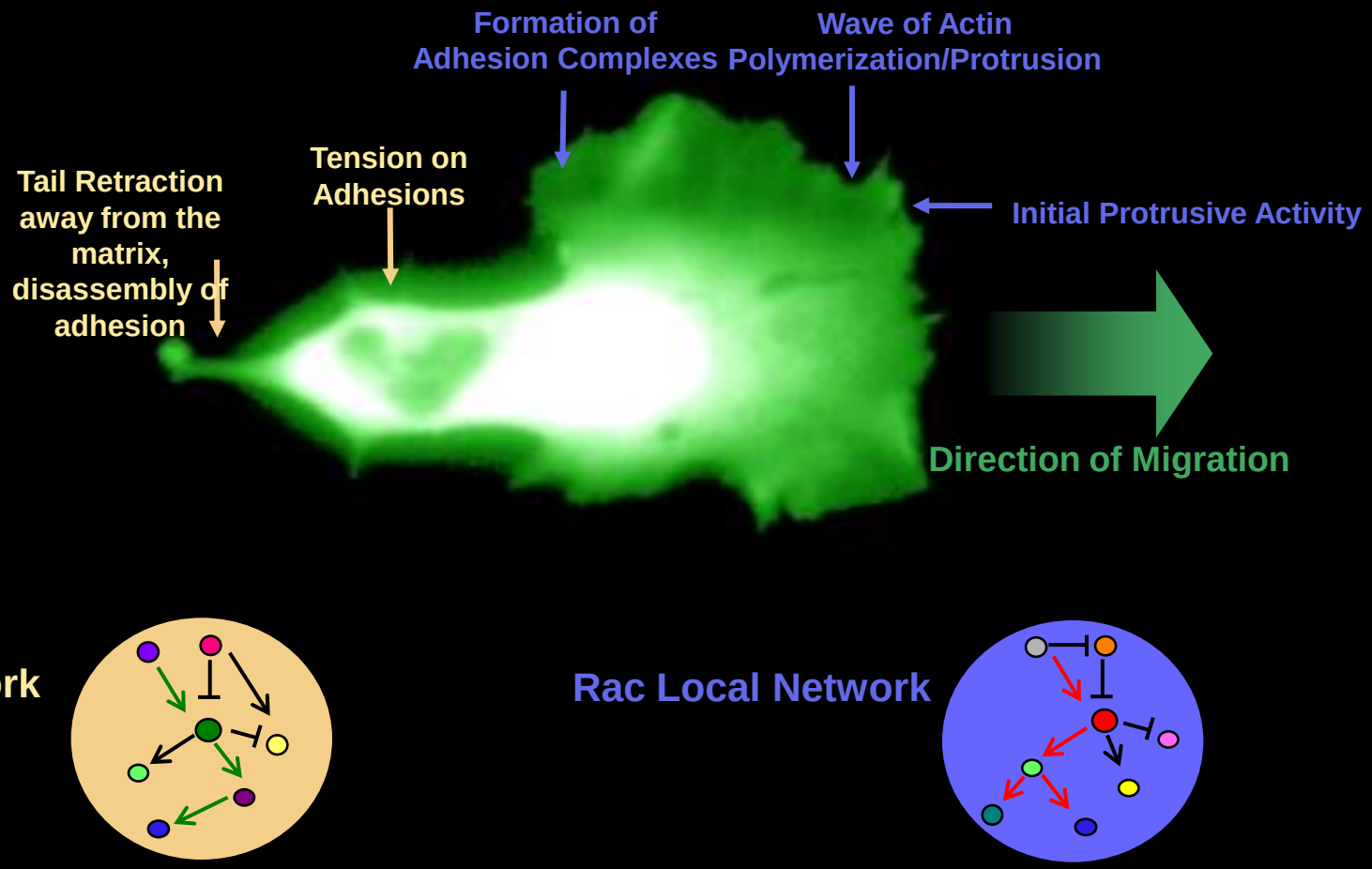


Luciferase reporters



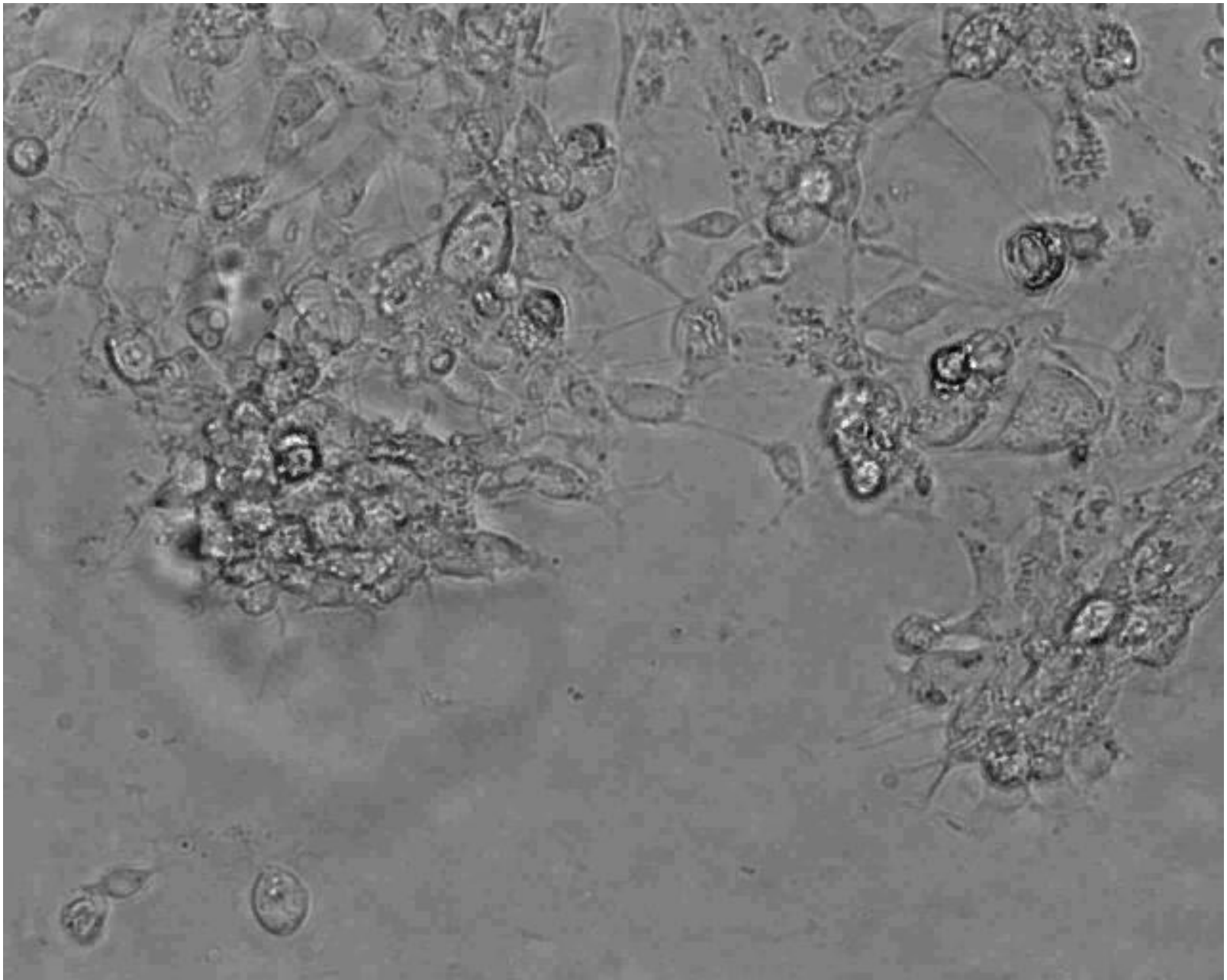
Simple image-based screen

The Activity Of Local Signaling Networks Regulates Dynamic Changes In Cell Morphology

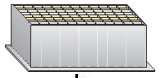


Chris Bakal

Scratch Assay in mBG2 Drosophila Cell Line

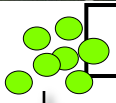


Experimental Design For A System-level Phenotypic Profiling

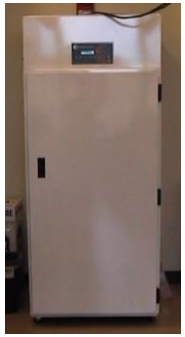


dsRNA/compound plating (from library of 15,000+ dsRNAs)

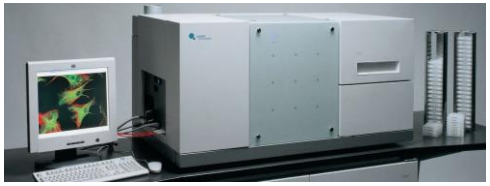
Add cells

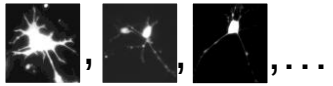


Cell Plating (384 well plates)

