

Article

Social Network Clustering Analysis for Detection of Associated Genetic Co-Mutations in Patients with Actionable Driver Mutations in NSCLC

Abed Agbarya ^{1,2,*}, Haitham Nasrallah ^{3,†}, Kamel Mhameed ³, Edmond Sabo ⁴, Walid Shalata ^{5,6}, Esti Liani ⁴, Salam Mazareb ⁴, Mohammad Sheikh-Ahmad ⁷, Leonard Saiegh ⁷, Dejan Radonjic ⁸, Viktor Sebek ⁹ and Dan Levy-Faber ^{1,10}

¹ The Ruth and Bruce Rappaport Faculty of Medicine, Technion-Israel Institute of Technology, Haifa 3109601, Israel; danfaber1@hotmail.com (D.L.-F.)

² Department of Oncology, Bnai-Zion Medical Center, Haifa 3339419, Israel

³ The Joseph Fishman Oncology Center, Rambam Medical Center, Haifa 3525409, Israel; h_nasrallah@rmc.gov.il (H.N.); k_mhameed@rmc.gov.il (K.M.)

⁴ Pathology Institute, Lady Davis Carmel Medical Center, Haifa 3436212, Israel; edmondsa@clalit.org.il (E.S.); esti.liani@clalit.org.il (E.L.); mazareb_salam@clalit.org.il (S.M.)

⁵ The Legacy Heritage Cancer Center, Dr. Larry Norton Institute, Soroka Medical Center, Beer Sheva 8457108, Israel; walid.shalata@gmail.com

⁶ Faculty of Health Sciences, Ben Gurion University of the Negev, Beer Sheva 8410501, Israel

⁷ Institute of Endocrinology, Bnai-Zion Medical Center, Haifa 3339419, Israel; mohammad.ahmad@b-zion.org.il (M.S.-A.); leonard.saiegh@gmail.com (L.S.)

⁸ Faculty of Medicine and Dentistry, Palacky University, Hnevotinska Street 3, 77515 Olomouc, Czech Republic; dejan_radonjic@msn.com

⁹ Department of Pharmacology, Faculty of Medicine and Dentistry, Palacky University, Hnevotinska Street 3, 77515 Olomouc, Czech Republic; sebek.viktor@gmail.com

¹⁰ Department of Cardiothoracic Surgery, Lady Davis Carmel Medical Center, Haifa 3436212, Israel

* Correspondence: abed.agbarya@b-zion.org.il

† These authors contributed equally to this work.

Abstract

Non-small cell lung cancer (NSCLC) exhibits genomic heterogeneity that affects tumor immunogenicity and PD-L1 expression. Patient clustering based on shared mutational profiles using social network analysis (SNA) has been narrowly explored. The study aimed to identify subgroups of NSCLC patients with similar somatic mutation profiles using network-based modularity clustering, and to compare these groups with respect to PD-L1 expression, Tumor mutation burden (TMB), and clinical variables. Data of patients with stage 4 (metastatic) NSCLC, whose tumor tissue samples were collected between 2022 and 2024, were analyzed. This retrospective study included NSCLC patients harboring actionable driver mutations in genes such as *EGFR*, *KRAS*, *ALK*, *BRAF*, *MET*. A social network of 129 patients was constructed. Two distinct genomic clusters were identified. Cluster 2 ($n = 55$) showed a higher prevalence of *KRAS*, *TP53*, *BRAF*, *STK11* and additional mutations, while cluster 1 ($n = 74$) displayed a limited number of driver mutations. Cluster 2 had significantly higher PD-L1 expression (29.8% vs. 13.7%, $p = 0.001$) and higher TMB (7.8 vs. 5.8, $p = 0.021$). In multivariate logistic regression, both PD-L1 and TMB were associated with cluster assignment ($p < 0.05$). Mutation-based SNA clustering delineated two biologically distinct subgroups of NSCLC patients. The highly mutated cluster displayed higher PD-L1 expression and TMB, a profile consistent with a more immunogenic phenotype. This method offers a novel integrative approach that requires prospective validation before clinical implementation.

Academic Editor: Seth Frieze

Received: 13 May 2026

Revised: 20 June 2026

Accepted: 25 June 2026

Published: 26 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

Keywords: non-small cell lung cancer (NSCLC); actionable driver mutations; social network analysis (SNA); co-mutations; programmed death-ligand (PD-L1); Tumor mutation burden (TMB)

1. Introduction

Non-small cell lung cancer (NSCLC) accounts for about 85% of all lung cancer cases and is considered the leading cause of cancer mortality worldwide, with approximately 1.35 million new cases and 1.18 million deaths each year [1]. About 70% of patients are diagnosed at an advanced stage (stage 4), with a five-year survival rate of less than 5% [2].

Cancer cells are known to harbor genetic alterations, some of which drive cancer progression and oncogenesis. Detection of the driver mutation is important for diagnosis, prognosis and planning therapy [3]. Advances in genomic testing have fundamentally changed the treatment approach to lung cancer [4]. Mutations in *EGFR*, alterations in *ALK*, *ROS1*, *BRAF*, *MET*, and *KRAS* serve as therapeutic targets and are identified in a cumulative rate of about 20–40% of patients [5]. Targeted therapies (tyrosine kinase inhibitors, TKIs) have led to improvements in lung cancer/tumor response and time to disease progression; however, resistance usually develops within a year [6].

Additional progress in lung cancer treatment has occurred thanks to immunotherapy, particularly the blockade of the programmed cell death protein-1/programmed death-ligand-1 (PD-1/PD-L1) axis [7]. About 25–30% of pulmonary tumors express high levels of PD-L1 (presented as tumor proportion score (TPS) $\geq 50\%$), and respond better to treatment with drugs such as pembrolizumab, nivolumab, and atezolizumab [8]. Studies such as KEYNOTE-024 demonstrated a significant improvement in survival with pembrolizumab compared to chemotherapy in NSCLC patients expressing high levels of PD-L1 [8].

In addition, a high tumor mutational burden (TMB) in lung cancer is a biomarker associated with a higher likelihood of response to immunotherapy [9]. However, tumors with mutations in *EGFR* or *ALK* tend to have a suppressed immune microenvironment and a lower response to immunotherapy [10], in contrast to mutations in *KRAS*, *MET*, and *BRAF*, which may be more immunogenic [11].

PD-L1 expression varies greatly among groups based on mutational profile [12]. For example, over 50% of tumors with an *EGFR* mutation express PD-L1 [13].

