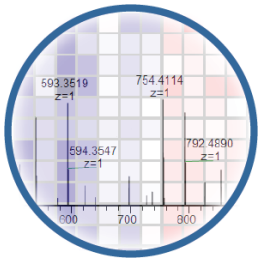
Library of Integrated Network-based Cellular Signatures



Proteomic Characterization Center for Signaling and Epigenetics

Strange and Wonderful Observations from the World of Proteomic Profiling

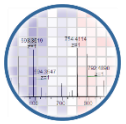
September 19, 2016

Jacob D. Jaffe – Broad Institute, PI

Li-Huei Tsai – MIT, co-investigator

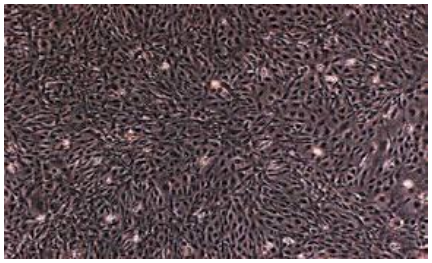
Michael MacCoss – U. of Washington, co-investigator

Outline



- Profiling and Infrastructure Progress
- Assessment of dataset quality
- Vignettes
- Extension into neuronal models
- Unlocking deeper signaling information through next-gen MS
- Outreach Activities

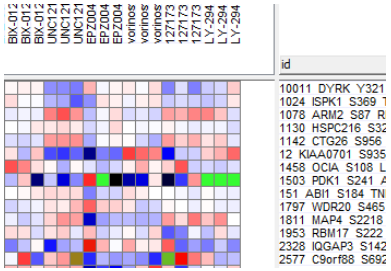
PCCSE Center Scientific Overview



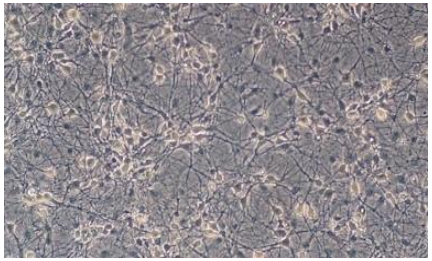
Cell Lines

Drugs

P100 / DIA
(phospho)



Signatures



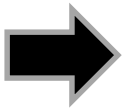
Neural Lineages

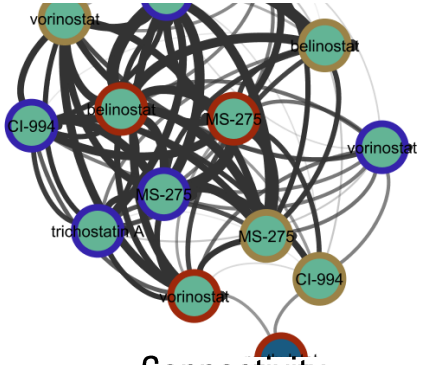
X

Genes

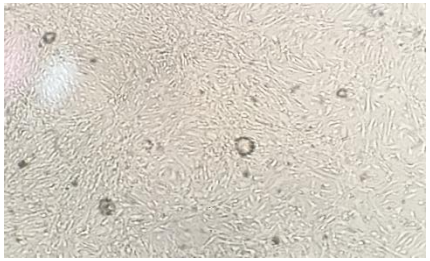
X

GCP
(histone marks)

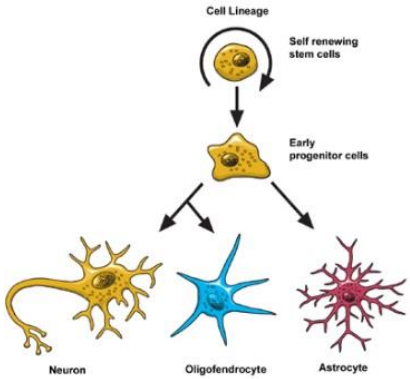




Connectivity



Primary Cells

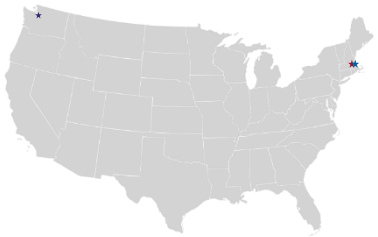


Fates

L1000
(mRNA)

