Investigation of the Prevalence of Associated Genetic Mutations (Co-Mutations) in Patients with Actionable Driver Mutations in Lung Cancer: A Retrospective Study

Abed Agbarya [*,†], Walid Shalata [†], Edmond Sabo, Leonard Saiegh, Yuval Shaham, Haitam Nasrallah, Kamel Mhameed, Salam Mazareb, Mohammad Sheikh-Ahmad, Dan Levy Faber

Posted Date: 10 February 2026

doi: 10.20944/preprints202602.0780.v1

Keywords: lung cancer; co-mutations; genetic mutations; actionable mutations; driver mutations; genomic cluster

 doi:10.20944/preprints202602.0780.v1

Article

Investigation of the Prevalence of Associated Genetic Mutations (Co-Mutations) in Patients with Actionable Driver Mutations in Lung Cancer: A Retrospective Study

Abed Agbarya [1,2,*,†], Walid Shalata [3,4,†] Edmond Sabo [5], Leonard Saiegh [6], Yuval Shaham [1], Haitam Nasrallah [7], Kamel Mhameed [7], Salam Mazareb [5], Mohammad Sheikh-Ahmad [1,6,†] and Dan Levy Faber [1,8,†]

[1] The Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, 1 Efron street, Haifa 3109601, Israel
[2] Department of Oncology, Bnai-Zion Medical Center, 47 Eliyahu Golomb Avenue, Haifa 3339419, Israel
[3] The Legacy Heritage Cancer Center, Dr. Larry Norton Institute, Soroka Medical Center, 151 Yitzhack Rager, Beer Sheva, 84101, Israel
[4] Faculty of Health Sciences, Ben Gurion University of the Negev, 1 Ben Gurion Avenue, Beer Sheva, 8410501, Israel
[5] Pathology Institute, Lady Davis Carmel Medical Center, 7 Michal Street, Haifa 3436212, Israel
[6] Institute of Endocrinology, Bnai-Zion Medical Center, 47 Eliyahu Golomb Avenue, Haifa 3339419 Israel
[7] The Joseph Fishman Oncology Center, Rambam Medical Center, 6 Efron street, Haifa, 3525409, Israel
[8] Department of Cardiothoracic Surgery, Lady Davis Carmel Medical Center, 7 Michal St., Haifa 3436212, Israel
* Correspondence: abed.agbarya@b-zion.org.il; Tel.: +97248359016
† These authors contributed equally to this work.

Abstract

Background/Objectives: Lung cancer remains the leading cause of cancer-related mortality globally. Approximately 45% of these tumors harbor oncogenic mutations that drive carcinogenesis and are amenable to targeted therapies. Other predictive biomarkers— e.g., PD-L1, TMB, and MSI—play a crucial role in patients' management. This study aims to investigate the existence of mutation clusters (co-mutations) and evaluate the correlation of these clusters with various clinical and laboratory parameters. **Methods**: A retrospective study was conducted utilizing pathological samples from lung cancer patients harboring mutations in *EGFR, KRAS, ALK, BRAF, MET, HER2, ROS1, NTRK*, and *NRG1*. Data were collected from the Institute of Pathology at Carmel Medical Center between the years 2022 and 2024. Patients were stratified using a Two-Step Cluster Analysis algorithm based on actionable mutations and co-mutations. Heatmaps and dendrograms were generated to assess the correlation between these genomic clusters, clinical metrics, and predictive biomarkers. **Results**: The study cohort included 129 patients with actionable mutations. Five distinct clusters were identified: Clusters 1,2, and 3 exhibited a high expression of *STK11* and *TP53* co-mutations alongside *KRAS* drivers (n=38, n=12 and n=23 respectively). Clusters 4 and 5 demonstrated high expression of *ALK* alterations and tumor suppressor gene mutations (n=31, n=25 respectively). Multivariate analysis demonstrated statistically significant differences between clusters regarding age, gender, PD-L1 expression, and Tumor Mutational Burden. No significant associations were found regarding ethnicity or Microsatellite Instability status. **Conclusions**: By constructing clusters based on the aggregate of genomic alterations in patients with actionable mutations, it is possible to predict associations with distinct demographic and clinical characteristics. Future research should apply this analytical approach to larger cohorts to further characterize these subgroups and investigate potential correlations with therapeutic efficacy.

 doi:10.20944/preprints202602.0780.v1

Keywords: lung cancer; co-mutations; genetic mutations; actionable mutations; driver mutations; genomic cluster

1. Introduction

Lung cancer is a primary cause of cancer mortality worldwide, resulting in high death rates among both men and women [1–3]. Despite increasing awareness regarding early diagnosis, lung cancer is frequently detected at advanced stages where therapeutic options are limited, resulting in a poor prognosis [1]. Consequently, frequent screening of high-risk individuals is essential to facilitate early detection and significantly improve survival rates. Accurate and early diagnosis of pulmonary lesions is critical to tailoring the most appropriate therapeutic regimen and improving patient outcomes [1].

Lung cancer is defined as the uncontrolled growth of abnormal cells originating in the lungs, leading to tumor formation and compromised pulmonary function. It originates in the epithelium of the large and small bronchi and may metastasize via the lymphatic system or hematogenous spread, with common metastatic sites including the liver, adrenal glands, bone, and the central nervous system [1,4,5].

The incidence of lung cancer is steadily rising. In 2020, it was the second most diagnosed cancer (following breast cancer), accounting for 11.4% of 19.3 million cases. By 2022, it became the most diagnosed malignancy, with nearly 2.5 million new cases, representing 12.4% of all global cancer cases (one in eight new cancers). Regarding mortality, lung cancer is the leading cause of cancer death, rising from 18% in 2020 to 18.7% in 2022 [2,3]. The primary risk factor is tobacco smoking, though non-smokers exposed to second-hand smoke are also at risk. Additional risk factors include occupational exposure to carcinogens (e.g., asbestos, radon), air pollution, chronic lung diseases (e.g., COPD), and hereditary cancer syndromes [1].

Lung malignancies are classified into two main histological groups: Non-Small Cell Lung Carcinoma (NSCLC) and Small Cell Lung Carcinoma (SCLC). NSCLC accounts for 80% of cases, while SCLC comprises nearly 20%. This study focuses on NSCLC, which is further subdivided into histological subtypes, primarily Squamous Cell Carcinoma (SCC) and Adenocarcinoma (ADC), each accounting for 30-40% of cases. Large Cell Carcinoma (LCC) is a less common subtype, representing just under 10% of cases [5].

