USING THE LIBRARY OF INTEGRATED NETWORK-BASED CELLULAR SIGNATURES (LINCS) TO CHARACTERIZE THE MECHANISM OF ACTION OF SMALL-MOLECULE THERAPEUTICS,

1. LINCS Center for Transcriptomics (data)

2. LINCS-BD2K Center (analysis)

ARAVIND SUBRAMANIAN
MAY 28, 2015
BROAD INSTITUTE OF MIT AND HARVARD

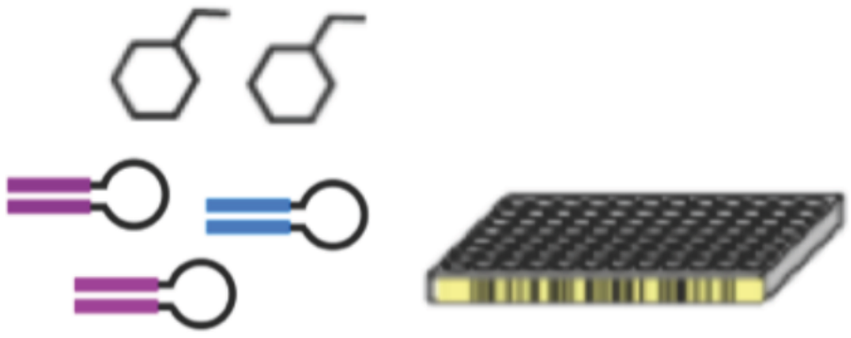
ARAVIND@BROADINSTITUTE.ORG

BROAD INSTITUTE

LINCS experiment

1. Perturbagen

e.g compound, shRNA, CRISPR



3. Read-out

e.g mRNA

Typical applications

- Mechanism of Action (MoA) of compounds and drugs
- Structure-activity relationships
- Functional annotation of genes
- Functional annotation of genetic variants
- Parsing Lists into “complimentation groups” of related function

Library of Integrated Network Based Cellular Signatures (LINCS)

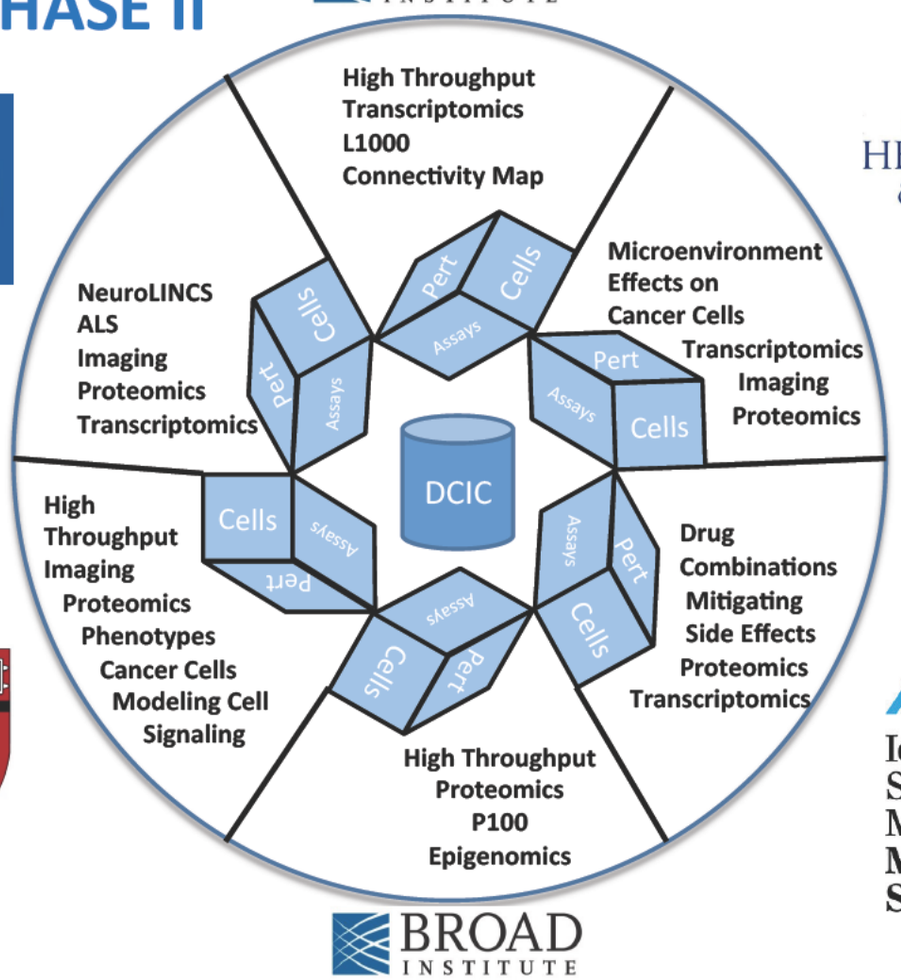
New NIH Program (launched FY2015)

LINCS PHASE II

BROAD INSTITUTE

UCI

OREGON HEALTH & SCIENCE UNIVERSITY OHSU



VERITAS

Icahn School of Medicine at Mount Sinai

BROAD INSTITUTE

[redacted]

Broad Institute - Connectivity Map (CMap)

a resource of perturbational profiles

Aim 1

Construct a **comprehensive reference database** of perturbational signatures

Aim 2

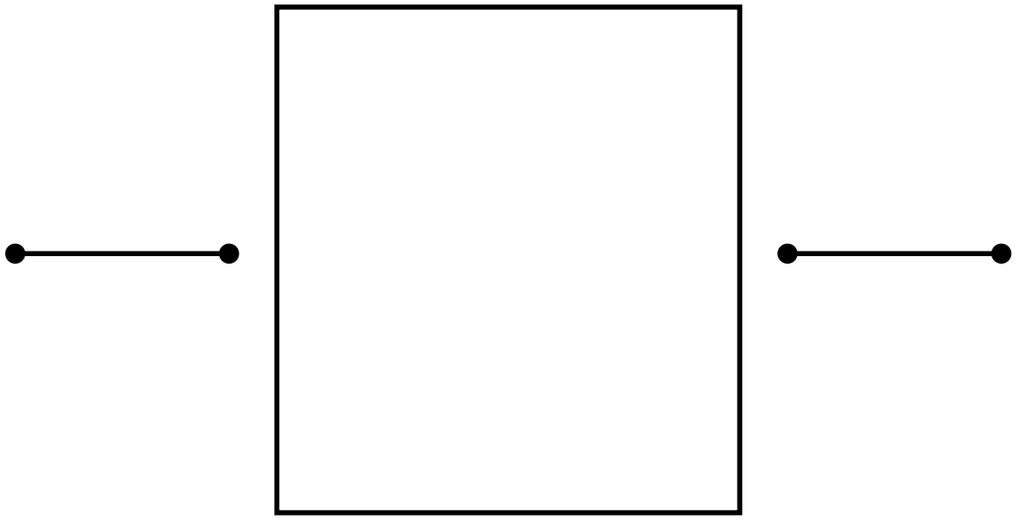
Develop **algorithms and software** to make these data and results accessible to the entire biomedical community

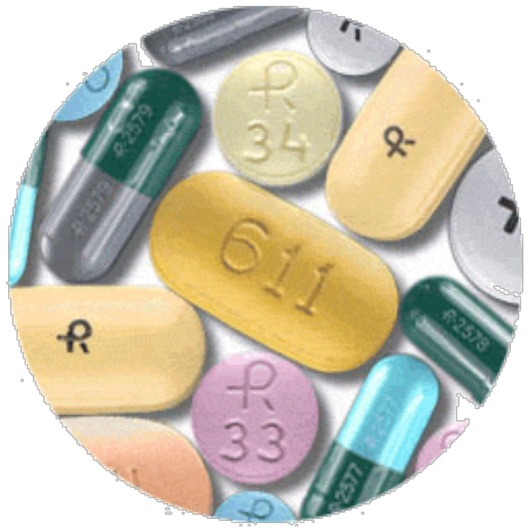
Aim 3

Utilize these resources to make **biological and therapeutic discoveries**

The Connectivity Map

Connecting disease to genes and drugs





mRNA Expression Database

PILOT

564 Affymetrix profiles 16,000+ users

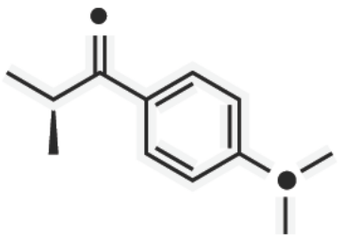
164 drugs 916 citations

(build02 expanded to ~7,000 profiles)

Lamb *et al*, Science (2006)

LINCS - desired dataset (~10M profiles)

- **CMap is limited by profiling cost**
- **low-cost, high-throughput method would enable …**



small-molecule compounds
Drugs, tool compounds, natural products, diverse synthesized libraries

genes (knock-down and gain-of-function)
shRNAs, CRISPRs, cDNAs, variants (natural and synthetic)

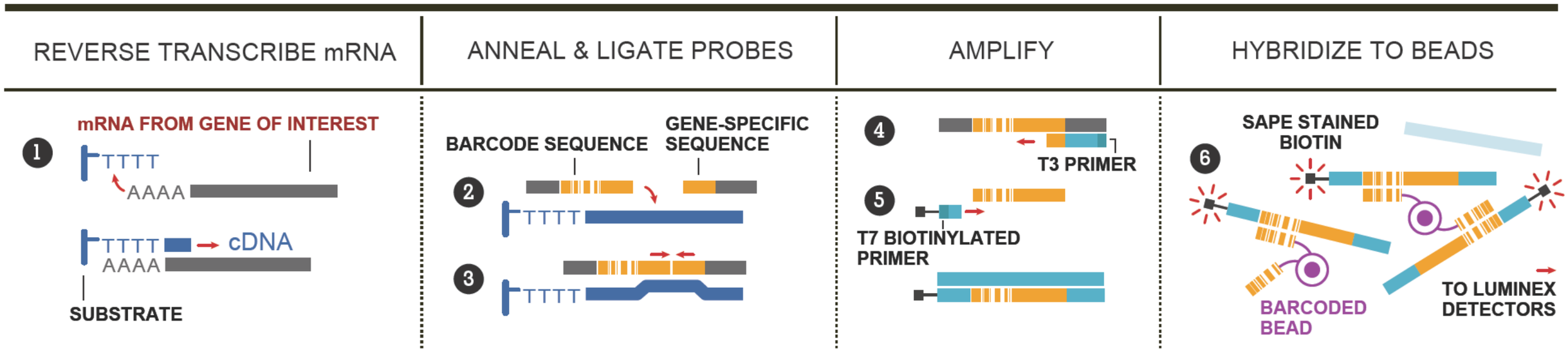
cellular contexts
diverse, multiple lineages, culture conditions, genetics

treatment parameters contexts
concentrations, durations, combinations

L1000 ASSAY

The L1000 Assay

ligation-mediated 1000-plex amplification



Measure 1,000 transcripts with 500-color Luminex beads

384-well plate format compatible with HTS

Cell lysates (not purified RNA)

Dramatic cost reduction

Reduced representation of transcriptome

using ‘landmarks’ to infer the rest

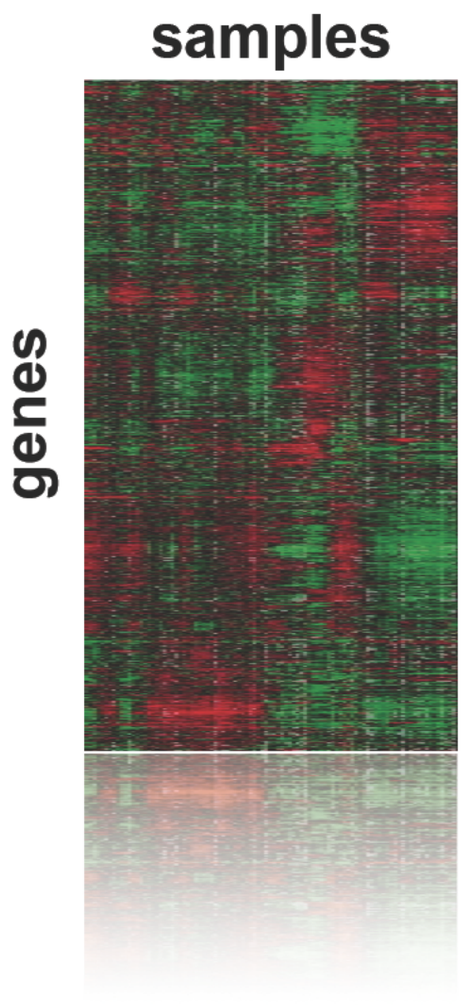
Reduced representation of transcriptome “Landmark genes”

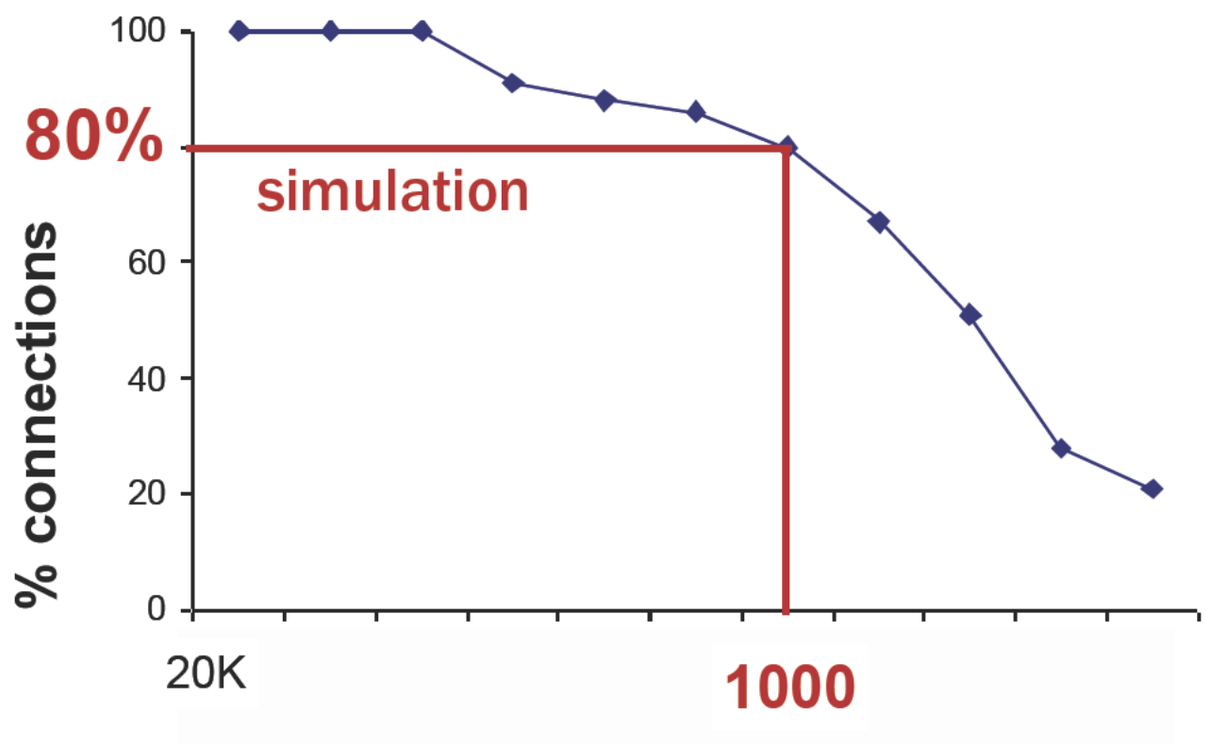
+

Computational inference model

=

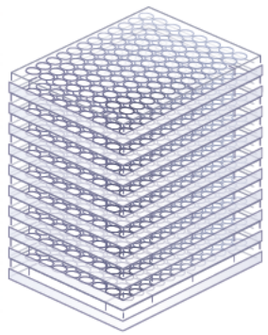
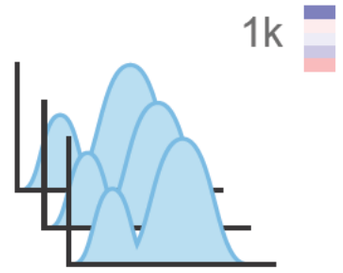
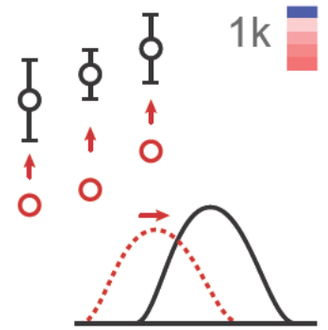
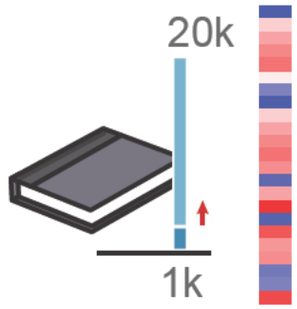
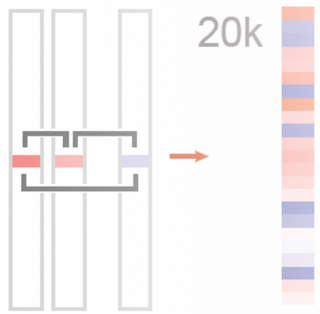
Genome wide expression





Processing L1000 data

Data flow: from scans to signatures

SCAN	DECONVOLUTION	NORMALIZATION	INFERENCE	DIFFERENTIAL EXPRESSION
				
Scan 384-well plates	Deconvolute 500 colors to 1000 genes	Scale 80 contol genes	Use landmarks to infer entire transcriptome	Compare replicate treatments to control
LXB	GEX	QNORM	INF DATA	SIG

$$y = ax^b + c$$

$$x_{norm} = F_2(F_1(x))$$

$$g_k = w_0 + \sum_{i=1}^{978} w_i l_i$$

$$z_i = \frac{x_i - median($$

$$MODZ = \frac{w_a \cdot z_a + w_b \cdot z_b +}{w_a + w_b +}$$

Signature generation

raw profiles to differentially expressed, ranked gene lists

Experiment

Signature

1 perturbagen in 1 cell type at 1 dose and for 1 duration of treatment

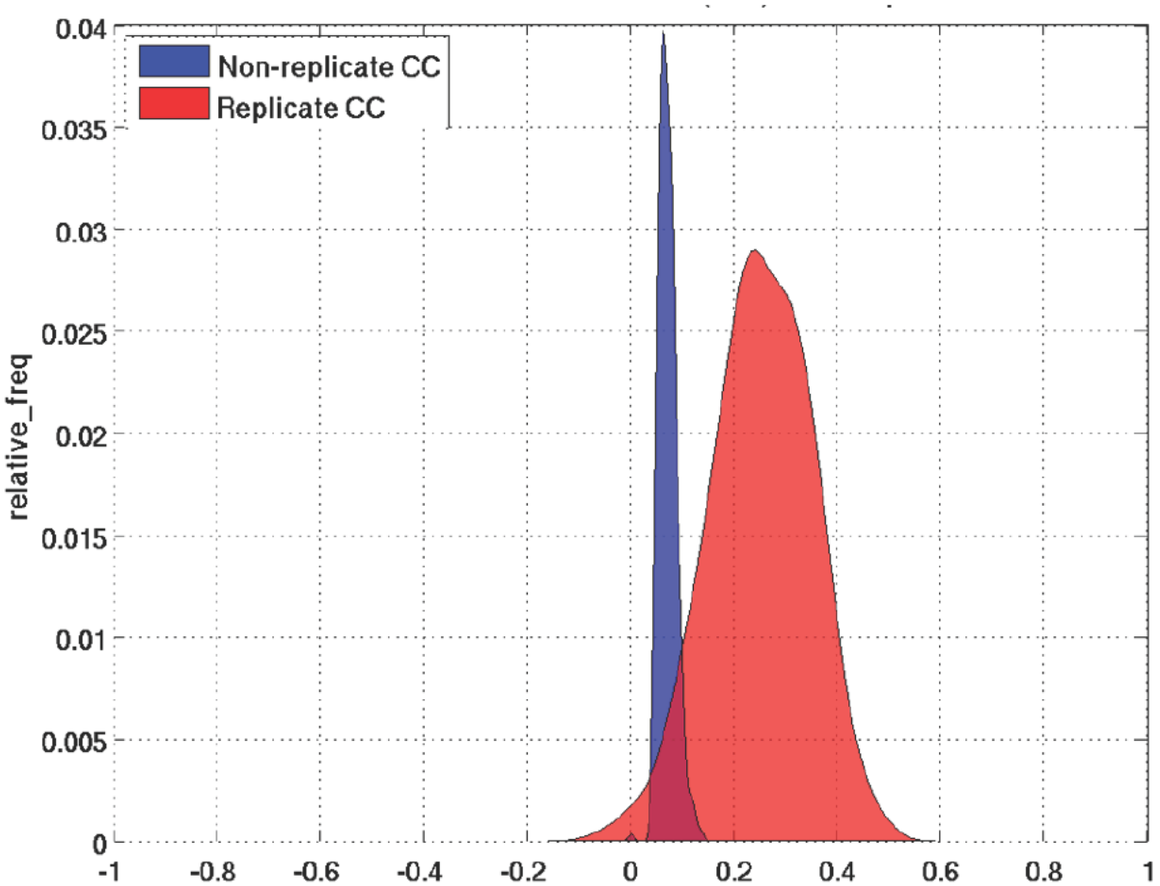
Database

Perturbagen	Cell Types	Doses	Duration	Replicates	# signatures
Compound 1	12	6	1	3	12 x 6 = 72
Compound 2	12	6	1	3	12 x 6 = 72
Compound 3	12	6	1	3	12 x 6 = 72
...					
Compound 348	12	6	1	3	12 x 6 = 72
TOTAL					72 x 348=**25,056**

L1000 performance - over batches

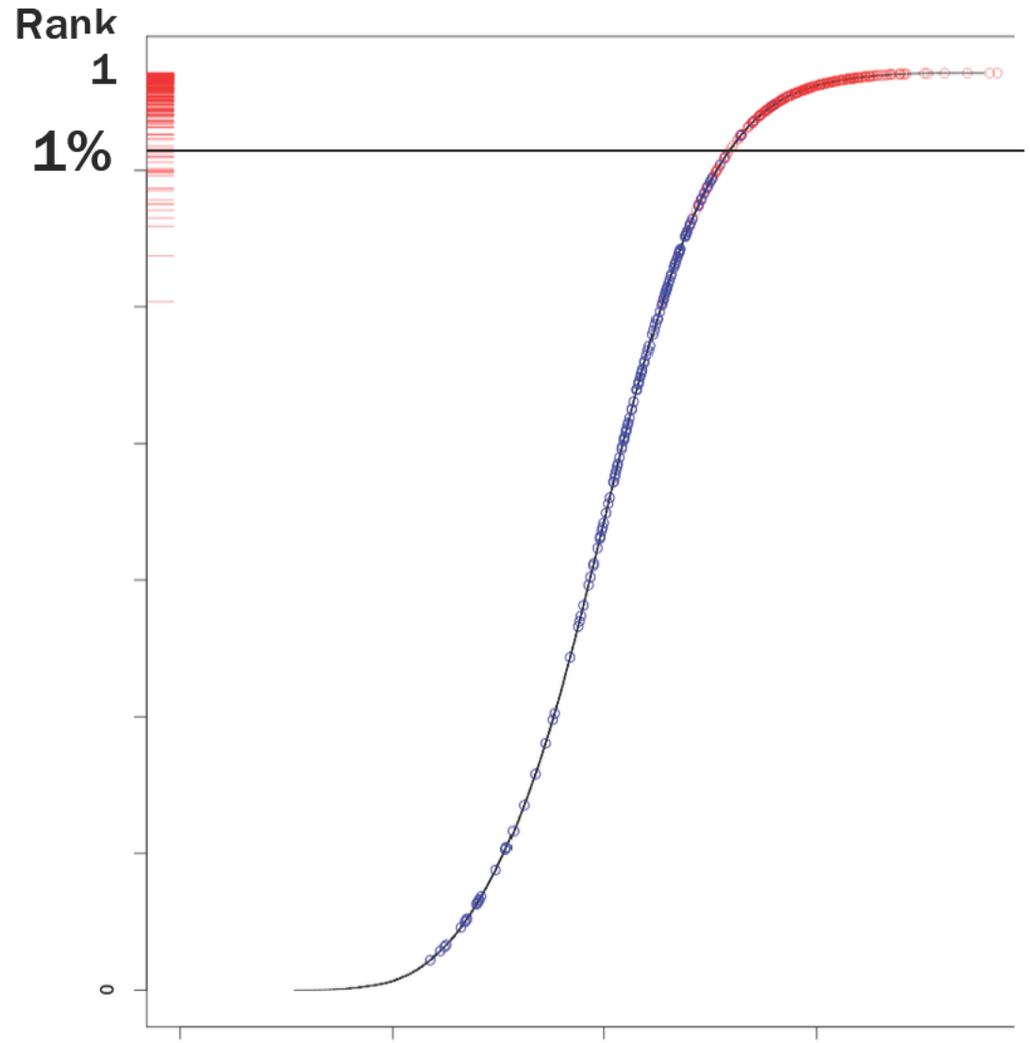
High connectivity between past & new profiles - data is robust

