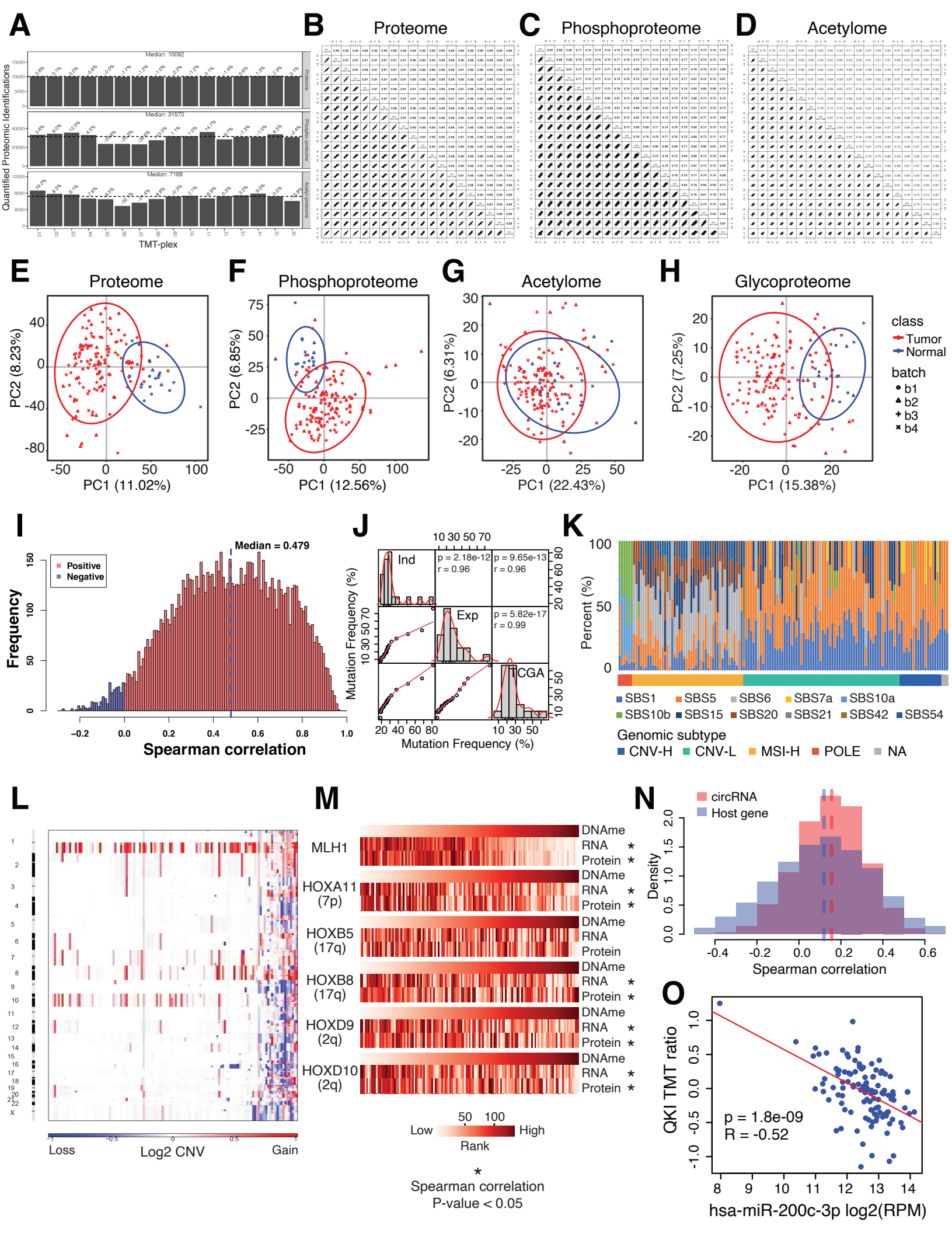

A
Median: 10092
Median: 31570
Median: 7188
Quantified Proteomic Identifications
Proteome
Phospho-proteome
Acetyl-proteome
TMT-plex
B
Proteome
C
Phosphoproteome
D
Acetylome
E
Proteome
PC2 (8.23%)
PC1 (11.02%)
F
Phosphoproteome
PC2 (6.85%)
PC1 (12.56%)
G
Acetylome
PC2 (6.31%)
PC1 (22.43%)
H
Glycoproteome
PC2 (7.25%)
PC1 (15.38%)
class
Tumor
Normal
batch
b1
b2
b3
b4
I
Median = 0.479
Positive
Negative
Frequency
Spearman correlation
J
Ind
Exp
TCGA
p = 2.18e-12
r = 0.96
p = 9.65e-13
r = 0.96
p = 5.82e-17
r = 0.99
Mutation Frequency (%)
K
Percent (%)
SBS1
SBS5
SBS6
SBS7a
SBS10a
SBS10b
SBS15
SBS20
SBS21
SBS42
SBS54
Genomic subtype
CNV-H
CNV-L
MSI-H
POLE
NA
L
Loss
Log2 CNV
Gain
M
MLH1
HOXA11 (7p)
HOXB5 (17q)
HOXB8 (17q)
HOXD9 (2q)
HOXD10 (2q)
DNAme
RNA
Protein
Low
High
Rank
Spearman correlation
P-value < 0.05
N
circRNA
Host gene
Density
Spearman correlation
O
QKI TMT ratio
p = 1.8e-09
R = -0.52
hsa-miR-200c-3p log2(RPM)

Figure S1: QC metrics for the proteogenomic data, related to Figure 1.

(A) Bar graphs of quantified proteins (top), phosphosites (middle), and acetylation sites (bottom). X-axis corresponds to the TMT-plex and the y-axis corresponds to the number of quantified proteomic identifications. Median is denoted by the dashed line and percentages indicate deviation from the median.

(B-D) Longitudinal performance was tested by repeated proteome (B), phosphoproteome (C) and acetylome (D) analysis of aliquots of the same patient-derived xenograft QC samples in standalone TMT-11 plexes, along with the EC study samples. Scatterplots and Pearson correlations comparing individual replicate measurements are shown.

(E-H) Principal component analysis (PCA) plots for whole proteome (E), phosphoproteome (F), acetylome (G), and glycoproteome (H), colored by tumor (red) and normal (blue) samples. Shapes of plot points represent data collection batches.

(I) Histogram of mRNA and protein level Spearman correlation values. Values above 0 are shown in red, while values less than 0 are shown in blue.

(J) Density plots showing correlations of mutation frequencies of top mutated genes across independent (Ind), exploratory (Exp), and TCGA cohorts. P-values derived from Pearson Correlation.

(K) Mutational signatures of the independent cohort.

(L) Heatmap of the CNV landscape of the independent cohort across the genome (rows) per sample (columns).

(M) Heatmap of correlations between DNA methylation and mRNA/protein levels for *MLH1* and *HOX* family genes. Spearman correlation p-value < 0.05 denoted by stars.

(N) Barplot of correlation between pairs of circRNAs and host genes.

(O) Correlation between hsa-miR-200c-3p expression and protein level of QKI, a circRNA regulator.