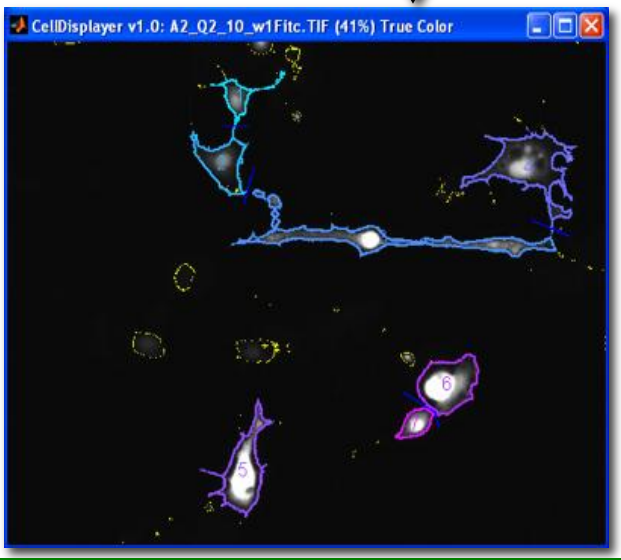


3-5 day incubation

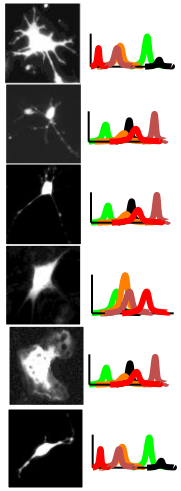




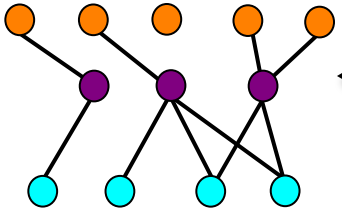
Automated Image Acquisition



Semi-Automated Image Analysis



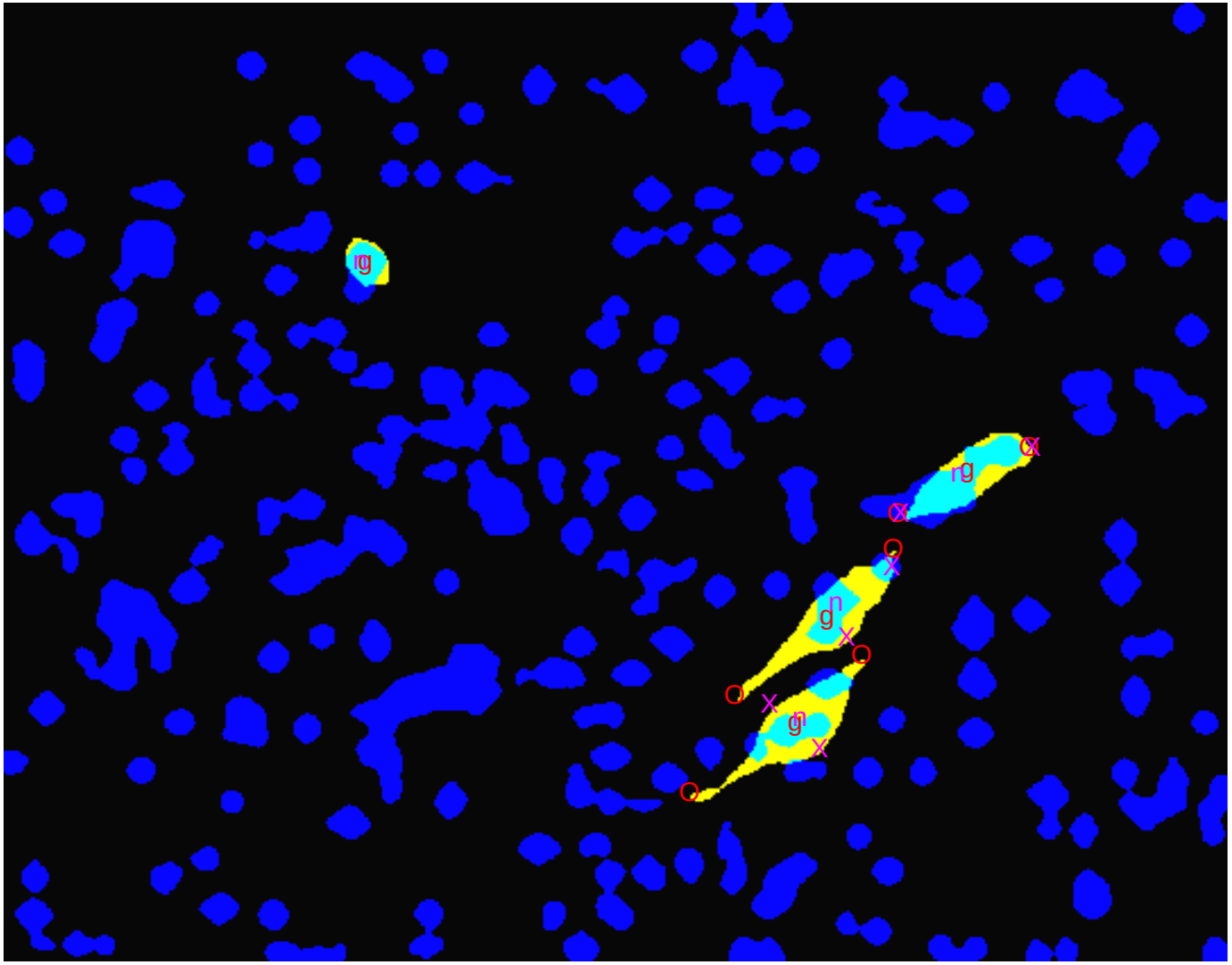
Data Management



Network Modeling

Bakal, Aach, Church and Perrimon
Science 2007

A small population of GFP-expressing cells are mixed to non-labeled cells to facilitate the segmentation analysis



Feature Extractions From Images

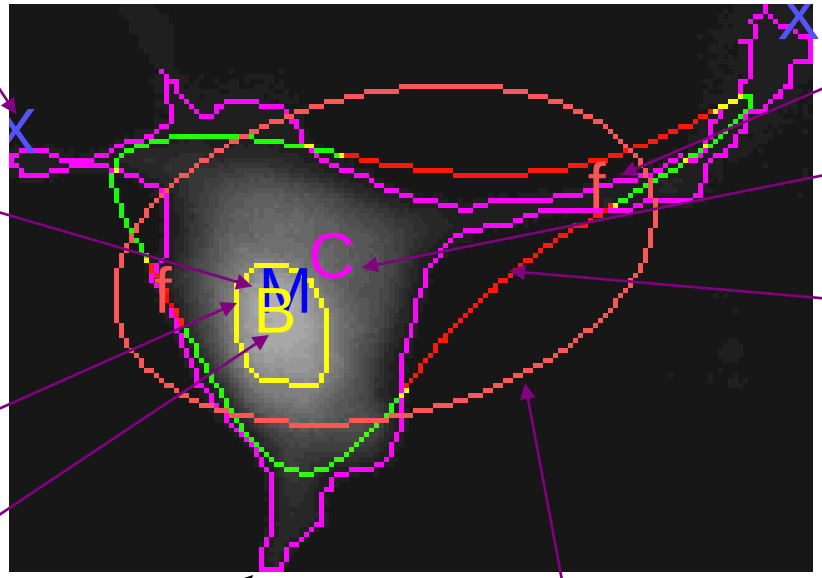


Image and cell number

Degree of smoothing

Sharpest process (X) (1 of 2) as determined at current smoothing level

Cell center of mass (M)

"Bright spot" boundary

"Bright spot" centroid (B)

Original CellSegmenter-determined boundary

Best-fit ellipse focus (f)

Cell geometric centroid (C)

