Last part: Putting it all together

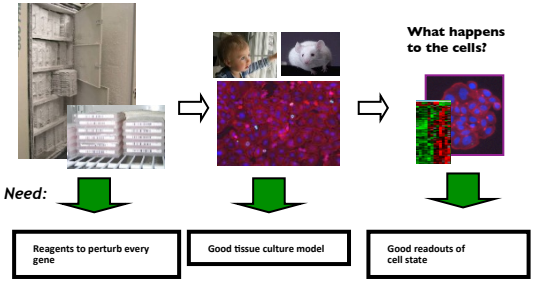
Define project

- Define question: "Find genes that do XYZ"
- Define biological model system
- Define assays to read out phenotypes of interest

Primary screen – feasibility and execution

- Optimize model system, assay(s); positive and negative controls
- Select gene set to interrogate

  Execute pilot and primary screen – select hits



Follow up on interesting genes/pathways

- Confirm assay result
- Confirm target gene specificity – multiple RNAi reagents, target KD
- Elaborate the biological effects,

  e.g. mechanism generality/context, biomedical sig?

Focus on follow-up ‘Figures 3-7’

- Project timelines
- Paths to ‘validation’ – what does validation mean?
- More detailed follow-up studies

Timelines: be realistic!!

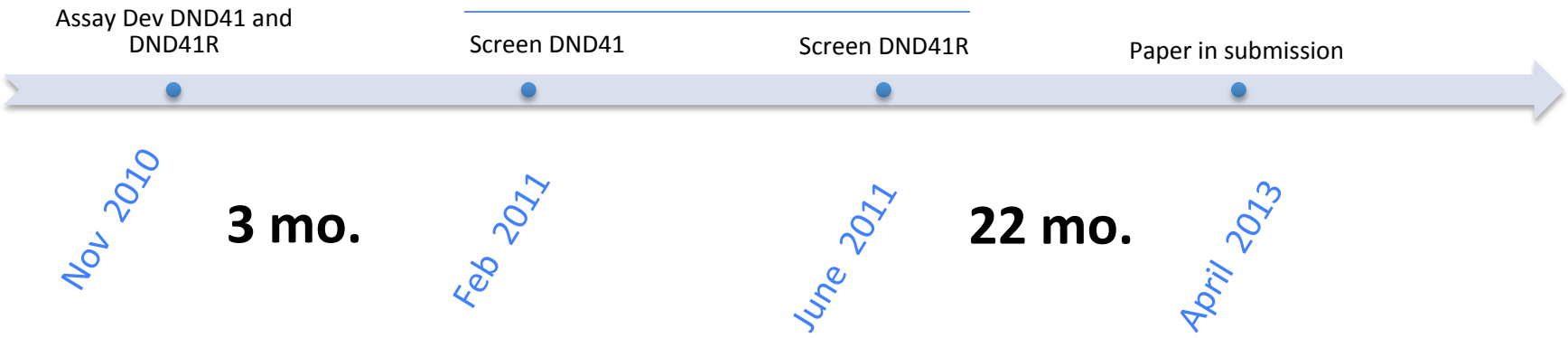
A typical project lasts 18-36 months - the screen itself taking up a small percentage of the time.

Most time-consuming part of a screening project, by far, is validation and follow up after the screen.

2-4 months 0.5 – 2 months 2– 6 months 12-24 months

Real examples

Screen chromatin regulators in 2 cell lines



ORF genome-wide screen in 1 cell line 4 drugs

Screen ‘validation’

Paths to 'validation' – what does 'validation' mean?
….two kinds of things to validate:

The phenotype

- Repeat original assay
- In different 'orthogonal' assay(s) for phenotype
- In different cell model systems

VALIDATION

The gene(s)

- Is phenotype due to expected effect of the shRNA or ORF?

e.g. discount off-target effects, non-specific effects

FOLLOW-UP

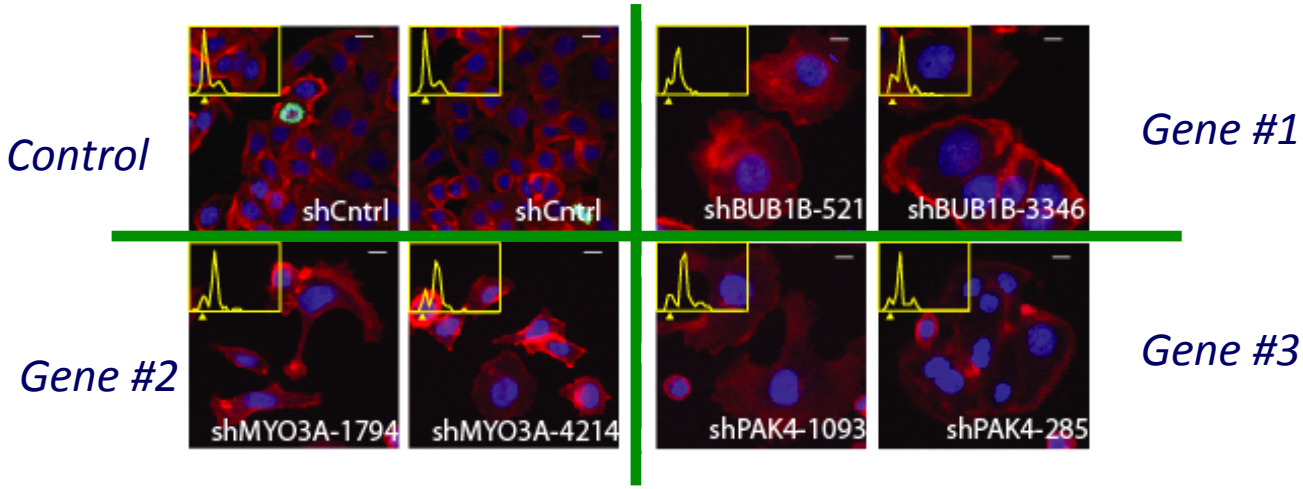
Mechanism, flesh-out process,
Relate to disease, etc.

Gene dose v. phenotype,
Enzymatic activity, structural feature?
Isoforms, protein modifications

"On-Target" confirmation for RNAi hits

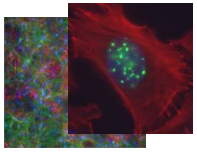
-> How to confirm that phenotype is due to target gene knockdown?

Multiple effective shRNA sequences!

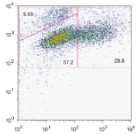


Moffat et. al. Cell 2006

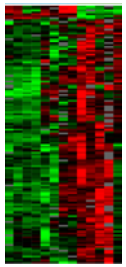
Complex phenotype characterization



- High-content (Image-based)



- FACS

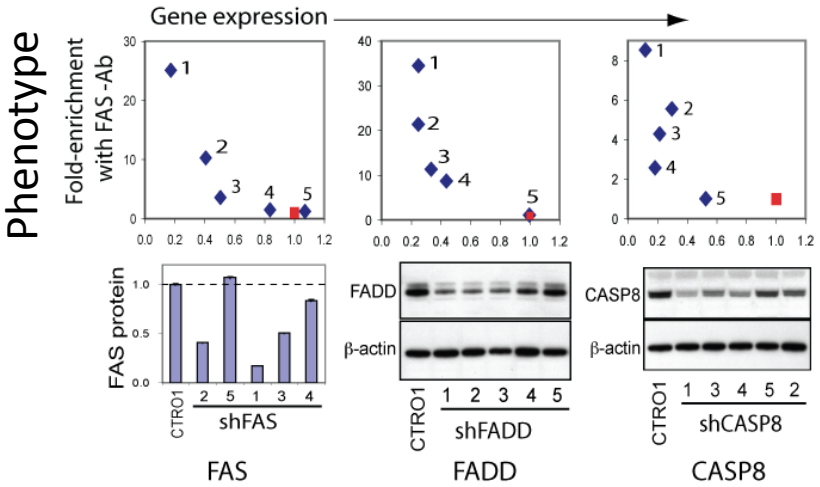


- GE-HTS

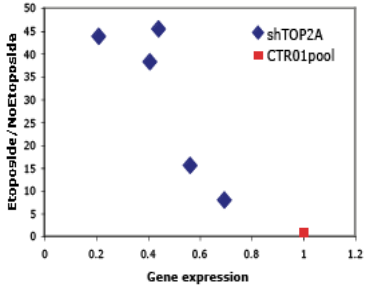
"On-Target" confirmation for RNAi hits

-> How to confirm that phenotype is due to target gene knockdown?

KD vs. phenotype correlation



Perfect Perfect Not perfect



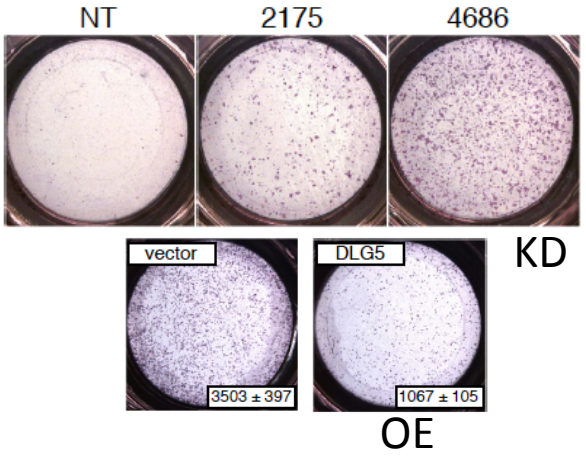
Top hit – TOP2A

BUT THIS KD v. PT RELATIONSHIP DOESN'T ALWAYS HOLD

"On-Target" confirmation for RNAi hits

-> How to confirm that phenotype is due to target gene knockdown?

Perturb gene in more than one way



RNAi v. Overexpression

DLG5 knockdown increases cell migration.

→ DLG5 overexpression decreases cell migration

RNAi v. Targeted small molecule

shRSK1 blocks shDLG5-induced cell migration

RSK inhibitor (BI-D1870) also blocks shDLG5-induced cell migration

"On-Target" confirmation for RNAi hits

-> How to confirm that phenotype is due to target gene knockdown?

1. Require multiple shRNAs for same gene to induce same phenotype. Obtain more shRNAs, siRNAs if needed.

2. Expand phenotypic characterization to show detailed agreement among hairpins targeting same hit gene (see example)

3. Determine if knockdown of target gene correlates with phenotype across the multiple hairpins…helpful when true, but not always true.

4. Perturb gene in other ways (small molecule, overexpression, genome engineering).

5. Perturb known 'relatives' of the hit gene if known (e.g. genes in same pathway).

6. cDNA rescue

Paths to 'follow – up':

What does it mean to 'learn a gene's function'?

What does it mean to 'define the genes invovled in a process'

The phenotype

- Repeat original assay
- Different 'orthogonal' assay(s) for phenotype
- Different cell model systems

VALIDATION

The gene(s)

- Is phenotype due to expected effect of the shRNA or ORF?

e.g. discount off-target effects, non-specific effects

Detailed nature of the process

Context? Tissue types, in vivo

What defines/governs process?

- more players, further assays

Mechanism' – biochemistry

Relationship to disease, etc.

Cross to other data sets

FOLLOW-UP

Detailed nature of the protein(s) and the proximal effects:

Gene dose v. phenotype,

Enzymatic activity, structural feature?

Isoforms, protein modifications

Immediate substrates, binding partners

The LONGEST part of a screening project!

Putting it all together: Project Examples

Emphasis on what's in Figures 3-7

A pooled screen for cell migration

Gromek Smolen, Daniel Haber

Smolen et. al., Genes & Dev. 2010

Endogenous negative regulators of cell migration — a pooled screen approach

Gromek Smolen
Daniel Haber Lab

CELL MIGRATION

Roles in normal development:

patterning during gastrulation
neural crest migration
heart valve formation…

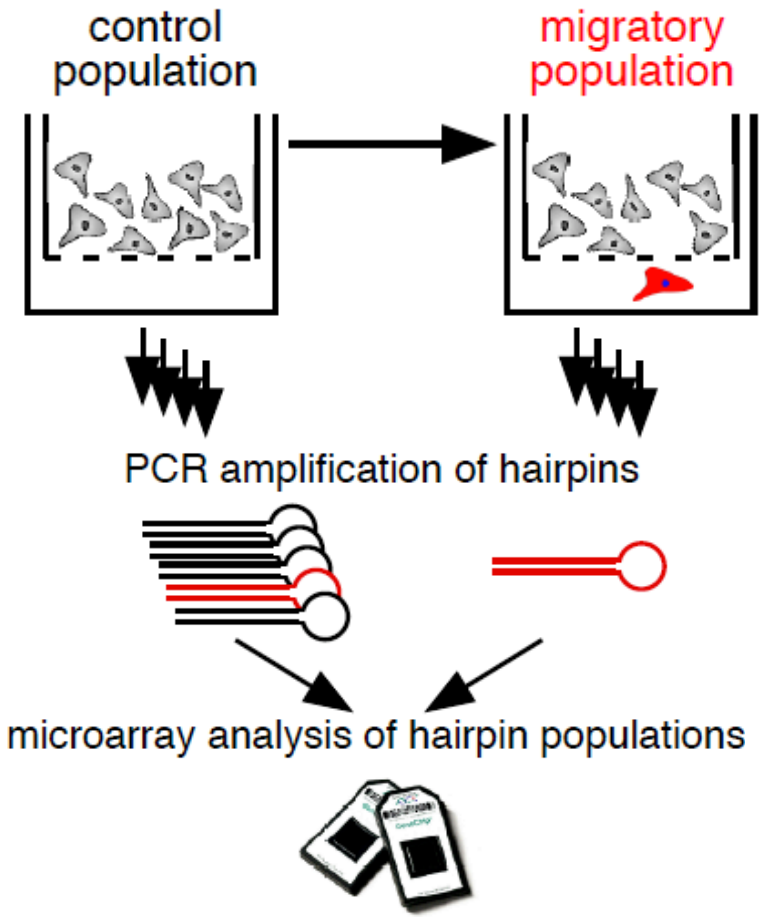
Roles in disease:

congenital birth defects
cancer metastasis…

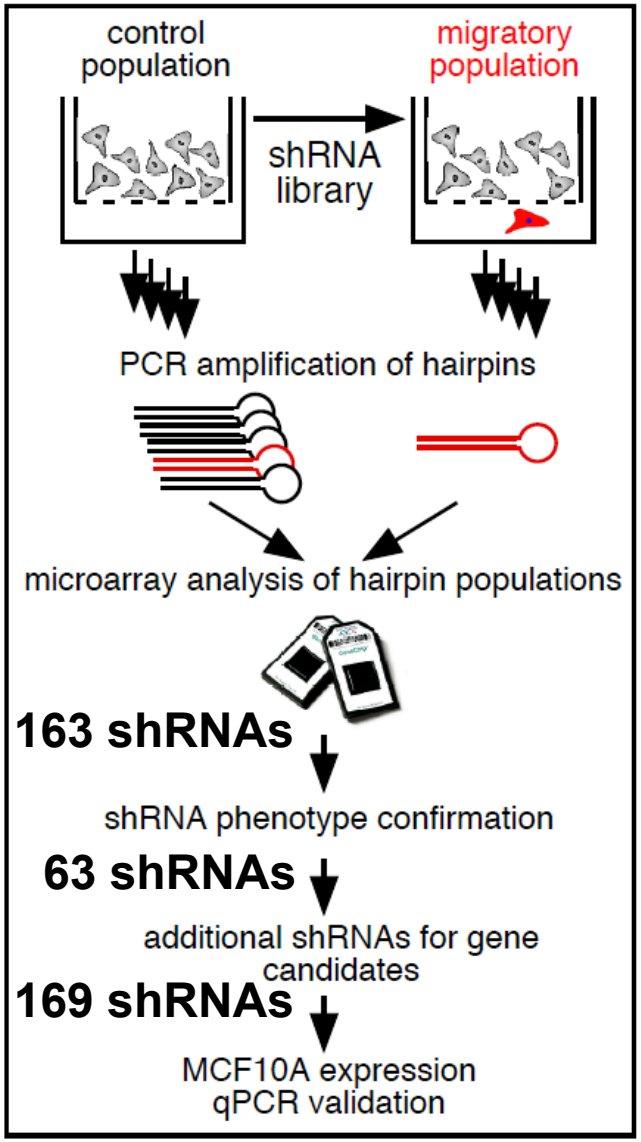
Determinants of cell migration :

morphology
ECM architecture
cytoskeleton
gene regulatory network…

Breast Cancer: MCF10A cells
Non-migratory cells



Regulators of Cell Migration Pooled Screen: Hit Gene Confirmation



SCREEN SUMMARY

55,000 shRNA constructs targeting 11,000 genes

Candidate selection criteria:

- top 1000 most enriched shRNAs in each replicate
- shRNAs enriched in at least 2 replicates
- genes with at least 2 non-overlapping shRNAs

163 shRNAs identified in the screen; 63 retested positive (39%)

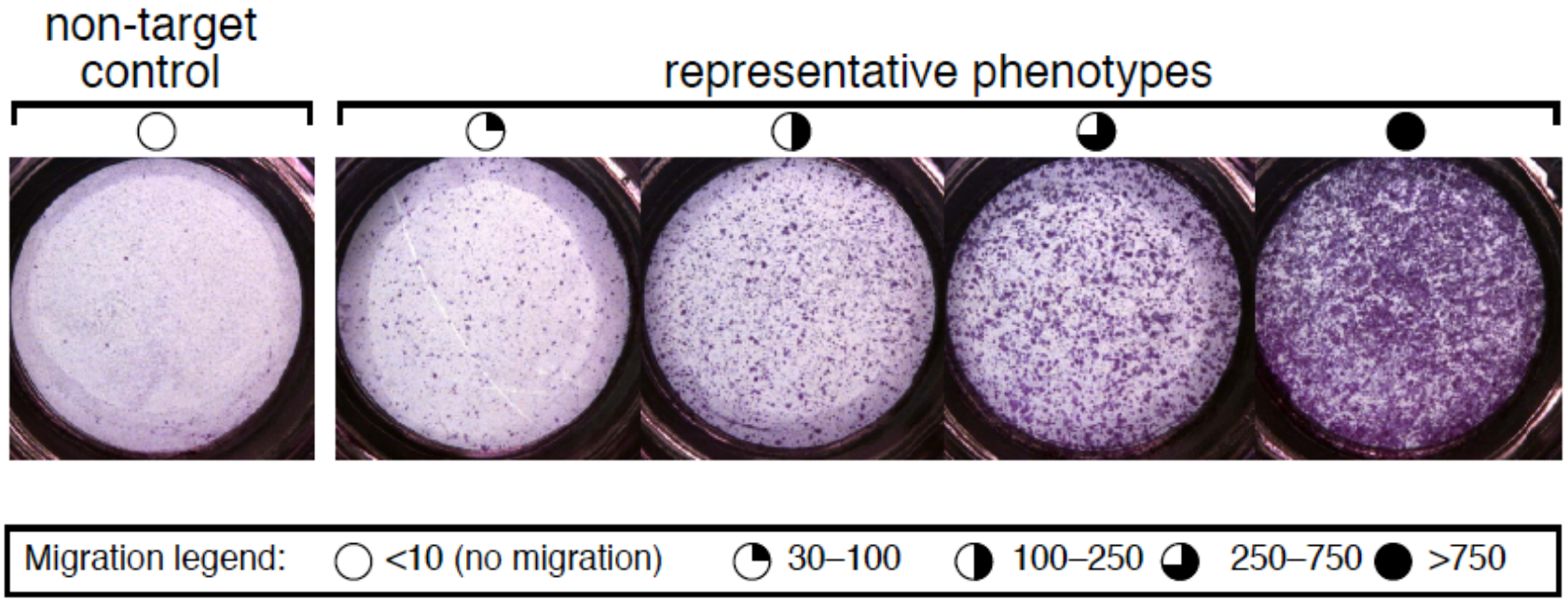
270 additional shRNAs for gene candidate tested — total 433 shRNAs; 106 positive (24%).

31 / 34 (91%) of candidate genes are expressed in MCF10A cells and show evidence of knockdown

31 hit genes

Representative migration phenotypes

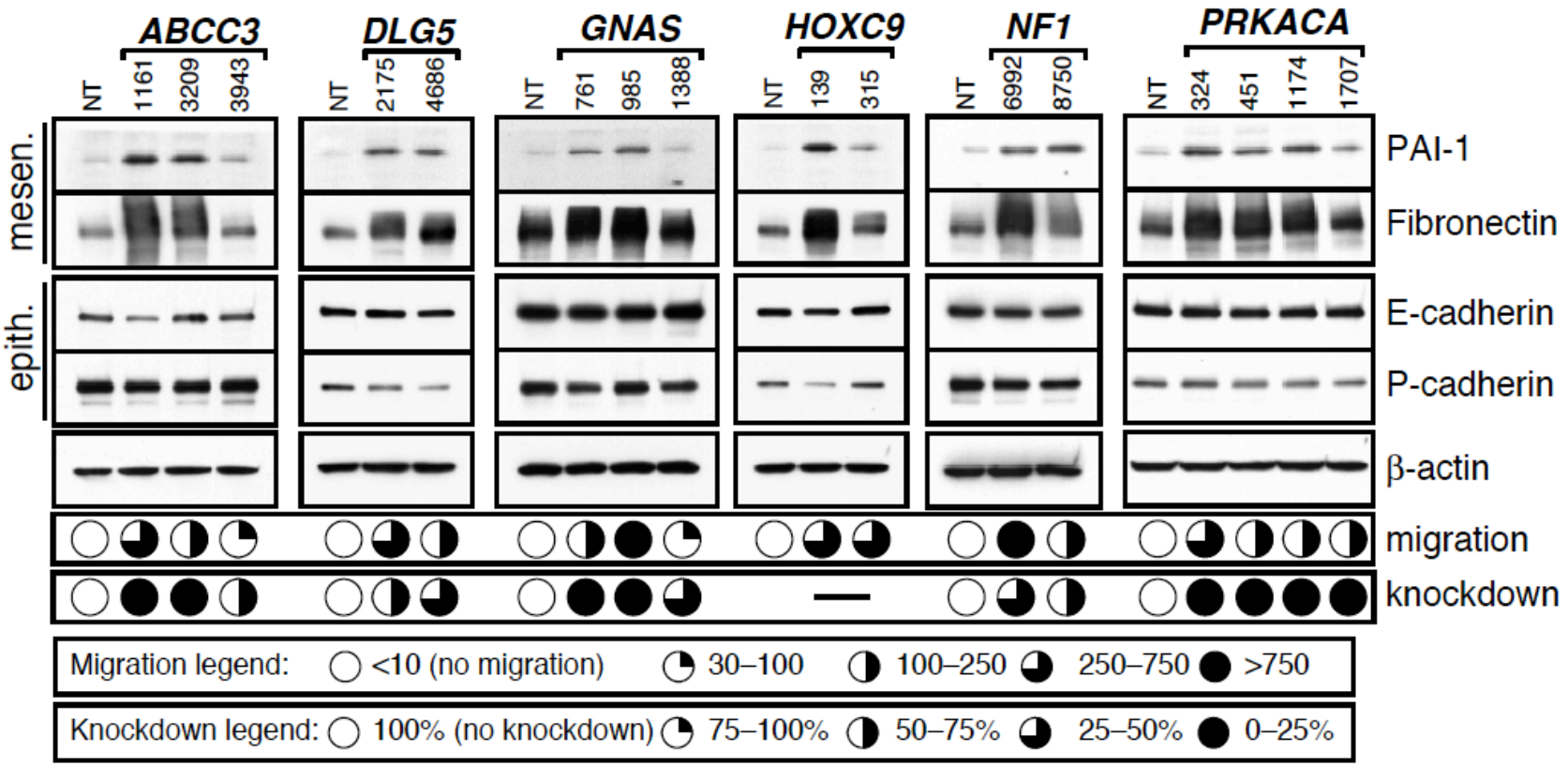
Diverse gene annotations,
Test one at a time: strong hits!



Gromek Smolen
Daniel Haber Lab

EMT marker analysis

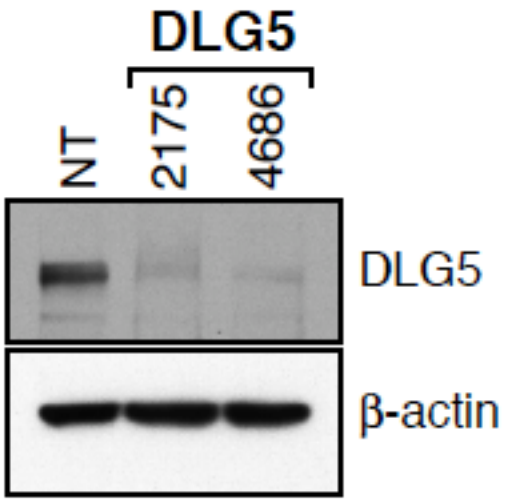
Several hits increase fibronectin but produce inconsistent reduction of epithelial markers