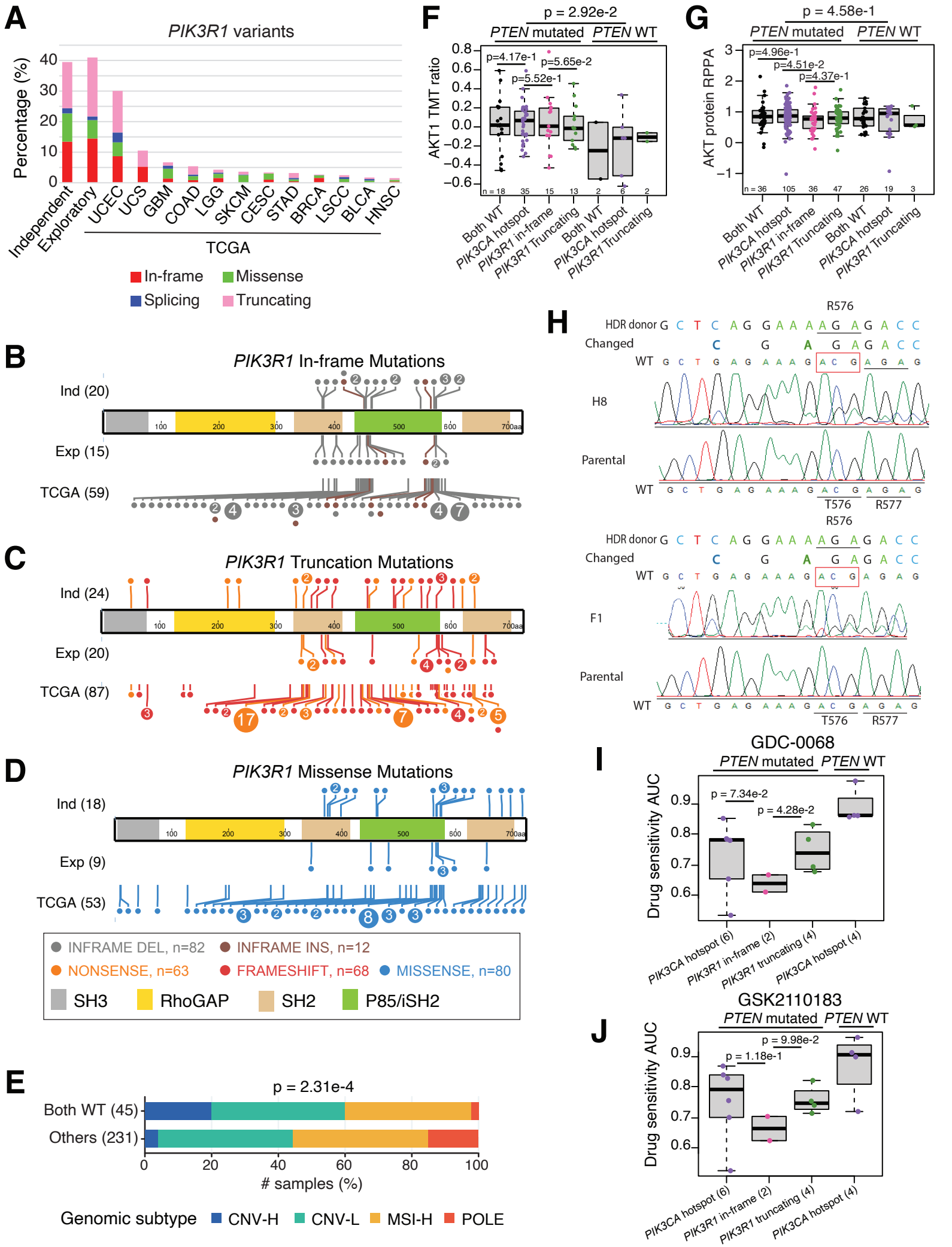

A
PIK3R1 variants
Percentage (%)
Independent
Exploratory
UCEC
UCS
GBM
COAD
LGG
SKCM
CESC
STAD
BRCA
LSCC
BLCA
HNSC
TCGA
In-frame
Missense
Splicing
Truncating
B
PIK3R1 In-frame Mutations
Ind (20)
Exp (15)
TCGA (59)
C
PIK3R1 Truncation Mutations
Ind (24)
Exp (20)
TCGA (87)
D
PIK3R1 Missense Mutations
Ind (18)
Exp (9)
TCGA (53)
INFRAME DEL, n=82
INFRAME INS, n=12
NONSENSE, n=63
FRAMESHIFT, n=68
MISSENSE, n=80
SH3
RhoGAP
SH2
P85/iSH2
E
p = 2.31e-4
Both WT (45)
Others (231)
samples (%)
Genomic subtype
CNV-H
CNV-L
MSI-H
POLE
F
p = 2.92e-2
PTEN mutated
PTEN WT
AKT1 TMT ratio
p=4.17e-1
p=5.65e-2
p=5.52e-1
Both WT
PIK3CA hotspot
PIK3R1 in-frame
PIK3R1 Truncating
G
p = 4.58e-1
AKT protein RPPA
p=4.96e-1
p=4.51e-2
p=4.37e-1
H
R576
HDR donor
Changed
WT
H8
Parental
T576
R577
F1
I
GDC-0068
p = 7.34e-2
p = 4.28e-2
Drug sensitivity AUC
PIK3CA hotspot (6)
PIK3R1 in-frame (2)
PIK3R1 truncating (4)
J
GSK2110183
p = 9.98e-2
p = 1.18e-1

Figure S2: *PIK3R1* in-frame indels are clustered and associated with sensitivity to AKT inhibitors, related to Figure 2.

(A) Barplot of *PIK3R1* variant frequencies across varying tumor types in both CPTAC EC cohorts, TCGA EC cohort, and other TCGA cohorts.

(B-D) Lollipop plots of *PIK3R1* in-frame, truncating, and missense mutations across three cohorts.

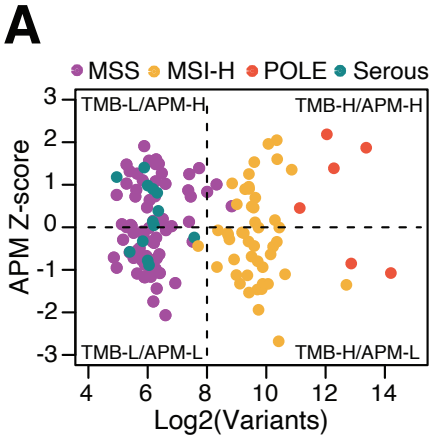
(E) Genomic subtype distribution across TCGA *PTEN* mutated tumors.

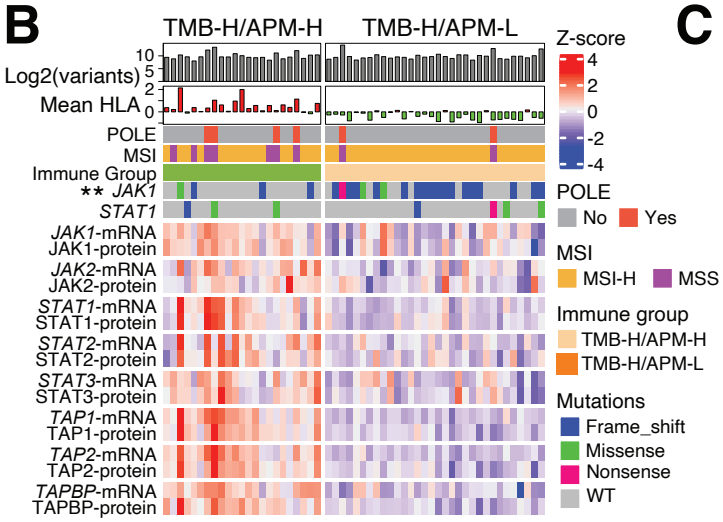
(F-G) Boxplots comparing AKT1 protein levels across *PTEN*, *PIK3CA*, and *PIK3R1* variants for the Independent cohort (F) and TCGA cohort (G). P-values derived from Student's t-test.

