NÖHÜ Müh. Bilim. Derg. / NOHU J. Eng. Sci., 2024; 13(4), 1545-1552

Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi
Niğde Ömer Halisdemir University Journal of Engineering Sciences

Araştırma makalesi / Research article

www.dergipark.org.tr/tr/pub/ngumuh / www.dergipark.org.tr/en/pub/ngumuh

Proteogenomic profiling of lung adenocarcinoma reveals therapeutic targets for precision medicine

Akciğer adenokarsinomunun proteogenomik analizi: hassas tıp için terapötik hedeflerin belirlenmesi

Abdulkadir Elmas[1,*]

[1] *Giresun University, Department of Electrical and Electronics Engineering, Giresun 28200, Türkiye*
[1] *Icahn School of Medicine at Mount Sinai, Department of Genetics and Genomic Sciences, New York, NY 10029, USA*

Abstract

Lung cancer is the top cause of cancer-related fatalities worldwide, impacting both men and women. A major challenge is its frequent diagnosis at advanced stages, which limits treatment options. While genomic and transcriptomic analyses have traditionally been used to identify potential drug targets, there remains an unexplored potential in targeting protein-level anomalies. This study systematically investigates the proteomic landscape of 109 primary lung adenocarcinoma (LUAD) tumors using comprehensive mass-spectrometry (MS) proteomics data. By focusing on kinases, the key actors in oncogenic signaling pathways, we aim to find new therapeutic targets for LUAD. Through intricate analyses encompassing tumor-normal differentials and inter-tumor variations, our study identifies notable overexpressed targets, including PLAU, MET, ERBB2, EGFR, PDK1 kinases, and THBS2, CRABP2, INPP4B proteins, many of which present no evidence of transcriptomic alteration. Several targets we identified through proposed approaches have corresponding inhibitor drugs, including ERBB2 kinase (Afatinib) and VEGF-A protein (Bevacizumab). Our findings validate known therapeutic markers in lung cancer and reveal candidate protein targets specific to LUAD, underscoring the efficacy of proteomic methodologies in advancing precision medicine for cancer.

Keywords: Weighted k-nearest neighbor (KNN) algorithm, Lung adenocarcinoma (LUAD), Proteomics, Targeted therapy, Precision oncology

Öz

Akciğer kanseri, dünya genelinde kanserle ilişkili ölümlerin başlıca nedeni olup, hem erkekleri hem de kadınları etkilemektedir. En büyük zorluklardan biri, genellikle hastalığın ileri evrelerinde teşhis edilmesidir; bu da tedavi seçeneklerini kısıtlamaktadır. Genomik ve transkriptomik analizler geleneksel olarak potansiyel ilaç hedeflerini belirlemede kullanılmıştır; ancak protein düzeyindeki anormallikleri hedeflemekte henüz keşfedilmemiş bir potansiyel bulunmaktadır. Bu çalışma, 109 birincil akciğer adenokarsinomu (LUAD) tümörünün proteomik profilini kapsamlı kütle spektrometrisi (MS) verileri kullanarak sistematik bir şekilde incelemektedir. Onkogenik sinyal yolarında kritik rol oynayan kinazlara odaklanarak, LUAD için yeni terapötik hedefler bulmayı amaçlıyoruz. Tümör-normal farklılıkları ve tümörler arası varyasyonları içeren ayrıntılı analizler sonucunda, PLAU, MET, ERBB2, EGFR, PDK1 kinazları ve THBS2, CRABP2, INPP4B proteinleri gibi önemli aşırı ekspres edilen hedefler belirlenmiştir. Bu hedeflerin çoğunda transkriptomik değişim kanıtı bulunmamaktadır. Önerilen yaklaşımlar aracılığıyla belirlediğimiz bazı hedefler için mevcut inhibitör ilaçlar geliştirilmiştir, ERBB2 kinazı (Afatinib) ve VEGF-A proteini (Bevacizumab) gibi. Bulgularımız, akciğer kanserindeki bilinen terapötik belirteçleri doğrulamakta ve LUAD'e özgü aday protein hedeflerini ortaya koyarak, proteomik yöntemlerin kanser tedavisinde kişiselleştirilmiş tıbbın ilerletilmesindeki etkinliğini vurgulamaktadır.

Anahtar kelimeler: Ağırlıklı k-en yakın komşu (KNN) algoritması, akciğer kanseri, proteomik metodolojiler, hedefe yönelik tedavi, hassas onkoloji

1 Introduction

Lung cancer remains the foremost cause of cancer-related mortality globally, affecting both men and alike [1]. The high mortality rate is largely attributable to the disease's tendency to be diagnosed at an advanced stage, often limiting the efficacy of conventional treatment options. Recent advances in cancer research emphasize the need for novel therapeutic strategies that can address this challenge by targeting molecular aberrations specific to the tumor [2]. Historically, treatment strategies for lung adenocarcinoma (LUAD) have predominantly focused on identifying potential drug targets through genomic and transcriptomic analyses [3]. However, these methods may overlook important therapeutic opportunities present at the protein level. Proteomic analysis, which focuses on the expression and functional alterations of proteins, offers a complementary approach that can reveal

* Sorumlu yazar / Corresponding author, e-posta / e-mail: abdulkadir.elmas@giresun.edu.tr (A. Elmas)
Geliş / Recieved: 03.09.2024 Kabul / Accepted: 03.10.2024 Yayımlanma / Published: 15.10.2024
doi: 10.28948/ngumuh.1542153

critical therapeutic targets that are not apparent through transcriptomic data alone [4-6].

