NEU Proteomics Capstone Module:
Case Studies in Quantitative Proteomics

Jacob D. Jaffe, Lev Litichevski, Jarrett Egerston,
Mike MacCoss, Meena Choi, Olga Vitek
w/ special guest: Ruedi Aebersold

Goals for this course

- Understand the concepts and value of a proteomic sentinel assay
- Understand the considerations in moving from discovery proteomics data to targeted proteomics
  - Practical: Selection of probes for a targeted proteomics sentinel assay
- **Understand considerations for treatment and normalization of research grade targeted proteomics assays**
  - Practical: Construct your own data processing pipeline
- Introduction to next-gen MS methods to expand targeted panels
- Understand the conceptual differences between using markers and profiles of activity
- Introduction to the concept of connectivity and exploration of the dataset in this space
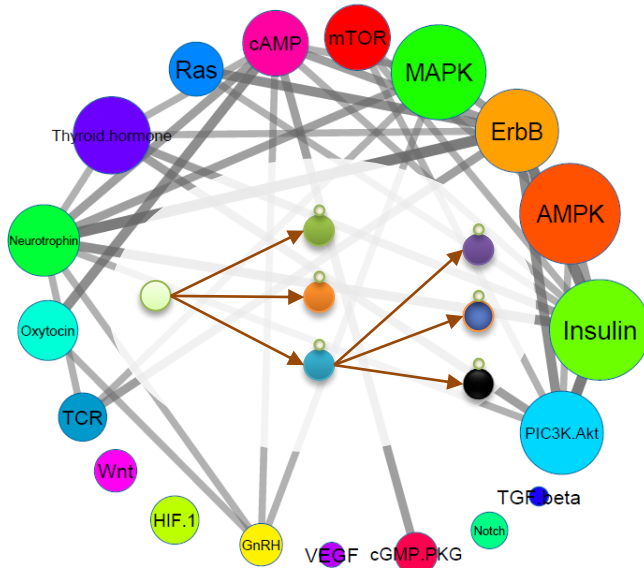
Lecture 2: The Targeted P100 Assay in Practice and Considerations for Reproducible Research

Jacob D. Jaffe

The P100 is a "Research Grade" Targeted PRM Assay

P100 Sentinel Assay; 96 phospho probes commonly observed and modulated in diverse cell types
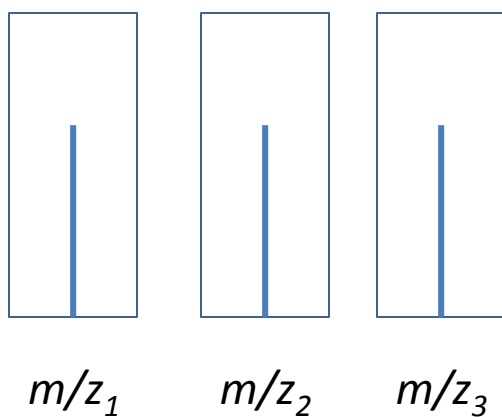
- 96 phosphopeptides monitored in 60 minute LCMS method

- Synthetic peptide internal standards with heavy amino acids

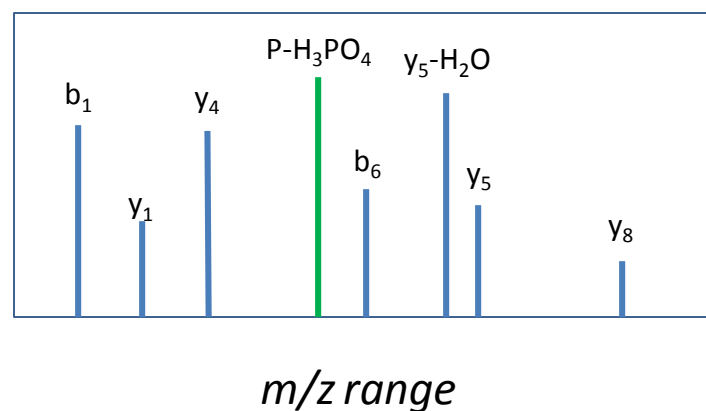- Can be acquired as scheduled PRM or unscheduled DIA

Definition of PRM

- PRM = MRM-HR = HR-MRM = Targeted Full Scan MS/MS

- Closest spiritual cousin is triple-quad based MRM/SRM, but:

SRM
Q3 mass analyzer

PRM
mass analyzer

m/z_1 m/z_2 m/z_3

m/z range

Discrete transitions,
hardware selected

Full scan

Definition, continued

- Assay is completely deterministic

- Precursor m/z (list) is specified
  - Possibly scheduled
  - Quadrupole or ion trap selection/isolation

- Fragmentation is performed
  - Any kind is OK

- Full MS/MS spectrum is recorded
  - Any analyzer: Orbitrap, TOF, scanning quad, ion trap, etc.

- Usually a full scan MS spectrum is also periodically recorded
 - Two chances to verify and quantify!

When do PRMs make sense?

- Exquisite selectivity required
  - Unit (quadrupole) vs. ppm (hi-res)

- Post-translational modification localization is required
  - GVDQ(pS)PLTPAGGK vs. GVDQSPL(pT)PAGGK

- Rapidly convert discovery data to targeted assay
  - Stay within platform

- You don't have a triple quad!
  - But still want the benefits of targeted proteomics

PTM Localization – shared ions, differential ions
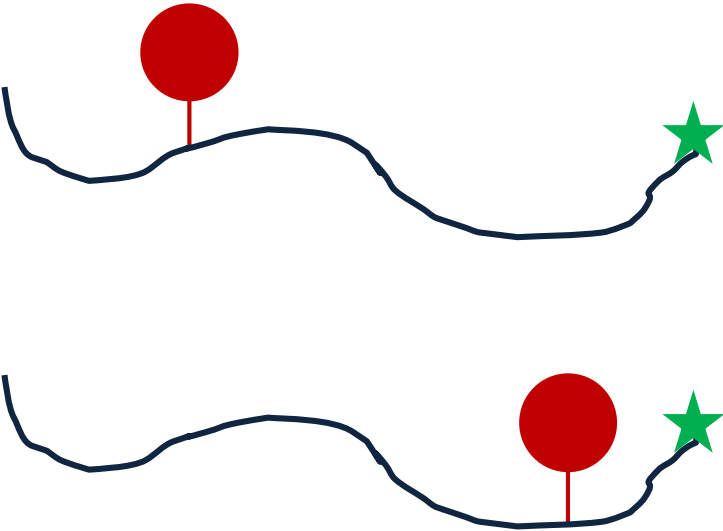
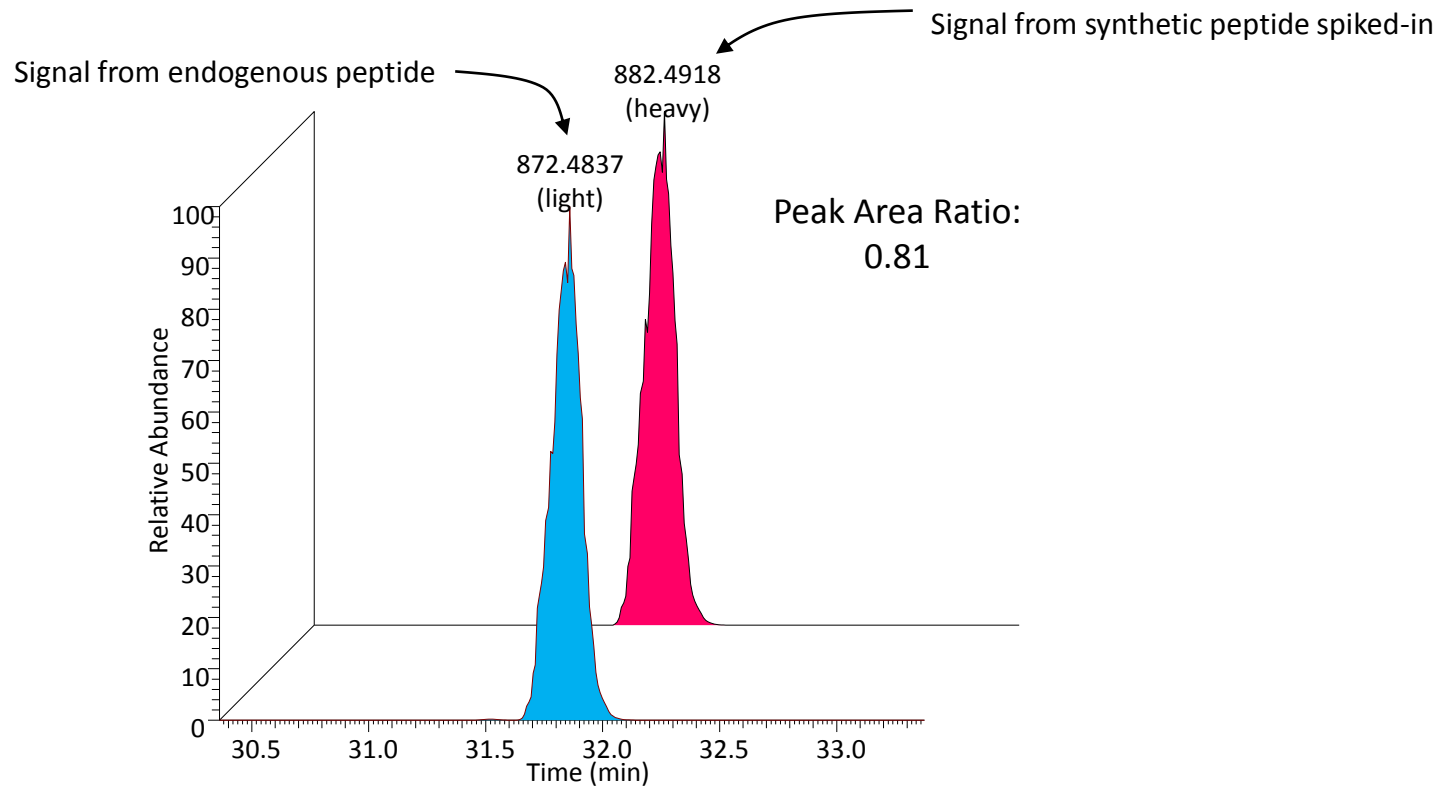Differentially Phosphorylated Peptides
With Same Base Sequence
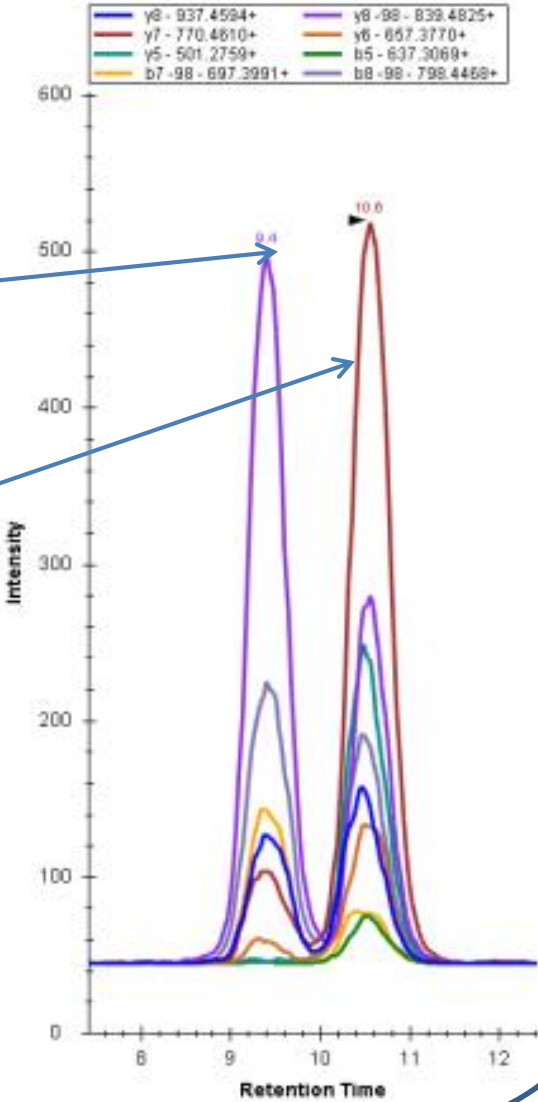
"Research Grade" PRM Concept

- A quantitative, targeted proteomics assay suitable for "everyday" use

- Ideally standardized with synthetic peptides (or SILAC)

- Rapid design cycle using discovery data/platform

- Enables longitudinal comparisons across days, months, years

- Output useful for rapidly guiding biology

- NOT:
  - Obsessed with LOD/LOQ
  - Suitable for clinical deployment

Peak area ratio: the atomic unit of P100 quantitative data

Signal from synthetic peptide spiked-in

Signal from endogenous peptide

882.4918
(heavy)

872.4837
(light)

Peak Area Ratio:
0.81

Relative Abundance

100
90
80
70
60
50
40
30
20
10
0

30.5 31.0 31.5 32.0 32.5 33.0

Time (min)

- Output is a **ratio** of an <u>endogenous</u> peptide to a <u>synthetic internal</u> standard peptide (heavier due to 15N, 13C, etc.)
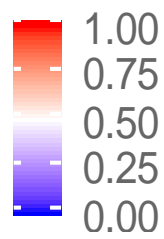 - *<u>Relative</u> abundance to standard of known concentration*

