Original Article

The causal effects of 2,821 protein level ratios on non-small cell lung cancer: a two-sample Mendelian randomization study

Xinyun Zou[1], Jinlan Shen[2], Xiaokai Li[1], Yong Diao[1], Ling Zhang[1]

[1]Department of Oncology, The General Hospital of Western Theater Command, Chengdu, China; [2]Department of Laboratory Medicine, The General Hospital of Western Theater Command, Chengdu, China

Contributions: (I) Conception and design: X Zou; (II) Administrative support: L Zhang; (III) Provision of study materials or patients: X Zou, J Shen; (IV) Collection and assembly of data: X Li; (V) Data analysis and interpretation: Y Diao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Ling Zhang, PhD. Department of Oncology, The General Hospital of Western Theater Command, No. 270 Rongdu Avenue, Jinniu District, Chengdu 610083, China. Email: zhangling71@163.com.

Background: Non-small cell lung cancer (NSCLC) has a complex etiology, making early diagnosis difficult and leading to high mortality rates, thus necessitating personalized treatment strategies. While protein level ratios have shown potential as biomarkers or therapeutic targets, their causal relationship with NSCLC remains unclear. This study aimed to investigate these causal links using Mendelian randomization (MR), providing insights into potential biomarkers and therapeutic avenues.

Methods: We executed an intricate two-sample MR study to explore the stochastic causal links between 2,821 protein level ratios and NSCLC. The genome-wide association study (GWAS) statistics for NSCLC and protein level ratios were sourced from the Finnish Database (version 10) and the UK Biobank, respectively. For the instrumental variables (IVs) related to protein level ratios, we selected IVs with a P value $<1.0\times10^{-5}$. Throughout this analysis, we applied five established MR techniques.

Results: Our study identified causal relationships between 142 protein level ratios and NSCLC. Notably, the AKR1B1/SUGT1 protein level ratio and the PLPBP/STIP1 protein level ratio demonstrated the most significant negative correlations with NSCLC risk. On the other hand, the ARHGEF12/IRAK4 protein level ratio and the BANK1/LBR protein level ratio exhibited the most significant positive correlations. Furthermore, sensitivity analyses did not reveal any significant heterogeneity or horizontal pleiotropy.

Conclusions: Studying specific protein level ratios in patients can reveal the molecular mechanisms and pathological processes of NSCLC, which has certain clinical significance for early diagnosis of NSCLC, understanding drug resistance mechanisms and developing personalized treatment strategies. However, these findings necessitate further validation through extensive clinical research.

Keywords: Protein level ratios; non-small cell lung cancer (NSCLC); Mendelian randomization (MR); causal inference; sensitivity

Submitted Aug 26, 2024. Accepted for publication Dec 17, 2024. Published online Jan 08, 2025.
doi: 10.21037/tcr-24-1523
View this article at: https://dx.doi.org/10.21037/tcr-24-1523

Introduction

Non-small cell lung cancer (NSCLC) significantly threatens global health with its high mortality and low survival rates. In 2024, worldwide lung cancer accounted for 234,580 new cases, or 12.4% of all new cancer diagnoses, and 125,070 deaths, which constituted 18.7% of all cancer-related deaths (1-3). NSCLC, the most common lung cancer type, comprises approximately 85% of all cases, often diagnosed at advanced stages due to subtle early symptoms (3,4).

NSCLC's etiology is multifaceted, shaped by genetic predispositions, environmental factors, and lifestyle choices, complicating early diagnosis (5). Chemotherapy resistance

is common among NSCLC patients, posing additional challenges to treatment (6). The significant variation in genetic background (7), molecular tumor characteristics (8), and clinical manifestations among patients requires personalized treatment approaches, presenting substantial challenges. Despite extensive research, the pathogenesis and chemotherapy resistance mechanisms of NSCLC (6), which involve complex regulatory processes, are still not fully understood (9). This significant knowledge gap highlights the urgency of continued research to devise more effective diagnostic and treatment strategies, aiming to enhance patient outcomes.

In biomedical research, scientists employ methods like mass spectrometry and other quantitative techniques to measure protein levels in diverse samples (10). Calculating the relative abundance of specific proteins allows researchers to establish protein level ratios. These data are crucial for identifying potential biomarkers that uncover diseases' molecular mechanisms and pathological processes (11). These insights are vital for early diagnosis, prognosis evaluation, and treatment efficacy monitoring.

NSCLC is governed by complex regulatory mechanisms (12). Leveraging protein level ratios enhances our grasp of NSCLC pathophysiology, paving the way for more effective and personalized treatments that improve patient outcomes. Studies have shown that elevated baseline levels of lactate dehydrogenase (LDH) and low levels of C-reactive protein (CRP) are associated with poor progression-free survival (PFS) in patients with advanced NSCLC (13). Additionally, a high baseline neutrophil-to-lymphocyte ratio (NLR) has been identified as a predictor of worse prognosis (14). Research also indicated that elevated baseline CRP levels are not only linked to a higher incidence of immune-related adverse events (irAEs) but are also associated with poorer overall outcomes (15). Furthermore, a retrospective study revealed that histone deacetylase 11 (HDAC11) is significantly overexpressed in NSCLC tumor tissues compared to adjacent non-cancerous tissues, highlighting its role as a critical prognostic biomarker (16). Numerous studies have explored the relationship between protein levels and the etiology of NSCLC, identifying several biomarkers with potential diagnostic, prognostic, or therapeutic implications (17,18). However, the causal relationship between specific protein level ratios and NSCLC remains controversial. These discrepancies arise from several factors, including confounding variables, reverse causality, and limitations of observational study designs. For instance, protein expression can be influenced by underlying inflammation, tumor microenvironment changes, or secondary effects of malignancy, making it challenging to discern whether alterations in protein levels drive cancer progression or merely reflect its presence (19).