Query space: 1 plate (384 profiles)



Spearman rank correlation

Query space: Full database (1.3M profiles)



Connectivity Score

L1000 performance - compared to Affymetrix

experimental L1000 data (*not* simulation)

79 connections from 7200 Affymetrix profiles

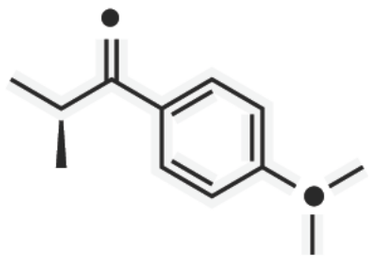
↓

Repeat Affymetrix profiles on L1000

↓

71/79 (90%) connections recovered

Current CMap dataset (2M (...10M) profiles)



~5,000 Compounds
- FDA approved drugs
- Bioactive tool compounds
- Screening hits

~3,000 Genes (shRNA & cDNA)
- Targets/pathways of approved drugs
- Candidate disease genes (e.g. recurrent from TCGA)
- Community nominations

15 Cell types
- Banked primary cell types
- Cancer cell lines (8)
- Primary hTERT-immortalized
- Patient-derived iPS cells

ANALYSIS METHODS

Design goals of L1000 data and analytics

- Comprehensive
- Information-rich readouts (not optimized to particular questions)
- Easy to look things up (like Google)
- Easy to compare to non-L1000 data
- Accessible to biologists and computationalists

Bookmarks Window Users Help

Query

apps.lincscloud.org/query

lincscloud

Query

Match user-defined gene sets to L1000 signatures

take a tour

cflynn

Name your query

Enter **Up-regulated** genes

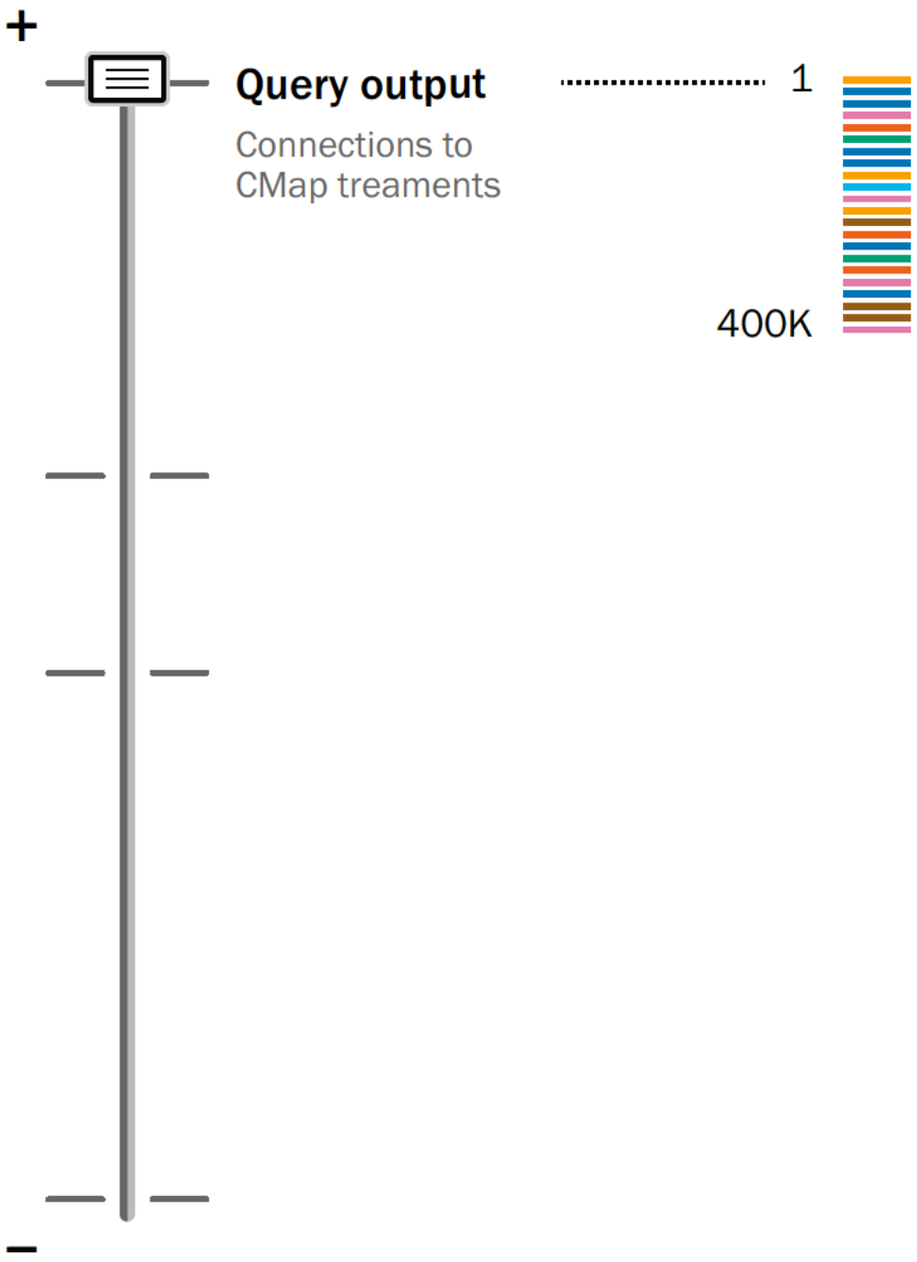
Enter one gene symbol or Affymetrix U133A probe ID per line or drag and drop a plain text file here.

Enter **Down-regulated** genes

Enter one gene symbol or Affymetrix U133A probe ID per line or drag and drop a plain text file here.

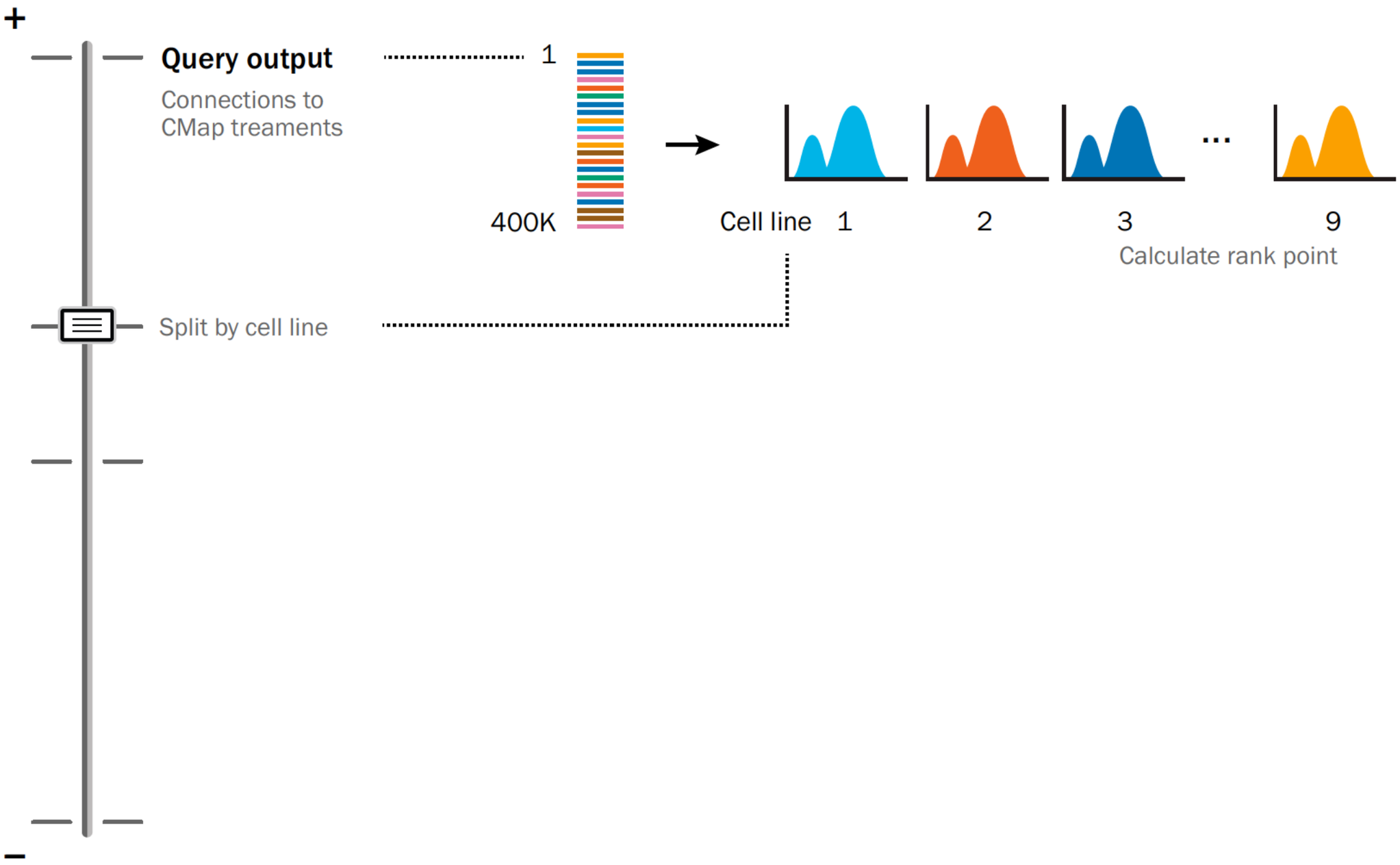
CMap result analysis

Many views of a single query



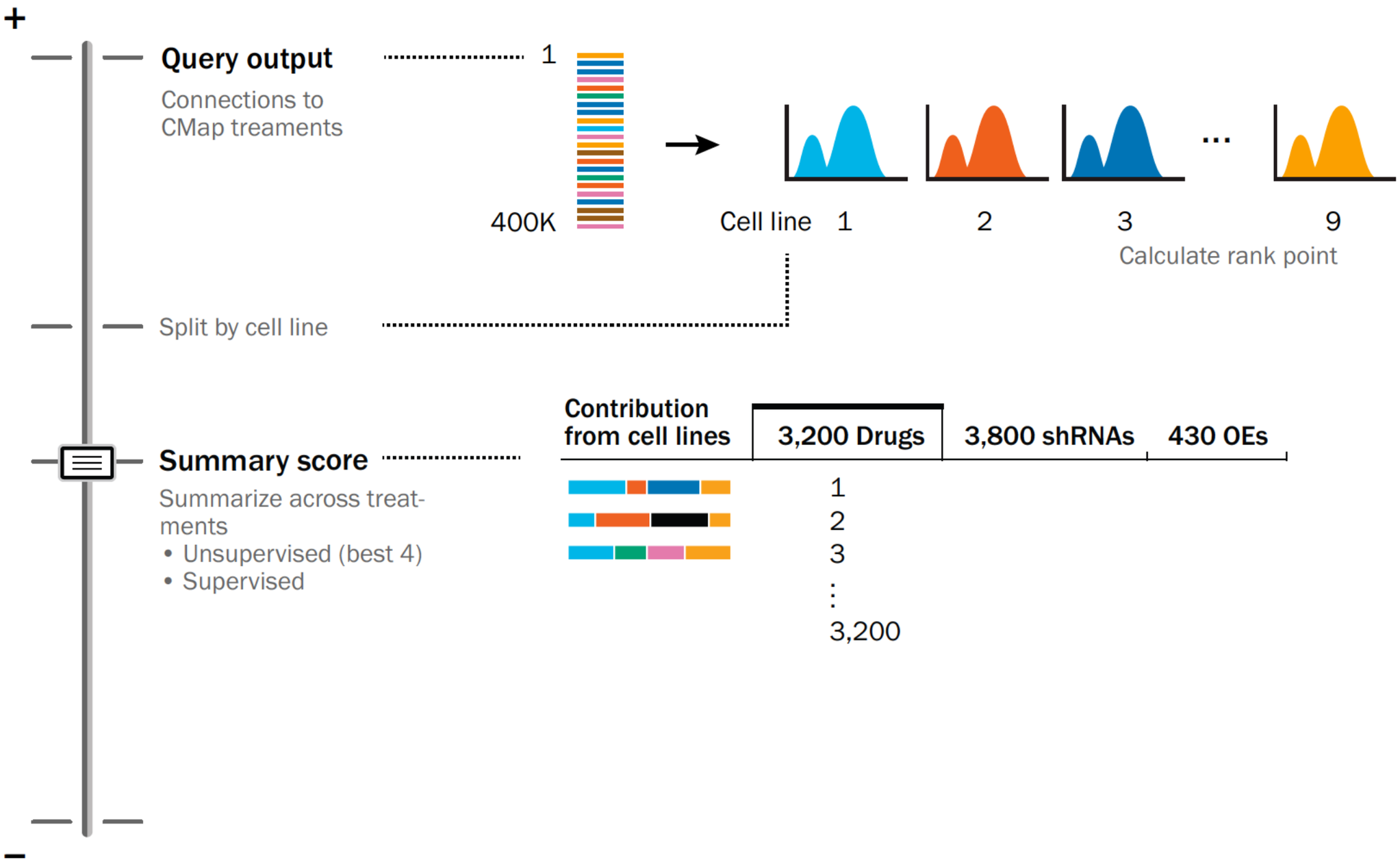
CMap result analysis

Many views of a single query



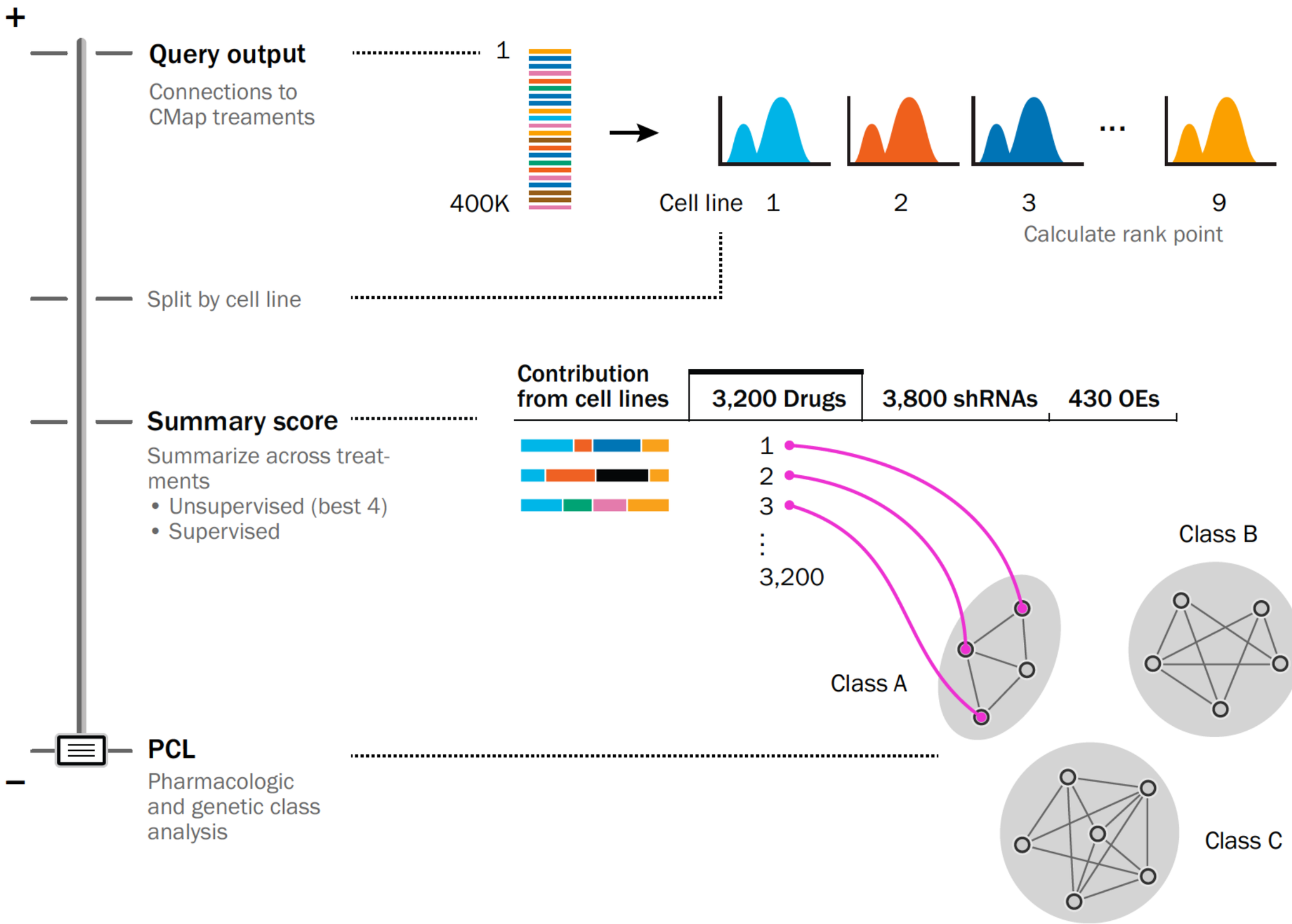
CMap result analysis

Many views of a single query



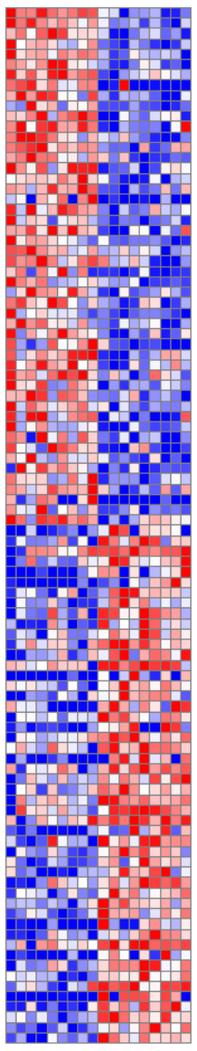
CMap result analysis

Many views of a single query



Query example

GEO Signature: HUVEC cells treated with pitavastatin



Rank	Compound
1	lovastatin
2	simvastatin
3	mevastatin
10	atorvastatin
29	pitavastatin
36	fluvastatin

3,200

Results

Given the mRNA expression signature of a TEST perturbagen

1. BIOACTIVITY

Does the perturbagen elicit a gene expression response?

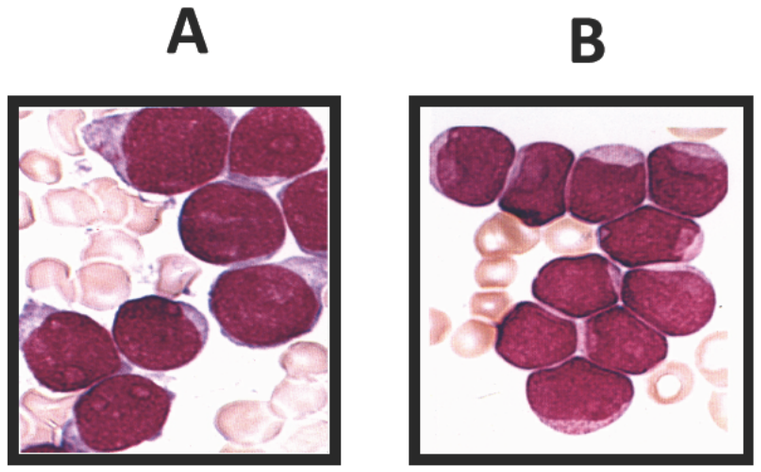
2. PATHWAYS

What gene pathways (or “networks”) are modulated?

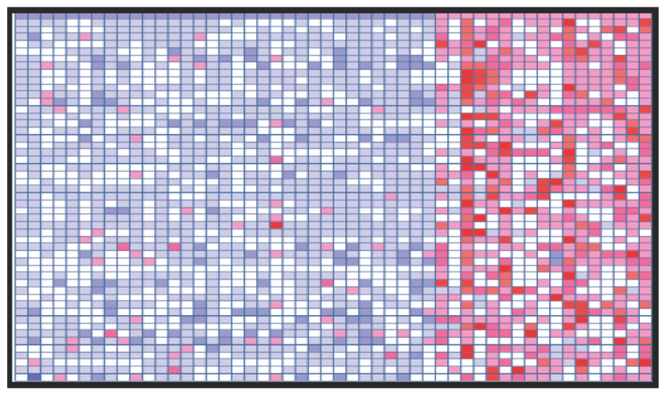
3. PHARMACOLOGICAL CLASSES

Which (previously characterized) pharmacological activities match the test perturbagen?