P100 Sample Prep: Automation allows for scalability

Quality Control Steps
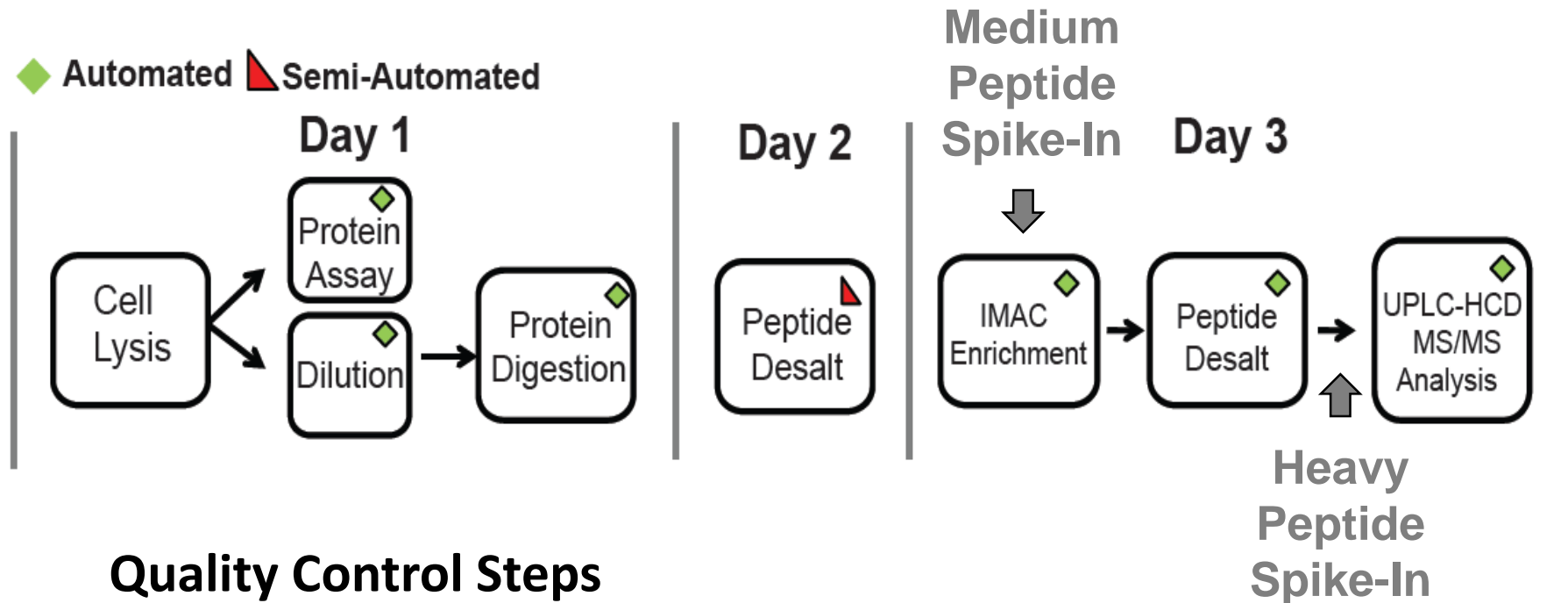
- Add standards to monitor enrichment
  - <u>recovery</u>

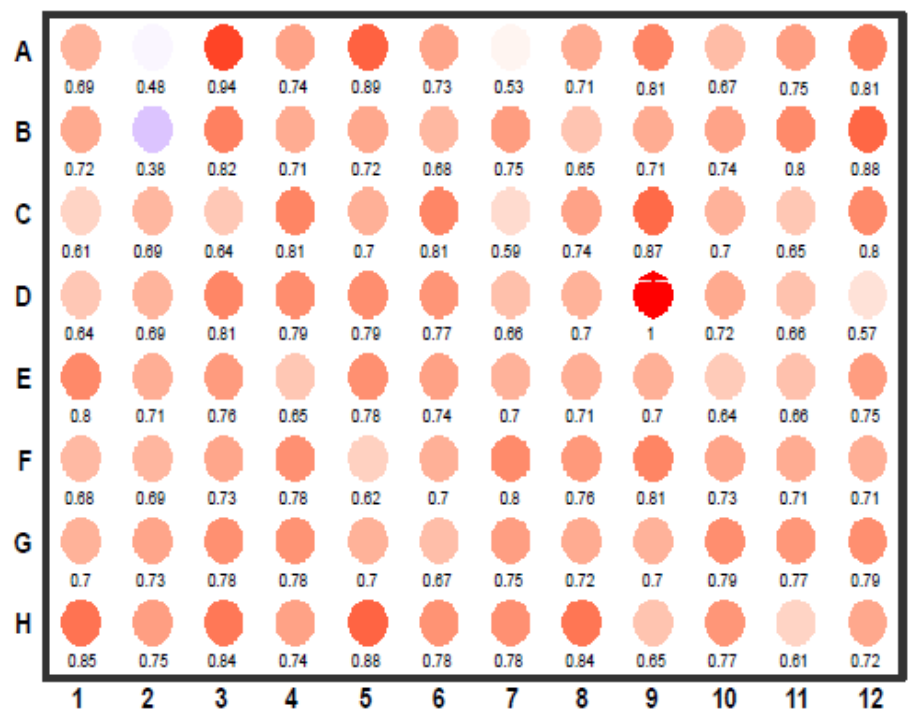# QC is important: Phosphopeptide Enrichment Recovery

% Recovery

Example of a 96-Well Plate Map of Phosphopeptide Enrichment Efficiency
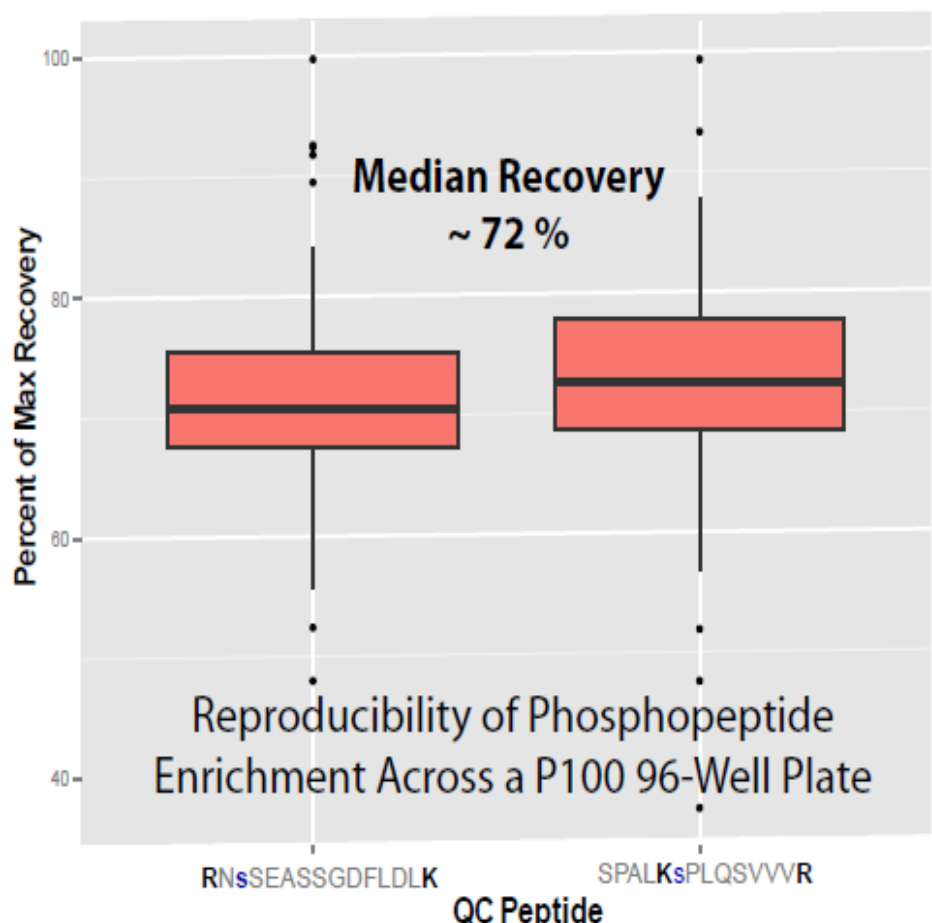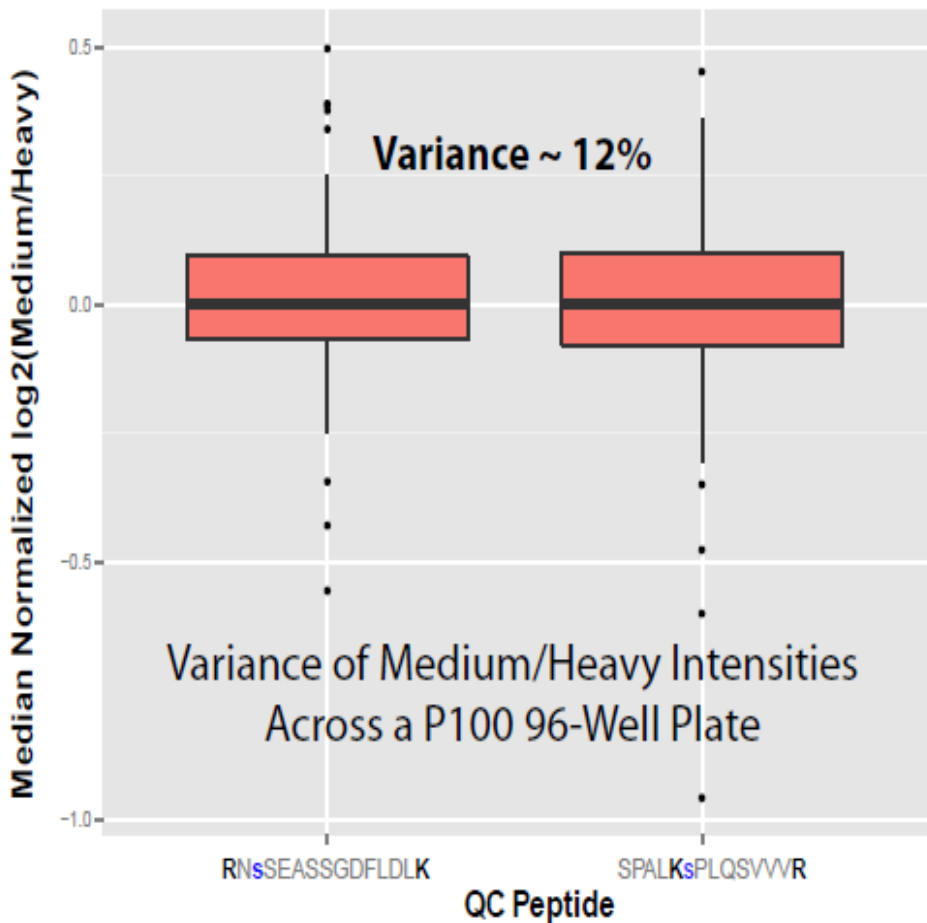
Automated Phosphopeptide Enrichment Characteristics

P100 Assay: 96-Well Plate Automation Evaluation

Revisiting the final goals of P100 assay development

? Drug responsiveness

? Universal (-ish)

✓ Ser/Thr focused – IMAC enrichment

✓ Compact – 60 minutes analytical time

✓ Operates at scale – Automated sample prep

✓ Longitudinally comparable – Synthetic standards with defined levels

Where we last left off – Will it work?

✓ Are there a set of phosphosites that we can (almost) always observe?

✓ Do these phosphosites behave in interesting ways in response to perturbations?

✓ Do we gain anything over gene expression measurements?

✓ Can we pick a (small) set of highly informative phosphopeptides?