TA correlation was found between different mutations and the expression of PD-L1 in tumors, for example: mutations that increase PD-L1 expression are *TP53*, *PTEN* LOSS, *MYC*, *KRAS*, *BRAF*, *MET* (indirectly) [14]. Mutations that suppress PD-L1 expression are *EGFR*, *ALK*, *STK11*, *KEAP1*, *NFE2L2* [15–19]. Moreover, there are mutations with variable effects on PD-L1 expression such as *PIK3CA*, *ERBB2*, *ROS1*, *RET* [20,21].

Possible clinical implications arise from the correlations between mutations and PD-L1 expression, for example, tumors harboring *KRAS* and *TP53* mutations have high PD-L1 expression, implying an inflammatory microenvironment, thus they have a high likelihood of response to immune checkpoint inhibitor (ICI)-based therapies. Tumors harboring *KRAS* and *STK11* mutations or *KRAS* and *KEAP1* were found to have low PD-L1 expression (cold immune environment), thus poor response to immunotherapy. Tumors driven by *EGFR* or *ALK* sometimes express PD-L1 but usually without significant immune infiltration; therefore, they have a limited benefit from PD-1/PD-L1 inhibitors [20–23].

The standard methods of genetic analysis often focus on a single mutation or a single biomarker [24]. However, a systemic approach that integrates multi-dimensional genomic information allows a deeper understanding of the connections between patients. Social

network analysis (SNA) offers an innovative methodology for identifying links between patients based on their somatic mutation profiles [25]. This approach enables the creation of groups of patients with genomic similarity—which may reveal subgroups with different biological and immune characteristics.

This approach is based on the pioneering work of Prof. Albert-László Barabási, who was among the first to apply the mathematics of social networks to biological systems. Barabási demonstrated how genes, proteins, and enzymes operate as part of a complex biological network, rather than as isolated individuals. These insights laid the foundation for the field of “Network Medicine”, which offers innovative ways to analyze diseases—and in this case, to study lung cancer [26].

The current study aims to examine whether the aforementioned correlations hold up in reality among NSCLC stage 4 patients at the Carmel Medical Center (CMC). To this end, a method that constructs a social network of patients according to shared mutations (Figure 1) is proposed and based on this network performs clustering of patients into groups with shared mutations. The PD-L1 expression in the different tumors will be compared.

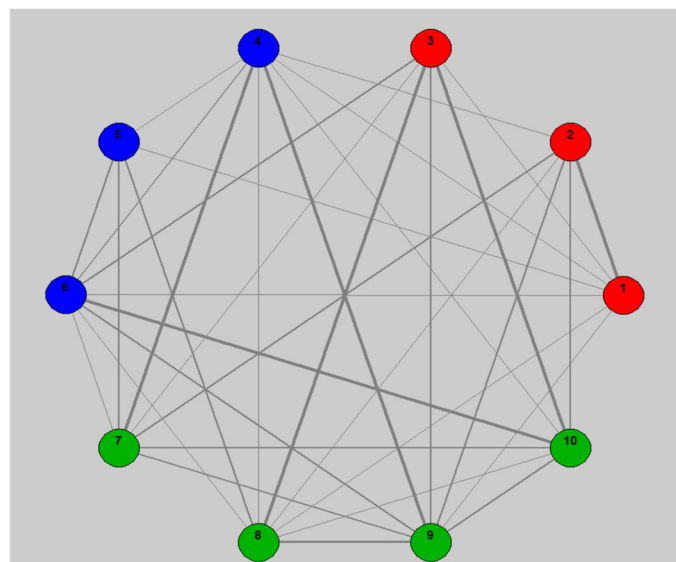


Figure 1. Illustration of a simulated example of a social network of patients linked by shared mutations. The color of the patients reflects the group, i.e., the shared cluster.

Research Hypothesis

Patients with metastatic NSCLC stage 4 lung cancer who are grouped according to shared somatic mutation profiles will show different PD-L1 expression profiles and different demographic characteristics, indicating variability in tumor immunogenicity and disease traits

Research Objectives

1. To identify groups of patients with similar genomic profiles using social network and modularity analysis.
2. To compare these groups according to PD-L1 expression levels in order to characterize the immunological differences between them.
3. To examine differences between the groups in demographic and clinical variables (age, gender, smoking, histology).

2. Materials and Methods

2.1. Study Design and Population

This single-center non-interventional retrospective study included 129 patients, aged 18 and older, diagnosed with stage 4 NSCLC who underwent genetic sequencing (NGS) and PD-L1 testing by immunohistochemistry at the Institute of Pathology of Carmel Medical Center [27]. Another inclusion criterion was patients having medical records data with demographics, clinical information, and lung tumor biopsy samples collected between 2022 and 2024. Exclusion criteria were patients aged <18 years old, incomplete medical records information or missing information of NSCLC tumor sample sequencing or PD-L1 testing.

2.2. Data Collection

Data Collection was composed of a patient identifier (pseudonym), somatic mutation profile (gene level such as KRAS, TP53, BRAF, or STK11), TPS result for PD-L1, age, gender, and ethnic origin. PD-L1 expression was assessed by immunohistochemistry and recorded as the tumor proportion score (TPS), defined as the percentage of viable tumor cells showing partial or complete membranous staining. Tumor mutational burden (TMB) and the somatic mutation profile were derived from the same next-generation sequencing (NGS) assay performed on the tumor biopsy samples.

2.3. Ethics

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 23 September 2024. Patient consent was waived due to the retrospective nature of the study.

2.4. Construction of the Social Network

The social network was constructed by the following steps: First, each patient was defined as a node. Second, an edge was created between any two patients sharing at least one somatic mutation. Third, the edge weight was determined by the number of shared mutations, so that patients with more genomic similarity received a stronger connection. Fourth, a weighted undirected network was obtained that represented the genomic proximity between patients. Somatic alterations were considered at the gene level: multiple variants within the same gene were collapsed to a single gene-level call, and copy-number changes and variants of uncertain relevance were not included in the similarity matrix. Edge weight was defined solely as the count of genes shared between two patients, without weighting for the functional importance, pathway membership, or predicted impact of individual mutations; consequently, driver and passenger gene-level calls contributed equally to network connectivity. This deliberately simple formulation is acknowledged as a limitation in Section 4.

