

From Whole Genome Sequencing toward precision medicine: A preliminary case study in EGFR



Wanlin Zheng
Josh Bloomstein
Eric Danziger
Stephanie Kinnunen

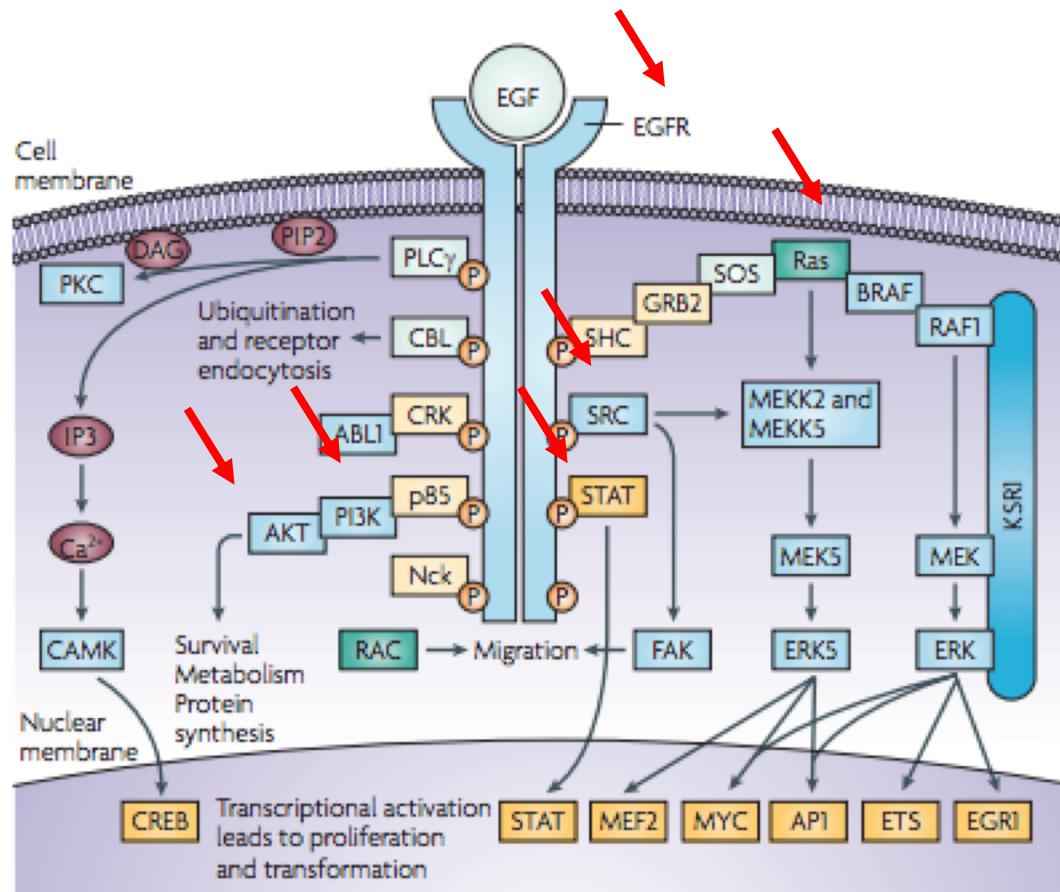


Team: HIF1a is not an oncogene

Identification of mutations using our patient dataset SNP-FF

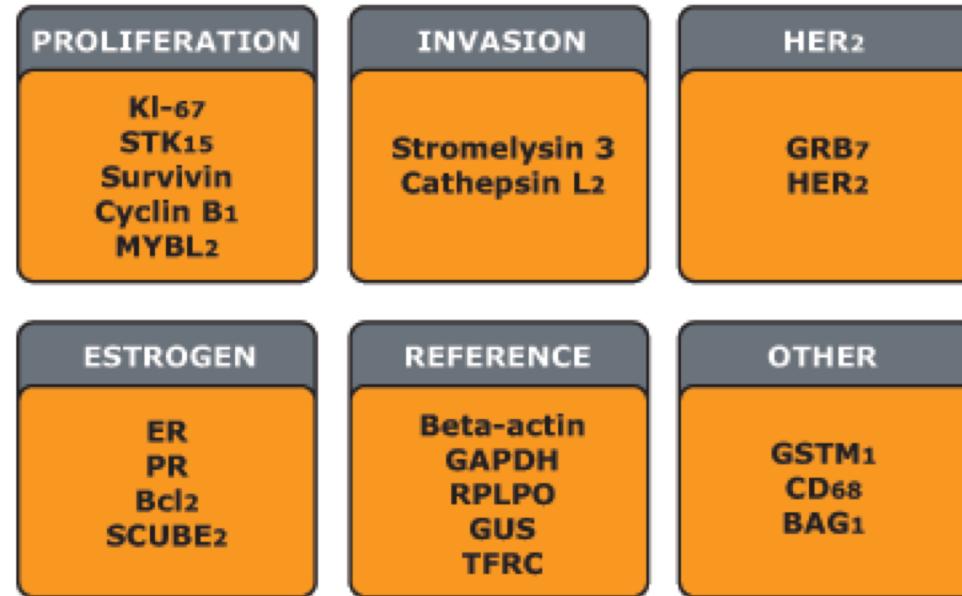
RBPJ	CSF1	MAPK9	EGFR	DTX4	TERT	FANCF
NUMB	HIF1A	STAT2	TSC1	SLC19A1	STAT5B	ATM
PTK2	MAP3K3	DLL3	CDK2	PIK3R2	MAPK1	KDR
PSEN1	NF1	MTOR	MMP12	CSF1R	MDM4	CCND1
CHEK2	PIK3R1	BRCA2	HEYL	NOTCH2	RBL2	MMP25
NCOR2	DPH1	ITGB3	KRAS	JAG2	E2F1	
