



# ASBAMC

*Animal Systems Biology Analysis and Modeling Center*

## Computational Systems Biology Approach to Comparative Pathogenicity Transcriptome Analysis and Modeling

A Service provided by the Animal Systems Biology ASBAMC at UC Davis

**Huaijun Zhou, PhD**

ASBAMC Program Director (PI)  
Department of Animal Science  
University of California, Davis  
[hzhou@ucdavis.edu](mailto:hzhou@ucdavis.edu)

**Kenneth Drake, PhD**

ASBAMC (Co-PI)  
Chief Technology Officer  
Seralogix  
[drake@seralogix.com](mailto:drake@seralogix.com)

This presentation is supported by Agriculture and Food Research Initiative Competitive Grant no. 2011-67015-30196 from the USDA National Institute of Food and Agriculture

Computational tool development supported by National Institutes of Allergies and Infectious Diseases (SBIR grants 2R44AI058362-02 and R43AI084223-01) and National Human Genome Research Institute (SBIR grants 2R44HG004457-02)



# Presentation Outline



- Animal Systems Biology Analysis and Modeling Center
  - Rationale and mission
- In Silico System Biology Analysis and Modeling Overview
  - Computational pipeline in an integrated analysis and modeling platform
- Case study: comparative pathogenicity analysis and modeling
  - Chickens inoculated with *Campylobacter jejuni* (*C. jejuni*)
    - Broiler chickens: Resistant (Line A) and susceptible (Line B)
  - *C. jejuni* colonization in ceca
    - Wild type and mutant strains
- Important Contact Information



# The ASBAMC at UC Davis

A joint university and industrial program

- Rationale
  - Exponential growth of “omic” data past and future
    - Microarray, next generation sequencing, metabolomics, and proteomics
    - Significant challenges to animal molecular biologists/researchers
  - Systems biology approach is essential for capitalization on “omic” data
    - Breakthroughs in animal infection, nutrition, reproduction, physiology, etc.
  - ASBAMC to provide systems biology computational services
    - Make available to broader base of animal researchers
    - Designed to aid organization that lack computational resources
- Mission
  - To make systems biology computational analyses and expertise accessible to the agricultural research community in a manner that expedites the realization of breakthroughs in functional genomics and proteomics discovery
- No cost to selected biological projects
  - 8 to 10 high impact biological projects per year selected by a scientific steering committee
- Seralogix is industry affiliate
  - Providing computational pipeline leveraging tools developed by NIAID and NHGRI sponsorships
- Funding by USDA National Institute of Food and Agriculture
  - Three year program

# Service Center Concept

Biological Project Clients



ASBAMC Portal

Data Submission | Results Access

Site Information & "Omic" Data Submission Manager

Omic Data Submission Staging

Web Servers

Database Servers

Analysis, Modeling & Visualization Manager

Analysis & Modeling Results Staging

Systems Biology Computational Pipeline

Focus of this presentation



- Data Importation
- Experimental Data Mgmt
- Annotation Update
- Data Preprocessing
- Analysis and Modeling

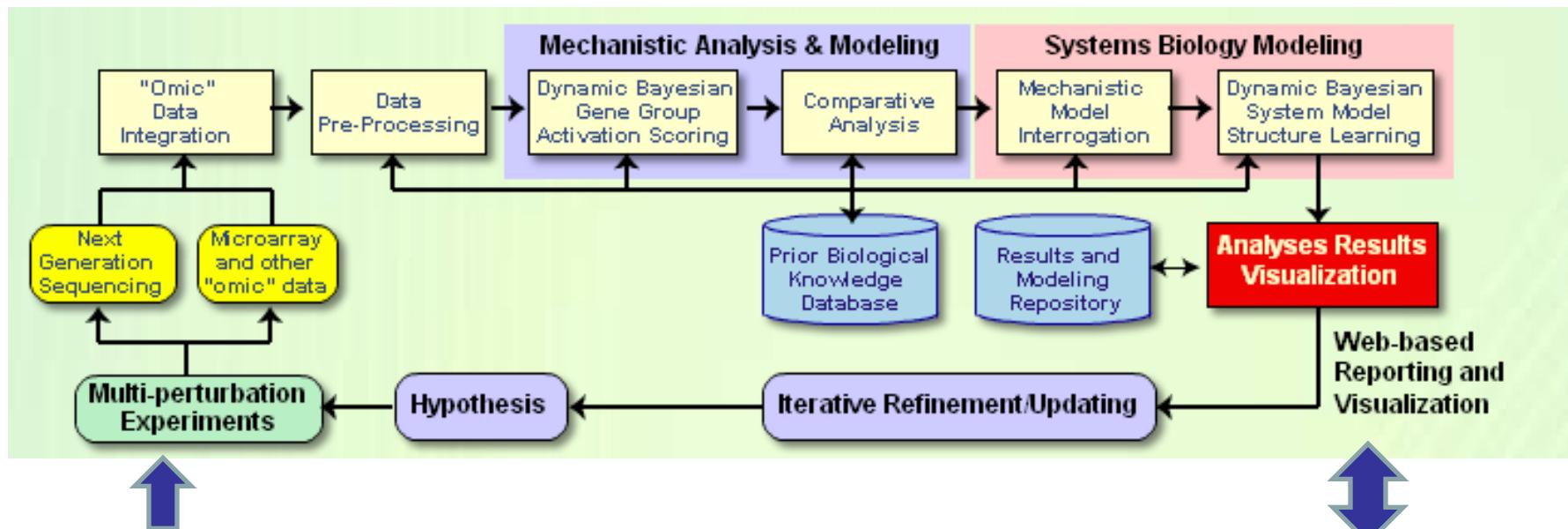
- Client Reporting Mgmt
- Knowledge Curation
- Results Preparation
- Client Training
- Client Support

ASBAMC Computational Capabilities



# *Experiment ↔ Computation Iterations* The Systems Biology Cycle

- Supports multi-conditional perturbation experiments
- “omic” data integration
- Incorporates prior biological knowledge
- Comparative analysis
  - Pathways, gene ontologies, genes
- System Modeling
  - Network learning
- Pattern discovery
- Storage of results
- Web-based report
  - Results and visualization
  - User driven interrogation of results



Clients data via ASBAMC Portal

Client accessible via ASBAMC Portal

# Web Reports by Rich Internet Application Presents Integrated Results and Visualization



## Classical Statistical Analysis



## Multi-conditional Comparative Analysis



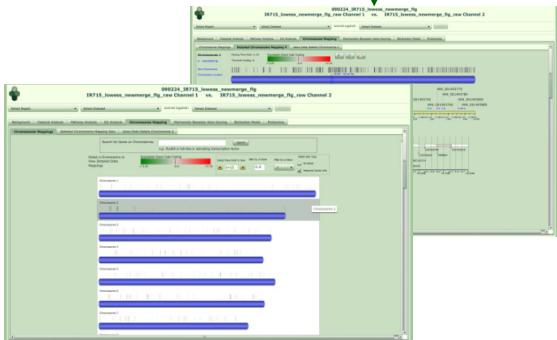
## Dynamic Visualization



## Proteomics



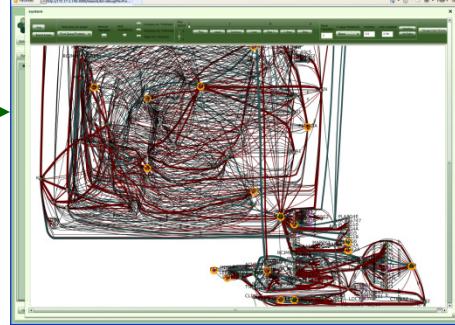
## Visualization & Reporting System



## Gene Ontology DBGGA Analysis



## Systems Biology Model



## Mechanistic Genes



## Chromosome Mapping

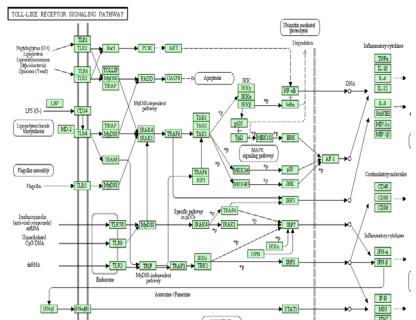


# Dynamic Bayesian Network Modeling

## Exploits the power of Dynamic Bayesian Networks (DBN)

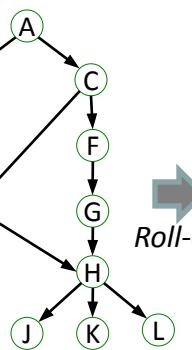
- models in terms of probabilistic relationships
- Models dynamic processes (change over time)
- Continuous variables (no discretization)
- learning (machine learning using perturbed data)
- biologically understandable networks (unlike neural nets)
- models heterogeneous data
- deals with missing/hidden data
- fusing with prior knowledge
- inference (predictions and what-ifs)
- pattern recognition

Toll-like Receptor Pathway



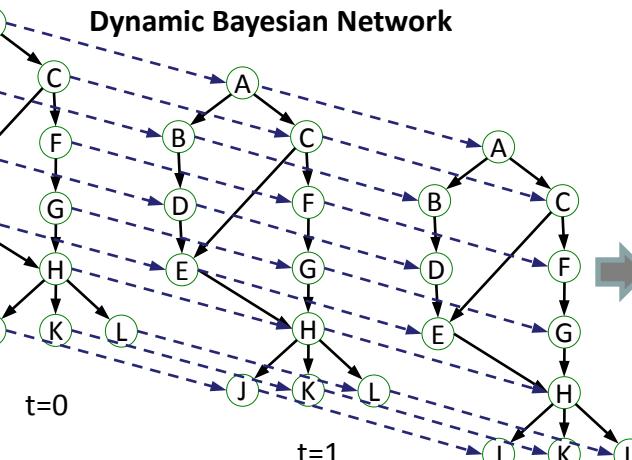
Prior Knowledge

Bayesian Network



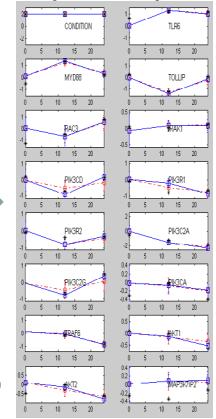
Causal Static Network Model

Dynamic Bayesian Network



Dynamic Network Model

Model  
Interrogation and  
Simulation  
(what-if)



Example Model Structure from Known Biological Knowledge



# DBN Gene Group Activation (DBGGA)

- DBGGA developed by Seralogix for multi-perturbation, time-course experimental studies
  - Comprehensive biological system analysis
    - thousands of biological processes (Gene Ontology)
    - hundreds of pathways and subnets
  - Identification of mechanistic genes/proteins
    - In context of genetic relationships
    - Sensitive to subtle changes and trends
- Systems Biology model
  - Structure learning
    - perturbed data and a priori gene-gene/protein-protein relations
  - Captures biosignatures (patterns)
  - Supports Interrogation and simulation



# Case Study: Comparative Pathogenicity

- Broiler chickens: Resistant (Line A) and susceptible (Line B)
- *Campylobacter jejuni* (*C. jejuni*) colonization in the ceca
  - wildtype (WT) and mutant (MT) strains
- Microarray gene expression data (44k Agilent chicken arrays)
  - 4 infected and non-infected (NI) biological replicates for each line
  - 4 time points post inoculation (1hr, 4hr, 24hr, 36hr)
  - Dye swap experimental design
- Comparative pathogenicity experimental design
  - Line A: WT vs NI
  - Line A: MT vs NI
  - Line A: MT vs WT
  - Line B: WT vs NI
  - Line B: MT vs NI
  - Line B: MT vs WT
  - Line A and B: A-NI vs B-NI



# *Analysis and Modeling Conducted*

- For each comparative condition
  - Classical statistical analysis (z-score with variance Bayesian estimator and Fold)
  - DBGGA on 141 signaling and metabolic pathways
  - DBGGA on 3291 gene ontology groups (GO)
  - Individual gene scores from DBGGA models on pathways and GO groups
  - Chromosome position mapping of gene scores (observed)
  - Systems biology model learned from data and prior biological knowledge
  - Comprehensive web report
- Pathogenicity comparative analyses
  - Pathways
  - Gene ontology
  - Systems model



