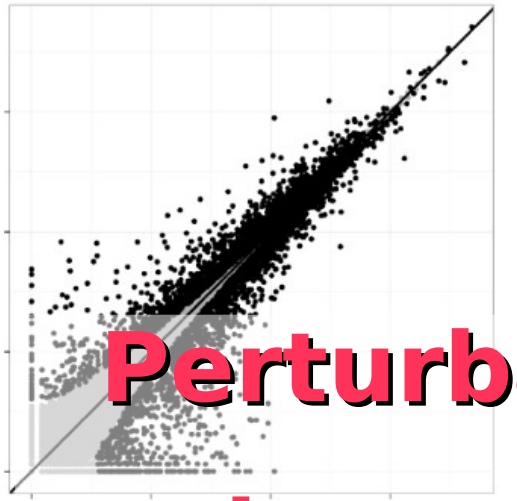


Perturbations of PIP3 signalling trigger a global remodelling of mRNA landscape and reveal a transcriptional feedback loop

Nicolas Le Novère
Babraham Institute,
n.lenovere@gmail.com

Babraham Institute

Babraham Institute

eulife

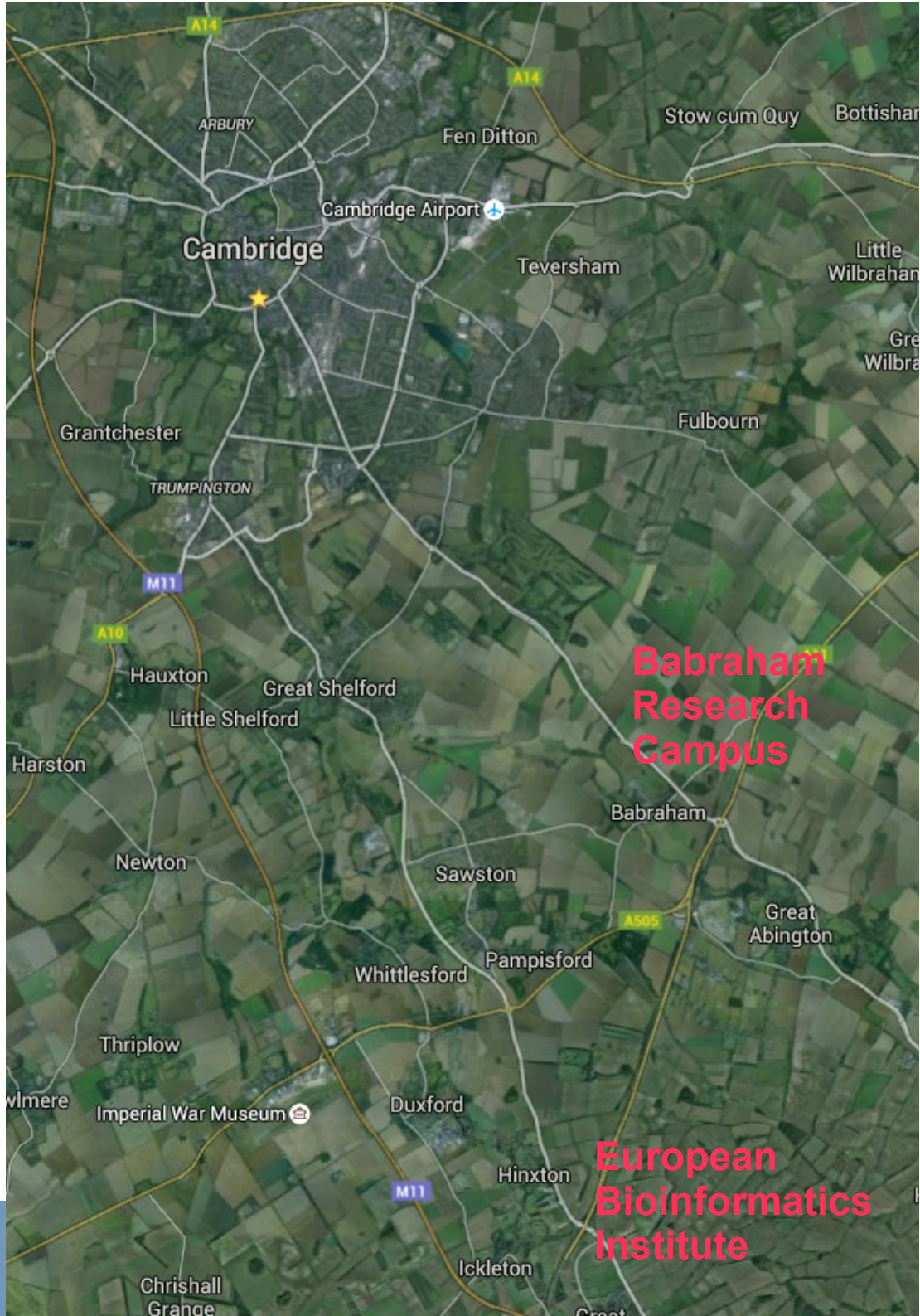
BBSRC
bioscience for the future

Programmes in:

- **Signalling**
- **Immunology**
- **Epigenetics**
- **Nuclear Dynamics**

Plateforms including:

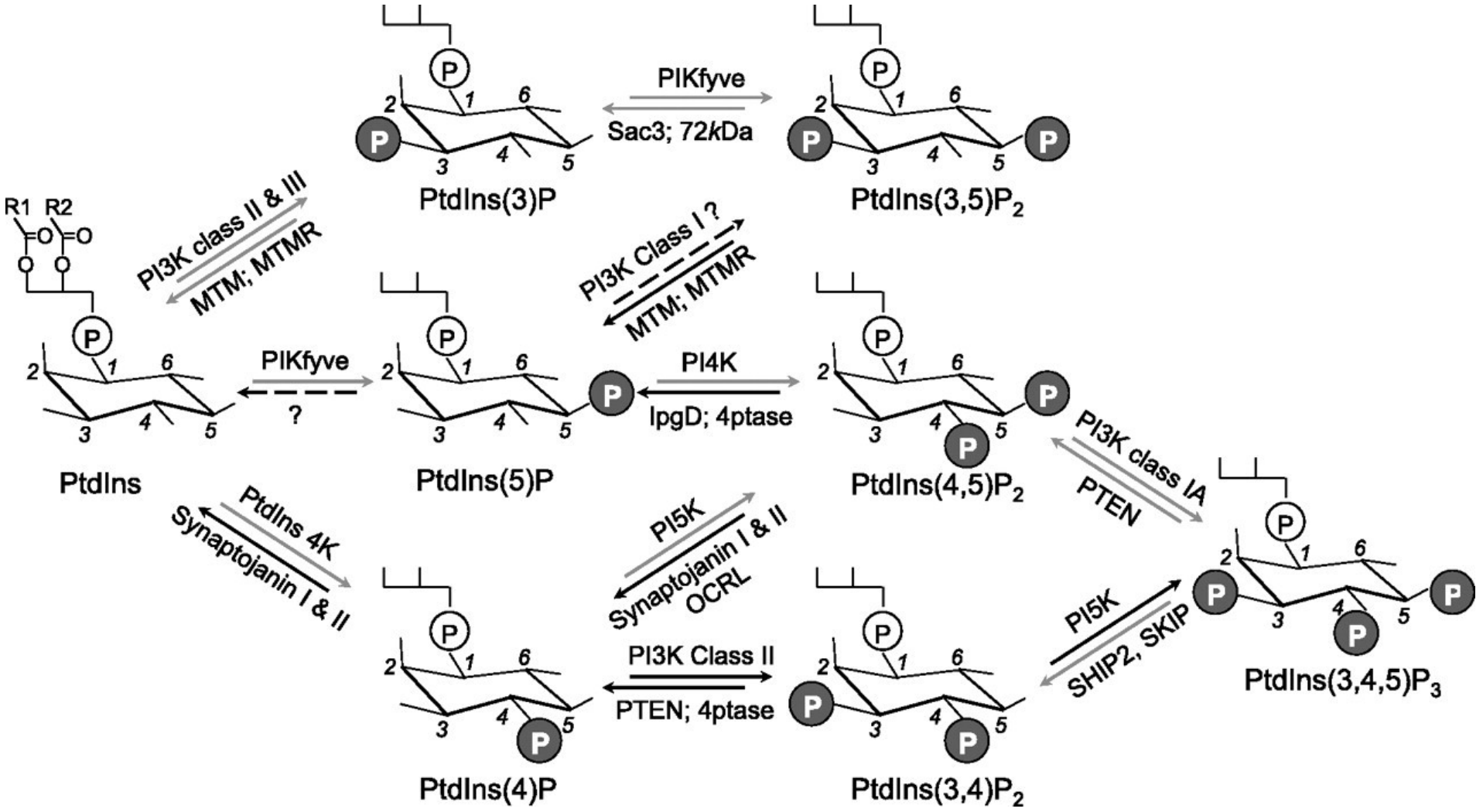
- **Bioinformatics**
- **Imaging**
- **FACS**
- **Lipidomics**
- **Mouse facilities**
- **Sequencing**

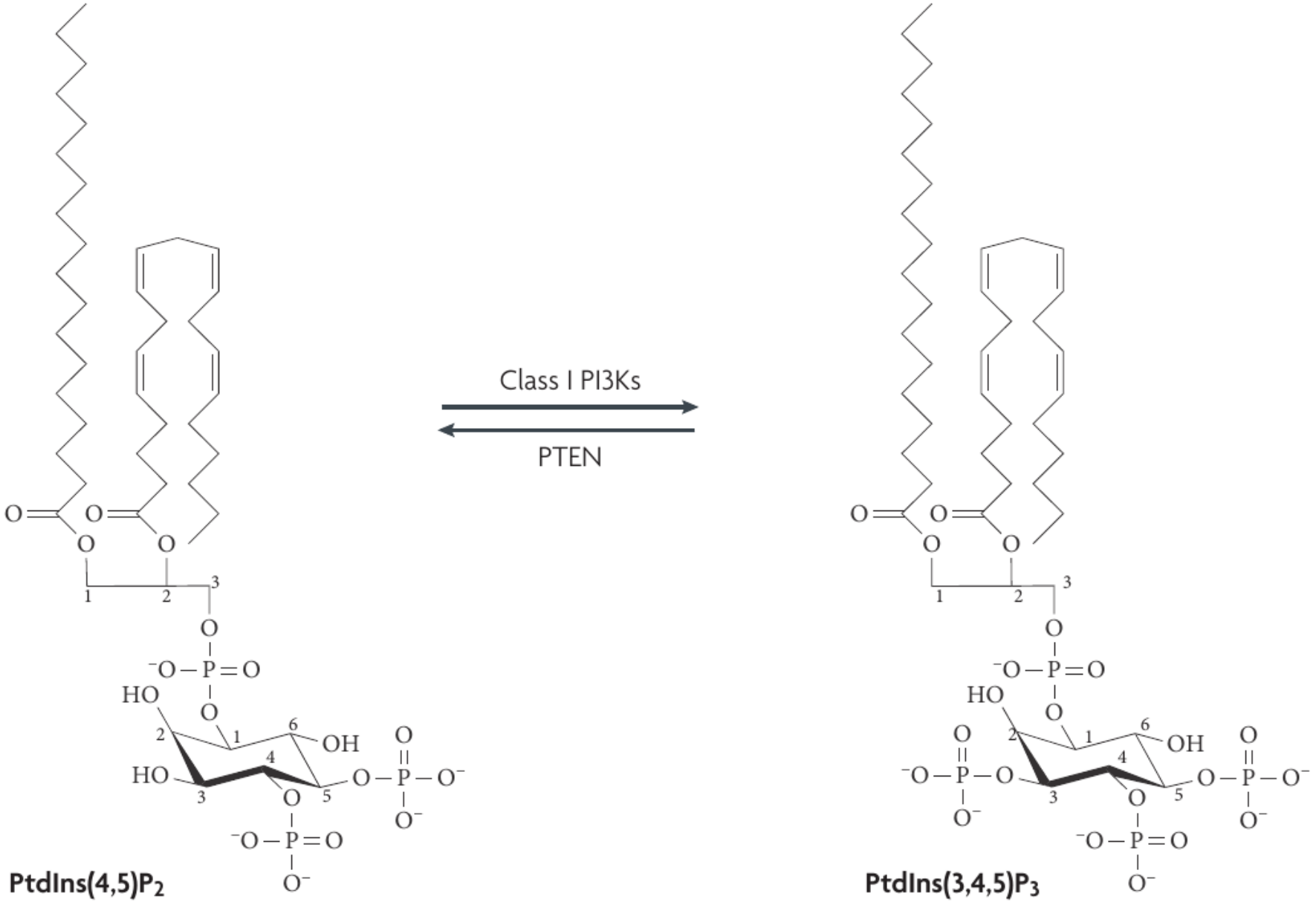


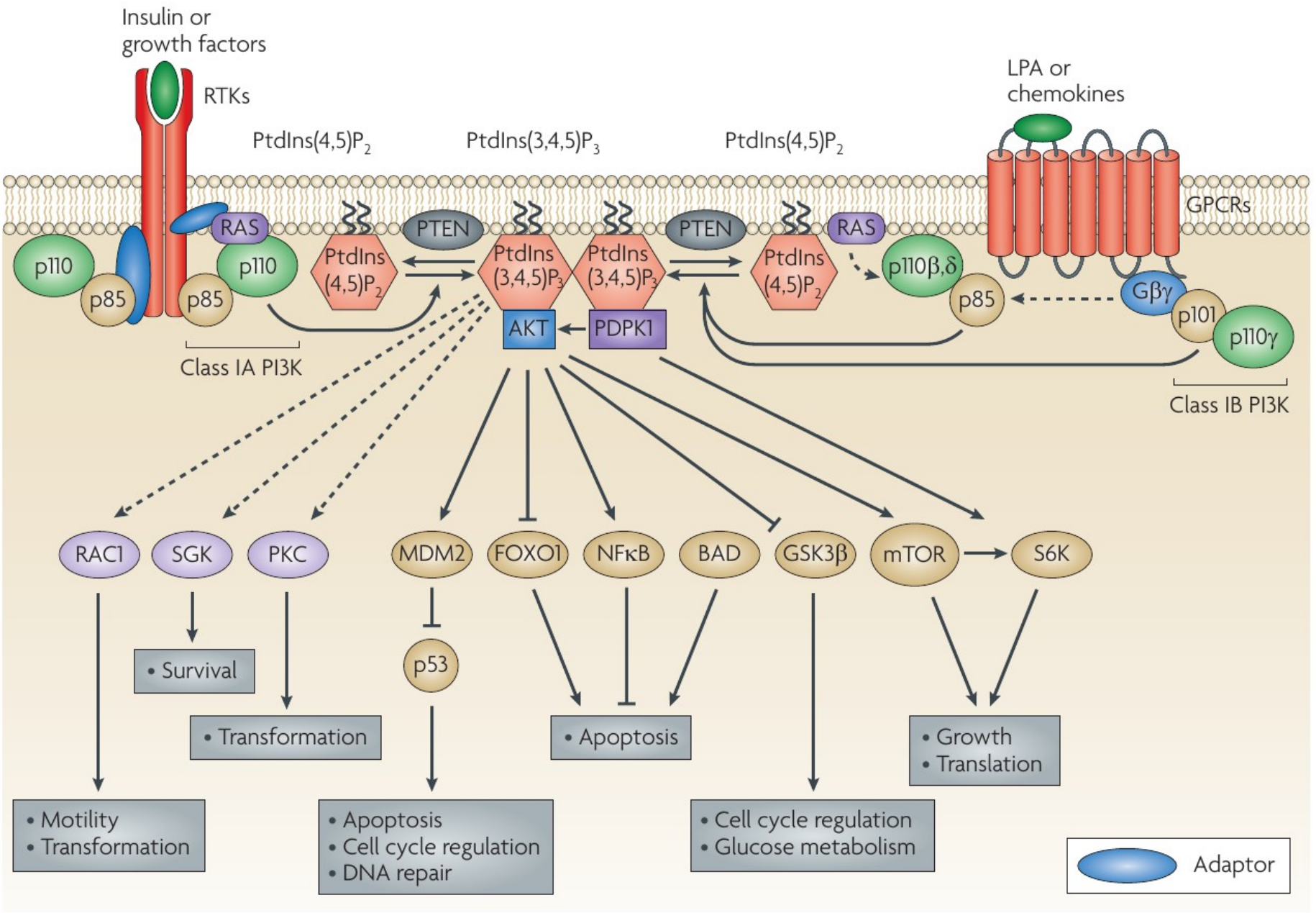
The Babraham Institute and the (phospho)lipids

- Discovery of the liposome
  Bangham AD, Standish MM, Watkins JC (1965) Diffusion of univalent ions across the lamellae of swollen phospholipids. *J Mol Biol* 13, 238–252.
- Discovery of IP3 signalling
  Berridge MJ and Irvine RF (1984) Inositol trisphosphate, a novel second messenger in cellular signal transduction. *Nature* 312, 315 – 321
- Phosphorylation of PIP2 into PIP3 by PI3K
  P.T. Hawkins, T.R. Jackson, L.R. Stephens (1992) Platelet-derived growth factor stimulates synthesis of PtdIns(3,4,5)P3 by activating a PtdIns(4,5)P2 3-OH kinase. *Nature* 358, 157-159
- PIP3-dependent activation of PKB by PDK1
  Stokoe D, Stephens LR, Copeland T, Gaffney PR, Reese CB, Painter GF, Holmes AB, McCormick F, Hawkins PT (1997) Dual role of phosphatidylinositol-3,4,5-trisphosphate in the activation of protein kinase B. *Science* 277, 567-570.
  Stephens L.R., Anderson K., Stokoe D., Erdjument-Bromage H., Painter G.F., Holmes A.B., Gaffney P.R.J., Reese C.B., McCormick F., Tempst P., Coadwell J., Hawkins P.T. (1998) Protein Kinase B Kinases That Mediate Phosphatidylinositol 3,4,5-Trisphosphate-Dependent Activation of Protein Kinase B. *Science* 279, 710-714

