Cancer proteogenomics in evolution: Assessing targets, therapy and resistance

Michael A Gillette, MD, PhD

Broad Institute of MIT and Harvard

Cancer and Evolution Symposium

14 October, 2020

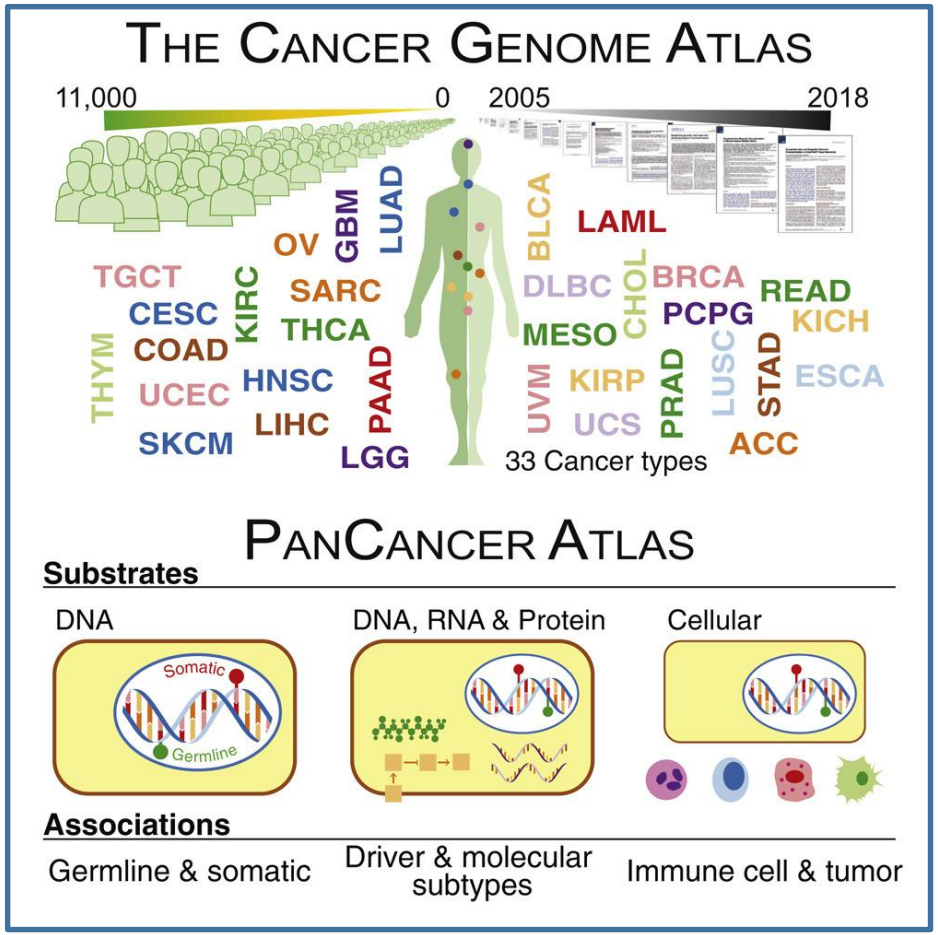
CPTAC CLINICAL PROTEOMIC TUMOR ANALYSIS CONSORTIUM

NIH NATIONAL CANCER INSTITUTE Office of Cancer Clinical Proteomics Research

Proteogenomics is an important tool in the campaign to analyze cancer from an evolutionary perspective

- An evolutionary perspective on cancer requires
  - Awareness of disease taxonomy
  - Knowledge of the environment
    - Tumor / microenvironment interactions
    - Special interest in the immune landscape
  - Detailed understanding of the biological repertoire of cancer
    - Sensitivity to individual differences is essential for personalized / precision medicine
  - Assessment of the effect of specific selective pressures
    - Direct analyses under therapeutic perturbations can give critical insights into response and resistance
- Like evolution in general, the unit of evolution can be variously defined
  - Cancer subtype, individual tumor, subclone, cell
  - Proteomic techniques and instrumentation are rapidly evolving; however, single cell global proteomics and especially PTM analyses are currently aspirational
- Full definitions of cancer taxonomy, tumor microenvironment and biological repertoire are fostered by comprehensive molecular characterization
- Determining impact of selective pressures requires approaches that can be serially applied, ideally in a clinical setting, as tumors evolve under perturbation

The Cancer Genome Atlas (TCGA) illuminated the cancer genome… but coverage of the proteome was sparse



11,000 cancers ~ 33 cancer types

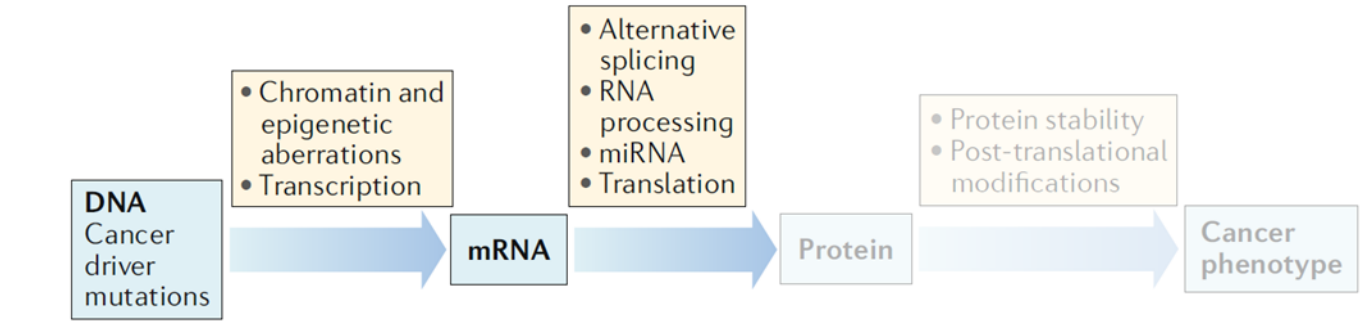
RPPA: 181 Abs; ~ 130 proteins / phosphosites

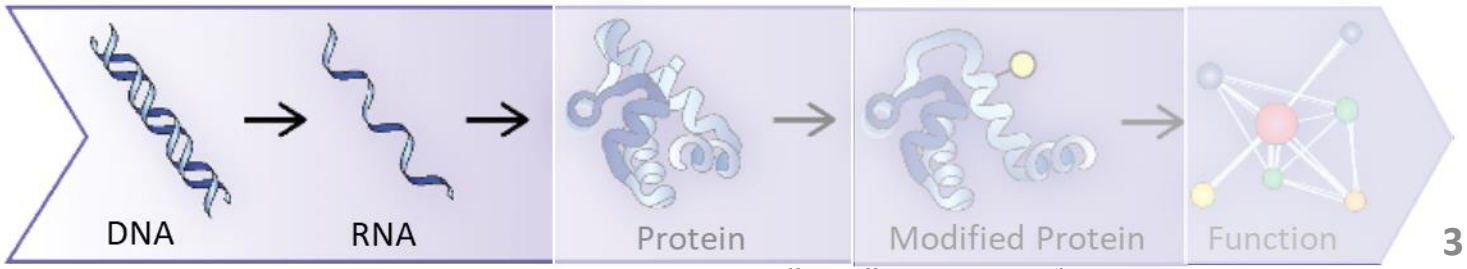
Ding et al, Cell, 2018

"Having a complete picture of every genomic change associated with each tumor can help us make personalized treatment decisions."

www.foundationmedicine.com

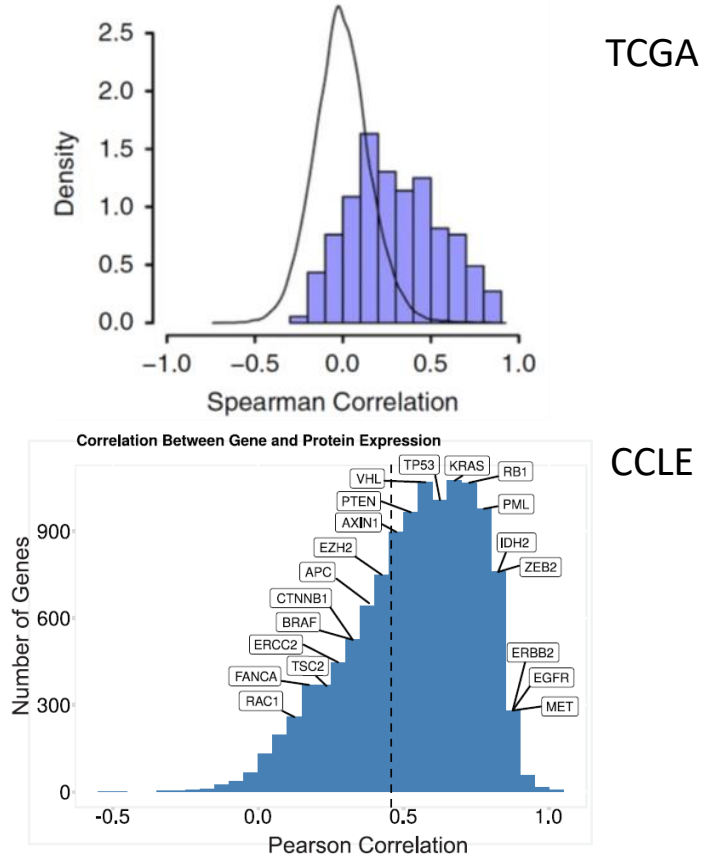
Many processes downstream of the genome can affect the tumor phenotype





Ellis, Gillette, Carr et al, Cancer Discovery, 2013

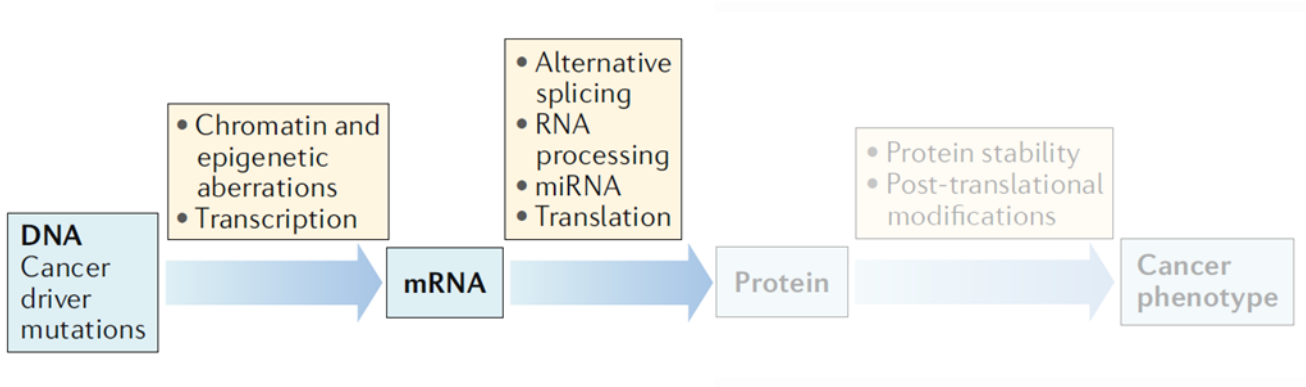
The Cancer Genome Atlas (TCGA) illuminated the cancer genome… but coverage of the proteome was sparse



"Having a complete picture of every genomic change associated with each tumor can help us make personalized treatment decisions."

www.foundationmedicine.com

Many processes downstream of the genome can affect the tumor phenotype



Akbani et al., *A Pan-cancer Proteomic Perspective on The Cancer Genome Atlas* Nat Comm 2014

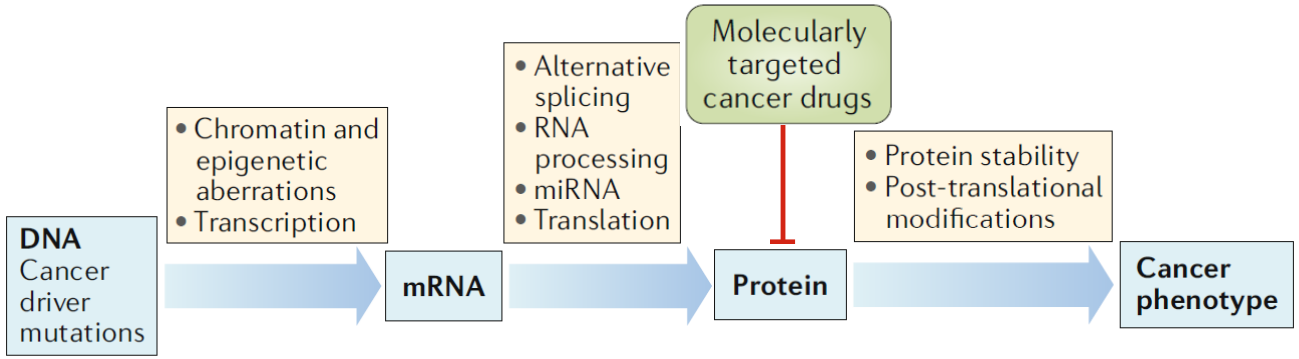
Nusinow et al., *Quantitative Proteomics of the Cancer Cell Line Encyclopedia* Cell 2020

Cancer proteogenomics supports integrated multi-omic analyses for more complete characterization of tumors and adjacent normal tissues

Use genomic, transcriptomic, and proteomic platforms simultaneously to gain a comprehensive understanding of human cancer in order to improve cancer diagnosis and treatment.

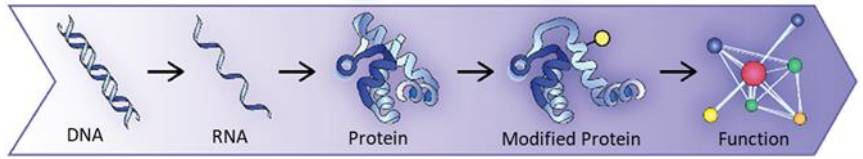
"Having a complete picture of every **molecular** change associated with each tumor can help us make personalized treatment decisions."