Smoothed boundary: Green = positive curvature, red = negative curvature region

Best ellipse fit to smoothed boundary

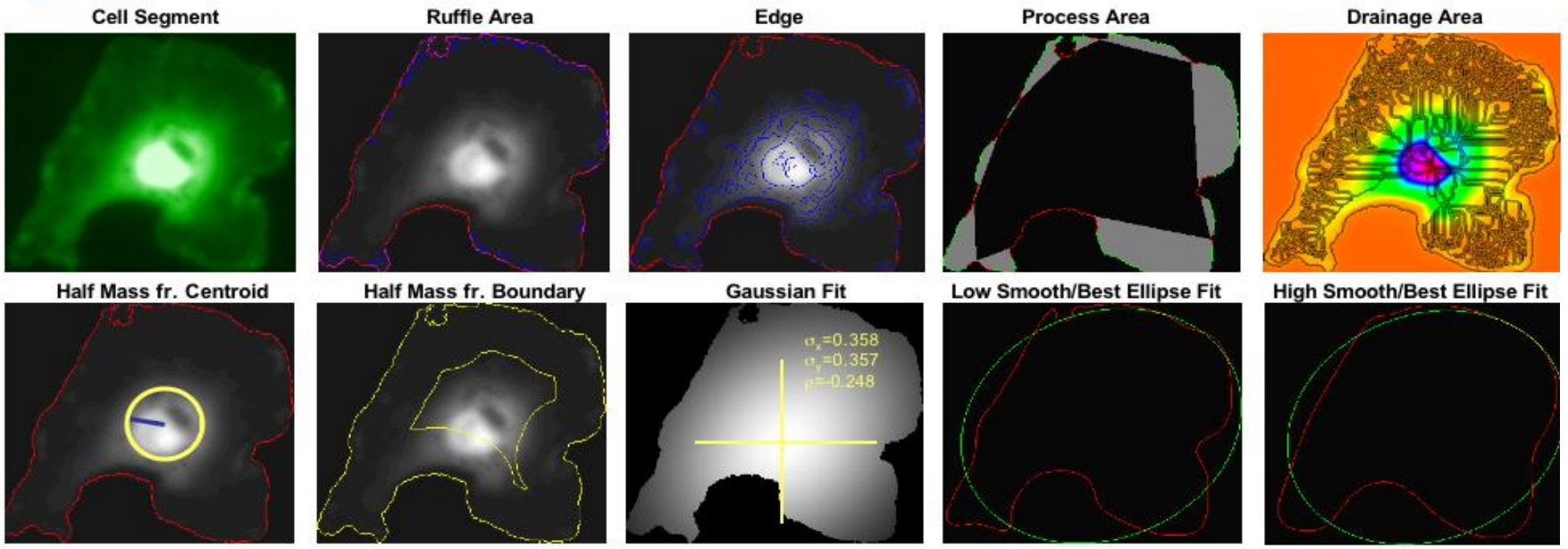
<u>150 distinct pieces of information</u>

- Basic geometric features (e.g., area, solidity, major axis length)