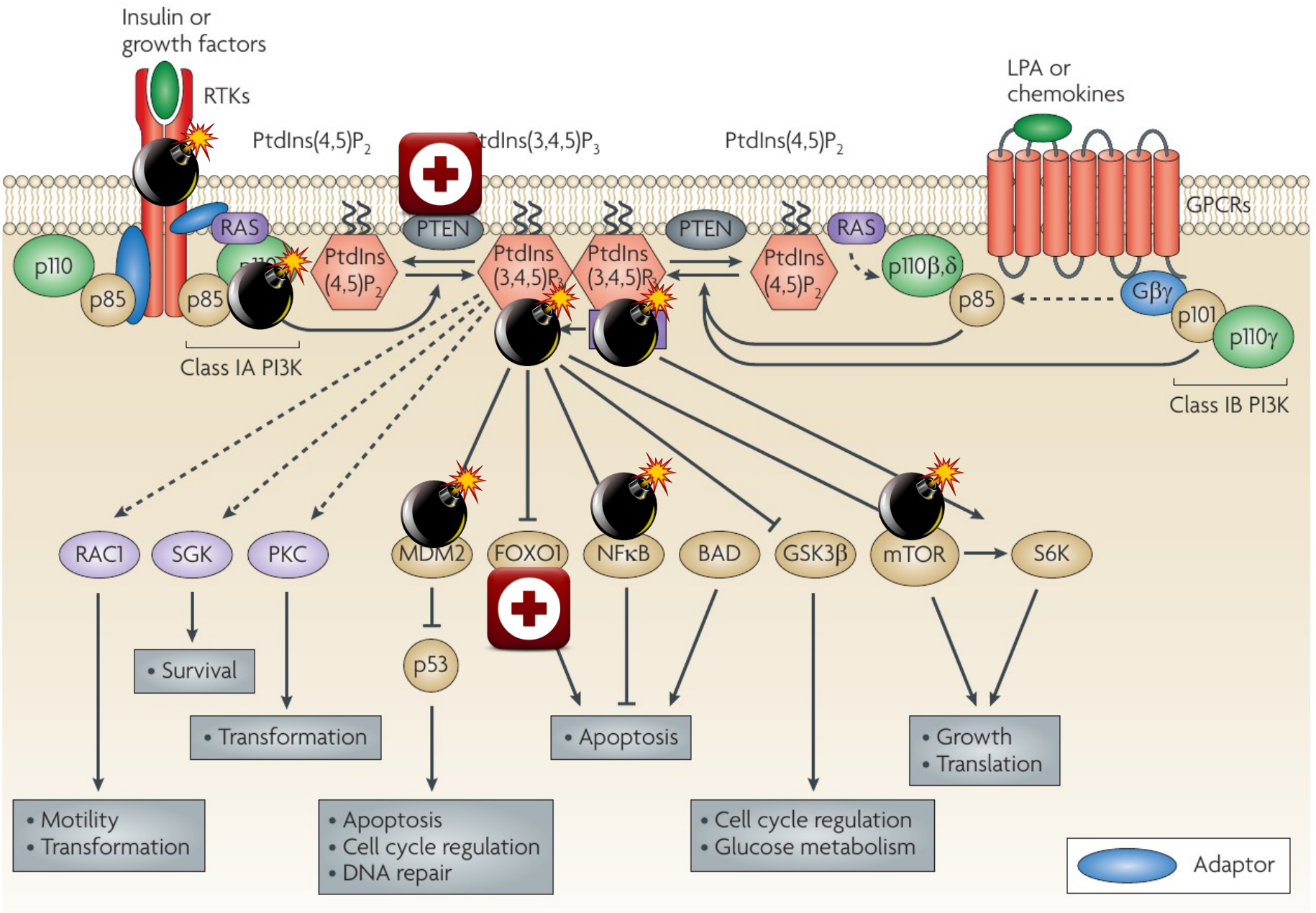
The phosphoinositides



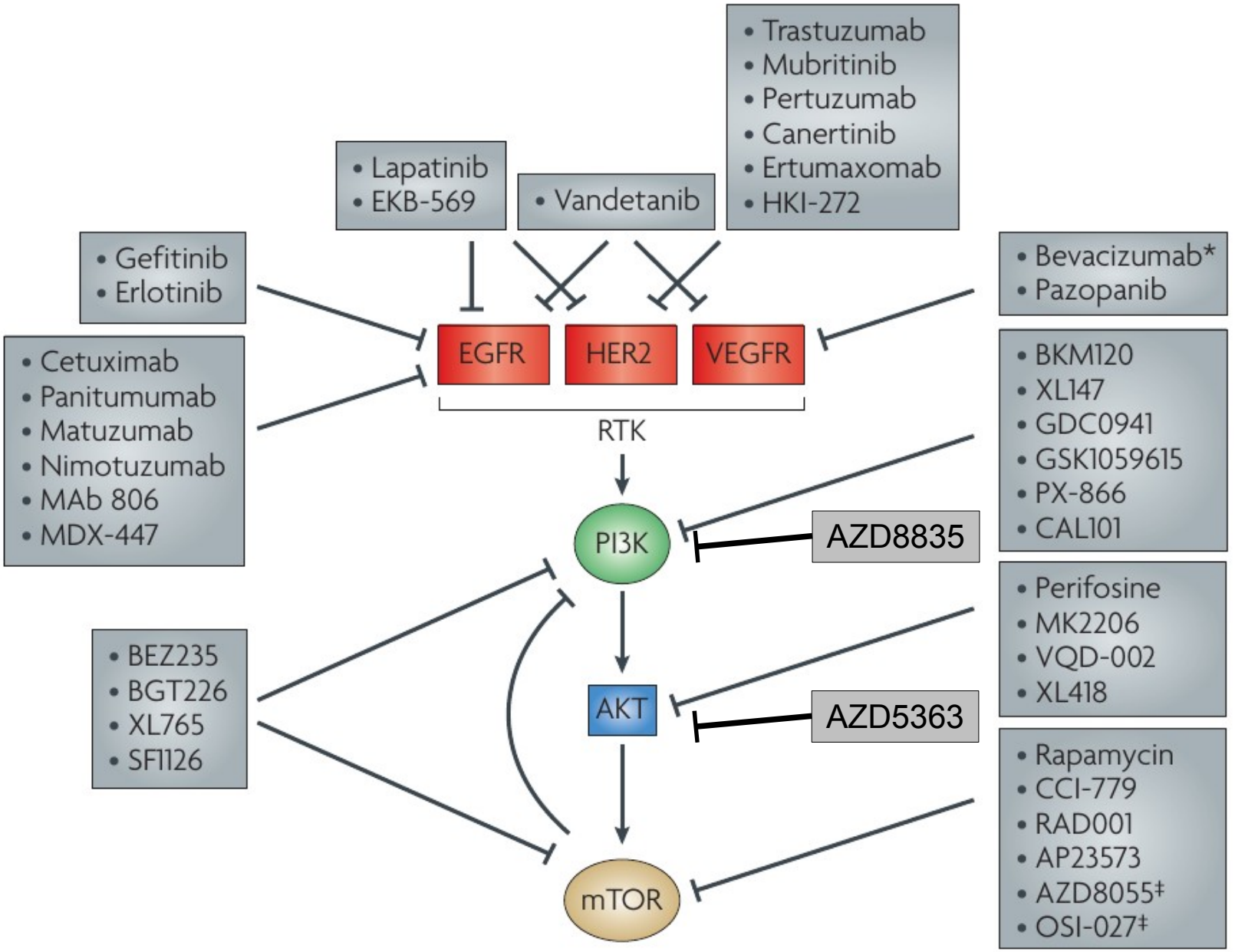

Class I PI3Ks
PTEN
PtdIns(4,5)P2
PtdIns(3,4,5)P3



Liu *et al* (2009) *Nat Rev Drug Discov*

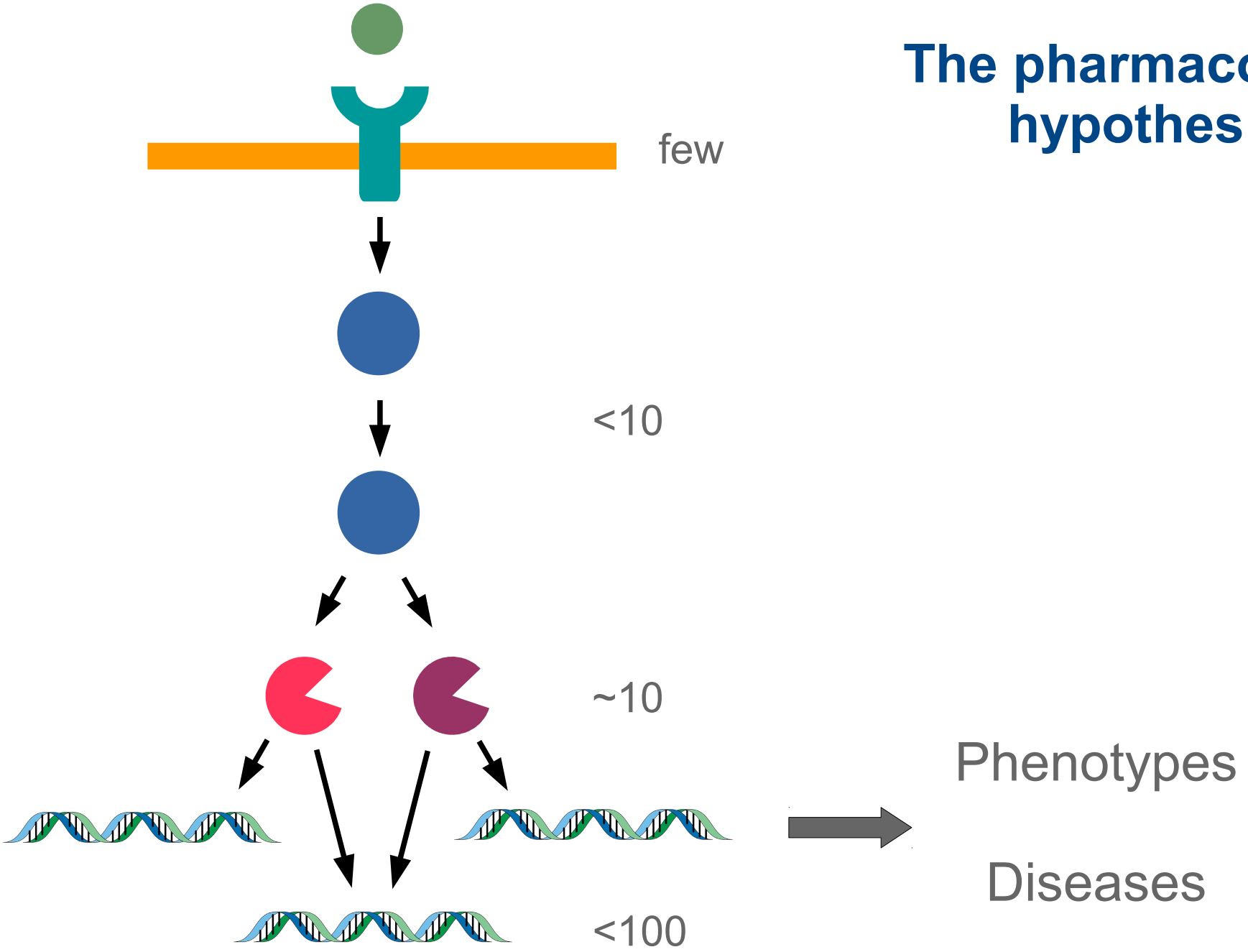


Liu *et al* (2009) *Nat Rev Drug Discov*

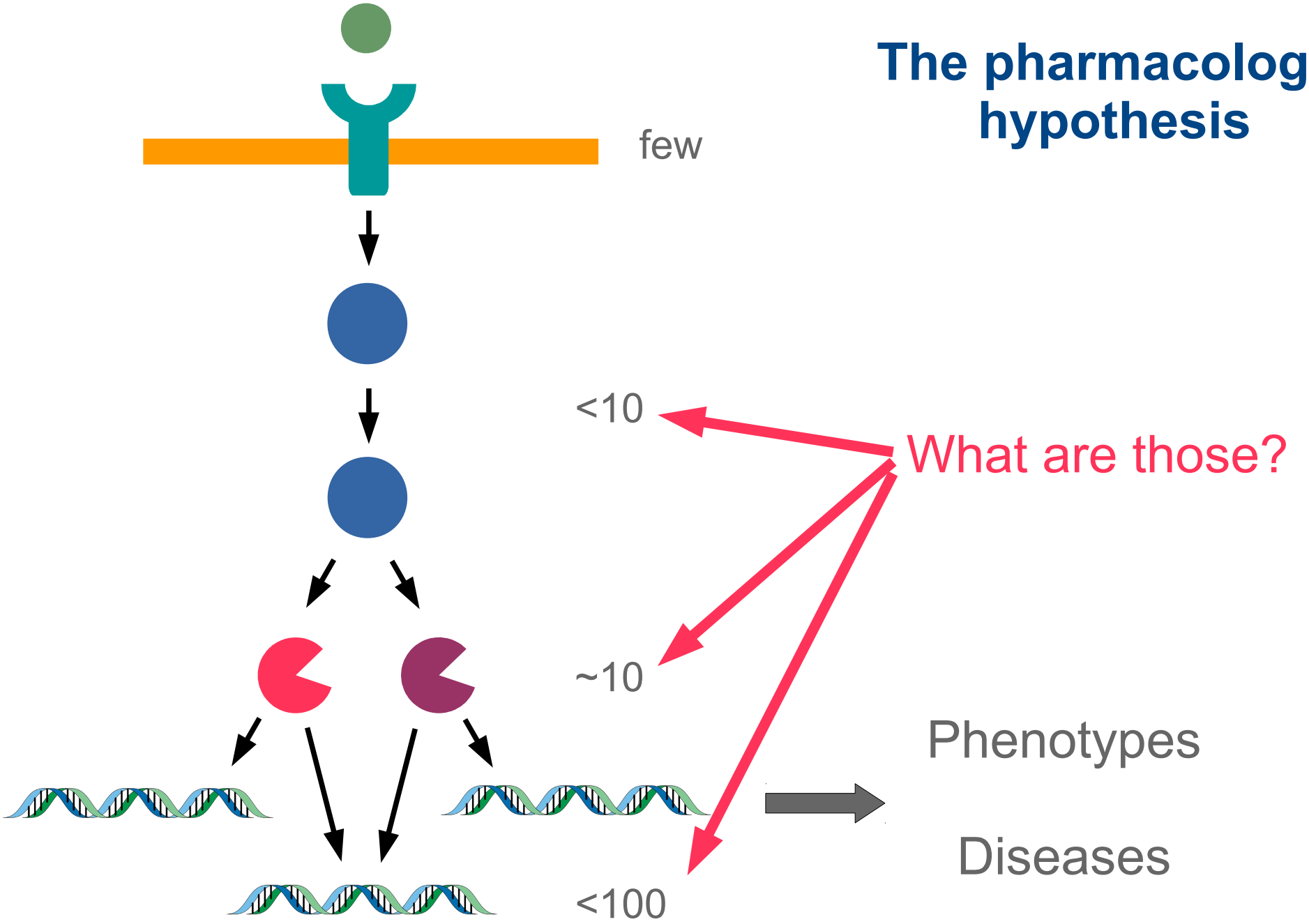


Liu *et al* (2009) *Nat Rev Drug Discov*

The pharmacology hypothesis

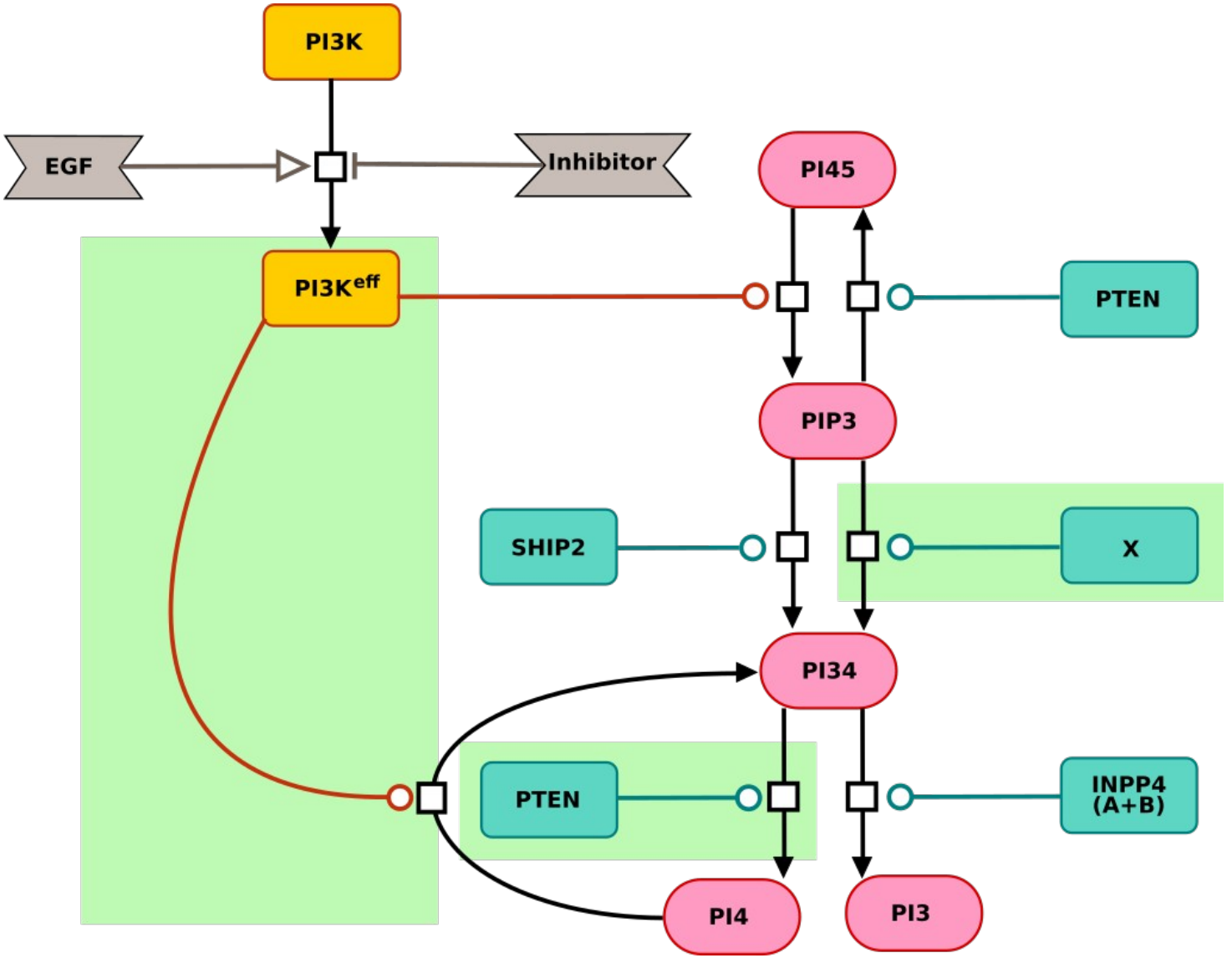


The pharmacology hypothesis



	experimental	computational
signalling	lipidomics (phosphoinositides mass-spectrometry)	chemical kinetic modelling
gene expression	transcriptomics (messenger RNA RNA-Seq)	clustering enrichment promoter analysis

Lipidomics and chemical kinetic modelling



Vladimir Kiselev

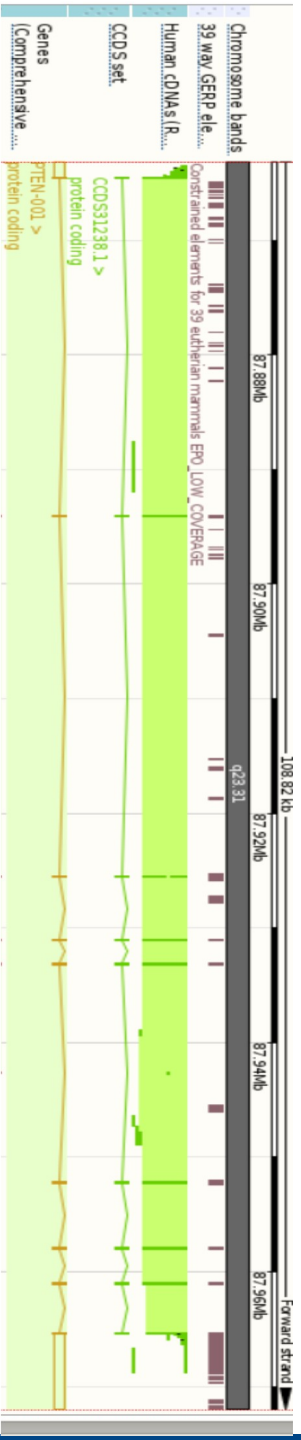
Véronique Juvin

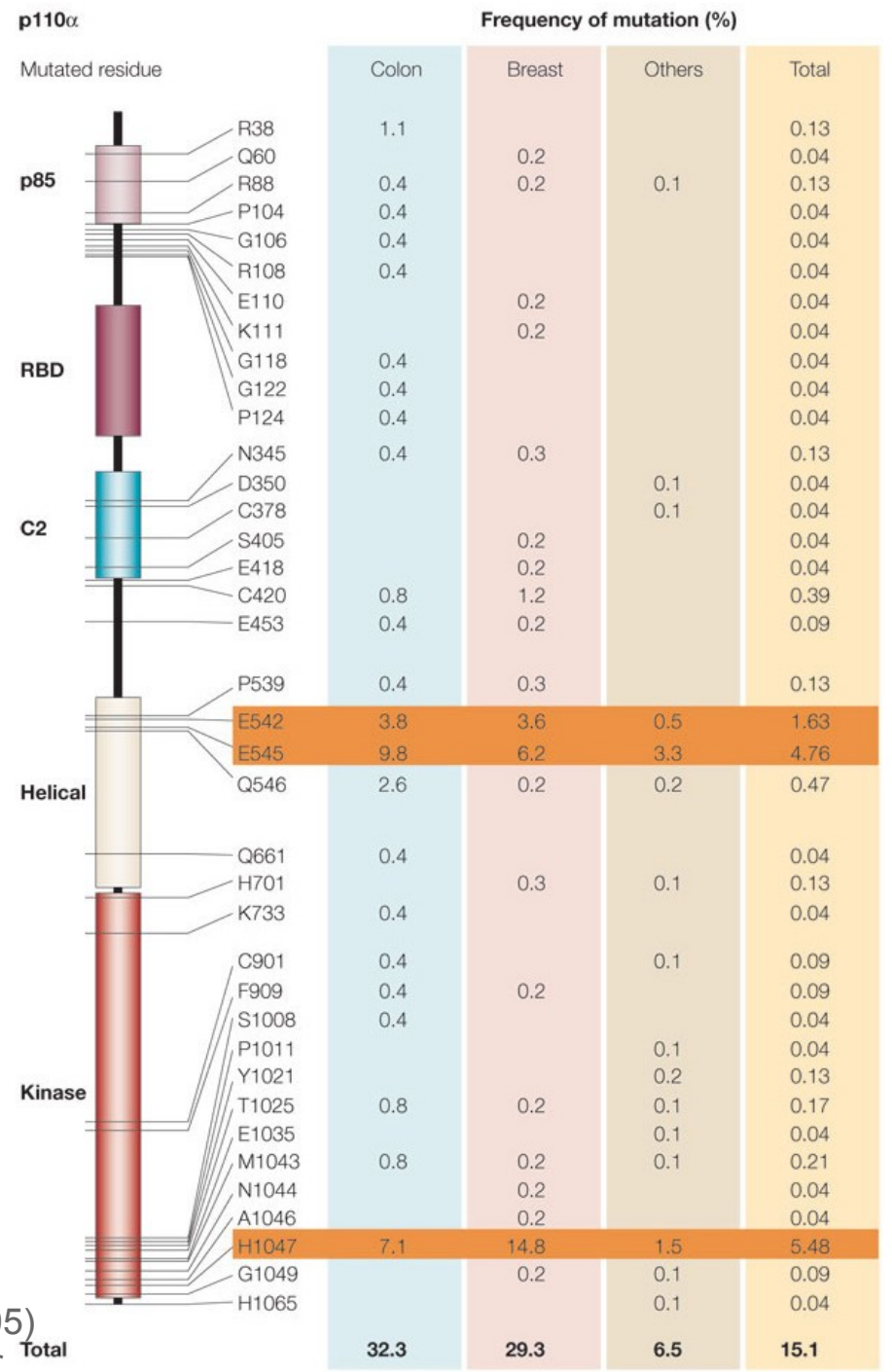
Kiselev, Juvin *et al* (2015) *Nucleic Acids Res*

PTEN

PI3K

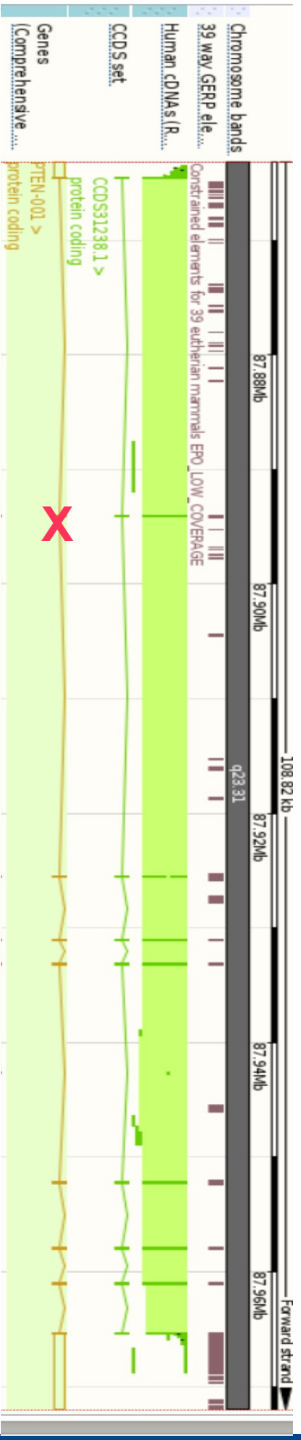
Isogenic MCF10a cell lines





Bader *et al* (2005)
Nat Rev Cancer

Babraham Institute



PTEN

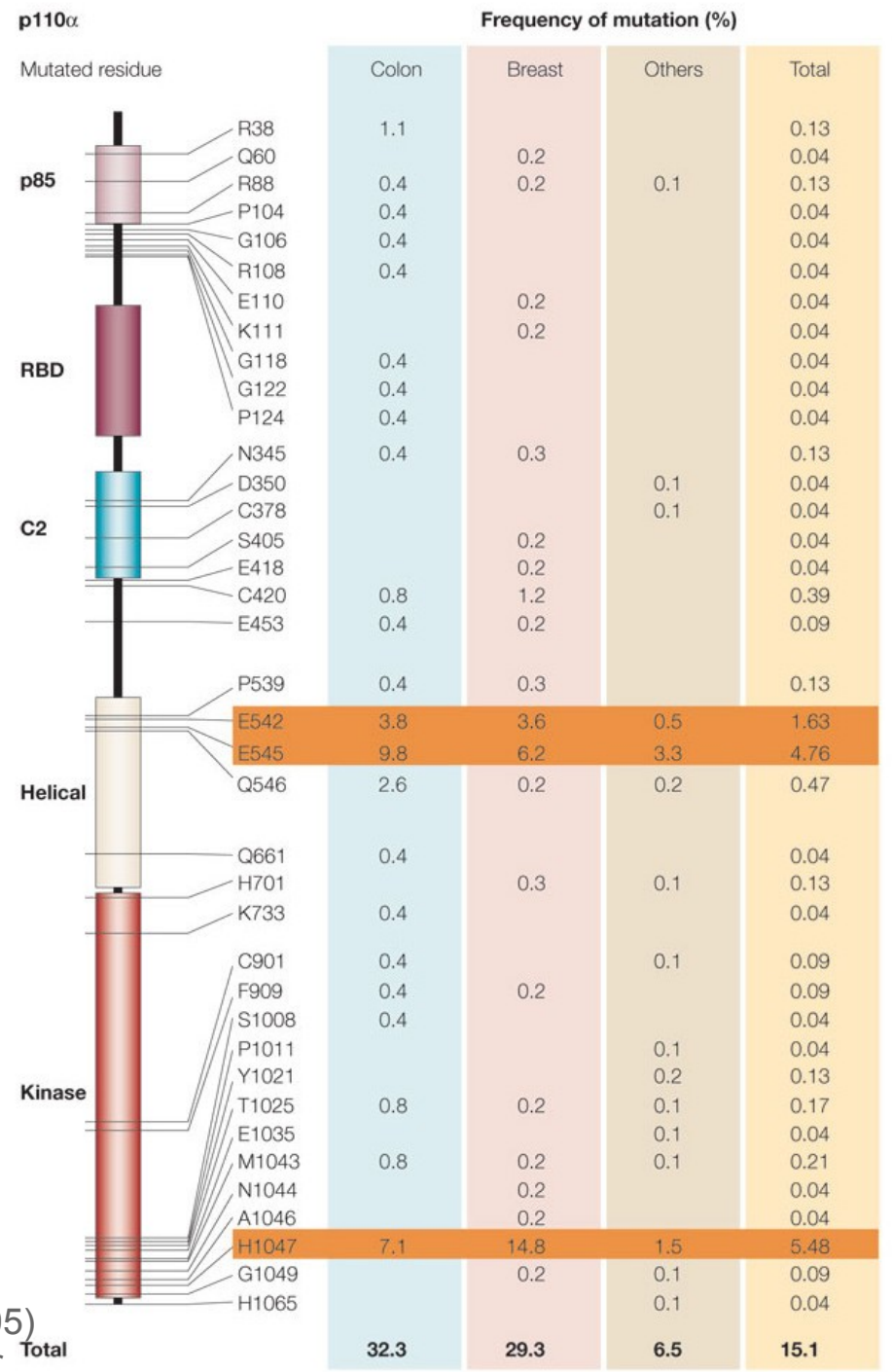
PI3K

Isogenic MCF10a cell lines

X

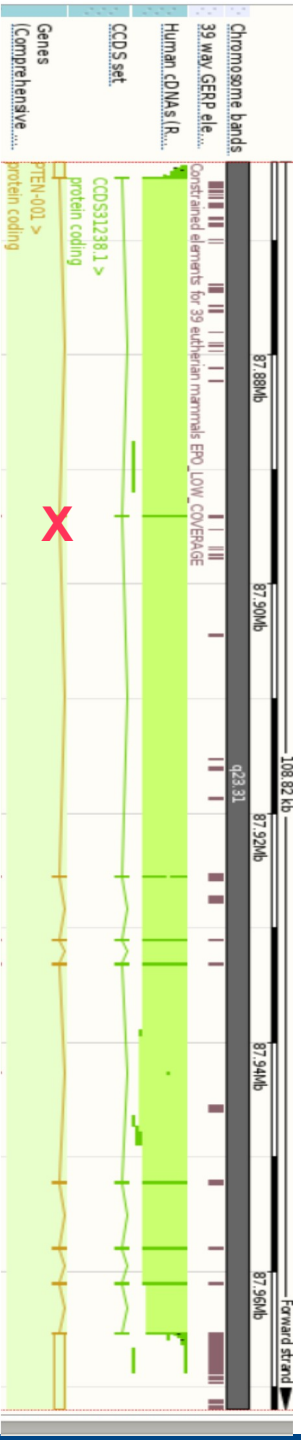
phosphatase domain

inactivating bi-allelic mutation



Bader *et al* (2005)
Nat Rev Cancer

Babraham Institute



PTEN

PI3K

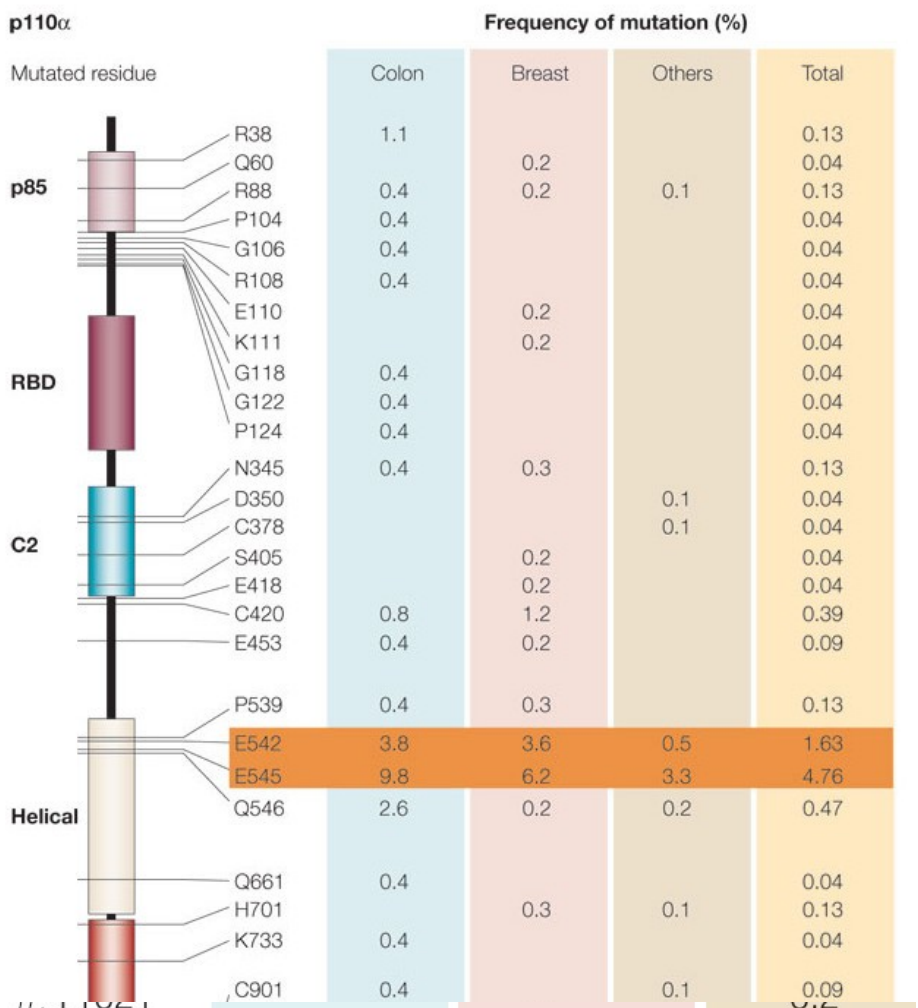
Isogenic MCF10a cell lines

X

phosphatase domain

inactivating bi-allelic mutation

activating mono-allelic mutation →



Mutated residue	Frequency of mutation (%)			
	Colon	Breast	Others	Total
R38	1.1			0.13
Q60		0.2		0.04
R88	0.4	0.2	0.1	0.13
P104	0.4			0.04
G106	0.4			0.04
R108	0.4			0.04
E110		0.2		0.04
K111		0.2		0.04
G118	0.4			0.04
G122	0.4			0.04
P124	0.4			0.04
N345	0.4	0.3		0.13
D350			0.1	0.04
C378			0.1	0.04
S405		0.2		0.04
E418		0.2		0.04
C420	0.8	1.2		0.39
E453	0.4	0.2		0.09
P539	0.4	0.3		0.13
E542	3.8	3.6	0.5	1.63
E545	9.8	6.2	3.3	4.76
Q546	2.6	0.2	0.2	0.47
Q661	0.4			0.04
H701		0.3	0.1	0.13
K733	0.4			0.04
C901	0.4		0.1	0.09
T1025	0.8	0.2		0.1
E1035				0.1
M1043	0.8	0.2		0.1
N1044		0.2		
A1046		0.2		
H1047	7.1	14.8		1.5
G1049		0.2		0.1
H1065				0.1