Current diagnostic and therapeutic strategies for lung cancer rely on histological diagnosis and molecular profiling [6]. Advanced Next-Generation Sequencing (NGS) allows for high-quality detection of a broad spectrum of mutations [7]. A significant proportion of NSCLC tumors harbor actionable mutations—genetic alterations treatable with targeted biological therapies rather than cytotoxic chemotherapy. The most prevalent actionable mutations discussed herein include Epidermal growth factor receptor (*EGFR*), *Kirstem rat sarcoma virus (KRAS) G12C*, and *ALK* [8–11].

The *KRAS G12C* mutation is the most common alteration in the *KRAS* gene, which encodes a key protein in cell proliferation pathways. *KRAS* mutations are present in 15-20% of NSCLC cases. This specific mutation is targetable by two biological agents, sotorasib and adagrasib, which selectively bind covalently to the KRAS protein, inhibiting its activity [12–14].

EGFR (Epidermal Growth Factor Receptor) mutations occur in approximately 15% of NSCLC cases in Western Europe and the USA and are frequently associated with non-smokers. *EGFR* encodes a transmembrane tyrosine kinase receptor involved in cell proliferation cascades. These mutations are treated with Tyrosine Kinase Inhibitors (TKIs) such as Osimertinib [8,10,11,14,15].

ALK (Anaplastic Lymphoma Kinase) gene rearrangements result from translocation, typically with *EML4*. Found in 3-7% of NSCLC cases, this translocation leads to overexpression of the ALK tyrosine kinase receptor, driving cell growth. This alteration is managed with TKIs such as crizotinib, ceritinib, lorlatinib, and alectinib [16]. Additional less common actionable mutations include *BRAF, MET, HER2, RET, ROS1, NTRK*, and *NRG1* [11,17].

 doi:10.20944/preprints202602.0780.v1

Beyond molecular profiling, immunotherapeutic biomarkers such as PD-L1 expression, Tumor Mutational Burden (TMB), and Microsatellite Instability (MSI) aid in treatment selection for NSCLC [18–20]. PD-L1 (Programmed Death-Ligand 1) expression on tumor cells suppresses the immune response by interacting with PD-1 on T-cells, inhibiting effector T-cells and promoting regulatory T-cells. Immune Checkpoint Inhibitors (ICIs) block this pathway to restore anti-tumor immunity [18]. TMB quantifies mutations within tumor DNA; high TMB correlates with better response to immunotherapy [19]. MSI refers to hypermutable DNA microsatellites due to defective mismatch repair (MMR) mechanisms and serves as a predictive marker for immunotherapy response [20].

This study utilized statistical analysis to determine whether specific clusters of co-mutations accompany actionable driver mutations. We specifically examined significant co-mutations including *TP53*, *KRAS*, *STK11*, *KEAP1*, and *CDKN2A*, alongside others of unclear significance. The existence of such clusters may influence prognosis and the efficacy of targeted therapies. We hypothesized that such clusters exist and aimed to identify them to optimize future therapeutic strategies. Additionally, we assessed the correlation of these clusters with biomarkers (PD-L1, TMB, MSI) and demographic parameters (age, gender, ethnicity) to identify patterns that may inform future studies on clinical efficacy and toxicity

2. Materials and Methods

2.1. Study Design and Population

This retrospective, non-interventional study included 129 patient samples archived at the Institute of Pathology database at Carmel Medical Center, Haifa. The cohort consisted of patients diagnosed with NSCLC who possessed a complete molecular profile containing at least one of the following actionable mutations: EGFR, KRAS, ALK, BRAF, MET, HER2, ROS1, NTRK, or NRG1. Lung tumor biopsy samples collected between 2022 and 2024 were included. Patients were screened based on the presence of specific actionable mutations defined in the inclusion criteria.

2.2. Data Collection

Following the initial screening of 129 patients, relevant data were extracted using LabOS software, which provided access to detailed pathological reports. Subsequently, statistical analysis and results interpretation were performed. All samples were sourced from the existing repository at the Institute of Pathology, Carmel Medical Center. Following data collection and screening, statistical analysis was performed on an anonymized dataset (identified only by serial numbers).

2.3. Ethics

The study was approved by the Helsinki Committee of Carmel Medical Center protocol CMC-0073-24 on September 23, 2024.

2.4. Statistical Analysis

Cluster analysis was performed based on actionable mutations, co-mutations, and Copy Number Variations (CNVs) using a Two-Step Cluster Analysis algorithm. This hybrid method is designed for categorical and/or continuous data and operates in two phases:

2.4.1. Pre-Clustering

The algorithm scans data to create small sub-clusters using a log-likelihood distance measure suitable for categorical (binary) variables. Since our variables are binary (presence or absence of a mutation), a log-likelihood distance metric estimates similarity based on the likelihood that data points will appear in the same cluster.

 doi:10.20944/preprints202602.0780.v1

2.4.2. Hierarchical Clustering

Sub-clusters are merged into final clusters based on model fit criteria (BIC or AIC). The algorithm determines the optimal number of clusters automatically, though manual specification is possible.

Five clusters were generated for all patients. Heatmaps were created for visualization using ArrayGen (Undri, Pune, Maharashtra, India). Heatmaps visually represent the relationships between mutation types across patients. The color gradient ranges from light yellow (0; absence/low presence of a mutation) to dark brown (+2; high intensity/prevalence of the mutation). It is important to note that in binary data clustering, color intensity reflects the prevalence or significance of a mutation within the context of the cluster.

2.4.3. Dendrograms

A dendrogram was constructed for each heatmap to visualize the hierarchical structure. It illustrates the similarity between patient mutation profiles or the co-occurrence of specific mutations. Branch height indicates the degree of dissimilarity; shorter distances imply high similarity.

2.4.4. Comparative Statistics

Clusters were compared based on demographics (age, gender, ethnicity) and biomarkers (PD-L1 through immunohistochemistry, TMB, and MSI/MSS via NGS). Continuous variables were tested for normality using the Kolmogorov-Smirnov test and analyzed using Student's t-test or Mann-Whitney U test as appropriate. Bonferroni correction was applied to address multiple hypothesis testing. Categorical variables were analyzed using Chi-square or Fisher's exact tests. A two-tailed p-value of ≤0.05 was considered statistically significant. Analysis was performed using IBM SPSS version 26 (IBM, Armonk, NY, U.S.A).

3. Results

To analyze binary categorical data (presence/absence of mutations), a Two-Step Cluster Analysis was performed on 129 patients' samples. This unsupervised algorithm grouped patients based on mathematical proximity. Five clusters were generated (Cluster 1, *n*=38; Cluster 2, *n*=12; cluster 3, *n*=23; Cluster 4, *n*=31 and cluster 5 *n*=25).