2.5. Louvain Algorithm and Modularity Measure

The Louvain algorithm is a fast hierarchical algorithm for detecting clusters in large networks [28]. It operates by maximizing modularity—the measure of link density within a cluster compared to links between clusters. High modularity means “natural” clusters where there is significant similarity among cluster members and few external links. The algorithm iteratively applies merging and splitting of clusters until the optimal partition is found. In this study, two distinct patient clusters were obtained.

2.6. Statistical Analysis

2.6.1. Between-Group Comparisons

Continuous variables (age, PD-L1, TMB) were compared using a two-sample *t*-test (Student's *t*-test). Categorical variables (gender, ethnicity) were compared using χ^2 tests to examine between-group differences. A *p*-value of ≤ 0.05 was considered statistically significant. Owing to the limited number of pre-specified comparisons, *p*-values are reported without formal correction for multiple testing and are interpreted as exploratory. Analysis was performed using IBM SPSS version 26 (IBM, Armonk, NY, USA).

2.6.2. Logistic Regression

A multivariate logistic regression was performed, in which the dependent variable was cluster membership (1 or 2), and the independent variables were: age, gender, ethnicity, PD-L1, and TMB.

The goal was to identify variables associated with patient assignment to the genomic cluster.

During the preparation of this manuscript/study, the authors used the AI tools ChatGPT version 4.5, deep research to help identify mutation patterns within algorithmic clusters. The authors have reviewed and edited the output and take full responsibility for the content of this publication. During manuscript preparation, no patient-level, identifiable, or case-level genomic data were entered into the AI system. All clustering, statistical analyses, and interpretation of genomic patterns were derived from the statistical software outputs and the underlying dataset.

3. Results

3.1. Social Network Clusters

Following the construction of the social network and the operation of the Louvain algorithm, the patients were divided into two clusters based on the mutations they shared. Figure 2a illustrates the social network of the patients as nodes sharing mutations. In addition, Figure 2b shows the mutations as nodes sharing patients.

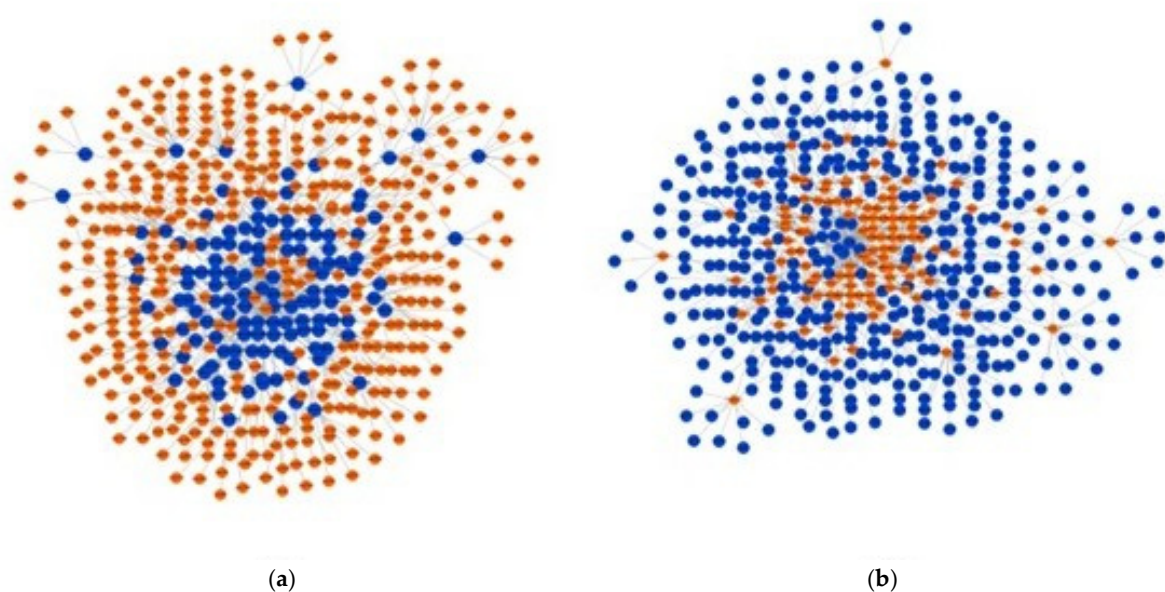


Figure 2. (a) A social network of patients (marked in blue), connected to each other through shared mutations (marked in orange). (b) A social network of mutations (marked in blue), connected to each other through shared patients (marked in orange).

The clustering process, that is, the grouping into two clusters of patients according to the mutations and co-mutations they share is illustrated in Figure 3.