▪ **Could profiling these peptides reproduce the observations of a deeper experiment?**

▪ **Is the assay generally useful?**

BROAD
INSTITUTE

Creating test datasets…

1. 3 cell lines x 16 compounds x 3 replicates
   - MCF7, PC3, HL-60
2. 3 cell lines x 6 doses of a single compound x 2 replicates
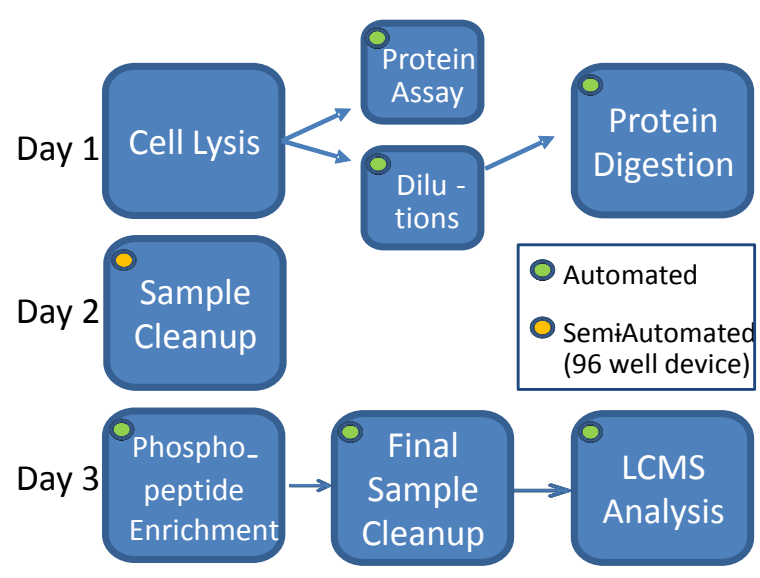3. ES cells – 12 compounds x 3 replicates, NPCs 6 cmpds x 3 reps

Cell growth

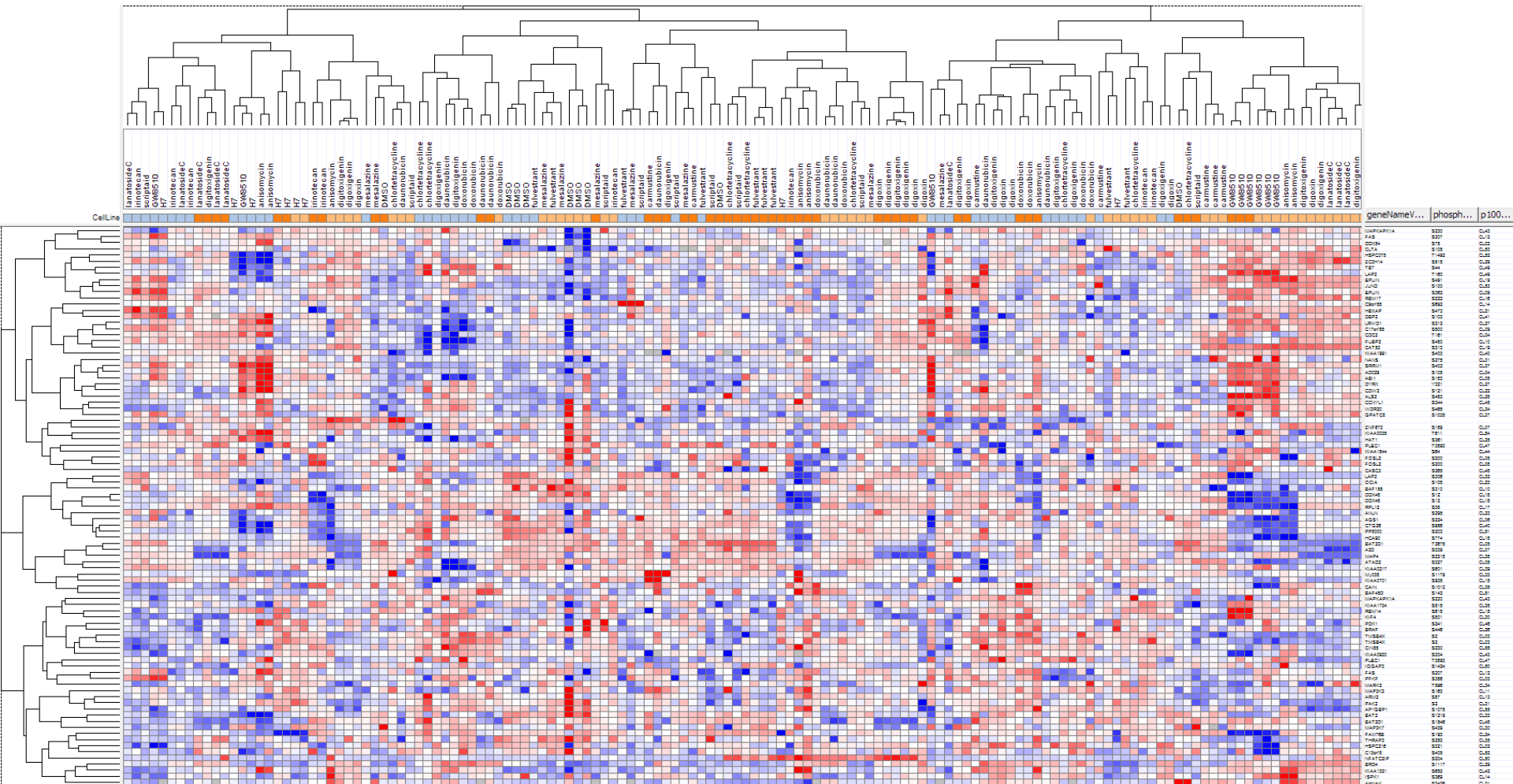| Timing Estimates	MCF7	HL-60
	PC3	

Week 1   Week 4

Lysis/Digestion

Timing Estimates	MCF7-PC3	HL-60/ #2

Week 5   Week 6

Data Acquisition

Timing Estimates	MCF7-PC3	HL-60/ #2

Week 5-6   Week 6-7

Day 1 Cell Lysis → Protein Assay
Cell Lysis → Dilu - tions → Protein Digestion

Day 2 Sample Cleanup

Day 3 Phospho- peptide Enrichment → Final Sample Cleanup → LCMS Analysis

- ● Automated
- ● Semi-Automated (96 well device)

Validation Data Set Overview

- Original: 144 samples x 97 probes
- Post filtering: 141 samples x 92 probes

Z-score
-3   3

Cell line
HL60
MCF7
PC3

Local structure reproduces deeper profiling
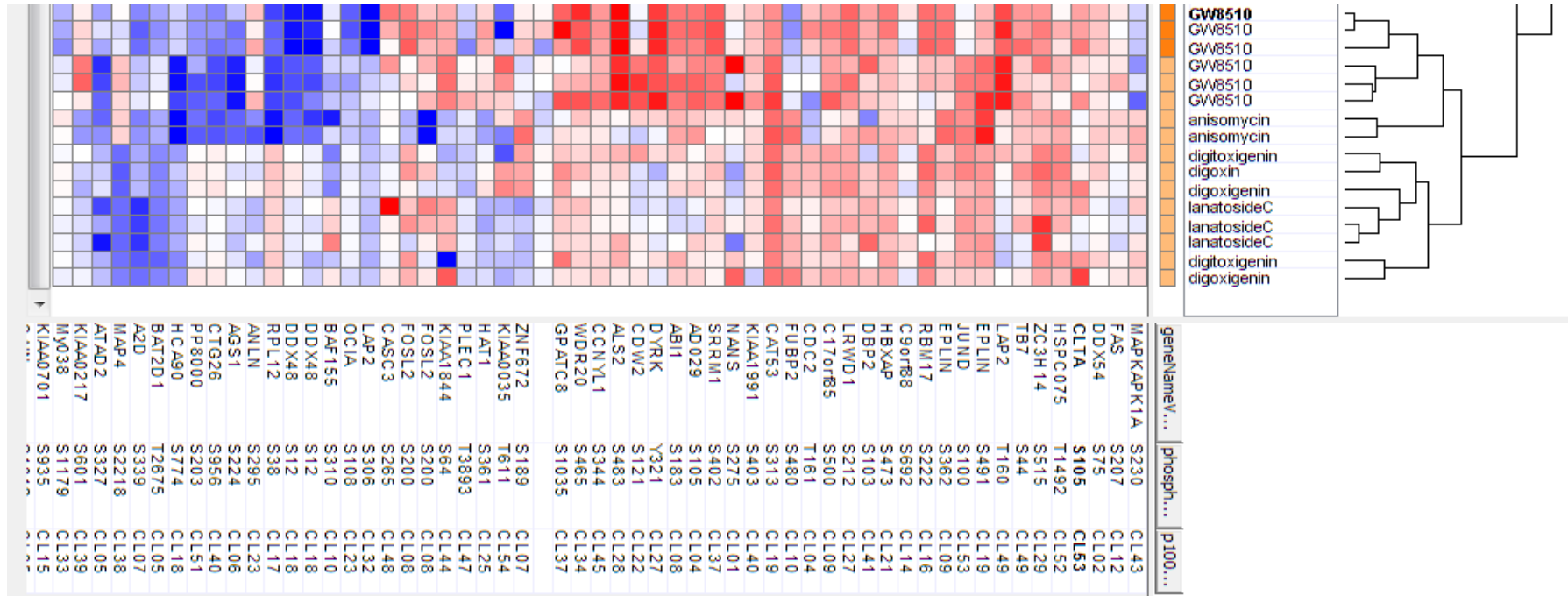
- Lineage independent signatures
  - Lineage-dependent components
- Replicates co-cluster
- Structural analogs co-cluster
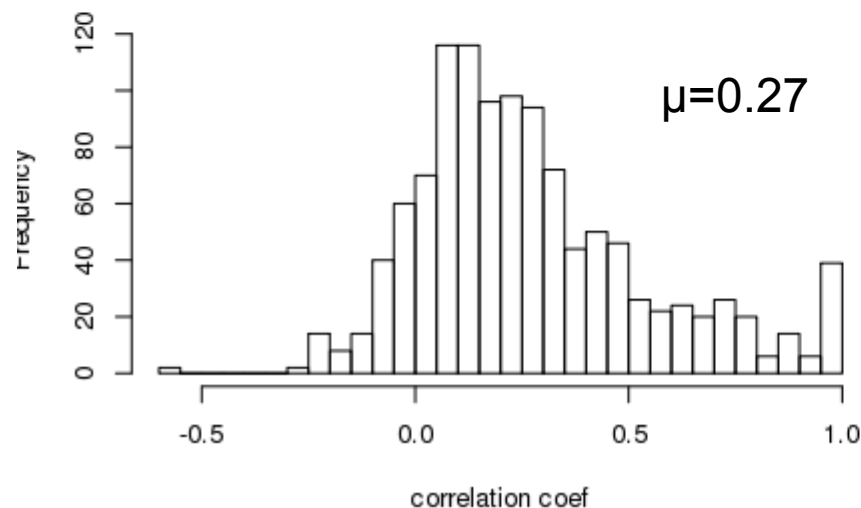- Results mirror those of deeper profiling

Summary Statistics Demonstrate Assay Performance

- Assessment of global assay performance for making connections
- Using all probes (and only measured probes, not imputed sites)
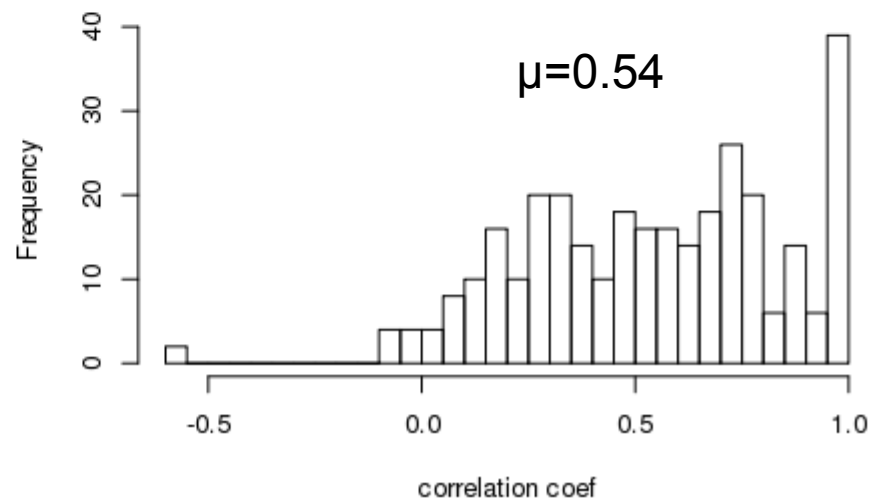- No attempt to optimize signatures for treatments

All Non-self Pairwise Correlations

$\mu=0.02$
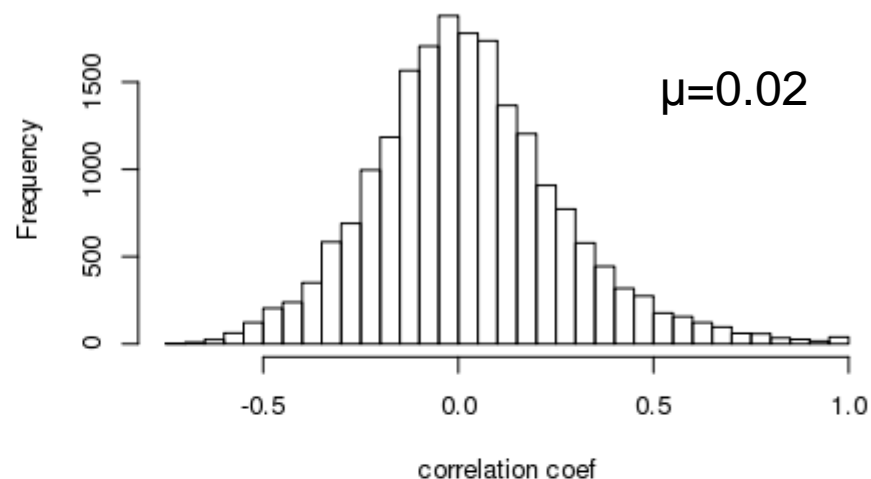
All Non-self Same Cmpd Correlations

$\mu=0.27$

All Non-self Same Line/Cmpd Correlations
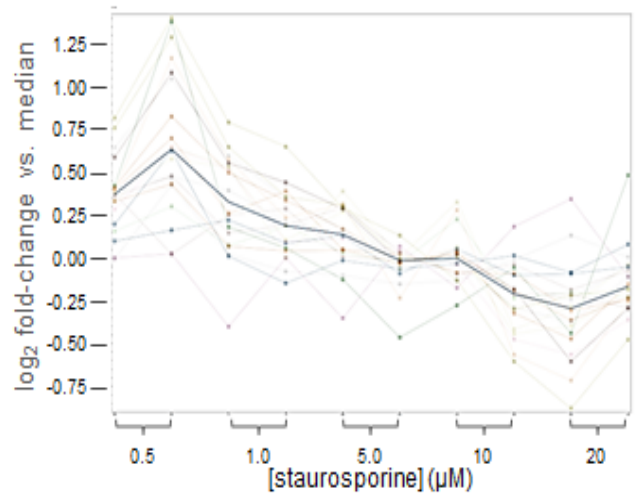
$\mu=0.54$

Will it work?