Each cluster was visualized using a dendrogram and a heatmap. In the heatmaps, mutations (x-axis) were color-coded to distinguish between primary mutations, co-mutations, and CNVs. Patients (y-axis) were listed by serial number. AI tools (Cloud, ChatGPT deep research) were utilized to assist in identifying mutation patterns within the algorithmic clusters.

3.1. Clustering of Actionable Mutations and Co-Mutations Cluster Patterns

3.1.1. Cluster 1 (n=38)

Characterized by KRAS mutations (various subtypes) alongside STK11 co-mutations, a high frequency of TP53 mutations, and a high concentration of MET alterations (exon 14 skipping + amplification) (Figure 1).

 doi:10.20944/preprints202602.0780.v1

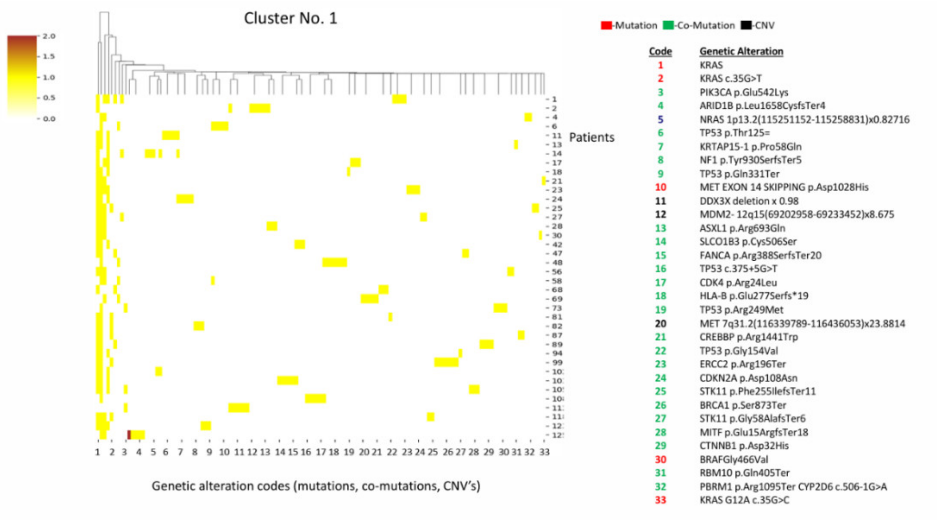

Figure 1. Heatmap and cluster 1 pattern. Key alterations: KRAS mutations (multiple, including G12A, G12A, G12D, G12V); STK11 mutations; TP53 mutations (several loss-of-function); MET exon 14 skipping + MET amplification; MDM2 amplification; NF1 and PIK3CA mutations. Pattern: high frequency of TP53 mutations; Frequent KRAS/STK11 co-mutations; MET alterations (exon 14 skipping + amplification).

3.1.2. Cluster 2 (n=12)

Similar KRAS/STK11/TP53 triad to Cluster 1, but distinguished by a higher frequency of EGFR CNV amplifications and mutations in chromatin remodeling genes (KDM6A, ATRX, SMARCA4) (Figure 2).

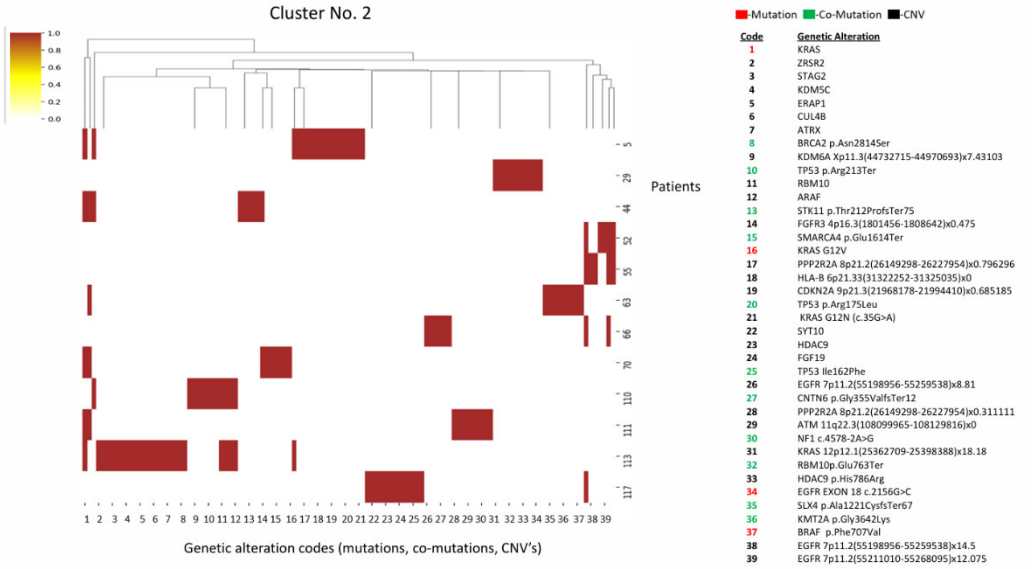

Figure 2. Heatmap and cluster 2 pattern. Key alterations: KRAS mutations (G12V, G12N); STK11 mutations; EGFR amplification (CNV), STK11, TP53 mutations, multiple chromatin modifiers: KDM6A, ATRX, SMARCA4; HLA-B deletion. Pattern: KRAS/STK11/TP53 triad similar to Cluster 1, but with more CNVs; EGFR CNV amplifications; Chromatin remodeling gene alterations.

3.1.3. Cluster 3 (n=23)

High presence of KRAS and STK11 co-mutations, but notably lower TP53 frequency compared to Clusters 1 and 2. Contains mutations in DNA repair pathways (Figure 3).