(H) Sanger sequencing confirmation of CRISPR-Cas9 created T576 deletion.

(I-J) Boxplots comparing responses of DepMap EC cell lines to (I) GDC-0068 (Ipatasertib) and (J) GSK2110183 (Afuresertib). P-values derived from Student's t-test.

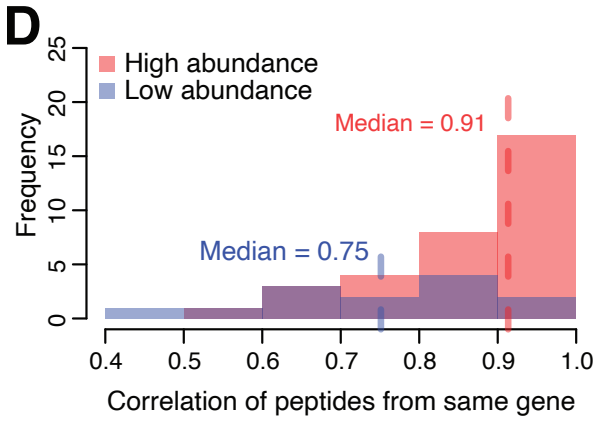
Boxplots: Box portion represents Interquartile range (IQR), midline corresponds to the median, and whiskers range from the minimum (bottom) and maximum (top) variability outside the first and third quartiles (Q1 and Q3). Outliers are shown as points above whiskers.

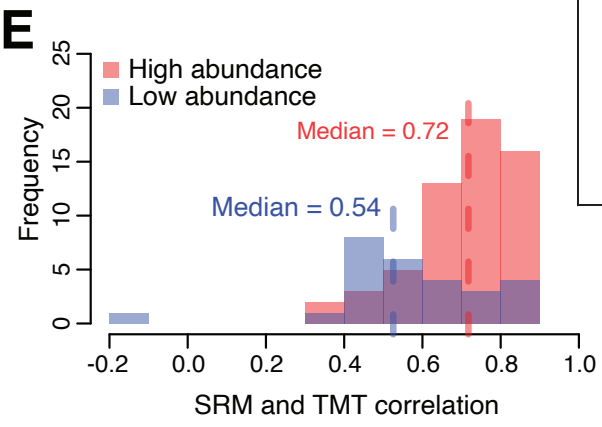




C

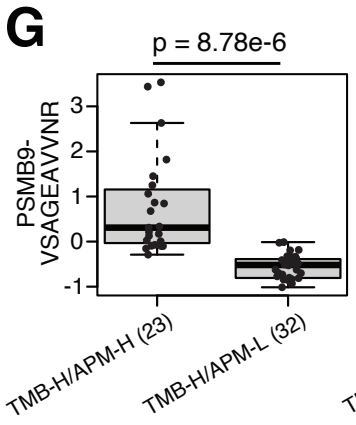
Pathway	Protein	Peptide
JAK/ STAT	JAK1	JAK1-DLVQGEHLGR
		JAK1-NVLVESEHQVK
	JAK2	JAK2-EVGDYGQLHETEVLLK
		JAK2-YDPLQDNTGEVVAVK
	STAT1	STAT1-ELSAVTFPDIIR
		STAT1-LLGPNASPDGLIPWTR
	STAT2	STAT2-EVLQSLPLTEIIR
		STAT2-HYQLLTEENIPENPLR
	STAT3	STAT3-GLSIEQLTTLAEK
		STAT3-LLGPGVNYSGCQITWAK
APM	B2M	B2M-IQVYSR
		B2M-VEHSDLSFSK
	ERAP1	ERAP1-ILASTQFEPTAAR
		ERAP1-TQEFPQILTLIGR
	ERAP2	ERAP2-FLQGVFQEDPEWR
		ERAP2-ILAVTDFEPTQAR
	HLA.A	HLA.A-FDSDAASQR
		HLA.A-FIAVGYVDDTQFVR
	HLA.B	HLA.B-APWIEQEGPEYWDR
	HLA.C	HLA.C-AYLEGTCVEWLR
		HLA.C-WAAVVVPSGQEQR
	PSMB10	PSMB10-LPFTALGSGQDAALAVLEDR
		PSMB10-YHFVPGTTAVLTQTVK
	PSMB8	PSMB8-FQHGVIAAVDSR
		PSMB8-VESTDVSDLLHQYR
	PSMB9	PSMB9-VILGNELPK
		PSMB9-VSAGEAVVNR
	TAP1	TAP1-ELISWGAPGSADSTR
		TAP1-QVAAVGQEPQVFGR
	TAP2	TAP2-EQLFSSLLR
		TAP2-NNIAYGLQSCEDDK
	TAPBP	TAPBP-IHHPSLPASGR
		TAPBP-WASGLTPAQNCPR