Circuitry

Extensive progress building the library

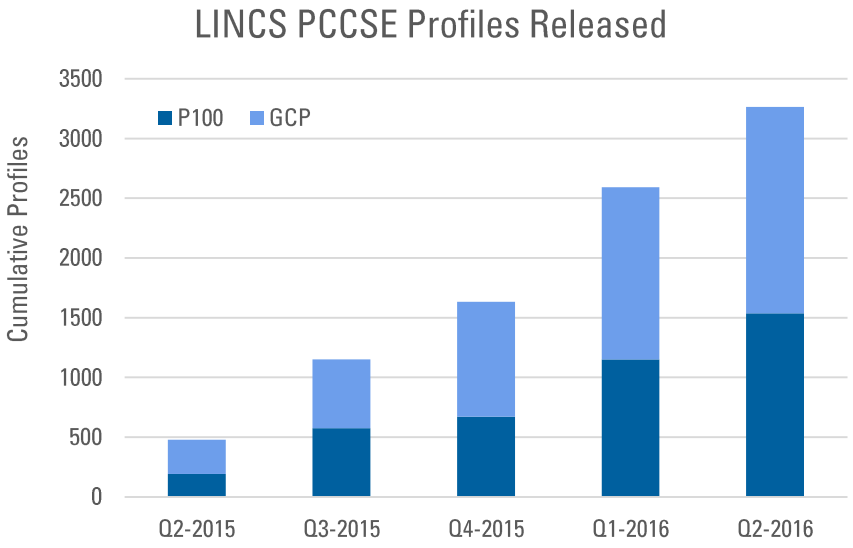


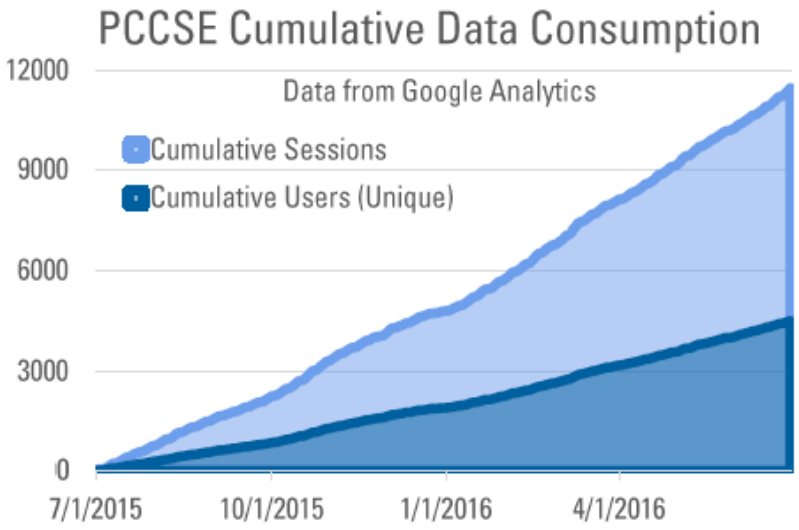
BROAD INSTITUTE

Massachusetts Institute of Technology

UNIVERSITY of WASHINGTON

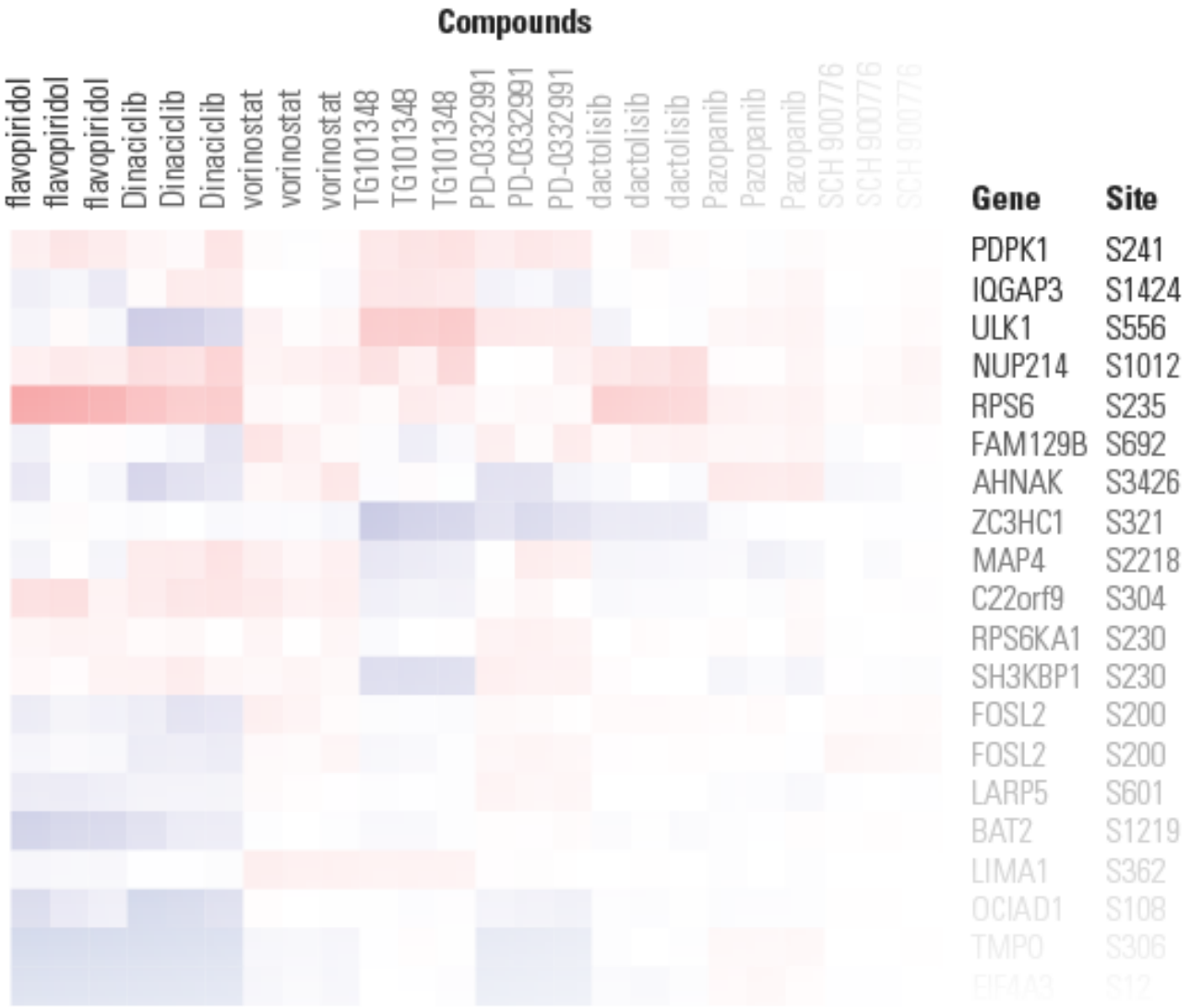
- 3400+ biological samples generated spanning:
  - 90 compounds in triplicate spanning many MoA
  - Beginning CRISPR/Cas9-based gene disruptions
  - 6 cellular models (breast, lung, skin, prostate, and pancreatic cancers; neuronal precursors)
  - 2 assay platforms (+L1000 in collaboration)
- Key infrastructure:
  - Public data repository with built-in signature visualization (bit.ly/PCCSEData)
  - Documented workflows; automated analytical pipeline for reproducibility





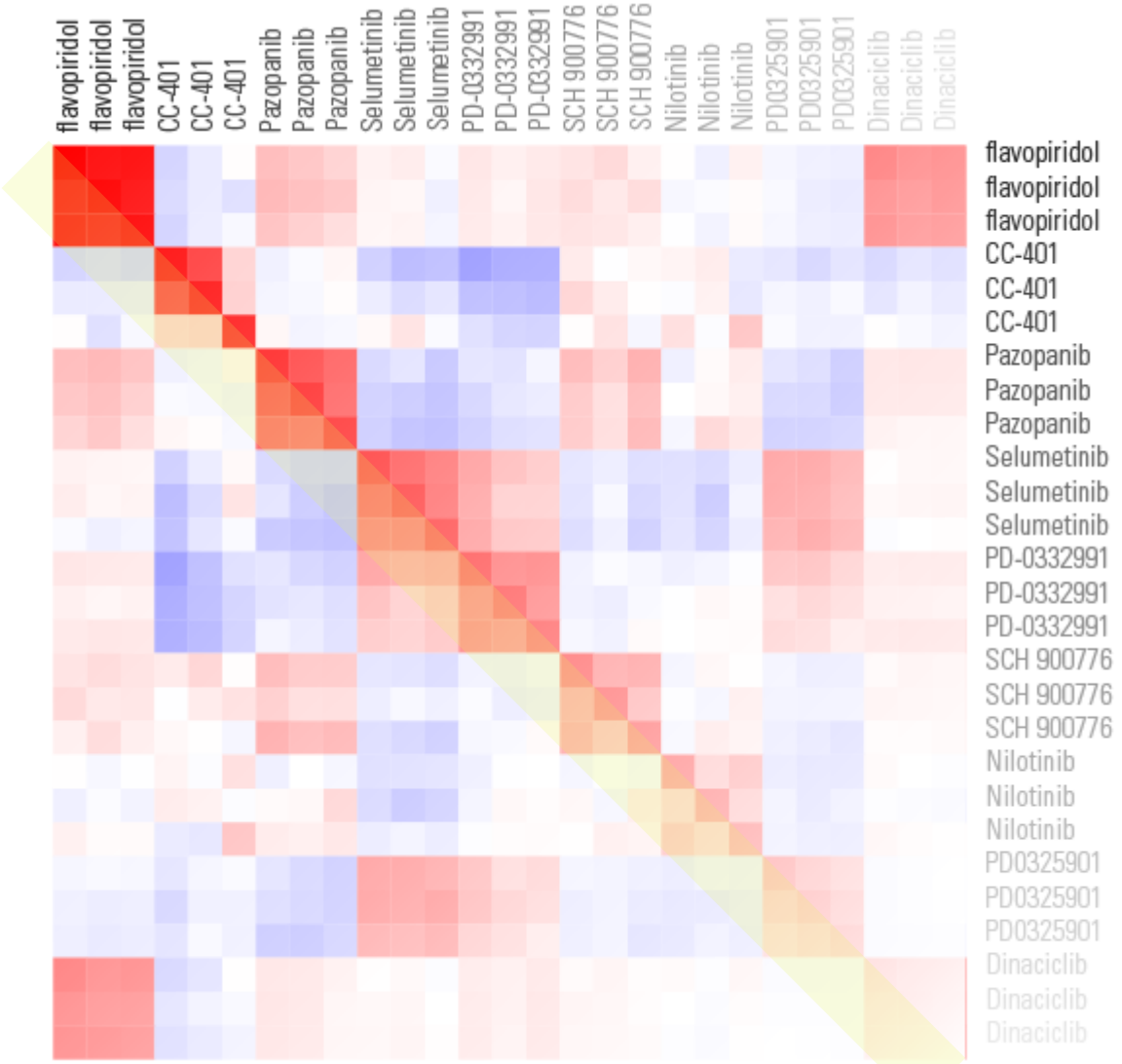
Turning signatures into connections: Signatures

Signatures are groups of related profiles



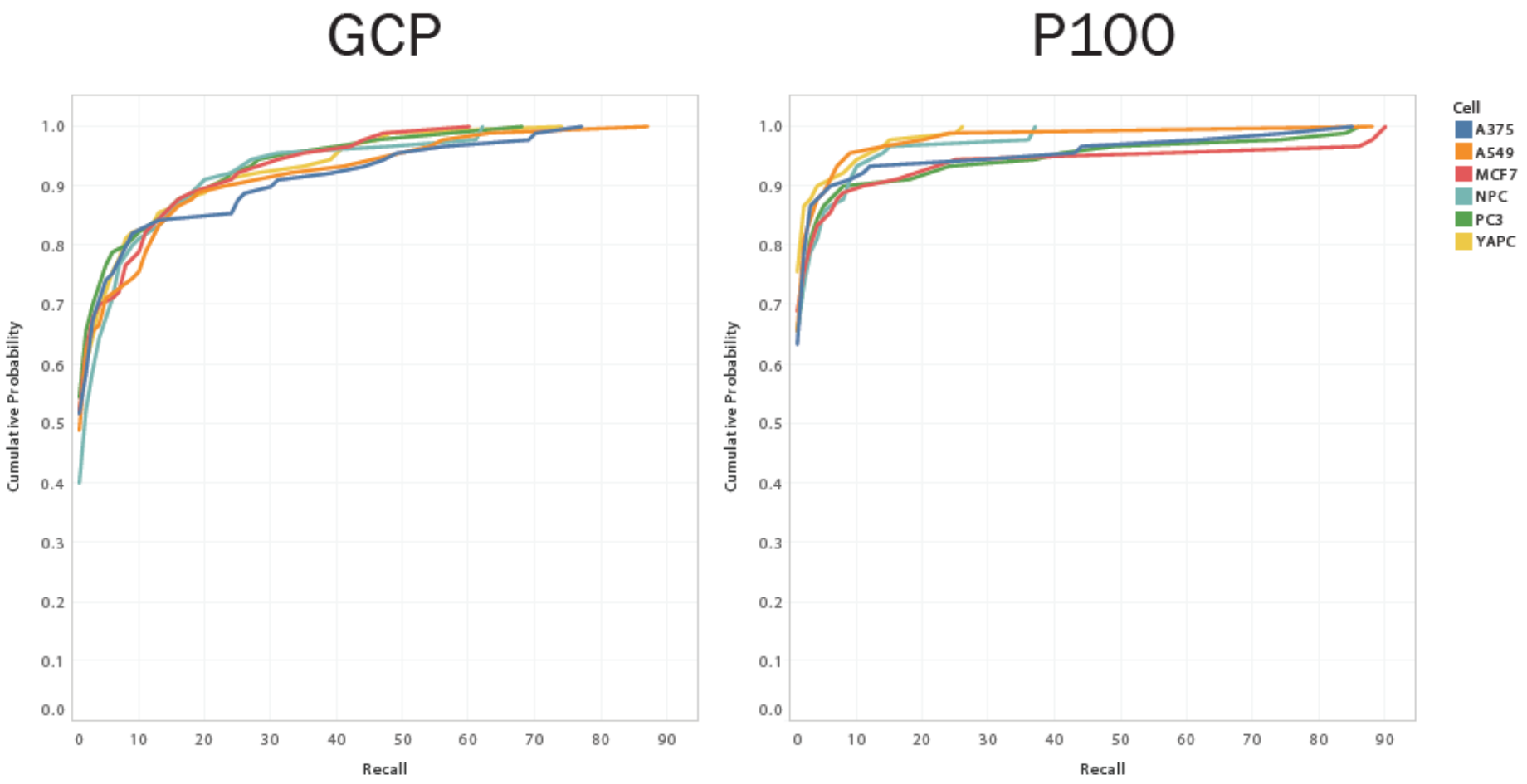
Turning signatures into connections: Similarities