Babraham Institute

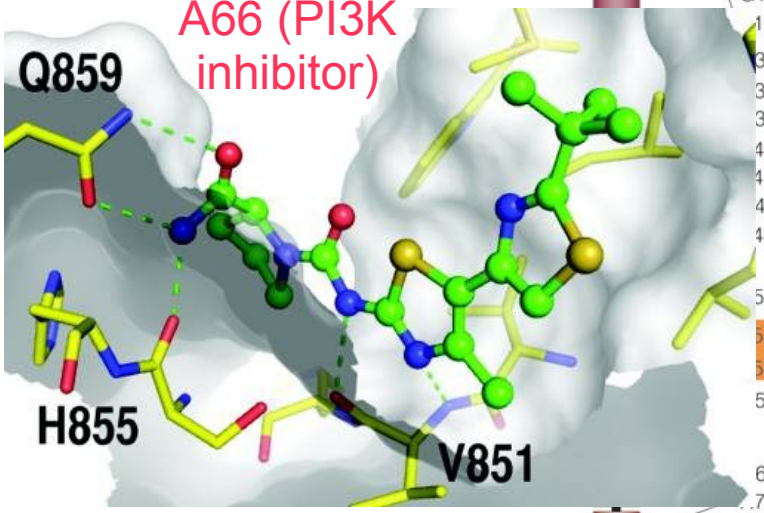
PTEN **PI3K**

Isogenic MCF10a cell lines

x

phosphatase domain

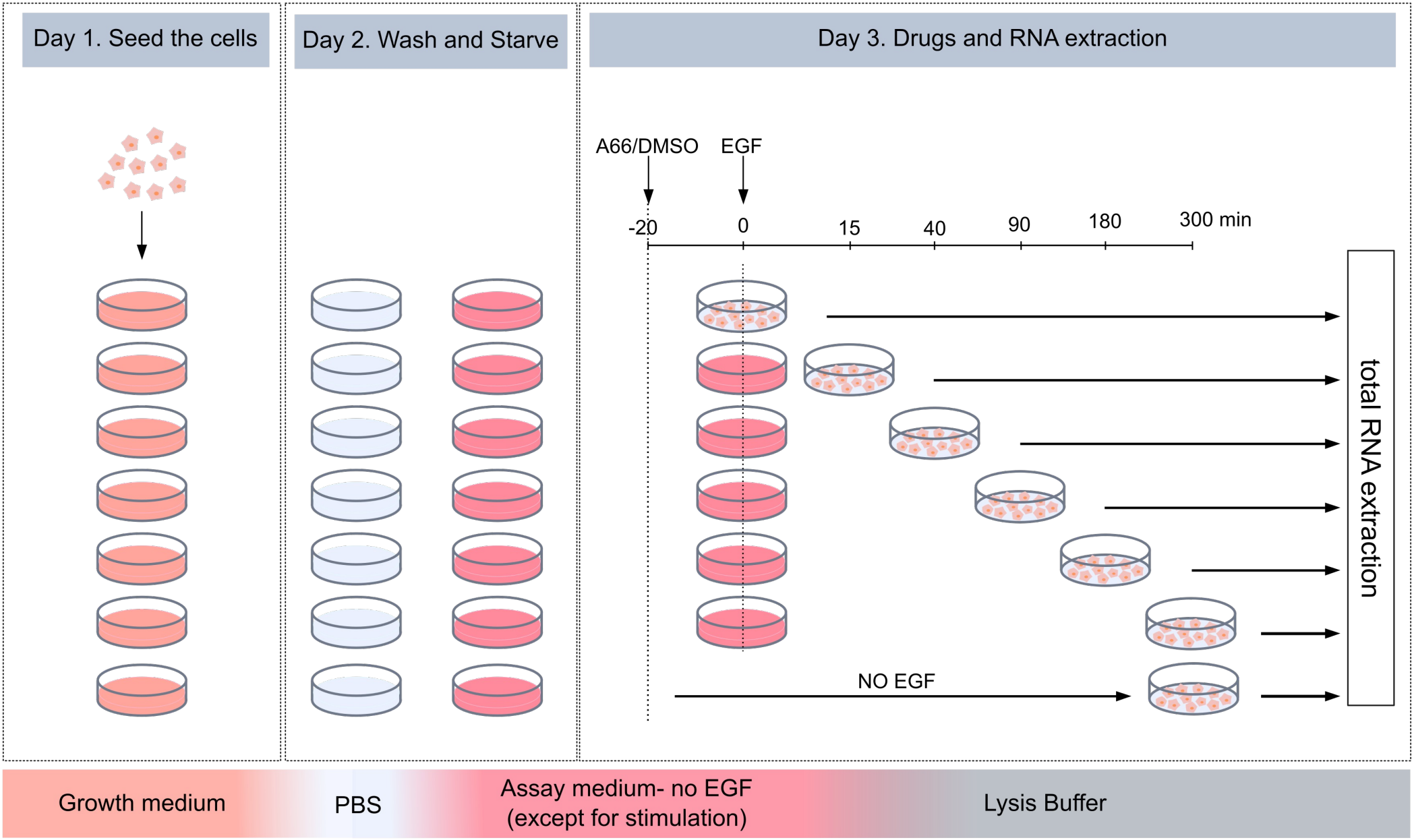
inactivating bi-allelic mutation



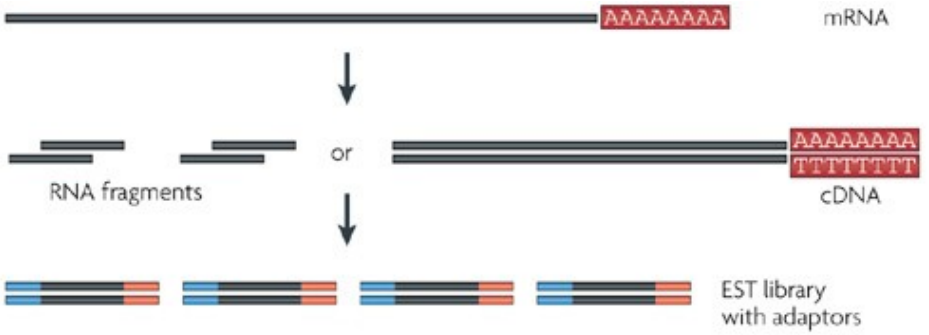
activating mono-allelic mutation →

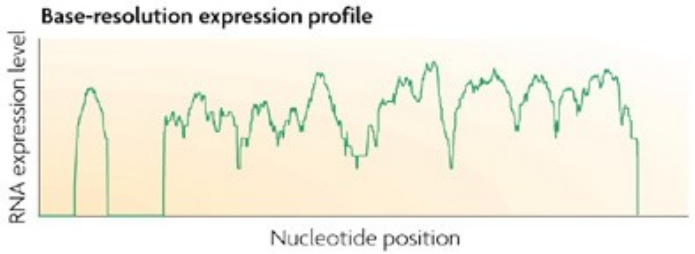
p110α Mutated residue		Frequency of mutation (%) Colon	Breast	Others	Total
p85	R38	1.1			0.13
	Q60		0.2		0.04
	R88	0.4	0.2	0.1	0.13
	P104	0.4			0.04
	G106	0.4			0.04
	R108	0.4			0.04
	E110		0.2		0.04
	K111		0.2		0.04
RBD	G118	0.4			0.04
	G122	0.4			0.04
	124	0.4			0.04
	345	0.4	0.3		0.13
	350			0.1	0.04
	378			0.1	0.04
	405		0.2		0.04
	418		0.2		0.04
	420	0.8	1.2		0.39
	453	0.4	0.2		0.09
	539	0.4	0.3		0.13
	542	3.8	3.6	0.5	1.63
	545	9.8	6.2	3.3	4.76
	546	2.6	0.2	0.2	0.47
	661	0.4			0.04
	701		0.3	0.1	0.13
	K733	0.4			0.04
	C901	0.4		0.1	0.09
Kinase	T1025	0.8	0.2		0.1
	E1035				0.1
	M1043	0.8	0.2		0.1
	N1044		0.2		
	A1046		0.2		
	H1047	7.1	14.8		1.5
	G1049		0.2		0.1
	H1065				0.1

Babraham Institute

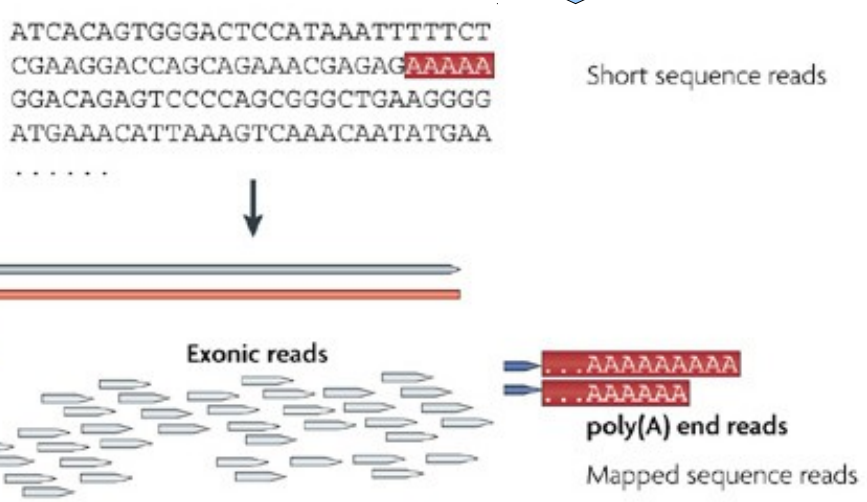

Day 1. Seed the cells
Day 2. Wash and Starve
Day 3. Drugs and RNA extraction
A66/DMSO
EGF
-20
0
15
40
90
180
300 min
total RNA extraction
NO EGF
Growth medium
PBS
Assay medium- no EGF
(except for stimulation)
Lysis Buffer

Measuring gene expression with RNA-Seq

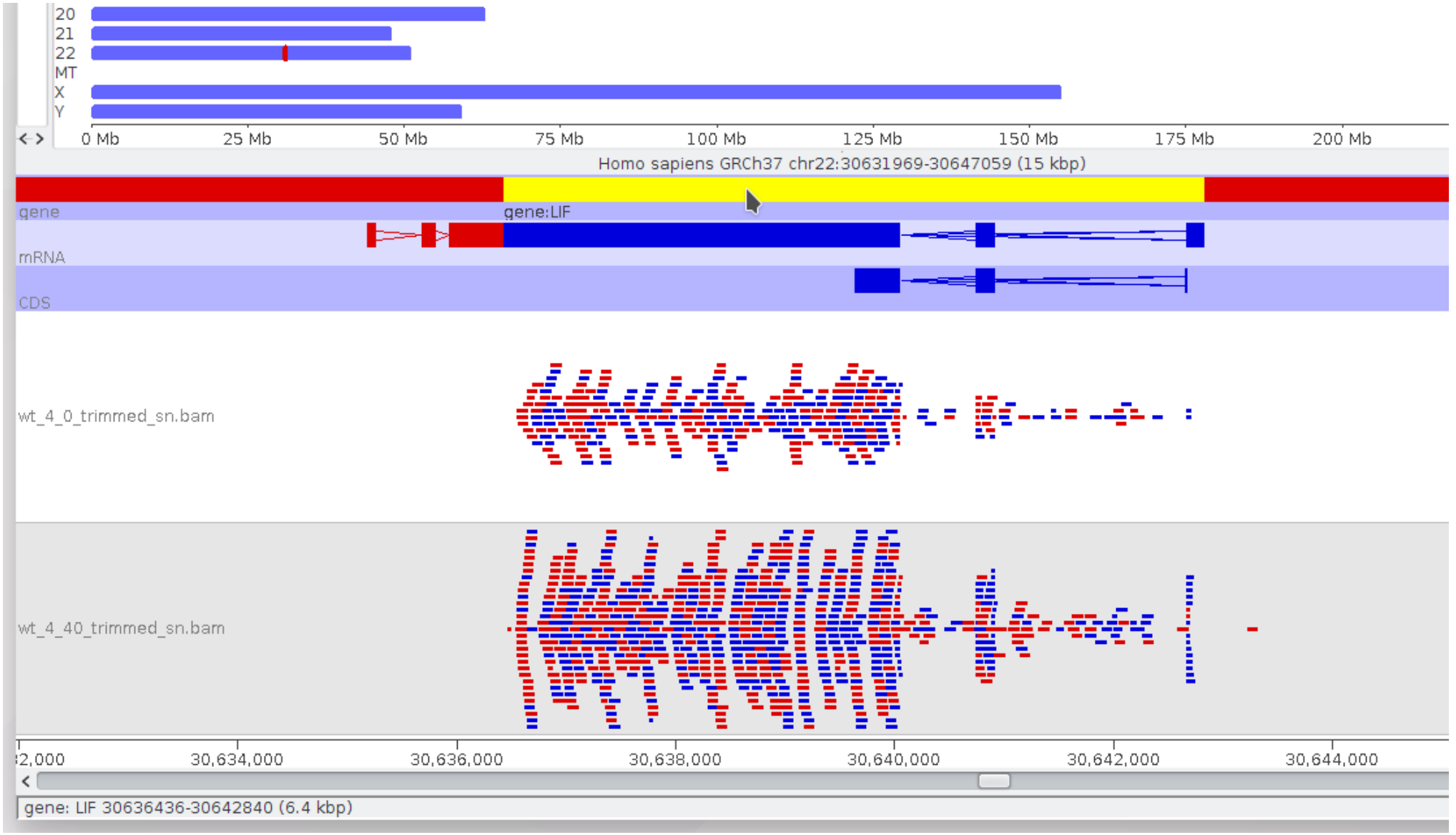




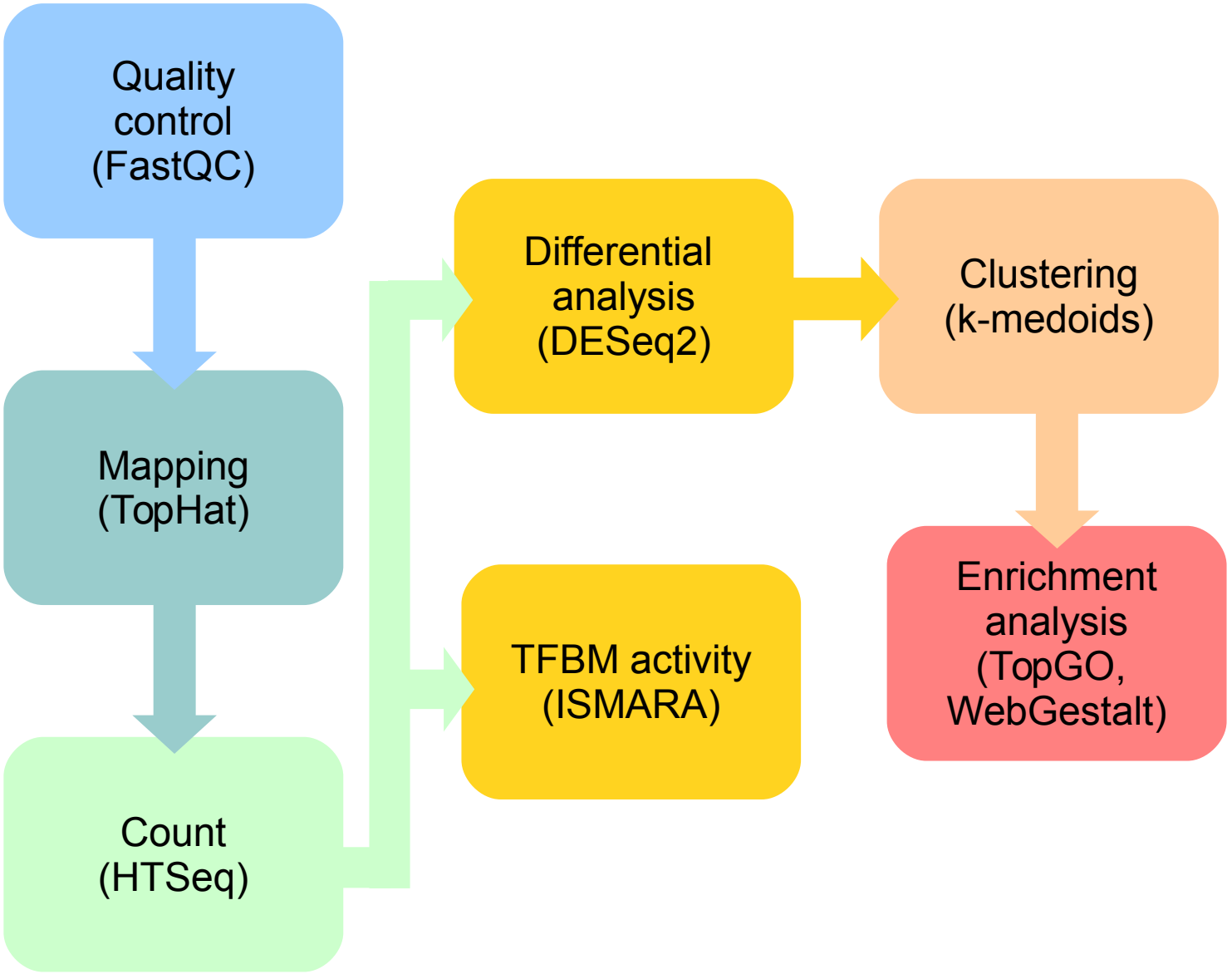
normalised gene expression	
Gene 1	1.65
Gene 2	2037
Gene 3	126
Gene 4	89.5
Gene 5	32045

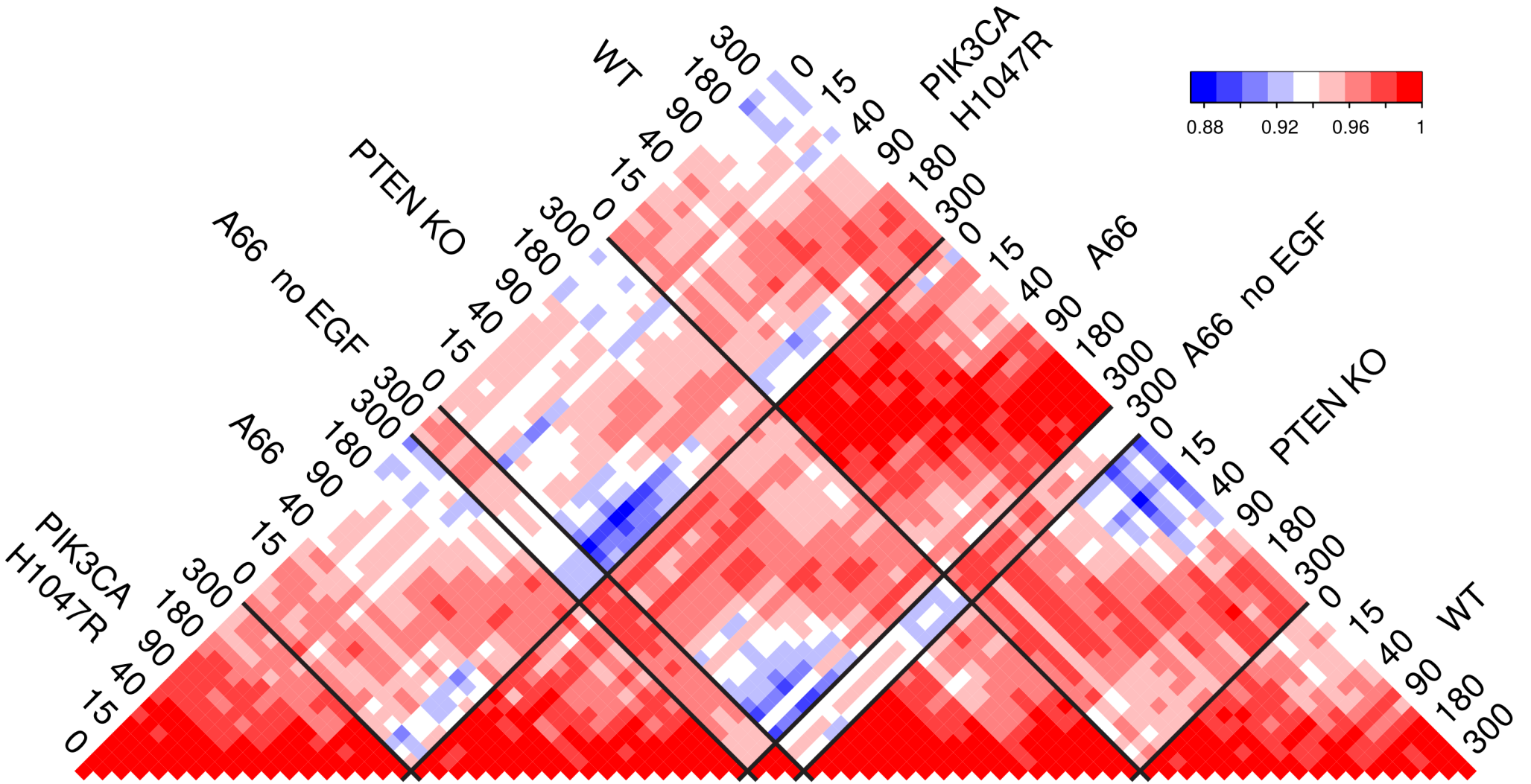


Wang *et al* (2009) *Nat Rev Genet*



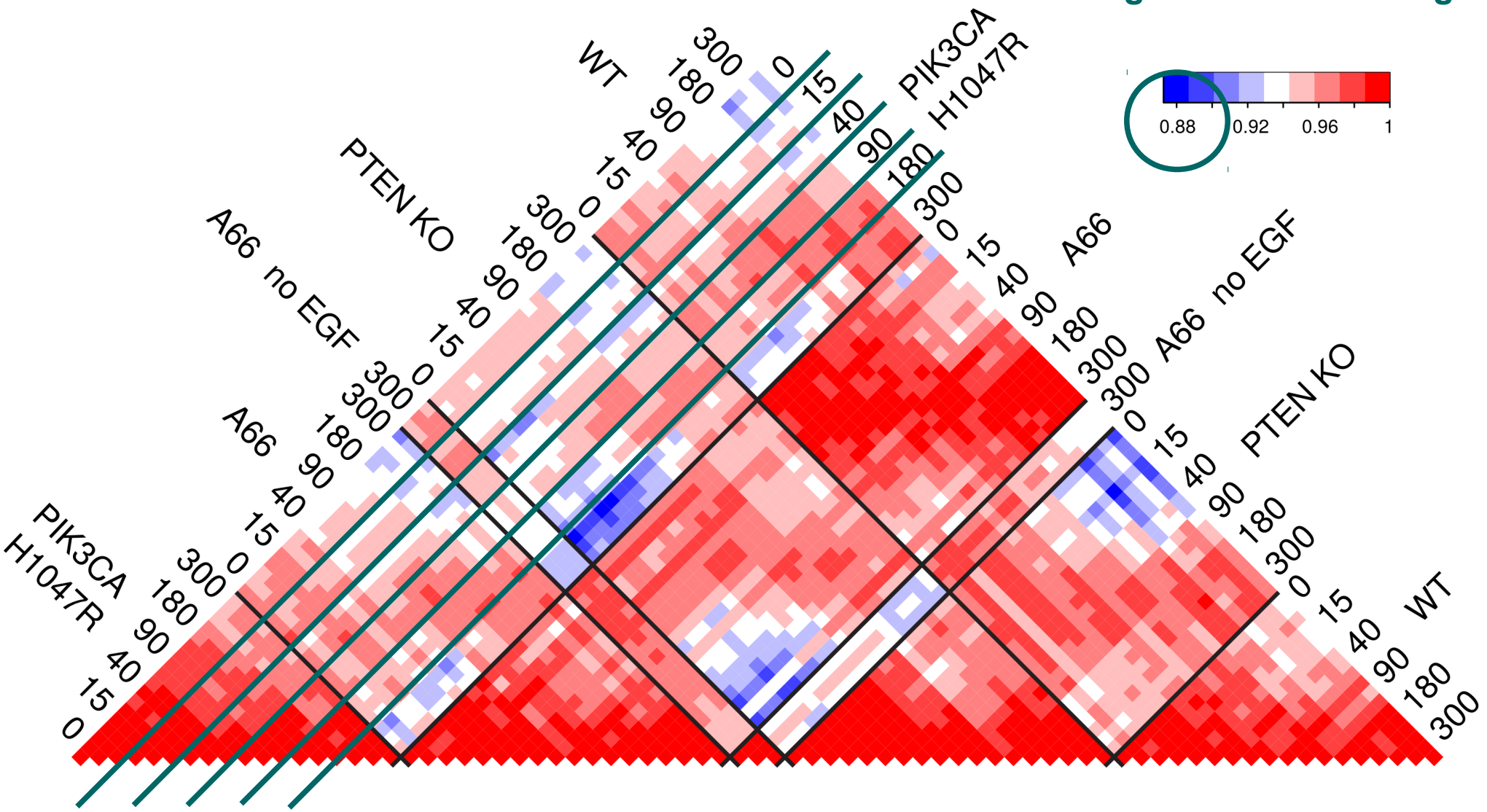
Babraham Institute


Quality control (FastQC)
Mapping (TopHat)
Count (HTSeq)
Differential analysis (DESeq2)
TFBM activity (ISMARA)
Clustering (k-medoids)
Enrichment analysis (TopGO, WebGestalt)



Babraham Institute

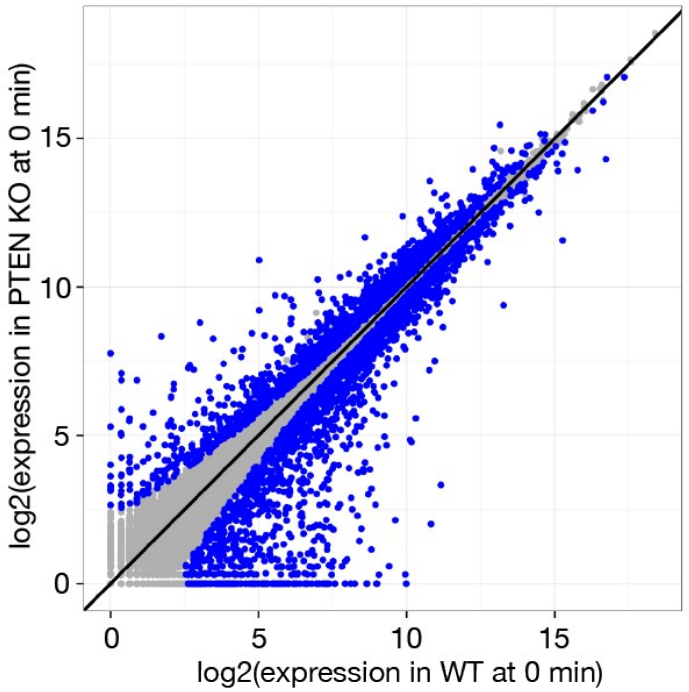
Most genes do not change

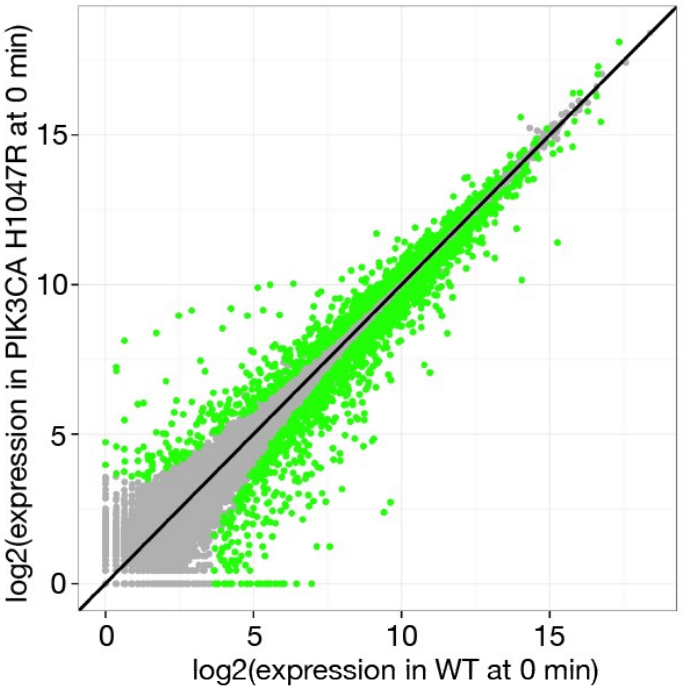


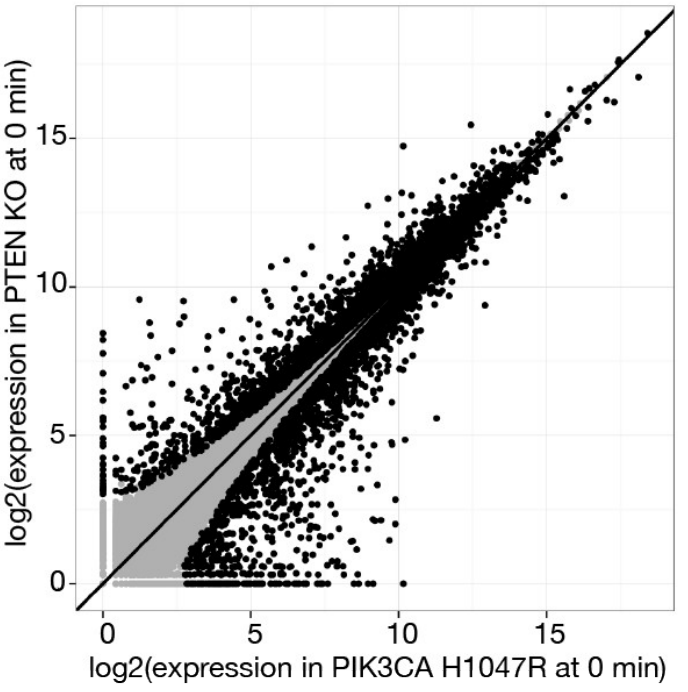
replicates are OK

Babraham Institute

But quite a few are affected nevertheless







4725 genes affected by A66

1543 genes affected by H1047R

2244 genes affected by PTEN-/-

The butterfly effect in cancer: A single base mutation can remodel the cell

Jonathan R. Hart[a], Yaoyang Zhang[b], Lujian Liao[b], Lynn Ueno[a], Lisa Du[a], Marloes Jonkers[a], John R. Yates III[b], and Peter K. Vogt[a,1]

Departments of [a]Molecular and Experimental Medicine and [b]Chemical Physiology, The Scripps Research Institute, La Jolla, CA 92037

Contributed by Peter K. Vogt, December 15, 2014 (sent for review August 11, 2014)

We have compared the proteome, transcriptome, and metabolome of two cell lines: the human breast epithelial line MCF-10A and its mutant descendant MCF-10A-H1047R. These cell lines are derived from the same parental stock and differ by a single amino acid substitution (H1047R) caused by a single nucleotide change in one allele of the PIK3CA gene, which encodes the catalytic subunit p110α of PI3K (phosphatidylinositol 3-kinase). They are considered isogenic. The H1047R mutation of PIK3CA is one of the most frequently encountered somatic cancer-specific mutations. In MCF-10A, this mutation induces an extensive cellular reorganization that far exceeds the known signaling activities of PI3K. The changes are highly diverse, with examples in structural protein levels, the DNA repair machinery, and sterol synthesis. Gene set enrichment analysis reveals a highly significant concordance of the genes differentially expressed in MCF-10A-H1047R cells and the established protein and RNA signatures of basal breast cancer. No such concordance was found with the specific gene signatures of other histological types of breast cancer. Our data document the power of a single base mutation, inducing an extensive remodeling of the cell toward the phenotype of a specific cancer.

