MODELS IN BIOLOGY: STATISTICAL, MATHEMATICAL AND COMPUTATIONAL

Dr. Dhiraj Kumar

Immunology Group

FORMULATE THE PROBLEM

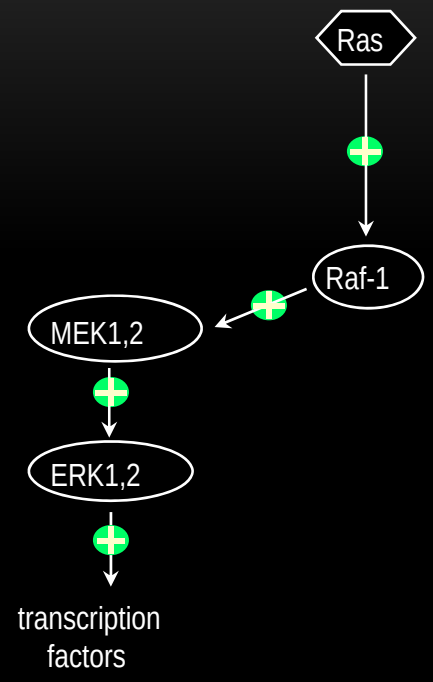
Diabetic/Obese Mice

Infected Mice

Normal Mice

Diseased Mice

What molecular events lead to these phenotypes??



Can certain features be monitored to understand the eventual phenotype

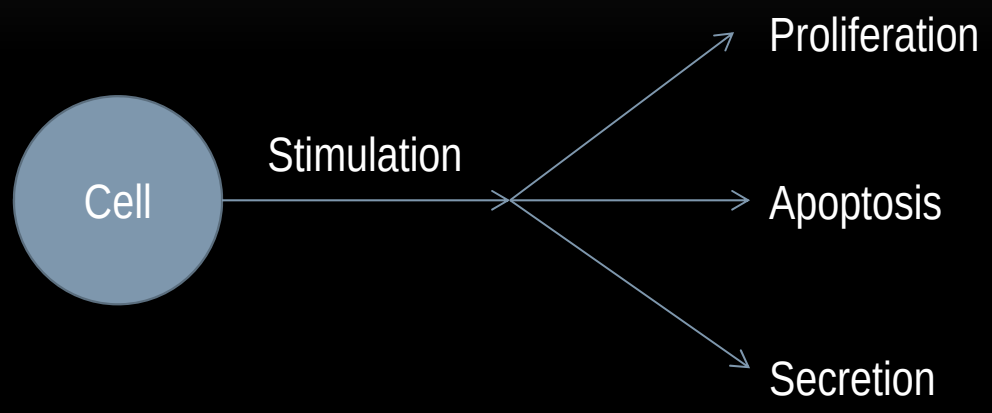
FEATURES THAT MAY BE MONITORED

- Gene expression profile (genome-wide/selective)
- Proteomic profiling
- Phosphorylation events etc.

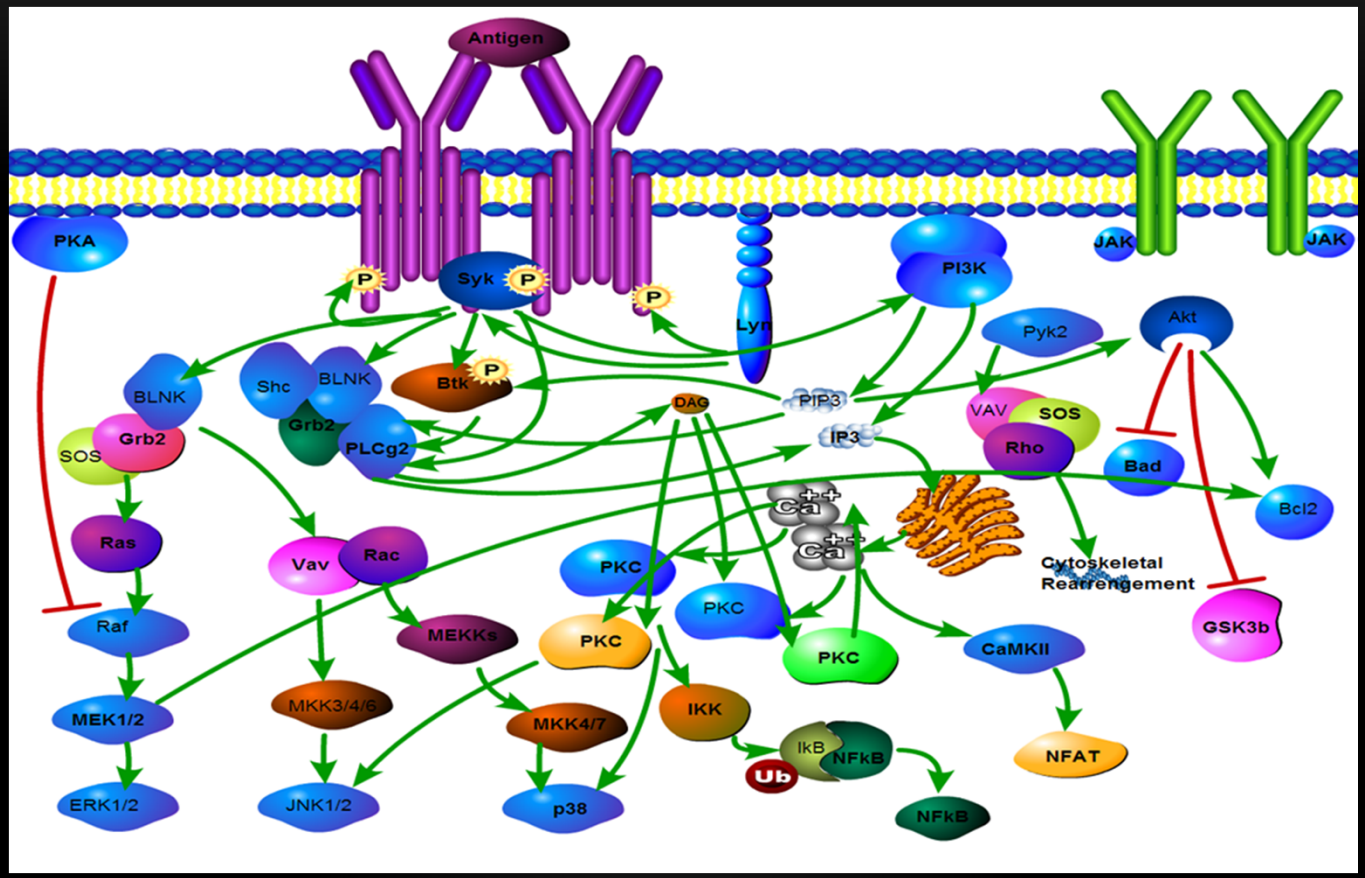
Question: Can we specifically monitor select fraction of the feature and predict what the cellular responses might be?

How to identify those select fraction of features: Molecular Signatures?