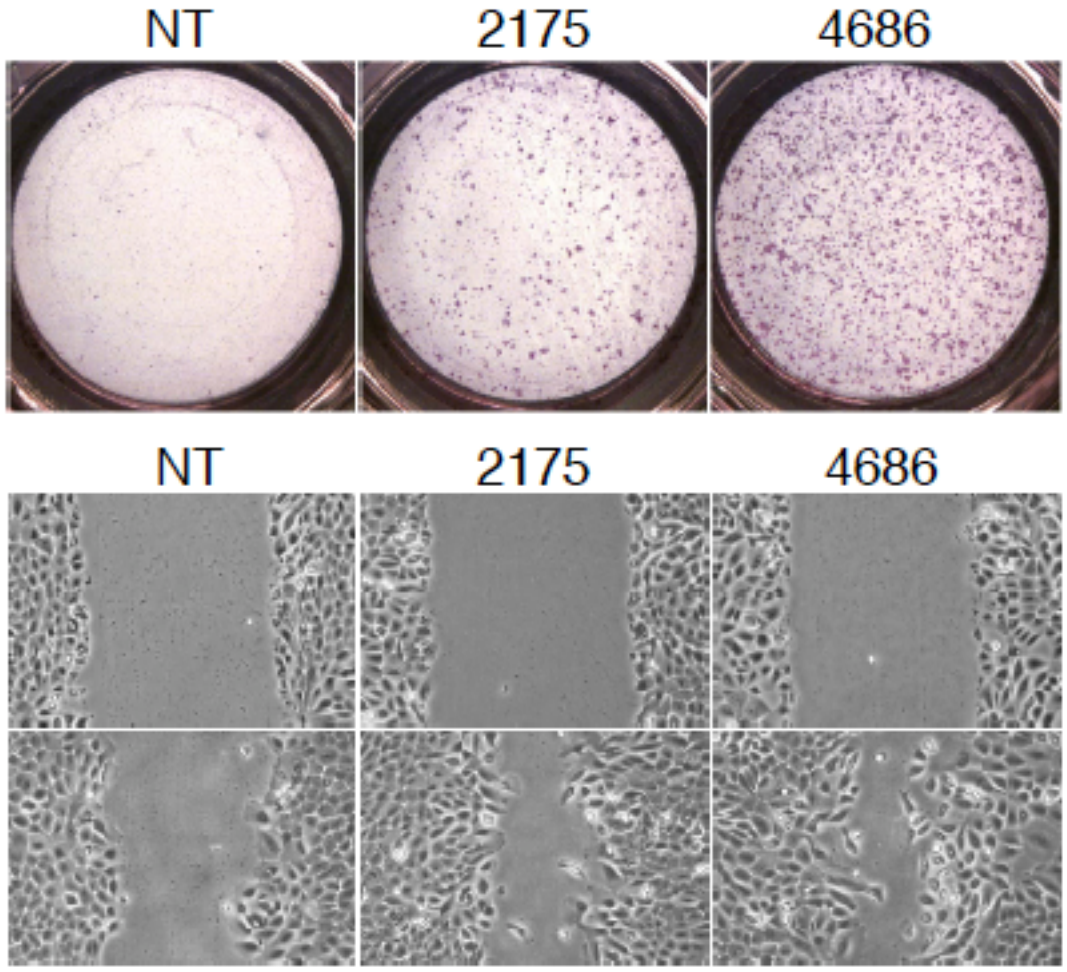


Gromek Smolen
Daniel Haber Lab

DLG5 knockdown increases cell migration

DLG: novel migration gene, downregulated upon YAP-induced migration

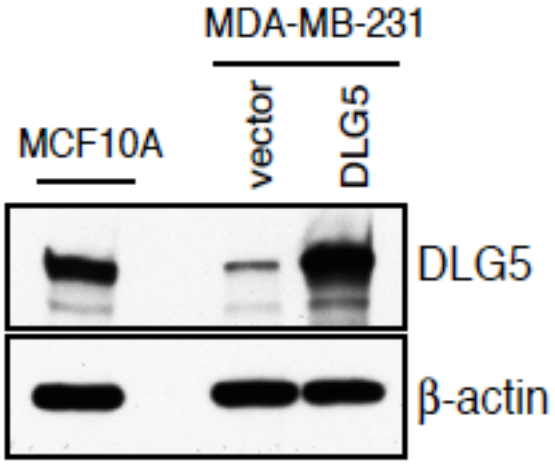


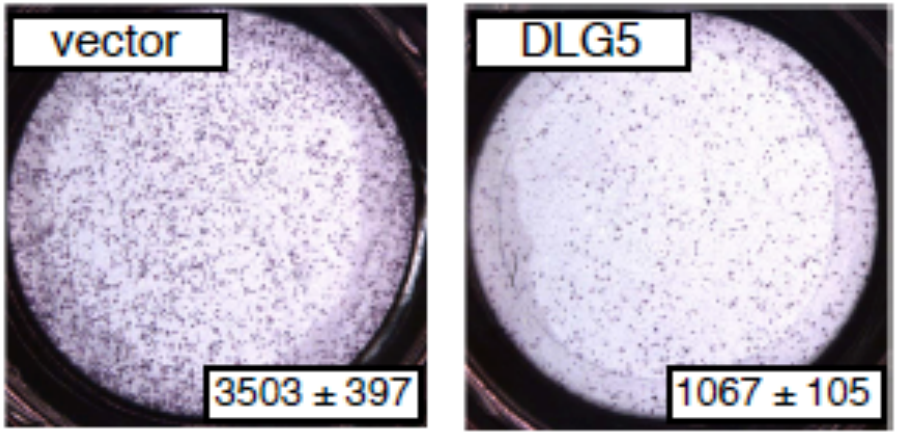


Gromek Smolen
Daniel Haber Lab

DLG5 overexpression decreases cell migration

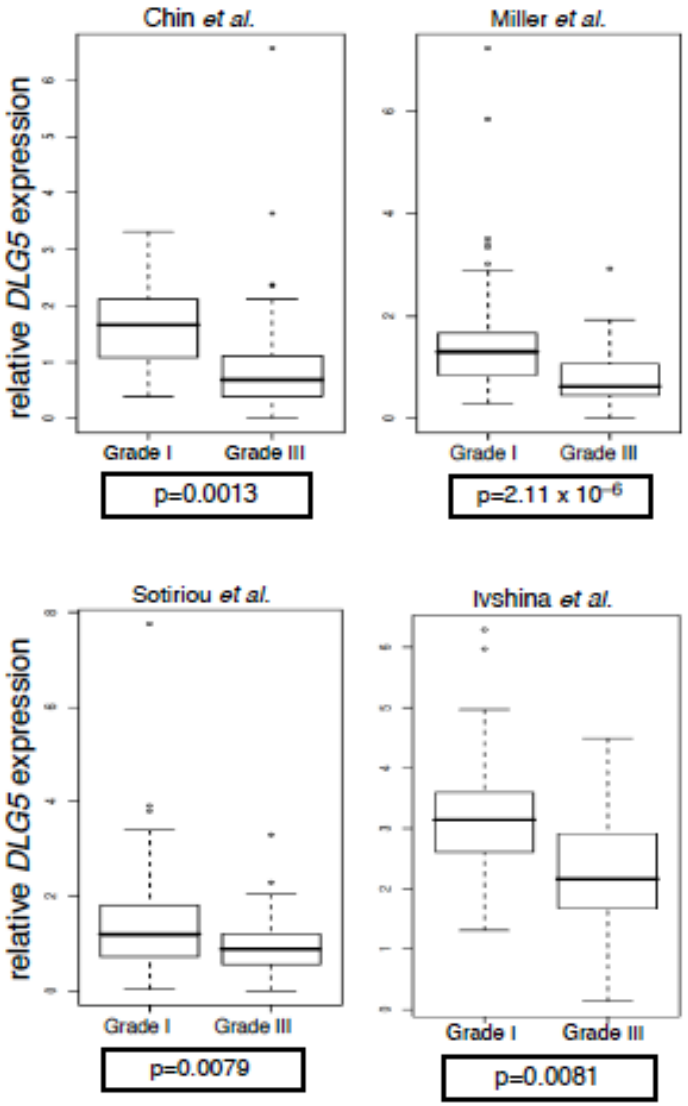
Use MDA-MB-231
Migratory breast cancer line



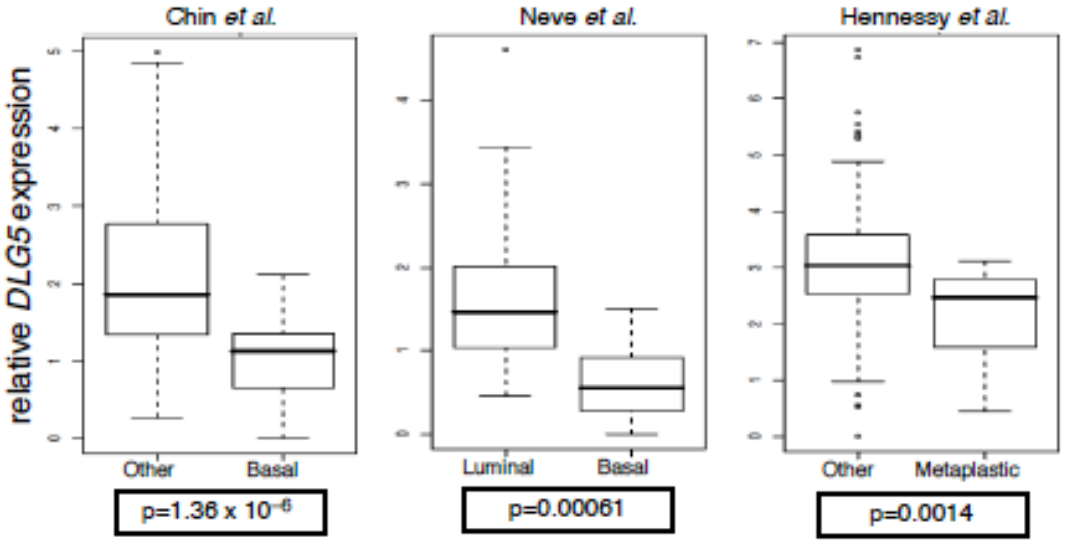


Gromek Smolen
Daniel Haber Lab

DLG5 expression in breast tumors



Compare to ‘orthogonal’ data

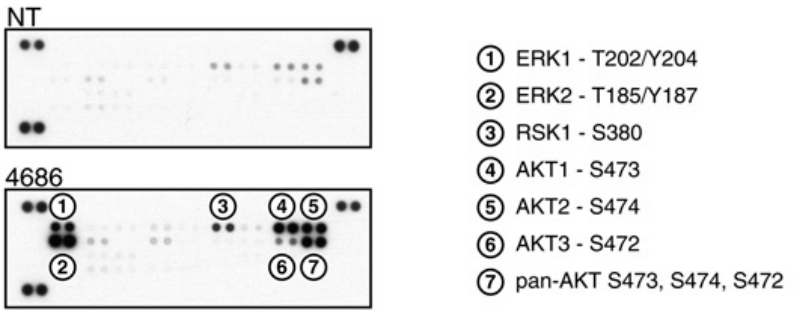


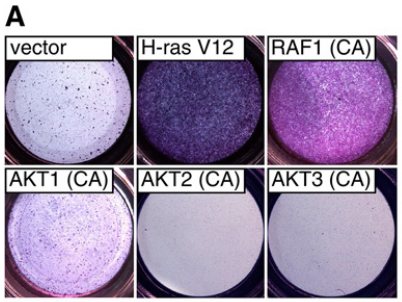
Gromek Smolen
Daniel Haber Lab

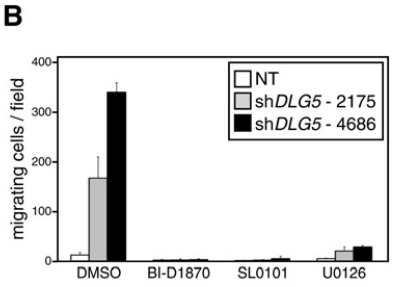
More mechanism??

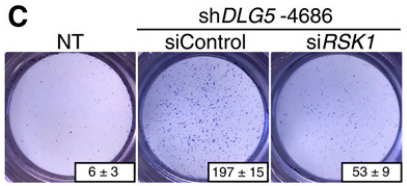
Knockdown of DLG5 → Use phospho-Ab array to monitor signaling pathways

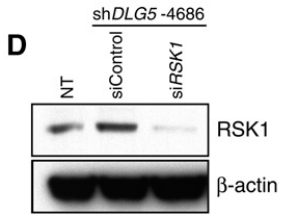
Prominent increases in phosphorylation of ERK1 and ERK2,
a downstream kinase, RSK1,
and AKT1–3 were observed











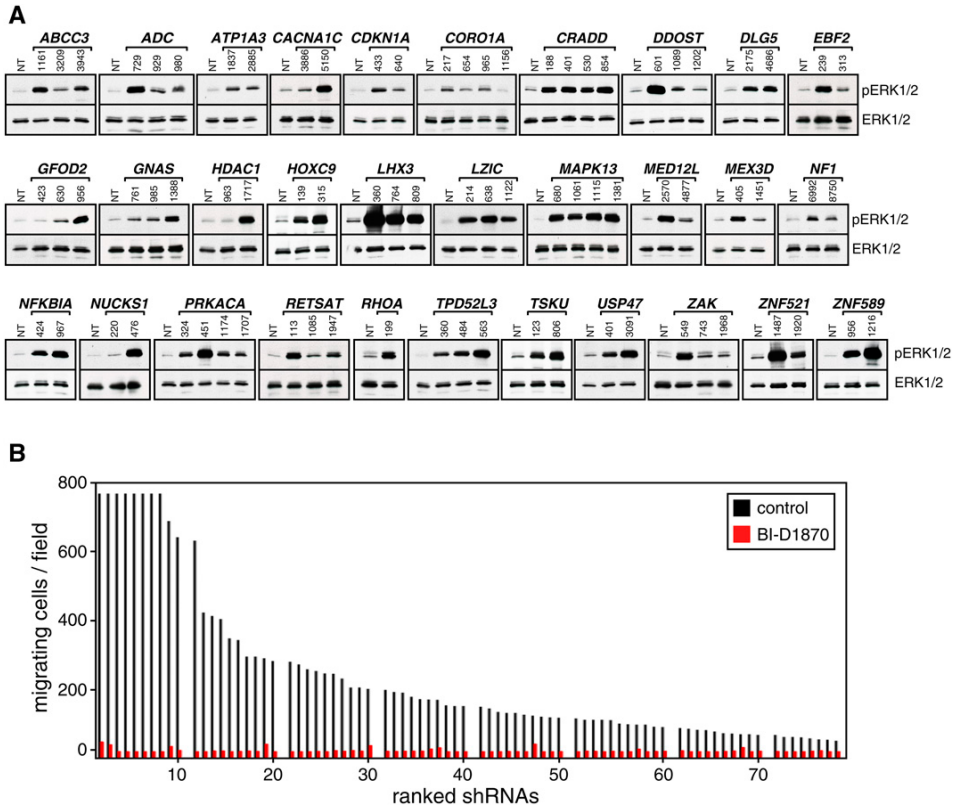
Activation of ERK1/2 caused migration.
Activated AKT1–3 did not.

