nature portfolio

Peer Review File

Proteogenomic Analysis Reveals Non-small Cell Lung Cancer Subtypes Predicting Chromosome Instability, and Tumor Microenvironment

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author): expertise in non-small cell lung cancer proteomics

This manuscript describes the proteogenomic analysis of a cohort of non small cell lung cancer (NSCLC) tumor samples obtained from Korean patients. One notable aspect of the paper is the integration of previous studies of tumors collected from patients primarily in the United States that mostly have European ancestry. The thorough comparison of their current dataset, generated to the same high standards and high content, to these previously published datasets is a significant strength. The resulting molecular classification has strong interest, and the initial connections between the subgroups and the selection of appropriate therapy (e.g. selinexor) is initially investigated and discussed, providing additional novel insights and potential long-term benefit for these patients. This manuscript is expected to be of high interest and can serve as a model for research to increase the diversity of the patient populations examined with proteogenomics to address health disparities. As an example, these strategies can be extrapolated to future datasets for tumors from patients with African ancestry or Hispanic ethnicity as well as other underserved populations.
The methods are sound, and the descriptions are sufficient to enable replication in the labs that have the resources.

Minor points need to be addressed in revision, as described below. Minor editing for spelling, grammar, and usage is needed.

1. In several instances, the authors describe the use of custom databases for searching for peptides from specific protein groups (e.g. immunoglobulins). The point should be made that without concatenation with all of the human sequences, a small percentage of tandem mass spectra can be incorrectly assigned or assigned to different peptide sequences in each custom database search. Have the authors made an effort to address this issue?

Specific Comments on Text and Figures

p. 4 Line 21: References 17 and 18 do not seem to match the discussion of multi-omics datasets as they discuss pathology review of H&E stained slides.

p. 5 line 60: More granularity is needed for the 27 tumors of other types. A supplemental

table with individual patient information should be referenced here in the text.
p. 6 lines 90-95: It would be interesting to the readers to describe how much each dataset contributes to the classification signatures. In other words, where is the most value derived? Is the addition of acetylation a key experiment even if the use of deacetylase inhibitors is not discussed. This commentary belongs in the discussion, rather than in the results.
p. 26 line 580: The use of the word "identified" may be misinterpreted, because it creates the expectation that peptides presented by MHC complexes were captured and analyzed with LC-MS/MS. Other verbs (posited, proposed, inferred, etc.) would be more appropriate; description of these peptides as candidate antigens would also be correct.
p. 29 lines 652-3: Additional discussion is needed to connect the histology of the tumors. As an example, the oncogenic drivers are more clearly defined in adenocarcinoma than squamous cell carcinoma, so that association may be expected by the reader.
p. 30 Lines 675-678: Can the authors propose a biomarker panel that could be used to select patients for Selinexor treatment? How does this potential therapy impact the patients included in each of the studies? LSCC has such high need for novel therapies that it may be able to proceed quickly to clinical investigation.
p. 32 line 741: Provide a reference or protocol for buffy coat enrichment of blood cells.
p. 38 line 942: Provide additional details for the peptide concentration with NanoDrop One.
p. 39: TMT labeling sections with protocol steps should be written as full sentences.
p. 39 Line 968; Text is unclear; use of interrogate is incorrect.
p. 40 and 41: Convert CAN to ACN.
p. 43 and throughout the manuscript: Check spelling of Lehtio.
p. 53 line 1405: The evidence of neo-antigens in MS needs to be clarified. Do the authors mean detection of the mutant peptide containing the amino acid variant or just detection of any peptides from the protein of origin?

Figure Legends: Full definitions of abbreviations should be available in the captions.
Figure 1 describes the presence of different drivers. Were there enough patients to get clear signal for categorization of these tumors? In other words, do ALK fusions always belong to one of the clusters defined in the manuscript? I think this point bears more discussion either in the body of the manuscript or in a supplement.
Figure 2b. The third panel (right) is unclear. The reclassification from the two other papers

makes sense, but the combination and reclassification seems to find mostly the same groups. Better explanation is needed to for the few samples that are reclassified in the third panel.
Figure 2C: The significant kinases are assessed by their estimated activity from phosphoproteomics. Correlation of the kinase expression and estimated activity should be investigated.
Figure 2D: The distributions of the prognostic markers are overlapping. Another metric, like Bhattacharya or Hellinger distance, is recommended to evaluate whether these metrics can have clinical utility. Alternatively, receiver operating characteristic (ROC) curves could be applied to the data.
Figure 3B: Increased magnification or arrow indicators should be used to highlight the parts of the tumor that are most relevant. The details are lost at this size in the figure.
Figure 4I: State clearly that the kinases are inferred from the phosphoproteomics results in each instance in the manuscript.
Figure 5F: These results may be driven by the high outliers in the smaller population being compared.

Extended Data Figure 1: Include the definitions for abbreviations in the caption.
Extended Data Figure 2: Correct spelling of phospho- and indicate that kinase activity is inferred from phosphoproteomics.
Extended Data Figure 4: Panels A and B should be flipped to start with the original classification and end with the WGD status. In panels H & I, correlation of the kinase expression and estimated activity should be investigated.

Reviewer #2 (Remarks to the Author): expertise in NSCLC genomics
Review Comments to Song et al. 2023

In this manuscript Song and coworkers has performed a proteogenomics analysis of a Korean cohort of 229 NSCLC cases. Multiple omics methods were used to collect the data including genomics, transcriptomics, proteomics, and PTM analysis (phosphorylation and acetylation). This comprehensive analysis contributes an additional resource for

investigating the biology of NSCLC. I compliment the authors on their massive work! Further, the authors use their data to define 5 subtypes of NSCLC and characterize the subtypes in relation to clinical data and clinically relevant areas such as prognostication and cancer therapy prediction. Four of the identified subtypes overlap with previously described NSCLC subtypes, while one (subtype 4) is described as a novel subtype. Although the analysis is based on in-depth data and potentially outputs some new findings, several questions need to be addressed before considering publication in Nature Communications. Most importantly, the potential non-biology related factors that may drive the novel subtype 4 needs to be investigated (see below). Further, the analysis of PTM data is in my view underdeveloped and could be much improved, potentially resulting in new findings. In addition, the potential value of XPO1 as a therapeutic target in WGD LSCC needs further validation. Some analyses and interpretations are unclear and needs further clarification (see specific comments below). In summary, I do not recommend publishing the manuscript in Nature Communications in its current form.

General comments

• Some of the figures and fonts are very small and difficult to read. This should be adjusted to increase the readability of the manuscript.
• Why is the LSCC samples so unbalanced between sexes (97% male)? How could this affect the interpretation of the analysis? Any LSCC finding may be confounded by gender. Test for independence if needed.
• It is unclear how imputation might affect the findings reported. A plot showing the percent imputation per sample, and an analysis to show if some subtypes have more imputed values would help. This would be relevant for total proteomics as well as phosphor and acetyl PTM data.
• To my understanding, the PTM data (phosphor and acetylation) is not normalized to total protein levels. If so, differences in total protein levels will drive the analysis and interpretation of phosphor and acetylation data. As an example, Figure 4e and Extended Data Figure 4f and 4g are almost identical, which is most likely a result of differential protein expression, and not differential phosphorylation or acetylation. In a similar way, the kinase activity prediction analysis would be impacted by total protein levels, and therefore

potentially biased. To really investigate the impact of PTMs on biology, which is an important topic, normalization of the PTM data against total protein levels can output new information that cannot be found through total protein level analysis.

Specific comments and questions

1. Subtype identification and characterization.
In general, enrichment analysis based on broad biology as in for example the mutSigDB Hallmarks genesets is only providing a very general and often not very specific output. To strengthen the characterization of the different subtypes it would be good to provide additional support based on the generated data (see below).
a. Subtype 1 is described as metabolic based on GSEA Hallmarks enrichment in Oxidative phosphorylation, mitochondrial matrix, and cellular respiration. Are there other supports for this interpretation? Do you see enrichment of STK11 mutations? Activation of the mTOR pathway?
b. Subtype 2 is described as early tumorigenesis and early-stage disease based on enrichment in IL-33 and Notch pathway. Other supports for this interpretation? Was TNM stage significantly different between Subtype 1 and 2?
c. Subtype 3 is characterized by male gender, LSCC, smokers, TP53mutations and WGD events. Are these findings independent in multi variate analysis? It looks like the subtype core is almost exclusively males with LSCC. In other words, Subtype 3 seems to be driven by histology and should perhaps be annotated as LSCC rather than proliferative.
d. For subtype 4, enrichment was found for hypoxia. In the text (row 125-126 in manuscript) it is written that this enrichment was found in the phospho data. In the corresponding figure (Figure 1e) it is written “protein”. Which data was used for the enrichment analysis? The same question for Neutrophil degranulation (text: phospho, figure: protein).
e. Subtype 4 is described as a “chromosomally stable tumor suppressor-driven mesenchymal subtype” based on the enrichment analysis. It is difficult to follow how this interpretation was done based on the analysis. What is meant by “tumor suppressor driven”? Is there additional support for this interpretation?
f. Was enrichment analysis performed for individual oncogenes and tumor suppressors defined here as “other” in a similar way that was performed for EGFR and TP53? Were there

any specific enriched oncogenes and tumor suppressors?

2. Characterization of Subtype 4 described as a novel NSCLC subtype.

a. Subtype 4 is almost exclusively driven by the phospho-proteomics data (96% of NMF features, row 183). A concern here is that this subtype could be driven by the quality of the generated phospho-data.

b. Were there significant differences in the TMT channel quant distribution of the phospho-proteomics data between the samples/subtypes? Boxplots showing the phospho quant distribution in each sample across the cohort would help evaluating this.

c. Were there differences in imputation of phospho-data between subtypes?

d. Was meta-data collected for the sample-collection and sample prep pipeline? The integrity of phosphorylations is very dependent on sample handling, and differences can impact the results of the analysis.

e. For subtype 4, 18/43 samples (42%) were missing transcriptomics data. For the other subtypes transcriptomics data were only missing in very few samples (subtype 1: 2/55 samples, subtype 2: 1/45, subtype 3: 0/52, subtype 5: 3/34). Please explain this difference. If the quality of the sample was too poor for transcriptomics analysis in a large part of the subtype 4 samples, this could indicate that also the proteome/PTM analysis was affected by sample quality. Was RNA quality used for selection of samples for transcriptomics? What is the distribution of RNA quality metrics across the full cohort? Is there a statistical difference in RNA quality between the subtypes?

f. Was the quantitative phospho-data normalized to total protein levels before the differential analysis was performed? If not, the phospho-analysis (and any other PTM analysis such as acetylation) will be driven largely by the total protein levels.

g. How many of the Subtype 4 features were substrates of CSNK2A1?

h. Was the total protein level of SLK or LRRFIP1 prognostic in the current cohort and in the CPTAC cohort?

3. Cellular landscape of the five subtypes

a. Subtype 5 is largely driven by immune infiltration. Yet about 25% of the Subtype 5 samples has low tumor infiltrating immune cell components (Figure 3k). What other features could explain the clustering of these non-infiltrated samples into Subtype 5?

4. Proteogenomic features underlying whole genome doubling in NSCLC.

a. One of the main findings is related to whole genome doubling (WGD, more than half of the chromosomes are gained) which is found specifically enriched in subtype 3 and subtype 1. The WGD analysis was based on whole exome sequencing and not whole-genome sequencing, which according to the authors may be a limitation in the study. A validation for a subset of cases using whole-genome sequencing could strengthen this finding.

b. Due to the high overlap between subtype 3, LSCC and genome doublings (WGD), specifically in the subtype 3 core, it is important to investigate the various subtype 3 findings in relation to histology. It is difficult to evaluate what is driving the findings without such analysis. The same is true for the validation of findings against CPTAC data where the authors have combined LSCC and LUAD histologies. Specifically:

i. Is TP53 more commonly mutated in LSCC cases from subtype 3 than in LSCC cases from other subtypes?

ii. Are WGD more common in LSCC/LUAD cases from subtype 3 than in LSCC/LUAD cases from other subtypes?

iii. Is XPO1 expression higher in LSCC than in LUAD in general?

iv. Is XPO1 expression higher in in LSCC cases from subtype 3 than in LSCC cases from other subtypes?

c. The XPO1 inhibitor screen was not performed in LSCC cases without WGD, and therefore it is impossible to judge if the sensitivity depends on histology or WGD. For assessing this, the experiment should be complemented with several LSCC organoids without WGD.

5. Immune landscape in NSCLC

a. Due to lack of transcriptomics data, relatively few subtype 4 samples were included in the immune cluster analysis. Further, 17/25 (68%) of Subtype 4 samples cluster with NAT (normal adjacent tissue). What drive the clustering of NAT, and why are the subtype 4 samples overrepresented in this cluster?

6. Multiomics profiling of neoantigens and immune clusters

a. To “confirm” a neoantigen candidate by proteomics support, a rule was set so that at least 20% of the total intensity of all reporter ions for the neoantigen comes from the sample with the corresponding somatic mutation (row 1301-1302). It would be important to

also assess if the sample with the mutation has the overall highest intensity in the set. In principle, the mutated sample should have a clear outlier pattern in the quantification, and all non-mutated samples should have “background” values.

b. Were the cryptic peptides confirmed using MS data?

c. Cryptic MAPs were positively correlated with immune infiltration which was associated with better survival (Figure 5c). Were the cryptic MAPs prognostic independently of immune infiltration?

Reviewer #3 (Remarks to the Author): expertise in computational muti-omics analysis

Song et al. present a thorough and comprehensive multi-omics analysis of a large Korean NSCLC cohort leading to the identification of 5 subtypes by using non-negative matrix factorization (NMF) clustering. Their approach to look at all NSCLC histologies contrasts with other large studies that have tended to focus separately on LUAD or LUSC histological subtypes. Their approach addresses the considerable overlap that can exist between the two major histological subtypes and more minor histologies that are less well characterized.

They analyzed 229 NSCLC patient tumors and a "replication" cohort of 462 patients from published studies. Their data includes a large-scale single-cell RNAseq dataset. Also, they did a histological review of patient data related to TILs, identified potential neoantigens in the tumor microenvironment and observed varying efficacy of adjuvant therapies between subtypes. Their data include genomic (WES), transcriptomic, proteomic, phosphoproteomic, and acetylprotoemic datasets. Depth of coverage is excellent. For example, in their proteomics studies they list more than 10K proteins, 40K phosphoproteins, and approx. 6K acetyl proteins in at least 30% of samples. Relative to other large-scale multi-omics/proteogenomics studies of lung cancer, this study matches if not surpasses them for its comprehensive generation and integration of data with clinical features and outcomes. Their comparative analysis is generally in agreement with and extends the insights described in earlier reports from TCGA and CPTAC studies. They note that worse outcome associated with their subtype 4 does not extend to the CPTAC cohort and thus may reflect an ethnicity impact.

As is the nature of such large-scale analyses, there is considerable conjecture on the role of differentially expressed genes and proteins and protein features. They describe evidence for selinexor sensitivity in patient-derived organoids associated with whole genome duplications (Subgroup 3). While a preliminary result, it represents an experimental test of an emerging hypothesis. Overall, the vast amounts of data and data analysis are convincing and of high quality. The paper is well written and conclusions clearly presented.

Comments:

1. The figures are extremely difficult to read due to small fonts. In Fig 2a the symbols (dot, faint rectangle, blank) are not defined.
2. The results associated with Fig 2 are confusing as written.
3. Line 170 reads: "four subtypes" ...should this be five subtypes?
4. The meaning of the many semi-transparent connections shown in Fig 2b are confusing and not explained.
5. What is the statistical significance of the correlation/overlaps indicated in Fig 2b?
6. Regarding the description of Subtype 4 and its lack of enrichment of features....is this referring to LUSC only, i.e., middle panel of Fig 2b?
7. How was kinase activity measured (line 184)?
8. How specifically are the indicated FDR values for CSNK2A1 and GSK3B (line 186) supporting the conclusions drawn on the activation of these kinases?
9. What are the specific phosphorylation sites quantified for CSNK2A1 and GSK3B , and does the literature support that these modifications are activating as opposed to inhibitory? Are they known to be sites that are regulated?
10. They describe phosphorylation at position S347 in SLK as "significantly upregulated" in subtype 4 (line 192). Have they quantified the level of SLK protein in comparison with pS347 peptides to justify the conclusion that the stoichiometry of phosphorylation is actually increased, rather than an increased expression of SLK with no change in the level of phosphorylation at S347.
11. Is there evidence that S347 is phosphorylated by CSNK2A1?

Reviewer #4 (Remarks to the Author): expertise in NSCLC neoantigens

The authors conduct a study of comprehensive multiomic analysis of 229 patients in Korea with NSCLC, performing whole exome sequencing, bulk RNA sequencing, and global/phospho/acetyl proteomic analyses. Global proteomic, phosphoproteomic, and acetylproteomic data were integrated and non-negative matrix factorization clustering performed to define 5 subtypes of NSCLC. The authors benchmarked their proposed subtype classification against previously published classifications and performed comparative analyses (Figure 2). The investigators glean important biological insights from their subtype classification and propose some potential therapeutic candidates.