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 February 2026 doi:10.20944/preprints202602.0780.v1

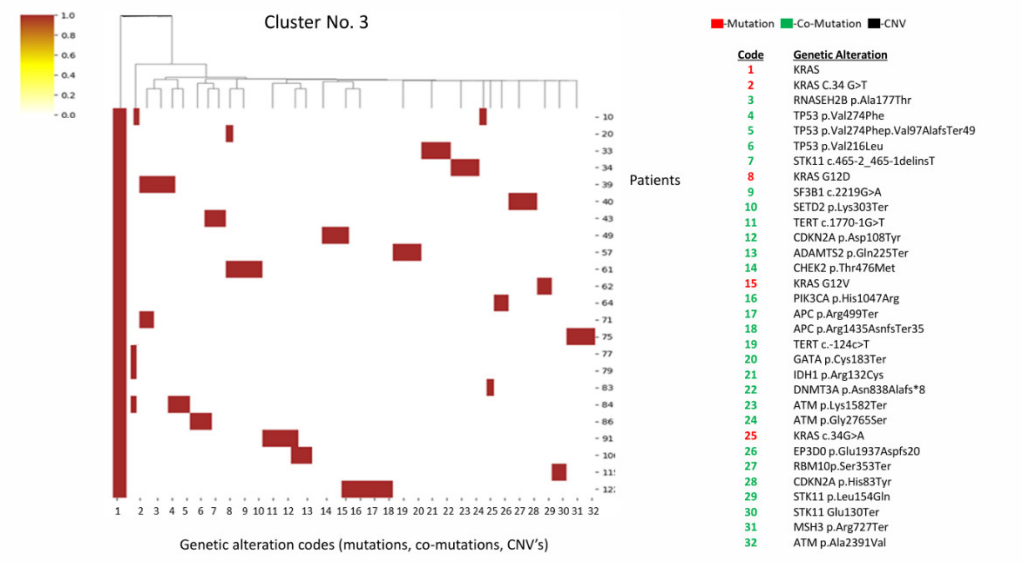

Figure 3. Heatmap and cluster 3 pattern. Key alterations: KRAS mutations (G12D, G12V, c.34G>A); Co-occurring STK11 mutations; TP53, CDKN2A, ATM, TERT, SETD2, and APC. Pattern: Strong presence of KRAS + STK11 mutations, but less TP53 activity than Cluster 1; Presence of DNA repair pathway genes: ATM, CHEK2, DNMT3A, IDH1.

3.1.4. Cluster 4 (n=31)

Defined by EGFR activating mutations co-occurring with ALK fusions. This represents an oncogene-driven profile with less genomic instability (Figure 4).

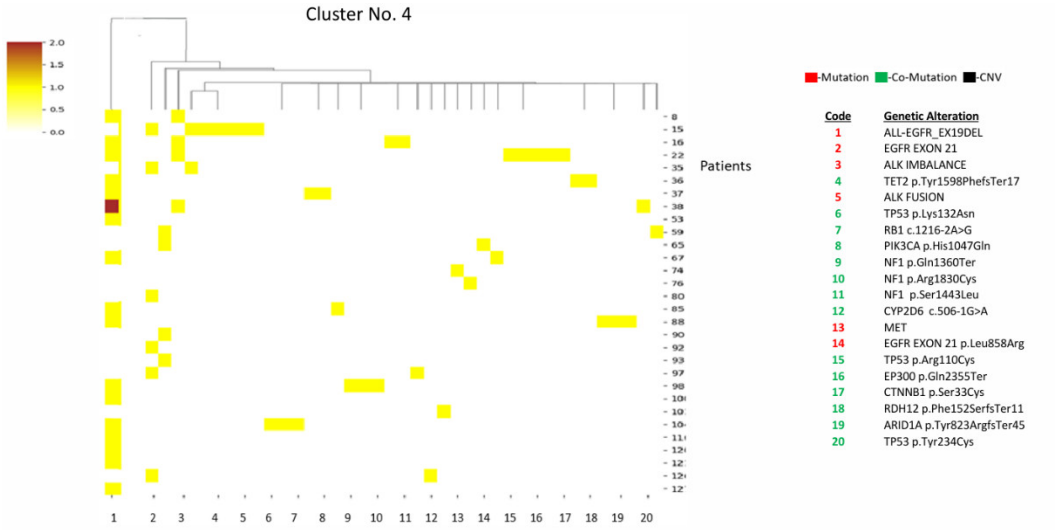

Figure 4. Heatmap and cluster 4 pattern. Key alterations: EGFR activating mutations (EX19DEL, EX21 L858R); ALK fusions; TP53, ARID1A, PIK3CA. Pattern: EGFR/ALK oncogene-driven tumors; Less genomic instability, more oncogene-addicted.

3.1.5. Cluster 5 (n=25)

Characterized by EGFR exon 20 mutations/insertions. Unlike Cluster 4, these tend to co-occur with Tumor Suppressor Loss (e.g., NF1, TP53, PTEN) (Figure 5).

 doi:10.20944/preprints202602.0780.v1

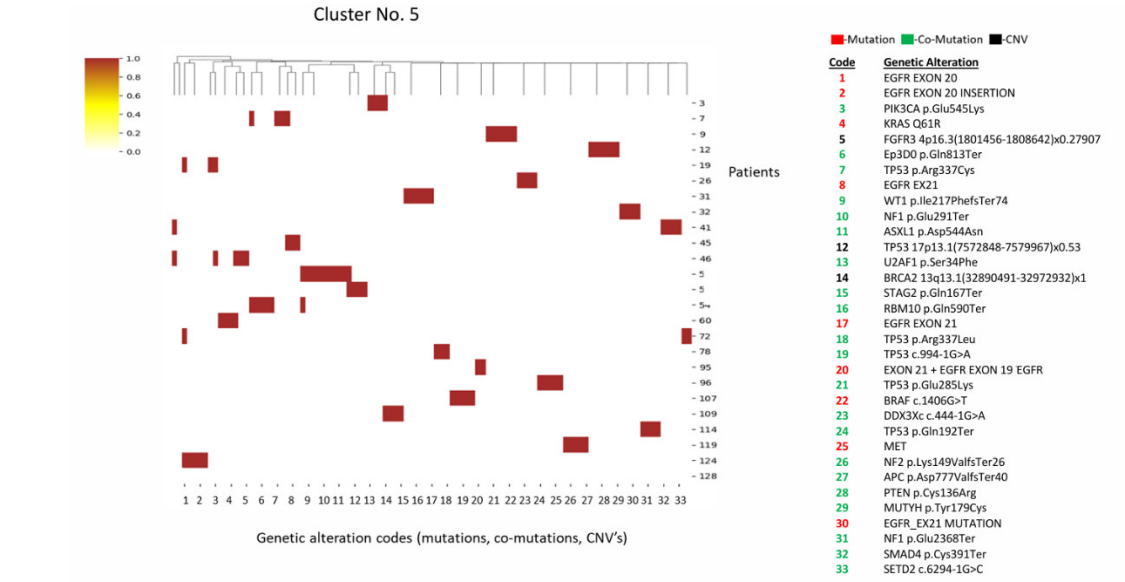

Figure 5. Heatmap and cluster 5 pattern. Key alterations: EGFR exon 20 mutations/insertions; NF1, NF2, TP53, PTEN, SETD2; PIK3CA, BRAF, KRAS (Q61R); similar to cluster 4, but with more diverse tumor suppressor loss. Pattern: EGFR exon 20 mutations; Tumor suppressor mutations.

