VIA MEDICA

ORIGINAL ARTICLE

Oncology in Clinical Practice
DOI: 10.5603/ocp.98228

ISSN 2450-1654
eISSN 2450-6478

Higher neoantigen load correlates with better overall survival in Chinese lung adenocarcinoma patients

Xuemei Liu[1], Yongfeng Yu[1], Wenhui Song[2], Xiao Deng[2], Kaili Gu[2], Mengdi Yuan[2], Zeyu Jiang[2], Yu Wang[2,*], Yafei Zhang[2]

[1]Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China
[2]MEDx (Suzhou) Translational Medicine Co., Ltd, Suzhou, China

Abstract

Introduction. Neoantigen load (NAL) has been extensively studied as a promising biomarker for immunotherapy. Recently it was also reported that NAL is associated with lung cancer patient survival, but the results were not consistent.

Material and methods. To further evaluate the prognostic value of NAL in lung cancer, we analyzed NAL in a cohort of 96 lung adenocarcinoma (AD) and 83 lung squamous cell carcinoma (SQ) patients from the Cancer Genome Atlas (TCGA). We found that high NAL correlates with better overall survival (OS) of AD patients but with worse OS of SQ patients. Next, we collected a total of 25 NSCLC patient samples and explored whole exome sequencing (WES) and a large targeted gene panel (Med1CDx panel containing 579 genes) for NAL and tumor mutation burden (TMB) analysis.

Results. We found that patients with both higher NAL and TMB, who underwent chemotherapy combined with immunotherapy, showed better OS and progression-free survival (PFS) in both AD and SQ subgroups. We also compared the concordance of NAL and TMB between WES and the Med1CDx panel. The R^2 for concordance of NAL and TMB prediction by WES and our Med1CDx panel was 0.81 and 0.86, respectively.

Conclusions. In this study, we showed that NAL is a useful biomarker for lung cancer OS prediction at least in the AD cohort. Furthermore, considering the high cost of WES, large targeted gene-panel-based NAL and TMB analysis could be a good alternative in clinical practical settings.

Keywords: neoantigen load, overall survival, lung adenocarcinoma, Chinese patients

Introduction

Lung cancer has been the leading cause of death worldwide and the 2nd common cancer type in 2020, accounting for 1.8 million cases of 10 million deaths in 2020 [World Health Organization (WHO) website] [1]. Advanced molecular diagnostics and recognition of targetable oncogenic driver alterations have led to dramatic changes in non-small cell lung cancer (NSCLC) treatment in recent years. Many new effective targeted agents were developed and the treatment of some oncogene-addicted NSCLC, such as *EGFR*-mutated or *ALK*-rearranged NSCLC is well-established [2, 3]. Although these drugs have revolutionized clinical practice, only a fraction of susceptible patients will benefit, and acquired resistance to these agents remains a challenge. Individualized vaccines targeting neoantigens would be a good option for lung cancer therapy in the future. Neoantigens are protein fragments derived only from cancer cells. With this unique property, targeting neoantigen allows the patient's immune system to detect and attack cancer cells instead of attacking healthy cells [4, 5]. Neoantigens are classified into two types: shared and personalized neoantigens. While shared neoantigens are not specific to an individual or tumor type, personalized neoantigens are highly specific to individual

*Correspondence: Yu Wang, PhD, MEDx (Suzhou)Translational Medicine Co., Ltd, Suite 901, B5 Building, 218 Xinghu St., Suzhou Industrial Park, 215123 Suzhou, China (yu.wang@medxtmc.com)
Received: 16 November 2023; Accepted: 16 February 2024; Early publication: 8 April 2024

tumors and are the basis of personalized neoantigen vaccines.

Identifying prognostic markers in cancer patients is essential because it allows the recognition of patient subpopulations that might anticipate different outcomes or might benefit from different types of therapies. Novel molecular prognostic biomarkers include such genes as *TP53*, *VEGF*, *TUBB3*, *Ki-67*, etc. However, despite an enormous amount of data available on molecular biomarkers, results are often not reproducible, partially due to the heterogeneity of study designs, techniques used, and data interpretation. Therefore, many molecular prognostic markers to date, have not managed to make their way into routine clinical use [6, 7]. Tumor mutation burden (TMB) and neoantigen load (NAL) have been extensively studied as promising biomarkers for predicting the anti-tumor effects of immune checkpoint inhibitors (ICIs) [8–12]. Several studies have reported the prognostic effect of TMB on the clinical benefits for patients with resected early-stage NSCLC, but the results are inconsistent. Two studies showed that a high TMB is associated with a favorable outcome in resected NSCLC patients [13, 14], but other reports demonstrated that TMB is not associated with overall survival of early-stage NSCLC patients, implying that TMB is not sufficient to predict NSCLC prognosis [15, 16]. However, utilizing computational tools to predict tumor NAL based on whole exome sequencing (WES) data has been confirmed to be a potentially useful method [17]. Recently, several studies have shown that NAL has good potential as a prognosis biomarker although the results were also contradictory. Gong et al. [18] showed that higher NAL exhibited better disease-free survival (DFS) for stage II/III Chinese lung squamous cell carcinoma (SQ) patients. However, another report demonstrated that high neoantigen burden was associated with significantly longer overall survival (OS) in the lung adenocarcinoma (AD) cohort of patients from Cancer Genome Atlas (TCGA) [19].

Tumor mutation burden and NAL analysis by WES is complicated and expensive due to large genomic space sequencing. Recent studies have shown that TMB can be accurately measured by smaller gene panels [9, 10, 20]. In this study, we assessed TMB and NAL by both WES and a large targeted gene panel and identified the correlation of TMB and NAL with clinical outcomes. The concordance of TMB and NAL measurements by WES and the large targeted gene panel was also determined. We expected to be able to provide more insights into biomarker discovery and identification for the prognosis of Chinese lung cancer patients. Personalized neoantigens predicted by WES were also compared with those from online public databases.

Material and methods

Cancer Genome Atlas data retrieval, neoantigen load calculation, and survival analysis

Thorsson et al. [21] presented an immunogenomic analysis of more than 10 000 tumors comprising 33 diverse cancer types by utilizing data compiled by TCGA. Clinical data for lung cancer patients was accessed and downloaded from https://gdc.cancer.gov/about-data/publications/panimmune. Predicted single nucleotide variant (SNV) and Indel neoantigen counts are available in this dataset. So we used predicted SNV, Indel neoantigen counts, and the sum of these two (NeoAll) to correlate with OS. Survfit objects were generated by surv_categorize() and survfit.formula() from R package survminer and survival. The ggsurvplot() function from R package ggplot2 was used to plot the Kaplan-Meier survival curve.

Patient cohort

A total of 25 patients with pathologically confirmed NSCLC, including 17 AD and 8 SQ patients, were enrolled between 2017 Feb and 2018 Nov. Clinical data were retrieved from the electronic medical records. Data acquisition was in line with relevant legislation and institutional review board guidelines. All patients were followed up regularly in the Shanghai Chest Hospital until recurrent or last follow-up. All procedures involving human participants performed in this study were in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional committee board of the Shanghai Chest Hospital and informed consent was taken from all the participants.