While their methods are similar to those in previously published cohorts, they perform their analysis in a Korean population (representing a lung cancer population enriched in adenocarcinoma and exhibiting low tumor mutation burden of 2.7 TMG that is distinct from the population studied in a Western population) and offer a valuable multiomic meta-analysis comparing to previously published data.

Strengths:

-Some of the biological insights they offer include the following: propose XPO1 as a potential druggable target with Selinexor for subtype 3, highlight a subtype 4 with potential targets in PI3K, VEGF, HIF pathways and neutrophil pathway requiring further characterization, and demonstrate cryptic MAPs as a features of subtype 2 which may be a potential vaccine target. They also show that SMARCA4 mutations are associated with increased expression of SLAMF7 in subtype 5, which might have therapeutic implications.

-Authors perform an analysis of both conventional and cryptic MAPs

Limitations:

-A limitation of the study is that the subtype classification is mainly pertinent to adenocarcinomas (subtypes 1, 2, 4, 5), as essentially all the lung squamous cell carcinomas fall into subtype 3. In addition, a limitation for interpreting how subtype classification impacts clinical outcomes (Figure 2g) is the heterogeneity in stage across the various subtypes, which would be expected to independently impact survival. A stage by stage comparison would be required to compare clinical outcomes across subtypes. Kaplan Meyer curves should include the number of patients being included at various time points.

-This study generates multiple new hypotheses that will require further evaluation in future studies that are beyond the scope of this study

Suggestions:

-The conclusions from the findings of adjuvant chemotherapy/adjuvant radiation clinical outcomes according to multiomic subsets appear to be overstated. Only half of patients received adjuvant chemo or radiation, and there are only 34 patients in subtype 5, really limiting the interpretation of the kaplan meyer curves in Figure 6g. I would like to see the impact of adjuvant chemotherapy on survival for the overall population (subtypes 1 – 5), as these treatments are typically associated with a 5% improvement in overall survival when analyzing large datasets of patients.

-In Figure 2h, outcomes for patients with and without metastases are compared in subtype 4, and the conclusion that metastasis is not the sole mechanism leading to poorer survival is a reasonable hypothesis, but not supported by the data which is not adequately powered to see a difference with only a handful of patients with metastasis.

-For the methods for calling cryptic MAPs, why was 3 frame translation (rather than 6 frame translational) used for searching for cryptic MAPs?

-Of the conventional and cryptic MAPs identified in this study, what was the distribution of predicted HLA binding and does this reflect the expected HLA distribution of a Korean population? Also, I would like to see the peptide length distribution, HLA allele distribution, and single nucleotide polymorphism data presented for cryptic MAPs vs conventional MAPs and how the results compared to those previously reported.

Clarifications:

-Line 720 – 726: Please clarify the method by which the 250 patients were selected among the 408 NSCLC patients. Manuscript states that these were selected based on patients with locally advanced (which I'm interpreting to be stage II – III) and metastatic (IV), but it's not clear to me that all the patients that are excluded are stage I patients per Fig 1A.

-In Figure 6B, is the legend mislabelled? Should the curve in blue be patients with recurrent cryptic MAPs? How is the term "recurrent cryptic MAPs" defined for the purposes of Fig 6B and is this different that the term "confirmed cryptic MAP"?

-Line 281: It is not feasible in clinical practice for adjuvant treatment to be based on

multiomics data. However, there are lessons learned from multiomic data studies that can have important implications for clinical practice.

-Line 685 – 686: ADAURA trial was not statistically designed to look for a difference in outcome for patients with and without chemotherapy, so this is not the appropriate interpretation of the trial results.

REVIEWER COMMENTS

Reviewer #1 (Remarks to the Author): expertise in non-small cell lung cancer proteomics

This manuscript describes the proteogenomic analysis of a cohort of non small cell lung cancer (NSCLC) tumor samples obtained from Korean patients. One notable aspect of the paper is the integration of previous studies of tumors collected from patients primarily in the United States that mostly have European ancestry. The thorough comparison of their current dataset, generated to the same high standards and high content, to these previously published datasets is a significant strength. The resulting molecular classification has strong interest, and the initial connections between the subgroups and the selection of appropriate therapy (e.g. selinexor) is initially investigated and discussed, providing additional novel insights and potential long-term benefit for these patients. This manuscript is expected to be of high interest and can serve as a model for research to increase the diversity of the patient populations examined with proteogenomics to address health disparities. As an example, these strategies can be extrapolated to future datasets for tumors from patients with African ancestry or Hispanic ethnicity as well as other underserved populations.
The methods are sound, and the descriptions are sufficient to enable replication in the labs that have the resources.

Minor points need to be addressed in revision, as described below. Minor editing for spelling, grammar, and usage is needed.

1. In several instances, the authors describe the use of custom databases for searching for peptides from specific protein groups (e.g. immunoglobulins). The point should be made that without concatenation with all of the human sequences, a small percentage of tandem mass spectra can be incorrectly assigned or assigned to different peptide sequences in each custom database search. Have the authors made an effort to address this issue?

Response: To reliably identify novel peptides, we used a multi-stage strategy: 1) the identification of canonical peptides, 2) the identification of modified peptides allowing up to 2,355 dynamic modifications as defined in Unimod, and 3) the identification of novel peptides. For each stage, the FDR was calculated separately using a target-decoy strategy, and identifications were obtained at 1% FDR. Only unidentified spectra from the previous stage were subjected to the subsequent stage. For novel peptide identification, we used two search tools, MS-GF+ and Comet, and rejected PSMs conflicting by both tools (i.e., identical spectra but different peptides assigned). In addition, the novel peptides were searched using BLAST and filtered out if there were peptide sequence matches in the various reference protein databases such as UniProt, RefSeq, and Gencode, allowing no more than a single amino acid substitution. We have added this content to the "Identification of Novel Peptides" section of our manuscript. The sections pertaining to immunoglobulins have been removed as antibody peptides were not analyzed in this study.

Specific Comments on Text and Figures

p. 4 Line 21: References 17 and 18 do not seem to match the discussion of multi-omics datasets as they discuss pathology review of H&E stained slides.

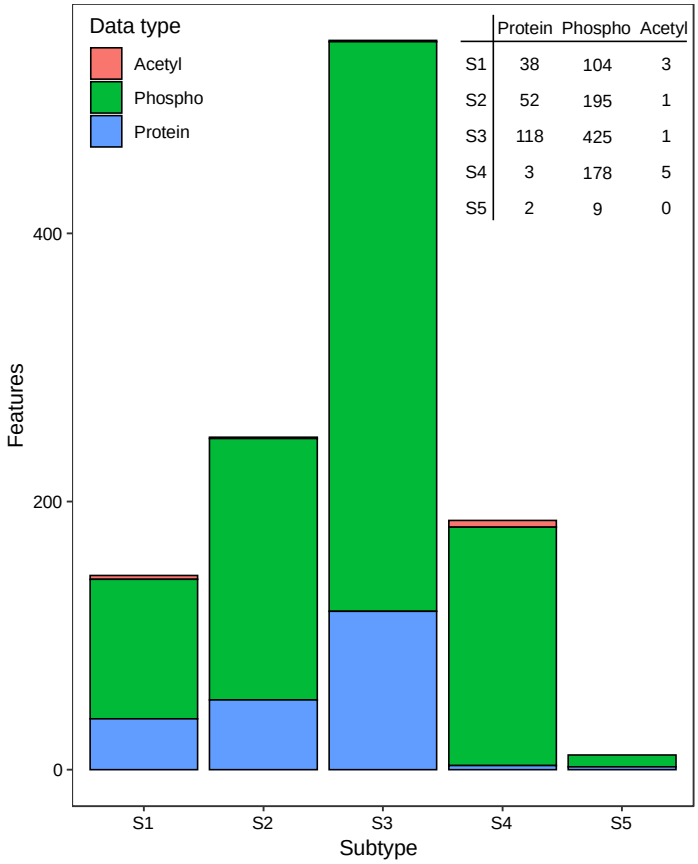
Response: The main insight from the referenced studies is that the pathological classification of lung cancer can be ambiguous and discrepant in many more cases than we expected. We believe that prior genomic or proteogemonic studies have concentrated on the specific subtypes of lung cancer. However, this approach may result in exclusion of cases where pathological subtyping was challenging or inconsistent, potentially resulting in the formation of distorted cohort. We have revised the sentences surrounding these references to more accurately convey these points.
(Page 4, Line 23-26)

p. 5 line 60: More granularity is needed for the 27 tumors of other types. A supplemental table with individual patient information should be referenced here in the text.

Response: We apologize for any confusion caused by our phrasing. By 'other types,' we referred to entries within the 'Path.Dx' column in **Supplementary Table 1a** that do not correspond to either adenocarcinoma or squamous cell lung carcinoma. To clarify this point, we have included an additional citation at the end of the sentence.
(Page 5, Line 61)

p. 6 lines 90-95: It would be interesting to the readers to describe how much each dataset contributes to the classification signatures. In other words, where is the most value derived? Is the addition of acetylation a key experiment even if the use of deacetylase inhibitors is not discussed. This commentary belongs in the discussion, rather than in the results.

Response: Thank you for your comment. Among the datasets analyzed, the phosphoproteome data stands out as the most valuable. Our NMF input included 16,370 proteins, 52,578 phosphosites, and 5,628 acetylsites. Among them, a total of 1,134 features were selected for the subtype signature, with 80.3% of these features originating from the phosphoproteome dataset. Analyzing the proportion of input features in each category, the global proteome dataset represented 1.3% (213 out of 16,370 proteins), the phosphoproteome dataset contributed 1.73% (911 out of 52,578 phosphosites), and the acetylome dataset contributed 0.18% (10 out of 5,628 acetylsites), respectively. Furthermore, the accompanying Reviewer-only figure 1 illustrates subtype signatures for the five subtypes identified. The phosphoproteome dataset's significant contribution is evident across all subtypes, with percentages ranging from 71.7% to 95.7%.



Reviewer-only Figure 1

Protein acetylation plays a crucial role in regulating gene transcription and signal transduction pathways[1]. A newly identified subtype, Subtype 4, exhibits significant prognostic features related to the HIF-1 signaling pathway. This includes a notable increase in acetylation at the K65 position of TIMP3, although it has not been identified as a signature feature of Subtype 4. TIMP3 is implicated in ANGPT2-regulated hypertension which could lead to poor prognosis[2], and its significant elevation in Subtype 4, compared to other subtypes, suggests a contribution to the HIF-1 signaling pathway.

Thus, we modified the text in Discussion as follows (Page 30, Lines 685-688):
The phosphoproteome dataset was the most informative for subtype identification, contributing 80% (911 out of 1,134) of the features, while global proteome and acetylome data also played crucial roles in decoding signaling pathways across the identified subtypes.

p. 26 line 580: The use of the word "identified" may be misinterpreted, because it creates the expectation that peptides presented by MHC complexes were captured and analyzed with LC-MS/MS. Other verbs (posited, proposed, inferred, etc.) would be more appropriate; description of these peptides as candidate antigens would also be correct.

Response: We corrected the description of the word as the reviewer suggested as follows:
(Page 27, Line 606) We inferred 85,430 neoantigen candidates and 775 cryptic MAPs

(**Supplementary Fig. 6a**) and annotated the origin of the cryptic MAPs based on the matched transcripts (**Supplementary Fig. 6b** and **Supplementary Table 6a**).

p. 29 lines 652-3: Additional discussion is needed to connect the histology of the tumors. As an example, the oncogenic drivers are more clearly defined in adenocarcinoma than squamous cell carcinoma, so that association may be expected by the reader.

Response: We briefly added some descriptions that several histological parameters were correlated with the multiomics subtypes, and stressed that the present subtype may delineate the grades of LUADs, targetable oncogenic drivers, metastatic potential, and tumor immune response.
(Page 30, Lines 691-694)

p. 30 Lines 675-678: Can the authors propose a biomarker panel that could be used to select patients for Selinexor treatment? How does this potential therapy impact the patients included in each of the studies? LSCC has such high need for novel therapies that it may be able to proceed quickly to clinical investigation.

Response: We showed that XPO1 expression levels vary across subtypes with WGD, specifically between subtype 1 (all LUAD, characterized by lower XPO1 expression) and subtype 3 (predominantly LSCC, with higher XPO1 expression). Furthermore, we found that only LSCC organoids with WGD responds to Selinexor treatment. Therefore, as the reviewer stated, XPO1 inhibition could represent a candidate of novel therapy for LSCC with WGD. Possibly high XPO1 expression, measurable through IHC or RT-PCR, to serve as a biomarker for treatment response, but determining the cutoff predicting the clinical benefit requires further investigation to validate the relationship. We have updated the discussion section to reflect this perspective on XPO1.
(Page 31, Lines 710-716)

p. 32 line 741: Provide a reference or protocol for buffy coat enrichment of blood cells.

Response: We added the book reference, now listed as reference 68 in our manuscript. (Page 33, Line 784).

p. 38 line 942: Provide additional details for the peptide concentration with NanoDrop One.

Response: We described the additional details for the peptide concentration using NanoDrop One.
(Pages 39-40, Lines 986-989)

p. 39: TMT labeling sections with protocol steps should be written as full sentences.

Response: We understand the concern raised regarding the phrase "according to the manufacturer's protocol" (Page 40, Line 1003). This might have been interpreted as omitting the description of the experimental process. Following that sentence, all

experimental procedures conducted in accordance with the manufacturer's protocol are already described.
(Page 40, Lines 1001-1005)

p. 39 Line 968; Text is unclear; use of interrogate is incorrect.

Response: We apologize for the confusion, we have removed the phrase "which interrogate analysis".
(Page 40, Line 1014)

p. 40 and 41: Convert CAN to ACN.

Response: We appreciate the thorough review. We have fixed all the notations correctly.
(Page 42, Lines 1065-1066)

p. 43 and throughout the manuscript: Check spelling of Lehtio.

Response: We appreciate the thorough review. We have fixed all the spelling of Lehtio correctly.
(Page 45, Line 1153)

p. 53 line 1405: The evidence of neo-antigens in MS needs to be clarified. Do the authors mean detection of the mutant peptide containing the amino acid variant or just detection of any peptides from the protein of origin?

Response: In the process of inferring neoantigen candidates, our focus was primarily on mutant peptides derived from somatic mutations. Conversely, for the inference of cryptic MAPs, we considered any peptides that arose from 3'/5'-UTRs, pseudogenes, lncRNAs, and abnormal splicing events, specifically those expressed in tumor samples. We have detailed this methodology in our **Supplementary Fig. 6a**, which provides a comprehensive overview of the process for inferring both neoantigen candidates and cryptic MAPs.

Figure Legends: Full definitions of abbreviations should be available in the captions. Figure 1 describes the presence of different drivers. Were there enough patients to get clear signal for categorization of these tumors? In other words, do ALK fusions always belong to one of the clusters defined in the manuscript? I think this point bears more discussion either in the body of the manuscript or in a supplement.

Response: We agreed with the reviewer's suggestion and have now included a bar plot as **Supplementary Fig. 1d** to illustrate the distribution of driver mutations and subtype. In addition to the significant enrichment of EGFR mutations in subtypes 1 and 2, we also identified a significant enrichment of KRAS mutations in subtype 5. We also observed that the trends of enrichment of HER2 exon 20 insertion mutations in subtype 1, inclusions of ALK / ROS1 fusions exclusively in subtype 1 and 5, and RET fusion only in subtype 2,

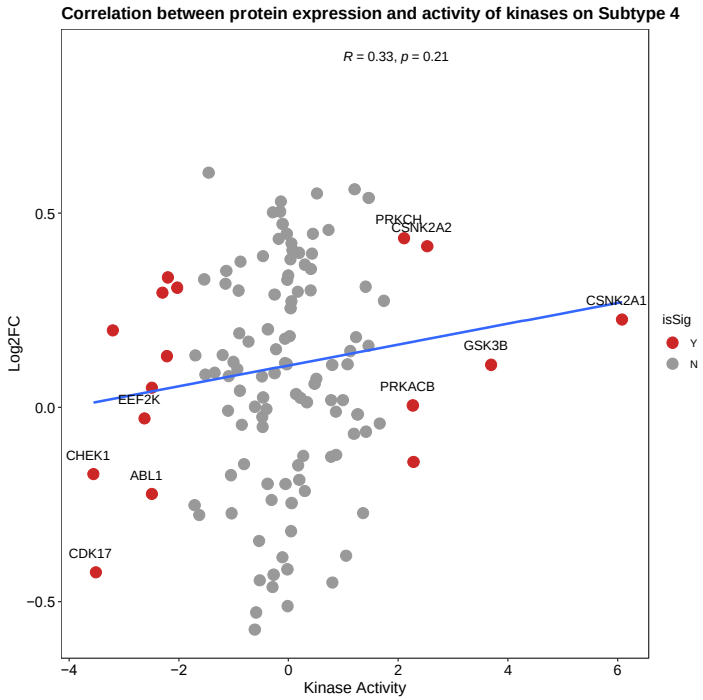
3, and 4. However, all these trends were not statistically significant. We included the comment about the subtype 5 enrichment of KRAS mutation in **Results** section.
(Page 8, Lines 141-145)

Figure 2b. The third panel (right) is unclear. The reclassification from the two other papers makes sense, but the combination and reclassification seems to find mostly the same groups. Better explanation is needed to for the few samples that are reclassified in the third panel.