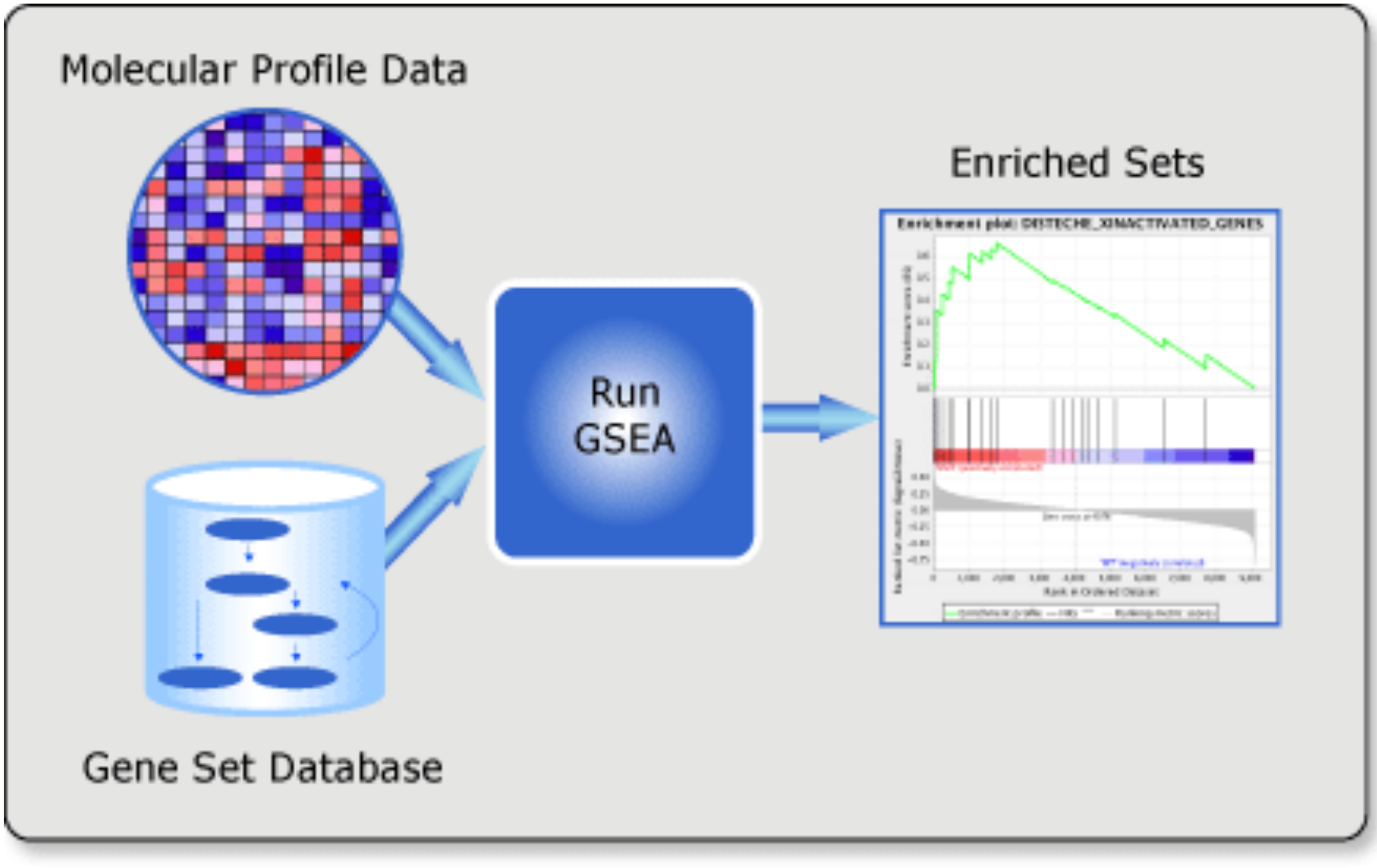
Biological States



mRNA signature

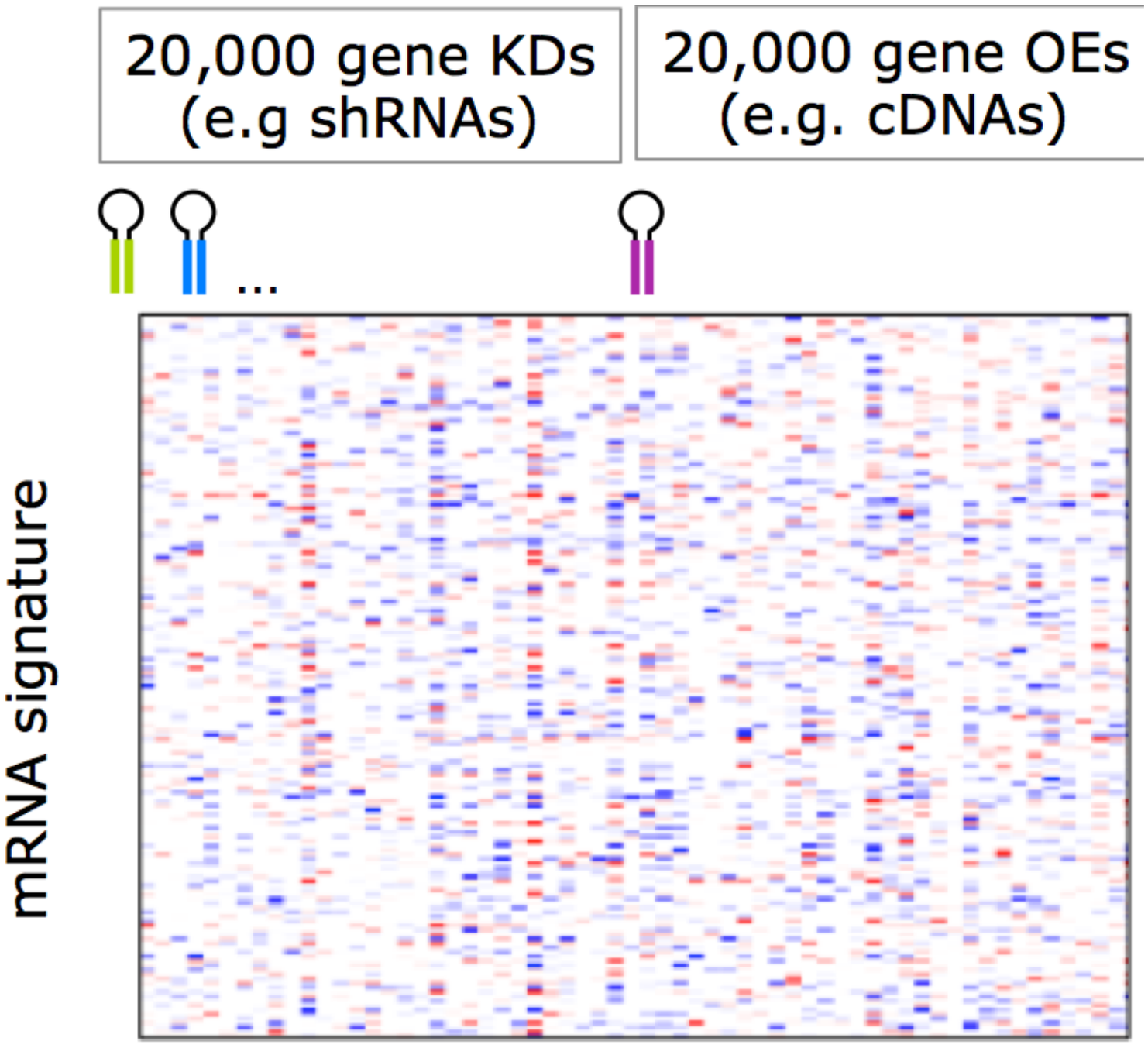


Interpretation

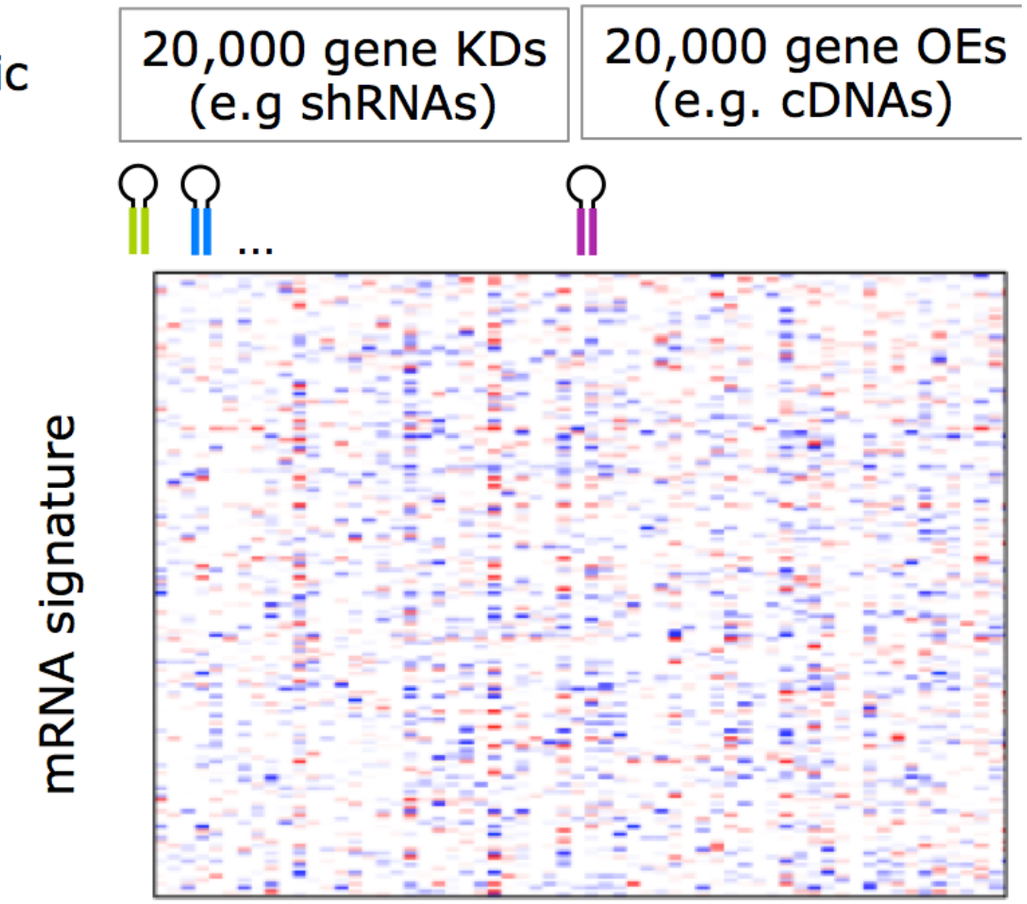


key challenge: Better gene sets

Solution: Systematic Known-down and over-expression of all genes



Hypothesis: Creating a systematic dataset of gene knock-down amd over-expression signatures



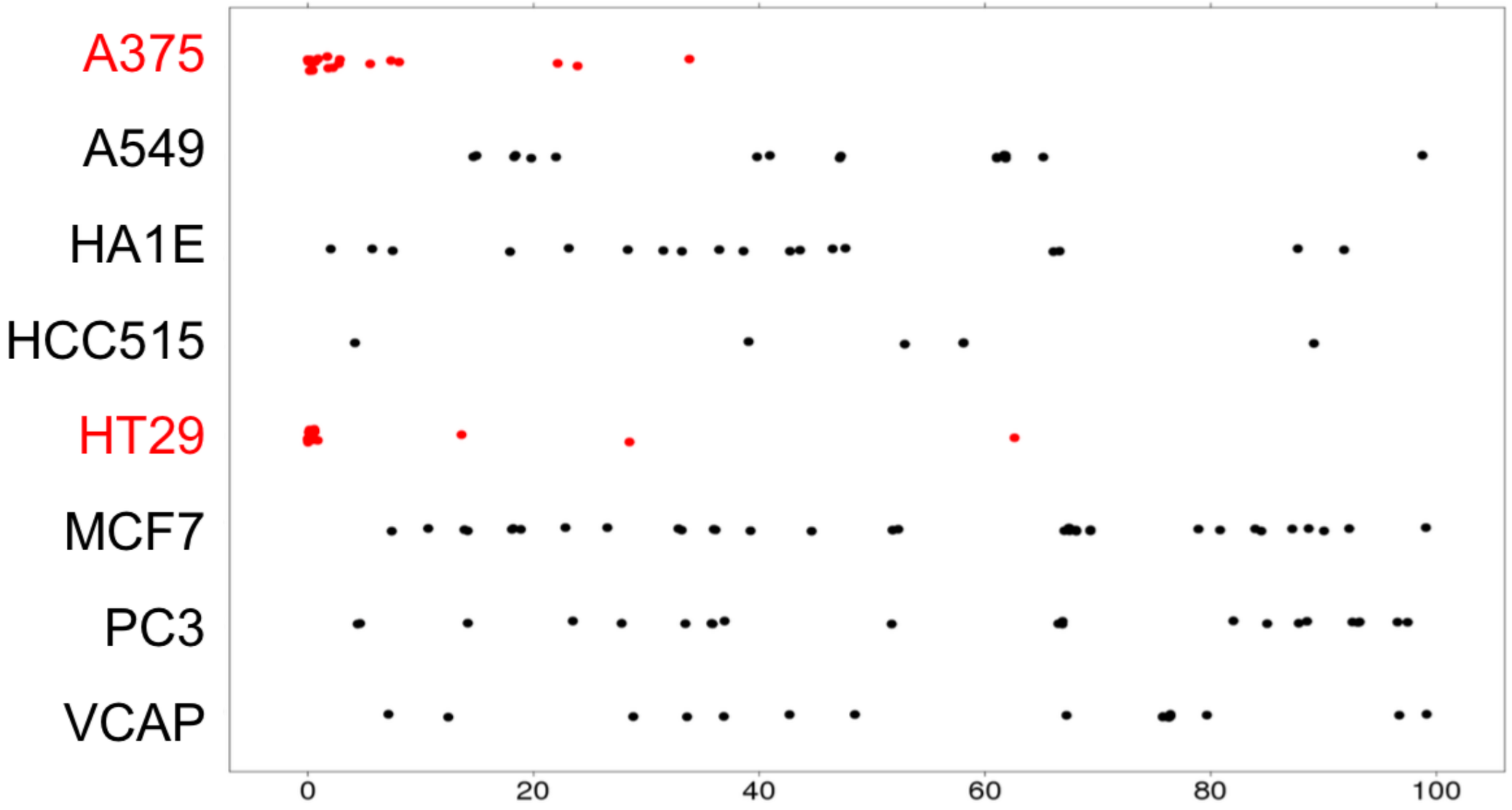
Would be useful for:

1) defining the "**core signature**" of a gene's function
2) identifying **gene-gene** functional connections
3) identifying **pathway** members
4) identifying **targets/pathways of small-molecules**

20,000 GENES x 2 (KD+OE) x replicates x cell types (10-20)
=> **5M** profiles (~1M so far)

Query with Vemurafenib, a small molecule inhibitor of BRAF

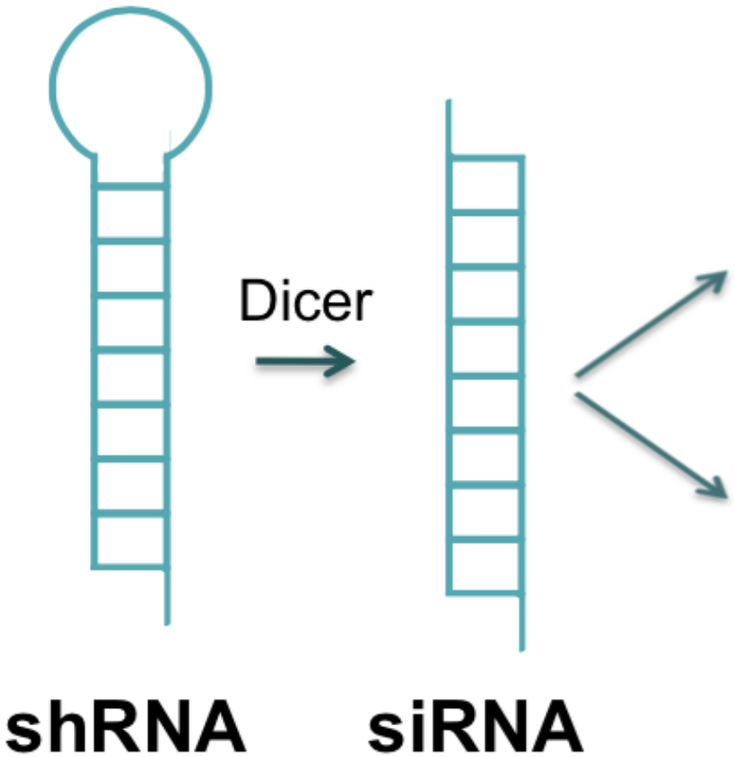
Each dot is an individual shRNA targeting BRAF



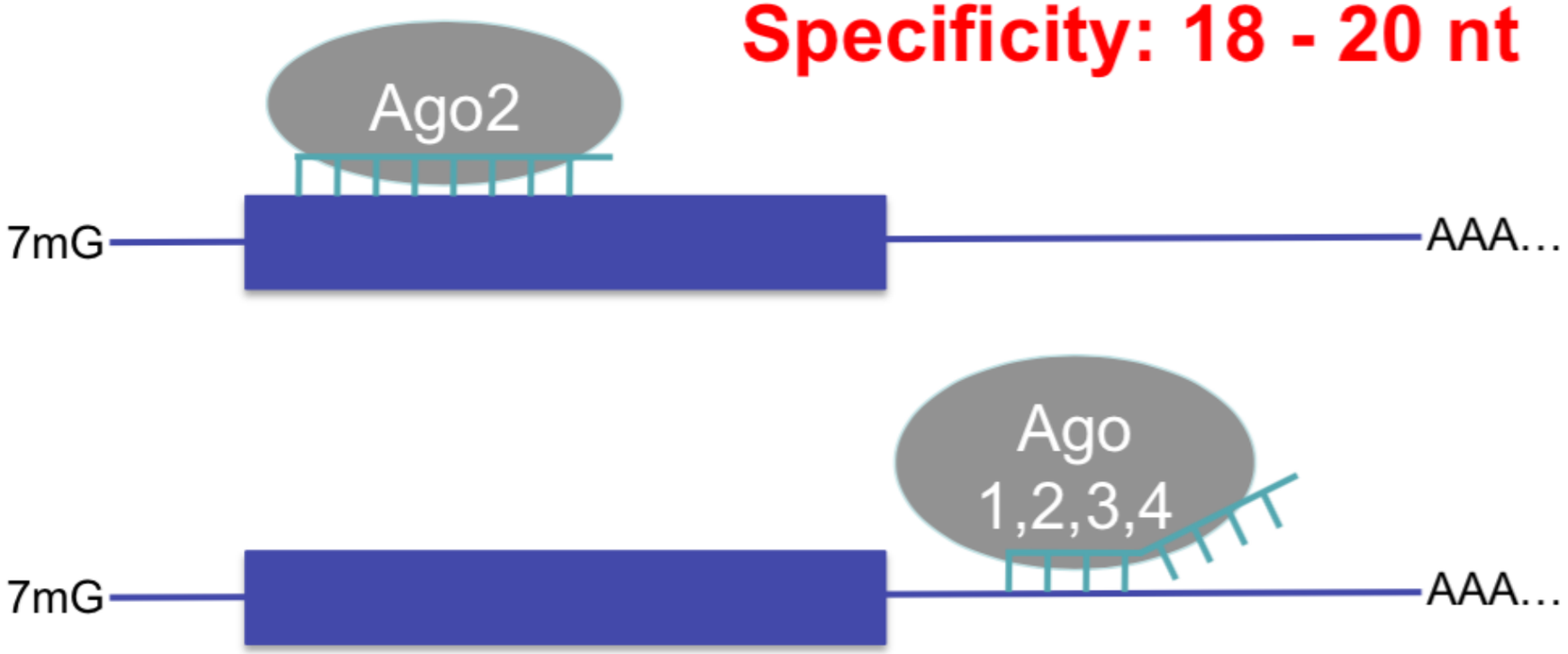
Two paths for shRNAs: how important are off-target effects?

RNAi

Perfect match to mRNA:
on-target gene knockdown
Specificity: 18 - 20 nt



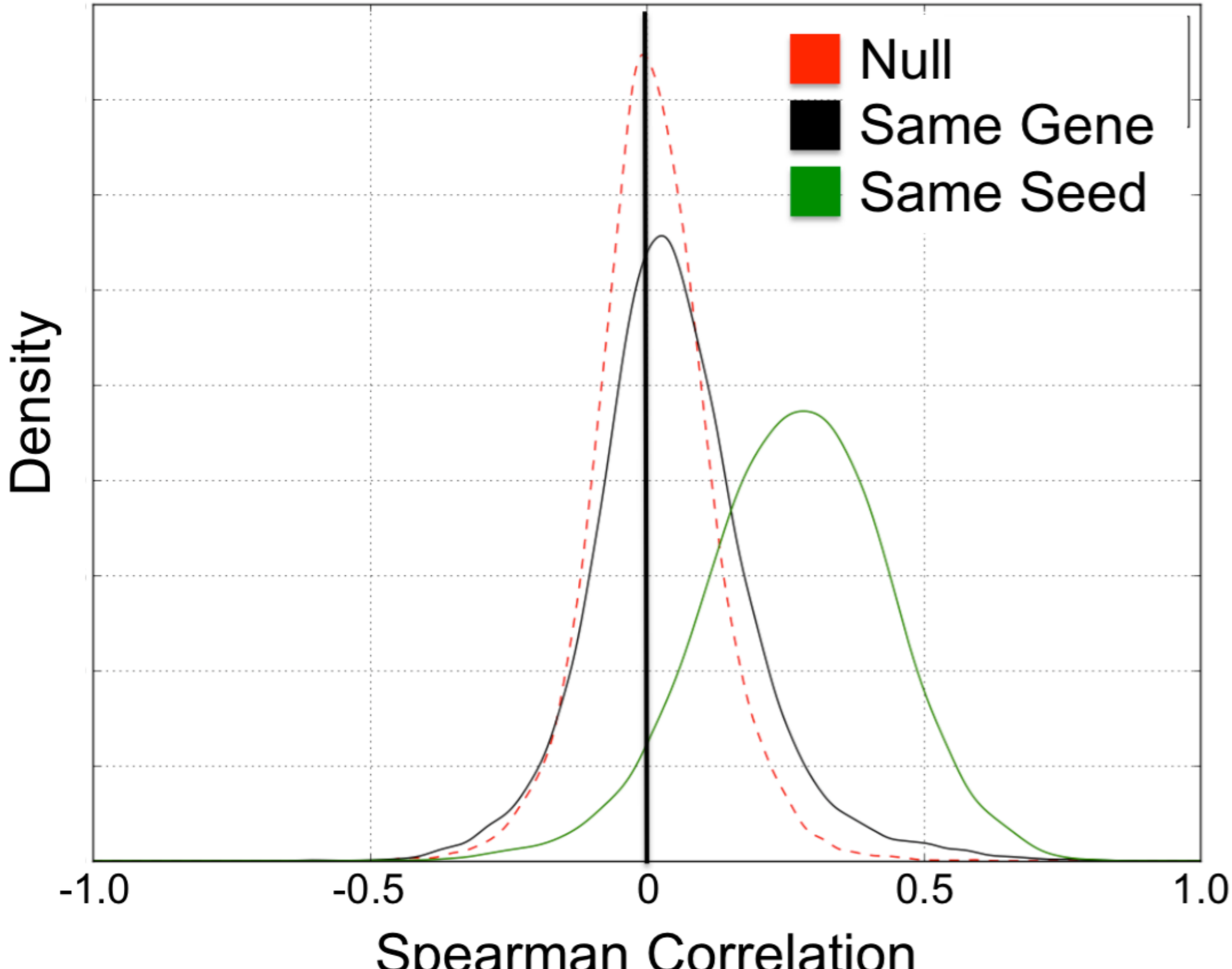
shRNA **siRNA**



‘Seed’ region binds to many 3’UTRs:
off-target effects
Specificity 6 - 8 nt

microRNA

shRNAs sharing a seed sequence correlate better than same-gene shRNAs



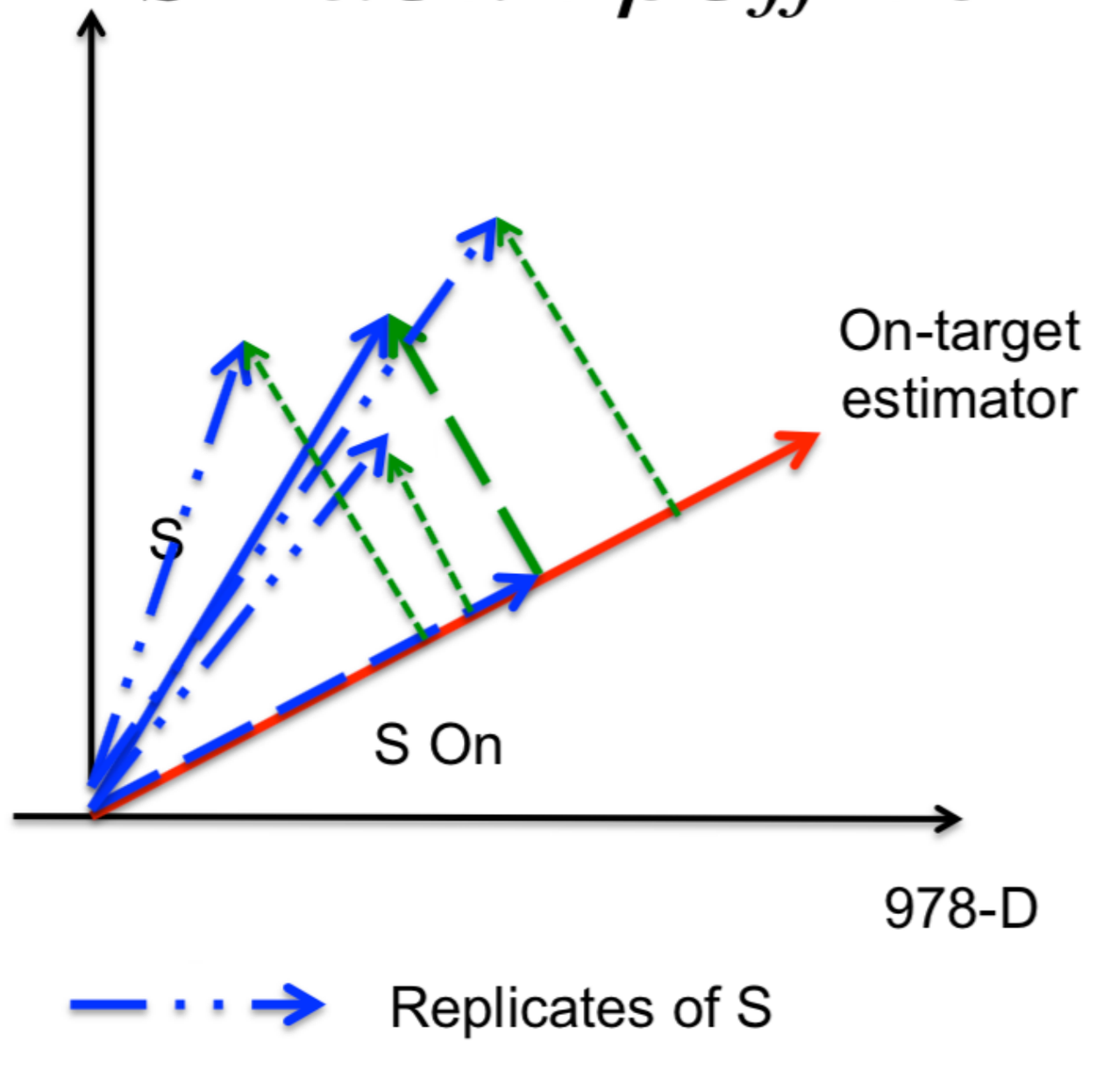
Projection Method decomposes on- and off-targets without prior knowledge of off-target mechanism

Method: Projection

All signatures are a combination of:

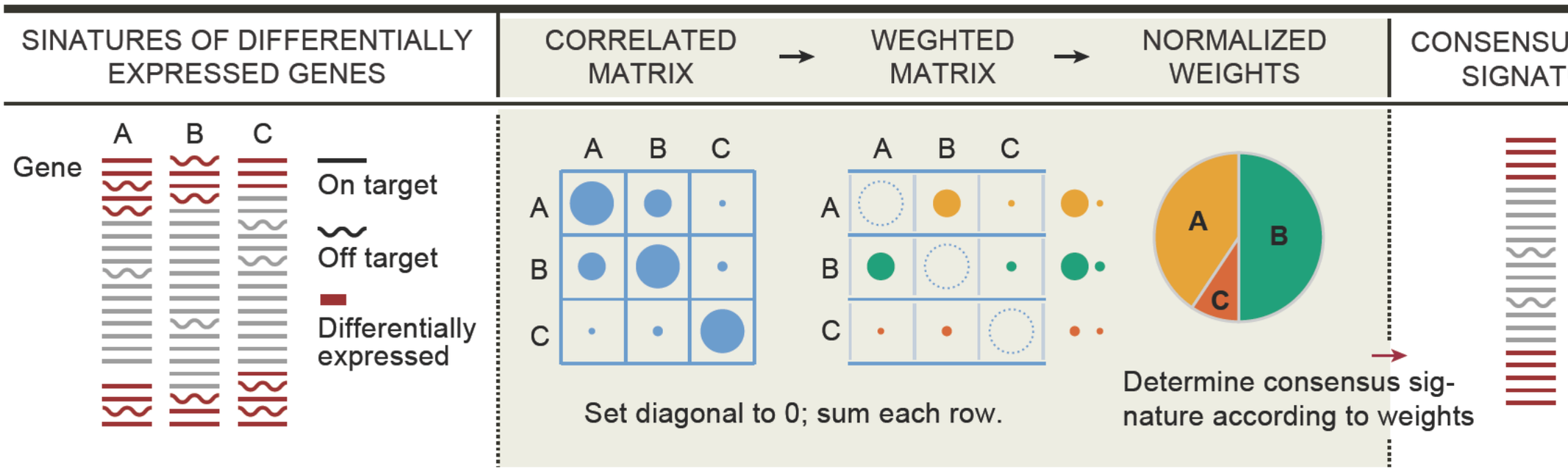
- Assay noise
  - Model with technical replicates
- On-target effects
  - Model with Consensus Gene Signature
- Off-target effects
  - Whatever is left

$$S = \alpha\vec{On} + \beta\vec{Off} + \varepsilon$$

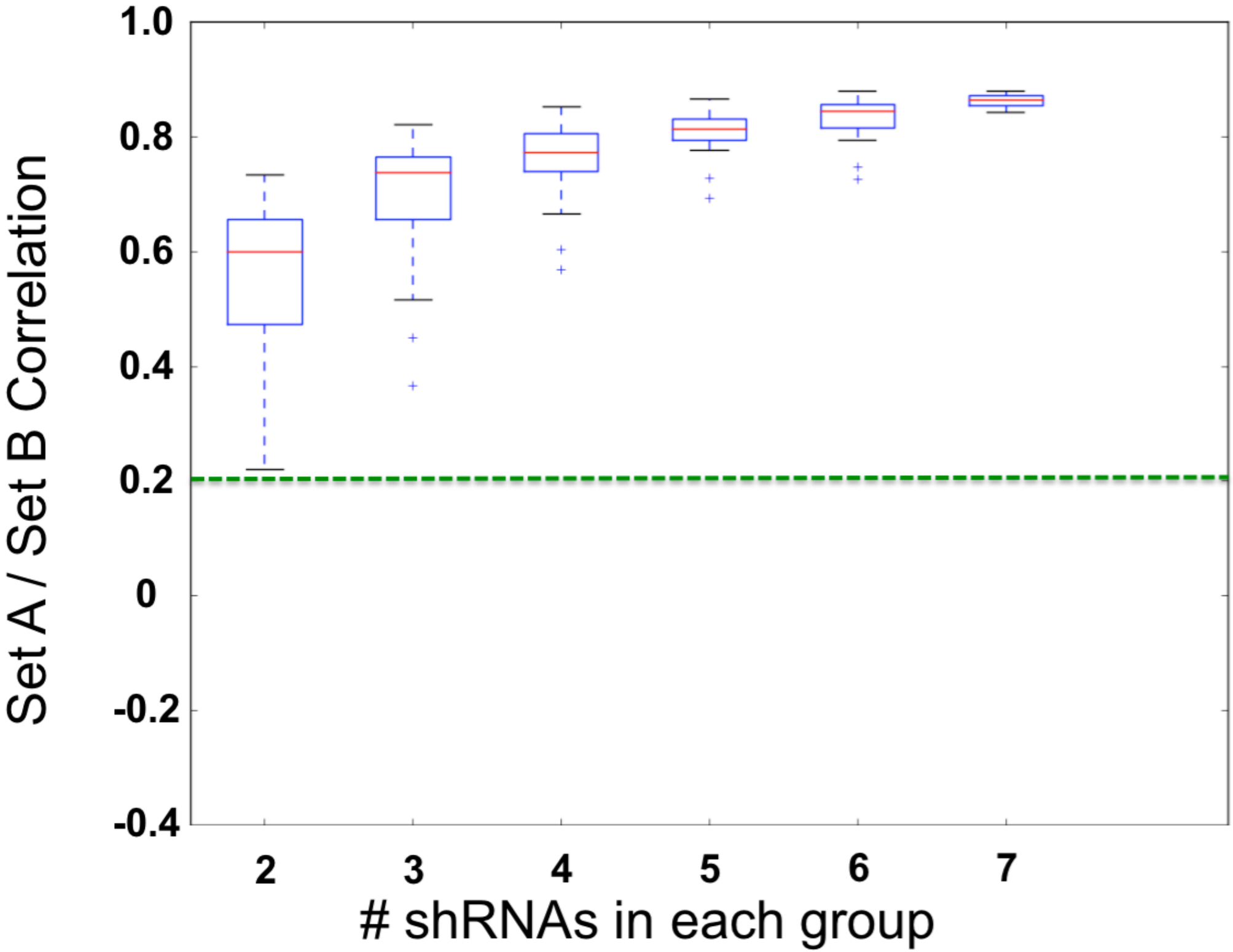


Consensus gene signatures (CGS)

Moderated Z procedure

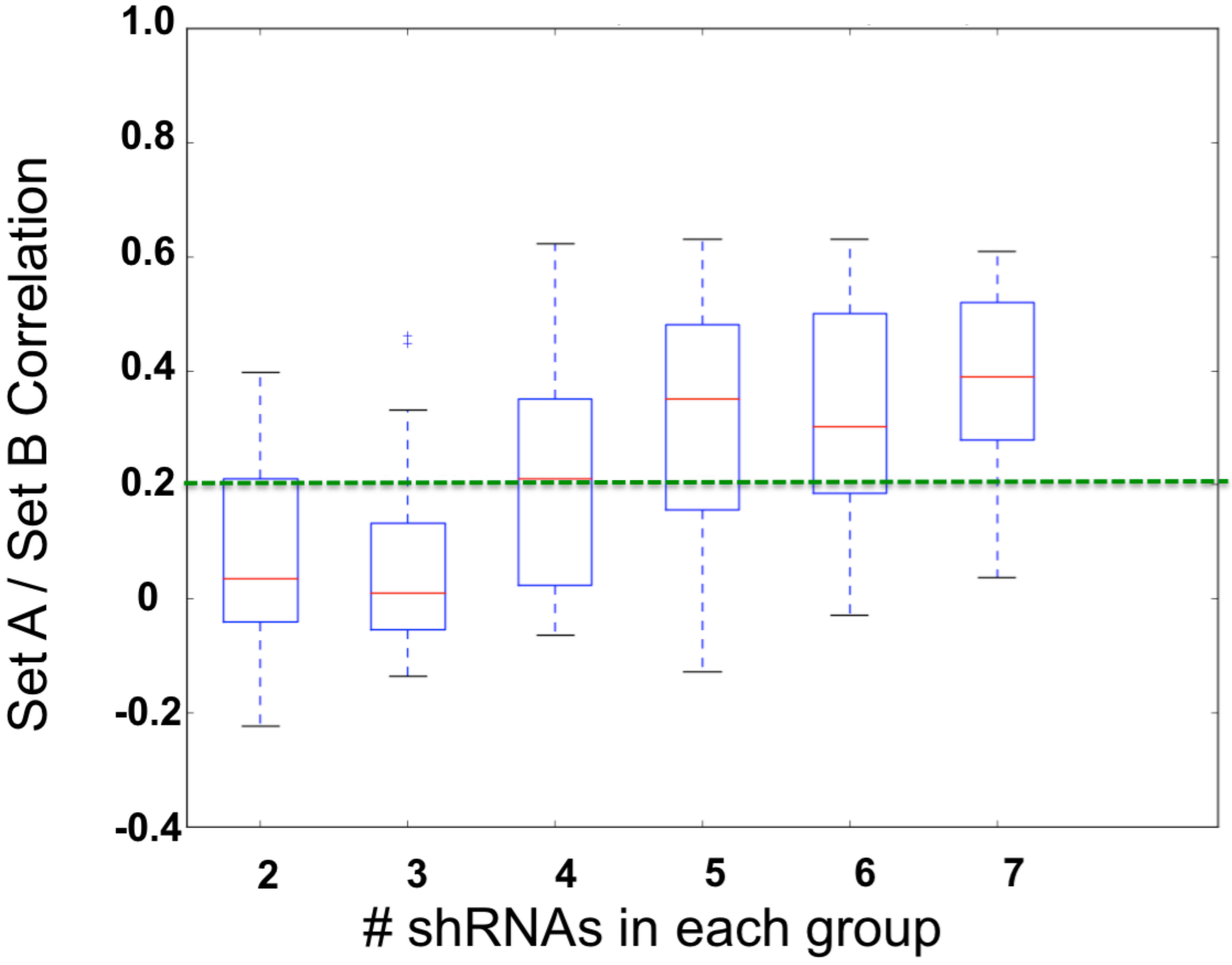


BRAF shRNAs converge in HT29 cell



BRAF shRNAs in HT29 cells, with BRAF V600E mutation

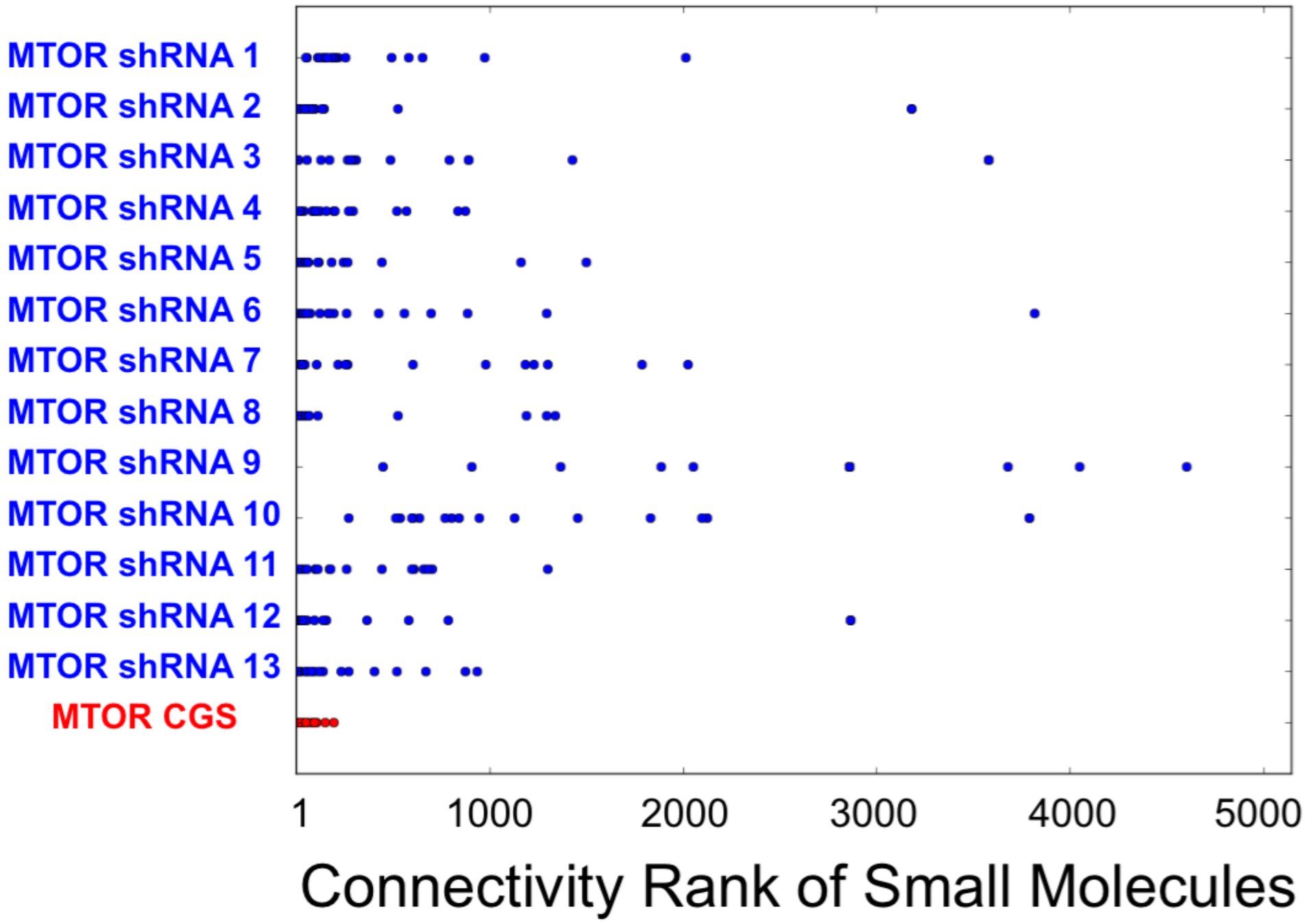
But not in HEPG2 cells (BRAF wt)



BRAF shRNAs in HepG2 cells, with wildype BRAF

Combining information from multiple shRNAs improves connections to compounds

Each dot is a dose / timepoint of rapamycin



Compound-knockdown connections

BEZ235 (PI3K/MTOR inhibitor) treatment

Compounds

rank	compound	target
1	AZD-8055	MTOR
2	KU-0063794	MTOR
3	PP-110	PI3K
4	BEZ235	MTOR, PI3K
5	PI-103	MTOR, PI3K
6	PI-828	PI3K
7	TGX-115	PI3K
9	WYE-354	MTOR
10	PIK-90	PI3K
12	OSI-027	MTOR
13	GDC-0941	PI3K
14	temsirolimus	MTOR
18	wortmannin	PI3K
19	GSK-1059615	MTOR, PI3K
20	WYE-125132	MTOR
⋮		
3,200		

Knockdown

rank	gene
1	MTOR
3	RPTOR
6	SGK1
12	KRAS
⋮	
3,800	

Compound-knockdown connections

Trametinib (MEK inhibitor) treatment

Compounds

rank	compound	target
1	selumetinib	MEK
3	PD-0325901	MEK
4	U0126	MEK
5	PD-0325901	MEK
6	MEK1-2-inhibitor	MEK
7	AS-703026	MEK
8	AZ-628	BRAF, RAF1
9	U-0126	MEK
11	PD-184352	MEK
15	ERK-inhibitor-11E	ERK
16	U-0126	MEK
18	neratinib	EGFR, ERBB2
23	PD-198306	MEK

⋮

3,200

Knockdown

rank	gene
3	MEK
7	KIT
21	EGFR
27	KRAS

⋮

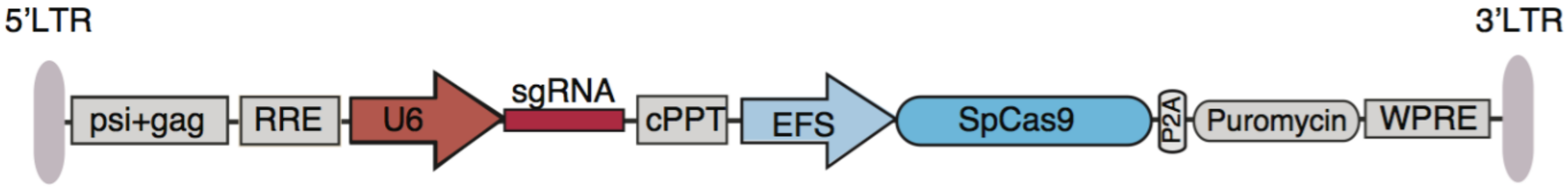
3,800

Overexpression

rank	gene
2	DUSP4
3	DUSP6
419	RAF1
423	IGF2
430	SRC

430

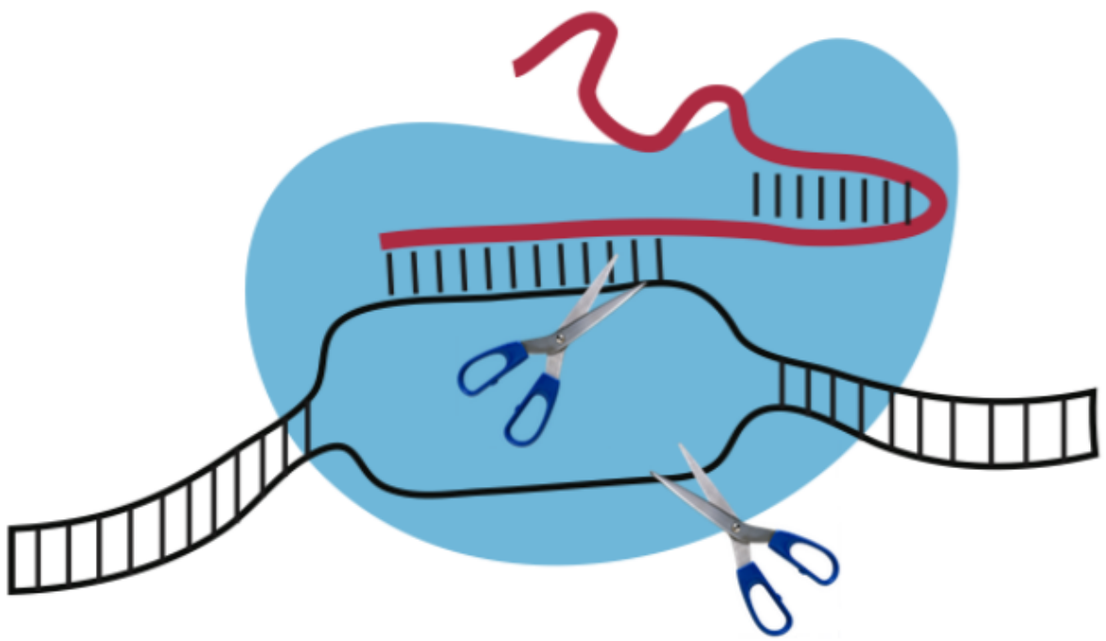
CRISPR system for lentiviral-based loss of function screening



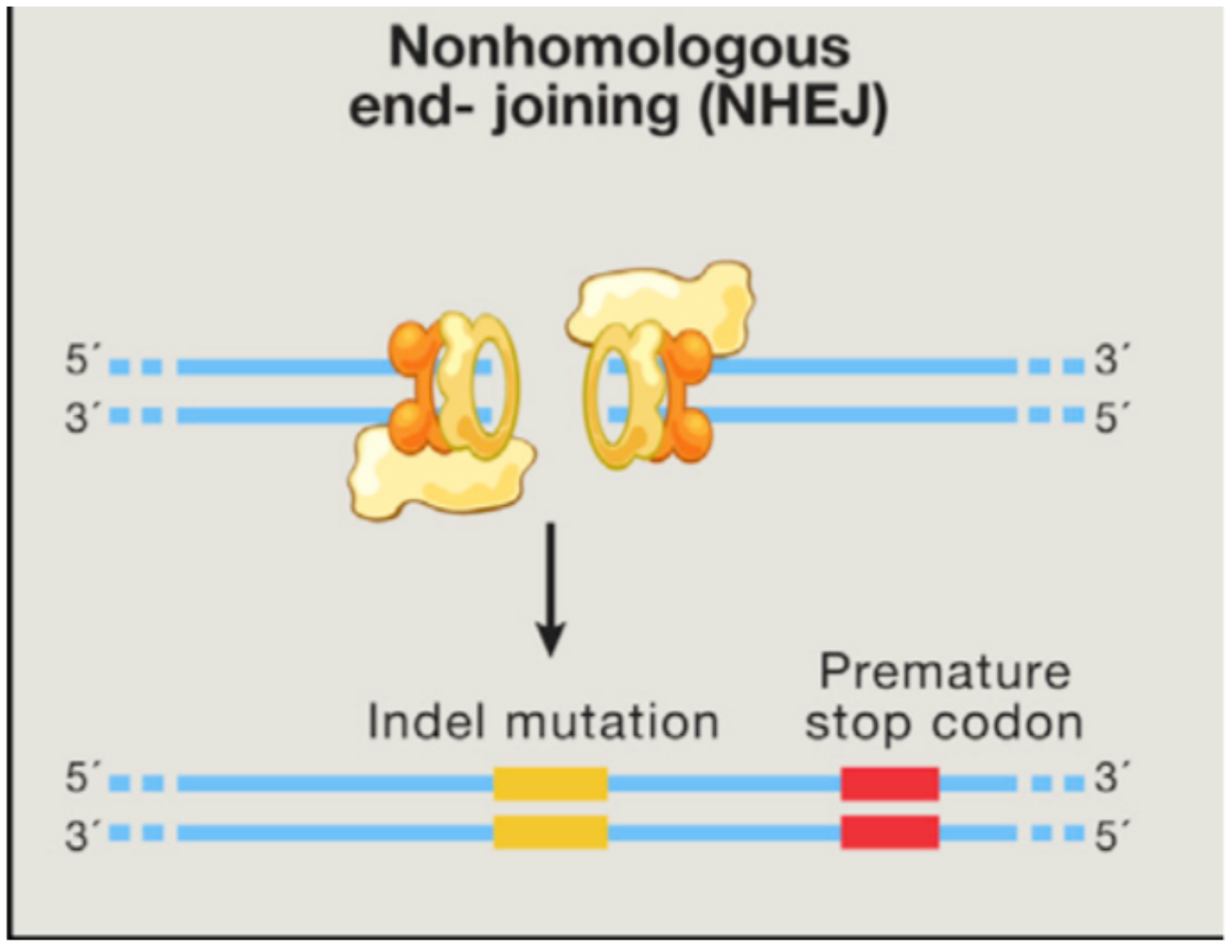
sgRNA: crRNA & tracrRNA

Program with 20nts of specificity

Binds Cas9



Cas9 endonuclease

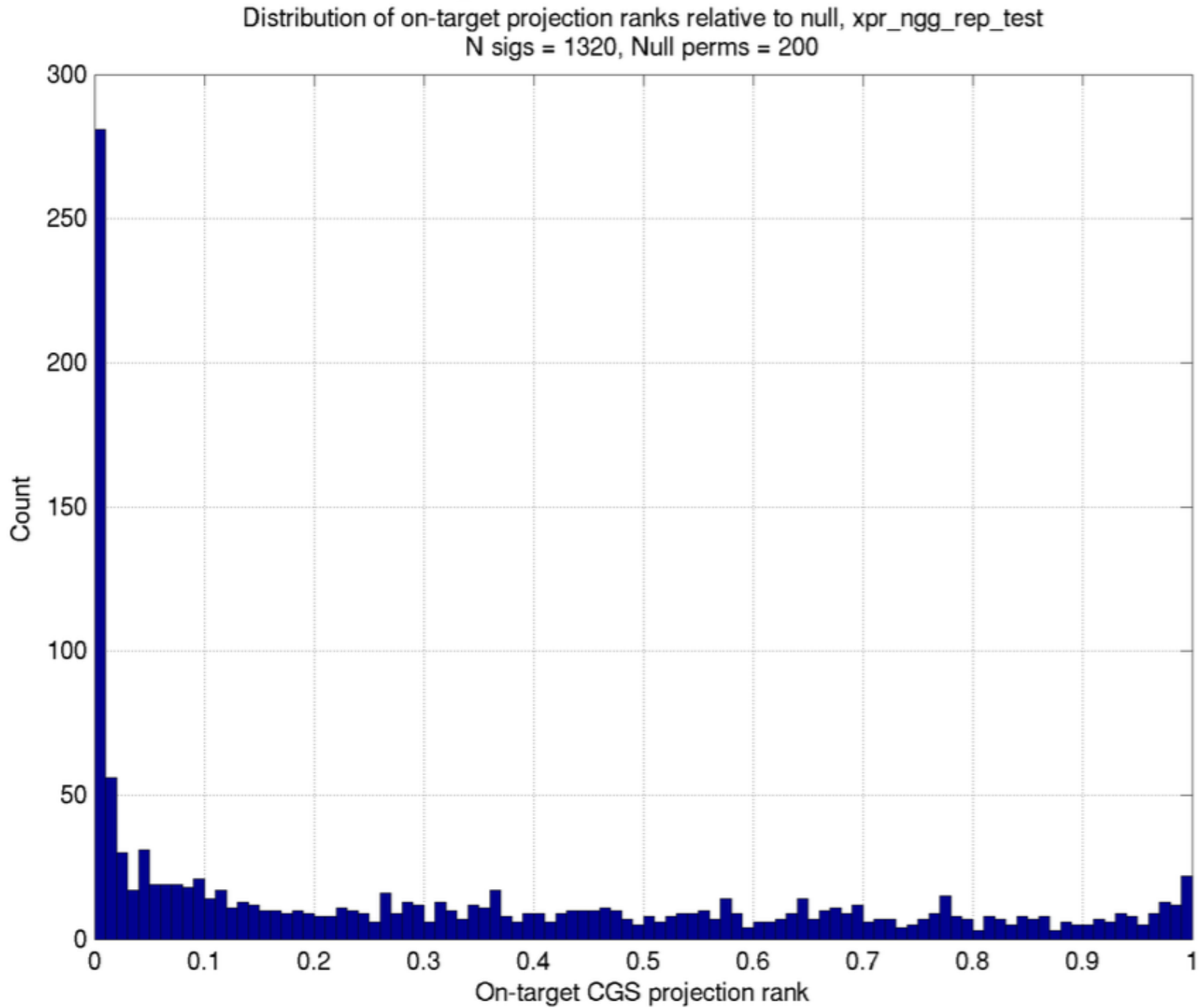


sgRNA On-target projection ranks

Validating the projection method: About 25% of CRISPR signatures find their on-target CGS

Procedure:

- Project sgRNA signature onto CGS
- Find rank CGS projection relative to 200 null CGS projections
- 257 of 1320 sgRNA signatures rank their CGS in the top 1%



Proper use of shRNAs

- Combine information from multiple shRNAs targeting a gene to reach conclusions
- shRNAs have off-target effects due to seed
  - More widespread than generally appreciated
  - All screens can benefit from explicitly examining and correcting for seed effects
- Combination of on-target effects and seed-based effects explain a very high percentage of shRNA activity
  - **We understand the proper use and analysis of these reagents**
- Explicit seed-matched controls can be used at small-scale to confirm on-target efficacy
  - Stop using your 'favorite' EGFP shRNA as a control!
- Orthogonal reagents to validate phenotype, e.g. cDNAs, CRISPRs

PHARMACOLOGICAL CLASSES

BROAD INSTITUTE

Given the mRNA expression signature of a TEST perturbagen

1. BIOACTIVITY

Does the perturbagen elicit a gene expression response?