Whole exome and large targeted panel sequencing and tumor mutation burden calculation

Next-generation sequence (NGS) was performed using genomic DNA isolated from formalin-fixed paraffin-embedded (FFPE) samples with WES and a large targeted panel (Med1CDx panel including full coding sequences (CDS) regions of 579 genes, selected introns of 37 genes, which is designed for fusion calling). Deep sequencing was performed by the Hiseq X10 or NovaSeq platforms with a mean depth of 5000X. The bioinformatics workflow utilized a customized variant calling method based on GATK4 and Varscan, which contains SNP/InDel/CNV/SV calling.

Tumor mutation burden was determined as the average number of coding mutations per megabase (Mb) of genome examined following the method of FoundationOne panel [22]. All single nucleotide variations and Indels in the coding region including the synonymous alternation of targeted genes were counted for TMB calculation. Alterations listed as known somatic alterations in COSMIC hot spots were excluded, and truncations in tumor suppressor genes were not

counted. Sites presented in 1000G, ESP6500, and gnomAD with ≥ 1% frequency and synonymous SNV sites were filtered. The threshold for high TMB was determined by the 25th and 75th percentiles method used by the FoundationOne panel [22].

Human leukocyte antigen typing, neoantigen prediction, and statistics analysis

Four-digit human leukocyte antigen (HLA) class I (HLA-A, HLA-B, and HLA-C) alleles of each patient were identified from WES data using Optitype (version 1.3.1, default parameters). The pvactools (version 1.5.8, default parameters) tool was used to predict binding of 8- to 11-mer mutant peptides to the patients' HLA alleles. Neoantigens predicted in this study were compared with those from the TSNAdb (version 4.0) database. Neoantigen load was calculated by dividing the total number of neoantigens in each sample by the CDS length (Mb) of the WES or Med1CDx panel. The Kaplan-Meier method and the log-rank test were performed to correlate survival of patients with genomic alterations and NAL. A $p \leq 0.05$ was considered statistically significant.

Mutation spectra analysis and comparison between lung adenocarcinoma and lung squamous cell carcinoma patients

Raw sequence variants were called from WES data according to the GATK best practice analysis pipeline. Variant calling datasets were annotated by ANNOVAR. Screening of gene mutations from the bulk of raw variants sites followed the below analysis criteria: (1) > 25× coverage in the variant site; (2) variant allele frequency ≥ 5% and at least 5 individual mutant reads; (3) filter variants only observed on positive-strand or negative-strand; (4) filter sites presented in 1000 Genome Project with ≥ 1% frequency, NHLBI-ESP project with 6500 exomes with ≥ 1% frequency and the Genome Aggregation Database (gnomAD) with ≥ 1% frequency; (5) filter sequence variation frequency ≥ 5% and variants site < 20× coverage in normal control samples. Vcf files were converted to mutation annotation format (MAF) by vcf2maf and vep. The R packages maftools (https://bioconductor.org/packages/release/bioc/vignettes/maftools/inst/doc/maftools.html) were used to summarize, analyze, annotate, and visualize somatic MAF files. The tool deconstructSigs (https://rdrr.io/cran/deconstructSigs/) was used to identify signatures present in tumor samples.

Results

Cancer Genome Atlas neoantigen load analysis and correlation with overall survival

With the clinical data from TCGA, we were able to determine the cutoff value of predictive SNV, Indel, and total neoantigen counts. Patient samples and mutation numbers used to determine the cutoff values were: 96 AD and 83 SQ patients; 525 and 72 SNV neoantigen counts, 5 and 10 Indel neoantigen counts, 174 and 284 all neoantigens counts from AD and SQ, respectively. In this TCGA dataset, we showed that in AD, when the same survival probability was applied, high SNV/Indel/total neoantigen correlated with better overall survival. But in SQ, the results showed the opposite trend, high SNV/Indel/total neoantigen correlated with worse overall survival (Fig. 1A–F).

Patient characteristics

The clinical characteristics of the patients are shown in Table 1. There were 17 AD patients and 8 SQ patients with a median age of 64 years (ranging between 34 and 83 years old). A total of 16 patients were former or current smokers: 53% (9/17) of AD patients versus 87.5% (7/8) of SQ patients (Tab. 1). The median cigarette consumption was 23 and 35 packs/year in AD and SQ patients, respectively. *EGFR* mutation occurred in 35% (6/17) of AD patients, but only in 12.5% (1/8) of SQ patients. For the treatment regime, 12 patients (12/25, 48%) received standard chemotherapy for at least one cycle.

Tumor mutation burden concordance between the 579 gene panel and whole exome sequencing

Evidence has suggested that the TMB and tumor-specific neoantigens are potential determinants of the response to ICIs and can influence patient outcomes in immunotherapy [8–12]. Whole exome sequencing allows a direct measurement of TMB. However, routine implementation of WES in clinical practice is unsuitable because of high costs, labor and time intensiveness, and extensive data management. To test the potential utility of our in-house 579 gene Med1CDx panel in clinically predictive TMB estimates, DNA from 16 samples (including 11 AD and 5 SQ) was profiled. The Med1CDx panel showed a good correlation with WES for TMB estimation, with R^2 correlation values of 0.86 for all mutations.

Analysis of personalized neoantigens derived from whole exome sequencing and Med1CDx panel

Whole exome sequencing and Med1CDx panel were applied to profile neoantigen spectra in 16 patients. Mutations occurring in the patients were assessed for predicted binding affinity to HLA alleles, and potential neoantigens were identified. In the WES data, 1733 neoantigens with binding affinities < 500 nM were identified. Of these, only 52 (3%) were observed through the Med1CDx panel, demonstrating that the large targeted gene panel failed to identify a broad spectrum of neoantigens as compared to WES. We found the most frequent shared missense mutation

