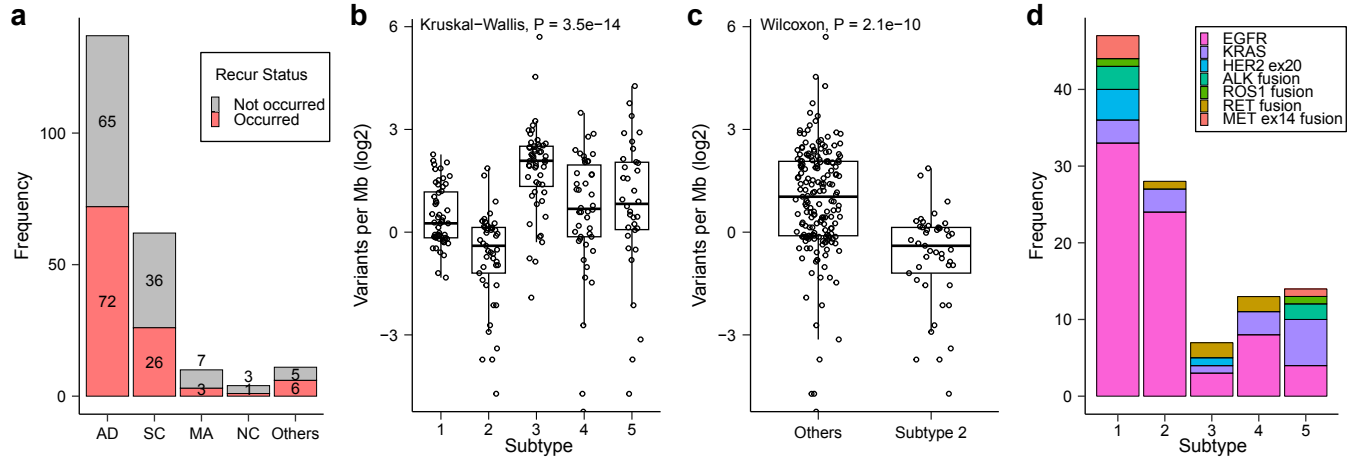
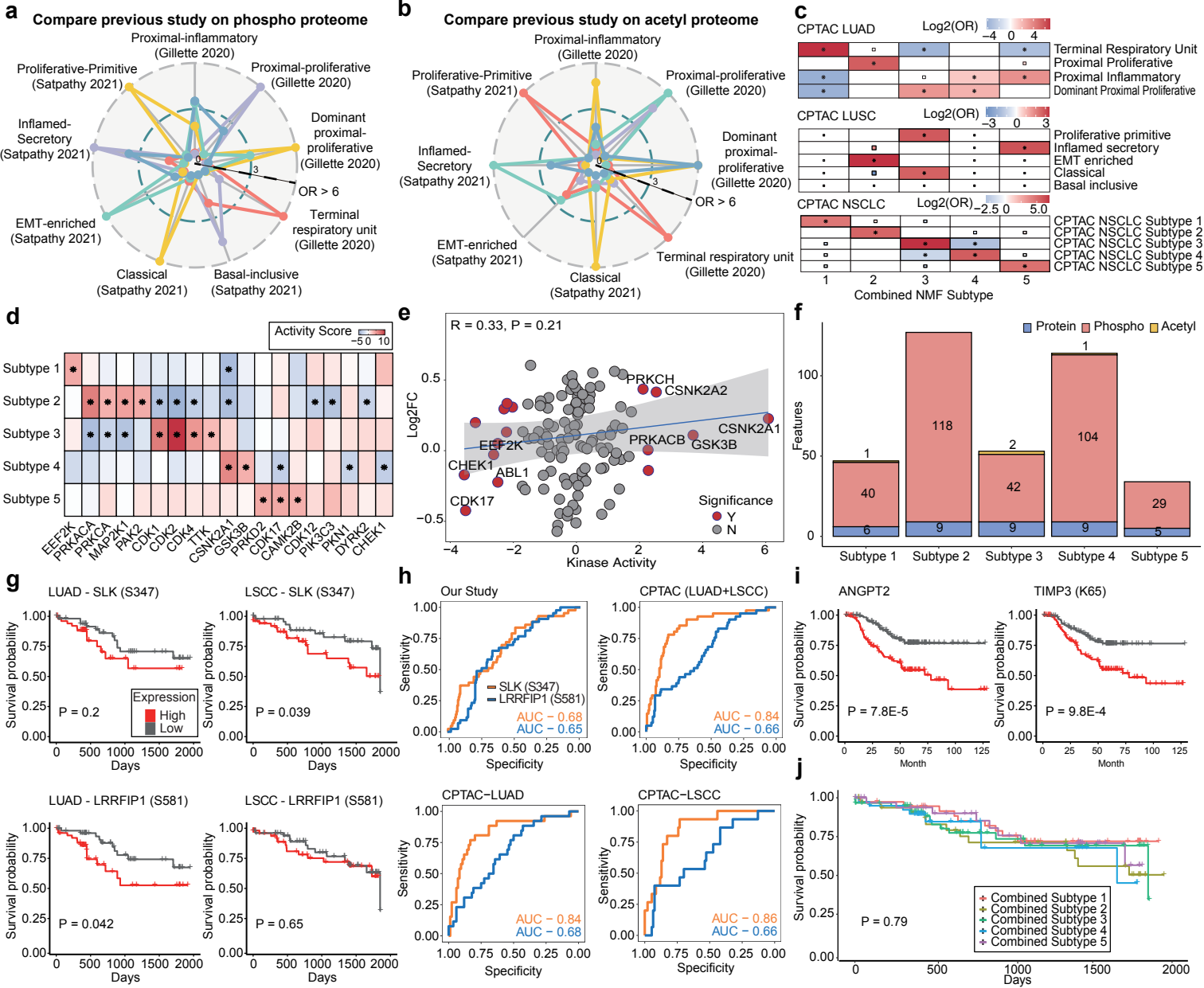
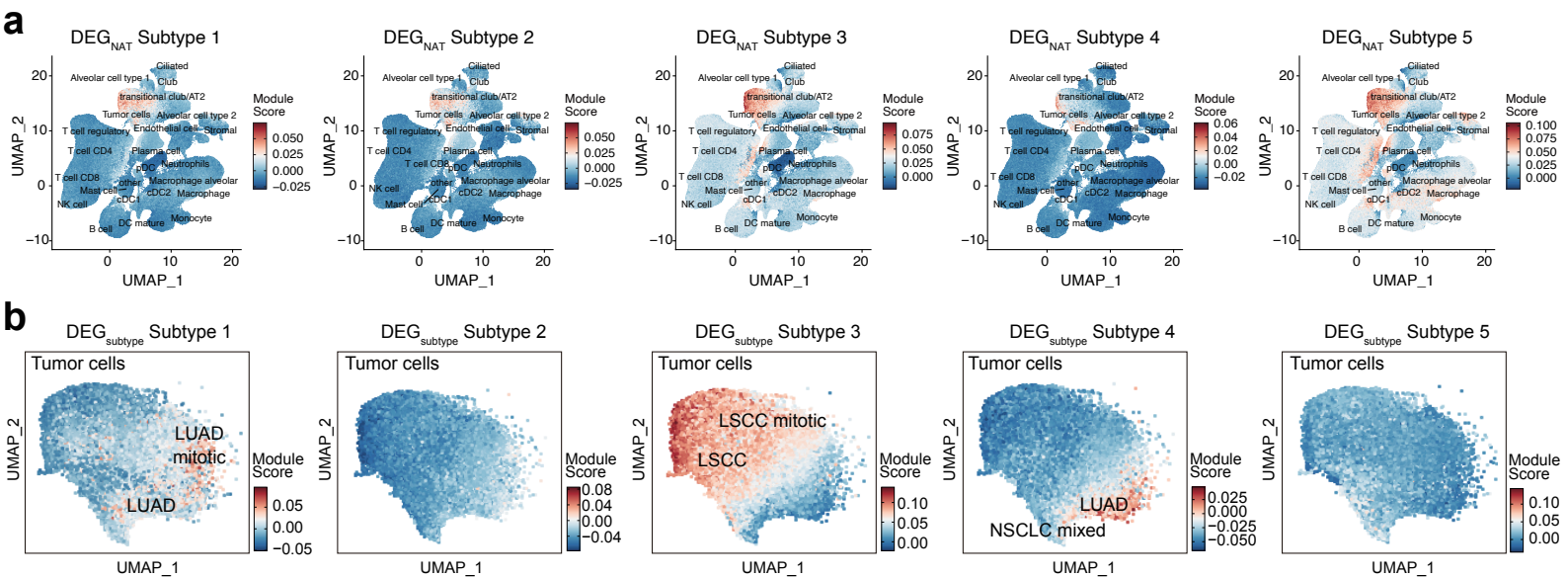