ERK/RSK inhibitors blocked shDLG5-induced migration

shRSK1 blocked shDLG5-induced migration

Mechanistic convergence of 31 migration hit genes →

- All increase phospho-ERK1/2
- RSK inhibitor (BI-D1870) blocks migration induced by all 31 gene knockdowns



phospho-ERK1/2 up

RSK inhibitor suppresses Migration by all hit shRNAs

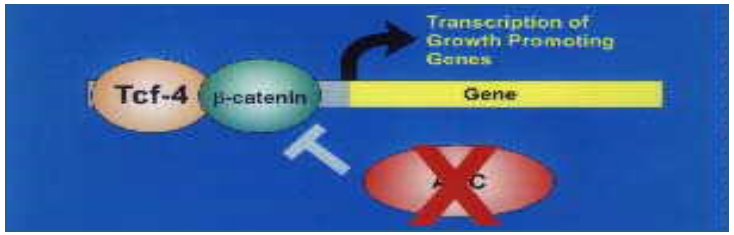
Integrated functional genomic approach to cancer:

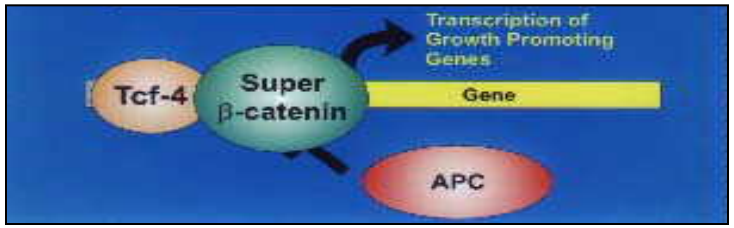
Oncogene discovery in colon cancer

WNT signaling activation in nearly all colon cancer

1. APC loss (85% of colon cancers)

2. GOF mutations in β-catenin (5-10%)

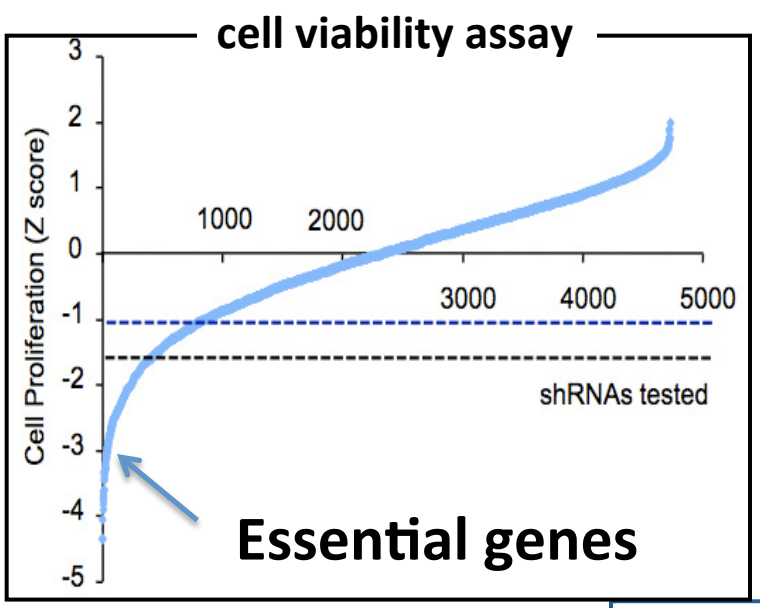




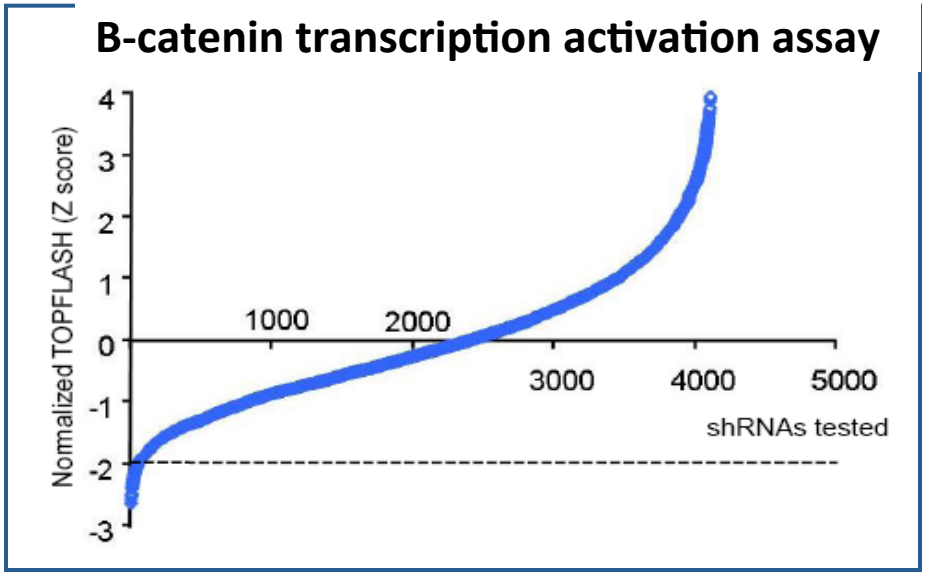
Ron Firestein, Bill Hahn
Nature Sept. 2008

B-catenin assay and cell viability assay – Two phenotypes

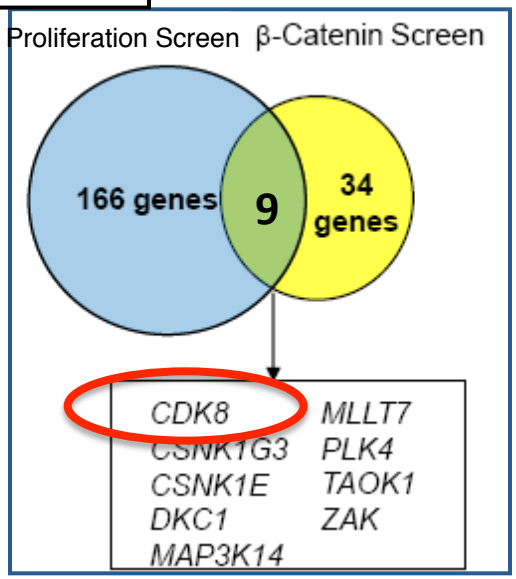
1000 genes – 5,000 shRNAs - 95% of all human kinases



So Young Kim
Ian Dunn



Ron Firestein



Proliferation screen

β-catenin dependent cells (HCT116)

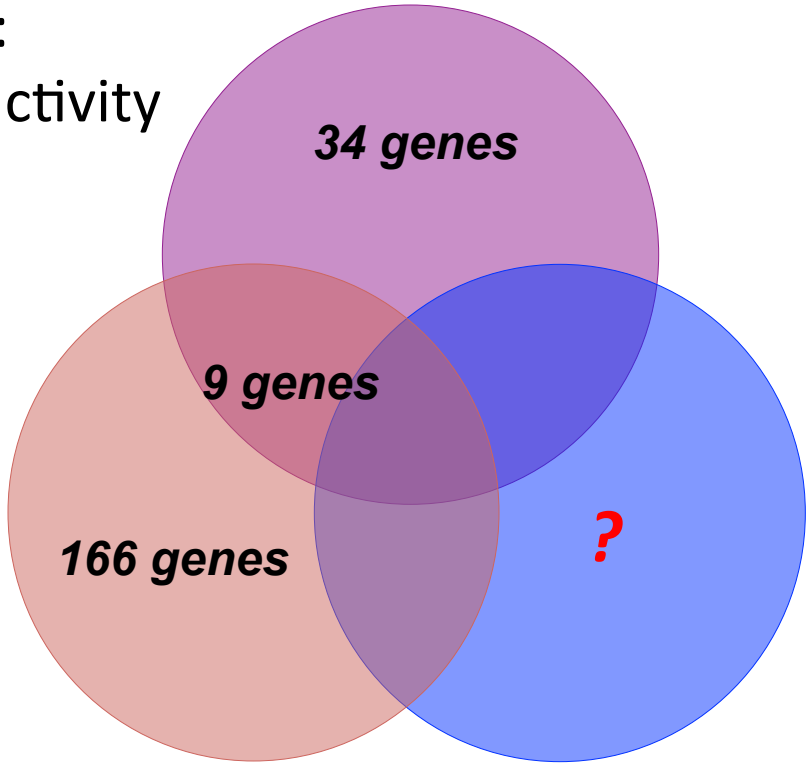
B-catenin screen

β-catenin dependent cells (DLD-1)

Firestein, Hahn

Compare to 3rd dataset - very different type

RNAi:
β-catenin activity

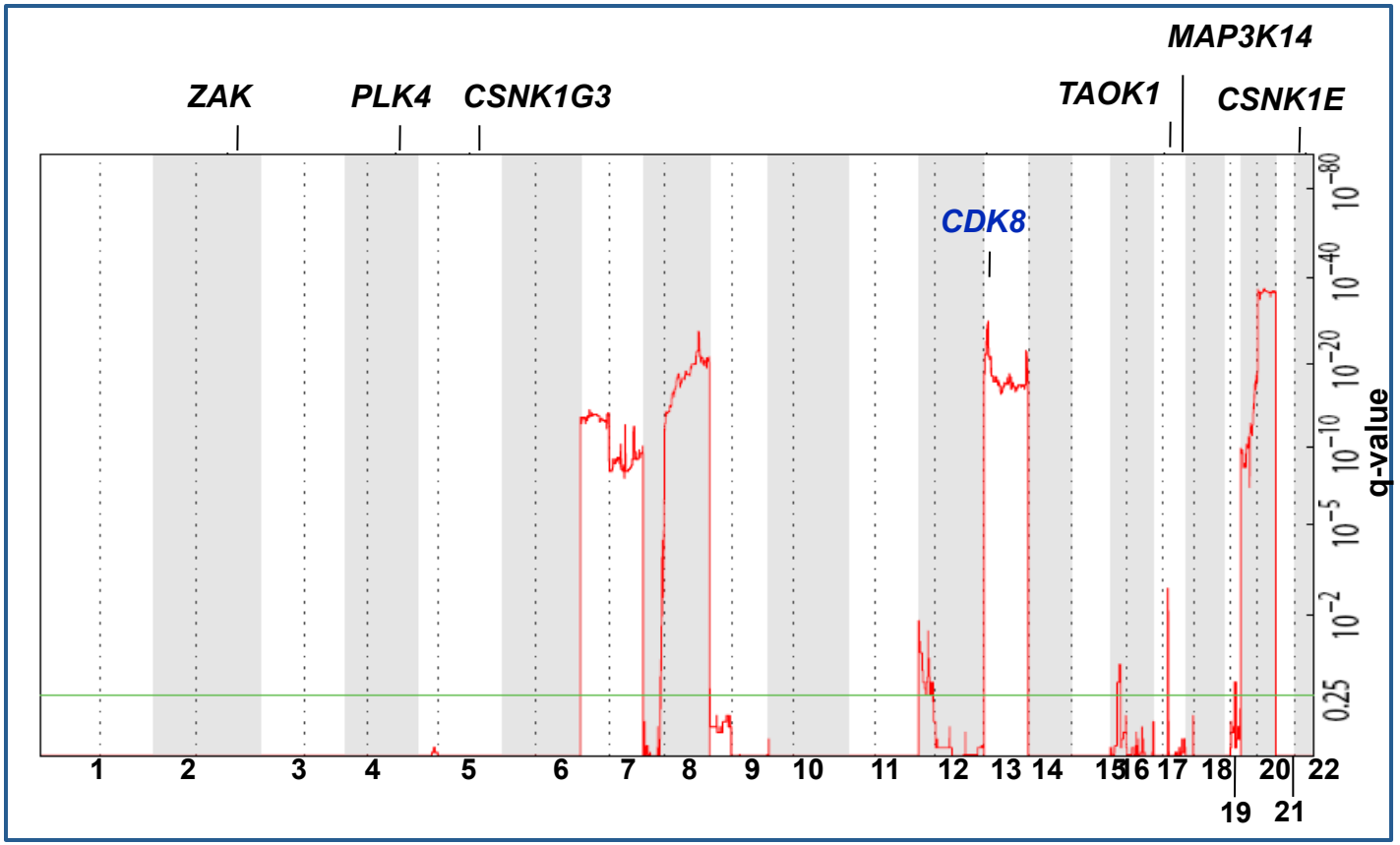


RNAi: proliferation/viability

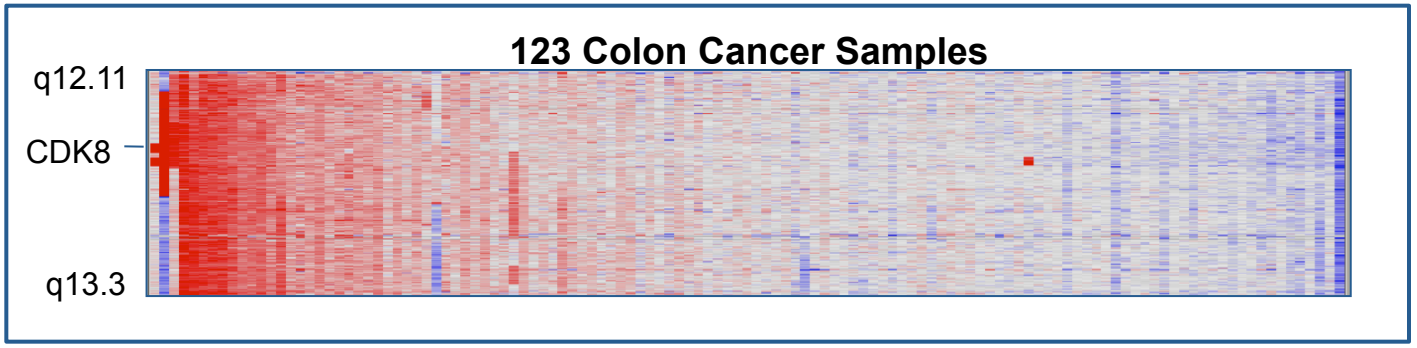
Amplified in colon tumors

CDK8 is amplified in colon cancers

SNP ARRAY (GISTIC)