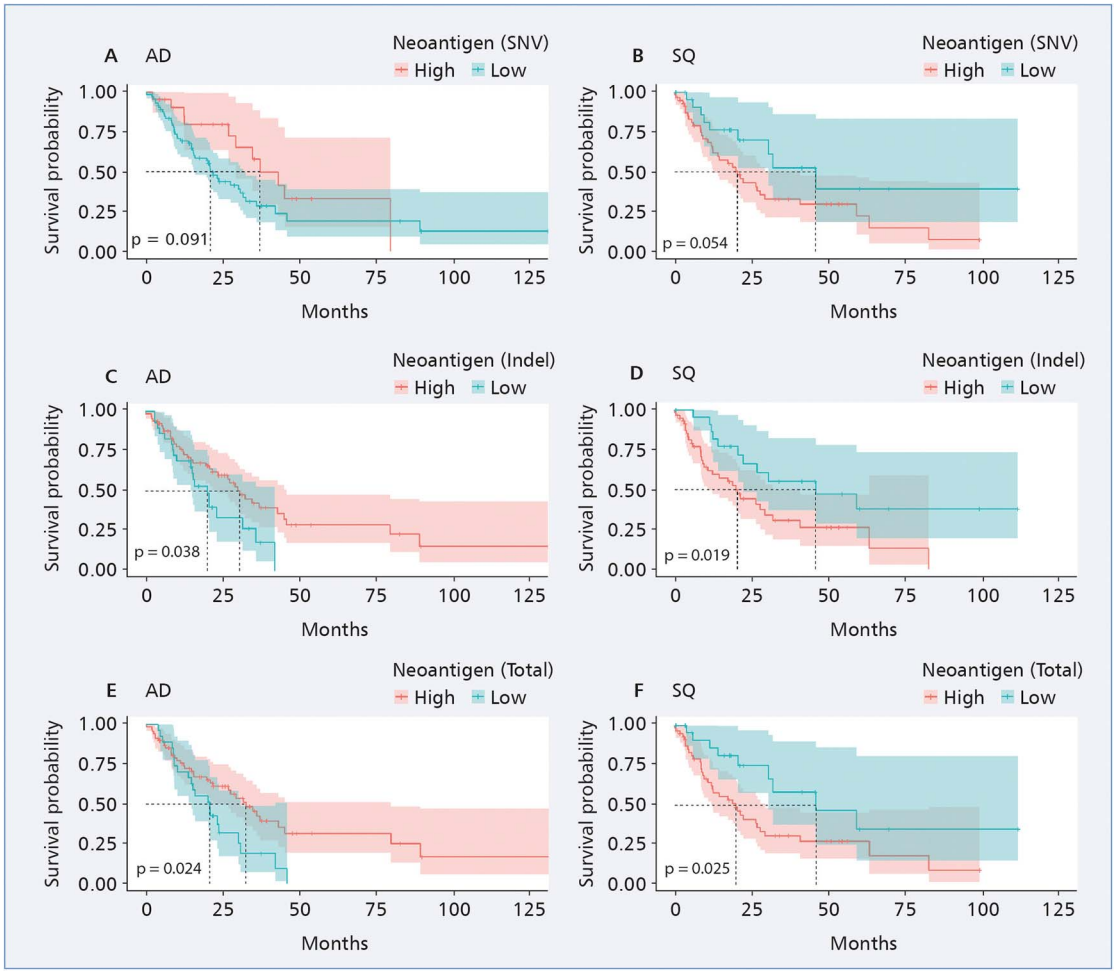


Figure 1. Correlation of neoantigen load with overall survival (OS) using Cancer Genome Atlas (TCGA) clinical data. Correlation of predicted single nucleotide variant (SNV) neoantigen (**A, B**), predicted Indel neoantigen (**C, D**), predicted total neoantigen (**E, F**) with OS

Table 1. Clinical characteristics of patients

Characteristics	Number (%)
Sex	
Male	18 (72)
Female	7 (28)
Age [years]	
Range	34–83
Medium	64
< 65	13 (52)
≥ 65	12 (48)
Smoking status	
Never	9 (36)
Former/current	16 (64)
Stage	
III	5 (20)
IV	20 (80)
Histologic diagnosis	
Adenocarcinoma	17 (68)
Squamous cell carcinoma	8 (32)

detected by WES were *TTN*, *MUC16*, and *TP53*, occurring in 11,11, and 9 patients, respectively. The neoantigens derived from WES were compared with those in the online public database which contains 1.1 million neoantigens. Twelve neoantigens from 9 patients were identical to those from the public database, even though the HLA types are different (Tab. 2). Among these neoantigens, neoantigens derived from *TP53* gene mutation were the most frequent (3 samples), and 2 neoantigens were derived from the *PCDHA4* and *FFAR2* genes (Tab. 2).

Neoantigen load analysis and correlation with progression-free survival and overall survival

The median number of NAL in all patients determined by WES was 3.0 neoantigens per Mb, while the median number called by the Med1CDx panel was 4.0. There was a linear relationship between neoantigens recovered from the Med1CDx panel and WES ($R^2 = 0.81$) (Fig. 2A, B). In the WES data, our data demonstrated that TMB and NAL were higher in the SQ subgroup than in the AD subgroup (Fig. 2C, D).

Table 2. Shared neoantigens between whole exome sequencing (WES) results and online public database of lung cancer

Neoantigen (frequency)	Gene in database	HLA type in samples	HLA type in database
TYSPALIKM (3)	*TP53*	HLA-C*07:02	HLA-A*23:01 HLA-A*24:02 HLA-C*04:01
RAFGRGLHV (2)	*FFAR2*	HLA-C*12:03	HLA-B*51:01 HLA-C*14:02
VRDGGSPSL (2)	*PCDHA4*	HLA-C*06:02	HLA-C*07:02 HLA-B*27:05
IIHTGEKPY (1) RTYTGEKPY (1)	*ZNF121* *ZNF559*	HLA-C*03:03	HLA-B*15:01
LTRPVHNAAR (1) ASHDERFKR (1)	*CDKN2A* *KDM2A*	HLA-A*31:01	HLA-A*33:03 HLA-A*11:01
AMLKNTVTI (1)	*MAGEC1*	HLA-A*02:01	HLA-A*02:01

HLA — human leukocyte antigen

Tumors with higher TMB carry higher NAL. Higher NAL was associated with improved overall survival using 10.3 neoantigens per Mb as a cutoff point (median OS not reached *versus* 11.0 months, log-rank p = 0.016) and progression-free survival (PFS) (median not reached *versus* 3.3 months, log-rank p = 0.03, 10.3 neoantigens per Mb as a cutoff point) (Fig. 2E, F).

Mutation spectra and signatures in lung adenocarcinoma and lung squamous cell carcinoma samples

In the SNV class, transitions of C>T, and C>A were significantly mutated in both AD and SQ samples. In addition to these two variations, the prevalence of T>C transitions was observed in AD and C>G in SQ. The top 20 mutated genes were quite different between the AD and SQ subgroups, but *TP53* and *TTN* were in the top 3 genes (Fig. 3A, C). In AD, genes with the highest mutation rate were *TP53* (9/17,53%), *MUC16* (7/17, 41%), *TTN*, and *EGFR* (6/17, 35%). *TP53* and *TTN* were the most frequent mutations in SQ (7/8,88%). Other mutations that presented frequently in SQ included *KMT2D* and *RYR2* (5/8, 62%), *MUC16*/*GOLGA6L2*/*SYNE1*/*NCAM1*/*OBSCN*/*TPTE* (4/8, 50%) (Fig. 3B, D). *TP53* and *TTN* were also among the top 3 mutations in tobacco smokers, with *MUC16* and *KMT2D* as other frequent mutations in AD and SQ smokers, respectively. The genes mutated in non-smokers in both AD and SQ subtypes were very diverse, and the mutation rates were not high.