2. PATHWAYS

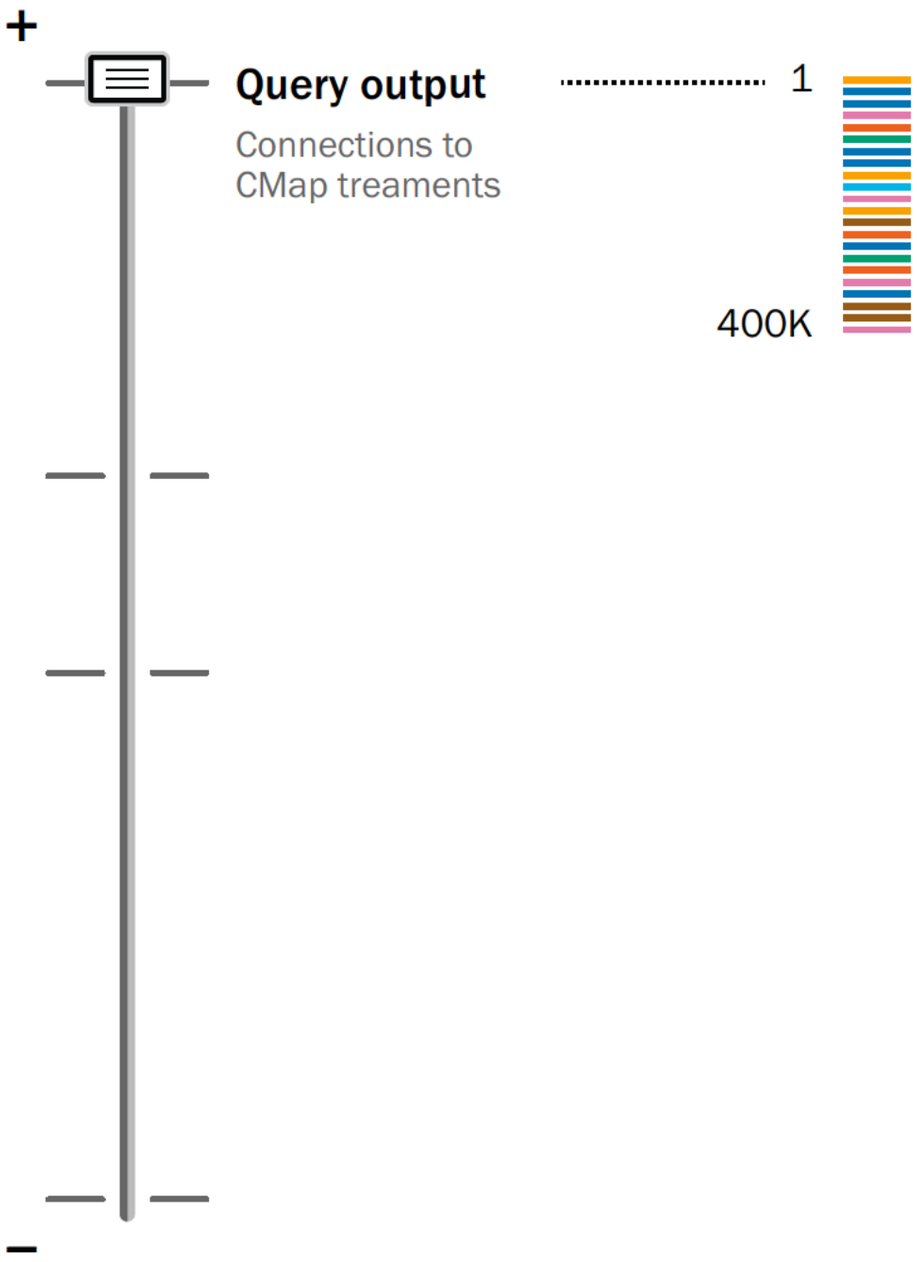
What gene pathways (or “networks”) are modulated?

3. PHARMACOLOGICAL CLASSES

Which (previously characterized) pharmacological activities match the test perturbagen?