Mendelian randomization (MR) offers a robust method to address these challenges by leveraging genetic variants as instrumental variables (IVs) to infer causality (20). Unlike traditional epidemiological approaches, MR minimizes confounding and is less susceptible to reverse causation. Given the limitations of prior studies, applying MR to investigate the causal relationship between protein level ratios and NSCLC could provide novel insights into the underlying biological mechanisms, identify potential therapeutic targets, and refine biomarker strategies for risk stratification and early intervention. We present this article in accordance with the STROBE-MR reporting checklist (available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-1523/rc).

Highlight box

Key findings

- This study identified causal relationships between 142 protein level ratios and non-small cell lung cancer (NSCLC). Notably, the AKR1B1/SUGT1 and PLPBP/STIP1 protein level ratios were most strongly negatively associated with NSCLC risk, while the ARHGEF12/IRAK4 and BANK1/LBR ratios exhibited the strongest positive correlations with the disease.

What is known and what is new?

- NSCLC presents with a complex etiology, making early diagnosis difficult and contributing to high mortality rates. Previous research has suggested a potential link between protein level ratios and NSCLC, but the causal mechanisms remain unclear.
- This manuscript presented the first large-scale Mendelian randomization analysis of 2,821 protein level ratios, identifying significant causal relationships between specific protein ratios and NSCLC risk. This study offered new insights into the molecular mechanisms of NSCLC, which may be useful for early diagnosis and personalized treatment strategies.

What is the implication, and what should change now?

- The findings suggested that specific protein level ratios could serve as biomarkers for NSCLC risk, with potential applications in early diagnosis and treatment personalization. These results warranted further clinical validation to confirm their applicability in diverse populations. Future studies should explore the development of diagnostic tools and targeted therapies based on these protein level ratios.

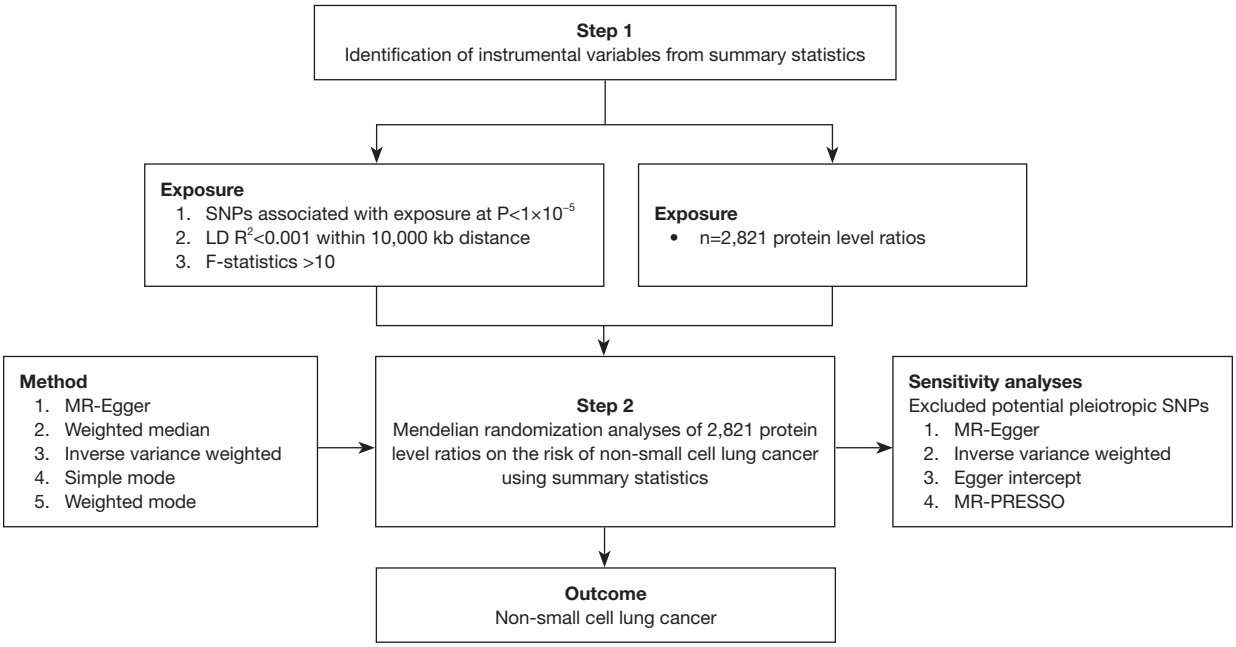

Figure 1 The study design of the associations of protein level ratios and NSCLC. SNP, single nucleotide polymorphism; LD, linkage disequilibrium; MR, Mendelian randomization; PRESSO, pleiotropy residual sum and outlier; NSCLC, non-small cell lung cancer.

Methods

Study design

This study uses genetic variants as IVs for MR analysis. The selection of these IVs and the entire research process are illustrated in *Figure 1*. MR analysis relies on three key assumptions (21): (I) relevance: genetic variants must be strongly associated with the exposure. (II) Independence: variants should be free from confounding variables. (III) Exclusion: variants must influence the outcome solely through the exposure. This study employs a two-sample MR design, utilizing distinct datasets for protein level ratios (exposure) and NSCLC (outcome).