Recent progress in mass spectrometry (MS) technology have significantly expanded global proteomic datasets, allowing for the quantification of nearly all proteins expressed in primary tumor cohorts [7,8]. This offers valuable opportunities to explore protein-level abnormalities, which may serve as prognostic biomarkers or therapeutic targets. Despite the extensive characterization of genomic aberrations, protein aberrations have historically been less well-defined, highlighting the urgent need for systematic analyses to identify potential targets [9].

In this work, we explore the proteogenomic landscape of LUAD by leveraging comprehensive mass-spectrometry (MS) proteomics data derived from 109 primary LUAD tumors and 102 matched normal samples [10]. Through the integration of these proteomic insights with genomic data and established drug-target relationships from the Drug-Gene Interaction database (DGIdb) [11], we aim to uncover novel protein targets that could be harnessed as effective therapeutic candidates. Our approach includes a detailed analysis of differentially expressed proteins (DEPs) to pinpoint those with significant alterations in tumor versus normal tissue. We also focus on kinases, which are often key players in oncogenic signaling pathways and are established therapeutic targets in various cancers [12]. Through this comprehensive analysis, we seek to uncover proteins that are overexpressed in LUAD and may serve as actionable targets for new or repurposed therapies. By comparing our proteomic findings with transcriptomic data, we also explore the relationship between protein-level overexpression and underlying genetic alterations, investigating whether protein overexpression is a result of genomic changes such as copy-number amplifications or post-transcriptional modifications, which can be missed in RNA-based analyses. This study seeks to advance precision medicine by focusing on protein-level abnormalities in LUAD and identifying potential druggable targets that could lead to improved patient outcomes. The findings have the potential to guide future therapeutic strategies and contribute to more personalized treatment approaches for lung cancer.

2 Materials and methods

2.1 Data sources, download, and standardized normalization

The proteomic and transcriptomic datasets for the LUAD cohort samples were obtained from The National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) [10]. The study involved 109 tumor specimens and 102 corresponding normal controls, obtained from 111 patients (34.6% female) with average onset age of 62.7 years. The dataset consisted of 11,029 quantified distinct proteins, including 507 kinases. The RNA-seq analysis generated gene expression profiles for the LUAD study group, using the tophat-cufflinks pipeline, identifying 35,220 protein-coding genes that showed FPKM > 1 in multiple samples, including 640 kinase-encoding genes. Quantile and log2 normalization were applied to the RNA-seq counts (FPKM-normalized), and genes not expressed in 20% or more of the samples were filtered out. The data distribution for the proteomics group was assessed, followed by the application of a uniform normalization method. The samples in this group were adjusted using their Median Absolute Deviation (MAD), standardizing them so that all samples across the different datasets shared a consistent (unit) MAD value. Additionally, protein markers with a significant amount of missing data (20% or more) were omitted from the analysis.

2.2. Identification of differentially-expressed proteins

A paired analysis was performed on the LUAD cohort to pinpoint proteins with differential expression by evaluating tumor samples against their corresponding normal tissues using the R package "limma" (version 3.42.2) [13]. Corrections were made for potential confounding variables, including batch effects (such as sequencing center/operator/date, TMT batch) and demographic factors (gender, age). The Benjamini-Hochberg (BH) procedure was utilized to correct p-values for multiple comparisons, aiming to manage the false discovery rate (FDR). Generally, no significant confounding influences were detected between protein expression levels and clinical variables like gender and age.

2.3. Detection of overexpressed proteins/genes

The OPPTI method [9] was employed to identify overexpressed markers. The method involves comparing marker expression levels to a predicted value for each tumor sample, which is derived using a k-nearest neighbor (KNN) algorithm. This algorithm calculates the predicted value based on the expression levels of other (nearest) co-expressed markers. The OPPTI method employs a permutation test to assess the statistical significance of a marker's calculated overexpression (dysregulation). Within the dataset, scores of dysregulation are shuffled among proteins in each sample, and baseline overexpression occurrences are estimated from these shuffles. This shuffling procedure is repeated several times, and the overall baseline overexpression occurrences from these repetitions are used to generate the permutation distribution.

3. Results and discussion

3.1. LUAD proteomics cohort

We curated genomic and comprehensive mass-spectrometry (MS) proteomics data from 111 lung adenocarcinoma patients [10], consisting of 109 tumor cases and 102 matched normal samples (Figure 1a). A standardized normalization process was applied, and stringent quality-control measures were followed (Methods), leading to the quantification of 11,029 proteins for our analyses. We also compiled a list of drug compound-associated genes from the Drug-Gene Interaction database (DGIdb) [11]. By aligning the measured proteins with the DGIdb list of druggable genes, we found that 1,834 proteins in our dataset are currently targetable by existing drugs.

a

Study	LUAD
PMID	32649874
Sample Size (Patients/Tumor/Normals)	P: 111 T: 109 N: 102
Female Percentage	34.6%
Average Age at Onset	62.7
Tumor Stage Distribution	1: 53.5% 2: 27.5% 3: 18.5% 4: 0.5%
# of quantified proteins	11,029
# of quantified DGIdb-druggable proteins	1,834
# of quantified kinases	507
# of quantified DGIdb-druggable kinases	322
# of quantified RNA-seq expressions (FPKM > 1)	35,220