# Line A MT vs WT Analysis and Modeling Walk-through

## Summary

		<b>1 hr</b>	<b>4 hrs</b>	<b>12 hrs</b>	<b>36 hrs</b>
<b>Gene Regulation</b>	<b>up</b>	<b>763</b>	<b>537</b>	<b>60</b>	<b>7</b>
	<b>down</b>	<b>291</b>	<b>83</b>	<b>6</b>	<b>0</b>
<b>Pathways</b>	<b>Activated*</b>	<b>44</b>	<b>18</b>	<b>0</b>	<b>0</b>
	<b>Suppressed</b>				
	*	<b>1</b>	<b>0</b>	<b>0</b>	<b>0</b>
<b>Biological Processes</b>	<b>Activated</b>	<b>850</b>	<b>248</b>	<b>0</b>	<b>21</b>
	<b>Suppressed</b>	<b>7</b>	<b>9</b>	<b>0</b>	<b>0</b>
<b>Cellular Component</b>	<b>Activated</b>	<b>147</b>	<b>27</b>	<b>0</b>	<b>1</b>
	<b>Suppressed</b>	<b>1</b>	<b>2</b>	<b>0</b>	<b>0</b>
<b>Molecular Function</b>	<b>Activated</b>	<b>216</b>	<b>57</b>	<b>0</b>	<b>3</b>
	<b>Suppressed</b>	<b>3</b>	<b>0</b>	<b>0</b>	<b>0</b>

\*Activated: majority of gene scores are up regulated

\*Suppressed: majority of gene scores are down regulated

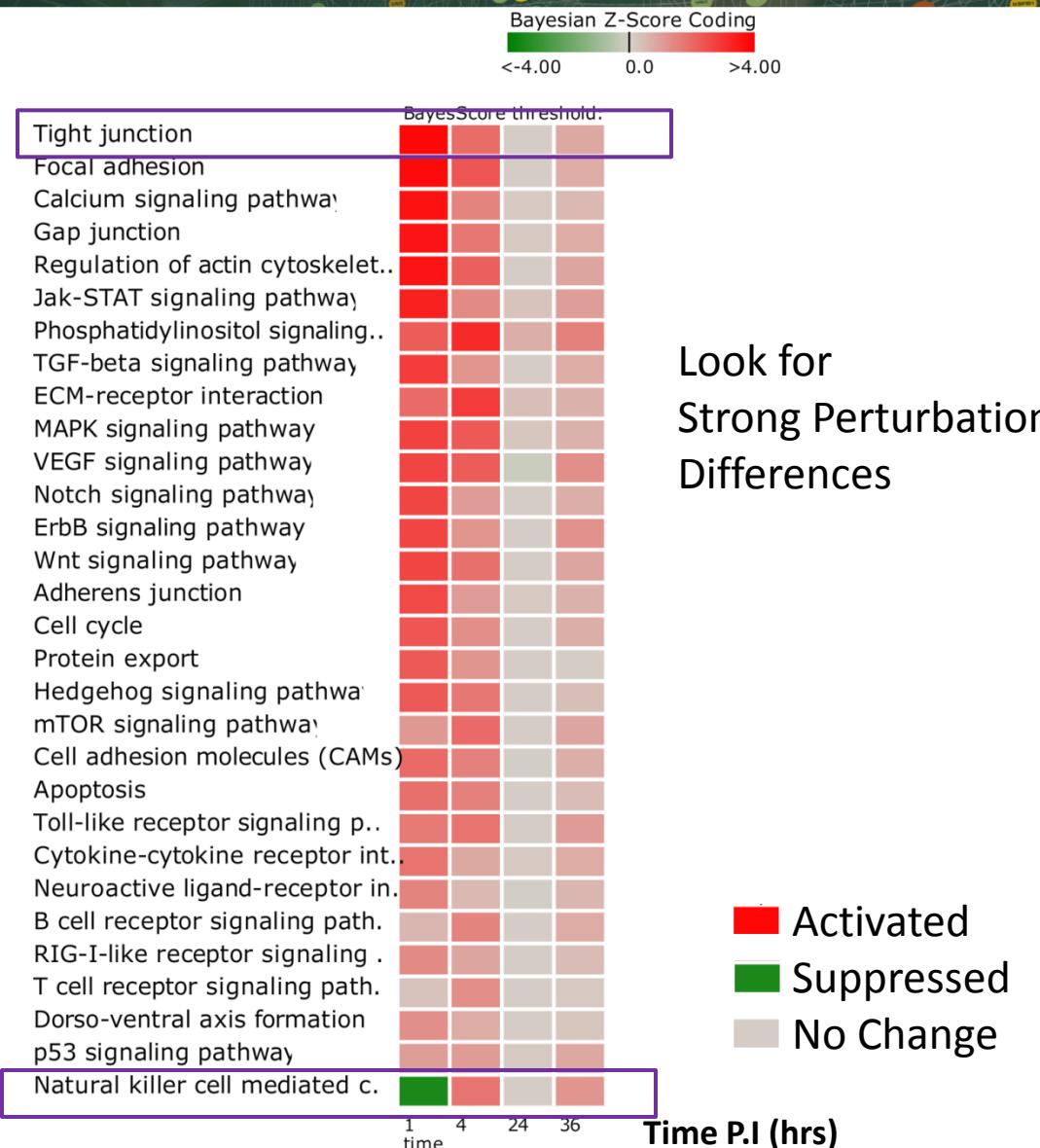


# DBGGA Pathway Scores Line A MT vs WT

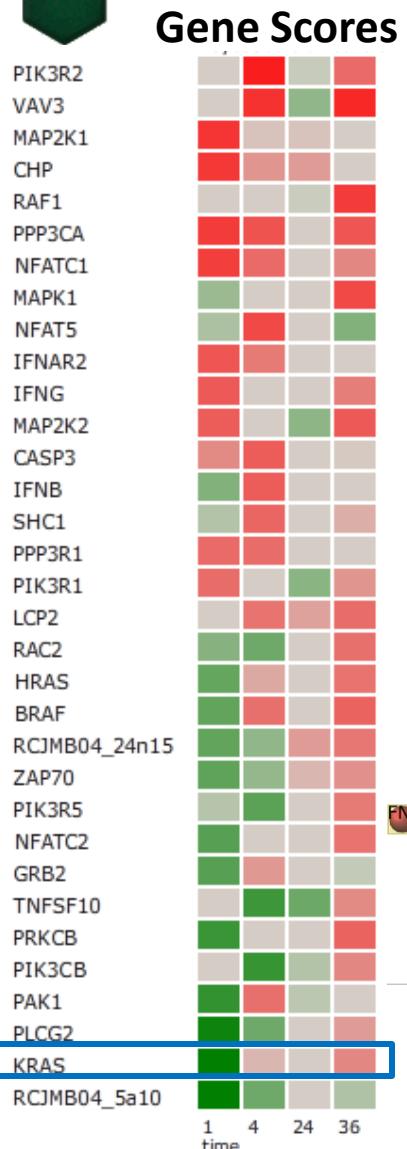
Subset of Pathways In These Categories

- Cell Communication
- Signaling Molecules & Interaction
- Signal Transduction
- Cell Growth and Death
- Membrane Transport
- Cell Motility
- Immune System
- Development

DBGGA Scores are transformation of log likelihood to a Bayesian z-score



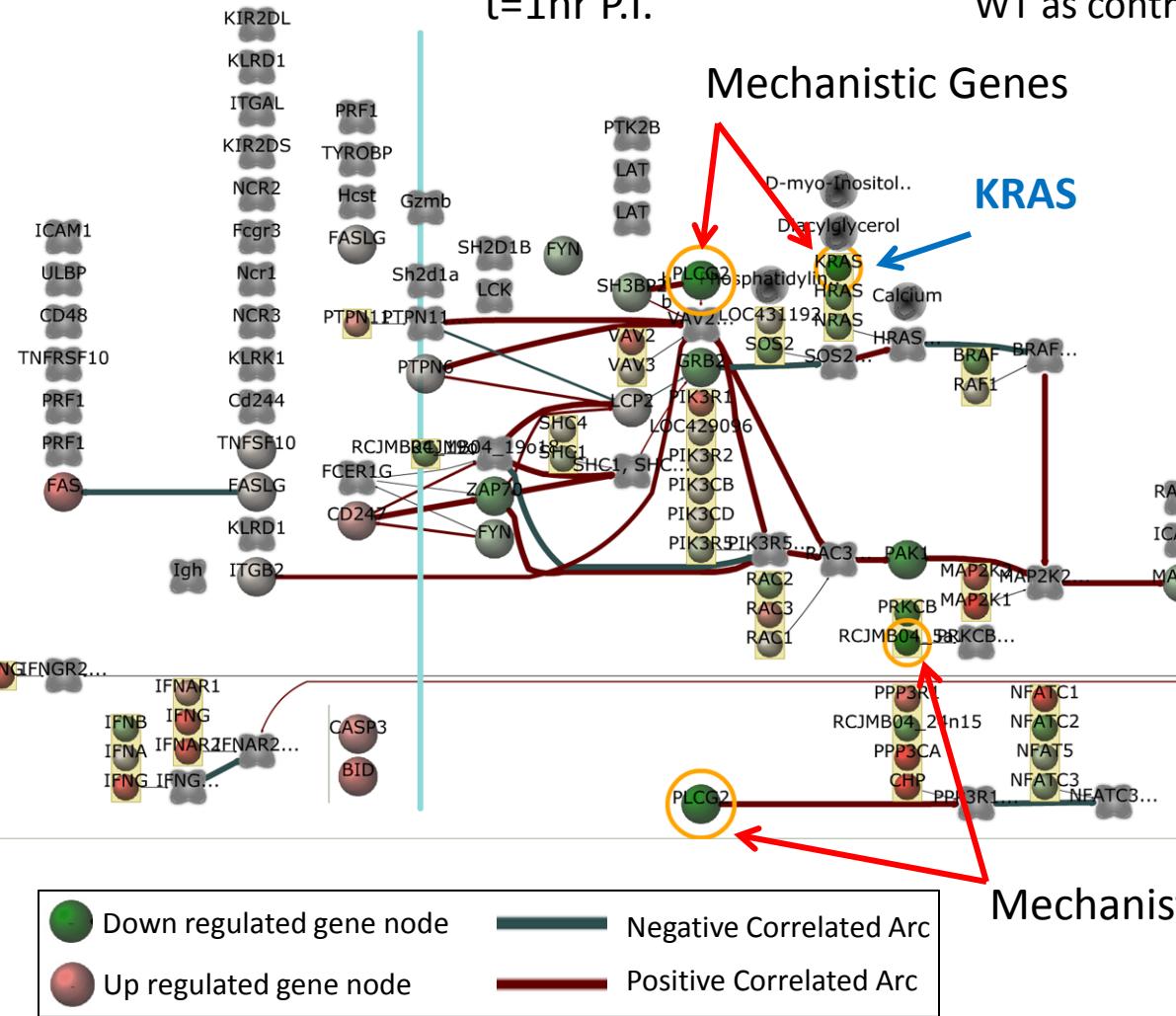
# Natural Killer Cell Mediated Cytotoxicity DBGGA Gene Scores and Visualization Model



