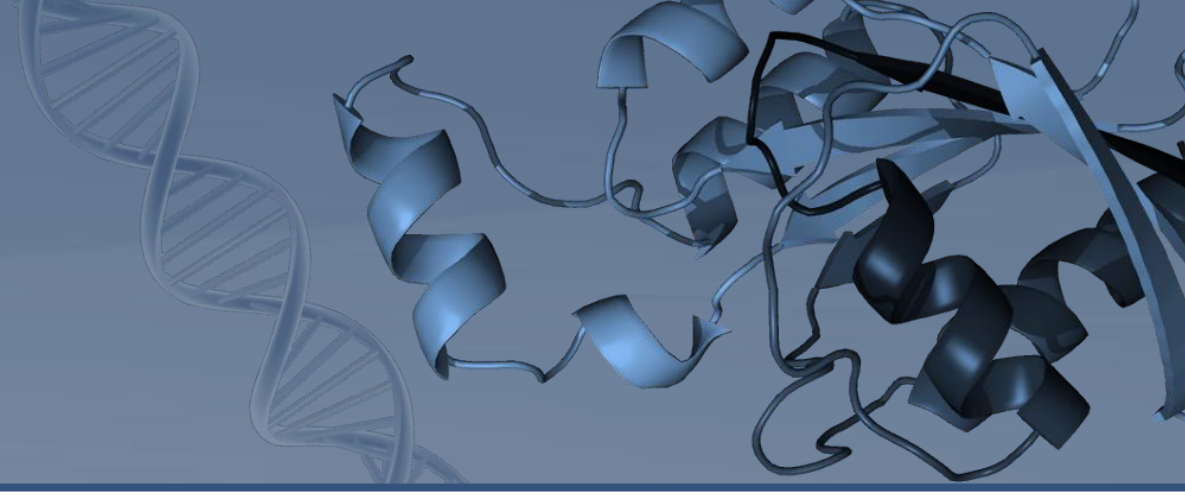
OFFICE OF CANCER CLINICAL PROTEOMICS RESEARCH



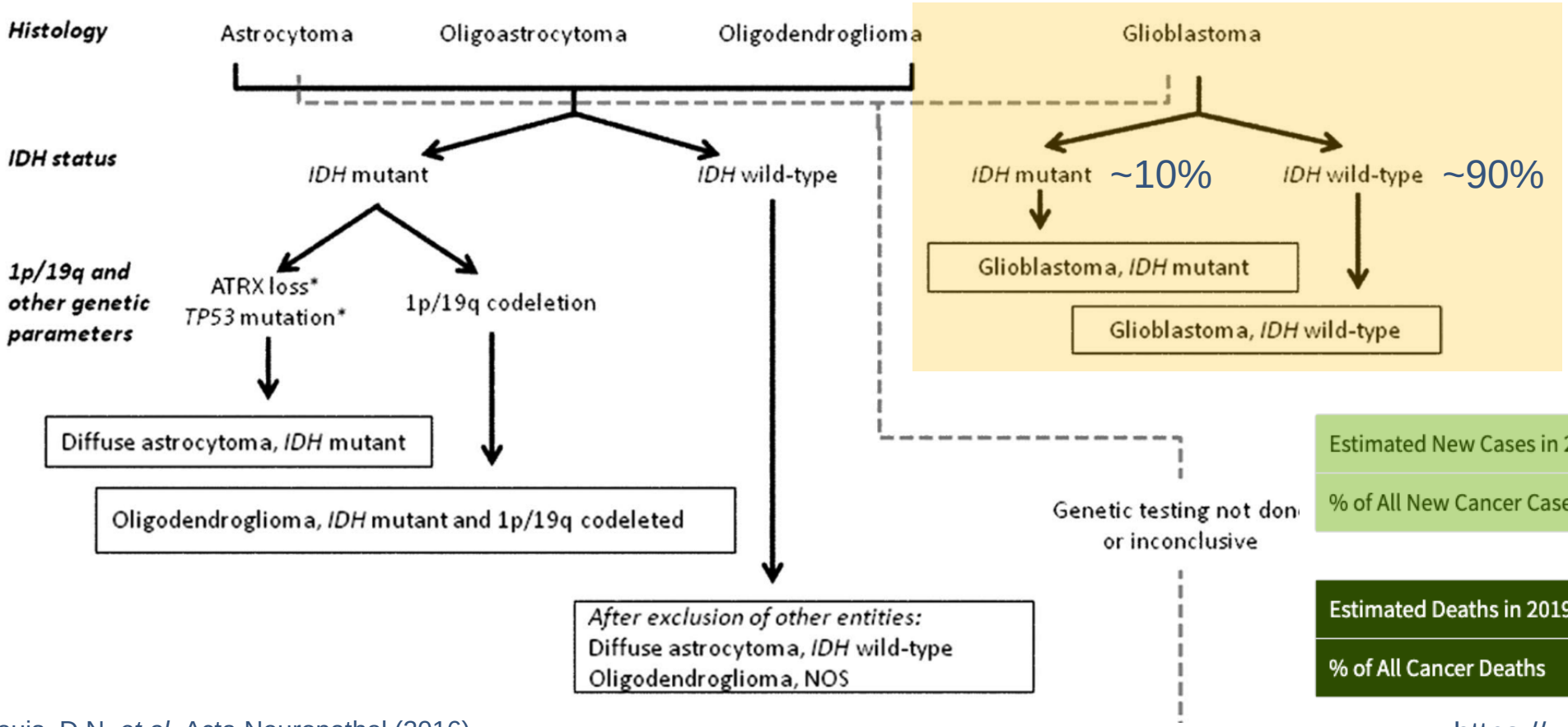
Proteogenomic analysis of Glioblastoma

Liang-Bo Wang
Ding Lab
Washington University in St. Louis

Runyu Hong
Fenyo Lab
NYU School of Medicine

TCIA CPTAC SIG Webinar

Glioblastoma is a rare disease with no personalized treatment



Estimated New Cases in 2019	23,820
% of All New Cancer Cases	1.4%

Estimated Deaths in 2019	17,760
% of All Cancer Deaths	2.9%

Louis, D.N. *et al*. Acta Neuropathol (2016)

https://www.cancer.gov/types/brain

- Poor prognosis: median overall survival: 14–17 month; 5-year survival rate ~10%
- Existing subtypes within IDH WT tumors has yet benefited the GBM treatment

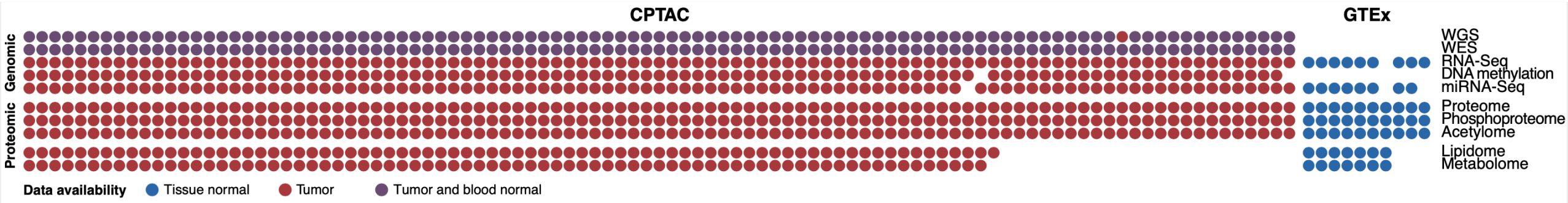
Goals of GBM proteogenomic characterization

Comprehensive characterization of GBM tumors

Discover clinically beneficial patient stratifications

Identify novel markers to be potential therapeutic targets

Data overview of 99 GBM tumors



(manuscript in revision)

Prospective collection of treatment naïve GBM tumors (n = 99)

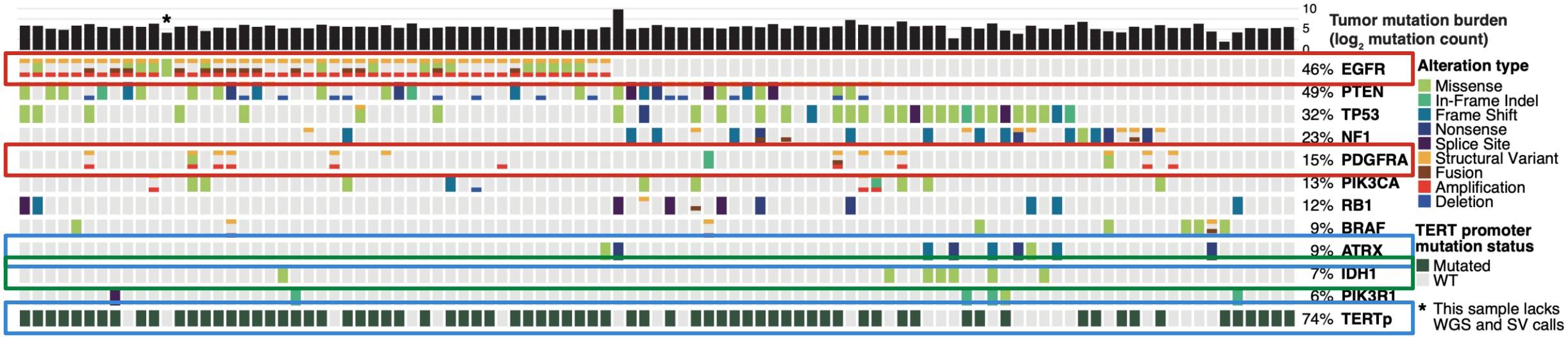
Normal frontal cortex samples from GTEx project (n = 10)

10 different data types on the same cryopulverized segment:

- Available in TCGA: WGS, WXS, RNA-Seq, DNA methylation array, and miRNA-Seq
- 5 novel mass spectrometry assays:
  proteome, phosphoproteome, acetylome, lipidome, and metabolome

Data type	# features
Proteome	11,141
Phosphoproteome	101,266 (86,554 unique sites)
Acetylome	18,767
Metabolome	134
Lipidome	581

Mutational landscape



Top 12 significantly mutated genes: mutations, structural variants (SV), fusions, and CNV