a
Recur Status
Not occurred
Occurred
Frequency
100
50
0
65
72
36
26
7
3
3
1
5
6
AD
SC
MA
NC
Others
b
Kruskal−Wallis, P = 3.5e−14
Variants per Mb (log2)
6
3
0
−3
1
2
3
4
5
Subtype
c
Wilcoxon, P = 2.1e−10
Variants per Mb (log2)
6
3
0
−3
Others
Subtype 2
d
EGFR
KRAS
HER2 ex20
ALK fusion
ROS1 fusion
RET fusion
MET ex14 fusion
Frequency
40
30
20
10
0
1
2
3
4
5
Subtype

Supplementary Fig. 1| Description of 229 Korean patients with NSCLC. a, Recurrent status according to histology (DX) for the patients with NSCLC (adenocarcinoma, "AD"; squamous cell carcinoma, "SC"; mucinous adenocarcinoma, mixed mucinous and non-mucinous adenocarcinoma, "MA"; large cell neuroendocrine carcinoma, "NC"; adenosquamous carcinoma, combined small cell carcinoma and adenocarcinoma, enteric adenocarcinoma, large cell carcinoma, pleomorphic carcinoma, "Others"). The color of bar represents recurrence status with each patient. **b-c**, Multiple comparison of tumor mutation burden (TMB) status between five subtypes and comparison between Subtype 2 and other subtypes. TMB were calculated by variants per Mb using whole exome sequencing data. The p-value was calculated using the two-sided Kruskal-Wallis test (b) and wilcoxon test (c). For box-plots, middle line, median: box edges, 25^{th} and 75^{th} percentiles; whiskers, most extreme points that do not exceed ±1.5 × IQR. **d**, Mutation frequency according to five subtypes.