MT as experiment  
WT as control



KRAS



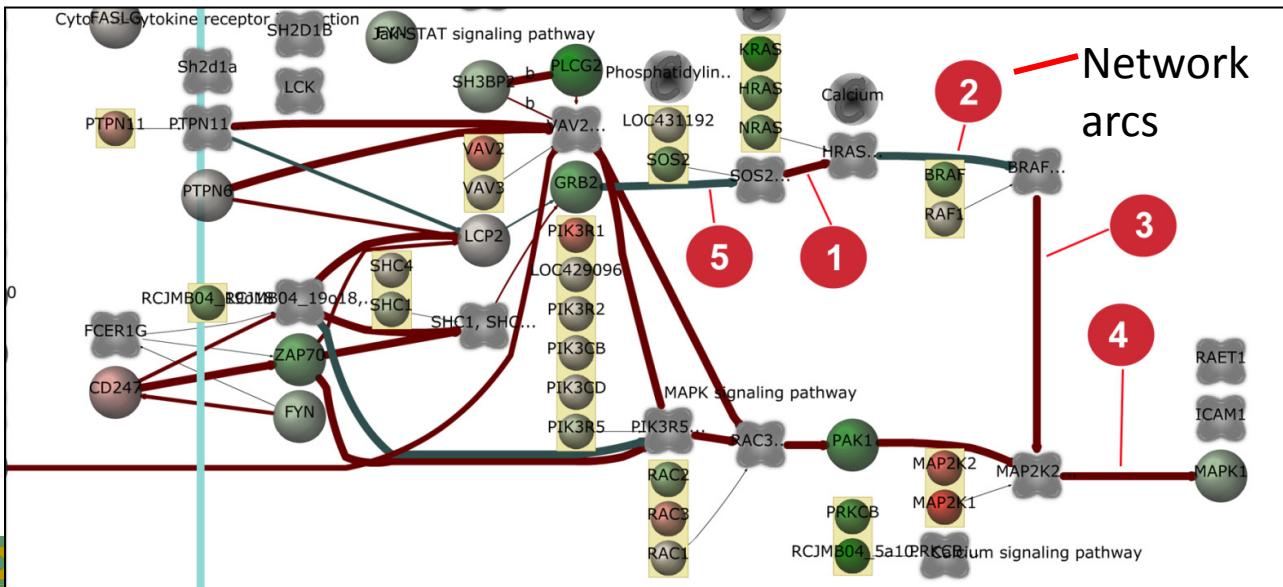
# Mechanistic Genes

# Gene Relationships: Model Interrogation

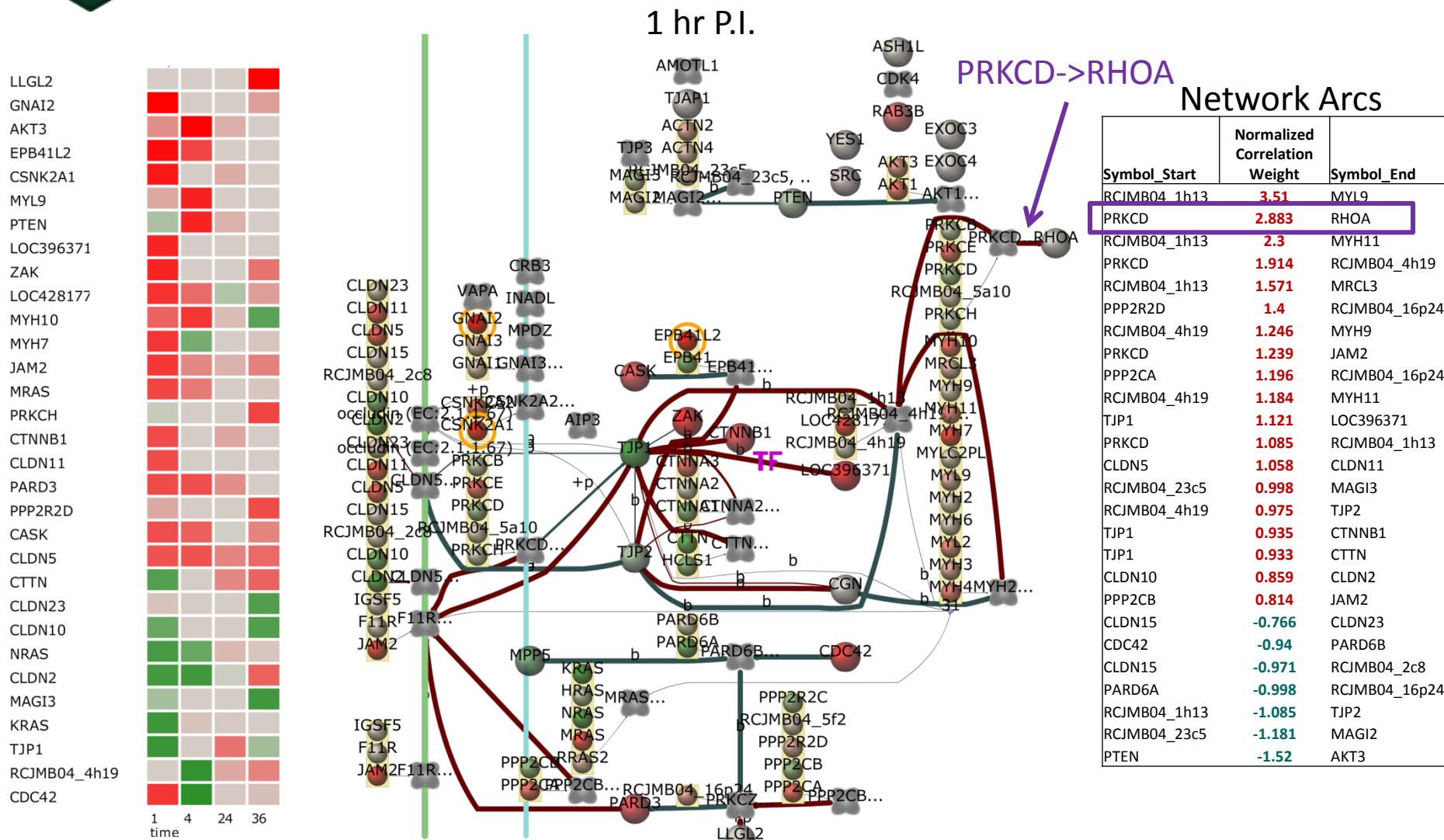
Table of Correlated Arc Gene-Gene Relationships (partial list)

Network Arc #	Symbol Start	Description Start Gene	Normalized Correlation Weight	Symbol End	Description End Gene	Function
1	SOS2	Son of sevenless homolog 2 (Drosophila)	1.895	HRAS	V-Ha-ras Harvey rat sarcoma viral oncogene homolog	activation
1	SOS2	Son of sevenless homolog 2 (Drosophila)	0.804	KRAS	V-Ki-ras2 Kirsten rat sarcoma viral oncogene homo.	activation
2	HRAS	V-Ha-ras Harvey rat sarcoma viral oncogene homolog	0.731	RAF1	V-raf-1 murine leukemia viral oncogene homolog 1	activation
2	NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog	0.573	RAF1	V-raf-1 murine leukemia viral oncogene homolog 1	activation
3	RAF1	V-raf-1 murine leukemia viral oncogene homolog 1	0.501	MAP2K2	Mitogen-activated protein kinase kinase 2	activation
1	SOS2	Son of sevenless homolog 2 (Drosophila)	0.489	NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog	activation
4	MAP2K1	Mitogen-activated protein kinase kinase 1	0.322	MAPK1	Mitogen-activated protein kinase 1	activation
5	GRB2	Growth factor receptor-bound protein 2	-0.245	SOS2	Son of sevenless homolog 2 (Drosophila)	activation
2	KRAS	V-Ki-ras2 Kirsten rat sarcoma viral oncogene homo.	-0.262	BRAF	V-raf murine sarcoma viral oncogene homolog B1	activation
2	NRAS	Neuroblastoma RAS viral (v-ras) oncogene homolog	-0.446	BRAF	V-raf murine sarcoma viral oncogene homolog B1	activation
5	GRB2	Growth factor receptor-bound protein 2	-0.476	LOC431192	Similar to alternate SOS1	activation
2	HRAS	V-Ha-ras Harvey rat sarcoma viral oncogene homolog	-0.765	BRAF	V-raf murine sarcoma viral oncogene homolog B1	activation

Zoom in on  
subnetwork  
Of Natural Killer  
Cell Mediated  
Cytotoxicity  
Pathway



# Tight Junction Pathway

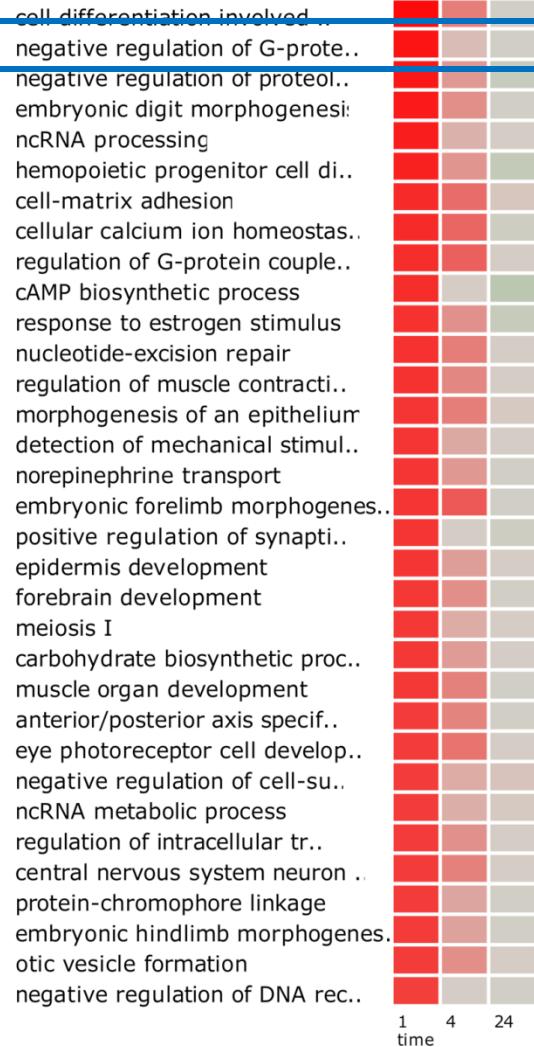




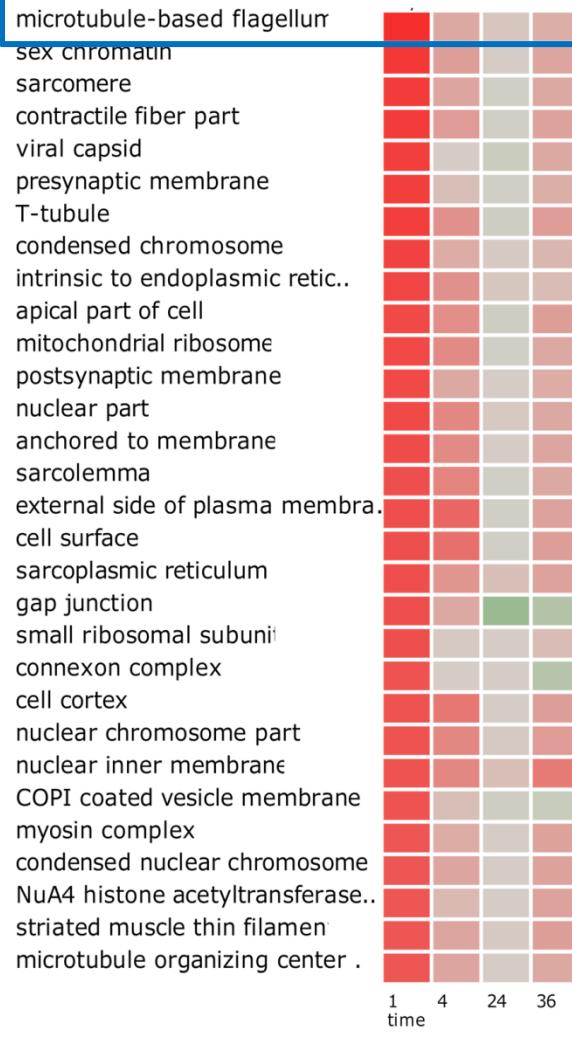
# Gene Ontology (GO) DBGGA Analysis (Line A MT vs WT)