Two receptor tyrosine kinases (RTKs) are highly altered: EGFR and PDGFRA

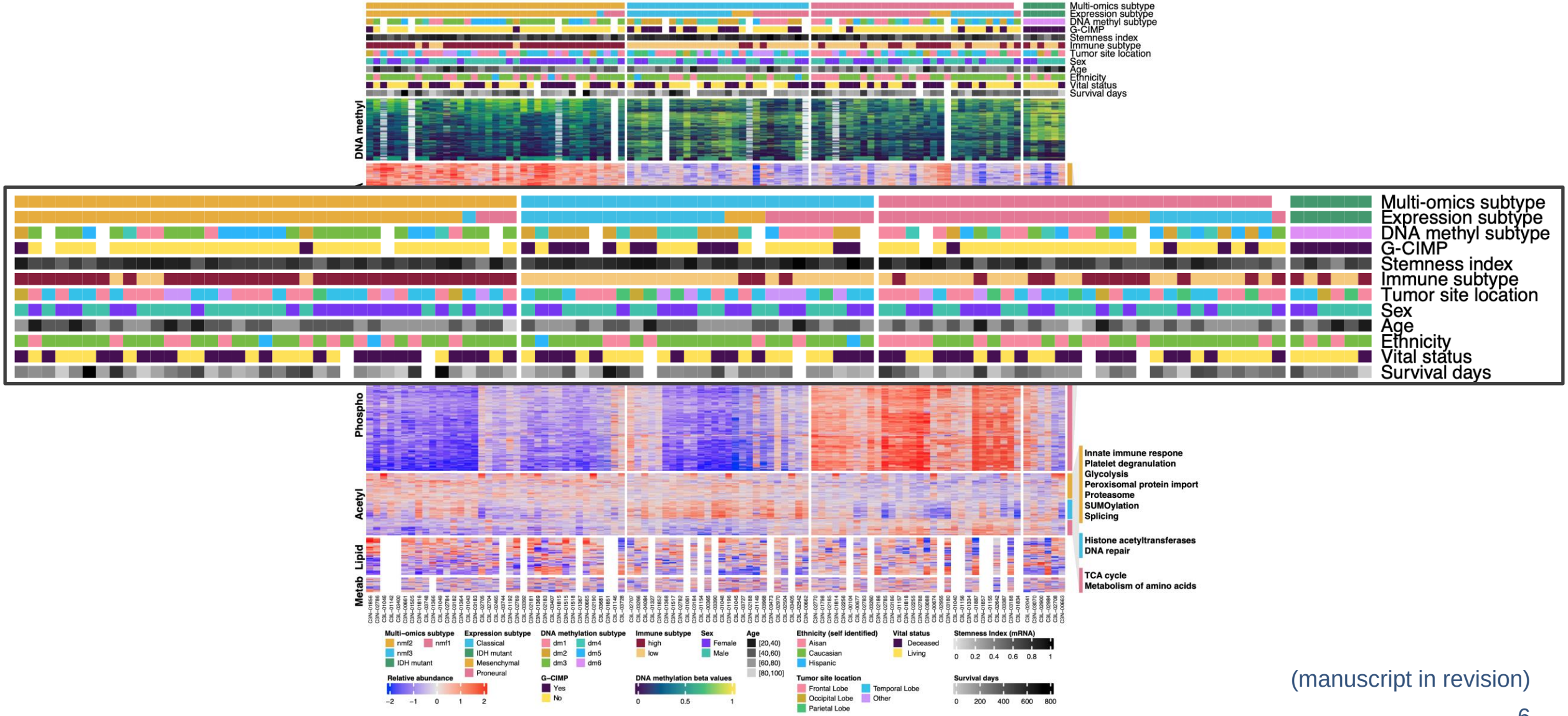
- Co-occurrence of structural variants and amplification

Two genes related to telomerase activity are highly altered:

- ATRX: associated with alternative lengthening of telomeres (ALT) phenotype
- TERT: promoter mutations using WGS

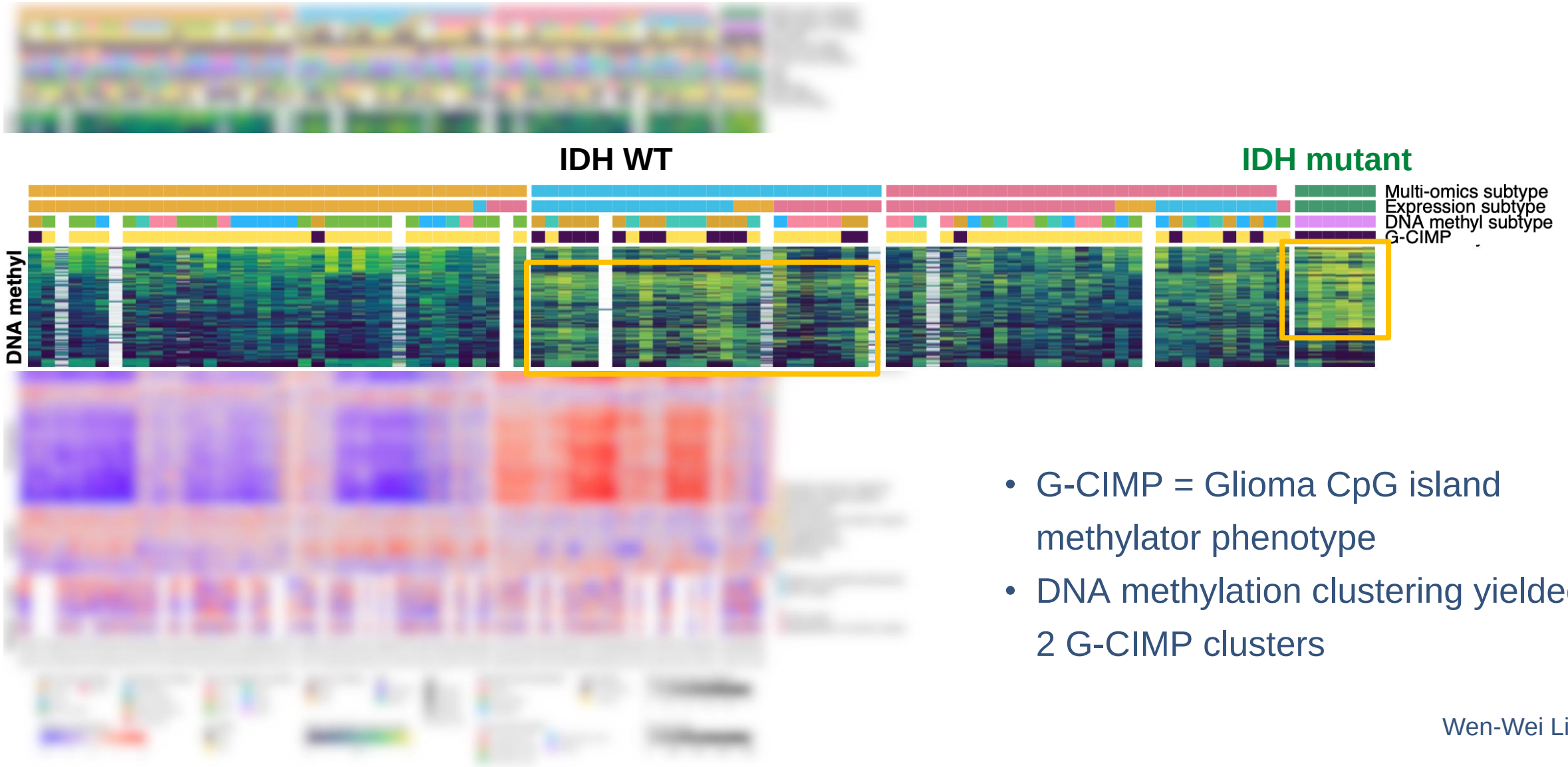
Song Cao
Alla Karpova

Overview of all tumor data



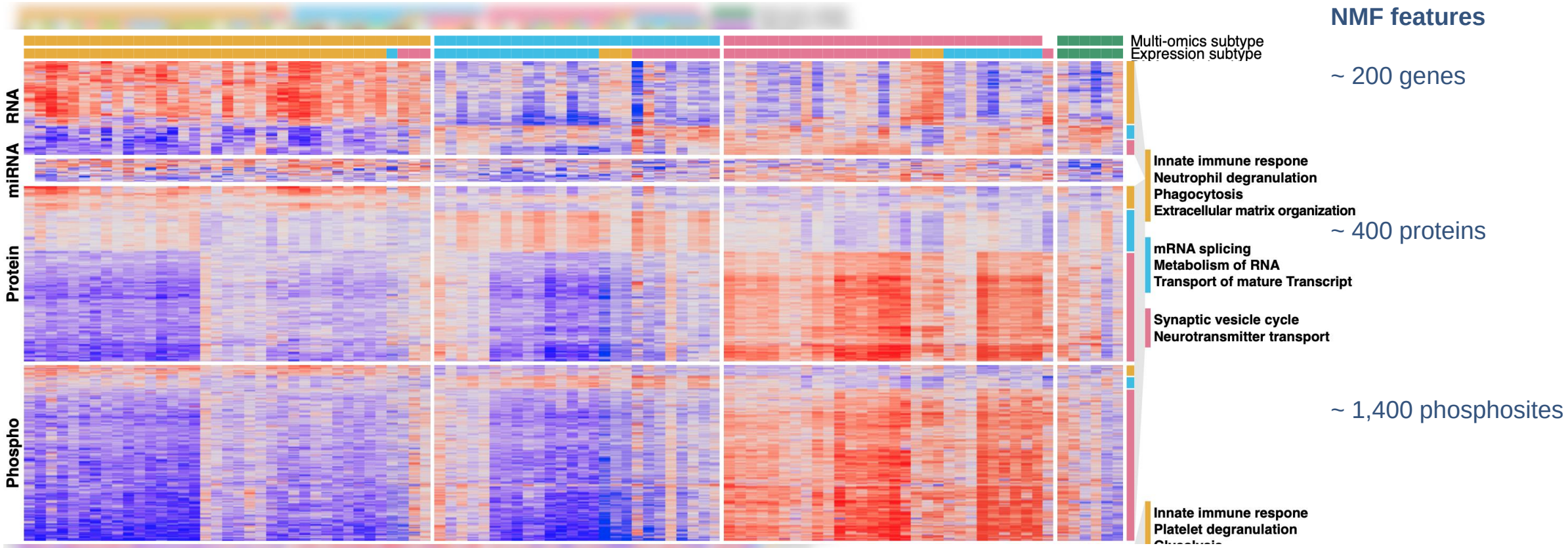
(manuscript in revision)

IDH mutant and G-CIMP phenotype



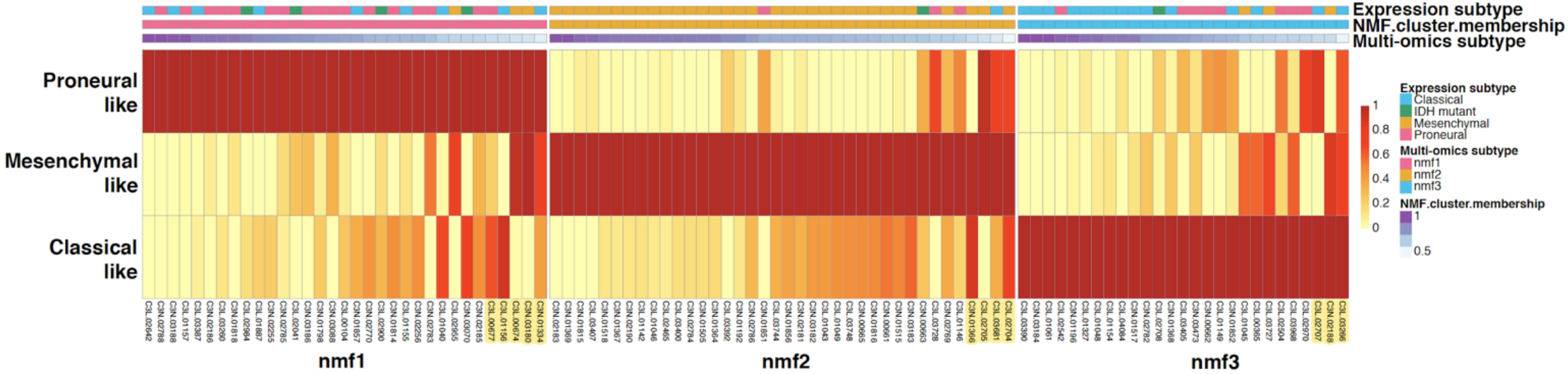
- G-CIMP = Glioma CpG island methylator phenotype
- DNA methylation clustering yielded 2 G-CIMP clusters