Genome-wide association study (GWAS) data sources for NSCLC

The GWAS statistical data for NSCLC were sourced from version 10 of the Finnish database (22), which integrates genetic and health registry data from 224,737 Finnish participants. Detailed information and the download link can be found in *Table 1*. NSCLC cases in the FinnGen database are identified by International Classification of Diseases (ICD) codes, diagnoses are based on pathological analysis and clinical assessment, and tumor classification is performed using the tumor-node-metastasis (TNM) classification system. Healthy controls for the FinnGen dataset were participants without any recorded cancer diagnoses or significant chronic diseases, as defined by exclusion criteria in their health registry data. In the FinnGen cohort, the clinical characteristics of NSCLC patients included age, sex, smoking history, tumor stage, and histological subtype. Most NSCLC cases were diagnosed at advanced stages, reflecting the typical clinical presentation.

Protein level ratios GWAS data sources

The GWAS data for protein level ratios were sourced from the UK Biobank (23); details and the download address are listed in *Table 1*. Utilizing Olink proteomics, data for 1,463 proteins across more than 54,000 UK Biobank samples revealed 4,248 associations involving 2,821 ratios between protein levels [ratio quantitative trait locis (rQTLs)]. For the UK Biobank, controls were selected as participants without diagnosed lung cancer or other malignancies at baseline. The UK Biobank cohort included participants aged 40–69 years, with detailed demographic, clinical, and lifestyle data. Protein levels were quantified by Olink's proximity extension assay (PEA), and protein level ratios were calculated by comparing the relative abundance of

Table 1 Details of the GWASs and datasets used in our analyses

Exposure or outcome	Sample size	Ancestry	Links for data download	PMID
NSCLC	314,193 controls and 5,315 NSCLC cases	FinnGen Project	https://storage.googleapis.com/finngen-public-data-r10/summary_stats/finngen_R10_C3_LUNG_NONSMALL_EXALLC.gz	36653562
Protein level ratios	2,821 phenotypes	UK Biobank	http://ftp.ebi.ac.uk/pub/databases/gwas/summary_statistics/; GCST90313126–GCST90315946	38412862

GWAS, genome-wide association study; NSCLC, non-small cell lung cancer.

specific proteins.

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). It is important to note that the data used in this study were derived exclusively from the GWAS summary level. Ethical approval and informed consent were obtained for the use of this data during the initial investigations.

Selection of IVs

To ensure high-quality results, we implemented rigorous quality control procedures to screen type IV genes aligned with MR assumptions. For IVs with protein level ratios, we selected IVs with P values less than 1.0×10^{-5}. Our methodology included using a linkage disequilibrium clustering algorithm with $R^2<0.001$ and a window size of 10,000 kb. To ensure consistency of effector alleles, we harmonized exposure and outcome datasets by removing single nucleotide polymorphisms (SNPs) with intermediate allele frequencies and those with inconsistent alleles. We calculated the F-statistic for each SNP using: $F = (R^2 \times (n - k - 1)) / (k \times (1 - R^2))$, where R^2 represents the proportion of variations explained by the IV (24). Given that the F-statistic for all SNPs exceeded 10, indicating strong instruments, only SNPs passing strict screening will be used as IVs in future analyses. N denotes the sample size, and K is the number of SNPs analyzed. We utilized the PhenoScanner website to check for potential confounders linked to the SNPs and found none. The SNPs that have undergone rigorous statistical analysis and significance threshold screening are shown in table available at https://cdn.amegroups.cn/static/public/tcr-24-1523-1.xlsx.

Statistical analysis

All analyses were conducted using R software, version 4.3.0, available at R project. To assess causal relationships between 2,821 protein level ratios and NSCLC, we employed five methods from the 'TwoSampleMR' package: inverse variance weighting (IVW), weighted median, simple mode, MR-Egger, and weighted mode. For all statistical analyses, P values less than 0.05 were considered statistically significant. We used Cochran's Q statistic (25) and corresponding p-values to assess heterogeneity among the IVs. The MR-Egger method (26) was used to detect horizontal pleiotropy, signaled by a statistically significant intercept term. Furthermore, the MR pleiotropy residual sum and outlier (MR-PRESSO) method (27) from the same package identified and excluded potential horizontal pleiotropic outliers that could significantly affect the results. Scatter plots and funnel plots were generated to validate the findings. The scatter plots confirmed no impact from outliers, and the funnel plots highlighted the robustness of the correlations and the uniformity of data (28).

Results

In this study, we employed the two-sample MR analysis to examine the random causal relationships between 2,821 protein level ratios and NSCLC, exploring the potential influence of these ratios on the development of the cancer. We utilized five analytical methods: MR-Egger, weighted median, inverse variance weighted, simple mode, and weighted mode. After comprehensive statistical analysis and finding P values less than 0.05 via the IVW mainstream statistical method, we identified 142 Protein level ratios that significantly affected NSCLC (details in table available at https://cdn.amegroups.cn/static/public/tcr-24-1523-2.xlsx). Each of the SNP's phenotypes associated with these 142 protein level ratios had an F-value greater than 10, indicating that the IVs were strongly valid (details in table available at https://cdn.amegroups.cn/static/public/tcr-24-1523-1.xlsx). Subsequently, we ranked the OR values from the IVW analysis, selecting 10 protein level ratios each with positive and negative correlations to NSCLC, as shown in *Figure 2*. The AKR1B1/SUGT1 protein level ratio [odds ratio (OR)