Many processes downstream of the genome can affect the tumor phenotype



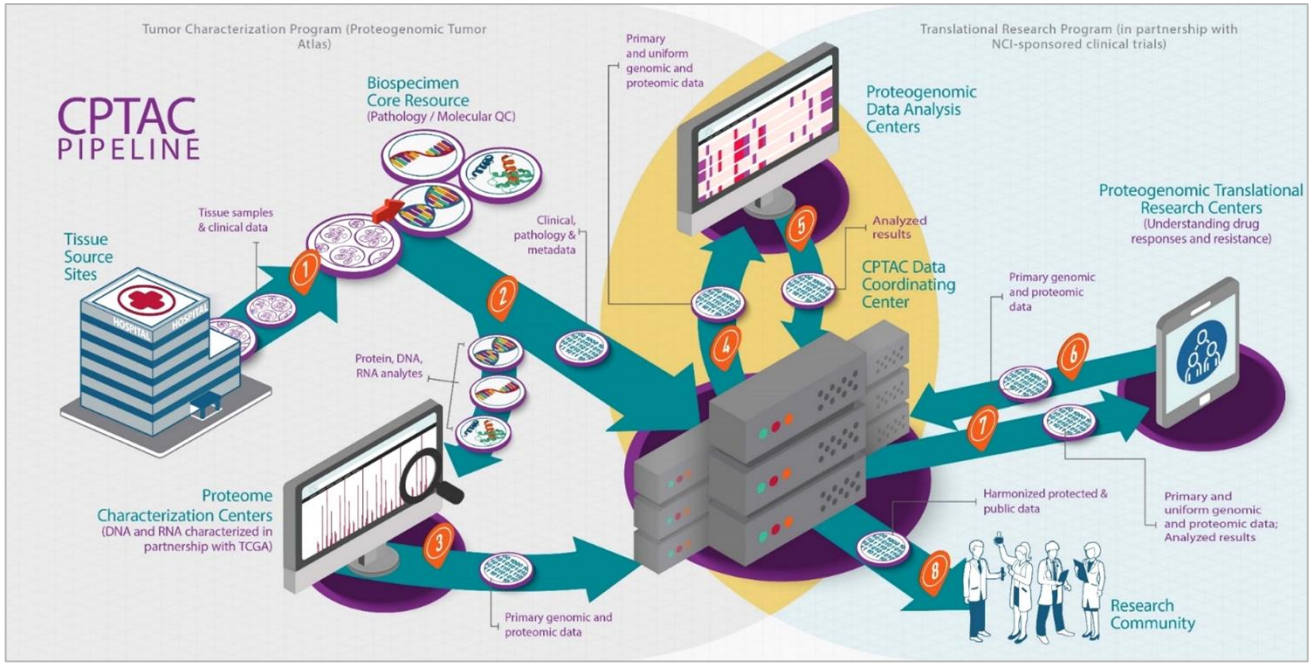
Zhang et al., *Clinical Potential of Mass spectrometry-based proteogenomics*
Nat Rev Clin Oncol, 2019

CLINICAL PROTEOGENOMICS TUMOR ANALYSIS CONSORTIUM



Goals

- Accelerate understanding of cancer biology
- Proteogenomically characterize tumors
- Produce public resources (data, assays, images, reagents) for hypothesis-driven science
- Support clinically relevant research projects



Achieved through

TUMOR CHARACTERIZATION

Proteome Characterization Centers

Proteogenomic Data Analysis Centers

TRANSLATIONAL RESEARCH

Proteogenomic Translational Research Centers

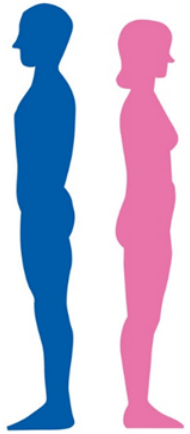
- pre-clinical and clinical trial samples
- Mechanisms, response, resistance, toxicity

Integrated research consortium that applies standardized comprehensive proteomics and genomics workflows, strict biospecimen collection protocols (optimized for genomics and proteomics) – ensuring rigor & reproducibility

Lung cancer is the leading cause for cancer-associated death in the US and worldwide
Among women, breast cancer leads incidence and is the second leading cause of death

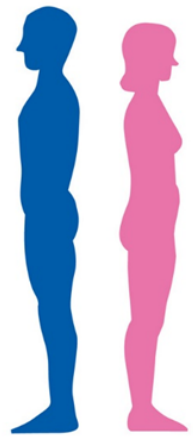
Estimated New Cases

Males			Females		
Prostate	191,930	21%	Breast	276,480	30%
Lung & bronchus	116,300	13%	Lung & bronchus	112,520	12%
Colon & rectum	78,300	9%	Colon & rectum	69,650	8%
Urinary bladder	62,100	7%	Uterine corpus	65,620	7%
Melanoma of the skin	60,190	7%	Thyroid	40,170	4%
Kidney & renal pelvis	45,520	5%	Melanoma of the skin	40,160	4%
Non-Hodgkin lymphoma	42,380	5%	Non-Hodgkin lymphoma	34,860	4%
Oral cavity & pharynx	38,380	4%	Kidney & renal pelvis	28,230	3%
Leukemia	35,470	4%	Pancreas	27,200	3%
Pancreas	30,400	3%	Leukemia	25,060	3%
All Sites	**893,660**	**100%**	**All Sites**	**912,930**	**100%**



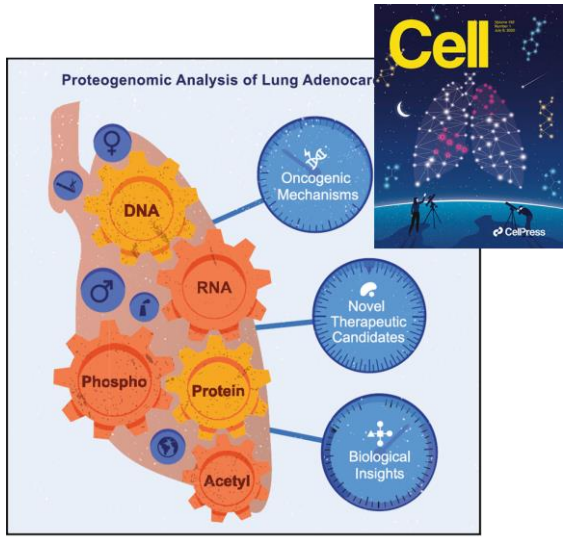
Estimated Deaths

Males			Females		
Lung & bronchus	72,500	23%	Lung & bronchus	63,220	22%
Prostate	33,330	10%	Breast	42,170	15%
Colon & rectum	28,630	9%	Colon & rectum	24,570	9%
Pancreas	24,640	8%	Pancreas	22,410	8%
Liver & intrahepatic bile duct	20,020	6%	Ovary	13,940	5%
Leukemia	13,420	4%	Uterine corpus	12,590	4%
Esophagus	13,100	4%	Liver & intrahepatic bile duct	10,140	4%
Urinary bladder	13,050	4%	Leukemia	9,680	3%
Non-Hodgkin lymphoma	11,460	4%	Non-Hodgkin lymphoma	8,480	3%
Brain & other nervous system	10,190	3%	Brain & other nervous system	7,830	3%
All Sites	**321,160**	**100%**	**All Sites**	**285,360**	**100%**

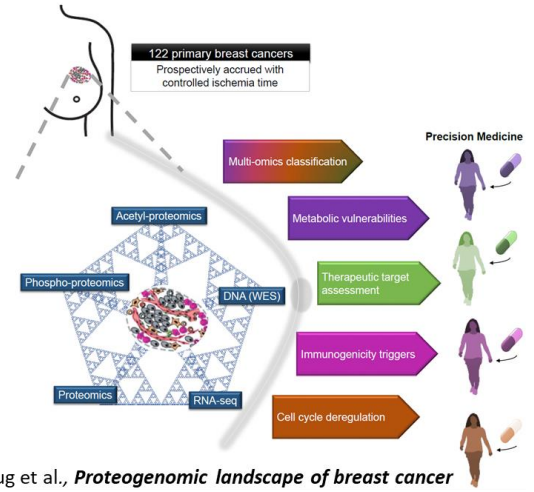


Bray CA Cancer Journal for Clinicians 2018

Siegel CA Cancer Journal for Clinicians, 2020



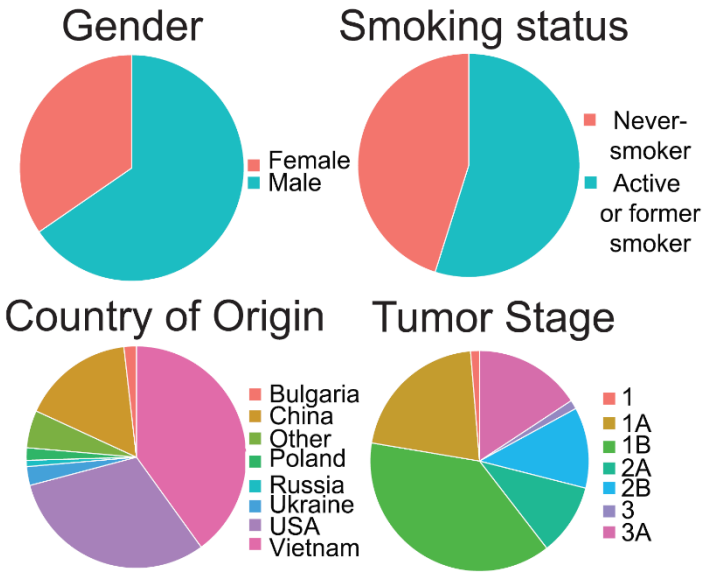
Gillette et al., ***Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma*** Cell 2020



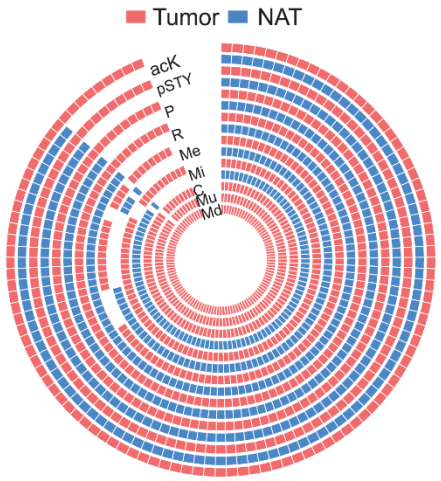
Krug et al., ***Proteogenomic landscape of breast cancer tumorigenesis and targeted therapy*** Cell 2020 (in press)

Microscaled proteogenomic methods for precision oncology Satpathy, et al. 2020

LUAD Discovery samples represent diverse country of origin, smoking status and stage
Genomics and proteomics profiles nearly complete for 110 LUADs & 101 NATs*

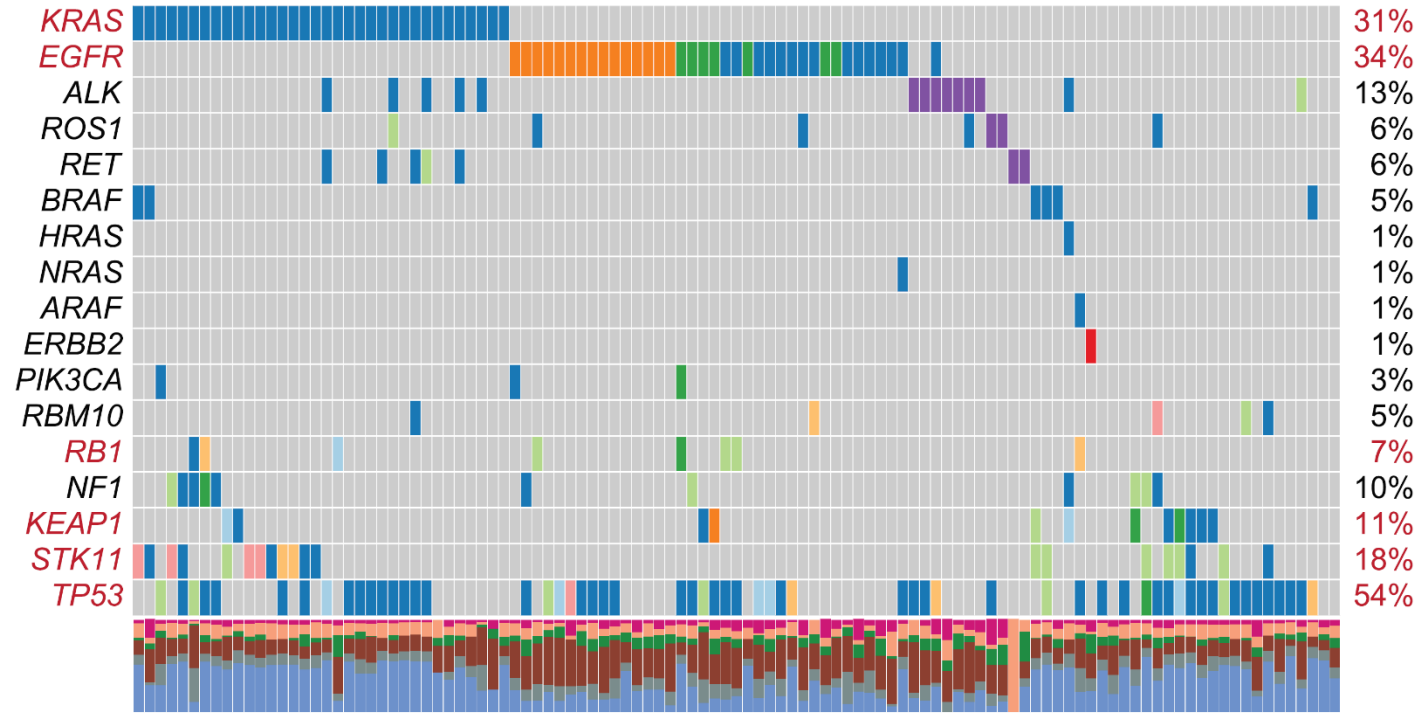


Platform	Data type	Features
Clinical metadata		15
WXS WGS	Germline mutations	16,660
	Somatic mutations	32,250
	CNA	19,267
Methylation array	DNA methylation	16,478
RNA-seq	mRNA	18,099
miRNA-seq	miRNA	2,585
TMT Proteomics	Proteins	10,699
	Phosphorylation	41,188
	Acetylation	6,906



Dataset (T/NAT)

acK (acetylome) 110/101
pSTY (phospho) 110/101
P (proteome) 110/101
R (RNA) 110/101
Me (Methylation) 100/87
Mi (miRNA) 107/100
C (CNA) 109
Mu (Mutation) 109
Md (Metadata) 110



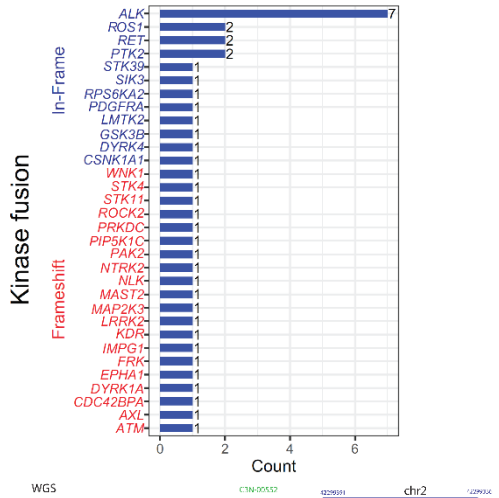
In_Frame_Del
Frame_Shift_Del
Frame_Shift_Ins
Multi_Hit
Fusion
In_Frame_Ins
Splice_Site
Nonsense_Mutation
No Somatic Mutation