Multi-omics subtypes on IDH WT tumors



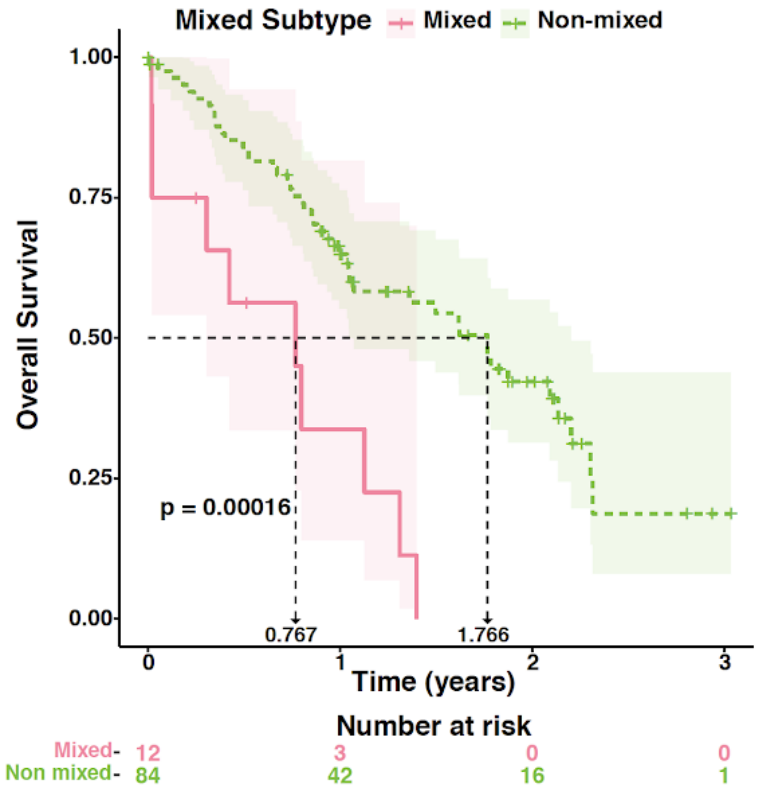
- Non-negative matrix factorization (NMF) using RNA, protein, and phospho
- Our multi-omics subtypes were based mostly on proteomic features

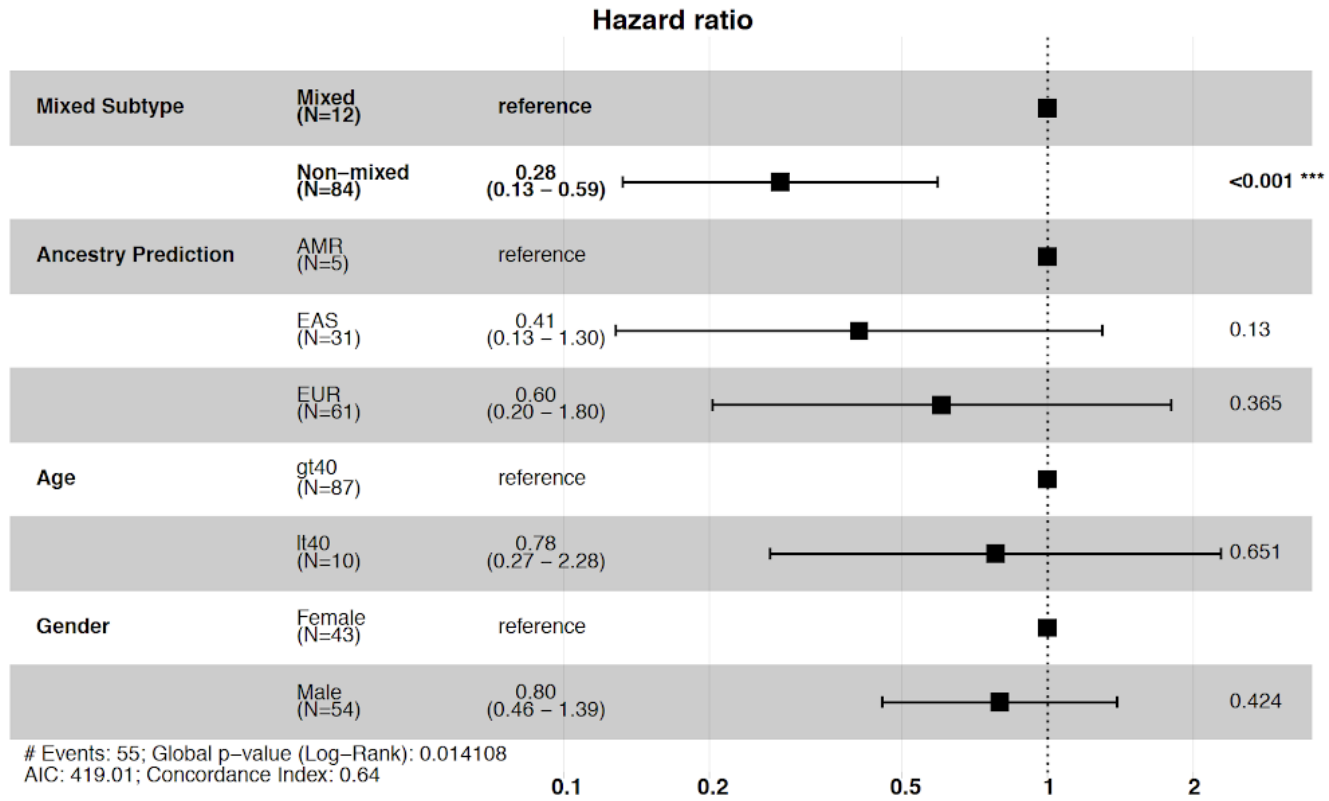
Mixed multi-omics subtype



- Mixed multi-omics subtype: tumors with signatures of more than 1 subtype (nmf membership < 0.55)
- Non-mixed multi-omics subtype: tumors with signature with mainly in 1 subtype (nmf membership ≥ 0.55)
- 12 tumors classified as mix subtype showing poor survival

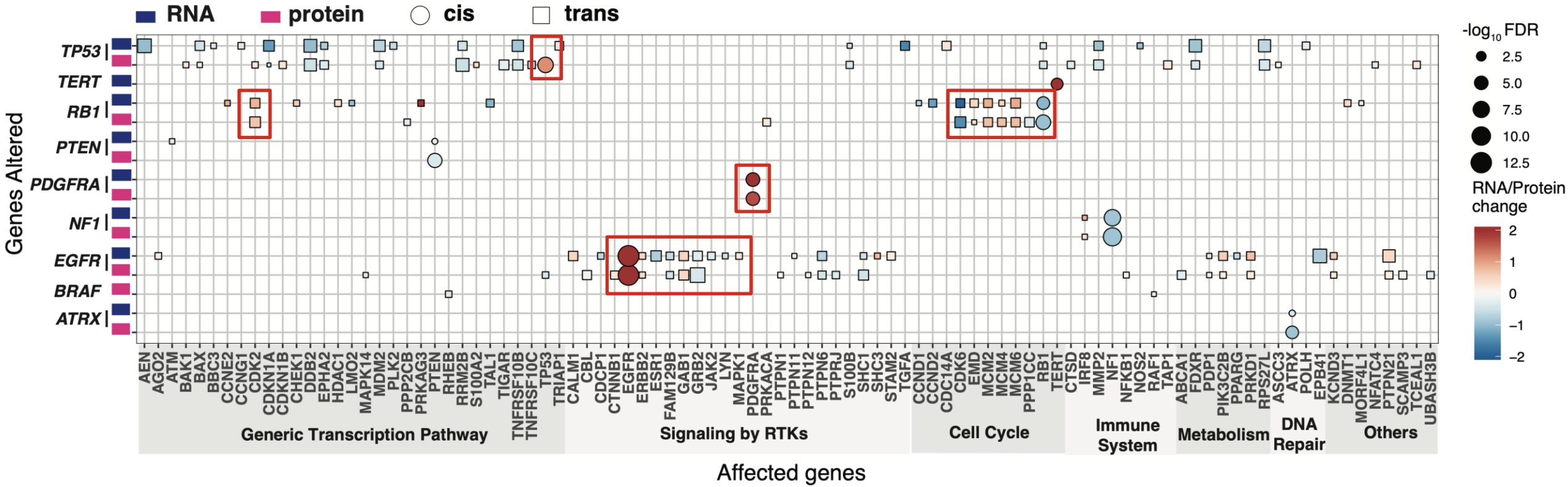
Mixed subtype tumors are associated with poor survival





Cox regression model of overall survival outcome between 12 mixed and 84 non-mixed tumor

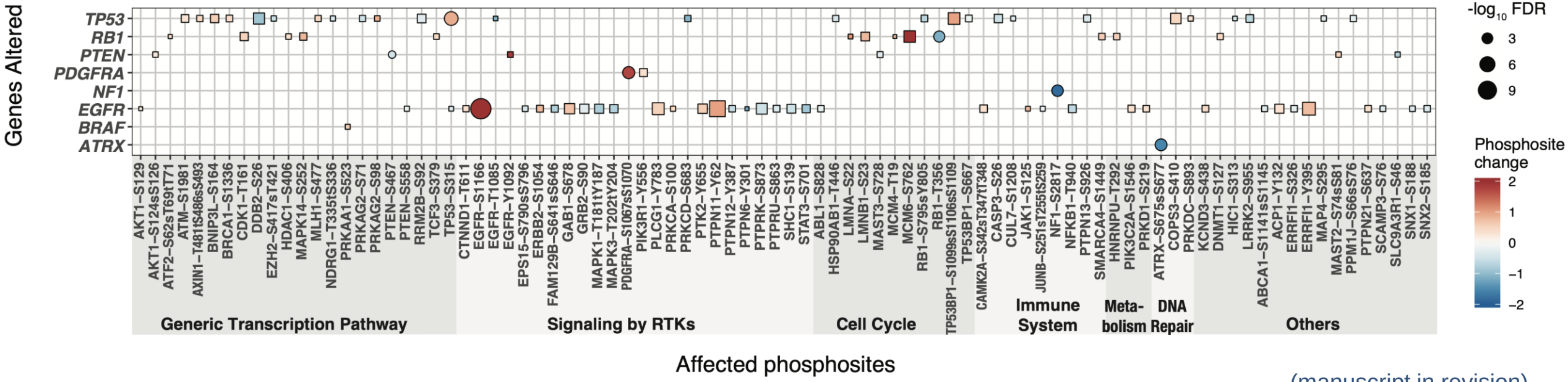
Genetic alteration impact on RNA and protein