Trait	N.SNPs	F (mean)	OR (95% CI)	P value	Heterogeneity.P	Pleiotropy.P	Cochran's Q test (I^2)
AKR1B1/SUGT1 protein level ratio	14	32.723	0.6849 (0.4898–0.9577)	0.0269	0.220	0.406	16.558 (21.49%)
PLPBP/STIP1 protein level ratio	14	41.127	0.7623 (0.6005–0.9679)	0.0259	0.514	0.478	12.167 (6.84%)
CERT/FHIT protein level ratio	11	68.263	0.7644 (0.6204–0.9418)	0.0116	0.751	0.317	6.731 (48.56%)
ITGAV/ITGB5 protein level ratio	21	49.782	0.7681 (0.6215–0.9492)	0.0146	0.242	0.812	24.004 (16.68%)
PPME1/TBC1D23 protein level ratio	10	80.852	0.7838 (0.6412–0.9579)	0.0174	0.527	0.624	8.069 (11.54%)
STX4/TBC1D23 protein level ratio	15	92.835	0.7955 (0.6832–0.9262)	0.0032	0.520	0.820	13.077 (7.06%)
CSF1/CXCL10 protein level ratio	37	36.736	0.8098 (0.6891–0.9518)	0.0105	0.488	0.554	35.578 (1.18%)
GLOD4/MIF protein level ratio	25	62.095	0.8134 (0.6911–0.9574)	0.0130	0.218	0.968	29.062 (17.42%)
RNF41/SH2B3 protein level ratio	22	54.864	0.8178 (0.6924–0.9659)	0.0179	0.994	0.409	8.183 (156.61%)
FGF21/LEP protein level ratio	27	70.891	0.8213 (0.6990–0.9650)	0.0167	0.329	0.722	28.618 (9.15%)
ARHGEF12/RAK4 protein level ratio	14	29.479	1.4564 (1.0408–2.0379)	0.0283	0.073	0.847	20.986 (38.05%)
BANK1/LBR protein level ratio	12	36.499	1.4306 (1.0529–1.9439)	0.0221	0.155	0.801	15.637 (29.66%)
FHIT/GLOD4 protein level ratio	15	60.415	1.3299 (1.0810–1.6361)	0.0070	0.222	0.833	17.674 (20.79%)
BANK1/RAK4 protein level ratio	20	46.724	1.3029 (1.0634–1.5962)	0.0107	0.075	0.464	28.447 (33.21%
PTPN6/TXNDC5 protein level ratio	16	47.549	1.2869 (1.0478–1.5805)	0.0161	0.418	0.494	15.467 (3.02%)
RWDD1/SULT1A1 protein level ratio	10	111.181	1.2818 (1.0784–1.5235)	0.0049	0.801	0.732	5.368 (67.66%)
NPDC1/PTGDS protein level ratio	23	53.768	1.2590 (1.0640–1.4897)	0.0073	0.418	0.547	22.721 (3.17%)
ELOA/POLR2F protein level ratio	18	47.442	1.2570 (1.0347–1.5270)	0.0212	0.647	0.443	14.285(19.01%)
CIAPIN1/PLPBP protein level ratio	13	49.039	1.2559 (1.0027–1.5730)	0.0473	0.893	0.675	6.427 (86.71%)
BANK1/TXLNA protein level ratio	16	67.784	1.2526 (1.0820–1.4500)	0.0026	0.672	0.667	12.086 (24.11%)

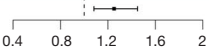

Figure 2 Forest plots showed the exploration of the causal effect of protein level ratios onset on NSCLC by using MR-Egger, weighted median and IVW. SNP, single nucleotide polymorphism; OR, odds ratio; CI, confidence interval; NSCLC, non-small cell lung cancer; MR, Mendelian randomization; IVW, inverse variance weighting.

=0.685; 95% confidence interval (CI): 0.490–0.958; P=0.03] and the PLPBP/STIP1 protein level ratio (OR =0.762; 95% CI: 0.601–0.968; P=0.03) exhibited the strongest negative correlations with NSCLC risk. Conversely, the ARHGEF12/IRAK4 protein level ratio (OR =1.456; 95% CI: 1.041–2.038; P=0.03) and the BANK1/LBR protein level ratio (OR =1.431; 95% CI: 1.053–1.944; P=0.02) showed the strongest positive correlations. These findings were consistent across all five analytical methods, as illustrated in the SNP scatter plot in *Figure 3*. Figures S1-S4 demonstrated the stability of the ratios of the four protein levels through forest plots, scatter plots, and funnel plots.

The MR-Egger regression intercept method and MR-PRESSO analysis showed no horizontal pleiotropy in the MR study (P>0.05; table available at https://cdn.amegroups.cn/static/public/tcr-24-1523-2.xlsx for details). Cochran's Q test, MR-Egger, and inverse variance weighted methods were used to test the heterogeneity of SNPs in the analysis results (table available at https://cdn.amegroups.cn/static/public/tcr-24-1523-2.xlsx for details). The results of the "leave-one-out" analysis demonstrated the reliability of the MR analysis (the null line was not within the total CI of the SNPs; table available at https://cdn.amegroups.cn/static/public/tcr-24-1523-3.xlsx for details). The above sensitivity analyses all demonstrated the stability of our results.