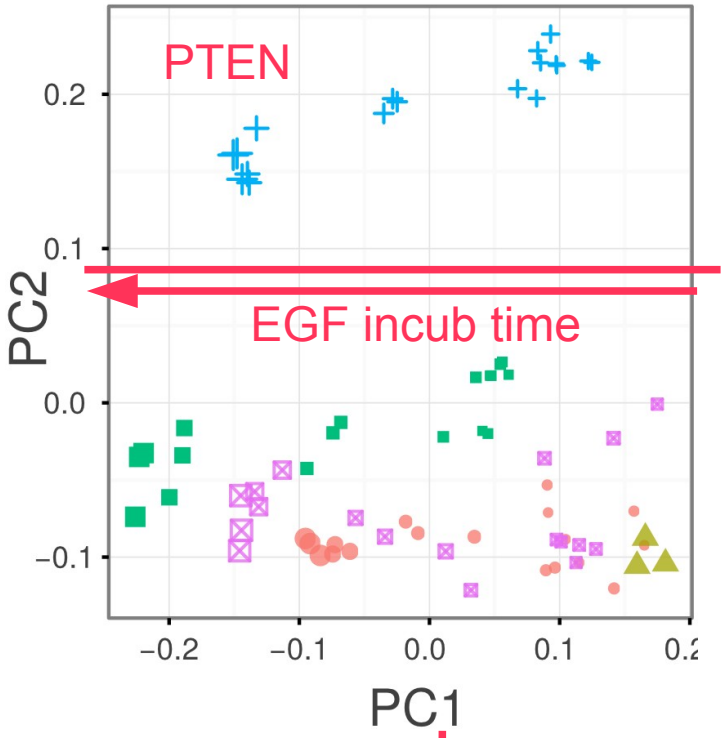
RNAseq | SILAC | knock-in | molecular signature | basal breast cancer

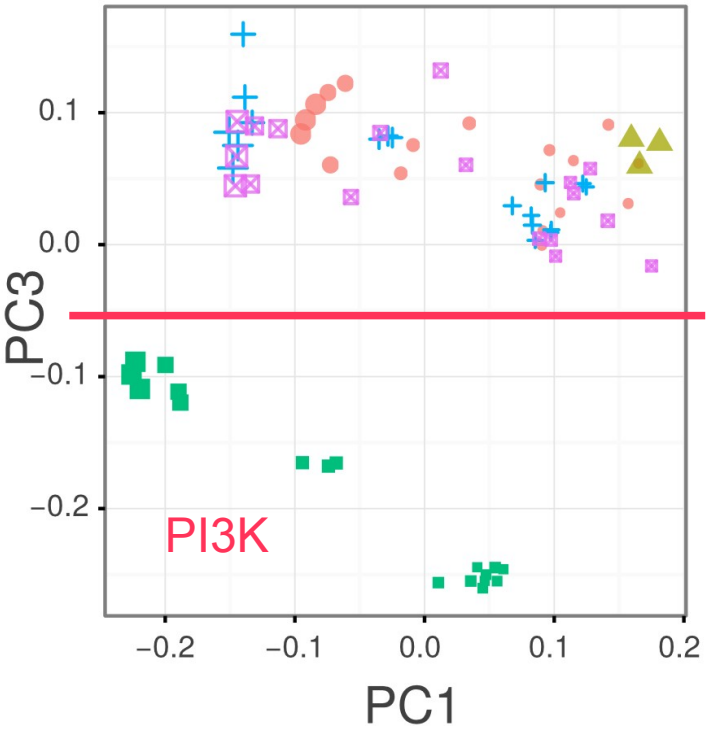
MCF-10A and MCF-10A-H1047R can grow in chemically defined, serum-free medium, facilitating the amino acid substitutions required by SILAC and avoiding the variability introduced by the use of serum in the culture medium (7, 9).

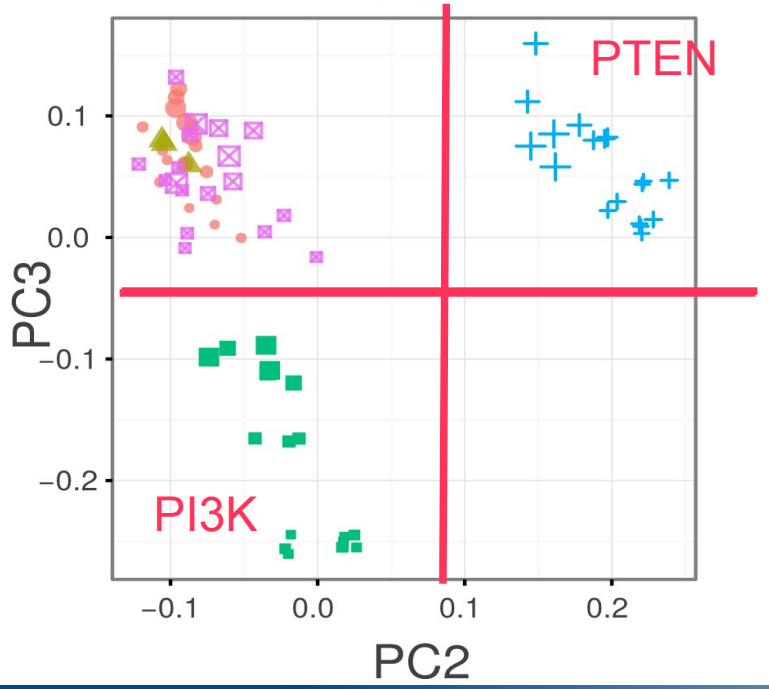
The changes induced in protein and RNA expression by the H1047R mutation document a comprehensive reorganization of the cell, including a shift of the expression patterns toward the signature of basal breast cancer.

Results

Genetic Comparison of the MCF-10A and MCF-10A-H1047R Cell Lines. MCF-10A and MCF-10A-H1047R are considered isogenic, except for the knock-in mutation of H1047R in one allele of PIK3CA. However, during the creation of the H1047R knock-in or in the course of the subsequent culture, other mutations in cancer-relevant genes could have been introduced or selected for. To investigate this possibility, both cell lines were studied by whole-exome sequencing. The procedures used for exome sequencing are described in *SI Materials and Methods*. This sequence information was used to determine variant SNPs (single-nucleotide polymorphisms) and insertions and deletions, as well as copy number variations. Variants that are significantly different between the two cell lines are shown in Table S1. Other





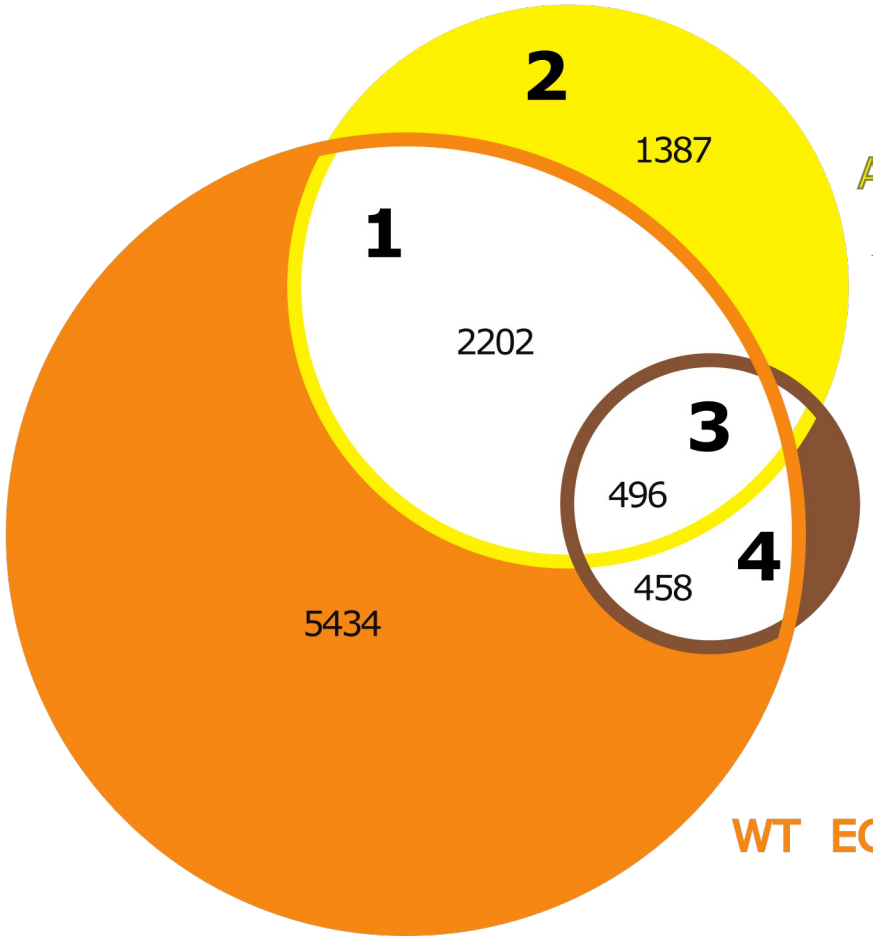


Time, min
- 0
- 100
- 200
- 300

Condition
- WT
- A66
- A66 no EGF
- PIK3CA H1047R
- PTEN KO

Babraham Institute

Effect of acute PI3K inhibition



A66 no EGF, T=0 vs T=300

A66 effect

WT vs A66 EGF-stimulation time-courses

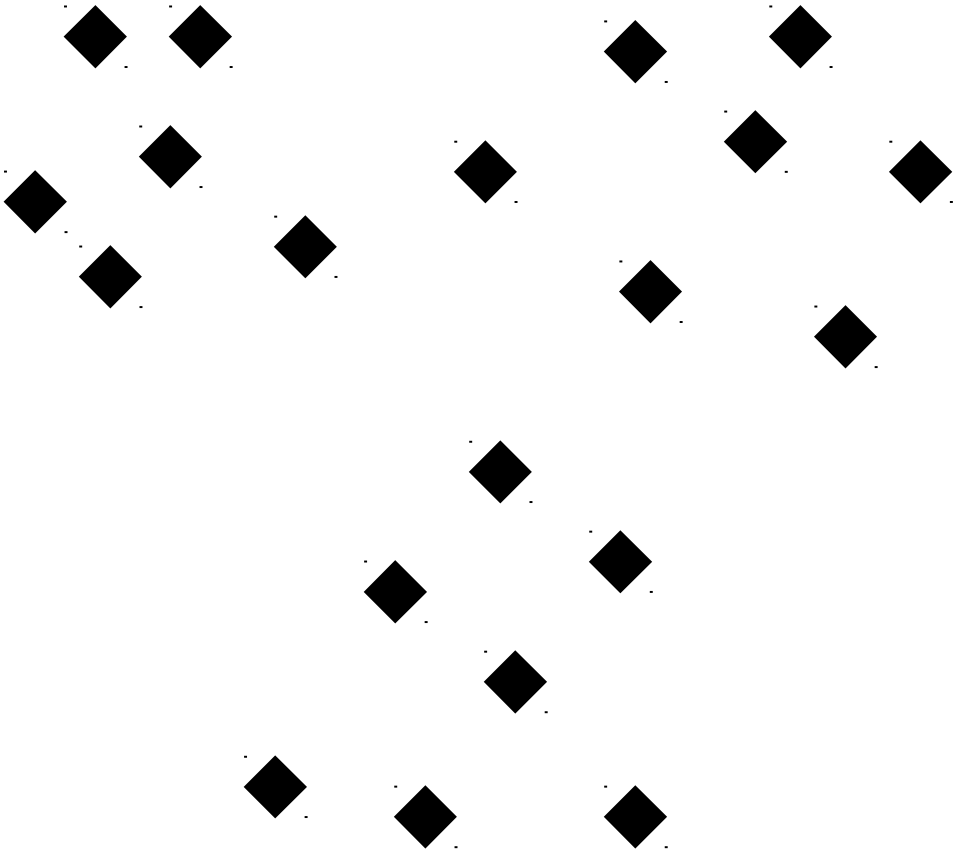
WT EGF-stimulation time-course

Most EGF effects are not PI3K-dependent (MAPK etc.)

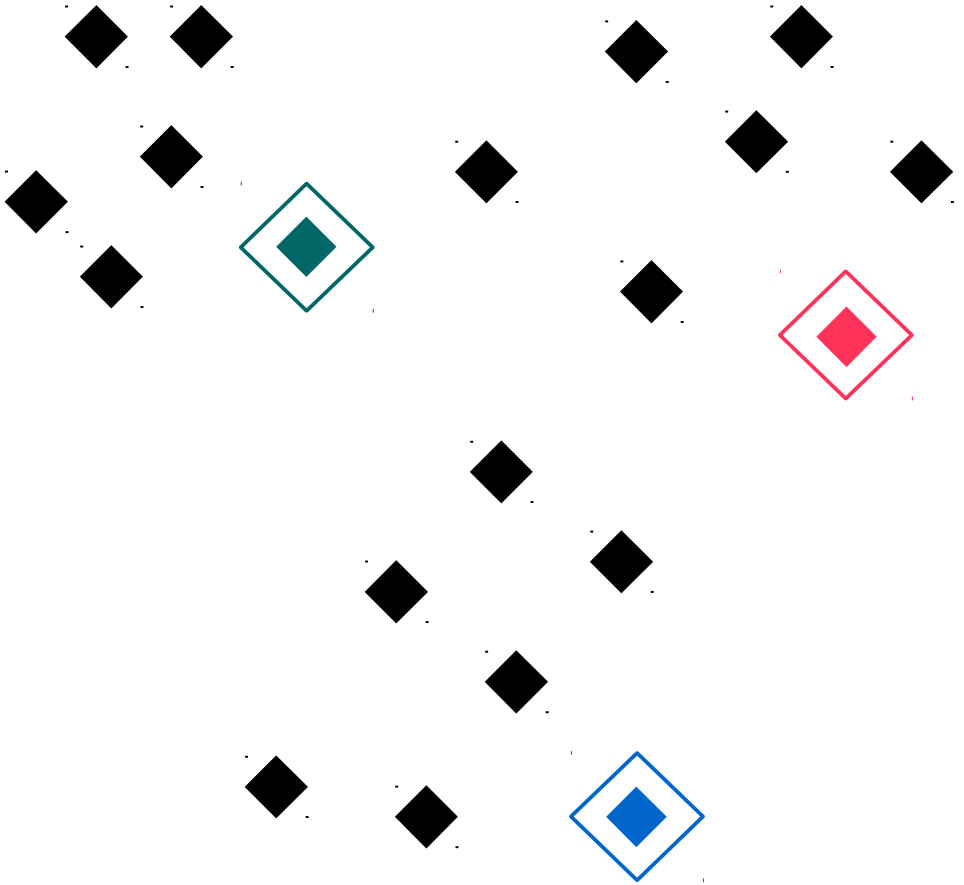
Clustering by K-medoid method

- Try 2 to n medoids, n depending on the number of samples
- Compute the stability of clusters with 100 data bootstrappings

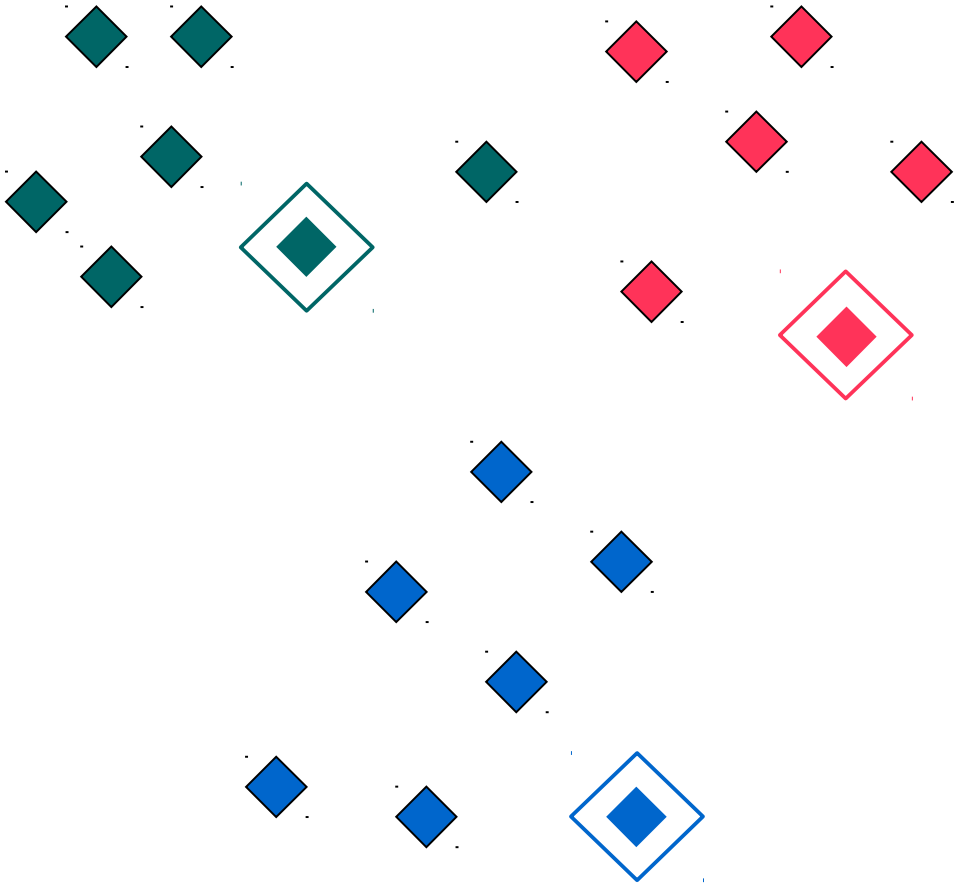
K-medoid clustering method



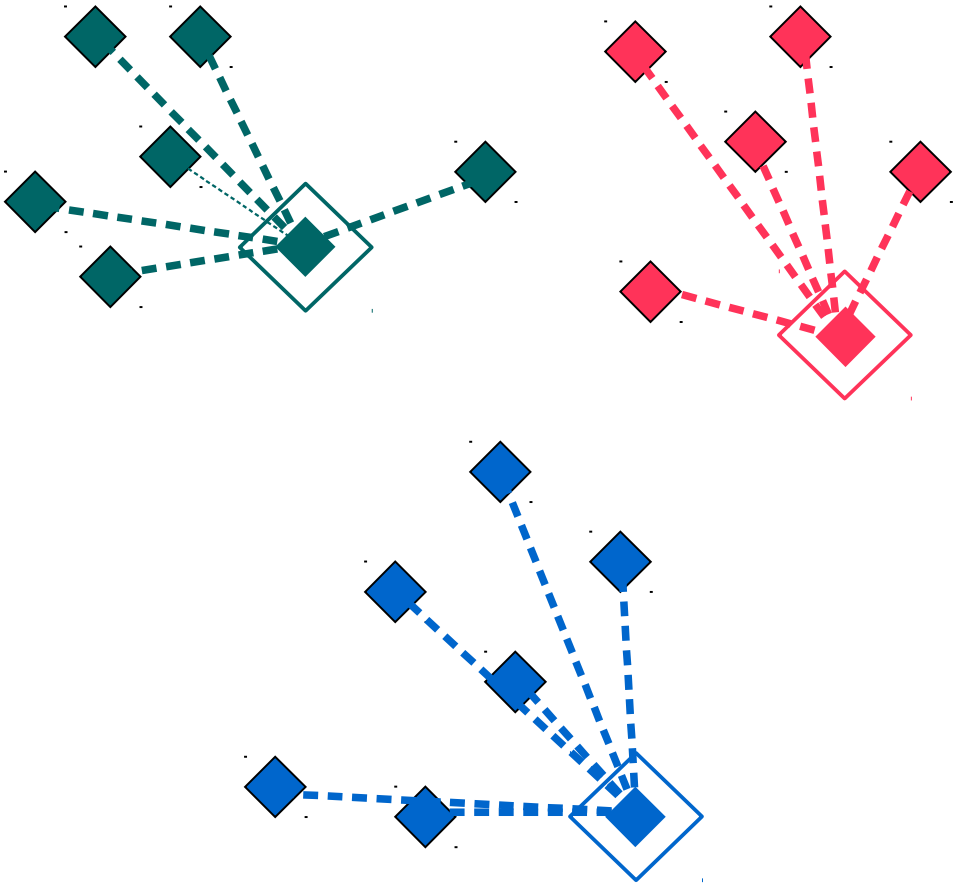
K-medoid: choose k medoids



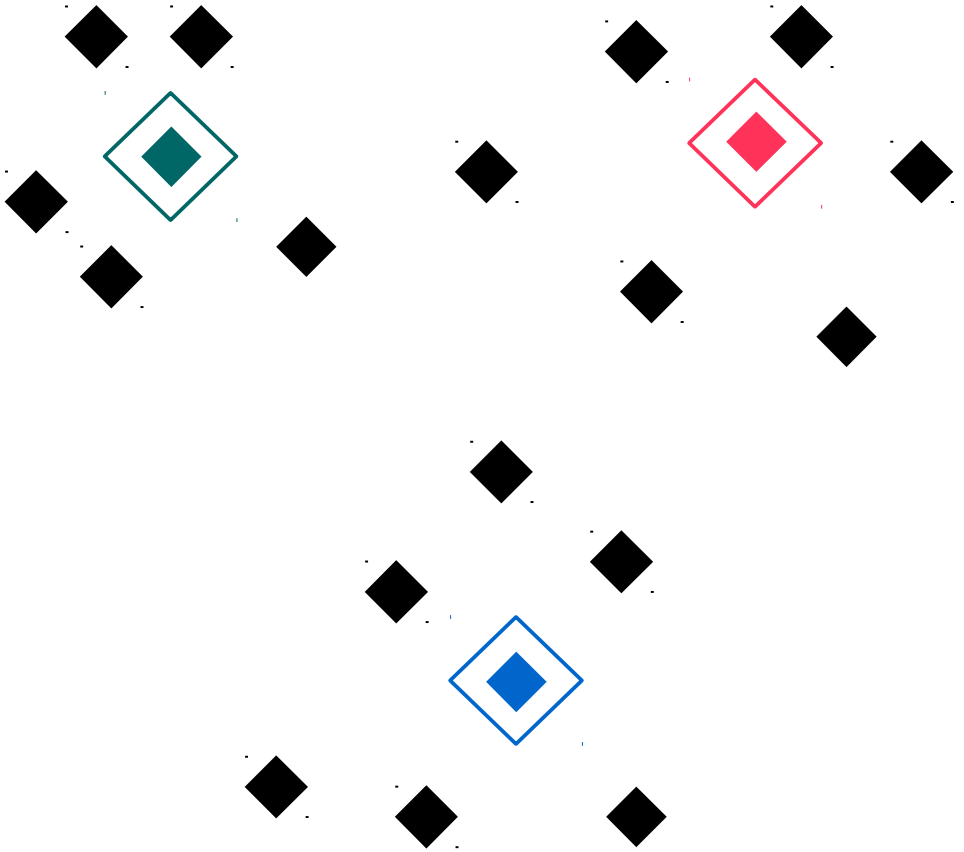
K-medoid: assign each dataset to the closest medoid