- ✓ **Are there a set of phosphosites that we can (almost) always observe?**

- ✓ **Do these phosphosites behave in interesting ways in response to perturbations?**

- ✓ **Do we gain anything over gene expression measurements?**

- ✓ **Can we pick a (small) set of highly informative phosphopeptides?**

- ✓ **Could profiling these peptides reproduce the observations of a deeper experiment?**

- ▪ **Is the assay generally useful?**
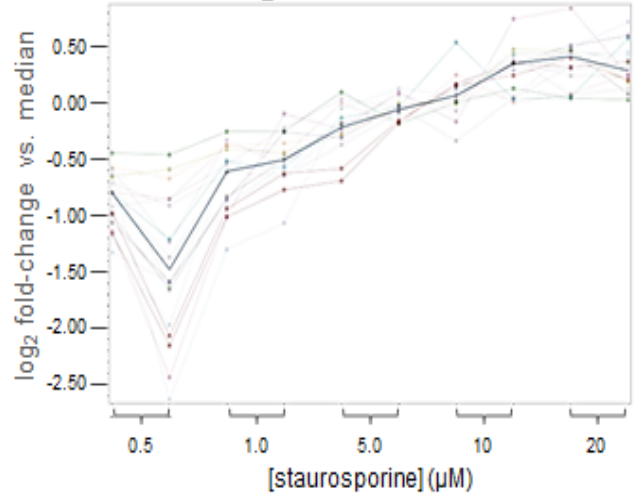
Dose-responsive signatures to kinase inhibitors
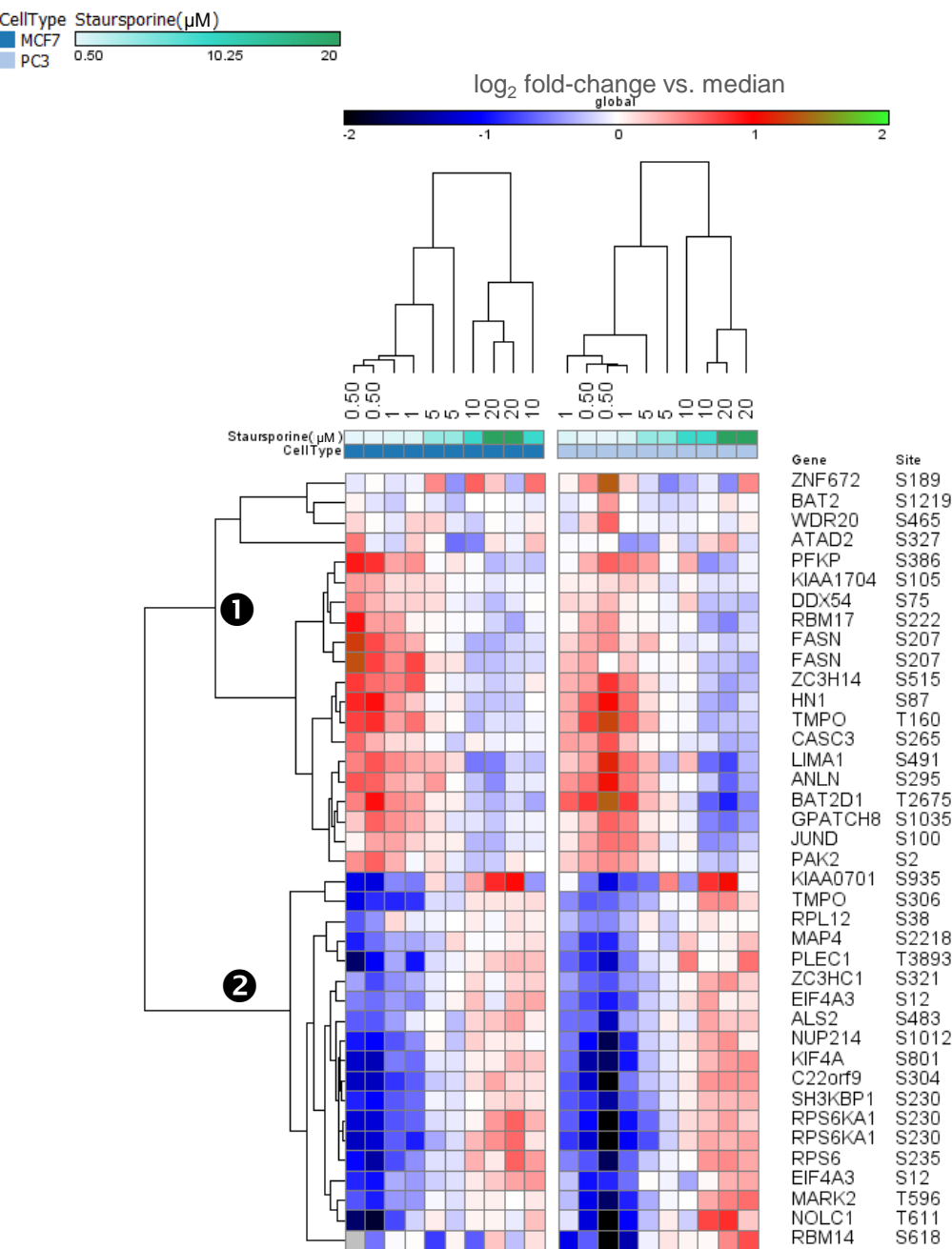
- Subset of P100 probes shown
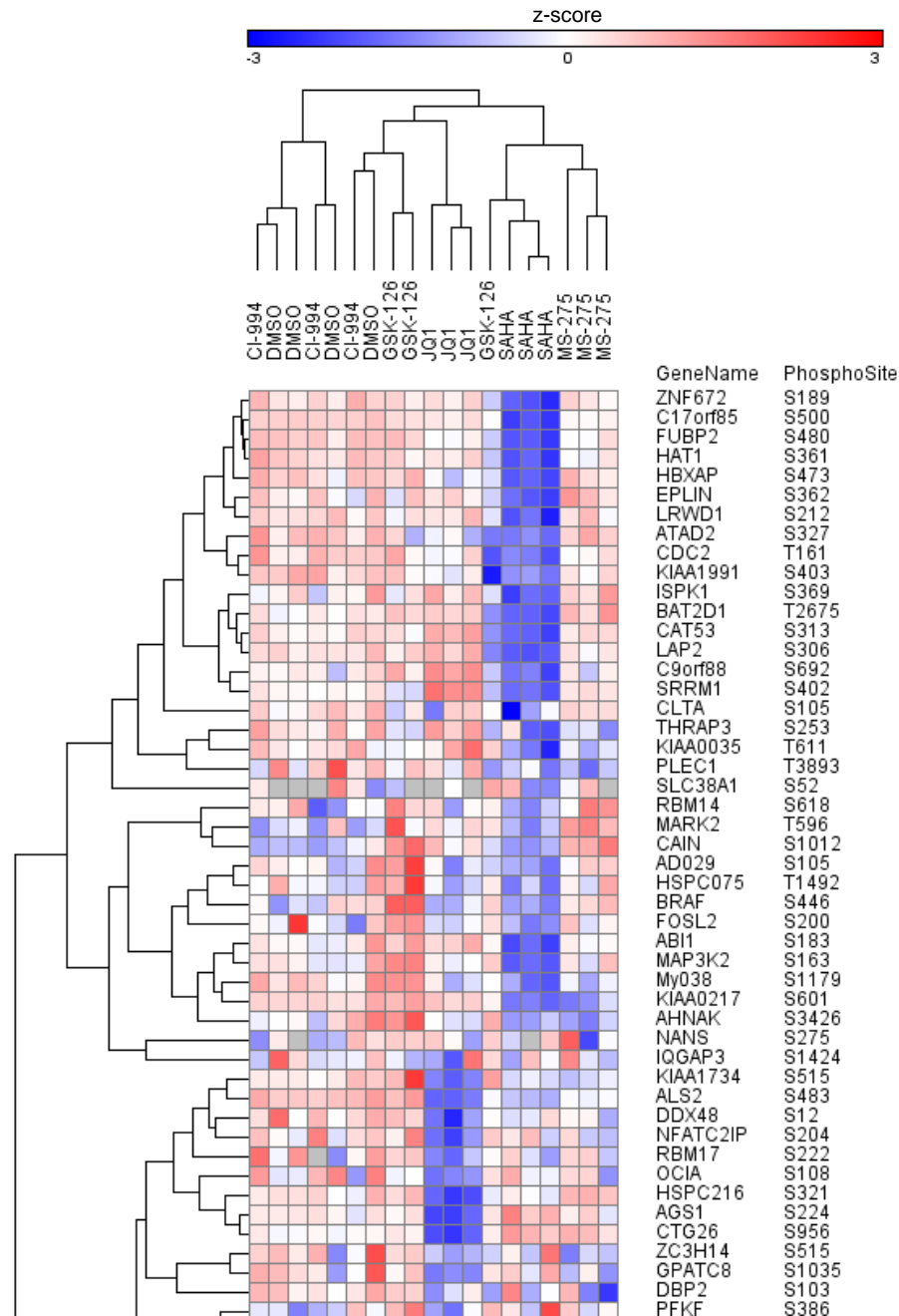- Unsupervised clustering puts doses in the ~right order

# Assay is robust across distant biological paradigms

- Cell system: Neural Progenitors derived from human ES cells
  - Using Dr. Evan Snyder's "neuralization" protocol

- Treatments: epigenetically directed compounds
  - HDACi, EZH2i, BRD4i

- No compounds or cell types in common with pilot

- Assay still generates robust signal
  - High rate of probe detection
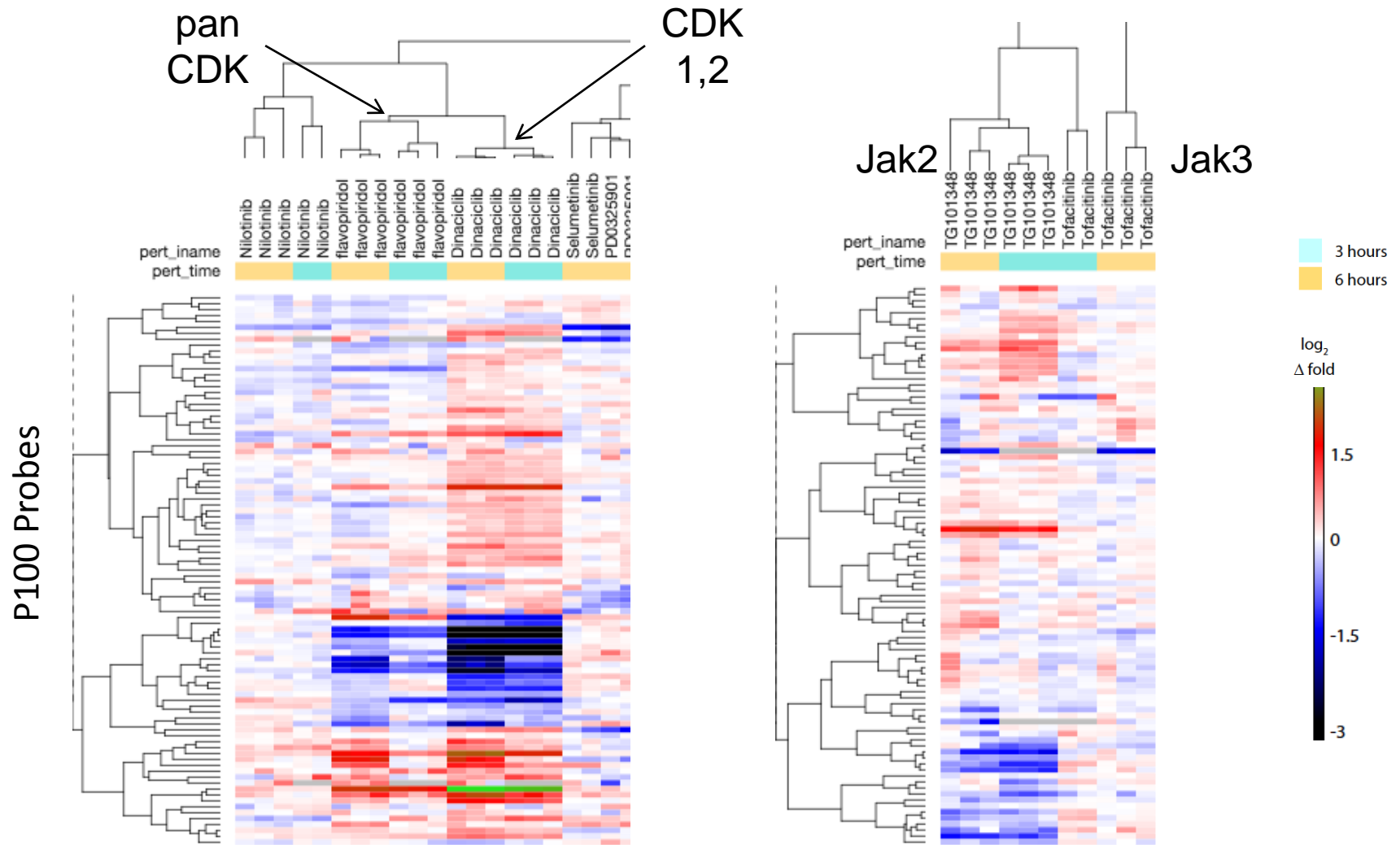 - Compound-specific signatures

Can P100 be Considered a Sentinel Assay?