Discussion

AKR1B1 is an enzyme belonging to the aldo-keto reductase superfamily, it catalyzes the reduction of aldehydes and ketones and is expressed in various tissues (29). AKR1B1 has been found to be abnormally expressed in many tumor types, playing a significant role in cancer initiation, progression, and drug resistance (30). Study has found that AKR1B1 helps breast cancer cells survive under oxidative stress by regulating cellular redox status and antioxidant defense mechanisms (31). Research indicated that inhibiting AKR1B1 activity can reduce the growth and invasion capabilities of colon cancer cells (32). SUGT1 is a highly conserved nuclear protein involved in spindle function and essential for the G1/S and G2/M transitions of the cell cycle, it plays a crucial role in cell cycle transitions and muscle regeneration (33). Studies show that SUGT1 expression is typically higher in colorectal cancer tissues than in non-tumor tissues, and its overexpression is associated with higher recurrence rates and poorer prognosis (34). Current research data do not directly address the role or significance of SUGT1 in NSCLC. Our research indicated that the AKR1B1/SUGT1 protein level ratio was negatively correlated with the risk of NSCLC, suggesting that a decrease in SUGT1 protein levels relative to AKR1B1 protein levels may reduce the risk of NSCLC. This finding provides a new perspective for the diagnosis and treatment of NSCLC.

PLPBP is a protein that binds pyridoxal phosphate (PLP). PLP, the active form of vitamin B6, is involved in various enzyme-catalyzed reactions, affecting the metabolism of amino acids and keto acids, which are crucial for maintaining metabolic balance in the body (35). Researchers

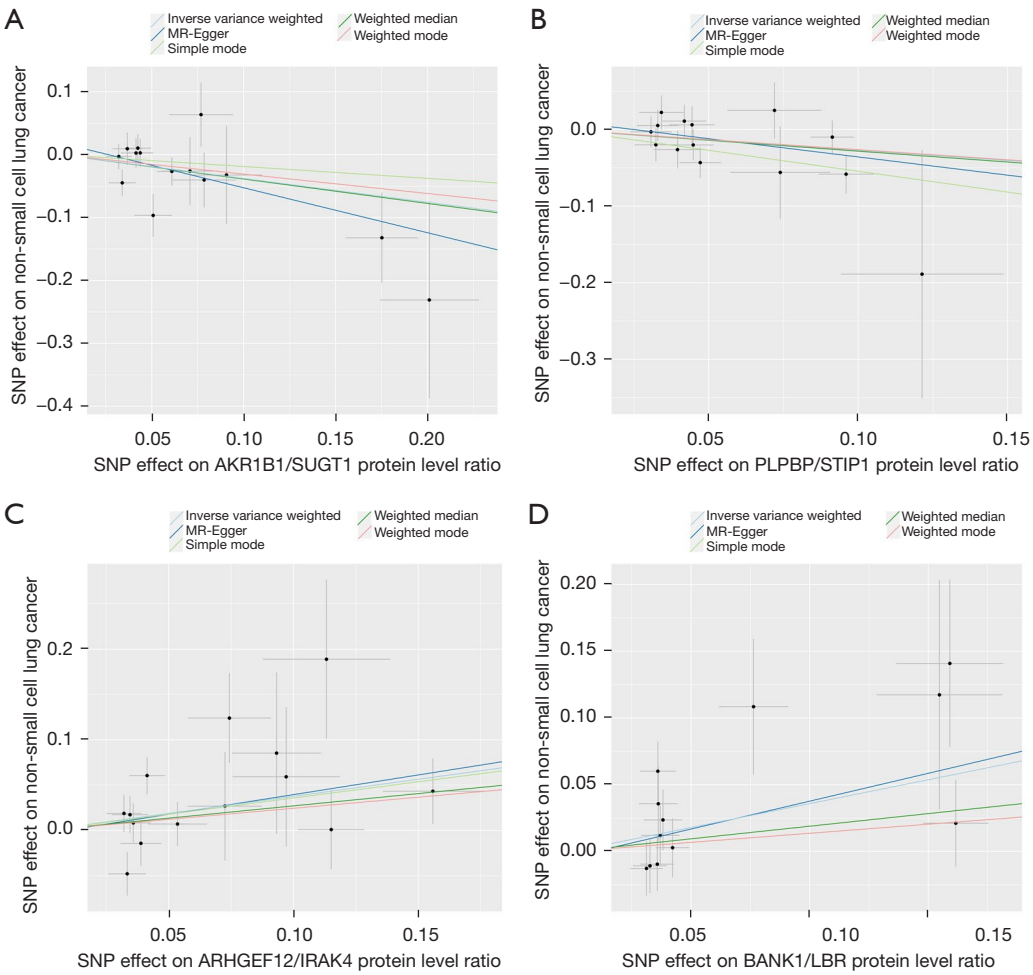

Figure 3 Scatter plot of the associations of genetic variants with four protein level ratios and the risk of NSCLC. (A) AKR1B1/SUGT1. (B) PLPBP/STIP1. (C) ARHGEF12/IRAK4. (D) BANK1/LBR. MR, Mendelian randomization; SNP, single nucleotide polymorphism; NSCLC, non-small cell lung cancer.