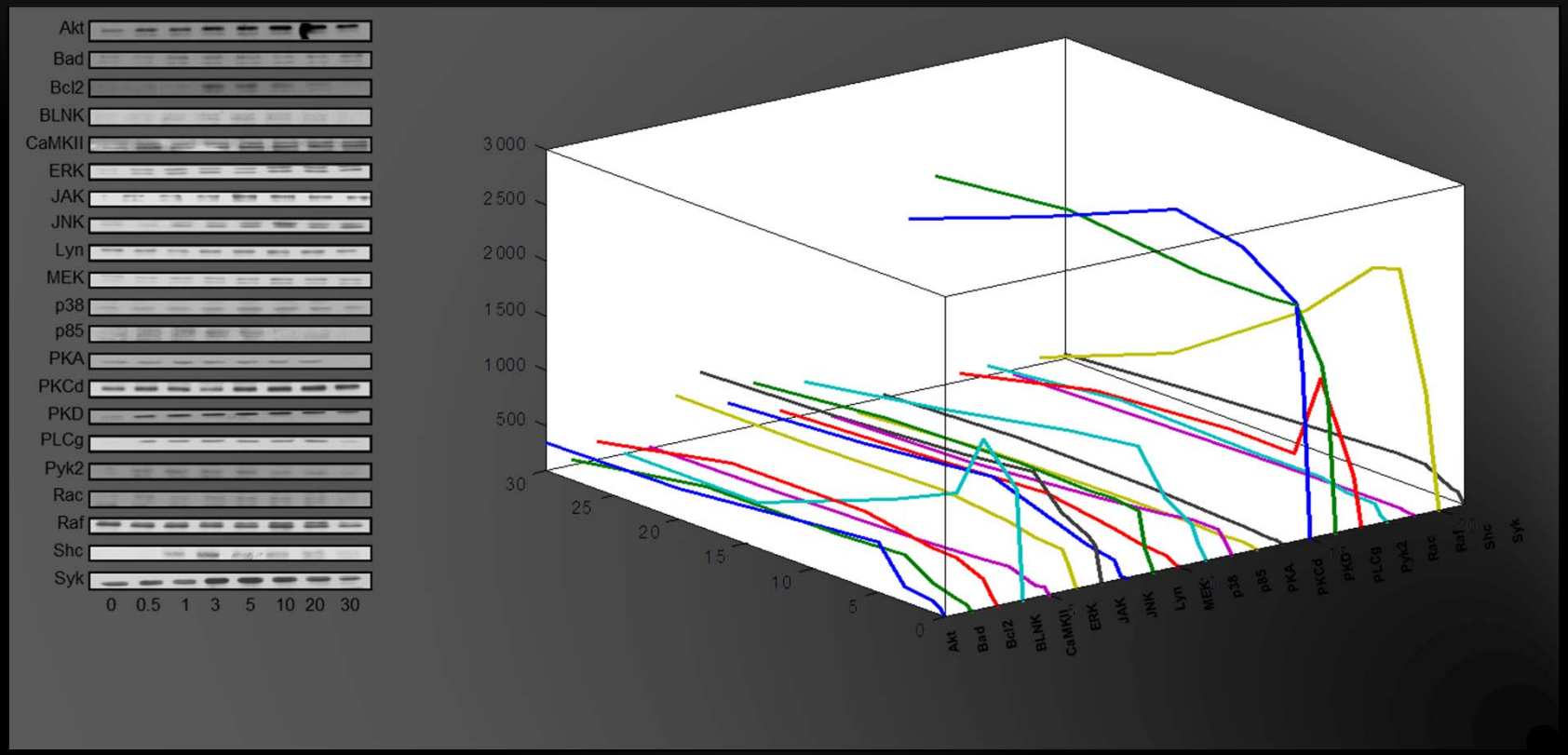
MOLECULAR SIGNATURES FOR A RECEPTOR DEPENDENT SIGNALING NETWORK



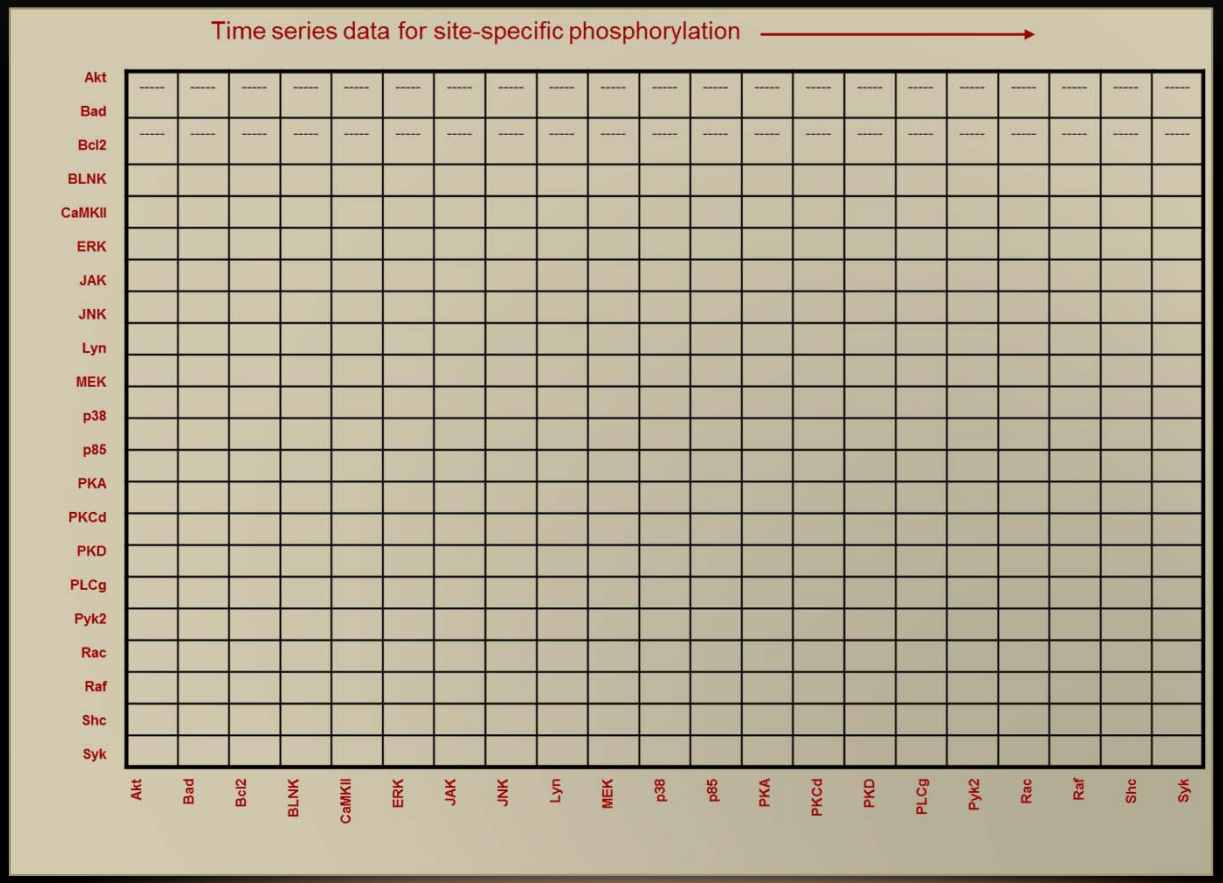
BCR SIGNALING NETWORK



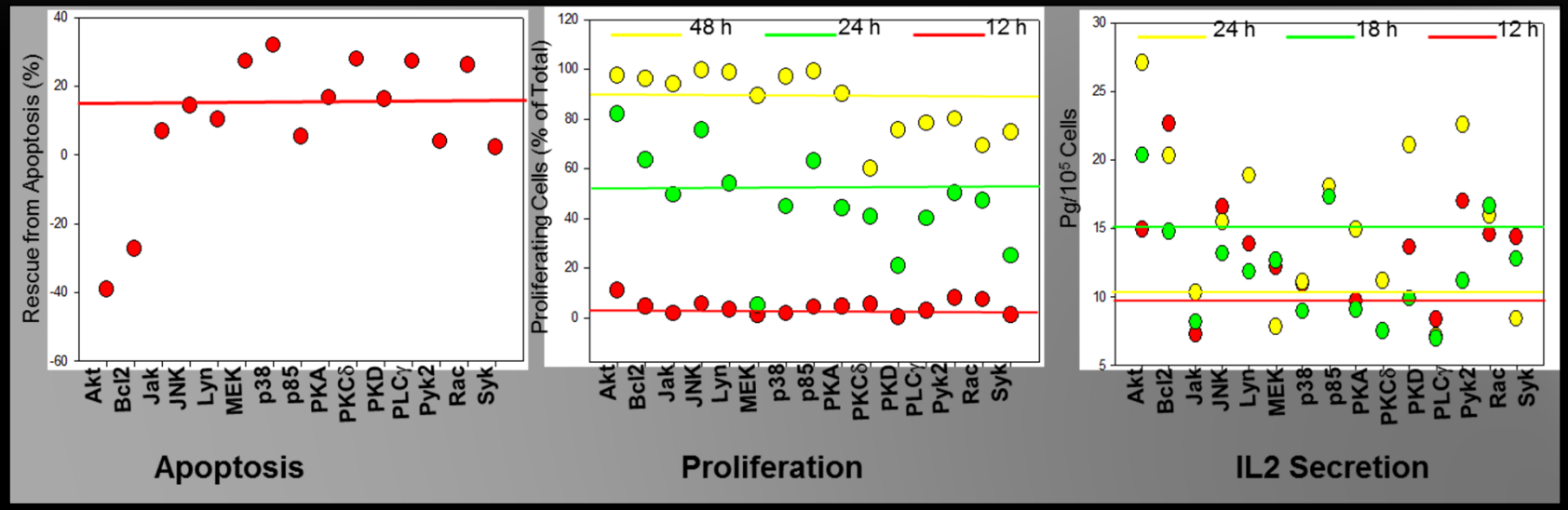
SIGNALING EVENTS UPON BCR STIMULATION (KEY MOLECULES)



MONITOR SIGNALING EVENTS UNDER VARIETY OF CONDITIONS



ALSO MEASURE EFFECT OF CONDITION ON CELLULAR RESPONSES

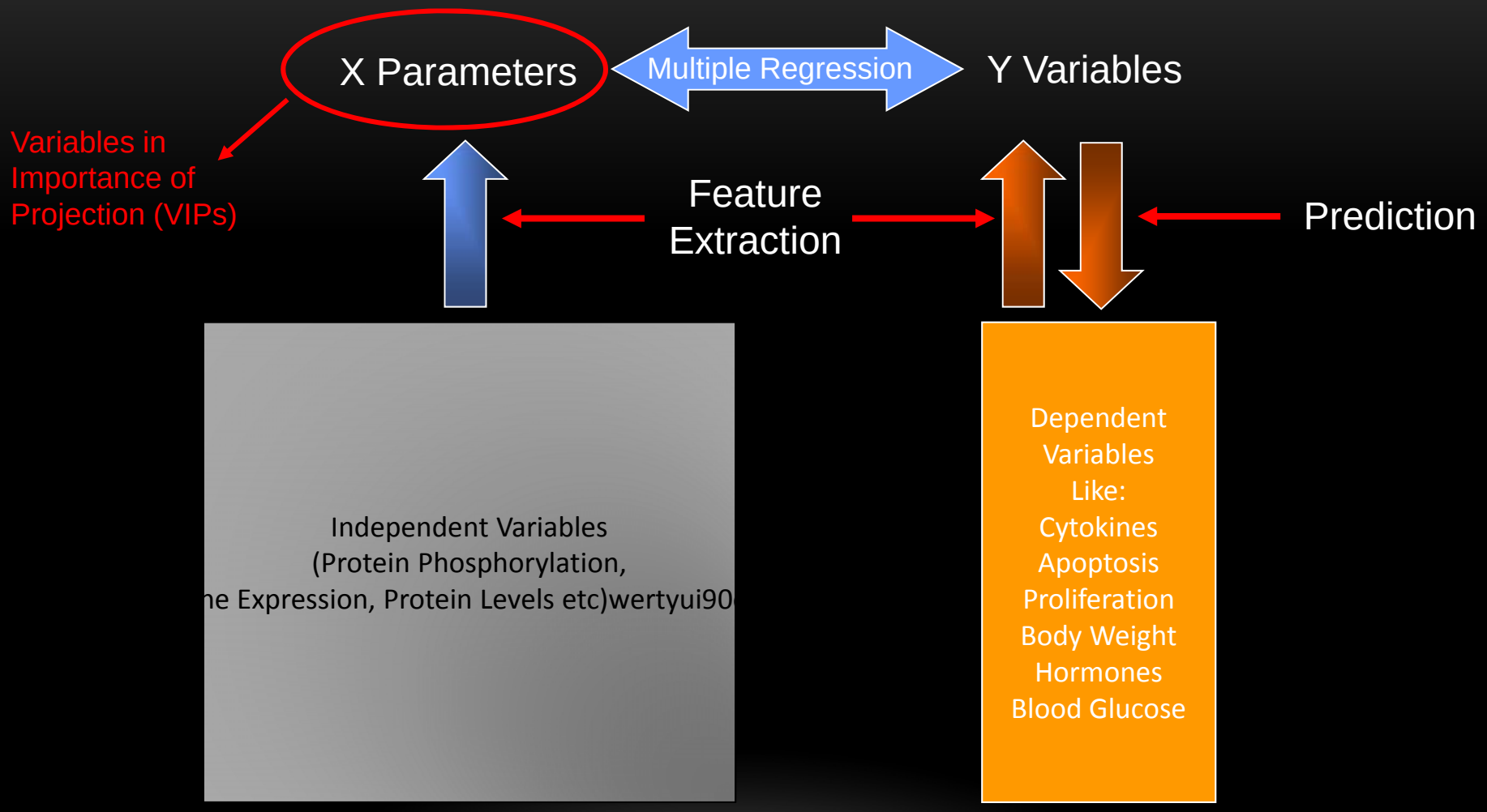


CAN WE ESTABLISH A RELATIONSHIP BETWEEN SIGNALING EVENTS AND CELLULAR RESPONSE?

How changes in signaling event leads to differential cellular responses?

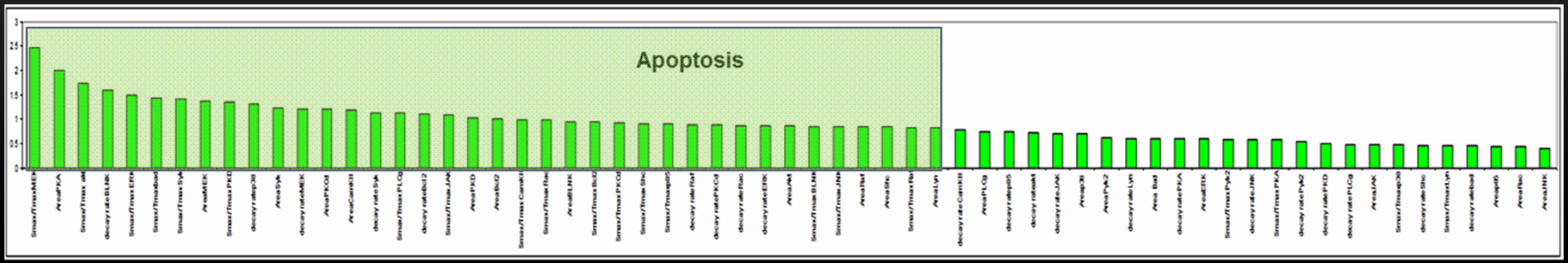
Are there components which have higher influence on deciding the cellular responses?