FGFR1	E2F4	ITCH	TP53BP1	PARP1	MMP21	
NCSTN	AKT3	RPS6KB2	MAML3	MLH1	MMP24	
ERBB2	ADAM10	MMP2	MYC	MAP2K2	MMP16	
ARRDC1	CCNB1	MMP17	MAP3K5	IGF1R	KDM5A	
MLST8	MAPK4	SNW1	ATR	E2F2	MAPK6	
SRC	OPCML	CDH1	RBL1	MAPK12	CDK4	
MAPK14	RPS6KA2	MMP14	STAT5A	JAG1	MAP3K1	
MAP3K4	MMP28	MAPK8	RICTOR	MAML2	MMP9	
AKT1	IGF1	TSC2	MMP19	MMP7	HDAC1	

Multiple mutations in EGFR signaling pathway



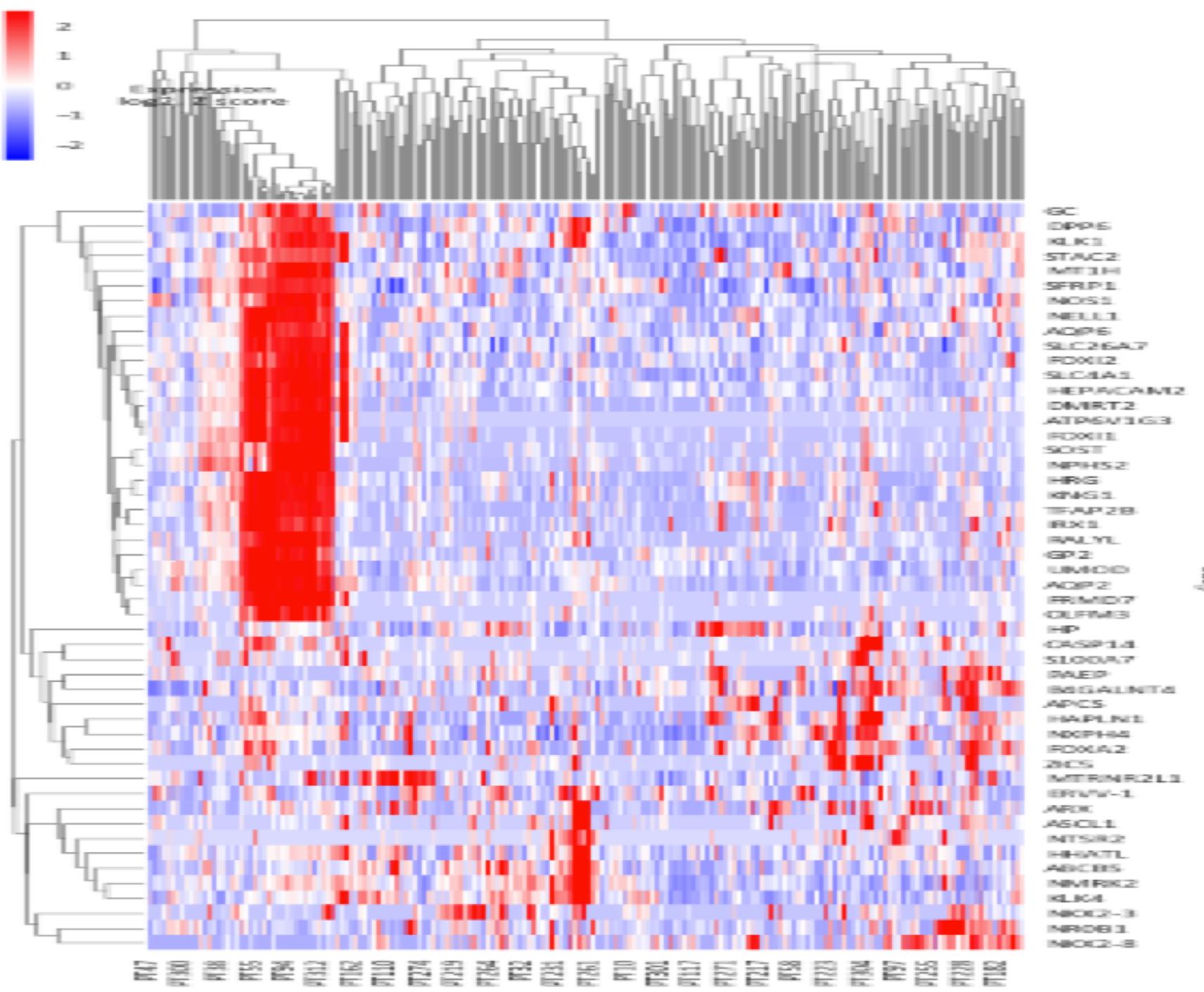
Functional Proteomics to dissect tyrosine kinase signalling pathways in cancer, Kolch et al., Nature Reviews Cancer, 2010.