3.2. Immunotherapeutic Biomarker Analysis

3.2.1. PD-L1 Expression

Comparison among clusters according to the ordinal categorical variable showed that the mean PD-L1 expression was lowest in cluster 4 samples (Figure 6). Cluster 4 had significantly lower PD-L1 expression compared to Cluster 1 (p=0.001) and Cluster 3 (p=0.032). Between cluster 1 and 5 there is a difference with a significance level of 0.011. Between cluster 3 and 4 there is a difference with a significance level of 0.032 (Figure 6).

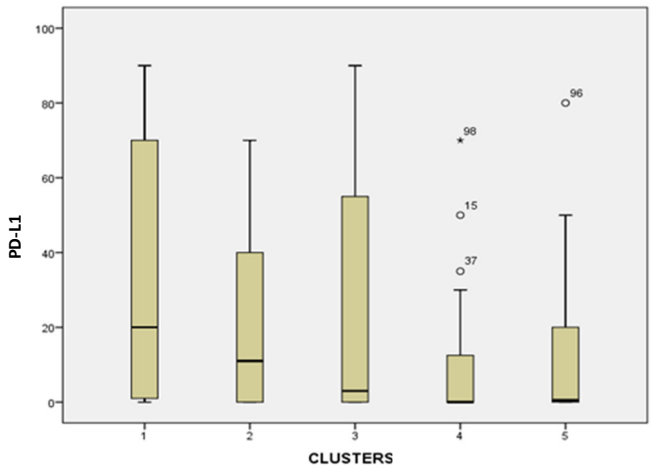

Figure 6. Boxplot graph of PD-L1 expression across the different clusters. The analysis to create this graph was performed on a continuous variable with values ranging from 0 to 100. Between cluster 1 and 4 there is a difference with a significance level of p=0.001. Between cluster 1 and 5 there is a difference with a significance level p=0.011. Between cluster 3 and 4 there is a difference with a significance level p=0.032. Between cluster 2 and 4 there is a non-significant difference, p=0.061.

 doi:10.20944/preprints202602.0780.v1

3.2.2. Tumor Mutational Burden (TMB)

TMB Statistical Significance: Cluster 4 had significantly lower TMB than Clusters 1 (p=0.012), 2 (p=0.001), and 3 (p=0.009). Cluster 2 had significantly higher TMB than Clusters 1(p=0.013), 4(p=0.001), and 5 (p=0.007) (Figure 7).

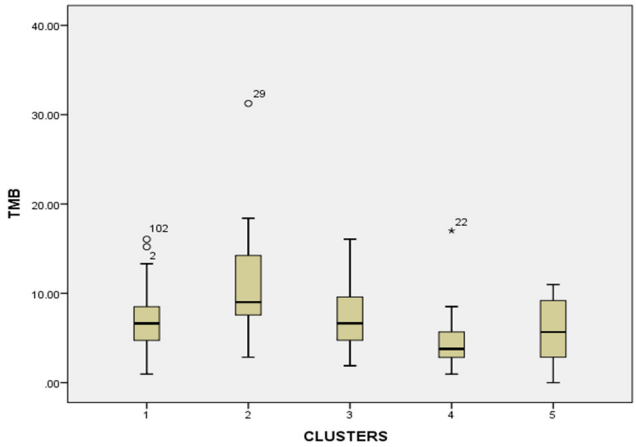

Figure 7. Boxplot graph of Tumor mutational burden (TMB) expression across the different clusters. The analysis for creating this graph was conducted on a continuous variable.

3.2.3. MSI/MSS

No statistically significant differences were found between clusters in their Microsatellite instability; all evaluable patients were Microsatellite Stable (MSS).

3.3. Demographic Analysis

3.3.1. Age

Age of patients analysis (Figure 8) yielded significant differences were between Cluster 1 vs. 2 (p=0.026), Cluster 1 vs. 5 (p=0.031), Cluster 2 vs. 4 (p=0.039), and Cluster 4 vs. 5 (p=0.044). Clusters 1 and 4 represented a younger demographic population.

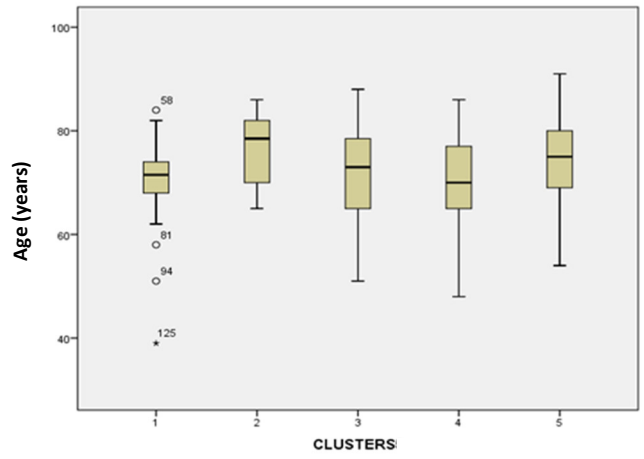

Figure 8. Boxplot graph of Age (years) variable across the different clusters.

 doi:10.20944/preprints202602.0780.v1

3.3.2. Gender

Figure 9 presents the comparison between genders among the clusters. Cluster 4 comprised of significantly more women than men, while clusters 1,2,3,5 had less women than men. The distribution of men and women is without statistically significant difference between cluster 1 and cluster 4. Findings of the analysis showed that between Cluster 2 and cluster 4 there is statistically significant difference (p=0.01), as well as between cluster 3 and cluster 4 (p=0.036) and between cluster 4 and cluster 5 (p=0.036).

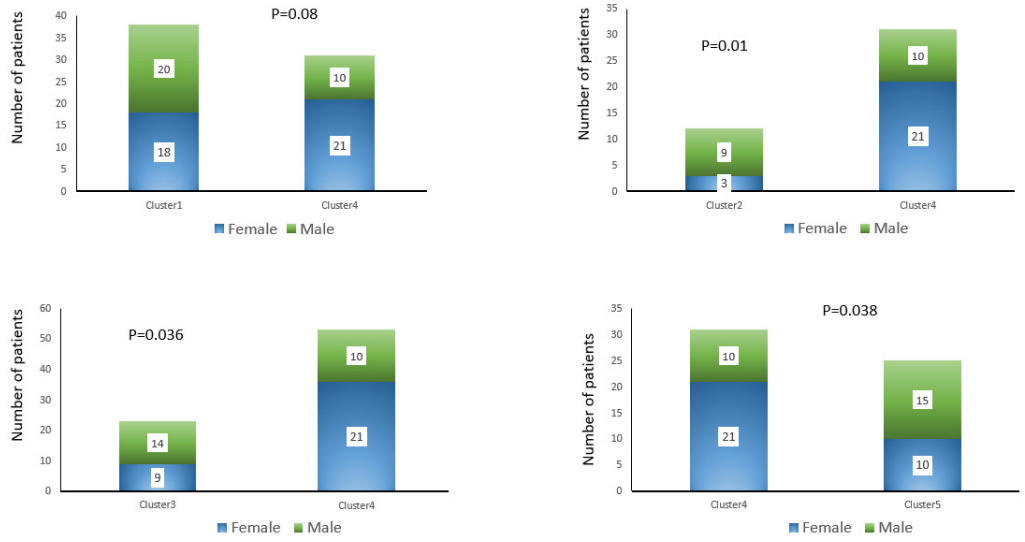

Figure 9. Comparison between genders among the clusters.