have discovered that highly active PLP can directly transfer from donors to receptors, a process involved in numerous reactions related to cancerous changes, including those associated with disease prognosis (35,36). According to The Human Protein Atlas, data on PLPBP expression in cancer reveal that PLPBP is expressed in the cytoplasm across all tissues, with variations in expression patterns depending on the type and stage of cancer (37). STIP1 protein is an adaptor protein that coordinates the functions of molecular chaperones HSP70 and HSP90 during protein folding, aiding in the proper folding and stability of proteins (38). STIP1 is involved in signaling pathways related to cell growth and death, supporting tumor cell survival, enhancing migration and invasion capabilities, and participating in resistance mechanisms to anticancer drugs (39). Research has shown that STIP1 promotes gastric cancer metastasis via the Wnt/β-catenin signaling pathway (40). Other studies have identified STIP1's role in driving tumor cell metastasis through various biological processes and pathways, including potential regulation of lung adenocarcinoma proliferation and migration via the JAK2/STAT3 pathway (41). Currently, there are no reports on the PLPBP/STIP1 protein level ratio in tumors. We proposed that this ratio is negatively correlated with the risk of NSCLC, suggesting that higher PLPBP expression relative to STIP1 may reduce the risk of NSCLC. This hypothesis warranted further experimental and clinical research.

ARHGEF12 plays a crucial role in cellular signal transduction. It primarily functions by binding to G proteins and activating RhoA GTPase, thus regulating cell

morphology and mobility (42). Expressed in various tissues including the brain, liver, and lungs, ARHGEF12 is also associated with diseases such as acute myeloid leukemia (AML) and large cell neuroendocrine carcinoma of the lung (43). A study suggested that targeting ARHGEF12 can promote differentiation in neuroblastoma, reduce MYCN protein levels, and decrease tumor malignancy (44). IRAK4 is a serine/threonine protein kinase that plays a crucial role in initiating innate immune responses, it signals through Toll-like receptors (TLRs) and IL-1 receptor pathways, aiding in pathogen recognition and the initiation of inflammatory responses (45). Deficiencies in IRAK4 can lead to immune deficiencies, increasing susceptibility to infections (46). Research indicated that IRAK4 often participates in pathways activated by the MYD88-L265P mutation, common in activated B-cell-like diffuse large B-cell lymphoma (ABC-DLBCL) and marginal zone lymphoma (47). Inhibitors targeting IRAK4, such as emavusertib (CA-4948), have shown potential in preclinical models of ABC-DLBCL (47). Studies have shown that IRAK4's involvement in the TLR pathway plays a role in the abnormal immune signaling observed in myelodysplastic syndromes (MDS) and AML (48). Activation of these pathways can affect disease progression and treatment responses. Experiments involving mice exposed to smoke via a smoke machine suggested that carcinogens in smoke lead to increased IL-1β and overexpression of IRAK4 in lung tissue (49). IRAK4 inhibitors have been shown to reduce the invasiveness of lung cancer cell lines, indicating IRAK4 as a potential target for lung cancer treatment. Both ARHGEF12 and IRAK4 have been studied in lung cancer and other tumors. Our research showed that an increase in the ARHGEF12/IRAK4 protein level ratio is associated with a higher risk of developing NSCLC, suggesting this ratio as a potential new target for NSCLC.

BANK1 is a protein primarily expressed in B cells. It plays a critical role in B cell receptor signaling, CD40-related signaling, TLR signaling, and the release of calcium ions from intracellular stores triggered by B cell receptor activation (50). Mutations in the *BANK1* gene are associated with diseases such as systemic lupus erythematosus (51), underscoring the protein's crucial role in immune regulation and disease development. Some research has explored the tumor-suppressive role of BANK1 in precursor B-cell acute lymphoblastic leukemia (52), although overall, its study in other tumors remains limited. Lamin B receptor (LBR) protein is located on the inner side of the nuclear membrane and anchors nuclear lamina and heterochromatin, helping to maintain the structural integrity of the nucleus, Additionally, LBR protein possesses Δ14-sterol reductase activity, participating in the biosynthesis of cholesterol. Mutations in the *LBR* gene are associated with genetic disorders such as Greenberg skeletal dysplasia and Pelger-Huet anomaly (53). Therefore, LBR protein plays a crucial role in maintaining nuclear structure, regulating gene expression, and its association with certain genetic disorders. Research indicated that the absence of LBR in colorectal cancer cells led to chromosomal instability and tumorigenesis (54). Currently, there is no research linking LBR to the development of NSCLC. Our research showed that the BANK1/LBR protein level ratio is positively correlated with the risk of developing NSCLC, potentially providing a new target for diagnosis and treatment.

Our MR analysis indicated that 142 protein level ratios influence NSCLC. Due to space limitations in the paper, we only discussed the four most impactful ratios, leaving the data on other ratios available for further research by other scholars.

Conclusions

In conclusion, this two-sample MR study investigated the impact of protein level ratios on NSCLC genetically, identifying 142 protein level ratios that significantly influence NSCLC. This finding has not been reported in previous studies. From a genomics perspective, these findings offer a new perspective on potential therapeutic targets for NSCLC and support personalized treatment strategies for the disease.

Acknowledgments

We thank the FinnGen research project for developing and curating their data resources of GWAS and we also thank Karsten Suhre for providing protein level ratios GWAS Data Sources for our analyses. We thank all of the participants in this study for their contributions.

Footnote

Reporting Checklist: The authors have completed the STROBE-MR reporting checklist. Available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-1523/rc

Peer Review File: Available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-1523/prf

Funding: This study was supported by grants from the People's Liberation Army General Hospital of Western Theater Command.

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://tcr.amegroups.com/article/view/10.21037/tcr-24-1523/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). It is important to note that the data used in this study were derived exclusively from the GWAS summary level. Ethical approval and informed consent were obtained for the use of this data during the initial investigations.