Similarities are correlations computed from signatures



Replicate consistency describes similarity on the lower diagonal

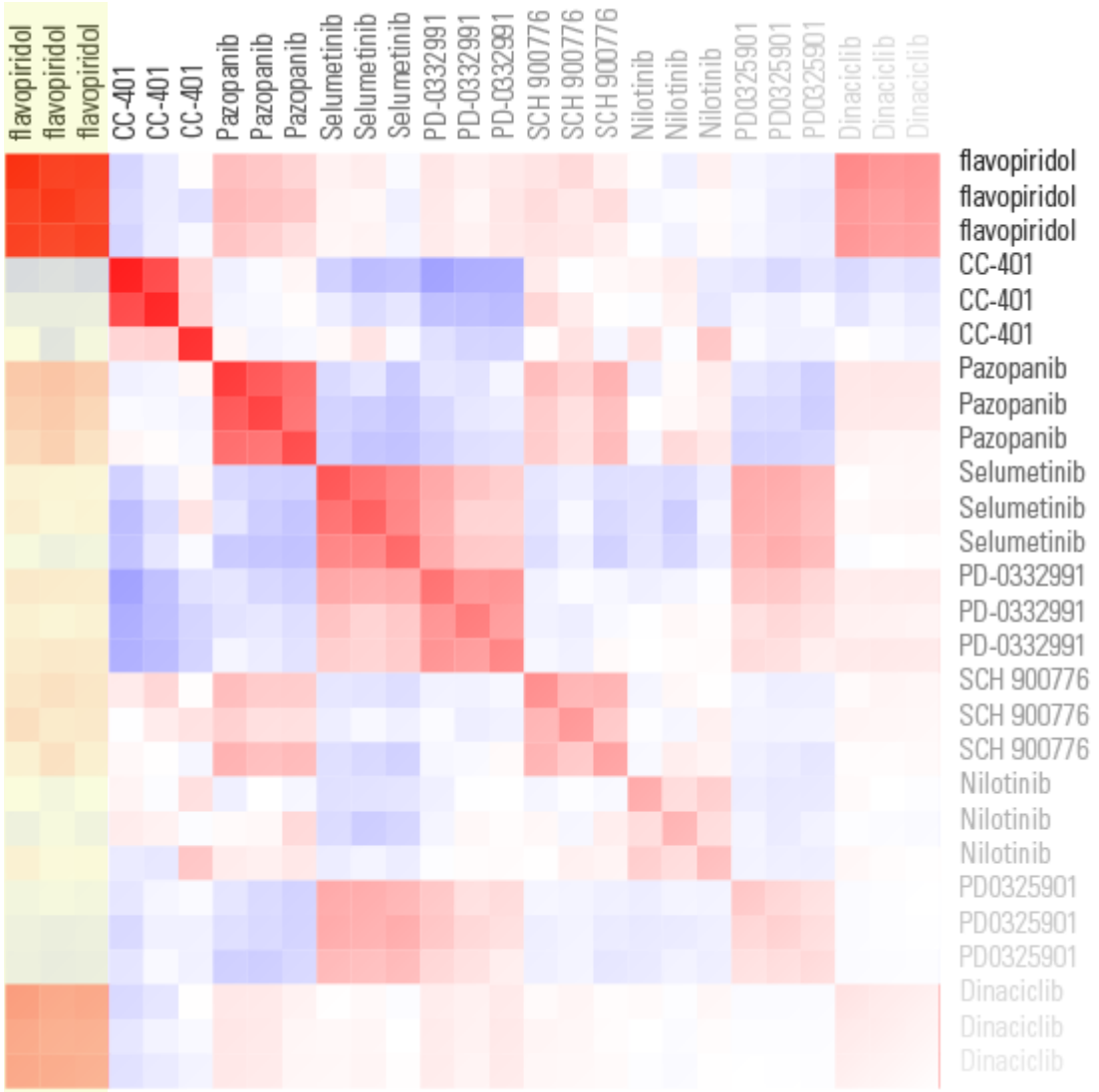
Consistency rank via similarity assesses dataset quality



- Average AUC: GCP ~92%, P100 ~96%
- 66% of compounds have signal in GCP (within top 3 recall rank)
- 83% of compounds have signal in P100 (within top 3 recall rank)

Turning signatures into connections: Connectivity

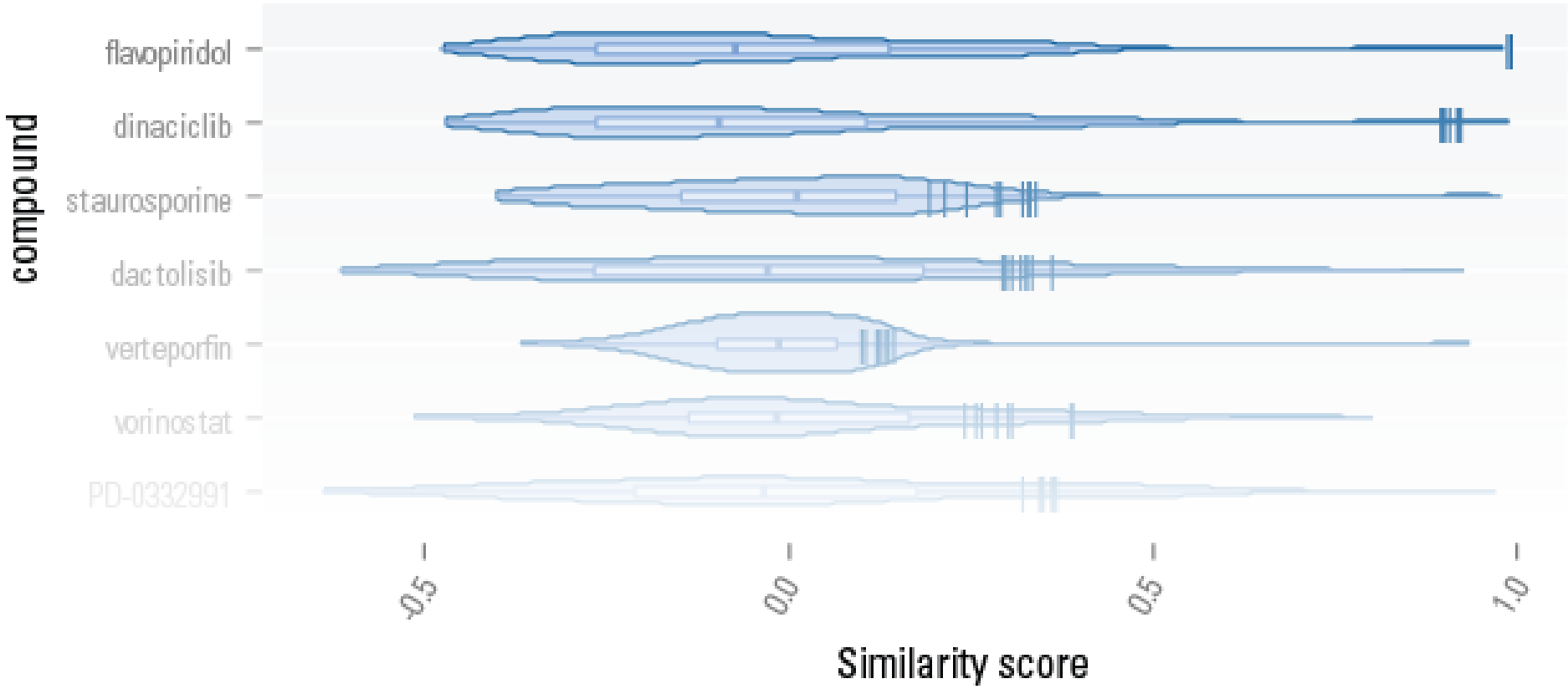
Similarities are correlations computed from signatures