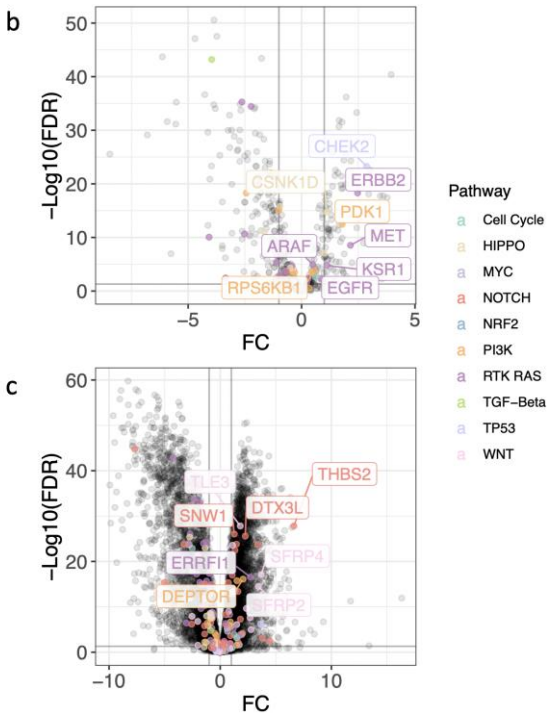

Figure 1. Overview of the study and differential protein expression in LUAD cohort. (a) Summary of the proteogenomic datasets from the human lung cancer cohort examined in this research. (b) Volcano plot illustrating the differential expression of kinase proteins between tumor versus normal samples, with the most significant kinases from oncogenic signaling pathways highlighted with labels. (c) The same analysis results in panel b for non-kinase proteins

Additionally, we selected kinase proteins for further examination, specifically concentrating on those known to be effective therapeutic targets across various cancer types. From a previously compiled set of 683 human kinase proteins [12], 507 were well-quantified within the LUAD dataset. Furthermore, these proteins were annotated by referencing ten oncogenic signaling pathways compiled by the TCGA PanCanAtlas, including PI3K, NOTCH, MYC, HIPPO, TGFβ, RTK/RAS/MAP-Kinase, β-catenin/WNT, Cell Cycle, oxidative stress response/NRF2, and TP53 signaling pathways [12].

3.2. Differentially expressed proteins

We conducted a paired analysis comparing tumor and normal tissues to identify proteins with differential expression (tumor-DEPs), while controlling for potential confounding factors such as age, gender, and ethnicity, utilizing the limma package [13] in R (version 3.42.2) (Methods). Our analysis revealed 405 significant kinase DEPs through differential expression testing, applying empirical Bayes moderation to the t-statistics [13] and enforcing a stringent false discovery rate (FDR) threshold of < 0.05. Of these DEPs, 42 were associated with oncogenic signaling pathways, and 6 kinases demonstrated more than 2-fold up-regulation in tumor samples, i.e., CHEK2 (log2-fold-change [FC] = 2.9, FDR = 6e-24), ERBB2 (FC = 2.5, FDR = 6e-19), PDK1 (FC = 1.8, FDR = 4e-13), MET (FC = 2.2, FDR = 3e-09), CSNK1D (FC = 1.1, FDR = 2e-15), and KSR1 (FC = 1.1, FDR = 1.7e-05) (Figure 1b).

Among the non-kinase proteins, a total of 8,219 differentially expressed proteins (DEPs) were identified (FDR < 0.05). Among these, 144 were associated with oncogenic signaling pathways, with 38 showing more than 2-fold up-regulation in tumor samples. Notable markers include THBS2 (FC = 6.6, FDR = 2e-28), DTX2 (FC = 3.4, FDR = 2e-24), SFRP4 (FC = 3.7, FDR = 8e-18), DTX3L (FC = 2.3, FDR = 3e-26), TLE3 (FC = 1.8, FDR = 2e-28), and SFRP2 (FC = 3.4, FDR = 4e-15) (Figure 1c). This differential expression analysis highlights a broad array of proteins that are elevated in tumor samples relative to normal tissues, necessitating further investigations to accurately identify and validate therapeutic targets among these DEPs.

3.3. Overexpression of currently-druggable proteins

To pinpoint potential therapeutic targets among overexpressed proteins in lung adenocarcinoma (LUAD), we employed OPPTI algorithm [9], a tool specifically designed for the detection of overexpressed proteins in global mass spectrometry (MS) proteomic datasets with varying quantitative distributions. This approach allows us to identify proteins that may exhibit efficacy upon inhibition and are not restricted by technical platform differences. We found 62 kinases with notable overexpression enrichment (permutation test by OPPTI for elevated markers, FDR < 0.05), of which 35 are DGIdb-listed druggable genes, including PLAU (Protein overexpression rate [PRO] = 27.5%, FDR = 2.3e-05), WNK2 (PRO = 27.4%, FDR = 3.2e-05), BMX (PRO = 26.2%, FDR = 8.5e-05), and CKM (PRO = 24.8%, FDR = 0.00021) (Figure 2a, 2b). We also identified