Response: We apologize for any confusion regarding the interpretation. In the comparison between the two classifications in revised **Figure 2c**, our goal was to assess the influence of the Korean cohort dataset on the classification by conducting the classification without the Korean cohort dataset (left, named as "Gillette et al., 2020+Satpathy et al., 2021") and with the Korean cohort dataset (right, named as "Combined NMF"). The results show a 77% concordance rate, suggesting that although the Korean cohort constitutes half of the data (229 out of 462) but not being a critical driving factor in the final classification.
Approximately half of the new datasets were integrated and analyzed alongside the original dataset, potentially leading to variations in sample classification. Although it seems to be influenced by the Korean cohort dataset, we anticipate that it will not significantly affect the classification, given the 89% concordance level demonstrated in the aforementioned response.

Figure 2C: The significant kinases are assessed by their estimated activity from phosphoproteomics. Correlation of the kinase expression and estimated activity should be investigated.

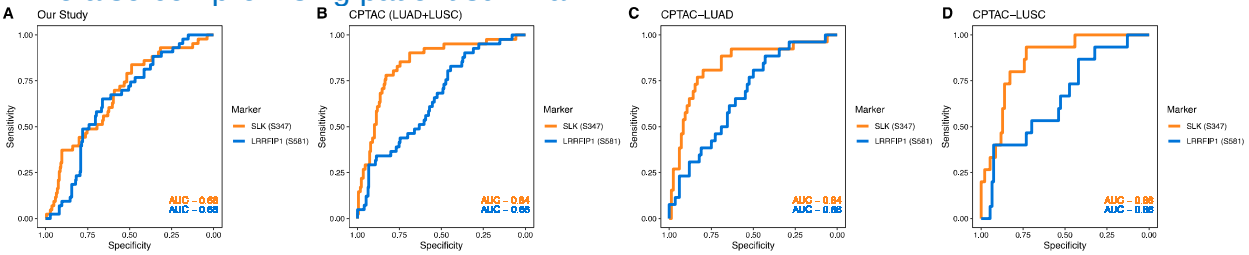
Response: Out of the 168 kinases predicted for activity, 121 have been observed through our protein data analysis. We conducted a comparison between the kinase activity scores and the expression levels (Log_2FC) of these 121 kinases as shown in the Reviewer-only figure 2. Focusing solely on the kinases that exhibited significant activity for the correlation calculation, we identified a moderate positive correlation of 0.33 (P =0.21).



Reviewer-only Figure 2

Figure 2D: The distributions of the prognostic markers are overlapping. Another metric, like Bhattacharya or Hellinger distance, is recommended to evaluate whether these metrics can have clinical utility. Alternatively, receiver operating characteristic (ROC) curves could be applied to the data.

Response: Thank you for your suggestion. As suggested, we examined the ROC curves for the SLK (S347) and LRRFIP1 (S581) markers in various datasets: Korean NSCLC (A), CPTAC LUAD+LSCC (B), CPTAC LUAD (C), and CPTAC LUSC (D) as shown in the Reviewer-only figure 3. In the Korean NSCLC dataset, the two markers exhibited AUC values of 0.68 (SLK (S347)) and 0.65 (LRRFIP1 (S581)), respectively. In contrast, in the CPTAC LUAD+ LSCC dataset, the AUC values were 0.84 for SLK (S347) and 0.66 for LRRFIP1 (S581). While our study dataset did not demonstrate high AUC values, possibly due to the limited sample size, the SLK (S347) marker in CPTAC data showed a performance of 0.84. This indicates that this marker is more effective not only in the Korean cohort but also in overall lung cancer. Furthermore, we wish to report that these markers clearly distinguish patients belonging to Subtype 4 from those in other subtypes, while also compromising patient survival.



Reviewer-only Figure 3

Figure 3B: Increased magnification or arrow indicators should be used to highlight the parts of the tumor that are most relevant. The details are lost at this size in the figure.

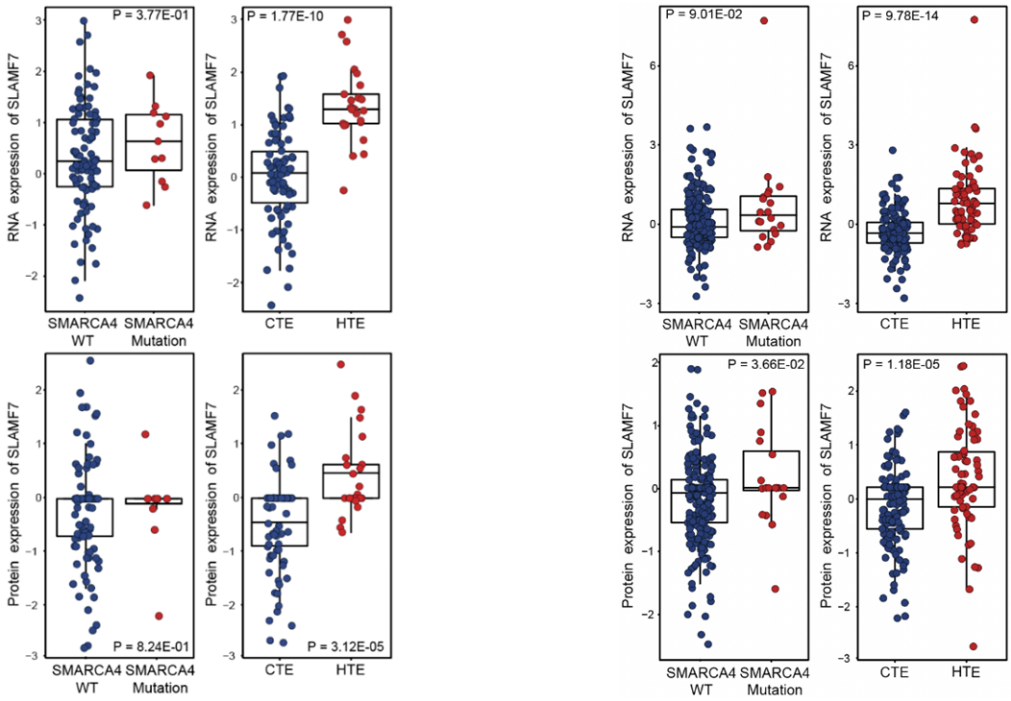
Response: For **Figure 3b**, we have increased the magnification and added the indicators for tumor and stromal components to highlight the findings shown in the manuscript. The legends for **Figure 3b** was also amended for proper explanation.

Figure 4I: State clearly that the kinases are inferred from the phosphoproteomics results in each instance in the manuscript.

Response: We corrected the description of the word as follows (Page 21, Line 459):
"Significantly upregulated kinases are highlighted with red triangles (FDR < 0.05) and mutations are shown in green boxes. Kinase activity scores are estimated from phosphoprotein expression."

Figure 5F: These results may be driven by the high outliers in the smaller population being compared.

Response: We concur with the reviewer's observation that our results had limited statistical power due to small sample size. To address this, we validated our findings by examining whether the trend we identified was also evident in another NSCLC multiomics cohort[3]. In this independent cohort, we consistently found a the positive correlation between SMARCA4 mutation and SLAMF7 expression, which similarly showed a positive correlation with the status, although this correlation did not reach statistical significance as provided in revised **Supplementary Fig. 5h**. Furthermore, we performed an analysis using an integrated dataset from both cohorts to increase statistical power. This approach yielded significant results in the correlation analysis between SLAMF7 protein expression and the presence of SMARCA4 mutation. These findings have been incorporated into the revised manuscript as follows (Page 24, Lines 536-537):
Among these immunomodulators, the SMARCA4 mutation was positively correlated with the expression of SLAMF7 at both the RNA and protein levels, and positively correlated with HTE status in both our study cohort and the independent cohort[3] (**Figure 5f**, **Supplementary Fig. 5h**, and **Supplementary Table 5k**).
<The analysis for the independent cohort> <The analysis for the integrated cohort>



Extended Data Figure 1: Include the definitions for abbreviations in the caption.

Response: We included definitions for all abbreviations used in the **Supplementary Fig. 1a**.
(Page 73, Line 3-7)

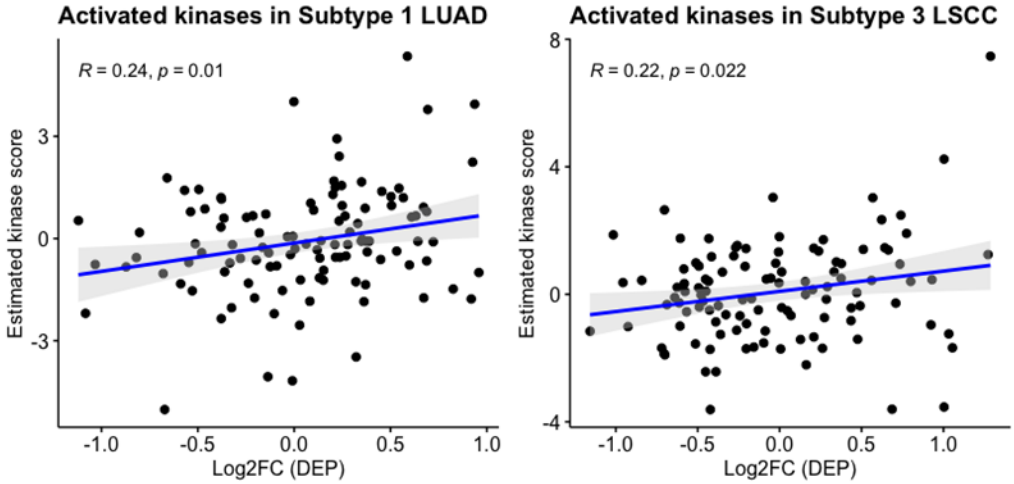
Extended Data Figure 2: Correct spelling of phospho- and indicate that kinase activity is inferred from phosphoproteomics.

Response: We have corrected spelling of “phosphor-“ to “phospho-“.
(Page 74, Line 18)
We have included **“derived from phospho proteome data”** for indicating that kinase activity is inferred from phosphoproteomics.
(Page 74, Line 19)

Extended Data Figure 4: Panels A and B should be flipped to start with the original classification and end with the WGD status. In panels H & I, correlation of the kinase expression and estimated activity should be investigated.

Response: We agree with the reviewer’s suggestion and have converted the order of the subtype and WGD status accordingly. Please refer to the revised **Supplementary Fig. 4a and 4b.**
Our analysis indicates that the estimated kinase scores and their corresponding protein expressions of the kinases in each subtype exhibit a significant positive correlation ($p < 0.05$), as shown in the Reviewer-only figure 4. Nonetheless, the strength of correlation is moderate. To provide a comprehensive view, we have displayed the expression levels of RNA, protein, and phosphoprotein for each kinase in **Figure 4i**. In addition, we have

added the Log2 fold-change values derived from differentially expressed protein analysis have been incorporated into **Supplementary Fig. 4h and 4i**. Except for RPS6KA3 in LUAD and EEF2K in LSCC, all kinases identified as significantly activated kinases were also significantly upregulated at the protein level.



Reviewer-only Figure 4

Reviewer #2 (Remarks to the Author): expertise in NSCLC genomics

Review Comments to Song et al. 2023

In this manuscript Song and coworkers has performed a proteogenomics analysis of a Korean cohort of 229 NSCLC cases. Multiple omics methods were used to collect the data including genomics, transcriptomics, proteomics, and PTM analysis (phosphorylation and acetylation). This comprehensive analysis contributes an additional resource for investigating the biology of NSCLC. I compliment the authors on their massive work! Further, the authors use their data to define 5 subtypes of NSCLC and characterize the subtypes in relation to clinical data and clinically relevant areas such as prognostication and cancer therapy prediction. Four of the identified subtypes overlap with previously described NSCLC subtypes, while one (subtype 4) is described as a novel subtype. Although the analysis is based on in-depth data and potentially outputs some new findings, several questions need to be addressed before considering publication in Nature Communications. Most importantly, the potential non-biology related factors that may drive the novel subtype 4 needs to be investigated (see below). Further, the analysis of PTM data is in my view underdeveloped and could be much improved, potentially resulting in new findings. In addition, the potential value of XPO1 as a therapeutic target in WGD LSCC needs further validation. Some analyses and interpretations are unclear and needs further clarification (see specific comments below). In summary, I do not recommend publishing the manuscript in Nature Communications in its current form.

General comments

• Some of the figures and fonts are very small and difficult to read. This should be adjusted to increase the readability of the manuscript.

Response: Thank you for the valuable comment. We have thoroughly updated all figures for better readability according to the Figure guideline of Nature Communications.

• Why is the LSCC samples so unbalanced between sexes (97% male)? How could this affect the interpretation of the analysis? Any LSCC finding may be confounded by gender. Test for independence if needed.

Response: Strong male predilection of the LSCC subgroup in the present cohort is one of the characteristics of the Korean cohort, which have been reported in several articles of Korean LSCC[4,5]. Although the degree of predilection is lower, male predilection in LSCC has also been well known in Western cohorts. Hence, we can stress that the cohort is not balanced but rather reflects the population characteristics of LSCC. The cause of this predilection in the Korean population is uncertain, but the smoking rate of Korean females might contribute to the phenomenon.
We briefly included the comment that the finding was also reported in other Korean LSCC cohorts.
(Page 5, Line 63-64)

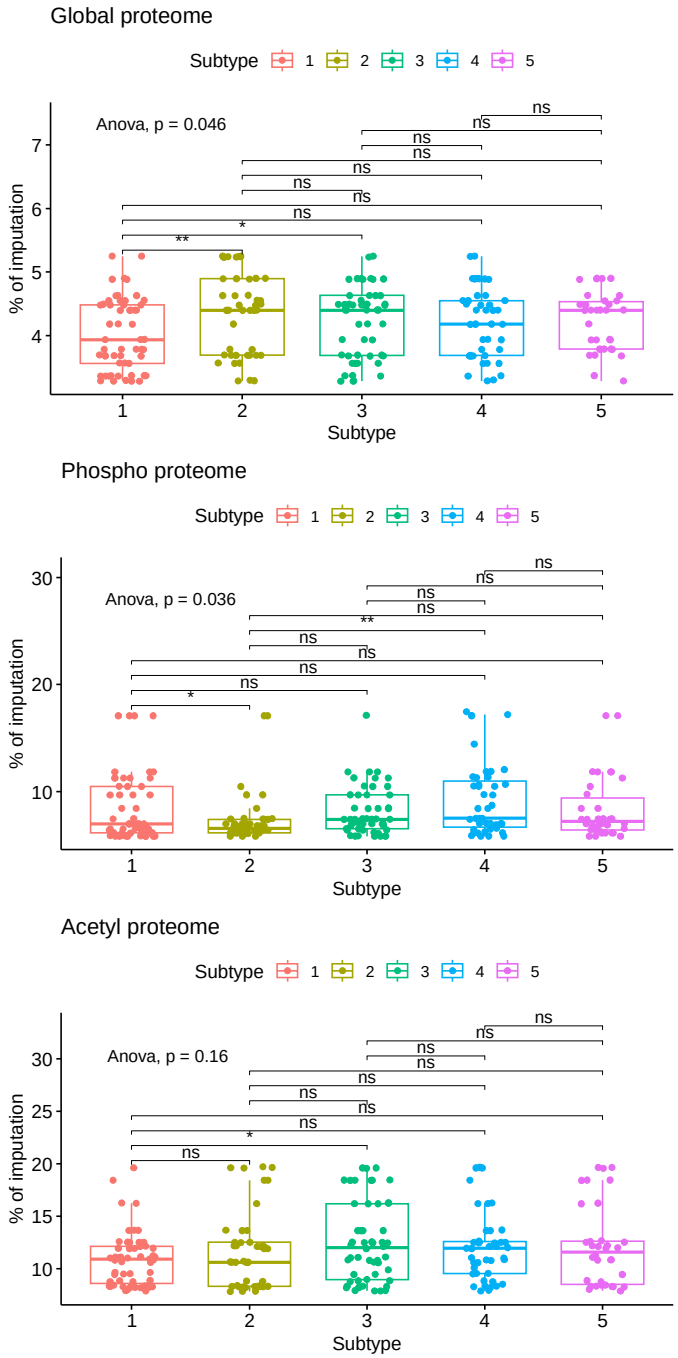
• It is unclear how imputation might affect the findings reported. A plot showing the percent imputation per sample, and an analysis to show if some subtypes have more imputed values would help. This would be relevant for total proteomics as well as phosphor and acetyl PTM data.

Response: Thanks for the comment. We agree with the potential errors that may arise when imputation might be skewed towards specific subtypes, as the reviewer mentioned. Therefore, we calculated the percentage of samples with imputation applied to our dataset (global proteome, phospho proteome, acetyl proteome) for all 229 individuals as shown in the Reviewer-only figure 5. For the global proteome, we observed that imputation was applied to less than 5% of all proteins analyzed in all samples, indicating a balanced distribution across subtypes. Similarly, we noted no skewed imputation across

subtypes in the PTM proteome, which includes phosphorylation and acetylation.