References

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2024;74:229-63.
2. World Cancer Research Fund International. Global lung cancer incidence and mortality statistics. 2022. Available online: https://www.wcrf.org/cancer-trends/lung-cancer-statistics/
3. Riely GJ, Wood DE, Ettinger DS, et al. Non-Small Cell Lung Cancer, Version 4.2024, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw 2024;22:249-74.
4. American Cancer Society. Lung Cancer Statistics. 2024. Available online: https://www.cancer.org/cancer/types/lung-cancer/about/key-statistics.html
5. González-Ruíz J, A Baccarelli A, Cantu-de-Leon D, et al. Air Pollution and Lung Cancer: Contributions of Extracellular Vesicles as Pathogenic Mechanisms and Clinical Utility. Curr Environ Health Rep 2023;10:478-89.
6. Min HY, Lee HY. Mechanisms of resistance to chemotherapy in non-small cell lung cancer. Arch Pharm Res 2021;44:146-64.
7. Adib E, Nassar AH, Abou Alaiwi S, et al. Variation in targetable genomic alterations in non-small cell lung cancer by genetic ancestry, sex, smoking history, and histology. Genome Med 2022;14:39.
8. Shen F, Guo W, Song X, et al. Molecular profiling and prognostic biomarkers in chinese non-small cell lung cancer cohort. Diagn Pathol 2023;18:71.
9. Zhang Z, Westover D, Tang Z, et al. Wnt/β-catenin signaling in the development and therapeutic resistance of non-small cell lung cancer. J Transl Med 2024;22:565.
10. Birhanu AG. Mass spectrometry-based proteomics as an emerging tool in clinical laboratories. Clin Proteomics 2023;20:32.
11. Sinha A, Mann M. A beginner's guide to mass spectrometry-based proteomics. The Biochemist 2020;42:64-9.
12. Zhang X, Chen R, Huo Z, et al. Blood-based molecular and cellular biomarkers of early response to neoadjuvant PD-1 blockade in patients with non-small cell lung cancer. Cancer Cell Int 2024;24:225.
13. Wei Y, Xu J, Huang X, et al. C-reactive protein and lactate dehydrogenase serum levels potentially predict the response to checkpoint inhibitors in patients with advanced non-small cell lung cancer. J Thorac Dis 2023;15:1892-900.
14. Sung M, Jang WS, Kim HR, et al. Prognostic value of baseline and early treatment response of neutrophil-lymphocyte ratio, C-reactive protein, and lactate dehydrogenase in non-small cell lung cancer patients undergoing immunotherapy. Transl Lung Cancer Res 2023;12:1506-16.
15. Onodera R, Chiba S, Nihei S, et al. High level of C-reactive protein as a predictive factor for immune-related adverse events of immune checkpoint inhibitors in non-small cell lung cancer: a retrospective study. J Thorac Dis 2023;15:4237-47.
16. Lu D, Ma Z, Huang D, et al. Clinicopathological characteristics and prognostic significance of HDAC11 protein expression in non-small cell lung cancer: a retrospective study. Transl Lung Cancer Res 2022;11:1119-31.
17. Baran K, Waśko J, Kryczka J, et al. The Comparison of