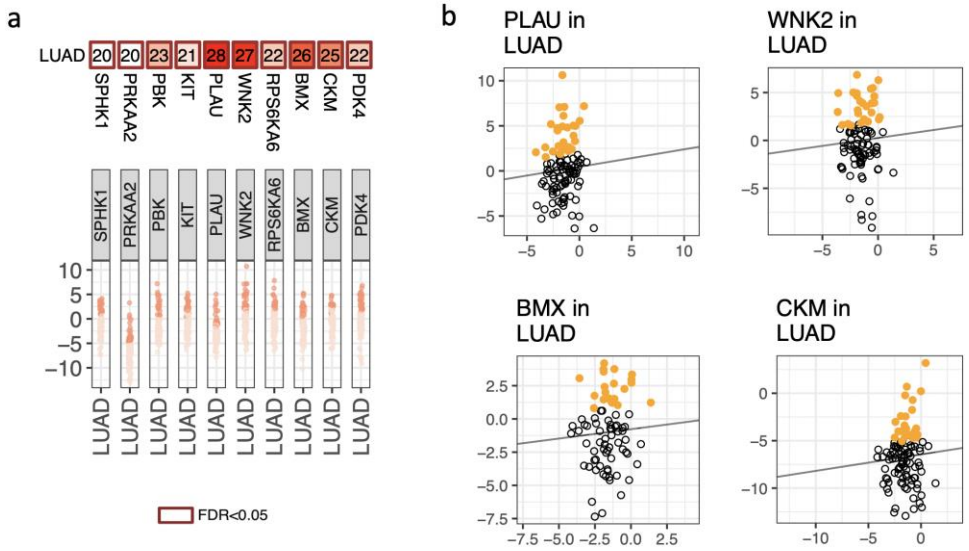

Figure 2. Overexpressed kinases in human LUAD tumors. (a) Druggable kinases exhibiting significant overexpression in LUAD cohort, detected by OPPTI. (b) Kinase overexpression at the sample level for the markers indicated in panel a, showing the observed expressions' (y-axis) deviation from their expected background values (x-axis), with the threshold value of overexpression not depicted

significant overexpression of 1,714 non-kinase proteins (OPPTI permutation test, FDR < 0.05), of which 379 are DGIdb-listed druggable genes, offering potential therapeutic targets in lung adenocarcinoma (LUAD), including CRP (PRO = 34.9%, FDR = < 1e-100), MMP12 (PRO = 34.9%, FDR = < 1e-100), MPO (PRO = 25.7%, FDR = 1e-4), and FCGR3B (PRO = 25.3%, FDR = 2e-4) (Figure S1a, S1b).

To enhance the confidence in identifying therapeutic targets based on expression, we compared differentially expressed proteins (DEPs) with overexpressed markers. 313 kinases were identified from DGIdb druggable genes, with 100 exhibiting elevated levels in both protein overexpression and differential expression (Table S1). Of these, 9 kinases demonstrated both notable differential expression (as assessed by the limma test with empirical Bayes moderation of t-statistics [13], FC ≥ 1, and FDR < 0.05) and significant overexpression (determined by the OPPTI permutation test with FDR < 0.05), including PLAU (FC = 3.5; PRO = 27.5%), MET (FC = 2.2; PRO = 29.4%), WNK2 (FC = 2.6; PRO = 27.4%), and STK17A (FC = 3.5; PRO = 25.7%). (Figure 3). The RAS pathway, represented by MET, ERBB2 and EGFR kinases, displayed the highest levels of dysregulation, while other prominent kinases included CHEK2 from the TP53 pathway (FC = 2.9; PRO = 17.4%) and STK3 from the HIPPO pathway (FC = 1; PRO = 18.3%).

Within the non-kinase proteins, we identified 1512 druggable proteins (DGIdb) quantified in this cohort, with 570 of them exhibiting elevated levels in both protein overexpression and differential expression (Figure S2) (Table S2). Notably, 122 (non-kinase) proteins were significantly differentially-expressed (FC ≥ 1, FDR < 0.05) and overexpressed (FDR < 0.05), including THBS2 (FC = 6.6; PRO = 33%), CRABP2 (FC = 5.8; PRO = 41.3%), and COL11A1 (FC = 4.4; PRO = 33.3%). CRABP2 has been identified as a promising target because its inhibition not only reduces metastasis and invasion in lung cancer but also enhances the effectiveness of chemotherapy [14]. Furthermore, several targets we identified through both differential expression (DEP) and overexpression (OPPTI) approaches have approved corresponding inhibitor drugs, including ERBB2 kinase (FC = 2.5; PRO = 16.5%) (Afatinib [15]) and VEGF-A protein (FC = 3.3; PRO = 26.6%) (Bevacizumab/Endostatin [16]). The effectiveness of treatment strategies involving the inhibition of other targets we identified remains to be validated.

3.4. Comparison between transcriptomic and protein-level aberrations

Elevated protein levels can arise due to genomic changes like copy-number amplifications, but they can also occur through post-transcriptional mechanisms, making it undetectable at RNA level. To explore this possibility, we systematically compared the prevalence of protein overexpression in patients with the occurrence of transcriptomic abnormalities. To compare protein overexpression with mRNA overexpression, we applied OPPTI to the RNA-seq data obtained from the same LUAD samples (Methods). This analysis revealed 24 proteins, exhibiting substantial overexpression (>10%) at both the mRNA and protein levels, including SFRP4 (RNA = 22.6%; PRO = 29.4%), TP53 (RNA = 19.1%; PRO = 29.2%), CDKN2A (RNA = 10.4%; PRO = 29.7%), IRS2 (RNA = 20.9%; PRO = 25.7%) (Figure 4a) (Table S3).