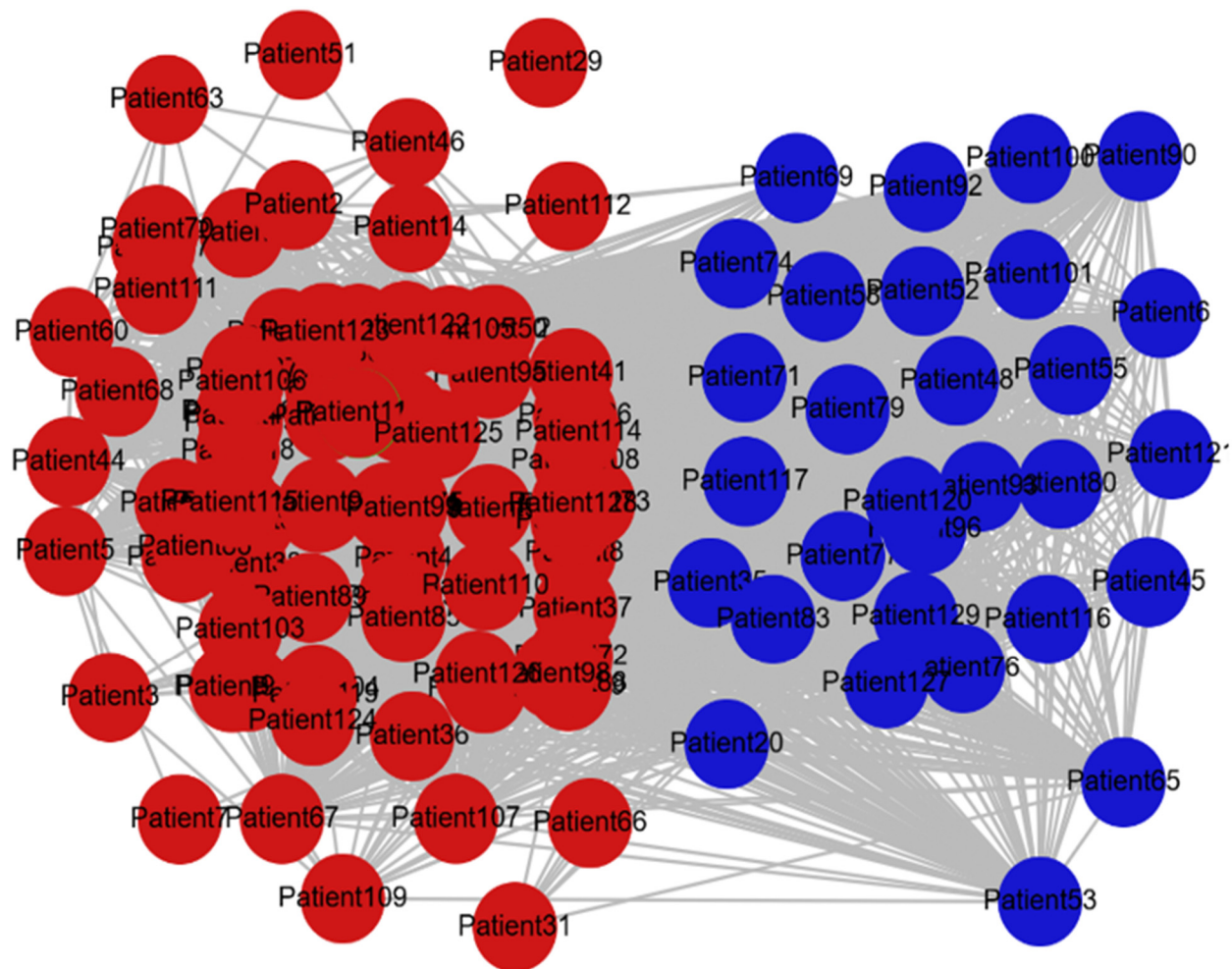


Figure 3. A network of patients connected through shared mutations. The thickness of the lines reflects the number of shared mutations. The colors indicate the two groups of the Louvain clustering. Cluster 1 is represented in red and Cluster 2 is represented in blue. The patients are connected to each other by the edges of the social network (the gray lines). The lines' thickness and number reflect the number of mutations or co-mutations linking the patients.

Table 1 describes two clusters of patients and the main mutations that are more relevant for treatment and/or for determining the prognosis of NSCLC. The patients grouped by the algorithm into cluster 1 show a smaller number of mutations compared to cluster 2. It should be noted that this table does not show the much larger number of additional genetic abnormalities that have been found, such as copy number variations and other mutations for which there is currently no direct information regarding their relevance in the somatic context of these cancers.

Table 1. Clusters of patients and their mutations.

Gene/Mutation	Cluster 1	Cluster 2
EGFR	2	22
ALK	2	4
ROS1	0	1
MET	2	15
RET	0	5

KRAS	0	38
BRAF	1	14
PIK3CA	2	7
TP53	4	50
STK11	0	25
KEAP1	0	1
SMARCA4	0	3
NRAS	0	1
DDR2	0	1
FGFR	0	6
PTEN	0	1
CDKN2	0	8
ARID	4	8
JAK2	1	1
Total relevant mutations	18	221

The tumors in cluster 2 are expected to be more immunogenic due to a larger number of mutations and co-mutations, and therefore they are expected to show higher levels of PD-L1 and TMB compared to the tumors from cluster 1, which have a smaller number of mutations and are therefore considered immunologically “cold,” and as a result, they are expected to show lower levels of PD-L1 and TMB.

3.2. Patients Demographics

Table 2 shows the demographic distribution of patients diagnosed with NSCLC after grouping the patients according to the mutations they share in common. There was no statistically significant difference in terms of age, gender, and ethnicity between the two clusters of mutations.

Table 2. Patient clusters demographics.

Variable	Cluster 1	Cluster 2	<i>p</i> -Value
Age (years), mean $\pm$ SD ¹	72 $\pm$ 1.1	70 $\pm$ 1.1	N.S. ¹ (>0.05)
Gender:			
Female (<i>n</i>)	39	22	N.S. ¹ (>0.05)
Male (<i>n</i>)	35	33	
Ethnicity:			
Arab (<i>n</i>)	16	12	N.S. ¹ (>0.05)
Jewish (<i>n</i>)	58	43	

¹ Abbreviations: SD, standard deviation; N.S. non-significant.

3.3. Expression of Biomarkers in Patients' Clusters

Table 3 shows the expression of PD-L1 and TMB values in the two clusters of patients according to the shared mutations. The clusters of each group of patients were formed by the social network algorithm, according to the common mutations shared by the patients. Cluster 2 patients expressed statistically significantly higher levels of both biomarkers, PD-L1 and TMB than cluster 1 patients.

Table 3. Biomarkers expression in co-mutation clusters of patients.

Variable	Cluster 1	Cluster 2	<i>p</i> -Value
PD-L1 ¹	13.7 $\pm$ 2.5 ²	29.8 $\pm$ 4.5	0.001
TMB ¹	5.8 $\pm$ 0.6 ³	7.8 $\pm$ 0.5	0.021

The values presented are mean $\pm$ SD. ¹ Abbreviations: PD-L1, Programmed Death-Ligand 1; TMB, Tumor Mutation Burden. ² The PD-L1 values in the table represent the percentage of positive staining in tumors. ³ TMB is presented as Mutation per megabase (Mut/Mb $\pm$ SD).

In clinical settings, biological treatment against PD-L1 is administered mainly according to a cutoff point of 50% [29]. Therefore, further analysis of PD-L1 expression levels was performed. Figure 4 shows a statistically significant difference between the two clusters upon differentiating/stratifying the PD-L1 index according to this 50% cutoff point. In cluster 2, a higher rate of patients had PD-L1 values above the 50% threshold compared to patients in cluster 1. This indicates that the tumors of cluster 2 are more immunogenic compared to cluster 1 tumors.