Reviewer-only Figure 5

Additionally, the Reviewer-only figure 6 presents the results of statistical analysis on the percentage of imputation by subtype, revealing some differences among certain subtypes (Subtype 1 vs. Subtype 2, Subtype 1 vs. Subtype 3 in the global proteome; Subtype 1 vs. Subtype 2, Subtype 2 vs. Subtype 4 in the phospho proteome; Subtype 1 vs. Subtype 3 in the acetyl proteome).



Reviewer-only Figure 6

• To my understanding, the PTM data (phosphor and acetylation) is not normalized to total protein levels. If so, differences in total protein levels will drive the analysis and interpretation of phosphor and acetylation data. As an example, Figure 4e and Extended Data Figure 4f and 4g are almost identical, which is most likely a result of differential protein expression, and not differential phosphorylation or acetylation. In a similar way, the kinase activity prediction analysis would be impacted by total protein levels, and therefore potentially biased. To really investigate the impact of PTMs on biology, which is

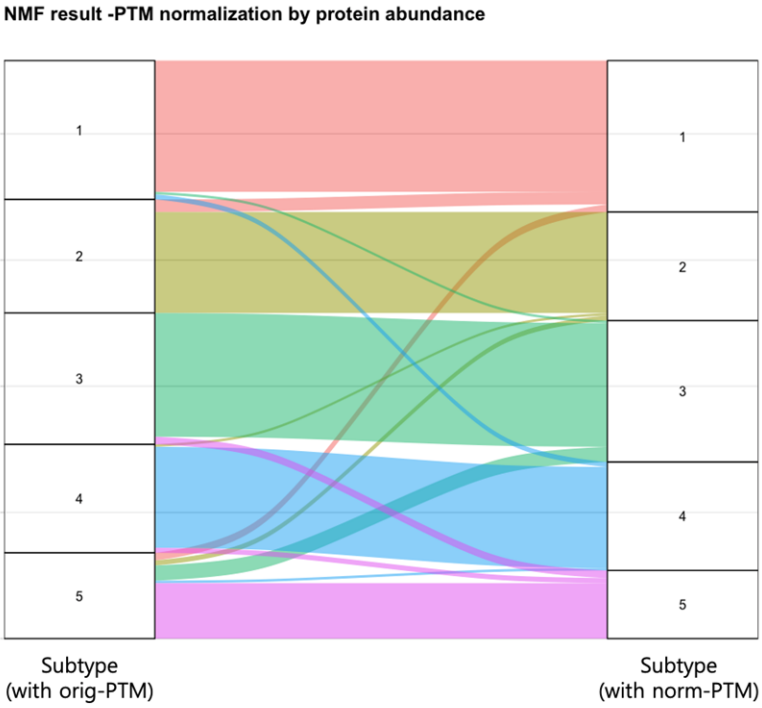
an important topic, normalization of the PTM data against total protein levels can output new information that cannot be found through total protein level analysis.

Response: We appreciate the valuable comment. As suggested by the reviewer, we normalized PTM data against protein abundance by employing the method published recently (Geffen, Yifat et al. Cell 2023). Then, we compared the original PTM data (denote 'orig-PTM') with this normalized PTM data(denote 'norm-PTM'). First of all, a portion of phosphorylation sites (n=9,027) and lysine acetylation sites (n=391) observed in this study were discarded due to a lack of corresponding protein abundance information. Then, we carried out NMF clustering and DE analysis. Reviewer-only figure 7 shows the result of the comparative NMF analysis of orig-PTM and norm-PTM. Most patients were found to be in the same cluster (i.e. 203/229), while 26 patients (~11%) changed their clusters. Comparative DE analysis showed that most DE PTMs were unchanged as shown in the Reviewer-only figure 8 (Pearson's r = ~0.97)

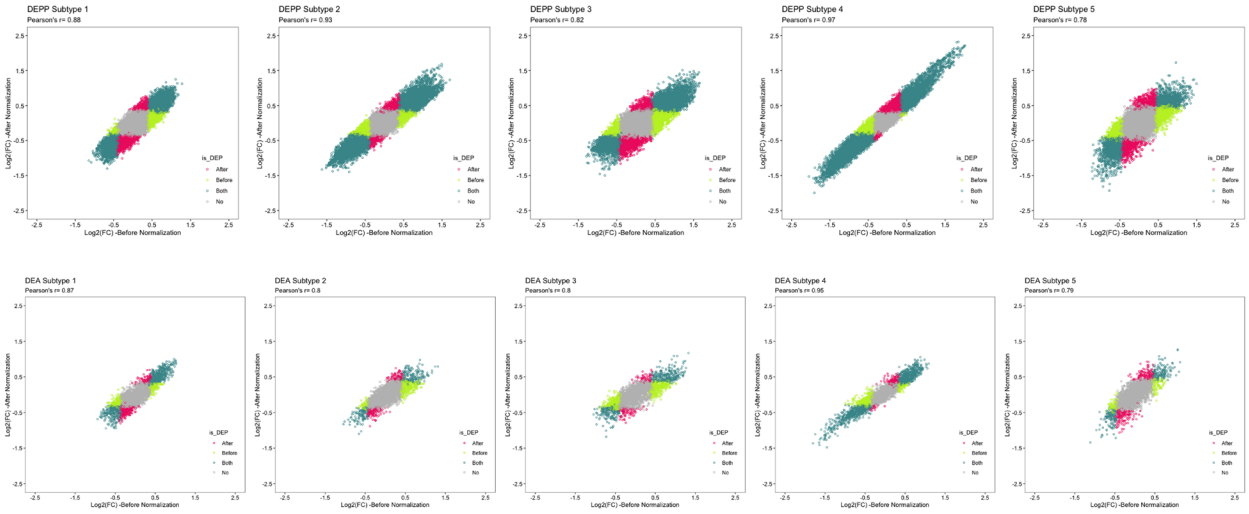
Further GSEA analysis and survival analysis showed also unchanged results. For example, the newly observed subtype 4 experienced the least alteration in its membership among all types. Furthermore, GSEA analysis revealed an enrichment of cell-cycle pathways in subtype 3, aligning with the presence of orig-PTMs (as shown in **Supplementary Figure 4f and 4g**). Despite subtype 5 exhibiting the most pronounced variances attributed to norm-PTMs, pathways linked to the immune system were still evident.

With our new analysis using norm-PTMs as suggested by the reviewer, we found that major results from orig-PTMs were consistent with those from norm-PTMs. Although considering protein abundance for PTM analysis may affect its results, this is not certainly the case in our result. We think that this may be because high correlation between orig-PTMs and norm-PTMs in both phosphoproteome (Pearson's r = 0.9) and acetylome (Pearson's r = 0.9). Taken together, we considered the reviewer's comments and

confirmed that PTM normalization did not contribute significantly to our findings.



Reviewer-only Figure 7



Reviewer-only Figure 8

Specific comments and questions

1. Subtype identification and characterization.
In general, enrichment analysis based on broad biology as in for example the mutSigDB Hallmarks genesets is only providing a very general and often not very specific output. To strengthen the characterization of the different subtypes it would be good to provide additional support based on the generated data (see below).

a. Subtype 1 is described as metabolic based on GSEA Hallmarks enrichment in Oxidative phosphorylation, mitochondrial matrix, and cellular respiration. Are there other supports for this interpretation? Do you see enrichment of STK11 mutations? Activation of the mTOR pathway?

Response: Among the 229 patients analyzed, only 12 exhibited STK11 mutations, and these mutations were not detected in subtype 1 (Subtype 2, n=2; Subtype 3, n=6; Subtype 4, n=2; Subtype 5, n=2). mTOR pathway is activated on patients who is Subtype 3 in protein and phospho data. Subtype 3 exhibits a pronounced activation of the cell cycle, indicating a close association with the mTOR signaling pathway. Consequently, it is likely that the mTOR pathway is upregulated in subtype 3.
The biologic mechanism of oxidative phosphorylation (OXPHOS) enrichment in subtype 1 is uncertain and needs further study. One hypothesis posits that this finding reflects the intrinsic characteristics of the terminal respiratory unit epithelium (alveolar pneumocytes) situated proximally to the tumor cells in Subtype 1. The first hypothesis is supported by the recent study showing the crucial role of mitochondrial electron transporting cascade and the mitochondrial dysfunction-associated stress response in proper development of alveolar pneumocytes (Han SH et al., Nature 2023). However, this theory does not elucidate why OXPHOS enrichment was not enriched in Subtype 2, despite its morphological similarities to normal alveolar epithelium. An alternative hypothesis suggests a connection with the EGFR signaling pathway, as partially evidenced in studies on intestinal stem cells (Zhang C et al., Curr Biol 2022). However, the hypothesis could be refuted by other experiment showing that the oxidative phosphorylation is reactivated upon the inhibition of EGFR signaling in non-small cell lung cancer (Rosa VD et al., Clin Cancer Res 2015).

b. Subtype 2 is described as early tumorigenesis and early-stage disease based on enrichment in IL-33 and Notch pathway. Other supports for this interpretation? Was TNM stage significantly different between Subtype 1 and 2?

Response: Assuming TNM stage I and II as early, 58% of subtype 1 patients are early, while 76% of subtype 2 patients are early. The results of the chi-square test indicate there is no significant difference (P=0.1069)

	Subtype 1	Subtype 2
Early (I, II)	32	34
Late (III, IV)	23	11

c. Subtype 3 is characterized by male gender, LSCC, smokers, TP53mutations and WGD events. Are these findings independent in multi variate analysis? It looks like the subtype core is almost exclusively males with LSCC. In other words, Subtype 3 seems to be driven by histology and should perhaps be annotated as LSCC rather than proliferative.

Response: As previously highlighted, the marked tendency for male smokers to fall within the LSCC subgroup is a distinctive feature of the Korean cohort. This characteristic has made it challenging to isolate the effects of male gender from those associated with LSCC in Subtype 3. However, we still think of Subtype 3 as proliferative since there are samples

with LUAD in Subtype 3—comprising 10 out of 52 samples in our cohort and 22 out of 64 samples in the CPTAC cohort. Additionally, it is noteworthy that within the LUAD samples, the incidence of WGD is enriched in Subtype 3 within LUAD samples. Please refer to the responses provided in your comment 4-b below for further details.

d. For subtype 4, enrichment was found for hypoxia. In the text (row 125-126 in manuscript) it is written that this enrichment was found in the phospho data. In the corresponding figure (Figure 1e) it is written "protein". Which data was used for the enrichment analysis? The same question for Neutrophil degranulation (text: phospho, figure: protein).

Response: We apologize for any confusion caused by discrepancies between the text and the figures presented. Hypoxia appears significantly as a of Subtype 4, not only in the context of proteins (adjusted $P < 0.05$, Benjamini-Hochberg adjustment) but also in the phosphorylation data (adjusted $P = 2.3 \times 10^{-}4$). Additionally, Neutrophil degranulation is also notably present as a significant characteristic of subtype 4 in both protein (adjusted $P = 4.6 \times 10^{-}6$) and phospho (adjusted $P = 1.7 \times 10^{-4}$) data (**Supplementary Table 1c**).

e. Subtype 4 is described as a "chromosomally stable tumor suppressor-driven mesenchymal subtype" based on the enrichment analysis. It is difficult to follow how this interpretation was done based on the analysis. What is meant by "tumor suppressor driven"? Is there additional support for this interpretation?

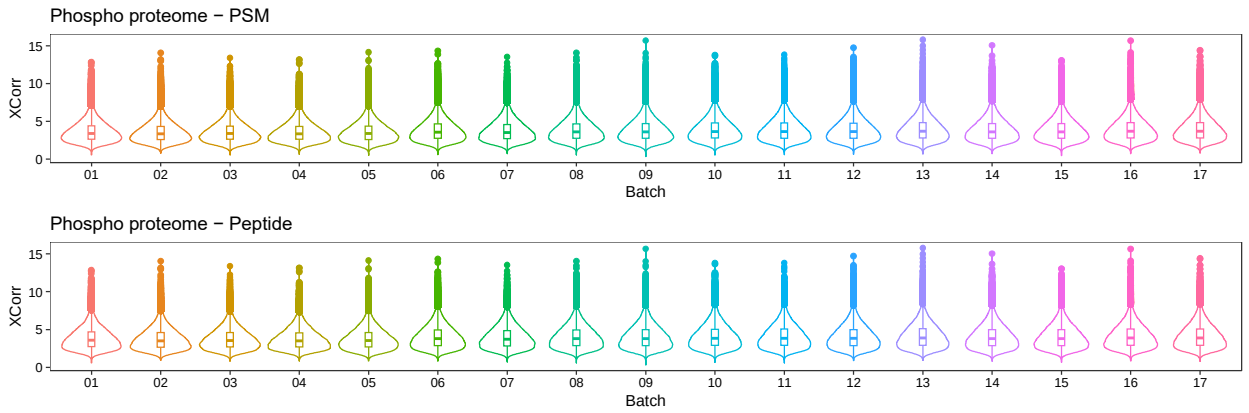
Response: As the reviewer pointed out, the rate of tumor suppressor gene alteration in subtype 4 is not prominent compared to the other subtypes. Also, we accept that the term "tumor suppressor-driven" is totally misleading. We have changed it as "**tumor suppressor-deficient**" for Subtype 3 (Page 7, Line 127), and for Subtype 4 we removed the term (Page 8, Line 135-136).

f. Was enrichment analysis performed for individual oncogenes and tumor suppressors defined here as "other" in a similar way that was performed for EGFR and TP53? Were there any specific enriched oncogenes and tumor suppressors?

Response: We missed describing the finding of significant enrichment of CDKN2A copy number loss in subtype 1, PIK3CA mutation in Subtype 3, and KRAS mutation in subtype 5. We have added the description about the findings in **Result** section (Page 7, Lines 101-102; Page 8, lines 141-145). The frequencies of oncogenic alterations are presented in **Supplementary Fig. 1d**. In contrast, any significantly enriched mutation within the other subtype has not been identified.

2. Characterization of Subtype 4 described as a novel NSCLC subtype.
a. Subtype 4 is almost exclusively driven by the phospho-proteomics data (96% of NMF features, row 183). A concern here is that this subtype could be driven by the quality of the generated phospho-data.

Response: Xcorr is a score function used in shotgun proteomics liquid chromatography tandem mass spectrometry (LC-MS/MS) experiments to obtain accurate peptide identifications. Higher scores indicate better quality mass spectrometry data. By examining the distribution of Xcorr in the results (peptide-spectrum Match (PSM) and peptide) of the produced phospho-proteome, we can observe that it represents fairly high-quality data between 17 batches as shown in the Reviewer-only figure 9. Since there are 5 subtypes evenly distributed across the 17 batches (**Supplementary Fig. 7**), the quality of the PSMs and peptides produced from the phospho-proteomics data is not expected to differ between subtypes.



Reviewer-only Figure 9

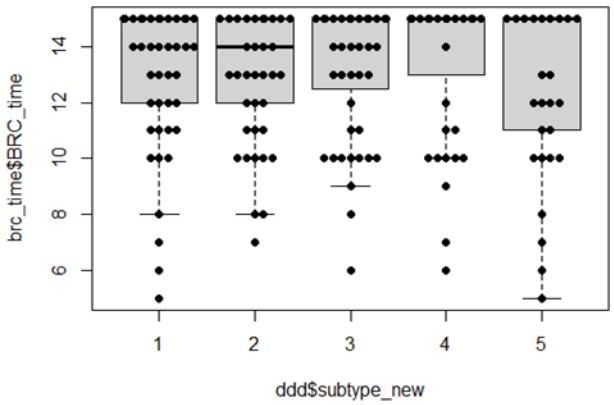
b. Were there significant differences in the TMT channel quant distribution of the phospho-proteomics data between the samples/subtypes? Boxplots showing the phospho quant distribution in each sample across the cohort would help evaluating this.
c. Were there differences in imputation of phospho-data between subtypes?

Response: As recommended by the reviewer, we have included boxplots in **Supplementary Fig. 7** to show the distribution of phospho quantification in each sample across the cohort. We also examined the quantification of the global proteome across the cohort and the proportion of subtypes across 17 batches and TMT channels excluding the reference channel. The quantification patterns of the phosphosite and global proteome were found to be similar, and the subtypes were evenly distributed across batches and TMT channels. There were no differences in the imputation of phospho-data between subtypes.

d. Was meta-data collected for the sample-collection and sample prep pipeline? The integrity of phosphorylations is very dependent on sample handling, and differences can impact the results of the analysis.

