

Application of Genomic Technology to Toxicology

Identifying Predictive Biomarkers and Assessing the
Impact of Chemicals on Cell Signaling Networks

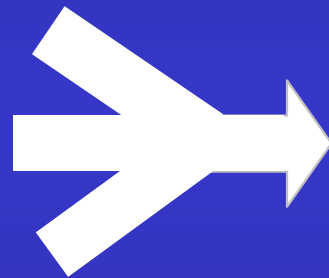
Rusty Thomas
CIIT Centers for Health Research
SOT Colgate Palmolive Lecture
San Diego, 2006

Is There a Role for Genomics in the 3Rs...

Reduce

Refine

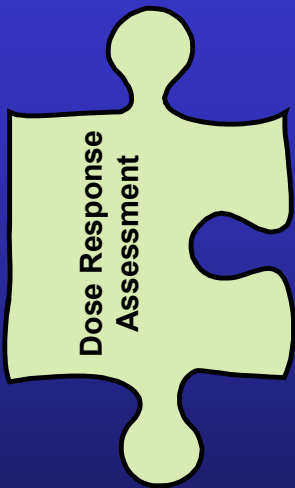
Replace



Is There a Role for Genomics in the 3Rs...



Application of Genomic and Metabonomic Technology to Identify Biomarkers Predictive of Rodent Cancer Bioassay



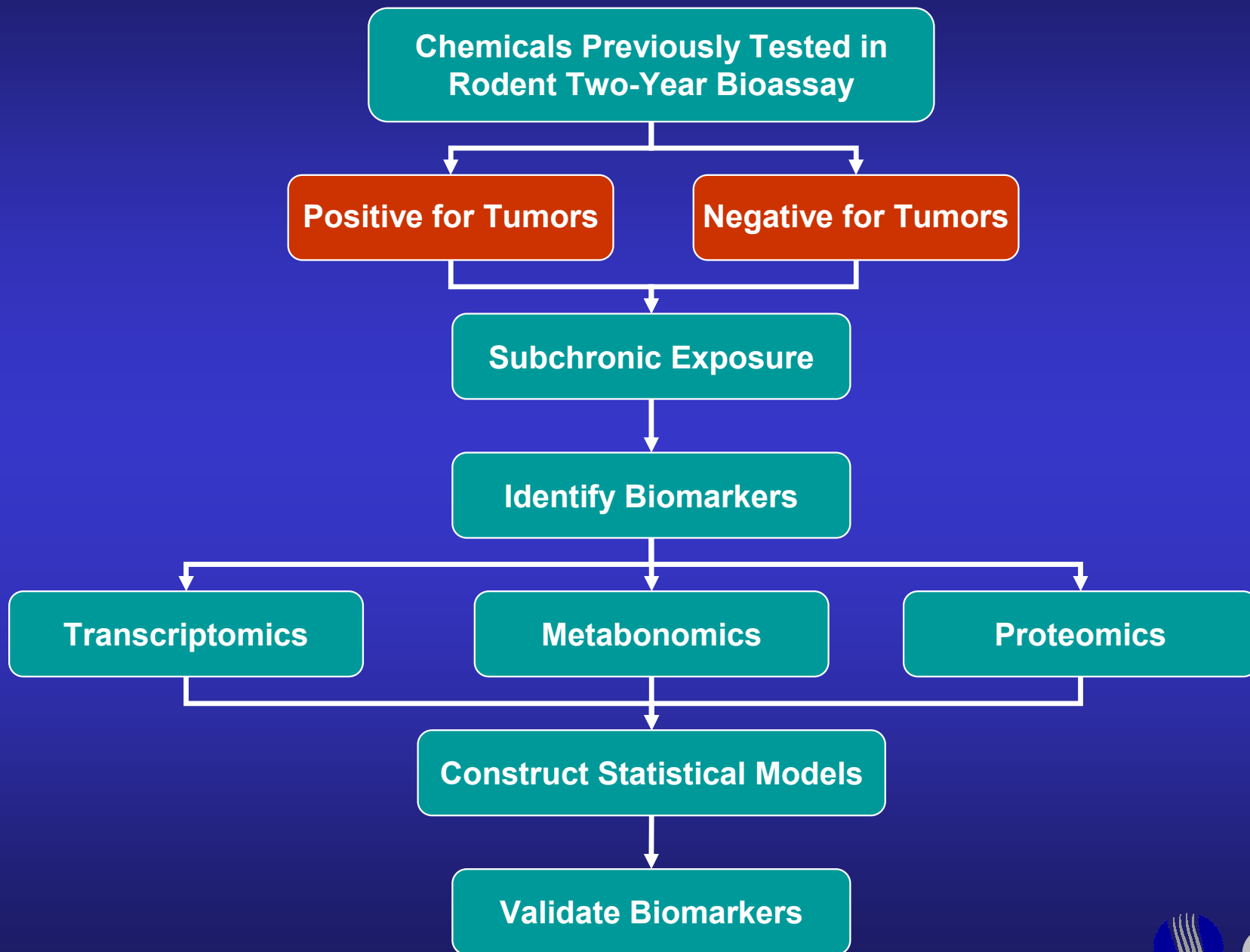
Application of High-Coverage Functional Genomic Screens to Dissect Cell Signaling Networks

Application of Genomics to Hazard ID

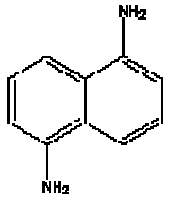
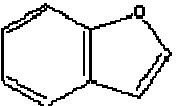
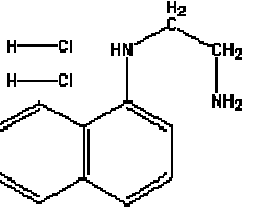
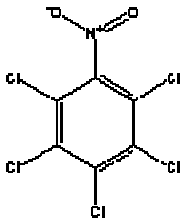
Background and Significance

- Two-year rodent bioassays play a central role in evaluating both the carcinogenic potential of a chemical and generating quantitative information on the dose-response behavior for chemical risk assessments.
- Due to the resource-intensive nature of these studies, each bioassay costs \$2 to \$4 million and takes over three years to complete.
- Since its inception in 1978, the number of chemicals currently tested by the National Toxicology Program (NTP) stands at 530 in long-term studies and 70 in short-term tests.
- Currently, there are approximately 80,000 chemicals registered for commercial use in the United States with 2,000 more added each year.

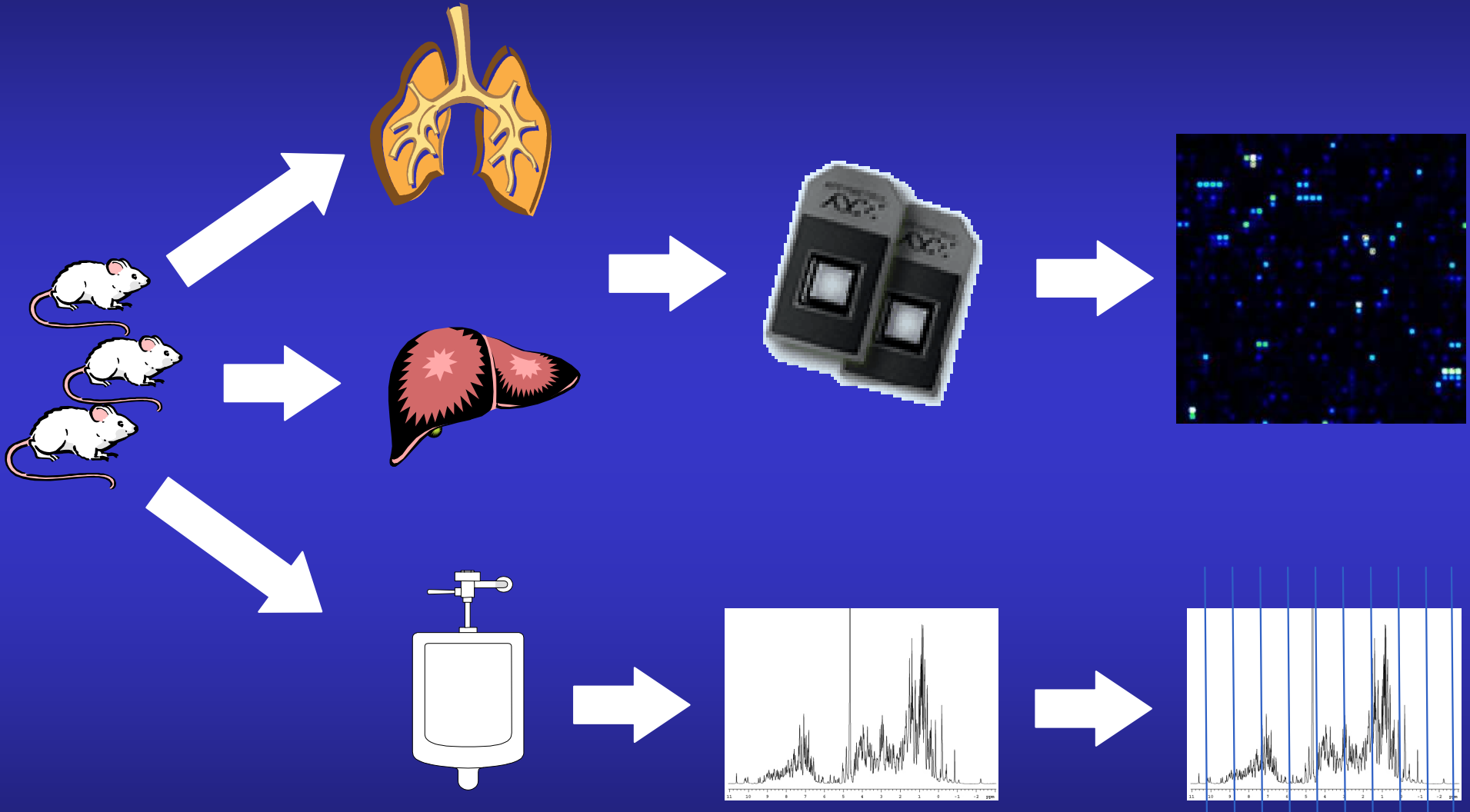
Experimental Flow Chart



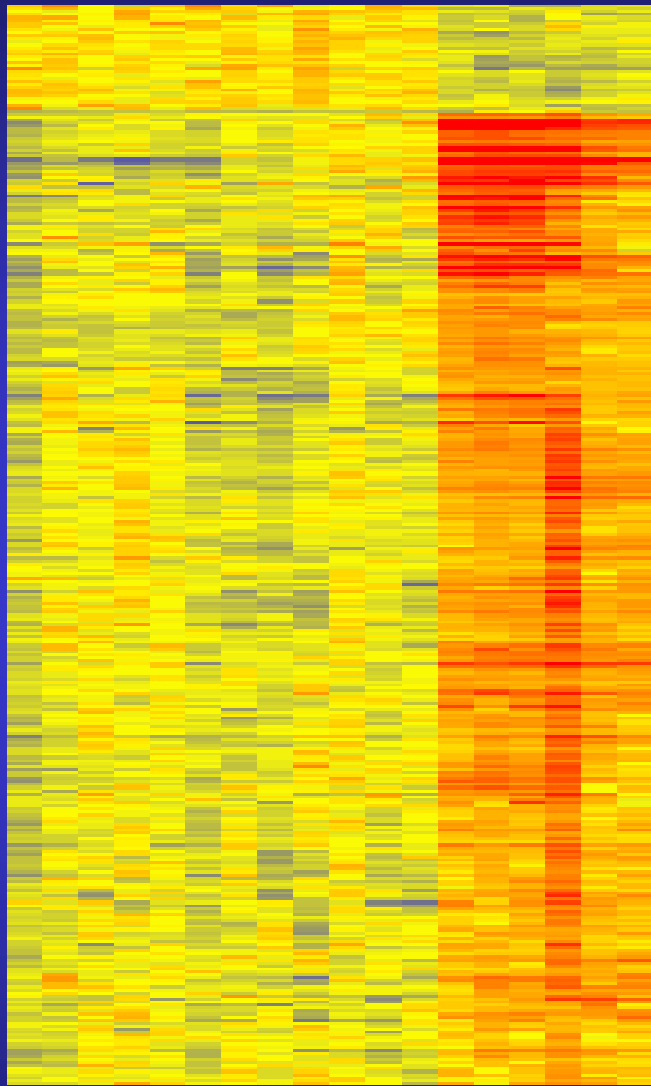
Experimental Design

Chemical	Genetic Tox	Exposure Route	MTD Dose	NTP Designation
 1,5-Naphthalenediamine (NAPD)	+	Feed	2,000 ppm	Positive
 Benzofuran (BFUR)	-	Gavage	240 mg/kg	Positive
 N-(1-naphthyl)ethylenediamine dihydrochloride (NEDD)	+	Feed	2,000 ppm	Negative
 Pentachloronitrobenzene (PCNB)	-	Feed	8,187 ppm	Negative
Corn oil (CCON)		Gavage		
Feed (FCON)				