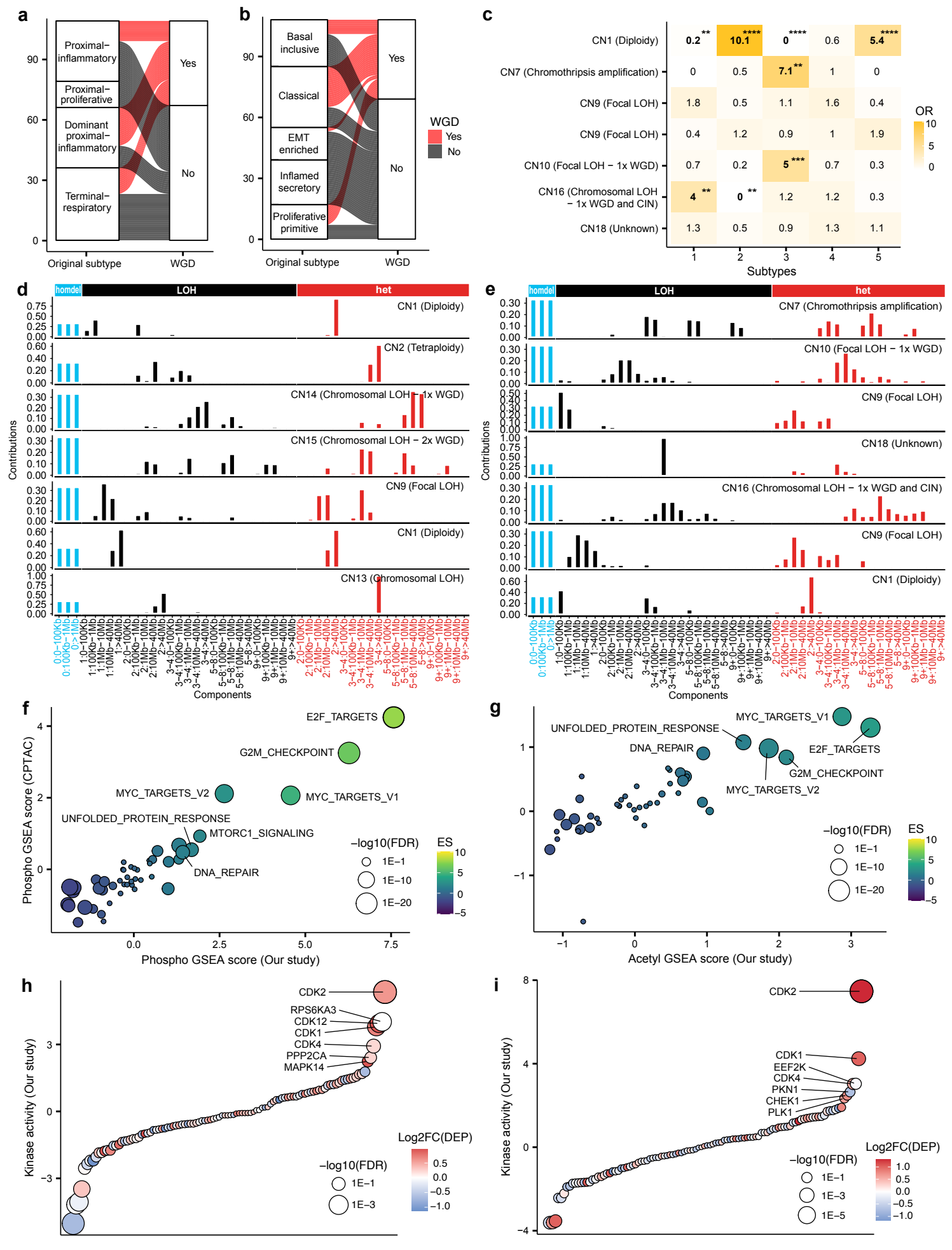

a
Compare previous study on phospho proteome
Proximal-inflammatory (Gillette 2020)
Proximal-proliferative (Gillette 2020)
Dominant proximal-proliferative (Gillette 2020)
Terminal respiratory unit (Gillette 2020)
Basal-inclusive (Satpathy 2021)
Classical (Satpathy 2021)
EMT-enriched (Satpathy 2021)
Inflamed-Secretory (Satpathy 2021)
Proliferative-Primitive (Satpathy 2021)
OR > 6
b
Compare previous study on acetyl proteome
c
CPTAC LUAD
CPTAC LUSC
CPTAC NSCLC
Log2(OR)
Terminal Respiratory Unit
Proximal Proliferative
Proximal Inflammatory
Dominant Proximal Proliferative
Proliferative primitive
Inflamed secretory
EMT enriched
Classical
Basal inclusive
CPTAC NSCLC Subtype 1
CPTAC NSCLC Subtype 2
CPTAC NSCLC Subtype 3
CPTAC NSCLC Subtype 4
CPTAC NSCLC Subtype 5
Combined NMF Subtype
d
Activity Score
Subtype 1
Subtype 2
Subtype 3
Subtype 4
Subtype 5
EEF2K
PRKACA
PRKCA
MAP2K1
PAK2
CDK1
CDK2
CDK4
TTK
CSNK2A1
GSK3B
PRKD2
CDK17
CAMK2B
CDK12
PIK3C3
PKN1
DYRK2
CHEK1
e
R = 0.33, P = 0.21
Log2FC
Kinase Activity
PRKCH
CSNK2A2
CSNK2A1
EEF2K
PRKACB
GSK3B
CHEK1
ABL1
CDK17
Significance
Y
N
f
Protein
Phospho
Acetyl
Features
Subtype 1
Subtype 2
Subtype 3
Subtype 4
Subtype 5
g
LUAD - SLK (S347)
LSCC - SLK (S347)
LUAD - LRRFIP1 (S581)
LSCC - LRRFIP1 (S581)
Survival probability
Days
Expression
High
Low
P = 0.2
P = 0.039
P = 0.042
P = 0.65
h
Our Study
CPTAC (LUAD+LSCC)
CPTAC−LUAD
CPTAC−LSCC
Sensitivity
Specificity
SLK (S347)
LRRFIP1 (S581)
AUC − 0.68
AUC − 0.65
AUC − 0.84
AUC − 0.66
AUC − 0.84
AUC − 0.68
AUC − 0.86
AUC − 0.66
i
ANGPT2
TIMP3 (K65)
Survival probability
Month
P = 7.8E-5
P = 9.8E-4
j
Survival probability
Days
P = 0.79
Combined Subtype 1
Combined Subtype 2
Combined Subtype 3
Combined Subtype 4
Combined Subtype 5

Supplementary Fig. 2| Characteristics of subtypes compared with CPTAC cohort and features related with poor prognosis. a-b, Overlap of subtype features between the five NMF subtypes in this study and subtypes identified in previous NSCLC multiomics studies on phospho-, acetyl proteome data. **c**, Heatmap of overlaps related to **Fig. 2b**. A statistical significance calculated using two-sided Fisher's exact test with Benjamini-Hochberg adjustment (a-c). Full rectangle and asterisk indicate significant overlaps (Adjusted P $\leq$ 0.05); faint rectangle indicates overlaps that pass only the nominal P value (Fisher's exact test P $\leq$ 0.05, adjusted P > 0.05); and blank indicates overlaps which is not significant (Fisher's exact test P > 0.05). **d**, Heatmap of significant kinases on subtypes. The color represents kinase activity derived from phosphoproteome data at each subtype and asterisk were displayed with significant results (*, adjusted P < 0.05). P-value was obtained from multivariate linear model and adjusted p-value was calculated using the Benjamini–Hochberg adjustment. **e,** Correlation of the kinase expression and estimated activity using two-sided Pearson correlation method. The color of points indicates statistical significance **f**, Distribution of unfavorable prognostic factor which overlapped with differentially expressed on subtypes. The color of bar plot represents data type of features. **g**, Survival curves of poor prognosis markers which contains *SLK* (S347) and *LRRFIP1* (S581) on CPTAC LUAD cohort and LSCC cohort. The p-value was calculated with the log-rank test. **h**, ROC curve analysis results for our cohort and the CPTAC cohort using *SLK* (S347) and *LRRFIP1* (S581). The color of lines indicates each features. **i**, Survival curves of features which are up-regulated

on Subtype 4 and related with HIF-1 signaling pathway which contributes poor prognosis on our cohort were represented. The p-value was calculated with the log-rank test. **j**, Survival plot of CPTAC NSCLC cohort between combined subtypes. The color illustrated each combined subtype.