3.3.3. Ethnic

No significant differences were observed between Jewish and Arab patients' ethnicity across clusters

4. Discussion

This study sought to identify prevalent associations between co-mutations and actionable driver mutations in lung cancer patients, and to examine their correlations with immunotherapeutic biomarkers and demographics. Our goal was to establish a foundation for future research investigating the clinical implications of these genomic subgroups to enhance precision medicine in lung cancer and to contribute to improve the prognosis of lung cancer patients.

The cluster analysis revealed distinct genomic patterns. *KRAS*-driven tumors (Clusters 1 and 3) frequently co-occurred with *STK11*. This finding aligns with recently publications on the biomarkers *STK11* co-mutation with *KRAS* and the implication as per therapeutic outcome with neoadjuvant treatment using ICI (immune checkpoint inhibitors) for NSCLC patients [21–23]. Furthermore, Clusters 1 and 3 exhibited varying degrees of *TP53* involvement which is in line with Schabath 2015 study [24]. Skoulidis et al. [25] observed that in lung adenocarcinoma PD-L1 expression plays a role in response of *STK*11 and *TP53* co-mutations with *KRAS*. In the current study Cluster 1 displayed high PD-L1 expression, suggesting potential immunogenicity. Suzawa et al., [26] findings indicated that *KRAS* mutation is common in NSCLC tumors harboring MET alteration. Similarly in the present study, Cluster 1 was associated with high levels of MET alterations. These co-existing genomic changes could each be targeted by specific agents, however due to reported therapeutic resistance [26], it is important to know the molecular profiling of each tumor while proposing a treatment plan.

In contrast, *EGFR*-driven tumors (Clusters 4 and 5) exhibited distinct co-mutation profiles. NSCLC patients having both mutated *EGFR* and ALK fusions have been described in a review by Kemper [27]. In the current study Cluster 4, was characterized by *EGFR* and *ALK* fusions, and

 doi:10.20944/preprints202602.0780.v1

demonstrated significantly lower PD-L1 which is in line with previous publication by Gainor et al. [28]. As well, Cluster 4 had the lowest TMB levels, compared to other clusters. Taken together Cluster 4 composition is indicative of an "oncogene-addicted" phenotype driven by signaling pathways (i.e., *EGFR*) rather than genomic instability. Cluster 5, involving *EGFR* exon 20 insertions, which are relatively rare [29,30], was associated with tumor suppressor loss [31,32] Cluster 5 patients bearing low TMB are predicted to have shorter survival than those with high TMB (i.e., cluster 2) [33].

Statistically, Cluster 4 (EGFR/ALK) presented with significantly lower biomarkers (PD-L1, TMB) compared to the *KRAS*-dominant clusters (clusters 1,2,3). Clusters 1,2,3 which include mutations KRAS, STK11, and TP53 may affect the immunogenicity of the patient by increasing PD-L1 expression [34]. *KRAS, STK11* and *TP53* mutations may also affect the presence of accompanying mutations related to genes involved in stabilizing the chromatin (cluster 2) and DNA repairs pathways (cluster 3) as similarly found in pancreatic cancer genomic mutations [35]. This aligns with the hypothesis that *KRAS, STK11*, and *TP53* mutations may drive immunogenicity and genomic instability (higher TMB), whereas *EGFR*/*ALK* alterations (clusters 4,5) rely on oncogenic signaling, resulting in "colder" tumors immunologically (with decreased PD-L1 and TMB expressions). This means that lung cancer tumors harboring EGFR/ALK mutations rely less on genomic instability.

Hence, proposed future research would deepen the investigation on this topic. i.e.,; as PD-L1 and TMB have clinical significance; and explore the creation of clusters that are in correlation with PD-L1 and TMB. Further studies should assess the effect of immune-therapy related biomarkers on the clustering in these patients and check the clinical significance by outcomes of the treatments such as efficiency and toxicity.

Frille et al. 2024 [36] presented interactions between four co-mutations in NSCLC: *KRAS, STK11, TP53* and *KEAP1*. They reported that combined *KRAS G12C* and *TP53* co-occurrence may be targeted by immunotherapy administered to patients diagnosed with advanced tumor stage in stage IV.

Demographically, Cluster 4 contained a significantly higher proportion of women and younger patients compared to other clusters, aligning with known epidemiological data regarding *EGFR* mutations [37]. Although cluster 4 consisted of higher rate of women in than in other clusters, it is difficult to draw a clear clinically significant conclusion it is plausible that the differences were arbitrary or as a result of other variables that were not assessed during the present study.

Further research should be conducted with a bigger sample and analysis of clinical results could assess if there is an effect of these characteristics on the mutations profile in the different clusters.

The study has several limitations, its retrospective design and a modest sample size of 129 patients, as well as heterogeneity in cluster sizes (e.g., n=12 in cluster 2 vs. n=38 in cluster 1), which may affect statistical power. Moreover, the tumor stage was not part of the present study although different stages (early vs. late stage) may be related to multiple coexisting mutations. Additionally, the lack of behavioral data (smoking status or drinking alcohol) in the database, limits the ability to control for environmental confounders while creating a more clinically accurate cluster. The use of an unsupervised algorithm, while mathematically robust, groups purely by proximity without inherent clinical weighting.

Future prospective studies are recommended with larger cohorts to validate these clusters and assess their clinical utility regarding treatment efficacy and toxicity. Specifically, investigating the impact of these co-mutation clusters on survival outcomes and resistance mechanisms is warranted.