## Selection of top scoring (perturbed GO groups)

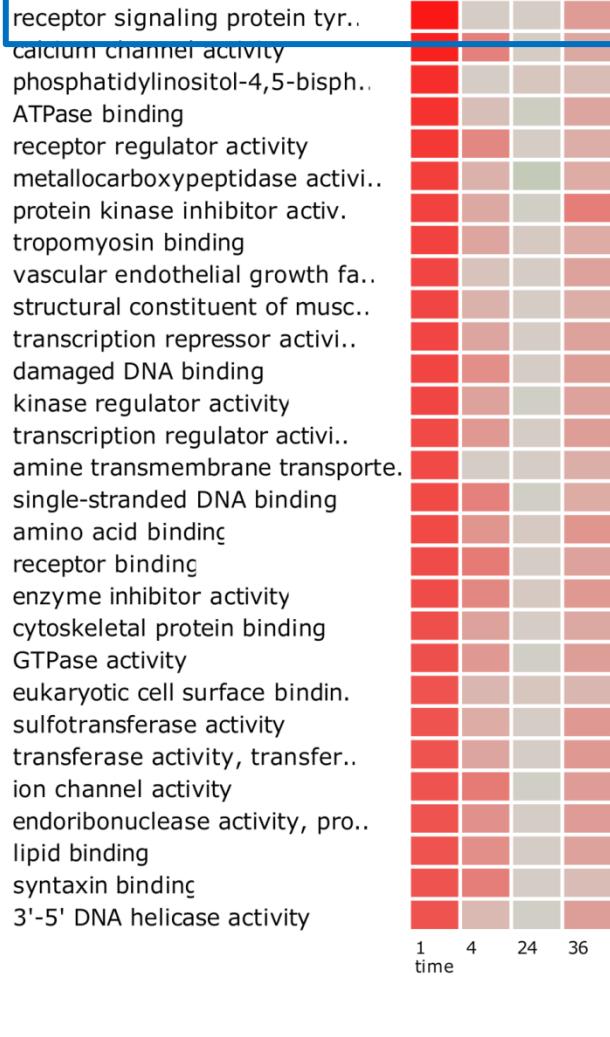
### Biological Process



### Cellular Component



### Molecular Function

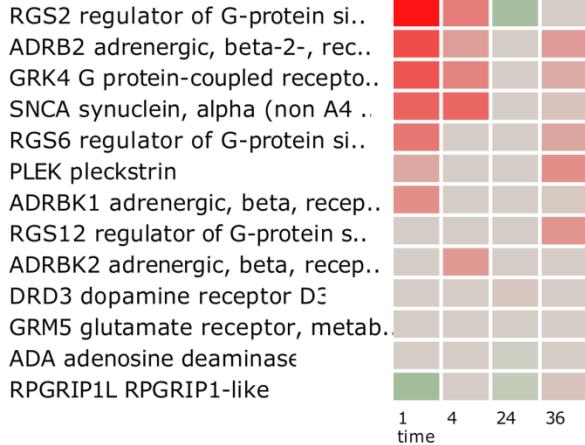


1  
time  
4  
24  
36

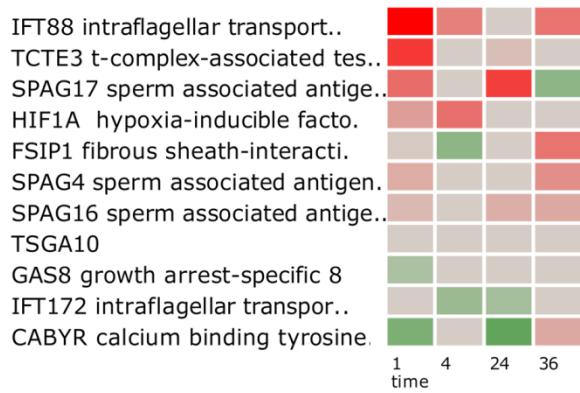
# Gene DBGGA score Heatmaps of Selected GO groups



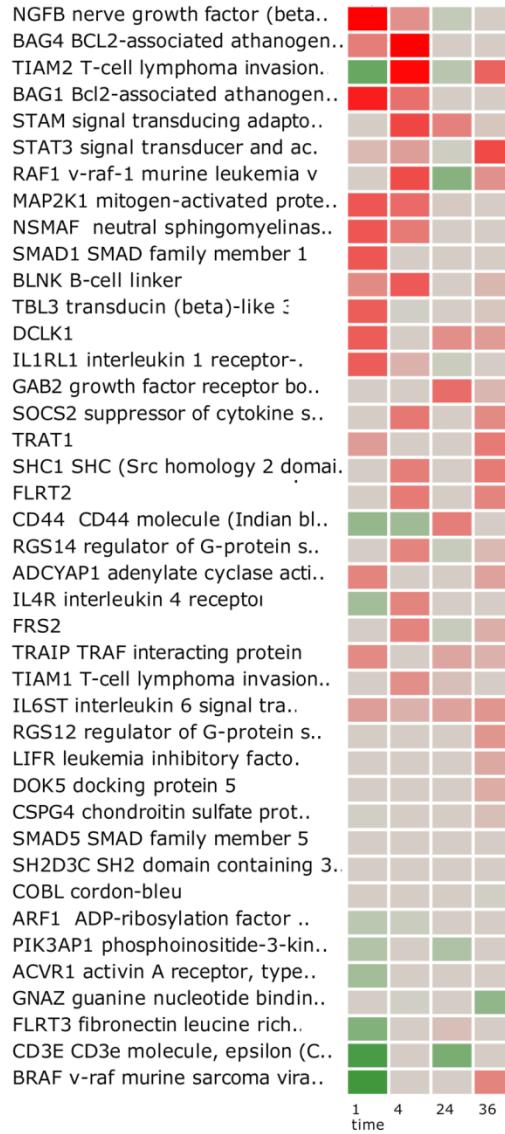
## negative regulation of G-protein coupled receptor protein signaling pathway



## microtubule-based flagellum



## Receptor signaling protein activity





# Comparative Pathogenicity Analysis Example

- Comparative pathogenicity study
  - Line A: WT vs NI
  - Line A: MT vs NI
  - Line A: MT vs WT
  - Line B: WT vs NI
  - Line B: MT vs NI
  - Line B: MT vs WT
  - Line A and B: A-NI vs B-NI
- Pathogenicity comparative analyses
  - Pathways
  - Gene ontology
  - Systems model
    - Candidate points of pathogen interaction

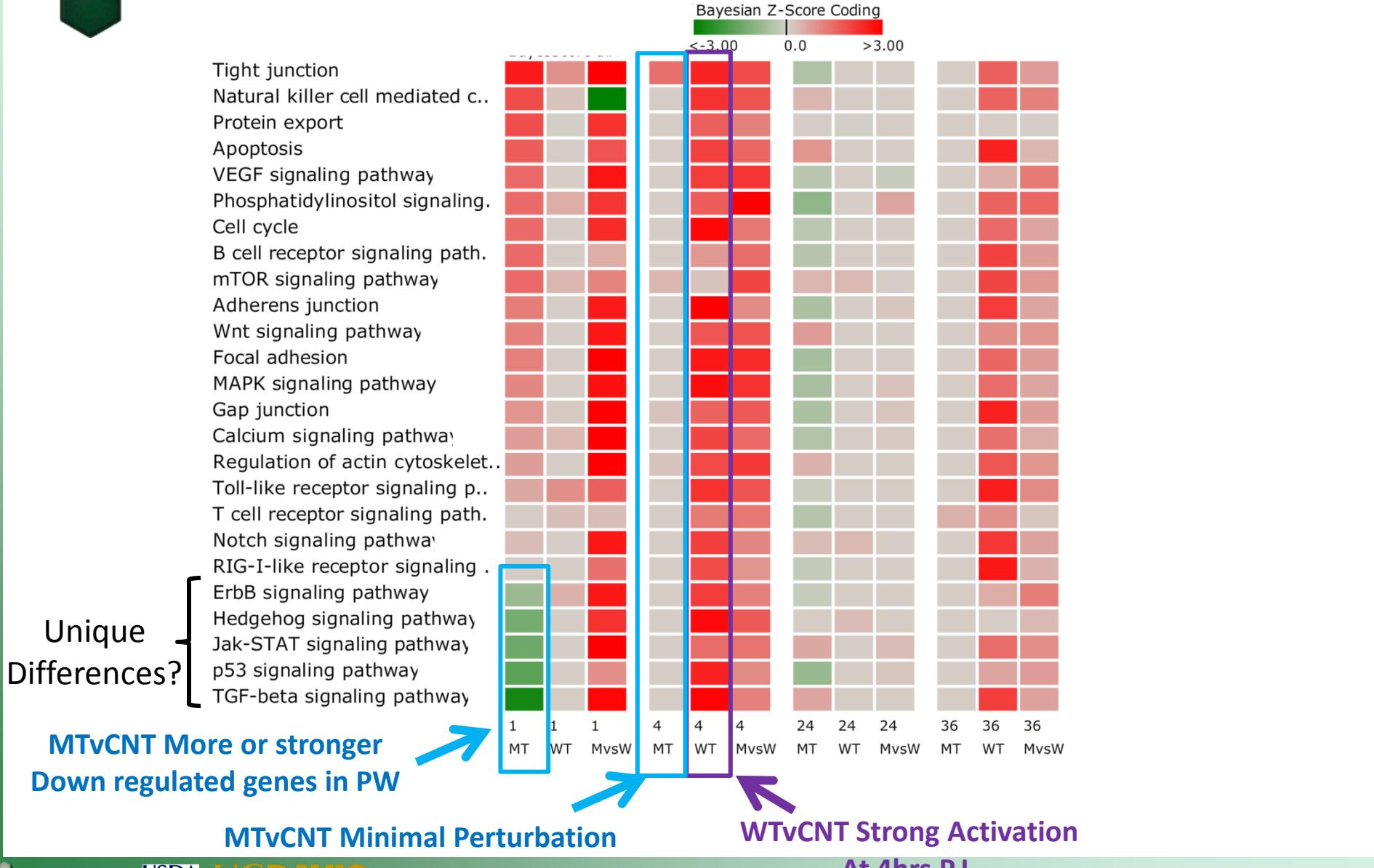


# Classical Z-score Summary Comparative Profiles

$|Z\text{-score threshold}| > 2.24$

				time P.I. (hrs)			
	<u>Condition</u>	<u>regulation</u>	<u>t=1</u>	<u>t=4</u>	<u>t=24</u>	<u>t=36</u>	
<b>Resistant</b>	Line A MT vs CNT	up	189	10	942	204	Lesser Perturbed Response Later
		down	65	75	81	65	
	Line A WT vs CNT	up	75	88	380	338	More Perturbed Response And earlier
		down	40	34	0	83	
<b>Susceptible</b>	Line A MT vs WT	up	763	537	60	7	More Perturbed Response And earlier
		down	291	83	6	0	
	Line B MT vs CNT	up	3559	541	334	235	More Perturbed Response And earlier
		down	897	0	165	49	
	Line B WT vs CNT	up	22	837	334	825	More Perturbed Response And earlier
		down	3	3	7	502	
	Line B MT vs WT	up	473	94	30	96	More Perturbed Response And earlier
		down	79	2	25	0	

# Line A Signaling Pathway Comparative Modeling

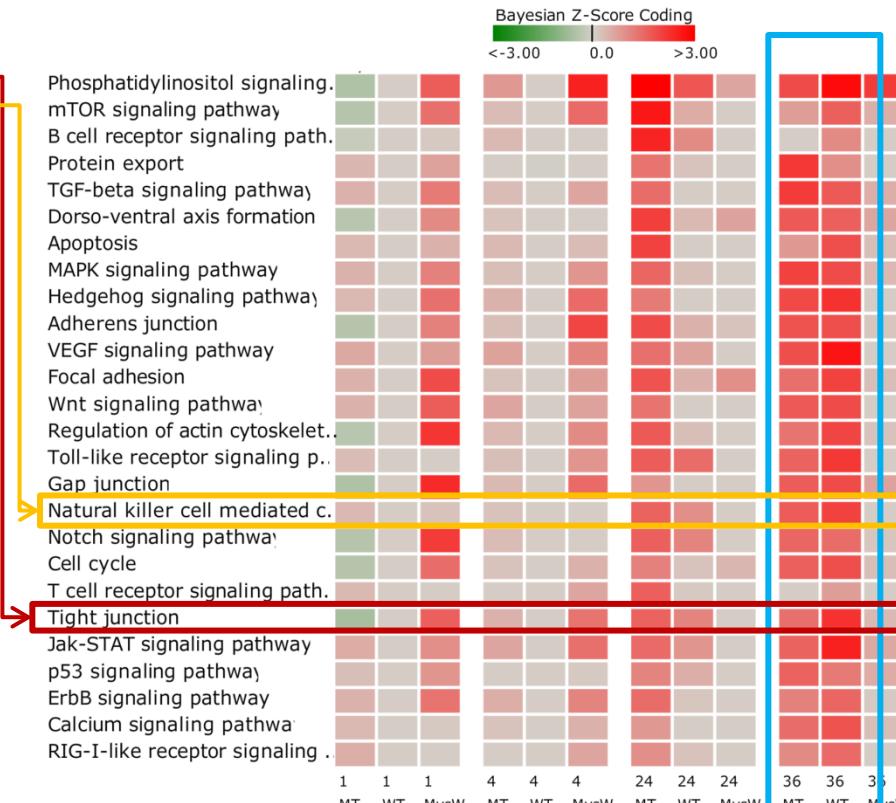


# Line A contrast to Line B Signaling Pathways

Line A(MT, WT, MvsW)