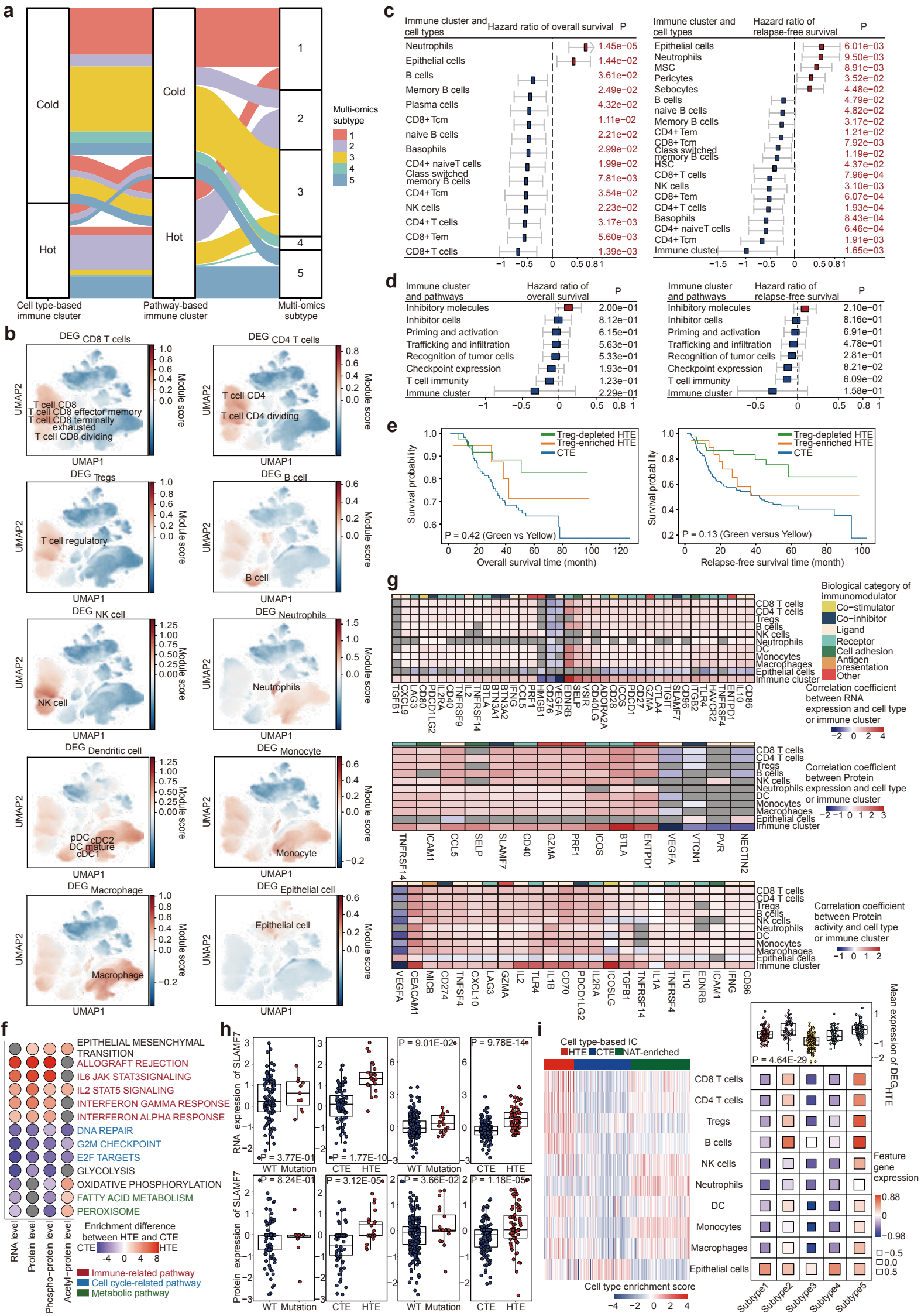

a
DEG_NAT Subtype 1
DEG_NAT Subtype 2
DEG_NAT Subtype 3
DEG_NAT Subtype 4
DEG_NAT Subtype 5
Ciliated
Alveolar cell type 1
Club
transitional club/AT2
Tumor cells
Alveolar cell type 2
T cell regulatory
Endothelial cell
Stromal
T cell CD4
Plasma cell
pDC
Neutrophils
T cell CD8
other
Macrophage alveolar
Mast cell
cDC2
Macrophage
cDC1
NK cell
Monocyte
B cell
DC mature
UMAP_1
UMAP_2
Module Score
b
DEG_subtype Subtype 1
DEG_subtype Subtype 2
DEG_subtype Subtype 3
DEG_subtype Subtype 4
DEG_subtype Subtype 5
Tumor cells
LUAD mitotic
LUAD
LSCC mitotic
LSCC
LUAD
NSCLC mixed
UMAP_1
UMAP_2
Module Score

Supplementary Fig. 3| UMAP of single-cell type specific Subtype 1 to. a, The results of single-cell type specific Subtype 1 to 5 using each DEG_{NAT}. Each point color represents the module score of each cell, and the more relevant it is to the cell types, the higher score is represented in red color. **b**, Umap of single-cell type specific tumor cell with $DEG_{subtype}$. UMAP information was obtained from the original study[23].


a
Proximal−inflammatory
Proximal−proliferative
Dominant proximal−inflammatory
Terminal−respiratory
Yes
No
Original subtype
WGD
b
Basal inclusive
Classical
EMT enriched
Inflamed secretory
Proliferative primitive
WGD
Yes
No
c
CN1 (Diploidy)
CN7 (Chromothripsis amplification)
CN9 (Focal LOH)
CN9 (Focal LOH)
CN10 (Focal LOH − 1x WGD)
CN16 (Chromosomal LOH − 1x WGD and CIN)
CN18 (Unknown)
OR
Subtypes
d
homdel
LOH
het
CN1 (Diploidy)
CN2 (Tetraploidy)
CN14 (Chromosomal LOH − 1x WGD)
CN15 (Chromosomal LOH − 2x WGD)
CN9 (Focal LOH)
CN1 (Diploidy)
CN13 (Chromosomal LOH)
Contributions
Components
e
CN7 (Chromothripsis amplification)
CN10 (Focal LOH − 1x WGD)
CN9 (Focal LOH)
CN18 (Unknown)
CN16 (Chromosomal LOH − 1x WGD and CIN)
CN9 (Focal LOH)
CN1 (Diploidy)
f
E2F_TARGETS
G2M_CHECKPOINT
MYC_TARGETS_V2
MYC_TARGETS_V1
UNFOLDED_PROTEIN_RESPONSE
MTORC1_SIGNALING
DNA_REPAIR
−log10(FDR)
ES
Phospho GSEA score (CPTAC)
Phospho GSEA score (Our study)
g
MYC_TARGETS_V1
UNFOLDED_PROTEIN_RESPONSE
DNA_REPAIR
E2F_TARGETS
G2M_CHECKPOINT
MYC_TARGETS_V2
−log10(FDR)
ES
Acetyl GSEA score (Our study)
h
CDK2
RPS6KA3
CDK12
CDK1
CDK4
PPP2CA
MAPK14
Log2FC(DEP)
−log10(FDR)
Kinase activity (Our study)
i
CDK2
CDK1
EEF2K
CDK4
PKN1
CHEK1
PLK1
Log2FC(DEP)
−log10(FDR)
Kinase activity (Our study)