(manuscript in revision)

Song Cao
Alla Karpova

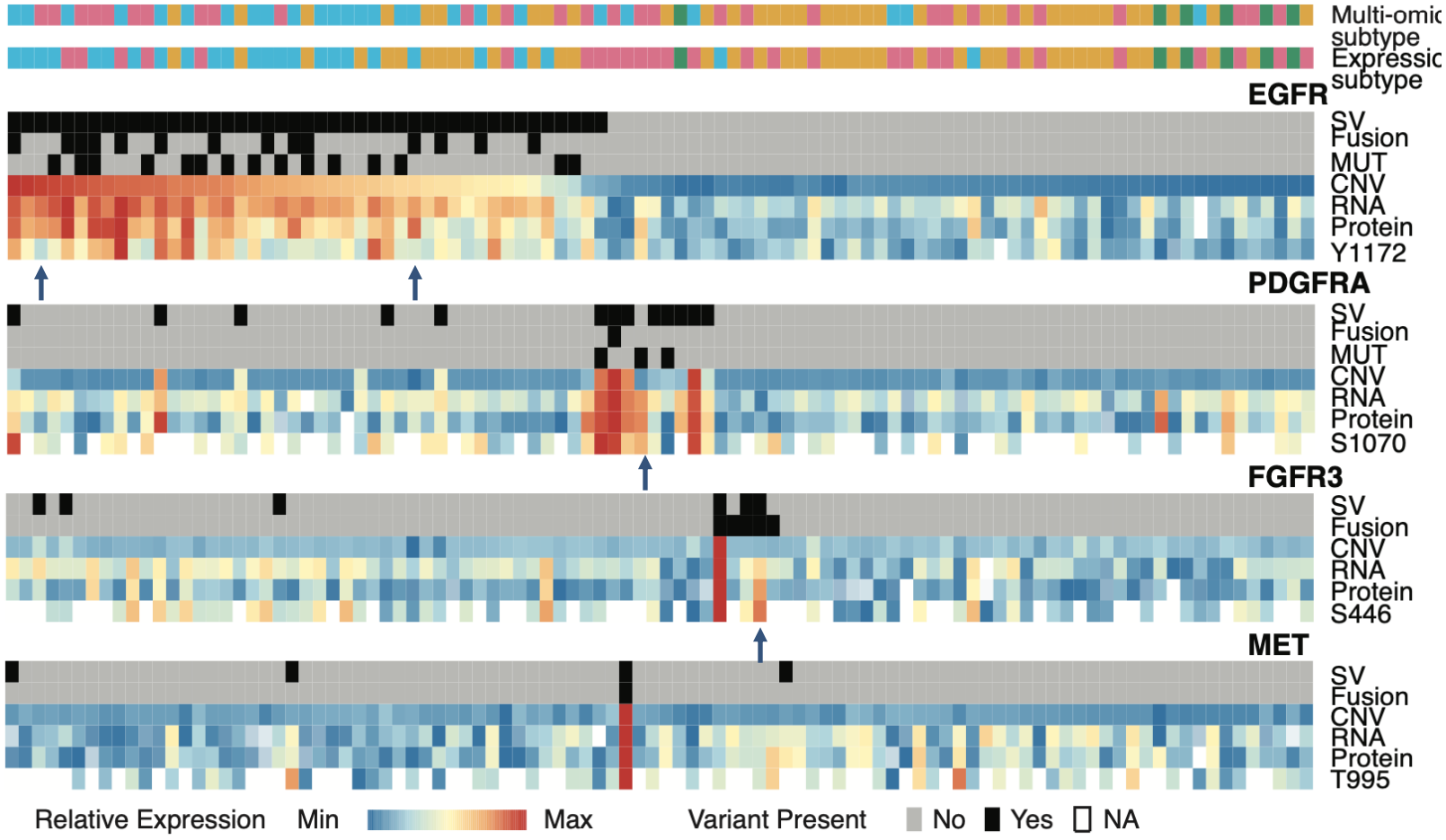
Genetic alteration impact on phosphoprotein



(manuscript in revision)

Song Cao
Alla Karpova

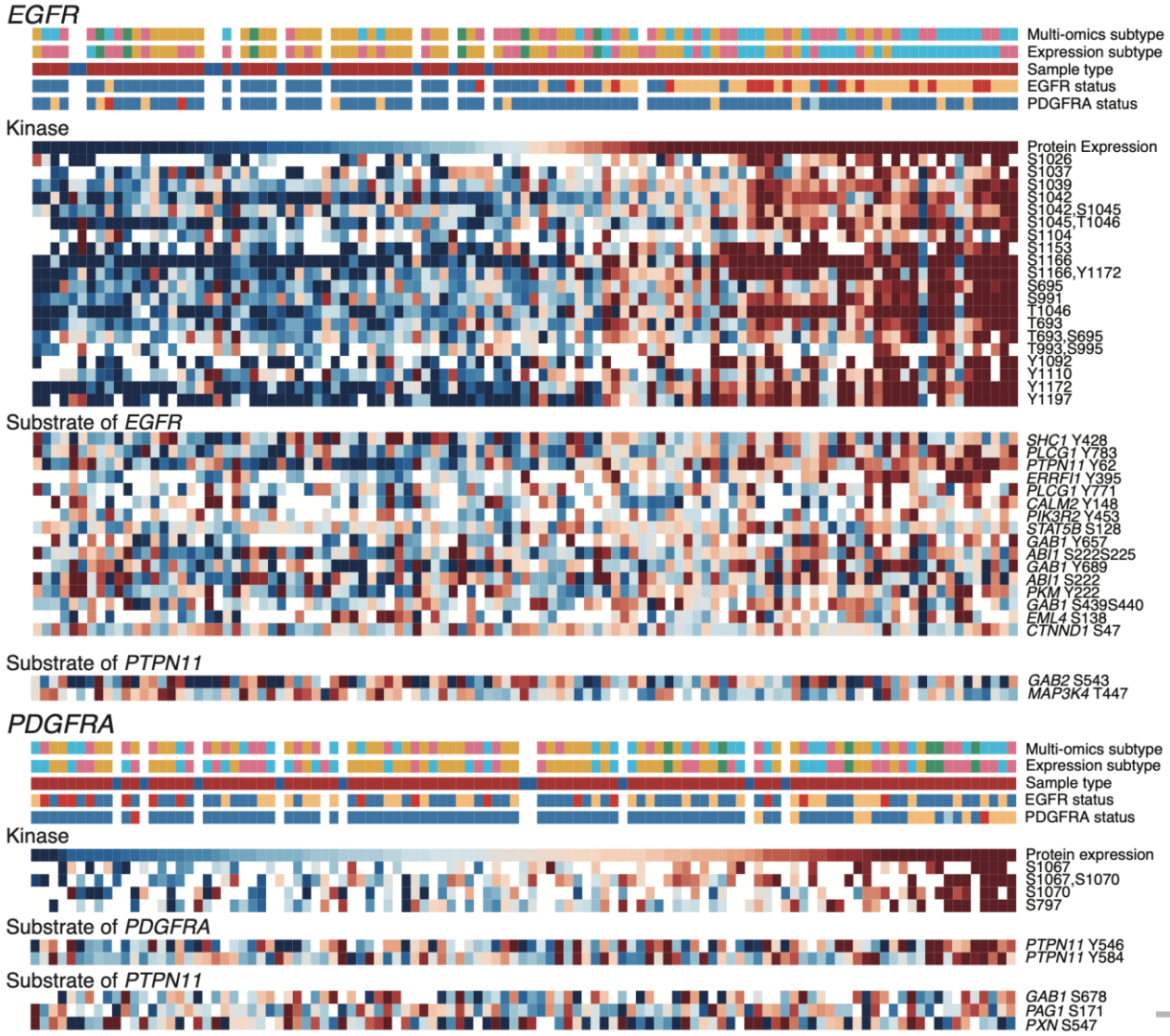
Proteomic and phosphoproteomic RTK activities can better stratify patients

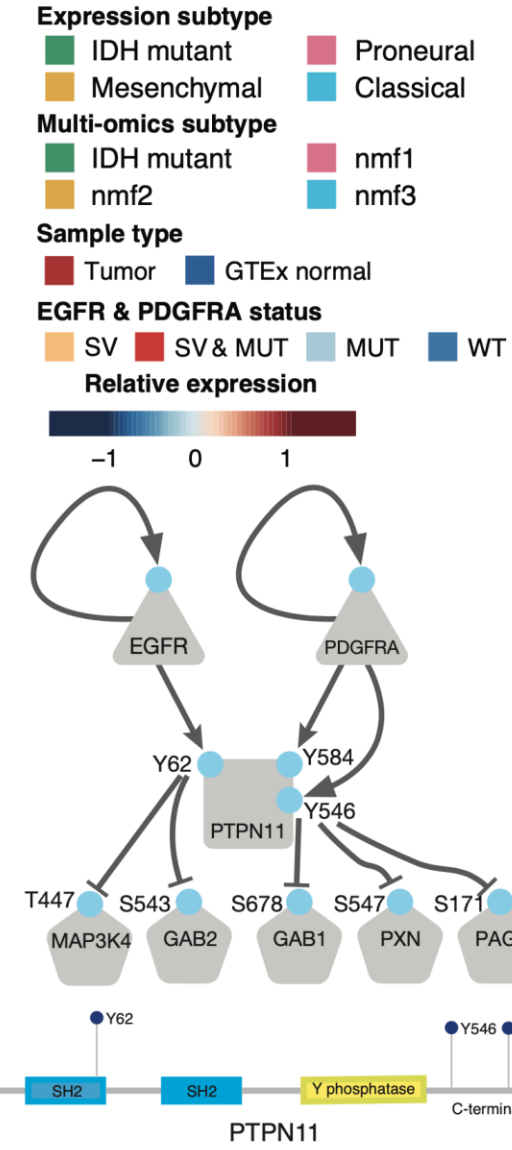


- RTK alterations are mutually exclusive, suggesting one hit is enough for GBM
- Altered but not active RTKs may not benefit from inhibition
- Potential new therapeutic target in the signal transducing cascade