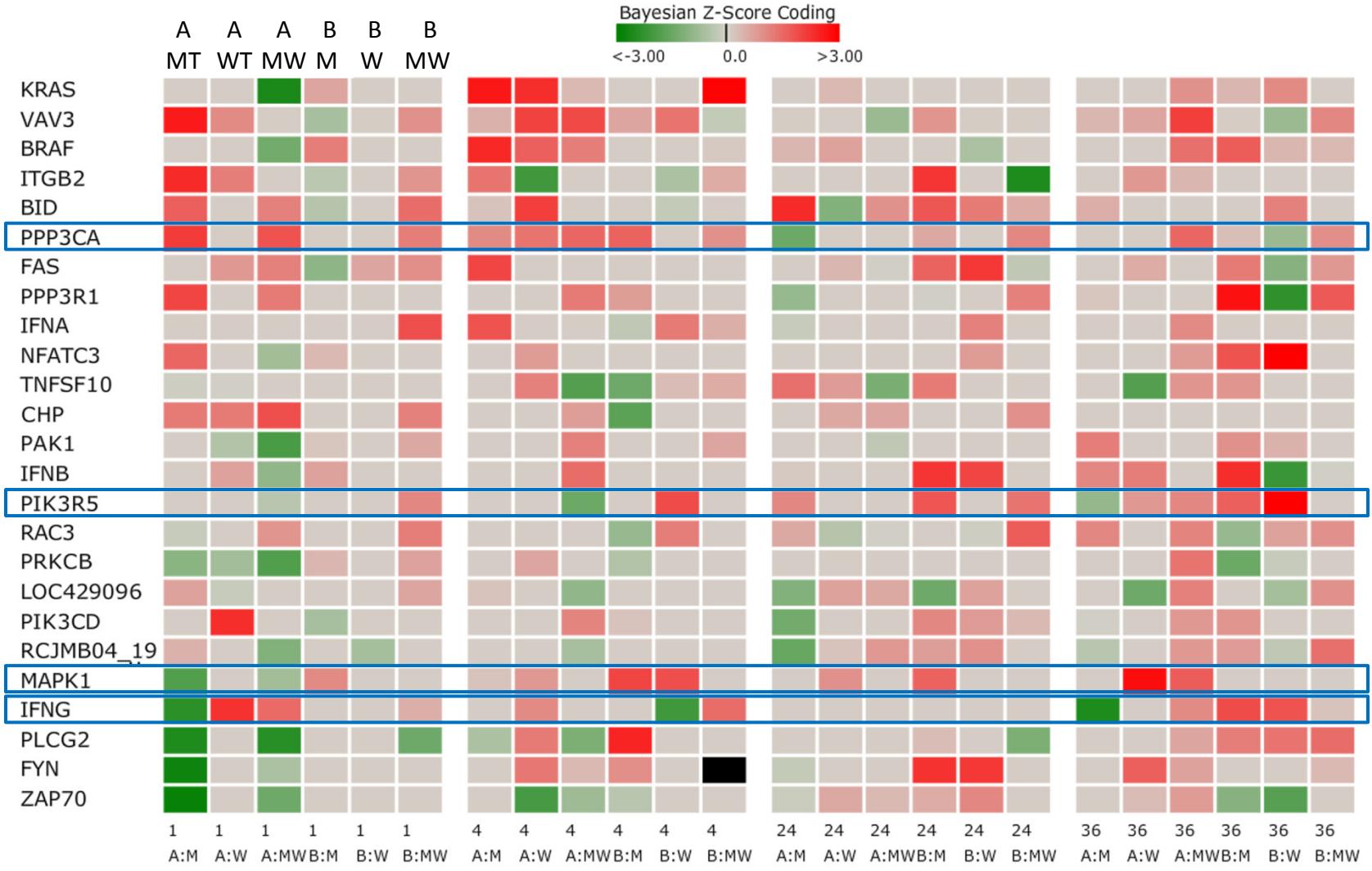


Line B (MT, WT, MvsW)





# Line A and B Natural Killer Cell Pathway





# Line A Metabolic Pathway Comparative Modeling

Unique  
Differences? {

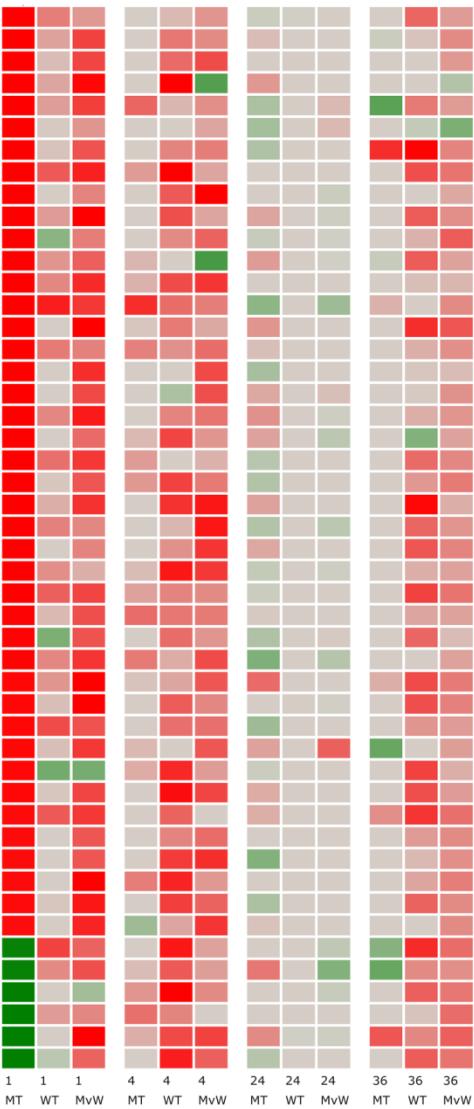


# Line A and B GO Comparative Modeling



cellular response to lipopolys..  
 molting cycle  
 regulation of epidermis develo..  
 response to pain  
 response to hexose stimulus  
 asymmetric cell division  
 SMAD protein signal transducti..  
 catecholamine biosynthetic pro..  
 mammary gland morphogenes..  
 regionalization  
 hair follicle development  
 response to starvation  
 dorsal/ventral axis specifcat.  
 nephron epithelium development  
 somitogenesis  
 bone morphogenesis  
 purine ribonucleotide biosynth..  
 substrate-bound cell migrator  
 neurotransmitter transport  
 forebrain regionalization  
 oxygen and reactive oxygen spe..  
 epithelial to mesenchymal tran..  
 positive regulation of osteobl..  
 response to acid  
 mesoderm morphogenesis  
 positive regulation of mesench..  
 ribonucleoside metabolic proce..  
 neurotransmitter uptake  
 ATP synthesis coupled electron..  
 epithelial tube morphogenesis  
 positive regulation of insulin..  
 hindbrain development  
 purine ribonucleoside monophos..  
 positive regulation of ion tra..  
 peptidyl-lysine modification  
 limb morphogenesis  
 negative regulation of cytokin..  
 regulation of MAPKKK cascade  
 regulation of RNA splicing  
 T cell apoptosis  
 anterior/posterior pattern for..  
 vascular endothelial growth fa..  
 interleukin-6 biosynthetic pro..  
 neutral amino acid transpor..  
 SMAD protein nuclear transloca..  
 regulation of epithelial to me..  
 calcium-independent cell-cell ..  
 transmembrane receptor protein.

**Line A**



**Top scoring Biological Processes**

calcium-independent cell-cell ..  
 artery morphogenesis  
 polyol biosynthetic process  
 glycine metabolic process  
 negative regulation of calcium..  
 respiratory burst  
 positive regulation of glucose..  
 inorganic anion transport  
 cell projection morphogenesis  
 axon ensheathment  
 regulation of glycogen metabol..  
 bile acid metabolic process  
 respiratory system development/  
 DNA damage response, signal tr..  
 late endosome to vacuole trans..  
 somatic recombination of immun..  
 regulation of glycogen biosynt..  
 regulation of gliogenesis  
 positive regulation of intrace..  
 protein monoubiquitinatior  
 negative regulation of B cell ..  
 regulation of bone mineralizat..  
 neuron projection regeneration  
 lung alveolus development  
 somatic stem cell maintenanc..  
 myeloid cell apoptosis  
 positive regulation of amine t..  
 fibrinolysis  
 activin receptor signaling pat..  
 flagellum assembly  
 inner cell mass cell prolifera..  
 negative regulation of transcr..  
 negative regulation of cytokin..  
 proteoglycan metabolic process  
 superoxide metabolic process  
 positive regulation of DNA rep..  
 regulation of cellular carbohy..  
 mesenchymal cell differentiati..  
 establishment of spindle local..  
 epidermal cell differentiatior..  
 regulation of Wnt receptor sig..  
 dopamine metabolic process  
 proteoglycan biosynthetic proc..  
 regulation of smooth muscle co..  
 regulation of cell fate commit..  
 response to tumor necrosis fac..  
 vesicle organization  
 cilium assemby

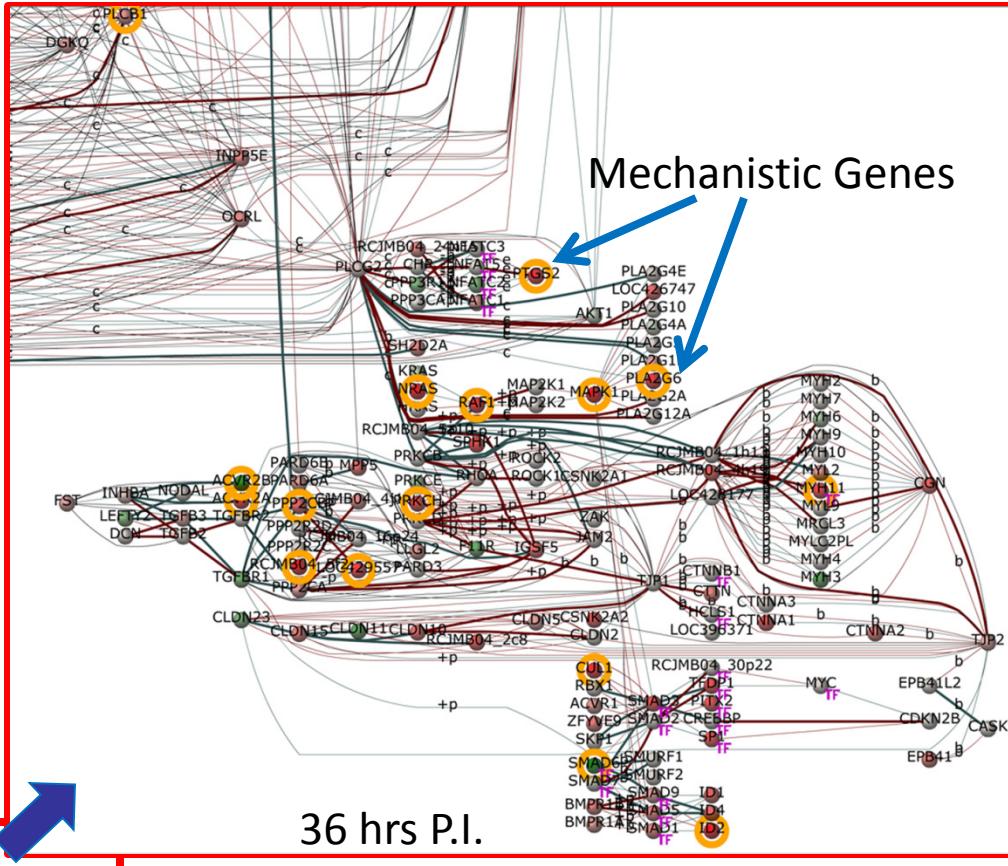
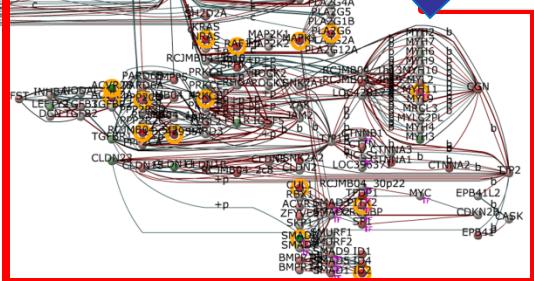
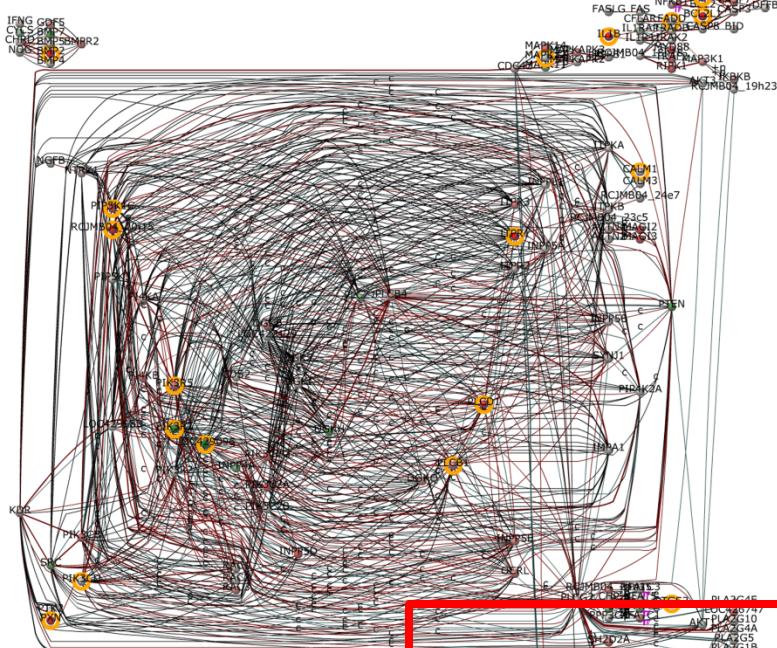
**Line B**