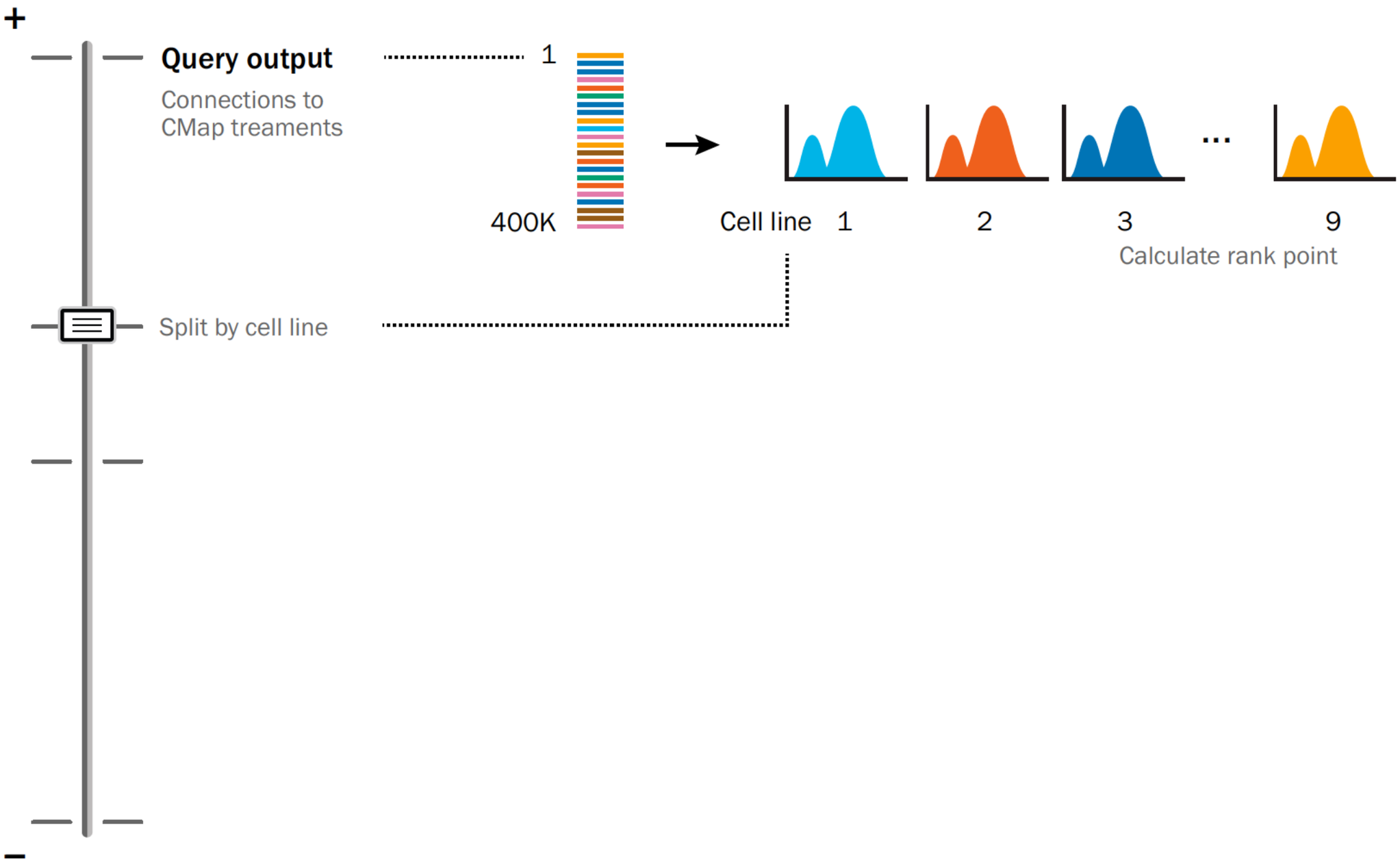
CMap result analysis

Many views of a single query



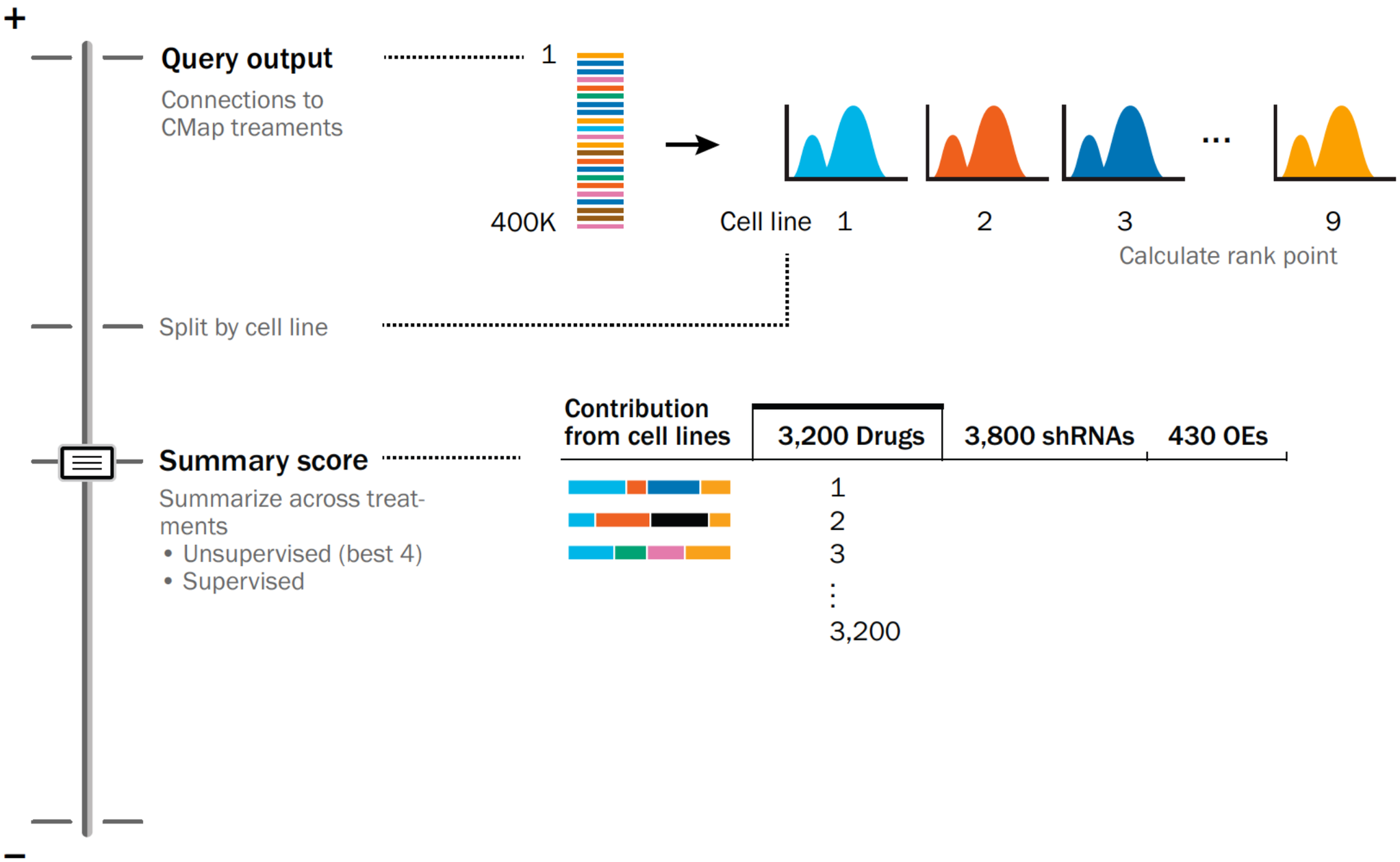
CMap result analysis

Many views of a single query



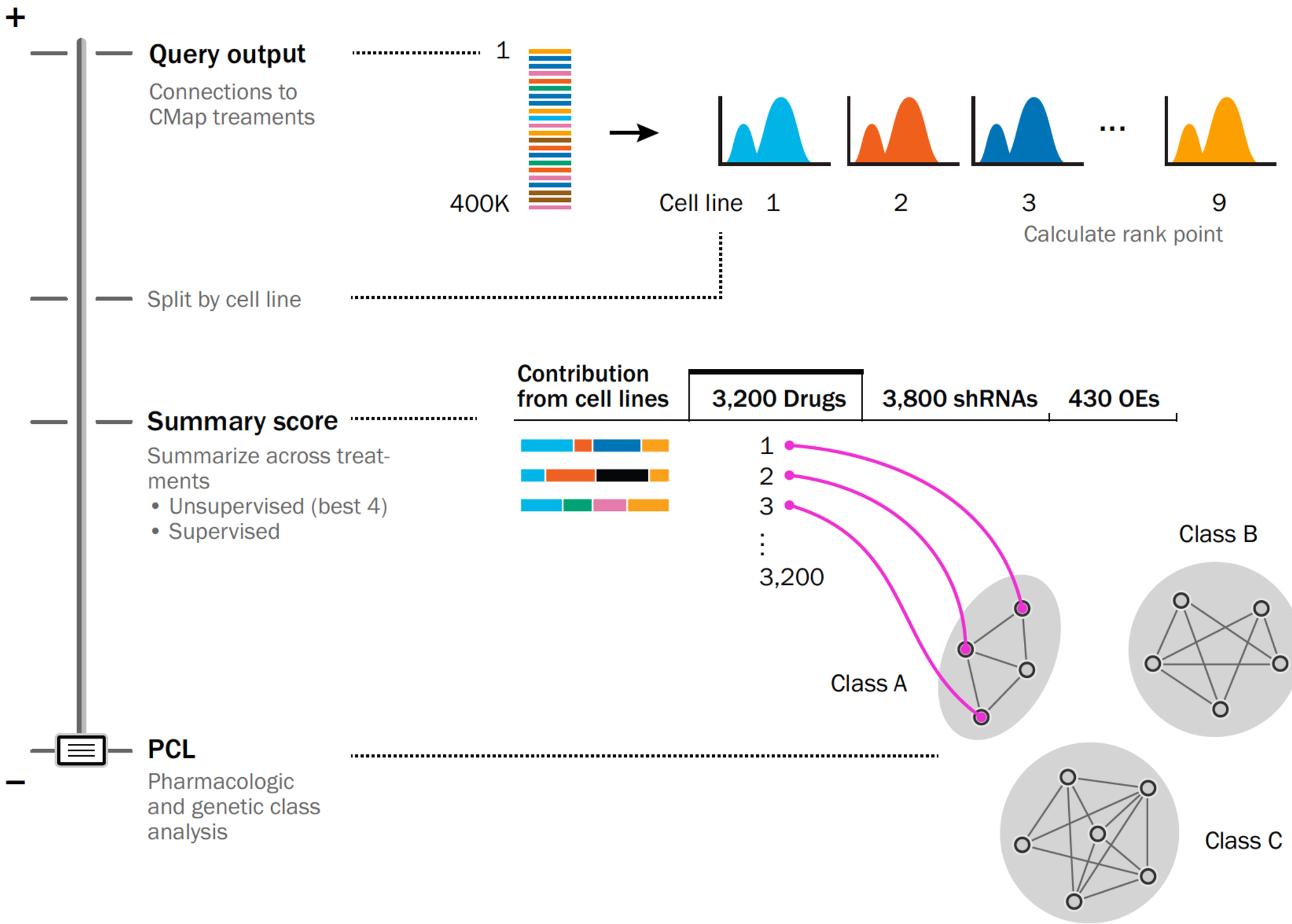
CMap result analysis

Many views of a single query



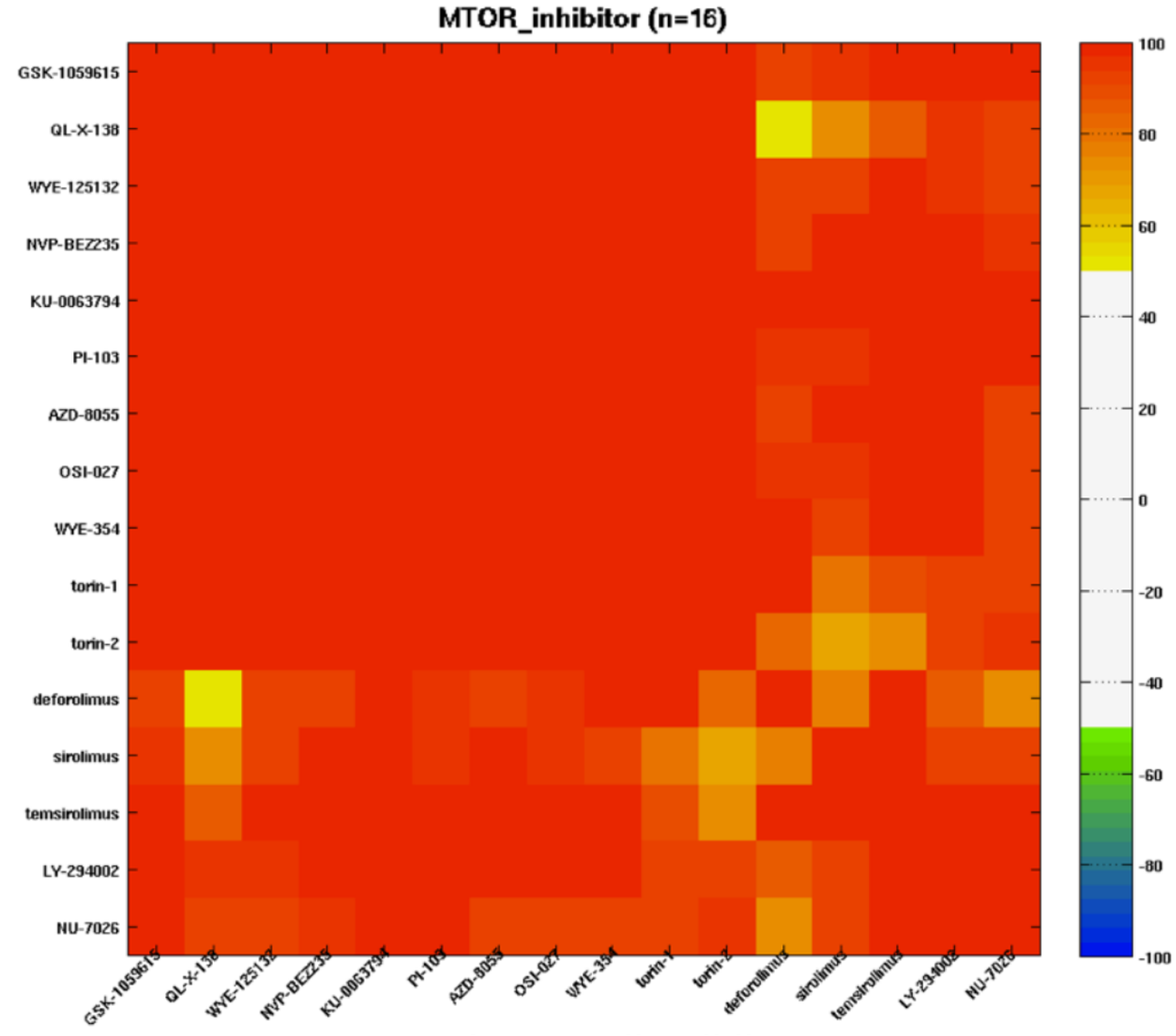
CMap result analysis

Many views of a single query

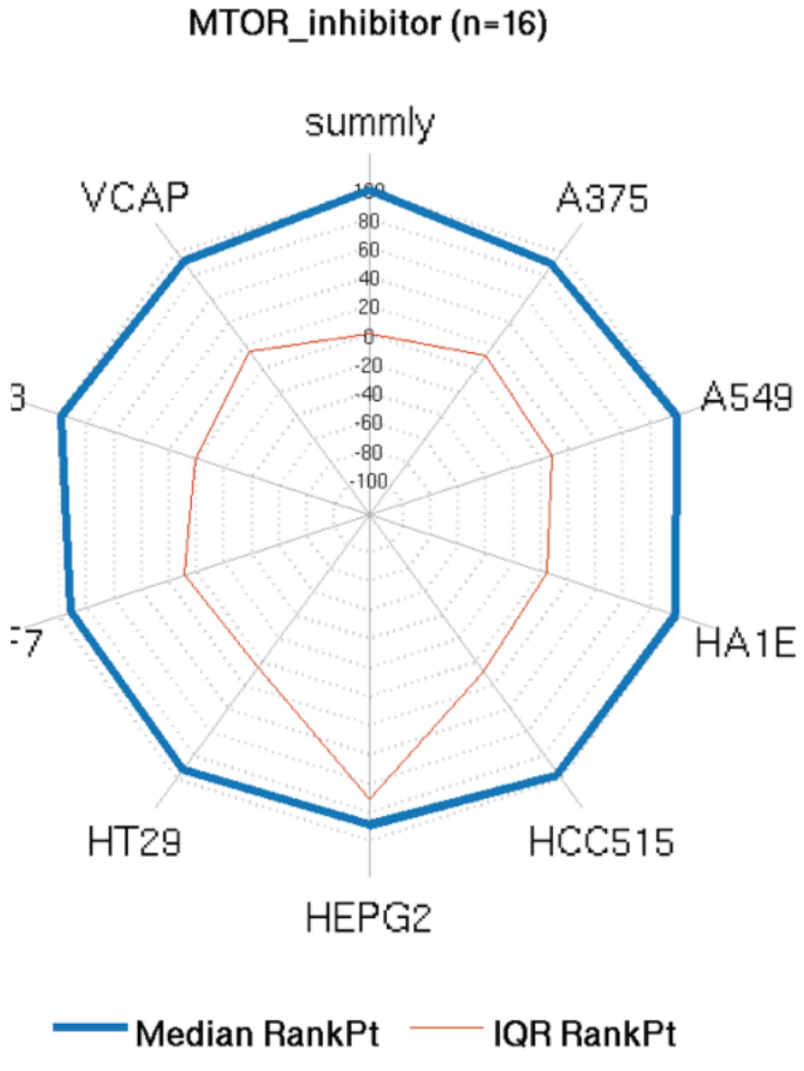


Robust PCI example

MTOR inhibitors

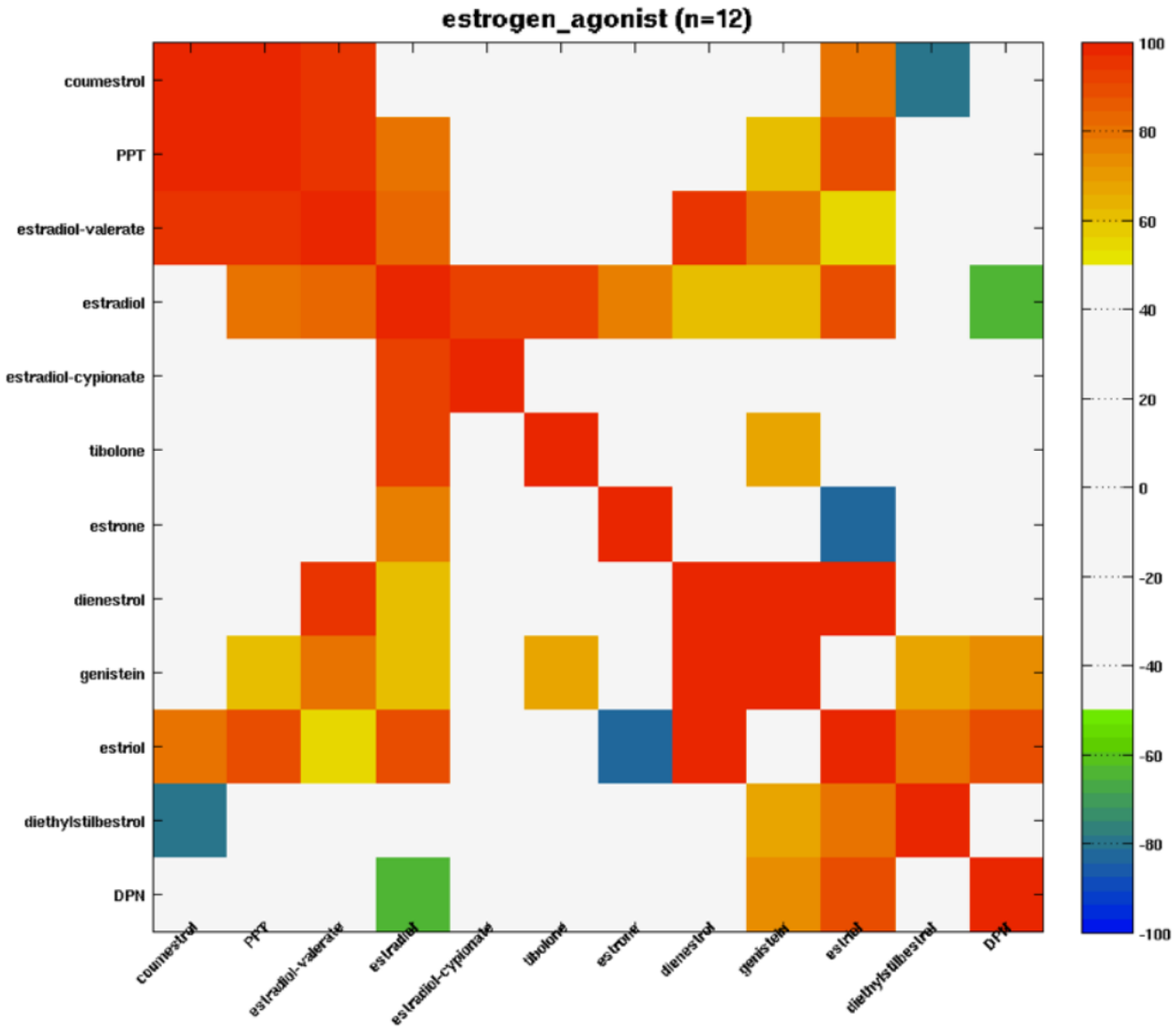


Across all cell lines

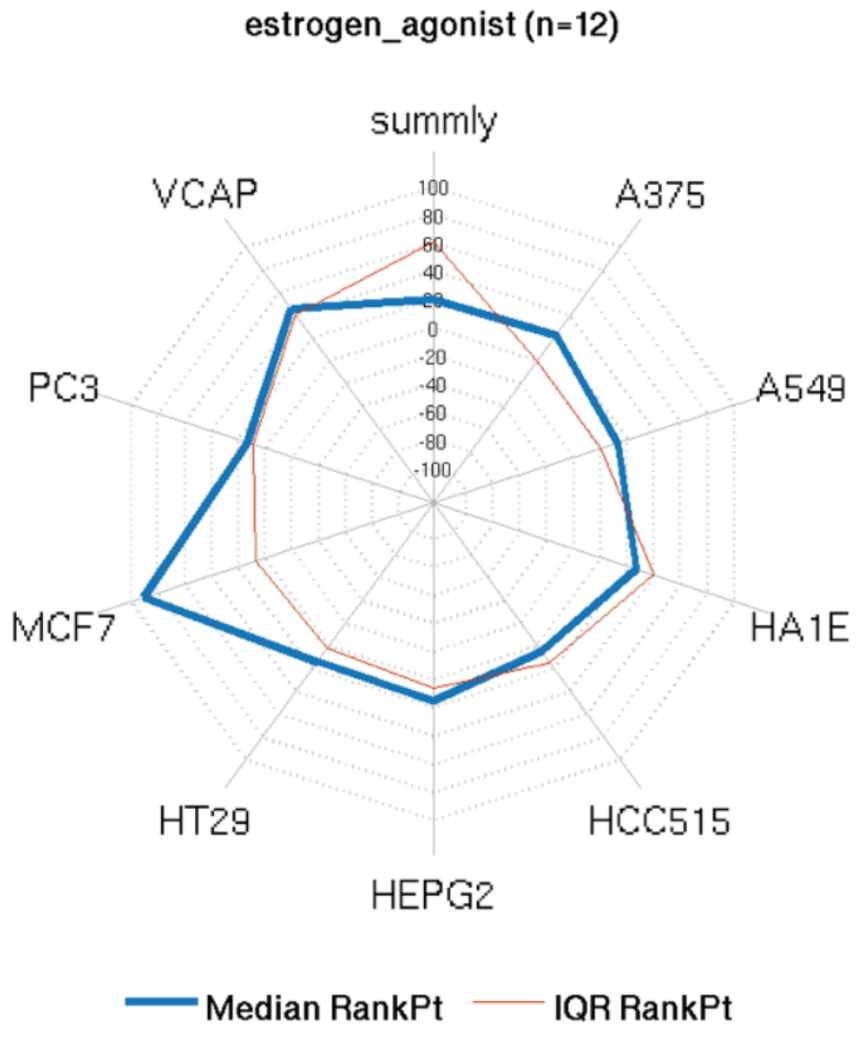


Cell type-specific PCI example

Estrogen signaling

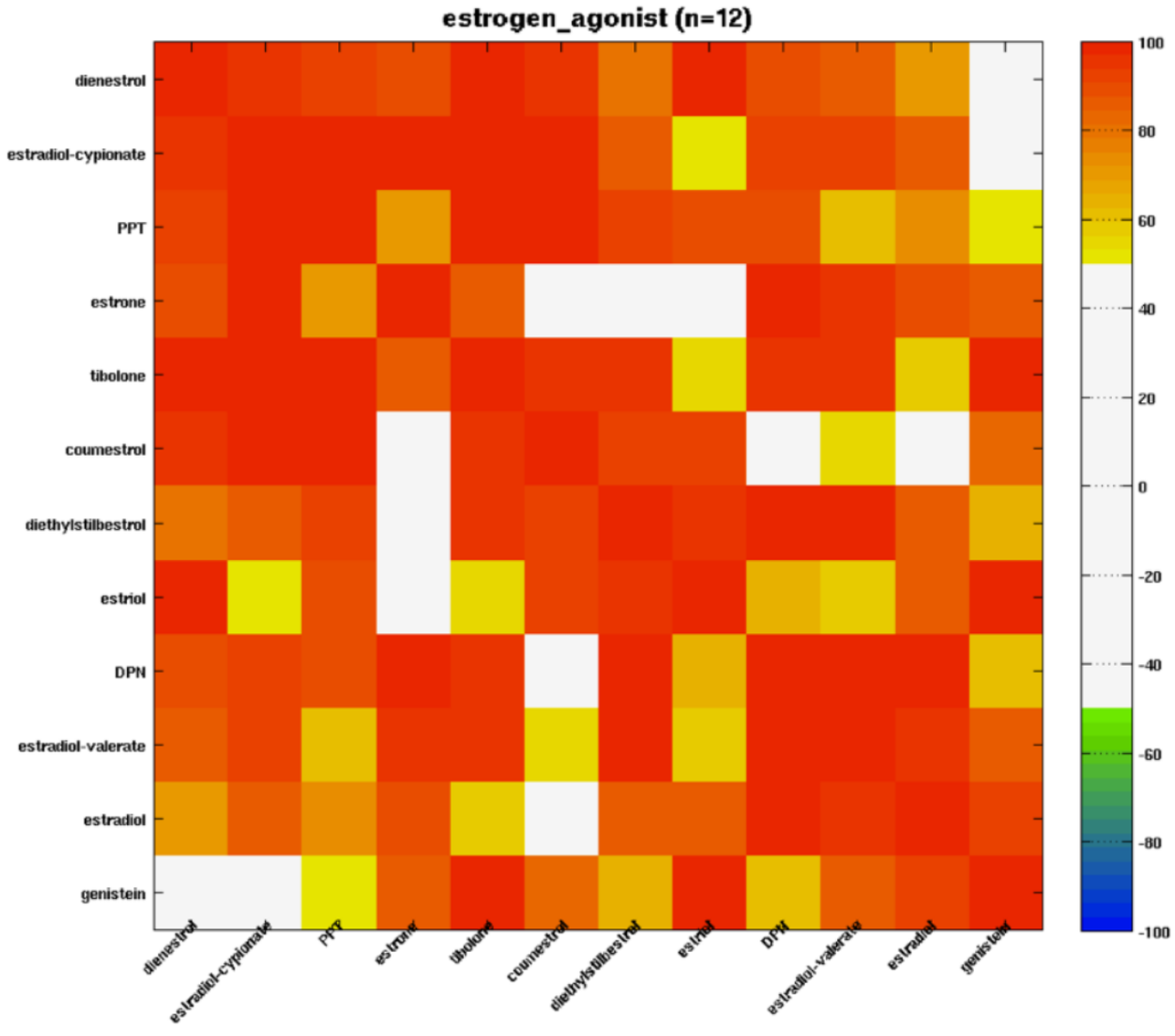


Across all cell lines

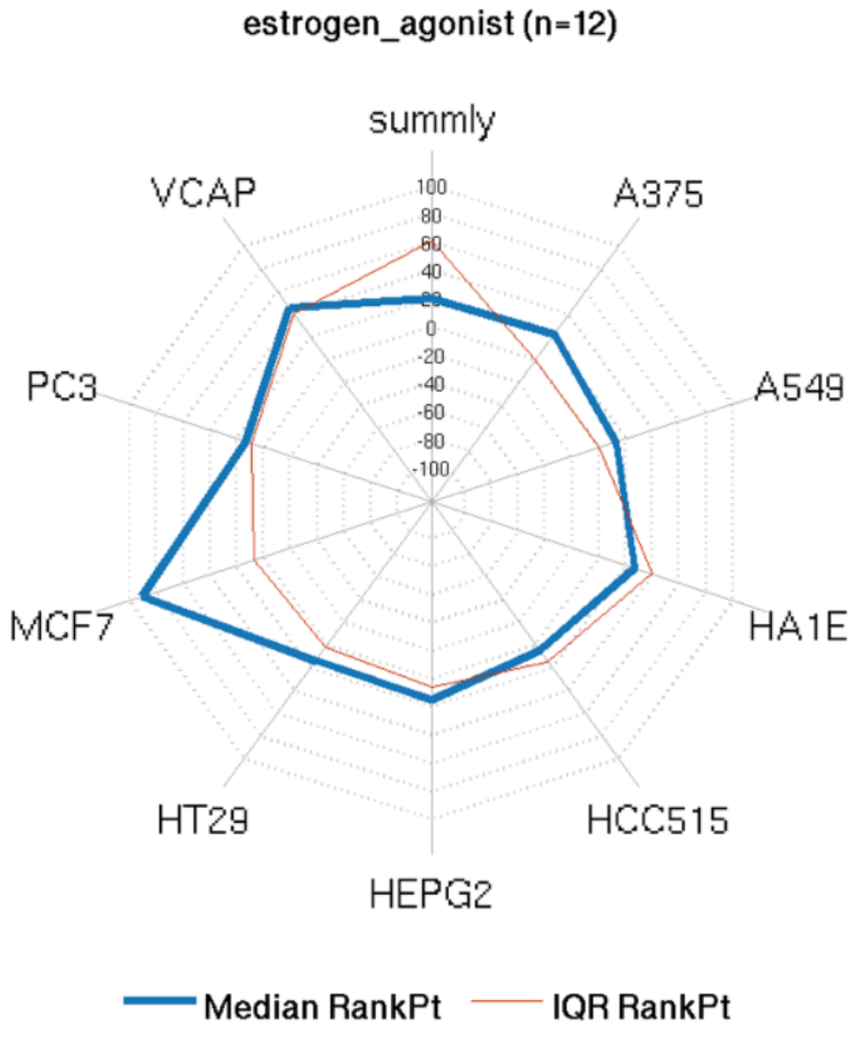


Cell type-specific PCL example

Estrogen signaling



MCF-7 only

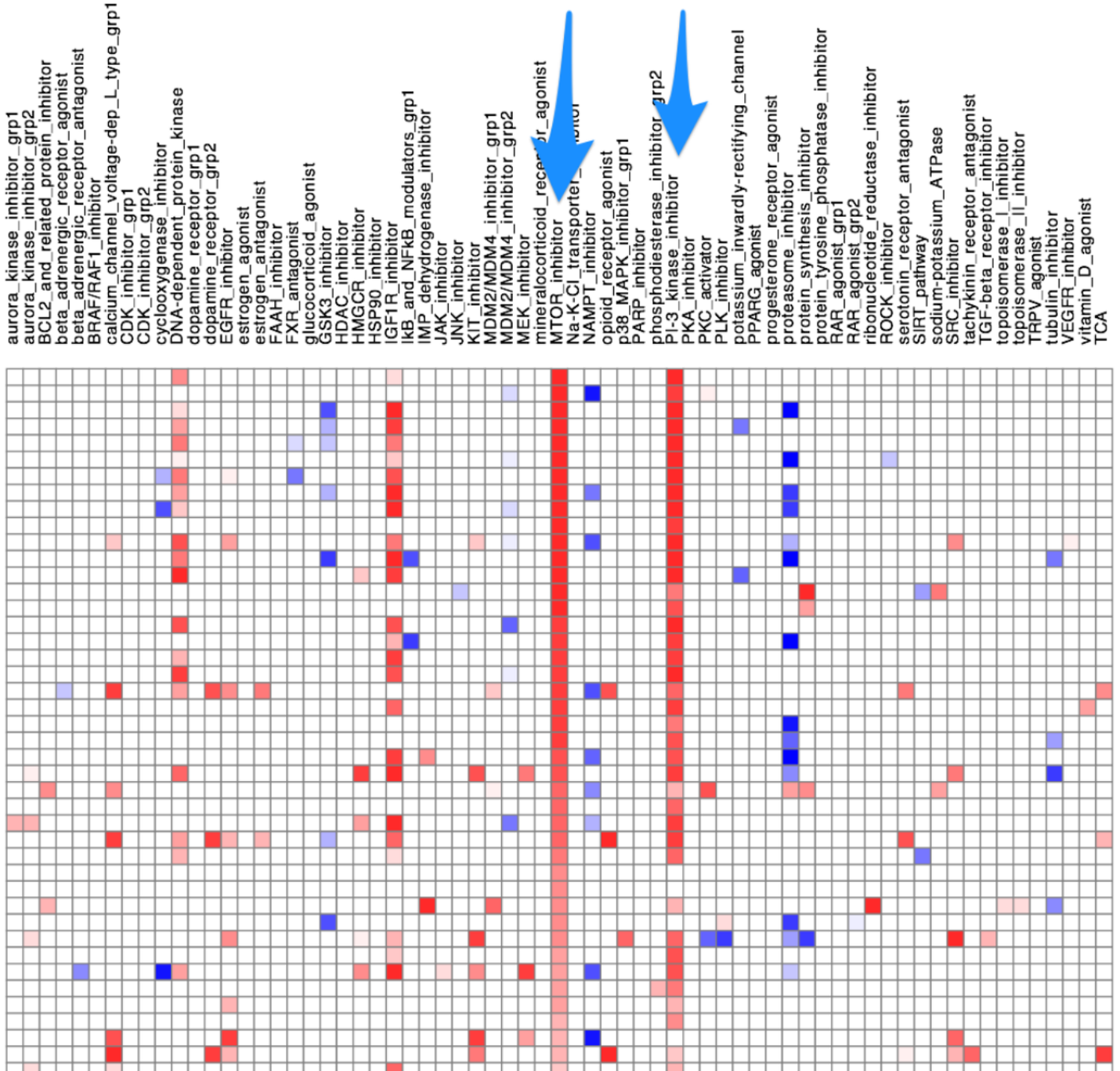


example result: Connectivity to 67 pharmacological classes - compounds sorted by matches to MTOR / PI3K inhibitors

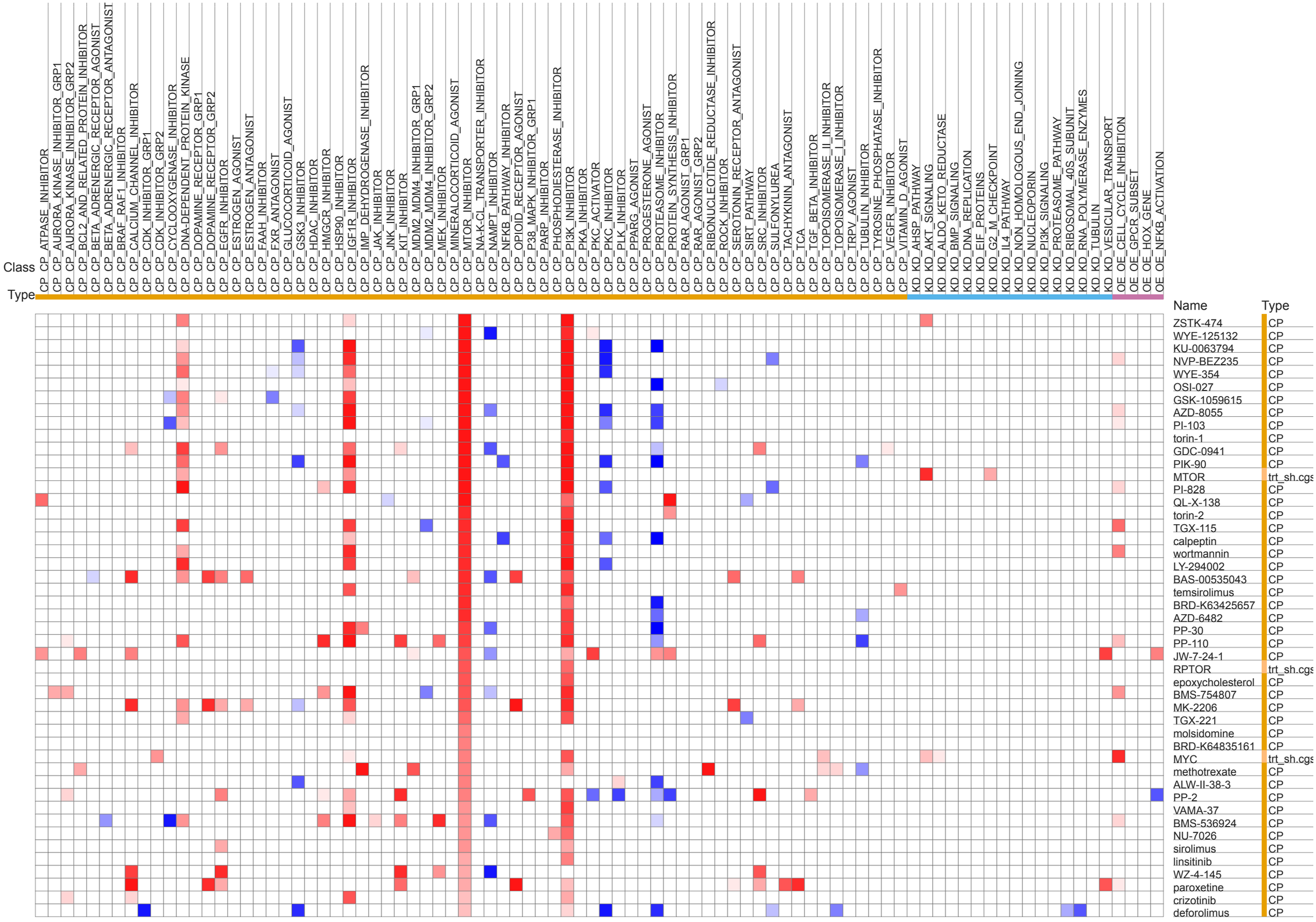
Columns: Pharmacological classes

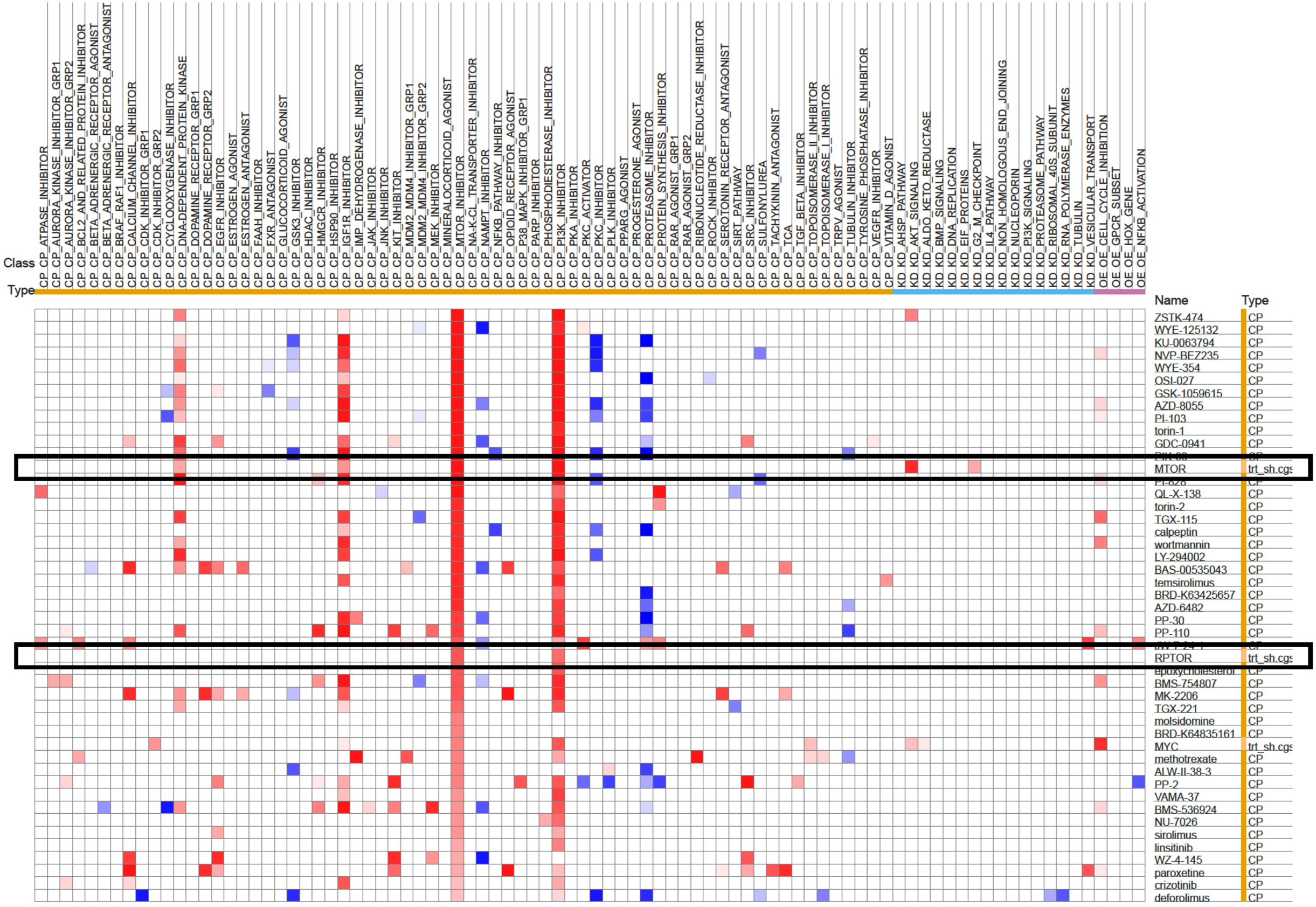
MTOR class

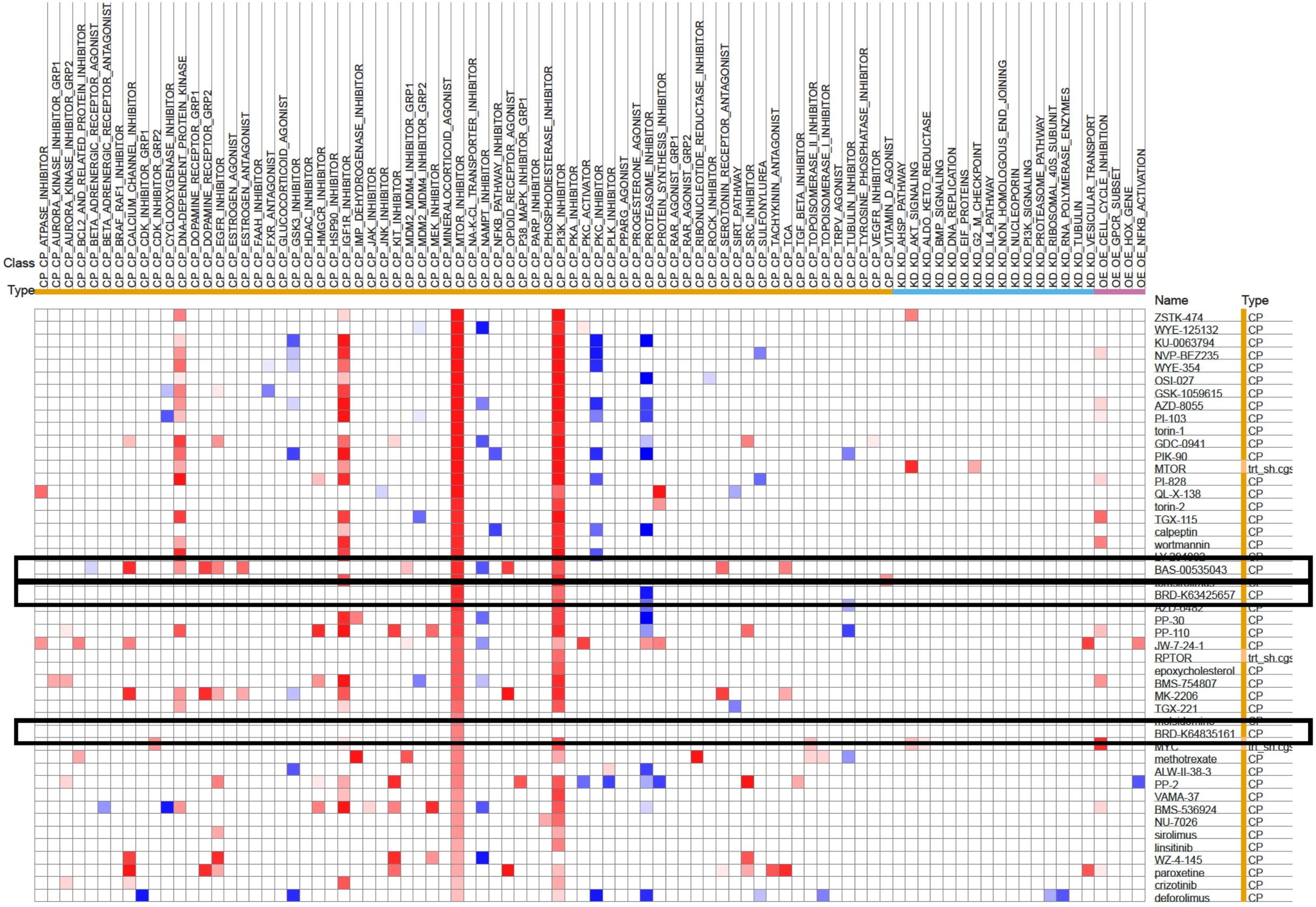
Pi3K class



Rows: Top 20 Compounds - know and novel (out of 4,000+ columns)