Oncotype DX 21 gene breast cancer panel



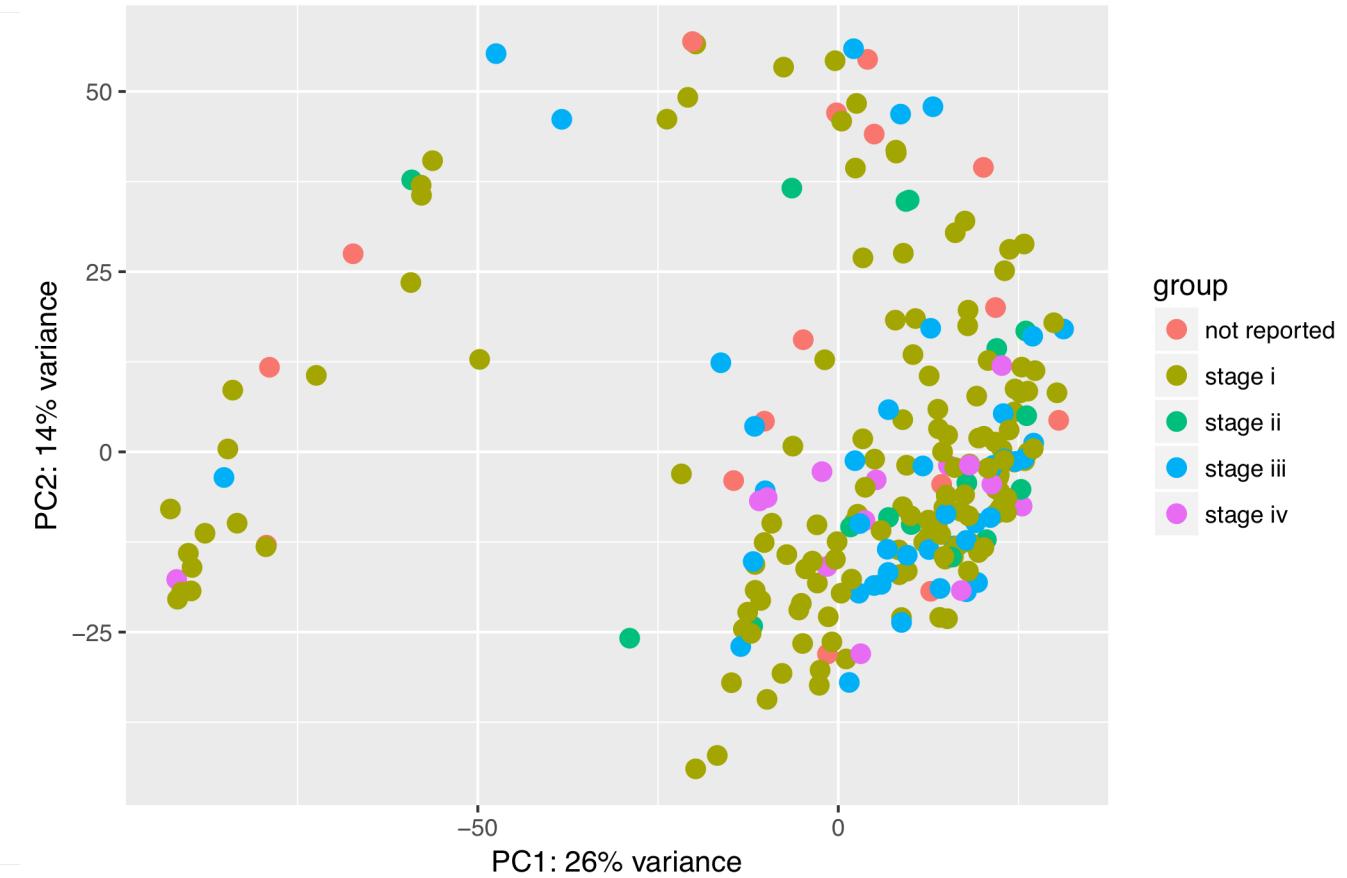
HER2 RT-PCR