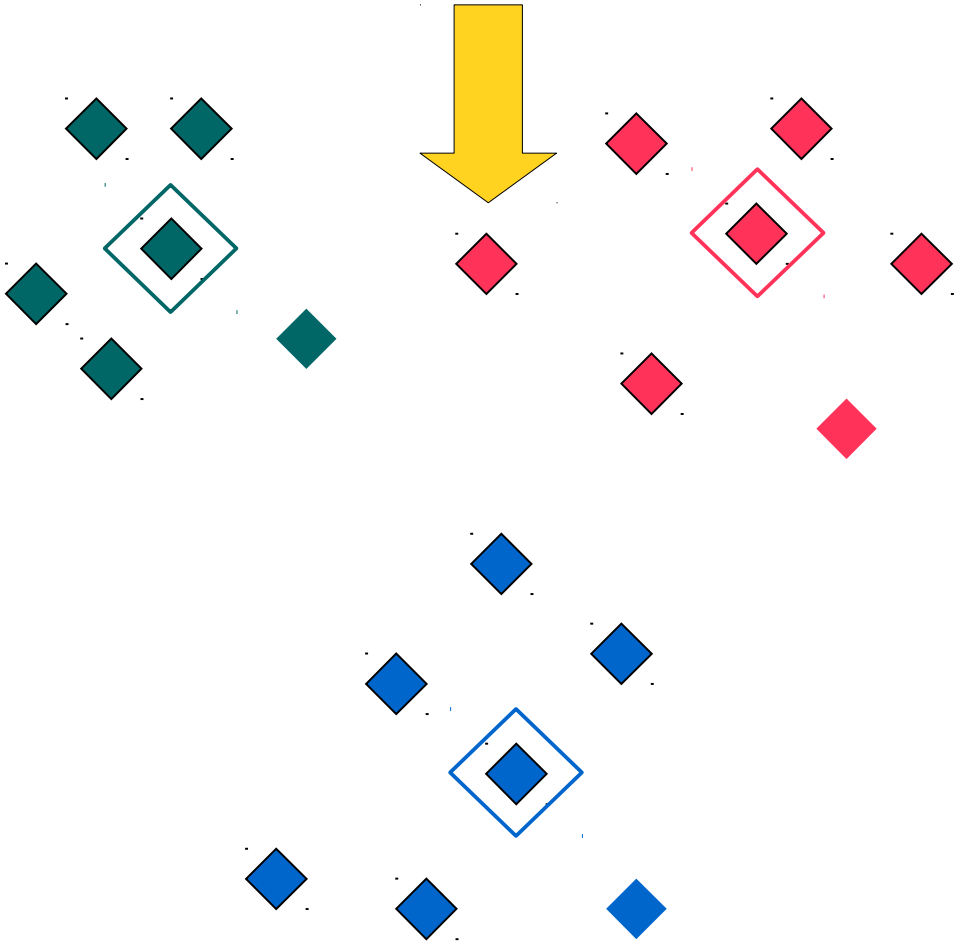
K-medoid: compute score



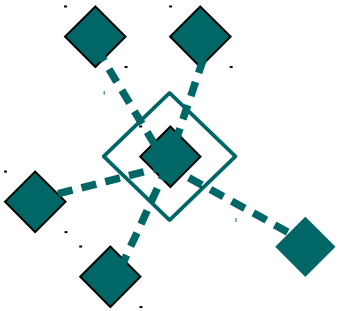
K-medoid: choose k non-medoids

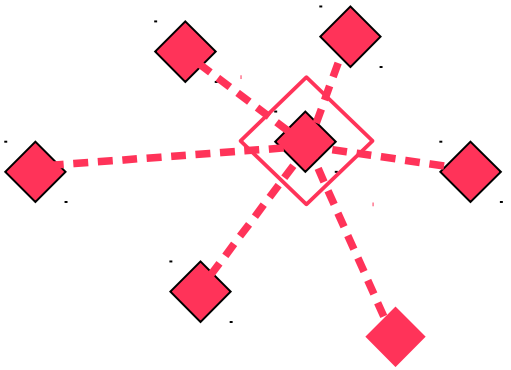


K-medoid: assign each dataset to the new medoids

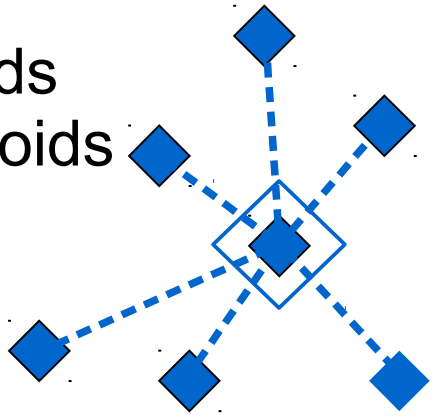


K-medoid: compute new scores

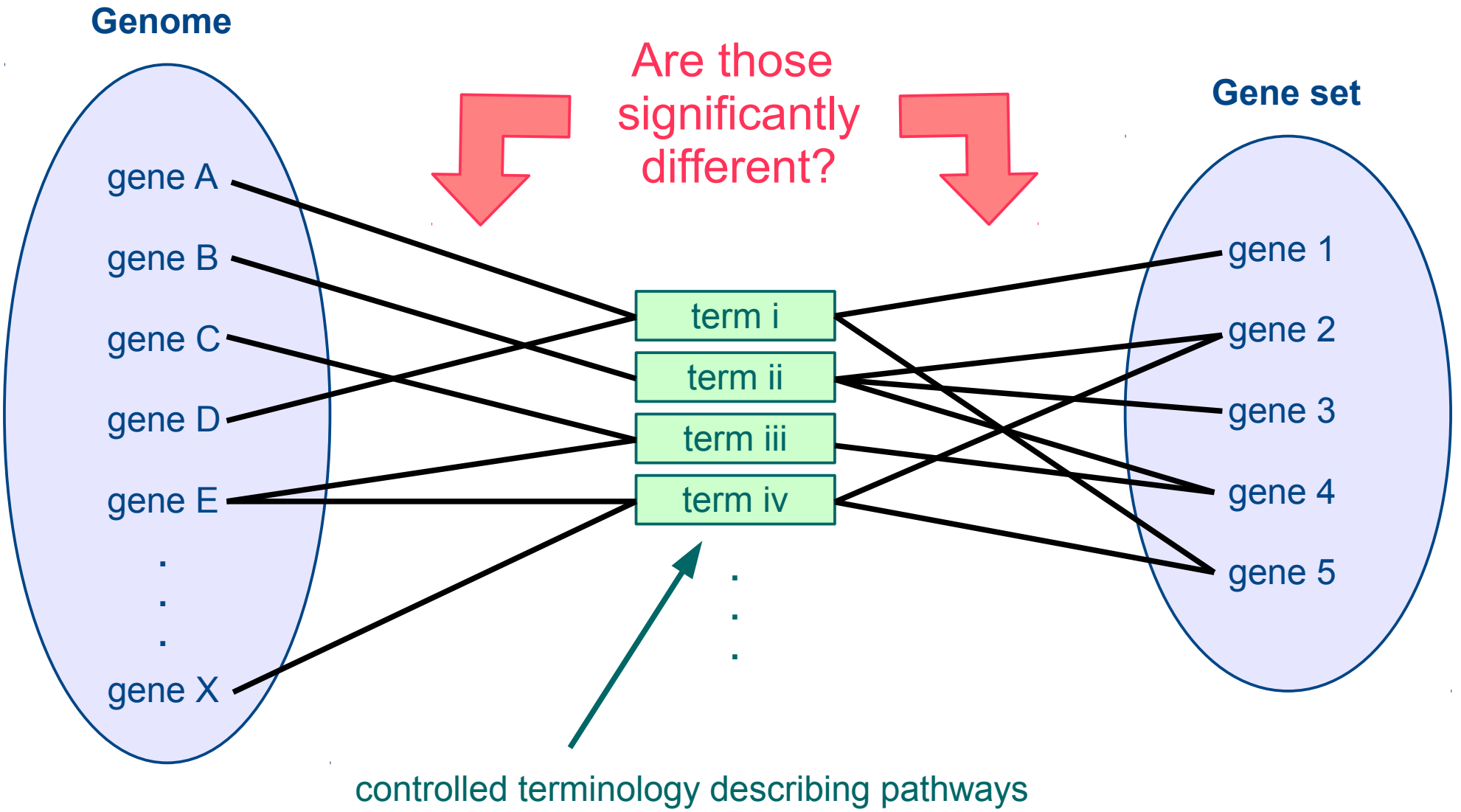




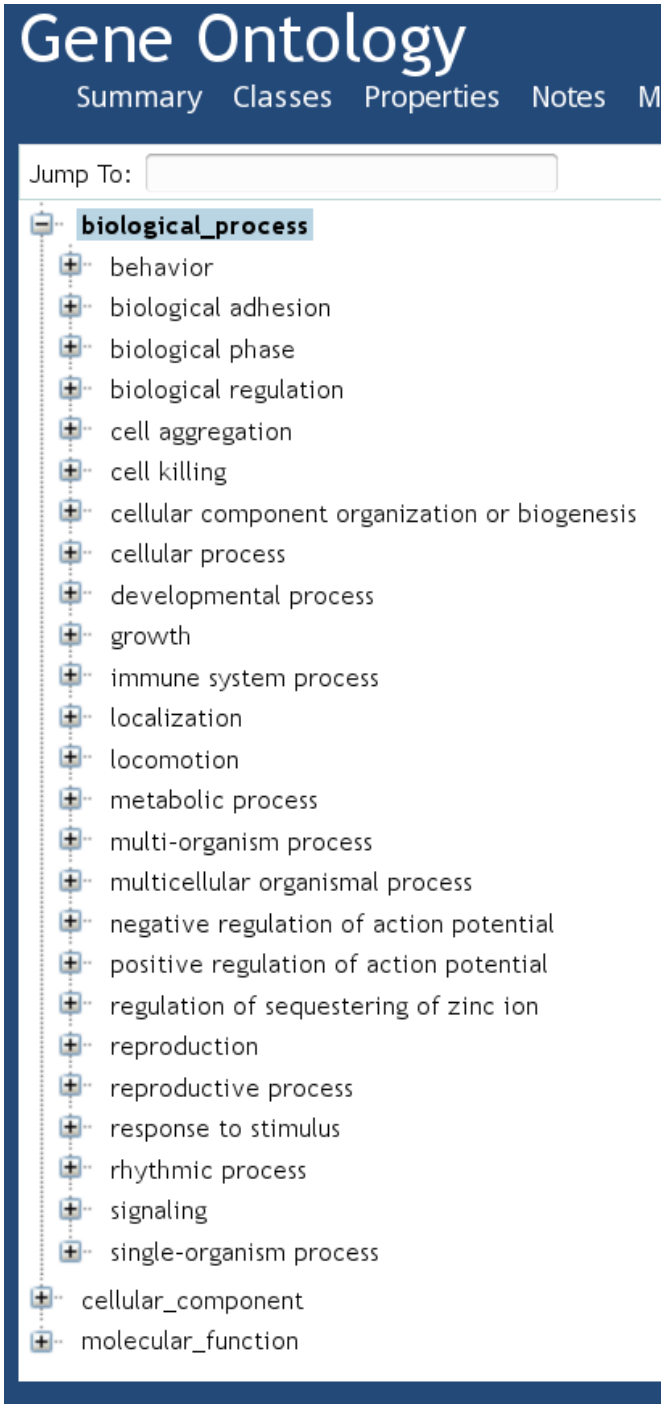
is score better?
yes → keep those medoids
no → keep previous medoids
Try new non-medoids



Principle of pathway enrichment



Gene Ontology annotations



Phosphatidylinositol 3-kinase regulatory subunit alpha

PIK3R1

Homo sapiens (Human)

View only features (sites, domains, PTMs ...)

Reviewed - Annotation score: ●●●●● - Experimental evidence at protein level[i]

BLAST | Align | Format | Add to basket | History

Function[i]

Binds to activated (phosphorylated) protein-Tyr kinases, through its SH2 domain, and acts as an adap
glucose uptake and glycogen synthesis in insulin-sensitive tissues. Plays an important role in signaling
(PubMed:17626883, PubMed:19805105, PubMed:7518429). Modulates the cellular response to ER stres
overloading in the liver and hence plays a role in glucose tolerance improvement (PubMed:20348923).

GO - Molecular function[i]

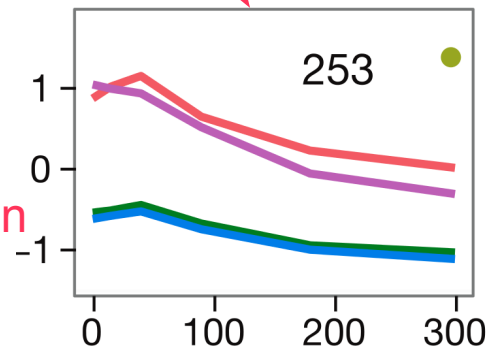
- 1-phosphatidylinositol-3-kinase regulator activity — Source: GO_Central
- ErbB-3 class receptor binding — Source: UniProtKB
- insulin binding — Source: UniProtKB
- insulin-like growth factor receptor binding — Source: UniProtKB
- insulin receptor binding — Source: UniProtKB
- insulin receptor substrate binding — Source: BHF-UCL
- neurotrophin TRKA receptor binding — Source: UniProtKB
- phosphatidylinositol 3-kinase binding — Source: BHF-UCL
- phosphatidylinositol 3-kinase regulator activity — Source: UniProtKB
- protein phosphatase binding — Source: UniProtKB
- transcription factor binding — Source: UniProtKB
- transmembrane receptor protein tyrosine kinase adaptor activity — Source: BHF-UCL

GO - Biological process[i]

- B cell differentiation — Source: Ensembl
- blood coagulation — Source: Reactome
- cellular glucose homeostasis — Source: UniProtKB
- cellular response to insulin stimulus — Source: UniProtKB
- cellular response to UV — Source: Ensembl
- epidermal growth factor receptor signaling pathway — Source: Reactome
- extrinsic apoptotic signaling pathway via death domain receptors — Source: Ensembl
- Fc-epsilon receptor signaling pathway — Source: Reactome

What am I showing you on the next slides?

one cluster,
comprising 253 genes



centred
normal
expression

time of EGF treatment

WT

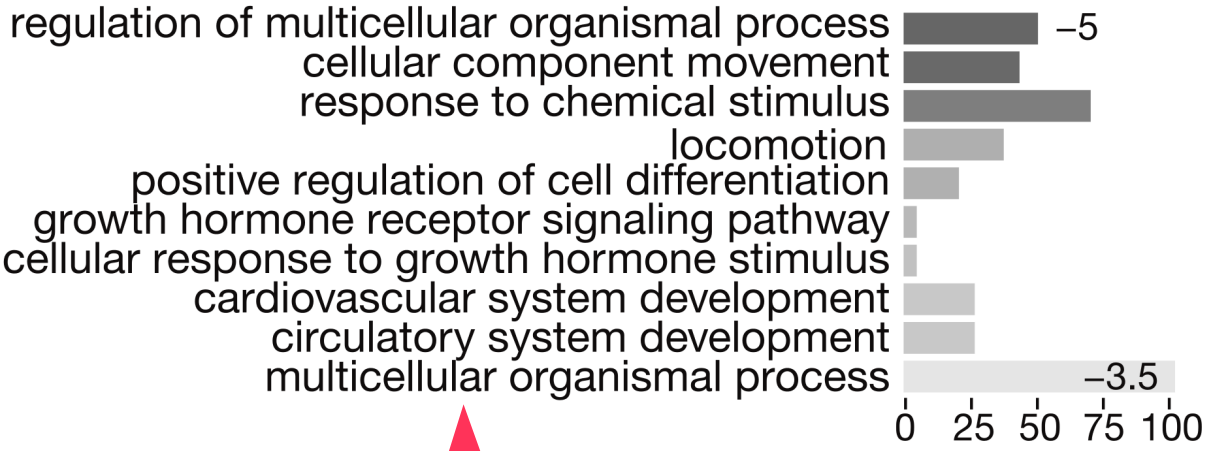
WT+A66

PI3K H1045R

PTEN -/-

A66 no EGF

log p-value
(darker = more significant)



Gene Ontology
Biological Process

number of genes
with this GO term

Babraham Institute

http://www.bioinformatics.babraham.ac.uk/shiny/kiselev-pip3-rna-seq-gene-profiles/

Gene expression profiles in MCF10A cells upon EGF stimulation

Gene HGNC symbol / Ensembl ID

LBH

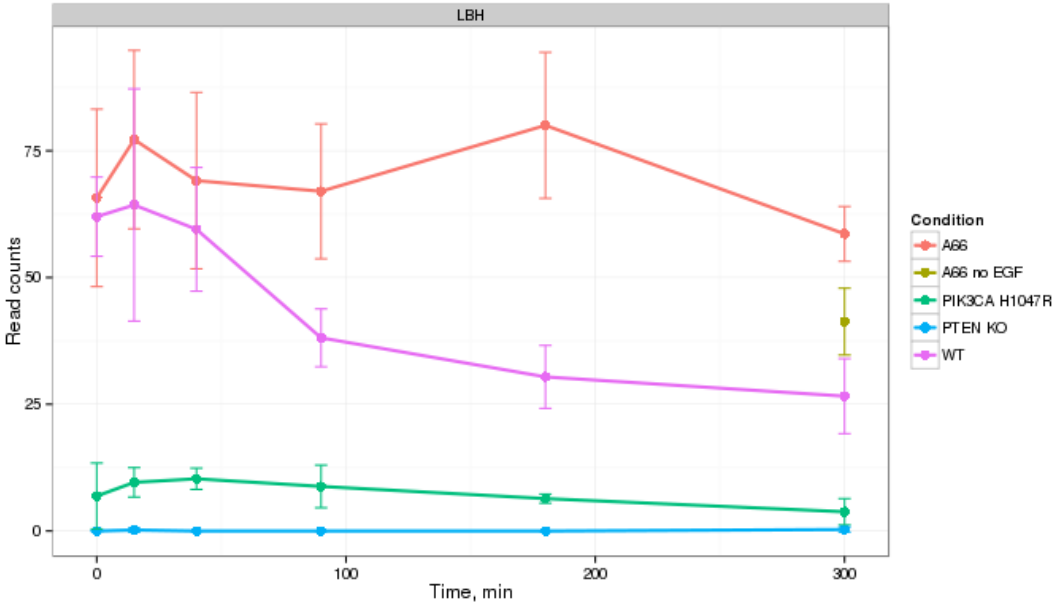
RPKM

Please click on links below only after the plot appears on the right.

Download Plot

Download Data

This is a Supplementary figure for the paper...
(read counts are normalized by library sizes)



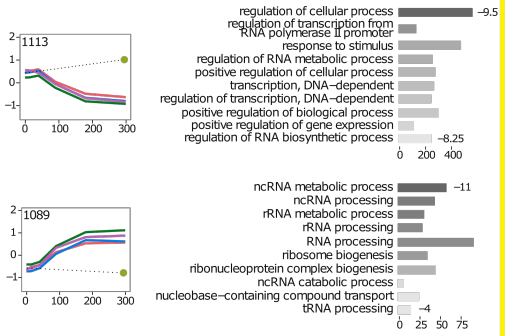
Babraham Institute

EGF

PI3K

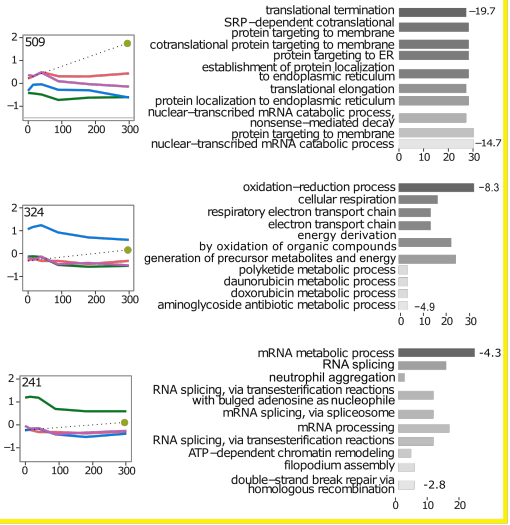
1

Regulation of transcription
Ribosome biogenesis
ncRNA/ rRNA/ RNA processing



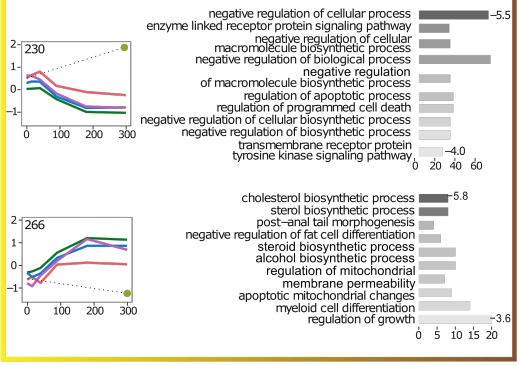
2

Cellular respiration/ Energy derivation
mRNA processing and splicing
Translation elongation/ termination



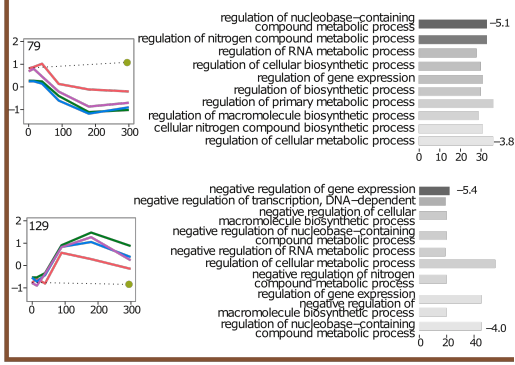
3

Apoptosis
Cell death
Sterol metabolism

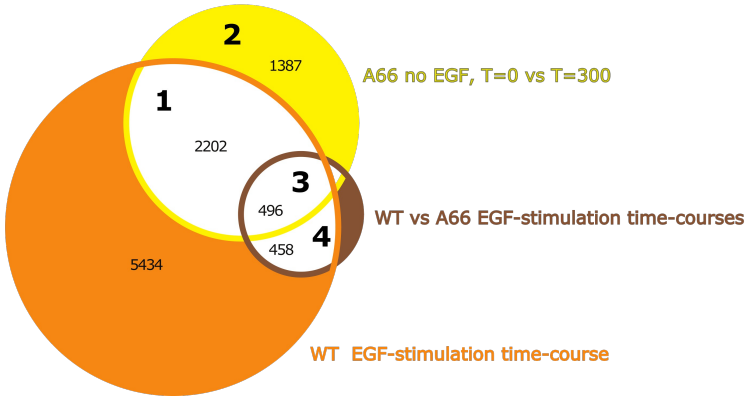


4

Regulation of RNA metabolism
Regulation of gene expression



WT
A66
A66 no EGF
PTEN KO
PIK3CA H1047R



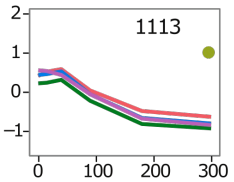
Babraham Institute

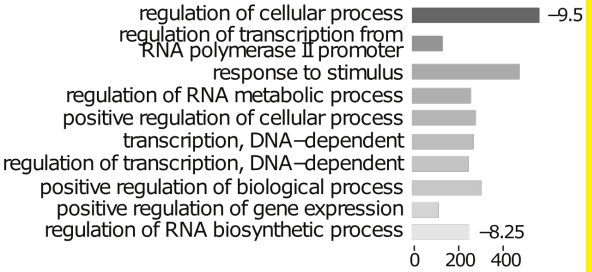
EC

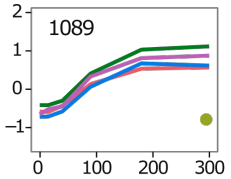
PI

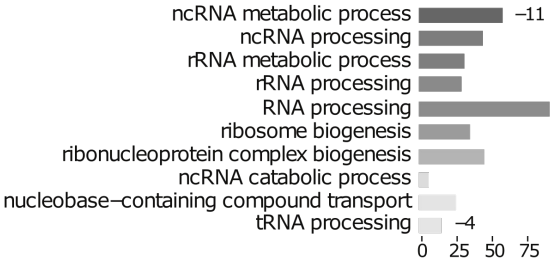
1

Regulation of transcription
Ribosome biogenesis
ncRNA/ rRNA/ RNA processing



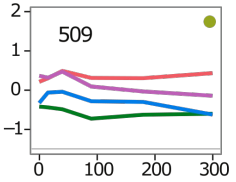


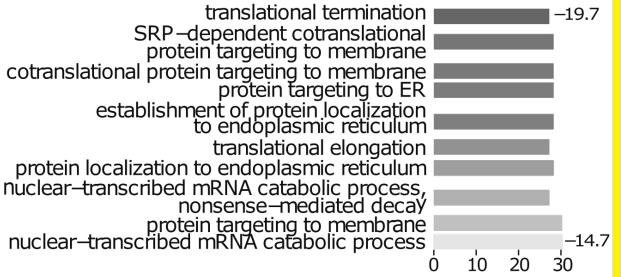


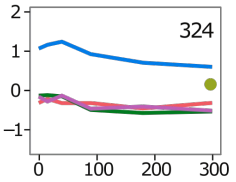


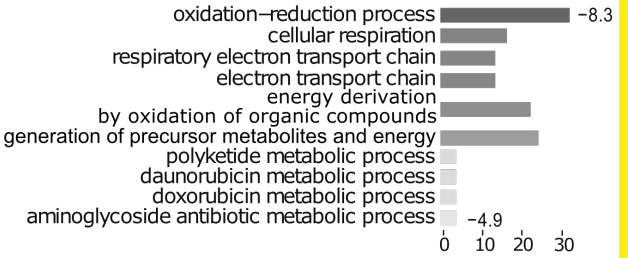
2

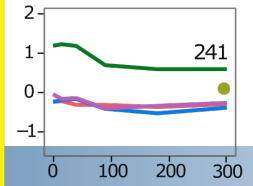
Cellular respiration/ Energy derivation
mRNA processing and splicing
Translation elongation/ termination