Furthermore, we identified 11 proteins that exhibited substantial up-regulation (≥10%) and had overexpression rates more than double those of their transcriptomic changes, including INPP4B (RNA = 13%; PRO = 33%) and THBS2 (RNA = 7%; PRO = 33%), and SFRP2 (RNA = 1.7%; PRO = 32%) (Figure 4b).

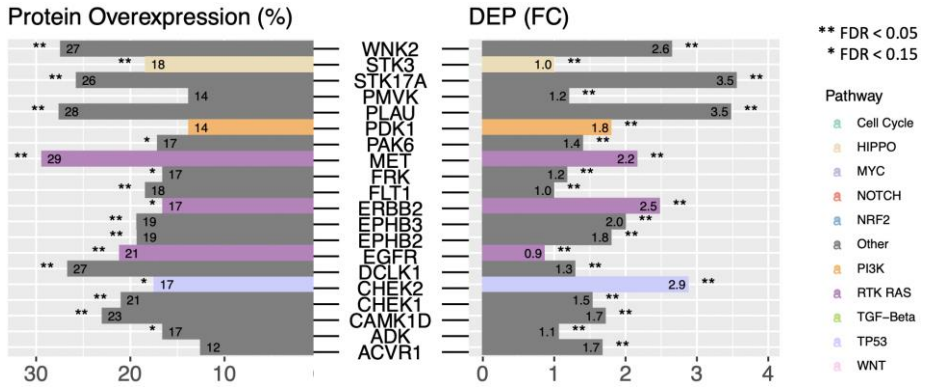

Figure 3. Potential LUAD kinase targets demonstrating protein overexpression and differential expression. Druggable kinases from DGIdb that exhibited substantially higher expression in tumors compared to normal tissues, as well as protein overexpression in the LUAD cohort

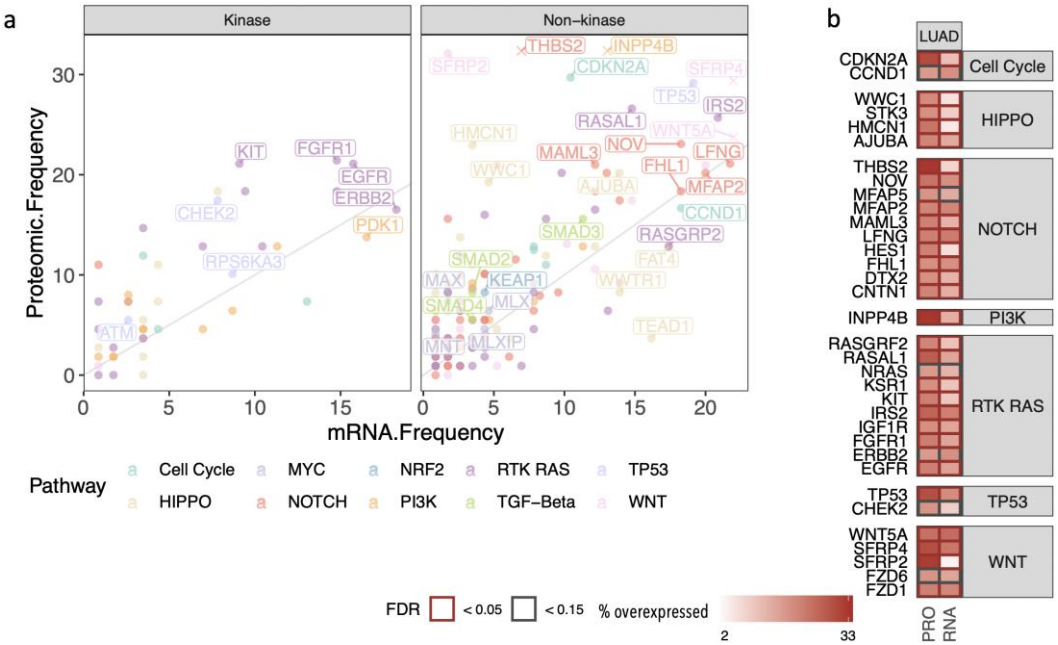

Figure 4. (a) Proportions of LUAD cases exhibiting protein and mRNA overexpression for genes involved in oncogenic signaling pathways. (b) Proteins from panel (a) that demonstrate significant enrichment for protein overexpression (FDR < 0.05) generally exhibit lower proportions of mRNA alterations

SFRP2 is a promising target for lung cancer because its overexpression inhibits the proliferation and metastasis of non-small cell lung cancer (NSCLC) cells by activating mitochondrial fission through the WNT signaling pathway, leading to reduced cell survival and increased apoptosis [17,18]. INPP4B (inositol polyphosphate 4-phosphatase type B) is also a potential therapeutic marker playing a dual role in both preventing tumor development by maintaining genome stability and inhibiting the PI3K-Akt-mTOR signaling pathway [19], and THBS2 is another prognostic marker in non-small cell lung cancer [20]. These findings underscore the limited presence of targets showing transcriptomic alterations in LUAD and illustrate that a proteomic approach could uniquely identify many overexpressed targets that show significant alterations at the protein level but are not easily detected at the mRNA level.