5. Conclusions

The current study findings indicate that actionable mutations consistently appear with specific co-mutations (e.g., KRAS with STK11/TP53; EGFR with ALK or Tumor Suppressor Loss). These associations likely influence tumor biology and response to therapy. The correlation between specific mutation clusters and immunotherapeutic biomarkers (PD-L1/TMB) suggests that genomic clustering could serve as a predictive tool for immunotherapy response.

 doi:10.20944/preprints202602.0780.v1

Author Contributions: Conceptualization, A.A., W.S., M.S.A, and D.L.F.; methodology, Y.S., and E.S.; software, Y.S., and E.S.; validation, A.A., W.S., M.S.A, and D.L.F. .; formal analysis, Y.S. and E.S.; investigation, Y.S. and E.S.; resources, A.A., W.S., M.S.A, D.L.F., L.S., Y.S., H.N., K.M., S.M. and E.S.; data curation, Y.S. and E.S.; writing—original draft preparation, Y.S.; writing—review and editing, A.A., W.S., M.S.A, D.L F., L.S., H.N., K.M., S.M. and E.S.; visualization, Y.S.; supervision, A.A., W.S., M.S.A, and D.L F.; project administration, A.A., W.S., M.S.A, and D.L.F. All authors A.A., W.S., M.S.A, D.L F., L.S., Y.S., H.N., K.M., S.M. and E.S. have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board Ethics Committee of Carmel Medical Center protocol code CMC-0073-24 on September 23, 2024.

Informed Consent Statement: Patient consent was waived due to retrospective nature of the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: During the preparation of this manuscript/study, the authors used AI tool, Cloud, ChatGPT deep research, for the purposes of assisting in identifying mutation patterns within algorithmic clusters. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

MSI	Microsatellite Instability
PD-L1	Programmed Death-Ligand 1
TMB	Tumor Mutational Burden
COPD	Chronic obstruction pulmonary disease
NSCLC	Non-Small Cell Lung Carcinoma
SCLC	Small Cell Ling Carcinoma
ADC	Adenocarcinoma
SCC	Squamous Cell Carcinoma
NGS	Next-Generation Sequencing
LCC	Large Cell Carcinoma
EGFR	Epidermal Growth Factor Receptor
TKI	Tyrosine Kinase Inhibitors
ALK	Anaplastic Lymphoma Kinase
ICI	Immune Checkpoint Inhibitors
CNV	Copy number variation
KRAS	Kirsten rat sarcoma virus
STK11	Serine threonine kinase 11
TP53	Tumor protein 53

References

1. World Health Organization. Lung Cancer. Available online: https://www.who.int/news-room/fact-sheets/detail/lung-cancer (accessed on 10 January 2026).
2. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin*. **2021**, *71(3)*, 209-249. doi:10.3322/caac.2166
3. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin*. **2024**, *74(3)*, 229-263. doi:10.3322/caac.21834

 doi:10.20944/preprints202602.0780.v1

4. Milovanovic, I.S.; Stjepanovic, M.; Mitrovic, D. Distribution patterns of the metastases of the lung carcinoma in relation to histological type of the primary tumor: An autopsy study. *Ann. Thorac. Med*. **2017**, *12(3)*, 191-198. doi:10.4103/atm.ATM_276_16
5. Kitamura, H.; Yazawa, T.; Okudela, K.; Shimoyamada, H.; Sato, H. Molecular and Genetic Pathogenesis of Lung Cancer: Differences Between Small-Cell and Non-Small-Cell Carcinomas. *The Open Pathology Journal*, **2008**, *2(1)*, 106–114. DOI:10.2174/1874375700802010106
6. Liu, S.; Xu, T.; Cao, X.; Li, H.; Jin, R. Histological transformation in lung cancer: Mechanisms, clinical characteristics, and therapeutic approaches. *Biochimica et Biophysica Acta (BBA) – Reviews on cancer*. **2025**, *1880*(5), Doi: 10.1016/j.bbaan.2025.189413
7. Satam, H.; Joshi, K.; Mangrolia, U.; et al. Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology (Basel)*. **2023**,*12(7)*,997. doi:10.3390/biology12070997
8. Raphael, A.; Dudnik, E.; Hershkovitz, D.; et al. *FGFR* Fusions as an Acquired Resistance Mechanism Following Treatment with Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors (EGFR TKIs) and a Suggested Novel Target in Advanced Non-Small Cell Lung Cancer (aNSCLC). *J Clin Med*. **2022**, *11*(9), 2475. doi:10.3390/jcm11092475
9. Raphael A, Onn A, Holtzman L, et al. The Impact of Comprehensive Genomic Profiling (CGP) on the Decision-Making Process in the Treatment of ALK-Rearranged Advanced Non-Small Cell Lung Cancer (aNSCLC) After Failure of 2nd/3rd-Generation ALK Tyrosine Kinase Inhibitors (TKIs). *Front Oncol*. **2022**, *12*, 874712. doi:10.3389/fonc.2022.874712
10. Mountzios, G.; Planchard, D.; Metro, G.; et al. Molecular Epidemiology and Treatment Patterns of Patients With EGFR Exon 20-Mutant NSCLC in the Precision Oncology Era: The European EXOTIC Registry. *JTO Clin Res Rep*. **2022**, *4(1)*, 100433. doi:10.1016/j.jtocrr.2022.100433
11. Le, X.; Elamin, Y.Y.; Zhang, J. New Actions on Actionable Mutations in Lung Cancers. *Cancers (Basel)*. **2023**, *15(11*), 2917. doi:10.3390/cancers15112917
12. Vijayalakshmi, R.; Krishnamurthy, A. Targetable "driver" mutations in non small cell lung cancer. *Indian J. Surg. Oncol*. **2011**, *2(3),* 178-188. doi:10.1007/s13193-011-0108-0
13. El Osta, B. *KRAS* G12C mutation: from undruggable target to potentially agnostic biomarker. *Transl. Lung Cancer Res*. **2023**, *12(6),* 1147-1151. doi:10.21037/tlcr-23-174
14. Wu, J.; Lin, Z. Non-Small Cell Lung Cancer Targeted Therapy: Drugs and Mechanisms of Drug Resistance. *Int. J. Mol. Sci*. **2022**, *23(23),* 15056. doi:10.3390/ijms232315056
15. Yoneda, K.; Imanishi, N.; Ichiki, Y.; Tanaka, F. Treatment of Non-small Cell Lung Cancer with EGFR-mutations. *J UOEH*. **2019**, *41(2),* 153-163. doi:10.7888/juoeh.41.153
16. Liao, B.C.; Lin, C.C.; Shih, J.Y.; Yang, J.C. Treating patients with ALK-positive non-small cell lung cancer: latest evidence and management strategy. *Ther. Adv. Med. Oncol*. **2015**, 7(5), 274-290. doi:10.1177/1758834015590593
17. Olmedo, M.E.; Cervera, R.; Cabezon-Gutierrez, L.; et al. New horizons for uncommon mutations in non-small cell lung cancer: *BRAF*, *KRAS*, *RET*, *MET*, *NTRK*, *HER2*. *World J Clin Oncol*. **2022**, *13(4*), 276-286. doi:10.5306/wjco.v13.i4.276
18. Lin, X., Kang, K., Chen, P. et al. Regulatory mechanisms of PD-1/PD-L1 in cancers. *Mol. Cancer* **2024**, *23*, 108. doi:10.1186/s12943-024-02023-w
19. Strickler, J. H.; Hanks, B. A; Khasraw, M. Tumor Mutational Burden as a Predictor of Immunotherapy Response: Is More Always Better? Clinical Cancer Research, **2021,** *27(5),* 1236–1241. doi:10.1158/1078-0432.CCR-20-3054
20. Li, K., Luo, H., Huang, L. et al. Microsatellite instability: a review of what the oncologist should know. *Cancer Cell Int*. **2020**, 20, 16. doi:10.1186/s12935-019-1091-8
21. Sisca, L.; Cascetta, P.; Aijaz, A.; et al. *KRAS* and *STK11* co-mutations in resectable non-small cell lung cancer: enduring prognostic value and impaired immunotherapy response. *Transl Lung Cancer Res*. **2025**, *14(7),* 2374-2382. doi:10.21037/tlcr-2025-463
22. Rosner, S.; Connor, S.; Sanber, K.; et al. Divergent Clinical and Immunologic Outcomes Based on STK11 Co-mutation Status in Resectable KRAS-Mutant Lung Cancers Following Neoadjuvant Immune Checkpoint Blockade. *Clin Cancer Res*. **2025**, *31*(2), 339-351. doi:10.1158/1078-0432.CCR-24-2983