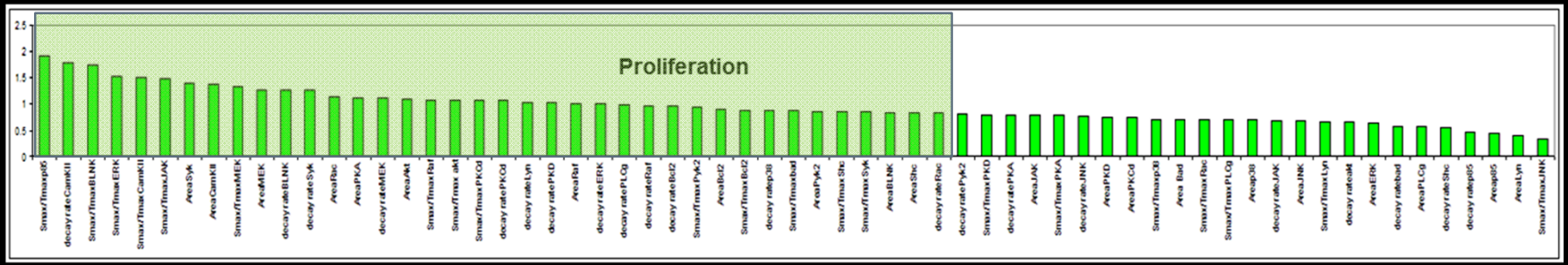
Partial Least Square Regression (PLSR) Modeling: Overview

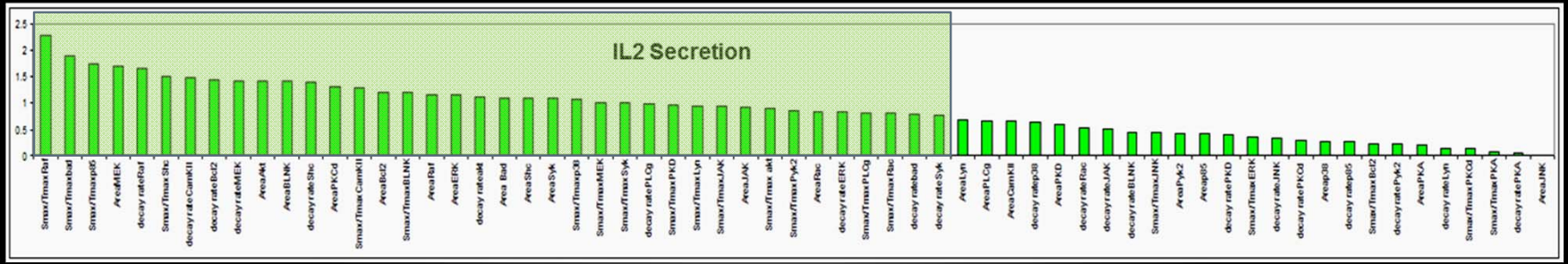


SIMCA P (Eumatrix)

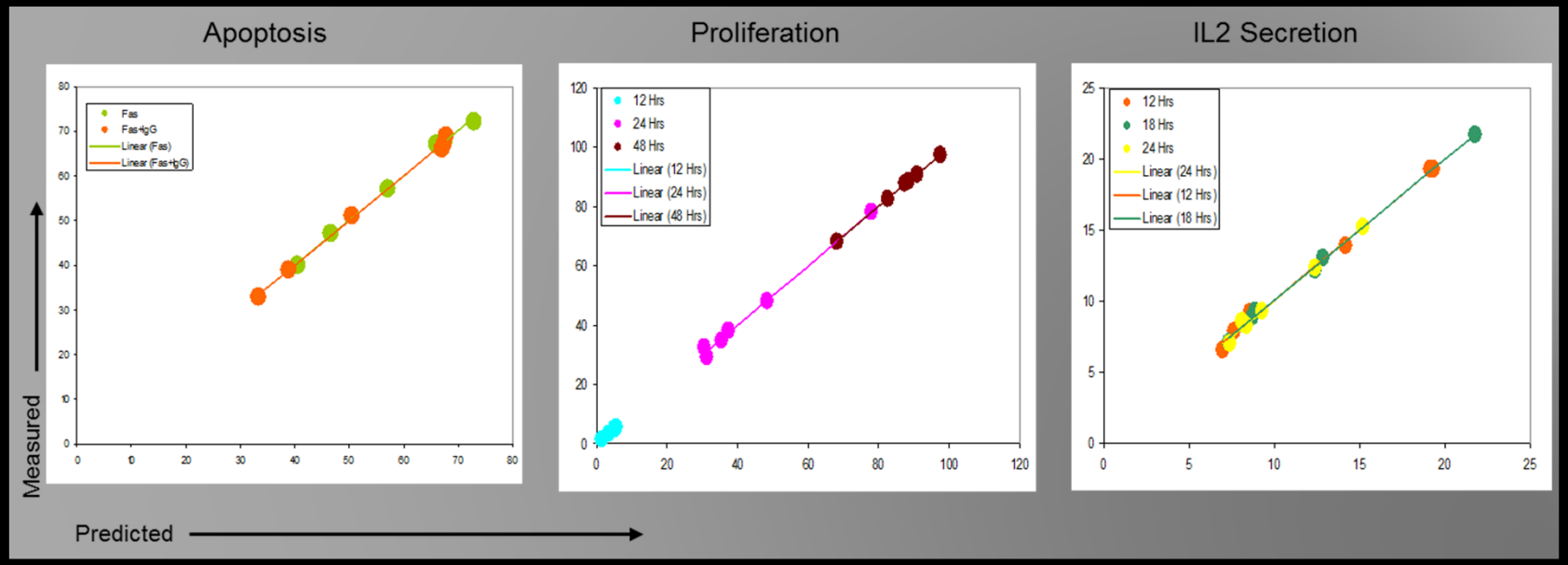
Variables in Importance of projection (VIPs)







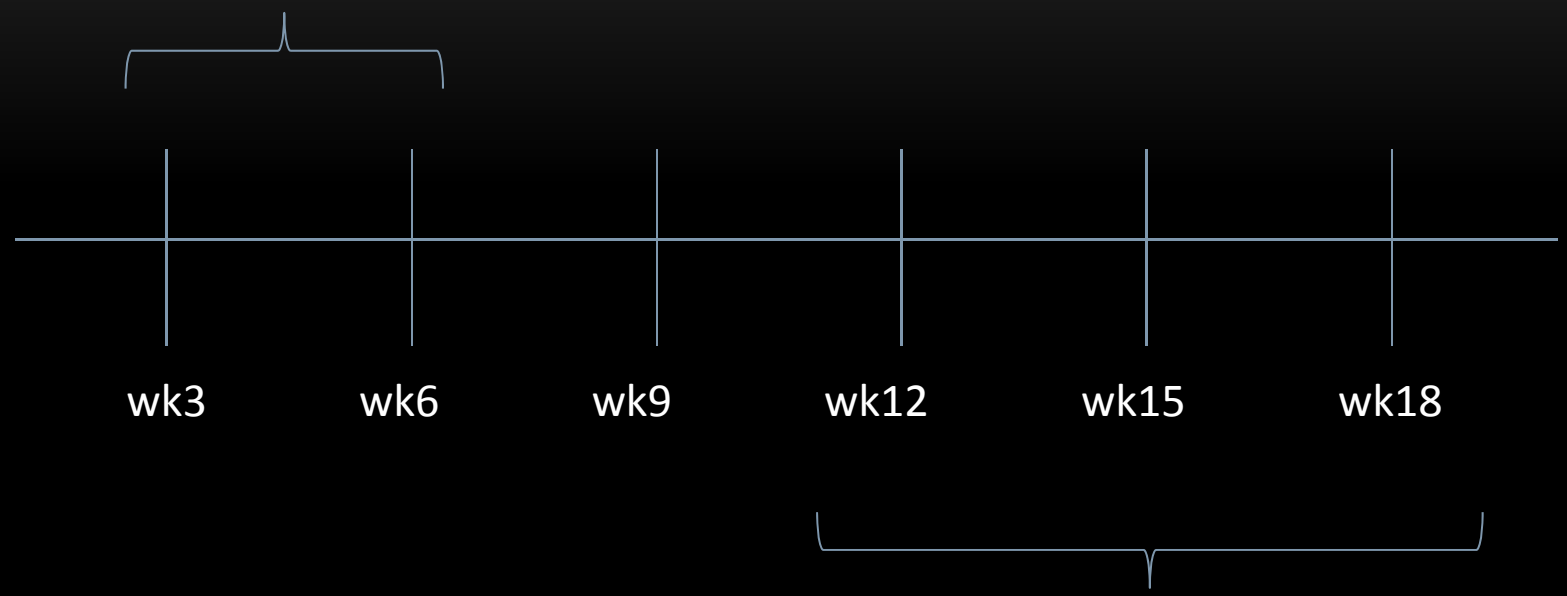
PREDICTION OF CELLULAR RESPONSES USING MOLECULAR SIGNATURES



PREDICTION OF OBESITY RESPONSE OF MICE FED ON HIGH CALORIE DIET

ADVANCE CELL AND MOLECULAR BIOLOGY
(ICG601)

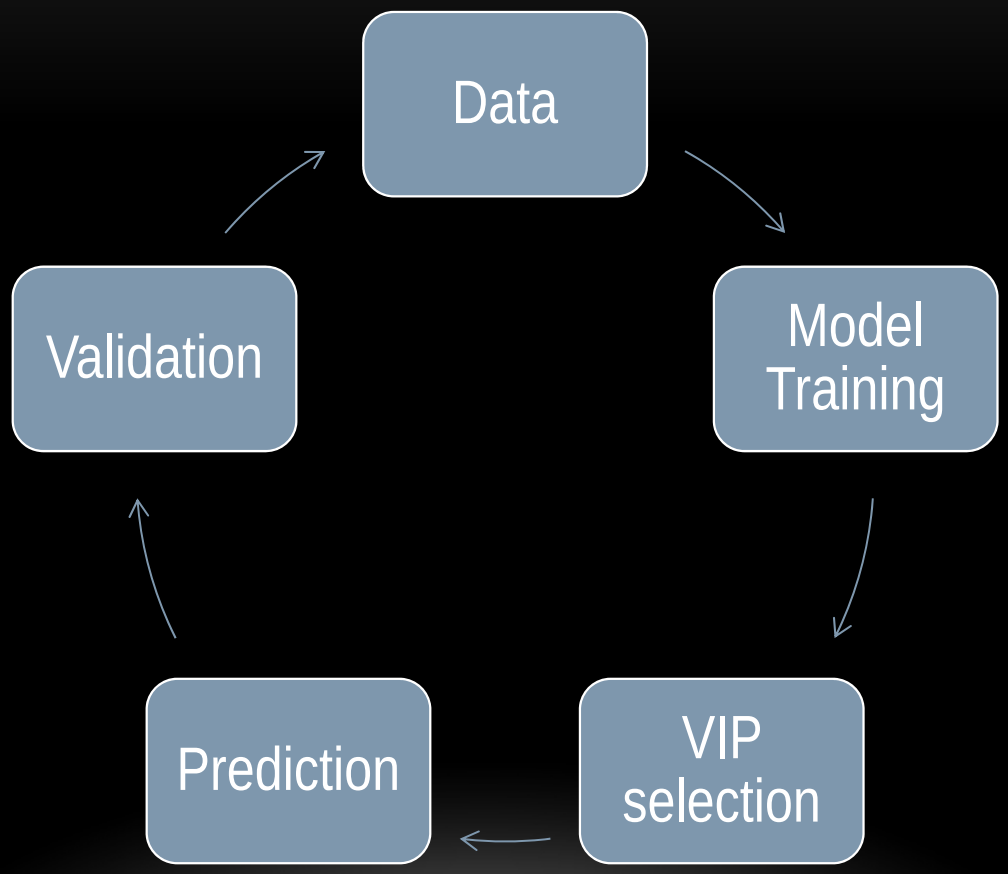
Expression data @
Early time point as
Independent variable