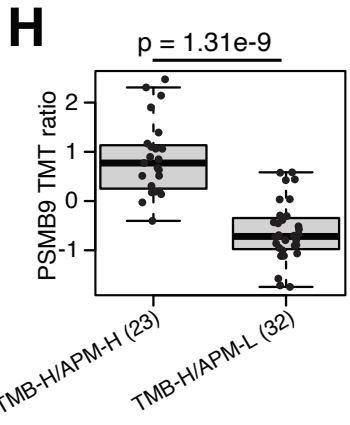


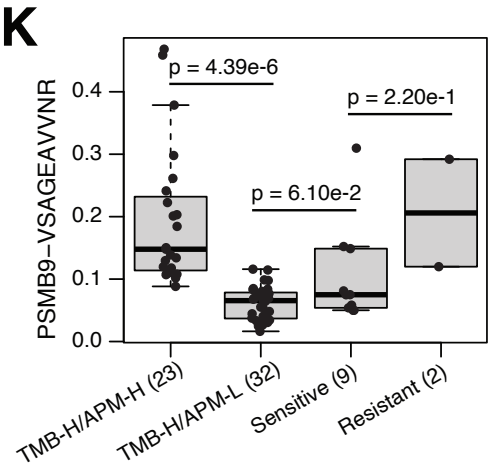


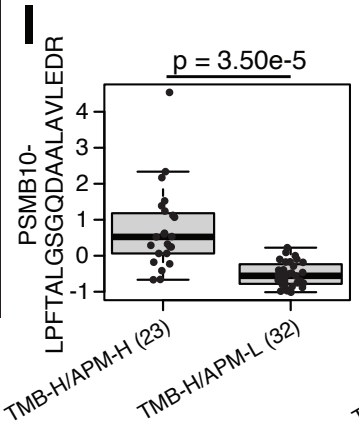
F

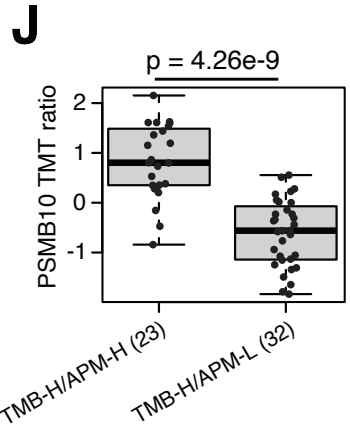
Importance	Peptide
1	PSMB10_LPFTALGSGQDAALAVLEDR
2	PSMB9_VSAGEAVVNR
3	PSMB10_YHFVPGTTAVLTQTVK
4	PSMB9_VILGNELPK
5	PSMB8_FQHGVIAAVDSR
6	B2M_IQVYSR
7	PSMB8_VESTDVSDLLHQYR
8	B2M_VEHSDLSFSK
9	TAP1_QVAAVGQEPQVFGR
10	HLA.A_FIAVGYVDDTQFVR
11	HLA.C_AYLEGTCVEWLR
12	STAT2_HYQLLTEENIPENPLR
13	STAT1_LLGPNASPDGLIPWTR
14	TAP1_ELISWGAPGSADSTR
15	TAP2_EQLFSSLLR
16	ERAP1_ILASTQFEPTAAR
17	TAPBP_WASGLTPAQNCPR
18	TAP2_NNIAYGLQSCEDDK
19	STAT3_GLSIEQLTTLAEK
20	ERAP1_TQEFPQILTLIGR
21	TAPBP_IHHPSLPASGR
22	HLA.A_FDSDAASQR
23	STAT3_LLGPGVNYSGCQITWAK
24	STAT2_EVLQSLPLTEIIR
25	HLA.C_WAAVVVPSGQEQR
26	ERAP2_FLQGVFQEDPEWR
27	JAK2_YDPLQDNTGEVVAVK
28	ERAP2_ILAVTDFEPTQAR
29	STAT1_ELSAVTFPDIIR
30	HLA.B_APWIEQEGPEYWDR
31	JAK1_DLVQGEHLGR
32	JAK2_EVGDYGQLHETEVLLK
33	JAK1_NVLVESEHQVK











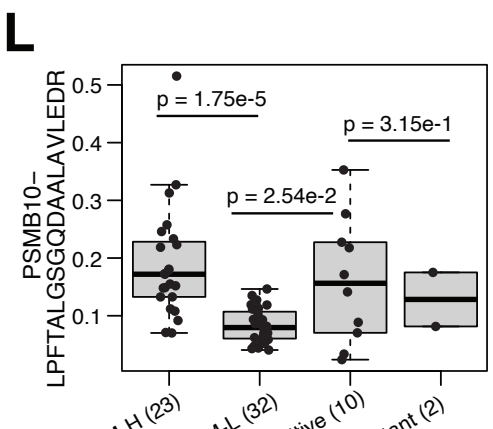


Figure S3: Selective reaction monitoring (SRM) assay for antigen presentation machinery (APM) status, related to Figure 3.

(A) Scatter plot of log2 transformed total number of variants (x-axis) versus APM z-score (y-axis). Quadrants represent the 4 TMB/APM subgroups.

(B) Heatmap of JAK-STAT and APM mRNA and protein levels in the TMB-H groups. Columns represent individual samples and rows represent genes.

(C) Table of proteins and corresponding peptides used for the targeted proteomic assay panel.

(D-E) Histograms showing the frequency of correlations between SRM peptides from the same genes (D) and SRM peptides with TMT-based protein levels (E) for the independent cohort.

(F) Table of peptides sorted by feature importance, selected from models in Figure 3F.

(G-J) Boxplots of SRM-based quantification of top two peptides (G,I) and global proteomics quantification of their host genes (H,J) in the TMB-H groups. P-values derived from Student's t-test.

(K-L) Boxplots of SRM-based quantification of top two peptides for the TMB-H groups compared to the treated cohort. P-values derived from Student's t-test.

Boxplots: Box portion represents Interquartile range (IQR), midline corresponds to the median, and whiskers range from the minimum (bottom) and maximum (top) variability outside the first and third quartiles (Q1 and Q3). Outliers are shown as points above whiskers.