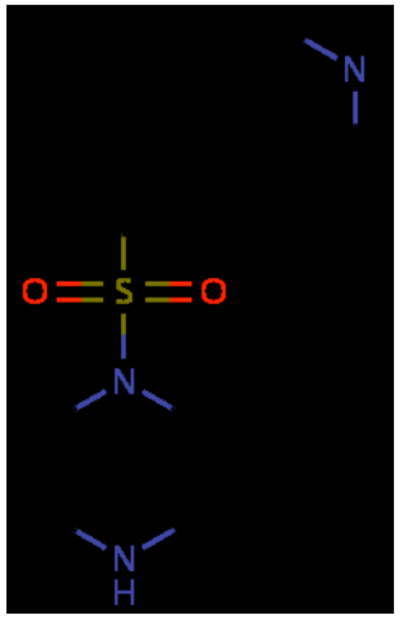

Class
Type
CP_ATPASE_INHIBITOR
CP_AURORA_KINASE_INHIBITOR_GRP1
CP_AURORA_KINASE_INHIBITOR_GRP2
CP_BCL2_AND_RELATED_PROTEIN_INHIBITOR
CP_BETA_ADRENERGIC_RECEPTOR_AGONIST
CP_BETA_ADRENERGIC_RECEPTOR_ANTAGONIST
CP_BRAF_RAF1_INHIBITOR
CP_CALCIUM_CHANNEL_INHIBITOR
CP_CDK_INHIBITOR_GRP1
CP_CDK_INHIBITOR_GRP2
CP_CYCLOOXYGENASE_INHIBITOR
CP_DNA-DEPENDENT_PROTEIN_KINASE
CP_DOPAMINE_RECEPTOR_GRP1
CP_DOPAMINE_RECEPTOR_GRP2
CP_EGFR_INHIBITOR
CP_ESTROGEN_AGONIST
CP_ESTROGEN_ANTAGONIST
CP_FAAH_INHIBITOR
CP_FXR_ANTAGONIST
CP_GLUCOCORTICOID_AGONIST
CP_GSK3_INHIBITOR
CP_HDAC_INHIBITOR
CP_HMGCR_INHIBITOR
CP_HSP90_INHIBITOR
CP_IGF1R_INHIBITOR
CP_IMP_DEHYDROGENASE_INHIBITOR
CP_JAK_INHIBITOR
CP_JNK_INHIBITOR
CP_KIT_INHIBITOR
CP_MDM2_MDM4_INHIBITOR_GRP1
CP_MDM2_MDM4_INHIBITOR_GRP2
CP_MEK_INHIBITOR
CP_MINERALOCORTICOID_AGONIST
CP_MTOR_INHIBITOR
CP_NA-K-CL_TRANSPORTER_INHIBITOR
CP_NAMPT_INHIBITOR
CP_NFKB_PATHWAY_INHIBITOR
CP_OPIOID_RECEPTOR_AGONIST
CP_P38_MAPK_INHIBITOR_GRP1
CP_PARP_INHIBITOR
CP_PHOSPHODIESTERASE_INHIBITOR
CP_PI3K_INHIBITOR
CP_PKA_INHIBITOR
CP_PKC_ACTIVATOR
CP_PKC_INHIBITOR
CP_PLK_INHIBITOR
CP_PPARG_AGONIST
CP_PROGESTERONE_AGONIST
CP_PROTEASOME_INHIBITOR
CP_PROTEIN_SYNTHESIS_INHIBITOR
CP_RAR_AGONIST_GRP1
CP_RAR_AGONIST_GRP2
CP_RIBONUCLEOTIDE_REDUCTASE_INHIBITOR
CP_ROCK_INHIBITOR
CP_SEROTONIN_RECEPTOR_ANTAGONIST
CP_SIRT_PATHWAY
CP_SRC_INHIBITOR
CP_SULFONYLUREA
CP_TACHYKININ_ANTAGONIST
CP_TCA
CP_TGF_BETA_INHIBITOR
CP_TOPOISOMERASE_II_INHIBITOR
CP_TOPOISOMERASE_I_INHIBITOR
CP_TRPV_AGONIST
CP_TUBULIN_INHIBITOR
CP_TYROSINE_PHOSPHATASE_INHIBITOR
CP_VEGFR_INHIBITOR
CP_VITAMIN_D_AGONIST
KD_AHSP_PATHWAY
KD_AKT_SIGNALING
KD_ALDO_KETO_REDUCTASE
KD_BMP_SIGNALING
KD_DNA_REPLICATION
KD_EIF_PROTEINS
KD_G2_M_CHECKPOINT
KD_IL4_PATHWAY
KD_NON_HOMOLOGOUS_END_JOINING
KD_NUCLEOPORIN
KD_PI3K_SIGNALING
KD_PROTEASOME_PATHWAY
KD_RIBOSOMAL_40S_SUBUNIT
KD_RNA_POLYMERASE_ENZYMES
KD_TUBULIN
KD_VESICULAR_TRANSPORT
OE_CELL_CYCLE_INHIBITION
OE_GPCR_SUBSET
OE_HOX_GENE
OE_NFKB_ACTIVATION
Name
Type
ZSTK-474 CP
WYE-125132 CP
KU-0063794 CP
NVP-BEZ235 CP
WYE-354 CP
OSI-027 CP
GSK-1059615 CP
AZD-8055 CP
PI-103 CP
torin-1 CP
GDC-0941 CP
PIK-90 CP
MTOR trt_sh.cgs
PI-828 CP
QL-X-138 CP
torin-2 CP
TGX-115 CP
calpeptin CP
wortmannin CP
LY-294002 CP
BAS-00535043 CP
temsirolimus CP
BRD-K63425657 CP
AZD-6482 CP
PP-30 CP
PP-110 CP
JW-7-24-1 CP
RPTOR trt_sh.cgs
epoxycholesterol CP
BMS-754807 CP
MK-2206 CP
TGX-221 CP
molsidomine CP
BRD-K64835161 CP
MYC trt_sh.cgs
methotrexate CP
ALW-II-38-3 CP
PP-2 CP
VAMA-37 CP
BMS-536924 CP
NU-7026 CP
sirolimus CP
linsitinib CP
WZ-4-145 CP
paroxetine CP
crizotinib CP
deforolimus CP

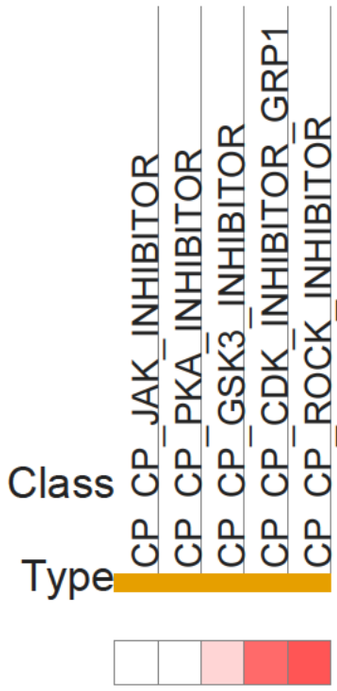

Class
Type
CP CP_ATPASE_INHIBITOR
CP CP_AURORA_KINASE_INHIBITOR_GRP1
CP CP_AURORA_KINASE_INHIBITOR_GRP2
CP CP_BCL2_AND_RELATED_PROTEIN_INHIBITOR
CP CP_BETA_ADRENERGIC_RECEPTOR_AGONIST
CP CP_BETA_ADRENERGIC_RECEPTOR_ANTAGONIST
CP CP_BRAF_RAF1_INHIBITOR
CP CP_CALCIUM_CHANNEL_INHIBITOR
CP CP_CDK_INHIBITOR_GRP1
CP CP_CDK_INHIBITOR_GRP2
CP CP_CYCLOOXYGENASE_INHIBITOR
CP CP_DNA-DEPENDENT_PROTEIN_KINASE
CP CP_DOPAMINE_RECEPTOR_GRP1
CP CP_DOPAMINE_RECEPTOR_GRP2
CP CP_EGFR_INHIBITOR
CP CP_ESTROGEN_AGONIST
CP CP_ESTROGEN_ANTAGONIST
CP CP_FAAH_INHIBITOR
CP CP_FXR_ANTAGONIST
CP CP_GLUCOCORTICOID_AGONIST
CP CP_GSK3_INHIBITOR
CP CP_HDAC_INHIBITOR
CP CP_HMGCR_INHIBITOR
CP CP_HSP90_INHIBITOR
CP CP_IGF1R_INHIBITOR
CP CP_IMP_DEHYDROGENASE_INHIBITOR
CP CP_JAK_INHIBITOR
CP CP_JNK_INHIBITOR
CP CP_KIT_INHIBITOR
CP CP_MDM2_MDM4_INHIBITOR_GRP1
CP CP_MDM2_MDM4_INHIBITOR_GRP2
CP CP_MEK_INHIBITOR
CP CP_MINERALOCORTICOID_AGONIST
CP CP_MTOR_INHIBITOR
CP CP_NA-K-CL_TRANSPORTER_INHIBITOR
CP CP_NAMPT_INHIBITOR
CP CP_NFKB_PATHWAY_INHIBITOR
CP CP_OPIOID_RECEPTOR_AGONIST
CP CP_P38_MAPK_INHIBITOR_GRP1
CP CP_PARP_INHIBITOR
CP CP_PHOSPHODIESTERASE_INHIBITOR
CP CP_PI3K_INHIBITOR
CP CP_PKA_INHIBITOR
CP CP_PKC_ACTIVATOR
CP CP_PKC_INHIBITOR
CP CP_PLK_INHIBITOR
CP CP_PPARG_AGONIST
CP CP_PROGESTERONE_AGONIST
CP CP_PROTEASOME_INHIBITOR
CP CP_PROTEIN_SYNTHESIS_INHIBITOR
CP CP_RAR_AGONIST_GRP1
CP CP_RAR_AGONIST_GRP2
CP CP_RIBONUCLEOTIDE_REDUCTASE_INHIBITOR
CP CP_ROCK_INHIBITOR
CP CP_SEROTONIN_RECEPTOR_ANTAGONIST
CP CP_SIRT_PATHWAY
CP CP_SRC_INHIBITOR
CP CP_SULFONYLUREA
CP CP_TACHYKININ_ANTAGONIST
CP CP_TCA
CP CP_TGF_BETA_INHIBITOR
CP CP_TOPOISOMERASE_II_INHIBITOR
CP CP_TOPOISOMERASE_I_INHIBITOR
CP CP_TRPV_AGONIST
CP CP_TUBULIN_INHIBITOR
CP CP_TYROSINE_PHOSPHATASE_INHIBITOR
CP CP_VEGFR_INHIBITOR
CP CP_VITAMIN_D_AGONIST
KD KD_AHSP_PATHWAY
KD KD_AKT_SIGNALING
KD KD_ALDO_KETO_REDUCTASE
KD KD_BMP_SIGNALING
KD KD_DNA_REPLICATION
KD KD_EIF_PROTEINS
KD KD_G2_M_CHECKPOINT
KD KD_IL4_PATHWAY
KD KD_NON_HOMOLOGOUS_END_JOINING
KD KD_NUCLEOPORIN
KD KD_PI3K_SIGNALING
KD KD_PROTEASOME_PATHWAY
KD KD_RIBOSOMAL_40S_SUBUNIT
KD KD_RNA_POLYMERASE_ENZYMES
KD KD_TUBULIN
KD KD_VESICULAR_TRANSPORT
OE OE_CELL_CYCLE_INHIBITION
OE OE_GPCR_SUBSET
OE OE_HOX_GENE
OE OE_NFKB_ACTIVATION
Name
Type
ZSTK-474 CP
WYE-125132 CP
KU-0063794 CP
NVP-BEZ235 CP
WYE-354 CP
OSI-027 CP
GSK-1059615 CP
AZD-8055 CP
PI-103 CP
torin-1 CP
GDC-0941 CP
MTOR trt_sh.cgs
QL-X-138 CP
torin-2 CP
TGX-115 CP
calpeptin CP
wortmannin CP
LY-294002 CP
BAS-00535043 CP
temsirolimus CP
BRD-K63425657 CP
AZD-6482 CP
PP-30 CP
PP-110 CP
RPTOR trt_sh.cgs
BMS-754807 CP
MK-2206 CP
TGX-221 CP
molsidomine CP
BRD-K64835161 CP
MYC trt_sh.cgs
methotrexate CP
ALW-II-38-3 CP
PP-2 CP
VAMA-37 CP
BMS-536924 CP
NU-7026 CP
sirolimus CP
linsitinib CP
WZ-4-145 CP
paroxetine CP
crizotinib CP
deforolimus CP

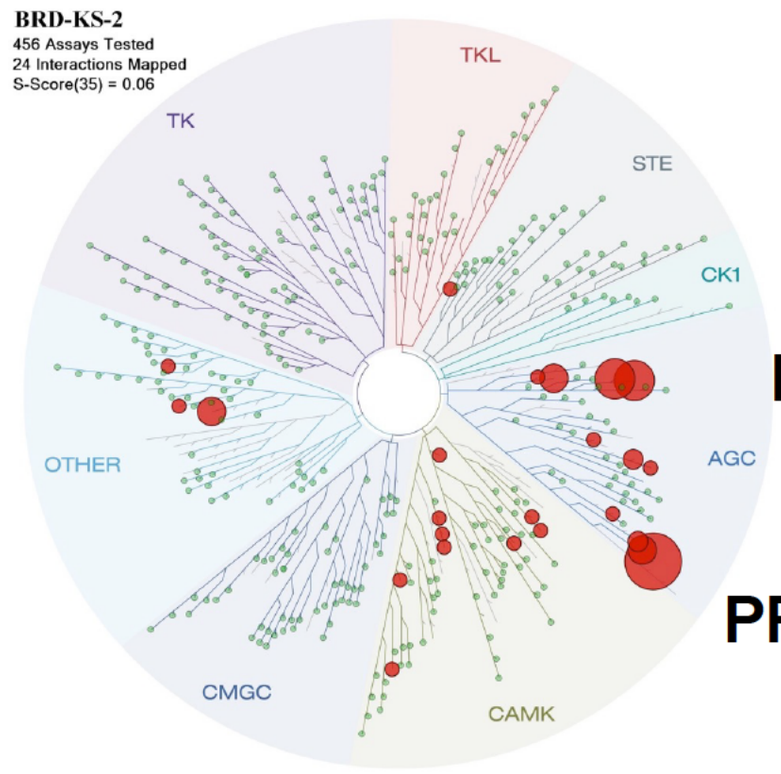

Class
Type
CP CP_ATPASE_INHIBITOR
CP CP_AURORA_KINASE_INHIBITOR_GRP1
CP CP_AURORA_KINASE_INHIBITOR_GRP2
CP CP_BCL2_AND_RELATED_PROTEIN_INHIBITOR
CP CP_BETA_ADRENERGIC_RECEPTOR_AGONIST
CP CP_BETA_ADRENERGIC_RECEPTOR_ANTAGONIST
CP CP_BRAF_RAF1_INHIBITOR
CP CP_CALCIUM_CHANNEL_INHIBITOR
CP CP_CDK_INHIBITOR_GRP1
CP CP_CDK_INHIBITOR_GRP2
CP CP_CYCLOOXYGENASE_INHIBITOR
CP CP_DNA-DEPENDENT_PROTEIN_KINASE
CP CP_DOPAMINE_RECEPTOR_GRP1
CP CP_DOPAMINE_RECEPTOR_GRP2
CP CP_EGFR_INHIBITOR
CP CP_ESTROGEN_AGONIST
CP CP_ESTROGEN_ANTAGONIST
CP CP_FAAH_INHIBITOR
CP CP_FXR_ANTAGONIST
CP CP_GLUCOCORTICOID_AGONIST
CP CP_GSK3_INHIBITOR
CP CP_HDAC_INHIBITOR
CP CP_HMGCR_INHIBITOR
CP CP_HSP90_INHIBITOR
CP CP_IGF1R_INHIBITOR
CP CP_IMP_DEHYDROGENASE_INHIBITOR
CP CP_JAK_INHIBITOR
CP CP_JNK_INHIBITOR
CP CP_KIT_INHIBITOR
CP CP_MDM2_MDM4_INHIBITOR_GRP1
CP CP_MDM2_MDM4_INHIBITOR_GRP2
CP CP_MEK_INHIBITOR
CP CP_MINERALOCORTICOID_AGONIST
CP CP_MTOR_INHIBITOR
CP CP_NA-K-CL_TRANSPORTER_INHIBITOR
CP CP_NAMPT_INHIBITOR
CP CP_NFKB_PATHWAY_INHIBITOR
CP CP_OPIOID_RECEPTOR_AGONIST
CP CP_P38_MAPK_INHIBITOR_GRP1
CP CP_PARP_INHIBITOR
CP CP_PHOSPHODIESTERASE_INHIBITOR
CP CP_PI3K_INHIBITOR
CP CP_PKA_INHIBITOR
CP CP_PKC_ACTIVATOR
CP CP_PKC_INHIBITOR
CP CP_PLK_INHIBITOR
CP CP_PPARG_AGONIST
CP CP_PROGESTERONE_AGONIST
CP CP_PROTEASOME_INHIBITOR
CP CP_PROTEIN_SYNTHESIS_INHIBITOR
CP CP_RAR_AGONIST_GRP1
CP CP_RAR_AGONIST_GRP2
CP CP_RIBONUCLEOTIDE_REDUCTASE_INHIBITOR
CP CP_ROCK_INHIBITOR
CP CP_SEROTONIN_RECEPTOR_ANTAGONIST
CP CP_SIRT_PATHWAY
CP CP_SRC_INHIBITOR
CP CP_SULFONYLUREA
CP CP_TACHYKININ_ANTAGONIST
CP CP_TCA
CP CP_TGF_BETA_INHIBITOR
CP CP_TOPOISOMERASE_II_INHIBITOR
CP CP_TOPOISOMERASE_I_INHIBITOR
CP CP_TRPV_AGONIST
CP CP_TUBULIN_INHIBITOR
CP CP_TYROSINE_PHOSPHATASE_INHIBITOR
CP CP_VEGFR_INHIBITOR
CP CP_VITAMIN_D_AGONIST
KD KD_AHSP_PATHWAY
KD KD_AKT_SIGNALING
KD KD_ALDO_KETO_REDUCTASE
KD KD_BMP_SIGNALING
KD KD_DNA_REPLICATION
KD KD_EIF_PROTEINS
KD KD_G2_M_CHECKPOINT
KD KD_IL4_PATHWAY
KD KD_NON_HOMOLOGOUS_END_JOINING
KD KD_NUCLEOPORIN
KD KD_PI3K_SIGNALING
KD KD_PROTEASOME_PATHWAY
KD KD_RIBOSOMAL_40S_SUBUNIT
KD KD_RNA_POLYMERASE_ENZYMES
KD KD_TUBULIN
KD KD_VESICULAR_TRANSPORT
OE OE_CELL_CYCLE_INHIBITION
OE OE_GPCR_SUBSET
OE OE_HOX_GENE
OE OE_NFKB_ACTIVATION
Name
Type
ZSTK-474 CP
WYE-125132 CP
KU-0063794 CP
NVP-BEZ235 CP
WYE-354 CP
OSI-027 CP
GSK-1059615 CP
AZD-8055 CP
PI-103 CP
torin-1 CP
GDC-0941 CP
PIK-90 CP
MTOR trt_sh.cgs
PI-828 CP
QL-X-138 CP
torin-2 CP
TGX-115 CP
calpeptin CP
wortmannin CP
BAS-00535043 CP
BRD-K63425657 CP
PP-30 CP
PP-110 CP
JW-7-24-1 CP
RPTOR trt_sh.cgs
epoxycholesterol CP
BMS-754807 CP
MK-2206 CP
TGX-221 CP
BRD-K64835161 CP
methotrexate CP
ALW-II-38-3 CP
PP-2 CP
VAMA-37 CP
BMS-536924 CP
NU-7026 CP
sirolimus CP
linsitinib CP
WZ-4-145 CP
paroxetine CP
crizotinib CP
deforolimus CP

BRD-2751

Confirmed ROCK1/2 inhibitor

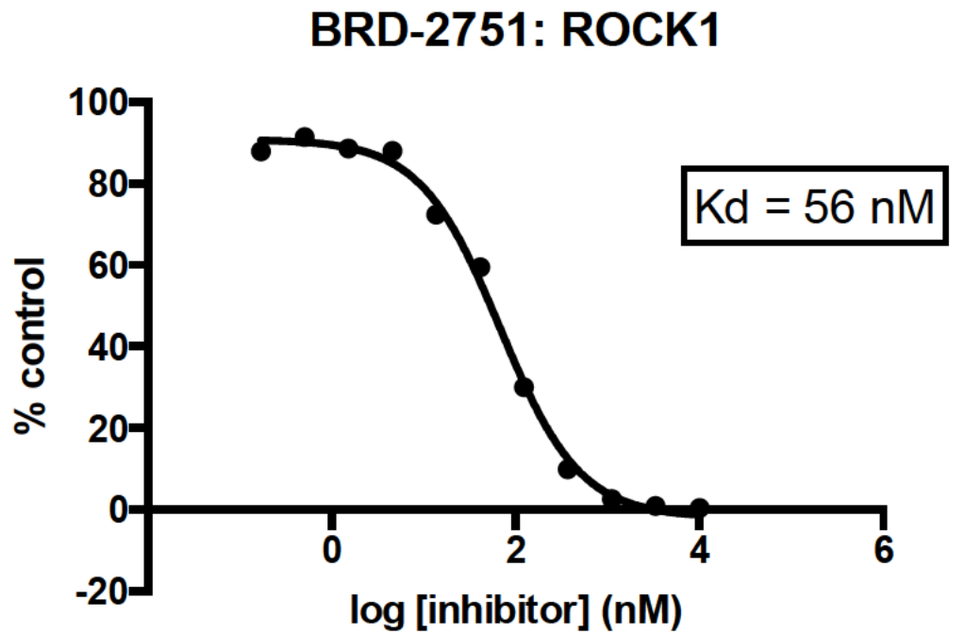






ROCK1/2

PRKCE

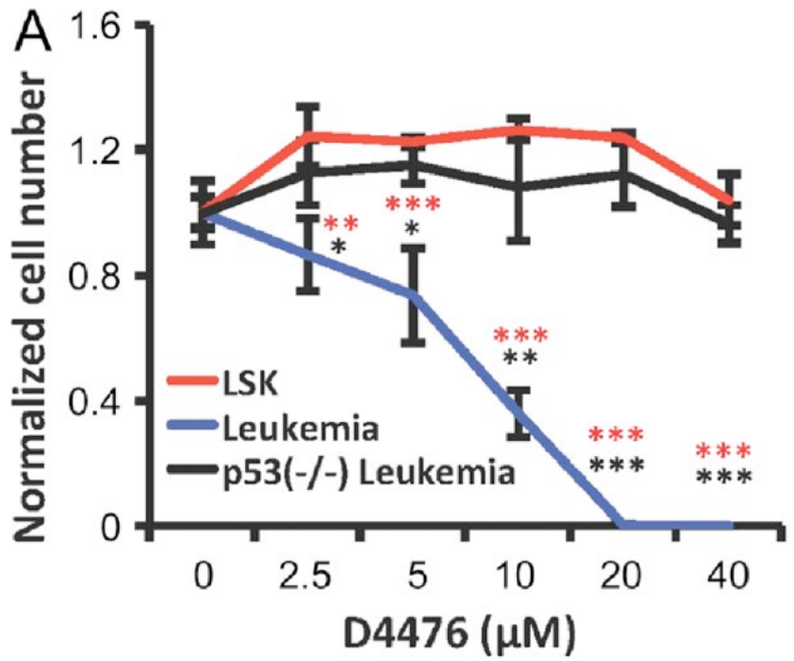


CSNK1A1 disease relevance

Investigated for pathogenic role in MDS and AML

Published March 10, 2014

JEM



Csnk1a1 inhibition has p53-dependent therapeutic efficacy in acute myeloid leukemia