Experimental Design

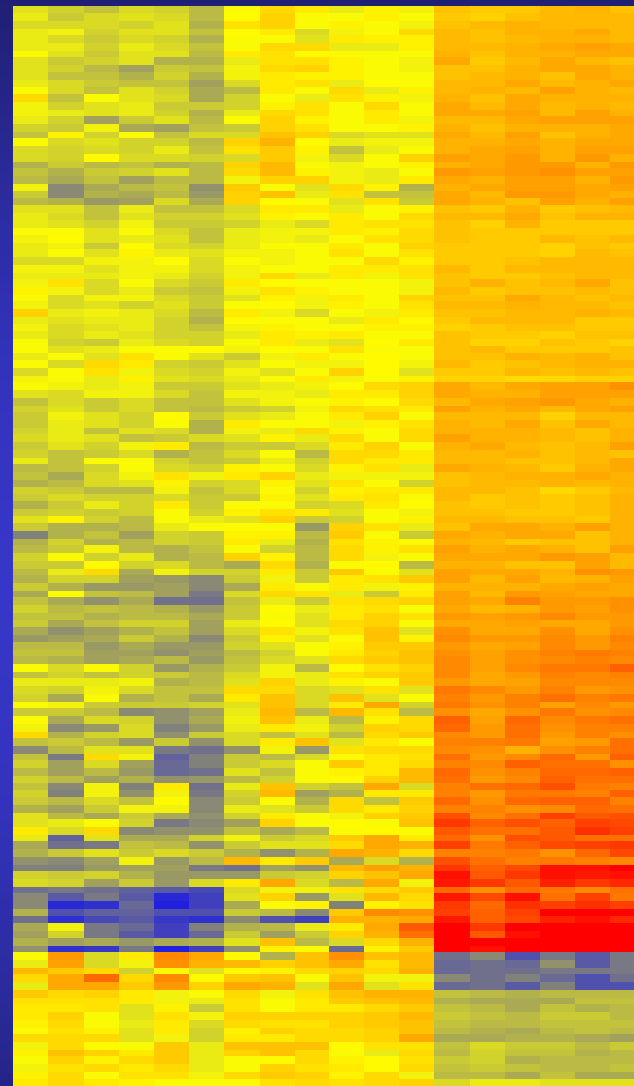


Experimental Results



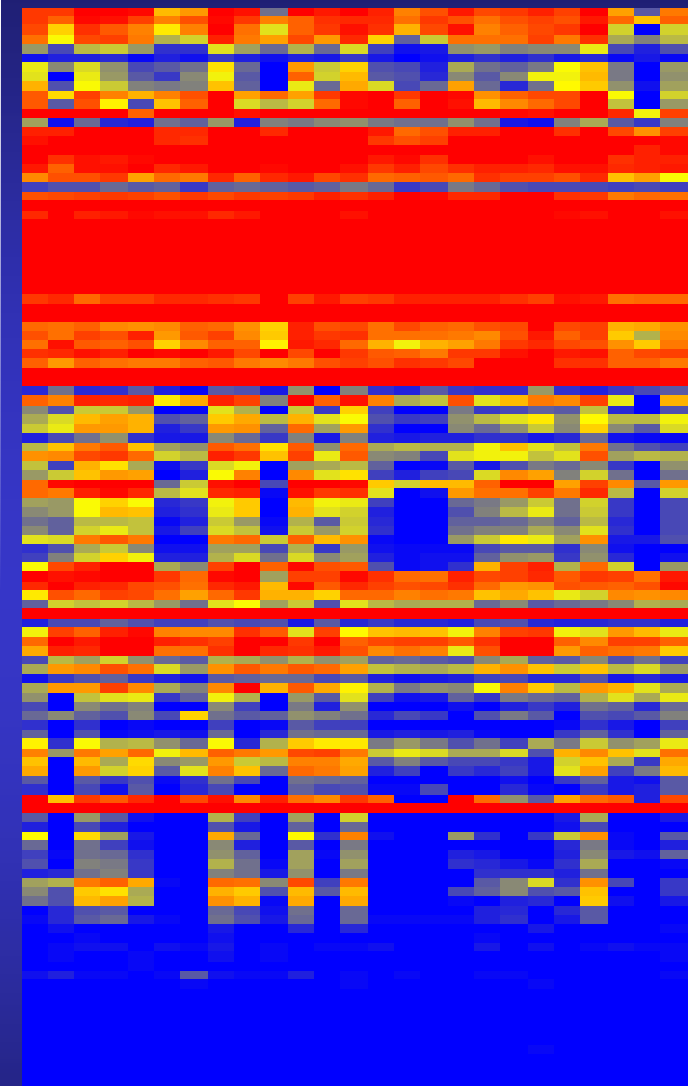
CCON FCON NEDD PCNB NAPD BFUR
← NONCARC → ← CARC →

Lung Gene Expression



CCON FCON NEDD PCNB NAPD BFUR
← NONCARC → ← CARC →

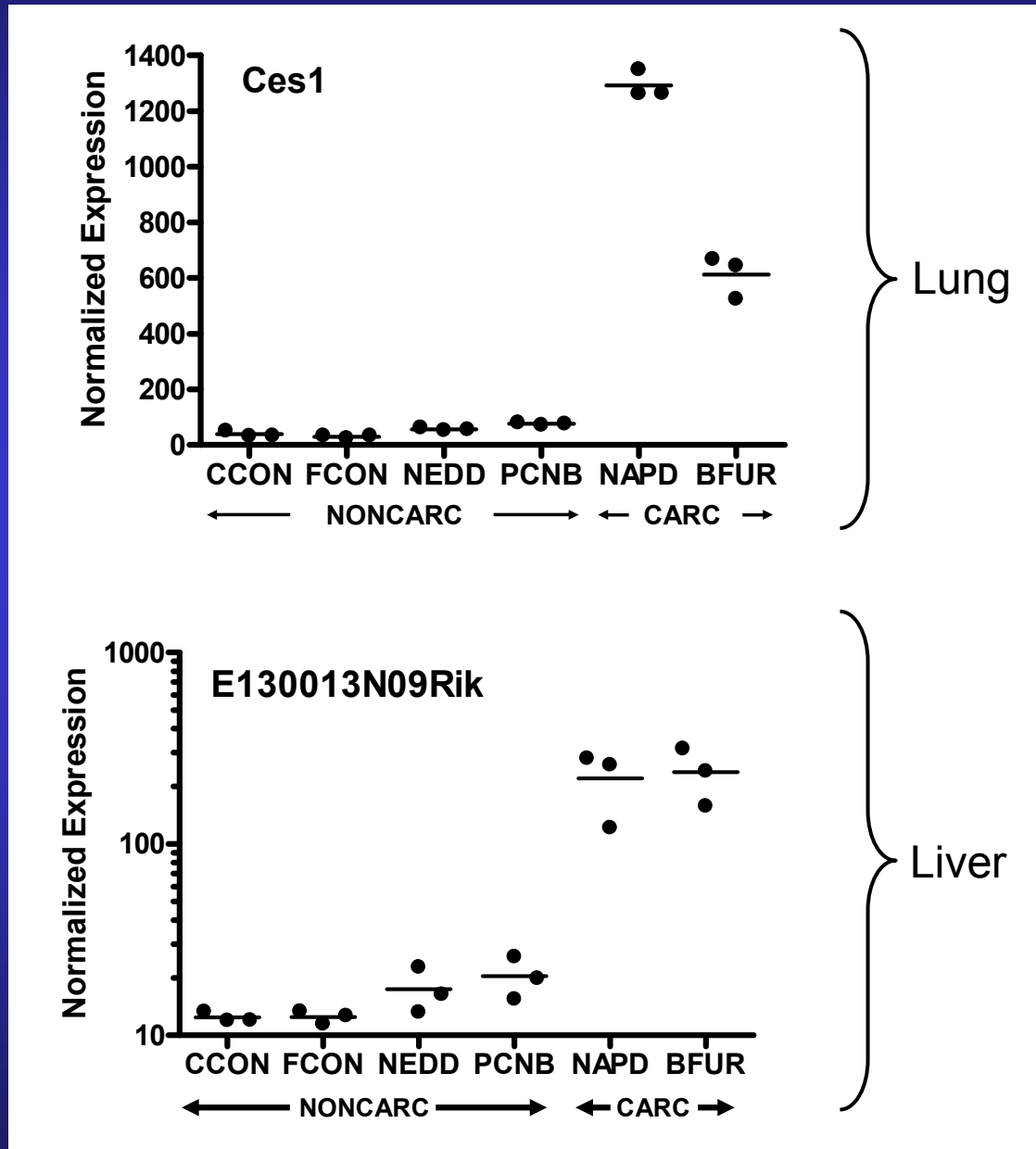
Liver Gene Expression



CCON FCON NEDD PCNB NAPD BFUR
← NONCARC → ← CARC →

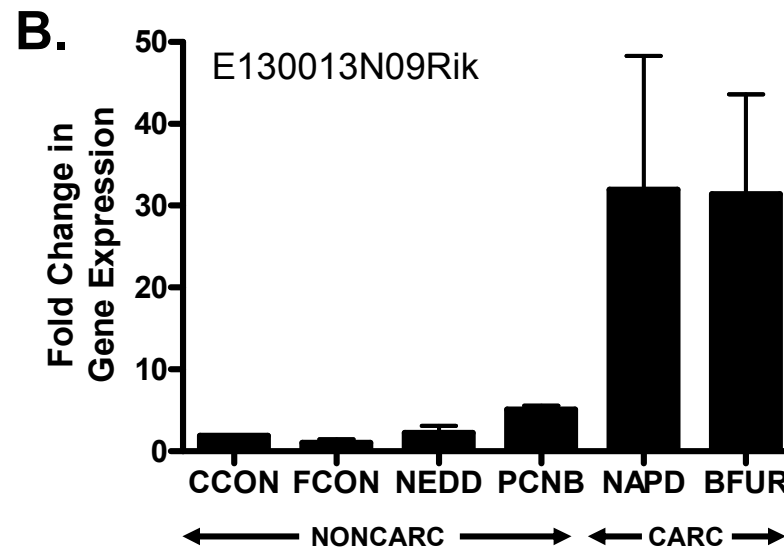
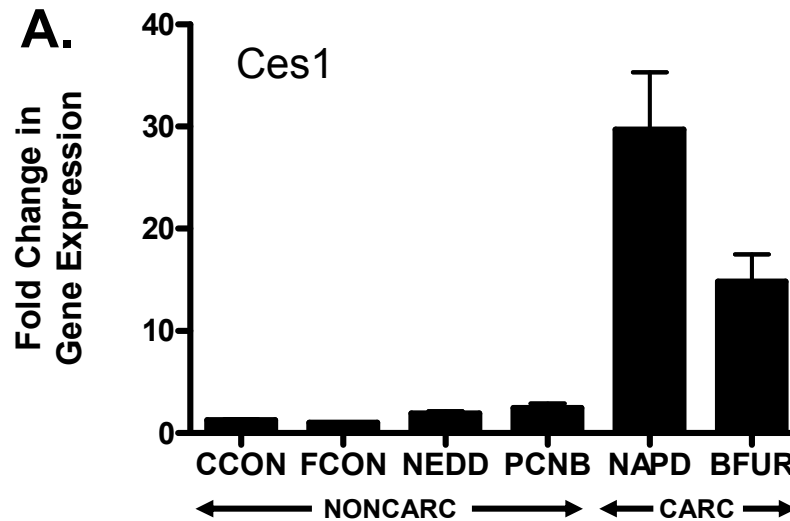
Urinary Metabolites

Experimental Results



Experimental Results

qRT-PCR Confirmation



Application of Genomics to Hazard ID

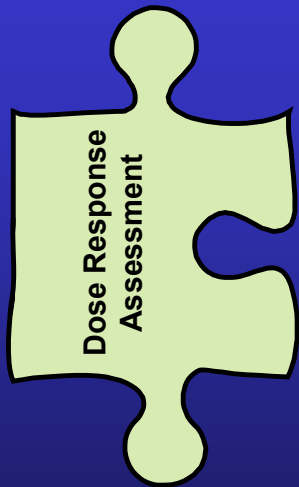
Results and Conclusions

- Statistical classification analysis of gene expression data in both the lung and liver samples exhibited 100% predictive accuracy based on three-fold cross-validation.
- Statistical classification analysis of urinary metabolites exhibited 88% predictive accuracy based on three-fold cross-validation.
- The two carcinogenic chemicals used in the study showed common changes despite different chemical structures, genotoxicity categories, and potential modes-of-action.

Is There a Role for Genomics in the 3Rs...