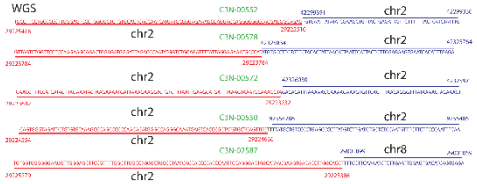
C>T C>G C>A T>A T>C T>G

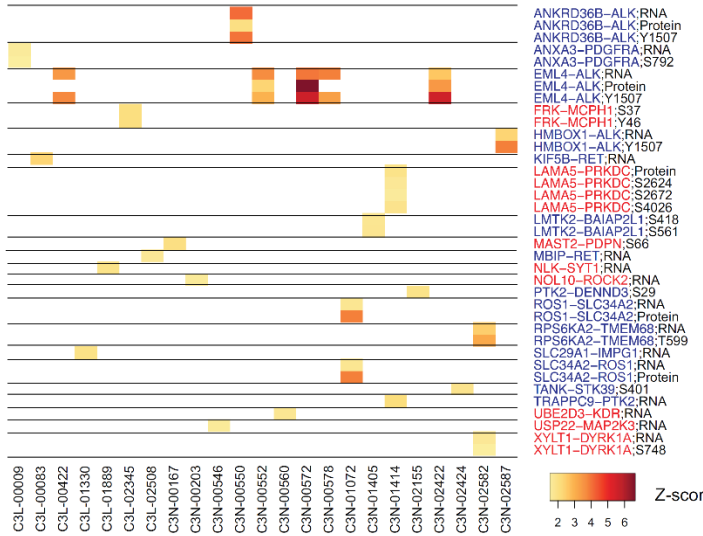
*NAT= Normal Adjacent Tissue

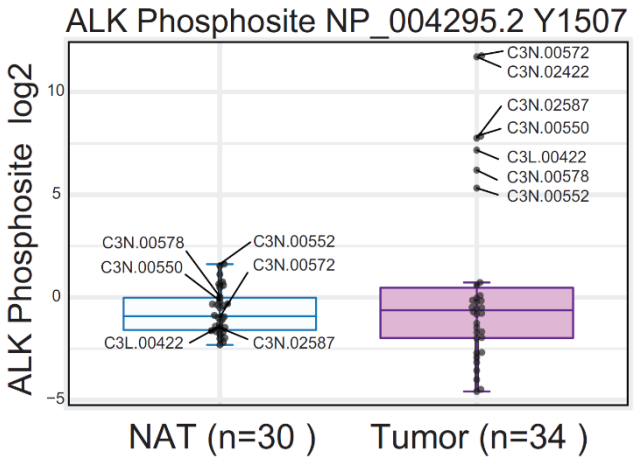
Gillette, et al., Cell 2020

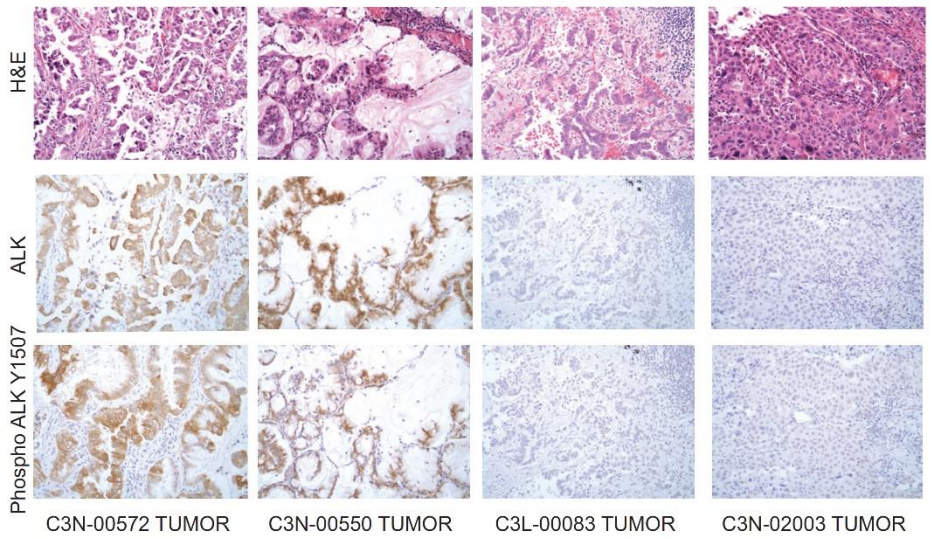
Global characterization of kinase fusions identified novel fusions, allowed assessment of likely functionality, nominated biomarkers and exposed fusion-driven biology

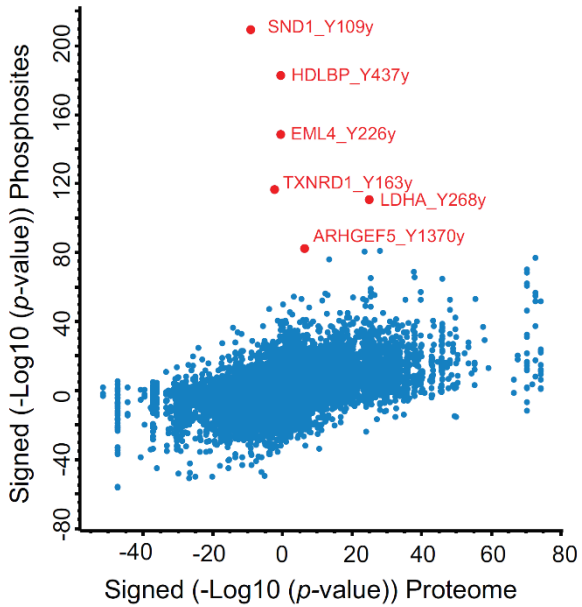






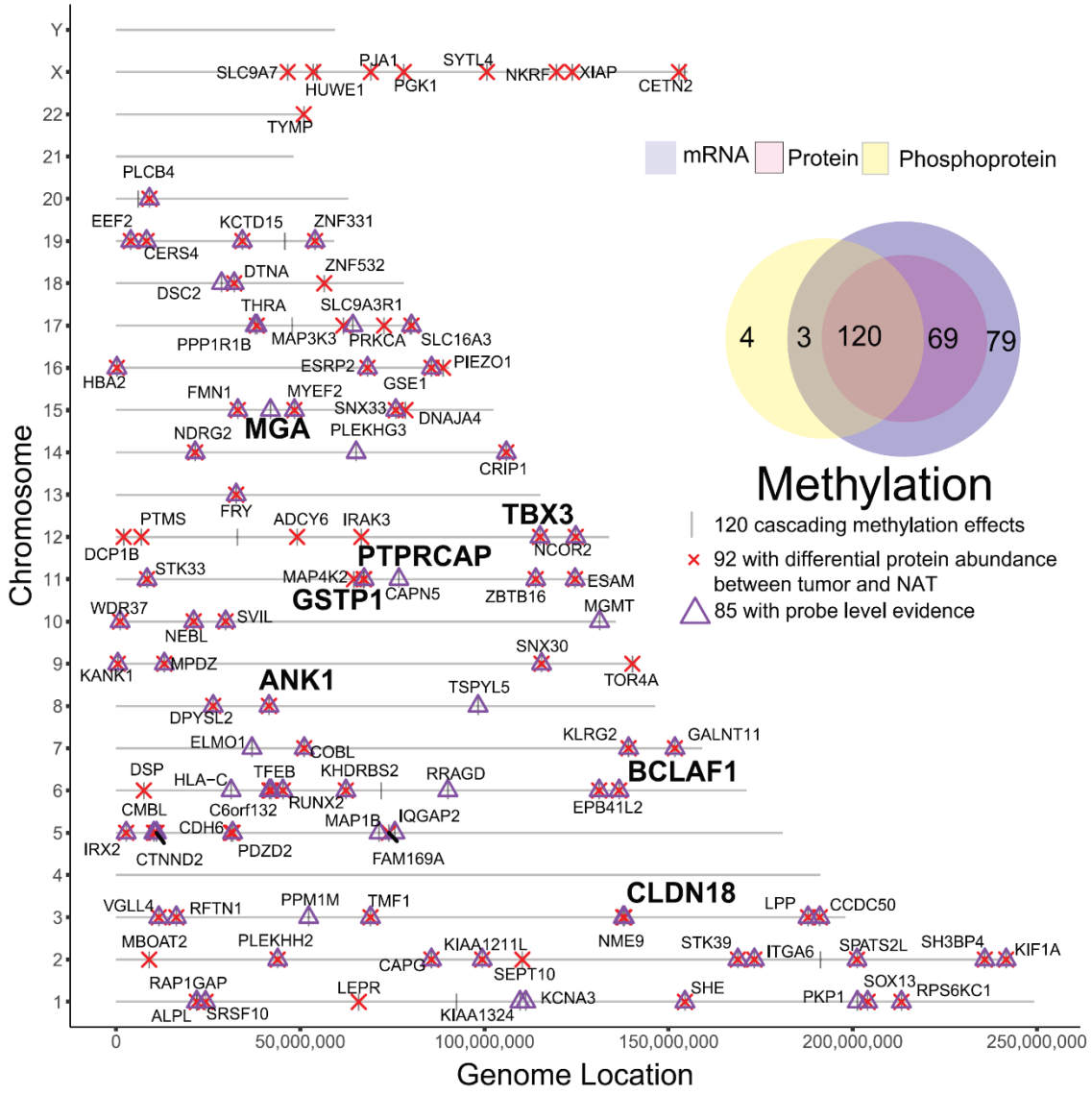


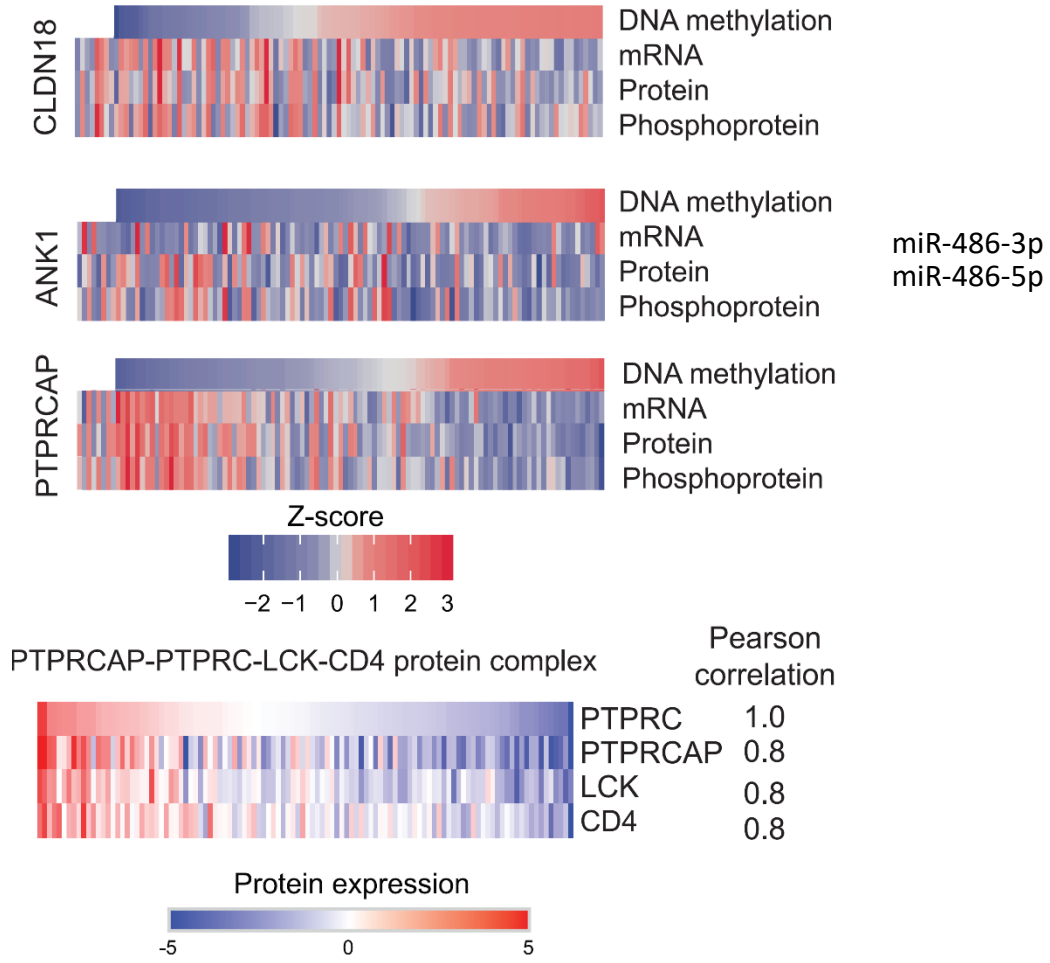




Gillette, et al., Cell 2020

120 genes had methylation associated with alterations of mRNA, protein and phosphosite expression, suggesting their possible functional significance



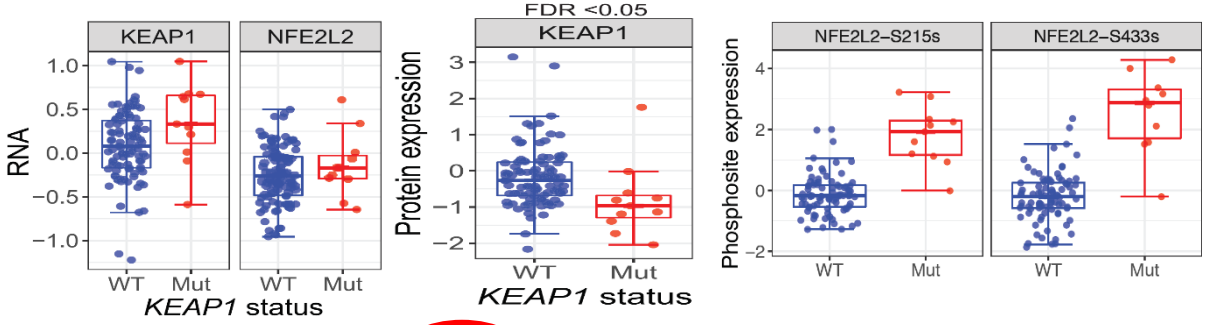


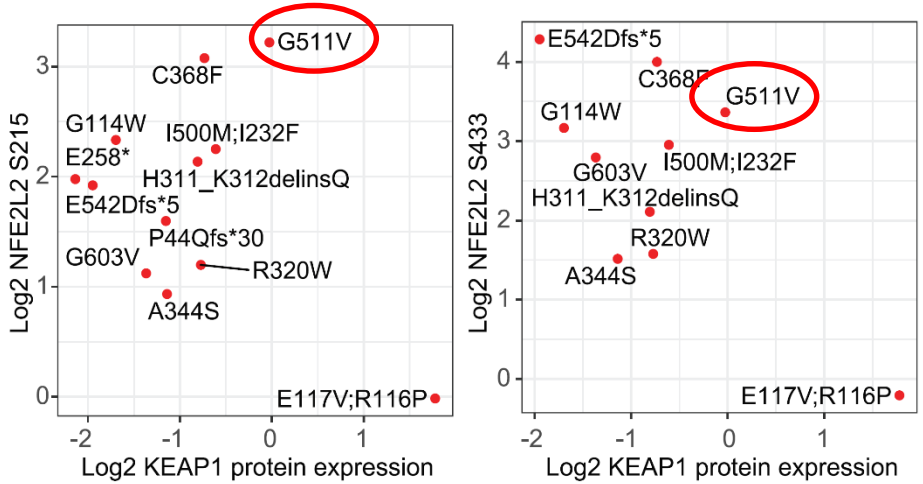
miR-486-3p
miR-486-5p

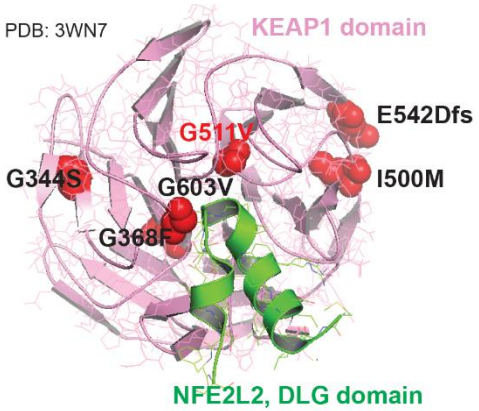
Proteins regulate antigen-specific T-cell activation