- Intensity-related features

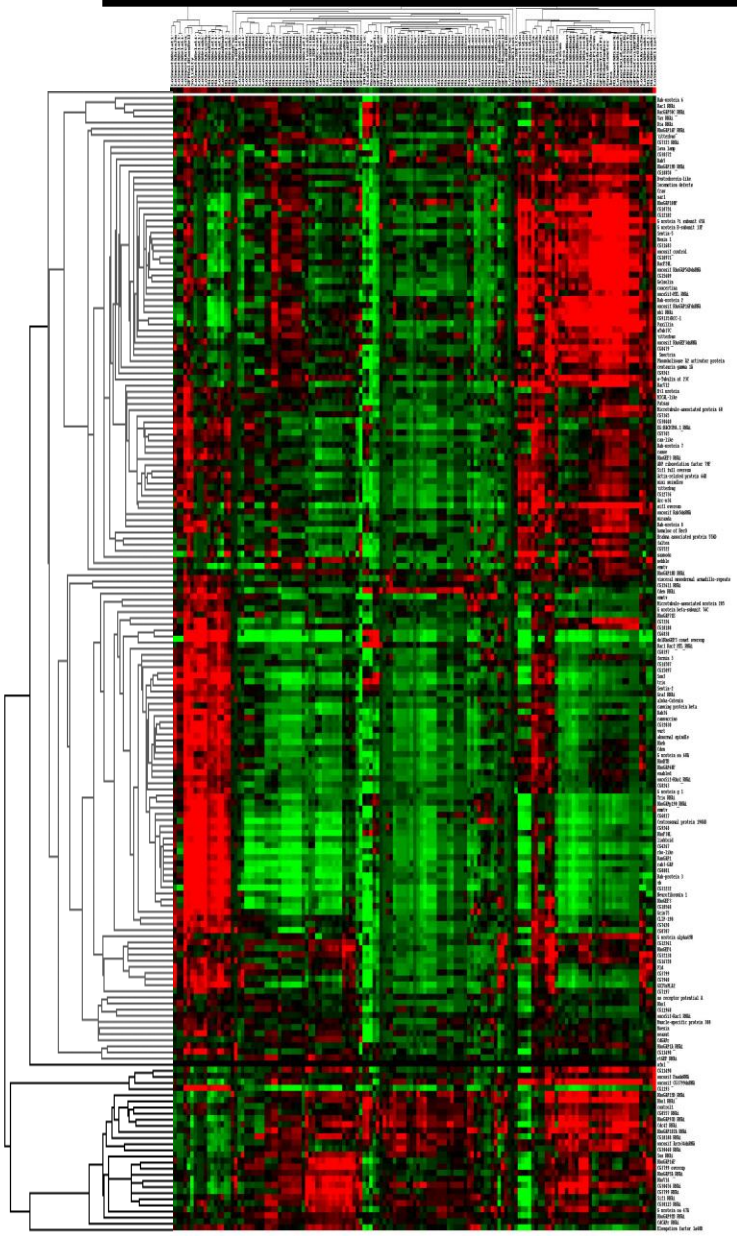
- Boundary-shape analysis

Feature Examples



Raw Morphological Data Are Not Interpretable