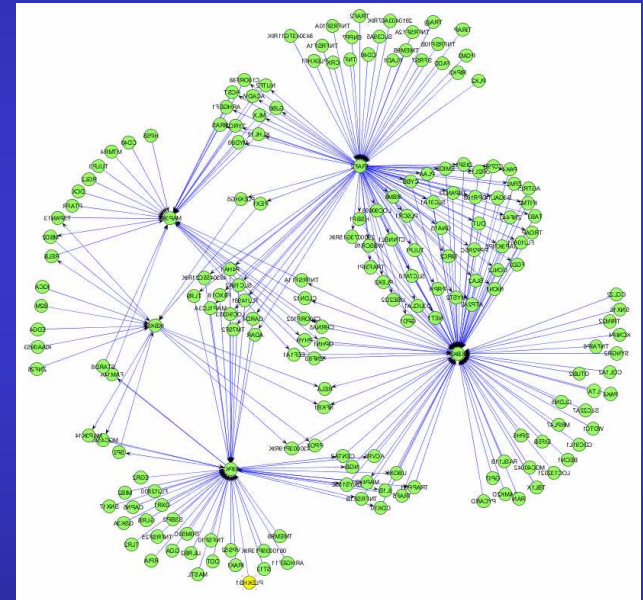


Application of Genomic and Metabonomic Technology to Identify Biomarkers Predictive of Rodent Cancer Bioassay



Application of High-Coverage Functional Genomic Screens to Dissect Cell Signaling Networks

Application of Genomics to Dose Response Assessment



Lessons from the Industrial Revolution...

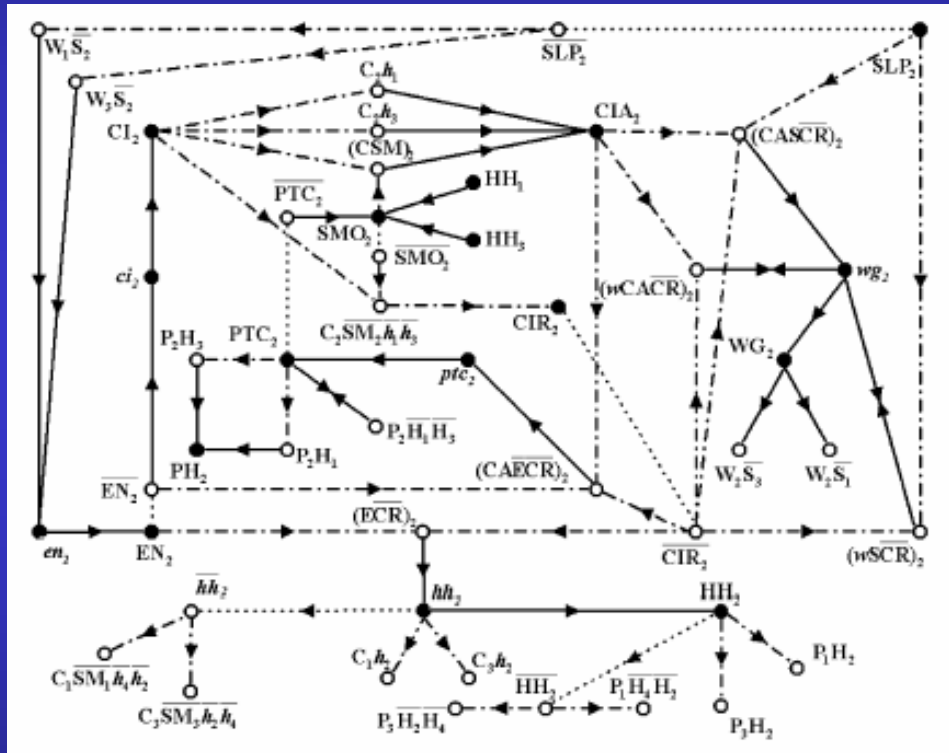
Application of Genomics to Dose Response Assessment

A fundamental understanding of the underlying biology will provide...

- potential molecular targets involved
- logic of the signaling network
- shape of the dose response curve
- how well this response is conserved across species

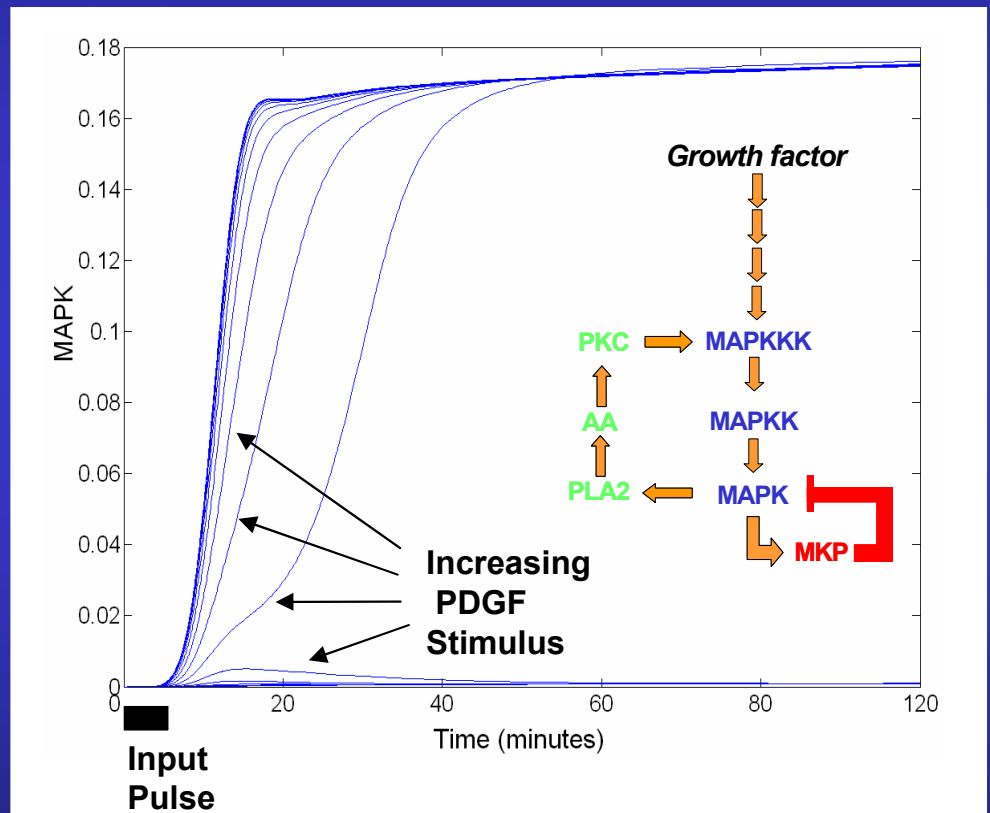
Examples With Drosophila and MAPK Signaling

Boolean Model of Drosophila Segment Polarity Gene Expression



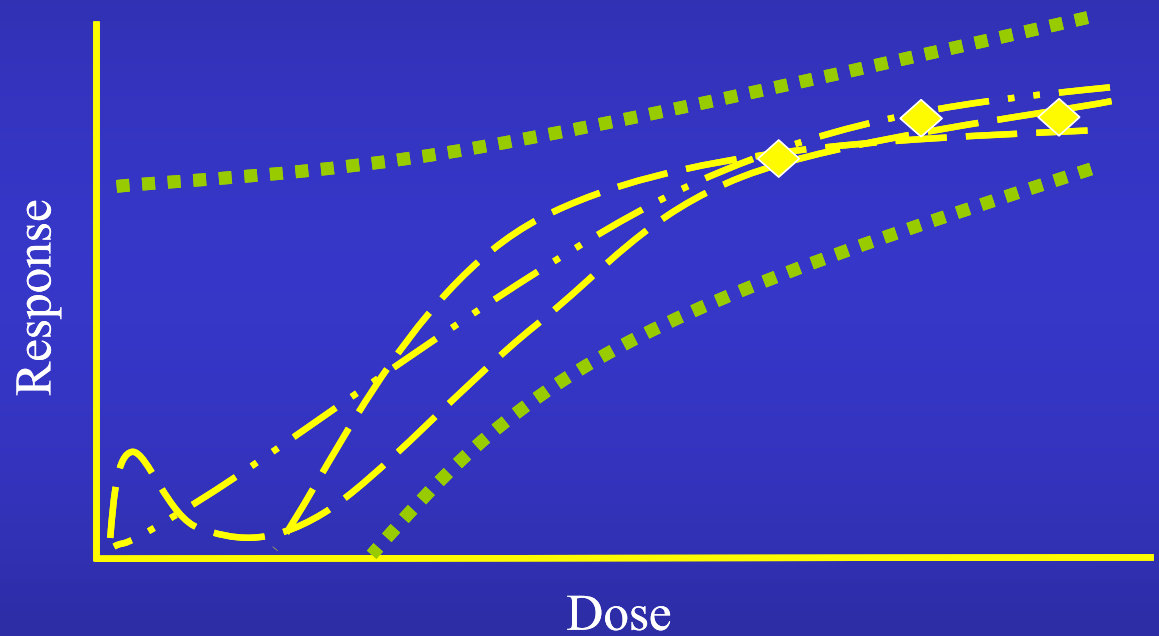
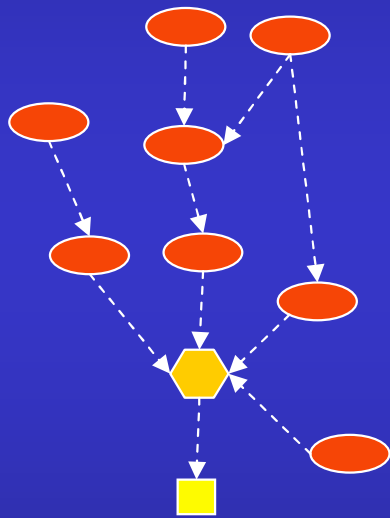
Albert R, Othmer HG. J Theor Biol. 223:1-18, 2003.

ODE Model of PDGF Stimulated MAPK



Bhalla, US *et al.* Science. 297:1018, 2002.
Conolly and Zhang, unpublished data.

Application of Genomics to Dose Response Assessment



How Do We Implement a Functional Genomics Approach to Dissect These Networks?

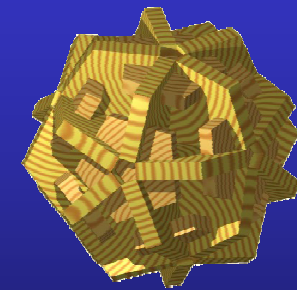
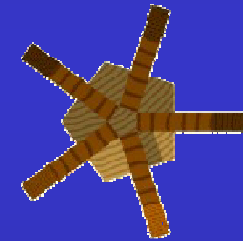
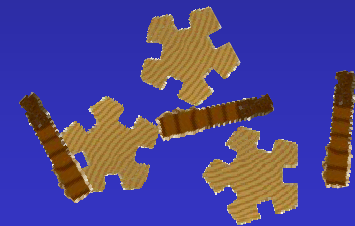
Assemble the “Parts List”



Identify How the Pieces Fit Together in Each Subsystem



Understand How the Subsystems Interconnect



How Do We Implement a Functional Genomics Approach to Dissect These Networks?

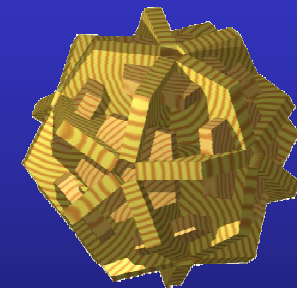
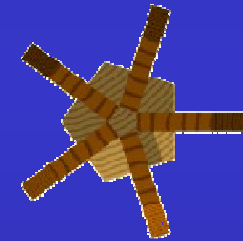
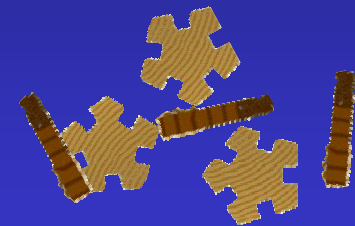
Assemble the “Parts List”