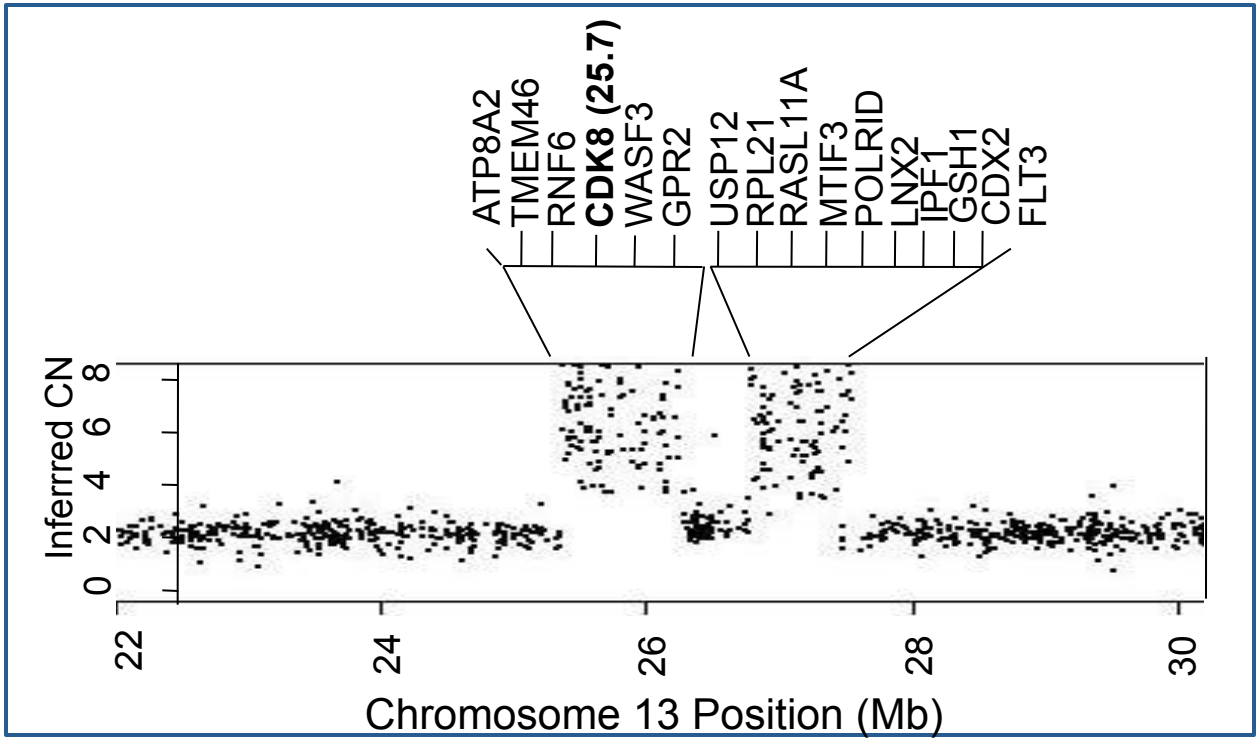
HEAT MAP AT
13q12.11-13.3



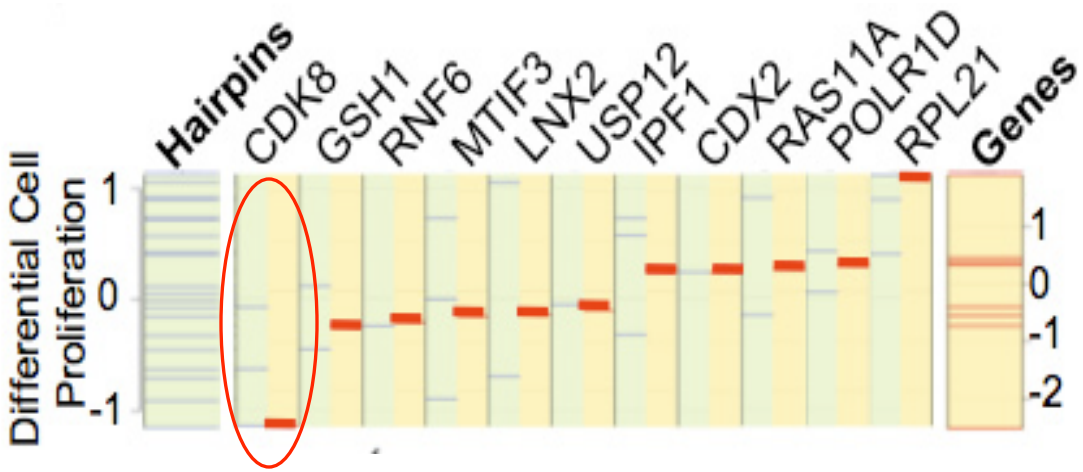
Adam Bass, Matthew Meyerson

Identification of Minimal Region of Copy Gain at 13q12

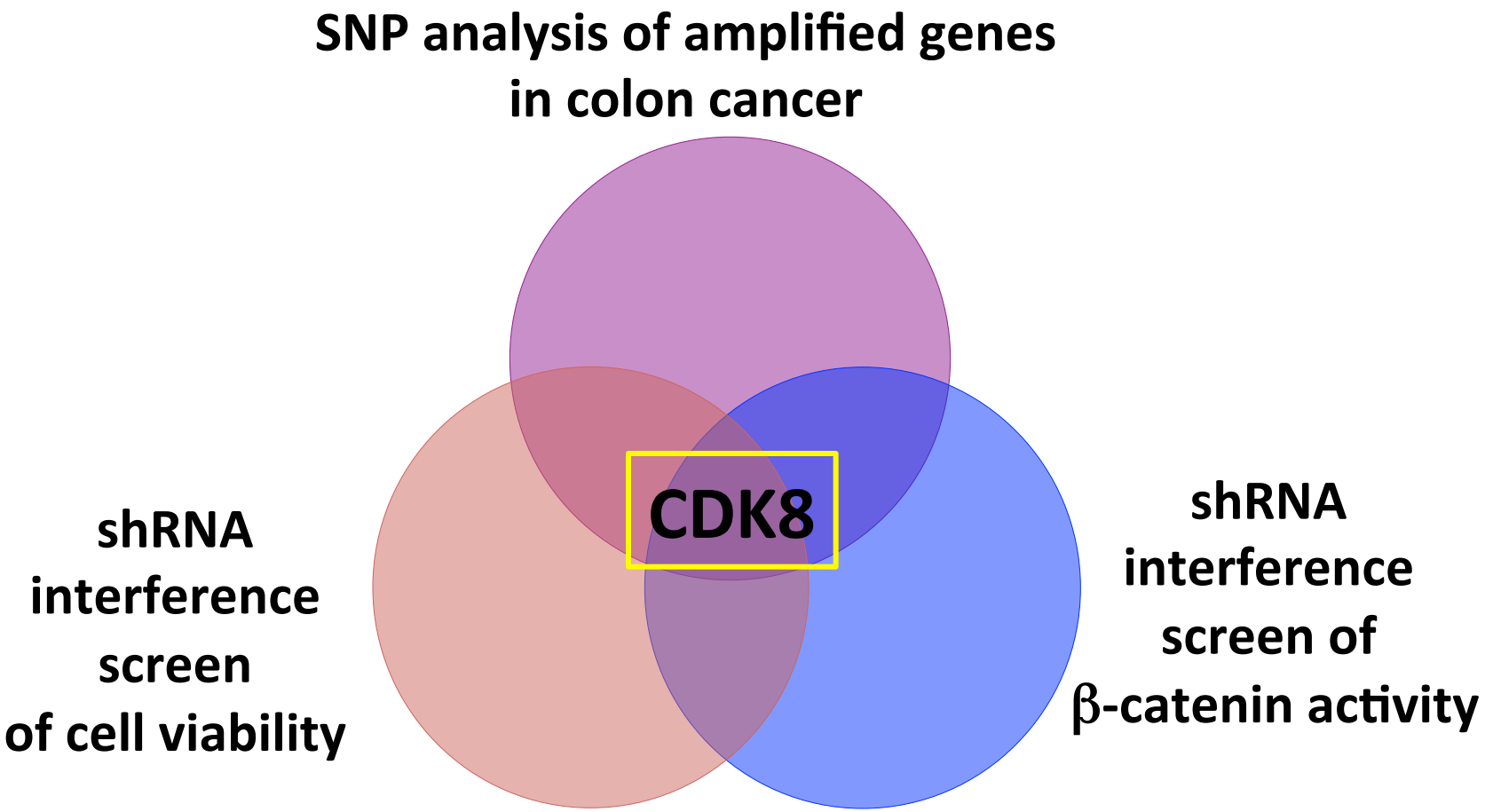
Copy number



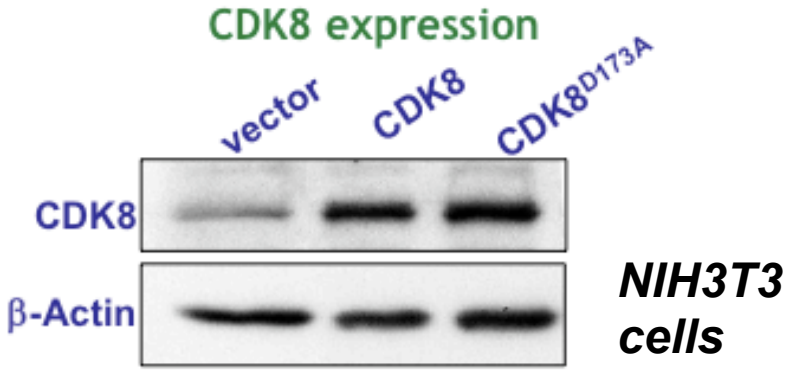
Proliferation/RNAi



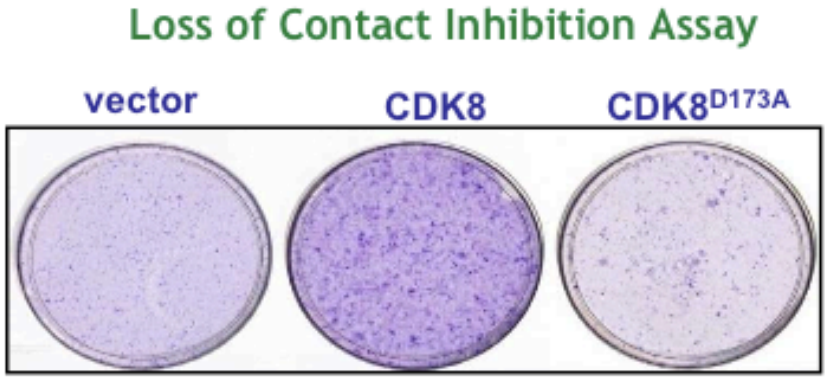
Novel oncogene CDK8

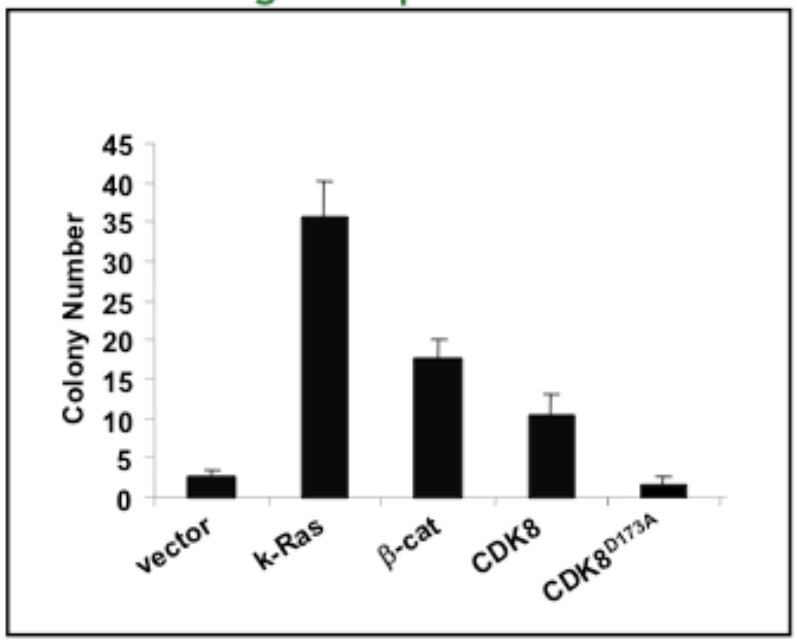


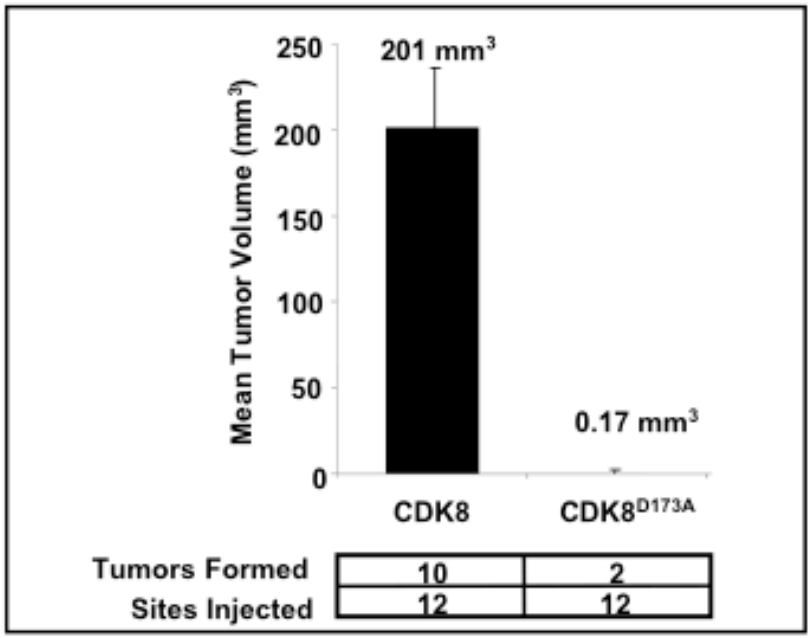
CDK8 overexpression drives transformation



NIH3T3 cells







CDK8 amplification promotes tumorgenicity via the pathway of β-catenin transcription activation