Common Name	MOA	Common Name	MOA
DMSO	Solvent control	vorinostat	HDAC inhibitor; general
Selumetinib	Mek1/2 inhibitor	CC-401	Jnk inhibitor
PD0325901	Mek1/2 inhibitor	VX-970	Rep. stress/ATR inhibitor
Everolimus	mTOR inhibitor	losmapimod	p38 MAPK inhibitor
vemurafenib	Raf inhibitor	PRI-724	Notch/Wnt/Hedgehog inhibitor
TG101348	Jak2 inhibitor	dactolisib	Ras/PI3K inhibitor
Tofacitinib	Jak3 inhibitor	afuresertib	Ras/AKT inhibitor
Pravastatin	Stat1 inhibitor	BYL719	Ras/PI3K-P110a inhibitor
SCH 900776	Rep. stress/CHK1 inhibitor	Pazopanib	PDGFR and VEGFR; Also c-KIT, FGFR, inhibitor
flavopiridol	CDK inhibitor	Nilotinib	Multikinase inhibitor
PD-0332991	CDK/4,6 inhibitor	lenalidomide	immunomodulator
Dinaciclib	CDK/1,2,5,9 inhibitor	AR A014418	GSK3 inhibitor
RO4929097	Notch/gamma secretase inhibitor	BMS-345541	IkKB inhibitor
BMS-906024	Notch/other inhibitor	IPI-145	Ras/PI3K-P110g,d inhibitor
Verteporfin	Hippo inhibitor	staurosporine	Kinase inhibitor; general
SP600125	Jnk inhibitor	PS-1145	IkK inhibitor

■ Sample Collection Strategy

Acute stimulation
(small molecule)
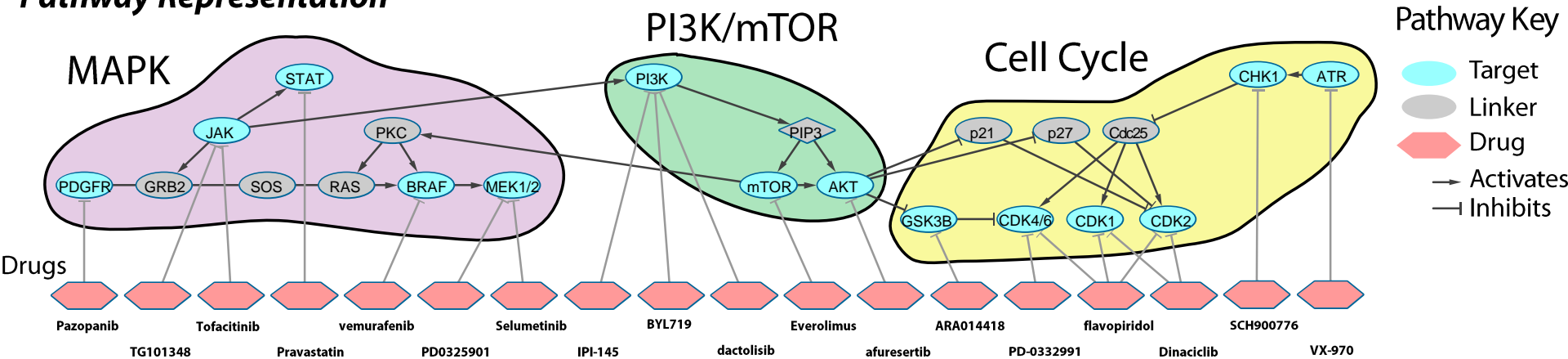
3 hours → P100   6 hours → P100   24 hours → P100
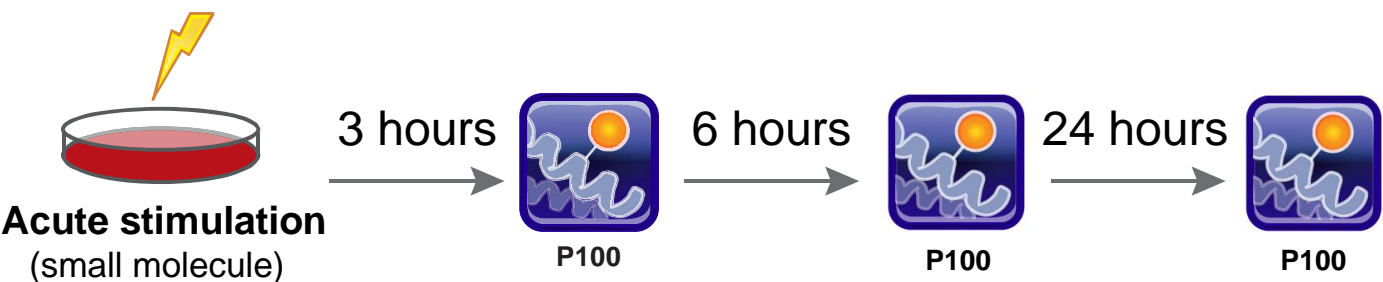
P100 Distinguishes Close Compound Class Members
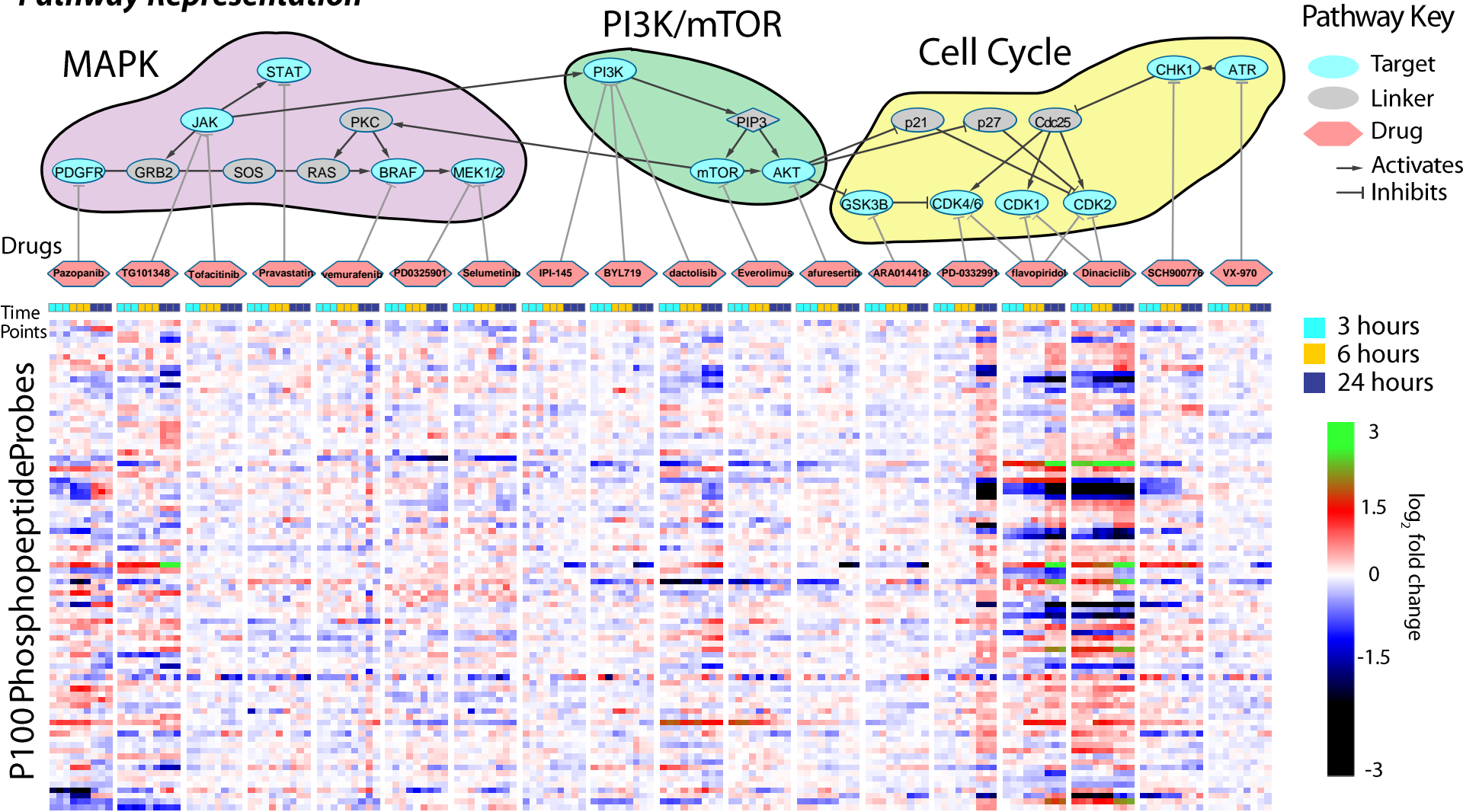
P100 as a pathway sentinel assay

Pathway Representation

- A "sentinel assay" informs on functional pathways without direct observation
  - Defined perturbagens can demonstrate sentinel functionality

- Treat MCF7 cells with various inhibitors along this pathway
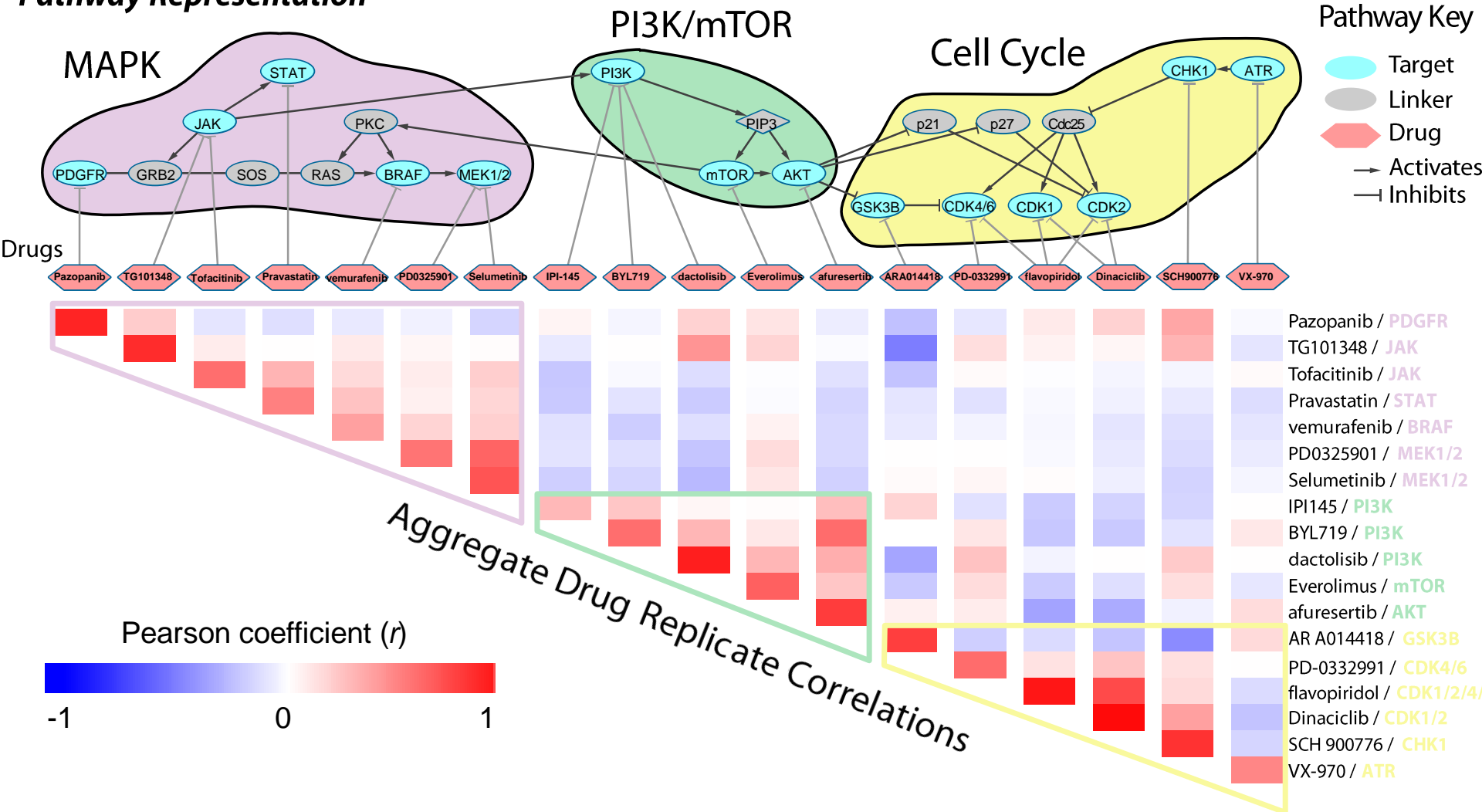
- Look for time-dependent effects

P100 dissects the linear pathway projection

P100 reveals local pathway modularity

Pathway Representation

Pathway Key

MAPK · PI3K/mTOR · Cell Cycle

Target · Linker · Drug · Activates · Inhibits

Aggregate Drug Replicate Correlations

Pearson coefficient (r)
-1   0   1

Pazopanib / PDGFR
TG101348 / JAK
Tofacitinib / JAK
Pravastatin / STAT
vemurafenib / BRAF
PD0325901 / MEK1/2
Selumetinib / MEK1/2
IPI145 / PI3K
BYL719 / PI3K
dactolisib / PI3K
Everolimus / mTOR
afuresertib / AKT
AR A014418 / GSK3B
PD-0332991 / CDK4/6
flavopiridol / CDK1/2/4/
Dinaciclib / CDK1/2
SCH 900776 / CHK1
VX-970 / ATR

Will it work?