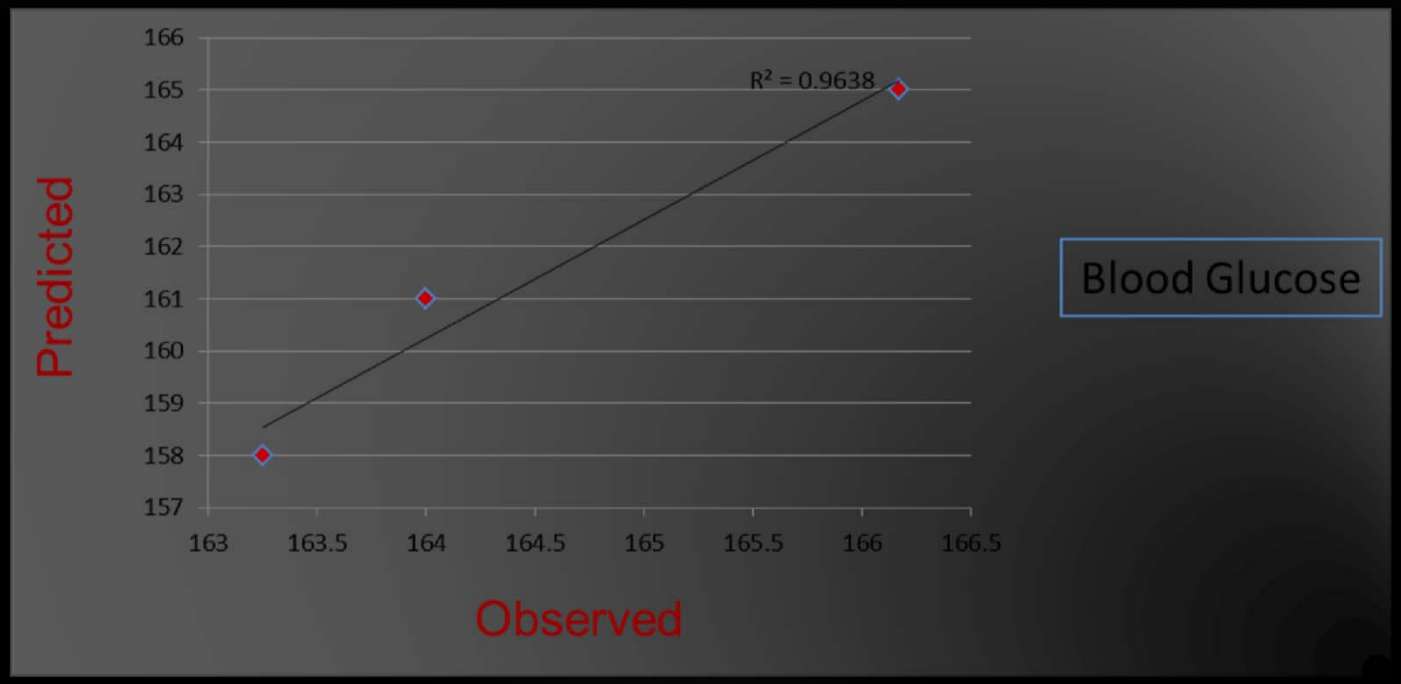
Predict phenotype for later time point
(dependent variable)

Establish state indicators (signatures)
@ early time point

PLS MODEL ITERATIVE FLOW-CHART



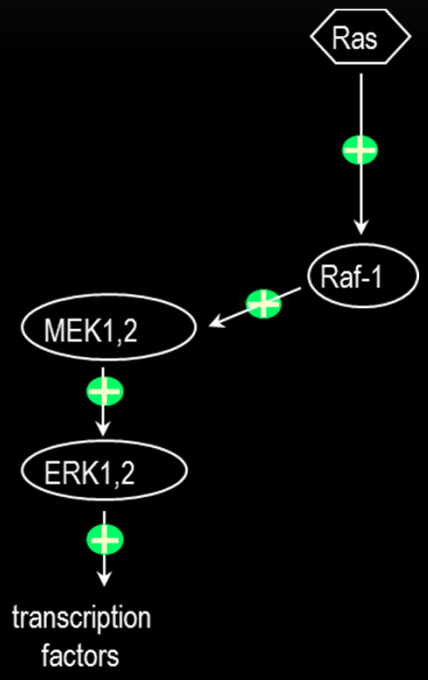
PREDICTING DIABETES IN MICE SEVERAL WEEKS BEFORE CHANGES IN BLOOD GLUCOSE LEVEL

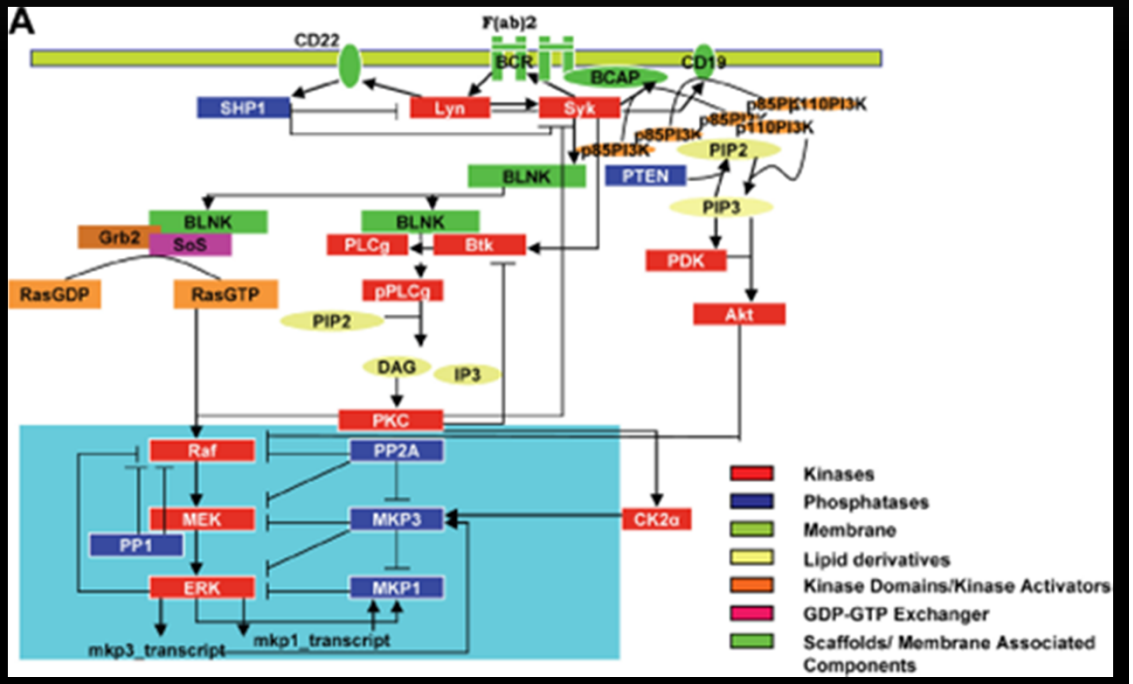


USES OF PLSR AND OTHER MACHINE LEARNING TECHNIQUES

- Establishing signatures of a response (molecular)
- Best suited for identification of Bio-markers!!!
- Help speed-up drug screening experiments

COMPUTATIONAL MODEL OF MAP KINASE PATHWAY FOR THEIR ROLE IN RESPONSE DIVERGENCE



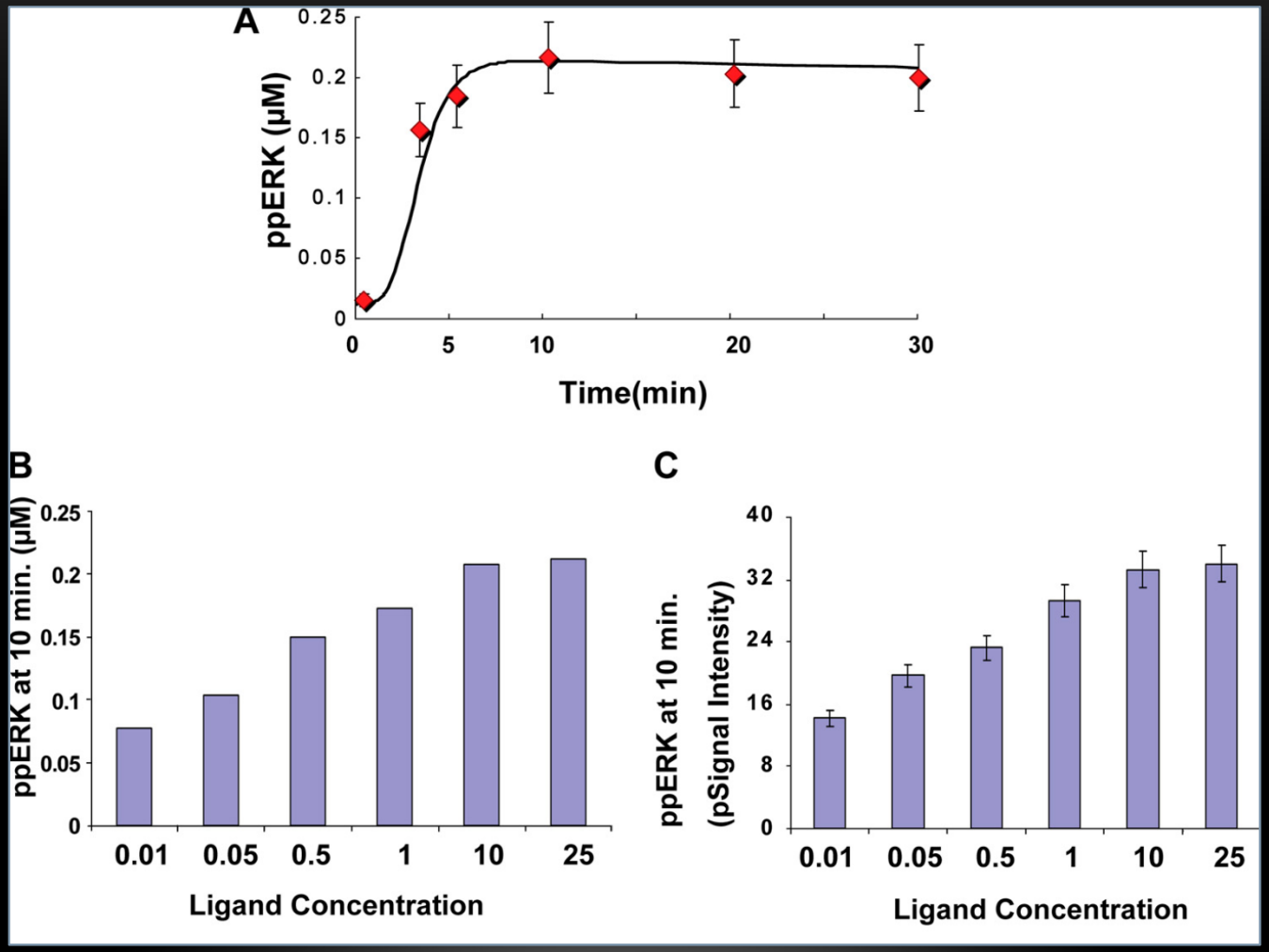


Role of phosphatases in shaping MAP kinase responses!!

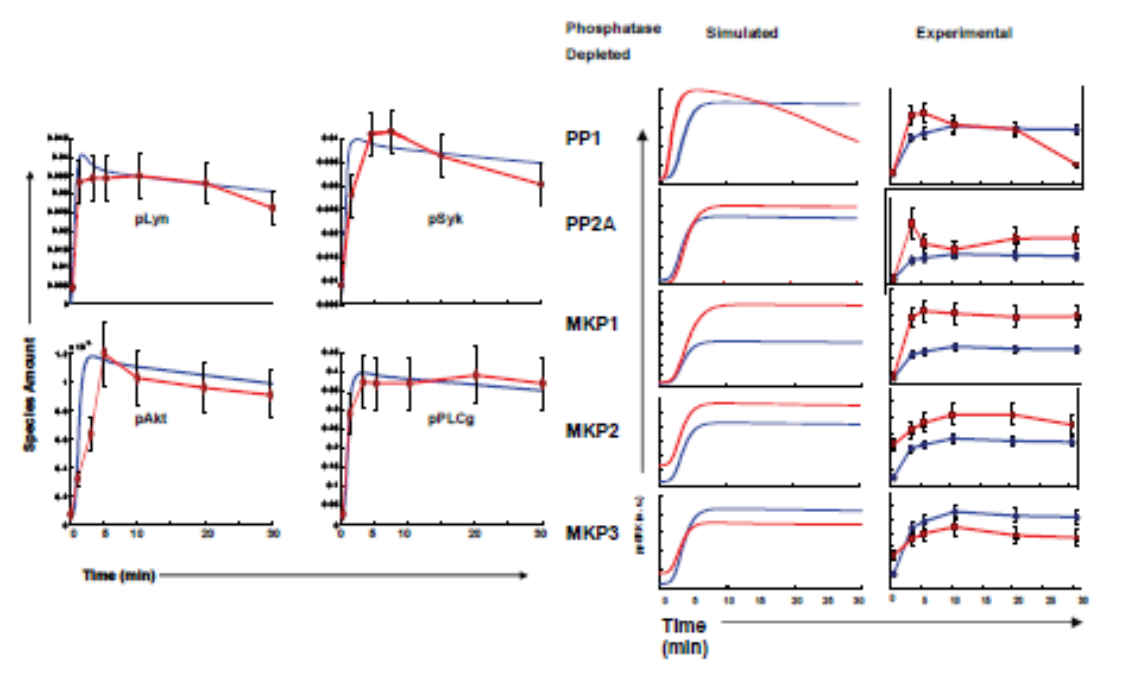
2. Reaction scheme and Reaction Rates:

	Reaction	Reaction Rate
1	L + BCR <-> [L:BCR1]	kon_1*L*BCR - koff_1*[L:BCR1]
2	[L:BCR] -> [L:BCRp]	kf_2*[L:BCR]/(0.25 + [L:BCR])
3	[L:BCRp] + Lyn <-> [L:BCRp:Lyn]	kon_3*[L:BCRp]*Lyn - koff_3*[L:BCRp:Lyn]
4	[L:BCRp:Lyn] -> [L:BCRp] + Lynp	kcat_4*[L:BCRp:Lyn]
5	Lynp + SHP1 <-> [Lynp:SHP1]	kon_5*Lynp*SHP1 - koff_5*[Lynp:SHP1]
6	[Lynp:SHP1] -> Lyn + SHP1	kcat_6*[Lynp:SHP1]
7	Lynp + CD22 -> Lynp + CD22p	vm_7*Lynp*CD22/(km_7+CD22)
8	CD22p + SHP1c <-> [CD22p:SHP1c]	kf_8*CD22p*SHP1c - kb_8*[CD22p:SHP1c]
9	[CD22p:SHP1c] -> CD22p + SHP1	kcat_9*[CD22p:SHP1c]
10	SHP1 -> SHP1c	kf_10*SHP1
11	CD22p -> CD22	kf_11*CD22p
12	Lynp + Syk <-> [Lynp:Syk]	kf_12*Lynp*Syk - kb_12*[Lynp:Syk]
13	[Lynp:Syk] -> Lynp + Sykp	kcat_13*[Lynp:Syk]
14	Sykp + SHP1 <-> [Sykp:SHP1]	kf_14*Sykp*SHP1 - kb_14*[Sykp:SHP1]
15	[Sykp:SHP1] -> Syk + SHP1	kcat_15*[Sykp:SHP1]
16	[L:BCRp] + Sykp <-> [L:BCRp:Sykp]	kf_16*[L:BCRp]*Sykp - kb_16*[L:BCRp:Sykp]
17	[L:BCRp:Sykp] -> [L:BCRpp] + Sykp	kcat_17*[L:BCRp:Sykp]
18	[L:BCRpp] + Lyn <-> [L:BCRpp:Lyn]	kf_18*[L:BCRpp]*Lyn - kb_18*[L:BCRpp:Lyn]
19	[L:BCRpp:Lyn] -> [L:BCRpp] + Lynp	kcat_19*[L:BCRpp:Lyn]
20	BCRpro -> BCR	kf_20*BCRpro
21	BCR -> BCRi	kf_21*BCR
22	2 [L:BCR1] <-> [L:BCR]	kon_22*[L:BCR1]^2 - koff_22*[L:BCR]
23	[L:BCR] -> BCRi + Li	kf_23*[L:BCR]
24	Btk + Lynp <-> [Btk:Lynp]	kf_24*Btk*Lynp - kb_24*[Btk:Lynp]
25	[Btk:Lynp] -> Btkp + Lynp	kcat_25*[Btk:Lynp]
26	Btk + Sykp <-> [Btk:Sykp]	kf_26*Btk*Sykp - kb_26*[Btk:Sykp]
27	[Btk:Sykp] -> Btkp + Sykp	kcat_27*[Btk:Sykp]
28	Btkp + BtkPase <-> [Btkp:BtkPase]	kf_28*Btkp*BtkPase - kb_28*[Btkp:BtkPase]
29	[Btkp:BtkPase] -> Btk + BtkPase	kcat_29*[Btkp:BtkPase]
30	Lynp + CD19 <-> [Lynp:CD19]	kf_30*Lynp*CD19 - kb_30*[Lynp:CD19]
31	[Lynp:CD19] -> Lynp + CD19p	kcat_31*[Lynp:CD19]
32	CD19p + p85PI3K <-> [CD19p:p85PI3K]	kf_32*CD19p*p85PI3K - kb_32*[CD19p:p85PI3K]
33	[CD19p:p85PI3K] + p110PI3K <-> [CD19p:p85PI3K:p110PI3K]	kf_33*[CD19p:p85PI3K]*p110PI3K - kb_33*[CD19p:p85PI3K:p110PI3K]
34	[CD19p:p85PI3K:p110PI3K] -> CD19p + PI3K	kcat_34*[CD19p:p85PI3K:p110PI3K]
35	Sykp + BCAP <-> [Sykp:BCAP]	kf_35*Sykp*BCAP - kb_36*[Sykp:BCAP]
36	[Sykp:BCAP] -> Sykp + BCAPp	kcat_36*[Sykp:BCAP]
37	BCAPp + p85PI3K <-> [BCAPp:p85PI3K]	kf_37*BCAPp*p85PI3K - kb_37*[BCAPp:p85PI3K]
38	[BCAPp:p85PI3K] + p110PI3K <-> [BCAPp:p85PI3K:p110PI3K]	kf_38*[BCAPp:p85PI3K]*p110PI3K - kb_38*[BCAPp:p85PI3K:p110PI3K]
39	[BCAPp:p85PI3K:p110PI3K] -> BCAPp + PI3K	kcat_39*[BCAPp:p85PI3K:p110PI3K]
40	BCAPp + Pase1 <-> [BCAPp:Pase1]	kf_40*BCAPp*Pase1 - kb_40*[BCAPp:Pase1]
41	[BCAPp:Pase1] -> BCAP + Pase1	kcat_41*[BCAPp:Pase1]
42	CD19p -> CD19	kf_42*CD19p
43	PI3K -> p85PI3K + p110PI3K	kf_43*PI3K
44	PIP2 + PI3K -> PIP3 + PI3K	vm_44 * PIP2 * PI3K/(km_44 + PIP2)
45	PIP3 -> PIP2	vm_45*PIP3/(km_45+PIP3)
46	PIP3 + Akte <-> [PIP3:Akte]	kf_46*PIP3*Akte - kb_46*[PIP3:Akte]
47	[PIP3:Akte] <-> Aktm	kf_47*[PIP3:Akte] - kb_47*Aktm
48	PIP3 + PDK <-> [PIP3:PDK]	kf_48*PIP3*PDK - kb_48*[PIP3:PDK]
49	[PIP3:PDK] -> PDKa	kf_49*[PIP3:PDK]
50	PDKa -> PDK + PIP3	kf_50*PDKa
51	PDKa + Aktm <-> [PDKa:Aktm]	kf_51*PDKa*Aktm - kb_51*[PDKa:Aktm]
52	[PDKa:Aktm] -> PDKa + Aktp	kcat_52*[PDKa:Aktm]
53	Aktp + PP2A <-> [Aktp:PP2A]	kf_53*Aktp*PP2A - kb_53*[Aktp:PP2A]
54	[Aktp:PP2A] -> Aktm + PP2A	kcat_54*[Aktp:PP2A]
55	Sykp + BLNK <-> [Sykp:BLNK]	kf_55*Sykp*BLNK - kb_55*[Sykp:BLNK]
56	[Sykp:BLNK] -> Sykp + BLNKp	kcat_56*[Sykp:BLNK]
57	BLNKp + Btkp <-> [BLNKp:Btkp]	kf_57*BLNKp*Btkp - kb_57*[BLNKp:Btkp]
58	BLNKp + Grb2 <-> [BLNKp:Grb2]	kf_58*BLNKp*Grb2 - kb_58*[BLNKp:Grb2]
59	BLNKp + Pase2 <-> [BLNKp:Pase2]	kf_59*BLNKp*Pase2 - kb_59*[BLNKp:Pase2]
60	[BLNKp:Pase2] -> BLNK + Pase2	kcat_60*[BLNKp:Pase2]
61	[BLNKp:Grb2] + Sos <-> [BLNKp:Grb2:Sos]	kf_61*[BLNKp:Grb2]*Sos - kb_61*[BLNKp:Grb2:Sos]
62	[BLNKp:Btkp] + PLCg <-> [BLNKp:Btkp:PLCg]	kf_62*[BLNKp:Btkp]*PLCg - kb_62*[BLNKp:Btkp:PLCg]
63	[BLNKp:Btkp:PLCg] -> [BLNKp:Btkp] + PLCgp	kcat_63*[BLNKp:Btkp:PLCg]
64	PLCgp + Pase3 <-> [PLCgp:Pase3]	kf_64*PLCgp*Pase3 - kb_64*[PLCgp:Pase3]
65	[PLCgp:Pase3] -> PLCg + Pase3	kcat_65*[PLCgp:Pase3]
66	IP3 -> ip3deg	kdeg_66*IP3
67	DAG -> dagdeg	kdeg_67*DAG
68	PLCgp + PIP2 -> PLCgp + IP3 + DAG	vm_68 * PIP2 * PLCgp/(km_68 +PIP2)
69	PKCcyt + DAG <-> [PKCcyt:DAG]	kf_69*PKCcyt*DAG - kb_69*[PKCcyt:DAG]
70	[PKCcyt:DAG] <-> PKCm	kf_70*[PKCcyt:DAG] - kb_70*PKCm
71	RasGTP -> RasGDP	kf_71*RasGTP
72	GEFp + RasGDP <-> [GEFp:RasGDP]	kf_72*GEFp*RasGDP - kb_72*[GEFp:RasGDP]
73	[GEFp:RasGDP] -> GEFp + RasGTP	kcat_73*[GEFp:RasGDP]
74	GEF + PKCm <-> [GEF:PKCm]	kf_74*GEF*PKCm - kb_74*[GEF:PKCm]
75	[GEF:PKCm] -> GEFp + PKCm	kcat_75*[GEF:PKCm]
76	PKCm + GAP <-> [PKCm:GAP]	kf_76*PKCm*GAP - kb_76*[PKCm:GAP]
77	[PKCm:GAP] -> PKCm + GAPp	kcat_77*[PKCm:GAP]
78	GEFp -> GEF	kf_78*GEFp
79	GAPp -> GAP	kf_79*GAPp
80	PKCm + Sykp <-> [PKCm:Sykp]	kf_80*PKCm*Sykp - kb_80*[PKCm:Sykp]
81	[PKCm:Sykp] -> PKCm + Sykpp	kcat_81*[PKCm:Sykp]
82	Sykpp -> Sykp	kf_82*Sykpp
83	GAP + RasGTP <-> [GAP:RasGTP]	kf_83*GAP*RasGTP - kb_83*[GAP:RasGTP]
84	[GAP:RasGTP] -> GAP + RasGDP	kcat_84*[GAP:RasGTP]
85	Raf + RasGTP <-> [Raf:RasGTP]	kf_85*Raf*RasGTP - kb_85*[Raf:RasGTP]
86	[Raf:RasGTP] + PKCm <-> [Raf:RasGTP:PKCm]	kf_86*[Raf:RasGTP]*PKCm - kb_86*[Raf:RasGTP:PKCm]
87	[Raf:RasGTP:PKCm] -> [Rafp:RasGTP] + PKCm	kcat_87*[Raf:RasGTP:PKCm]
88	Raf + Aktp <-> [Raf:Aktp]	kf_88*Raf*Aktp - kb_88*[Raf:Aktp]
89	[Raf:Aktp] -> Rafser259 + Aktp	kcat_89*[Raf:Aktp]
90	Rafser259 + PP1 <-> [Rafser259:PP1]	kf_90*Rafser259*PP1 - kb_90*[Rafser259:PP1]
91	[Rafser259:PP1] -> Raf + PP1	kcat_91*[Rafser259:PP1]
92	Rafp + Aktp <-> [Rafp:Aktp]	kf_92*Rafp*Aktp - kb_92*[Rafp:Aktp]
93	[Rafp:Aktp] -> Rafpp + Aktp	kcat_93*[Rafp:Aktp]
94	Rafpp + PP1 <-> [Rafpp:PP1]	kf_94*Rafpp*PP1 - kb_94*[Rafpp:PP1]
95	[Rafpp:PP1] -> Rafp + PP1	kcat_95*[Rafpp:PP1]
96	Rafp + PP2A <-> [Rafp:PP2A]	kf_96*Rafp*PP2A - kb_96*[Rafp:PP2A]