Identify How the Pieces Fit Together in Each Subsystem

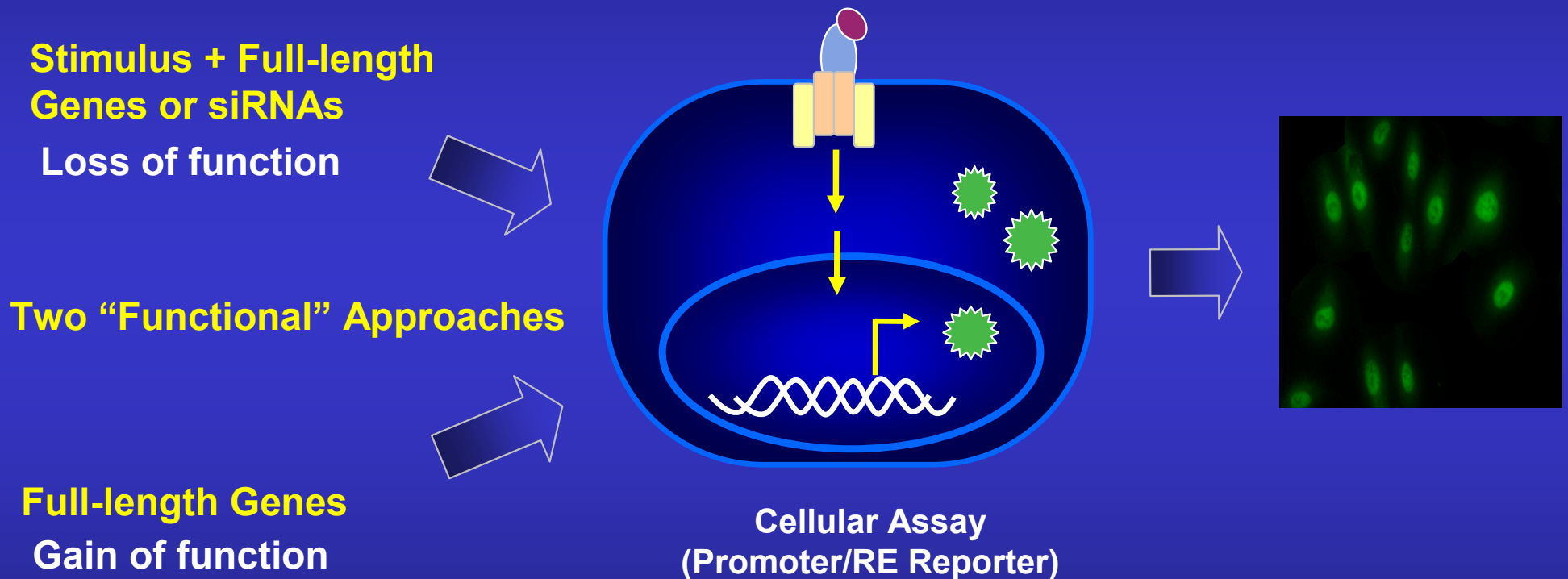


Understand How the Subsystems Interconnect



Assembling the “Parts List”

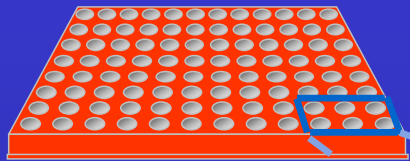
Anatomy of a Screen: Constructing The Assay



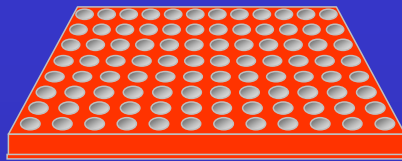
Assembling the “Parts List”

Anatomy of a Screen

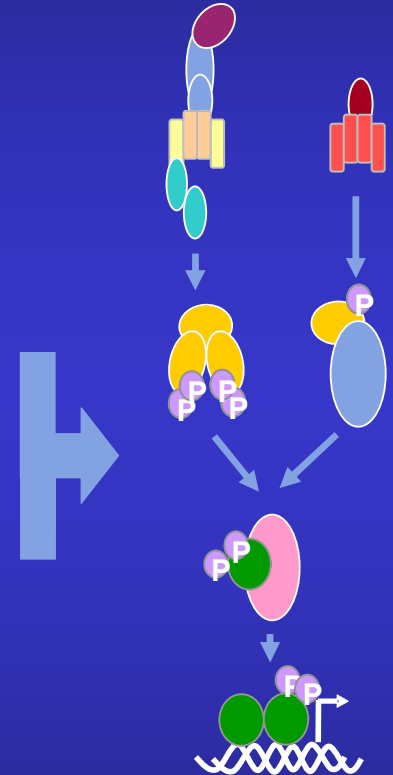
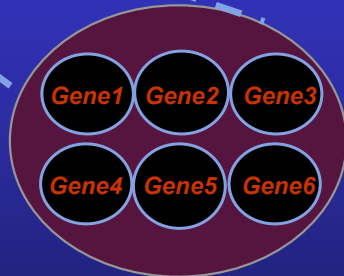
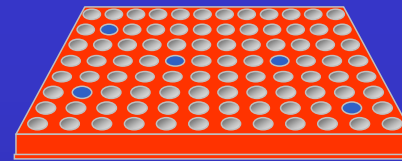
Arrayed, full-length gene or siRNA set in 384-well plates



Transfect genes into reporter cells



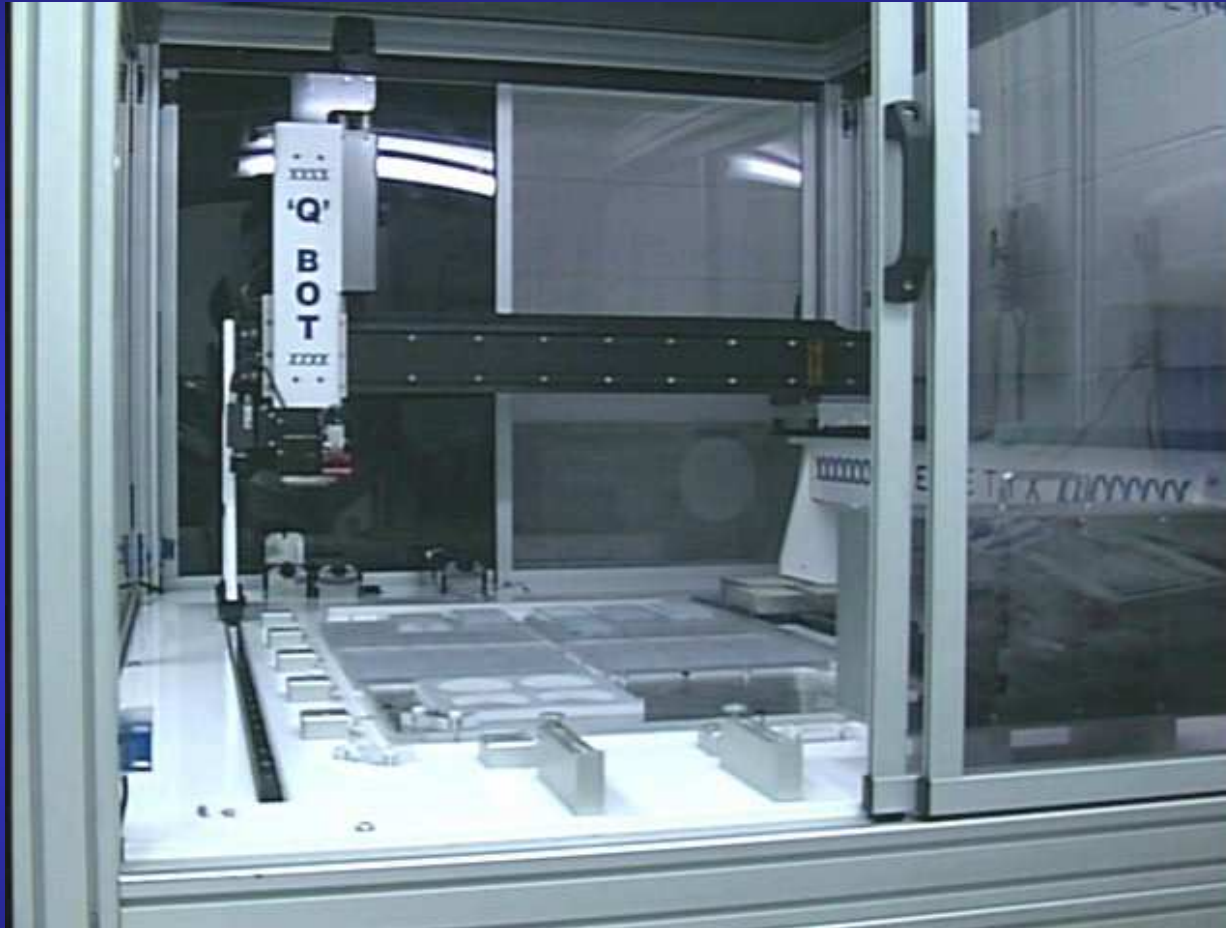
Identify hits



Identify components of the signaling pathway

Assembling the “Parts List”

Anatomy of a Screen



Movie Showing Robotic
Preparation of Full-length cDNA
library

Assembling the “Parts List”

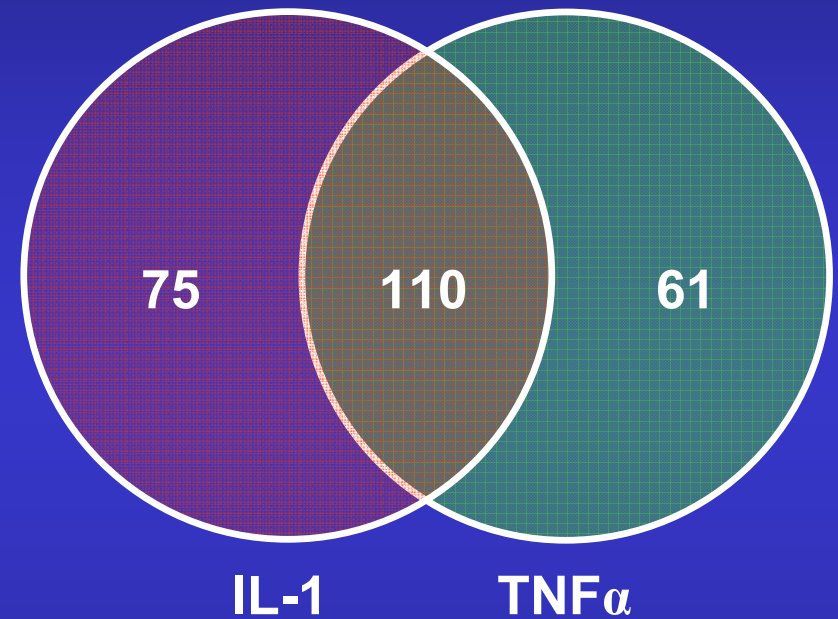
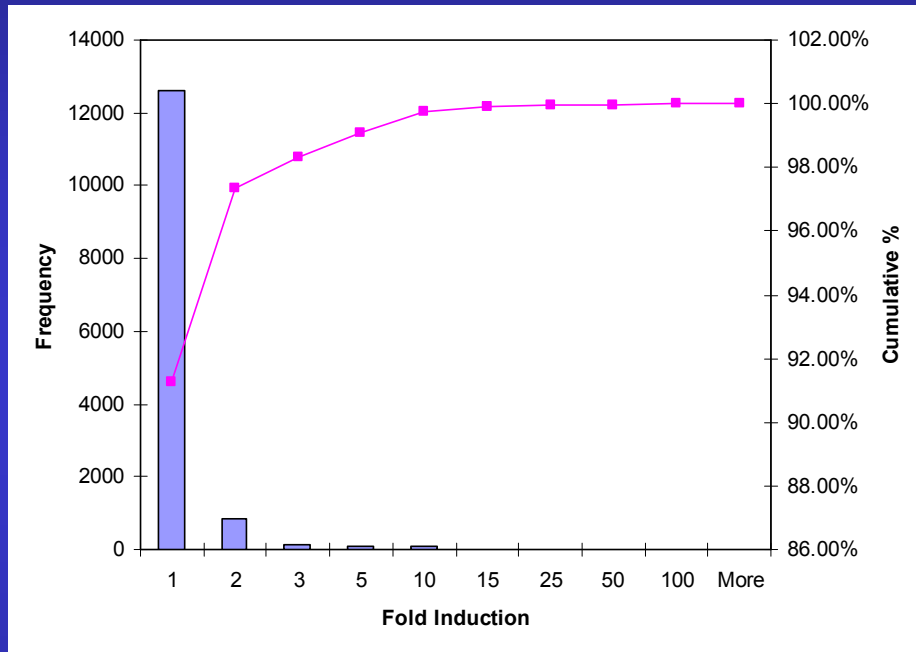
Anatomy of a Screen



Movie Showing Robotic
Screening of Full-length cDNA
library

Preliminary Results

NFkB Gain-of-Function and Loss-of-Function Screens



Screen Type: Full-length Gene
Gain-of-Function
Genes Screened: ~14,000

Screen Type: Full-length Gene
Loss-of-Function
Genes Screened: ~14,000

How Do We Implement a Functional Genomics Approach to Dissect These Networks?