Connectivity comes from analysis of the similarity matrix

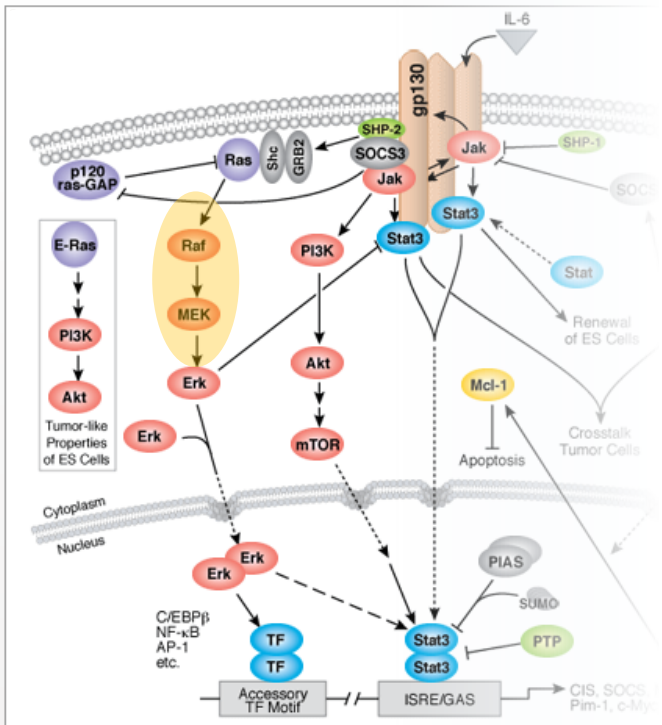
Turning signatures into connections

Connectivity queries put correlations in context

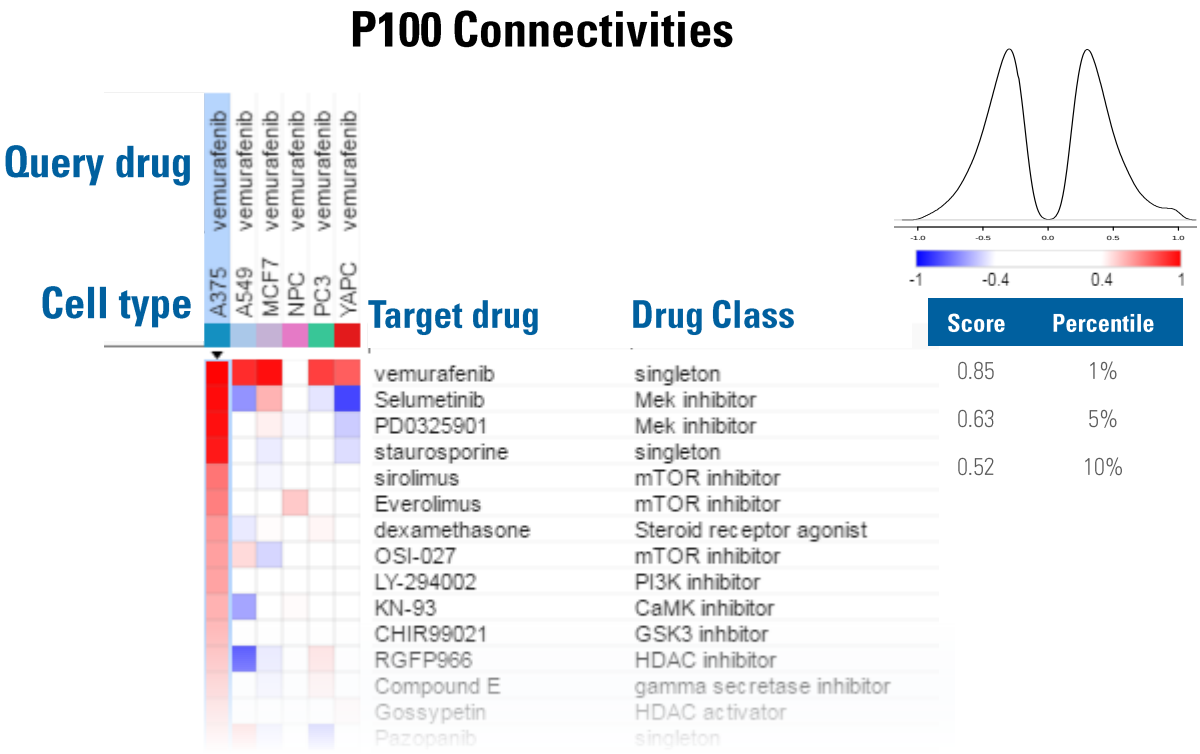


The **connectivity score** is a signed value of the KS-test

Vignette 1: What's a BRAF inhibitor to do when its target is not a dependency?

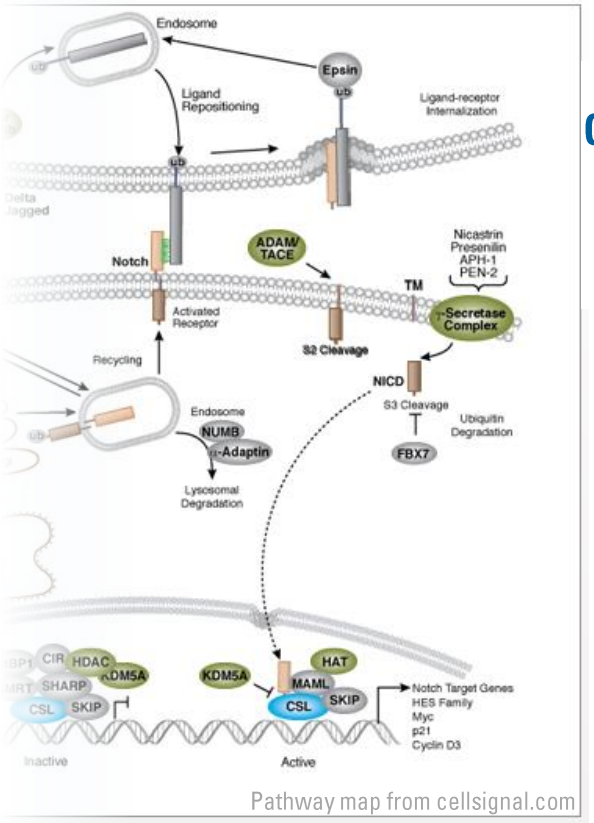


Pathway map from cellsignal.com



- **A375** has mutant BRAF (V600E) and is dependent on it
- Vemurafenib is the top ranked connection ALL **cancer** lines
- **NPCs** do not seem to respond at all
- *So what does it do in the other cell lines?*

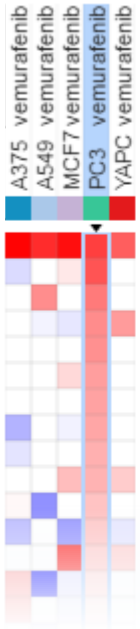
Vignette 1: What's a BRAF inhibitor to do when its target is not a dependency?