Marcus Järås,[1,2] Peter G. Miller,[1] Lisa P. Chu,[1] Rishi V. Puram,[1] Emma C. Fink,[1] Rebekka K. Schneider,[1] Fatima Al-Shahrour,[1] Pablo Peña,[2] L. Jordan Breyfogle,[1] Kimberly A. Hartwell,[1,3] Marie E. McConkey,[1] Glenn S. Cowley,[3] David E. Root,[3] Michael G. Kharas,[4,5] Ann Mullally,[1] and Benjamin L. Ebert[1,3]

CSNK1A1 inhibitor: unmet clinical need

Predicted CMap inhibitors

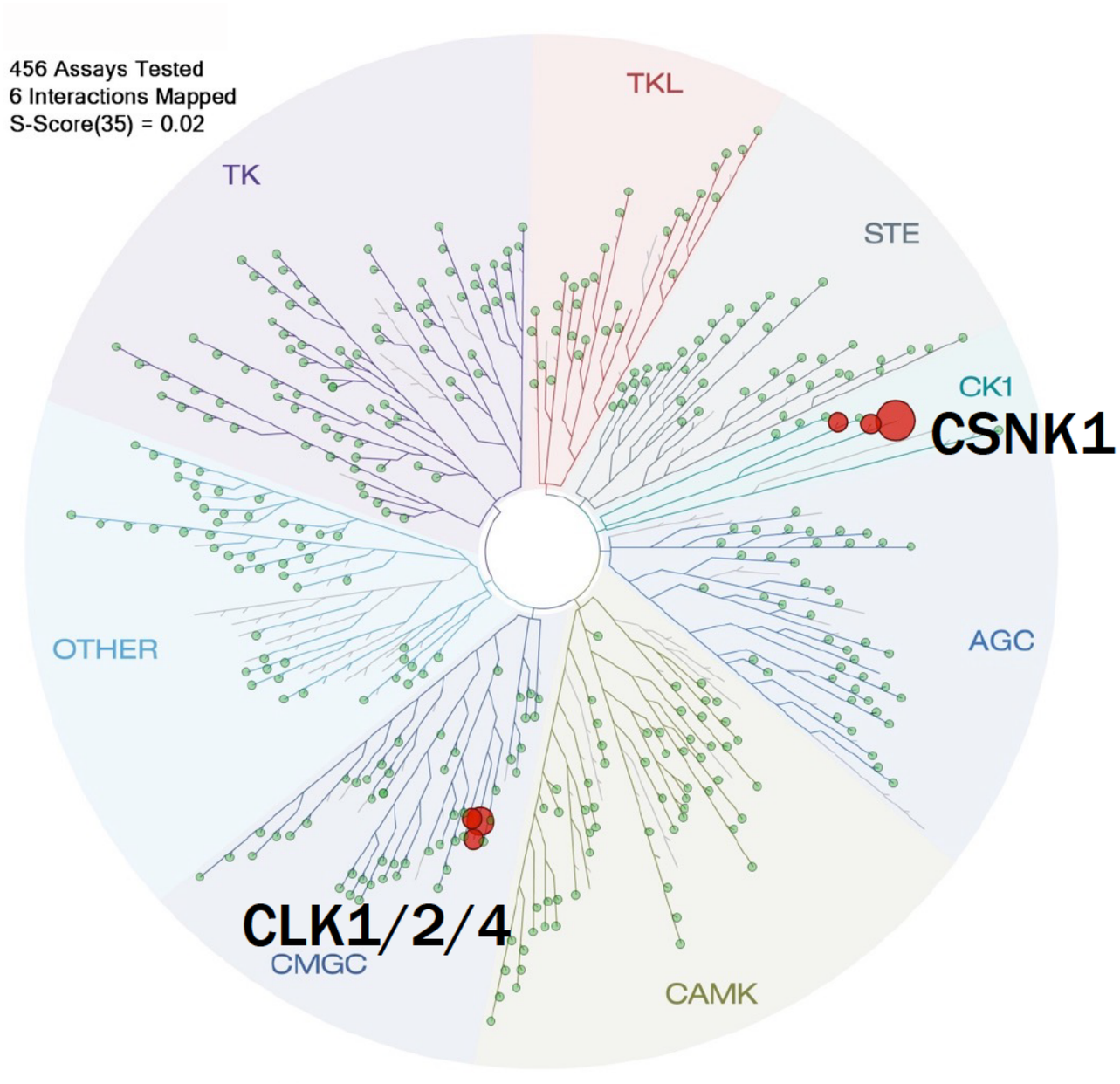
rank	compound
1	BRD-1868
2	BRD-8715
3	azithromycin
4	BRD-4657
5	BRD-1433
6	BRD-1742
7	phenazopyridine
8	caffeine
9	tivozanib
10	BRD-0220

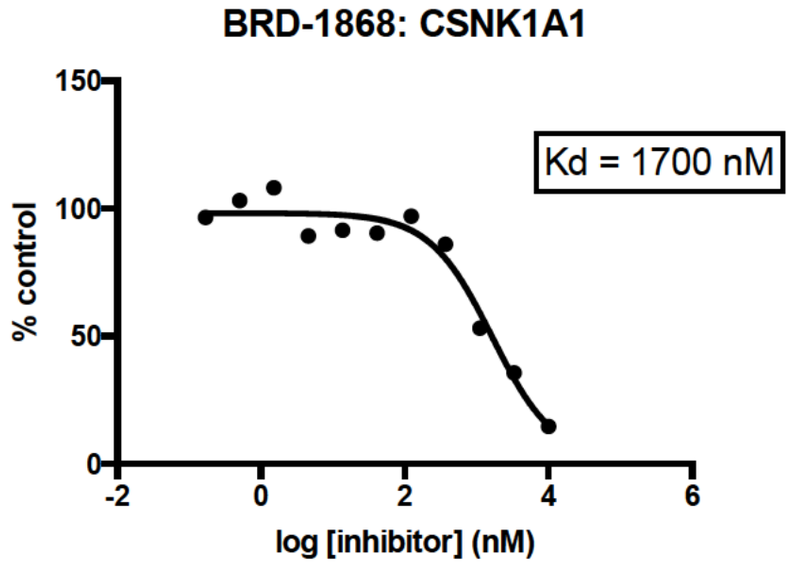
rank	compound
1	BRD-1868
2	BRD-4657
3	ENMD-2076
4	LY-456236
5	salmeterol
6	BRD-4053
7	RS-102895
8	BMS-754807
9	MBCQ
10	DM-55-3

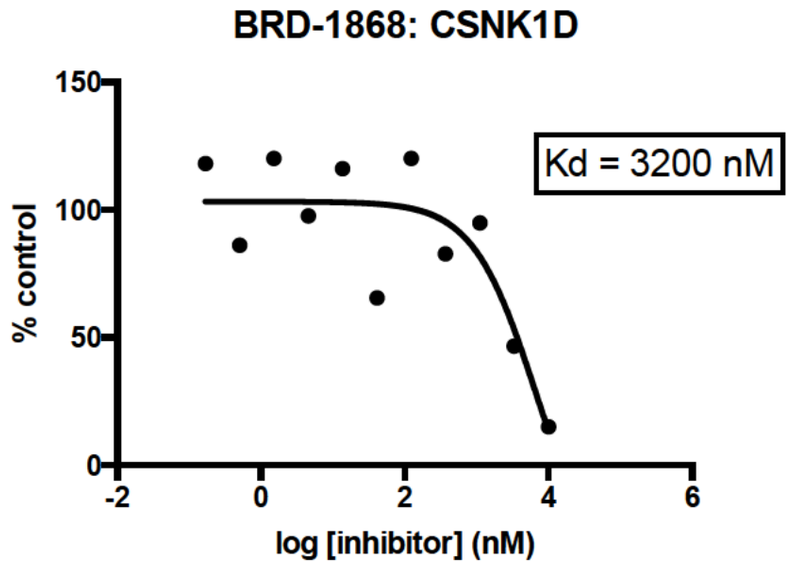
Query	CSNK1A1 shRNA	All compounds
Result	Compound rank	CSNK1A1 rank

Discovery of novel CSNK1A1 inhibitor

CMap prediction confirmed

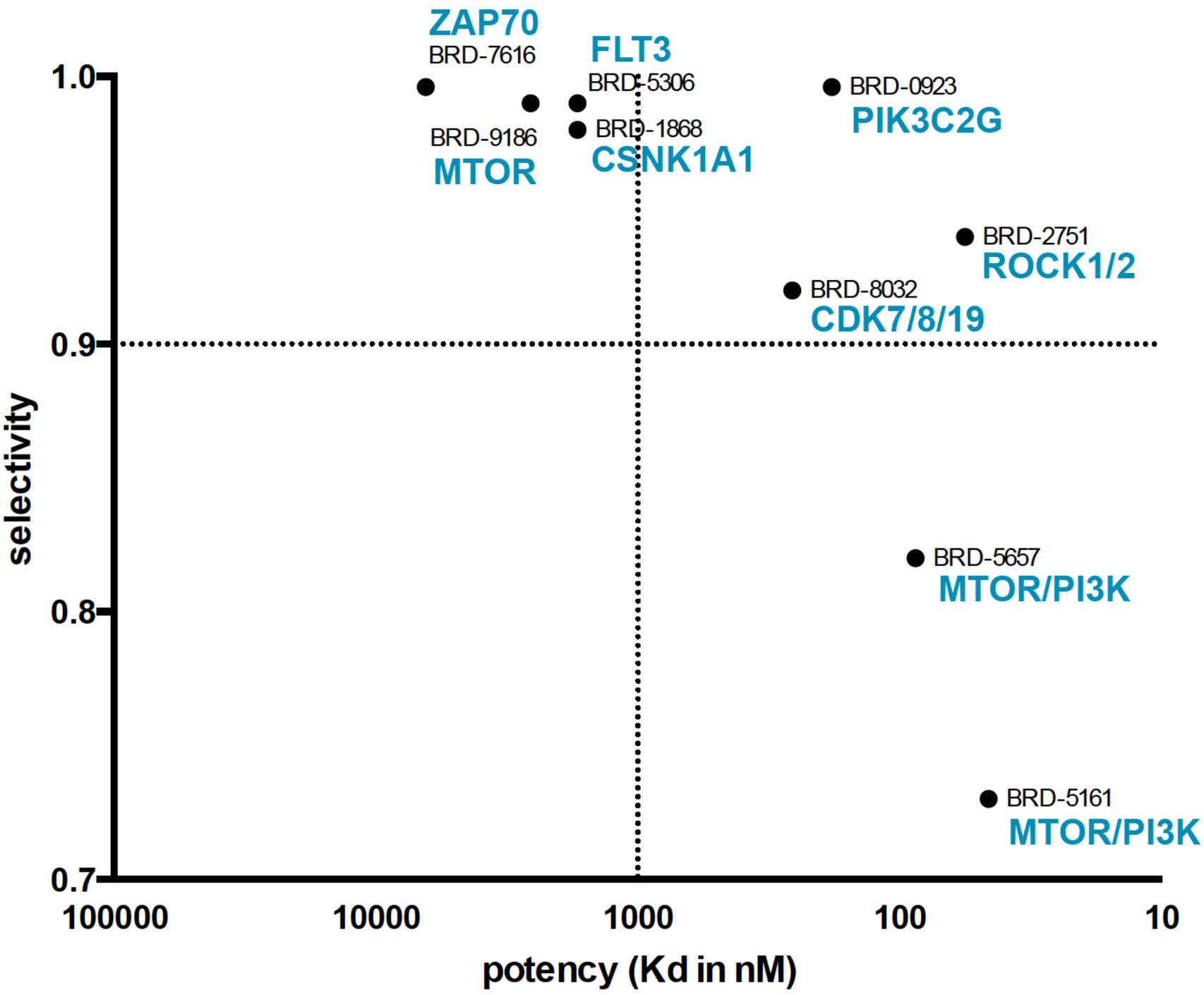






Kinase inhibitor hits

Potency vs selectivity



GEO signatures
Pharmacologic classes
Genetic classes

Growth factors
TGF-alpha
EGF
Heregulin
IL4R
KIT
RTK
KDR
EGFR
GRB2
SOS
NF1
RAS
RAF
MEK
ERK
FOS
CDC42
RAC
RHO
GPCR ligand
7-TMR
G-PROT
ADCY
PKA
CREB
TACR1
ADRB1
DRD2
DPR1
GPCRs
Hormones
Estrogen
Androgens
TNFA
TNFR
Survival factors
IGF1
RTK
IGFR
RTKs
Cytokine receptor
Cytokines
IFN-gamma
JAKs
STATs
PIAS
NIK
IKB
IKK
NFKB
PSMB1
PI3K
AKT
AKKA
IMPDH
MTOR
S6K
JDR
NR3C1
ESR1
RAR
NR3C2
FYN
SHC
FAK
SRC
PLC
PKC
CDC42
PI3K
RAC
CAS
CRK
MOS
MKK
DUSP4
JNK
JUN
Integrins
ECM
WNT
Frizzled
Disheveled
GSK3B
APC
B-Catenin
Cells
E-Cadherin
Vesicular transporter
HMGCS
HMGCR
cholesterol
Anti-growth factors
TGFB
TGFBR
BMPR
SMADs
P16
CyclD:CDK4
RB
E2F
P15
P21
P27
P53
MDM2
ARF
CyclE:CDK2
CHEK
Cell Proliferation (Cell Cycle)
Changes in Gene Expression
SERCA
Protein synthesis
AURK
PLR
PPARG
TOP
HDACs
HSP90
HOX
EIF
Nucleoporin
TUBB
PARP
RRM
DNA replication
Na+-K+ ATP ase
Na+-K+-Cl2- transporter
K+ inward rectifying channel
Ion channels
ROCK
MLCP
COX
FAAH
TRPV1
SIRT
NAMPT

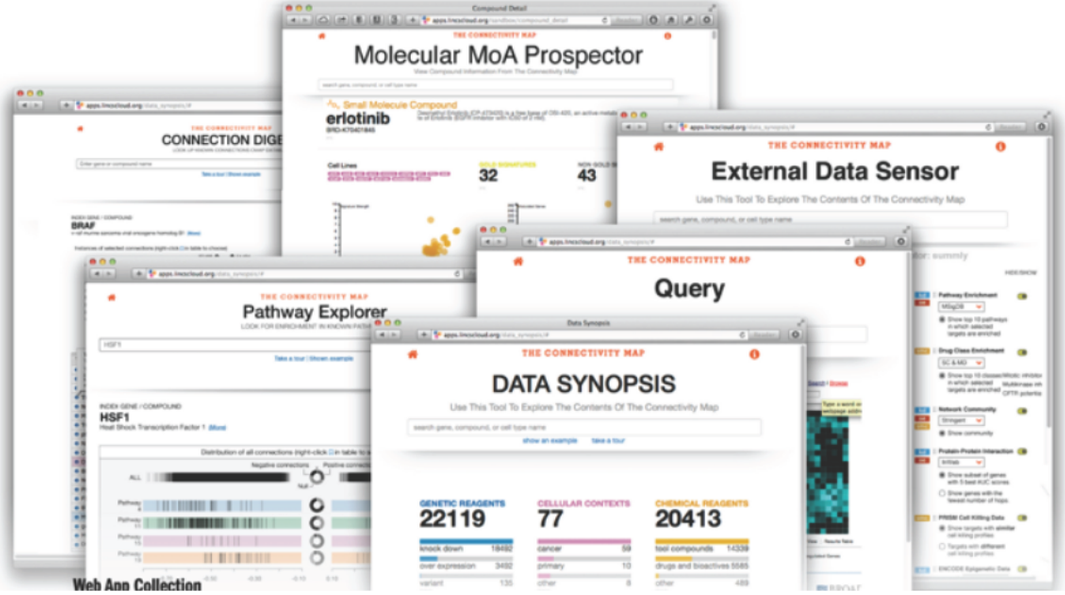
METHODS & RESOURCES

BROAD INSTITUTE

Assay: L1000

- Measure 978 transcripts with 500-color Luminex beads
- 384-well plate format compatible with HTS
- Cell lysates (not purified RNA)
- Dramatic cost reduction

Algorithms and tools (under development)



http://lincsproject.org
http://lincs-dcic.org/
http://lincscloud.org

Why repurpose existing drugs?

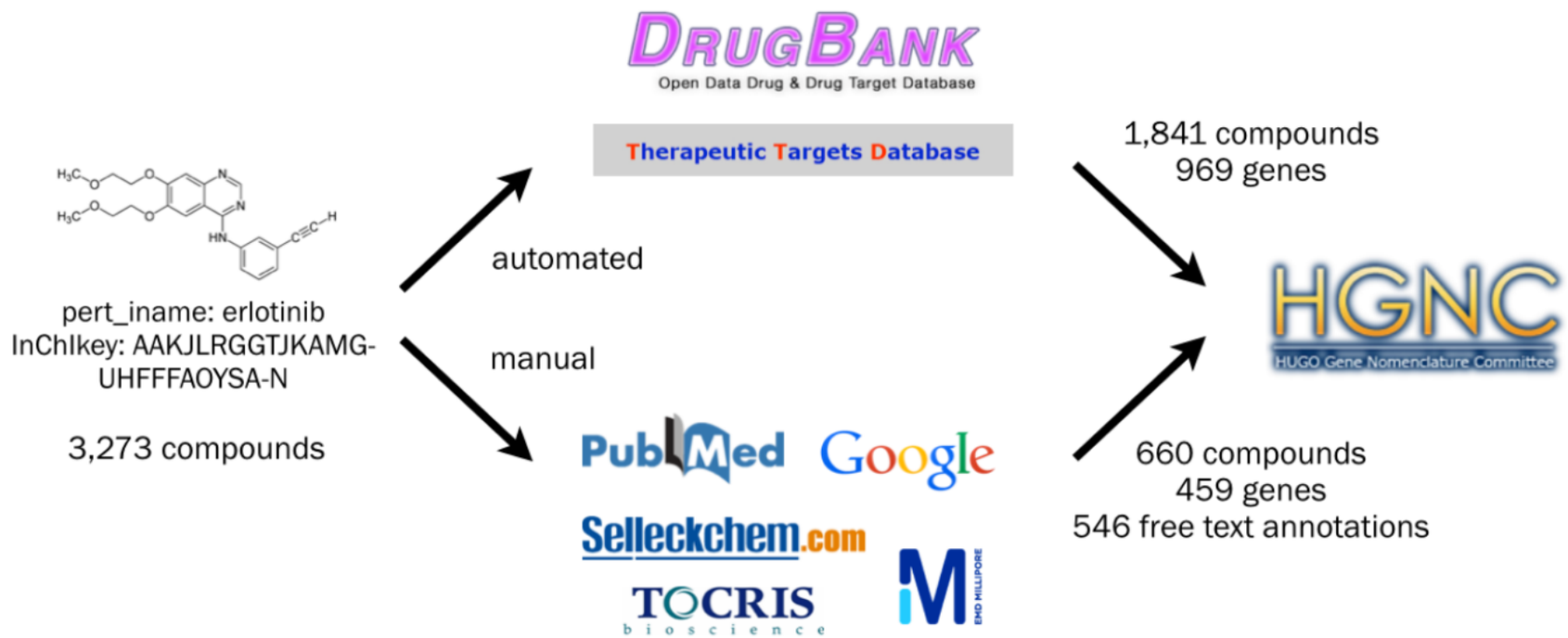
Motivation

- Bioactive compounds have long been known to act on multiple targets, potentially enabling additional uses and explaining toxicities
- Developing a new drug takes up to 10-15 years and several hundred million dollars
- Thousands of “orphan” or rare diseases exist where the market size alone may not justify a dedicated discovery program

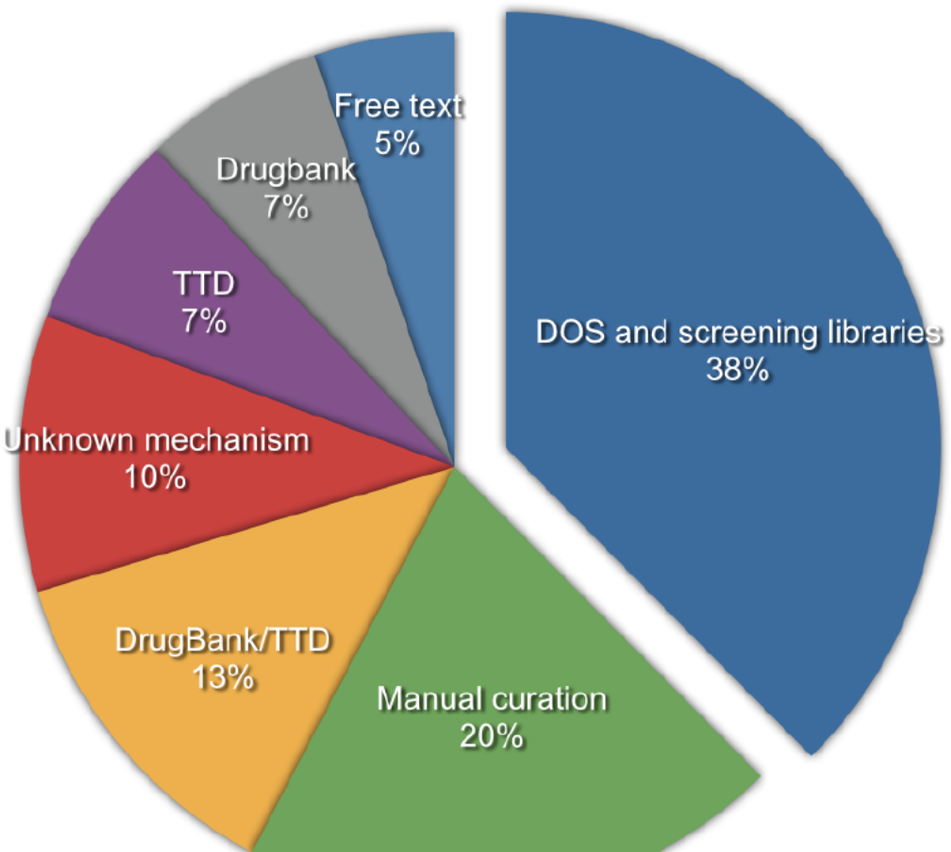
Compound annotation

Integrating high-quality source information

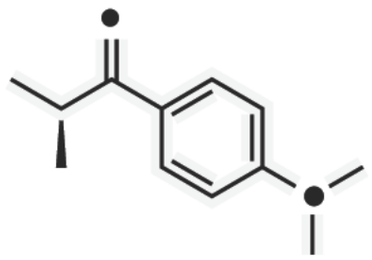
Curation process



Current Database



Current CMap dataset (2M profiles)



~5,000 Compounds

- FDA approved drugs
- Bioactive tool compounds
- Screening hits

~3,000 Genes (shRNA & cDNA)

- Targets/pathways of approved drugs
- Candidate disease genes (e.g. recurrent from TCGA)
- Community nominations

15 Cell types

- Banked primary cell types
- Cancer cell lines (8)
- Primary hTERT-immortalized
- Patient-derived iPS cells

Design goals of CLUE

Cloud-based

No need to download data (tools to data)
Elastic compute (fast, scale-on-demand)

Harmonized

Centralized annotations and meta data
Interoperable with other perturbational datasets (e.g PRISM, P100)

Reproducible

Prepackaged Box with - all of the data + pre-processing pipeline + analsis tools

Customizable

Detach'able from Broad infrastructure
Companies can securely mix proprietary datasets with public data

Access at: clue.io
(early access phase till July 2015; full version in August)

Looking forward

resources coming soon …

- Genome-scale perturbations
- CRISPR/Cas9 knock-outs
- Small-molecule analogs and libraries

- Multiple readouts at scale

- Diverse cell types
- Patient-derived iPS cells

- Next generation biologist-focused tools

- **Data integration (GTEx, GEO, Tox21)**

Team @Broad Institute

Jacob Asiedu
Josh Bittker
Steven Corsello
Desiree Davidson
John Davis
John Doench
Oana Enache
Corey Flynn
Joshua Gould
Jodi Hirschman
David Lahr

Daniel Lam
Xiaodong Lu
Federica Piccioni
Rajiv Narayan
Ted Natoli
David Peck
Jackie Rosains
Ian Smith
Bang Wong

David Root
Stuart Schreiber
Lucienne Ronco

Todd Golub

BROAD INSTITUTE

Questions?

aravind @ broadinstitute.org

BROAD INSTITUTE