- ✓ **Are there a set of phosphosites that we can (almost) always observe?**

- ✓ **Do these phosphosites behave in interesting ways in response to perturbations?**

- ✓ **Do we gain anything over gene expression measurements?**

- ✓ **Can we pick a (small) set of highly informative phosphopeptides?**

- ✓ **Could profiling these peptides reproduce the observations of a deeper experiment?**

- ✓ **Is the assay generally useful?**

Revisiting the final goals of P100 assay development

- ✓ Drug responsiveness – many compound classes

- ✓ Universal (-ish) – works even in NPCs

- ✓ Ser/Thr focused – IMAC enrichment

- ✓ Compact – 60 minutes analytical time

- ✓ Operates at scale – Automated sample prep

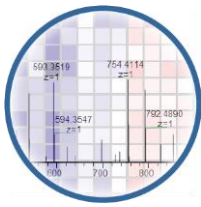- ✓ Longitudinally comparable – Synthetic standards with defined levels
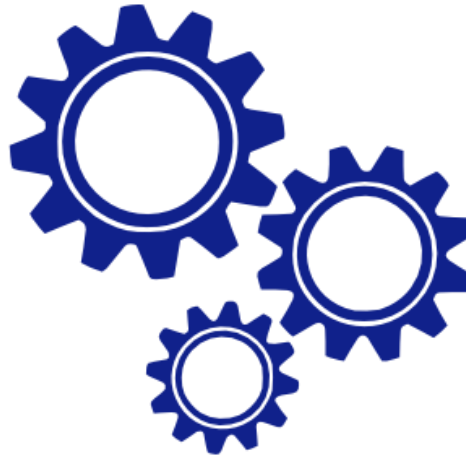
Sentinel assays in a production environment

Proteomic data pipelines for reproducible, shareable research

Introduction to the LINCS Consortium
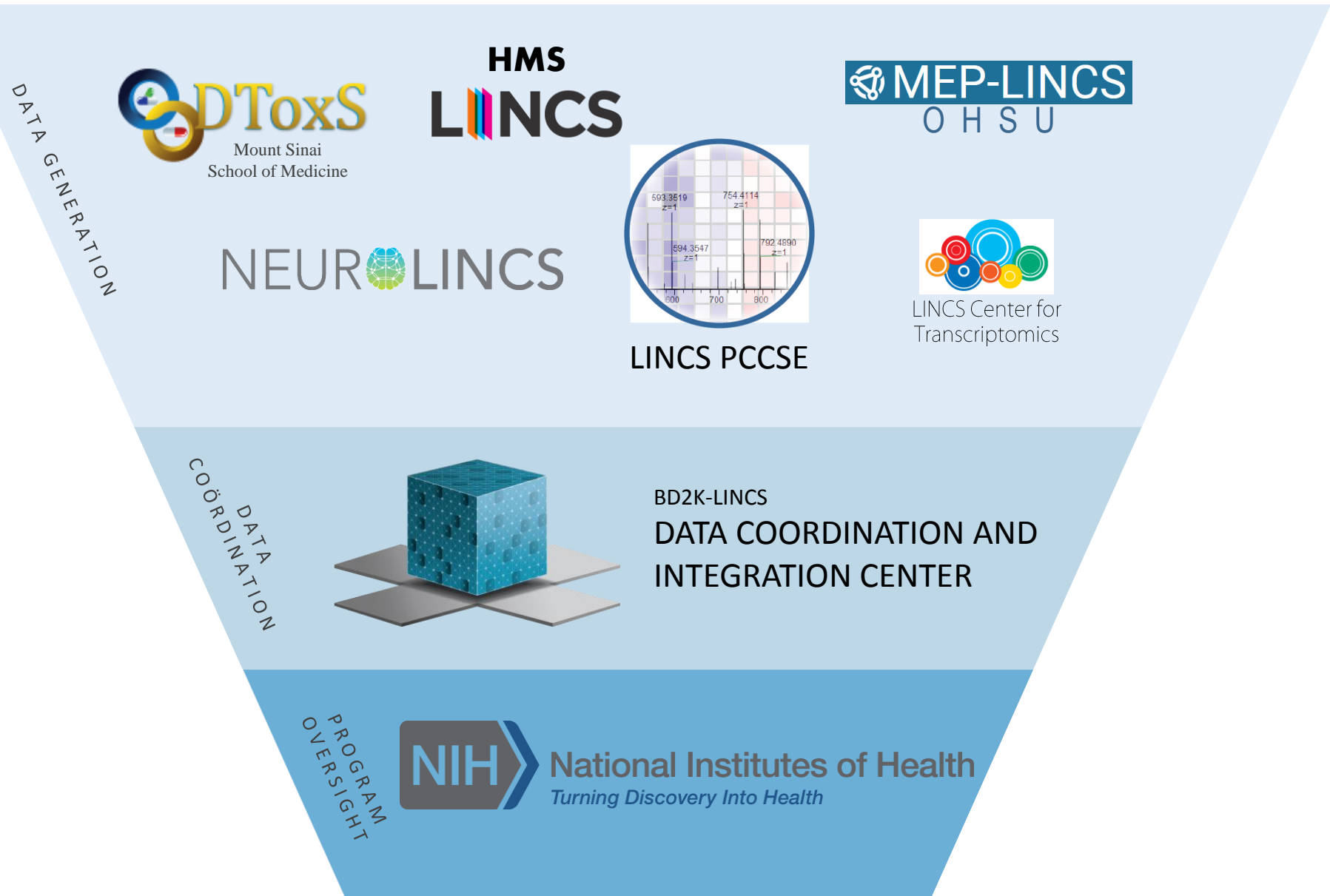
A truly amazing team…

Fatema Abdurrob   Mukta Bagul   Joel Blanchard   Amanda Creech   Desiree Davison   Katherine DeRuff   Jarrett Egertson   Shawn Egri   Todd Golub

Jake Jaffe   Tak Ko   Dave Lahr   Dan Lam   Lev Litichevskiy   Xiaodong Lu   Mike Mac Coss   Brendan MacLean   Ted Natoli
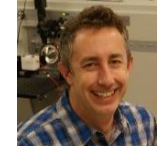
Adam Officer   Ryan Peckner   Malvina Papanastasiou   Brian Searle   Vagisha Sharma   Aravind Subramanian   Sonia Ting   Li-Huei Tsai   Sebastian Vaca   Jennie Young

NIH LINCS PROGRAM

BROAD INSTITUTE

Massachusetts Institute of Technology

UNIVERSITY of WASHINGTON

Proteomics and Biomarker Discovery

Phosphosignaling is informative and targetable

Phosphosignaling

θ

ψ

Gene Expression

ϕ

Epigenetics

- >270k catalogued phosphosites[1]
  - >84% are Ser/Thr

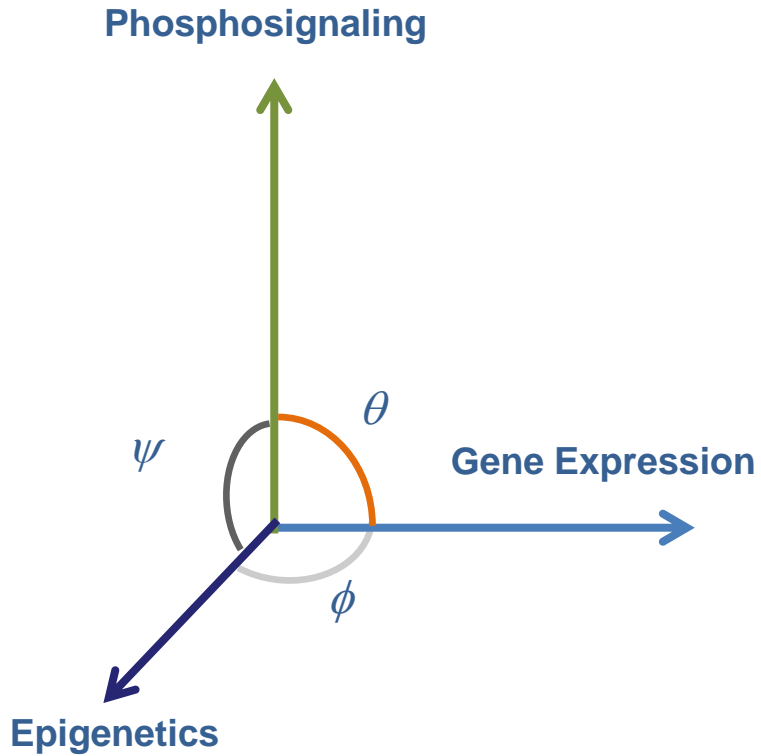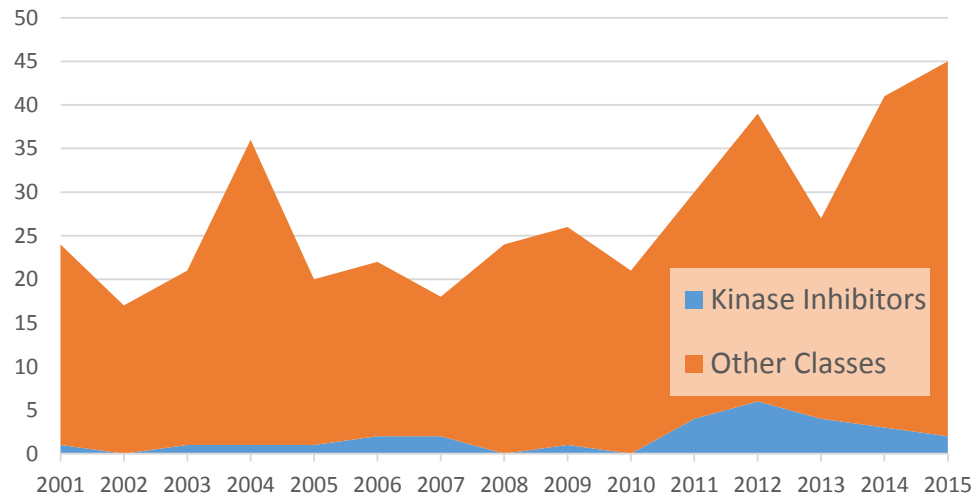- Expanding therapeutic class that targets phosphosignaling[2]
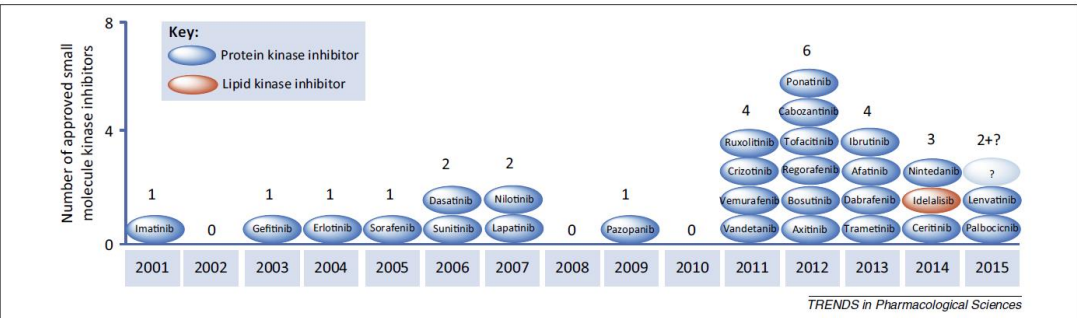
Figure 2. FDA-approved small-molecule kinase inhibitors (at April 2015).