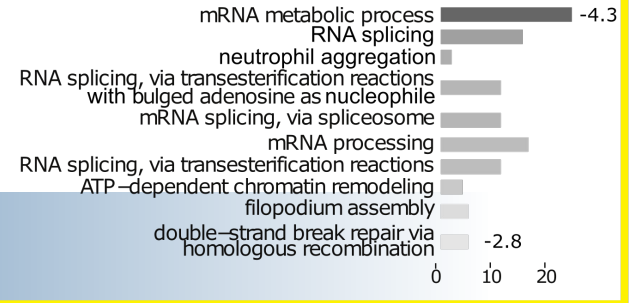




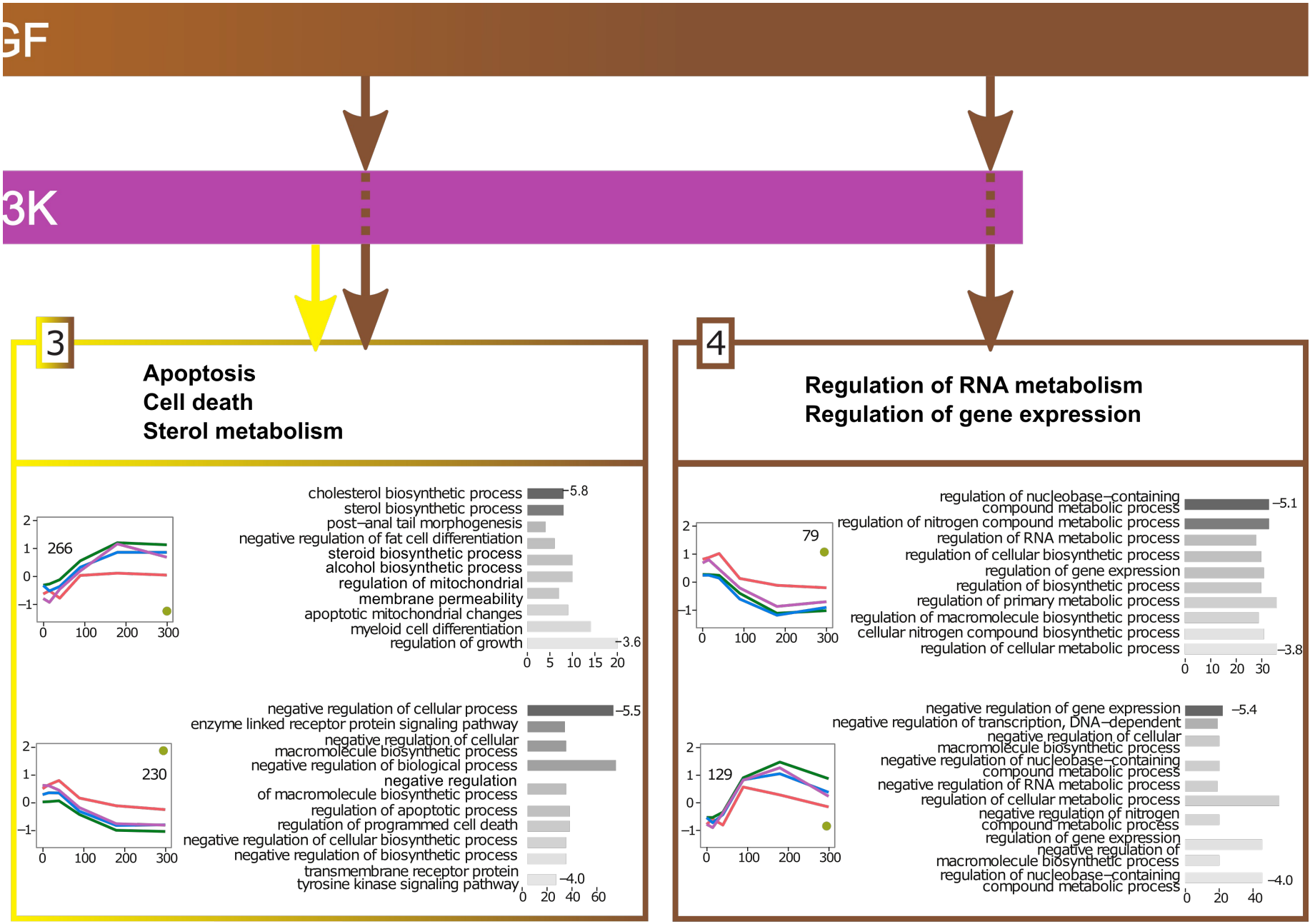






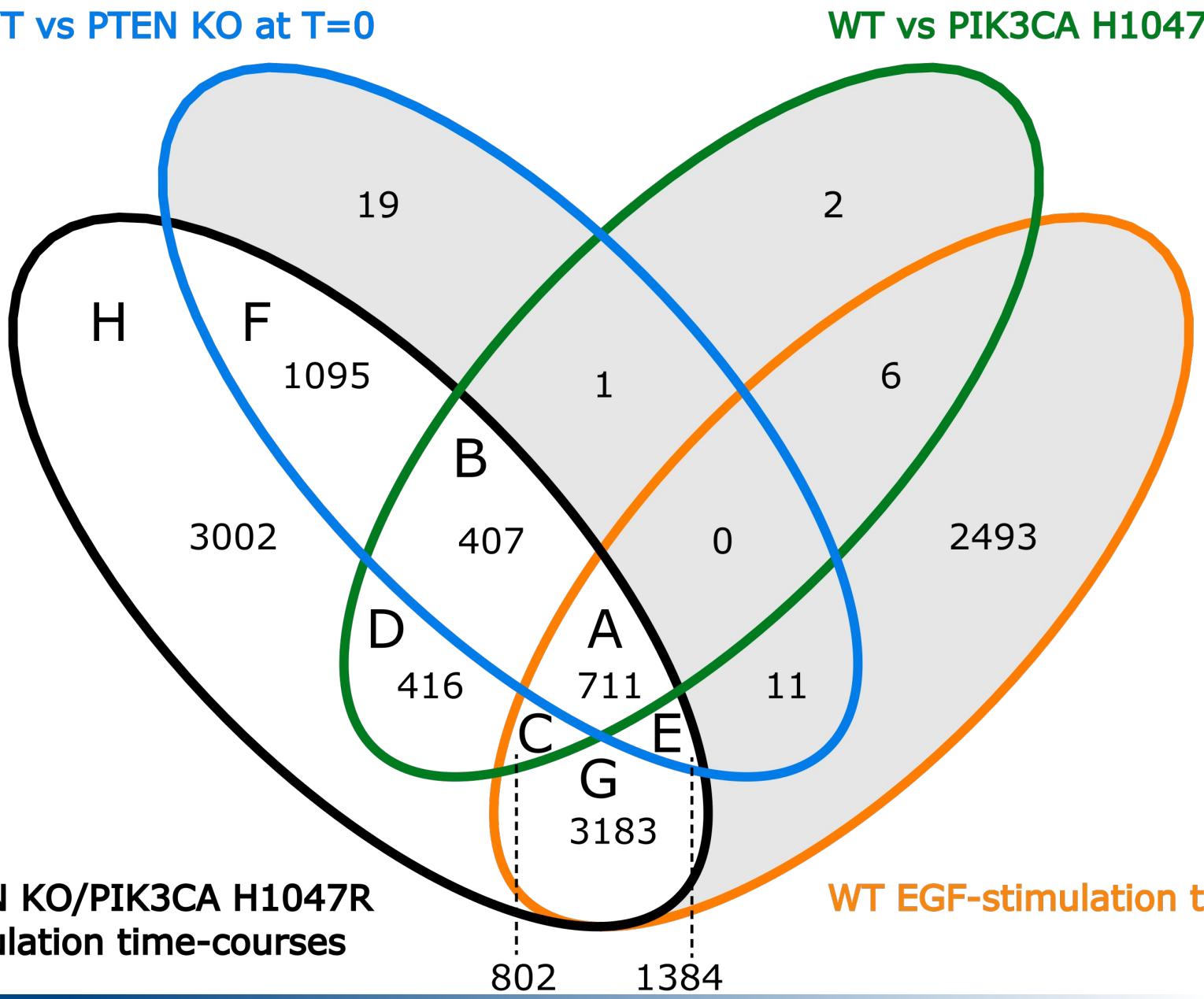


Babraham Institute



Babraham Institute

Effects of constitutive mutations

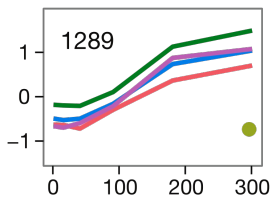


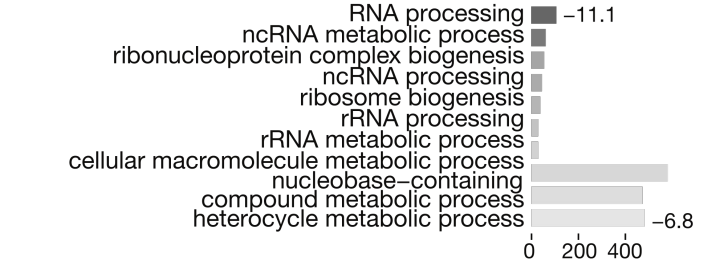
Babraham Institute

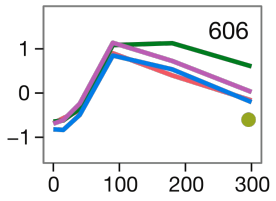
PIK3CA H1047R + PTEN KO

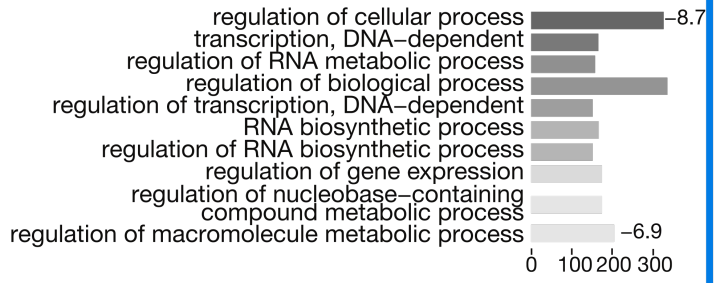
G

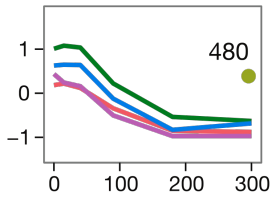
ncRNA/ rRNA/ RNA processing
Ribosome biogenesis
Transcription

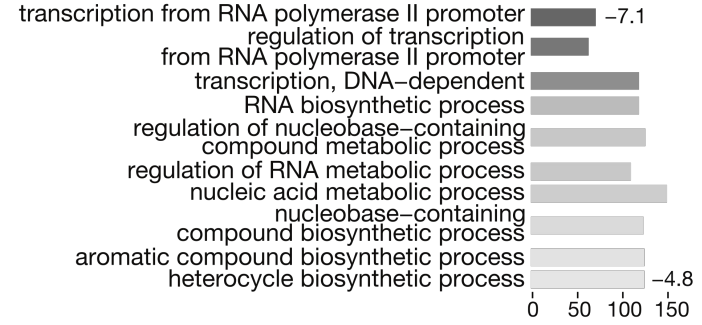






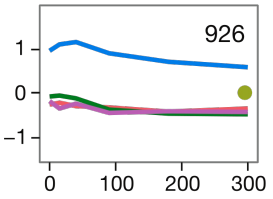


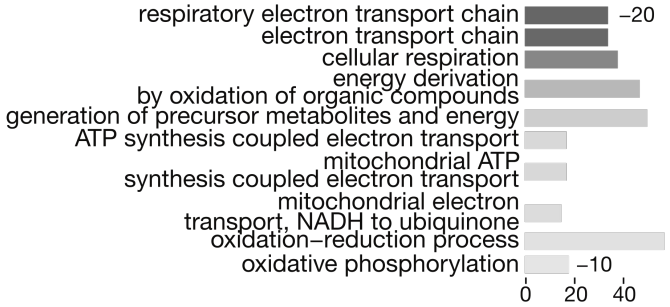


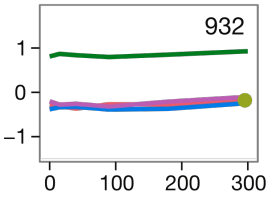


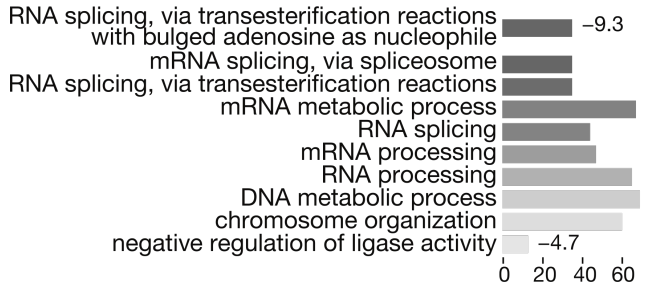
H

Cell respiration
Energy derivation
mRNA splicing/ processing



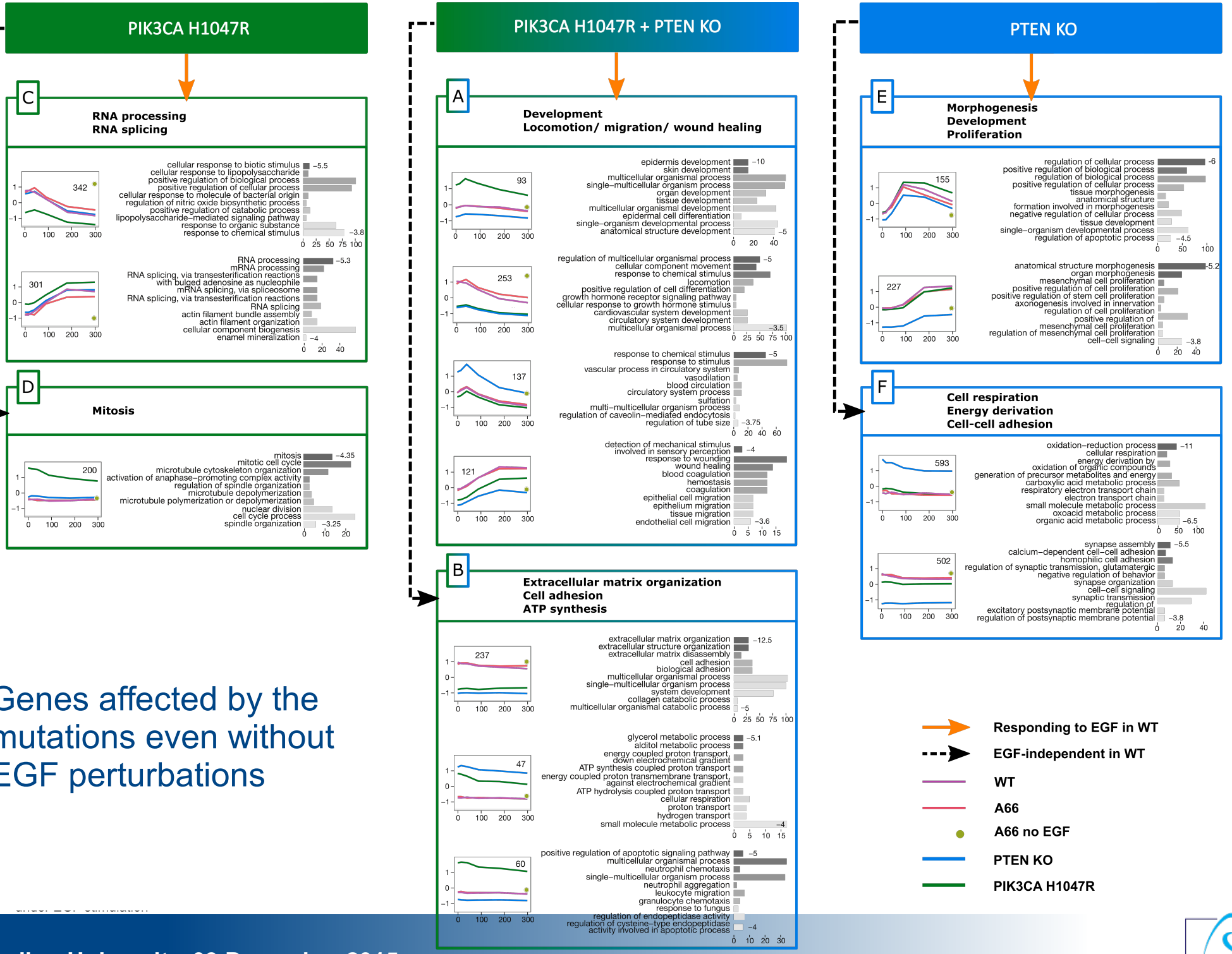






PI3K responding to EGF in mutants (~ groups 1 and 2 of A66)

Babraham Institute



Genes affected by the mutations even without EGF perturbations

Babraham Institute

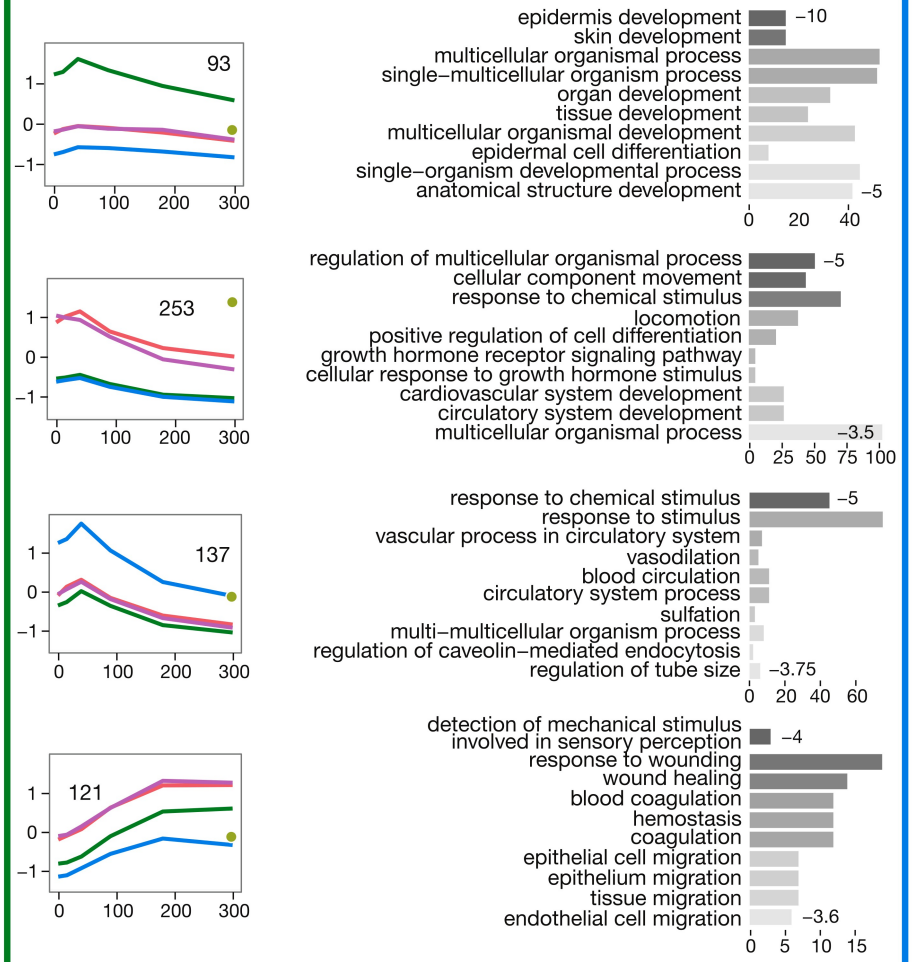
responding to EGF in WT

PIK3CA H1047R + PTEN KO

no response to EGF in WT

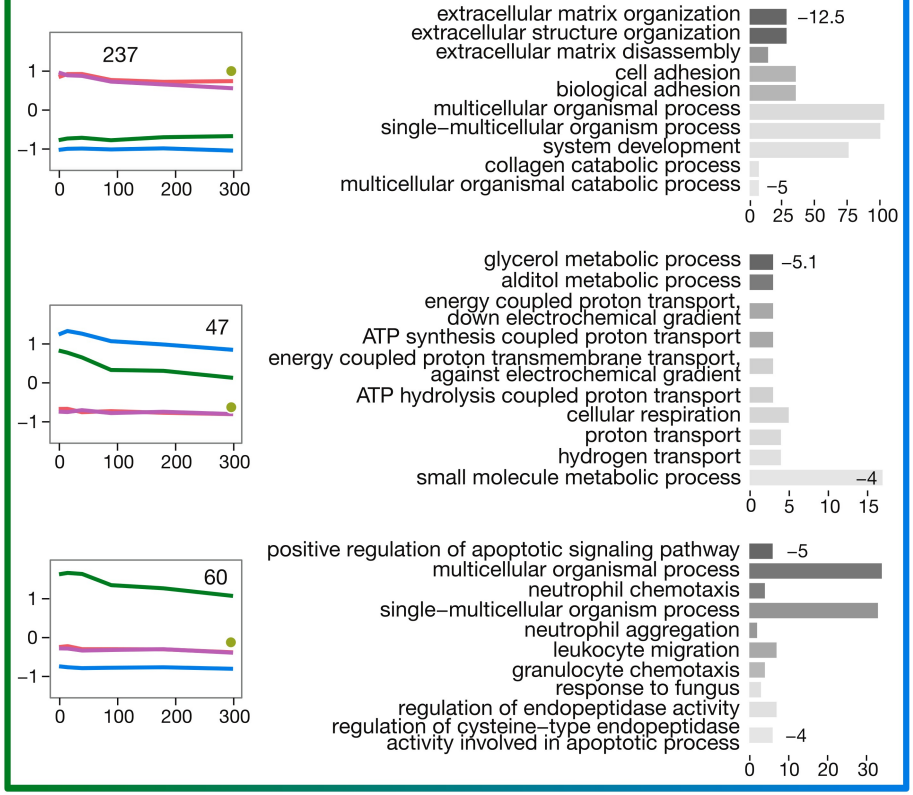
A

Development
Locomotion/ migration/ wound healing



B

Extracellular matrix organization
Cell adhesion
ATP synthesis



Genes responding to both mutations in a coherent fashion; affected by EGF in the mutants

Babraham Institute

no response to EGF in WT

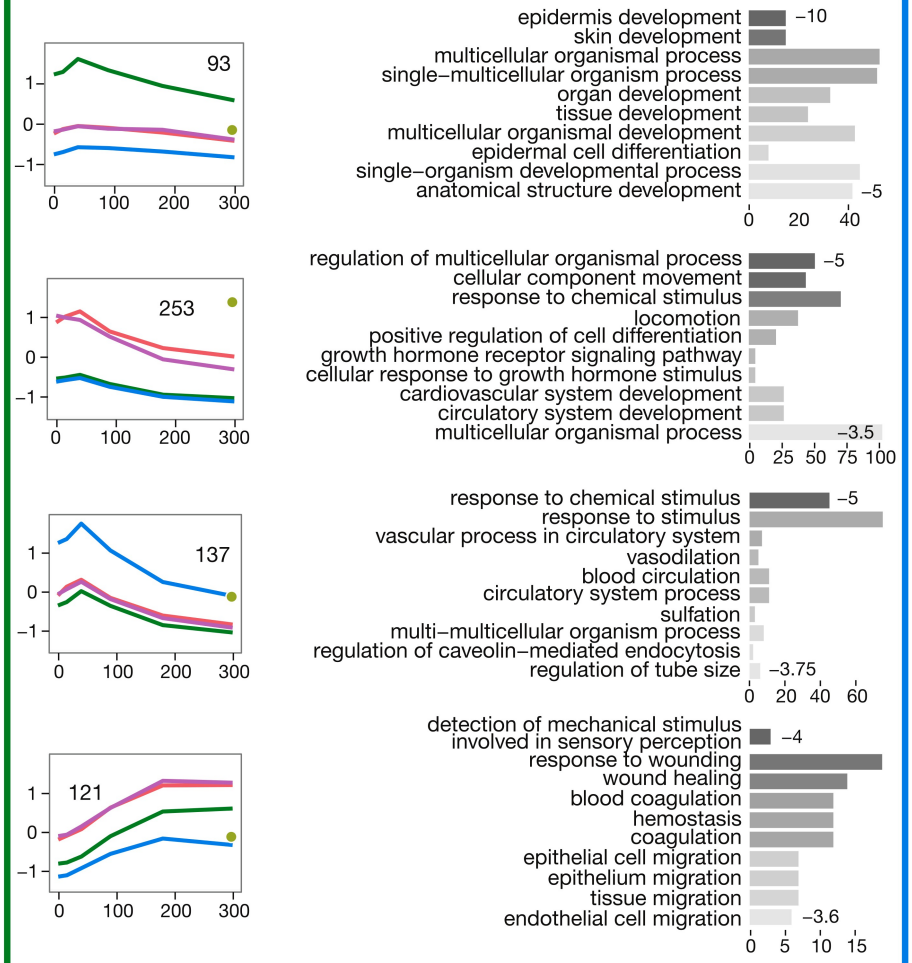
PIK3CA H1047R + PTEN KO

responding to EGF in WT

DYNAMIC

A

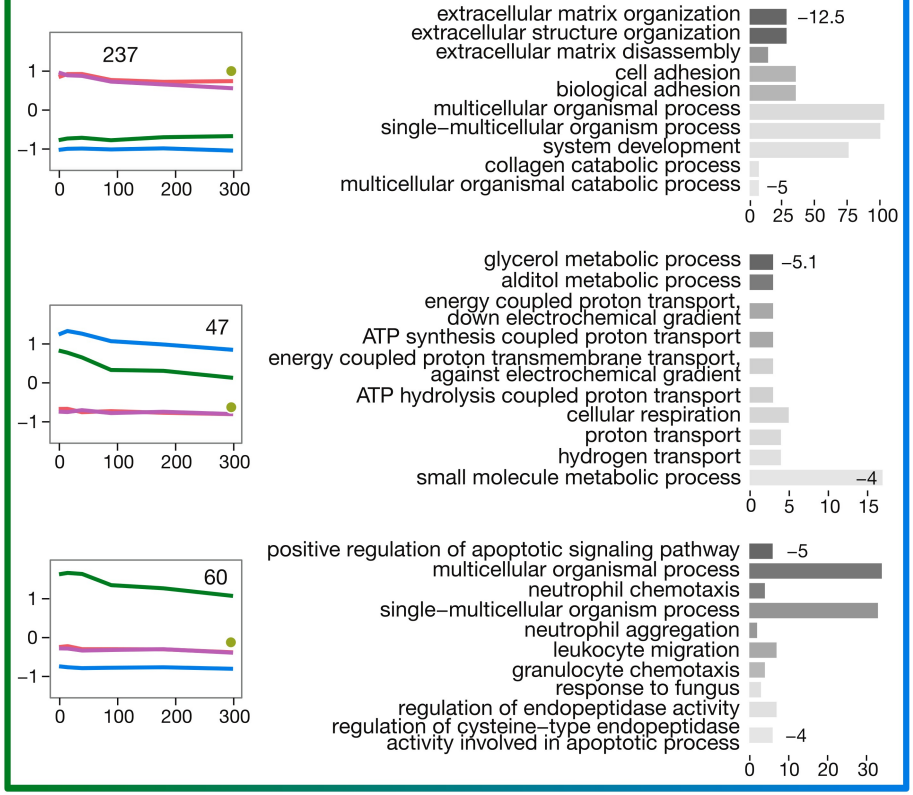
Development
Locomotion/ migration/ wound healing



STATIC

B

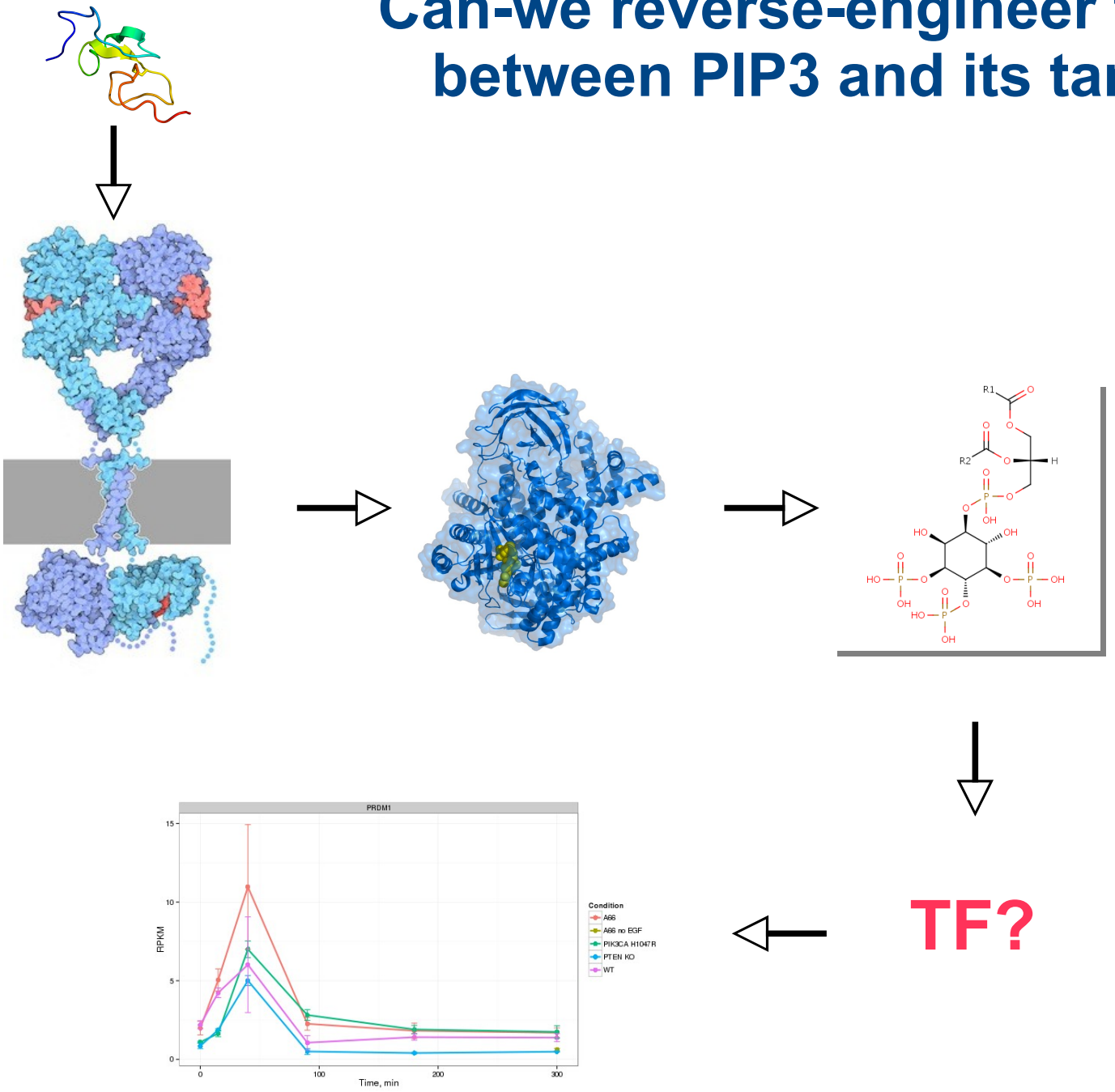
Extracellular matrix organization
Cell adhesion
ATP synthesis