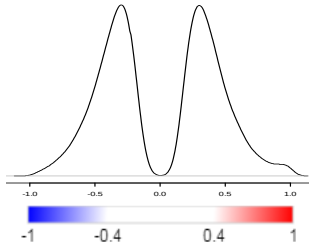


P100 Connectivities

Query drug

Cell type

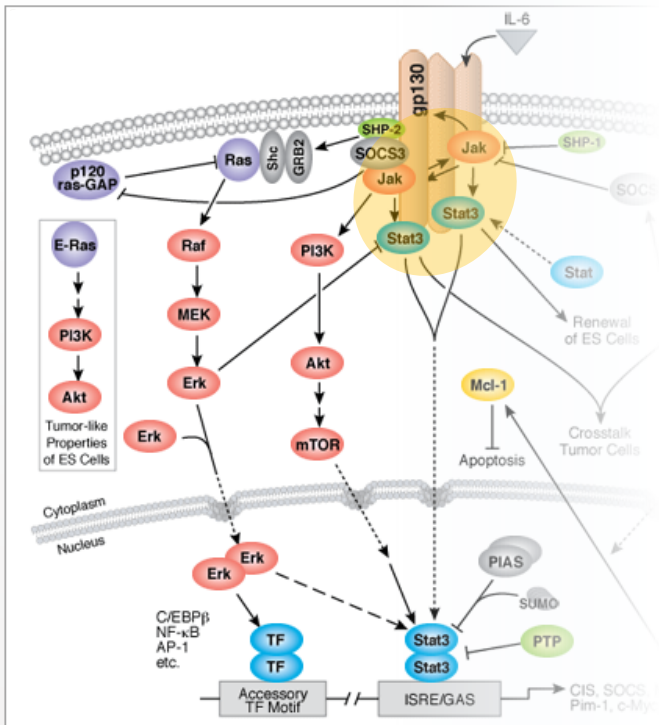




Target drug	Drug Class	Score	Percentile
vemurafenib	singleton	0.85	1%
BMS-906024	Notch pathway inhibitor *	0.63	5%
Ginkgetin	HDAC activator *	0.52	10%
1271738-62-5	Acetyltransferase inhibitor *		
calpain inhibitor II	singleton		
Tofacitinib	Jak/Stat inhibitor		
UNC1215	singleton		
IPI145	PI3K inhibitor		
RO4929097	Notch pathway inhibitor *		
Verteporfin	singleton		
Olaparib	singleton		
Rolipram	singleton		
Pravastatin	Jak/Stat inhibitor		
RGFP966	HDAC inhibitor		
Compound E	gamma secretase inhibitor		

- Does vemurafenib inhibit the **Notch** pathway in **PC3**?

Vignette 1: What's a BRAF inhibitor to do when its target is not a dependency?

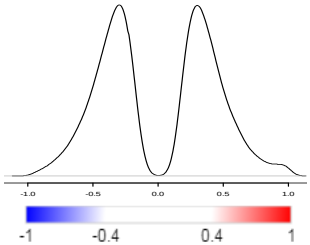


Pathway map from cellsignal.com

P100 Connectivities

Query drug

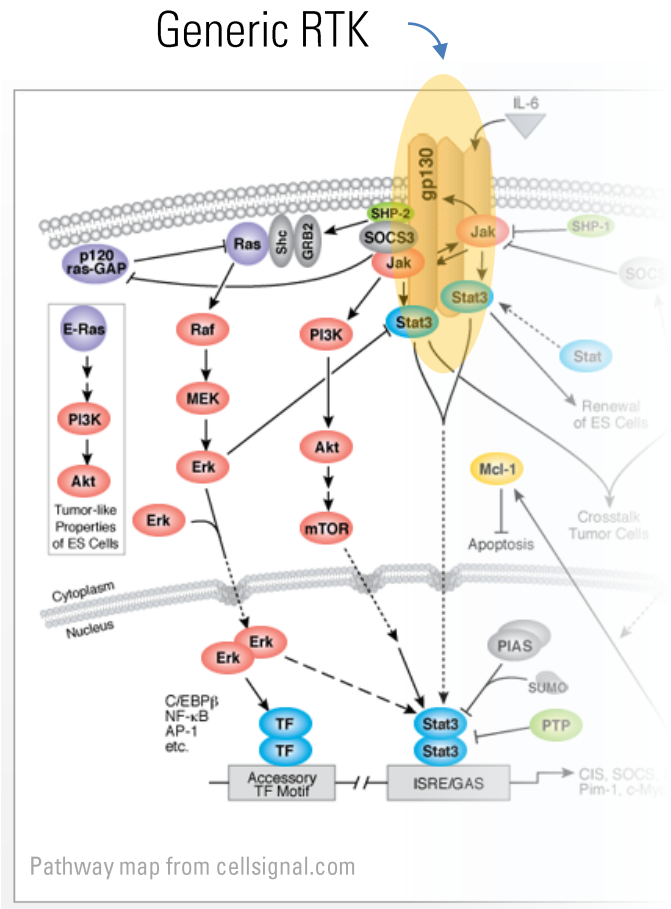
Cell type



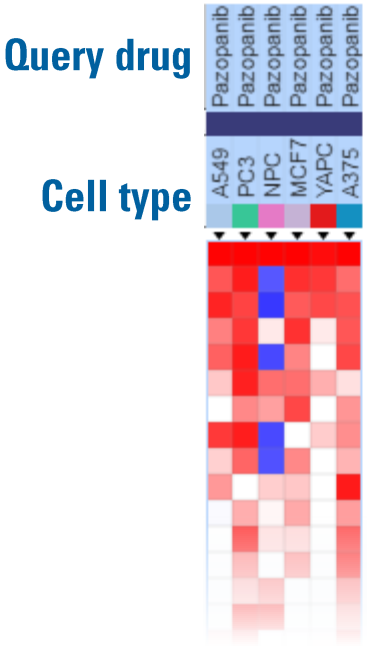
Target drug	Drug Class	Score	Percentile
vemurafenib	singleton	0.85	1%
JQ1-S	BRD inhibitor *		
Pravastatin	Jak/Stat inhibitor	0.63	5%
GSK525762A	BRD inhibitor *		
tacrolimus	singleton	0.52	10%
Selumetinib	Mek inhibitor		
GSK1210151A	BRD inhibitor *		
Verteporfin	singleton		
ruxolitinib	Jak/Stat inhibitor		
decitabine	DNA methyltransferase inhibit		
KU-55933	Cell cycle inhibitor		
Etoposide	singleton		
Tofacitinib	Jak/Stat inhibitor		
GSK-J4	Lysine demethylase inhibitor		
BMS-906024	Notch pathway inhibitor		

- Connections to **BRD inhibitors** seem enriched in **MCF7**?
- Many lines seem to have some residual effects in Jak/Stat, which is ~upstream of RAF

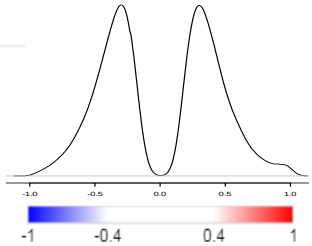
Vignette 2: The strange case of Pazopanib in NPCs



P100 Connectivities



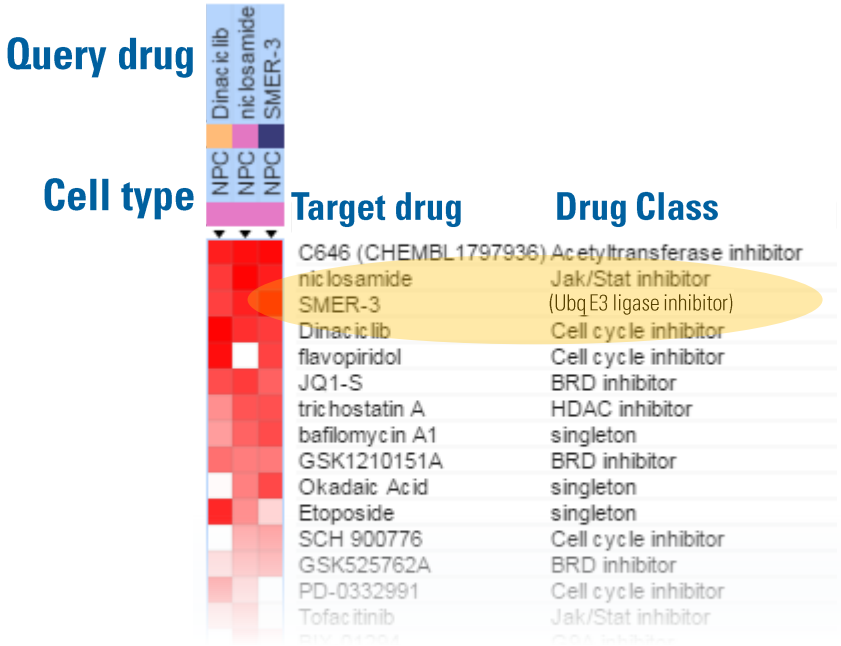
Target drug	Drug Class	Score	Percentile
Pazopanib	singleton	0.85	1%
SMER-3	(Ubq E3 ligase inhibitor)	0.63	5%
niclosamide	Jak/Stat inhibitor	0.52	10%
Resveratrol	Sirtuin inhibitor		
Dinaciclib	Cell cycle inhibitor		
Nilotinib	(Bcr-Abl inhibitor)		
SP600125	Jnk inhibitor		
C646 (CHEMBL1797936)	Acetyltransferase inhibitor		
SCH 900776	Cell cycle inhibitor		
dactolisib	PI3K inhibitor		
momelotinib	Jak/Stat inhibitor		
Roscovitine	Cell cycle inhibitor		
OSI-027	mTOR inhibitor		
LY-294002	PI3K inhibitor		
staurosporine	singleton		