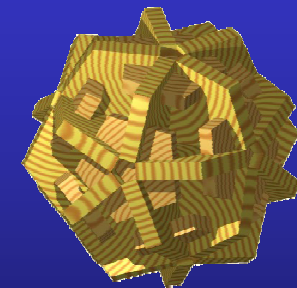
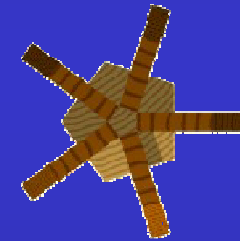
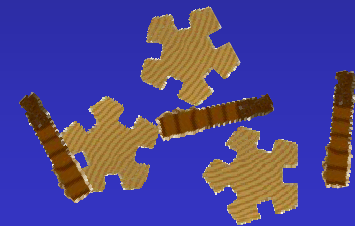
Assemble the “Parts List”



Identify How the Pieces Fit Together in Each Subsystem

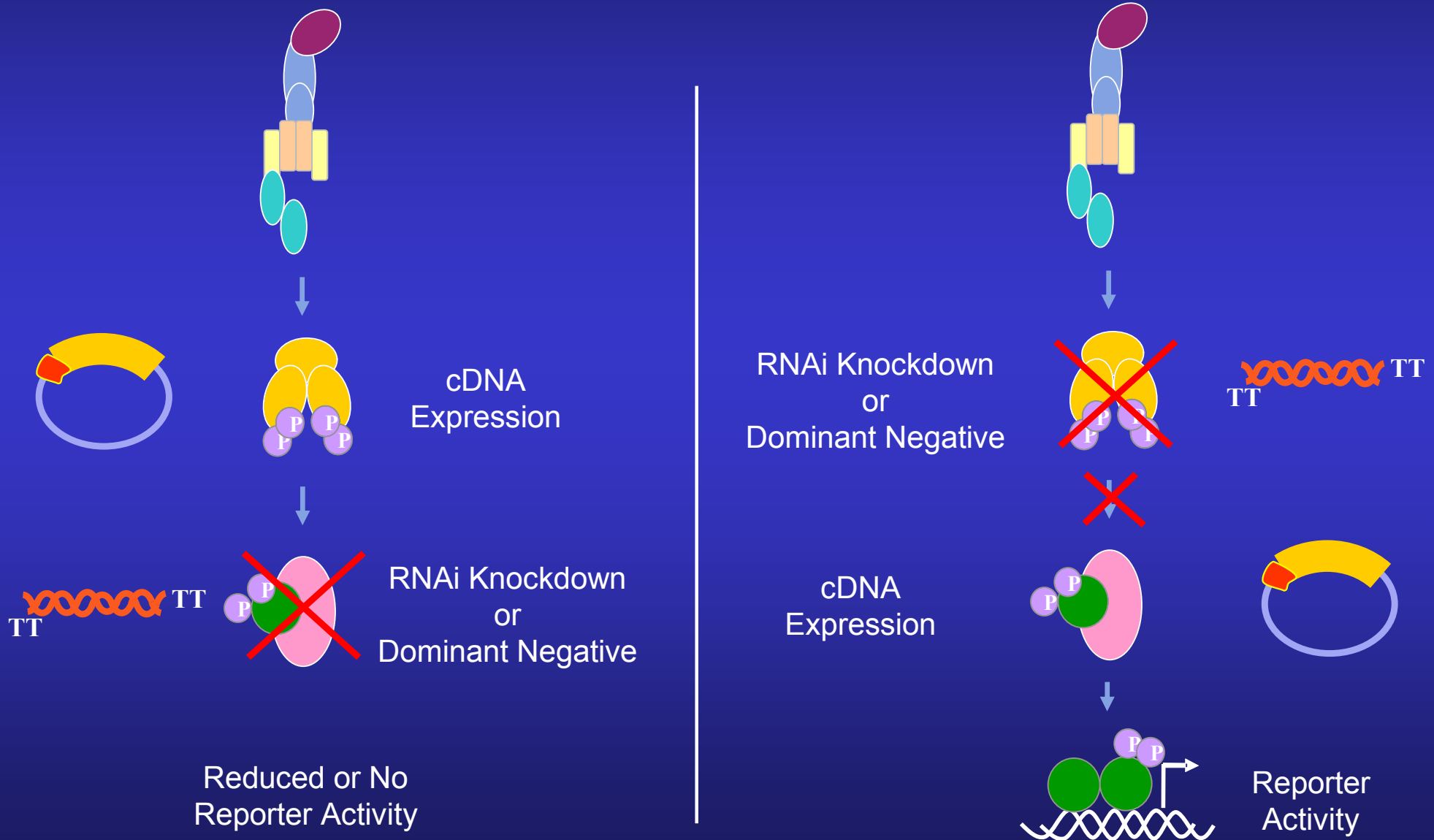


Understand How the Subsystems Interconnect



Identify How the Pieces Fit Together

Anatomy of a Screen: Organizing the Pathway



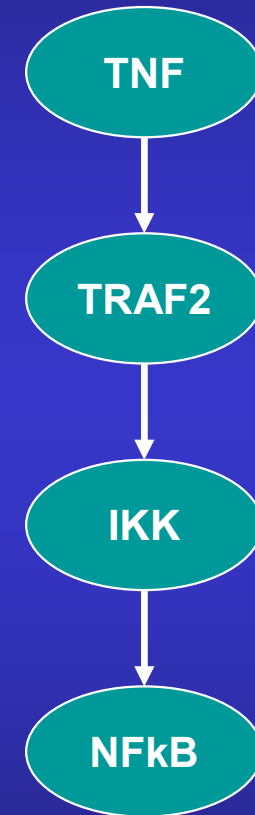
Identify How the Pieces Fit Together

Anatomy of a Screen: Organizing the Pathway

TRAF2	IKK	1
NFKB	IKK	-1
TNF	IKK	1

IKK	TRAF2	-1
NFKB	TRAF2	-1
TNF	TRAF2	1

TRAF2	IKK	1
TRAF2	NFKB	1
NFKB	IKK	-1
TNF	TRAF2	1
TNF	IKK	1
NFKB	TNF	-1



Individual Screen Matrices

Combined Matrix

Functional Network

Identify How the Pieces Fit Together

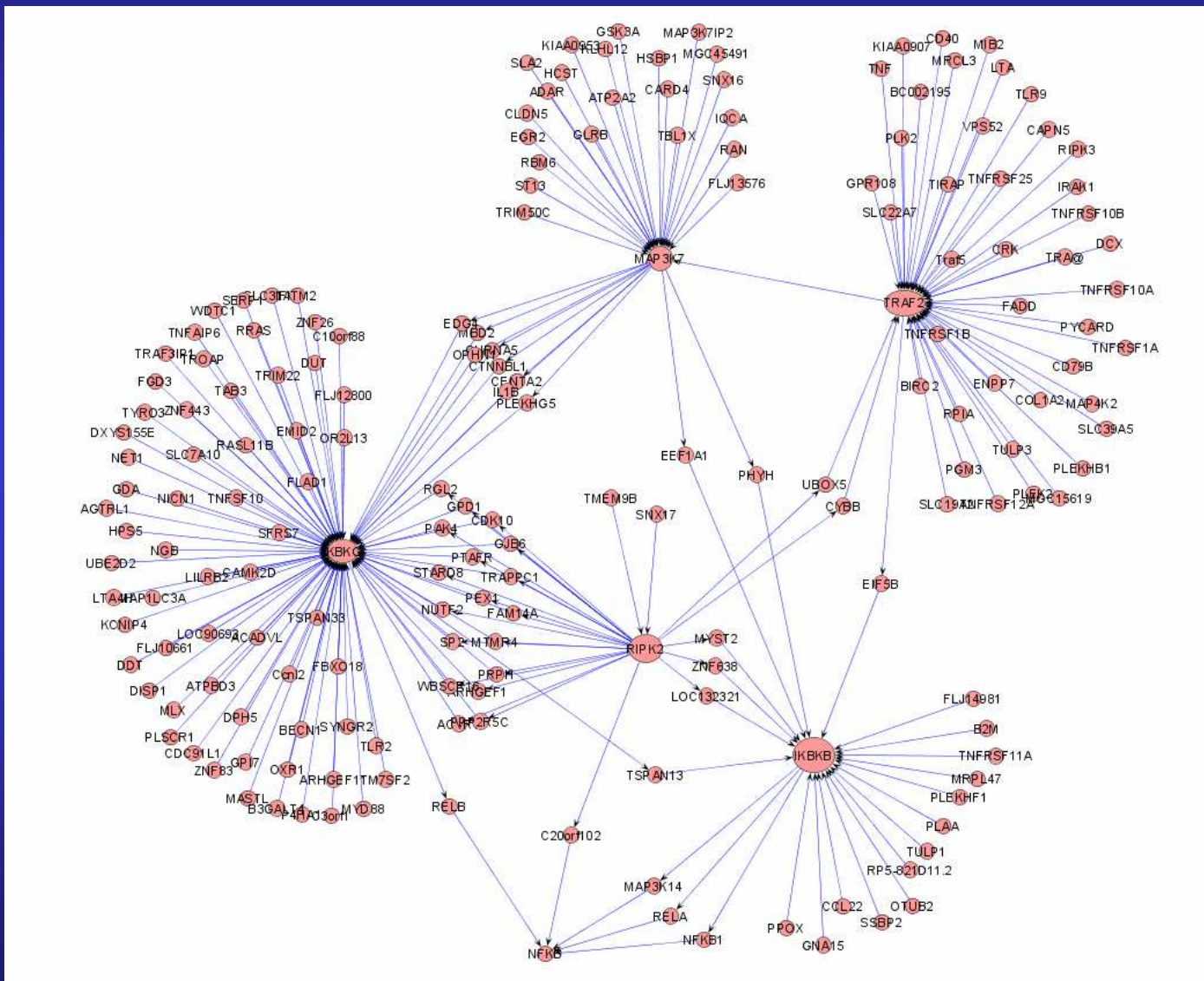
Anatomy of a Screen: Organizing the Pathway

The screenshot displays the CDNA Screen Data System interface. The main window shows a table with columns for File, Luc, Ds, L, Ds, Nor, Ratio, GenB, Gene Sy, and GenBan. A dialog box titled "Interaction Analysis of Screens" is open, showing settings for Data Column (Ratio), Gene Identifier (GenBank Accession), Upper Bound Cutoff (80%), Lower Bound Cutoff (30%), Reference Screen (Analysis 21), and Genetically Modified Screens (Analysis 20, 21, 22, 28). Below the dialog, an "Adjacency Matrix" window shows a grid of interactions between genes, with columns for Name, ACADVL, ACVR1, ADAR, AGTRL1, ARHGEF1, ARHGEF11, ATP2A2, ATPBD3, B2M, B3GALT4, and BBB. The matrix cells contain terms like "Upstream", "Downstream", "Marginal", and "Upstream".

Name	ACADVL	ACVR1	ADAR	AGTRL1	ARHGEF1	ARHGEF11	ATP2A2	ATPBD3	B2M	B3GALT4	BBB
ACADVL	Upstream	Downstream	Marginal	Upstream	Downstream	Downstream	Upstream	Marginal	Downstream	Marginal	Marginal
ACVR1	Downstream	Marginal	Marginal	Upstream	Downstream	Downstream	Upstream	Marginal	Downstream	Upstream	Downstream
ADAR	Upstream	Marginal	Upstream	Upstream	Downstream	Downstream	Upstream	Marginal	Downstream	Upstream	Downstream
AGTRL1	Upstream	Downstream	Upstream	Downstream	Marginal	Downstream	Upstream	Marginal	Downstream	Upstream	Downstream
ARHGEF1	Upstream	Downstream	Upstream	Upstream	Upstream	Upstream	Marginal	Downstream	Upstream	Upstream	Marginal
ARHGEF11	Marginal	Marginal	Downstream	Upstream	Upstream	Upstream	Marginal	Downstream	Upstream	Upstream	Marginal
ATP2A2	Downstream	Downstream	Upstream	Downstream	Upstream	Upstream	Upstream	Marginal	Upstream	Marginal	Downstream
ATPBD3	Marginal	Downstream	Marginal	Downstream	Upstream	Upstream	Marginal	Marginal	Downstream	Marginal	Downstream
B2M	Downstream	Upstream	Upstream	Downstream	Marginal	Marginal	Downstream	Marginal	Marginal	Marginal	Downstream
B3GALT4	Upstream	Marginal	Upstream	Upstream	Marginal	Upstream	Upstream	Upstream	Marginal	Marginal	Marginal
BBB	Upstream	Upstream	Downstream	Upstream	Upstream	Upstream	Downstream	Marginal	Downstream	Marginal	Marginal
BC002195	Downstream	Upstream	Upstream	Upstream	Upstream	Upstream	Downstream	Marginal	Downstream	Marginal	Downstream
BECN1	Upstream	Upstream	Marginal	Upstream	Upstream	Upstream	Marginal	Marginal	Marginal	Marginal	Downstream
BIRC2	Downstream	Upstream	Marginal	Downstream	Marginal	Upstream	Marginal	Downstream	Marginal	Marginal	Marginal
C10ORF88	Marginal	Marginal	Marginal	Upstream	Marginal	Marginal	Upstream	Downstream	Marginal	Marginal	Marginal
C20ORF102	Marginal	Marginal	Marginal	Upstream	Marginal	Marginal	Upstream	Downstream	Marginal	Upstream	Marginal
C3ORF1	Marginal	Upstream	Upstream	Downstream	Downstream	Upstream	Upstream	Upstream	Upstream	Upstream	Marginal
CAMK2D	Upstream	Upstream	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Downstream	Marginal	Marginal
CAPN5	Marginal	Downstream	Marginal	Upstream	Marginal	Marginal	Marginal	Marginal	Downstream	Marginal	Downstream
CARD4	Downstream	Downstream	Marginal	Marginal	Marginal	Upstream	Downstream	Upstream	Upstream	Upstream	Marginal
CCC	Downstream	Downstream	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal
CCL22	Marginal	Upstream	Downstream	Marginal	Marginal	Upstream	Marginal	Downstream	Marginal	Upstream	Marginal
CCNL2	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Upstream	Upstream	Upstream	Upstream	Downstream
CD40	Downstream	Marginal	Marginal	Upstream	Downstream	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal
CD79B	Upstream	Marginal	Marginal	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Marginal
CDC91L1	Upstream	Upstream	Upstream	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal
CDK10	Upstream	Upstream	Upstream	Downstream	Marginal	Downstream	Marginal	Upstream	Upstream	Marginal	Marginal
CENTA2	Upstream	Upstream	Upstream	Downstream	Marginal	Marginal	Marginal	Marginal	Marginal	Downstream	Marginal
CHRNA5	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream
C10M12	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream	Upstream