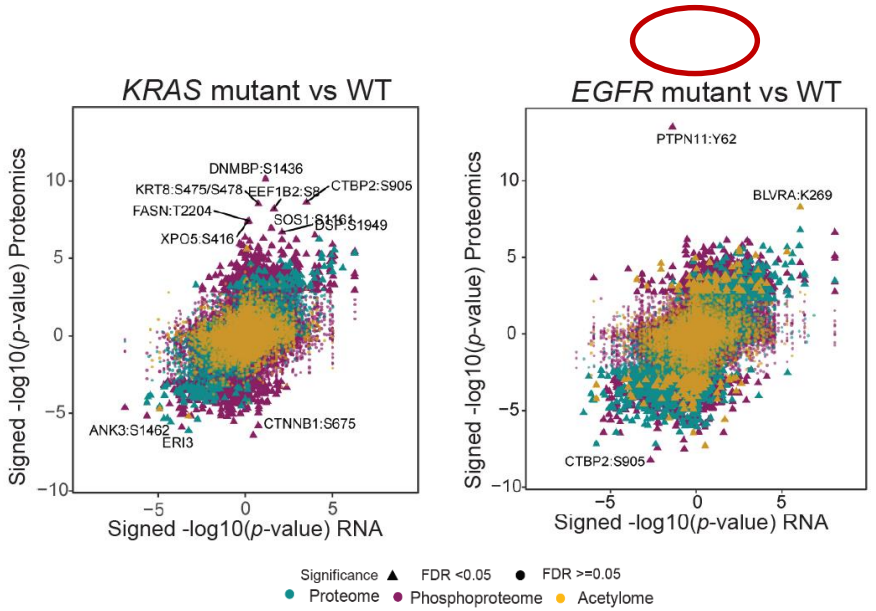
Gillette, et al., Cell 2020

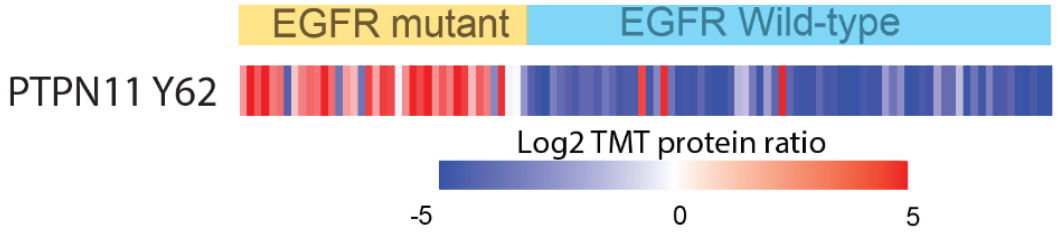
Proteogenomics exposes KEAP1 / NFE2L2 (NRF2) biology and a putative novel regulatory mechanism
Mutation association analysis highlights important outliers seen only in the phosphosite data





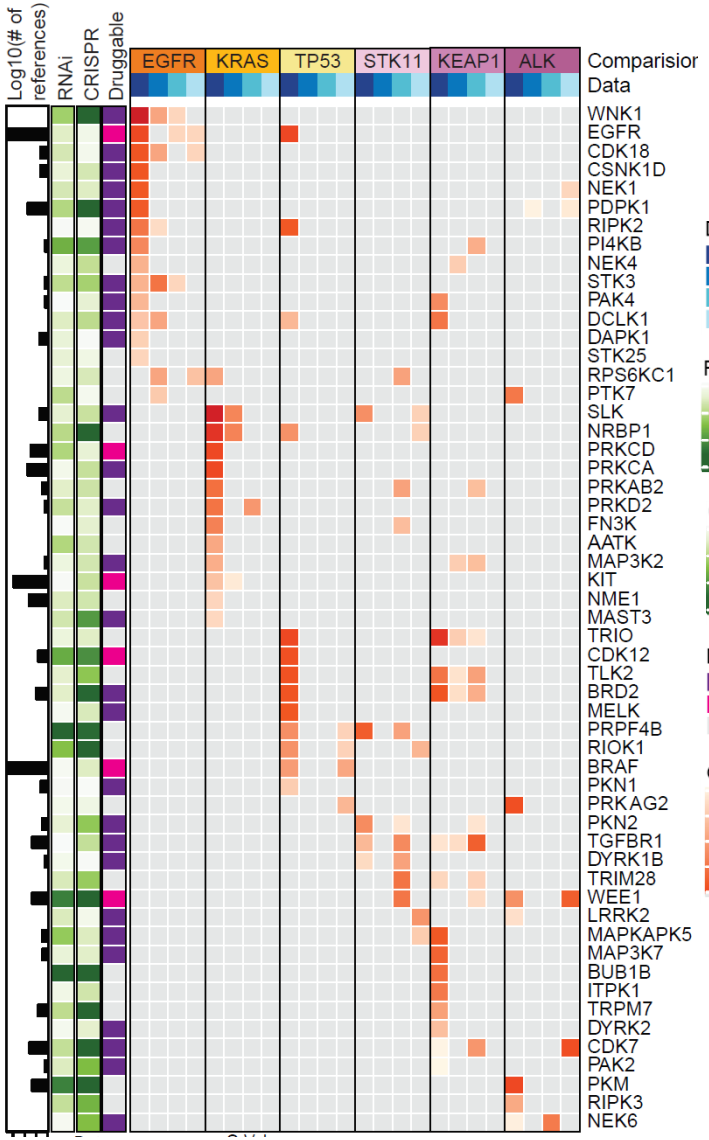


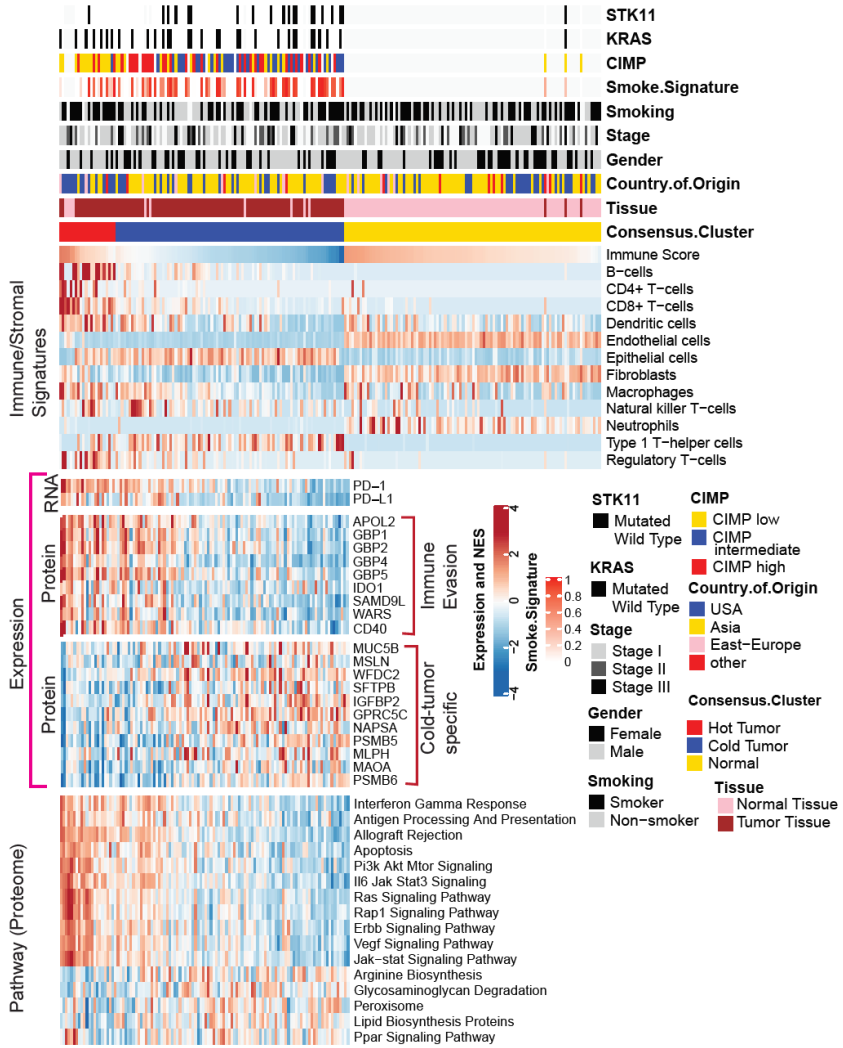


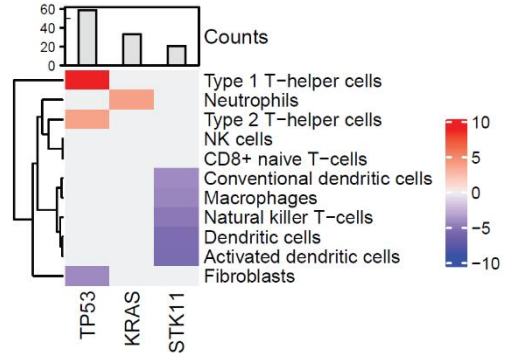


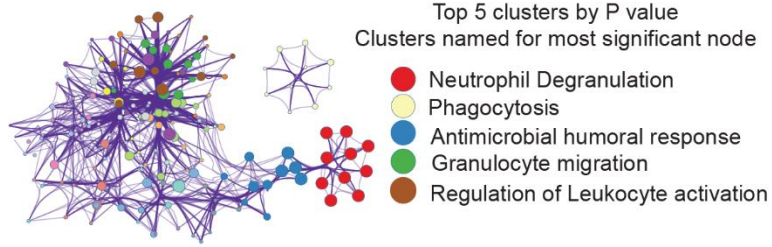
Gillette, et al., Cell 2020

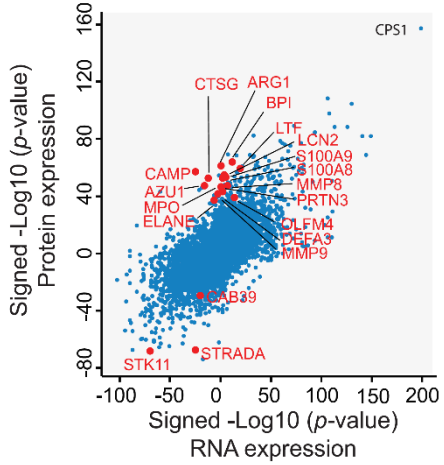
Kinase outlier analyses nominate candidate therapeutic targets
The immune landscape of LUAD shows regulated “cold” and “hot” tumor clusters
STK11 mutant tumors are especially “cold” and associated with neutrophil degranulation







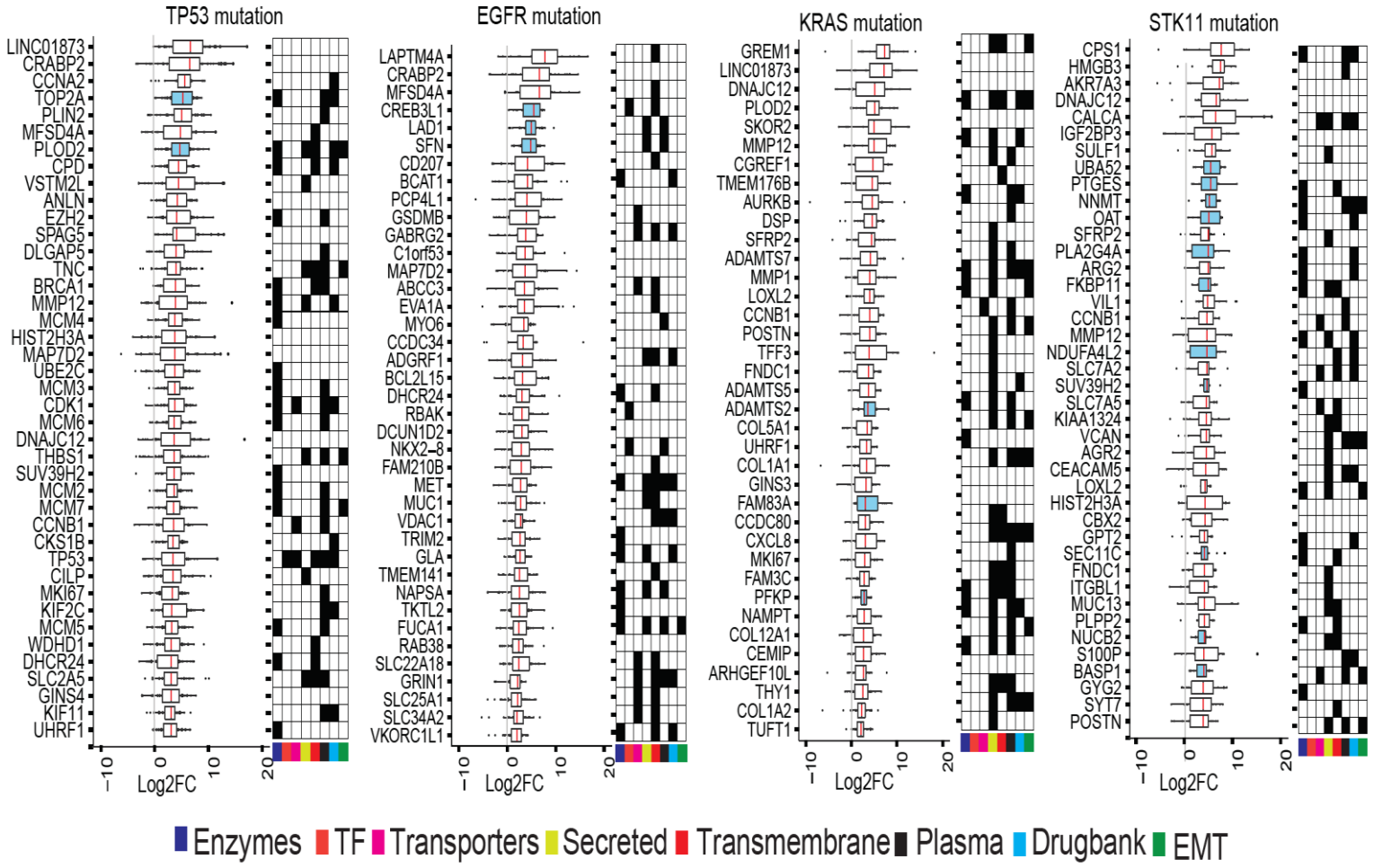


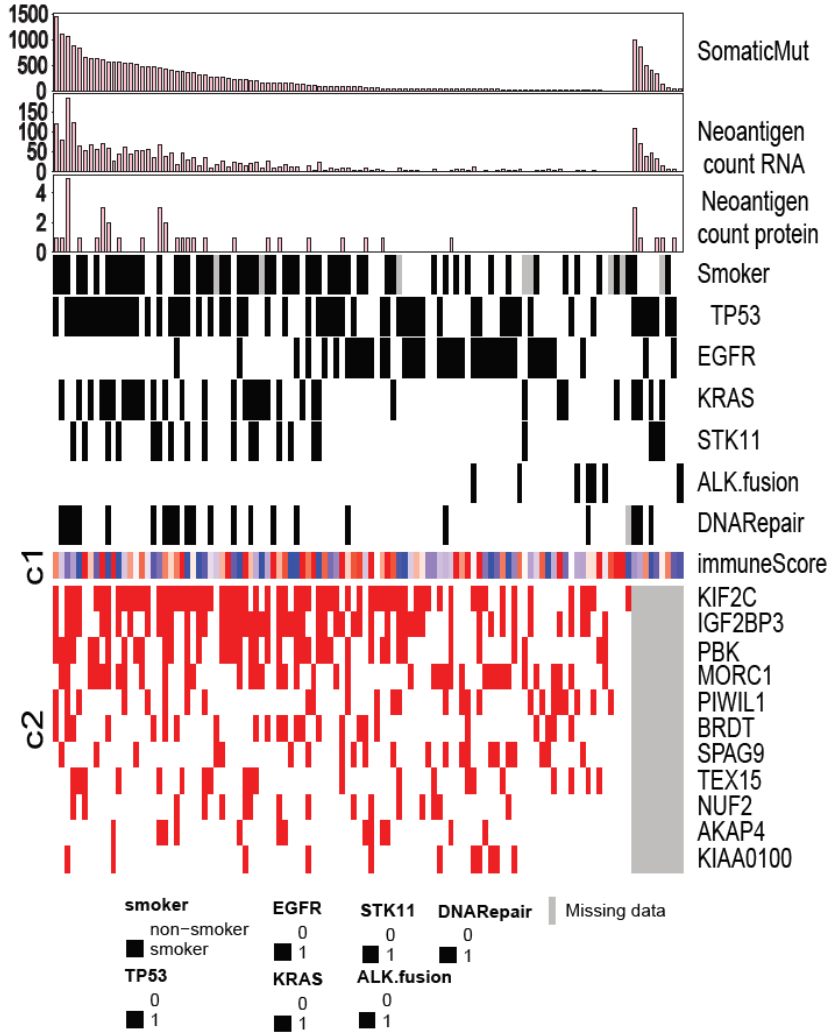


Gillette, et al., Cell 2020

Data provide a resource for global and subtype-specific LUAD biomarker development