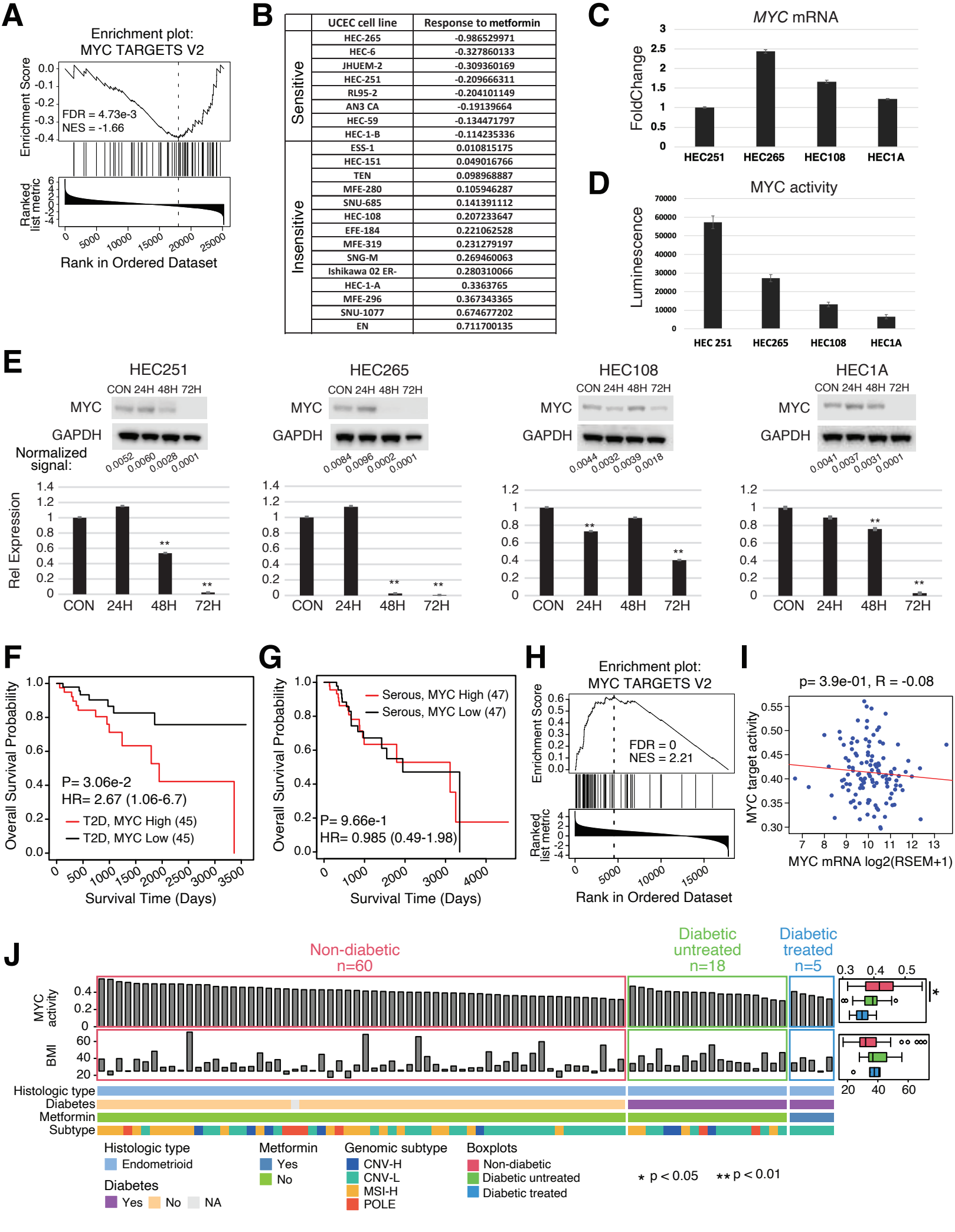

A
Enrichment plot:
MYC TARGETS V2
Enrichment Score
FDR = 4.73e-3
NES = -1.66
Ranked list metric
Rank in Ordered Dataset
B
UCEC cell line
Response to metformin
Sensitive
HEC-265 -0.986529971
HEC-6 -0.327860133
JHUEM-2 -0.309360169
HEC-251 -0.209666311
RL95-2 -0.204101149
AN3 CA -0.19139664
HEC-59 -0.134471797
HEC-1-B -0.114235336
Insensitive
ESS-1 0.010815175
HEC-151 0.049016766
TEN 0.098968887
MFE-280 0.105946287
SNU-685 0.141391112
HEC-108 0.207233647
EFE-184 0.221062528
MFE-319 0.231279197
SNG-M 0.269460063
Ishikawa 02 ER- 0.280310066
HEC-1-A 0.3363765
MFE-296 0.367343365
SNU-1077 0.674677202
EN 0.711700135
C
MYC mRNA
FoldChange
HEC251
HEC265
HEC108
HEC1A
D
MYC activity
Luminescence
HEC 251
HEC265
HEC108
HEC1A
E
HEC251
HEC265
HEC108
HEC1A
CON 24H 48H 72H
MYC
GAPDH
Normalized signal:
0.0052 0.0060 0.0028 0.0001
0.0084 0.0096 0.0002 0.0001
0.0044 0.0032 0.0039 0.0018
0.0041 0.0037 0.0031 0.0001
Rel Expression
CON 24H 48H 72H
F
Overall Survival Probability
P= 3.06e-2
HR= 2.67 (1.06-6.7)
T2D, MYC High (45)
T2D, MYC Low (45)
Survival Time (Days)
G
Overall Survival Probability
Serous, MYC High (47)
Serous, MYC Low (47)
P= 9.66e-1
HR= 0.985 (0.49-1.98)
Survival Time (Days)
H
Enrichment plot:
MYC TARGETS V2
Enrichment Score
FDR = 0
NES = 2.21
Ranked list metric
Rank in Ordered Dataset
I
p= 3.9e-01, R = -0.08
MYC target activity
MYC mRNA log2(RSEM+1)
J
Non-diabetic
n=60
Diabetic
untreated
n=18
Diabetic
treated
n=5
MYC activity
BMI
Histologic type
Diabetes
Metformin
Subtype
Histologic type
Endometrioid
Diabetes
Yes No NA
Metformin
Yes
No
Genomic subtype
CNV-H
CNV-L
MSI-H
POLE
Boxplots
Non-diabetic
Diabetic untreated
Diabetic treated
* p < 0.05 ** p < 0.01

Figure S4: Metformin may target MYC in EC, related to Figure 4.

(A) MYC Targets V2 enrichment plot from pathway analysis comparing metformin-treated versus untreated patients with Type2 Diabetes (T2D) in the exploratory cohort.

(B) Table of DepMap EC cell lines treated with metformin and their corresponding responses. Cell lines are separated into “sensitive” and “insensitive” groups.

(C-D) Barplots of *MYC* mRNA expression (C) and MYC activity (D) in 4 EC cell lines.

(E) Western blots of MYC expression in metformin treated EC cell lines, with corresponding barplots (bottom panel) of quantified relative expression (y-axis) across timepoints (x-axis).

(F-G) Survival analysis of TCGA T2D patients (F) and serous tumors (G) with high and low MYC activity. P-values derived from log-rank test.

(H) MYC Targets V2 enrichment plot from pathway analysis comparing MYC-high and low TCGA tumors at the protein RPPA level.

(I) Scatter plot of MYC activity (y-axis) and *MYC* mRNA expression (x-axis).

(J) Heatmap of all endometrioid tumors in the exploratory cohort, sorted by MYC activity (top panel) and grouped by diabetes and metformin treatment status. Side boxplots (right) compare MYC activity (top) and BMI (bottom) across the diabetes/treatment groups. P-values derived from Student’s t-test

Boxplots: Box portion represents Interquartile range (IQR), midline corresponds to the median, and whiskers range from the minimum (bottom) and maximum (top) variability outside the first and third quartiles (Q1 and Q3). Outliers are shown as points above whiskers.

Barplots: Error bars represent the standard error of the mean (SEM). P-values derived from Student’s t-test. * $p < 0.05$, ** $p < 0.01$.