Preliminary Results

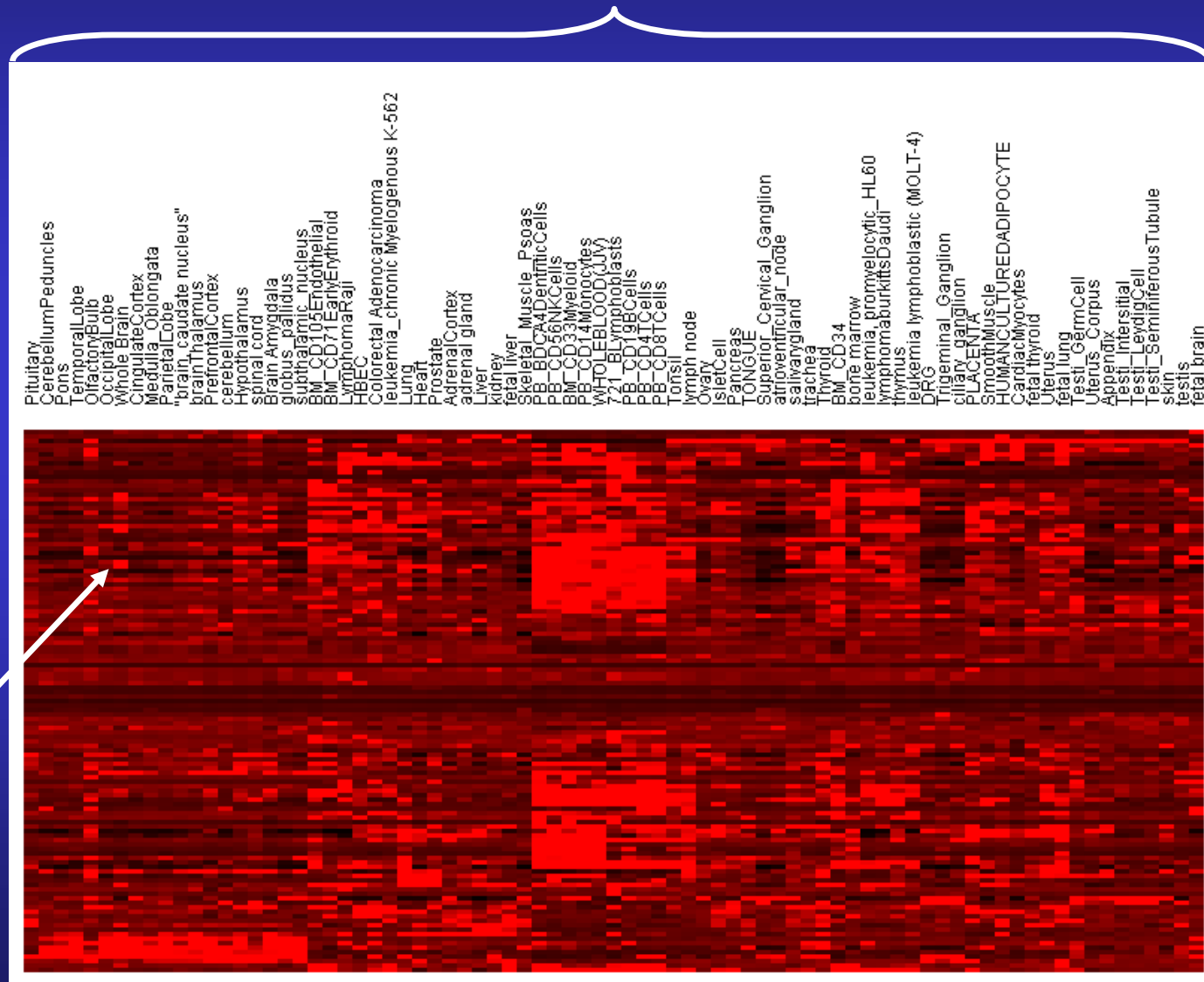
Functional Network Map



Preliminary Results

Tissue Specific Networks

79 Human Tissues



Intensity is the expression of that gene in the tissue (i.e., is it there?)

NFkB Functional Hits

*Based on expression data in www.symatlas.org

Preliminary Results

Tissue Specific Networks

Gene Atlas Expression Data Filter - 3_1_06

Organism: **human** Array Type: **HG-U133A** Data Type: **RMA** Retrieve

Expression Data

Gene Symbol	721 B-lymphobl...	Adipocyte	Adrenal cor...	Adrenal gland	Amygdala	Appendix	Atrioventricular n...	BM-
GSK3A	10.920856654644	11.4739586...	11.4551849...	11.5095354...	11.5132244...	11.4352694...	11.472749710083	11.0
ARHGGEF1	11.855183124542	11.9410393...	12.0958018...	11.9175513...	11.9492723...	11.9198535...	12.0368800958	11.7
IRAK1	12.270458698273	12.4555956...	12.1304766...	12.1156487...	11.9817431...	11.7652369...	11.739474376043	12.4
ADAR	12.689049720764	11.9281504...	12.0255355...	12.0841768...	12.4342433...	11.8998993...	11.699463923772	12.2
SYNGR2	12.570389270782	11.8406232...	12.0093310...	11.9890608...	12.0770318...	11.8421650...	11.838548660278	11.9
OPHN1	12.202203273773	12.7499779...	13.4319564...	12.3475886...	12.6480538...	13.1930155...	13.625313123067	12.8
EEF1A1	13.305665135384	13.0939081...	12.8625532...	12.5538527...	11.8778402...	12.6440111...	12.494581818581	12.8
OXR1	4.3904826641083	4.36766481...	4.57968831...	4.77729415...	6.91715002...	4.58815765...	4.4025285243988	4.60
SNX16	4.319258928299	4.41901516...	4.23261380...	4.37277865...	4.30990242...	5.1722702...	4.3965632915497	4.50
TRAF3IP1	4.3728847503662	4.43809413...	4.52264070...	4.30504083...	4.45959544...	4.15841984...	4.471364736557	4.51
ZNF83	4.4616777896881	4.50071144...	4.88819217...	4.39295458...	4.73030519...	4.66557049...	4.6383783817291	4.90
PLK2	4.3556568622589	4.41583466...	4.89693522...	4.48548293...	6.55782866...	4.82937073...	4.3621325492859	4.48

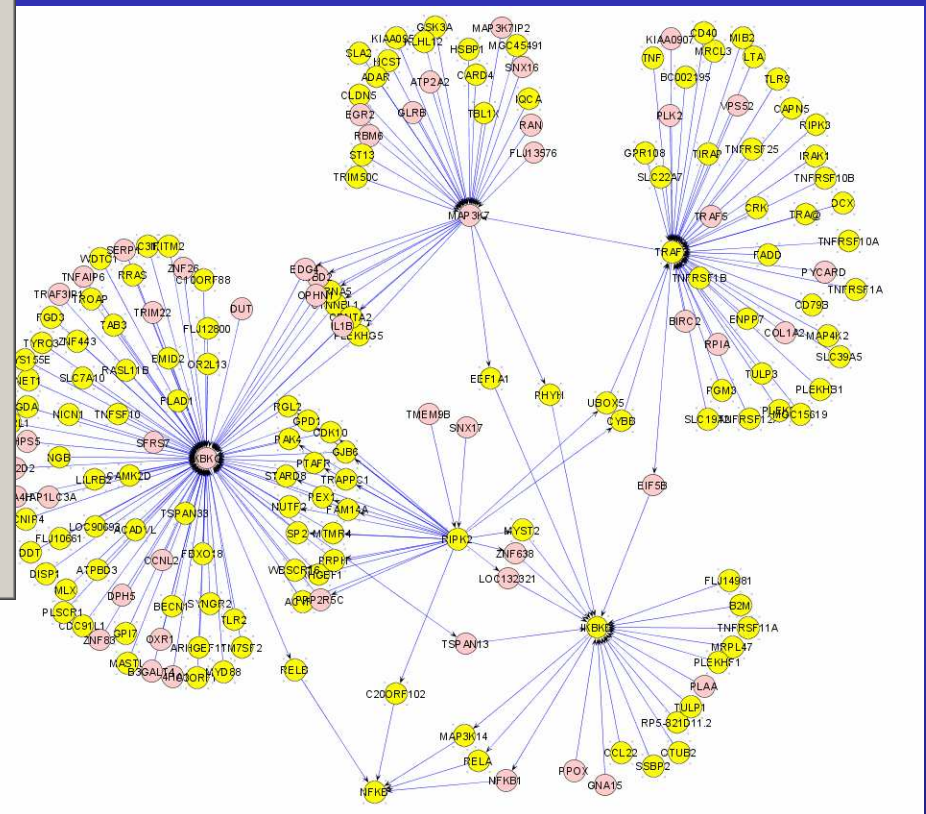
Filter Operation

Tissue Type: **Kidney**

Minimum: **4.144835** Maximum: **12.73332** Filter Value: **4.144835** Confidence Interval: Mean Median **99**

AP Call: **P**

Apply Filter New Network Close



*Expression values based on GNF SymAtlas (www.symatlas.org)

How Do We Implement a Functional Genomics Approach to Dissect These Networks?

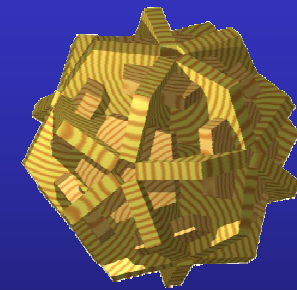
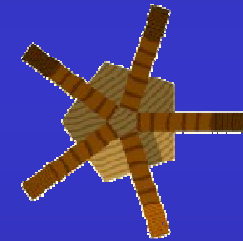
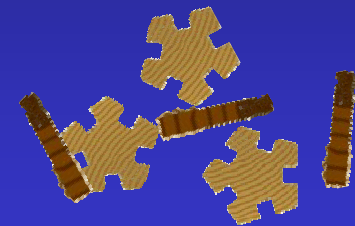
Assemble the “Parts List”



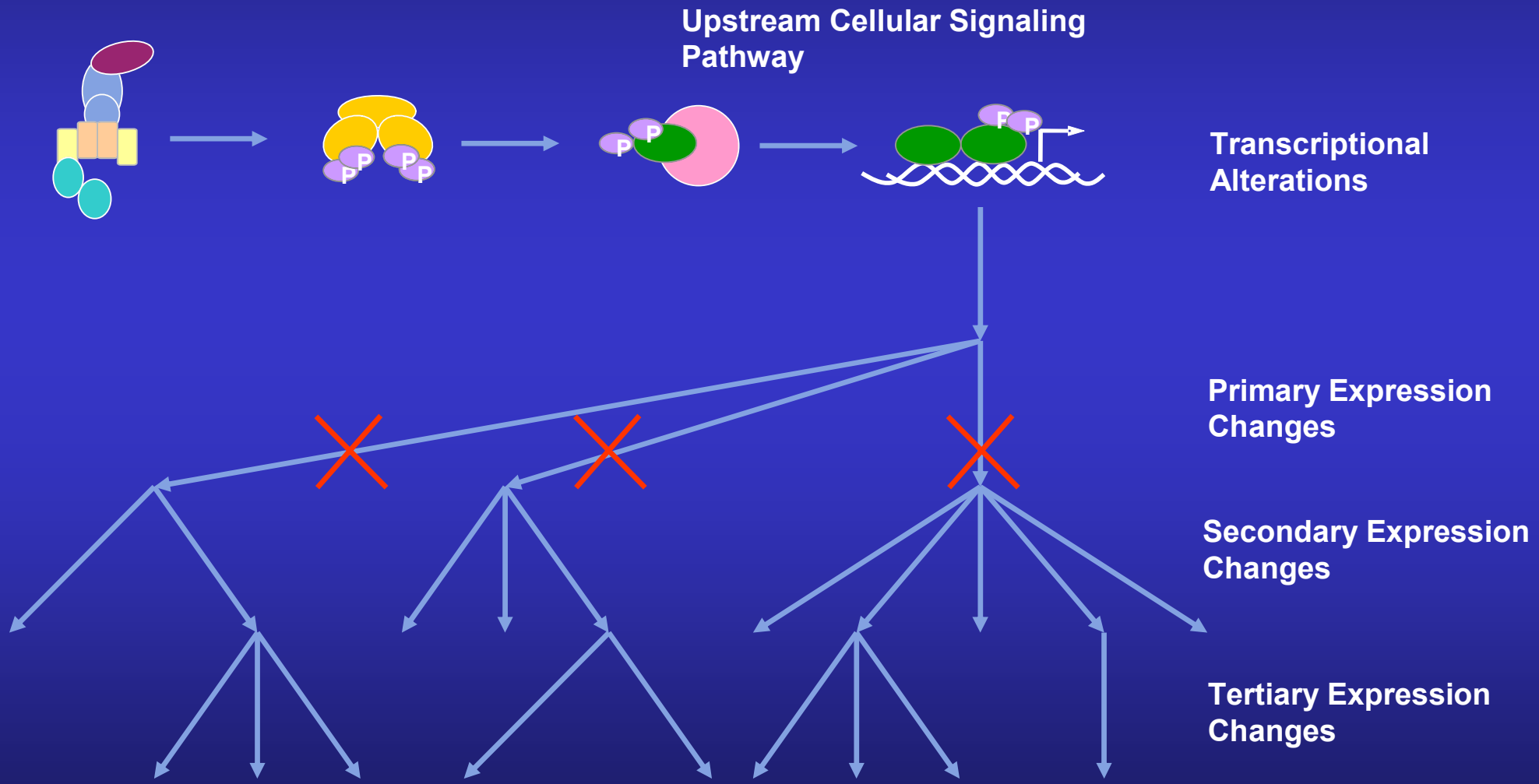
Identify How the Pieces Fit Together in Each Subsystem