Analysis of the mutational somatic substitutions using the COSMIC Mutational Signatures database (v2 — March 2015) demonstrated that the AD mutations were distributed in Signatures 1, 3, 4, while the SQ mutations were mainly Signature 3 and 4 related (Fig. 3E, F). Signature 1 is common in all cancer types and most cancer samples. Signature 3 was reported in breast, ovarian, and pancreatic cancers, our results confirmed that this signature set was also presented in lung cancer samples. Signature 4 has been found in lung adenocarcinoma and lung squamous carcinoma and is considered to be associated with smoking and tobacco mutagens.

Discussion

Neoantigen load, as a promising biomarker for predicting ICI efficacy, has been extensively studied, particularly in melanoma, lung cancer, and gynecological tumors [19, 23, 24]. In these types of cancer, NAL can predict anti-tumor effects of ICIs. However, research into neoantigens still faces challenges, such as the lack of an established standard protocol for neoantigen prediction or an optimized cutoff value for NAL. Previous research reported that large targeted panels are sufficient for most variant identification and NAL prediction [20, 21]. In this study, we compared a large targeted panel with WES to identify and profile NAL. Our data also showed that NAL measurement by the Med1CDx panel has a strong correlation with exome sequencing, which suggests that using a large gene panel for NAL prediction is feasible in NAL estimation. However, detailed neoantigen profiling demonstrated by the Med1CDx panel sequencing results failed to duplicate mutation estimated from WES data. Only 3% of the neoantigens identified by WES were observed through the Med1CDx panel, indicating that large targeted gene panels would not be appropriate for personalized neoantigen-based therapy development despite their convenience and advantages.

Currently, the association between NAL and genomic alterations is being studied to explore whether gene mutations can be utilized to estimate NAL for predicting the response to ICI therapies [11]. Common oncogene mutations can disrupt genome stability and alter immune status by creating novel antigens. Lyu et al. [25] found that that patients with mutant *TP53* exhibited enhanced tumor antigenicity and antigen presentation compared to those with wild-type *TP53*, and were more likely to benefit from ICI therapy. Besides common oncogenes, some rare gene mutations were also reported to cause an increase in NAL. Zhang et al. [26] reported that compared with patients with wild-type tumors, patients with *MUC16* mutant tumors have a significant increase in NAL, which is related to improved OS of patients with *MUC16* mutation containing NSCLC and melanoma. Research based on the TCGA indicated that *TP53*, *TTN*, and *MUC16* were the most frequently mutated genes in various cancers, including lung cancer [27]. Consistent with this, our study also found that these three genes have the highest mutation frequency in AD, while in SQ, *TP53* and *TTN* are the two genes