Serum Exosome Protein Profile in Diagnosis of NSCLC Patients. Int J Mol Sci 2023;24:13669.
18. Frey A, Martin D, D'Cruz L, et al. C-Reactive Protein to Albumin Ratio as Prognostic Marker in Locally Advanced Non-Small Cell Lung Cancer Treated with Chemoradiotherapy. Biomedicines 2022;10:598.
19. Zhao H, Wu L, Yan G, et al. Inflammation and tumor progression: signaling pathways and targeted intervention. Signal Transduct Target Ther 2021;6:263.
20. Burgess S, Swanson SA, Labrecque JA. Are Mendelian randomization investigations immune from bias due to reverse causation? Eur J Epidemiol 2021;36:253-7.
21. de Leeuw C, Savage J, Bucur IG, et al. Understanding the assumptions underlying Mendelian randomization. Eur J Hum Genet 2022;30:653-60.
22. Kurki MI, Karjalainen J, Palta P, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature 2023;613:508-18.
23. Suhre K. Genetic associations with ratios between protein levels detect new pQTLs and reveal protein-protein interactions. Cell Genom 2024;4:100506.
24. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. Stat Med 2021;40:5434-52.
25. Deng L, Fu S, Yu K. Bias and mean squared error in Mendelian randomization with invalid instrumental variables. Genet Epidemiol 2024;48:27-41.
26. Burgess S, Thompson SG. Interpreting findings from Mendelian randomization using the MR-Egger method. Eur J Epidemiol 2017;32:377-89.
27. Yuan S, Chen J, Ruan X, et al. Smoking, alcohol consumption, and 24 gastrointestinal diseases: Mendelian randomization analysis. Elife 2023;12:e84051.
28. Huang SJ, Li R, Xu S, et al. Assessment of bidirectional relationships between circulating cytokines and periodontitis: Insights from a mendelian randomization analysis. Front Genet 2023;14:1124638.
29. Ji J, Jin D, Xu M, et al. AKR1B1 promotes pancreatic cancer metastasis by regulating lysosome-guided exosome secretion. Nano Research 2022;15:5279-94.
30. Banerjee S. Aldo Keto Reductases AKR1B1 and AKR1B10 in Cancer: Molecular Mechanisms and Signaling Networks. Adv Exp Med Biol 2021;1347:65-82.
31. Di Benedetto C, Borini Etichetti C, Cocordano N, et al. The p53 tumor suppressor regulates AKR1B1 expression, a metastasis-promoting gene in breast cancer. Front Mol Biosci 2023;10:1145279.
32. Khayami R, Hashemi SR, Kerachian MA. Role of aldo-keto reductase family 1 member B1 (AKR1B1) in the cancer process and its therapeutic potential. J Cell Mol Med 2020;24:8890-902.
33. He Z, Chen X, Liu G, et al. Sugt1 loss in skeletal muscle stem cells impairs muscle regeneration and causes premature muscle aging. Life Medicine 2023;2:lnad039.
34. Ogi H, Sakuraba Y, Kitagawa R, et al. The oncogenic role of the cochaperone Sgt1. Oncogenesis 2015;4:e149.
35. Tremiño L, Llop A, Rubio V, et al. The Conserved Family of the Pyridoxal Phosphate-Binding Protein (PLPBP) and Its Cyanobacterial Paradigm PipY. Life (Basel) 2022;12:1622.
36. Xie M, Pei DS. Serine hydroxymethyltransferase 2: a novel target for human cancer therapy. Invest New Drugs 2021;39:1671-81.
37. Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209-49.
38. Bhattacharya K, Picard D. The Hsp70-Hsp90 go-between Hop/Stip1/Sti1 is a proteostatic switch and may be a drug target in cancer and neurodegeneration. Cell Mol Life Sci 2021;78:7257-73.
39. de Azevedo ALK, Gomig THB, Ribeiro EMSF. Stress-induced phosphoprotein 1: how does this co-chaperone influence the metastasis steps? Clin Exp Metastasis 2024;41:589-97.
40. Huang L, Zhai E, Cai S, et al. Stress-inducible Protein-1 promotes metastasis of gastric cancer via Wnt/β-catenin signaling pathway. J Exp Clin Cancer Res 2018;37:6.
41. Guo X, Yan Z, Zhang G, et al. STIP1 Regulates Proliferation and Migration of Lung Adenocarcinoma Through JAK2/STAT3 Signaling Pathway. Cancer Manag Res 2019;11:10061-72.
42. Nanda S, Calderon A, Sachan A, et al. Rho GTPase activity crosstalk mediated by Arhgef11 and Arhgef12 coordinates cell protrusion-retraction cycles. Nat Commun 2023;14:8356.
43. Panagopoulos I, Andersen K, Eilert-Olsen M, et al. Therapy-induced Deletion in 11q23 Leading to Fusion of KMT2A With ARHGEF12 and Development of B Lineage Acute Lymphoplastic Leukemia in a Child Treated for Acute Myeloid Leukemia Caused by t(9;11)(p21;q23)/KMT2A-MLLT3. Cancer Genomics Proteomics 2021;18:67-81.
44. Yang Y, Wang S, Cai J, et al. Correction: Targeting ARHGEF12 promotes neuroblastoma differentiation,

MYCN degradation, and reduces tumorigenicity. Cell Oncol (Dordr) 2023;46:1871-2.
45. Garcia-Manero G, Platzbecker U, Lim KH, et al. Research and Clinical Updates on IRAK4 and Its Roles in Inflammation and Malignancy: Themes and Highlights from the 1st Symposium on IRAK4 in Cancer. Front Hematol 2024;3:1339870.
46. Sacta MA, Lee J. Delayed Presentation of IRAK4 Deficiency. J Clin Immunol 2024;44:122.
47. Guidetti F, Arribas A J, Cannas E, et al. Abstract C139: The IRAK4 inhibitor emavusertib (CA-4948) synergizes with second generation BTK inhibitors acalabrutinib and zanubrutinib in MYD88-L265P mutated lymphoma cell lines. Mol Cancer Ther 2023;22:C139.
48. Uckelmann HJ, Klusmann JH. Double trouble: IRAK1/4 inhibitors in AML/MDS. Blood 2023;142:945-6.
49. Aggarwal RK, Sidoli S, Sahu S, et al. Smoking carcinogen-induced inflammation promotes lung carcinogenesis via IRAK4 activation. Cancer Res 2024;84:1464.
50. Gómez Hernández G, Morell M, Alarcón-Riquelme ME. The Role of BANK1 in B Cell Signaling and Disease. Cells 2021;10:1184.
51. Castillejo-López C, Delgado-Vega AM, Wojcik J, et al. Genetic and physical interaction of the B-cell systemic lupus erythematosus-associated genes BANK1 and BLK. Ann Rheum Dis 2012;71:136-42.
52. Auer F, Guillermo RH, González-Herrero I, et al. A Tumor Suppressor Role for Bank1 in B-Cell Precursor Acute Lymphoblastic Leukemia. Blood 2018;132:1333.
53. Nikolakaki E, Mylonis I, Giannakouros T. Lamin B Receptor: Interplay between Structure, Function and Localization. Cells 2017;6:28.
54. Patil S, Deshpande S, Sengupta K. Nuclear envelope protein lamin B receptor protects the genome from chromosomal instability and tumorigenesis. Hum Mol Genet 2023;32:745-63.

Cite this article as: Zou X, Shen J, Li X, Diao Y, Zhang L. The causal effects of 2,821 protein level ratios on non-small cell lung cancer: a two-sample Mendelian randomization study. Transl Cancer Res 2025;14(2):1101-1110. doi: 10.21037/tcr-24-1523