New Drugs Approved Each Year[3]

[1]Hornbeck, et al *Nuc Acids Res* (2015)
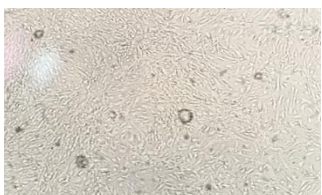[2]Wu et al *Trends in Pharm Sci* (2015)
[3]FDA (+ ref 2)

Proteomic Characterization Center for Signaling and Epigenetics: Center Scientific Overview
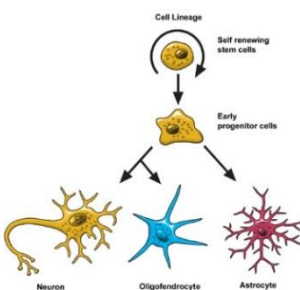

Cell Lines

Neural Lineages

Primary Cells
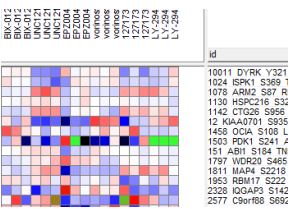
X

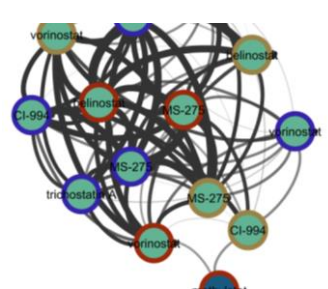Drugs

Genes

Fates

X

Targeted
MS DATA

P100 / DIA
(phospho)

GCP
(histone marks)

Matrix DATA

Signatures

Connectivity

Our Compounds

Set 1: Epigenetically Active

Common Name	MOA
DMSO	negative control
GSK1210151A	Bromodomain inhibitor; BRD2,3,4
GSK525762A	Bromodomain inhibitor; BRD2,3,4
jq1	Bromodomain inhibitor; BRD4
zebularine	DNMT1 inhibitor
decitabine	DNMT1 inhibitor
EPZ-5676	DOT1L inhibitor
EPZ004777	DOT1L inhibitor
gsk126	EHZ2 inhibitor
EPZ-5687	EHZ2 inhibitor
CPI-169	EHZ2 inhibitor
UNC0646	G9a inhibitor
UNC0321	G9a inhibitor
BIX 01338	G9a inhibitor
BIX 01294	G9a inhibitor; ? also kinase inhibitor (mTOR)
vorinostat	HDAC inhibitor; general
entinostat (MS275)	HDAC inhibitor; general
tacedinaline	HDAC inhibitor; general
trichostatin-a	HDAC inhibitor; general
belinostat	HDAC inhibitor; general
methylstat	JHDM inhibitor (Jumonji C)
GSK-J4	JMJD3 inhib
UNC1215	L3MBTL3 probe
MI-2	MLL inhibitor
OSI-027	mTOR inhibitor
resveratrol	SirT1 activator (?)
EX527	SirT1 inhib
salermide	SirT1 inhib; SirT2 inhib
geldanamycin	HSP90 inhibitor
rapamycin	mTOR inhibitor
staurosporine	Kinase inhibitor; general
LY-294002	Kinase inhibitor; pim-1, PI3k

Set 2: Neuroactive

| Common Name	MOA
DMSO	negative control
Etoposide	Top2B inhibitor, DNA DSB inducer
KU-55933	ATM inhibitor
SMER-3	E3 ligase inhibitor
ruxolitinib	JAK 1 & 2 inhibitor
niclosamide	STAT3 signaling inhibitor
VU0155056	inhibitor of phospholipase D1/D2
tacrolimus	inhibitor of calcineurin
tretinoin	Retinoic Acid Receptor agonist
curcumin	NFkB inhibitor
semagacestat	gamma secretase inhibitor
dexamethasone	Glucocorticoid receptor agonist
Olaparib	PARP inhibitor
bafilomycin A1	inhibitor of the vacuolar-type H+-ATPase
Exifone	Antioxidant
vorinostat	HDAC inhibitor; general
KN-62	CaMKIIalpha inhibitor
calpain inhibitor II	calpain inhibitor
C646	CBP/p300 inhibitor
Compound E	gamma secretase inhibitor
KN-93	CaMKII/CaMKIV inhibitor
Gossypetin	HDAC1 activator
Ginkgetin	HDAC1 activator
Okadaic Acid	PP1 and PP2a inhibitor
momelotinib	JAK1/2 inhibitor
TBB	Casein kinase II inhibitor
Roscovitine	Cdk5 inhibitor
Rolipram	PDE4 inhibitor
RGFP966	HDAC3 inhibitor
sotrastaurin	inhibitor of protein kinase C beta
staurosporine	Kinase inhibitor; general
CHIR99021	GSK3 inhibitor

Set 3: Ser/Thr kinase pathways

Common Name	MOA
DMSO	negative control
Selumetinib	Mek1/2 inhibitor
PD0325901	Mek1/2 inhibitor
Everolimus	mTOR inhibitor
vemurafenib	Raf inhibitor
TG101348	Jak2 inhibitor
Tofacitinib	Jak3 inhibitor
Pravastatin	Stat1 inhibitor
SCH 900776	Rep. stress/CHK1 inhibitor
flavopiridol	CDK inhibitor
PD-0332991	CDK/4,6 inhibitor
Dinaciclib	CDK/1,2,5,9 inhibitor
RO4929097	Notch/gamma secretase inhibitor
BMS-906024	Notch/other inhibitor
Verteporfin	Hippo inhibitor
SP600125	Jnk inhibitor
vorinostat	HDAC inhibitor; general
CC-401	Jnk inhibitor
VX-970	Rep. stress/ATR inhibitor
losmapimod	p38 MAPK inhibitor
PRI-724	Notch/Wnt/Hedgehog inhibitor
dactolisib	Ras/PI3K inhibitor
afuresertib	Ras/AKT inhibitor
BYL719	Ras/PI3K-P110a inhibitor
Pazopanib	PDGFR and VEGFR; multi RTKi
Nilotinib	Multikinase inhibitor
lenalidomide	immunomodulator
AR A014418	GSK3 inhibitor
BMS-345541	IkKB inhibitor
IPI-145	Ras/PI3K-P110g,d inhibitor
staurosporine	Kinase inhibitor; general
PS-1145	IkK inhibitor

Data are freely available

https://bit.ly/PCCSEData

Or just Google "LINCS panorama"
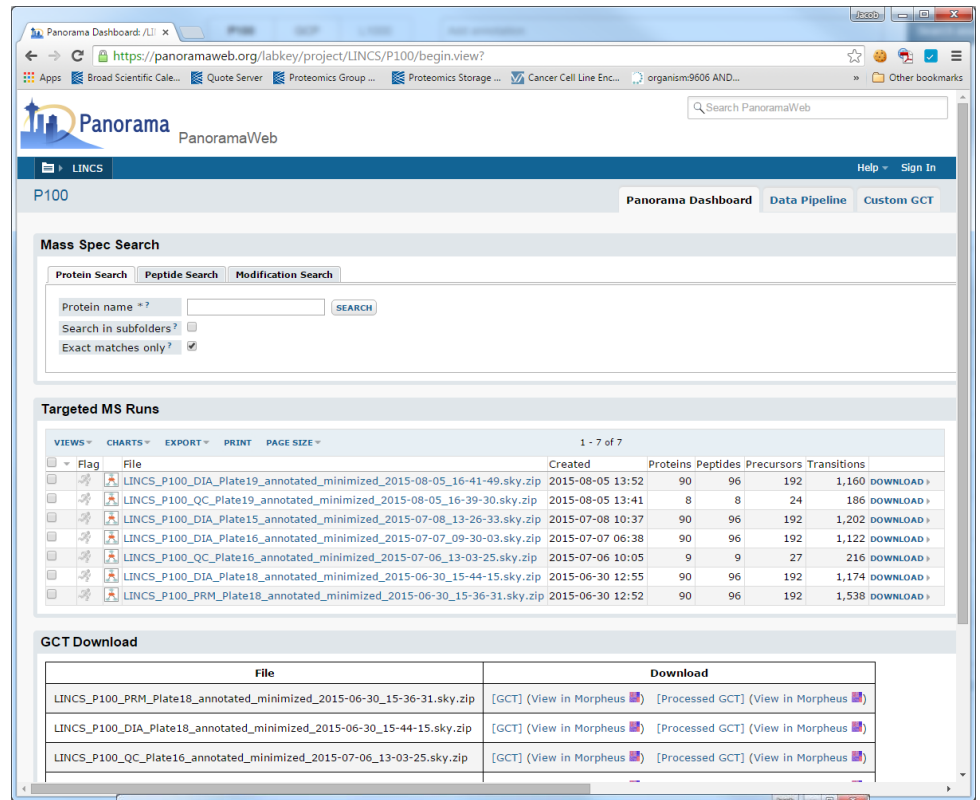
PCCSE Core Goals

- Create large-scale perturbation data with multiple proteomic readouts (~10,000 samples in 6 years)

- Have an impact on drug characterization and discovery of therapeutic indications

- Share data on a common platform with consortium members
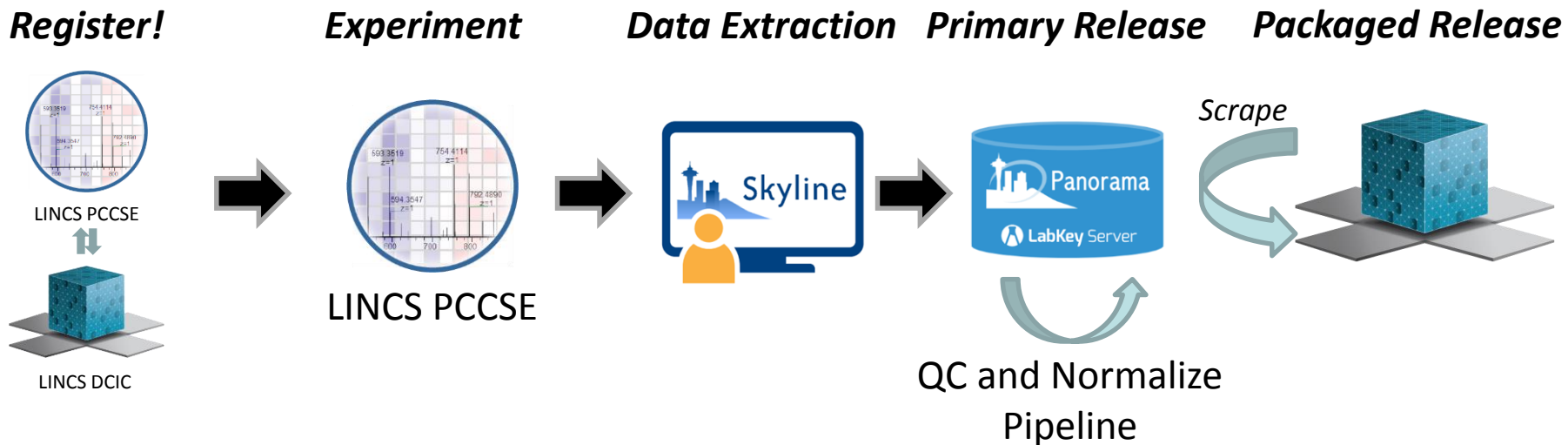
- Share data – and make it useful – with the world

How do we manage proteomic data on this scale with these goals in mind?

What would the ideal data pipeline look like?

- *Prêt-à-porter* philosophy: output is ready to consume

- Automated

- Standards-based

- Controlled & controllable

- Replicable by others

- Self-documenting

- Minimizes human errors

Data life cycle overview

Register! **Experiment** **Data Extraction** **Primary Release** **Packaged Release**

LINCS PCCSE

LINCS DCIC

LINCS PCCSE

Skyline

Panorama
LabKey Server

Scrape

QC and Normalize Pipeline

Official LINCS Data Levels

0: Raw MS Data   **1**: Raw Quant Profiles   **2**: Raw Numerical Data
3: Normalized Data

Embedded Metadata
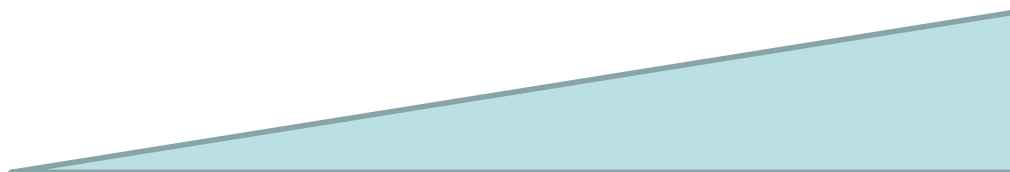
Metadata embedding and processing cycle in practice

- **Entity registration via templated documents**
  - Handshake with DCIC
  - Enables DCIC to add value, integrate

Internal ID#	Compound Name	SM_Center_Sample_ID	PubChem ID	Isomeric SMILES	LSM ID	Treatment Conc
12-1-1	staurosporine	BRD-K17953061	44259	C[C@@]12[C@@H]	LSM-1103	1
12-1-2	staurosporine	BRD-K17953061	44259	C[C@@]12[C@@H]	LSM-1103	1

Automated generation to reduce errors

- **Preformat metadata before beginning!**
  - Directly generate MS queue sequence
  - Prepare for further integration

Internal ID#	Compound Name	LSM ID	BRD# of Compound	Pubchem CID	Isomeric SMILES string
89-1-5	DMSO	LSM-36361	BRD-K08970894	679	CS(=O)C
89-1-5	DMSO	LSM-36361	BRD-K08970894	679	CS(=O)C
89-1-5	DMSO	LSM-36361	BRD-K08970894	679	CS(=O)C