Supplementary Fig. 4| Proteogenomic features underlying whole-genome doubling in NSCLC subtypes. **a**, Alluvial plot showing WGD-positive samples in original subtypes from the CPTAC LUAD cohort. **b**, Alluvial plot showing WGD-positive samples in original subtypes from the CPTAC LSCC cohort. **c**, Overlap of copy number signatures and the combined NMF subtypes in the CPTAC cohorts. The colors indicate the odds ratio of two-sided Fisher's exact test. The COSMIC v3 signature and etiology for each signature are indicated on the y-axis. Significance levels are denoted by the number of asterisks. *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. **d**, Mutational profile of CN signature in our cohort. **e**, Mutational profile of CN signature in CPTAC cohorts. The COSMIC v3 signature and etiology for each CN signature are written in each signature profile. LOH, loss of heterozygosity; 1x WGD, once-genome-doubled; 2x WGD, twice-genome-doubled; CIN, chromosome instability. **f**, Dot plot of phosphoprotein-level gene set enrichment analysis (GSEA) revealing upregulated pathways in Subtype 3 in both cohorts. **g**, Dot plot of acetyl protein-level GSEA revealing upregulated pathways in Subtype 3 in both cohorts. The x- and y-axis are enrichment scores (ES) from the current study and CPTAC NSCLC data, respectively. Labeled pathways are the top 6 upregulated pathways in Subtype 3. The Molecular Signatures Database (MSigDB) hallmark gene set v7.4 was used for GSEA. **h**, Enriched kinases in Subtype 1 LUAD compared to Subtype 2 LUAD in our cohort. The sizes of points indicate $-\log_{10}$(FDR) from kinase activity estimation. The colors of points indicate $\log_2$ fold changes of protein expression in Subtype 1 LUAD compared to Subtype 2 LUAD in

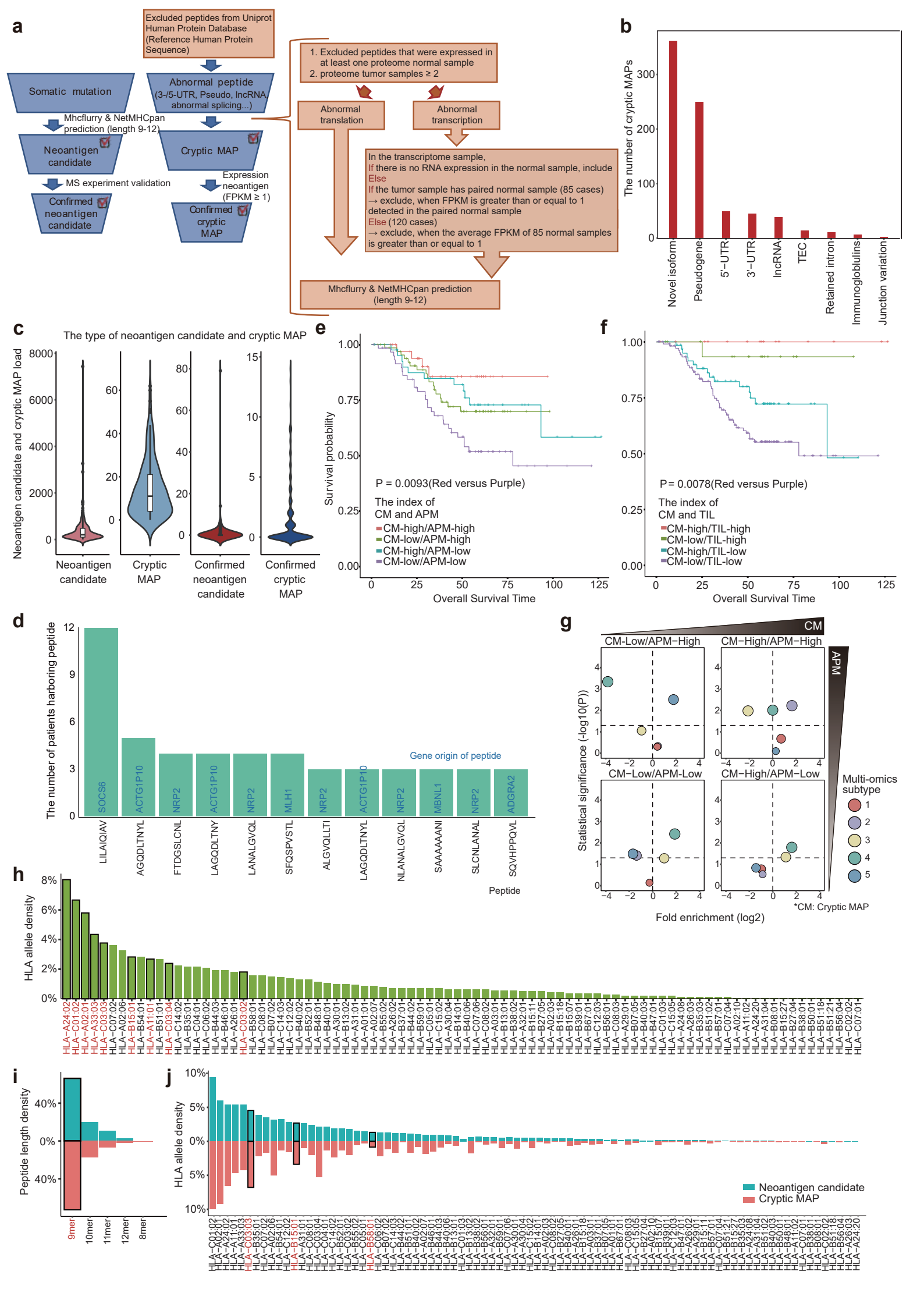
our cohort. Significantly upregulated kinases were labeled ($P < 0.05$). **I**, Enriched kinases in Subtype 3 LSCC compared to Subtype 4 and 5 LSCC in our cohort. The sizes of points indicate $-\log_{10}$(FDR) from kinase activity estimation. The colors of points indicate $\log_2$ fold changes in protein expression in Subtype 3 LSCC compared to Subtype 4 and 5 LSCC in our cohort. Significantly upregulated kinases were labeled ($P < 0.05$).