4. Conclusions

In this study, we conducted a comprehensive proteogenomic investigation of lung adenocarcinoma (LUAD) utilizing mass spectrometry (MS) data from 109 primary LUAD tumors and 102 corresponding normal samples. Our objective was to identify potential therapeutic targets that might not be evident through transcriptomic analyses alone (Figure 1). We highlighted multiple proteins with differential expression between tumor and normal tissues (Figure 1), including several that exhibited significant overexpression in tumors, such as ERBB2, EGFR, PDK1, PLAU, CRABP2 (Figure 2, 3, S2). Integration of mRNA and protein expressions allowed us to pinpoint numerous

proteins in key signaling pathways with no corresponding alterations at the transcriptomic level, such as INPP4B, THBS2, SFRP2 (Figure 4). This array of proteogenomic analyses has uncovered a list of important targets in LUAD (Table S1, S2, S3).

Among the differentially expressed proteins, we identified several kinases and non-kinase proteins involved in various oncogenic signaling pathways that exhibit significant up-regulation in LUAD tumors, including MET and EGFR kinases, which are already established targets for non-small-cell lung cancer [21]. Numerous proteins with significant overexpression did not exhibit corresponding changes at the mRNA level (ex. INPP4B), suggesting that post-transcriptional mechanisms, such as protein stabilization or altered protein degradation, may contribute to the observed protein-level abnormalities. Thus, integrating proteomics and transcriptomics data could be critical for understanding tumor biology and developing effective therapies, such as INPP4B protein [19]. Additionally, we discovered several kinases and non-kinase proteins with notable overexpression that are not currently targeted by existing therapies, suggesting opportunities for developing new drugs or repurposing existing ones to address these novel protein markers. For example, functional evidence highlights PLAU as a promising target in lung squamous cell carcinoma, given its crucial role in metastasis, and suggests its potential for early diagnosis and therapeutic intervention to inhibit disease progression [22]. PDK1 is another promising target because its upregulation in NSCLC promotes tumor growth and metastasis, making it a potential prognostic marker and therapeutic target for NSCLC treatment [23]. We also identified WNK2 overexpression as an important biomarker for combination therapy in lung cancer, as its suppression by CBX8 promotes invasion and migration, suggesting that restoring WNK2 activity may inhibit these processes and potentially limit metastasis. Nonetheless, the efficacy of such approaches will need to be validated through preclinical and clinical studies to establish their therapeutic potential.

In summary, our study offers an in-depth analysis of the proteogenomic landscape in LUAD, revealing numerous potential therapeutic targets that enhance the insights gained from existing genomic and transcriptomic data. By concentrating on protein-level alterations, we have identified targets that might be overlooked by conventional methods, setting the stage for future research into innovative therapeutic strategies. Subsequent studies should focus on validating the therapeutic potential of these identified targets and investigating their roles in LUAD pathogenesis. We plan to integrate *in vitro* screenings of anticancer compounds on human LUAD cell lines (or employ other functional data, e.g., CCLE [24]), and examine the relationship between targeted protein levels and the survival of cells following treatment, enabling the identification of "expression-driven" dependencies [25]. Additionally, integrating proteomic data with other omics approaches, such as metabolomics and epigenomics, could provide further insights into the complex molecular mechanisms underlying LUAD and lead to more effective and personalized treatment options.

Data and software availability

Data for LUAD cohort [10] can be found on The National Cancer Institute's Clinical Proteomic Tumor Analysis Consortium (CPTAC) resources: https://proteomic.datacommons.cancer.gov/pdc/study/PDC000153, and https://pdc.cancer.gov/pdc/.

The marker overexpression tool OPPTI can be accessed via https://github.com/Huang-lab/oppti. The analyses performed were scripted in R programming language (v3.6.2).

Acknowledgements

The author wishes to acknowledge LUAD cohort and its participating patients and families that generously contributed the data. The author would like to thank the members of the Huang lab (https://labs.icahn.mssm.edu/kuanhuanglab/members/) for their valuable discussions that contributed to the development of this work. The author has thoroughly edited and verified all final codes and texts. This work was supported by Giresun University funds to AE.

Conflict of interest

The author declares that there is no conflict of interest.

Similarity rate (iThenticate): 9%

References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, F. Bray, Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 71(3), 209–249, 2021. https://doi.org/10.3322/caac.21660.

[2] R. Ruiz-Cordero and W. P. Devine, Targeted therapy and checkpoint immunotherapy in lung cancer. Surgical pathology clinics, 13(1), 1, 17–33, 2020. https://doi.org/110.1016/j.path.2019.11.002.

[3] C. Swanton and R. Govindan, Clinical implications of genomic discoveries in lung cancer. New England Journal of Medicine, 374(19),19, 1864–73, 2016. https://doi.org/10.1056/NEJMra1504688.

[4] K.-L. Huang et al. Proteogenomic integration reveals therapeutic targets in breast cancer xenografts. Nature Communications, 8, 14864, 2017. https://doi.org/10.1038/ncomms14864

[5] K.-L. Huang et al., Spatially interacting phosphorylation sites and mutations in cancer. Nature Communications, 12(1), 1, 2313, 2021. https://doi.org/10.1038/s41467-021-22481-w.