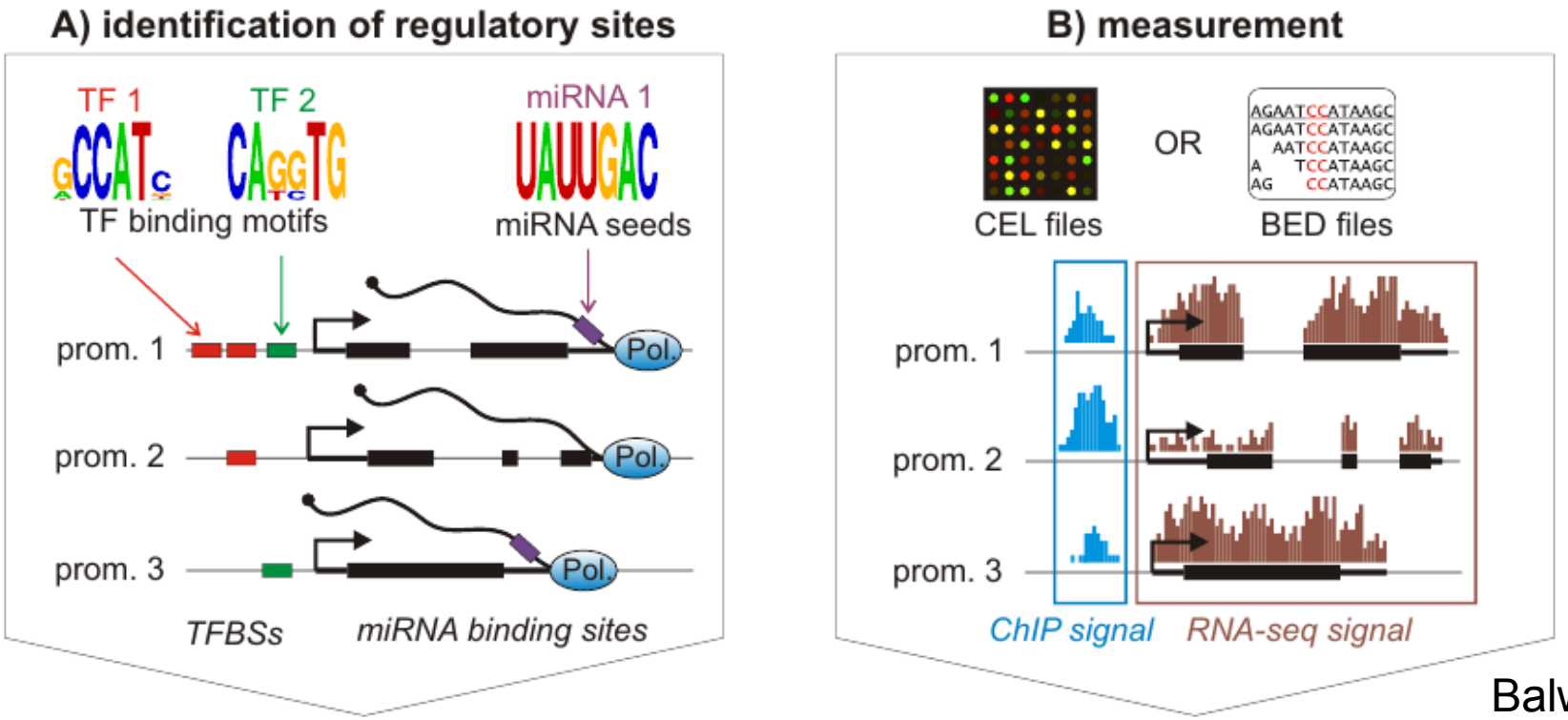


Genes responding to both mutations in a coherent fashion

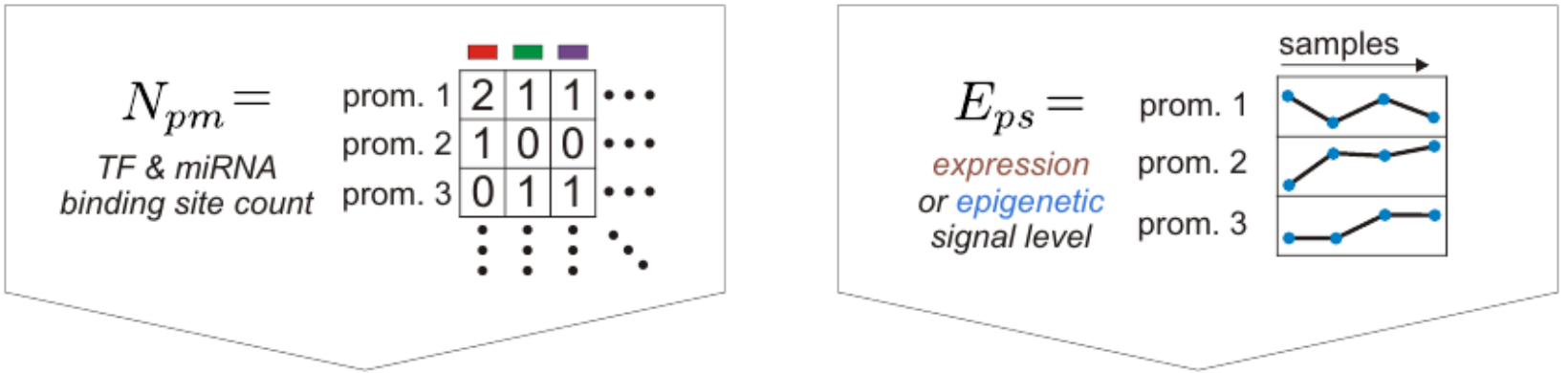
Babraham Institute

Can-we reverse-engineer the link between PIP3 and its targets?





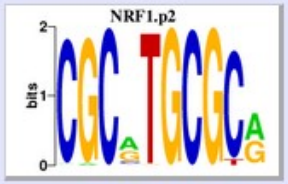
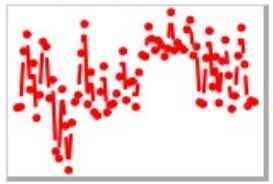
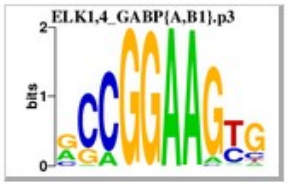
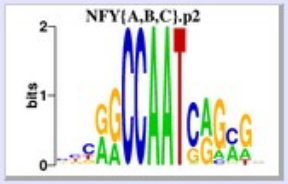
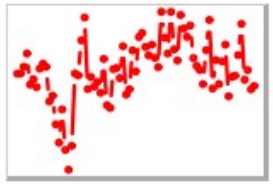
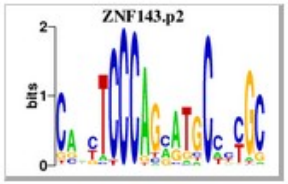
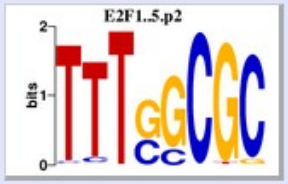
Balwiertz *et al* (2014) Genome Res

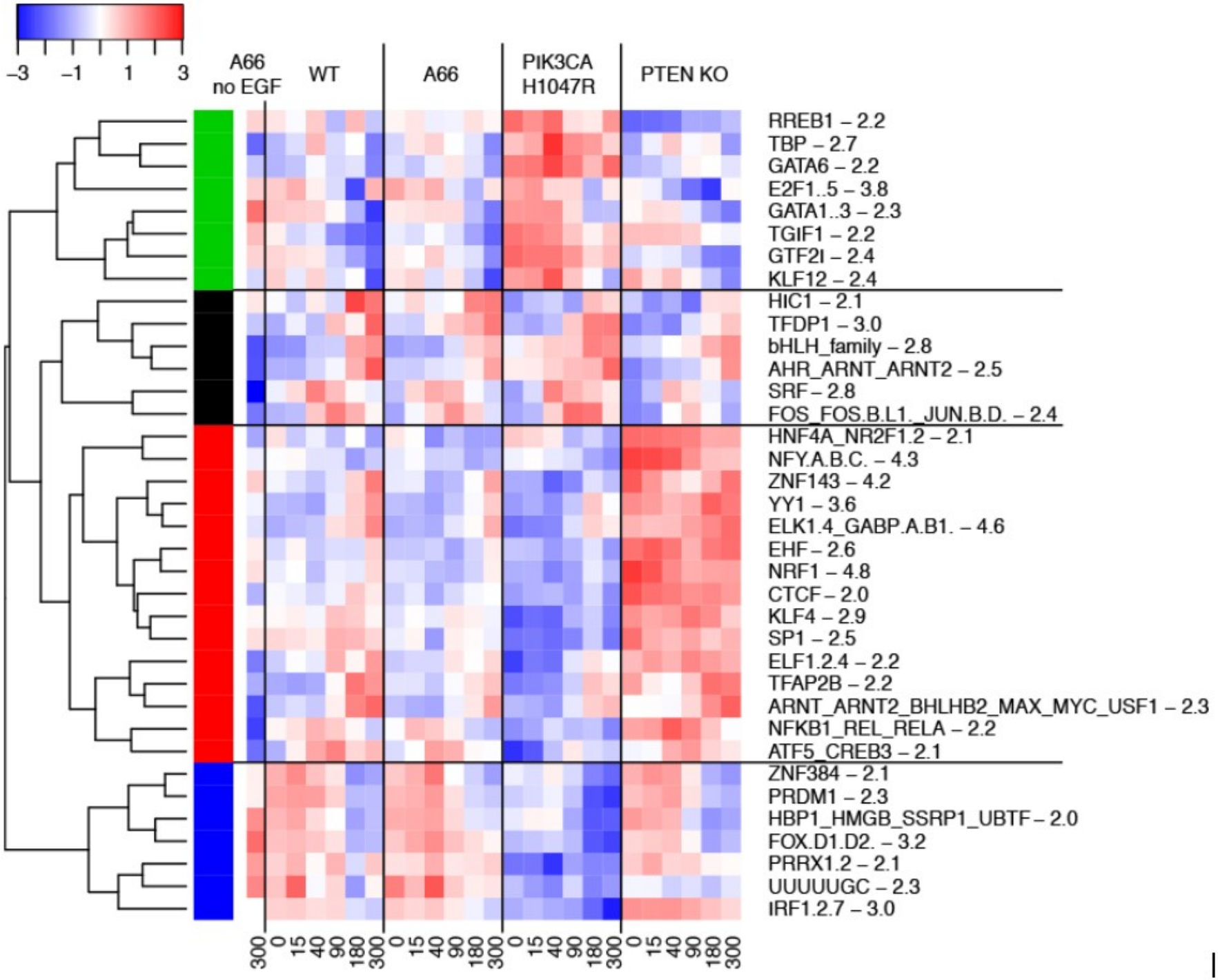


D) MARA model

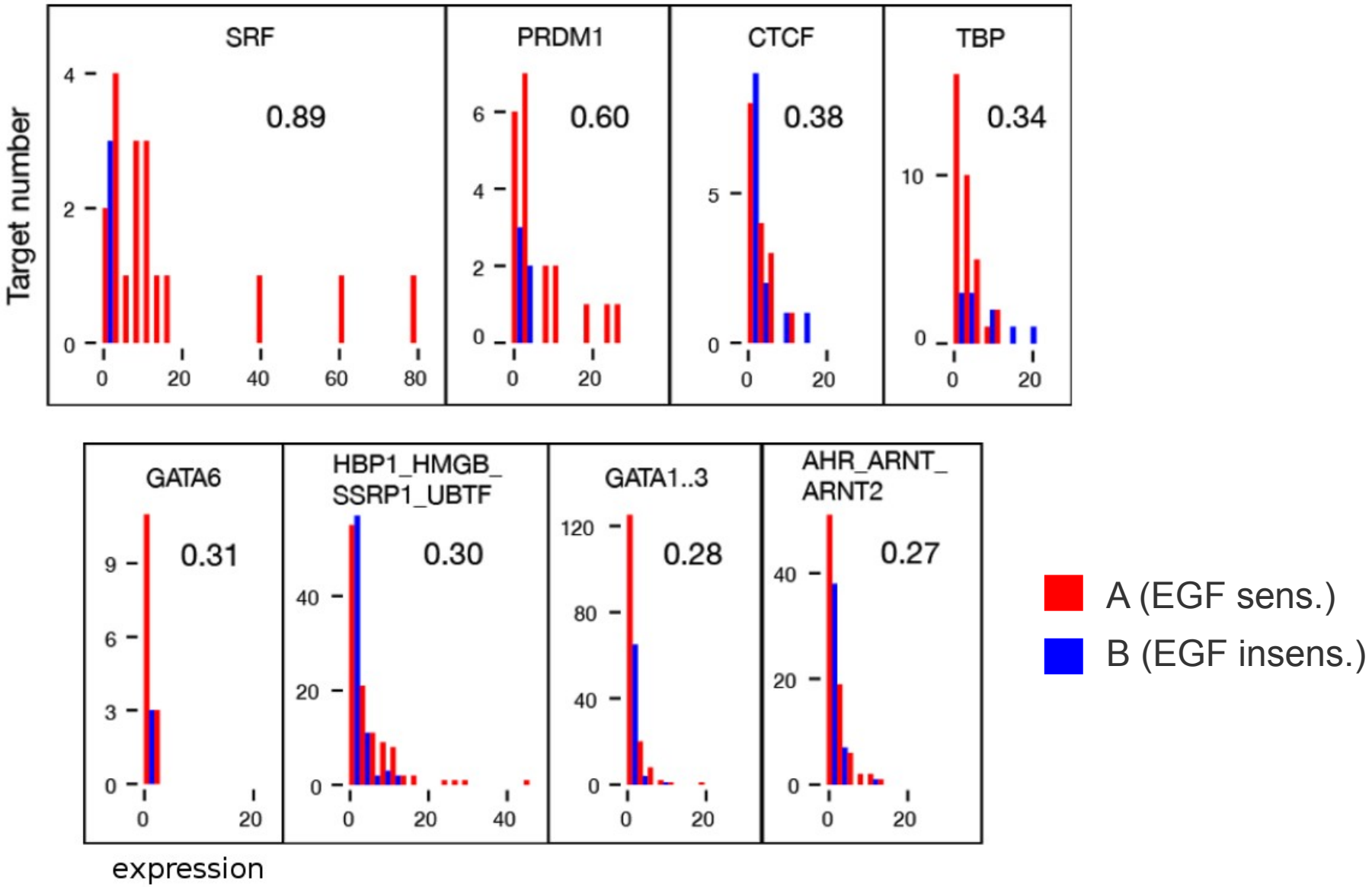
$$E_{ps} = \sum_m N_{pm} \cdot A_{ms} + c_p + \tilde{c}_s$$

Babraham Institute

Motif name	Z-value	Associated genes	Profile	Logo
NRF1.p2	4.755	NRF1 (EWG, ALPHA-PAL)		
ELK1,4_GABP{A,B1}.p3	4.633	GABPA (E4TF1A, NFT2, NRF2, E4TF1-60, NRF2A) GABPB1 (E4TF1-47, GABPB) ELK4 (SAP1) ELK1		
NFY{A,B,C}.p2	4.292	NFYC (CBF-C) NFYB (CBF-A, HAP3) NFYA (HAP2, CBF-B)		
ZNF143.p2	4.154	ZNF143 (SBF, pHZ-1, STAF)		
E2F1..5.p2	3.631	E2F4 (E2F-4) E2F5 E2F2 (E2F-2) E2F1 (RBP3) E2F3		
				

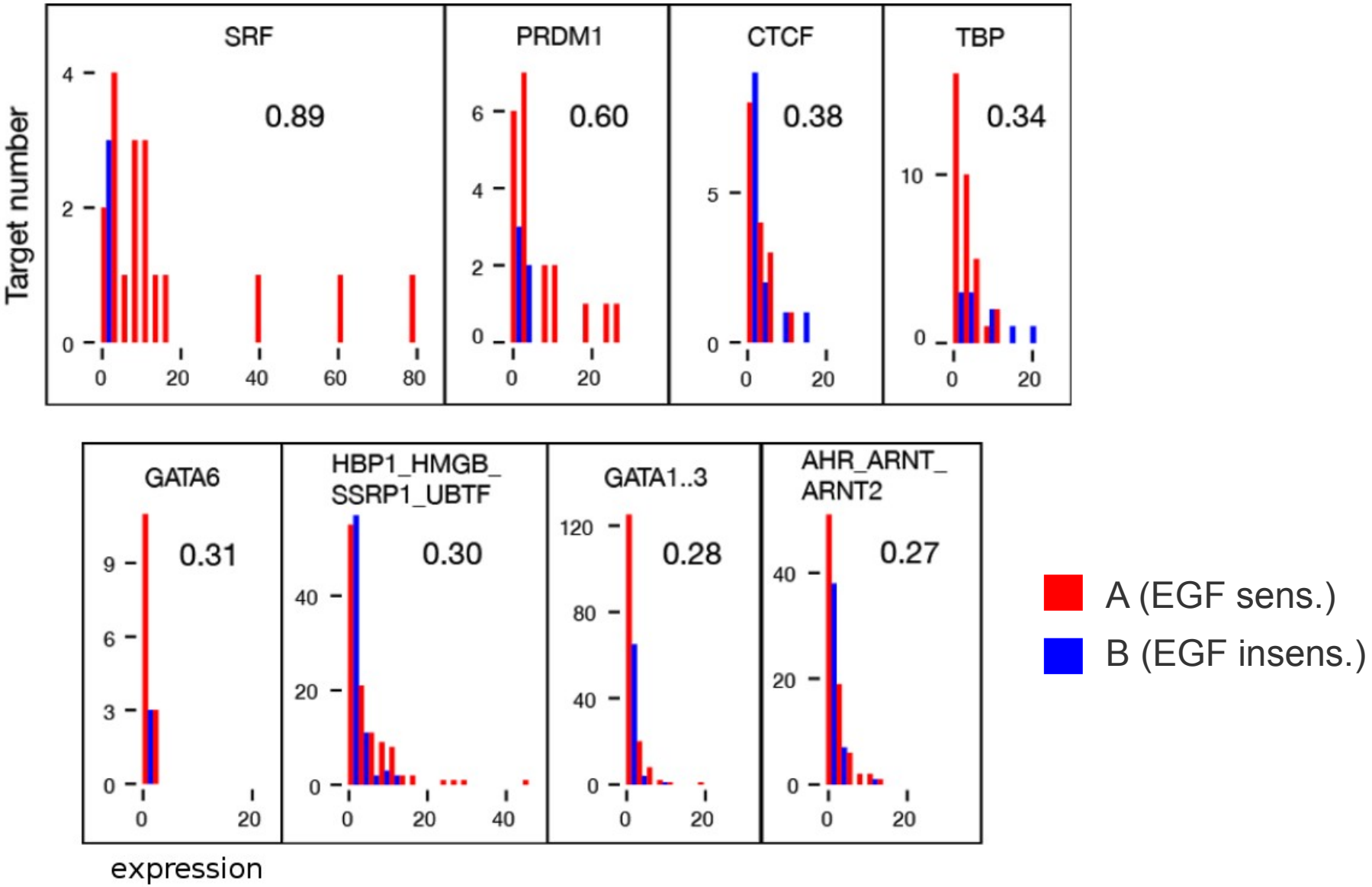


Babraham Institute



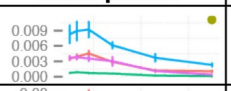
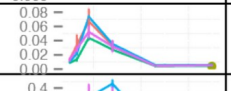
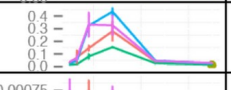
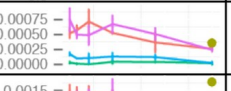
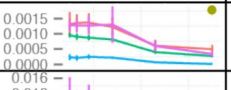
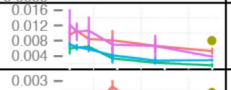
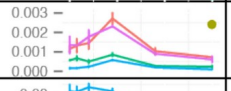
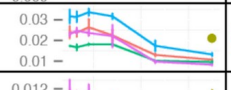
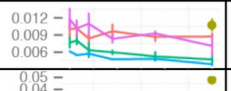
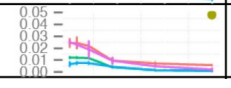
Babraham Institute

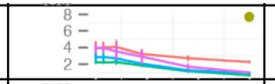
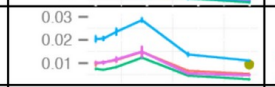
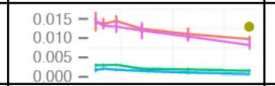
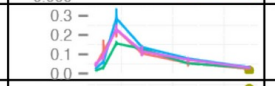
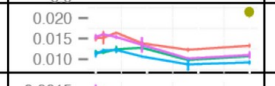
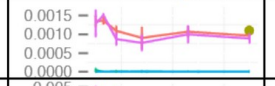
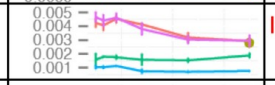
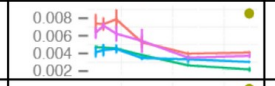
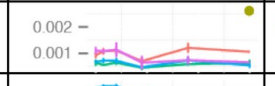
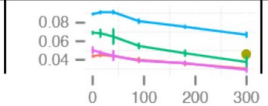
known new



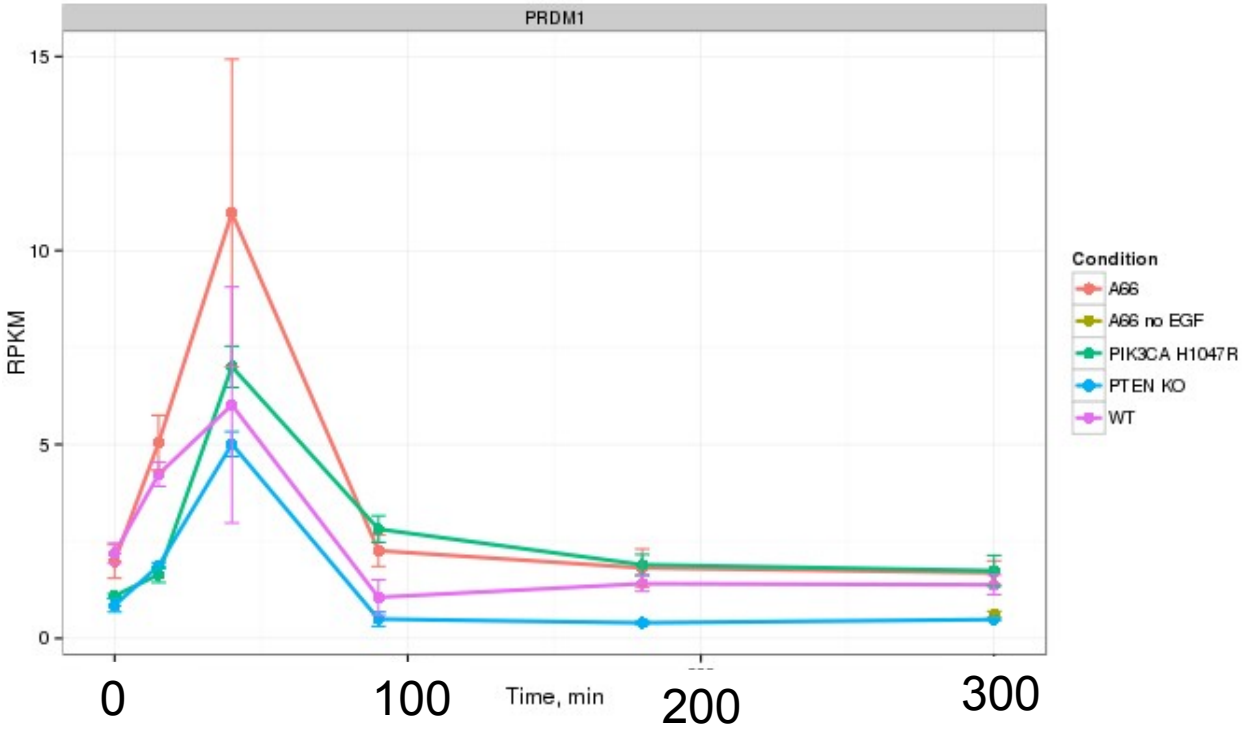
Babraham Institute

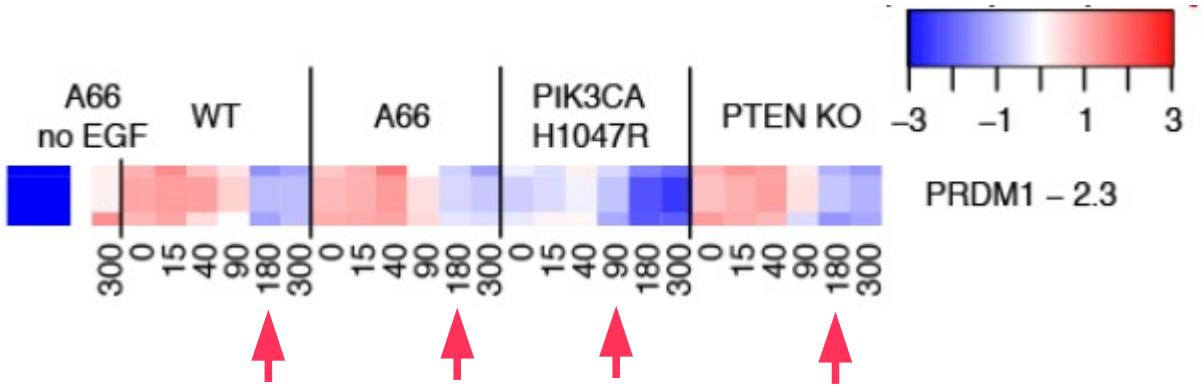
BLIMP1 (PRDM1) targets

Motif target gene	Sm	Gene profile	Description
KCNB1	28.0333		
NCOA7	23.8204		
LIF	20.1823		
ASPA	11.6217		
SORBS2	11.2664		
UBA7	9.8144		
LMCD1	9.23465		
CYLD	5.01778		Deubiquitination of AKT (Lim et al. 2012)
TAPBPL	4.94158		
PIK3IP1	4.6666		Negative regulator of hepatic PI3K activity (He et al. 2008)

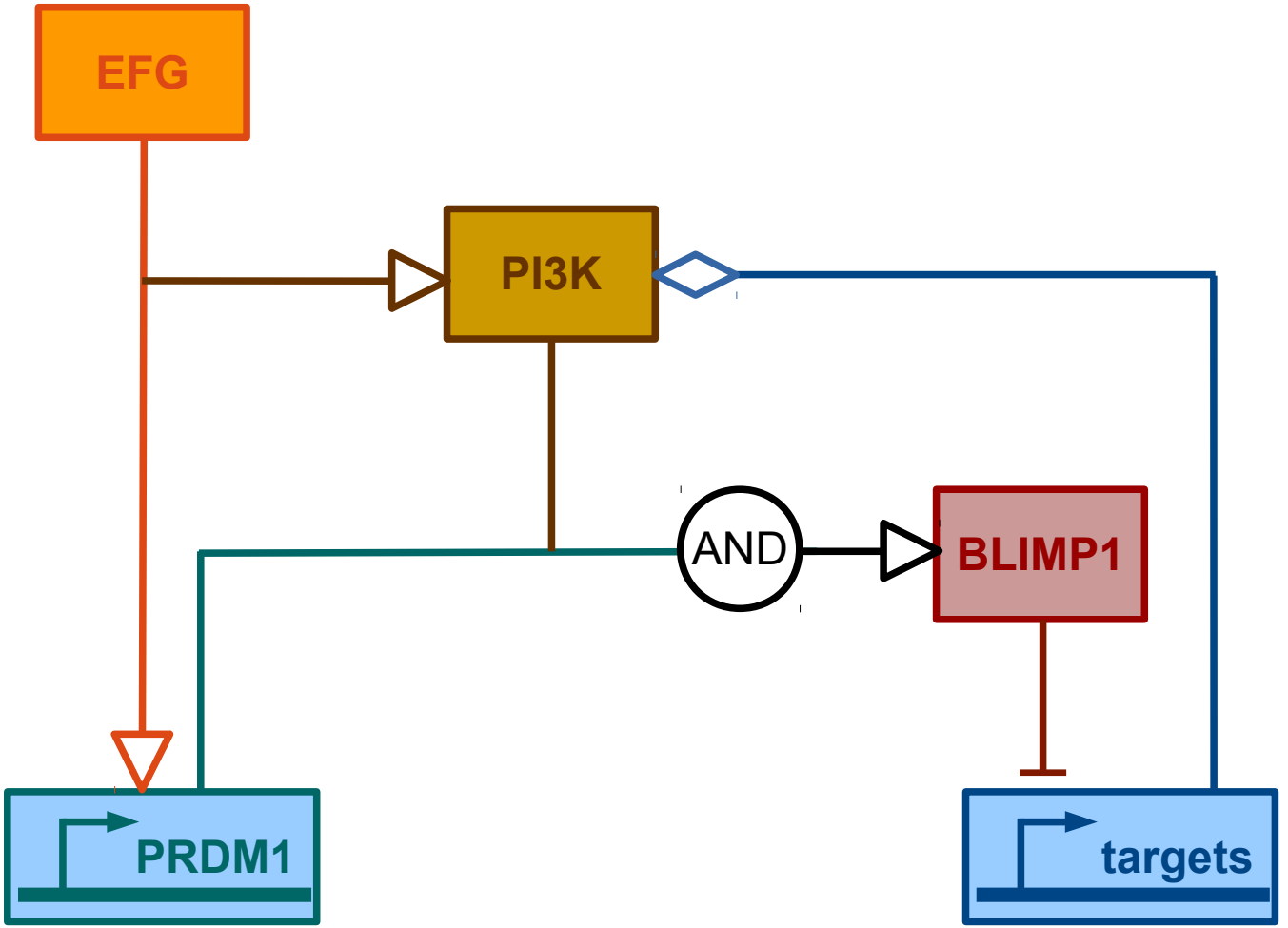
GLUL	4.56994		
NEDD4L	4.42406		Possible inhibition of PI3K phosphorylation (Kovacevic et al. 2013)
VWA5A	2.90631		Breast tumor suppressor
PPP1R3B	2.76543		
CIR1	1.69659		
ZNF737	1.37612		
LPXN	1.20711		Interacts with tyrosine kinases (upstream of PI3K)
FAM134B	1.06833		
GAB2	0.844677		Involved in the activation of PI3K (Gu et al. 2000)
EPHX1	0.657508		