Response: We already described in the manuscript that we selected the patients available with tumor and NAT samples, which were frozen within 15 minutes. The elapsed time from tissue sampling and sample freezing was not statistically different among the subtypes (Kruskal-Wallis test P = 0.357) as shown in the Reviewer-only figure 10. Therefore, it can be stressed that there is no significant difference in sample quality

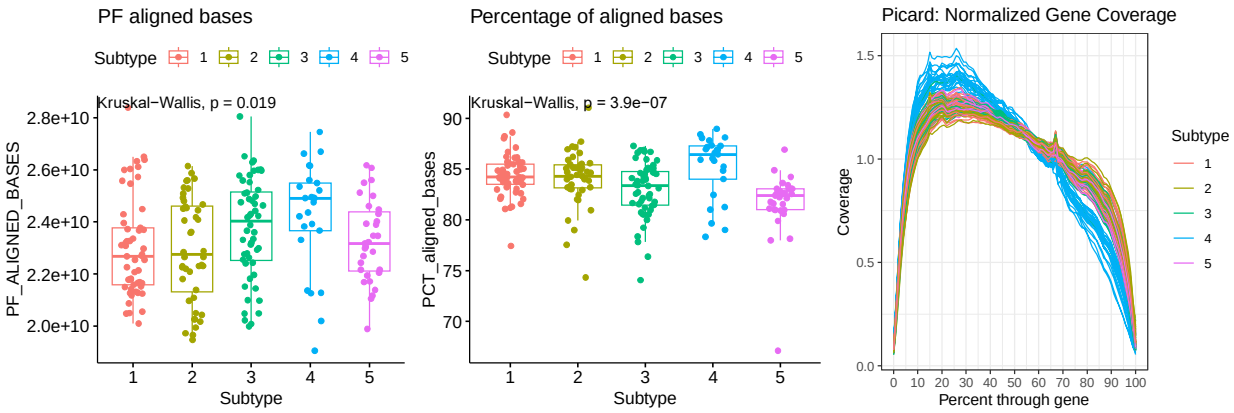
between the subtypes. We also checked the distribution of our phosphorproteome and found that it is not sample-specific result as shown in the **Supplementary Fig. 7a** (please refer to the responses provided in your 2-d comment above for further details).



Reviewer-only Figure 10

e. For subtype 4, 18/43 samples (42%) were missing transcriptomics data. For the other subtypes transcriptomics data were only missing in very few samples (subtype 1: 2/55 samples, subtype 2: 1/45, subtype 3: 0/52, subtype 5: 3/34). Please explain this difference. If the quality of the sample was too poor for transcriptomics analysis in a large part of the subtype 4 samples, this could indicate that also the proteome/PTM analysis was affected by sample quality. Was RNA quality used for selection of samples for transcriptomics? What is the distribution of RNA quality metrics across the full cohort? Is there a statistical difference in RNA quality between the subtypes?

Response: In the RNA library preparation, a total of 24 samples did not pass QC. This was due to two reasons: some samples did not have an adequate amount of RNA input for library preparation, while for others, we were unable to obtain the RNA integrity number. Many of these samples belonged to Subtype 4 (18/24). After excluding these samples, we performed RNA-seq library preparation using the Illumina kit ("Ribo-Zero H/M/R_Gold"), which is capable of generating reliable RNA-seq data from low-quality RNA. To ensure that the included samples did not bias towards any particular subtype, we measured various quality metrics for our RNA-seq data (Reviewer-only Figure 11). The 'PF_aligned_bases' metric showed a uniform distribution of the total number of aligned PF bases across sample subtypes. Similarly, the percentage of aligned bases demonstrated the uniformity of RNA-seq alignment results across subtypes. Furthermore, the 'normalized gene coverage' in Picard QC metrics did not show significant differences among subtypes, although some bias was observed in subtype 4. To assess the impact of the samples dropped from the transcriptome analysis on the PTM analysis, we examined their distribution in the clusters. As shown in **Supplementary Figure 7**, the excluded samples did not differ independently in the clusters, indicating that their exclusion had no significant impact on the proteome/PTM analysis.



Reviewer-only Figure 11

f. Was the quantitative phospho-data normalized to total protein levels before the differential analysis was performed? If not, the phospho-analysis (and any other PTM analysis such as acetylation) will be driven largely by the total protein levels.

Response: Please refer to the responses provided in your general comment above for further details.

g. How many of the Subtype 4 features were substrates of CSNK2A1?

Response: The NMF features of Subtype 4 consist of 3 proteins, 178 phosphosites, and 5 acetylsites. Among these 178 phosphosites, there are 11 substrates significantly associated with CSNK2A1.

h. Was the total protein level of SLK or LRRFIP1 prognostic in the current cohort and in the CPTAC cohort?

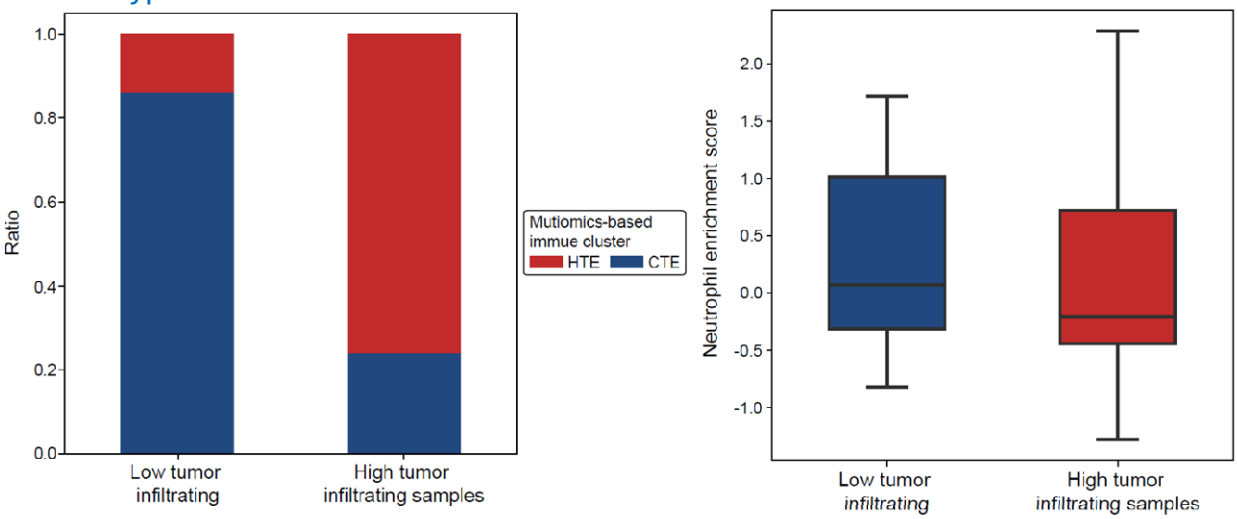
Response: Thanks for your suggestion. The protein levels of SLK and LRRFPI1 show no significant difference (SLK, p = 0.44; LRRFP1, p = 0.41) in our cohort. In the CPTAC cohort, we found the same pattern (SLK, p = 0.77; LRRFIP1, p = 0.36). This highlights the importance of phosphorylation for prognostic features.

3. Cellular landscape of the five subtypes

a. Subtype 5 is largely driven by immune infiltration. Yet about 25% of the Subtype 5 samples has low tumor infiltrating immune cell components (Figure 3k). What other features could explain the clustering of these non-infiltrated samples into Subtype 5?

Response: The immune cell infiltration was examined only by light microscopy examination of the tumor sections, which may result in an underestimation of immune cell infiltration on the examination. Furthermore, there is no light microscopy-based study perfectly predicting the response of immunotherapy (Adegoke NA et al., J Immunothera Cancer 2023). Possibly, these discrepancies suggest that the multiomics-based evaluation of tumor immune responses could be more sensitive compared to the light microscopy-based immune microenvironment subtyping. We confirmed the discrepancy

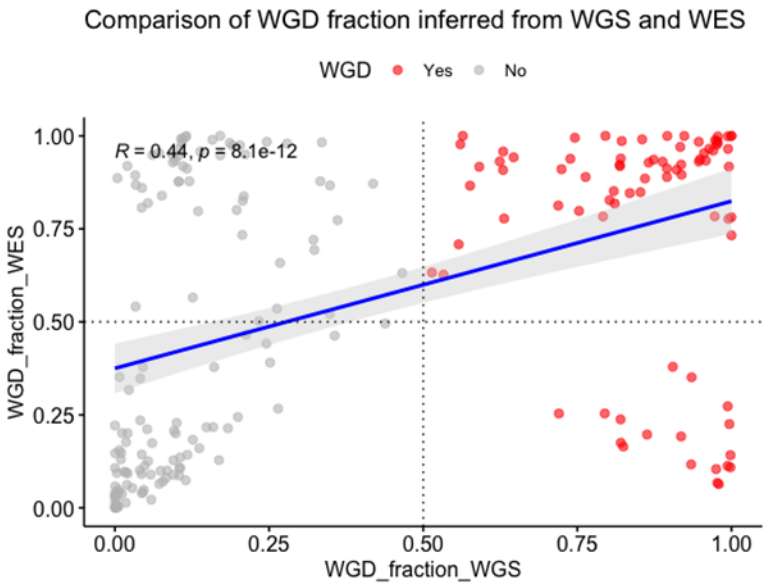
that a subset of samples with low levels of tumor-infiltrating immune cells were hot tumor-enriched tumors (HTEs) according to multiomics-based immune clustering as described in the Reviewer-only figure 12. Moreover, these samples showed higher enrichment of neutrophils, which were more prevalent in Subtype 5, potentially driving their classification into Subtype 5.



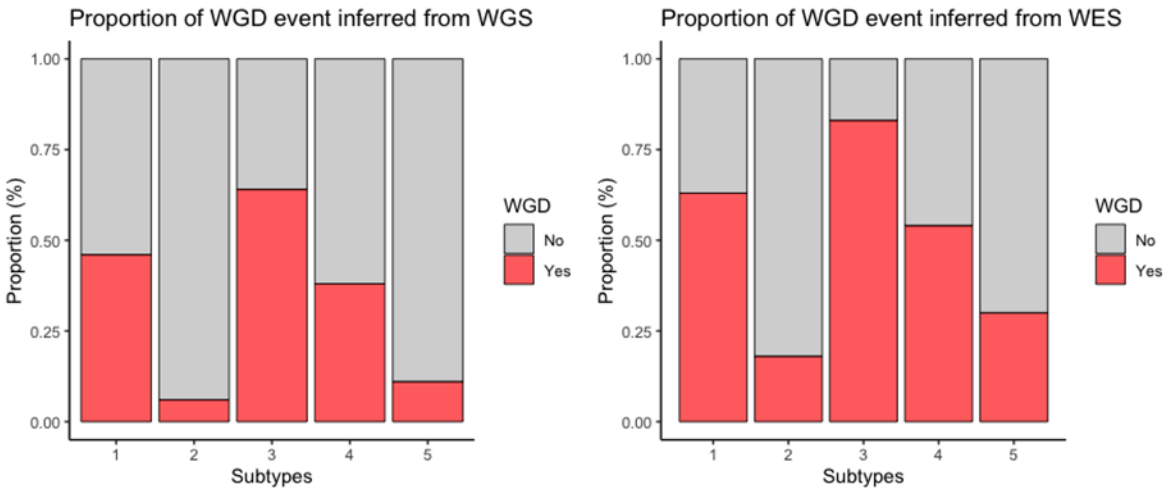
Reviewer-only Figure 12

4. Proteogenomic features underlying whole genome doubling in NSCLC.
a. One of the main findings is related to whole genome doubling (WGD, more than half of the chromosomes are gained) which is found specifically enriched in subtype 3 and subtype 1. The WGD analysis was based on whole exome sequencing and not whole-genome sequencing, which according to the authors may be a limitation in the study. A validation for a subset of cases using whole-genome sequencing could strengthen this finding.

Response: Thank you for your comment. We wish to highlight that the WGD status of the CPTAC samples was determined by whole genome sequencing (WGS) data. In our analysis, particularly reflected in **Figure 4**, we noted a substantial concordance in the rate of WGD within Subtype 3 when comparing our cohort with the CPTAC cohort. Furthermore, we compared the WGD status inferred from both WGS and whole-exome sequencing (WES) across all samples from the CPTAC cohort. The results, as illustrated in the Reviewer-only figure 13, reveal a significant correlation between the WGD fraction inferred by WGS and WES ($R = 0.44$, $p = 8.1 \times 10^{-12}$). While inferring WGD status using WES resulted in more cases of WGD-positive compared to WGS, the proportion of WGD samples still showed enrichment in Subtype 1 and 3 (Reviewer-only figure 14). Thus, although our study acknowledges the limitation of inferring WGD with WES, we believe this does not significantly impact the robustness of our results.



Reviewer-only Figure 13

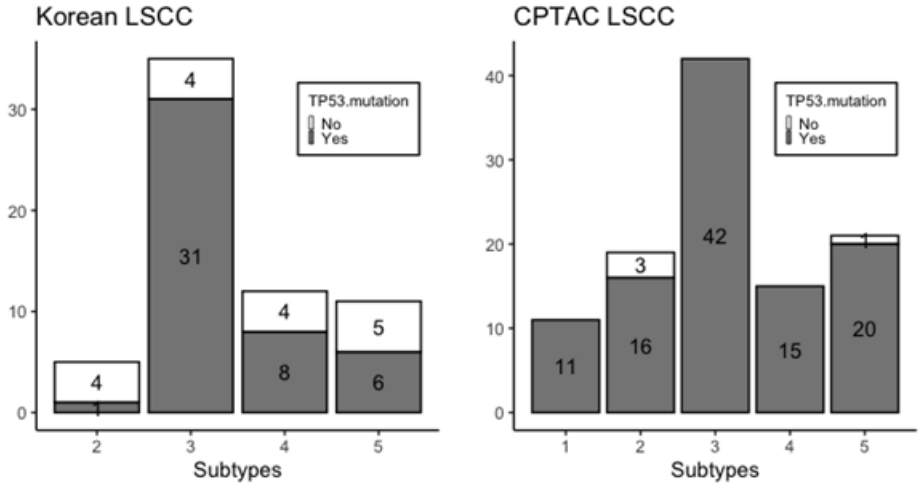


Reviewer-only Figure 14

b. Due to the high overlap between subtype 3, LSCC and genome doublings (WGD), specifically in the subtype 3 core, it is important to investigate the various subtype 3 findings in relation to histology. It is difficult to evaluate what is driving the findings without such analysis. The same is true for the validation of findings against CPTAC data where the authors have combined LSCC and LUAD histologies. Specifically:
i. Is TP53 more commonly mutated in LSCC cases from subtype 3 than in LSCC cases from other subtypes?

Response: In the CPTAC cohort, nearly all samples of LSCC exhibited TP53 mutations, making comparisons challenging. However, in our cohort, as depicted in the Reviewer-only figure 15, we observed a higher proportion of TP53 mutations in Subtype 3 within LSCC. While TP53 mutations alone may have been difficult to reproduce in the CPTAC

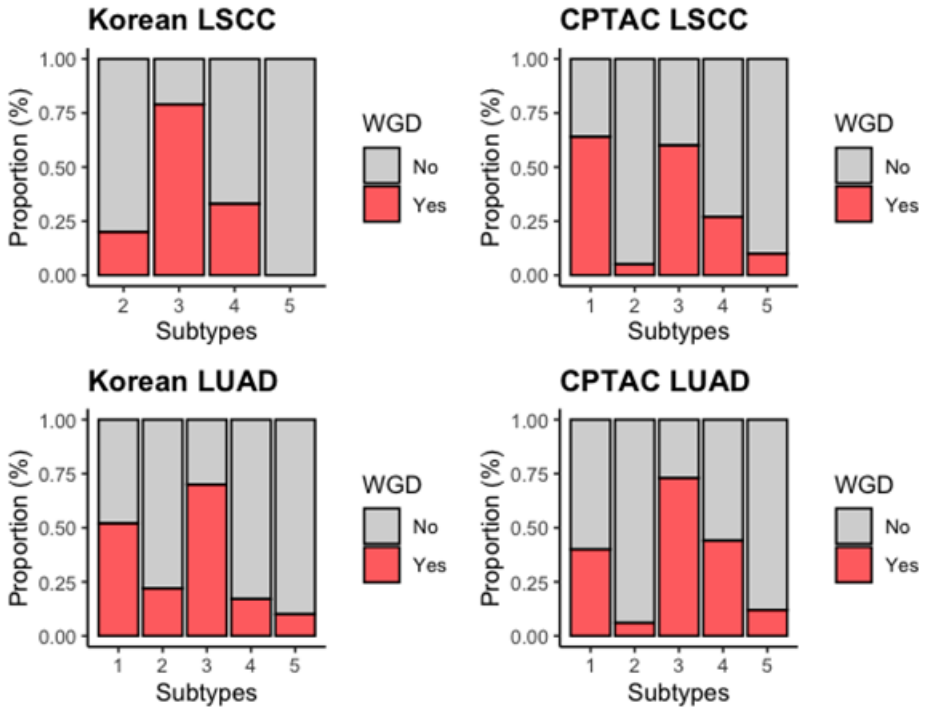
dataset. We observed that the occurrence rate of SOX2 amplification and TP53 co-mutation is enriched in Subtype 3, as shown in **Figure 4d**.



Reviewer-only Figure 15

ii. Are WGD more common in LSCC/LUAD cases from subtype 3 than in LSCC/LUAD cases from other subtypes?