145 phenotypic features



187 conditions

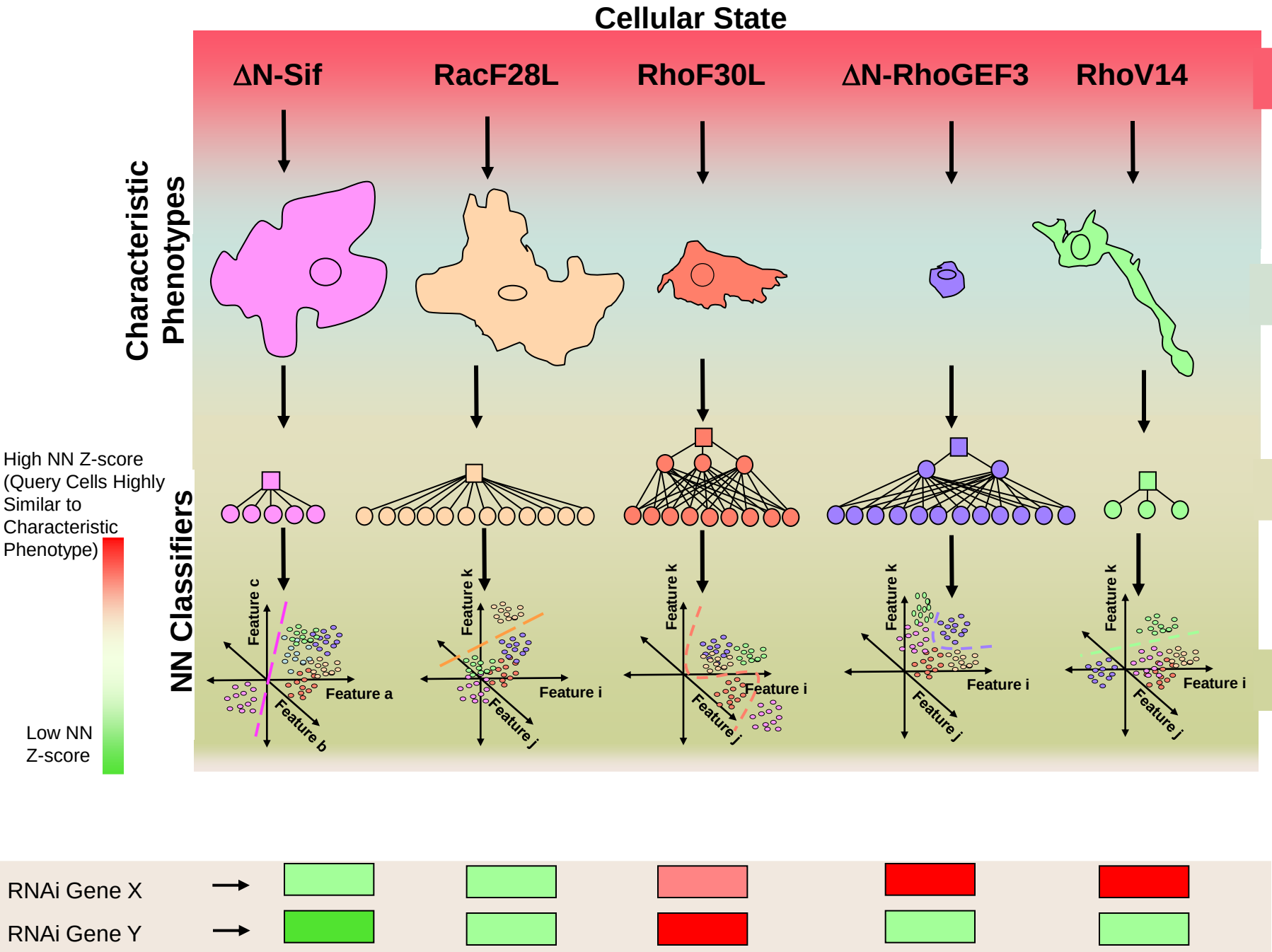
166 dsRNA

20 overexpression

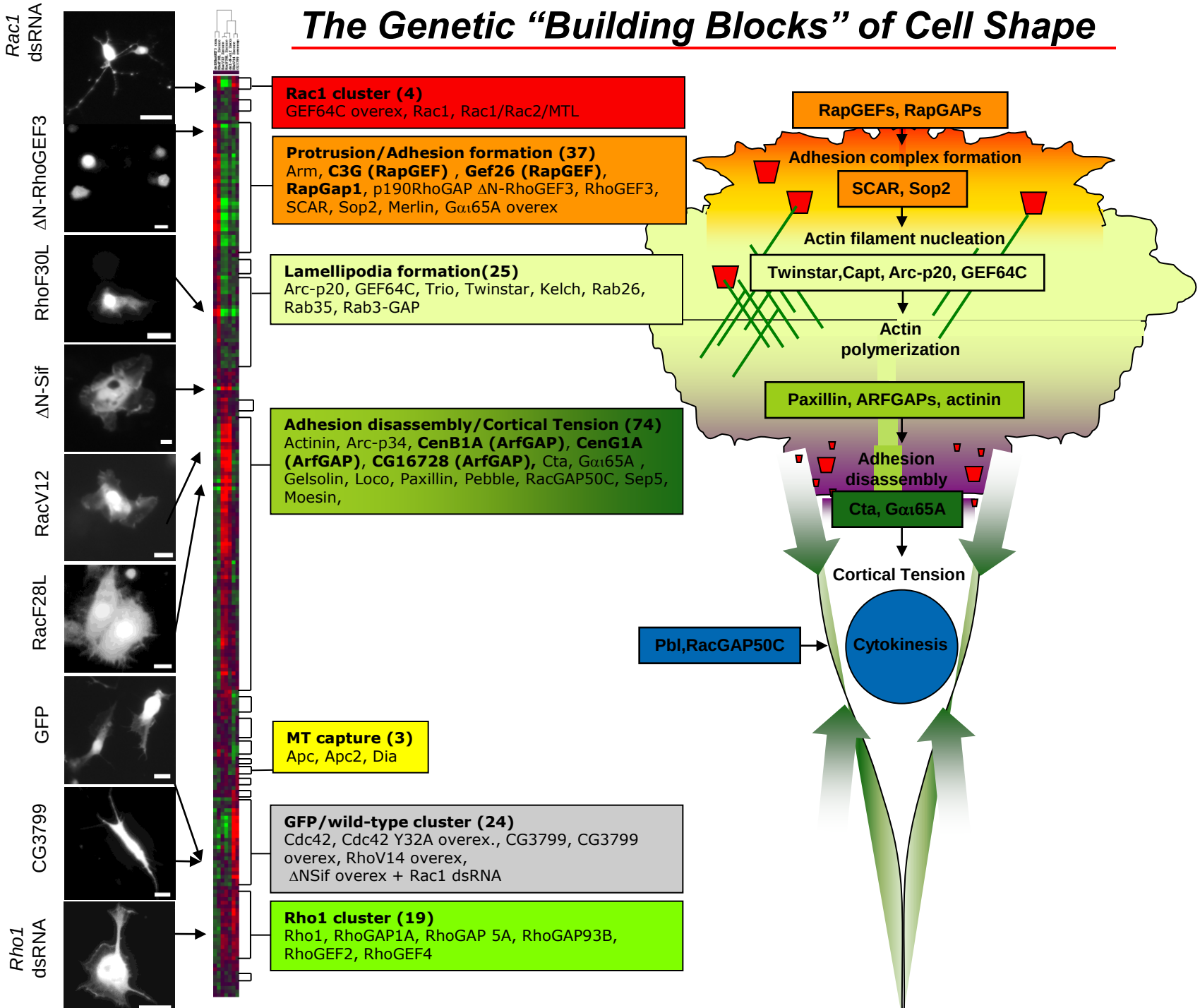