# Line A - WTvsCNT System Model(SM) Interrogation

## Seed pathways:

- TGF-beta signaling pathway
- Tight Junction
- Apoptosis
- VEGF signaling pathway
- Phosphatidylinositol signaling system



Only 233 genes, 1209 arcs



# System Model (SM) Interrogation (drilling down)

Comparing MT-SM with WT-SM and using info from MTvsWT-SM

Mechanistic Genes  
from MTvsWT-SM

Gene Arc Pair	Normalized Weight MT	Normalized Weight WT	Start Gene Mechanistic	End Gene Mechanistic	Gene Arc Pair	Normalized Weight MT	Normalized Weight WT	Start Gene Mechanistic	End Gene Mechanistic
RCJMB04_1h13->MRCL3	<b>0.359</b>	-0.146			PLCG2->PLA2G4A	<b>0.284</b>	-0.021		
PPP3R1->NFATC3	<b>0.341</b>	-0.008			RCJMB04_5a10->RCJMB04_4h19	-0.361	0.065		
CUL1->SMAD2	<b>0.318</b>	-0.14		✓	IL1R1->MYD88	-0.371	-0.074		
NFKB1->BIRC2	<b>0.306</b>	-0.026			CLDN15->CLDN11	-0.473	-0.078		
PLCG2->PLA2G4A	<b>0.284</b>	-0.021	✓		PLCZ1->PLCD1	-0.513	-0.162		
CLDN23->RCJMB04_2c8	<b>0.271</b>	-0.089			INHBA->ACVR2B	-0.517	-0.265		
CLDN15->CLDN5	<b>0.213</b>	-0.1		✓	TGFBR1->PPP2CA	-0.528	-0.162		
PRKCD->IGSF5	<b>0.188</b>	-0.071		✓	SMAD7->SMURF1	-0.589	-0.117		
KDR->PIK3R2	<b>0.132</b>	-0.024	✓		RBX1->SMAD3	-0.606	-0.279		
PLCZ1->INPP5B	-0.124	<b>0.006</b>	✓		RCJMB04_23c5->MAGI2	-0.615	-0.288		
TJP1->CTNNA3	-0.158	<b>0.054</b>			RCJMB04_5a10->SPHK1	-0.654	-0.12		
SMAD9->ID1	-0.239	<b>0.007</b>			SKP1->SMAD2	-0.695	<b>0.06</b>		
RCJMB04_5a10->RCJMB04_4h19	-0.361	<b>0.065</b>		✓	TGFBR1->RHOA	-0.696	-0.229		
SKP1->SMAD2	-0.695	<b>0.06</b>	✓		TGFBR1->PPP2CB	-0.699	-0.215		
PRKCD->RHOA	<b>1.98</b>	<b>1.071</b>			SKP1->SMAD3	-0.704	-0.259		
RCJMB04_1h13->MYL9	<b>1.357</b>	<b>0.585</b>			PLCG2->PLA2G4E	-0.706	-0.297		
BMPR1B->SMAD9	<b>1.335</b>	<b>1.026</b>			PARD6A->RCJMB04_16p24	-0.737	-0.432		
INHBA->TGFB2	<b>1.324</b>	<b>1.022</b>			RCJMB04_5a10->RCJMB04_1h13	-0.741	-0.48		
RCJMB04_18h18->TRADD	<b>1.271</b>	<b>0.738</b>		✓	PLCG2->RCJMB04_5a10	-0.78	-0.1		
RCJMB04_20j15->PI4KA	<b>1.202</b>	<b>0.921</b>	✓		PRKCB->RHOA	-0.794	-0.312		
BIRC2->CASP7	<b>1.095</b>	<b>0.131</b>			PLCG2->PLA2G1B	-0.816	-0.026		
PIK3CB->RCJMB04_20j15	<b>0.999</b>	<b>0.67</b>			ITPKA->RCJMB04_24e7	-0.862	-0.397		
PLCG2->PLA2G2A	<b>0.755</b>	<b>0.358</b>			PRKCB->SPHK1	-0.876	-0.609		
TRADD->CASP8	<b>0.642</b>	<b>0.344</b>			SMAD6->SMAD9	-0.911	-0.6		
SRC->PIK3CB	<b>0.507</b>	<b>0.175</b>			RCJMB04_18h18->TRAF2	-1.106	-0.64		
BIRC2->CASP3	<b>0.423</b>	<b>0.071</b>			EPB41L2->CASK	-1.112	-0.506		
IL1R1->TRADD	<b>0.388</b>	<b>0.115</b>			RCJMB04_5a10->RHOA	-1.268	-0.224		
PPP3R1->NFATC3	<b>0.341</b>	-0.008			SMAD9->ID4	-1.354	-0.823		
NFKB1->BIRC2	<b>0.306</b>	-0.026	✓		SMAD6->SMURF1	-1.398	-0.855		
PPP2CB->RCJMB04_16p24	<b>0.305</b>	<b>0.019</b>							



# Concluding Remarks

- ASBAMC can produce this and even more results
- Analyzing such results manually is not practical
  - NEED Integrated pipeline, Data management, processing, analysis, models, visualization
- Deciphering/interpreting the results takes time
  - Learning curve
  - Requires good techniques of visualization and ease of navigation
  - ASBAMC will provide training and support
- ASBAMC will soon be ready to accept biological project application
  - Please leave card or email Dr. Zhou or Dr. Drake to receive application instructions when they are available (contact info on next slide)





## Other info and contacts



- DEMO of software: Wed. Jan 18 at 10:20am, California Rm.
- Poster P0894: 10:00 - 11:30am Monday
  - Computational Systems Biology Approach to Comparative Pathogenicity Transcriptome Analysis and Modeling
- Poster P0997: 3:00 - 4:30pm Monday
  - Computational Approach for Learning an In vivo Interactome Model for Bovine Salmonellosis Pathogenesis
- ASBAMC portal will be online soon! [www.asbamc.org](http://www.asbamc.org)
  - Applications process, descriptions, announcements, etc.

Questions? Please feel free to contact us:

Ken Drake  
drake@seralogix.com  
512-828-7955 ext 1

Huaijun Zhou  
hzhou@ucdavis.edu  
530-752-1034



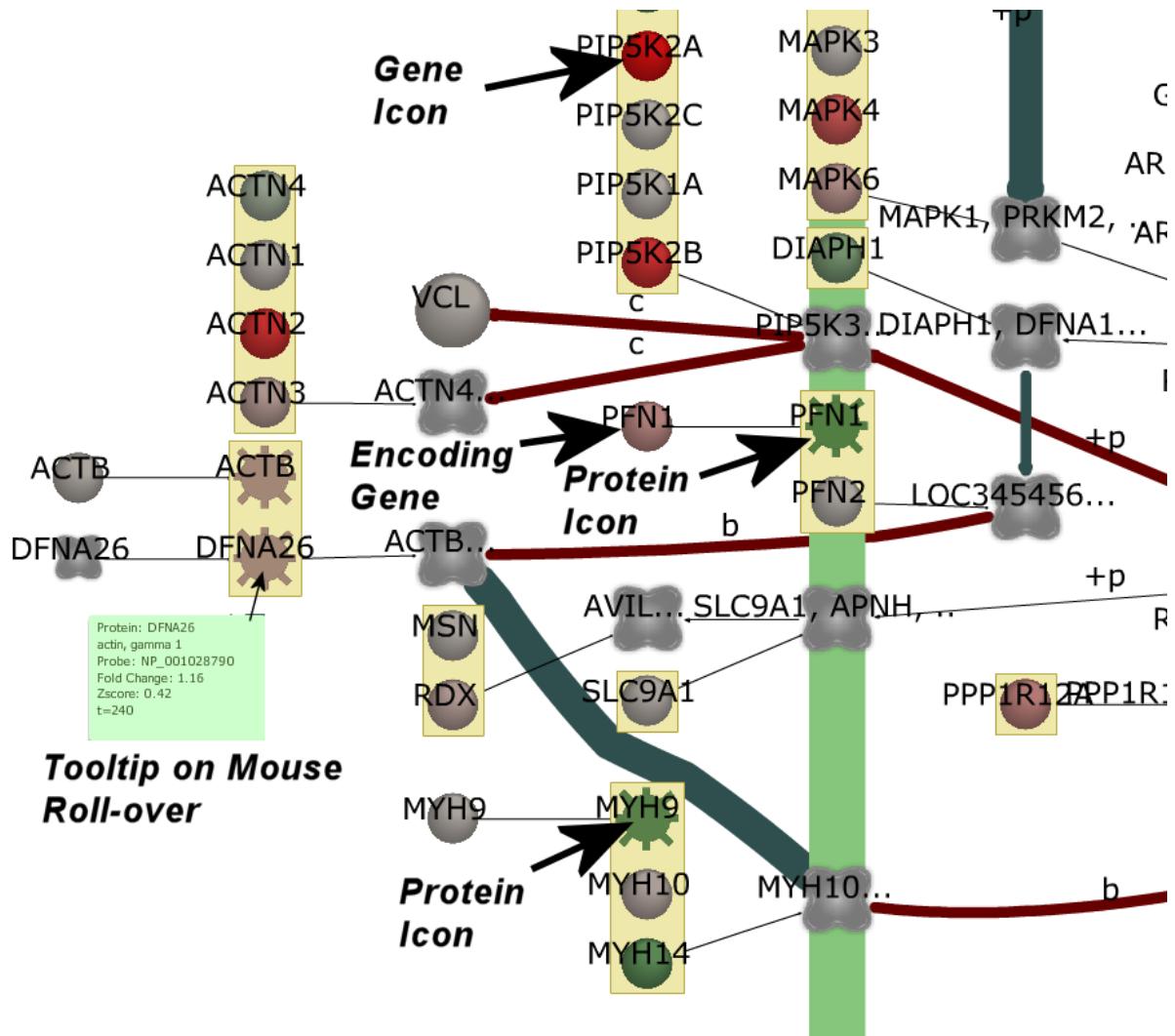
- Extra slides on interactome model not used in presentation show