Song Cao
Alla Karpova

Signal cascade of EGFR and PDGFRA identified shared downstream hub PTPN11

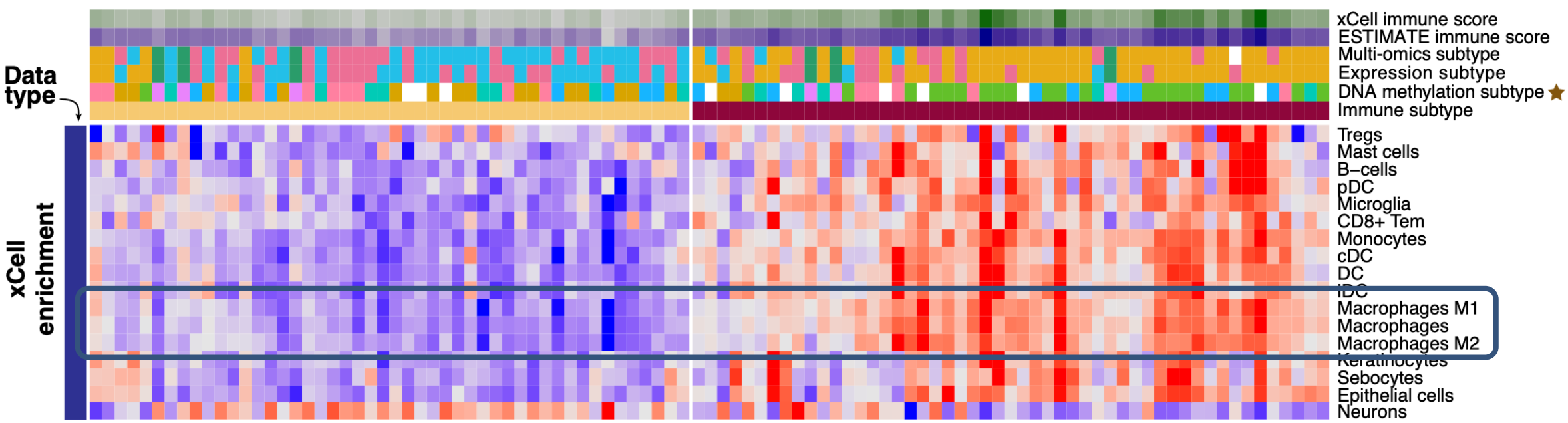


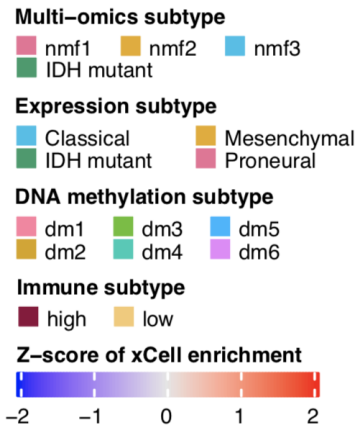


(manuscript in revision)

Song Cao
Alla Karpova

Enrichment for macrophages stratifies patients into immune high and low groups

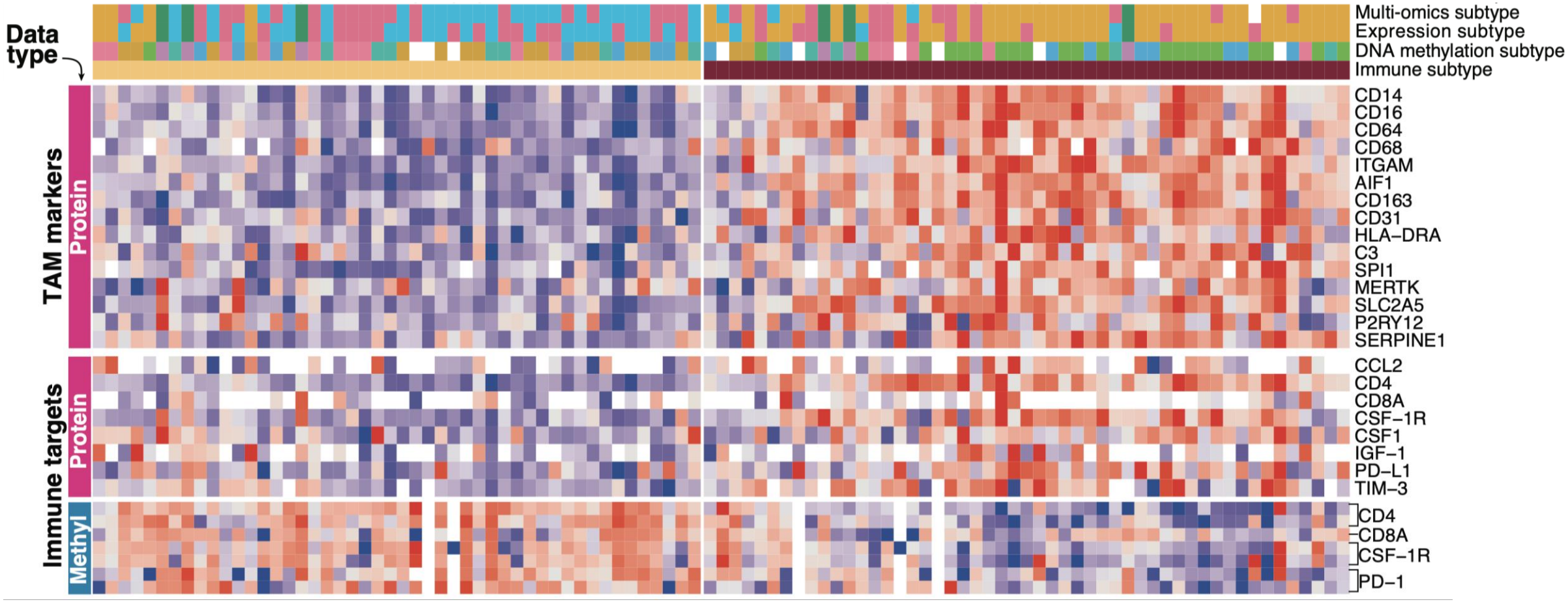




- Cell type enrichment by xCell based on bulk gene expression deconvolution
- T-cell oriented immunotherapy clinical trials have failed
- We identified two macrophage-based immune subtypes of GBM

Yize Li
Joseph Rothstein
Francesca Petralia

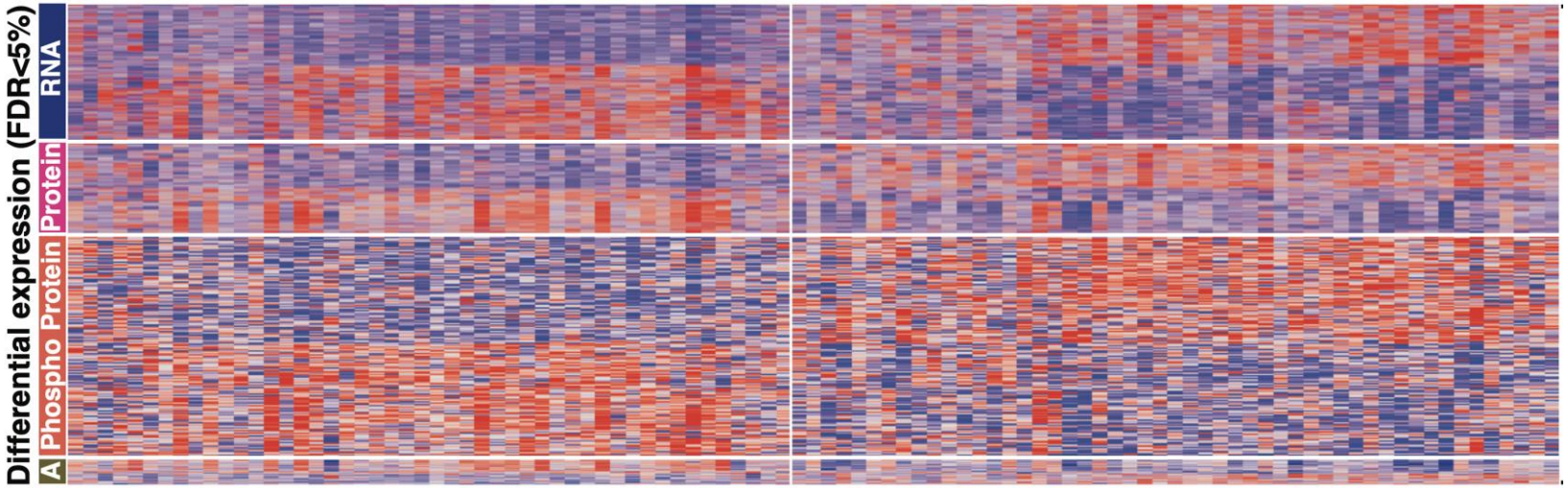
Distinctive protein and DNA methylation patterns between two immune subtypes



TAM = Tumor associated macrophages

Yize Li
Wen-Wei Liang

Consistent pathway enrichment from different data types between immune subtypes



Pathways enriched in immune low

- Cell cycle
- Chromatin modifying enzymes
- Chromatin organization
- DNA double-strand break repair
- G2M checkpoint
- Regulation of TP53 activity
- Spliceosome
- SUMOylation
- Telomere maintenance
- Regulation of PTEN transcription

Pathways enriched in immune high

- Apoptosis
- Exosome
- Hemostasis
- Immune system
- Innate immune system
- MAP2K and MAPK activation
- Neutrophil degranulation
- Platelet activation, signaling and aggregation
- Regulation of actin cytoskeleton
- Signaling by interleukins

Major takeaways

The most comprehensive proteogenomic characterization of GBM to date

Treatment naïve tumors to discover new therapeutic options and inform future clinical trial design

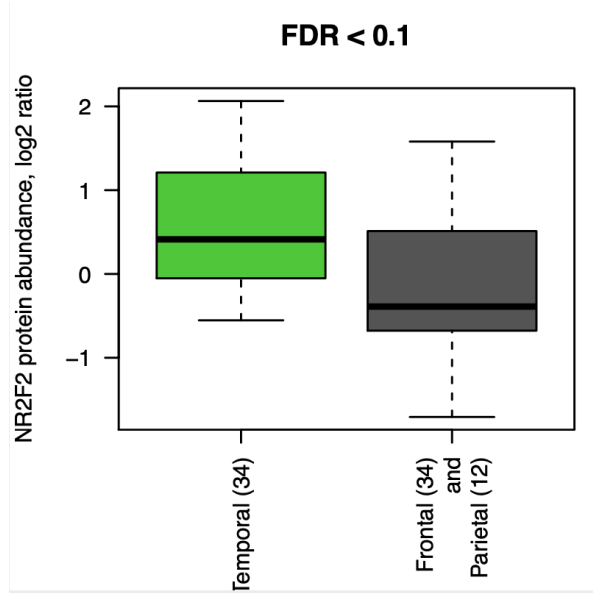
New biological insights and clinical implications captured at the protein level

- Three multi-omics subtypes with unique characteristics at the protein, phosphoprotein and acetylation levels
- Mixed subtype tumors are associated with worse prognosis
- Common regulatory hub PTPN11 by RTK
- Immune subtypes in relation to macrophage content