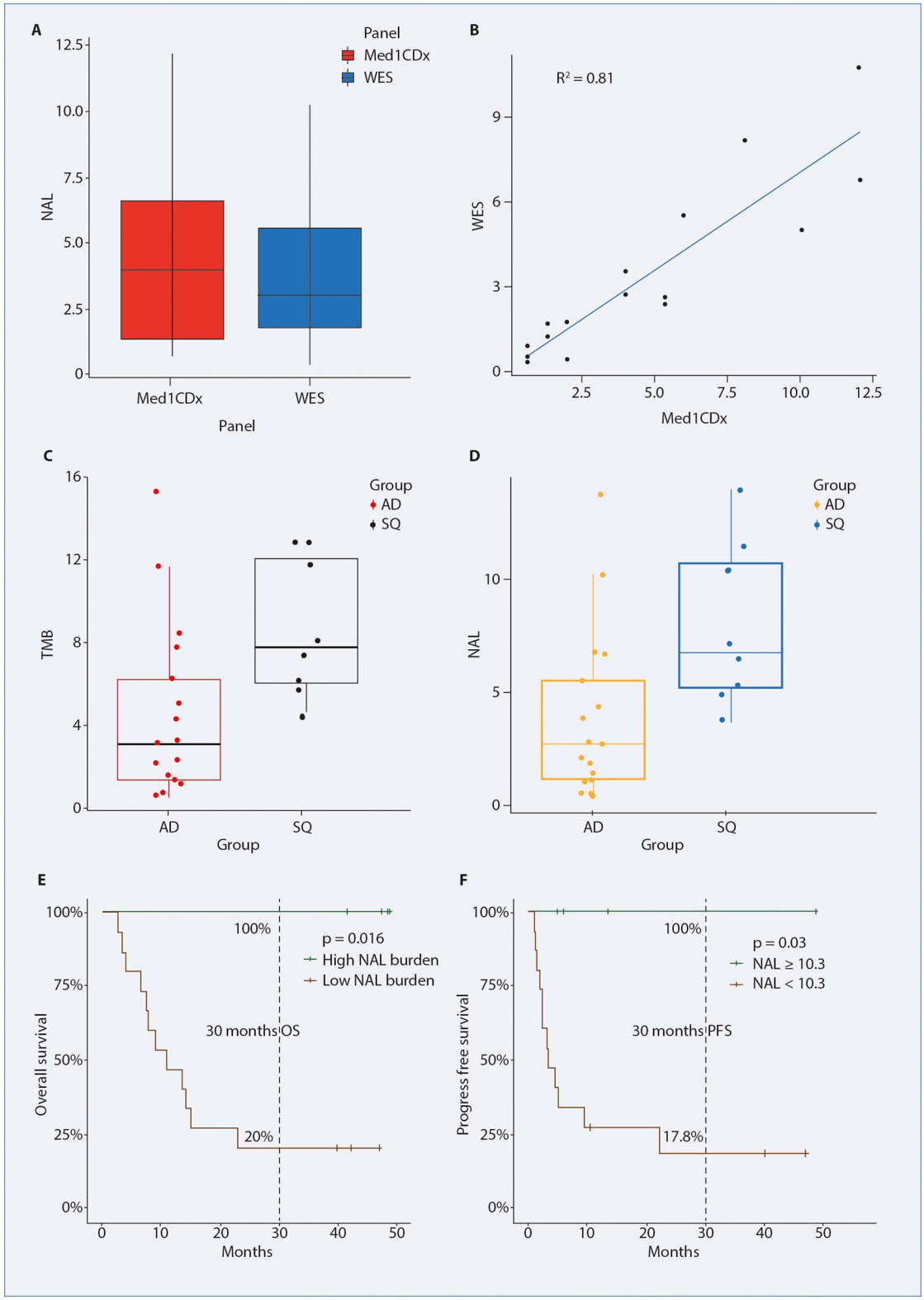
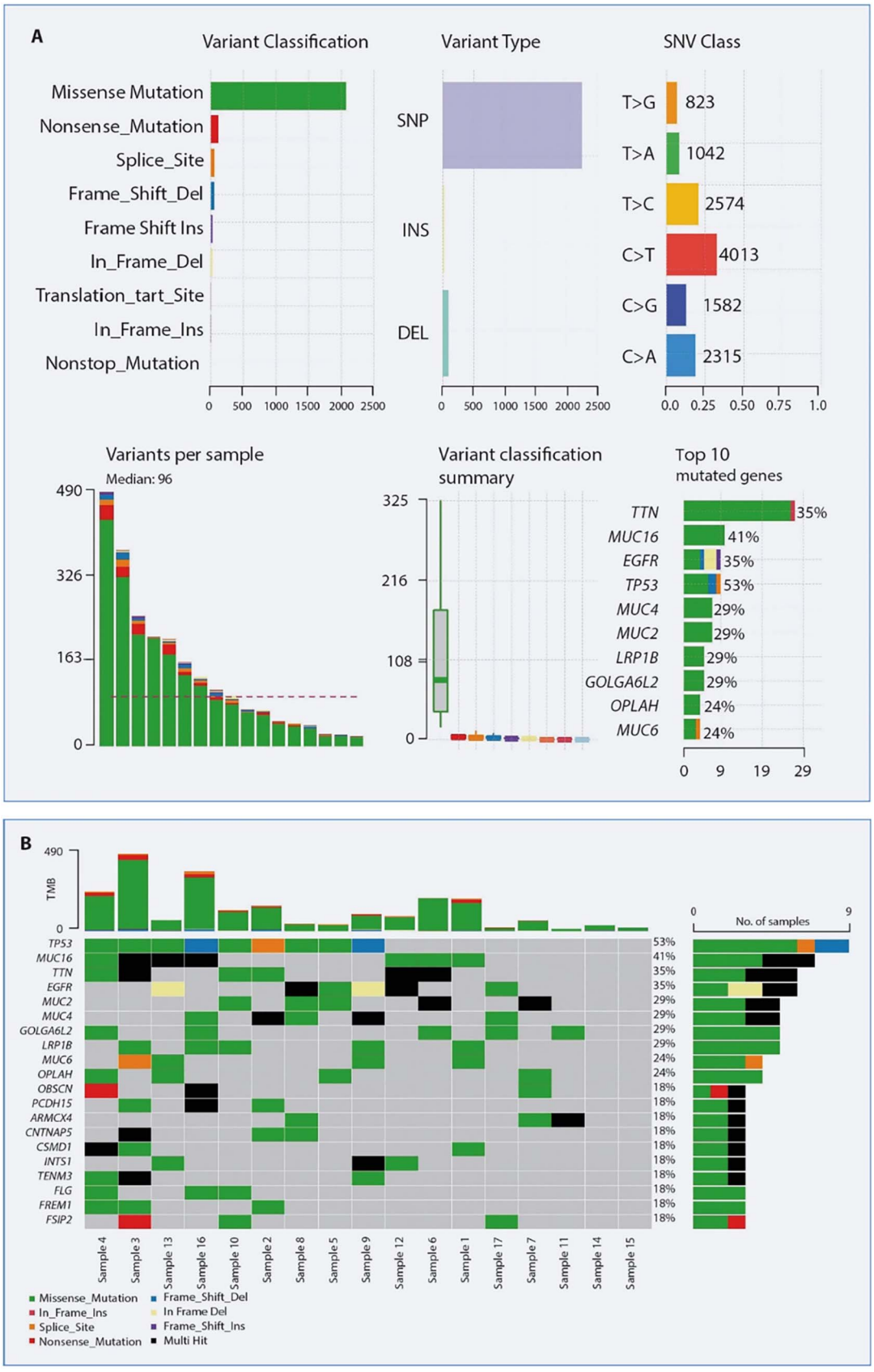
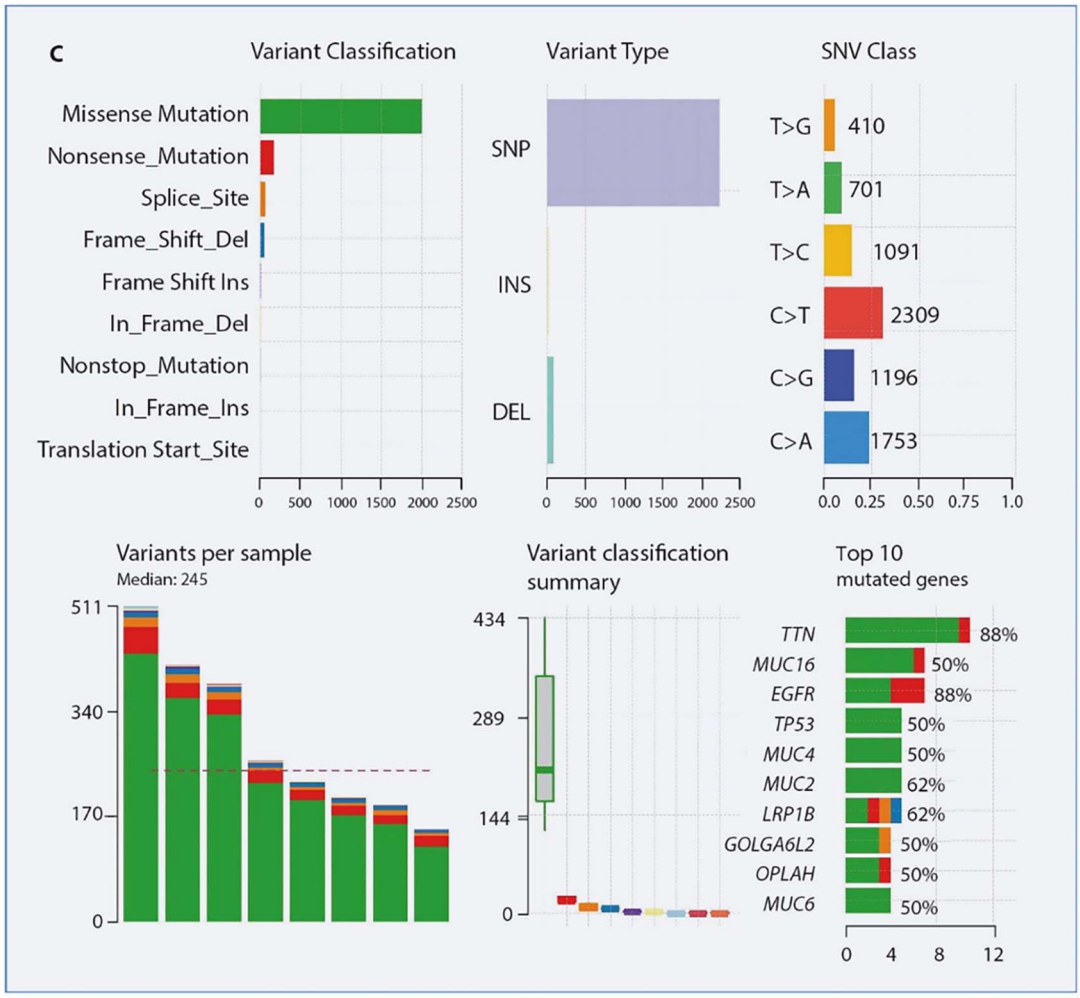