23. Manolakos, P.; Ward, L.D. A Critical Review of the Prognostic and Predictive Implications of *KRAS* and *STK11* Mutations and Co-Mutations in Metastatic Non-Small Lung Cancer. *J. Pers. Med.* **2023**, *13*, 1010. https://doi.org/10.3390/jpm13061010
24. Schabath, M.B.; Welsh, E.A.; Fulp, W.J.; et al. Differential association of STK11 and TP53 with KRAS mutation-associated gene expression, proliferation and immune surveillance in lung adenocarcinoma. *Oncogene*. **2016**, *35(24)*, 3209-3216. doi:10.1038/onc.2015.375
25. Skoulidis, F.; Goldberg, M.E.; Greenawalt, D.M.; et al. *STK11/LKB1* Mutations and PD-1 Inhibitor Resistance in *KRAS*-Mutant Lung Adenocarcinoma. *Cancer Discov*. **2018**, *8(7)*, 822-835. doi:10.1158/2159-8290.CD-18-009
26. Suzawa, K.; Offin M, Lu D, et al. Activation of KRAS Mediates Resistance to Targeted Therapy in MET Exon 14-mutant Non-small Cell Lung Cancer. *Clin Cancer Res*. **2019**, *25(4)*, 1248-1260. doi:10.1158/1078-0432.CCR-18-1640
27. Kemper, M.; Elges, S.; Kies, P.; et al. What do we know about the role of neoadjuvant targeted therapy in early-stage *EGFR*-mutant and *ALK*-fused non-small cell lung cancer?-a narrative review of the current literature. *Transl Lung Cancer Res*. **2024**, *13(10)*, 2813-2827. doi:10.21037/tlcr-24-359
28. Gainor, J.F.; Shaw, A.T.; Sequist, L.V.; et al. EGFR Mutations and ALK Rearrangements Are Associated with Low Response Rates to PD-1 Pathway Blockade in Non-Small Cell Lung Cancer: A Retrospective Analysis. *Clin Cancer Res*. **2016**, *22(18*), 4585-4593. doi:10.1158/1078-0432.CCR-15-3101
29. Garzón-Ibáñez, M., Reyes, R., Molina-Vila, M.Á. et al. Landscape and clinical implications of *EGFR* exon 20 insertions in non-small cell lung cancer patients. *Clin Transl Oncol* **27**, 3559–3569 (2025). https://doi.org/10.1007/s12094-025-03899-w
30. Seo, D.; Lim, J.H. Targeted Therapies for EGFR Exon 20 Insertion Mutation in Non-Small-Cell Lung Cancer. *Int J Mol Sci*. **2024**, *25(11*), 5917. doi:10.3390/ijms25115917
31. Canale, M.; Andrikou, K.; Priano, I.; et al. The Role of *TP53* Mutations in *EGFR*-Mutated Non-Small-Cell Lung Cancer: Clinical Significance and Implications for Therapy. *Cancers (Basel)*. **2022**, *14(5*, 1143. doi:10.3390/cancers14051143
32. Christopoulus, P.; Kluck, K.; Kichner, M.; et al. the impact of TP53 co-mutations and immunologic microenvironment on outcome of lung cancer with EGFR exon 20 insertions. *Eur J Cancer*. **2022**, *170*, 106-118. https://doi.org/10.1016/j.ejca.2022.04.020
33. Dong, Y.; Khan, L.; Yao, Y. Immunological features of EGFR-mutant non-small cell lung cancer and clinical practice: a narrative review. *J Natl. Cancer Cent*. **2024**, *4(4)*, 289-298. https://doi.org/10.1016/j.jncc.2024.06.004
34. Zhang, F., Wang, J., Xu, Y. et al. Co-occurring genomic alterations and immunotherapy efficacy in NSCLC. *npj Precis. Onc.* **2022, *6*,** 4. https://doi.org/10.1038/s41698-021-00243-7
35. Abraham, S.; et al. Relations between mutant KRAS and TP53 subtypes and other co-mutations in pancreatic cancer. J. Clin Oncol. **2023**, *41*, e16294. Doi: 10.1200/JCO.2023.41.16.supple.e16294
36. Frille, A.; Boeschen, M.; Wirtz, H.; Stiller, M.; Blaker, H.; von Laffert, M. *TP53* co-mutations n advanced lung adenocarcinoma: comparative bioinformatic analyses suggest ambivalent character on overall survival alongside *KRAS*, *STK11* and *KEAP1* mutations. *Front. Oncol*. **2024**, *14*. Doi: 10.3389/fonc.2024.1367583
37. Bell, D.W.; Brannigan, B.W.; Matsuo, K.; et al. Increased prevalence of EGFR-mutant lung cancer in women and in East Asian populations: analysis of estrogen-related polymorphisms. *Clin Cancer Res*. **2008**, *14(13*), 4079-4084. doi:10.1158/1078-0432.CCR-07-5030