a
Cold
Hot
Cold
Hot
1
2
3
4
5
Multi-omics subtype
Cell type-based immune clsuter
Pathway-based immune cluster
Multi-omics subtype
c
Immune cluster and cell types
Hazard ratio of overall survival
P
Neutrophils 1.45e−05
Epithelial cells 1.44e−02
B cells 3.61e−02
Memory B cells 2.49e−02
Plasma cells 4.32e−02
CD8+ Tcm 1.11e−02
naive B cells 2.21e−02
Basophils 2.99e−02
CD4+ naiveT cells 1.99e−02
Class switched memory B cells 7.81e−03
CD4+ Tcm 3.54e−02
NK cells 2.23e−02
CD4+ T cells 3.17e−03
CD8+ Tem 5.60e−03
CD8+ T cells 1.39e−03
Immune cluster and cell types
Hazard ratio of relapse-free survival
P
Epithelial cells 6.01e−03
Neutrophils 9.50e−03
MSC 8.91e−03
Pericytes 3.52e−02
Sebocytes 4.48e−02
B cells 4.79e−02
naive B cells 4.82e−02
Memory B cells 3.17e−02
CD4+ Tem 1.21e−02
CD8+ Tcm 7.92e−03
Class switched memory B cells 1.19e−02
HSC 4.37e−02
CD8+ T cells 7.96e−04
NK cells 3.10e−03
CD8+ Tem 6.07e−04
CD4+ T cells 1.93e−04
Basophils 8.43e−04
CD4+ naiveT cells 6.46e−04
CD4+ Tcm 1.91e−03
Immune cluster 1.65e−03
d
Immune cluster and pathways
Hazard ratio of overall survival
P
Inhibitory molecules 2.00e−01
Inhibitor cells 8.12e−01
Priming and activation 6.15e−01
Trafficking and infiltration 5.63e−01
Recognition of tumor cells 5.33e−01
Checkpoint expression 1.93e−01
T cell immunity 1.23e−01
Immune cluster 2.29e−01
Immune cluster and pathways
Hazard ratio of relapse-free survival
P
Inhibitory molecules 2.10e−01
Inhibitor cells 8.16e−01
Priming and activation 6.91e−01
Trafficking and infiltration 4.78e−01
Recognition of tumor cells 2.81e−01
Checkpoint expression 8.21e−02
T cell immunity 6.09e−02
Immune cluster 1.58e−01
b
DEG CD8 T cells
T cell CD8
T cell CD8 effector memory
T cell CD8 terminally exhausted
T cell CD8 dividing
DEG CD4 T cells
T cell CD4
T cell CD4 dividing
DEG Tregs
T cell regulatory
DEG B cell
B cell
DEG NK cell
NK cell
DEG Neutrophils
Neutrophils
DEG Dendritic cell
pDC
cDC2
DC mature
cDC1
DEG Monocyte
Monocyte
DEG Macrophage
Macrophage
DEG Epithelial cell
Epithelial cell
UMAP1
UMAP2
Module score
e
Treg-depleted HTE
Treg-enriched HTE
CTE
Survival probability
P = 0.42 (Green vs Yellow)
Overall survival time (month)
P = 0.13 (Green versus Yellow)
Relapse-free survival time (month)
g
Biological category of immunomodulator
Co−stimulator
Co−inhibitor
Ligand
Receptor
Cell adhesion
Antigen presentation
Other
CD8 T cells
CD4 T cells
Tregs
B cells
NK cells
Neutrophils
DC
Monocytes
Macrophages
Epithelial cells
Immune cluster
TGFB1 CXCL9 LAG3 CD80 PDCD1LG2 IL2RA CD40 TNFRSF9 IL2 TNFRSF14 BTLA BTN3A1 BTN3A2 IFNG CCL5 PRF1 HMGB1 CD276 VEGFA EDNRB SELP VSIR CD40LG ADORA2A CD28 ICOS PDCD1 CD27 GZMA CTLA4 TIGIT SLAMF7 CD96 ITGB2 TLR4 TNFRSF4 HAVCR2 ENTPD1 IL10 CD86
Correlation coefficient between RNA expression and cell type or immune cluster
TNFRSF14 ICAM1 CCL5 SELP SLAMF7 CD40 GZMA PRF1 ICOS BTLA ENTPD1 VEGFA VTCN1 PVR NECTIN2
Correlation coefficient between Protein expression and cell type or immune cluster
VEGFA CEACAM1 MICB CD274 TNFSF4 CXCL10 LAG3 GZMA IL2 TLR4 IL1B CD70 PDCD1LG2 IL2RA ICOSLG TGFB1 TNFRSF14 IL1A TNFRSF4 IL10 EDNRB ICAM1 IFNG CD86
Correlation coefficient between Protein activity and cell type or immune cluster
f
EPITHELIAL MESENCHYMAL TRANSITION
ALLOGRAFT REJECTION
IL6 JAK STAT3SIGNALING
IL2 STAT5 SIGNALING
INTERFERON GAMMA RESPONSE
INTERFERON ALPHA RESPONSE
DNA REPAIR
G2M CHECKPOINT
E2F TARGETS
GLYCOLYSIS
OXIDATIVE PHOSPHORYLATION
FATTY ACID METABOLISM
PEROXISOME
RNA level
Protein level
Phospho−protein level
Acetyl−protein level
Enrichment difference between HTE and CTE
CTE
HTE
Immune-related pathway
Cell cycle-related pathway
Metabolic pathway
h
RNA expression of SLAMF7
Protein expression of SLAMF7
P = 3.77E-01
P = 1.77E-10
P = 9.01E-02
P = 9.78E-14
P = 8.24E-01
P = 3.12E-05
P = 3.66E-02
P = 1.18E-05
WT Mutation
CTE
HTE
i
Cell type-based IC
HTE
CTE
NAT-enriched
Cell type enrichment score
P = 4.64E-29
Mean expression of DEG HTE
CD8 T cells
CD4 T cells
Tregs
B cells
NK cells
Neutrophils
DC
Monocytes
Macrophages
Epithelial cells
Feature gene expression
Subtype1
Subtype2
Subtype3
Subtype4
Subtype5