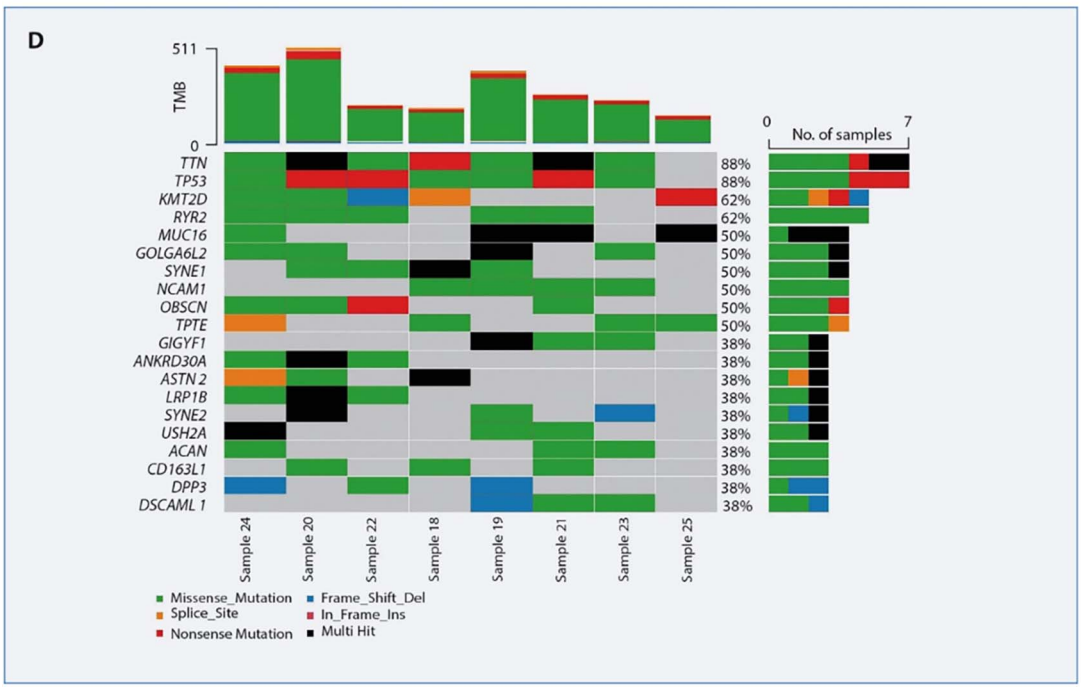
Figure 2. Tumor mutation burden (TMB) and tumor neoantigen load (NAL) analysis and association with clinical outcomes; **A.** NAL prediction by whole exome sequencing (WES) and Med1CDx panel; **B.** Correlation of NAL prediction by WES and Med1CDx panel; TMB value (**C**) and NAL (**D**) comparison in lung adenocarcinoma (AD) and lung squamous cell carcinoma (SQ) patients; Association of NAL with overall survival (OS) (**E**) and progression-free survival (PFS) (**F**)



→

Figure 3. Mutation spectra and signatures of lung adenocarcinoma (AD) and lung squamous cell carcinoma (SQ) patients; **A, B.** Mutation spectra of AD patients; **C, D.** Mutation spectra of SQ patients; Mutation signatures of AD patients (**E**) and SQ patients (**F**)





→

Figure 3 cont. Mutation spectra and signatures of lung adenocarcinoma (AD) and lung squamous cell carcinoma (SQ) patients; **A, B.** Mutation spectra of AD patients; **C, D.** Mutation spectra of SQ patients; Mutation signatures of AD patients (**E**) and SQ patients (**F**)

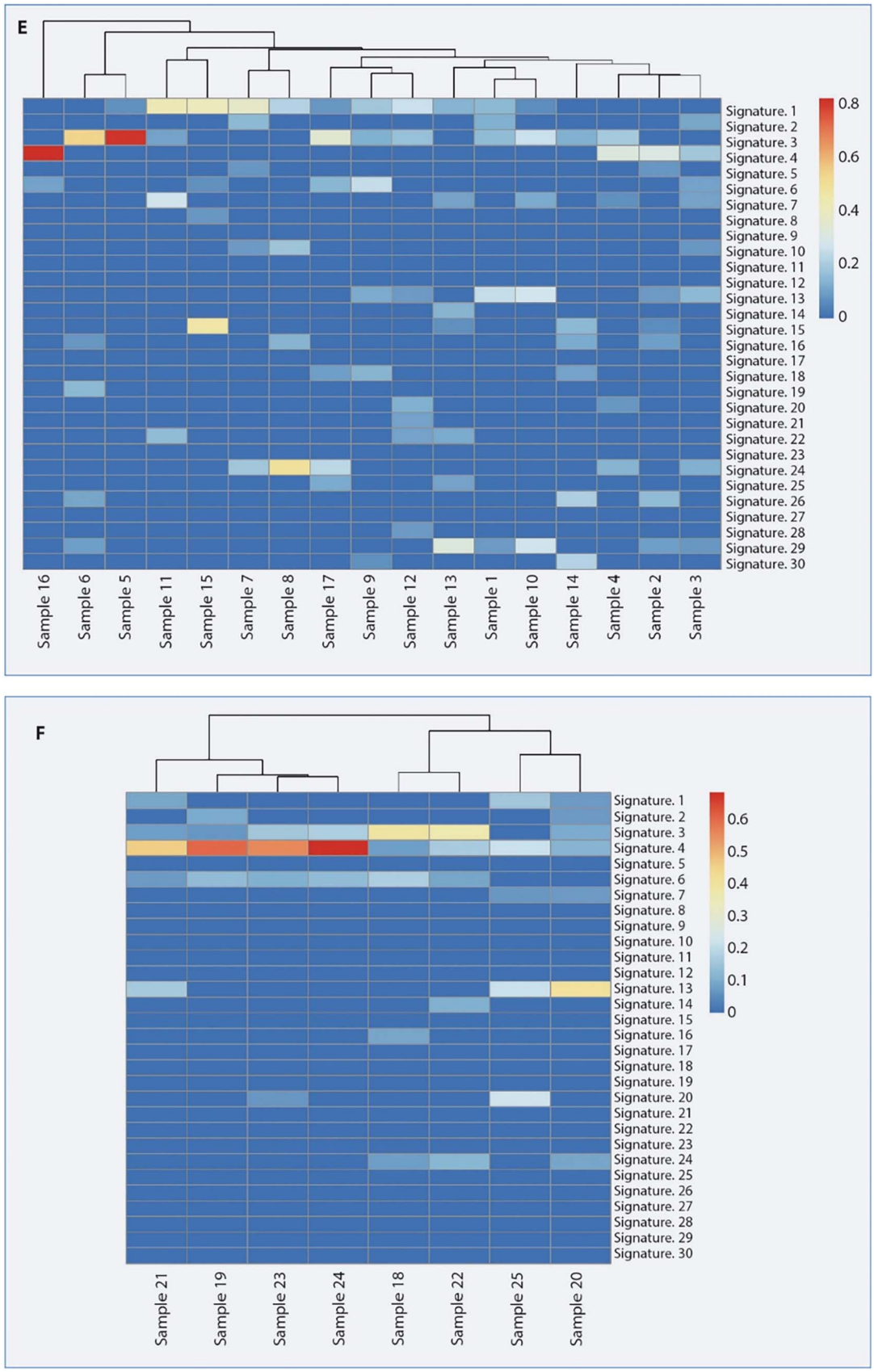


Figure 3 cont. Mutation spectra and signatures of lung adenocarcinoma (AD) and lung squamous cell carcinoma (SQ) patients; **A, B.** Mutation spectra of AD patients; **C, D.** Mutation spectra of SQ patients; Mutation signatures of AD patients (**E**) and SQ patients (**F**)

with the highest frequency. It is also possible that due to the high mutation frequency of *TP53* and *MUC16*, NAL is associated with the prognosis for AD and SQ in this study.