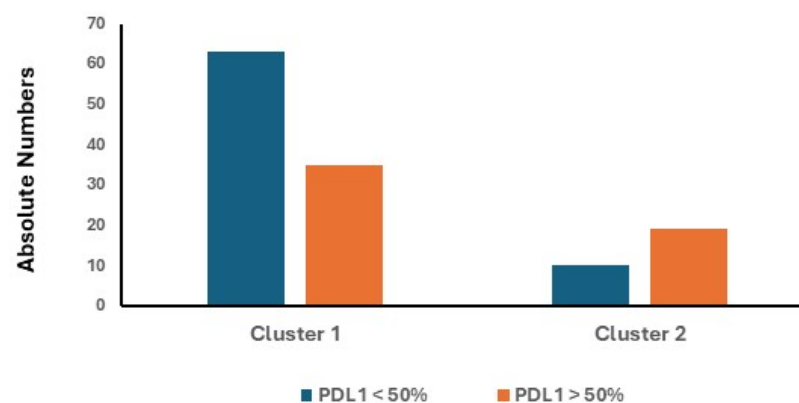


Figure 4. PD-L1 stratification across clusters. This graph refers to the number of patients who were found to have PD-L1 values higher or lower than the 50% cutoff point. $p = 0.004$. OR = 3.4.

A multivariable logistic regression analysis that included both the demographic details and the PD-L1 and TMB markers was performed in order to examine the association between the two clusters of patients according to the mutations they share (Table 4). In the multivariable analysis, the two molecular markers, PD-L1 and the TMB index are considered independent variables in predicting the two patient clusters.

Table 4. Multivariable logistic regression analysis.

Variable	Betas	<i>p</i> -Values
TMB ¹	0.107	0.032
PD-L1 ¹	0.016	0.022
Constant	−1.231	

¹ Abbreviations: PD-L1, Programmed Death-Ligand 1; TMB, Tumor Mutation Burden.

4. Discussion

In this study, we applied a network-based approach grounded in Network Medicine principles to classify patients with non-small cell lung cancer according to shared somatic mutation profiles rather than individual driver alterations [30,31]. Using social network analysis and modularity optimization via the Louvain algorithm, we identified two biologically distinct patient clusters characterized by markedly different genomic complexity and immunologic features. An “immunogenic” cluster harboring multiple mutations (cluster 2) versus an “immune-cold” cluster (cluster 1).

To our knowledge, this is among the first studies applying patient-level social network modeling to co-mutation structure in NSCLC. The study includes multiple

complex analyses; first, the patient-to-patient mutation links, second, biomarkers and combining the interactions into clusters.

The principal finding of this study is that patients grouped into the highly connected cluster (cluster 2), defined by a greater burden of co-mutations—including *KRAS*, *TP53*, *STK11*, and *BRAF*—exhibited significantly higher levels of PD-L1 expression and tumor mutational burden compared to patients in cluster 1. These findings suggest that genomic complexity, as captured by network connectivity, is associated with increased tumor immunogenicity.

This association should be interpreted with appropriate caution. The network was constructed from the count of shared somatic mutations, and TMB is itself a measure of mutational burden derived from the same sequencing assay. The higher TMB observed in the more densely connected cluster is therefore partly expected by construction rather than an independent confirmation of the clustering, and the relationship between network connectivity and TMB is not fully independent of the input data. The PD-L1 finding is less directly coupled to the network input, because PD-L1 expression was measured separately by immunohistochemistry; nevertheless, both biomarkers covary with the underlying mutational profile. Accordingly, the present results are best regarded as demonstrating internal consistency between the mutation-derived clustering and established immunogenicity markers, rather than as external validation of the clusters.

The current study findings are consistent with prior studies demonstrating that specific co-mutation patterns influence the tumor immune microenvironment [32]. In particular, pulmonary tumors harboring co-occurring mutational alterations in *KRAS* and *TP53* have been associated with immunogenic inflamed tumor phenotypes and elevated PD-L1 expression [33], whereas combinations such as *KRAS* with *STK11* are often linked to immunologically “cold” tumors and resistance to immune checkpoint inhibitors. Thus, the later tumors’ molecular profile presents low TMB and PD-L1 levels, indicating a lower response to immunotherapy [34].

The coexistence of these mutations within cluster 2 likely reflects a heterogeneous but overall more immunogenic genomic landscape, which may explain the higher PD-L1 and TMB levels observed.

Interestingly, although mutations such as *STK11* are typically associated with reduced PD-L1 expression, their presence within a broader co-mutational context did not negate the overall immunogenic phenotype of cluster 2. This highlights an important concept: the biological behavior of tumors may depend more on combinatorial mutational patterns than on single-gene effects, supporting the rationale for integrative analytical approaches such as social network analysis [35,36].

The fact that cluster 2, which includes the *STK11* mutation, had a high PD-L1 expression level, implies the dominance of co-mutation context rather than a single mutation impact such as *STK11*, which is usually affiliated with low PD-L1 [19]. The distinction between clusters was due solely to biological-molecular factors, and not clinical or social variables.

Another notable finding is the absence of significant differences in demographic variables between clusters. This strengthens the interpretation that the observed differences in PD-L1 expression and TMB are primarily driven by tumor-intrinsic genomic features rather than patient-related factors [34,37]. From a clinical perspective, this suggests that network-derived genomic classification [38] could provide independent and complementary information beyond standard clinical stratification.

A significant advantage of this network-based approach is the ability to characterize patients with complex co-mutation patterns that are not identified in single-mutation testing—for example, *KRAS* with *TP53* versus *KRAS* with *STK11*, which have opposing

immunological profiles. Networks of this type may enable better tailoring of immunotherapy treatments in the future.

The clinical implications of these findings are potentially significant. Both PD-L1 expression and TMB are established biomarkers used to guide immunotherapy decisions in NSCLC [21,39]. The ability of a network-based model to identify patient subgroups enriched for high PD-L1 and TMB raises the possibility that such approaches could improve patient selection for immune checkpoint inhibitor therapies. Moreover, this methodology may help refine current biomarker strategies by incorporating the complexity of co-mutation patterns rather than relying on single biomarkers alone.

Luo et al. [40] stratified NSCLC patients according to somatic mutations by using metabolic-based clustering of computational biology and network propagation algorithm. That approach was validated through classification of NSCLC patient subtypes evaluated for immunotherapy followed by survival benefits.