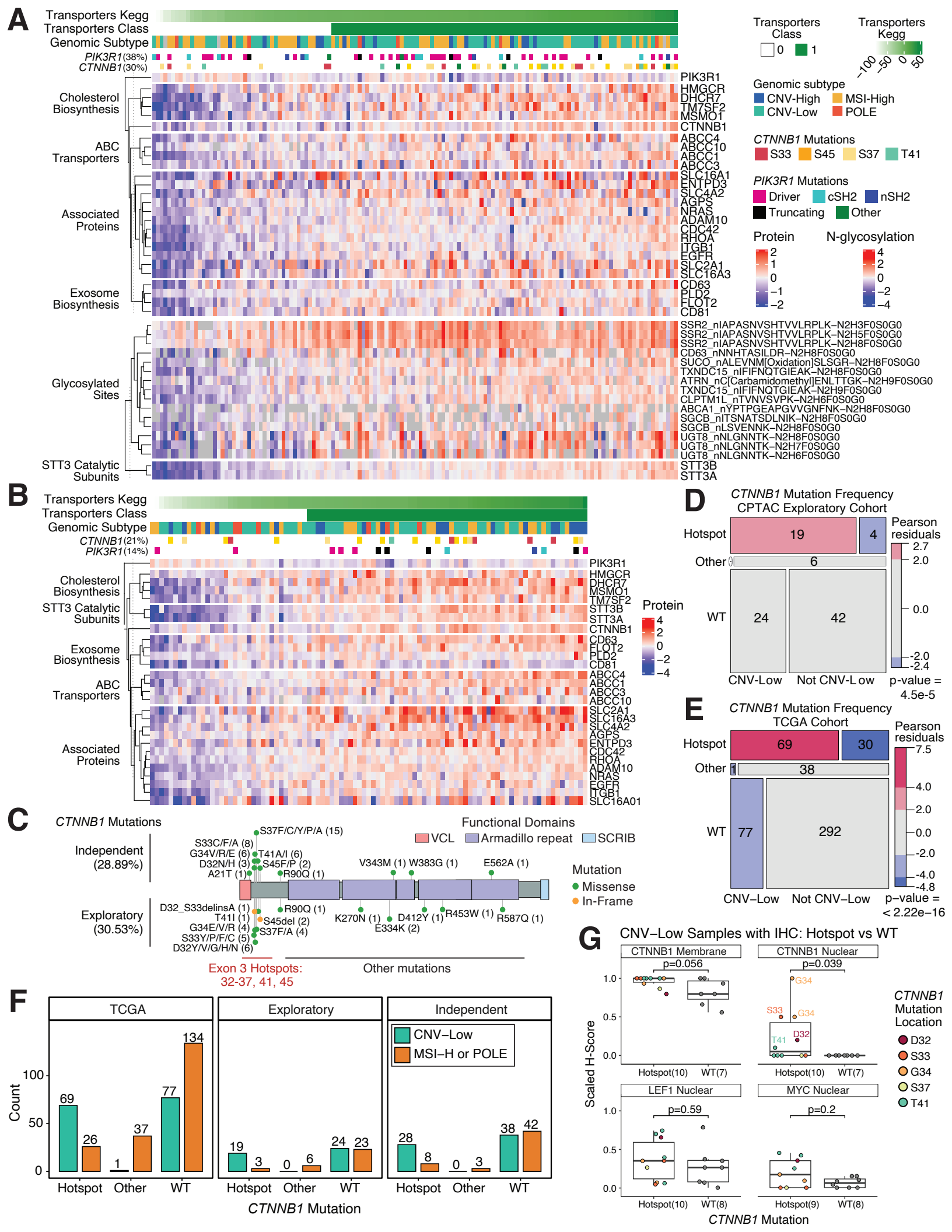

A
Transporters Kegg
Transporters Class
Genomic Subtype
PIK3R1 (38%)
CTNNB1 (30%)
Cholesterol Biosynthesis
ABC Transporters
Associated Proteins
Exosome Biosynthesis
Glycosylated Sites
STT3 Catalytic Subunits
PIK3R1
HMGCR
DHCR7
TM7SF2
MSMO1
CTNNB1
ABCC4
ABCC10
ABCC1
ABCC3
SLC16A1
ENTPD3
SLC4A2
AGPS
NRAS
ADAM10
CDC42
RHOA
ITGB1
EGFR
SLC2A1
SLC16A3
CD63
PLD2
FLOT2
CD81
SSR2_nIAPASNVSHTVVLRPLK–N2H3F0S0G0
SSR2_nIAPASNVSHTVVLRPLK–N2H5F0S0G0
SSR2_nIAPASNVSHTVVLRPLK–N2H8F0S0G0
CD63_nNNHTASILDR–N2H8F0S0G0
SUCO_nALEVNM[Oxidation]SLSGR–N2H8F0S0G0
TXNDC15_nIFIFNQTGIEAK–N2H8F0S0G0
ATRN_nC[Carbamidomethyl]ENLTTGK–N2H9F0S0G0
TXNDC15_nIFIFNQTGIEAK–N2H9F0S0G0
CLPTM1L_nTVNVSVPK–N2H6F0S0G0
ABCA1_nYPTPGEAPGVVGNFNK–N2H8F0S0G0
SGCB_nITSNATSDLNIK–N2H8F0S0G0
SGCB_nLSVENNK–N2H8F0S0G0
UGT8_nNLGNNTK–N2H8F0S0G0
UGT8_nNLGNNTK–N2H7F0S0G0
UGT8_nNLGNNTK–N2H6F0S0G0
STT3B
STT3A
Transporters Class
0
1
Transporters Kegg
−100
−50
0
50
Genomic subtype
CNV-High
MSI-High
CNV-Low
POLE
CTNNB1 Mutations
S33
S45
S37
T41
PIK3R1 Mutations
Driver
cSH2
nSH2
Truncating
Other
Protein
2
1
0
−1
−2
N-glycosylation
4
2
0
−2
−4
B
Transporters Kegg
Transporters Class
Genomic Subtype
CTNNB1 (21%)
PIK3R1 (14%)
Cholesterol Biosynthesis
STT3 Catalytic Subunits
Exosome Biosynthesis
ABC Transporters
Associated Proteins
PIK3R1
HMGCR
DHCR7
MSMO1
TM7SF2
STT3B
STT3A
CTNNB1
CD63
FLOT2
PLD2
CD81
ABCC4
ABCC1
ABCC3
ABCC10
SLC2A1
SLC16A3
SLC4A2
AGPS
ENTPD3
CDC42
RHOA
ADAM10
NRAS
EGFR
ITGB1
SLC16A01
Protein
4
2
0
−2
−4
C
CTNNB1 Mutations
Independent (28.89%)
Exploratory (30.53%)
S37F/C/Y/P/A (15)
S33C/F/A (8)
G34V/R/E (6)
T41A/I (6)
D32N/H (3)
S45F/P (2)
A21T (1)
R90Q (1)
V343M (1)
W383G (1)
E562A (1)
D32_S33delinsA (1)
T41I (1)
S45del (2)
G34E/V/R (4)
S37F/A (4)
S33Y/P/F/C (5)
D32Y/V/G/H/N (6)
R90Q (1)
K270N (1)
D412Y (1)
E334K (2)
R453W (1)
R587Q (1)
Functional Domains
VCL
Armadillo repeat
SCRIB
Mutation
Missense
In-Frame
Exon 3 Hotspots: 32-37, 41, 45
Other mutations
D
CTNNB1 Mutation Frequency
CPTAC Exploratory Cohort
Hotspot
Other
WT
19
4
6
24
42
CNV-Low
Not CNV-Low
Pearson residuals
2.7
2.0
0.0
-2.0
-2.4
p-value = 4.5e-5
E
CTNNB1 Mutation Frequency
TCGA Cohort
Hotspot
Other
WT
69
30
1
38
77
292
CNV−Low
Not CNV−Low
Pearson residuals
7.5
4.0
2.0
0.0
-2.0
-4.0
-4.8
p−value = < 2.22e−16
F
TCGA
Exploratory
Independent
CNV−Low
MSI−H or POLE
Count
100
50
0
69
26
1
37
77
134
19
3
0
6
24
23
28
8
0
3
38
42
Hotspot
Other
WT
CTNNB1 Mutation
G
CNV−Low Samples with IHC: Hotspot vs WT
CTNNB1 Membrane
p=0.056
CTNNB1 Nuclear
p=0.039
LEF1 Nuclear
p=0.59
MYC Nuclear
p=0.2
Scaled H-Score
1.0
0.5
0.0
Hotspot(10)
WT(7)
WT(8)
Hotspot(9)
G34
S33
T41
D32
CTNNB1 Mutation Location
D32
S33
G34
S37
T41
CTNNB1 Mutation

Figure S5: *CTNNB1* hotspot mutations block DKK-induced degradation, related to Figure 5.

(A) Heatmap of protein based transporter pathway activity and affected pathways at protein and glyco-peptide levels in the independent cohort. Top panels show transporter classes and *CTNNB1* and *PIK3R1* mutations.

(B) Same as (A) for the Exploratory cohort.

(C) Lollipop plots of *CTNNB1* mutations in the independent cohort (top) and exploratory cohort (bottom).

(D-E) Mosaic plots showing distribution of *CTNNB1* mutations across CNV-L tumors versus all other tumors in the exploratory cohort (D) and TCGA cohort (E). P-value determined from Chi-Square Test.