# Transcriptomic and proteomic fusion

Salmonella (STM) Condition



- Gene icon
- Protein icon
- Group or not measured gene icon
- b** binding
- +p** phosphorylation



Magnification of Actin Cytoskeleton Pathway Area (prior slide)



# *Interactome Model Learned from Prior Knowledge and Gene Perturbation Expression Data from Pathogen and Host*

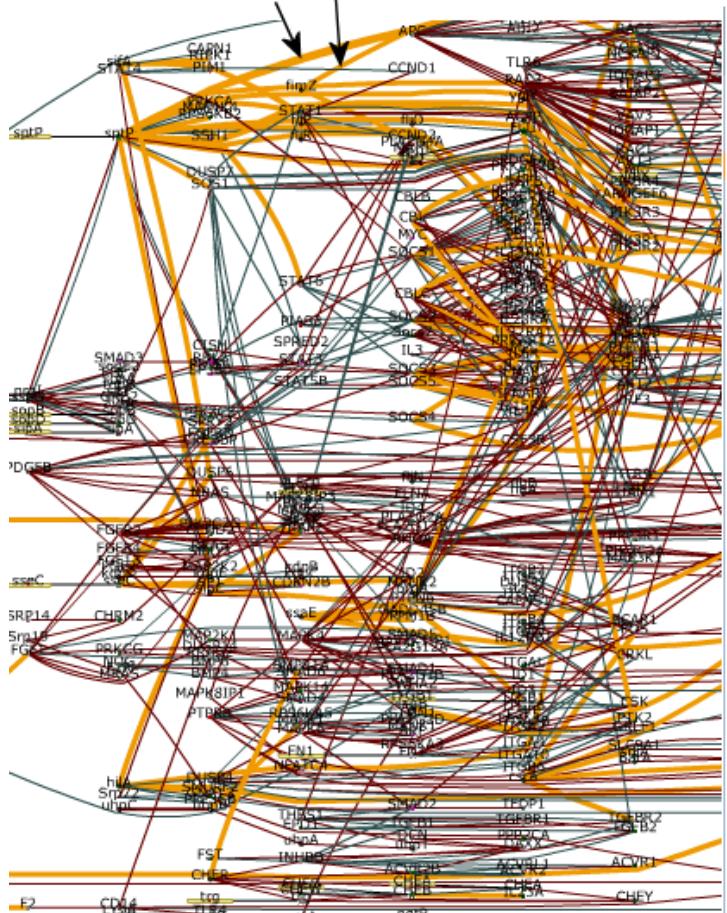


- Bayesian computational technique employs prior knowledge and gene expression data to establish probability of interaction
  - protein domain interactions
  - Sequence similarity to known interactions
  - Gene ontology association similarity
- Creates an interactome model comprised of >600 host genes
  - 9 host pathways
    - MAPK, ErbB, mTOR, Wnt, VEGF, Toll-like, GhRH
  - 6 pathogen pathways + 1 GO group
    - Protein export, Type III and IV secretion, Flagellar assembly, two component system, Bacterial chemotaxis, Pathogenesis GO

# Interactome Model Learned from Prior Knowledge and Gene Perturbation Expression Data (*Salmonella* and *Bovine* Host)

Interaction structure learning from pathogen and host gene expressions

Pathogen to Host Interactions (Orange Arcs)

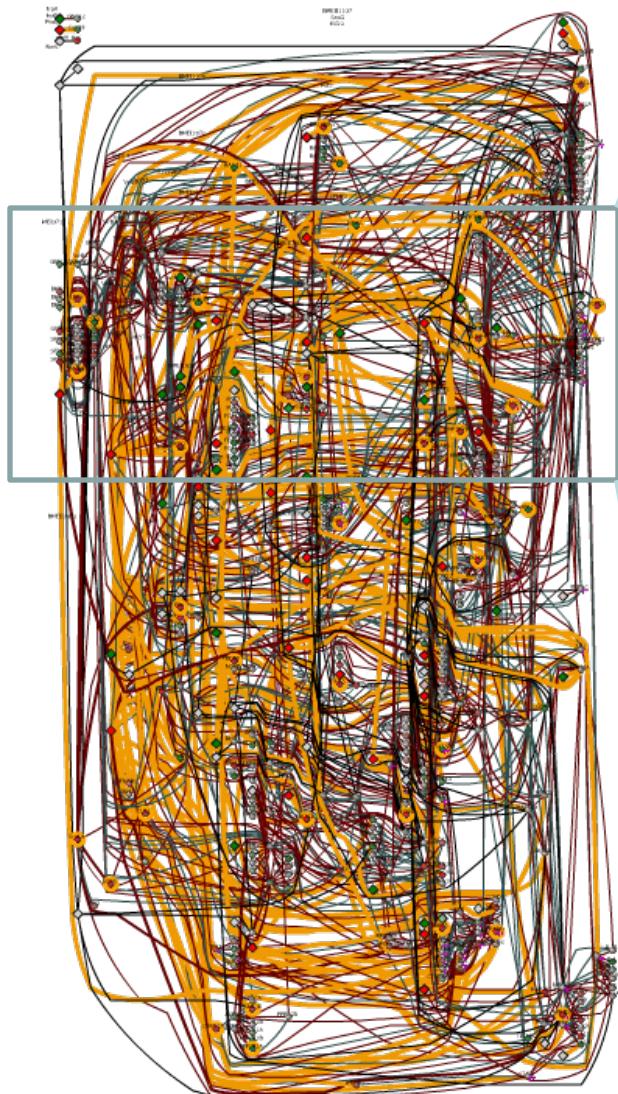


Bayesian Network Structure after Learning

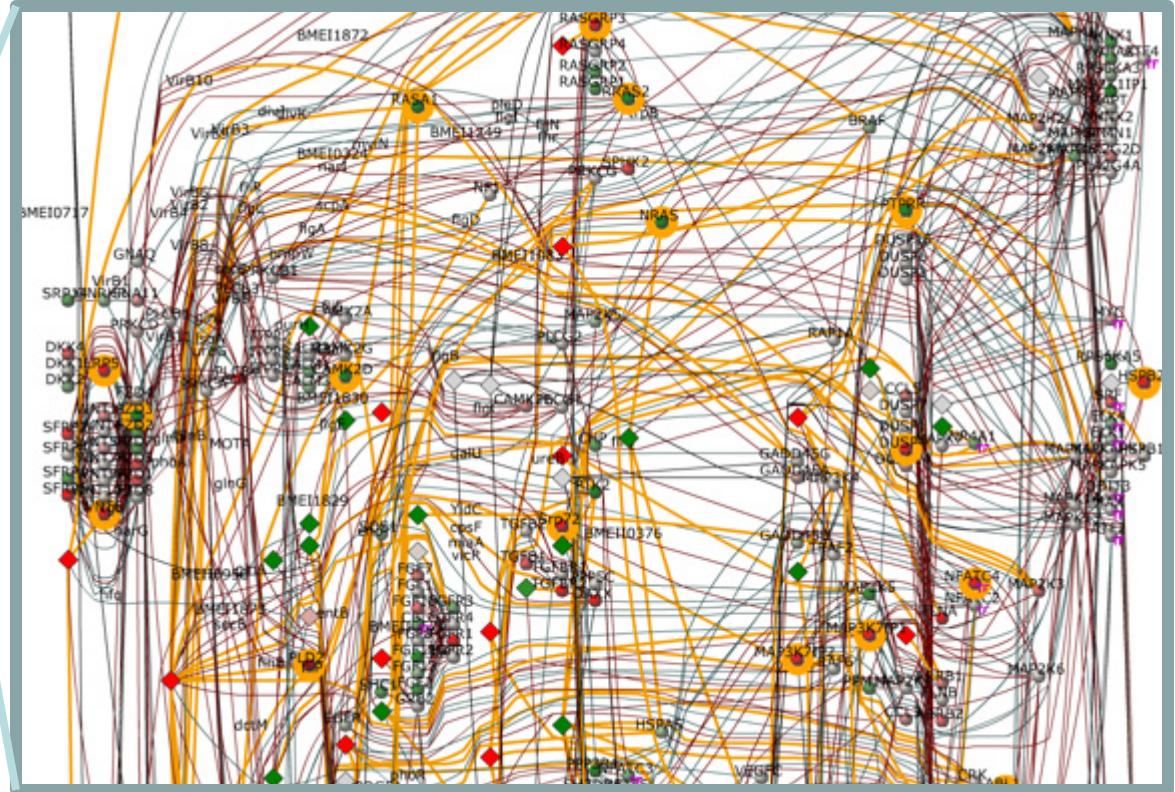
Top Pathogen-Host Protein Interaction Predictions

Pathogen Gene/Protein	Pathogen Gene/Protein Description	Correlation Score	Host Gene/Protein	Host Gene/Protein Description	Mechanistic Gene	Prediction Method
sptP_stm_prot	protein tyrosine phosphate	0.307	SSH1	slingshot homolog 1 ( <i>Drosophila</i> )	✓	PD
sptP_stm_prot	protein tyrosine phosphate	-0.306	DUSP7	dual specificity phosphatase 7	✓	PD
sptP_stm_prot	protein tyrosine phosphate	-0.201	MAP2K2	mitogen-activated protein kinase kinase 2	✓	PD
sptP_stm_prot	protein tyrosine phosphate	0.239	RPS6KB2_prot	ribosomal protein S6 kinase, 70kDa, polypeptide 2	✓	PD
sptP_stm_prot	protein tyrosine phosphate	0.348	MAPK7_prot	mitogen-activated protein kinase 7	✓	PD
sptP_stm_prot	protein tyrosine phosphate	-0.257	PAK1_prot	p21/Cdc42/Rac1-activated kinase 1	✓	PD
sptP_stm_prot	protein tyrosine phosphate	0.366	PRKACB	protein kinase, cAMP-dependent, catalytic subunit beta	✓	PD
sptP_stm_prot	protein tyrosine phosphate	-0.233	PRKCA	protein kinase C, alpha	✓	PD
sptP_stm_prot	protein tyrosine phosphate	-0.214	MAP4K4_prot	mitogen-activated protein kinase kinase kinase 4	✓	PD
sptP_stm_prot	protein tyrosine phosphate	0.264	MAPK3_prot	mitogen-activated protein kinase 3	✓	PD
sptP_stm_prot	protein tyrosine phosphate	-0.277	STAT1_prot	signal transducer and activator of transcription 1	✓	PD
sptP_stm_prot	protein tyrosine phosphate	0.376	STAT6	signal transducer and activator of transcription 6	✓	PD
sptP_stm_prot	protein tyrosine phosphate	0.2	RAC3	ras-related C3 botulinum toxin substrate 3	✓	PD
hilA_stm	invasion genes transcriptional regulator	-0.34	RAC2	ras-related C3 botulinum toxin substrate 2	✓	PD
hilA_stm	invasion genes transcriptional regulator	-0.293	GNA13	guanine nucleotide binding protein G(i)alpha 13	✓	PD
hilA_stm	invasion genes transcriptional regulator	-0.427	PPP3CB_prot	protein phosphatase 3 (formerly 2B) regulatory subunit 3	✓	PD
csrA_stm	carbon storage regulator	-0.256	TGFBR2	transforming growth factor, beta receptor 2	✓	PD
MPL_stm	myeloproliferative leukemia virus oncogene	0.211	THBS3	thrombospondin 3	✓	PD
MPL_stm	myeloproliferative leukemia virus oncogene	-0.142	CSF2RA	colony stimulating factor 2 receptor	✓	PD
pgtA_stm	Phosphoglycerate transport	-0.214	MAP3K8_prot	mitogen-activated protein kinase kinase kinase 8	✓	PD
glnL_stm	sensory kinase (phosphotransferase)	-0.285	ACTN4_prot	actinin, alpha 4	✓	PD
prgK_stm_prot	cell invasion protein; lipoprotein	-0.307	HSPA1A_prot	heat shock 70kDa protein 1A	✓	SS
prgK_stm_prot	cell invasion protein; lipoprotein	0.193	TRAF2	TNF receptor-associated factor 2	✓	SS
prgK_stm_prot	cell invasion protein; lipoprotein	0.187	CREBBP	CREB binding protein (Rubinstein-Taybi syndrome associated)	✓	SS
prgK_stm_prot	cell invasion protein; lipoprotein	-0.134	WASF1	WAS protein family, member 1	✓	SS
fimZ_stm	fimbrial protein Z, putative	-0.356	ARPC1A_prot	actin related protein 2/3 complex, subunit 1A	✓	SS
hilA_stm	invasion genes transcriptional regulator	-0.394	HSP90	heat shock 70kDa protein 5 (glucose 6-phosphate 1-kinase 3)	✓	SS
hilA_stm	invasion genes transcriptional regulator	-0.256	SMURF2_prot	SMAD specific E3 ubiquitin protein ligase 2	✓	SS
sptP_stm_prot	protein tyrosine phosphate	-0.306	CSF2RA	colony stimulating factor 2 receptor	✓	SS
sptP_stm_prot	protein tyrosine phosphate	0.222	IL2RG	interleukin 2 receptor, gamma (severe combined immunodeficiency)	✓	SS
zraP_stm	zinc-resistance associated protein	-0.365	IL13RA1_prot	interleukin 13 receptor, alpha 1	✓	SS
cher_stm	glutamate methyltransferase	0.204	WASL	Wiskott-Aldrich syndrome-like protein	✓	SS
sipC_stm	cell invasion protein	0.162	INHBB	inhibin, beta B	✓	SS
		0.251	CDKN2B	cyclin-dependent kinase inhibitor 2B		