This study has several strengths, including the use of real-world genomic and clinical data, and the application of an innovative analytical framework that captures relationships between patients rather than isolated variables. However, several limitations should be acknowledged. First, the retrospective single-center design may limit generalizability. Second, the sample size is modest, particularly for subgroup analyses. Third, the study lacks clinical outcome data, such as the response to immunotherapy or survival, which would be necessary to confirm the predictive value of the identified clusters. In addition, potential technical variability in PD-L1 assessment and TMB calculation may introduce measurement bias. Several further limitations warrant emphasis. The edge weight depended only on the number of shared gene-level alterations and did not incorporate the functional importance, pathway context, or predicted impact of individual mutations, so that driver and likely passenger alterations contributed equally to connectivity; this may influence cluster composition. Because the network and TMB were both derived from the same sequencing data, the higher TMB of the more connected cluster is partly inherent to the construction of the network and does not constitute independent validation. The proposed network-based clustering was not formally benchmarked against simpler or established alternatives—such as raw mutation count, hierarchical or Jaccard-distance clustering, or groups defined directly by KRAS, TP53, and STK11 status—so the incremental value of the social network approach over these comparators remains to be established. The smoking status and histologic subtype, although listed among the study objectives, were not consistently available in the source records and could not be incorporated into the demographic comparison or the regression model. Other potential confounders, including prior treatment, tissue source, tumor cellularity, and the timing of sampling, were likewise not adjusted for. The PD-L1 and TMB assays were not standardized across a single platform with documented inter-observer review, and the analysis was restricted to gene-level point mutations, without copy-number variation, gene expression, or epigenetic data. Finally, as a single-institution cohort drawn from one regional population, the findings may be sensitive to dataset composition and require external, multi-center validation.

Future research should focus on validating these findings in larger, multi-center cohorts and integrating clinical outcomes to assess the predictive utility of network-defined clusters. Expanding this approach to include additional genomic features—such as copy number variations, gene expression profiles, and epigenetic alterations—may further enhance its biological and clinical relevance. Furthermore, applying similar network-based methodologies across different tumor types could help to determine whether these patterns are tumor-specific or if they represent a broader principle in oncology. To establish that the social network approach adds value, future analyses should formally benchmark it against simpler comparators, including raw mutation

count, hierarchical and Jaccard-distance clustering, and groups defined directly by KRAS, TP53, and STK11 status. Furthermore, they should incorporate functional weighting of mutations and a formal assessment of cluster stability. To support reproducibility, the de-identified patient-by-gene matrix, cluster assignments, biomarker values, and the analysis code and clustering settings should be made available.

5. Conclusions

Building a social network based on shared mutations is a feasible approach for identifying genomic subgroups in NSCLC. The cluster with multiple mutations (cluster 2) was characterized by higher levels of PD-L1 and TMB, a profile consistent with a more immunogenic phenotype. In multivariable analysis, PD-L1 and TMB were associated with cluster assignment, whereas demographic features were not. Because the cohort was retrospective and lacked immunotherapy-response and survival data, these results are hypothesis-generating: a social network-based approach may, pending prospective validation, serve as a complementary tool for characterizing co-mutation patterns relevant to immunotherapy selection.

Author Contributions: Conceptualization, A.A., H.N. and D.L.-F.; methodology, K.M. and E.S.; software, W.S. and E.L.; validation, S.M., A.A., H.N. and D.L.-F.; formal analysis, K.M. and E.S.; investigation, W.S. and E.L.; resources, A.A., H.N., M.S.-A., L.S., D.L.-F., V.S., D.R. and E.S.; data curation, K.M. and W.S.; writing—original draft preparation, H.N. and A.A.; writing—review and editing, M.S.-A., L.S., D.L.-F., V.S., D.R. and E.S.; visualization, W.S. and E.L.; supervision, A.A., H.N., D.L.-F. and E.S.; project administration A.A., H.N., D.L.-F. and E.S. All authors 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 23 September 2024.

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

Data Availability Statement: The data that support the findings of this study are available from the Corresponding Author, [A.A.], upon reasonable request. To support reproducibility, a de-identified patient-by-gene matrix, the cluster assignments, the corresponding PD-L1 and TMB values, and the analysis code together with the network-construction and Louvain settings will be provided to qualified researchers upon reasonable request. Patient-level genomic data are not deposited in a public repository because of institutional and ethics-committee restrictions on potentially identifiable information.

Acknowledgments: During the preparation of this manuscript/study, the authors used the AI tools ChatGPT deep research to help identify mutation patterns within algorithmic clusters. The authors have reviewed and edited the output and take full responsibility for the content of this publication. During manuscript preparation, no patient-level, identifiable, or case-level genomic data were entered into the AI system. All clustering, statistical analyses, and interpretation of genomic patterns were derived from the statistical software outputs and the underlying dataset.

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

Abbreviations

The following abbreviations are used in this manuscript:

CMC	Carmel Medical Center
NSCLC	Non-Small Cell Lung Cancer
PD-1	Programmed cell death-1
PD-L1	Programmed Death-Ligand 1
SNA	Social Network Analysis
TKI	Tyrosine Kinase Inhibitor
TMB	Tumor Mutation Burden
TPS	Tumor Proportion Score