Understand How the Subsystems Interconnect

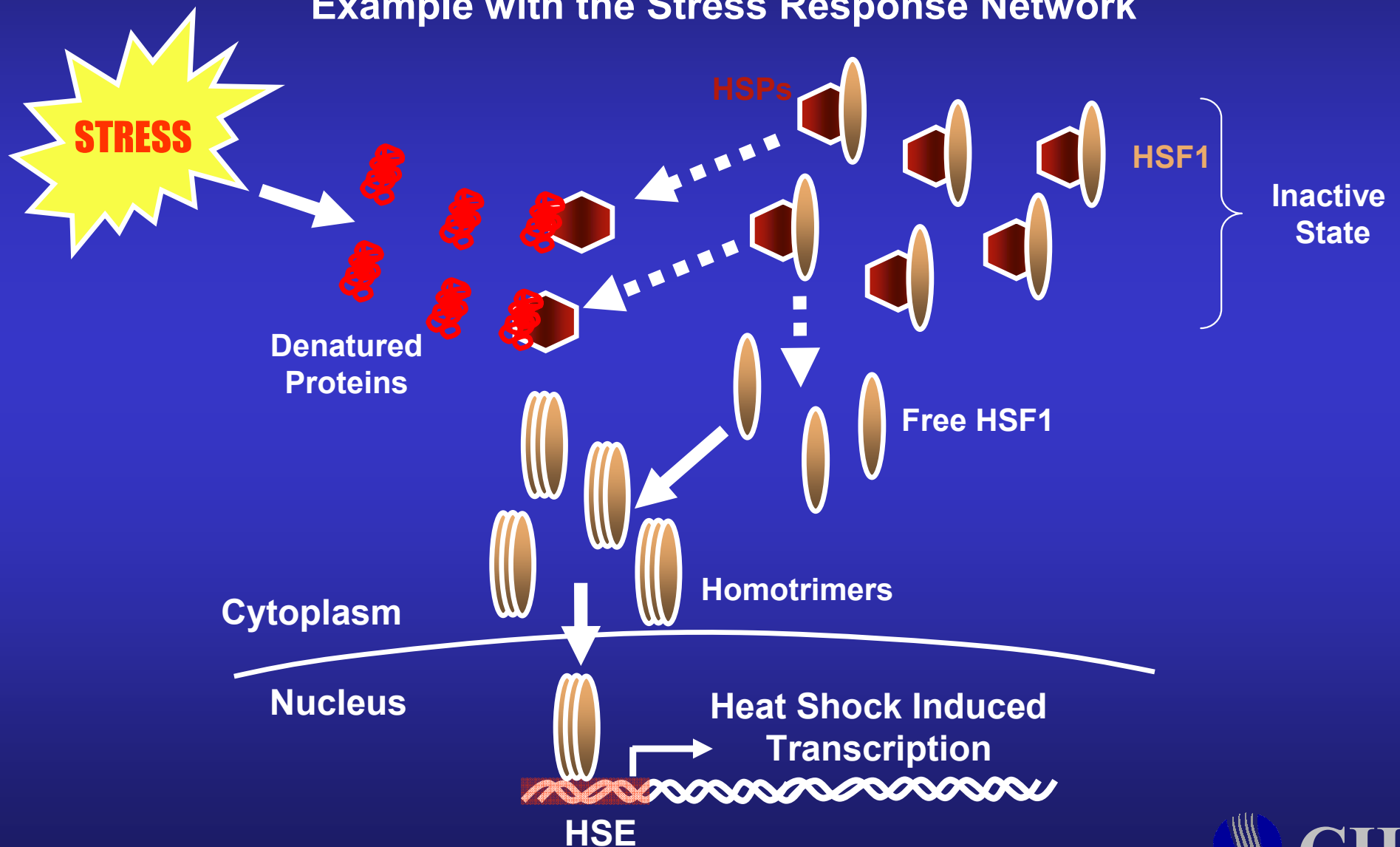


Understanding How the Subsystems Interconnect

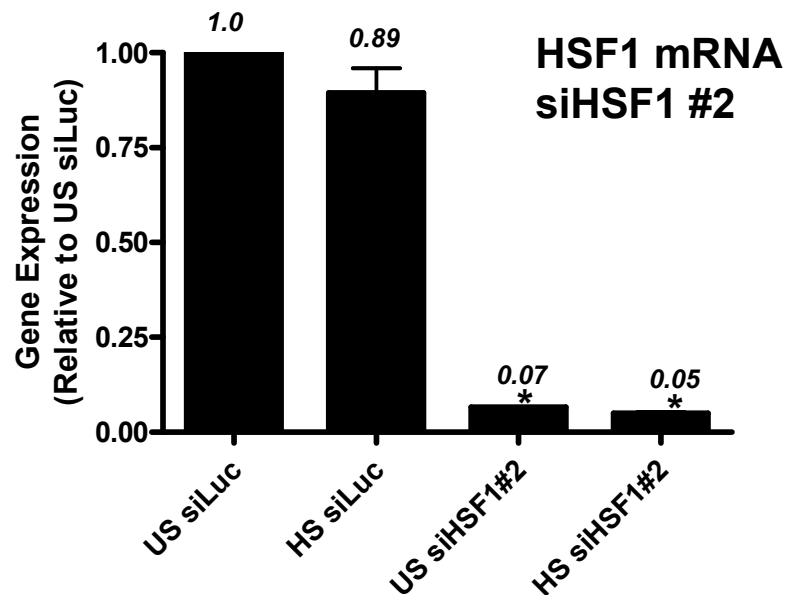
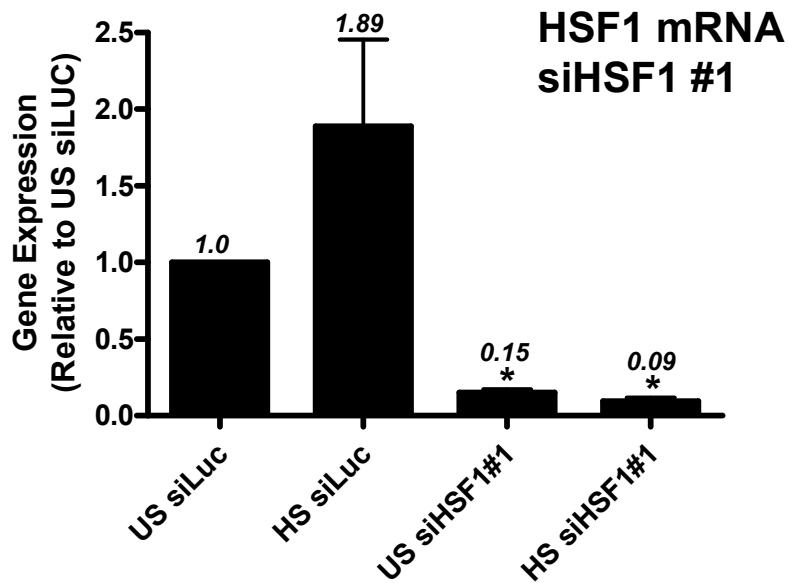


Understanding How the Subsystems Interconnect

Example with the Stress Response Network



Understanding How the Subsystems Interconnect

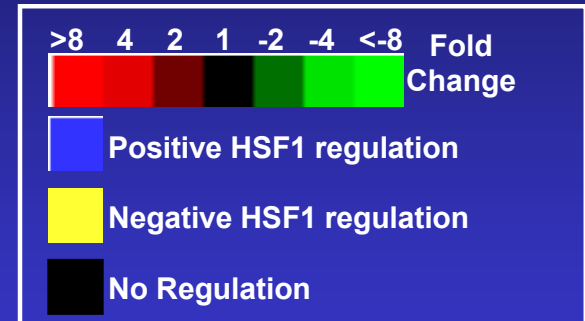
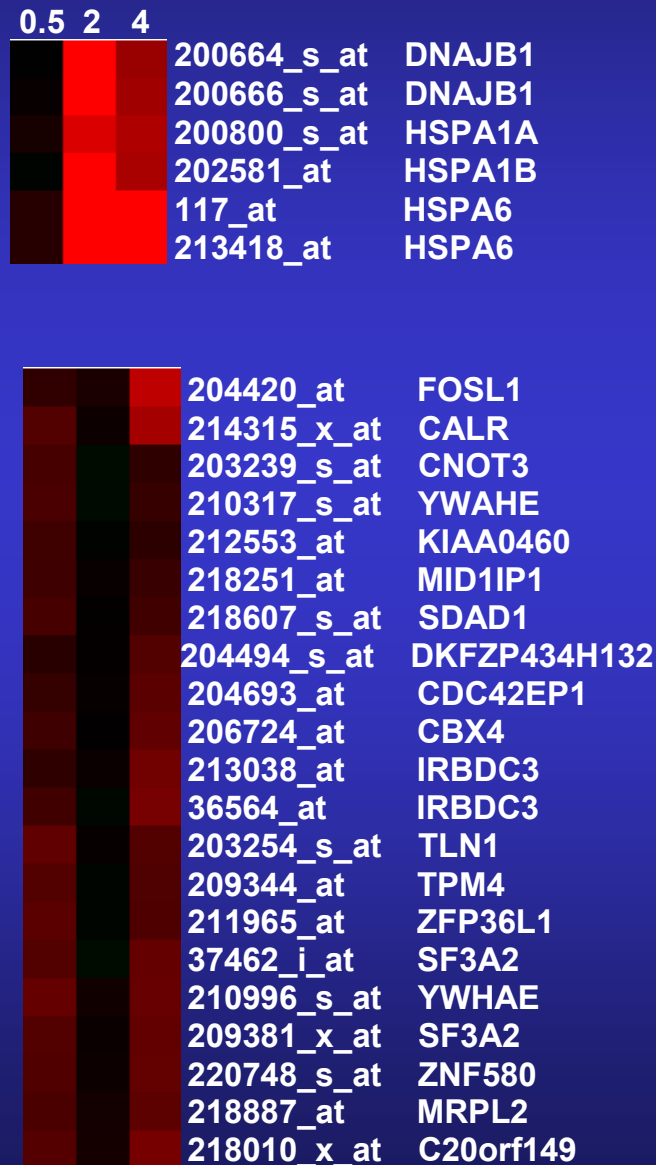
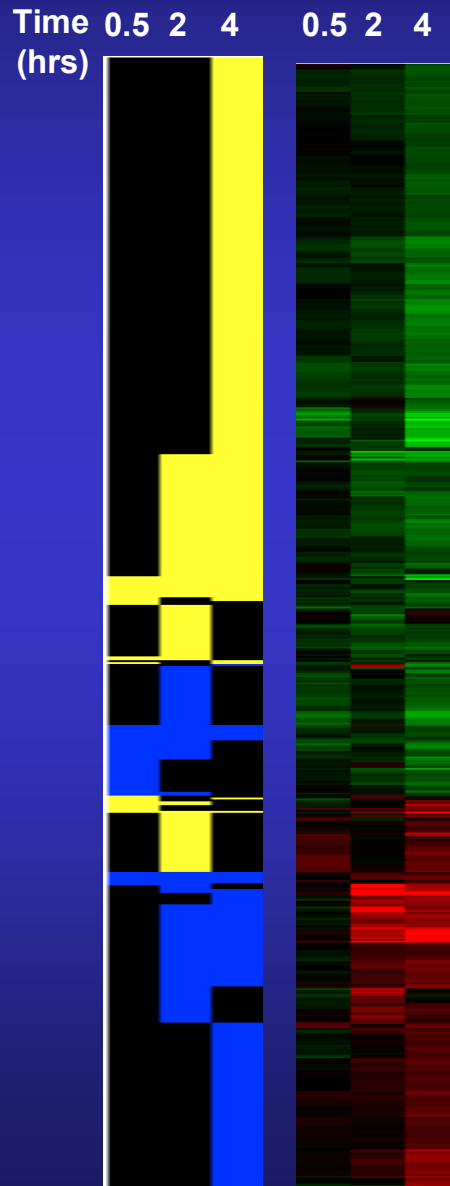