# Interactome Model Learned from Prior Knowledge and Gene Perturbation Expression Data (*B. Melitensis* and Bovine Host)



Pathogen Predicted Interaction Points with Host



Pathogen Gene/Protein

- ◆ Up regulated
- ◆ Down regulated

Arcs Connecting Pathogen  
To Host Protein  
Predicted Interactions



Host Gene/Protein

- Up regulated
- Down regulated

# Pathogen-Host Predicted Protein-Protein Interactions

(Sample of highly correlated and anti-correlated interactions)

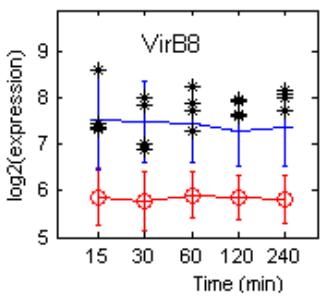
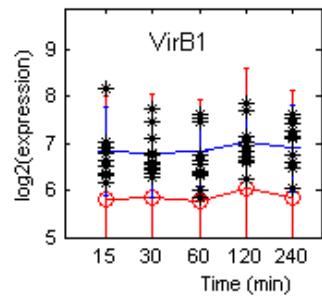
Pathogen Gene	Pathogen Gene Description	Normalized Correlation Wght.	Host Gene	Host Gene Description	Prediction Source
BMEI1829_bme	25 kDa outer-membrane immunogenic protein	0.588	BAD	BCL2-antagonist of cell death	Sequence Similarity
BMEI1829_bme	25 kDa outer-membrane immunogenic protein	0.559	NFKBIA	nuclear factor of kappa light polypeptide gene enhancer in B-cells, alpha	Sequence Similarity
BMEI1249_bme	25 kDa outer-membrane immunogenic protein	0.515	MAP2K2	mitogen-activated protein kinase kinase 2	Sequence Similarity
BMEI1249_bme	25 kDa outer-membrane immunogenic protein	0.446	TRAF2	TNF receptor-associated factor 2	Sequence Similarity
BMEI1872_bme	cell surface protein	0.367	JUN	jun oncogene	Sequence Similarity
scpA_bme	scpA	0.343	CRK	v-crk sarcoma virus CT10 oncogene homolog (avian)	Sequence Similarity
hfq_bme	host factor-I protein	0.319	IKBKE	inhibitor of kappa light polypeptide gene enhancer in B-cells, kinase epsilon	Sequence Similarity
BMEI1249_bme	25 kDa outer-membrane immunogenic protein	0.315	MAPK9	mitogen-activated protein kinase 9	Sequence Similarity
BMEI1249_bme	25 kDa outer-membrane immunogenic protein	0.304	MAP2K7	mitogen-activated protein kinase kinase 7	Sequence Similarity
flhB_bme	flagellar biosynthesis protein FlhB	0.294	PIK3C3	phosphoinositide-3-kinase, class 3	Sequence Similarity
BMEI1249_bme	25 kDa outer-membrane immunogenic protein	0.281	MAP2K3	mitogen-activated protein kinase kinase 3	Sequence Similarity
hfq_bme	host factor-I protein	0.265	CRKL	v-crk sarcoma virus CT10 oncogene homolog (avian)	Sequence Similarity
eno_bme	phosphopyruvate hydratase (EC:4.2.1.11)	0.243	MYD88	myeloid differentiation primary response gene 88	Sequence Similarity
BMEI0717_bme	22 kDa outer membrane protein precursor	0.226	MAP2K2	mitogen-activated protein kinase kinase 2	Sequence Similarity
glnL_bme	two-component system, NtrC family, nitrogen	0.224	RAC1	ras-related C3 botulinum toxin substrate 1 (rho)	Sequence Similarity
VirB5_bme	K03200 type IV secretion system	0.224	MAP2K6	mitogen-activated protein kinase kinase 6	Sequence Similarity
scpA_bme	scpA	0.202	PTK2	PTK2 protein tyrosine kinase 2	Sequence Similarity
glnL_bme	two-component system, NtrC family, nitrogen	0.2	MAP2K6	mitogen-activated protein kinase kinase 6	Sequence Similarity
dctM_bme	C4-dicarboxylate transporter, DctM subunit	-0.237	PPP2CA	protein phosphatase 2 (formerly 2A), catalytic subunit, type 1A	Sequence Similarity
fnr_bme	fnr	-0.239	NR4A1	nuclear receptor subfamily 4, group A, member 1	Sequence Similarity
scpA_bme	scpA	-0.241	LY96	lymphocyte antigen 96	Sequence Similarity
glnL_bme	two-component system, NtrC family, nitrogen	-0.244	HSPA1A	heat shock 70kDa protein 1A	Sequence Similarity
YidC_bme	preprotein translocase subunit YidC	-0.271	MAP4K3	mitogen-activated protein kinase kinase kinase 3	Sequence Similarity
BMEI1829_bme	25 kDa outer-membrane immunogenic protein	-0.325	SMAD2	SMAD family member 2	Sequence Similarity
BMEI1872_bme	cell surface protein	-0.339	PRKCG	protein kinase C, gamma	Sequence Similarity
trpD_bme	bifunctional anthranilate synthase/anthranilate	-0.373	Cab39	calcium binding protein 39	Domain Prediction
hfq_bme	host factor-I protein	-0.481	RHOA	ras homolog gene family, member A	Sequence Similarity

# Interactome Model Employed for Pathogen Gene KO Simulations (*B. Melitensis* in Bovine Host)

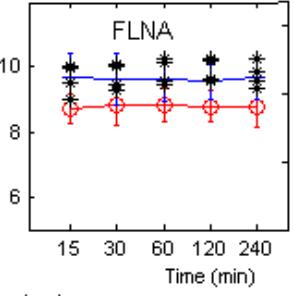
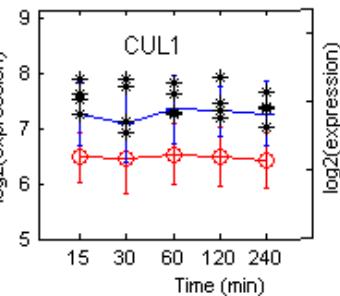
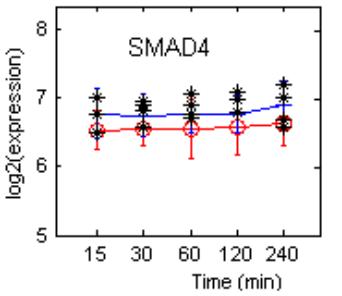
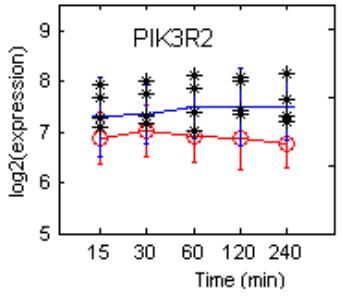
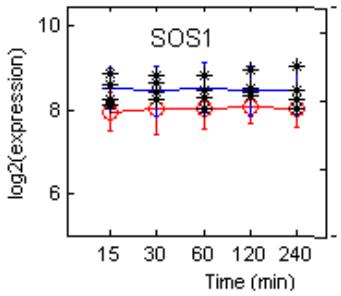
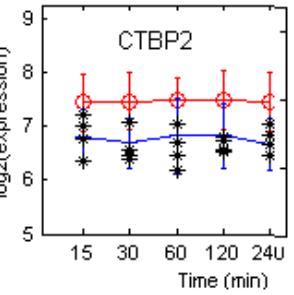
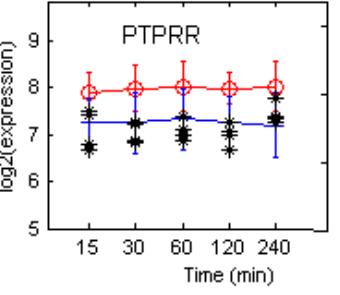
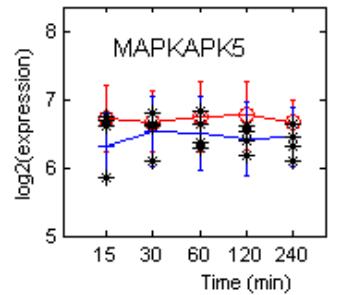
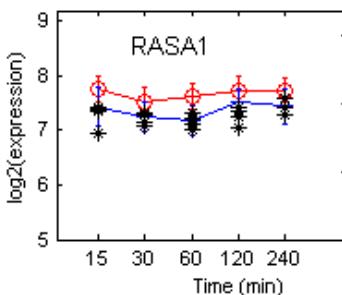
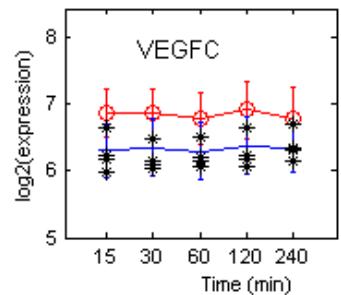
## VirB1 KO simulation

### Knockout

Pathogen



Host



### Significantly Modulated by KO

CTBP2	VirB8
TGFBR1	SOS1
MAPKAPK5	PPP2CA
SMAD4	DUSP16
PTPRR	PIK3CA
FLNA	MAPK7
KDR	RASA1
VEGFC	CUL1
IFNAR2	PIK3R2
CCND3	PPP5C

CTBP2

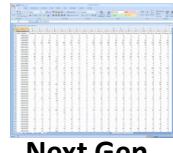
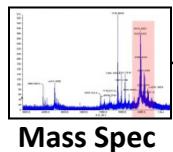
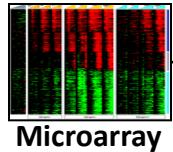
Knockout Simulation  
Pathogen WT Simulation

\* Real Host Expression Values  
Used to Train Model

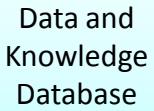
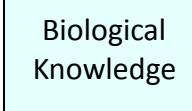
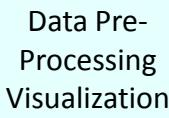


# Integrated platform Implements the Computational Pipeline (simplified overview)

## Data



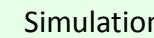
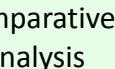
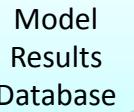
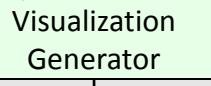
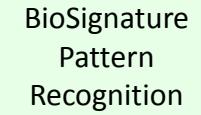
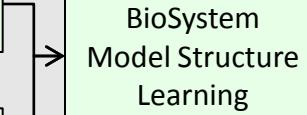
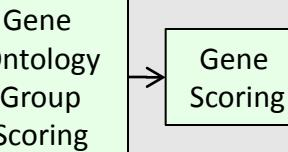
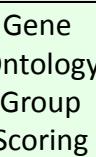
Microarray  
Mass Spec  
Next Gen Sequencing



Gene Ontologies Pathways

Genome/  
Proteome  
Annotations

## Bayesian Computational Methods (learning, scoring, comparing steps)



End User Visualization and Reporting System