References

1. Rocco, D.; Della Gravara, L.; Boccia, M.C.; Palazzolo, G.; Gridelli, C. Novel Combination of Therapeutic Approaches in Advanced NSCLC with EGFR Activating Mutations. *Targets* **2024**, *2*, 237–249. <https://doi.org/10.3390/targets2030014>.
2. American Cancer Society. *Lung Cancer Survival Rates, by Stage*. American Cancer Society. 2024. Available online: <https://www.cancer.org/cancer/types/lung-cancer/detection-diagnosis-staging/survival-rates.html> (accessed on 2 October 2025).
3. Ostroverkhova, D.; Przytycka, T.M.; Panchenko, A.R. Cancer driver mutations: Predictions and reality. *Trends Mol. Med.* **2023**, *29*, 554–566. <https://doi.org/10.1016/j.molmed.2023.03.007>.
4. Vicidomini, G. Current Challenges and Future Advances in Lung Cancer: Genetics, Instrumental Diagnosis and Treatment. *Cancers* **2023**, *15*, 3710. <https://doi.org/10.3390/cancers15143710>.
5. Chevallier, M.; Borgeaud, M.; Addeo, A.; Friedlaender, A. Oncogenic driver mutations in non-small cell lung cancer: Past, present and future. *World J. Clin. Oncol.* **2021**, *12*, 217–237. <https://doi.org/10.5306/wjco.v12.i4.217>.
6. Mok, T.S.; Wu, Y.-L.; Thongprasert, S.; Yang, C.-H.; Chu, D.-T.; Saijo, N.; Sunpaweravong, P.; Han, B.; Margono, B.; Ichinose, Y.; et al. Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. *New Engl. J. Med.* **2009**, *361*, 947–957. <https://doi.org/10.1056/NEJMoa0810699>.
7. Lahiri, A.; Maji, A.; Potdar, P.D.; Singh, N.; Parikh, P.; Bisht, B.; Mukherjee, A.; Paul, M.K. Lung cancer immunotherapy: Progress, pitfalls, and promises. *Mol. Cancer* **2023**, *22*, 40. <https://doi.org/10.1186/s12943-023-01740-y>.
8. Reck, M.; Rodríguez-Abreu, D.; Robinson, A.G.; Hui, R.; Csőszi, T.; Fülöp, A.; Gottfried, M.; Peled, N.; Tafreshi, A.; Cuffe, S.; et al. Pembrolizumab versus chemotherapy for PD-L1–positive non-small-cell lung cancer. *New Engl. J. Med.* **2016**, *375*, 1823–1833. <https://doi.org/10.1056/NEJMoa1606774>.
9. Hellmann, M.D.; Ciuleanu, T.-E.; Pluzanski, A.; Lee, J.S.; Otterson, G.A.; Audigier-Valette, C.; Minenza, E.; Linardou, H.; Burgers, S.; Salman, P.; et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. *Cancer Cell* **2018**, *33*, 345–354.e4. <https://doi.org/10.1056/NEJMoa1801946>.
10. To, K.K.W.; Fong, W.; Cho, W.C.S. Immunotherapy in Treating EGFR-Mutant Lung Cancer: Current Challenges and New Strategies. *Front. Oncol.* **2021**, *11*, 635007. <https://doi.org/10.3389/fonc.2021.635007>.
11. Skoulidis, F.; Goldberg, M.E.; Greenawalt, D.M.; Hellmann, M.D.; Awad, M.M.; Gainor, J.F.; Schrock, A.B.; Hartmaier, R.J.; Trabucco, S.E.; Gay, L.; et al. STK11/LKB1 mutations and PD-1 inhibitor resistance in KRAS-mutant lung adenocarcinoma. *Cancer Discov.* **2018**, *8*, 822–835. <https://doi.org/10.1158/2159-8290.CD-18-0099>.
12. Sarova, P.; Mosleh, B.; Zehetmayer, S.; Oberndorfer, F.; Widder, J.; Prosch, H.; Aigner, C.; Idzko, M.; Hoda, M.A.; Gompelmann, D. PD-L1 expression in patients with non-small-cell lung cancer is associated with sex and genetic alterations: A retrospective study within the Caucasian population. *Thorac. Cancer* **2024**, *15*, 1598–1606. <https://doi.org/10.1111/1759-7714.15336>.
13. Lee, C.K.; Man, J.; Lord, S.; Links, M.; Gebiski, V.; Mok, T.; Yang, J.C.-H. Checkpoint Inhibitors in Metastatic EGFR-Mutated Non-Small Cell Lung Cancer-A Meta-Analysis. *J. Thorac. Oncol.* **2017**, *12*, 403–407. <https://doi.org/10.1016/j.jtho.2016.10.007>.
14. Liu, Y.; Wu, A.; Li, X.; Wang, S.; Fang, S.; Mo, Y. A retrospective analysis of eleven gene mutations, PD-L1 expression and characteristics in non-small cell lung cancer patients. *Asian J. Surg.* **2022**, *45*, 367–375. <https://doi.org/10.1016/j.asjsur.2021.06.030>.
15. Li, J.; Shi, D.; Li, S.; Shi, X.; Liu, Y.; Zhang, Y.; Wang, G.; Zhang, C.; Xia, T.; Piao, H.-L.; et al. KEAP1 promotes anti-tumor immunity by inhibiting PD-L1 expression in NSCLC. *Cell Death Dis.* **2024**, *15*, 175. <https://doi.org/10.1038/s41419-024-06563-3>.
16. Gao, W.; Wang, L.; Zhao, Y.; Zhu, L. The role of PD-L1 in EGFR-mutant non-small cell lung cancer. *Discov. Oncol.* **2025**, *16*, 307. <https://doi.org/10.1007/s12672-025-02089-y>.