Widely expressed Cancer-testis (CT) antigens are prime candidates as both biomarkers and immunogenic targets



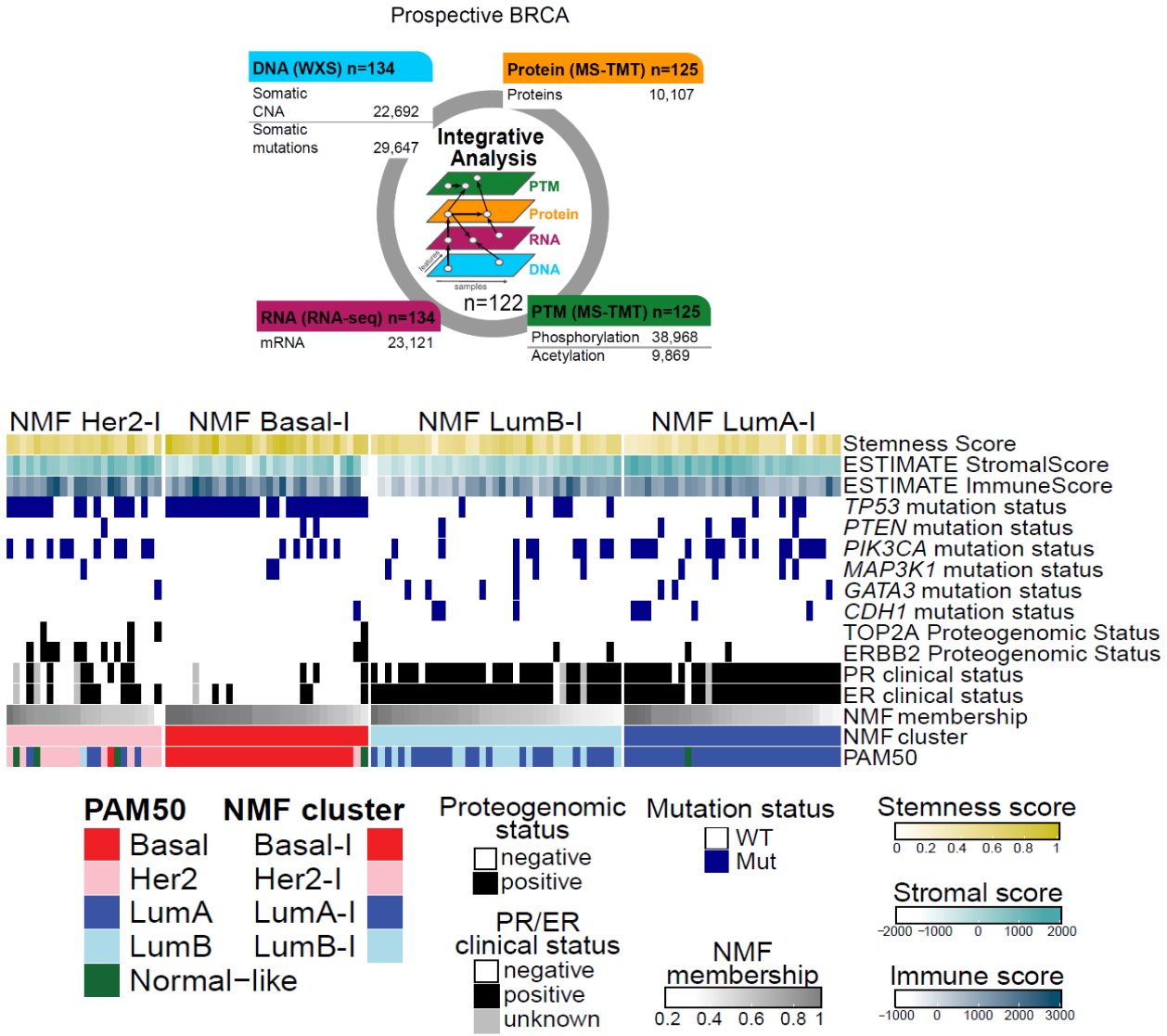


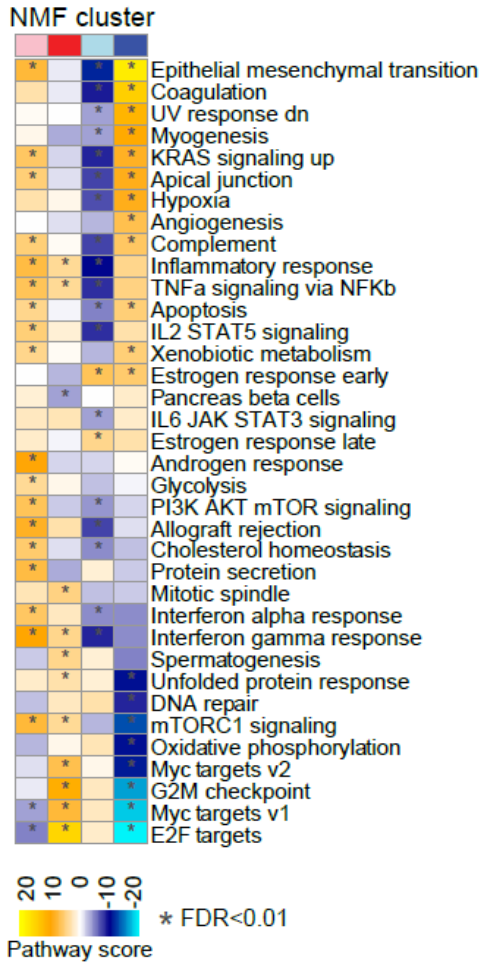
Gillette, et al., Cell 2020

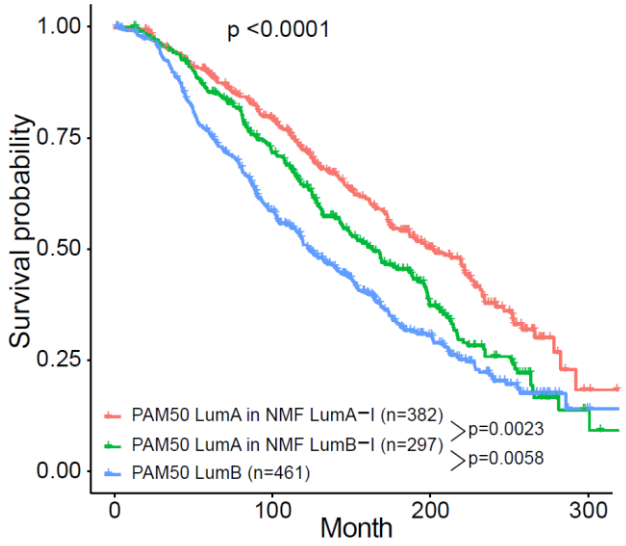
Non-negative matrix (NMF)-based multi-omics clustering defined 4 breast cancer clusters

Luminal clustering was discordant with PAM50 assignments

PAM50 Luminal A tumors assigned to the NMF Luminal B-enriched cluster had intermediate prognosis

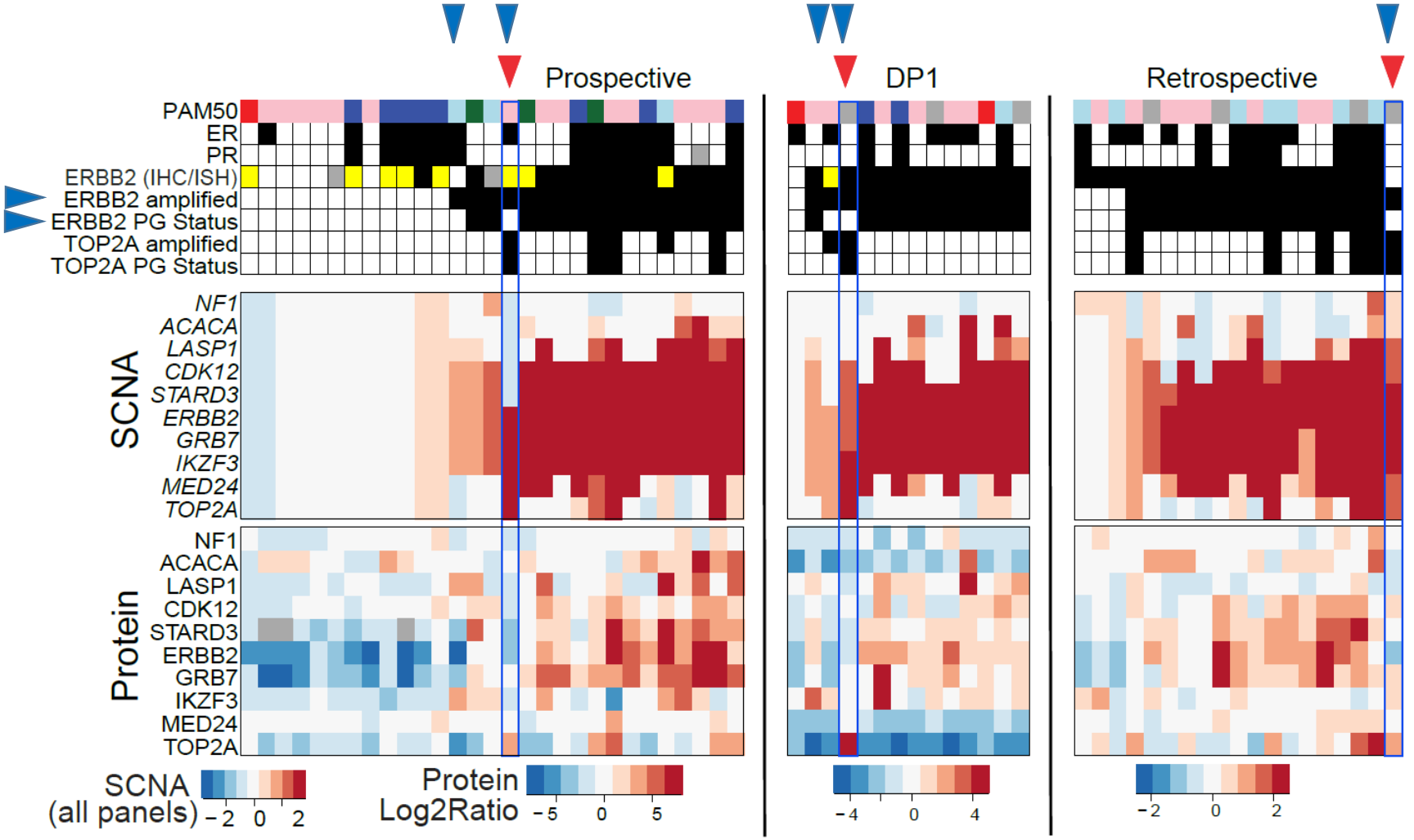


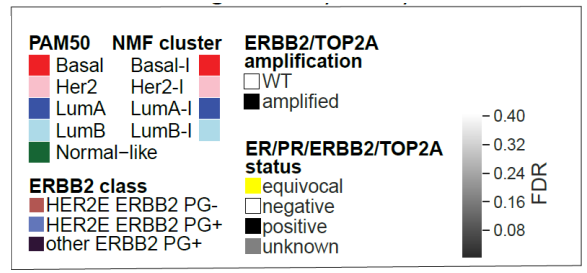




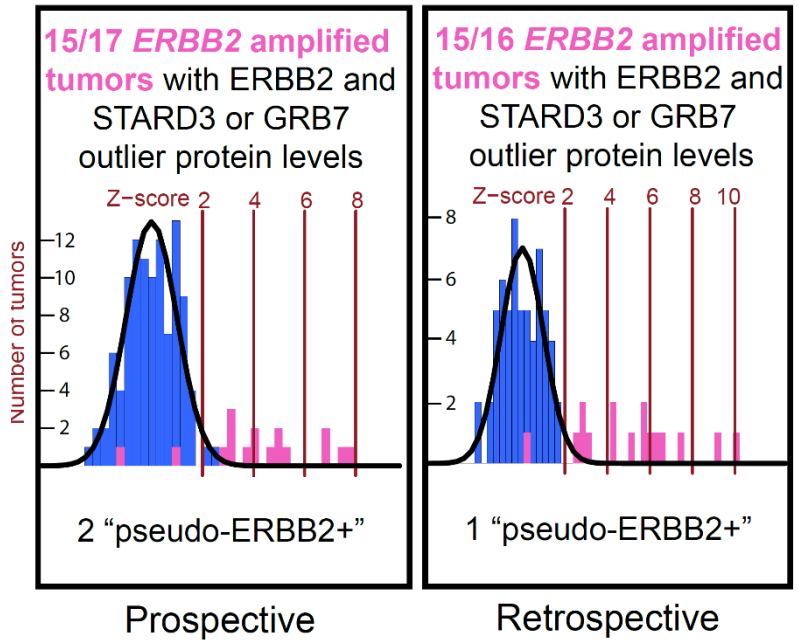
Krug, et al., Cell 2020 (in press)

Proteogenomic analysis of ERBB2 positive tumors shows “pseudo-ERBB2” samples with ERBB2 amplification but not protein expression. Some of these may have alternative 17q drivers.