Firestein, Hahn

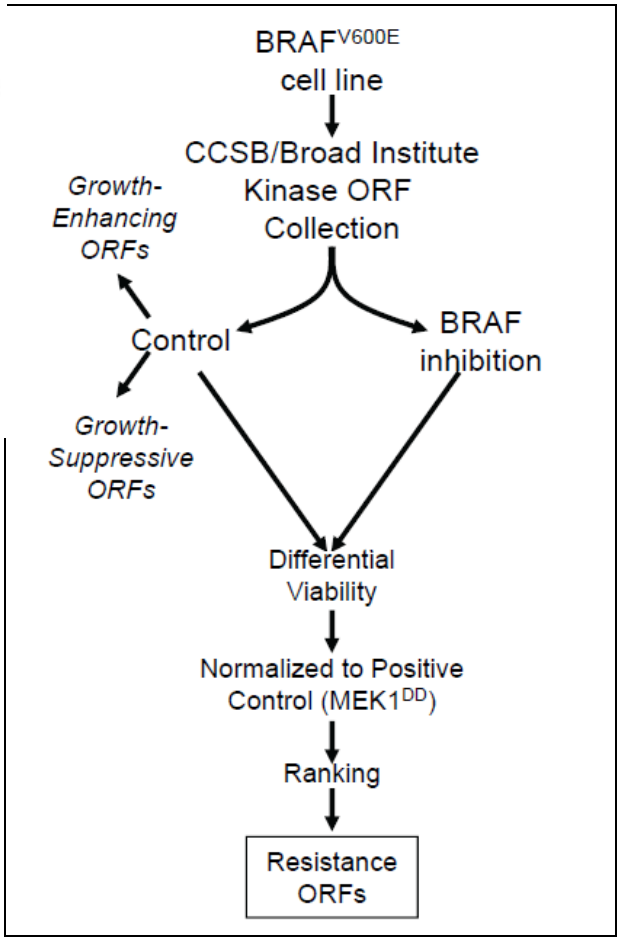
Drug resistance in BRAF-mutant melanoma

LETTER

doi:10.1038/nature09627

COT drives resistance to RAF inhibition through MAP kinase pathway reactivation

Cory M. Johannessen[1,2*], Jesse S. Boehm[1*], So Young Kim[1,2,3†], Sapana R. Thomas[1,2], Leslie Wardwell[2], Laura A. Johnson[1,2], Caroline M. Emery[2], Nicolas Stransky[1], Alexandria P. Cogdill[4], Jordi Barretina[1,2,5], Giordano Caponigro[6], Haley Hieronymus[1,7,8], Ryan R. Murray[3,9,10], Kourosh Salehi-Ashtiani[3,9,10], David E. Hill[3,9,10], Marc Vidal[3,9,10], Jean J. Zhao[9,11], Xiaoping Yang[1], Ozan Alkan[1], Sungjoon Kim[12], Jennifer L. Harris[12], Christopher J. Wilson[6], Vic E. Myer[6], Peter M. Finan[6], David E. Root[1], Thomas M. Roberts[9], Todd Golub[1,5,8], Keith T. Flaherty[4], Reinhard Dummer[13], Barbara L. Weber[6], William R. Sellers[6], Robert Schlegel[6], Jennifer A. Wargo[4], William C. Hahn[1,2,3,5] & Levi A. Garraway[1,2,5]

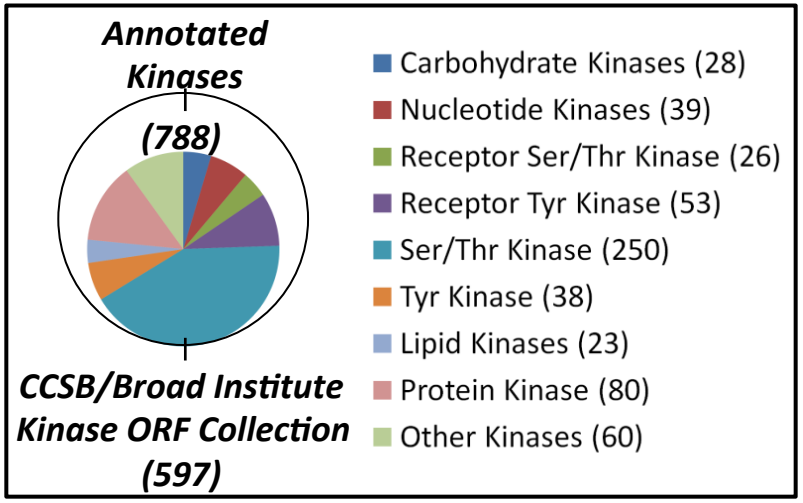


Cory Johannessen, Levi Garraway

Resistance genes via ORF based rescue screens

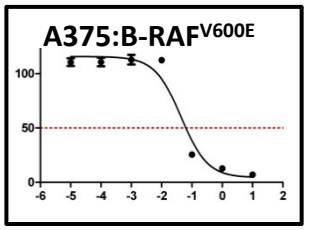
Cellular phenotype (drug sensitivity) → *Ectopic expression of human genes* → *Phenotype rescue (resistance)*

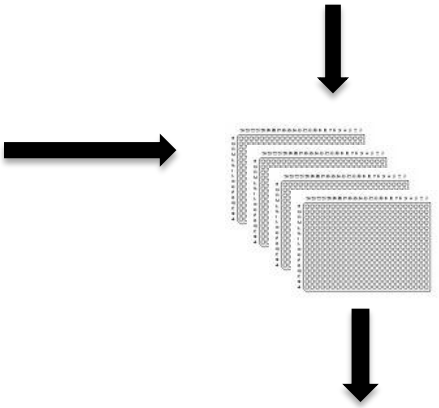
Resistance to



Rescue Screen:

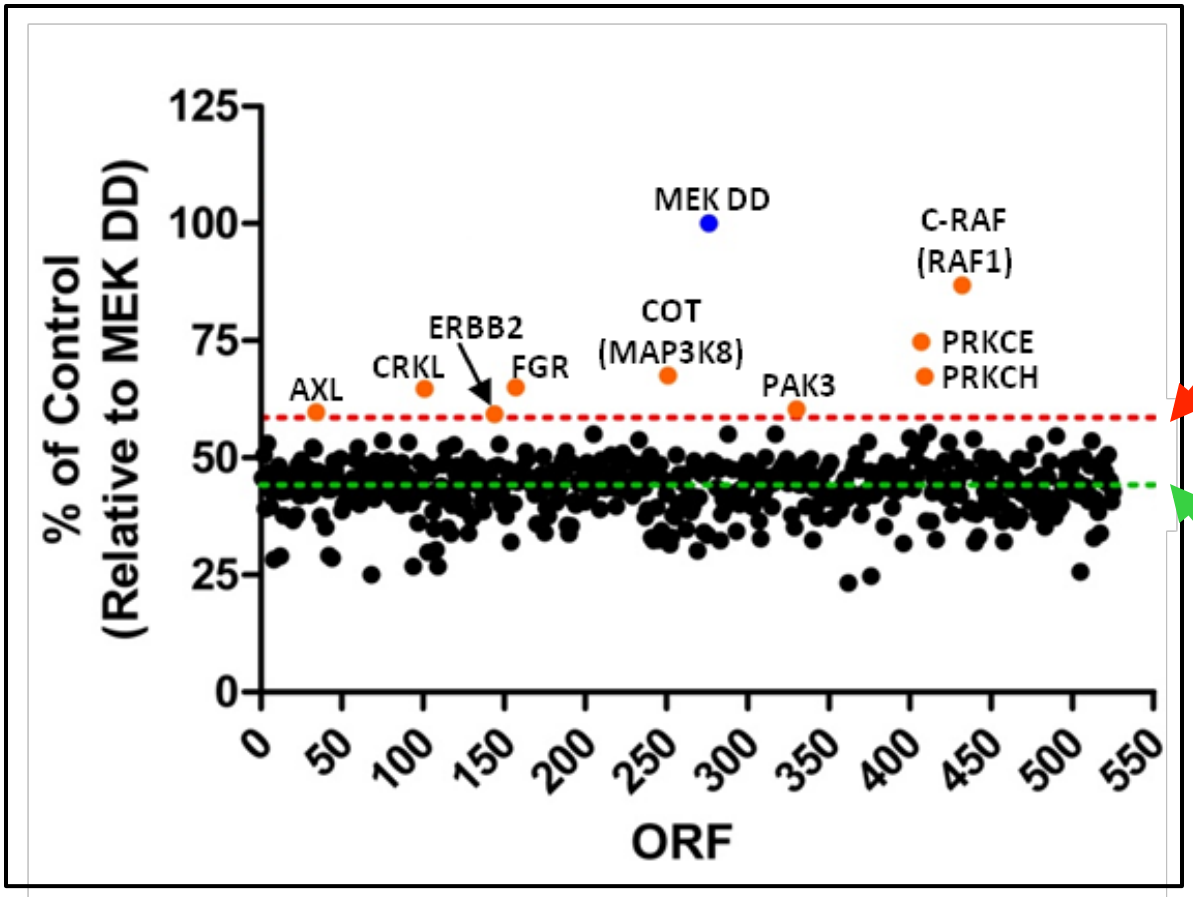
PLX4720 Sensitive Cell Line





Phenotypic Rescue of 1 µM PLX4720 (relative cell number)

A screen for kinases that bypass B-RAF inhibition



COT and **C-RAF**: candidate resistance kinases

Prioritization Screen
(2 cell lines, 8-point GI_{50})

Rank	Gene
1	*COT*
2	*C-RAF*
3	*CRKL*
4	*FGR*
5	*PRKCE*
6	*PRKCH*
7	*ERBB2*
8	*AXL*
9	*PAK3*

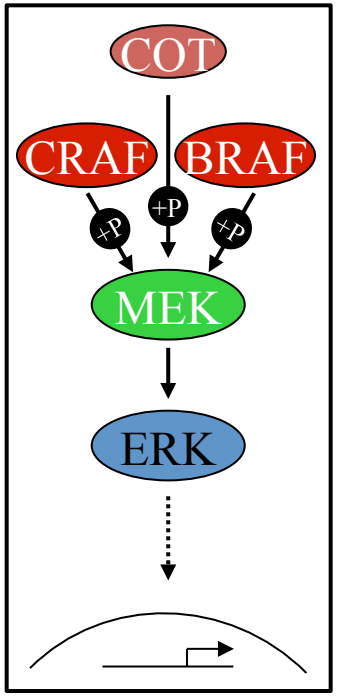
C-RAF/RAF1

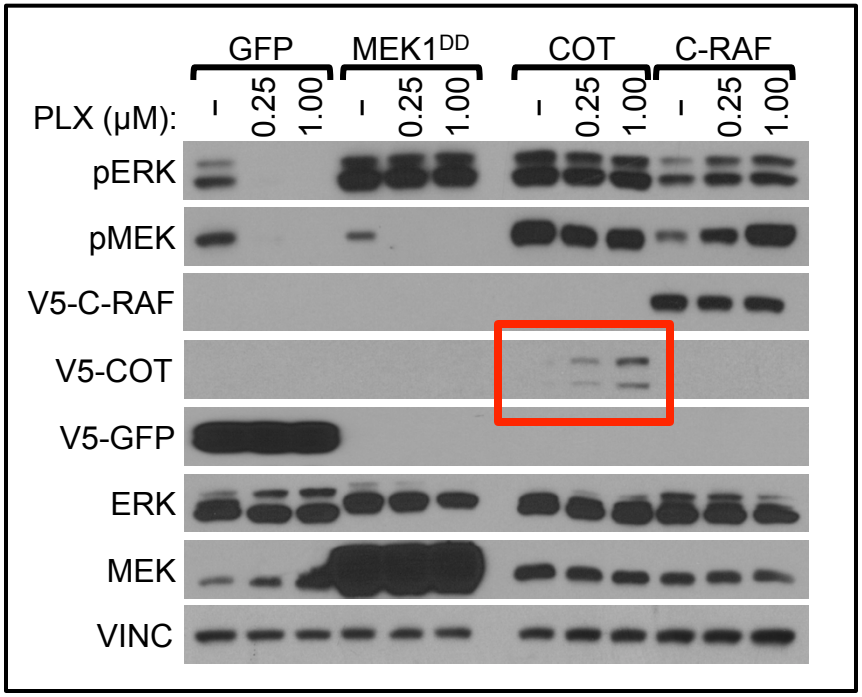
- Heterodimerizes with BRAF to activate the canonical MAPK signaling cassette
- Has been suggested to mediate resistance to RAF inhibition.

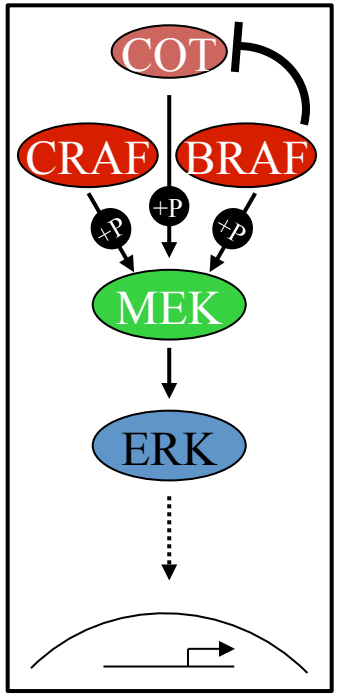
COT/TPL2/MAP3K8

- Like B- & C-RAF, COT is a **MAP3K**
- Has been shown to directly phosphorylate MEK1, activating ERK
- Not linked to melanoma

COT and C-RAF re-activate the MAPK pathway

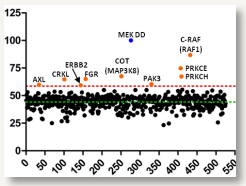




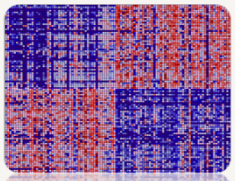


Identification of model systems to interrogate COT-mediated resistance

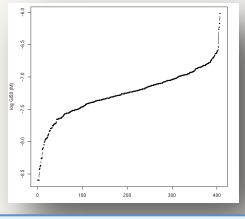
Candidate kinases



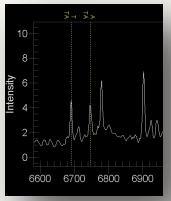
Broad-Novartis Cancer Cell Line Encyclopedia