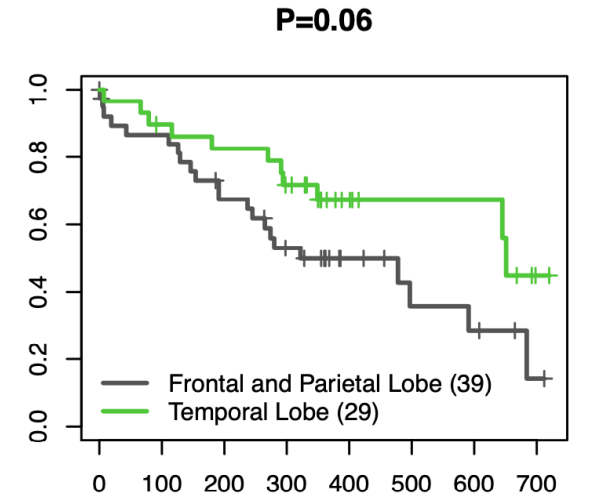
How TCIA images have helped the analysis

Curate the tumor laterality and location using MRI and CT imaging (Dmitry Rykunov, Alexis Demopoulos)

- NR2F2 showed significantly higher protein abundance in temporal tumors than in frontal and parietal tumors
- Potential temporal-specific tumor associated stem cell marker



Apply deep learning to identify histopathology image signatures between different molecular subgroups (Runyu Hong)



Identifying and visualizing molecular features on histopathology images

H&E imaging data from TCIA

- Image resolution: 20X, 10X, 5X
- Cutting each slide into 299x299 pixel tiles

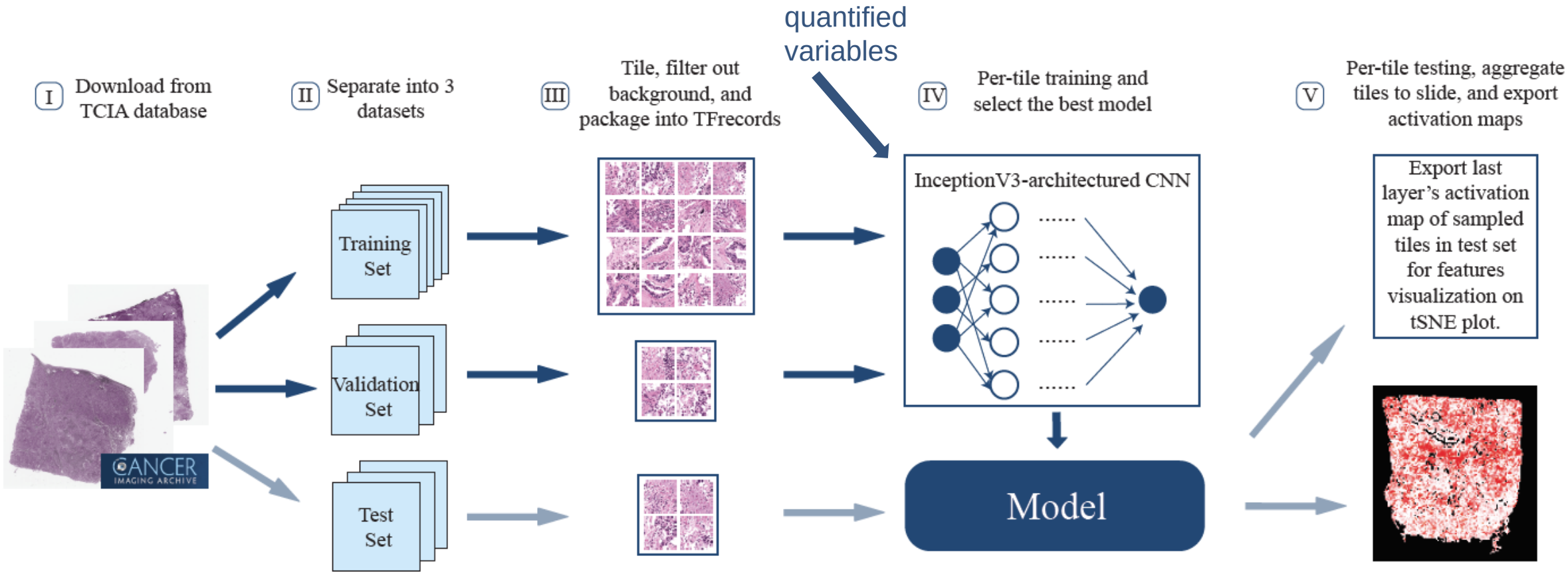
Modeling strategy

- 70% for training; 15% for validation; 15% for testing (patient-level separation)

Image stratifications

- G-CIMP: 113 slides, 86 patients
- Immune response: 129 slides, 99 patients
- Telomere length: 128 slides, 98 patients

H&E imaging analysis method overview



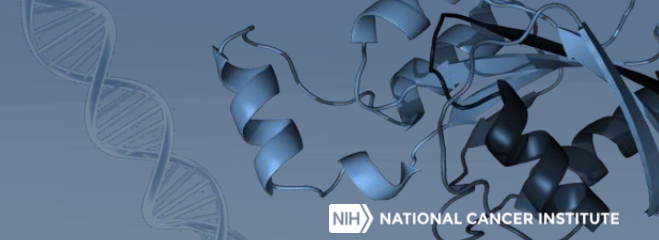
* CPTAC-LUAD Manuscript

Deep Learning architectures used:

Inception V1-V4; InceptionResnet V1&V2; Self-designed architecture (S1, F1, FS1)

F1 and FS1 takes quantified variables (cellularity, necrosis, tumor nuclei, age, tumor weight) in addition to the images.

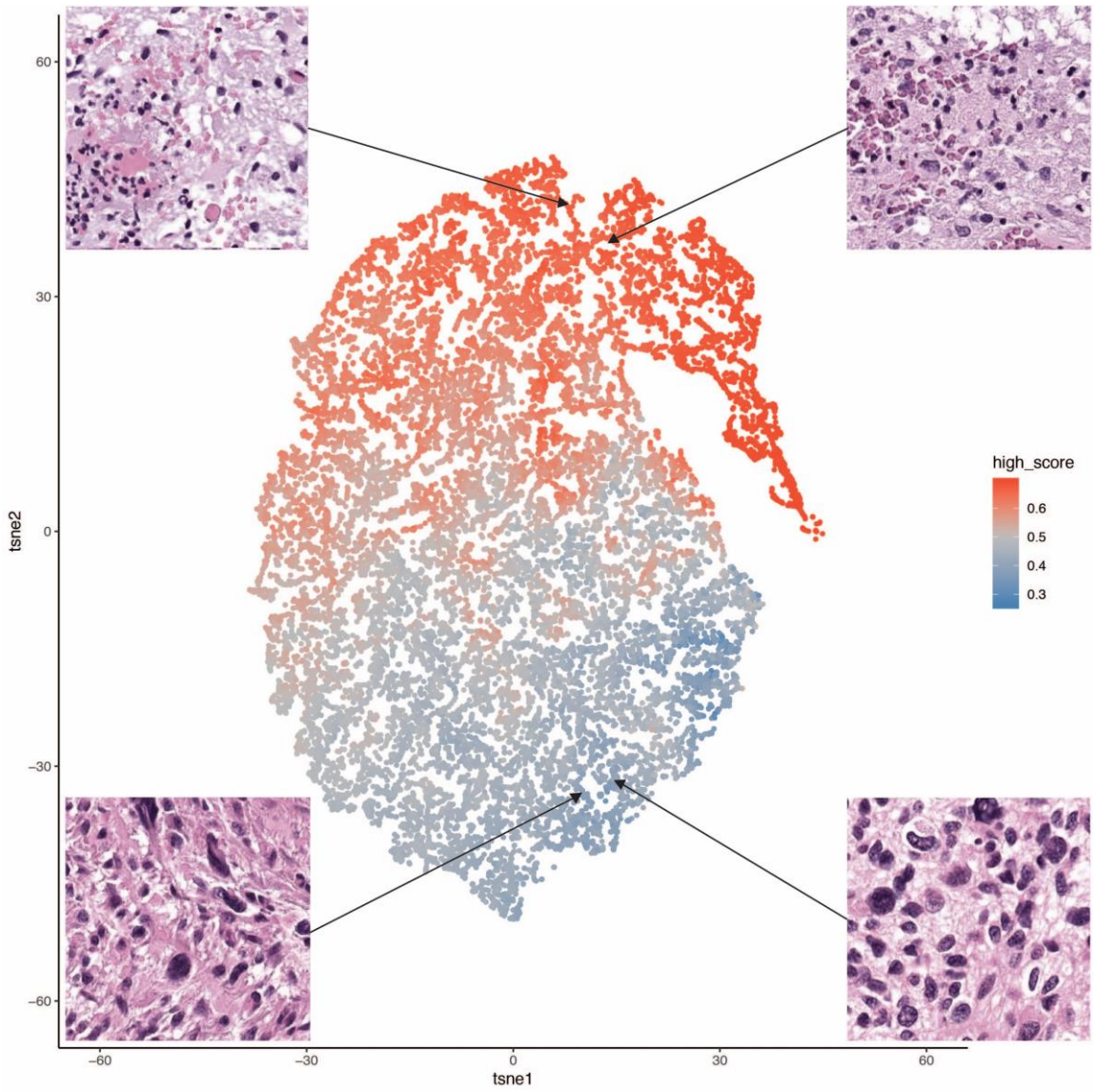
Promising results



Feature	Tile Resolution	Architecture	Patient AUROC	Tile AUROC
G-CIMP	10X	S1	1	0.89107
G-CIMP	20X	F1	0.81818	0.86861
G-CIMP	20X	S1	0.72727	0.68754
Immune	20X	F1	0.82143	0.65902
Immune	20X	FS1	0.71429	0.74064
Immune	10X	F1	0.75	0.65249

Feature	Tile Resolution	Architecture	Normal Patient AUROC	Short Patient AUROC	Long Patient AUROC	Normal Tile AUROC	Short Tile AUROC	Long Tile AUROC
Telomere	20X	F1	0.5	0.84615	0.84615	0.60721	0.6879	0.84768
Telomere	5X	FS1	0.77273	0.80769	0.53846	0.52926	0.62666	0.52052