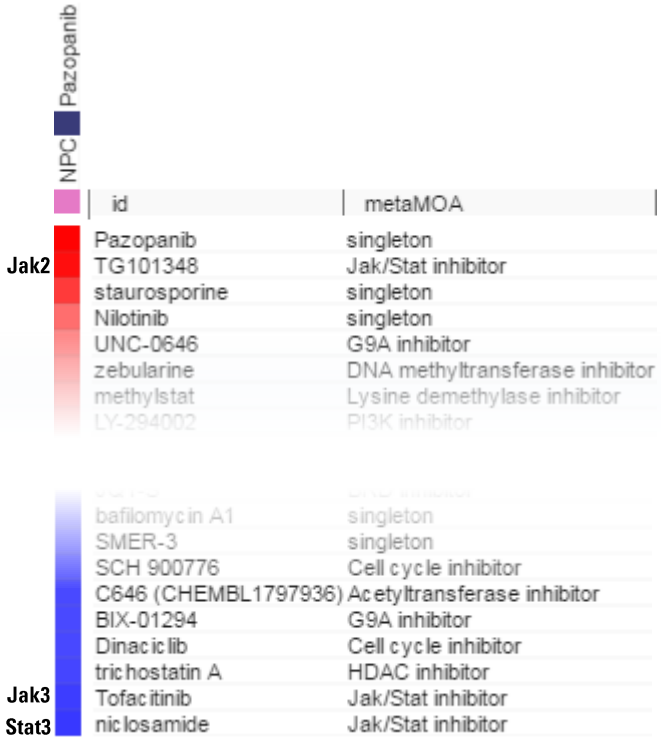
- Pazopanib is the top ranked connection **all lines**
- SMER3, niclosamide, and dinaciclib are strong (+) connections in **cancer** lines
- However, these 3 are **strong (-) connections in NPCs**.

Vignette 2: The strange case of Pazopanib in NPCs

Connectivity of Pazopanib Cancer Connections in NPCs

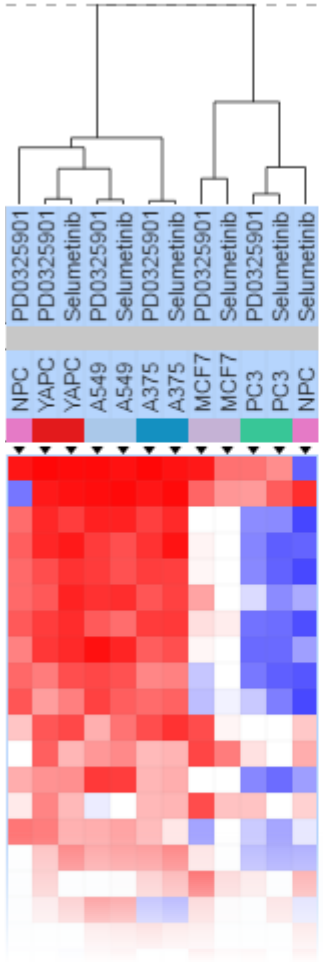


(+) and (-) Connectivity of Pazopanib in NPCs



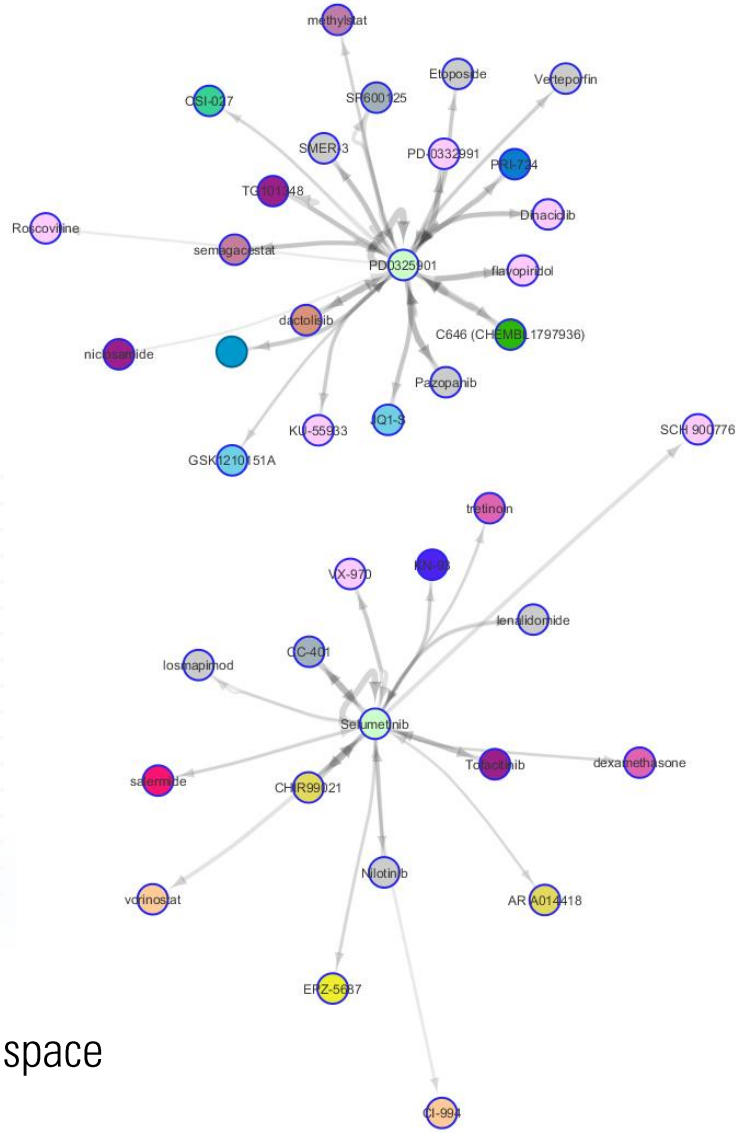
- But all of these things are strongly connected in NPCs, just not to pazopanib
- Perhaps Pazopanib has **Jak2** selectivity in NPCs? Or NPCs favor Jak2 signaling?

Vignette 3: Chromatin connections of MEK inhibitors?



GCP Connectivities

Target drug	Drug Class
PD0325901	Mek inhibitor
Selumetinib	Mek inhibitor
TG101348	Jak/Stat inhibitor
PD-0332991	Cell cycle inhibitor
Dinaciclib	Cell cycle inhibitor
Resveratrol	Sirtuin inhibitor
flavopiridol	Cell cycle inhibitor
Etoposide	singleton
SMER-3	singleton
C646 (CHEMBL1797936)	Acetyltransferase inhibitor
trichostatin A	HDAC inhibitor
MS-275	HDAC inhibitor
niclosamide	Jak/Stat inhibitor
belinostat	HDAC inhibitor
KU-55933	Cell cycle inhibitor
staurosporine	singleton
CI-994	HDAC inhibitor
geldanamycin	singleton
RGFP966	HDAC inhibitor