17. Liu, S.-Y.; Dong, Z.-Y.; Wu, S.-P.; Xie, Z.; Yan, L.-X.; Li, Y.-F.; Yan, H.-H.; Su, J.; Yang, J.-J.; Zhou, Q.; et al. Clinical relevance of PD-L1 expression and CD8+ T cells infiltration in patients with EGFR-mutated and ALK-rearranged lung cancer. *Lung Cancer* **2018**, *125*, 86–92. <https://doi.org/10.1016/j.lungcan.2018.09.010>.
18. Ruiz, G.; Enrico, D.; Mahmoud, Y.D.; Ruiz, A.; Cantarella, M.F.; Leguina, L.; Barberis, M.; Beña, A.; Brest, E.; Starapoli, S.; et al. Association of PD-L1 expression with driver gene mutations and clinicopathological characteristics in non-small cell lung cancer: A real-world study of 10441 patients. *Thorac. Cancer* **2024**, *15*, 895–905. <https://doi.org/10.1111/1759-7714.15244>.
19. Khalil, M.; Tsao, M.-S. Tumour- and Non-Tumour-Associated Factors That Modulate Response to PD-1/PD-L1 Inhibitors in Non-Small Cell Lung Cancer. *Cancers* **2025**, *17*, 2199. <https://doi.org/10.3390/cancers17132199>.
20. Chu, C.-H.; Huang, Y.-H.; Lee, P.-H.; Hsu, K.-H.; Chen, K.-C.; Su, K.-Y.; Yu, S.-L.; Tseng, J.-S.; Yang, T.-Y.; Chang, G.-C. Various impacts of driver mutations on the PD-L1 expression of NSCLC. *PLoS ONE* **2022**, *17*, e0273207. <https://doi.org/10.1371/journal.pone.0273207>.
21. Chen, Y.; Wang, P.; Lian, R.; Yuan, M.; Yu, P.; He, H.; Chen, P.; Zhou, H.; Chen, W.; Zhang, D.; et al. Comprehensive characterization of PD-L1 expression and immunotherapy-related genomic biomarkers in early- versus advanced-stage non-small cell lung cancer. *BMC Pulm. Med.* **2025**, *25*, 219. <https://doi.org/10.1186/s12890-025-03687-w>.
22. Tønnesen, E.M.T.; Stougaard, M.; Meldgaard, P.; Lade-Keller, J. Prognostic value of KRAS mutations, TP53 mutations and PD-L1 expression among lung adenocarcinomas treated with immunotherapy. *J. Clin. Pathol.* **2023**, *77*, 54–60. <https://doi.org/10.1136/jcp-2022-208574>.
23. Proulx-Rocray, F.; Routy, B.; Nassabein, R.; Belkaid, W.; Tran-Thanh, D.; Malo, J.; Tonneau, M.; El Ouarzadi, O.; Florescu, M.; Tehfe, M.; et al. The prognostic impact of KRAS, TP53, STK11 and KEAP1 mutations and their influence on the NLR in NSCLC patients treated with immunotherapy. *Cancer Treat. Res. Comm.* **2023**, *37*, 100767. <https://doi.org/10.1016/j.ctarc.2023.100767>.
24. Dan, A.; Burtavel, L.-M.; Coman, M.-C.; Focsa, I.-O.; Duta-Ion, S.; Juganaru, I.-R.; Zaruha, A.-G.; Codreanu, P.-C.; Strugari, I.-M.; Hotinceanu, I.-A.; et al. Genetic Blueprints in Lung Cancer: Foundations for Targeted Therapies. *Cancers* **2024**, *16*, 4048. <https://doi.org/10.3390/cancers16234048>.
25. Grewal, E.; Godley, J.; Wheeler, J.; Tang, K.L. Use of social network analysis in health research: A scoping review protocol. *BMJ Open* **2024**, *14*, e078872. <https://doi.org/10.1136/bmjopen-2023-078872>.
26. Barabási, A.L.; Gulbahce, N.; Loscalzo, J. Network medicine: A network-based approach to human disease. *Nat. Rev. Genet.* **2011**, *12*, 56–68. <https://doi.org/10.1038/nrg2918>.
27. Agbarya, A.; Shalata, W.; Sabo, E.; Saiegh, L.; Shaham, Y.; Nasrallah, H.; Mhameed, K.; Mazareb, S.; Sheikh-Ahmad, M.; Faber, D.L. Investigation of the Prevalence of Associated Genetic Mutations (Co-Mutations) in Patients with Actionable Driver Mutations in Lung Cancer: A Retrospective Study. *Diagnostics* **2026**, *16*, 1106. <https://doi.org/10.3390/diagnostics16071106>.
28. Blondel, V.D.; Guillaume, J.L.; Lambiotte, R.; Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.* **2008**, *2008*, P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>.
29. Koury, J.; Lucero, M.; Cato, C.; Chang, L.; Geiger, J.; Henry, D.; Hernandez, J.; Hung, F.; Kaur, P.; Teskey, G.; et al. Immunotherapies: Exploiting the Immune System for Cancer Treatment. *J. Immunol. Res.* **2018**, *2018*, 9585614. <https://doi.org/10.1155/2018/9585614>.
30. Cucchiara, F.; Crucitta, S.; Petrini, I.; Perez, D.d.M.; Ruglioni, M.; Pardini, E.; Rolfo, C.; Danesi, R.; Del Re, M. Gene-network analysis predicts clinical response to immunotherapy in patients affected by NSCLC. *Lung Cancer* **2023**, *183*, 107308. <https://doi.org/10.1016/j.lungcan.2023.107308>.
31. Wang, X. Gene mutation-based and specific therapies in precision medicine. *J. Cell Mol. Med.* **2016**, *20*, 577–580. <https://doi.org/10.1111/jcmm.12722>.
32. Ma, L.; Zhu, L.-J.; Xu, Z.-J.; Shu, Y.-Q. Co-mutations in tumor immune microenvironment and immunotherapy. *Chin. Med. J.* **2021**, *134*, 1055–1057. <https://doi.org/10.1097/CM9.0000000000001455>.
33. Gu, M.; Xu, T.; Chang, P. KRAS/LKB1 and KRAS/TP53 co-mutations create divergent immune signatures in lung adenocarcinomas. *Ther. Adv. Med. Oncol.* **2021**, *13*, 17588359211006950. <https://doi.org/10.1177/17588359211006950>.
34. Di Federico, A.; Alden, S.; Smithy, J.; Ricciuti, B.; Alessi, J.; Wang, X.; Pecci, F.; Lamberti, G.; Gandhi, M.; Vaz, V.; et al. Interpatient variation in PD-L1 expression and tumor mutational burden and the impact on outcomes to immune checkpoint inhibitor therapy in patients with non-small-cell lung cancer. *Ann. Oncol.* **2024**, *35*, 902–913. <https://doi.org/10.1016/j.annonc.2024.06.014>.
35. Jiang, L.; Yu, H.; Ness, S.; Mao, P.; Guo, F.; Tang, J.; Guo, Y. Comprehensive Analysis of Co-Mutations Identifies Cooperating Mechanisms of Tumorigenesis. *Cancers* **2022**, *14*, 415. <https://doi.org/10.3390/cancers14020415>.

36. Sinkala, M. Mutational landscape of cancer-driver genes across human cancers. *Sci. Rep.* **2023**, *13*, 12742. <https://doi.org/10.1038/s41598-023-39608-2>.
37. Yarchoan, M.; Albacker, L.A.; Hopkins, A.C.; Montesin, M.; Murugesan, K.; Vithayathil, T.T.; Zaidi, N.; Azad, N.S.; Laheru, D.A.; Frampton, G.M.; et al. PD-L1 expression and tumor mutational burden are independent biomarkers in most cancers. *JCI Insight* **2019**, *4*, e126908. <https://doi.org/10.1172/jci.insight.126908>.
38. Hofree, M.; Shen, J.P.; Carter, H.; Gross, A.; Ideker, T. Network-based stratification of tumor mutations. *Nat. Methods* **2013**, *10*, 1108–1115. <https://doi.org/10.1038/nmeth.2651>.
39. Dai, M.; Lucas, F.A.S.; Alvarez, H.; Ballester, L.; Chen, H.; Patel, K.P.; Rashid, A.; Rao, S.; Routbort, M.J.; Sura, G.; et al. Unraveling the nexus: Tumor mutational burden, PD-L1 expression, and oncogenic alterations in non-small cell lung cancer cytology specimens. *Cancer Cytopathol.* **2026**, *134*, e70090. <https://doi.org/10.1002/cncy.70090>.
40. Luo, X.; Zhang, X.; Su, D.; Li, H.; Zou, M.; Xiong, Y.; Yang, L. Deep Clustering-Based Metabolic Stratification of Non-Small Cell Lung Cancer Patients Through Integration of Somatic Mutation Profile and Network Propagation Algorithm. *Interdiscip. Sci. Comput. Life Sci.* **2025**, *17*, 948–969. <https://doi.org/10.1007/s12539-025-00699-2>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.