Visualization of FS1 Immune 20X



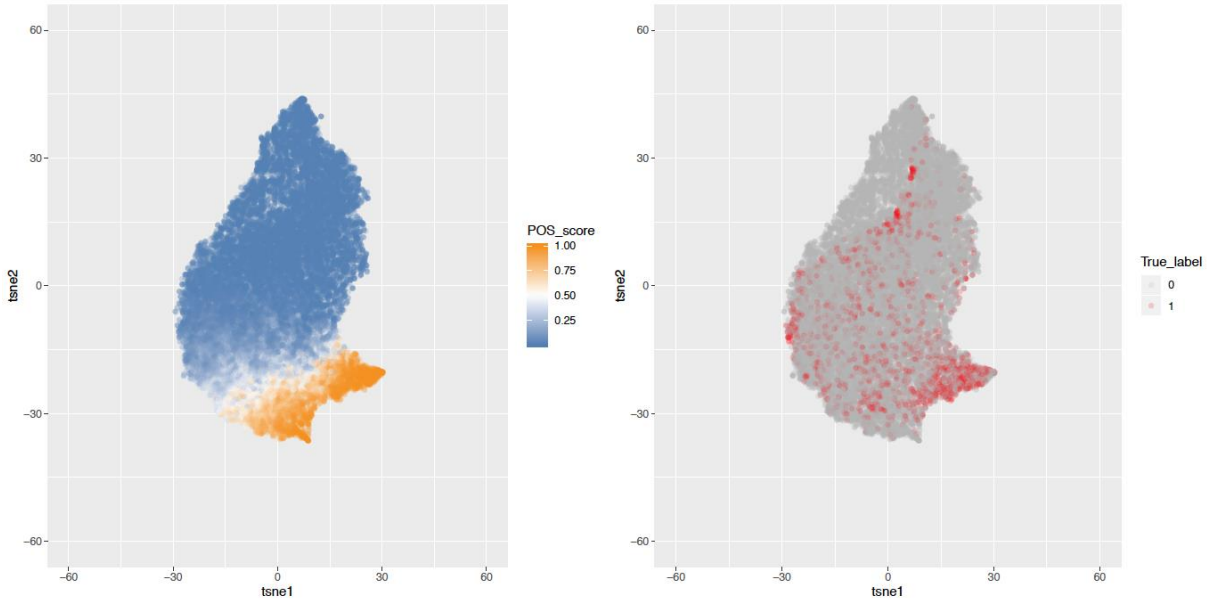
- 20,000 randomly sampled tiles from test set
- Activation maps (1,250-long vectors)

  -> 2D space
- 2 experienced pathologists examined the clusters
- Giant cells in immune low
- Inflammatory cells: 20% in high, 5% in low
- microcystic change and a few vascular structures in high

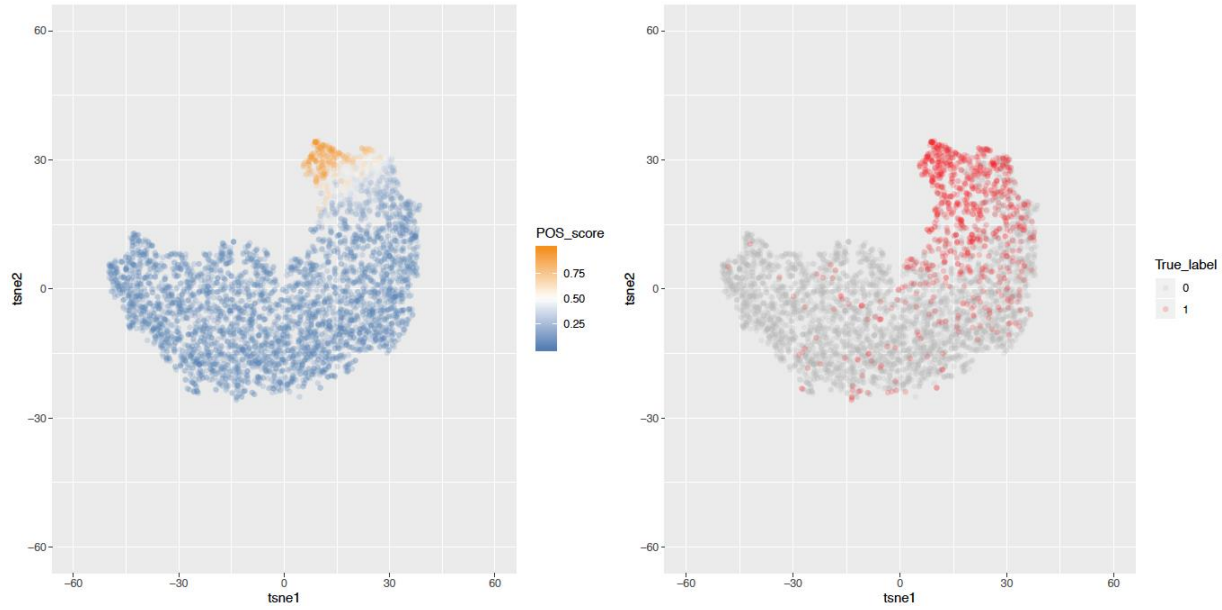
(manuscript in revision)

Visualization of G-CIMP

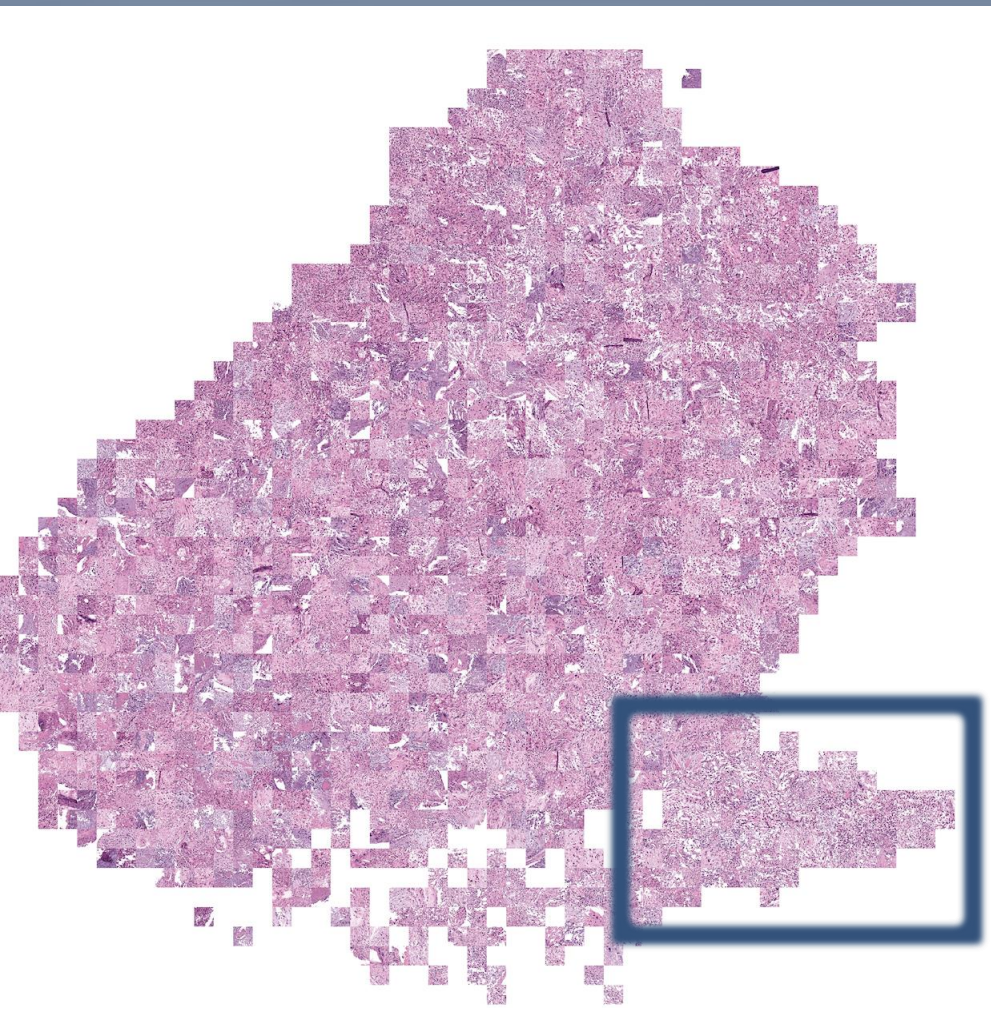
S1 G-CIMP 20X



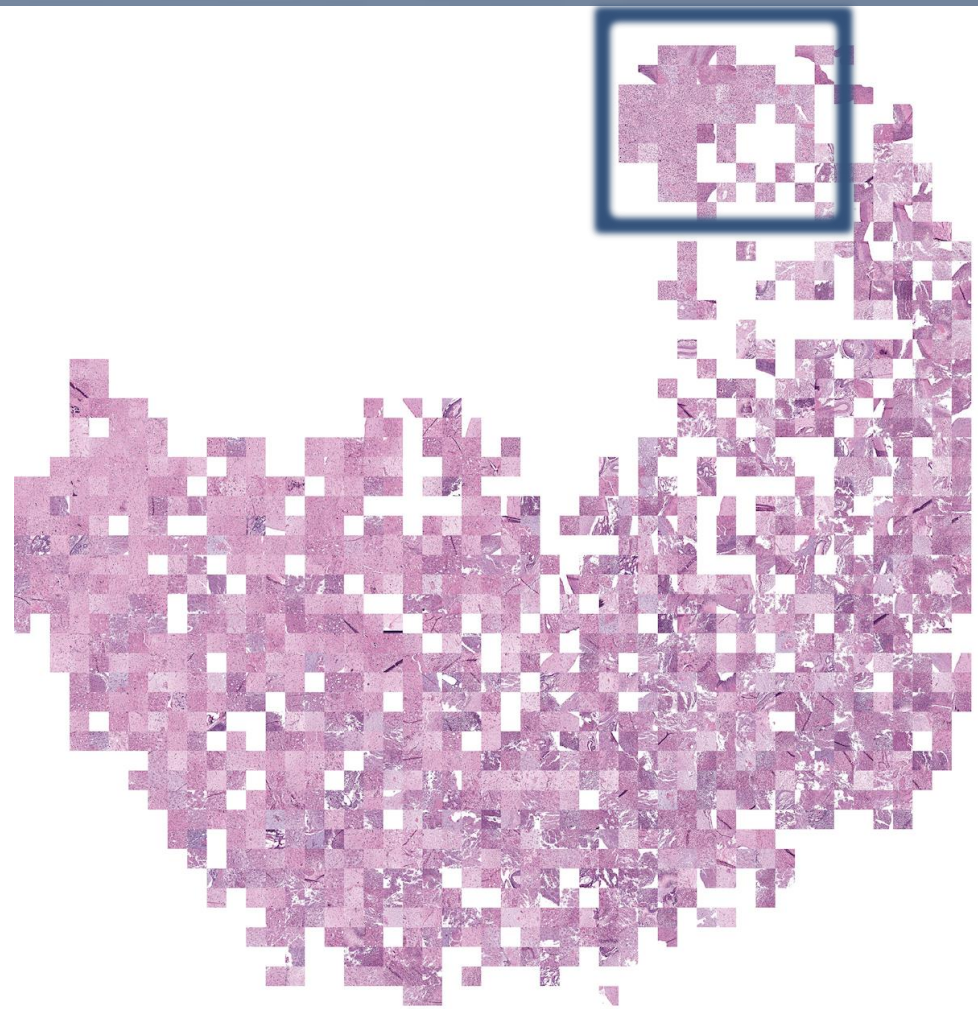
S1 G-CIMP 10X



Visualization of G-CIMP (continued)