Recent studies have shown that neoantigens are not only associated with the response to anti-programmed cell death protein 1 (anti-PD-1) therapy in NSCLC patients but also are useful biomarkers for lung cancer prognosis and prediction of responses to chemotherapy in Chinese patients. A previous study analyzed NSCLC samples collected from patients treated with pembrolizumab and reported that higher NAL in tumors was associated with improved objective response and PFS [28]. Chae et al. [29] analyzed mutations in DNA repair genes using TCGA samples and found that NAL correlated with the expression of PD-1 and programmed death-ligand 1 (PD-L1) and tended to increase OS of patients with lung adenocarcinoma. High NAL is linked to DNA repair mutations and an increased number of tumor-infiltrating lymphocytes [24]. Although prognosis indication for neoantigen in lung cancer is promising, more detailed studies are still needed because contradictory results were obtained on the correlation of neoantigens with clinical outcomes of lung cancer subtypes. McGranahan et al. [19] reported that high NAL (defined as the upper quartile of NAL) was associated with significantly longer overall survival in lung AD but not SQ. Gong et al. [18] showed that higher NAL (> 2 neoantigens/Mb) exhibited better DFS for SQ but not AD patients. A benefit from adjuvant chemotherapy was correlated with lower NAL ($\leq$ 2 neoantigens/Mb). In our study, we demonstrated that high NAL is correlated with better overall survival in the AD subgroup in the clinical data retrieved from TCGA. In the 25-patient cohort, NAL could predict lung cancer prognosis in both AD and SQ patient subgroups both of which underwent a treatment strategy combining chemotherapy and immunotherapy.

Our study has a few limitations. Firstly, the study involved a relatively small sample size of participants. Given the potential implications of NAL for clinical decision-making, it is necessary to conduct further validation within a larger, independent patient cohort. Second, the method for calculating neoantigen load in this study differs from that of TCGA, although the impact on the observed trends may not be significant. In TCGA, the identification of potential neoantigen peptides was conducted using NetMHCpan v3.0 (Nielsen and Andreatta, 2016), and the neoantigen load refers to the number of pMHCs (peptides predicted to bind with MHC proteins). In our study, neoantigen prediction was carried out using the pvactools (version 1.5.8, default parameters). The NAL was calculated by dividing the total number of neoantigens by the CDS length in Mb of the WES or Med1CDx panel. Moreover, these two neoantigen algorithms were all based on the HLA-I binding prediction and did not account for other aspects of neoantigen production, including the processing and presentation of antigens, integration into the genome, and immune recognition. Last, all patients in this study were in stages III/IV, and the treatment regimens were primarily chemotherapy combined with immunotherapy. Therefore, the findings about the correlation between neoantigen and OS of AD patients are generally applicable to patients and treatments within this scope.

Conclusions

In conclusion, our study demonstrated that NAL could be used as a useful prognosis marker to provide stratification for lung adenocarcinoma (LUAD) patient outcomes. Further studies with a larger cohort from multiple institutions are needed to validate the current data and confirm the prognostic role of NAL in different subtypes of lung cancer. The results of genomic alternation in this study also show that *TP53* and *MUC16* are among the genes with the highest mutation frequency in NSCLC and are involved in the correlation of NAL value and prognosis. In addition, we have shown that the Med1CDx panel (targeting 579 genes) can accurately assess TMB and NAL compared with WES, providing more evidence on the feasibility of using a large targeted gene panel in TMB and NAL analysis.

Article Information and Declarations

Data availability statement

All data are available.

Ethics statement

The study was approved by the institutional committee board of Shanghai Chest Hospital and informed consent was taken from all the participants.

Author contributions

X.L.: study design, results interpretation and writing of the manuscript; Y.Y.: study design, results interpretation and writing of the manuscript; W.S.: bioinformatics and statistical data analysis; X.D.: Bioinformatics and statistical data analysis; K.G.: patient clinical data analysis; M.Y.: neoantigens data analysis; Z.J.: revision of the manuscript; Y.W.: results interpretation and writing of the manuscript; Y.Z.: revision of the manuscript.

All authors have reviewed the manuscript and approved the final version.

Funding

This work was funded by the National Key R&D Program of China (2016YFC1303300 to S.L.), the National Natural Science Foundation of China (82030045 to S.L.; 82172633 to YF.Y; 82002423 to WX.J), the Shanghai Municipal Science & Technology Commission Research Project (19411950500 to S.L.), the Shanghai Shenkang Action Plan (16CR3005A to

S.L.), the Shanghai Chest Hospital Project of Collaborative Innovation (YJXT20190105 to S.L.) the Yang Fan plan of Shanghai Science and Technology Commission Research Project (19YF1407300 to WX. J), and the Beijing Xisike Clinical Oncology Research Foundation (Y-HR2020MS-0982 to YF.Y.).

Acknowledgements

We acknowledged MEDx (Suzhou) colleagues in technical advices.

Conflict of interest

Authors declare no conflict interest.

Supplementary material

There is no supplementary material.