Proteomic Assessment of *ERBB2* amplicon

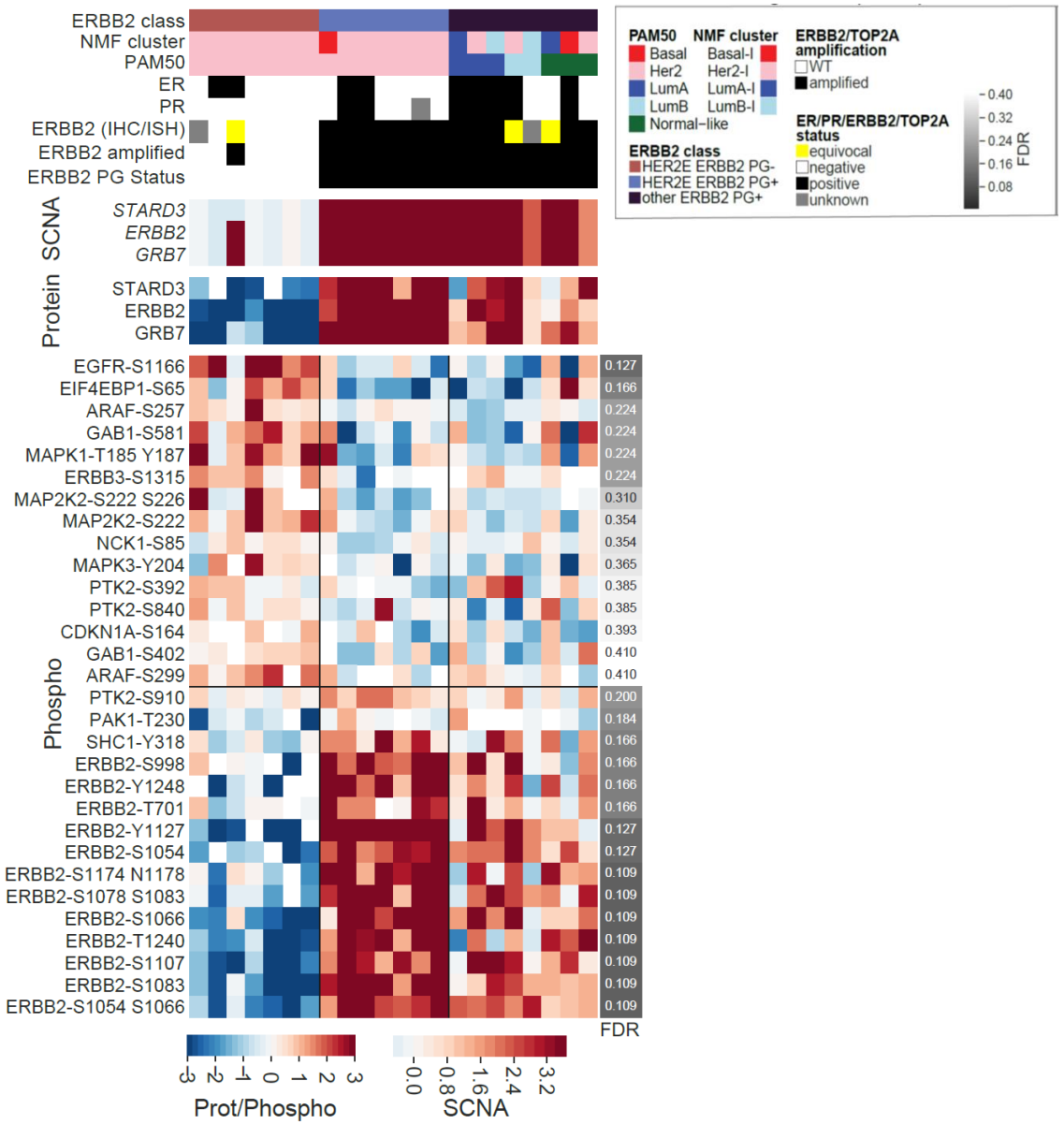


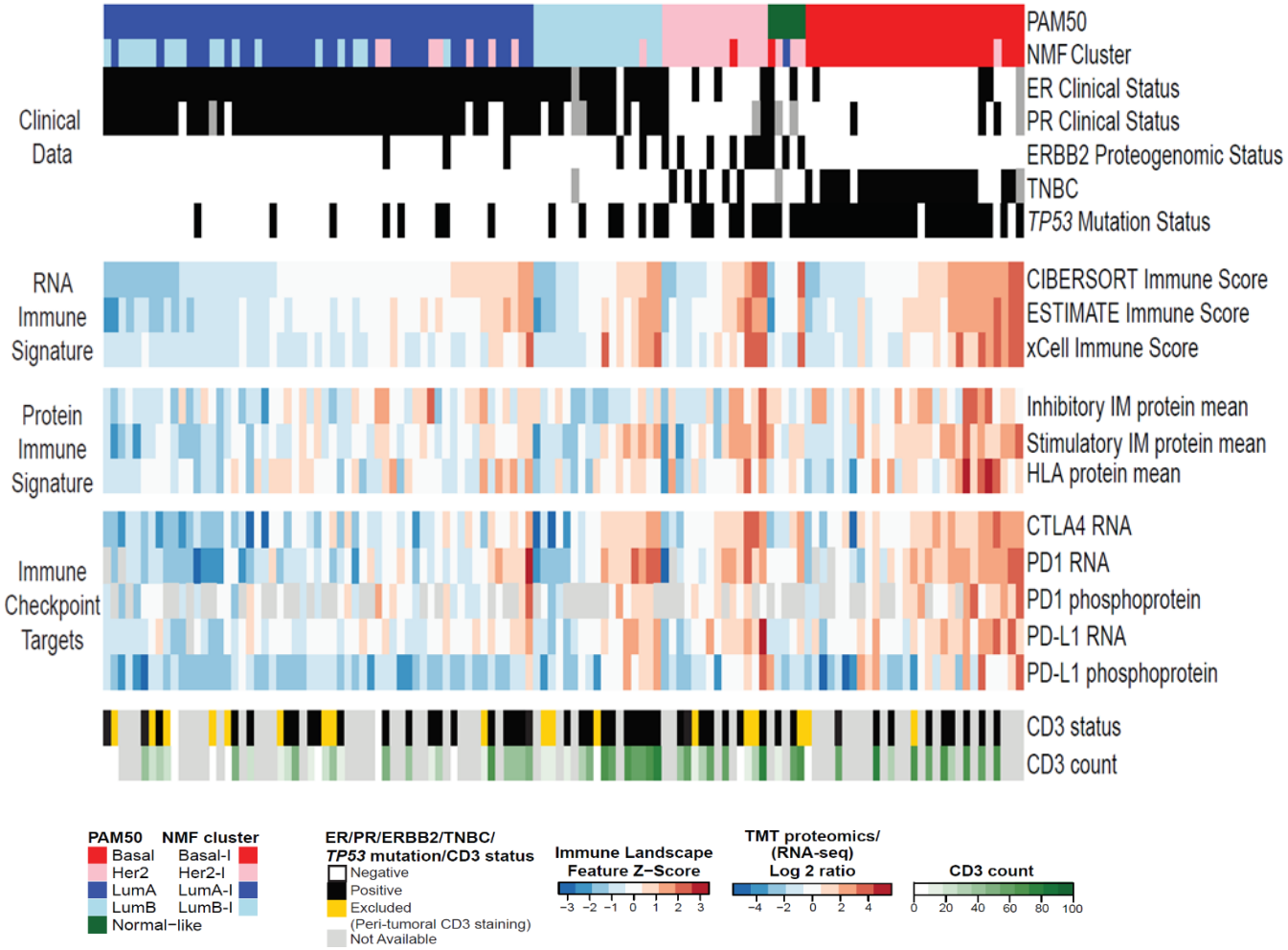
DP1: Satpathy et al.
Nature Communications 2020

Krug, et al., Cell 2020 (in press)

PAM50 HER2E, PG ERBB2-negative samples had phosphosite evidence of other ERBB and MAPK signaling

Proteogenomic analysis of the I-TME suggested broader applicability of immunotherapy in breast cancer



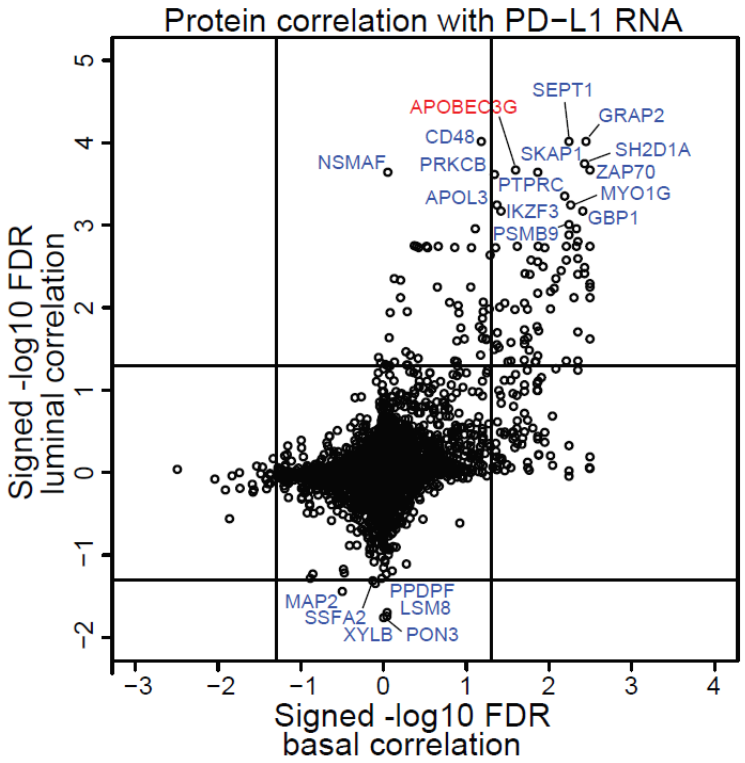


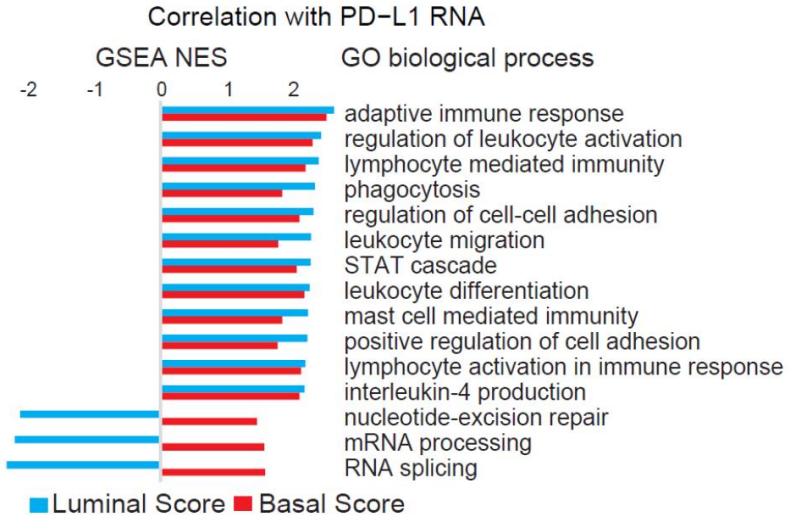
Krug, et al., Cell 2020 (in press)

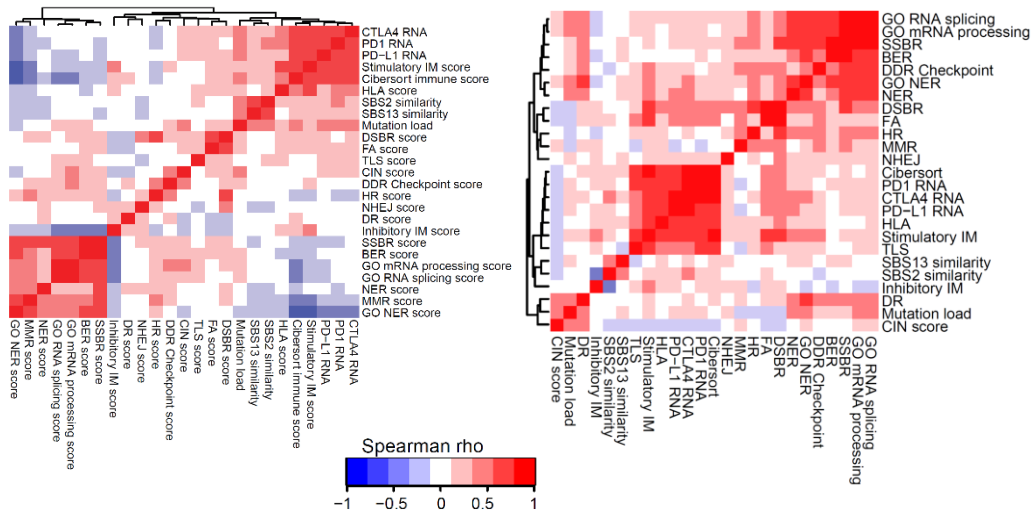
APOBEC-mediated mutagenesis correlates with an active I-TME in luminal breast cancer

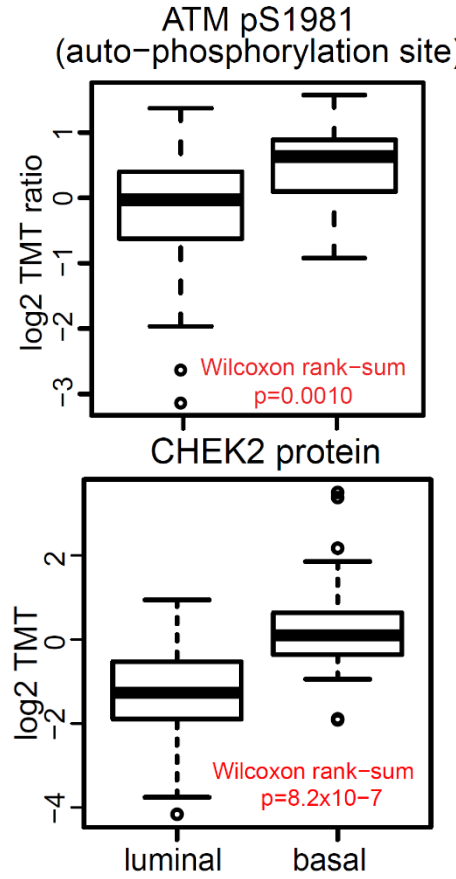
I-TME markers negatively correlated with NER, BER and MMR in luminal samples only

Phosphoproteomic data were consistent with suppressed DNA damage checkpoint activity in luminals



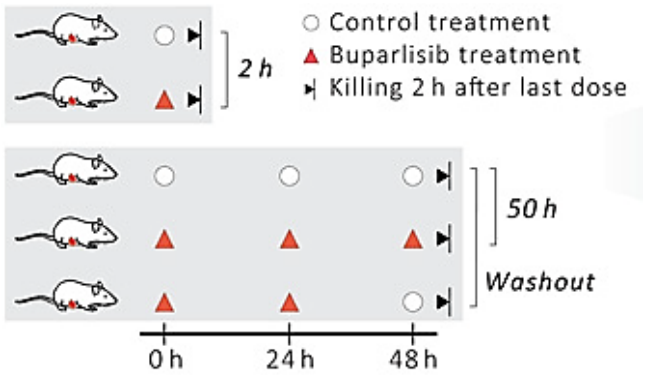


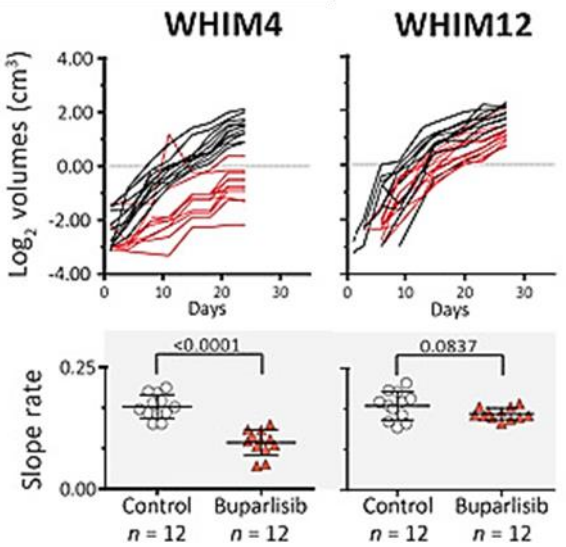


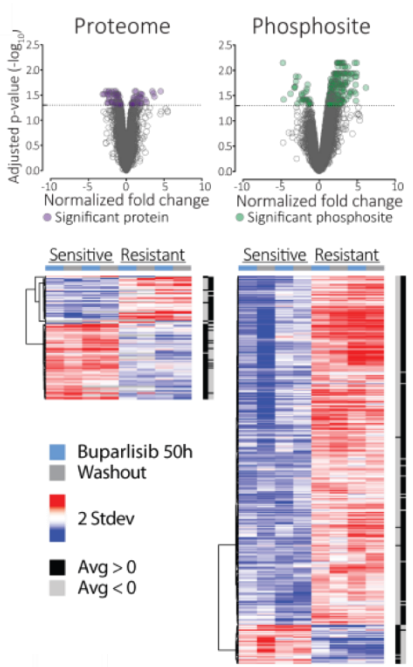


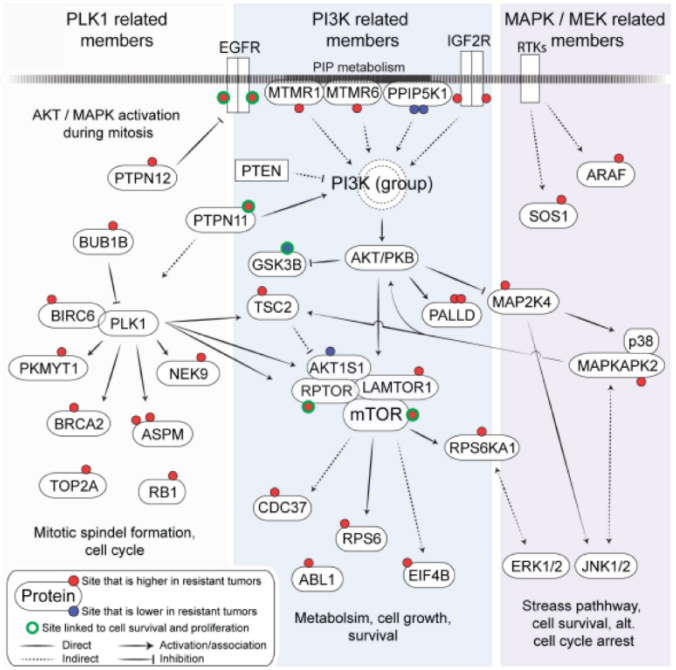
Krug, et al., Cell 2020 (in press)