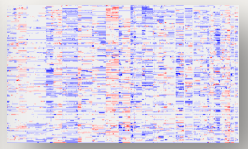
Expression profiling



Drug Sensitivity



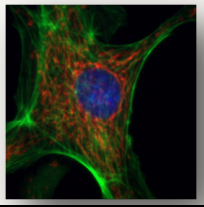
Mutation analysis



Copy number analysis

>1000 cancer cell lines

Model Cell Systems



+B-RAF^{V600E}

+/-COT amplification

Identify COT-amplified BRAFV600E mutant cell lines

Cancer cell lines: **752**

↓

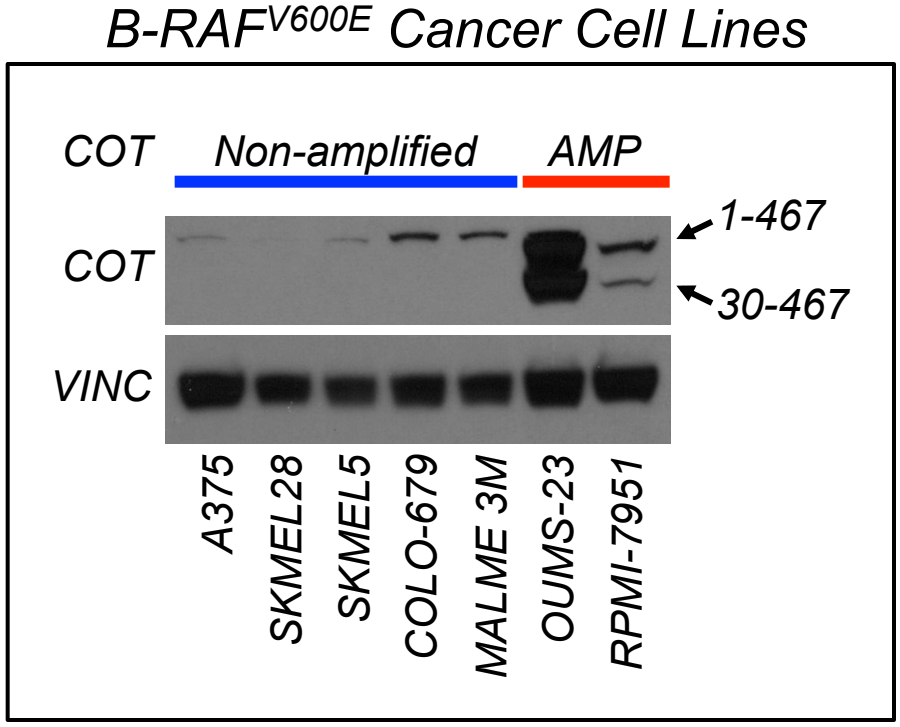
SNP/Oncomap: **534**

↓

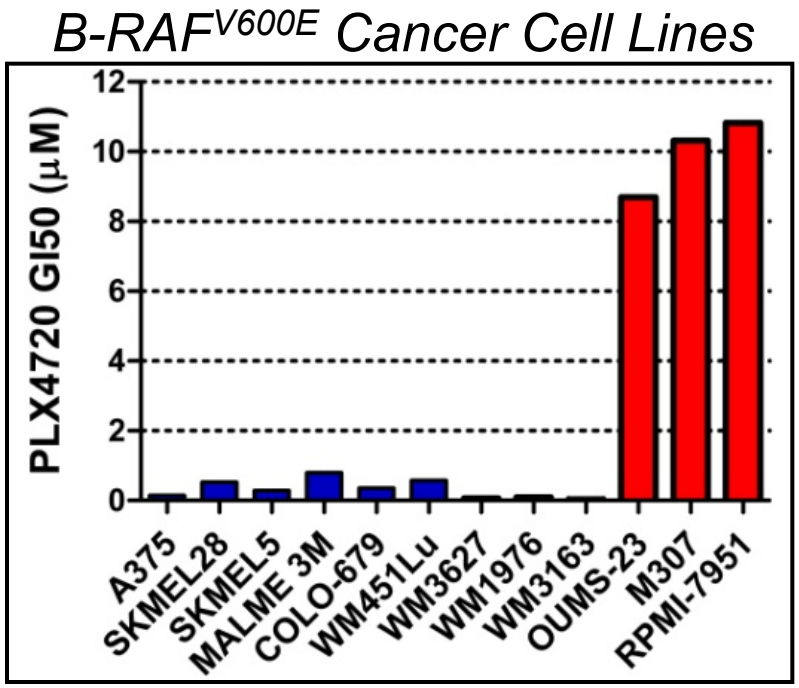
$B\text{-}RAF^{V600E}$: **38**

↓

$B\text{-}RAF^{V600E}$ + COT copy no. gains: **2**



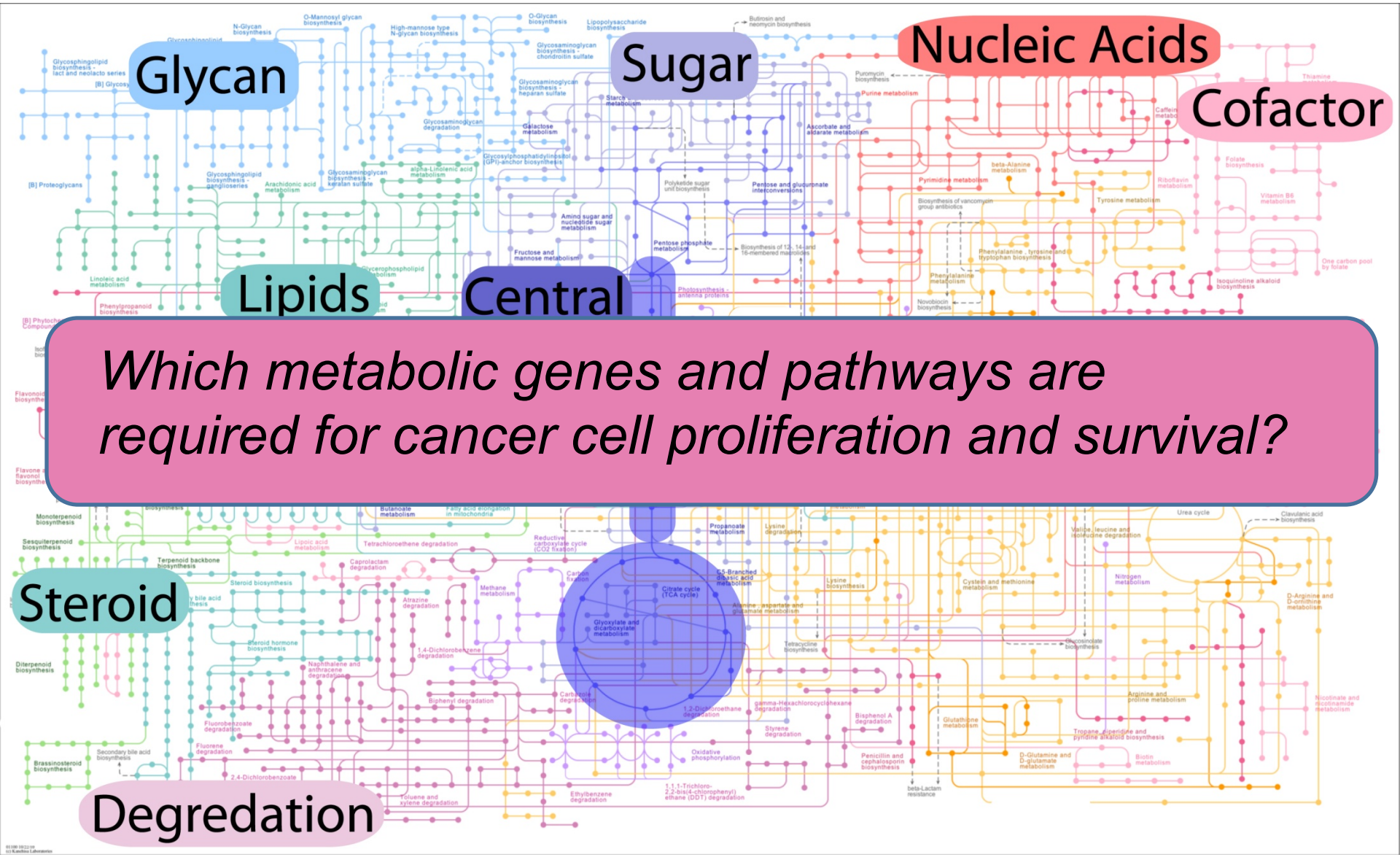
COT amplification predicts resistance in BRAFV600E cancer cell lines



A mouse in vivo screen

Rich Possemato, David Sabatini

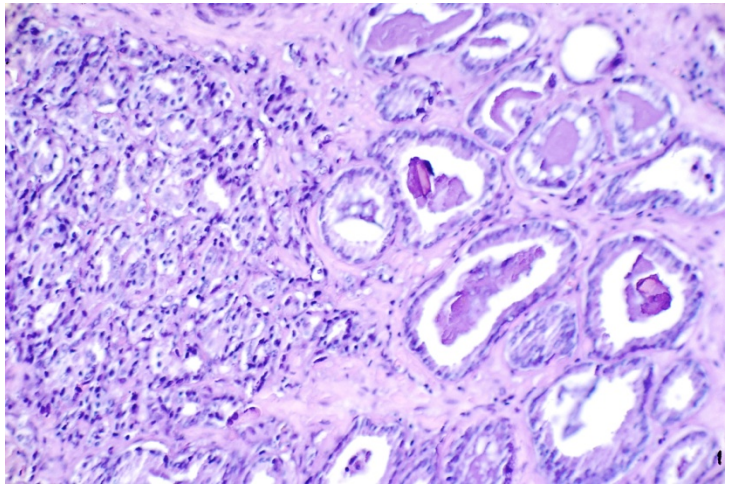
Effects of oncogene activation on all metabolism is poorly understo



Which metabolic genes and pathways are required for cancer cell proliferation and survival?

Adapted from KEGG

Cells in a tumor exist in a poorly understood environment

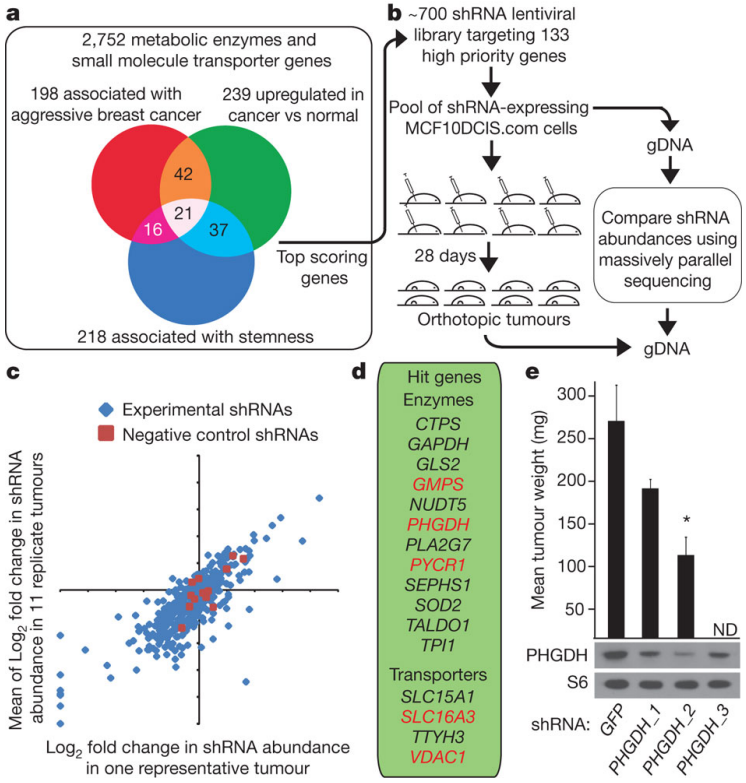


O. Brawley

2000 mg/L Glucose
300 mg/L Glutamine
30 mg/L Serine
10 mg/L Glycine
1 mg/L Folate
1 mg/L Glutathione
…
20 mg/L Glutamate
15 mg/L Histidine

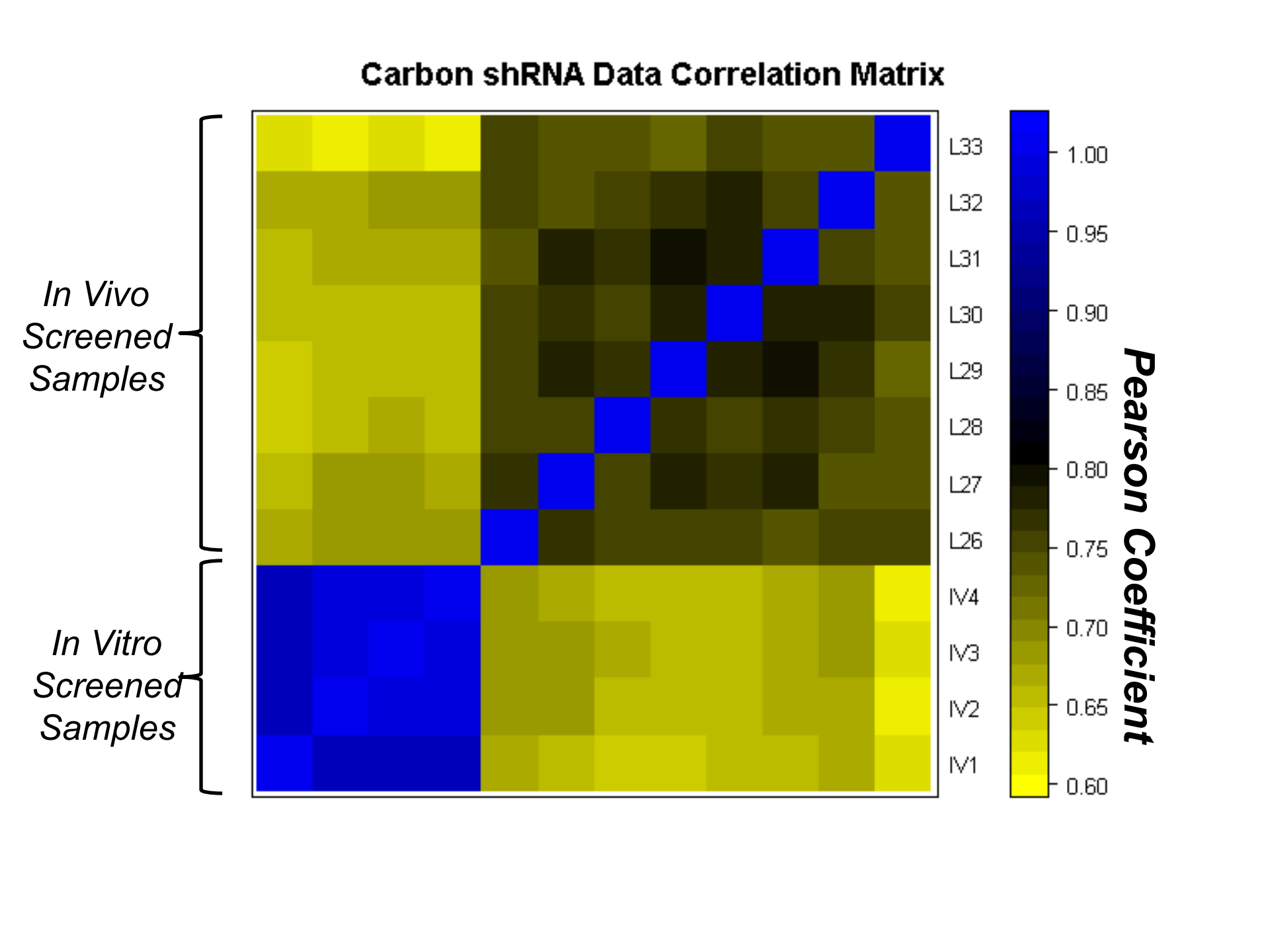
???? mg/L Glucose
??? mg/L Glutamine
?? mg/L Serine
?? mg/L Glycine
? mg/L Folate
? mg/L Glutathione
…
?? mg/L Glutamate
?? mg/L Histidine

Outline of *in vivo* pooled screening strategy identifying PHGDH as essential for tumorigenesis.