EGF → PRDM1 expression
PIP3 → BLIMP1 function





Babraham Institute


EFG
PI3K
AND
BLIMP1
PRDM1
targets

Summary

1) Most effects of EGF on gene expressions are not mediated by PIP3 (not surprising)

2) Expression of a very large number of genes is affected by PIP3 perturbations: “Butterfly effect”

3) Different perturbations affect different gene populations

4) Subset of coherent effects: “static” cellular functions are EGF-insensitive, while “dynamic” are EGF-sensitive

5) Blimp1 is identified as a new TF downstream of PIP3

6) Blimp1 targets form a transcriptional feedback loop on PIP3 signalling

EMBL-EBI

Le Novère group

Martina Froehlich
Vladimir Kiselev
Pınar Pir
Nicolas Rodriguez

UCL

Luscombe group

Elodie Darbo
Raphaelle Luisier
Kathi Zarnack

Babraham Institute

Stephens/Hawkins group

Véronique Juvin
Mouhannad Malek

BBSRC
bioscience for the future

Bioinformatics team

Simon Andrews
Anne Segond-Pichon

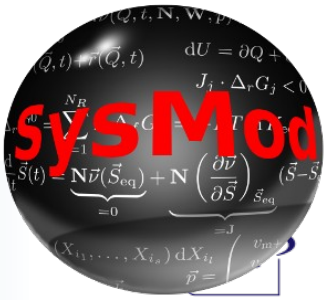
AstraZeneca

ISCB Community of Special Interest
Computational modeling of biological systems

SysMod aims at bridging the gap between bioinformatics and systems biology modeling

- Bioinformatics network inference to build models
- Transcriptomics and proteomics to parametrize/constraint models
- Communication between systems modellers and bioinformaticians to build models of whole cells, organs and organisms
- Successful PBPK, QSP/T will only work if pharmaco/toxicogenomics collaborate with drug discovery
- Precision medicine requires bioinformatics-based analysis of patient data and model-based predictions of treatments

http://sysmod.info **@cosi_sysmod**
sysmod-coord@googlegroups.com

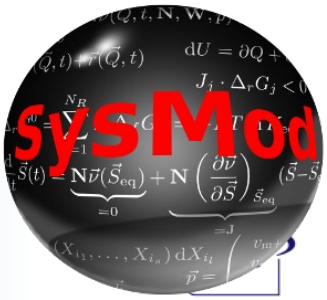


SIG meeting

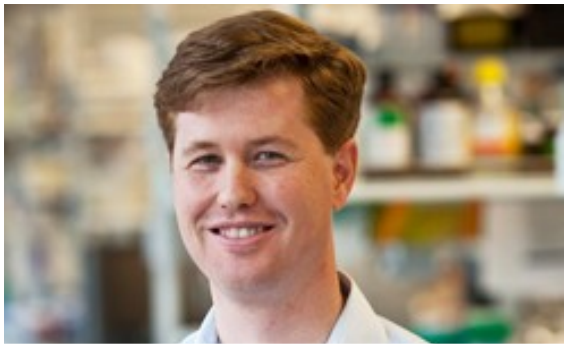
July 9th 2016; Lecture, poster and discussion sessions

Will welcome the presentation of all all types of modeling applied to any biological question. This includes, but is not limited to:

- chemical kinetics
- reaction-diffusion models
- constraint-based reconstruction and analysis
- Multi-agents
- qualitative models
- hybrid models
- multi-scale approaches
- PBPK/PD modeling
- efficient solvers and algorithms
- visualization techniques



Vassily Hatzimanikatis
(EPFL, Switzerland)
systems biotechnology, bioinformatics,
complexity of biological systems.

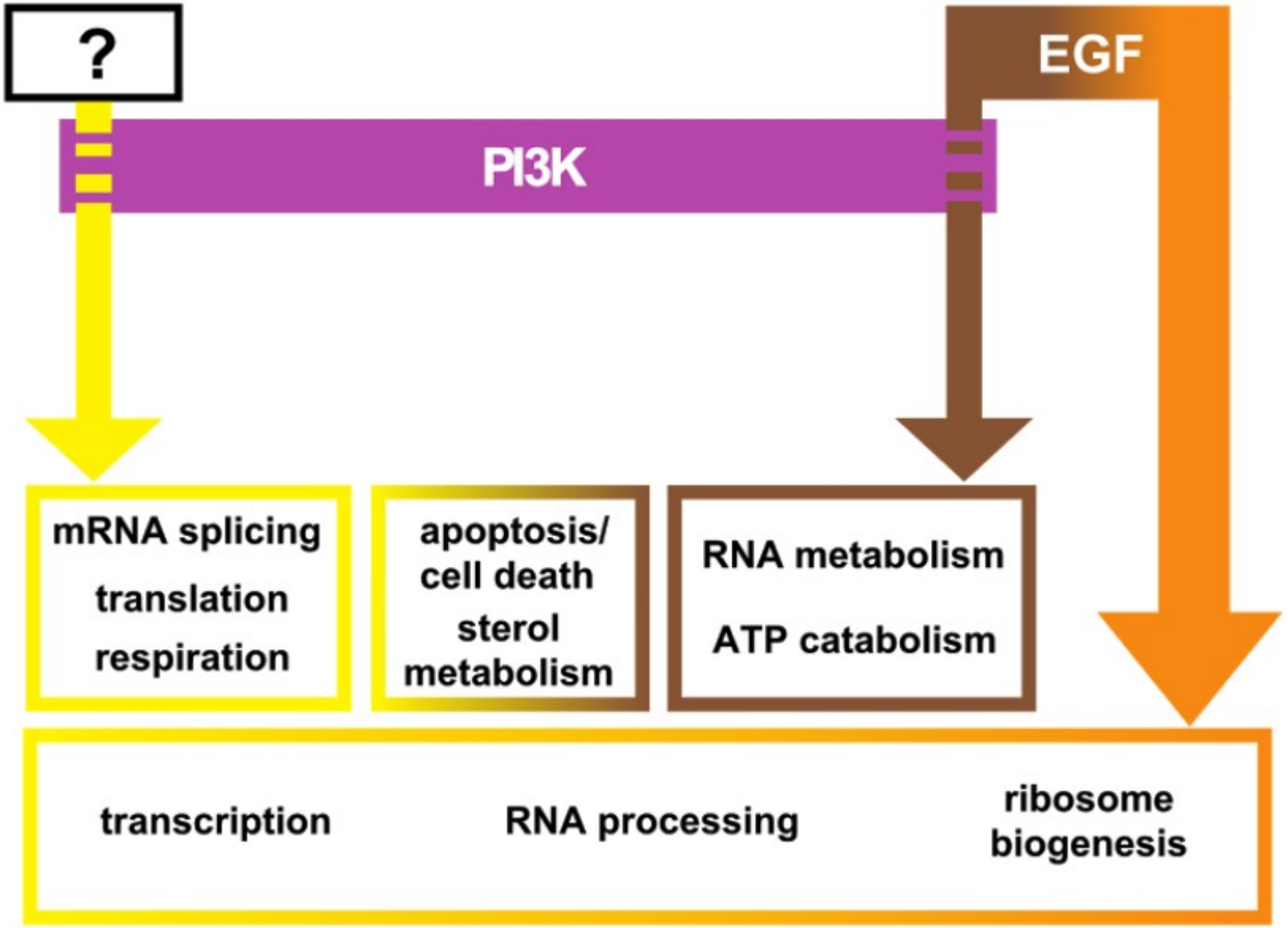


Nathan Price
(Institute for Systems Biology, Univ of Washington, USA)
integration of bioinformatics and systems modeling.
combining whole-genome network models with genomic data

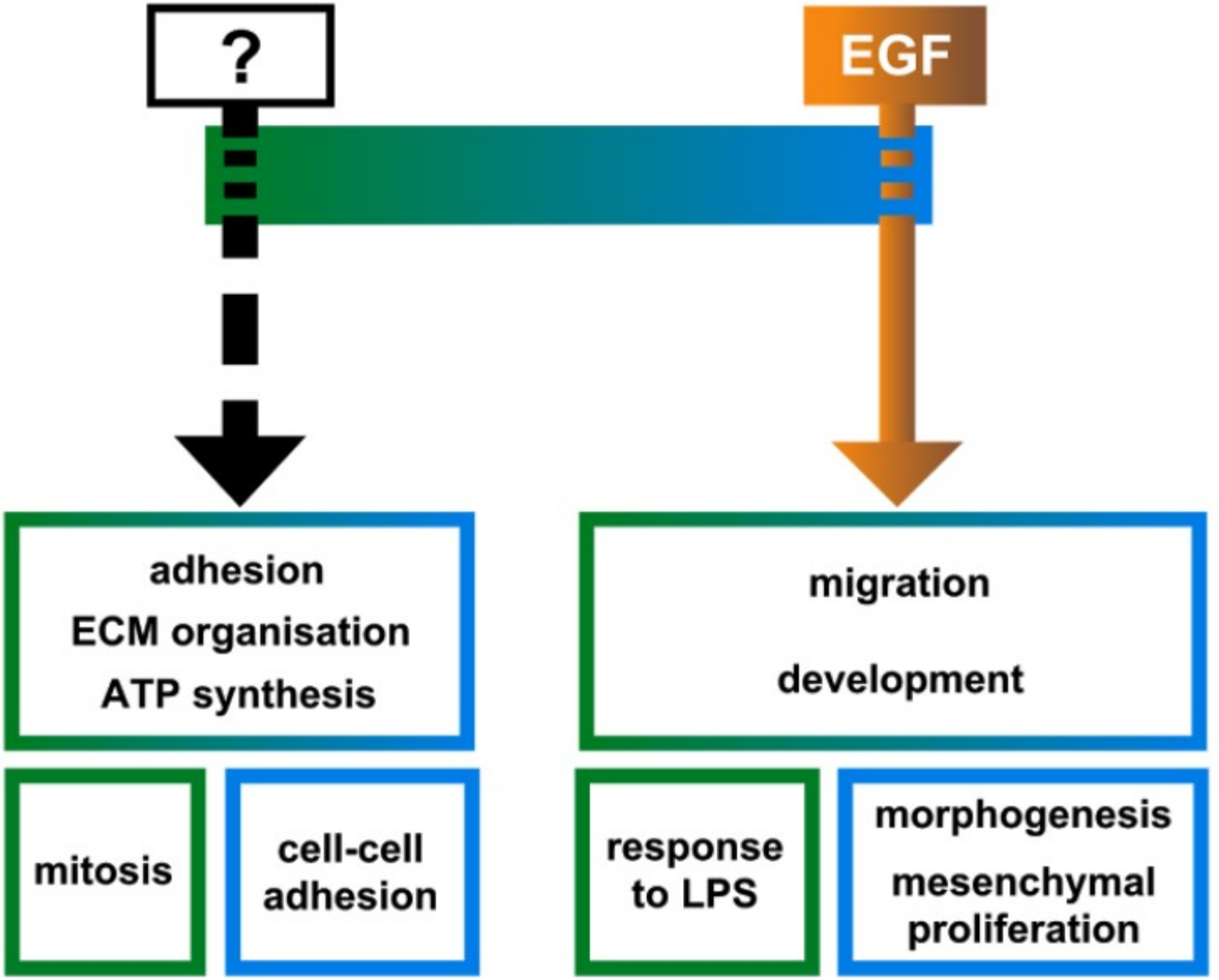
Vincent Danos
(CNRS, France and University of Edinburgh, UK)
rule-based modeling, creator of the Kappa language

Ioannis Xenarios
(Swiss Institute of Bioinformatics, University of Lausanne)
Combining logical modeling and bioinformatics.
Head of SwissProt database

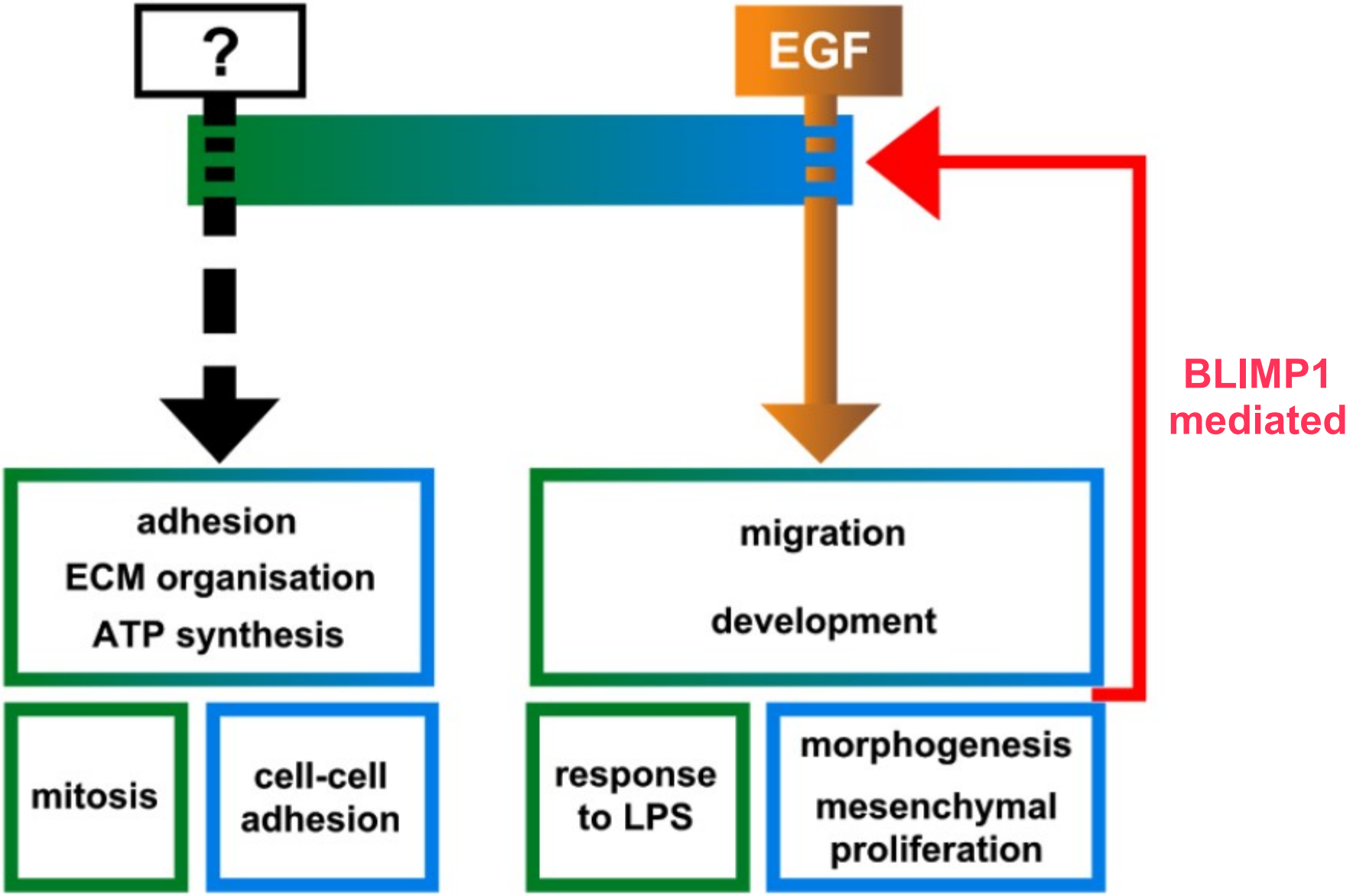
Acute perturbation of PIP3 signalling



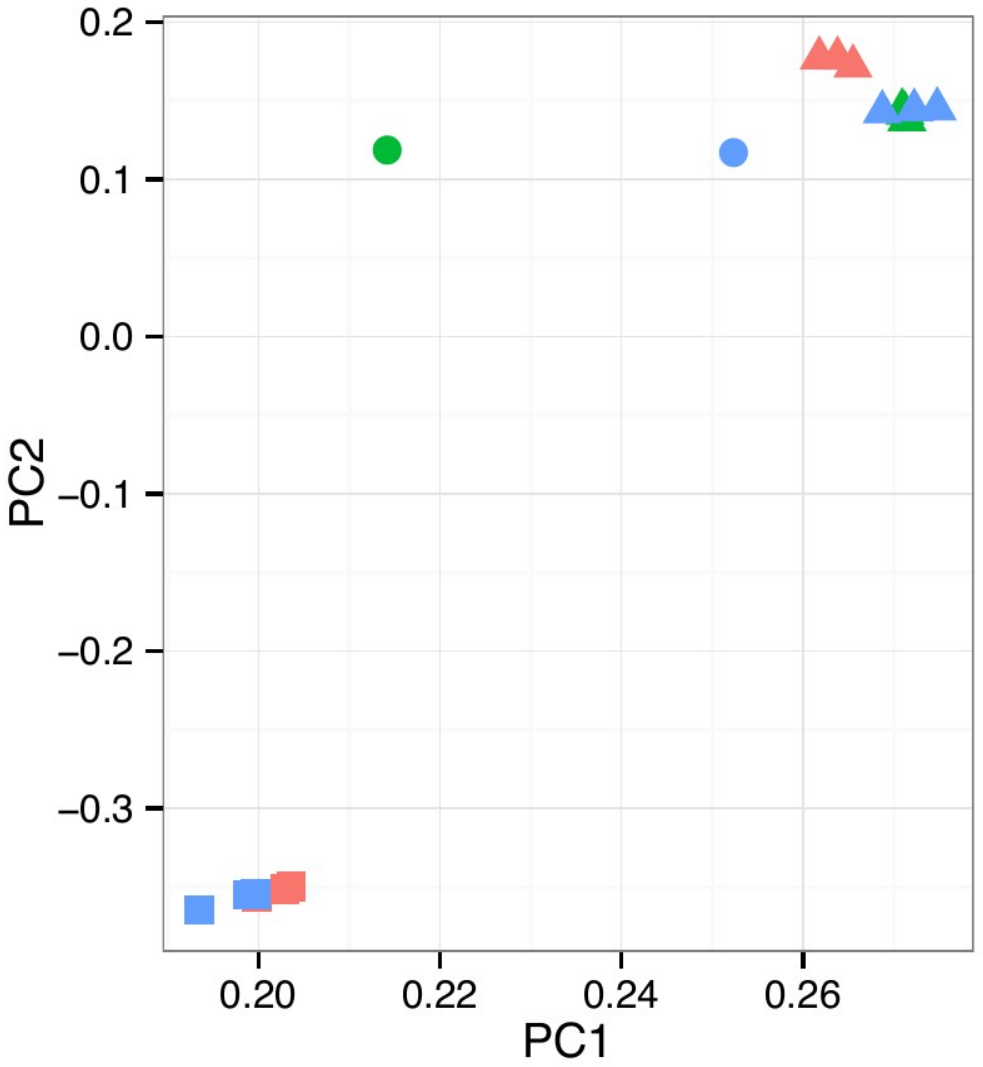
Chronic perturbation of PIP3 signalling



Chronic perturbation of PIP3 signalling



Biggest source of variability is the lab ...



Paper

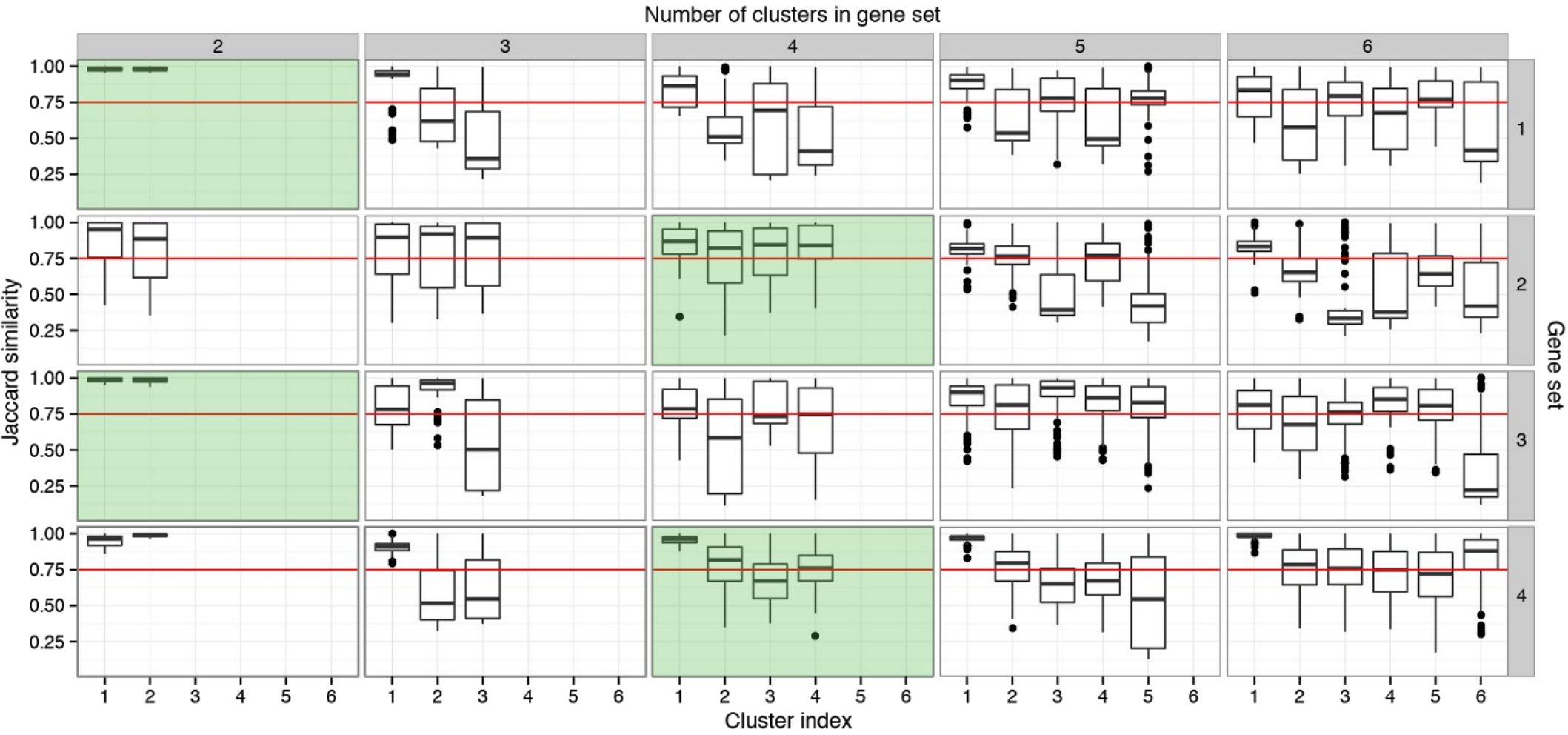
- Klijn *et al. Nat Biotechnol*. 2015
- Ours
- Hart *et al. PNAS* 2015

Condition

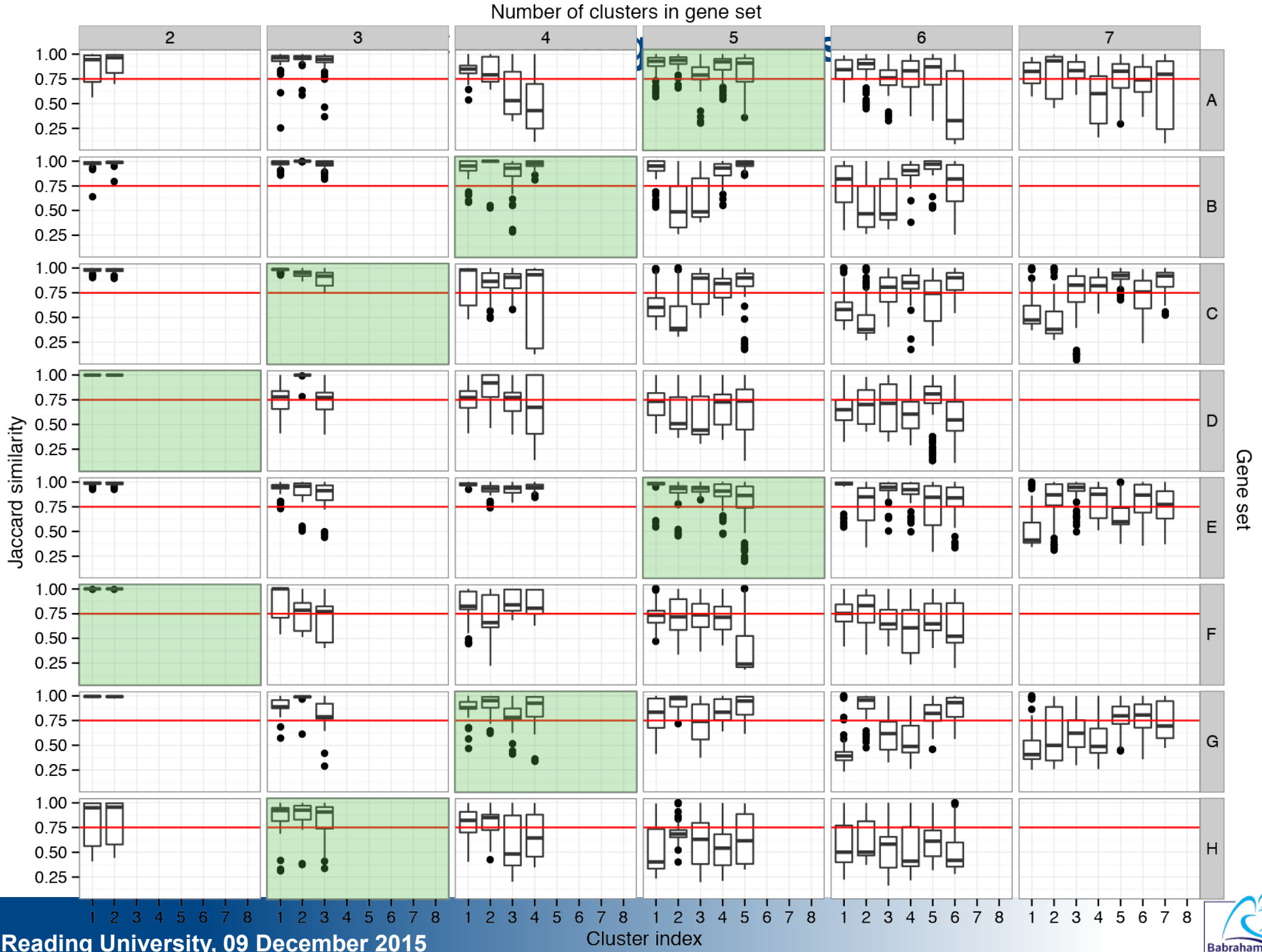
- H1047R
- PTEN–/–
- WT

All cells are isogenic MCF10a

Clustering A66



Babraham Institute



Babraham Institute