Acquisition Instrument	Curie
Acquisition LC	N
Acquisition Date	9/15/2016
Study Identifier	P100_Plate39b_HAoSMC_CardioOutreach
Path	C:\Curie_Public\2016_Sept\LINCS C:\Xcalibur\methods\LINCS\P100_DIA_11amu_Overlap_22a
Method	muwindow_60min_50msIT_400-1000_27loopcount

Unique Sample IDs generated, integrated into raw files

- **MS data acquisition**
  - Maximize redundancy and traceability
  - Unique ID directly incorporated into MS files

```
CM20160908_P100_Plate39b_HAoSMC_CardioOutreach_P-0039b_A01_acq_01
Type: Unknown ID: PHC-44828-001A01 Row: 0
Sample Name:
Comments:        BRD-K08970894_DMSO_0_uM_3_h_rep_01
```

Attachment of metadata and reverification

- **Embed metadata directly in Skyline**
  - Reattach metadata and integrate preliminary results

Document Grid: p100_samples_IMPORTCHECK

Views ▾ | ◀ 1 of 80 ▶ ▶| | Export... | Find: | 🄰ᵃ

Replicate Name	id	cell_id	det_plate	det_well	det_well_enricher	det_normalization_c	det_filename	pert_id	lsm_id	pubchem_cid
A01_acq_01	PHC-44828-001A...	HAoSMC	P-0039b	A1	0.73	1	CN20160915_P1...	BRD-K08970894	LSM-36361	679
A02_acq_01	PHC-44828-002A...	HAoSMC	P-0039b	A2	0.85	1	CN20160915_P1...	BRD-K08970894	LSM-36361	679
A03_acq_01	PHC-44828-003A...	HAoSMC	P-0039b	A3	0.8	1	CN20160915_P1...	BRD-K08970894	LSM-36361	679

Data export

- **Publish to Panorama**
  - Metadata carries from Skyline
  - Processing, QC, normalization happens on server

Panorama
LabKey Server

DCIC involved in curation of metadata for interoperability and future integration

Genetic Perturbagens
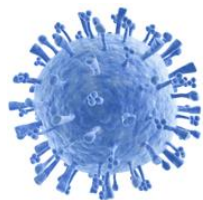
- PAM sequence
- Target gene information
- Internal ID

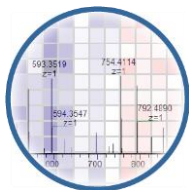**Small Molecule Perturbagens**
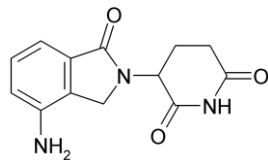
- Compound Name
- PubChem CID
- Batch information
- Internal ID
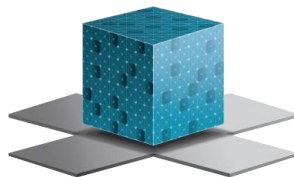
Information registered with DCIC who assign unique LSM/LNA identifiers

- Used in data integration across multiple assay types
- All metadata verified with the DCIC prior to sample annotation

Collect metadata for compounds and viruses, verify and register with DCIC and input into our databases

- Aim to minimize human intervention in this process
- Assign unique sample IDs

- More than 20 metadata fields are assigned to each sample prior to release
- Retain several more for internal records

QC, Normalization, and Pipeline Control

1. Extract quantitative ratios

2. Log_2 transform

3. Filter out samples with low detection rates

4. Filter out peptides with low detection rates

5. Sample load normalization (columns)

6. Sample outlier detection/removal

7. Peptide normalization and scaling (rows)

https://bitbucket.org/jaketeyjake/p100_production

pr_probe_normalization_group	pr_probe_suitability	pr_processing_params
1	TRUE	{"samplePctCutoff":0.7}
1	TRUE	{"samplePctCutoff":0.7}
1	TRUE	{"samplePctCutoff":0.7}
1	TRUE	{"samplePctCutoff":0.7}
1	TRUE	{"samplePctCutoff":0.7}
1	TRUE	{"samplePctCutoff":0.7}
1	FALSE	{"samplePctCutoff":0.7}

- Process controls can be embedded in Skyline document as JSON:

  - { "samplePctCutoff":0.7, "probePctCutoff":0.5 }

- Self-documenting pipeline via provenance codes, added to metadata:

  - PRM1+L2X+SF7+PF5+PSDF3+LLB+OF5+RMN

- Same code runs locally and on Panorama server

Ready to consume matrix format: GCT

https://github.com/cmap/l1ktools
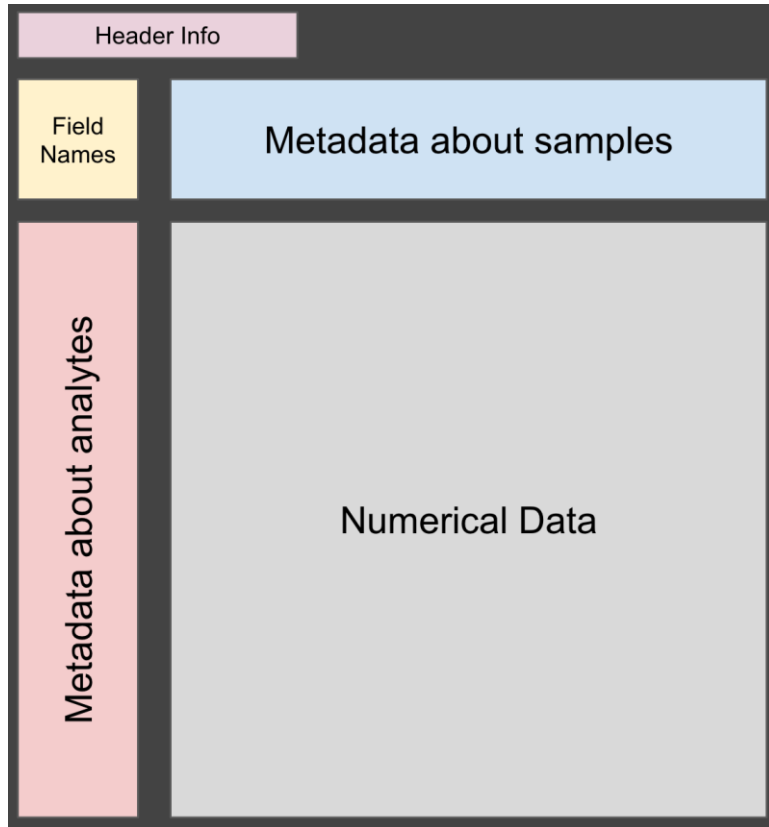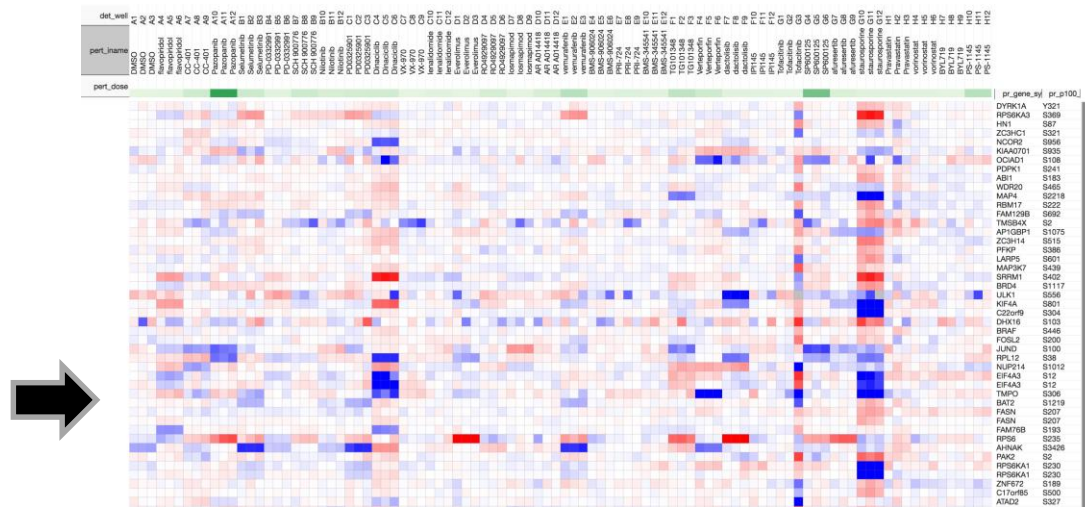
- Easily consumable by a growing number of tools
  - Morpheus, GSEA, various APIs (R, python, java, etc.)

- Keeps metadata associated with data

- Flexible but well-described
  - HDF option for "big" data

Panorama is one-stop shopping for data

http://bit.ly/PCCSEData

Panorama — PanoramaWeb

GCT Download (P100)

| *Level 1* — File	*Level 2* — Download *Level 3*
LINCS_P100_DIA_Plate41_annotated_minimized_2017-02-16_14-31-38.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate43_annotated_minimized_2017-02-15_11-53-50.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate39_annotated_minimized_2017-02-10_16-39-35.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate24_annotated_minimized_2017-01-10_12-37-57.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate38_annotated_minimized_2017-01-05_17-02-26.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate40_annotated_minimized_2016-11-18_16-14-26.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate34_annotated_minimized_2016-08-15_10-38-52.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_PRM_Plate31_annotated_minimized_2016-08-15_09-45-56.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate30_annotated_minimized_2016-08-05_10-41-17.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_PRM_Plate32_annotated_minimized_2016-07-22_11-29-42.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_DIA_Plate33_annotated_minimized_2016-06-29_12-24-03.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_PRM_Plate35_annotated_minimized_2016-06-10_11-43-55.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)
LINCS_P100_PRM_Plate25_annotated_minimized_2016-06-10_11-43-57.sky.zip	[GCT] (View in Morpheus) [Processed GCT] (View in Morpheus)

Panorama is the engine that drives reproducible automated analysis
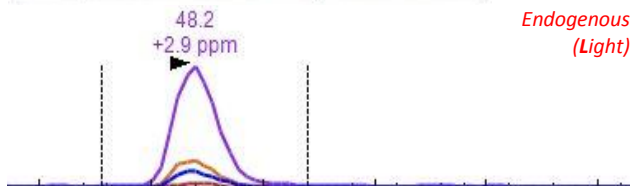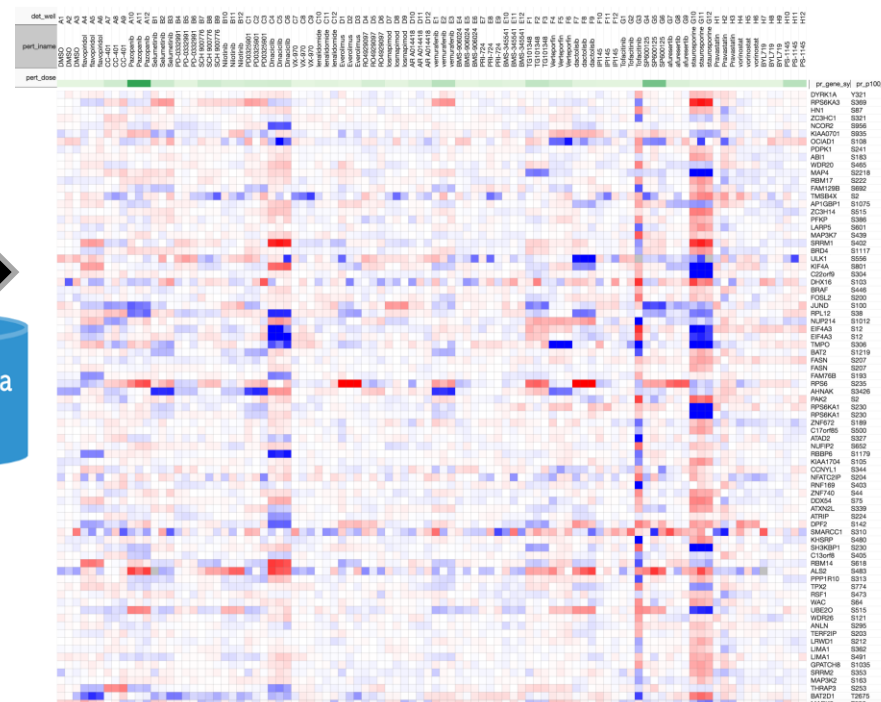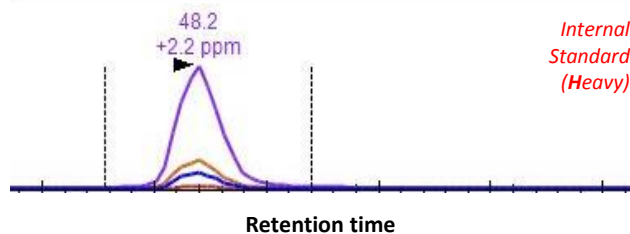
- Just upload Skyline document and Panorama/LabKey executes data pipeline
- Links out to downstream tools
- Custom data slicing
- API access to data

Key lessons learned for data pipelines

- Think of the downstream audience first

- Leverage suitable external resources, expertise, pre-existing architecture

- Determine what metadata is needed to make the data itself useful, mineable, discoverable, harmonizable

- Programs, not spreadsheets

- Embed and document