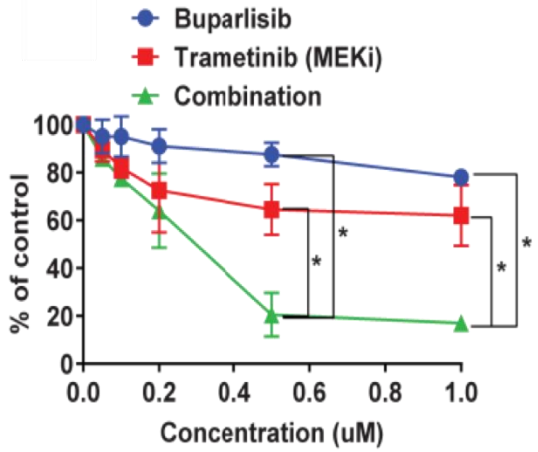
Phosphoproteogenomics in model systems is a powerful approach for assessing response and identifying mechanisms of resistance to targeted therapies

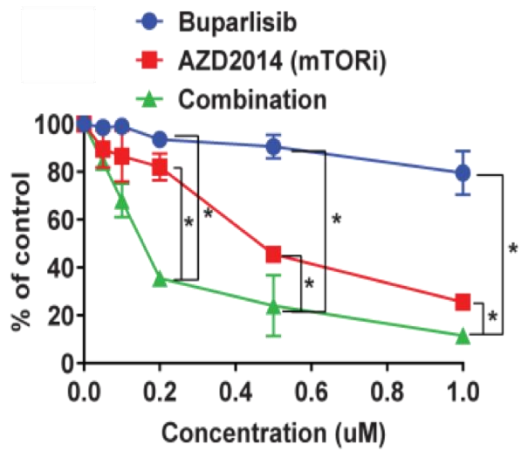




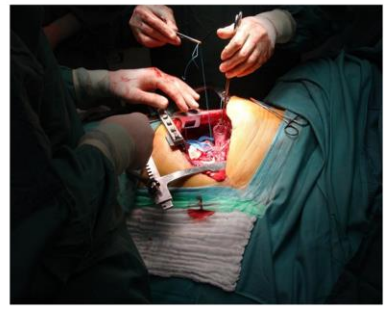






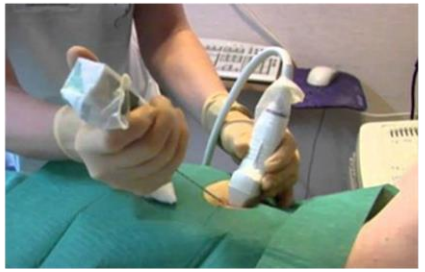


Bulk Tumor (gm-scale)



- Grams of wet weight tissue obtained from a surgical resection for analysis
- Samples of bulk tumors are cryopulverized to obtain uniform sample for DNA, RNA and protein processing

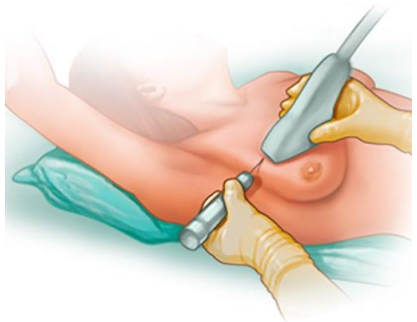
Core Needle Biopsy (mg scale)

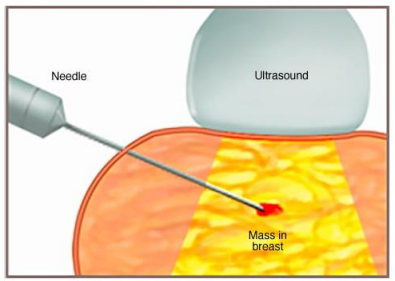


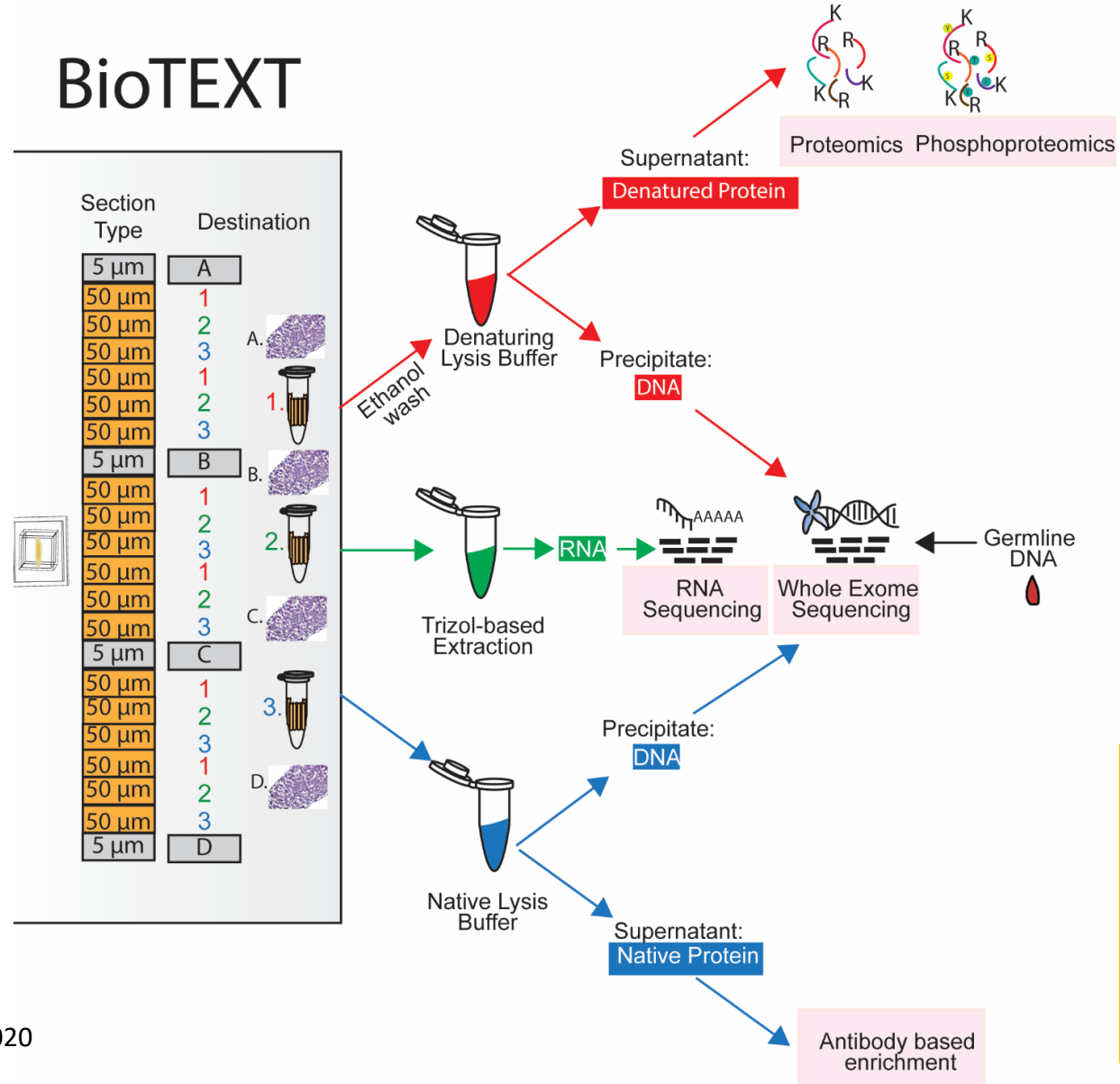
- Typically 1-2 cores per patient (10-20 mg wet weight tissue/core)
- Substantially lower DNA, RNA, protein yield
- Cryopulverization not feasible
- Often embedded in wax (OCT)

Mundt, et al. *Mass Spectrometry–Based Proteomics Reveals Potential Roles of NEK9 and MAP2K4 in Resistance to PI3K Inhibition in Triple-Negative Breast Cancers* Cancer Research 2018

Biopsy Trifecta Extraction (BioTExt) allows suite of full suite of genomic and proteomic analyses analyses from a single needle-core biopsy







Microscaled proteomics (MiProt)
Multiplex proteomic analysis
10 different patient cores/plex

Genomics
WGS
WXS
RNA-Seq
miRNA-Seq

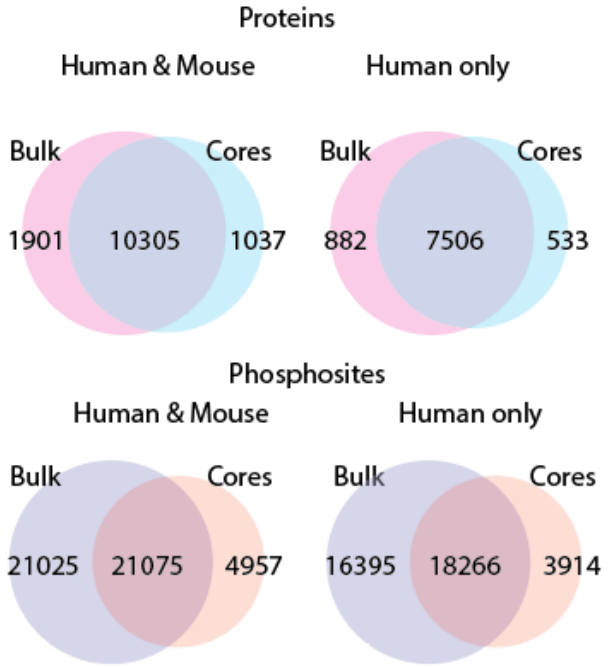
Reserved sample for targeted proteomics
MRM, PRM
Kinome Profiling

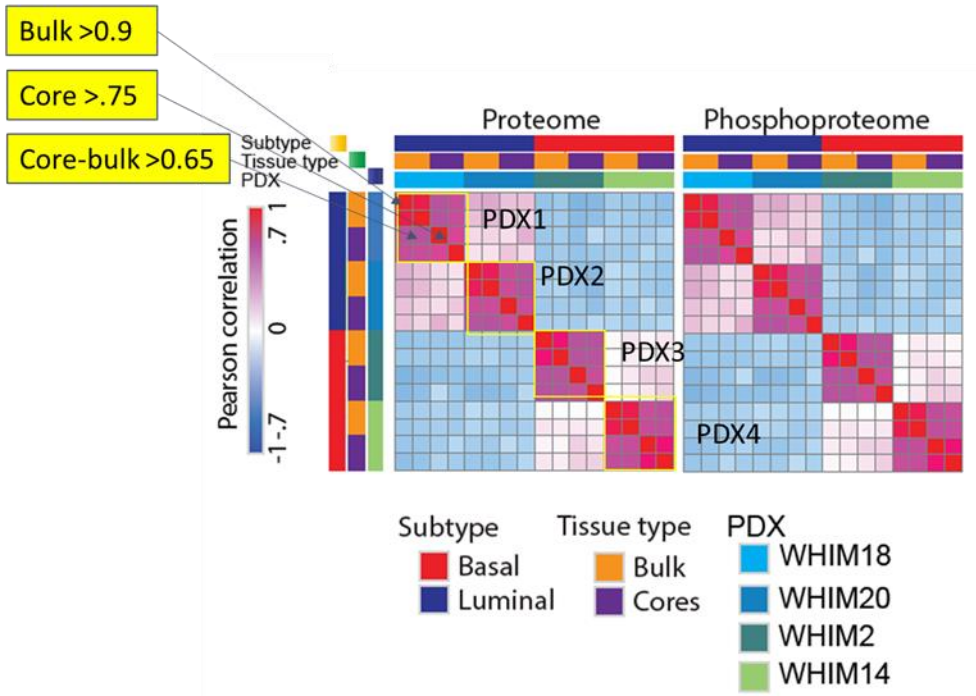
Satpathy et al., Nat Comm 2020

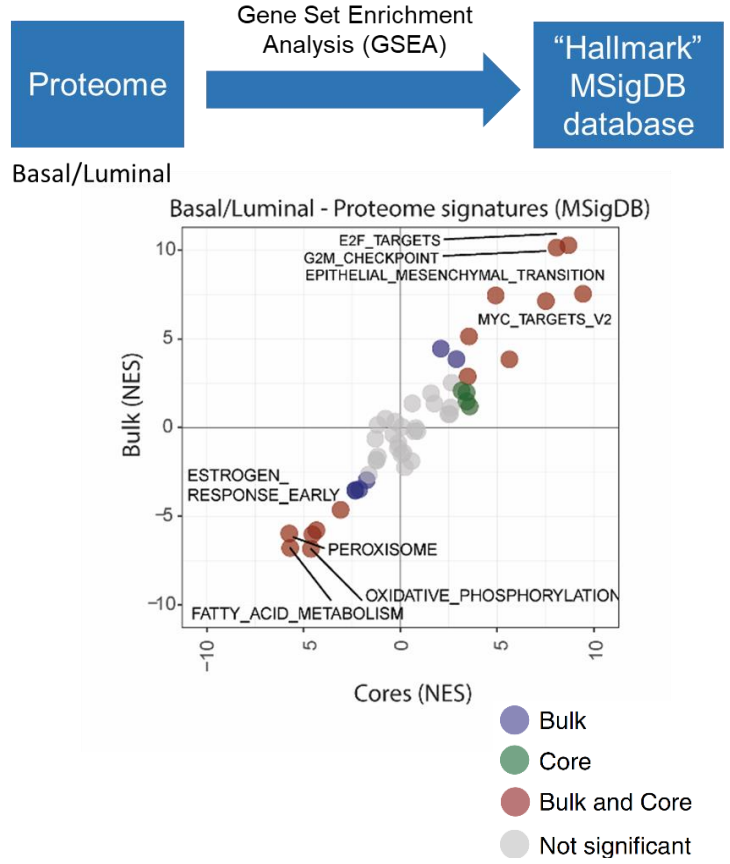
Proteome depth from cores is similar to bulk
Number of quantified phosphosites is reduced
Biology is preserved