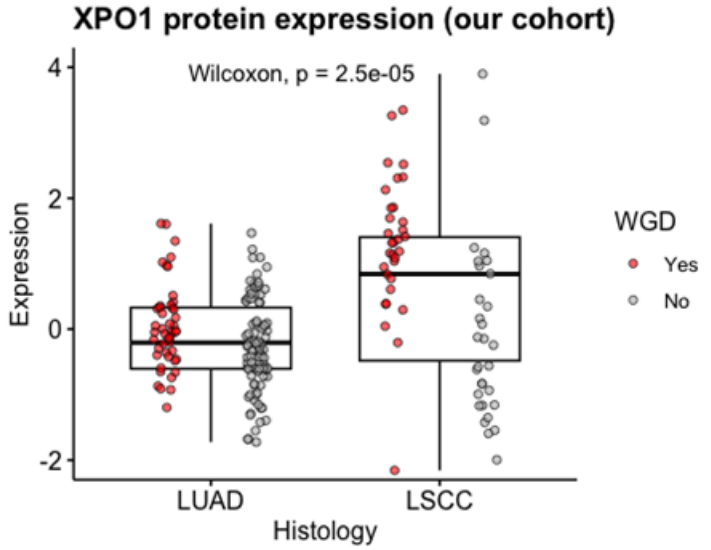
Response: Yes, when examining each LSCC and LUAD individually, we observed the high frequencies of WGD samples in Subtype 3 as well as in Subtype 1, which are two major subtypes characterized by WGD as shown in the Reviewer-only figure 16.



Reviewer-only Figure 16

iii. Is XPO1 expression higher in LSCC than in LUAD in general?

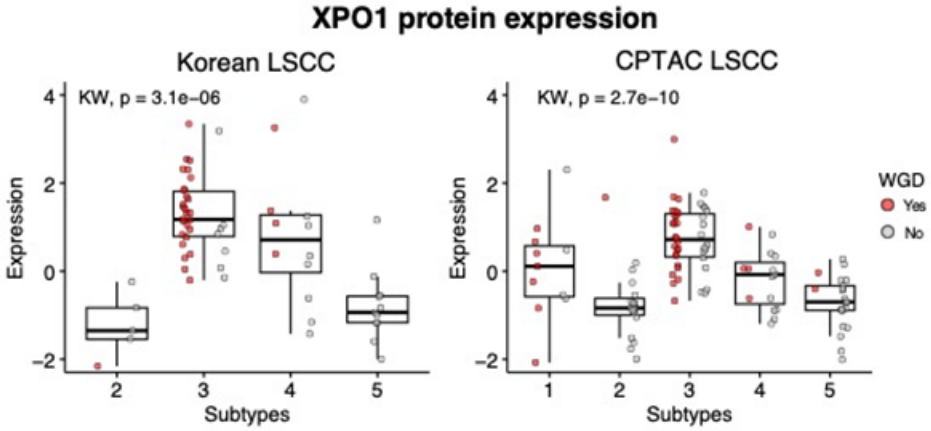
Response: In our cohort, we observed generally higher XPO1 expression levels in LSCC compared to LUAD, as shown in the Reviewer-only figure 17. We could not confirm this finding with the CPTAC LUAD cohort due to the lack of available protein expression data for XPO1.



Reviewer-only Figure 17

iv. Is XPO1 expression higher in in LSCC cases from subtype 3 than in LSCC cases from other subtypes?

Response: When examining only LSCC, we could still observe that the expression level of XPO1 in Subtype 3 was higher compared to other subtypes, as depicted in the Reviewer-only figure 18. This finding was consistent across both our cohort and the CPTAC cohort.



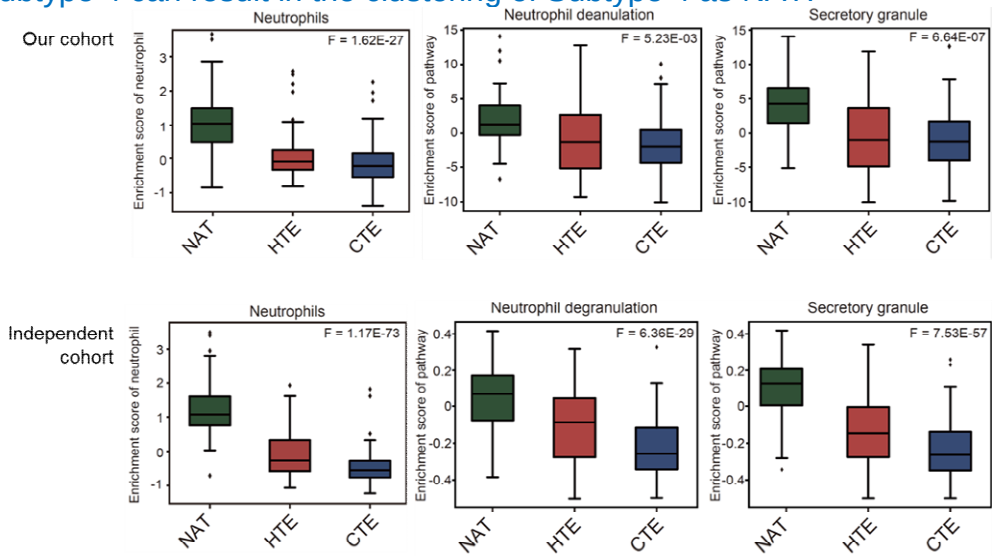
Reviewer-only Figure 18

c. The XPO1 inhibitor screen was not performed in LSCC cases without WGD, and therefore it is impossible to judge if the sensitivity depends on histology or WGD. For assessing this, the experiment should be complemented with several LSCC organoids without WGD.

Response: Unfortunately, we could not generate the LSCC organoid without WGD; therefore, direct comparison of XPO1 inhibitor efficacy according to WGD in the LSCC organoid is not feasible. However, we showed that XPO1 protein expression was significantly elevated according to WGD status in Subtype 3 NSCLC, and we identified that the XPO1 inhibitor consistently inhibited cell viability in three separated LSCC WGD organoids. Therefore, the XPO1 inhibitor may be clinically beneficial for LSCC with WGD. Nevertheless, we agree with you that further validation studies are required to characterize the benefits.
The comment that the finding requires validation is briefly added in **Results** section. (Page 19, Line 430)

5. Immune landscape in NSCLC
a. Due to lack of transcriptomics data, relatively few subtype 4 samples were included in the immune cluster analysis. Further, 17/25 (68%) of Subtype 4 samples cluster with NAT (normal adjacent tissue). What drive the clustering of NAT, and why are the subtype 4 samples overrepresented in this cluster?

Response: We agree with the reviewer in noting that Subtype 4 was overrepresented in NAT samples. It is important to acknowledge that both multiomics-based and immune-based clustering methods are not perfect in accurately distinguishing between patients and NAT consistently across previous NSCLC multiomics studies[3,6]. As the reviewer pointed out, we observed that specific cell compositions and pathway activities were significantly higher in the NAT group compared to both HTE and CTE tumors. These included neutrophils, the secretory granule pathway, and neutrophil degranulation, as illustrated in the Reviewer-only figure 19. Interestingly, we also observed the activation of neutrophils and these pathways in Subtype 4, as depicted in **Figure 3a** and **Figure 1a**. These patterns were consistent across an independent cohort comprised of two NSCLC multiomics cohorts[3,6]. Therefore, we propose that this enrichment in both NAT and Subtype 4 can result in the clustering of Subtype 4 as NAT.



Reviewer-only Figure 19

6. Multiomics profiling of neoantigens and immune clusters
a. To “confirm” a neoantigen candidate by proteomics support, a rule was set so that at least 20% of the total intensity of all reporter ions for the neoantigen comes from the sample with the corresponding somatic mutation (row 1301-1302). It would be important to also assess if the sample with the mutation has the overall highest intensity in the set. In principle, the mutated sample should have a clear outlier pattern in the quantification, and all non-mutated samples should have “background” values.

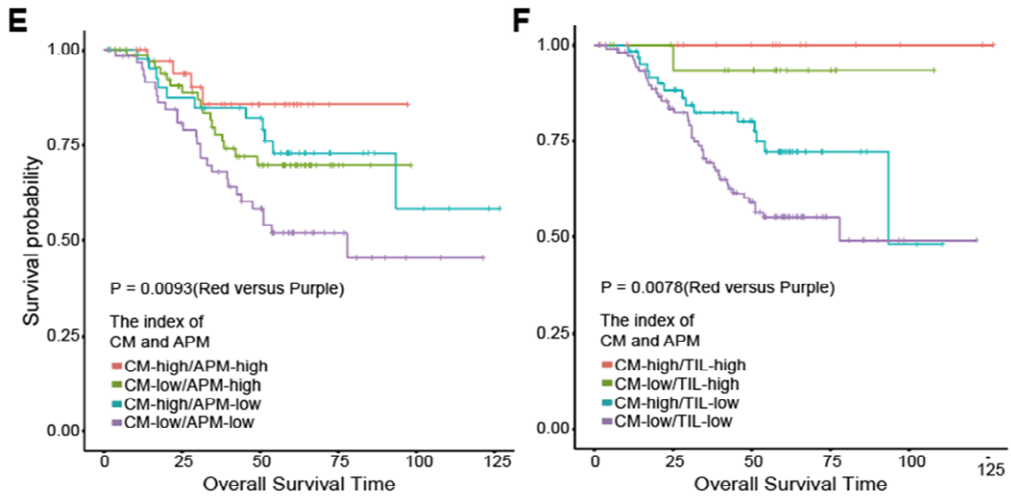
Response: As recommended by the reviewer, we also assessed whether the sample with the mutation had the highest intensity in the set. Of the mutations previously selected, only one did not meet the criterion. **Figure 6a and Supplementary Fig. 6c** have been updated using the new list.

b. Were the cryptic peptides confirmed using MS data?

Response: The cryptic peptides were confirmed by MS data as described in the "**Identification of novel peptides**" section.

c. Cryptic MAPs were positively correlated with immune infiltration which was associated with better survival (Figure 5c). Were the cryptic MAPs prognostic independently of immune infiltration?

Response: We thank the reviewer for pointing this out. In revision, we tested whether cryptic MAPs have prognostic value regardless of the status of immune infiltration. As a result, the high load of cryptic MAPs was associated with better survival in not only high but also low immune infiltration status, implying independence as a prognostic marker of cryptic MAPs from immune infiltration. This result, shown below, is provided in revised **Supplementary Fig. 6f**.



Reviewer #3 (Remarks to the Author): expertise in computational muti-omics analysis

Song et al. present a thorough and comprehensive multi-omics analysis of a large Korean NSCLC cohort leading to the identification of 5 subtypes by using non-negative matrix factorization (NMF) clustering. Their approach to look at all NSCLC histologies contrasts with other large studies that have tended to focus separately on LUAD or LUSC histological subtypes. Their approach addresses the considerable overlap that can exist between the two major histological subtypes and more minor histologies that are less well characterized.

They analyzed 229 NSCLC patient tumors and a "replication" cohort of 462 patients from published studies. Their data includes a large-scale single-cell RNAseq dataset. Also, they did a histological review of patient data related to TILs, identified potential neoantigens in the tumor microenvironment and observed varying efficacy of adjuvant therapies between subtypes. Their data include genomic (WES), transcriptomic, proteomic, phosphoproteomic, and acetylprotoemic datasets. Depth of coverage is excellent. For example, in their proteomics studies they list more than 10K proteins, 40K phosphoproteins, and approx. 6K acetyl proteins in at least 30% of samples. Relative to other large-scale multi-omics/proteogenomics studies of lung cancer, this study matches if not surpasses them for its comprehensive generation and integration of data with clinical features and outcomes. Their comparative analysis is generally in agreement with and extends the insights described in earlier reports from TCGA and CPTAC studies. They note that worse outcome associated with their subtype 4 does not extend to the CPTAC cohort and thus may reflect an ethnicity impact.

As is the nature of such large-scale analyses, there is considerable conjecture on the role of differentially expressed genes and proteins and protein features. They describe evidence for selinexor sensitivity in patient-derived organoids associated with whole genome duplications (Subgroup 3). While a preliminary result, it represents an experimental test of an emerging hypothesis. Overall, the vast amounts of data and data analysis are convincing and of high quality. The paper is well written and conclusions clearly presented.

Comments:
1. The figures are extremely difficult to read due to small fonts. In Fig 2a the symbols (dot, faint rectangle, blank) are not defined.

Response: Thanks for your comment. We have included the definition of these symbols. (Page 13, Lines 250-254)

2. The results associated with Fig 2 are confusing as written.

Response: We rewrite results associated with **Figure 2** for avoiding any confusion (Page 10, Lines 172-187).

3. Line 170 reads: "four subtypes" ...should this be five subtypes?

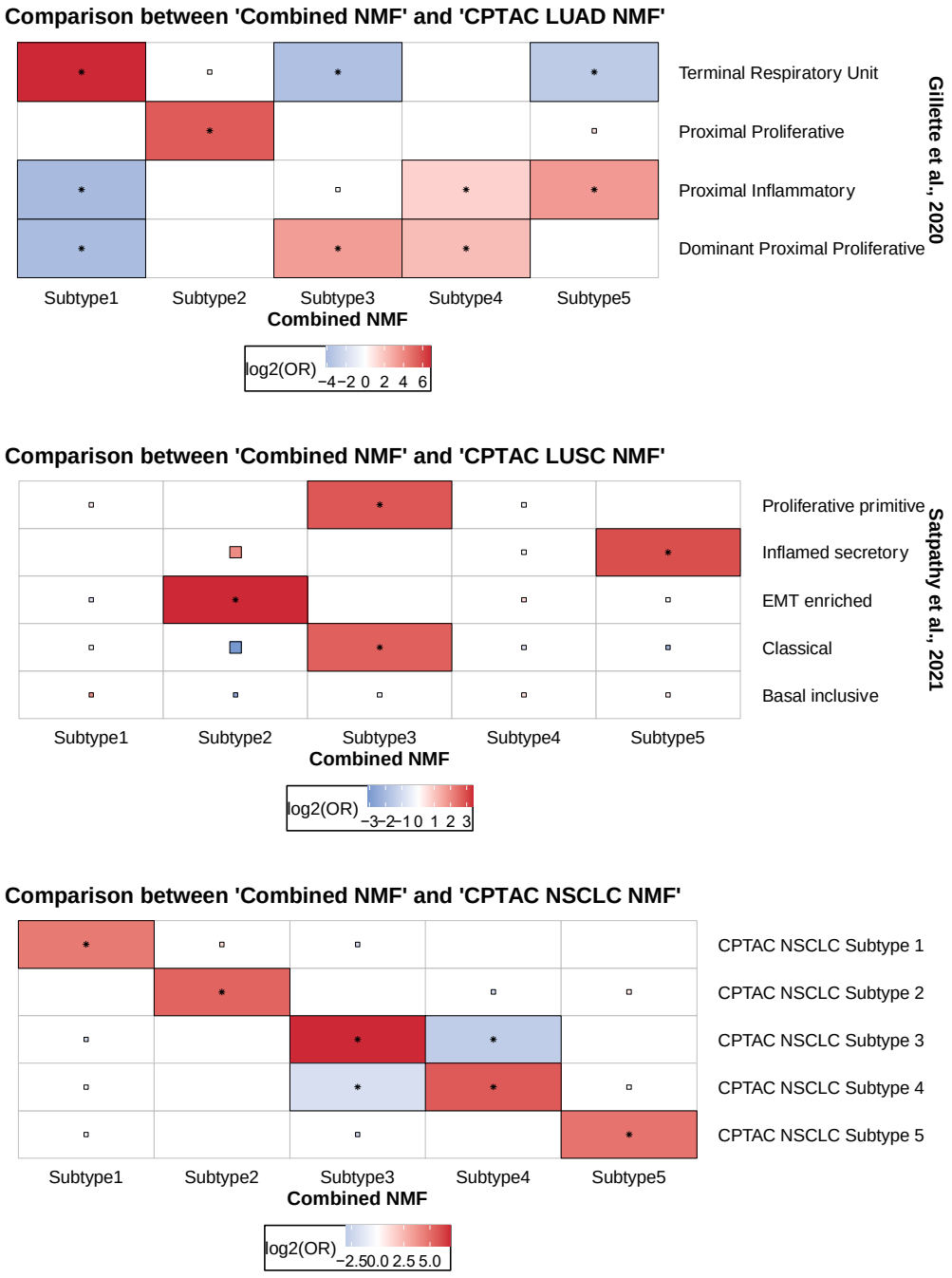
Response: We apologize for the confusion, but here we mean four. By 'four,' we meant to indicate that Subtypes 1, 2, 3, and 5 exhibit a close association with the non-small cell lung cancer NMF subtypes from previous studies[3,6]. Therefore, we have additionally specified Subtypes 1, 2, 3, and 5 in the text.
(Page 10, Line 180)

4. The meaning of the many semi-transparent connections shown in Fig 2b are confusing and not explained.

Response: We aimed to identify which of the five subtypes of 'Combined NMF' are connected to the main streams defined in previous studies[3,6] and represent this with bold solid lines and colors for clarity. Semi-transparent lines and colors depict other associated stream connections.