(F) Bar plots showing distribution of *CTNNB1* mutations across CNV-L and high mutation burden tumors (POLE and MSI-H) in the TCGA (left), exploratory (middle), and independent cohorts (right).

(G) Boxplots of IHC H-scores in *CTNNB1* hotspot mutants versus WT CNV-L tumors for independent cohort tumors with IHC data. P-values determined by Wilcoxon Rank Sum Test.

Boxplots: Box portion represents Interquartile range (IQR), midline corresponds to the median, and whiskers range from the minimum (bottom) and maximum (top) variability outside the first and third quartiles (Q1 and Q3). Outliers are shown as points above whiskers.

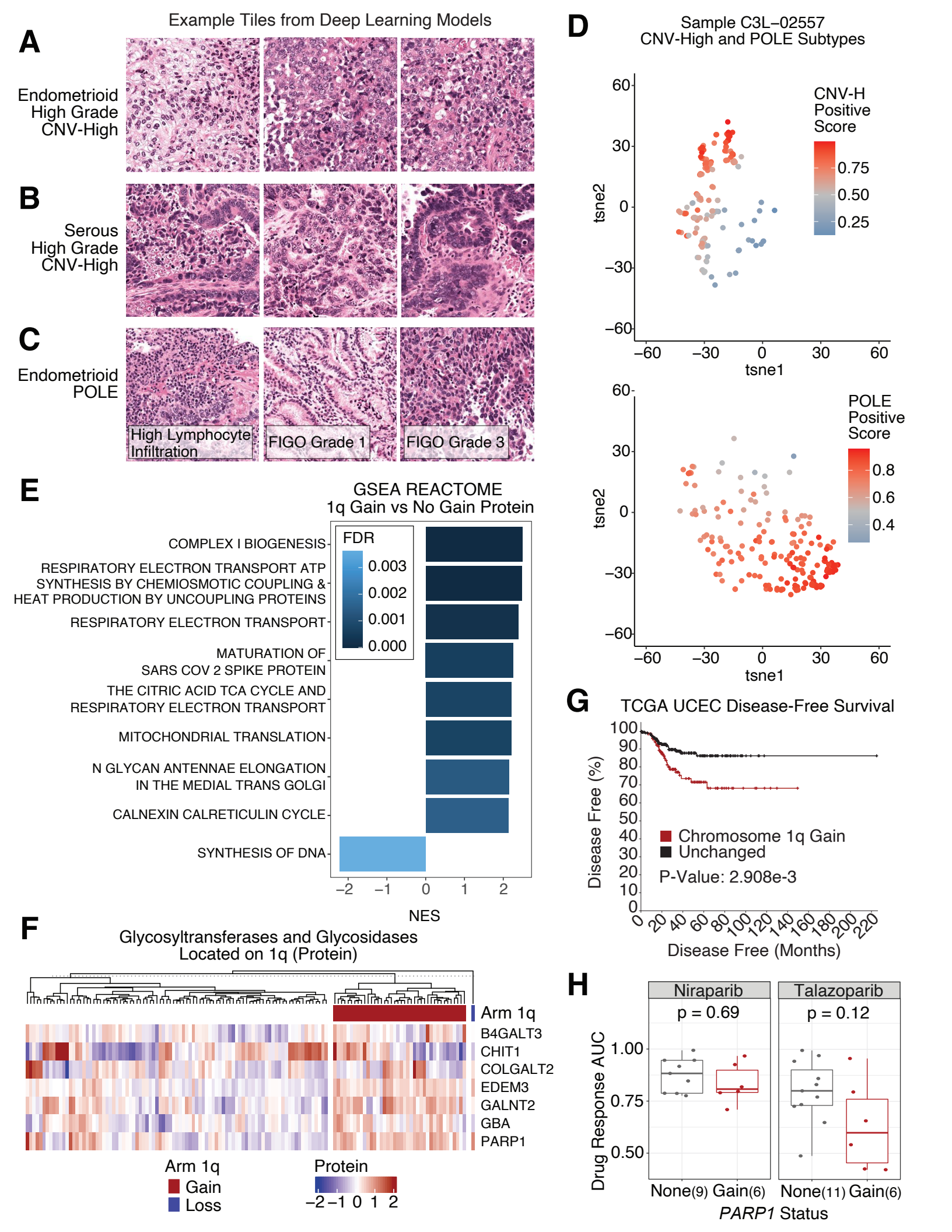

Example Tiles from Deep Learning Models
A
Endometrioid
High Grade
CNV-High
B
Serous
High Grade
CNV-High
C
Endometrioid
POLE
High Lymphocyte Infiltration
FIGO Grade 1
FIGO Grade 3
D
Sample C3L–02557
CNV-High and POLE Subtypes
CNV-H
Positive
Score
0.75
0.50
0.25
tsne2
tsne1
POLE
Positive
Score
0.8
0.6
0.4
E
GSEA REACTOME
1q Gain vs No Gain Protein
FDR
0.003
0.002
0.001
0.000
COMPLEX I BIOGENESIS
RESPIRATORY ELECTRON TRANSPORT ATP SYNTHESIS BY CHEMIOSMOTIC COUPLING & HEAT PRODUCTION BY UNCOUPLING PROTEINS
RESPIRATORY ELECTRON TRANSPORT
MATURATION OF SARS COV 2 SPIKE PROTEIN
THE CITRIC ACID TCA CYCLE AND RESPIRATORY ELECTRON TRANSPORT
MITOCHONDRIAL TRANSLATION
N GLYCAN ANTENNAE ELONGATION IN THE MEDIAL TRANS GOLGI
CALNEXIN CALRETICULIN CYCLE
SYNTHESIS OF DNA
NES
F
Glycosyltransferases and Glycosidases
Located on 1q (Protein)
Arm 1q
B4GALT3
CHIT1
COLGALT2
EDEM3
GALNT2
GBA
PARP1
Arm 1q
Gain
Loss
Protein
G
TCGA UCEC Disease-Free Survival
Disease Free (%)
Chromosome 1q Gain
Unchanged
P-Value: 2.908e-3
Disease Free (Months)
H
Niraparib
p = 0.69
Talazoparib
p = 0.12
Drug Response AUC
None(9) Gain(6)
None(11) Gain(6)
PARP1 Status

Figure S6: Deep learning models successfully classify molecular features, related to Figure 6.

(A-C) Example tiles from high-grade, CNV-H endometrioid (A), high-grade, CNV-H serous (B), and POLE with varying grade (C) samples.

(D) tSNE plots where each point is a tile, for sample C3L-02557, which is both CNV-H and POLE subtypes. Points are colored by predicted CNV-H score (top) and POLE score (bottom).

(E) Barplot of pathway enrichment analysis using protein levels comparing 1q gain and no gain in the independent cohort.

(F) Heatmap showing protein expression of glycosyltransferases and glycosidases located on chromosome 1q, separated by 1q copy number status.

(G) Disease-free survival comparing patients with and without 1q gain in the TCGA cohort.

(H) Boxplots of niraparib and talazoparib response in DepMap EC cell lines with and without *PARP1* amplification. P-values determined by Wilcoxon Rank Sum Test.

Boxplots: Box portion represents Interquartile range (IQR), midline corresponds to the median, and whiskers range from the minimum (bottom) and maximum (top) variability outside the first and third quartiles (Q1 and Q3). Outliers are shown as points above whiskers.