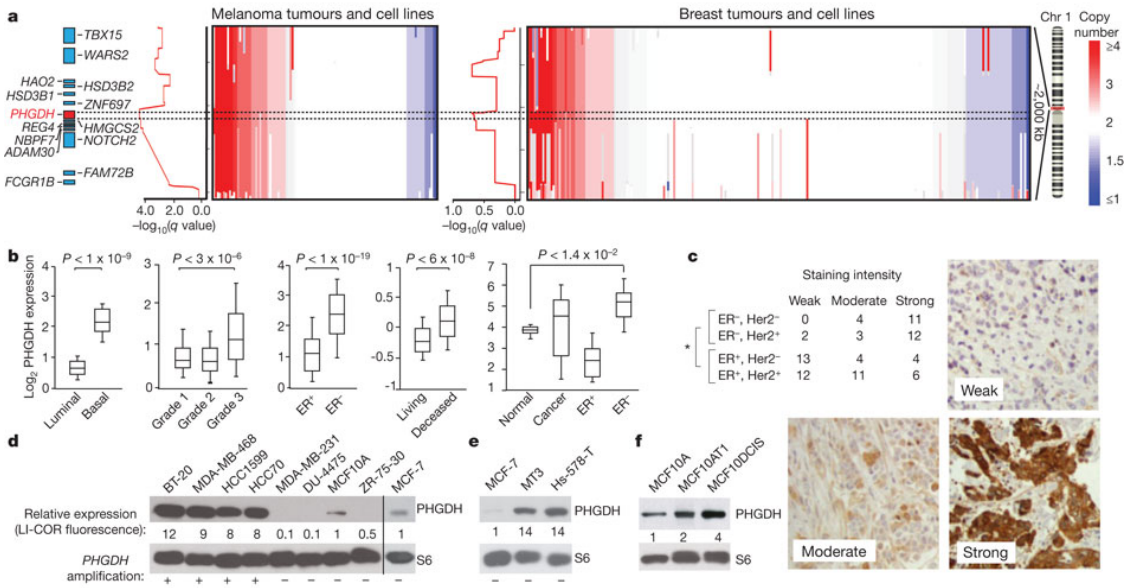


For five hit genes (*PHGDH*, *GMPS*, *SLC16A3*, *PYCR1* and *VDAC1*), two scoring shRNAs were tested for their effects on tumour formation. Each of these shRNAs suppressed expression of their targets in MCF10DCIS.com cells and reduced tumour-forming capacity. (Fig. 1e and Supplementary Fig. 2c). For reasons discussed later, *PHGDH* was of particular interest.

R Possemato *et al. Nature* **000**, 1-5 (2011) doi:10.1038/nature10350


Carbon shRNA Data Correlation Matrix
In Vivo Screened Samples
In Vitro Screened Samples
L33
L32
L31
L30
L29
L28
L27
L26
IV4
IV3
IV2
IV1
1.00
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
Pearson Coefficient

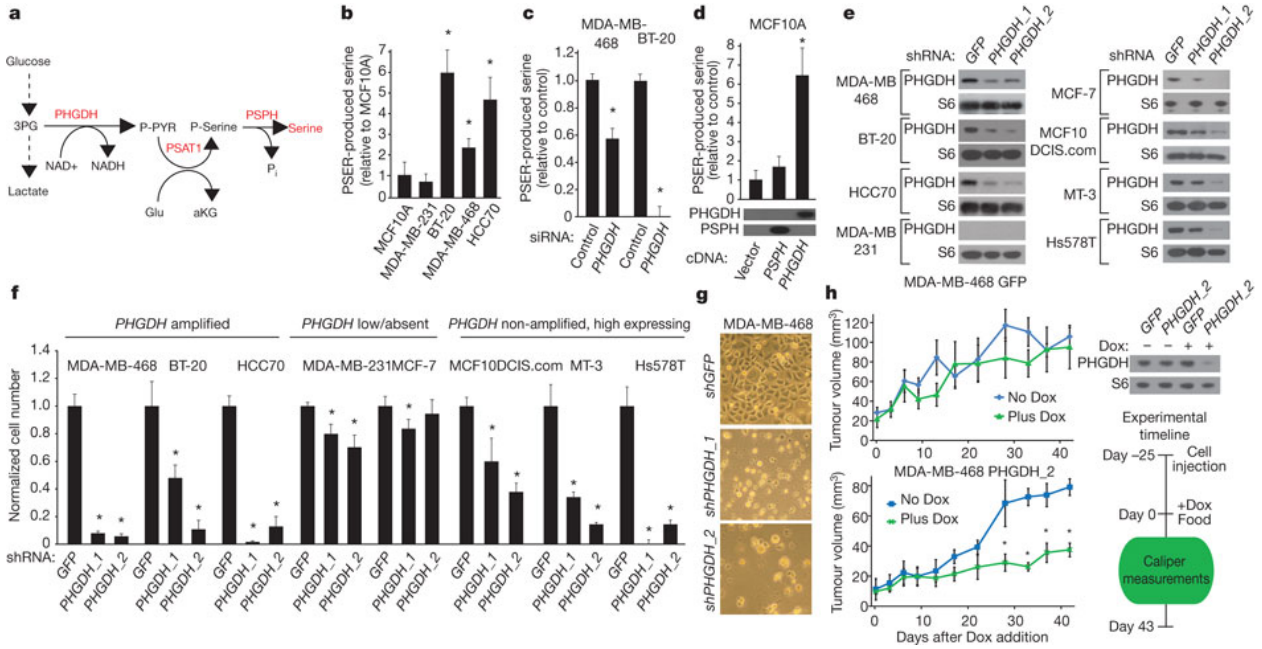
Genomic amplifications of *PHGDH* in cancer and association of PHGDH expression with aggressive breast cancer markers.



To prioritize genes for follow-up studies we consulted a recently available analysis of copy number alterations across cancer genomes. Indeed, *PHGDH* exists in a region of chromosome 1p commonly amplified in breast cancer and melanoma

R Possemato *et al. Nature* **000**, 1-5 (2011) doi:10.1038/nature10350

Cell lines with elevated PHGDH expression have increased serine biosynthetic pathway activity and are sensitive to PHGDH suppression.

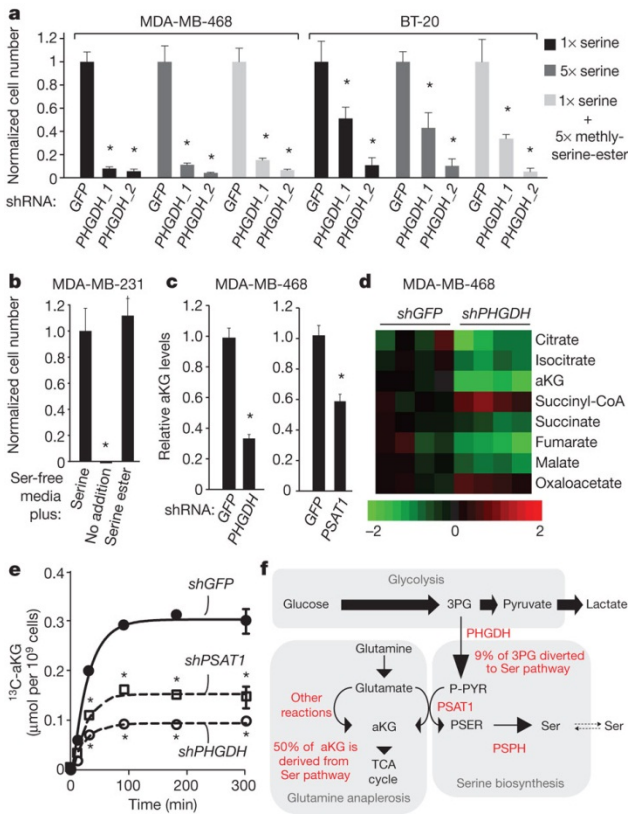


R Possemato *et al. Nature* **000**, 1-5 (2011) doi:10.1038/nature10350

To investigate whether PHGDH suppression can affect the growth of established tumours, we generated an inducible shRNA that, upon doxycycline treatment, reduced PHGDH protein levels in MDA-MB-468 cells.

(different cells)

Suppression of PHGDH results in a deficiency in anaplerosis of glutamine to aKG.



However, PHGDH suppression inhibited proliferation even in cells growing in media containing normal levels of extracellular serine, and supplementation with additional serine or a cell-permeable methyl-serine-ester did not blunt the effects of the PHGDH suppression

Uh-oh

In fact, of the major metabolites measured, aKG was the one with the most significant and largest change upon PHGDH suppression, whereas serine levels were not significantly changed

Ah, maybe ok…

Knock down PSAT1

R Possemato *et al. Nature* **000**, 1-5 (2011) doi:10.1038/nature10350

Paths to 'follow – up':

What does it mean to 'learn a gene's function'?

What does it mean to 'define the genes invovled in a process'

The phenotype

- Repeat original assay
- Different 'orthogonal' assay(s) for phenotype
- Different cell model systems

VALIDATION

The gene(s)

- Is phenotype due to expected effect of the shRNA or ORF?

e.g. discount off-target effects, non-specific effects

Detailed nature of the process

Context? Tissue types, in vivo

What defines/governs process?

- more players, further assays

Mechanism' – biochemistry

Relationship to disease, etc.

Cross to other data

FOLLOW-UP

Detailed nature of the protein(s) and the proximal effects:

Gene dose v. phenotype,

Enzymatic activity, structural feature?

Isoforms, protein modifications

Immediate substrates, binding partners

The LONGEST part of a screening project!

Some interesting quotes…..

QUESTIONING VALIDITY OF MODEL AND/OR READOUT

"In this highly artificial model, the authors identify potential enhancers and inhibitors of the loss of ASSAY staining when cells overexpressing GENE are ….TREATED WITH X. To my knowledge no one has shown that this phenomenon actually occurs in vivo with endogenous VERSION OF TREATMENT...."

"The authors have mainly used indirect measures for PHENOTYPE. All these phenotypes could have alternative explanations. Could the authors perform a TYPEOFASSAY assay to directly measure PHENOTYPE in CELL TYPE lines....?"

Since they use a model of GENEX overexpression for their screen it is difficult to know whether this modifier (HIT GENE) regulates endogenous GENEX. Moreover the CELLS PLUS TREATMENT model system utilized probably has little relevance to DISEASE pathogenesis."

"The study of more specific markers of PHENOTYPE that are exquisitely more sensitive to PERTURBATION may be informative."

NOT ENOUGH FOLLOW-UP

".... the mechanistic part of this study (biochemical and genetic follow up) is not sufficient to unequivocally support the presented hypothesis".

"The molecular mechanisms of action of these genes in PHENOTYPE are tested only very superficially and not in sufficient depth for a journal like JOURNAL. A strong focus on PHENOTYPE and no ASSOCIATED PHENOTYPE data obtained by quantitative techniques (such as ASSAYS) leaves the reader surprised and unsatisfied.

"....they chose to pursue just one hit, GENE, in much detail. They have now provided conclusive evidence of GENE's functional involvement in selective autophagy, but they have not yet understood its mechanistic role. In this light, the present contribution of this work is to generate a potentially valuable but largely unproven resource of selective PHENOTYPE genes, and to validate conclusively that ONE GENE is involved in PHENOTYPE while leaving its mechanistic role in this process unresolved."

NOT ENOUGH VALIDATION

".... the authors should dissipate any potential concern on off target effects by performing a phenotype rescue experiment with RNAi-resistant versions of the inactivated genes."

"The bioinformatics analysis is interesting, but speculative."

"...the authors select one of the genes, GENE, for in-depth analysis. ..... it is difficult to draw any conclusions about the involvement of the other (HIT) genes in PHENOTYPE based on this single example."

Screen projects

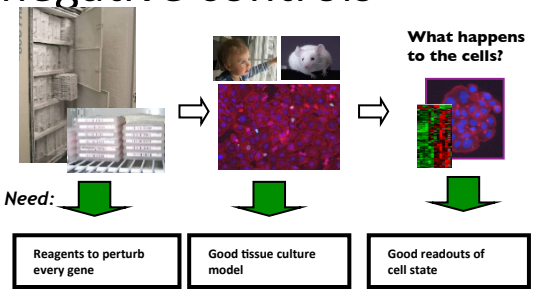
Define project

- Define question: "Find genes that do XYZ"
- Define biological model system
- Define assays to read out phenotypes of interest

Primary screen – feasibility and execution

- Optimize model system, assay(s); positive and negative controls
- Select gene set to interrogate

  Execute pilot and primary screen – select hits



Follow up on interesting genes/pathways

- Confirm assay result
- Confirm target gene specificity – multiple RNAi reagents, target KD
- Elaborate the biological effects,

  e.g. mechanism generality/context, biomedical sig?

Good luck! Come with questions.