NPC Connectivity networks

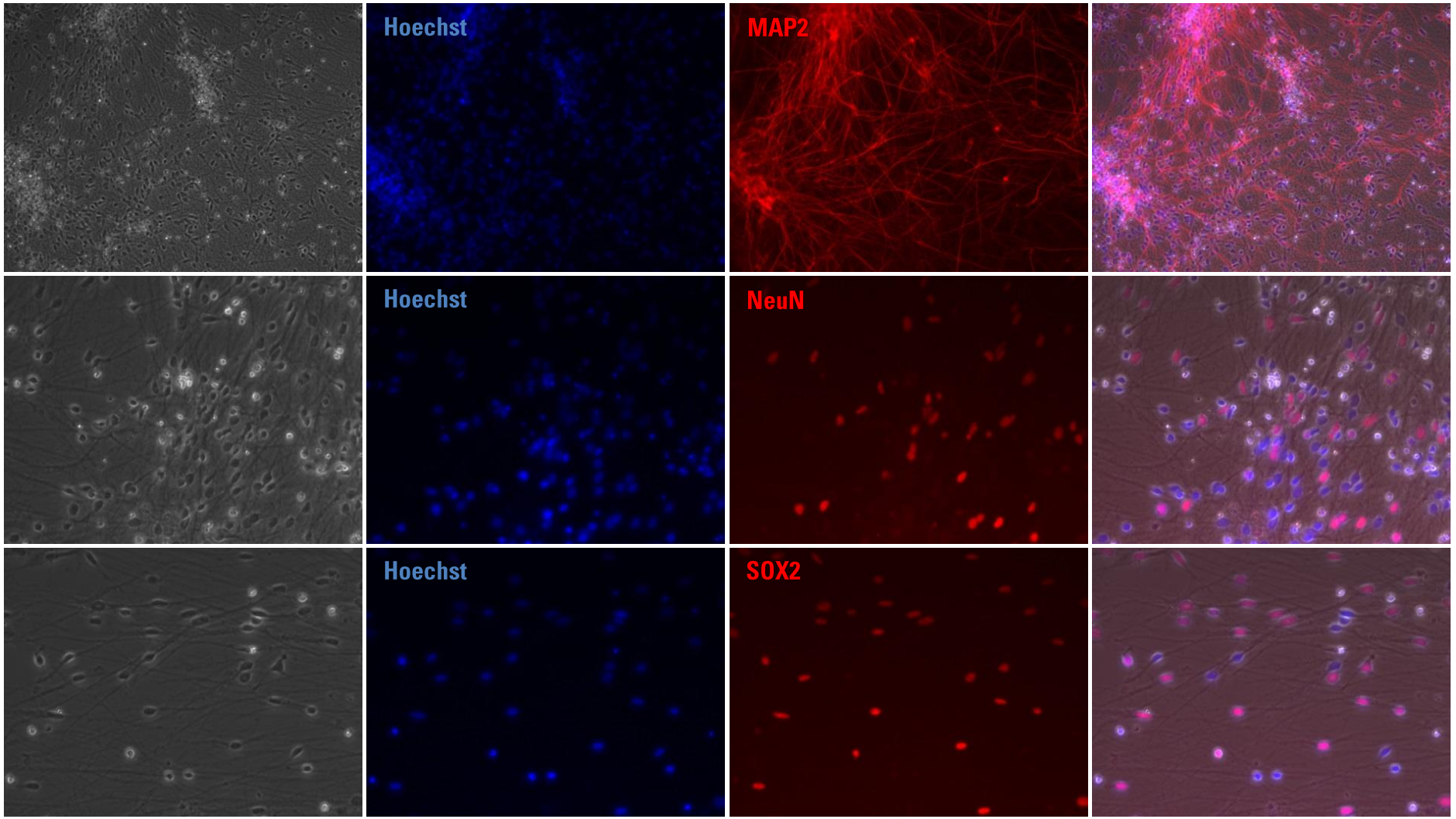
- Mek inhibitors have a “recallable” signature in chromatin space
- 3 basic classes of response as derived from connections
  - Connections to cell cycle drugs suggest arrest, neutrality, and pro-proliferation
  - Again, NPCs show interesting biology with differential connections of PD0325901 and Selumetinib

Preview of the coming year

- More analysis of data set and apps for interaction with the data
  - See poster from Lev Litichevsky and Ryan Peckner
- CRISPR/Cas9-based gene disruptions in multiple cell types
  - Already in progress
- Profiling of drug effects in ESC-derived neurons
  - The ESC system has inducible Cas9 which will allow for further manipulations
  - See poster from Jennie Young, Joel Blanchard, and Fatema Abdurrob
- Deeper mining of comprehensive MS data
  - See poster from Jarrett Egertson

Creating neurons for high-content proteomics assays

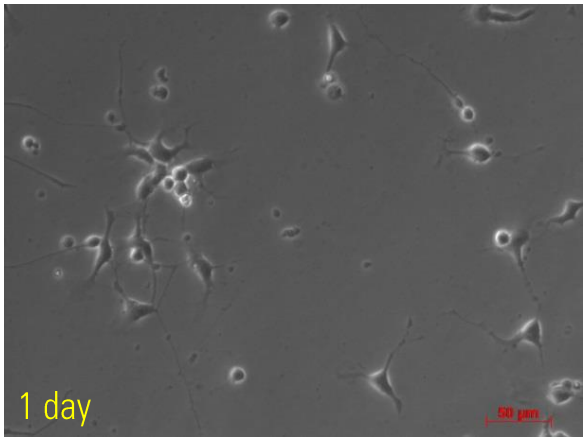
Passive differentiation via growth factor withdrawal (GFW)

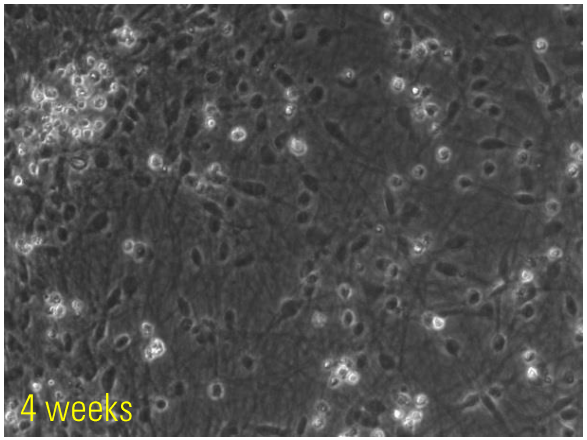


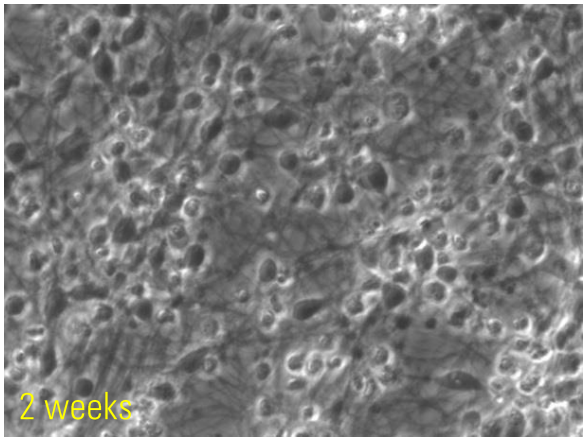
4 weeks

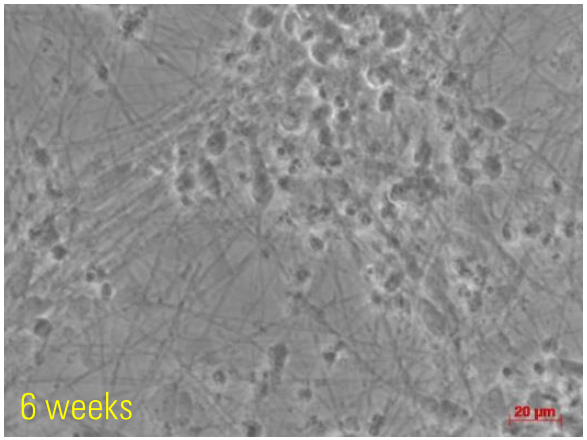
Creating neurons for high-content proteomics assays

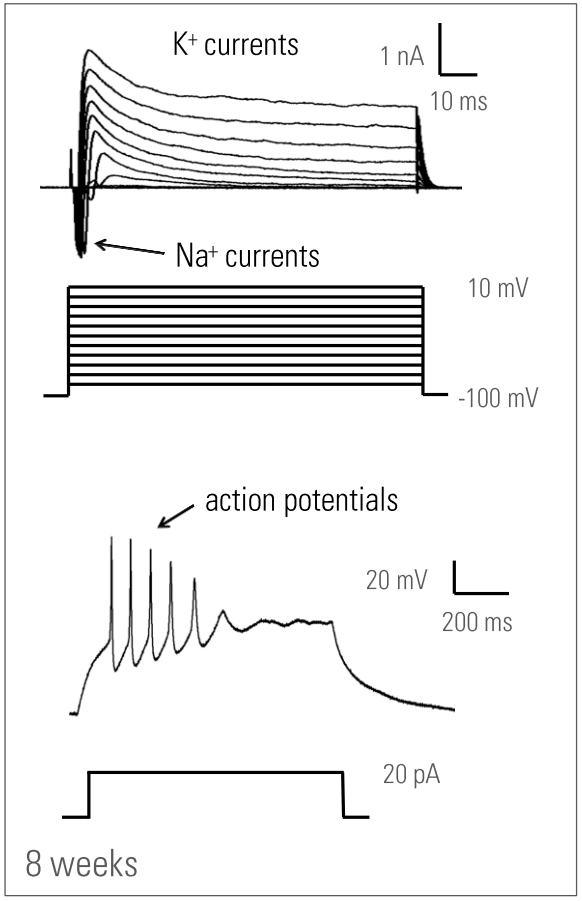
Passive differentiation via growth factor withdrawal (GFW)