Tracing phosphorylation profile of ERK1/2 experimentally and computationally

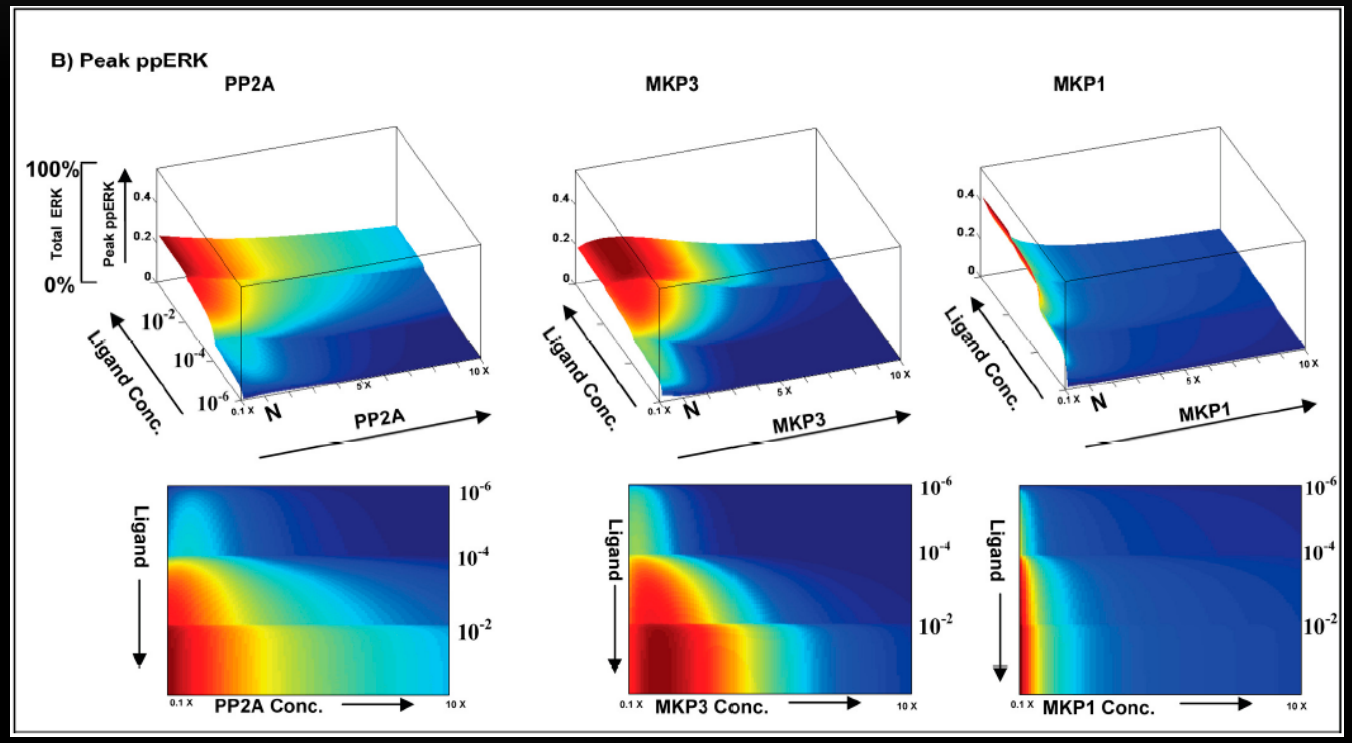


SIMULATIONS RECAPITULATE EXPERIMENTAL PROFILES



TRUE TO TYPE MODEL ALLOWS TO EXPLORE FURTHER

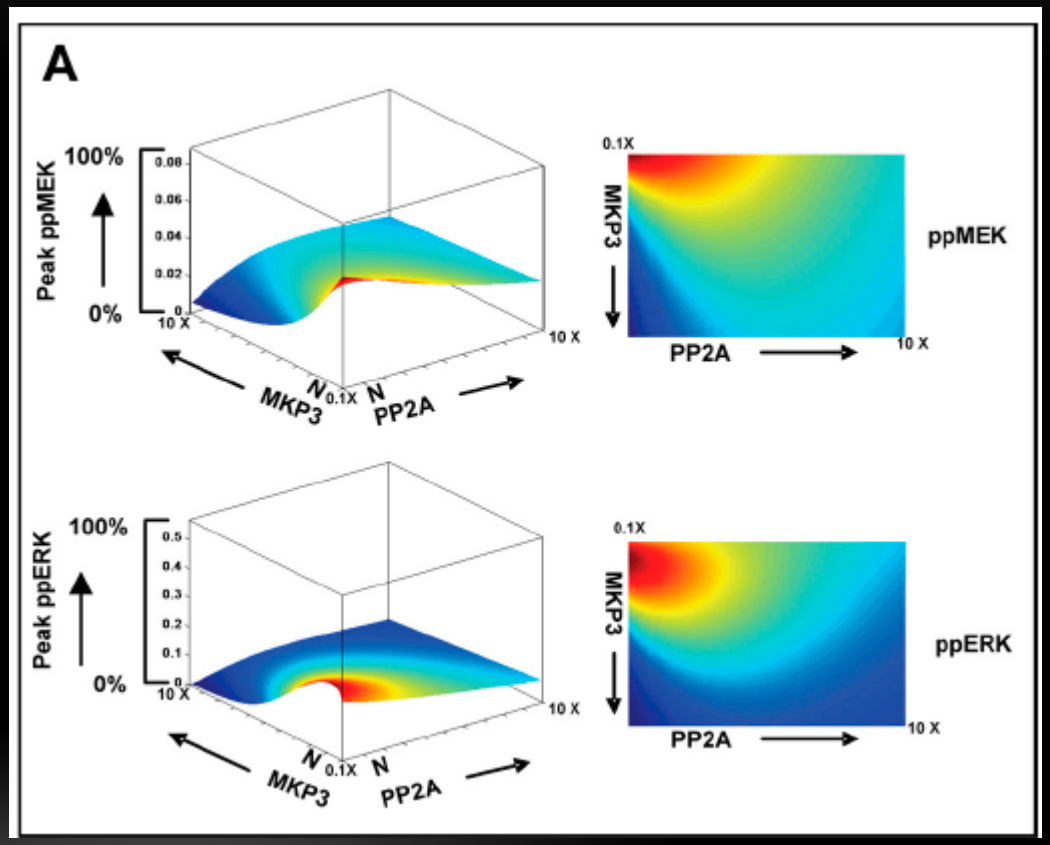
How concentration of phosphatases shape MAP kinase Response??



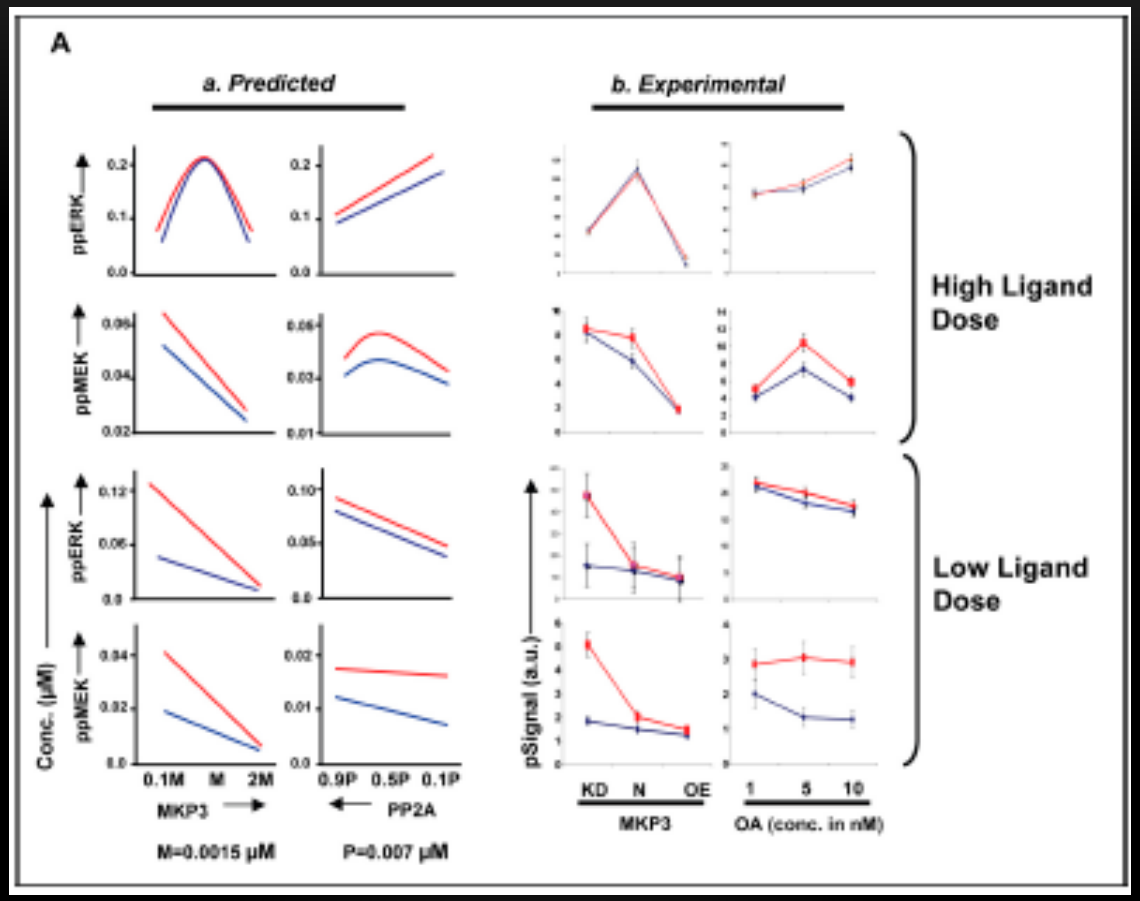
Ask difficult questions and try to address them!!

IDENTIFY CROSS-TALKS THAT ARE AS SUCH DIFFICULT TO IMAGINE OR EXPLAIN

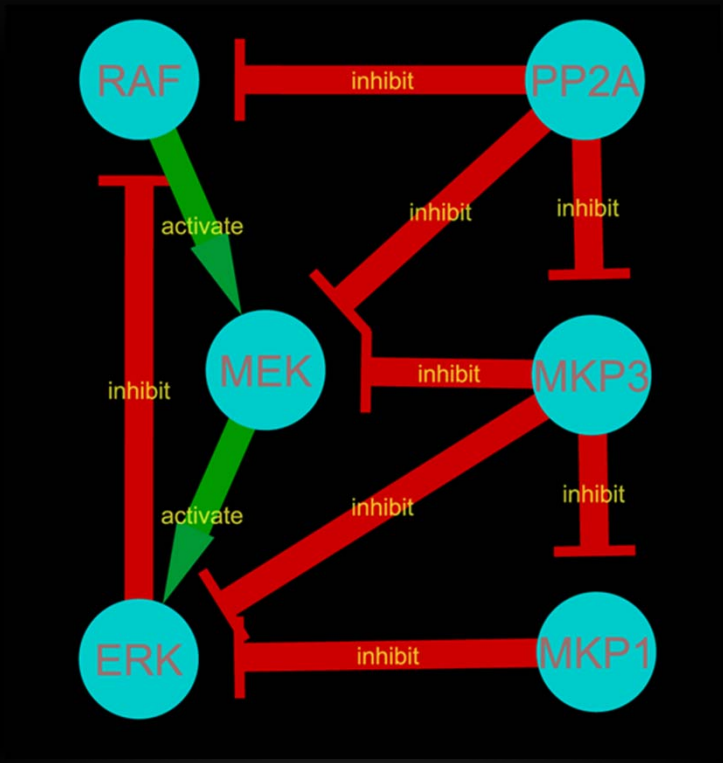
Can we understand cross-Regulation between two different Phosphatases?