[6] K.-L. Huang et al., Regulated phosphosignaling associated with breast cancer subtypes and druggability. Molecular & Cellular Proteomics, 18(8), 1630–1650, 2019. https://doi.org/10.1074/mcp.RA118.001243.

[7] M. J. Ellis et al., Connecting genomic alterations to cancer biology with proteomics: the NCI clinical proteomic tumor analysis consortium. Cancer Discovery, 3(10), 1108–12, 2013. https://doi.org/10.1158/2159-8290.CD-13-0219.

[8] K. V Ruggles et al., Methods, tools and current perspectives in proteogenomics. Molecular & Cellular Proteomics, 16(6), 959–981, 2017. https://doi.org/10.1074/mcp.MR117.000024.

[9] A. Elmas, S. Tharakan, S. Jaladanki, M. D. Galsky, T. Liu, and K.-L. Huang, Pan-cancer proteogenomic investigations identify post-transcriptional kinase targets. Communications Biology, 4(1), 1112, 2021. https://doi.org/10.1038/s42003-021-02636-7.

[10] M. A. Gillette et al., Proteogenomic characterization reveals therapeutic vulnerabilities in lung adenocarcinoma. Cell, 182(1), 200-225, 2020. https://doi.org/10.1016/j.cell.2020.06.013.

[11] K. C. Cotto et al., DGIdb 3.0: a redesign and expansion of the drug-gene interaction database. Nucleic Acids Research, 46(D1), D1068–D1073, 2018. https://doi.org/10.1093/nar/gkx1143.

[12] F. Sanchez-Vega et al., Oncogenic signaling pathways in the cancer genome atlas.," Cell, 173(2), 321-337.e10, 2018. https://doi.org/10.1016/j.cell.2018.03.035

[13] M. E. Ritchie et al., Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res, 43(7), e47, 2015. https://doi.org/10.1093/nar/gkv007.

[14] J. I. Wu, Y. P. Lin, C. W. Tseng, H. J. Chen, and L. H. Wang, Crabp2 Promotes Metastasis of Lung Cancer Cells via HuR and Integrin β1/FAK/ERK Signaling. Scientific Reports, 9(1), 2019. https://doi.org/10.1038/s41598-018-37443-4.

[15] A. C. Garrido-Castro and E. Felip, HER2 driven non-small cell lung cancer (NSCLC): potential therapeutic approaches. Translational Lung Cancer Research, 2(2), 122–7, 2013.https://doi.org/10.3978/j.issn.2218-6751.2013.02.02.

[16] B. Piperdi, A. Merla, and R. Perez-Soler, Targeting angiogenesis in squamous non-small cell lung cancer. Drugs, 74(4), 403–13, 2014. https://doi.org/10.1007/s40265-014-0182-z.

[17] P. Li, S. Zhao, and Y. Hu, SFRP2 modulates non-small cell lung cancer A549 cell apoptosis and metastasis by regulating mitochondrial fission via Wnt pathways. Molecular Medicine Reports, 20(2), 1925–1932, 2019. https://doi.org/10.3892/mmr.2019.10393.

[18] M. Montagner et al., Crosstalk with lung epithelial cells regulates Sfrp2-mediated latency in breast cancer dissemination. Nature Cell Biology, 22(3), 289–296, 2020.https://doi.org/10.1038/s41556-020-0474-3.

[19] Y. Sun et al., Loss of tumor suppressor inositol polyphosphate 4-phosphatase type B impairs DNA double-strand break repair by destabilization of DNA tethering protein Rad50. Cell Death & Disease, 11(4), 292, 2020. https://doi.org/10.1038/s41419-020-2491-3.

[20] Y. Bao, E. Yan, and N. Wang, Evaluation of GREM1 and THBS2 as prognostic markers in in non-small cell lung cancer. Journal of Cancer Research and Clinical Oncology, 149(10), 7849–7856, 2023. https://doi.org/10.1007/s00432-023-04746-7.

[21] N. Coleman, L. Hong, J. Zhang, J. Heymach, D. Hong, and X. Le, Beyond epidermal growth factor receptor: MET amplification as a general resistance driver to targeted therapy in oncogene-driven non-small-cell lung cancer. ESMO Open, 6(6), 100319, 2021. https://doi.org/10.1016/j.esmoop.2021.100319.

[22] P. Ning et al., PLAU plays a functional role in driving lung squamous cell carcinoma metastasis. Genes & Diseases, 11(2), 554–557, 2024. https://doi.org/10.1016/j.gendis.2023.04.010.

[23] T. Liu and H. Yin, PDK1 promotes tumor cell proliferation and migration by enhancing the Warburg effect in non-small cell lung cancer. Oncology Reports, 37(1), 193–200, 2017. https://doi.org/10.3892/or.2016.5253.

[24] J. Barretina et al., The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. Nature, 483(7391), 603–7, 2012. https://doi.org/10.1038/nature11003.

[25] S. K. Jaladanki, A. Elmas, G. S. Malave, and K.-L. Huang, "Genetic dependency of Alzheimer's disease-associated genes across cells and tissue types. Scientific Reports, 11(1), 12107, 2021. https://doi.org/10.1038/s41598-021-91713-2.