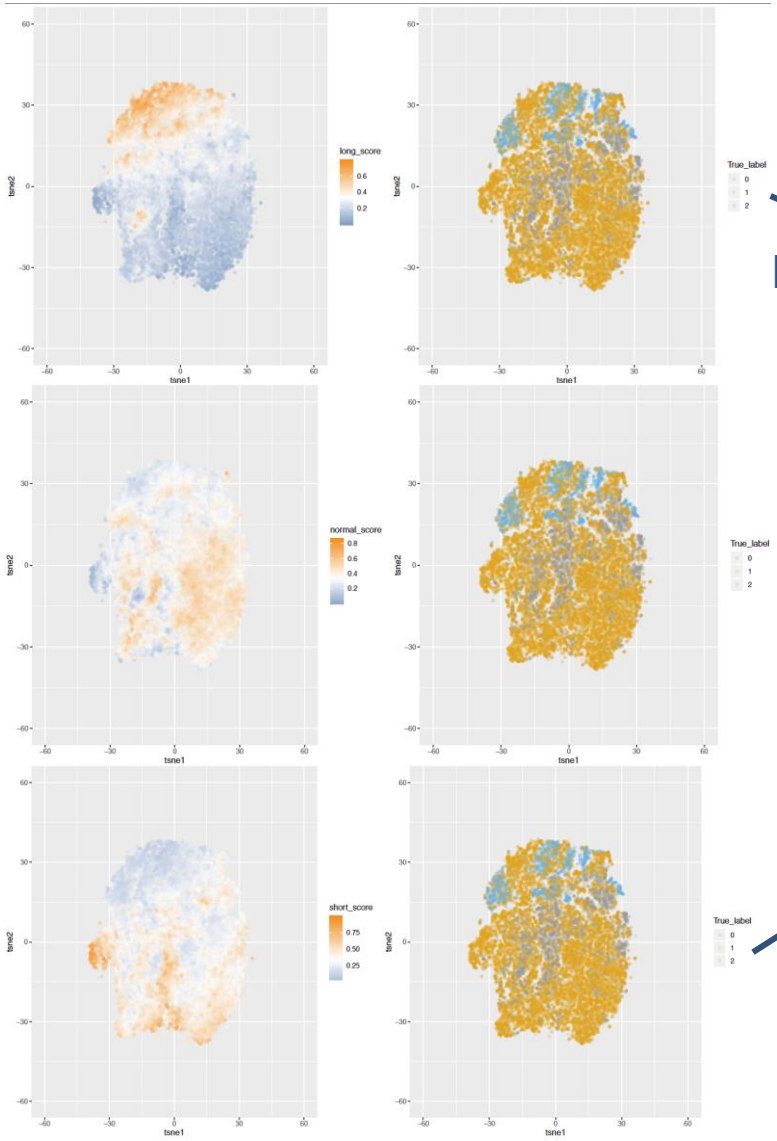
20X

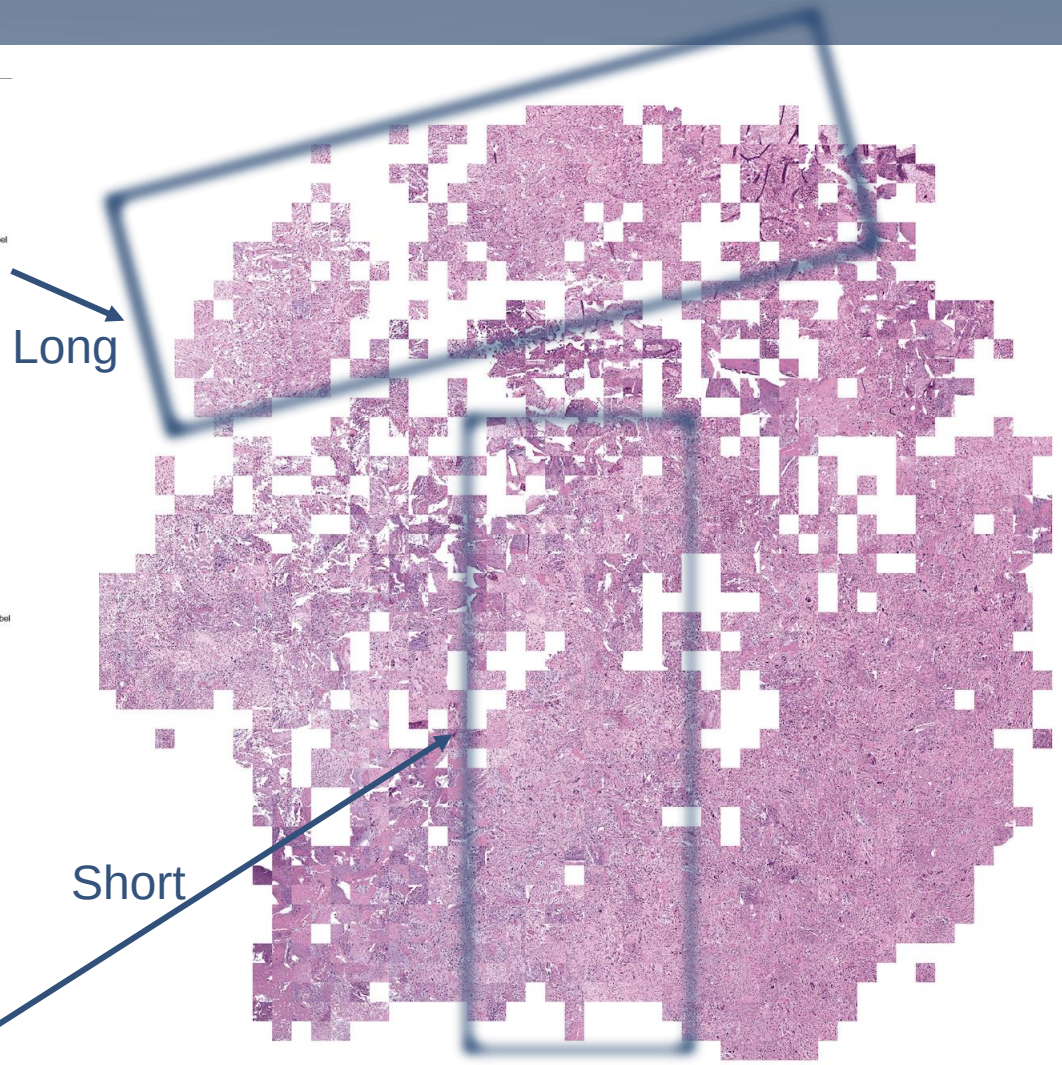


10X

- Focal vascular proliferation (glomeruloid vascular proliferation)
- High cellularity
- No microcystic changes

Visualization of F1 Telomere 20X





- Abundant vascular structures in long (likely glomeruloid vascular proliferation)
- focally larger cells and gemistocytes

Summary of the histopathology imaging model

- Found typical histopathology features in GBM corresponding to different subgroups
- Discovered novel histopathology features linking to biological and molecular responses
- Integrated deep learning models work better than image only models

Future/Ongoing directions

Validation using independent cohorts

Validation using orthogonal technologies

- Targeted MRM
- IHC
- Single nuclei RNA sequencing

New therapeutic targets based on the proteomic signatures

- *In silico* prediction using LINCS, CCLE, CTRP and PRISM datasets (Vasileios Stathias)

Imaging analysis

- Differentiation between additional molecular features (e.g. genetic alterations and subtypes)
- Molecular and imaging signatures in different tumor locations

Where to access the data

Genomic data:

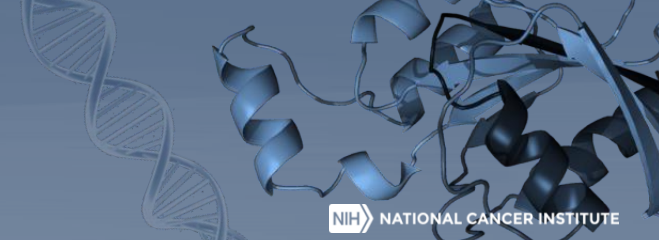
- Genomic Data Commons (GDC): https://portal.gdc.cancer.gov/

Proteomic data:

- CPTAC Glioblastoma (GBM) Discovery Study: https://cptac-data-portal.georgetown.edu/cptac/s/S048
- Proteomic Data Commons (PDC): https://pdc.cancer.gov/pdc/

	Radiology Imaging Statistics	Pathology Imaging Statistics
Modalities	CR, CT, MR, SC	Pathology
Number of Patients	66	189
Number of Studies	164	N/A
Number of Series	1,771	N/A
Number of Images	156,493	510
Images Size (GB)	39.8	112

GBM proteogenomic characterization team



PNNL
- Tao Liu
- Karin Rodland
- Richard D. Smith
- Jennifer Kyle
- Marina Gritsenko
- Chia-Feng Tsai
- Vladislav A. Petyuk
- Jamie Moon
- Rosalie K. Chu
- Karl K. Weitz
- Ronald J. Monroe
- Matthew E. Monroe
- Rui Zhao
- Kelly G. Stratton
- Lisa M. Bramer
- Erika Zink
- Sneha P. Couvillion
- Kent J. Bloodsworth

WashU
- Li Ding
- Milan G. Chheda
- Albert H. Kim
- Feng Chen
- Liang-Bo Wang
- Alla Karpova
- Song Cao
- Yize Li
- Yige Wu
- Wen-Wei Liang
- Michael C. Wendl
- Wagma Caravan
- Daniel Cui Zhou
- Xiaolu Yang
- Houxiang Zhu
- Matthew A. Wyczalkowski
- Shuangjia Lu
- Jessika Baral
- Lijun Yao

MSSM
- Pei Wang
- Dmitry Rykunov
- Joseph Rothstein
- Francesca Petralia
- Boris Reva
- Xiaoyu Song
- Jiayi Ji
- Weiping Ma
- Seungyeul Yoo
- Azra Krek
- Weiva Sieh

NYU
- David Fenyö
- Kelly V. Ruggles
- Runyu Hong
- MacIntosh Cornwell
- Emily Kawaler
- Wenke Liu

NCI
- Henry Rodriguez
- Mehdi Mesri
- Chelsea Newton
- Scott Jewell
- Mathangi Thiagarajan
- Tara Hiltke
- Ana I. Robles
- Chris R. Kinsinger
- Emily S. Boja
- Karen A. Ketchum

BYU
- Samuel H. Payne
- Lindsey K. Olsen
- Brittany Henderson
- Caleb M. Lindgren
- Hannah Boekweg

BCM
- Bing Zhang
- Yongchao Dou

Broad
- D. R. Mani
- Karsten Krug
- Shankha Satpathy

UMiami
- Steven Chen
- Antonio Colaprico
- Vasileios Stathias

GBM working group
- Jill S. Barnholtz-Sloan
- Adam C. Resnick
- Qing Kay Li
- Sandra Cottingham
- Alexey I. Nesvizhskii
- Antonio Iavarone
- Maciej Wiznerowicz
- Alexis Demopoulos
- Simina M. Boca