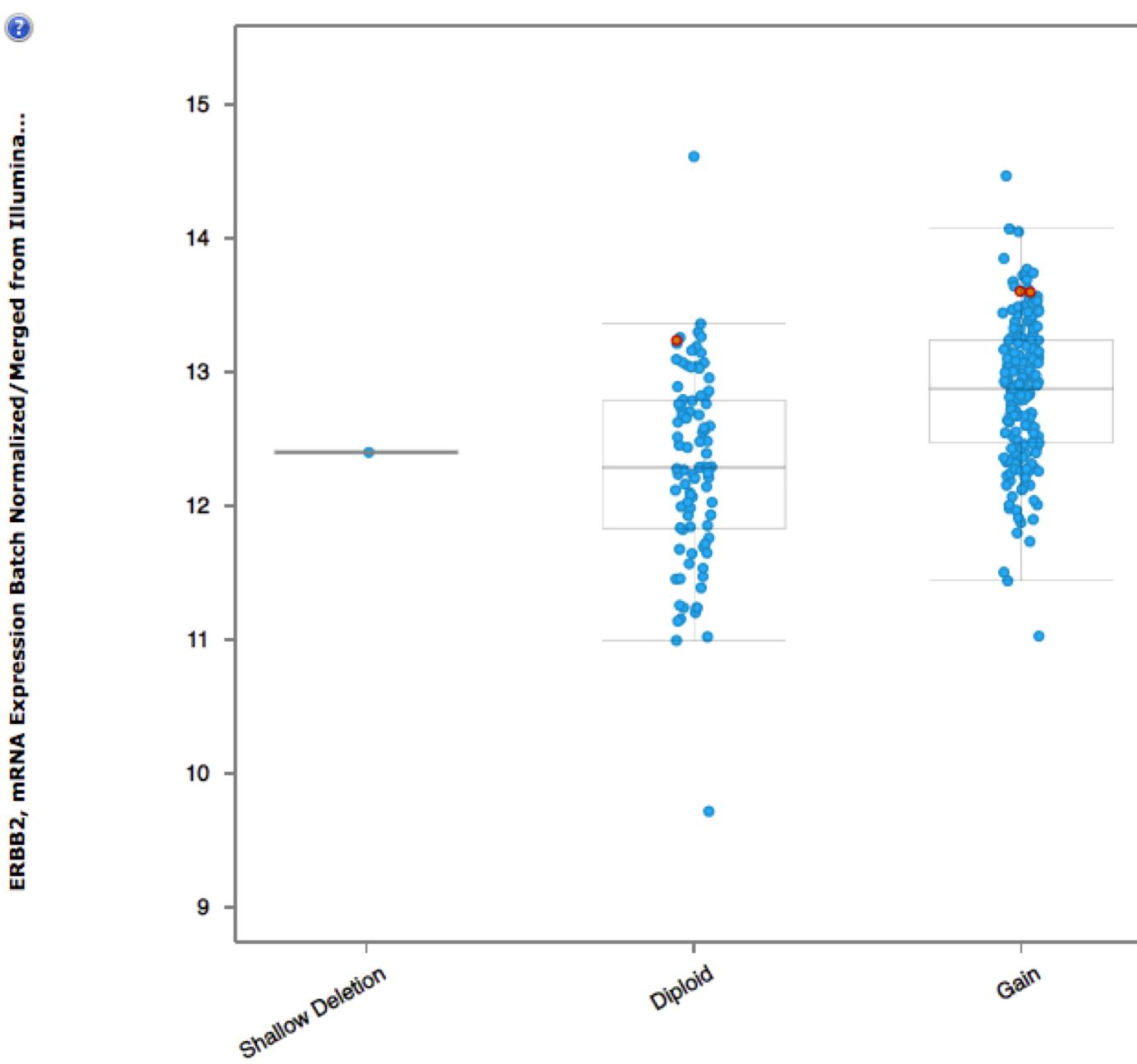
Small gene panels may not include mutation information on driver genes