Supplementary Tables

Table S1. DGIdb druggable (kinase) proteins exhibiting elevated levels in both protein overexpression and differential expression
https://www.columbia.edu/~ae2321/workspace/LUAD/TableS1.xlsx

Table S2. DGIdb druggable (non-kinase) proteins quantified in this cohort, with 570 demonstrating elevated levels in both protein overexpression and differential expression
https://www.columbia.edu/~ae2321/workspace/LUAD/TableS2.xlsx

Table S3. Proteins exhibiting substantial overexpression (>10%) at either mRNA or protein levels (or both levels)
https://www.columbia.edu/~ae2321/workspace/LUAD/TableS3.xlsx

Supplementary Figures

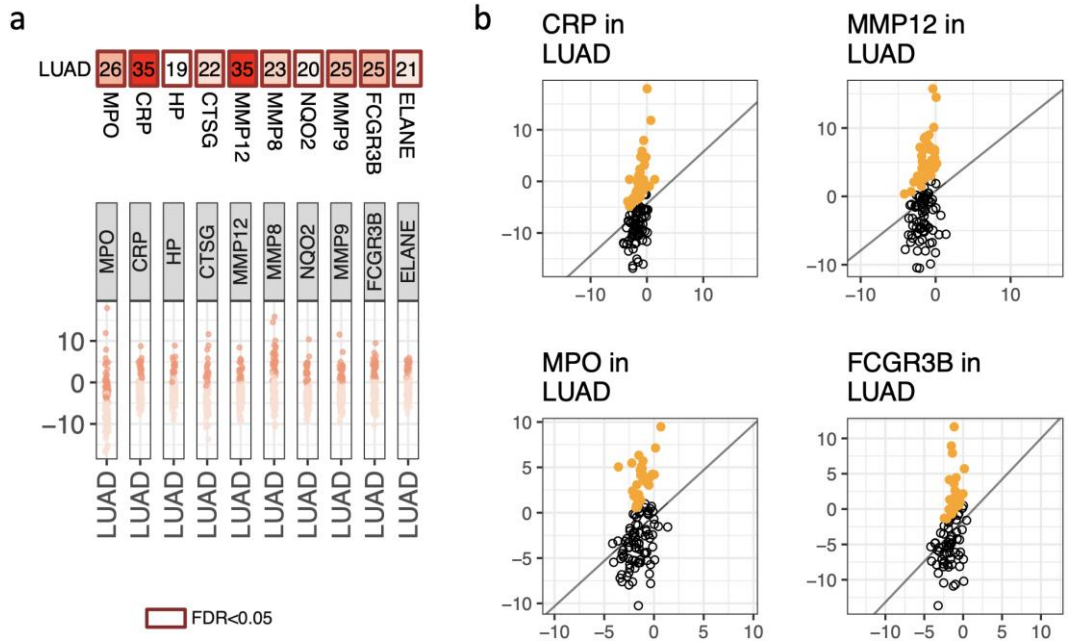

Figure S1. Non-kinase proteins overexpressed in human LUAD tumors. (a) The top ten druggable non-kinase proteins with the most pronounced overexpression in the LUAD tumors. (b) Marker overexpressions listed in panel a displayed at sample level.

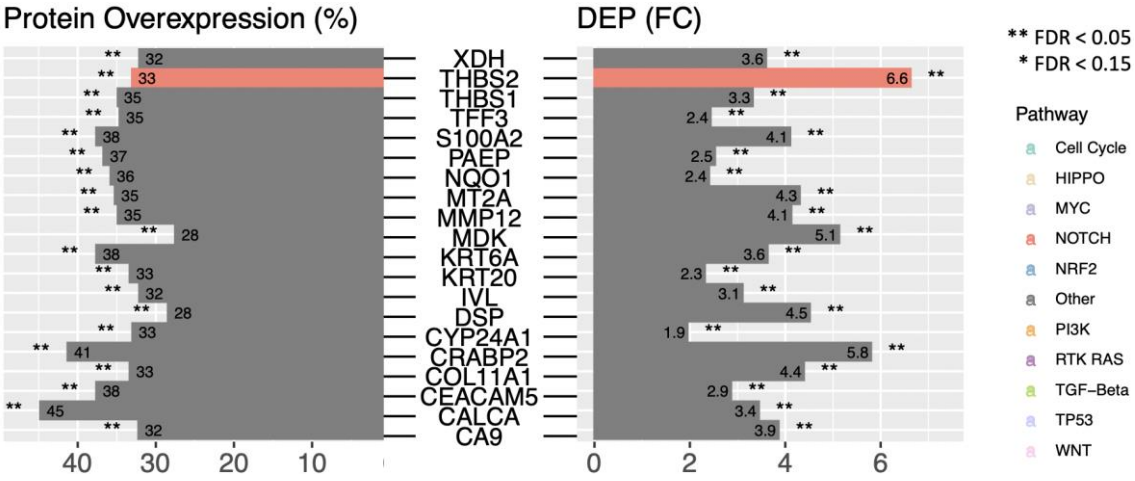

Figure S2. Non-kinase proteins that can be targeted by drugs, along with their associated compounds, demonstrated significantly elevated expression in tumor samples compared to normal tissues and exhibited protein overexpression within the LUAD cohort.