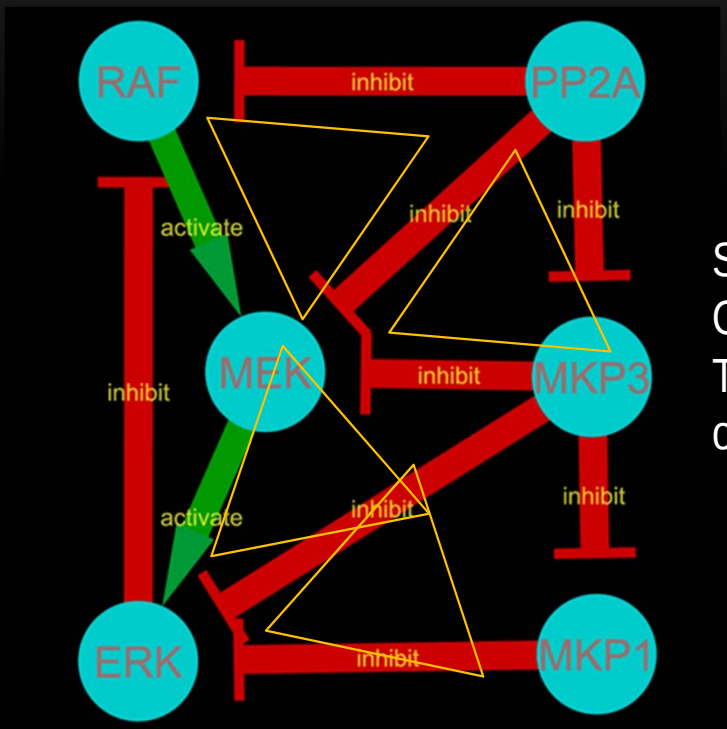
Experimentally verify the computationally observed cross-talk



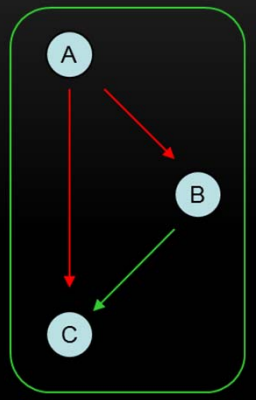
MAP KINASE PATHWAY IS INTEGRATED WITH A PARALLEL PHOSPHATASE CASCADE



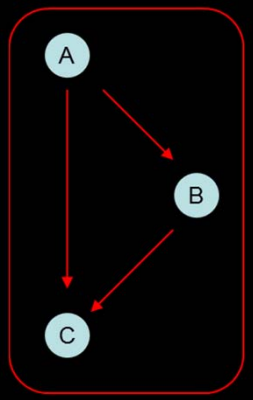
This integration helps regulate multitude of Cellular functions by this pathway!!!



Specific organization
Of Feed-forward loops
Through the integration
of phosphatases



Coherent II

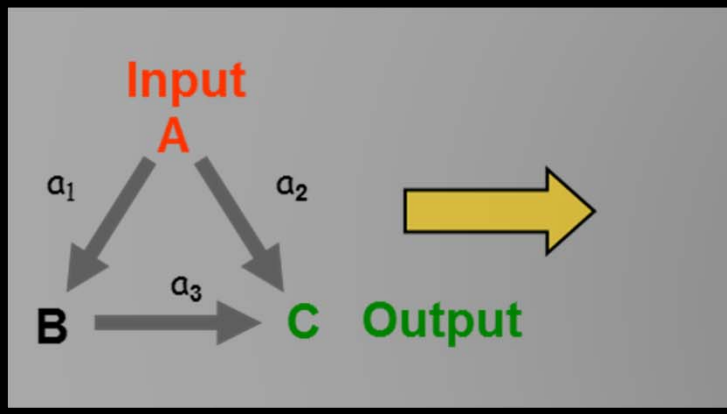


Incoherent II

COMPUTATIONAL (NUMERICAL) VERSUS MATHEMATICAL

- Beyond 3 variables, mathematical calculations become extremely difficult, but numerical calculations may be easily done
- A typical mathematical model does not have more than three nodes (there is no such limit for computational models)
- Results from a mathematical model are independent of the numerical value of parameters

MATHEMATICAL MODEL FOR A THREE NODE FEED-FORWARD LOOP

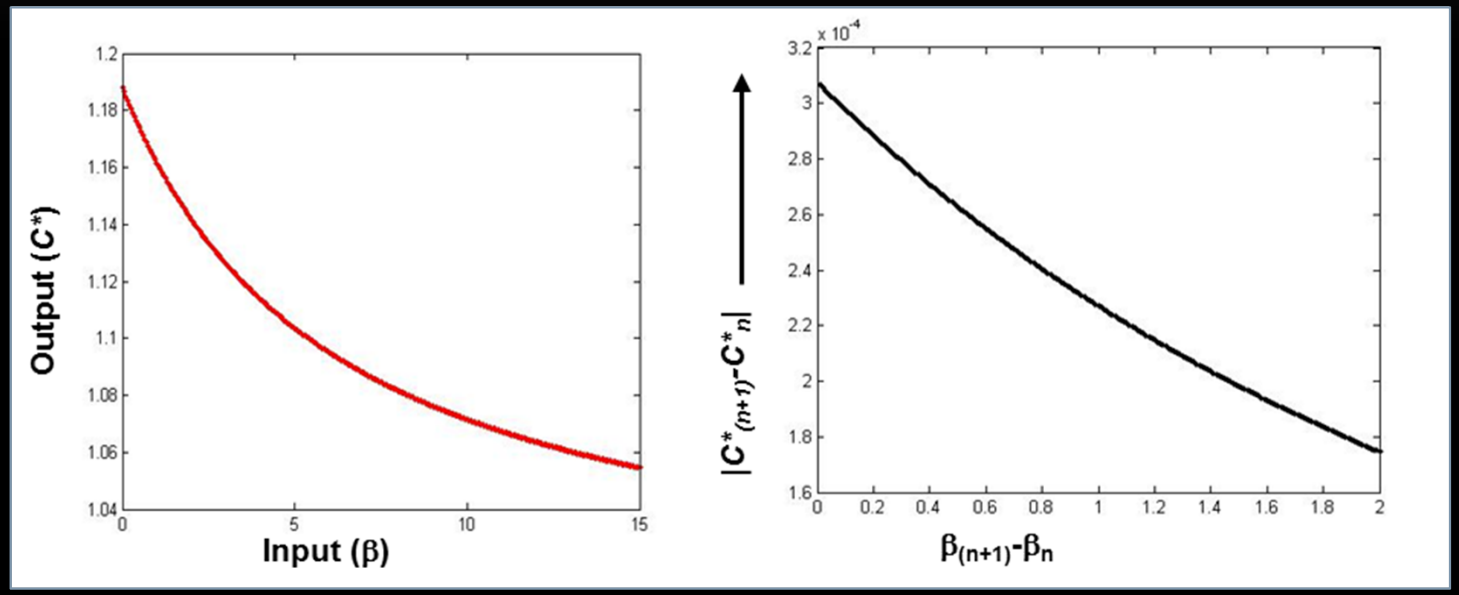


$$\frac{dA}{dt} = \alpha A\left(1 - \frac{A}{\beta}\right) \equiv F_1(A,B,C),$$

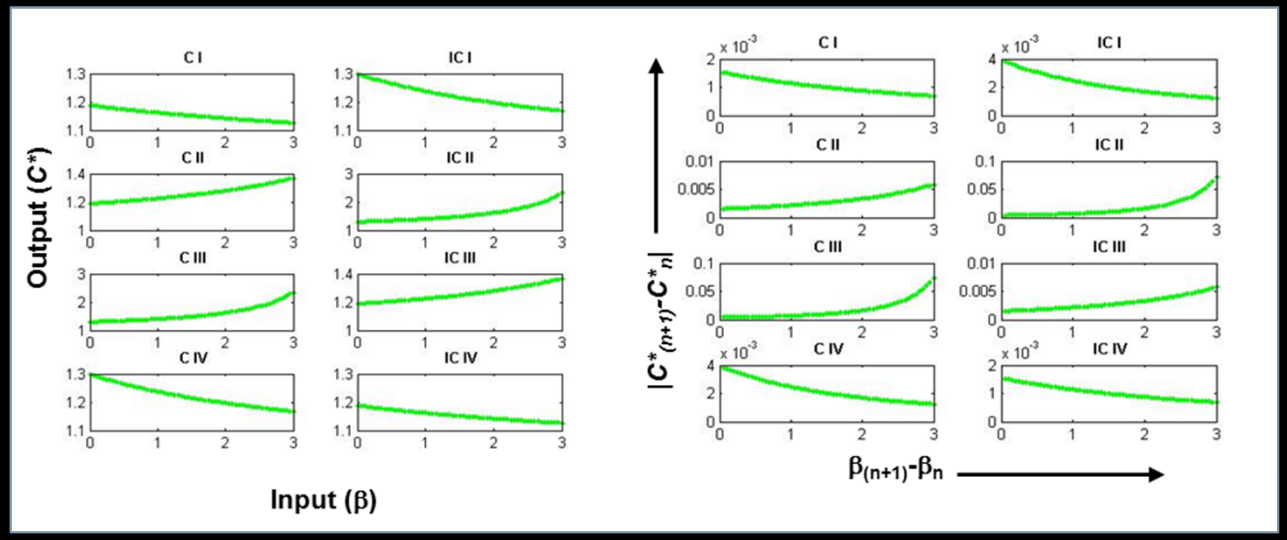
$$\frac{dB}{dt} = k_1 - \delta_1 B + a_1 A(1-B) \equiv F_2(A,B,C),$$

$$\frac{dC}{dt} = k_2 - \delta_2 C + a_2 A(1-C) + a_3 B(1-C) \equiv F_3(A,B,C),$$

STUDY INPUT-OUTPUT (I/O) RELATIONSHIP THROUGH A FFL



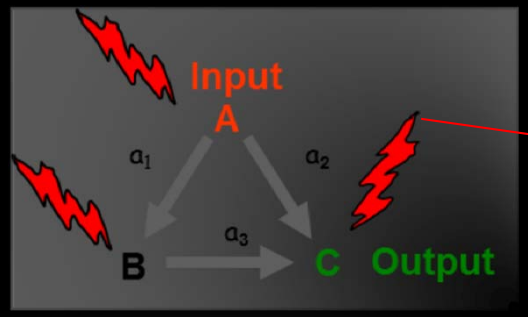
EFFECT OF FFL TYPES ON I/O RELATIONSHIP



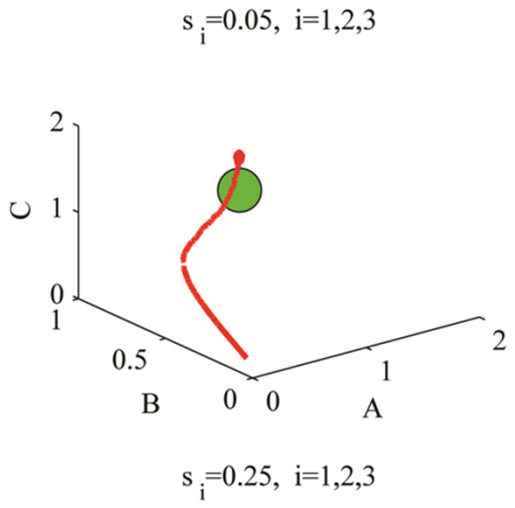
HOW DOES FFLS BEHAVE UNDER STOCHASTIC ENVIRONMENT

$$dA = F_1(A,B,C)dt + \sigma_1(A - A^*)d\xi_t^1$$
$$dB = F_2(A,B,C)dt + \sigma_2(B - B^*)d\xi_t^2$$
$$dC = F_3(A,B,C)dt + \sigma_3(C - C^*)d\xi_t^3$$