Supplementary Fig. 5| The characteristics of immune clusters among patients with NSCLC and their correlations with multiomics subtypes. a, Comparison of cell type-based immune cluster, pathway-based immune clusters, and multiomics subtypes. Most of the hot-tumor-enriched (HTE) tumors were included in multiomics subtype 2 and 5. In contrast, a large portion of cold-tumor-enriched (CTE) tumors were matched with multiomics subtype 1, 3, and 4. **b**, The DEG for each cell type was used to generate the UMAP plot of scRNA-seq data. Each point color represents the module score of the cell type-specific DEGs, and the higher score is expressed in red color. The UMAP information was obtained from the multiple NSCLC studies (Salcher et al., 2022). **c-d**, Forest plots showed hazard ratios of OS (left) and RFS (right) across 64 immune and stroma cells in addition to the cell type-based immune cluster (**c**), and 7 immune-related pathways in addition to the pathway-based immune cluster (**d**). Only the results that were statistically significant were described in section (**c**). A lower hazard ratio than zero (blue box) indicates that HTE or a high score of cell type and pathway were associated with longer survival. All hazard ratios (grey lines) are estimated with 95% confidence intervals (CIs). The p-value indicates statistical significance for the survival analysis with log-rank Mantel-Cox test. A red text indicates statistical significance for the survival analysis with log-rank Mantel-Cox test. **e**, Kaplan-Meier curve shows the survival pattern of the 3 groups of patients (n = 174) divided by the status of Tregs, and immune clusters. The p-value was obtained by survival comparison between the two groups of Treg-enriched and Treg-depleted HTE using the log-rank

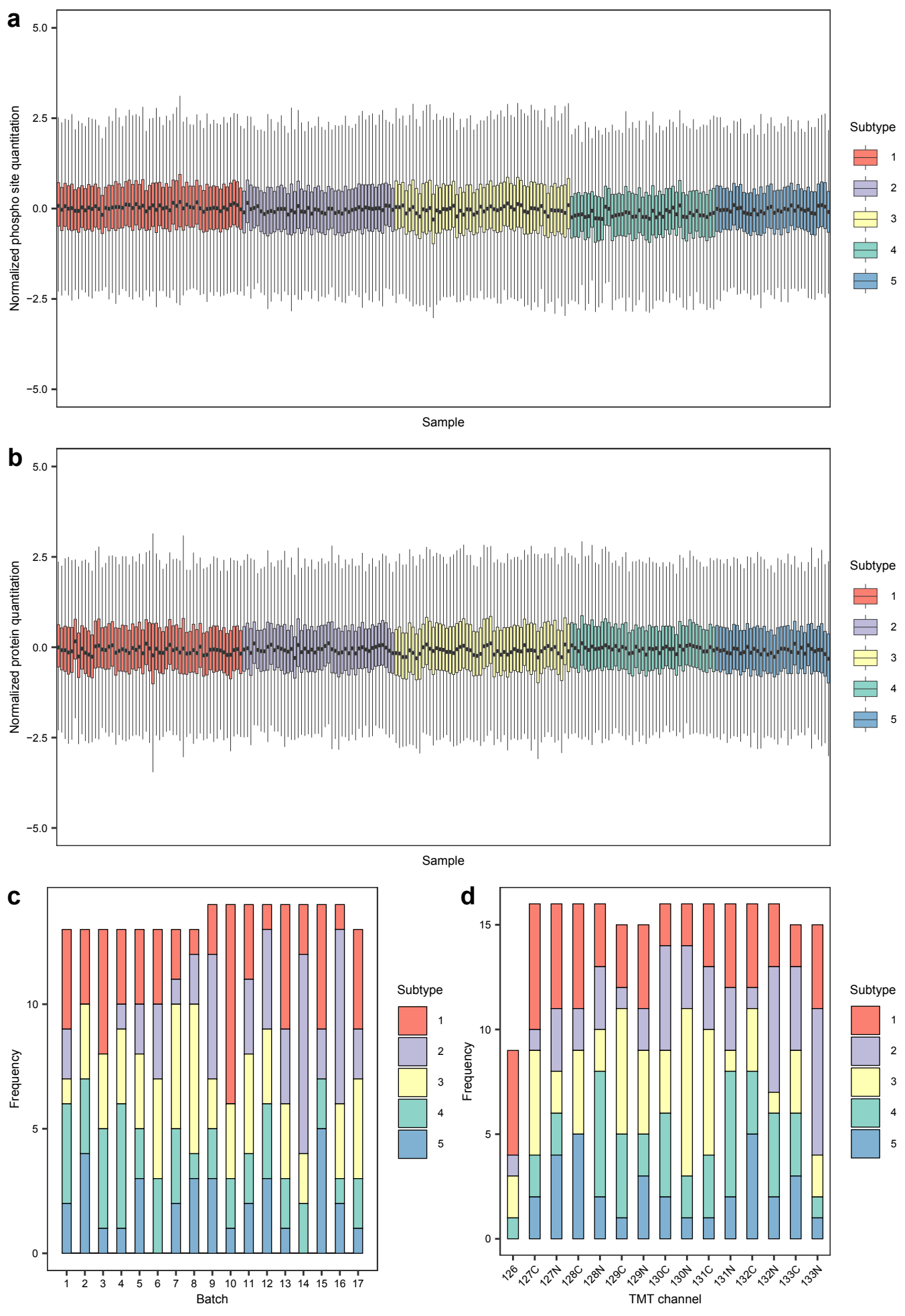
Mantel-Cox test. **f**, Pathway enrichment analyses were performed between HTE and CTE at transcriptome levels, proteome, phosphoproteome, and acetylproteome using single-sample GSEA and PTM-SEA. The positive and negative t-values of the relevant pathways are colored as red and blue circles. Each analysis result was described as a grey circle when it was not statistically significant ($P < 0.05$). The MSigDB hallmark gene set v7.4 was used for the pathway analysis. **g**, The expression (top and middle) or activity (bottom) of 40 immunomodulators that were known to be cancer cell ligands and immune cell inhibitory receptors[14,121] were tested for correlation with the enrichment score of cell types including the status of immune cluster. Only the immunomodulators having a significant correlation with at least a cell type and with the immune cluster were analyzed for their associations. The clustering was performed column-wise using Euclidean distance measurements. Each correlation was described as a grey color when it was not statistically significant ($P < 0.05$). Correlation coefficients and p-values were obtained from a generalized linear model (GLM). **h**, The left boxplots display the RNA and protein expression of *SLAMF7* according to the *SMARCA4* mutation status (top) and immune cluster (bottom) in two independent multi-omics cohorts derived from the studies of Satpathy et al. ($n = 202$) and Gillette et al. ($n = 211$). The right boxplots depict similar analysis results for the integrative cohort, which includes our cohort in addition to the two independent cohorts. A two-sided t-test was conducted to assess the differences in expression. The box represents the 25th and 75th percentiles, the central mark denotes the median, and the whiskers extend to the

most extreme points within ±1.5 × IQR. **i**, (left) Immune subtyping was performed based on the enrichment score of cell types across three multi-omics cohorts including our cohort as described in (h). The consensus clustering was performed in the integrated cohorts. (right) The box and balloon plot show the mean expression of feature genes of HTE and 10 cell types across the integrated multi-omics subtypes. The feature genes were defined as top 300 and 30 genes that were overexpressed in HTE samples and cell-type highly enriched samples, respectively. The center lines, edges, and whiskers of the boxplot indicate the medians, 25th and 75th percentiles, and the most extreme points within ±1.5 × IQR, respectively.