References

1. 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(3): 209–249, doi: 10.3322/caac.21660, indexed in Pubmed: 33538338.
2. Chan BA, Hughes BGM. Targeted therapy for non-small cell lung cancer: current standards and the promise of the future. Transl Lung Cancer Res. 2015; 4(1): 36–54, doi: 10.3978/j.issn.2218-6751.2014.05.01, indexed in Pubmed: 25806345.
3. Majeed U, Manochakian R, Zhao Y, et al. Targeted therapy in advanced non-small cell lung cancer: current advances and future trends. J Hematol Oncol. 2021; 14(1): 108, doi: 10.1186/s13045-021-01121-2, indexed in Pubmed: 34238332.
4. Blass E, Ott PA. Advances in the development of personalized neoantigen-based therapeutic cancer vaccines. Nat Rev Clin Oncol. 2021; 18(4): 215–229, doi: 10.1038/s41571-020-00460-2, indexed in Pubmed: 33473220.
5. Jiang T, Shi T, Zhang H, et al. Tumor neoantigens: from basic research to clinical applications. J Hematol Oncol. 2019; 12(1): 93, doi: 10.1186/s13045-019-0787-5, indexed in Pubmed: 31492199.
6. Šutić M, Vukić A, Baranašić J, et al. Diagnostic, Predictive, and Prognostic Biomarkers in Non-Small Cell Lung Cancer (NSCLC) Management. J Pers Med. 2021; 11(11), doi: 10.3390/jpm11111102, indexed in Pubmed: 34834454.
7. Thakur MK, Gadgeel SM. Predictive and Prognostic Biomarkers in Non-Small Cell Lung Cancer. Semin Respir Crit Care Med. 2016; 37(5): 760–770, doi: 10.1055/s-0036-1592337, indexed in Pubmed: 27732997.
8. Gubin MM, Artyomov MN, Mardis ER, et al. Tumor neoantigens: building a framework for personalized cancer immunotherapy. J Clin Invest. 2015; 125(9): 3413–3421, doi: 10.1172/JCI80008, indexed in Pubmed: 26258412.
9. Fancello L, Gandini S, Pelicci PG, et al. Tumor mutational burden quantification from targeted gene panels: major advancements and challenges. J Immunother Cancer. 2019; 7(1): 183, doi: 10.1186/s40425-019-0647-4, indexed in Pubmed: 31307554.
10. Büttner R, Longshore JW, López-Ríos F, et al. Implementing TMB measurement in clinical practice: considerations on assay requirements. ESMO Open. 2019; 4(1): e000442, doi: 10.1136/esmoopen-2018-000442, indexed in Pubmed: 30792906.
11. Zou XL, Li XB, Ke H, et al. Prognostic Value of Neoantigen Load in Immune Checkpoint Inhibitor Therapy for Cancer. Front Immunol. 2021; 12: 689076, doi: 10.3389/fimmu.2021.689076, indexed in Pubmed: 34992591.
12. Wang P, Chen Y, Wang C. Beyond Tumor Mutation Burden: Tumor Neoantigen Burden as a Biomarker for Immunotherapy and Other Types of Therapy. Front Oncol. 2021; 11: 672677, doi: 10.3389/fonc.2021.672677, indexed in Pubmed: 33996601.
13. Devarakonda S, Rotolo F, Tsao MS, et al. Tumor Mutation Burden as a Biomarker in Resected Non-Small-Cell Lung Cancer. J Clin Oncol. 2018; 36(30): 2995–3006, doi: 10.1200/JCO.2018.78.1963, indexed in Pubmed: 30106638.
14. Talvitie EM, Vilhonen H, Kurki S, et al. High tumor mutation burden predicts favorable outcome among patients with aggressive histological subtypes of lung adenocarcinoma: A population-based single-institution study. Neoplasia. 2020; 22(9): 333–342, doi: 10.1016/j.neo.2020.05.004, indexed in Pubmed: 32585428.
15. Jiang T, Shi J, Dong Z, et al. Genomic landscape and its correlations with tumor mutational burden, PD-L1 expression, and immune cells infiltration in Chinese lung squamous cell carcinoma. J Hematol Oncol. 2019; 12(1): 75, doi: 10.1186/s13045-019-0762-1, indexed in Pubmed: 31299995.
16. Yu H, Chen Z, Ballman KV, et al. Correlation of PD-L1 Expression with Tumor Mutation Burden and Gene Signatures for Prognosis in Early-Stage Squamous Cell Lung Carcinoma. J Thorac Oncol. 2019; 14(1): 25–36, doi: 10.1016/j.jtho.2018.09.006, indexed in Pubmed: 30253973.
17. Schumacher TN, Schreiber RD. Neoantigens in cancer immunotherapy. Science. 2015; 348(6230): 69–74, doi: 10.1126/science.aaa4971, indexed in Pubmed: 25838375.
18. Gong L, He R, Xu Y, et al. Neoantigen load as a prognostic and predictive marker for stage II/III non-small cell lung cancer in Chinese patients. Thorac Cancer. 2021; 12(15): 2170–2181, doi: 10.1111/1759-7714.14046, indexed in Pubmed: 34128337.
19. McGranahan N, Furness AJS, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. Science. 2016; 351(6280): 1463–1469, doi: 10.1126/science.aaf1490, indexed in Pubmed: 26940869.
20. Garofalo A, Sholl L, Reardon B, et al. The impact of tumor profiling approaches and genomic data strategies for cancer precision medicine. Genome Med. 2016; 8(1): 79, doi: 10.1186/s13073-016-0333-9, indexed in Pubmed: 27460824.
21. Thorsson V, Gibbs DL, Brown SD, et al. Cancer Genome Atlas Research Network, Cancer Genome Atlas Research Network. The Immune Landscape of Cancer. Immunity. 2018; 48(4): 812–830.e14, doi: 10.1016/j.immuni.2018.03.023, indexed in Pubmed: 29628290.
22. Chalmers ZR, Connelly CF, Fabrizio D, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med. 2017; 9(1): 34, doi: 10.1186/s13073-017-0424-2, indexed in Pubmed: 28420421.
23. Van Allen EM, Miao D, Schilling B, et al. Genomic correlates of response to CTLA-4 blockade in metastatic melanoma. Science. 2015; 350(6257): 207–211, doi: 10.1126/science.aad0095, indexed in Pubmed: 26359337.
24. Zhu Y, Meng X, Ruan X, et al. Characterization of Neoantigen Load Subgroups in Gynecologic and Breast Cancers. Front Bioeng Biotechnol. 2020; 8: 702, doi: 10.3389/fbioe.2020.00702, indexed in Pubmed: 32754579.
25. Lyu Q, Lin A, Cao M, et al. Alterations in TP53 Are a Potential Biomarker of Bladder Cancer Patients Who Benefit From Immune Checkpoint Inhibition. Cancer Control. 2020; 27(1): 1073274820976665, doi: 10.1177/1073274820976665, indexed in Pubmed: 33356494.
26. Zhang L, Han X, Shi Y. Association of MUC16 Mutation With Response to Immune Checkpoint Inhibitors in Solid Tumors. JAMA Netw Open. 2020; 3(8): e2013201, doi: 10.1001/jamanetworkopen.2020.13201, indexed in Pubmed: 32845327.
27. Kim N, Hong Y, Kwon D, et al. Somatic mutaome profile in human cancer tissues. Genomics Inform. 2013; 11(4): 239–244, doi: 10.5808/GI.2013.11.4.239, indexed in Pubmed: 24465236.
28. Rizvi NA, Hellmann MD, Snyder A, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science. 2015; 348(6230): 124–128, doi: 10.1126/science.aaa1348, indexed in Pubmed: 25765070.
29. Chae YK, Anker JF, Bais P, et al. Mutations in DNA repair genes are associated with increased neo-antigen load and activated T cell infiltration in lung adenocarcinoma. Oncotarget. 2018; 9(8): 7949–7960, doi: 10.18632/oncotarget.23742, indexed in Pubmed: 29487705.