5. What is the statistical significance of the correlation/overlaps indicated in Fig 2b?

Response: We calculated the statistical significance of the three comparisons depicted in **Fig. 2b** based on reviewers' comments using Fisher's exact test as shown in the Reviewer-only figure 20. Accordingly, we consider a strong association between two comparisons when the log2 odds ratio (OR) is positive and the adjusted p-value obtained through Benjamini-Hochberg adjustment for Fisher's exact test is below 0.05. This is also represented in **Figure 2b** by solid lines and distinct colors.



Reviewer-only Figure 20

Full rectangle and asterisk indicate significant overlaps (Fisher's test adjusted P ≤ 0.05, Benjamini-Hochberg adjustment); faint rectangle indicates overlaps that pass only the nominal P value (Fisher's test P ≤ 0.05, Fisher's test adjusted P > 0.05); and blank indicates overlaps which is not significant (Fisher's test P > 0.05).

6. Regarding the description of Subtype 4 and its lack of enrichment of features....is this referring to LUSC only, i.e., middle panel of Fig 2b?

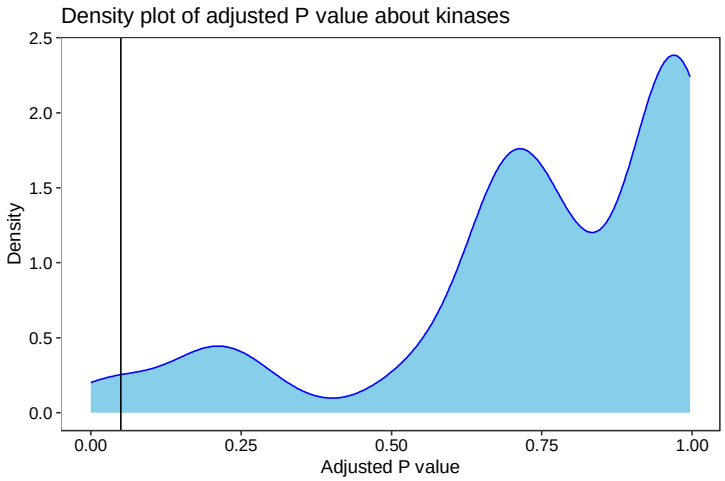
Response: We compared the top features of the NMF subtypes between our cohort and those of other studies[3,6-8] by datasets (protein, **Figure 2a**; phospho and acetyl, **Supplementary Fig. 2a and 2b**). Subtype 4 did not show any enrichment of the top protein features for previous subtypes, but contained some phosphorylation features associated with the EMT-enriched (LSCC[6]) subtype and acetylation features of the inflammatory subtypes of LUAD[3] and LSCC[6] (**Supplementary Fig. 2a and 2b**).
(Page 10, Line 165-175)

7. How was kinase activity measured (line 184)?

Response: The detailed method of kinase activity estimation is written on 'Kinase activity estimation based on phosphoproteomic data'.
(Page 49, Lines 1278-1294)

8. How specifically are the indicated FDR values for CSNK2A1 and GSK3B (line 186) supporting the conclusions drawn on the activation of these kinases?

Response: The indicated false discovery rate (FDR) values for CSNK2A1 and GSK3B, as shown on line 186, support the conclusions drawn on the activation of these kinases by indicating the level of confidence in the significance of their differential activation. A low FDR value suggests that the observed differences in the activation levels of these kinases between Subtype 4 and others are unlikely to be due to random chance. In other words, it indicates a high degree of statistical significance in the observed differences. Therefore, low FDR values for CSNK2A1 and GSK3B reinforce the conclusion that Subtype 4 indeed activates these kinases. It implies that the changes observed in the activation levels of these kinases are likely to be biologically meaningful and not simply a result of random variability or noise in the data.
Reviewer-only figure 21 shows the distribution of adjusted p-values from the Benjamini-Hochberg adjustment for all calculated kinase activity scores that were subjected to kinase activity estimation. Additionally, the solid line at x = 0.05 represents the threshold we consider indicative of estimated kinase activity scores that are not random outcomes. Specifically, this applies to CSNK2A1 and GSK3B.



Reviewer-only Figure 21

9. What are the specific phosphorylation sites quantified for CSNK2A1 and GSK3B, and does the literature support that these modifications are activating as opposed to inhibitory? Are they known to be sites that are regulated?

Response: We identified 221 and 332 substrates (phosphorylation sites) specifically regulated by CSNK2A1 and GSK3B, respectively, in our data. These were all obtained from known kinase-substrate Databases, please see my response to your comment #11 below for more details about this.
Furthermore, when comparing the fold change values obtained from comparing Subtype 4 and others for the phosphoproteome, with an absolute fold change greater than 1.5 and an adjusted p-value of 0.05 or less (Benjamini-Hochberg adjustment), CSNK2A1 and GSK3B were found to have 98 and 96 significant substrates (phosphosites) each. Among these, T366 phosphorylation of PTEN, a common substrate of CSNK2A1 and GSK3B, has been reported to have reduced biological activity in the regulation of PI3K-dependent signaling of PTEN[9], meaning that its activity in the PI3K to PIP2 dephosphorylation process is reduced, and PI3K activates the conversion of PIP2 to PIP3 and subsequent AKT signaling, which is not regulated by PTEN. It has also been reported that CSNK2A1 phosphorylates S129 of AKT1, activating it's activity[10]. Comprehensively, in subtype 4, these modifications (i.e. phosphorylation) derived from CSNK2A1 and GSK3B may support the activation of PI3K/AKT signaling pathway (Refer to revised **Figure 2f**).

10. They describe phosphorylation at position S347 in SLK as "significantly upregulated" in subtype 4 (line 192). Have they quantified the level of SLK protein in comparison with pS347 peptides to justify the conclusion that the stoichiometry of phosphorylation is actually increased, rather than an increased expression of SLK with no change in the level of phosphorylation at S347.

Response: We appreciate the reviewer's comment. As suggested by the reviewer, we normalized PTM data against protein abundance by employing the method published recently (Geffen, Yifat et al. Cell 2023). Interestingly, phosphorylation at position S347 in SLK showed more increase when corrected for protein abundance than before (Before: Log_2(fold change) = 0.47, $P < 0.01$; After: Log_2(fold change) = 0.61, $P < 0.001$). We also examined survival analysis result with normalized data, which indicated that high expression of phosphorylation at position S347 in SLK contributes to poor outcome ($P = 0.00093$).
While we acknowledge the potential influence of protein quantity on the quantification of its phosphorylation sites, we believe that even when considering this, phosphorylation at position S347 in SLK could still be considered a target significantly increased in Subtype 4 in our dataset.

11. Is there evidence that S347 is phosphorylated by CSNK2A1?

Response: We utilized the PHONEMeS package for kinase activity estimation. Quoting from the paper[11], kinase activity estimation is performed through a Prior Knowledge

Network (PKN) ('PHONEMeS combines identified phosphopeptide abundance measurements (e.g., from untargeted shotgun phosphoproteomic experiments) with a large Prior Knowledge Network (PKN)'). The PKN, by default, is constructed as a network of kinase/phosphatase-substrate relationships using the OmniPath R package's database. Through confirmation from this database, we ascertain that CSNK2A1 phosphorylates SLK S347, indicating its role in phosphorylation.

Reviewer #4 (Remarks to the Author): expertise in NSCLC neoantigens

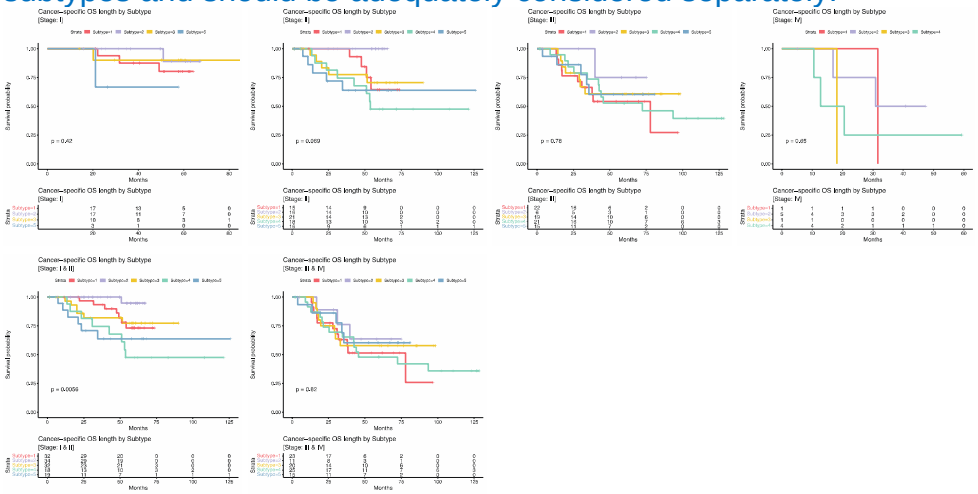
The authors conduct a study of comprehensive multiomic analysis of 229 patients in Korea with NSCLC, performing whole exome sequencing, bulk RNA sequencing, and global/phospho/acetyl proteomic analyses. Global proteomic, phosphoproteomic, and acetylproteomic data were integrated and non-negative matrix factorization clustering performed to define 5 subtypes of NSCLC. The authors benchmarked their proposed subtype classification against previously published classifications and performed comparative analyses (Figure 2). The investigators glean important biological insights from their subtype classification and propose some potential therapeutic candidates. While their methods are similar to those in previously published cohorts, they perform their analysis in a Korean population (representing a lung cancer population enriched in adenocarcinoma and exhibiting low tumor mutation burden of 2.7 TMG that is distinct from the population studied in a Western population) and offer a valuable multiomic meta-analysis comparing to previously published data.

Strengths:
-Some of the biological insights they offer include the following: propose XPO1 as a potential druggable target with Selinexor for subtype 3, highlight a subtype 4 with potential targets in PI3K, VEGF, HIF pathways and neutrophil pathway requiring further characterization, and demonstrate cryptic MAPs as a features of subtype 2 which may be a potential vaccine target. They also show that SMARCA4 mutations are associated with increased expression of SLAMF7 in subtype 5, which might have therapeutic implications.
-Authors perform an analysis of both conventional and cryptic MAPs.

Limitations:
-A limitation of the study is that the subtype classification is mainly pertinent to adenocarcinomas (subtypes 1, 2, 4, 5), as essentially all the lung squamous cell carcinomas fall into subtype 3. In addition, a limitation for interpreting how subtype classification impacts clinical outcomes (Figure 2g) is the heterogeneity in stage across the various subtypes, which would be expected to independently impact survival. A stage by stage comparison would be required to compare clinical outcomes across subtypes. Kaplan Meyer curves should include the number of patients being included at various time points.

Response: We agree with the reviewer's comment: “Kaplan Meyer curves should include the number of patients being included at various time points”. We have updated **Figure**

2g and 2h. In our study, we clustered subtypes based on global proteomic and PTM proteomic datasets, interpreting that the concentration of stages within specific subtypes carries clinical significance. Subtype 2, for instance, exhibits a higher composition of patients at relatively early stages (Stages I and II). However, acknowledging that stages may vary between subtypes and could impact survival, we compared survival across each stage and also conducted analyses for early (I and II) and late stages (III and IV) as shown in the Reviewer-only figure 22. Our findings revealed no significant differences between subtypes within each stage, which could be attributed to insufficient data when divided by stage. Notably, significant differences were observed among patients in the early stages (I and II) between subtypes (P = 0.0056), with Subtype 4 identified as the most vulnerable subtype. Therefore, since the subtypes were derived from proteomic data rather than clinical data (stage), the distribution across stages can reflect the characteristics of subtypes and should be adequately considered separately.



Reviewer-only Figure 22

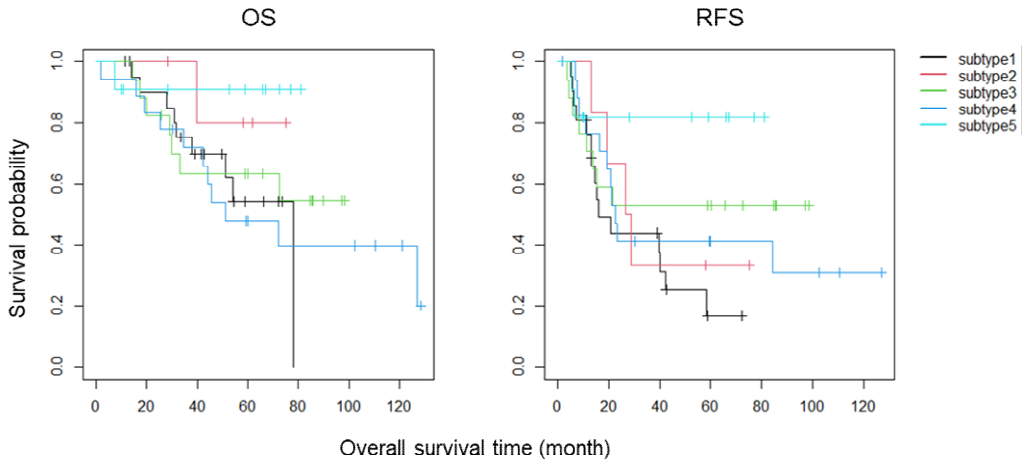
-This study generates multiple new hypotheses that will require further evaluation in future studies that are beyond the scope of this study

Response: In this paper, we have subdivided the subtypes of non-small cell lung cancer in Koreans and identified their clinical and biological characteristics. Furthermore, we have investigated the biological pathways contributing to these characteristics and identified key regulatory factors within these pathways, utilizing independently reported data (from CPTAC) for validation. We agree with the reviewer's comment regarding the need for additional evaluation of many of our assertions. Currently, we are in the process of studying targeted validation experiments for factors (proteins or post-translational modifications) that represent unique features corresponding to the hypotheses we have proposed.

Suggestions:
-The conclusions from the findings of adjuvant chemotherapy/adjuvant radiation clinical outcomes according to multiomic subsets appear to be overstated. Only half of patients received adjuvant chemo or radiation, and there are only 34 patients in subtype 5, really limiting the interpretation of the kaplan meyer curves in Figure 6g. I would like to see the

impact of adjuvant chemotherapy on survival for the overall population (subtypes 1 – 5), as these treatments are typically associated with a 5% improvement in overall survival when analyzing large datasets of patients.

Response: We agree with the reviewer's comment that the results presented in **Figure 6g** may be overstated due to the small sample sizes of each subtype. Following the reviewer's recommendation, we conducted a survival analysis incorporating adjuvant chemotherapy across the entire study population. This analysis revealed that subtype 5 exhibited the most favorable prognosis compared to other subtypes, although the statistical significance was marginal (p = 0.136 in OS and p = 0.062 in RFS) as shown in the Reviewer-only figure 23. This trend supports the conclusion drawn from **Figure 6g**, suggesting that immunogenic subtypes enhance the efficacy of adjuvant chemotherapy. We have included these findings for the overall population in the revised manuscript as follows: (Page 28, Lines 646-649). In contrast, no substantial improvement in survival was observed in patients with other subtypes who underwent adjuvant chemotherapy or chemoradiation therapy. Subtype 5 also demonstrated the most favorable prognosis compared to other subtypes in the overall population treated with adjuvant chemotherapy, despite the marginal statistical significance (data not shown). This underscores the clinical significance of Subtype 5 in predicting the response to adjuvant chemotherapy.



Reviewer-only Figure 23

-In Figure 2h, outcomes for patients with and without metastases are compared in subtype 4, and the conclusion that metastasis is not the sole mechanism leading to poorer survival is a reasonable hypothesis, but not supported by the data which is not adequately powered to see a difference with only a handful of patients with metastasis.
(KU)

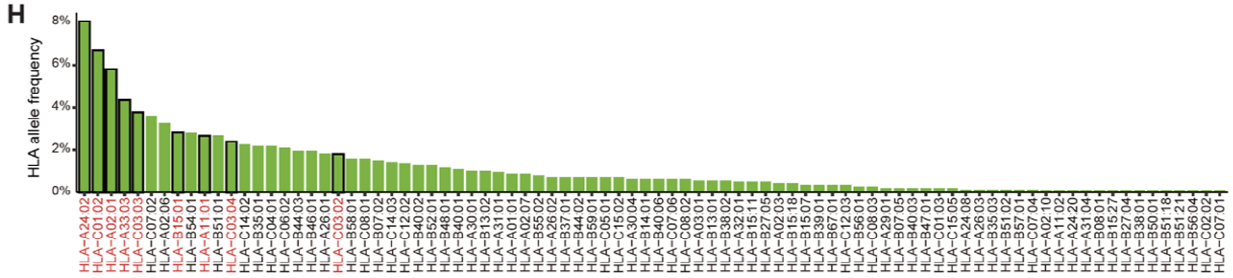
Response: In this study, downstream analysis was not conducted on the entire cohort of 229 patients; instead, it was specifically focused on one subtype (Subtype 4), which was further divided into groups based on the presence or absence of metastasis. As a result, the analysis involved 29 patients with metastasis and 12 patients without metastasis within Subtype 4. We have also added the corresponding population count information by time for each group in the revised **Figure 2**. We recognize that this may not be a large enough sample size for the reviewer's concerns, but we hope you understand that we wanted to report what we found within a limited cohort.