95% of Rho signaling pathway

10 synthetic (overexpression + dsRNA)

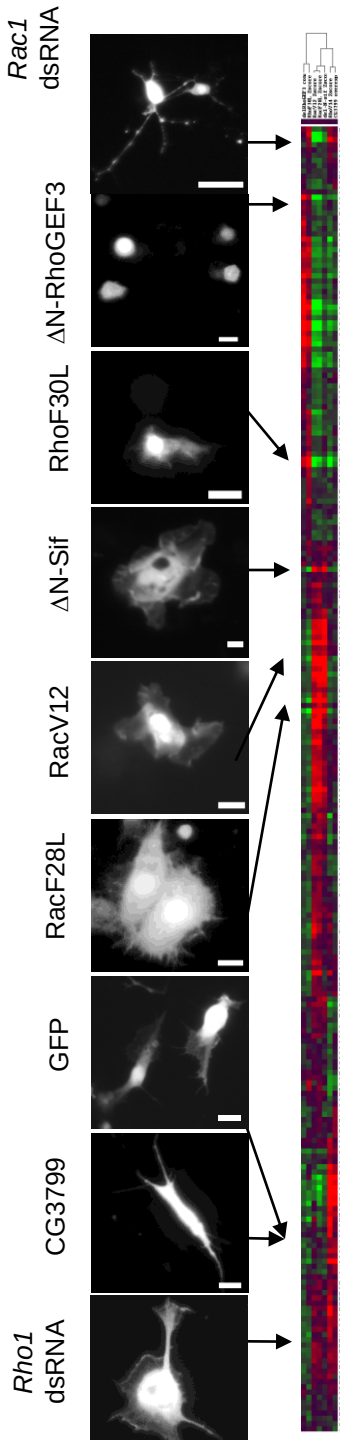
Using Neural Networks to Derive Morphological Signatures