HSF1 Protein



US HS US HS US HS
siHSF1#1 siHSF1#2 siLuc

Understanding How the Subsystems Interconnect

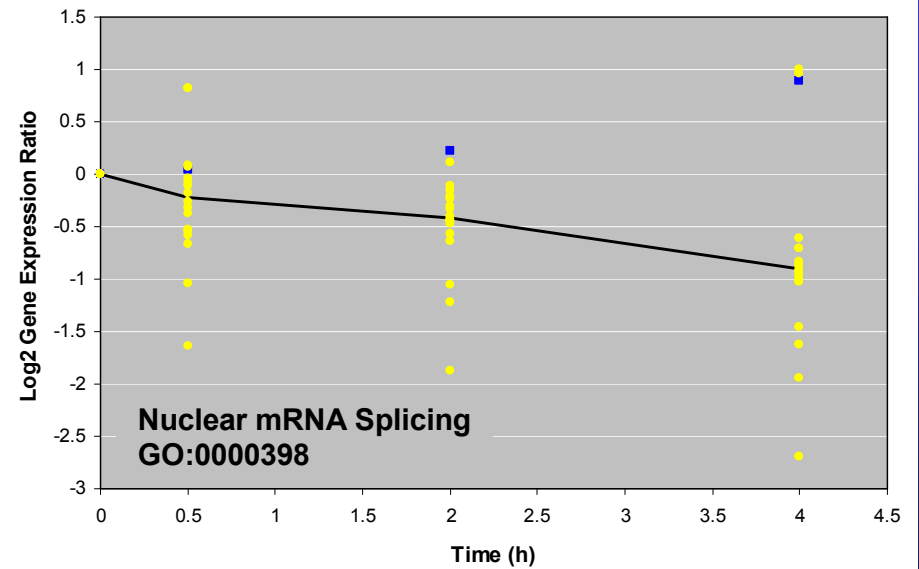
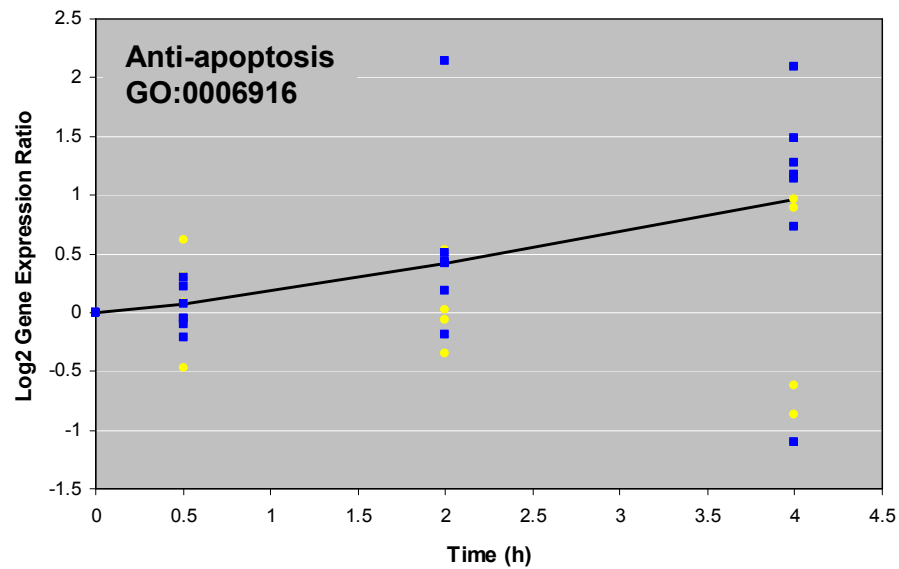
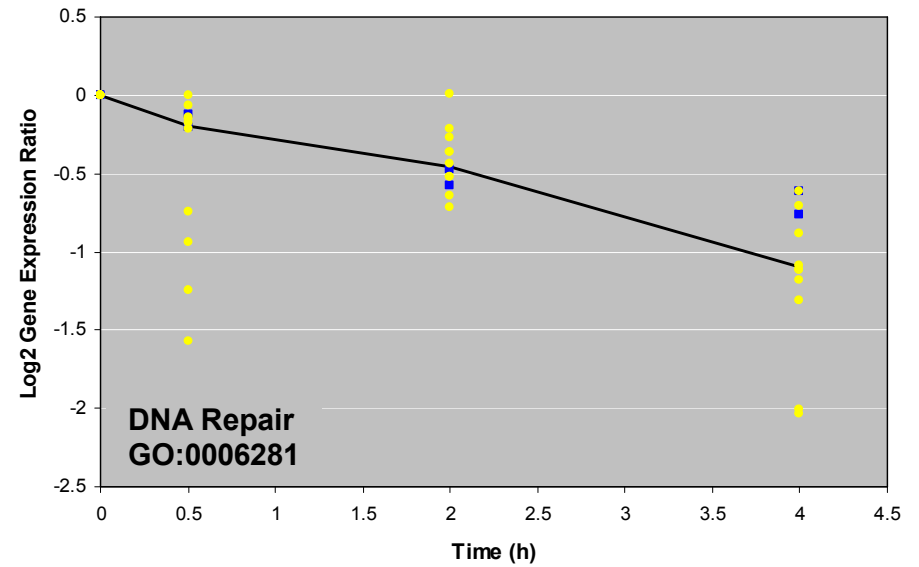
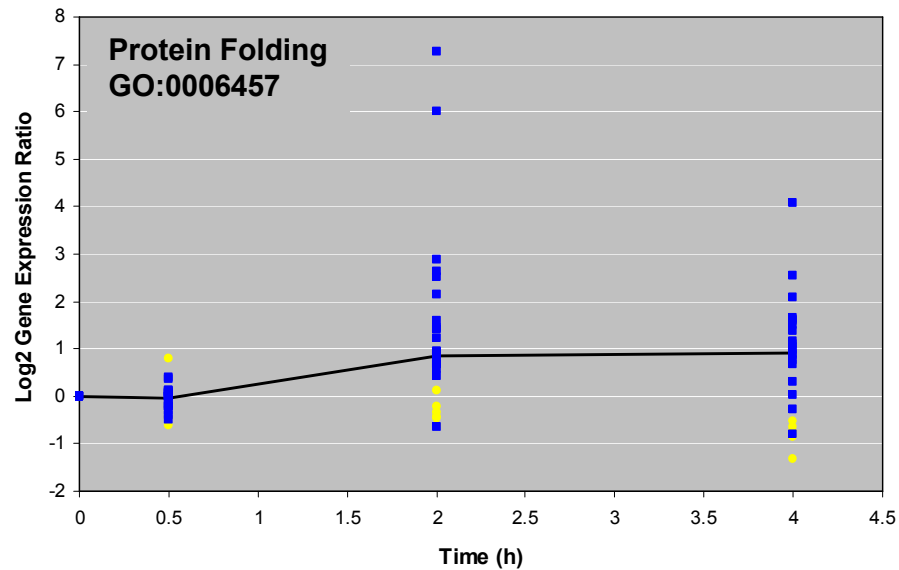


Understanding How the Subsystems Interconnect

Gene ontology (GO) analysis of genes with significant regulation by HSF1 and significant alterations in expression following proteotoxic stress.

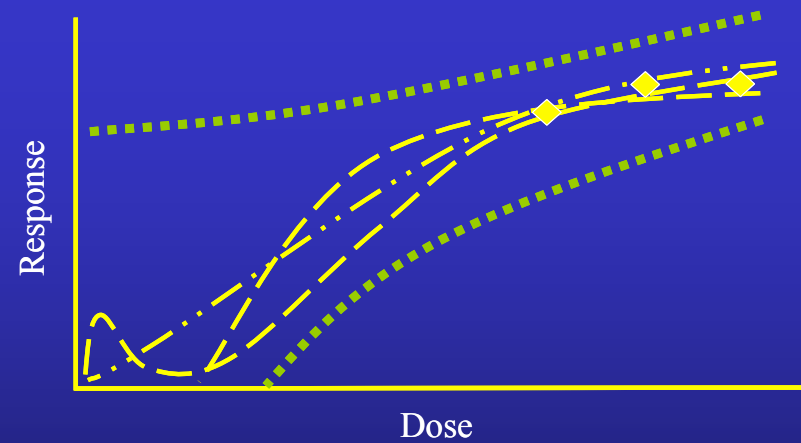
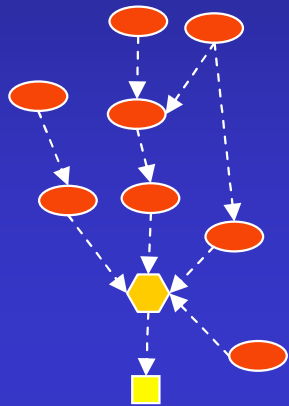
Biological Process GO Category	<i>p</i> -value	Mol. Function GO Category	<i>p</i> -value
Response to unfolded protein	0.000	DNA binding	0.000
Regulation of transcription, DNA-dependent	0.000	Metal ion binding	0.000
Protein folding	0.000	Unfolded protein binding	0.000
Transcription	0.000	Zinc ion binding	0.000
Chromosome organization and biogenesis	0.000	Heat shock protein binding	0.000
Nuclear mRNA splicing, via spliceosome	0.001	Nucleic acid binding	0.000
Anti-apoptosis	0.009	Small GTPase regulator activity	0.003
DNA repair	0.010	Ubiquitin thiolesterase activity	0.017
Protein ubiquitination	0.012	Cysteine-type endopeptidase activity	0.021
Response to virus	0.020	Transcription factor activity	0.023
Negative regulation of transcription	0.022	Chromatin binding	0.024
Response to stress	0.029	Binding	0.027
Regulation of cell cycle	0.029	Helicase activity	0.032
RNA splicing	0.031		

Understanding How the Subsystems Interconnect



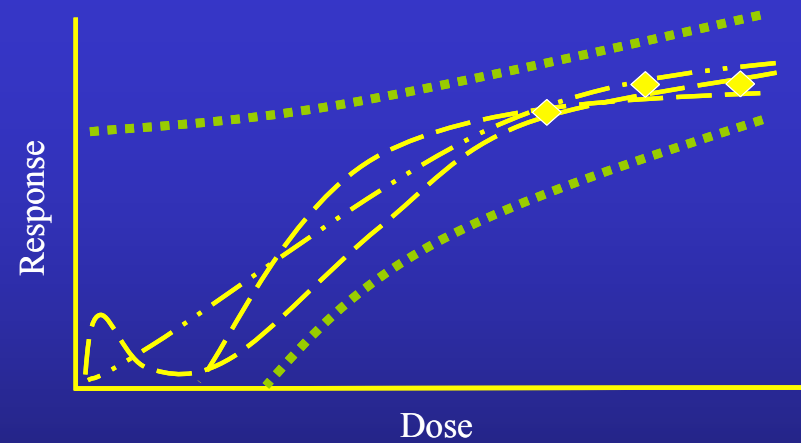
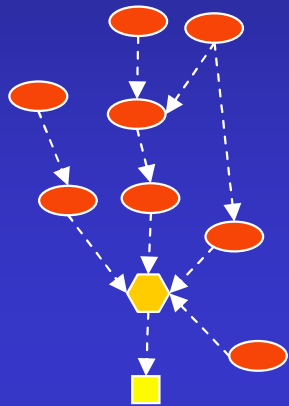
The Ultimate Goal

Apply Genomic Tools in an Integrated Approach



The Ultimate Goal

Apply Genomic Tools in an Integrated Approach



Acknowledgements

- Research Associates
 - Linda Pluta
 - Dana Stanley
- Postdocs
 - Tom Halsey
 - Todd Page
- Bioinformatics
 - Longlong Yang
- Senior Management
 - William Greenlee
 - Dave Dorman
 - Mel Andersen
- Funding
 - American Chemistry Council
 - NIEHS
- Collaborators
 - John Hogenesch (Scripps)
 - John Walker (GNF)
 - Russ Wolfinger (SAS)
 - Jeff MacDonald (UNC)
 - Tom Kodadek (UTSMS)
 - Devanjan Sikder (UTSMS)