Incorporation of systemic Noise into the model

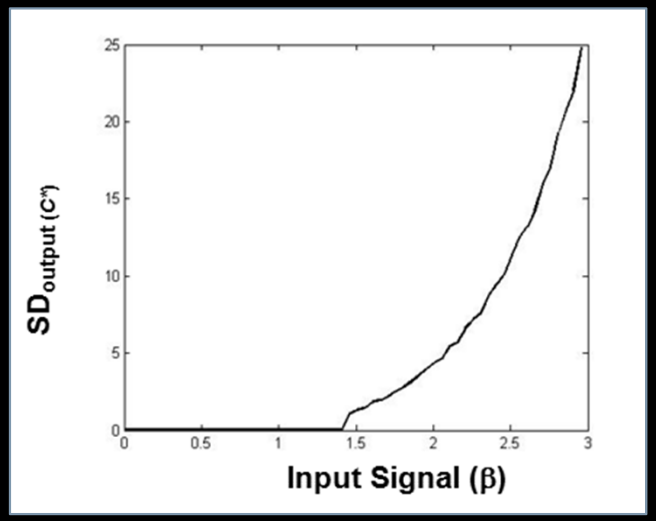


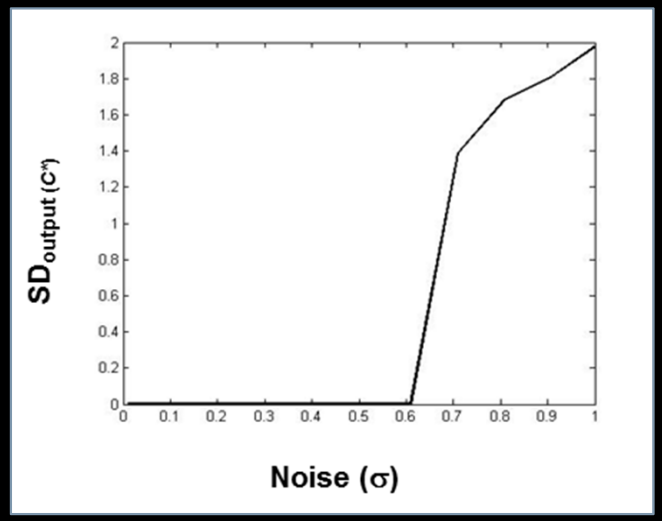
Cytokines
Growth Factors
Nutrients
Environment stress
Protein stability
Etc.



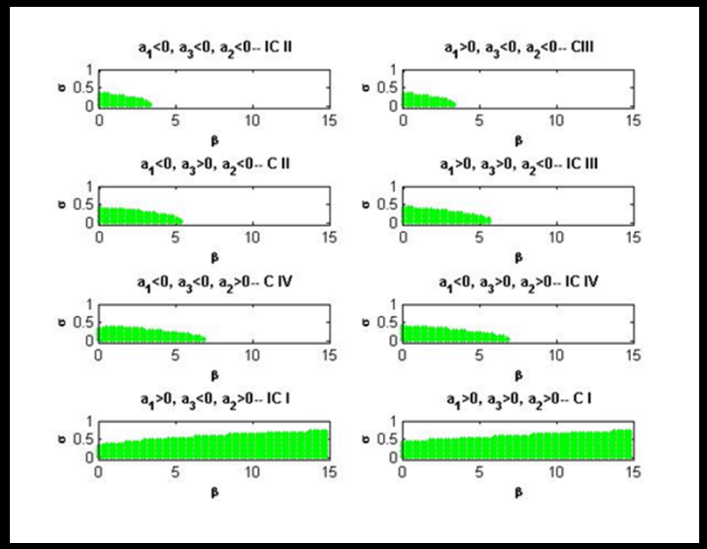
Output gets unstable at higher signal or noise in the presence of noise

HOW TO MEASURE INSTABILITY OF OUTPUT?



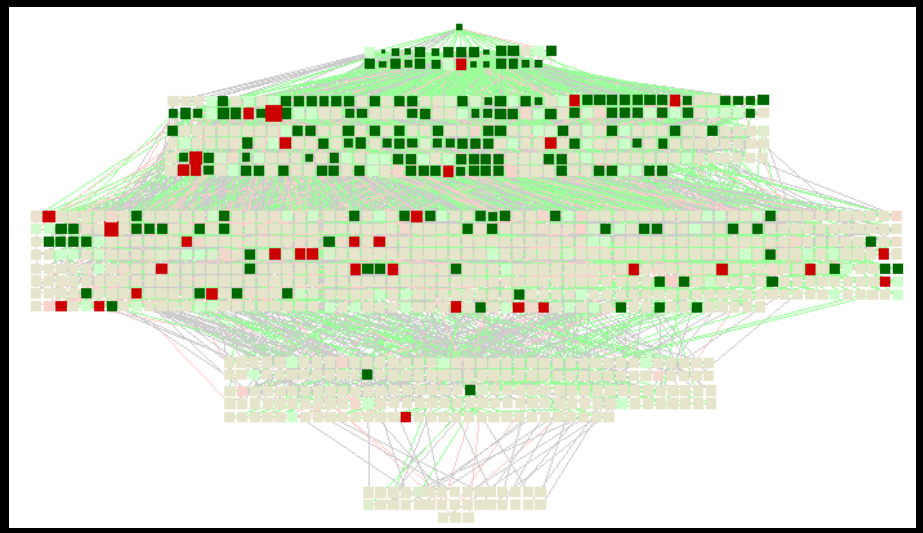


THE MEASURE OF INSTABILITY COULD BE USED TO RANK THE FFLS FOR THEIR ROBUSTNESS



Motif Type	Feature I	Area	Rank
Incoherent II	a1<0, a3<0, a2<0	73	1
Coherent III	a1>0, a3<0, a2<0	74	2
Coherent II	a1<0, a3>0, a2<0	137	3
Incoherent III	a1>0, a3>0, a2<0	148	4
Coherent IV	a1<0, a3<0, a2>0	162	5
Incoherent IV	a1<0, a3>0, a2>0	165	6
Incoherent I	a1>0, a3<0, a2>0	703	7
Coherent I	a1>0, a3>0, a2>0	721	8

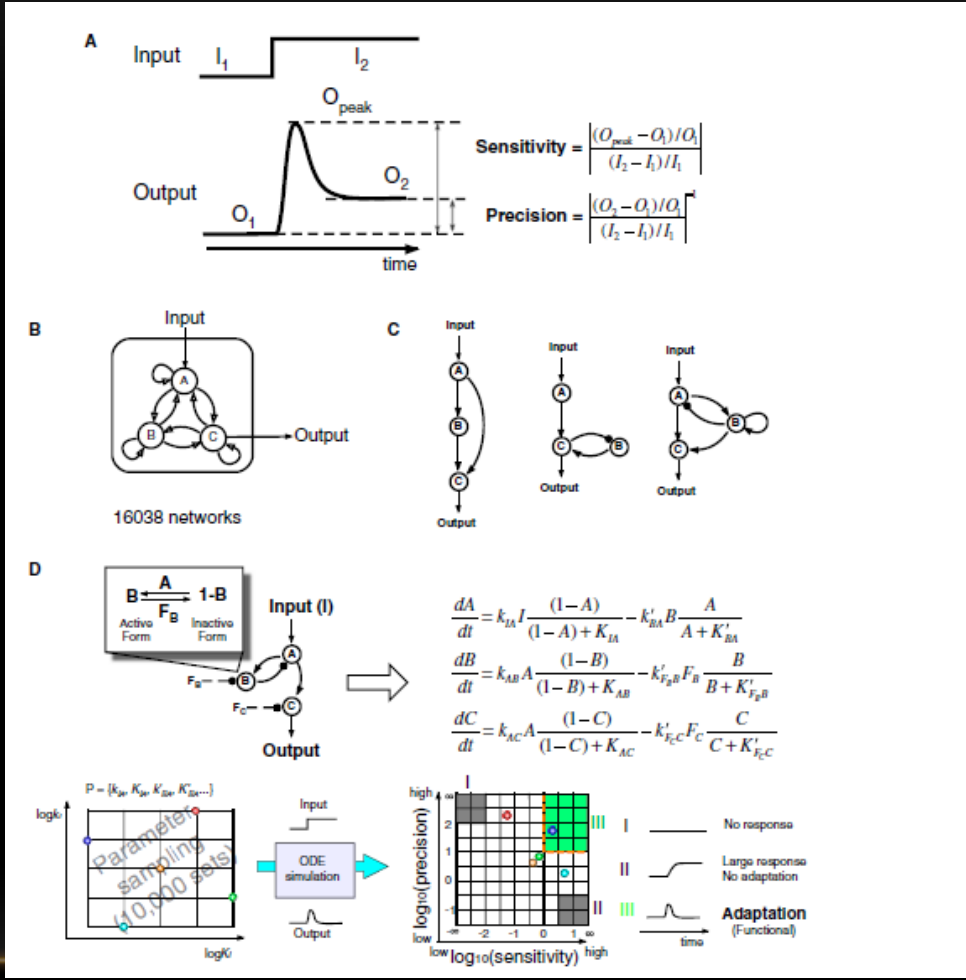
IDENTIFY ROBUSTNESS AND VULNERABILITY HOT-SPOTS IN THE SIGNALING NETWORK



Node Function	Robust Nodes (%)	Vulnerable Nodes (%)
Kinase	28.08	8
Transcription Factor	12.35	6.67
Receptor	9.87	8
Adaptor	6.79	4
Phosphatases	1.54	12
Apoptosis	3.08	22.67

Design principle of a signaling network???

NETWORK MOTIFS FOR BIOCHEMICAL ADAPTABILITY

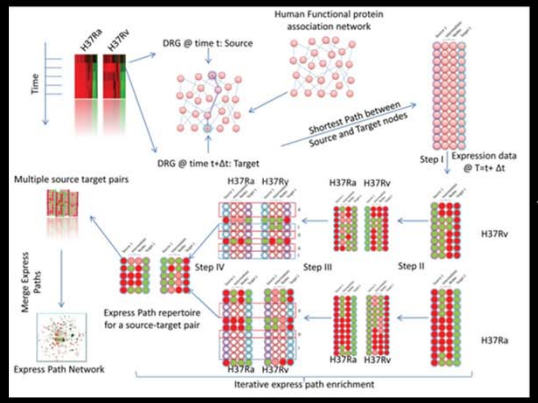


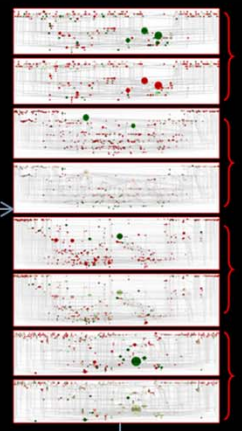
CONCLUSIONS

- Any approach that tries to understand biology in an inclusive manner is Systems Biology
- Experiments are equally important as computation and/or theory
- Use of different computational, theoretical (Mathematical) or statistical approach depends on the specific question being asked
- Network based approaches help understand intricacies of biological regulation
- Complexity of individual biological events bound to get unraveled in more detail, simple tools will not be enough to understand biology.

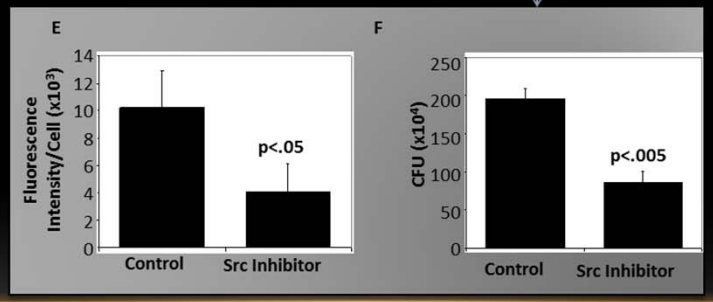
NETWORK BASED APPROACHES???

EPA





Identify Src as key Response regulator



Src inhibition selectively Targets virulent infections

ADVANCE CELL AND MOLECULAR BIOLOGY
(ICG601)