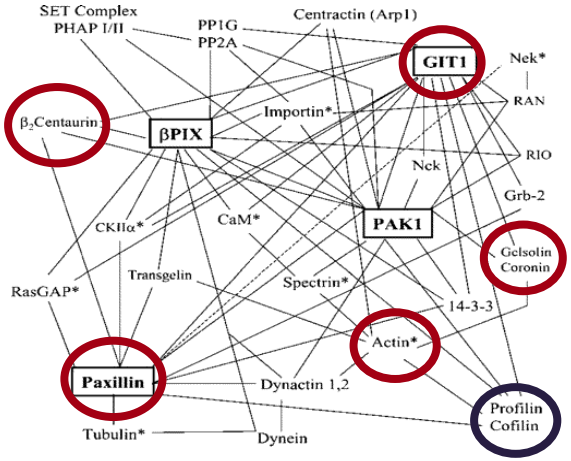
The Genetic "Building Blocks" of Cell Shape



Clustering Predicts Physical Complexes

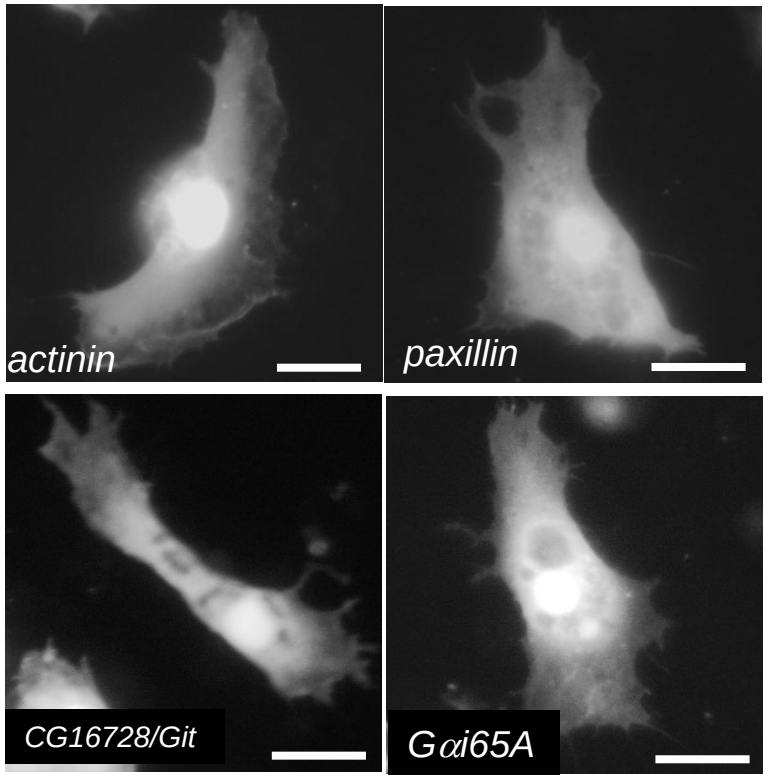


Adhesion disassembly/Cortical Tension (74)
Actinin, Arc-p34, **CenB1A (ArfGAP)**, **CenG1A (ArfGAP)**, **CG16728 (ArfGAP),** Cta, Gαi65A , Gelsolin, Loco, Paxillin, Pebble, RacGAP50C, Sep5, Moesin,



Mayhew et al., *J. Proteome Res. 2006*

Cluster 18



Inferring Genetic Architecture from "Systems Genetics" Studies

1.Building high confidence Networks

. RNAi
. Mass Spec
. Transcriptomics

2. Network analyses

. Protein complexes
. Genetic interaction signatures "Epistasis-map"
. Flow
- transcriptome signatures
- phosphorylation signatures