a
Excluded peptides from Uniprot Human Protein Database (Reference Human Protein Sequence)
Somatic mutation
Abnormal peptide (3-/5-UTR, Pseudo, lncRNA, abnormal splicing...)
Mhcflurry & NetMHCpan prediction (length 9-12)
Neoantigen candidate
Cryptic MAP
MS experiment validation
Expression neoantigen (FPKM ≥ 1)
Confirmed neoantigen candidate
Confirmed cryptic MAP
1. Excluded peptides that were expressed in at least one proteome normal sample
2. proteome tumor samples ≥ 2
Abnormal translation
Abnormal transcription
In the transcriptome sample,
If there is no RNA expression in the normal sample, include
Else
If the tumor sample has paired normal sample (85 cases)
→ exclude, when FPKM is greater than or equal to 1 detected in the paired normal sample
Else (120 cases)
→ exclude, when the average FPKM of 85 normal samples is greater than or equal to 1
Mhcflurry & NetMHCpan prediction (length 9-12)
b
The number of cryptic MAPs
Novel isoform
Pseudogene
5'-UTR
3'-UTR
lncRNA
TEC
Retained intron
Immunoglobulins
Junction variation
c
The type of neoantigen candidate and cryptic MAP
Neoantigen candidate and cryptic MAP load
Neoantigen candidate
Cryptic MAP
Confirmed neoantigen candidate
Confirmed cryptic MAP
e
Survival probability
P = 0.0093(Red versus Purple)
The index of CM and APM
CM-high/APM-high
CM-low/APM-high
CM-high/APM-low
CM-low/APM-low
Overall Survival Time
f
P= 0.0078(Red versus Purple)
The index of CM and TIL
CM-high/TIL-high
CM-low/TIL-high
CM-high/TIL-low
CM-low/TIL-low
Overall Survival Time
d
The number of patients harboring peptide
Gene origin of peptide
SOCS6
ACTG1P10
NRP2
ACTG1P10
NRP2
MLH1
NRP2
ACTG1P10
NRP2
MBNL1
NRP2
ADGRA2
LILAIQIAV
AGQDLTNYL
FTDGSLCNL
LAGQDLTNY
LANALGVQL
SFQSPVSTL
ALGVQLLTI
LAGQDLTNYL
NLANALGVQL
SAAAAAANI
SLCNLANAL
SQVHPPQVL
Peptide
g
CM
APM
CM-Low/APM-High
CM-High/APM-High
CM-Low/APM-Low
CM-High/APM-Low
Statistical significance (-log10(P))
Fold enrichment (log2)
Multi-omics subtype
*CM: Cryptic MAP
h
HLA allele density
i
Peptide length density
9mer
10mer
11mer
12mer
8mer
j
HLA allele density
Neoantigen candidate
Cryptic MAP

Supplementary Fig. 6| The process of identifying neoantigens and cryptic MAPs, along with their characteristics. **a**, Schematic diagram shows how to identify the type of neoantigen candidate and cryptic MAP. The main processes are shown in blue and detailed processes are shown in orange. **b**, Number of cryptic MAPs derived from novel isoforms, pseudo genes, untranslated regions (UTR), long noncoding RNAs (lncRNA), to be experimentally confirmed (TEC), retained introns, immunoglobulins, and junction variations is shown in the bar plot. **c**, Distribution of the neoantigen candidate load and cryptic MAP load across the patients was described accordingly by the type of neoantigen candidate and cryptic MAP. The middle lines, edges, whiskers, and black dots of the boxplot represent the medians, 25th and 75th percentiles, the most extreme points within ±1.5 × IQR, and outliers, respectively. **d**, Bar plot shows the number of confirmed cryptic MAPs observed in more than 3 patients in descending order. The character on the bar indicates the gene from which the peptide is derived. **e-f**, Kaplan-Meier curve shows the survival pattern of the four groups of patients (n = 174) divided by confirmed cryptic MAP load and antigen presenting machinery (APM, e) or TIL pattern (f). The p-value was obtained by curve comparison between the two groups having the largest difference with the two-sided log-rank Mantel-Cox test. **g**, Result of enrichment analysis for the four groups described in Figure S6E to the multi-omics subtypes. X- and Y-axis indicate an enrichment and a statistical significance calculated using two-sided Fisher's exact test with Benjamini-Hochberg adjustment. The size of the dot represents the statistical significance. **h**, The bar plots illustrate the frequency

of HLA alleles within our patient cohort. Red highlights denote HLA alleles observed at a frequency of more than 10% in the Korean population, as reported in a previous study[58,59]. **i**, The bar plot displays the frequency of the length of neoantigen candidates and cryptic MAPs in our binding prediction procedure. The red highlight indicates the most common peptide length observed in a large-scale peptidome study[60]. **j**, The frequency of HLA alleles harboring the peptide-MHC complex is depicted. Red highlights indicate HLA alleles that presented shared neoantigens in a previous lung cancer study[61].


a
Normalized phospho site quantitation
5.0
2.5
0.0
−2.5
−5.0
Sample
Subtype
1
2
3
4
5
b
Normalized protein quantitation
5.0
2.5
0.0
−2.5
−5.0
Sample
Subtype
1
2
3
4
5
c
Frequency
10
5
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Batch
Subtype
1
2
3
4
5
d
Frequency
15
10
5
0
126 127C 127N 128C 128N 129C 129N 130C 130N 131C 131N 132C 132N 133C 133N
TMT channel
Subtype
1
2
3
4
5

Supplementary Fig. 7| Quality assessment of phospho quantitation in 229 patients from a Korean NSCLC cohort. **a**, Distribution of normalized protein quantitation. **b**, Distribution of normalized phospho site quantitation. **c**, Proportion of subtypes for 17 TMT batches. **d**, Proportion of subtypes for TMT channels.