Hierarchical clustering of TCGA KIRP yields subtypes

TCGA All genes Principal component analysis





TCGA KIRP Copy Number
Variation gain may be
associated with higher mRNA
expression of HER2 (ERBB2)

TCGA KIRP RNA-Seq Cox Proportional Hazard Model

Surv(Death, Followup days) ~ Gene

Gene	Beta	p value
CDK4	-0.000423558	0.01192192
E2F2	-0.006678589	0.046164402
FGFR1	-0.000182627	0.007001034
MAPK8	0.000957689	0.043868428
MSH2	-0.001189108	0.020838839
PARP1	-0.000336744	0.003993623
STAT5B	0.000228466	0.047300526

Future directions

Creating a pathway differential expression score to apply to patients

Selected 5 gene sets from Broad Hallmarks database, within categories that patient has mutation

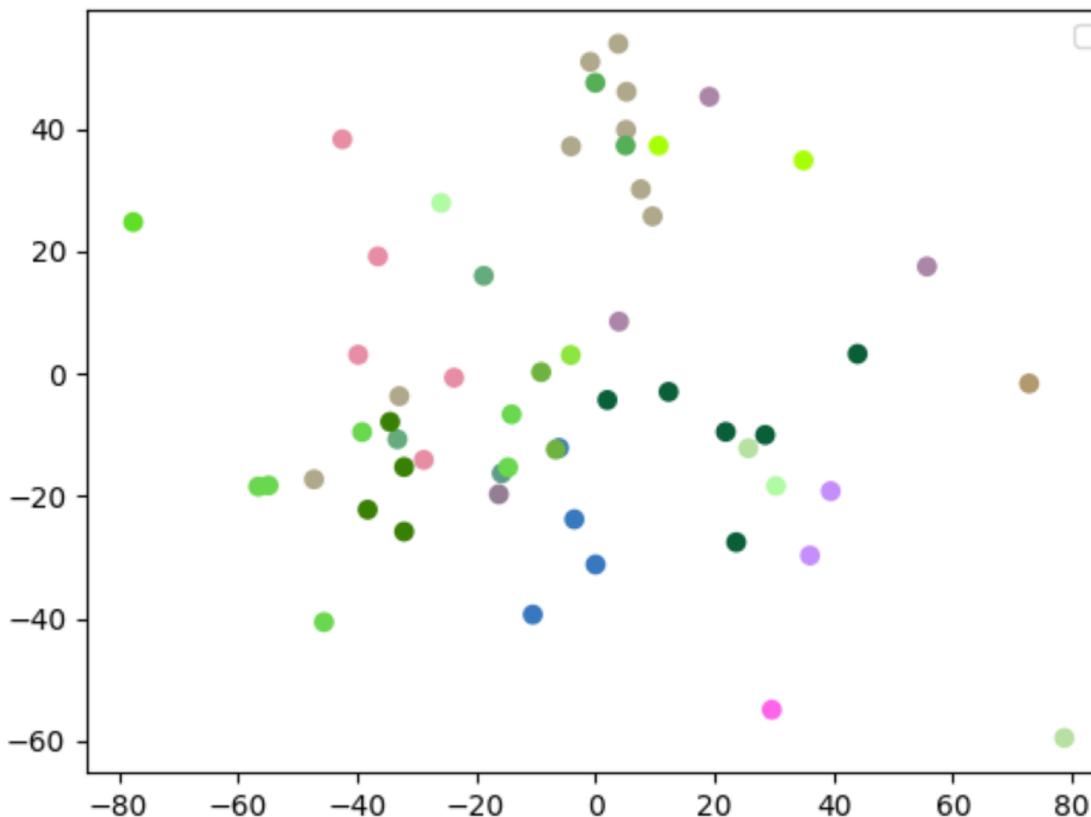
Pathways:

1. IL-6-JAK-STAT3-signaling
2. KRAS signaling up
3. p53 pathway
4. Hypoxia signaling
5. PI3K Akt mTOR signaling

$$\sum_{i=1}^j \frac{GeneExpression_i - \overline{GeneExpression}}{SD}$$

Principal Component Analysis of TCGA KIRP patients by pathway activation score

Low IL-6



Colors represent gross activation of
cancer-related pathways

Each group is a
combination of high/low
of 5 pathway sets

Use feature selection and machine learning to models for predicting clinical outcomes in KIRP