-For the methods for calling cryptic MAPs, why was 3 frame translation (rather than 6 frame translational) used for searching for cryptic MAPs?

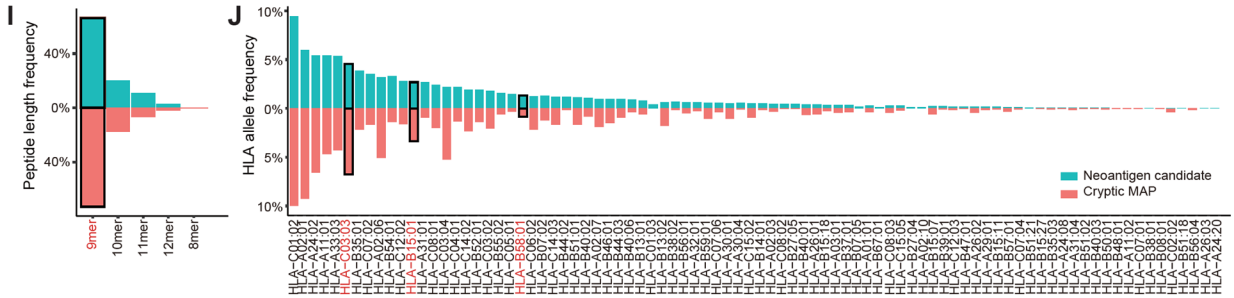
Response: We focused on identifying novel peptides with RNA evidence to reduce false positives (the details in "**Identification of Novel Peptides**"). Since our RNA-seq data is strand-specific, antisense transcripts were predicted using StringTie (GffCompare codes s and x, details available at https://ccb.jhu.edu/software/stringtie/gffcompare.shtml). However, when we searched against all predicted transcripts using our benchmarking datasets[12], some types of transcripts including antisense transcripts were rarely identified. Consequently, we chose to exclude infrequently translated predicted transcripts identified by StringTie (GffCompare codes e, o, s, x, I, y, p, r, and u) and opted for three-frame translation instead of six-frame translation.

-Of the conventional and cryptic MAPs identified in this study, what was the distribution of predicted HLA binding and does this reflect the expected HLA distribution of a Korean population? Also, I would like to see the peptide length distribution, HLA allele distribution, and single nucleotide polymorphism data presented for cryptic MAPs vs conventional MAPs and how the results compared to those previously reported.

Response: In response to the reviewer's suggestion, we analyzed the distribution of HLA alleles before and after binding prediction to ensure consistency with distributions previously reported in the Korean population. Our analysis showed that the top-ranked HLA alleles in our cohort align closely with the highly frequent HLA alleles (≥10%) reported in the Korean population, as highlighted by the bold-lined bars in **Supplementary Fig. 6h**.



Following the reviewer's comment, we analyzed the distribution of peptide length and HLA allele post-binding prediction. It is important to note that SNPs were filtered out before the mutation calling analysis during the binding prediction process, making an analysis of SNP distribution unavailable. Consequently, 9-mer peptides were predominantly predicted to bind with both neoantigen candidates and cryptic MAPs as previously reported. Additionally, some of the most frequently presented peptides by HLAs in our predictions have been documented in the previous study (indicated by bold-lined bars), although discrepancies may arise from differences in cohort selection and methodological approaches.



These results are presented in the revised **Supplementary Fig. 6h-j** and discussed in the manuscript.

(Page 28, Lines 651-654) Furthermore, the distribution of HLA alleles, both before and after binding prediction, was consistent with previous reports on the Korean population and lung cancer studies[13-16], supporting the validity of our results (**Supplementary Fig. 6h-j**).

Clarifications:

-Line 720 – 726: Please clarify the method by which the 250 patients were selected among the 408 NSCLC patients. Manuscript states that these were selected based on patients with locally advanced (which I'm interpreting to be stage II – III) and metastatic (IV), but it's not clear to me that all the patients that are excluded are stage I patients per Fig 1A.

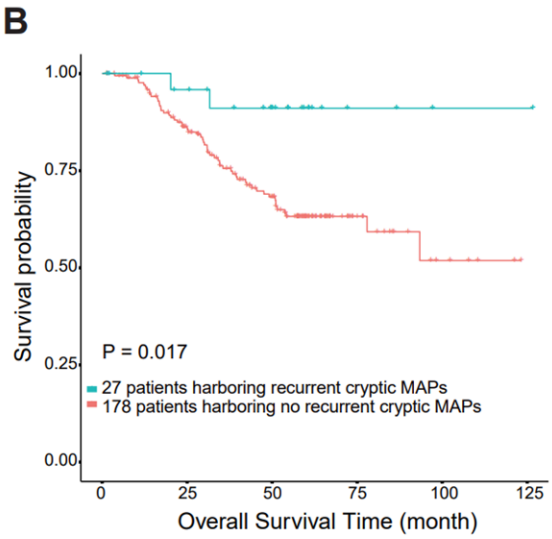
Response: We acknowledge that there were some confusing statements about the patient selection. Initially, we selected 408 patients for whom both fresh frozen tumor and NAT tissue which were frozen within 15 minutes. Subsequently, from this group, we included patients who were positive for nodal metastasis (N=137). Following this, we included the node-negative cases for remaining 113 cases. We have clarified the process of patient selection in Method section to avoid any confusion.

(Page 33, Lines 766-768)

-In Figure 6B, is the legend mislabelled? Should the curve in blue be patients with recurrent cryptic MAPs? How is the term "recurrent cryptic MAPs" defined for the purposes of Fig 6B and is this different that the term "confirmed cryptic MAP"?

Response: We apologize for any inconvenience caused by the mislabeling error in **Figure 6b**. We have now corrected the color of the curves accordingly; blue now accurately represents patients with recurrent cryptic MAPs. We define "recurrent cryptic MAPs" refers to confirmed cryptic MAPs that were predicted to be present in more than three patients. This clarification has been incorporated into the revised manuscript for better understanding.

(Page 27, Lines 615-618) Our analysis revealed 12 confirmed cryptic MAPs occurring in more than three patients, which we have termed recurrent cryptic MAPs. Notably, some of these MAPs originate from the same gene of origin (**Supplementary Fig. 6d** and **Supplementary Table 6c**).



-Line 281: It is not feasible in clinical practice for adjuvant treatment to be based on multiomics data. However, there are lessons learned from multiomic data studies that can have important implications for clinical practice.

Response: We acknowledge the discrepancy in the line number referenced by the reviewer, which likely corresponds to line 681. We agree with your points that the multiomics data is not appropriate for clinical application. We amended the paragraph in Discussion section to reflect the above points.
(Page 31, Lines 720-725)

-Line 685 – 686: ADAURA trial was not statistically designed to look for a difference in outcome for patients with and without chemotherapy, so this is not the appropriate interpretation of the trial results.

Response: The article we cited (Li et al., Transl Lung Cancer Res 2021) described no significant survival benefit from adjuvant therapy in the Osimertinib-treated group. Upon reflection, we agree that this reference may not robustly support our results; therefore, we replaced the references to the other articles[17,18].
(Page 31, Lines 727-729)

Reference

1 Xia, C., Tao, Y., Li, M., Che, T. & Qu, J. Protein acetylation and deacetylation: An important regulatory modification in gene transcription (Review). *Exp Ther Med* **20**, 2923-2940 (2020). https://doi.org:10.3892/etm.2020.9073

2 Zheng, Z. *et al.* Pulmonary Hypertension: A Predictor of Lung Cancer Prognosis? *Am J Respir Crit Care Med* **204**, 1112-1113 (2021). https://doi.org:10.1164/rccm.202105-1256LE

3 Gillette, M. A. *et al.* Proteogenomic Characterization Reveals Therapeutic Vulnerabilities in Lung Adenocarcinoma. *Cell* **182**, 200-225 e235 (2020). https://doi.org:10.1016/j.cell.2020.06.013

4 Kim, Y. *et al.* Integrative and comparative genomic analysis of lung squamous cell carcinomas in East Asian patients. *J Clin Oncol* **32**, 121-128 (2014). https://doi.org:10.1200/JCO.2013.50.8556

5 Lee, J. G., Kim, H. C. & Choi, C. M. Recent Trends of Lung Cancer in Korea. *Tuberc Respir Dis (Seoul)* **84**, 89-95 (2021). https://doi.org:10.4046/trd.2020.0134

6 Satpathy, S. *et al.* A proteogenomic portrait of lung squamous cell carcinoma. *Cell* **184**, 4348-4371 e4340 (2021). https://doi.org:10.1016/j.cell.2021.07.016

7 Xu, J. Y. *et al.* Integrative Proteomic Characterization of Human Lung Adenocarcinoma. *Cell* **182**, 245-261 e217 (2020). https://doi.org:10.1016/j.cell.2020.05.043

8 Lehtio, J. *et al.* Proteogenomics of non-small cell lung cancer reveals molecular subtypes associated with specific therapeutic targets and immune evasion mechanisms. *Nat Cancer* **2**, 1224-1242 (2021). https://doi.org:10.1038/s43018-021-00259-9

9 Maccario, H., Perera, N. M., Davidson, L., Downes, C. P. & Leslie, N. R. PTEN is destabilized by phosphorylation on Thr366. *Biochem J* **405**, 439-444 (2007). https://doi.org:10.1042/BJ20061837

10 Borgo, C., D'Amore, C., Sarno, S., Salvi, M. & Ruzzene, M. Protein kinase CK2: a potential therapeutic target for diverse human diseases. *Signal Transduct Target Ther* **6**, 183 (2021). https://doi.org:10.1038/s41392-021-00567-7

11 Gjerga, E., Dugourd, A., Tobalina, L., Sousa, A. & Saez-Rodriguez, J. PHONEMeS: Efficient Modeling of Signaling Networks Derived from Large-Scale Mass Spectrometry Data. *J Proteome Res* **20**, 2138-2144 (2021). https://doi.org:10.1021/acs.jproteome.0c00958

12 Song, K. J. *et al.* Assessing Long-Term Stored Tissues for Multi-Omics Data Quality and Proteogenomics Suitability. *bioRxiv*, 2024.2003.2013.583805 (2024). https://doi.org:10.1101/2024.03.13.583805

13 Choe, W. *et al.* Identification of 8-Digit HLA-A, -B, -C, and -DRB1 Allele and Haplotype Frequencies in Koreans Using the One Lambda AllType Next-Generation Sequencing Kit. *Ann Lab Med* **41**, 310-317 (2021). https://doi.org:10.3343/alm.2021.41.3.310

14 Lee, K. W., Oh, D. H., Lee, C. & Yang, S. Y. Allelic and haplotypic diversity of HLA-A, -B, -C, -DRB1, and -DQB1 genes in the Korean population. *Tissue Antigens* **65**, 437-447 (2005). https://doi.org:10.1111/j.1399-0039.2005.00386.x

15 Sarkizova, S. *et al.* A large peptidome dataset improves HLA class I epitope prediction across most of the human population. *Nat Biotechnol* **38**, 199-209 (2020). https://doi.org:10.1038/s41587-019-0322-9

16 Jiang, T. *et al.* Heterogeneity of neoantigen landscape between primary lesions and their matched metastases in lung cancer. *Transl Lung Cancer Res* **9**, 246-256 (2020). https://doi.org:10.21037/tlcr.2020.03.03

17 Isaka, T. *et al.* Efficacy of Platinum-Based Adjuvant Chemotherapy on Prognosis of Pathological Stage II/III Lung Adenocarcinoma based on EGFR Mutation Status: A

Propensity Score Matching Analysis. *Mol Diagn Ther* **23**, 657-665 (2019). https://doi.org:10.1007/s40291-019-00419-9

18 Kawaguchi, Y. *et al.* Epidermal growth factor receptor mutation subtype has differential effects on adjuvant chemotherapy for resected adenocarcinoma pathological stages II-III. *Oncol Lett* **18**, 6451-6458 (2019). https://doi.org:10.3892/ol.2019.11050

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed the queries from the first round of review very thoroughly. Some of the reviewer only figures could benefit the manuscript, if they were included as additional supporting information. Minor revisions are suggested below.

1. The database searching strategy is not included in the methods. It should be added on p. 43 of the current manuscript draft between the LC-MS/MS methods and the downstream data analysis. The response to reviewer 1 query 1 should also be included in the text to describe the sequential searches.

2. Reviewer only Figures 2 and 3 would be beneficial to include as supplemental figures in the manuscript.

Reviewer #2 (Remarks to the Author):

All of my questions have now been adequately answered by Song and co-workers, and I have no further questions.

Reviewer #3 (Remarks to the Author):

Assessing responses to concerns was confusing since the quoted page and line numbers do not correspond to the revised manuscript. Most concerns have been addressed. However:

Concern #4, which relates to Fig. 2b, has not been address by way of revision in the manuscript.

Concern #5, which relates to Fig. 2b has not been addressed by revision of the figure or in manuscript text. The reviewer-only Fig. 20 is helpful and should be added as a supplementary figure or included as a revised figure in the manuscript.

Reviewer #4 (Remarks to the Author):

I commend the authors for their detailed and thoughtful responses to the reviewers' comments.

They adequately addressed my questions about conventional and cryptic MAPS and appreciate the results presented in revised supplementary Fig 6h-j

A few minor points in the reporting of the results:

Reviewer only Figure 22: Acknowledge the limitations in interpreting survival differences by subtype when further broken down by stage due to insufficient numbers. I think it is worth noting in the reporting of the results on lines 307 – 310 that the prognosis of the biological subtypes may also impacted by stage distribution, as patients with subtype IV did not have stage I disease at diagnosis (and were all stage II and higher in this cohort).

Reviewer only Figure 23 and manuscript lines 715 – 757: would modify the language from "this underscores the clinical significance of subtype 5 in predicting the response to adjuvant chemotherapy" to be suggestive, rather than so definitive. For example, that subtype 5 in this cohort is associated with clinical benefit from adjuvant chemotherapy.

REVIEWERS' COMMENTS

Reviewer #1 (Remarks to the Author):

The authors have addressed the queries from the first round of review very thoroughly. Some of the reviewer only figures could benefit the manuscript, if they were included as additional supporting information. Minor revisions are suggested below.

1. The database searching strategy is not included in the methods. It should be added on p. 43 of the current manuscript draft between the LC-MS/MS methods and the downstream data analysis. The response to reviewer 1 query 1 should also be included in the text to describe the sequential searches.

Response: Thank you for the detailed guidance. We added a new section "The Database Search Strategy" to Page 33, Lines 934-943, and rearranged the subsequent paragraphs to describe the specific techniques for each database search method.

2. Reviewer only Figures 2 and 3 would be beneficial to include as supplemental figures in the manuscript.

Response: Thank you for the detailed guidance. We included the corresponding data in Supplementary Fig. 2e and 2h.

Reviewer #2 (Remarks to the Author):

All of my questions have now been adequately answered by Song and co-workers, and I have no further questions.

Reviewer #3 (Remarks to the Author):

Assessing responses to concerns was confusing since the quoted page and line numbers do not correspond to the revised manuscript. Most concerns have been addressed. However:

Concern #4, which relates to Fig. 2b, has not been address by way of revision in the manuscript.

Response: Thank you for the detailed guidance. I apologize for the lack of clarification. We added it to the Figure 2 legend to clarify (Page 63, Lines 27-28).

Concern #5, which relates to Fig. 2b has not been addressed by revision of the figure or in manuscript text. The reviewer-only Fig. 20 is helpful and should be added as a supplementary figure or included as a revised figure in the manuscript.

Response: Thank you for the detailed guidance. To help clarify the manuscript, we added it as Supplementary Fig. 2c.

Reviewer #4 (Remarks to the Author):

I commend the authors for their detailed and thoughtful responses to the reviewers' comments.

They adequately addressed my questions about conventional and cryptic MAPS and appreciate the results presented in revised supplementary Fig 6h-j

A few minor points in the reporting of the results:

Reviewer only Figure 22: Acknowledge the limitations in interpreting survival differences by subtype when further broken down by stage due to insufficient numbers. I think it is worth noting in the reporting of the results on lines 307 – 310 that the prognosis of the biological subtypes may also impacted by stage distribution, as patients with subtype IV did not have stage I disease at diagnosis (and were all stage II and higher in this cohort).

Response: Thank you for the helpful suggestion. We added reviewer's comment to Page 11, Line 224 for readers to reference and understand.

Reviewer only Figure 23 and manuscript lines 715 – 757: would modify the language from "this underscores the clinical significance of subtype 5 in predicting the response to adjuvant chemotherapy" to be suggestive, rather than so definitive. For example, that subtype 5 in this cohort is associated with clinical benefit from adjuvant chemotherapy.

Response: As suggested by the reviewer, we modified the tone of description about adjuvant chemotherapy on subtype 5 from definitive to suggestive.

(Page 20, Lines 513-515) This underscores that subtype 5 could be associated with a clinical benefit from adjuvant chemotherapy.