From 25 ug peptides/core

- >10,000 proteins
- >20,000 phosphosites

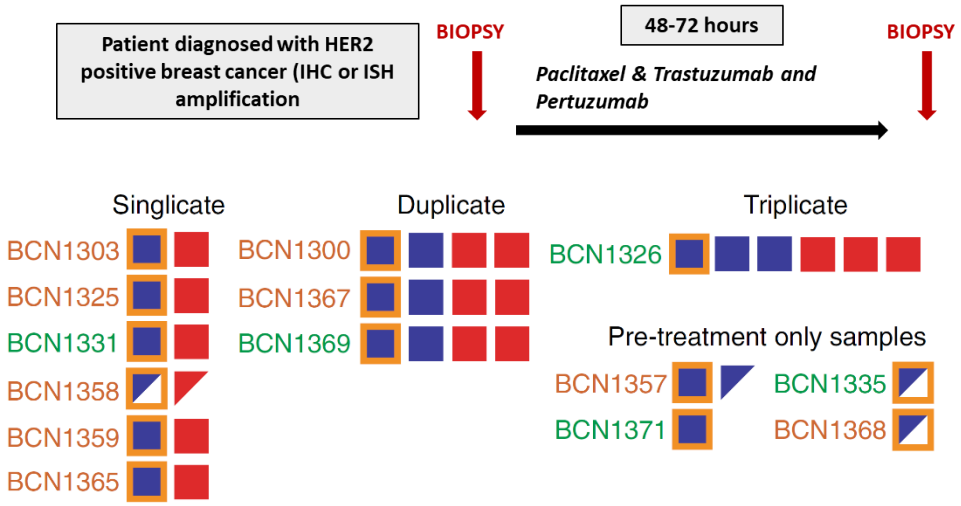




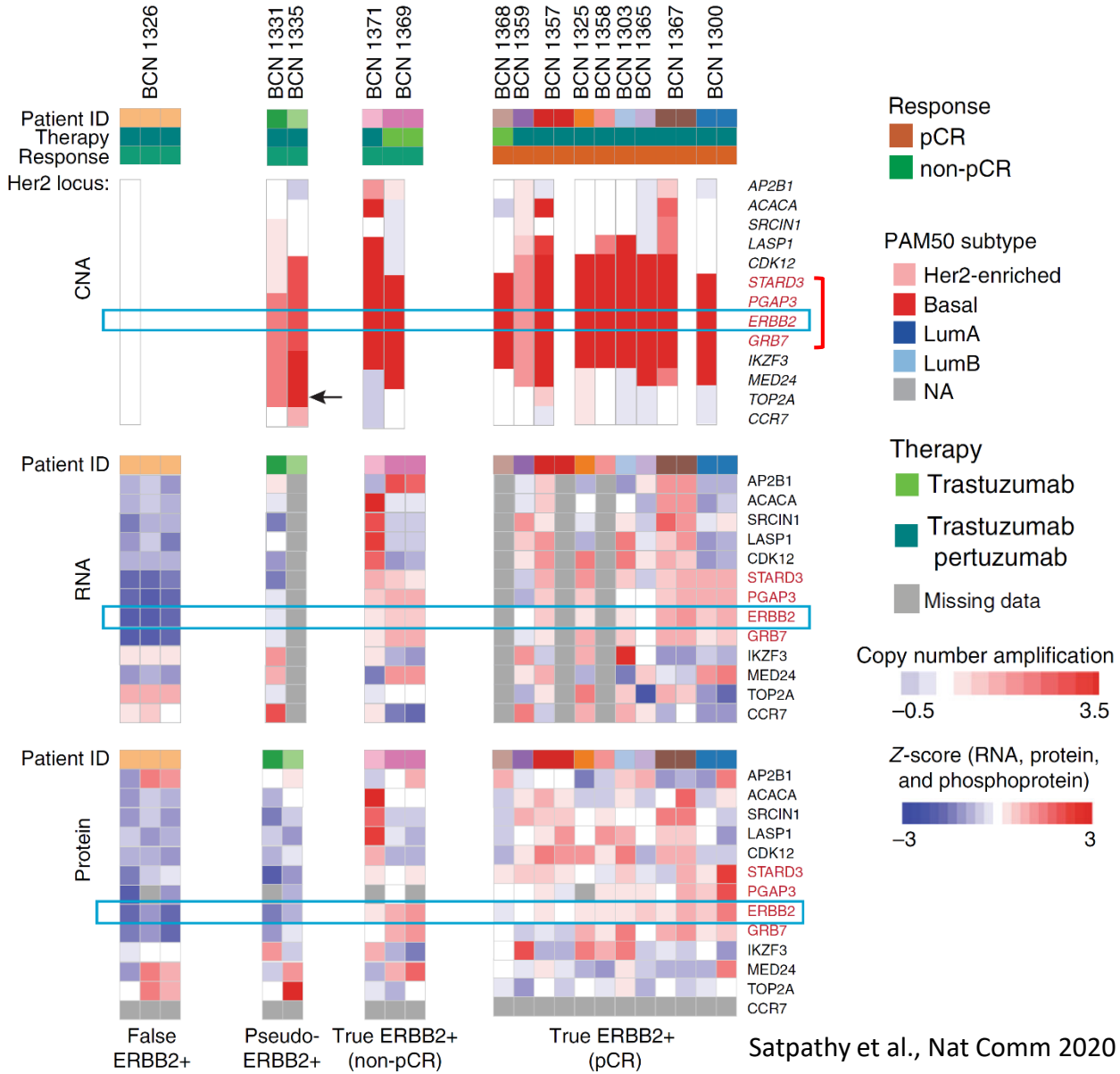


Satpathy et al., Nat Comm 2020

Microscaling technologies have been successfully applied to needle biopsy samples from clinical trials

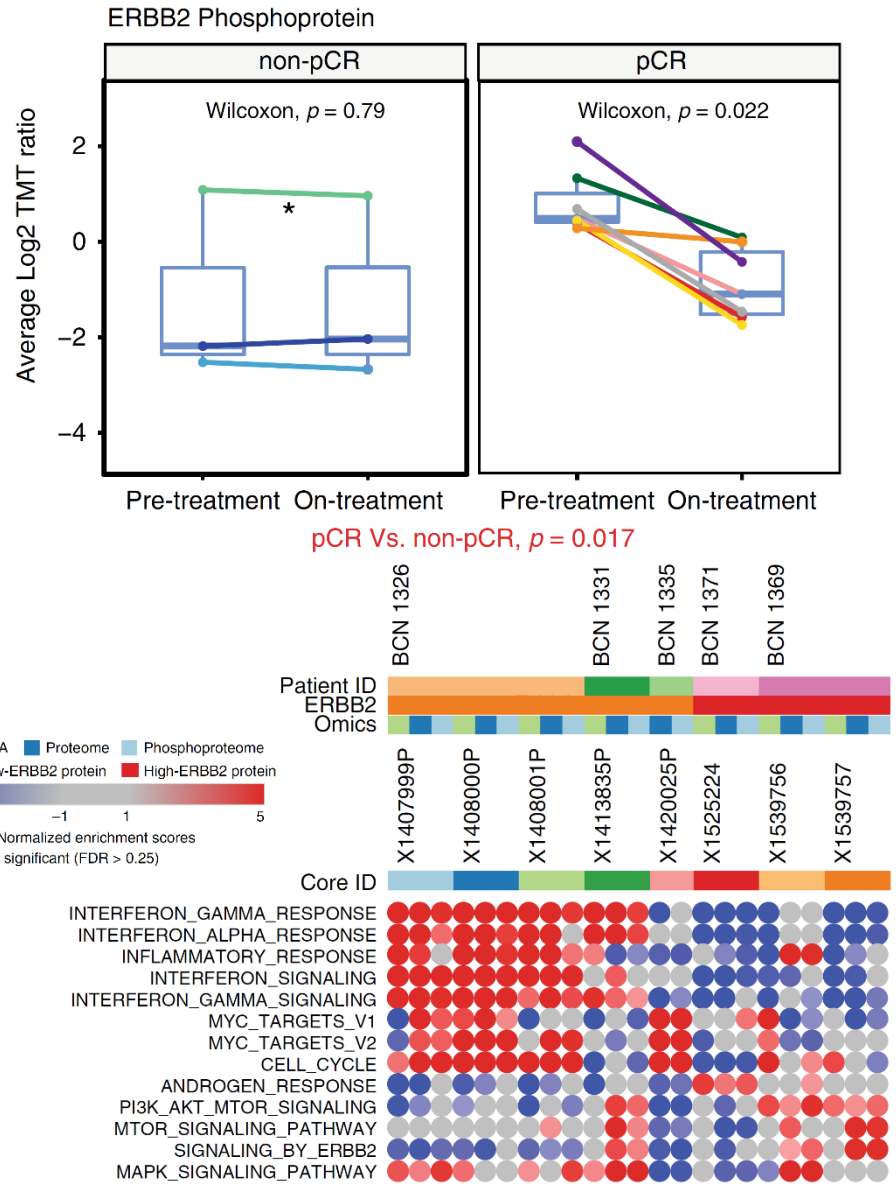


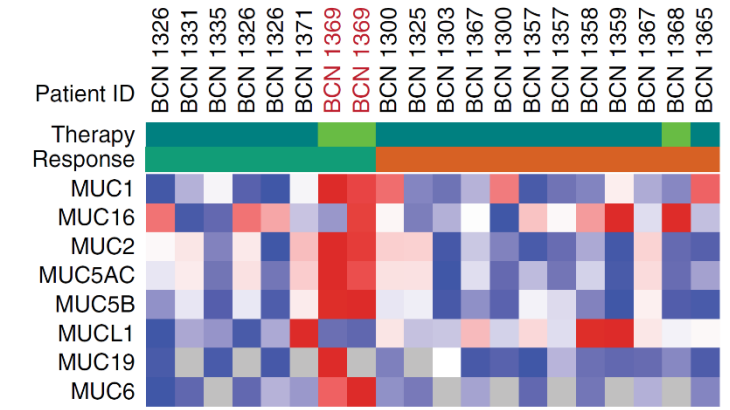
Analysis	Total observed data points	Average observed per sample
CNA (WXS)	27,217	27,217
Gene expression (RNA-seq)	23,549	19,492
Somatic mutation (WXS)	369	27
Proteins (TMT11 Proteomics)	11,657	10,333
Phosphopeptides (TMT11 Proteomics)	23,261	17,401

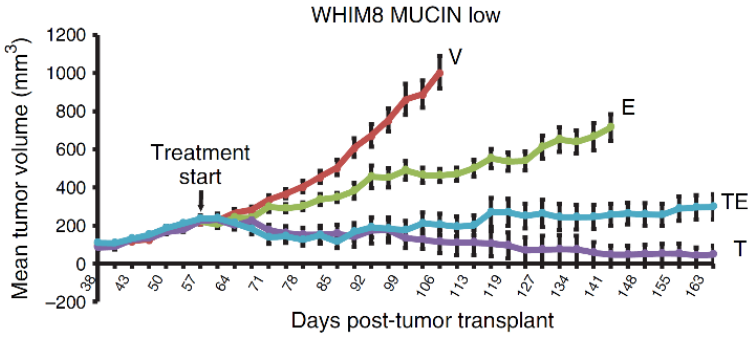


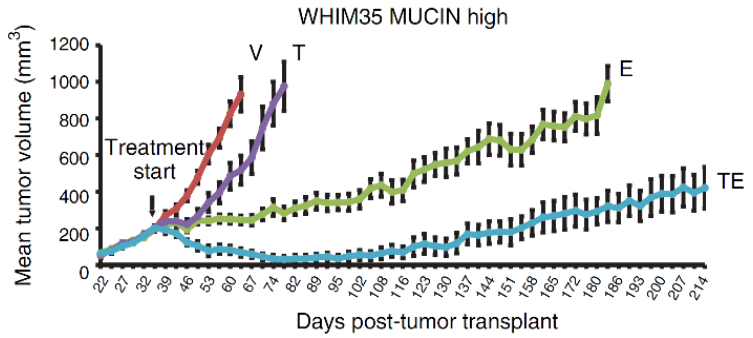
Satpathy et al., Nat Comm 2020

Measurement of ERBB2 phosphoprotein predicts response at 48-72 hours on treatment
Phosphoproteomics suggests mechanisms of resistance and therapeutic alternatives
Initial model-based verification data are encouraging









Satpathy et al., Nat Comm 2020

Summary

- **Proteogenomics provides a powerful, reproducible and complementary approach to characterizing cancer biology, exploring mechanisms of resistance and identifying potential therapeutic vulnerabilities**
- **Proteogenomics should be part of the armamentarium in programs designed to analyze cancer from an evolutionary perspective, helping illuminate**
  - ***Disease taxonomy***
    - Revised hormone receptor positive breast cancer subtype assignments
  - ***Knowledge of the environment***
    - Immune landscape gives insight into biology and therapeutic options in lung and breast cancer
      - STK11 tumors may be vulnerable to therapies targeting neutrophil degranulation proteins
      - Subsets of luminal tumors nominated for immune therapy
  - ***Detailed biological repertoire of cancer and individual tumors***
    - Proteomic and PTM associations with driver mutations, fusion events and promoter methylation
    - Sample-level characterization of vulnerabilities from phosphosite and kinase outlier analyses
    - Subtype- and sample-specific metabolic profiling leveraging acetylproteomics
    - Improved definition of clinically important marker status (ERBB2, Rb) with thereapeutic implications
  - ***Effects of specific selective pressures***
    - Model systems may improve understanding of mechanisms of resistance
    - Direct analyses of human tumors on treatment can give critical insights into response and resistance

Acknowledgments

BROAD INSTITUTE

- **PROTEOMICS**
  - **Karsten Krug**
  - **Shankha Satpathy**
  - Lauren Tang
  - Melanie MacMullan
  - Shayan Avanessian
  - Ramani Kothadi
  - Karl Clauser
  - Philipp Mertins
  - **D. R. Mani**
  - **Steve Carr**
- **GENOME ANALYSIS**
  - David Heiman
  - Mike Noble
  - Chet Birger
  - Yifat Geffen
  - Yosef Maruvka
  - Gad Getz

Washington University in St. Louis

- **Alla Karpova**
- Song Cao
- Steven Foltz
- Matthew Wyczalkowski
- Li Ding

NYU

- **Lili Blumenberg**
- Emily Kawaler
- Kelly Ruggles
- David Fenyo

University of Michigan

- **Mohan Dhanasekaran**
- Marcin Cieslik
- Alexey Nesvizhkii
- Gil Omenn
- Arul Chinnaiyan

Mount Sinai

- **Francesca Petralia**
- Xiaoyu Song
- Jiayi Ji
- Boris Reva
- Azra Krek
- Pei Wang

BCM Baylor College of Medicine

- **Eric Jaehnig**
- **Meenakshi Anurag**
- George Miles
- JonathanLei
- Suhas Vasaikar
- Chen Huang
- Bing Zhang
- Thomas (“Trey”) Westbrook
- **Matthew Ellis**

NIH NATIONAL CANCER INSTITUTE

- Emily Boja
- Tara Hiltke
- Chris Kinsinger
- Mehdi Mesri
- Ana Robles
- Henry Rodriguez

leidos

- Mathangi Thiagarajan
- Linda Hannick

CPTAC
National Cancer Institute
at the National Institutes of Health