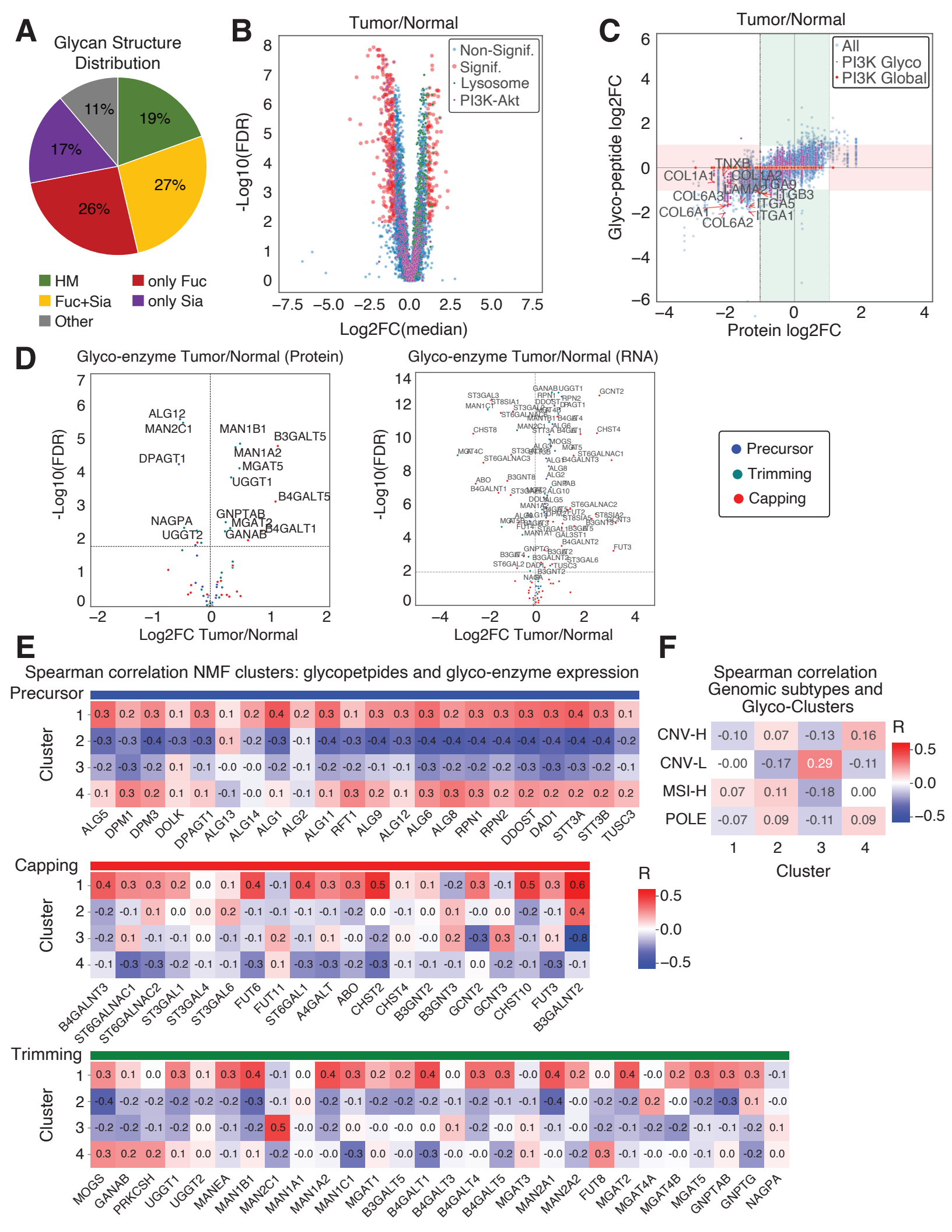

A
Glycan Structure Distribution
19%
27%
26%
17%
11%
HM
only Fuc
Fuc+Sia
only Sia
Other
B
Tumor/Normal
Non-Signif.
Signif.
Lysosome
PI3K-Akt
-Log10(FDR)
Log2FC(median)
C
Tumor/Normal
All
PI3K Glyco
PI3K Global
Glyco-peptide log2FC
Protein log2FC
TNXB
COL1A1
COL1A2
AMA2
ITGA9
ITGB3
COL6A3
ITGA5
COL6A1
ITGA1
COL6A2
D
Glyco-enzyme Tumor/Normal (Protein)
Glyco-enzyme Tumor/Normal (RNA)
-Log10(FDR)
Log2FC Tumor/Normal
ALG12
MAN2C1
MAN1B1
B3GALT5
MAN1A2
DPAGT1
MGAT5
UGGT1
B4GALT5
GNPTAB
NAGPA
MGAT2
B4GALT1
UGGT2
GANAB
Precursor
Trimming
Capping
E
Spearman correlation NMF clusters: glycopetpides and glyco-enzyme expression
Precursor
Capping
Trimming
Cluster
ALG5 DPM1 DPM3 DOLK DPAGT1 ALG13 ALG14 ALG1 ALG2 ALG11 RFT1 ALG9 ALG12 ALG6 ALG8 RPN1 RPN2 DDOST DAD1 STT3A STT3B TUSC3
B4GALNT3 ST6GALNAC1 ST6GALNAC2 ST3GAL1 ST3GAL4 ST3GAL6 FUT6 FUT11 ST6GAL1 A4GALT ABO CHST2 CHST4 B3GNT2 B3GNT3 GCNT2 GCNT3 CHST10 FUT3 B3GALNT2
MOGS GANAB PRKCSH UGGT1 UGGT2 MANEA MAN1B1 MAN2C1 MAN1A1 MAN1A2 MAN1C1 MGAT1 B3GALT5 B4GALT1 B4GALT3 B4GALT4 B4GALT5 MGAT3 MAN2A1 MAN2A2 FUT8 MGAT2 MGAT4A MGAT4B MGAT5 GNPTAB GNPTG NAGPA
F
Spearman correlation Genomic subtypes and Glyco-Clusters
CNV-H -0.10 0.07 -0.13 0.16
CNV-L -0.00 -0.17 0.29 -0.11
MSI-H 0.07 0.11 -0.18 0.00
POLE -0.07 0.09 -0.11 0.09
1 2 3 4
Cluster
R
0.5
0.0
−0.5

Figure S7: Multi-omic and glycoproteomic NMF clustering separates tumors into 4 clusters, related to Figure 7.

(A) Pie chart of glycan structure distribution among observed glycopeptides.

(B) Volcano plot of differential glycopeptide expression in tumor versus normal samples. X-axis shows log2 fold change and the y-axis shows -log10 FDR, which is determined by Student's T-test.

(C) Scatter plot comparing log2 fold change of tumor/normal at glyco-peptide (y-axis) levels and corresponding unmodified protein levels (x-axis).

(D) Volcano plots comparing tumor and normal glyco-enzymes at the protein (left) and RNA (right) levels. Points are colored by glyco-enzyme function: Precursor (blue), Trimming (green), and Capping (red)

(E) Heatmaps showing correlations of glyco-peptide and glyco-enzyme levels across glyco-clusters, separated by glyco-enzyme function.

(F) Heatmaps showing correlations between NMF clusters and genomic subtypes.