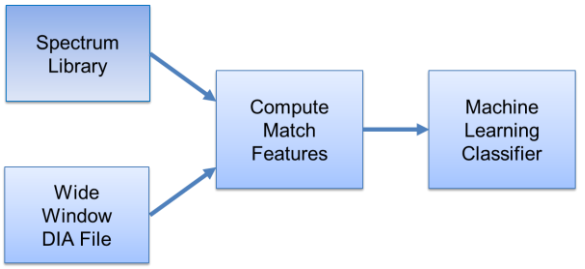


Next-gen comprehensive MS will extend the P100

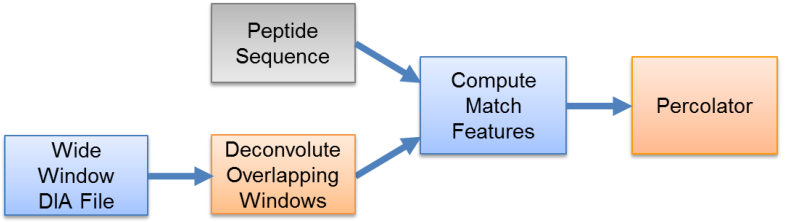
- P100 data have been collected using a “DIA” MS method
  - Pioneered by MacCoss Lab
- DIA has the potential to identify and quantify 1000s of phosphosites in our data
- A key challenge is developing the algorithms to “unlock” these data

Two innovative tools for comprehensive MS analysis

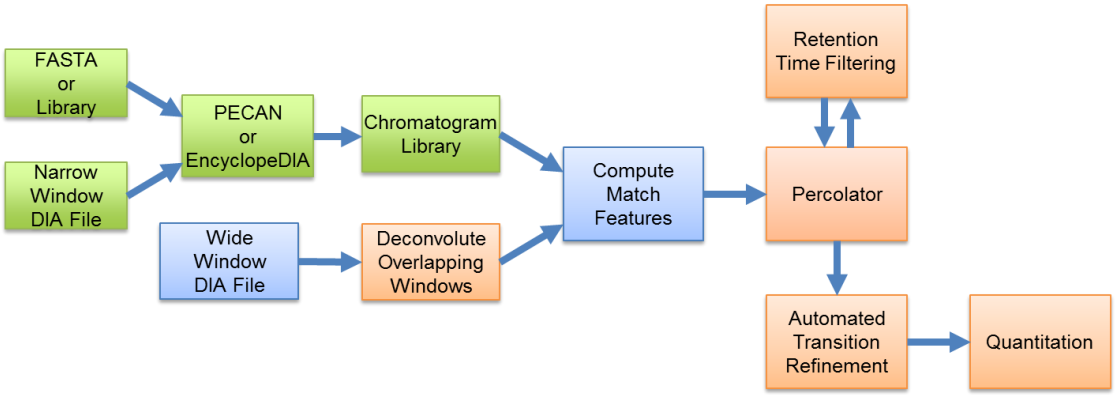
- Typical Workflow



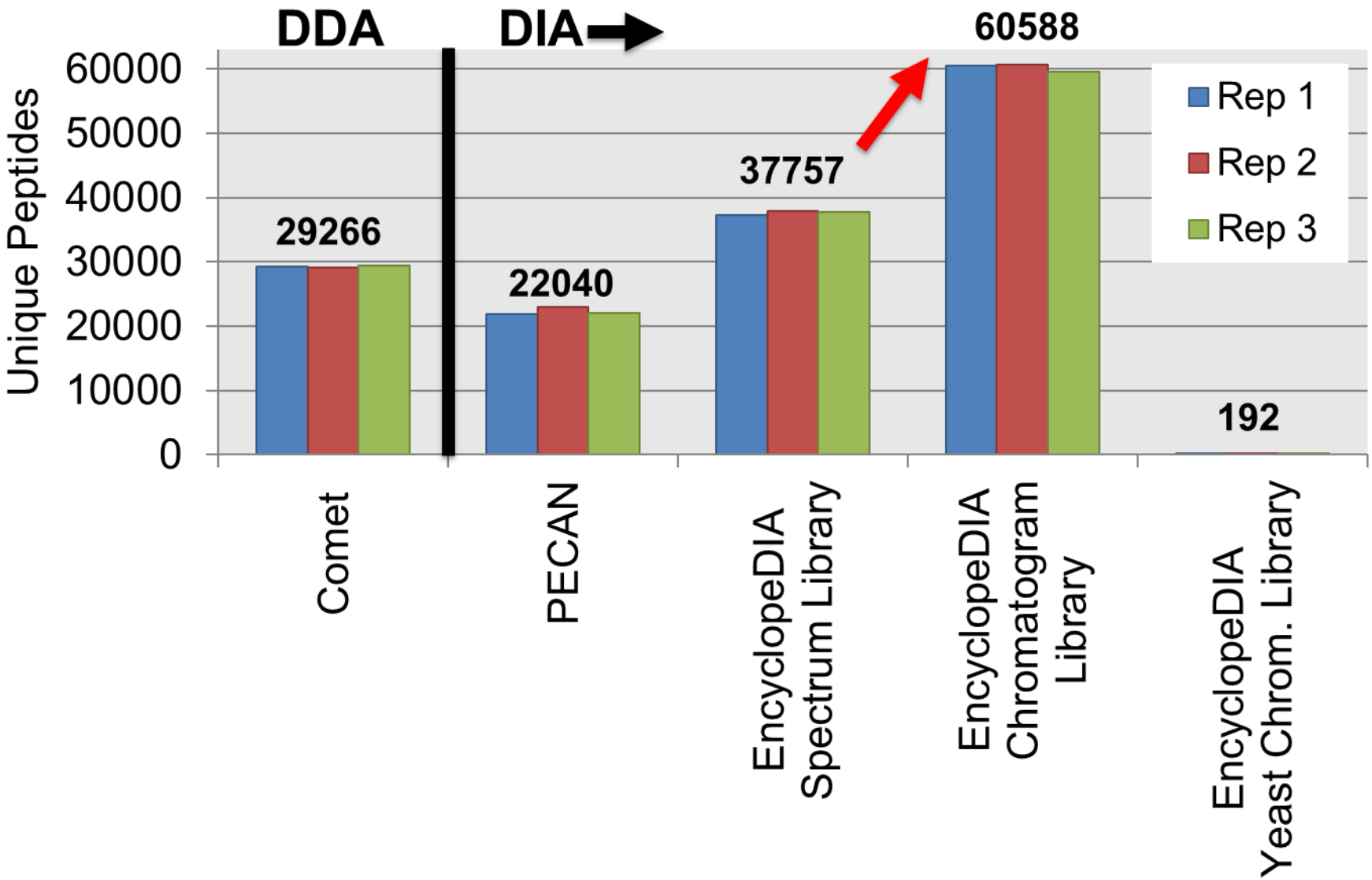
- Tool 1: Pecan
  - Uses only peptide sequences
  - No spectral or retention time information



- Tool 2: EncyclopeDIA
  - Uses peptide sequences
  - Spectrum or Chromatogram libraries
  - Positional isomers



Pioneering methods meeting or exceeding goals



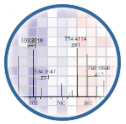
PCCSE Outreach and Intra-LINCS activities

- Signatures of cardiovascular hypertension induced by chemotherapeutic drugs
  - Molecular Cardiology Research Institute, Tufts Medical Center
  - See poster from Srila Gopal
- Signatures of neuropsychiatric phenotypes and their responses to drugs in patient-isolated iPS-derived NPCs and neurons
  - Massachusetts General Hospital
- Mapping of P100 probe-sets onto known pathway networks
  - Georgetown University and University of Delaware
  - Together with DCIC through LINCS EDSR
- Substrates to promote neuronal fates and phenotypes
  - Intra-LINCS with OHSU MEP LINCS Center

PCCSE Overall Summary

- Great progress has been made in the first two years in establishing our center
  - Vignettes illustrate the promise of our data
- Our neurobiology models are progressing nicely and we are excited to extend these further
- Next-Gen MS holds great promise for increasing the impact of our work
- We are poised to use our data "as is" for comprehensive connectivity analysis, but also as a springboard for comparison with new data to be made via:
  - Our continued efforts
  - Our outreach and collaboration efforts
  - Data made by third parties

Acknowledgments



Broad Institute – Jaffe Lab and Proteomics Platform

Amanda Creech
Shawn Egri
Adam Officer
Malvina Papanastasiou
Ryan Peckner
Sebastian Vaca
Janice Williamson
Steve Carr

University of Washington – MacCoss Lab

Jarrett Egertson
Brendan MacClean
Brian Searle
Vagisha Sharma
Sonia Ting

MIT – Tsai Lab & Stem Cell Core

Fatema Abdurrob
Tak Ko
Joel Blanchard
Jennie Young

Broad Institute – LINCS Common Core

Desiree Davison
David Lahr
Daniel Lam
Lev Litichevskiy
Xiaodong Lu
Ted Natoli
Mukta Bagul
Aravind Subramanian
Todd Golub

Key Collaboration Partners

Iris Jaffe – Tufts Medical Center
Srila Gopal – Tufts Medical Center
Stephen Haggarty – Mass General Hospital