

Genomic Variations and Immune-Related Features of TMB, PD-L1 Expression and CD8⁺ T Cell Infiltration in Chinese Pulmonary Sarcomatoid Carcinoma

Chenyue Zhang^{1,*}, Zhenxiang Li^{2,*}, Yanxiang Zhang³, Chenglong Zhao⁴, Hui Wang⁵, Jiamao Lin⁶, Cuicui Liu⁷, Xiaohui Wang⁸, Haiyong Wang⁹

¹Department of Integrated Therapy, Fudan University Shanghai Cancer Center, Shanghai Medical College, Shanghai, 200032, People's Republic of China; ²Department of Radiation Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, People's Republic of China; ³Berry Oncology Corporation, Beijing, 102206, People's Republic of China; ⁴Department of Pathology, The First Affiliated Hospital of Shandong First Medical University and Shandong Provincial Qianfoshan Hospital, Jinan, 250013, People's Republic of China; ⁵Department of Thoracic Surgery, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, 250021, People's Republic of China; ⁶Department of Traditional Chinese Medicine, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, 250117, People's Republic of China; ⁷Department of Oncology, Linyi People's Hospital, Linyi, 276000, People's Republic of China; ⁸Research Service Office, Shandong Liaocheng People's Hospital, Liaocheng, People's Republic of China; ⁹Department of Internal Medicine-Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, 250117, People's Republic of China

*These authors contributed equally to this work

Correspondence: Haiyong Wang, Department of Internal Medicine-Oncology, Shandong Cancer Hospital and Institute, Shandong First Medical University and Shandong Academy of Medical Sciences, Jinan, 250117, People's Republic of China, Tel +86 0531 67626332, Fax +86 0531 67626332, Email wanghaiyong6688@126.com

Background: Pulmonary sarcomatoid carcinoma (PSC) is a rare and distinct subtype of lung cancer characterized by its aggressiveness and dismal prognosis. However, genomic landscape and immune contexture have not been fully elucidated among PSC patients.

Methods: In the present study, whole-exome-sequencing (WES) analyses were performed to depict genomic landscape of 38 independent PSC samples. Tumor mutation burden (TMB) was calculated with the total number of non-synonymous SNVs and indel variants per megabase of coding regions. PD-L1 expression and CD8⁺ T cell density were evaluated by immunohistochemistry in PSC samples. Their associations with genomic mutation were further assessed in genes with most frequent mutation. Overall survival (OS) of PSC patients with top mutated genes and high and low TMB, PD-L1 and CD8⁺ TIL expressions were further compared. Subgroup analyses of OS stratified by morphology and pathological type were conducted. Their correlation with TMB, PD-L1 and CD8⁺ T cell were further assessed.

Results: We identified a cohort of genomic and somatic mutation in PSC patients. Subgroup patients with distinct clinicopathological features were found to harbor different genomic mutations and immunologic features. Besides, genomic profiles influenced outcomes, with SARS mutation associated with worsened prognosis.

Conclusion: Through the mapping of genetic and immunologic landscape, we find the heterogeneity among the subgroups of PSC. Our findings may provide opportunities for therapeutic susceptibility among Chinese PSC patients.

Keywords: pulmonary sarcomatoid carcinoma, mutational landscape, tumor mutational burden, programmed death ligand-1, CD8⁺ T cell

Introduction

Non-small-cell lung cancer (NSCLC) represents a majority of all lung cancers, and of these, pulmonary sarcomatoid carcinoma (PSC) only accounts for a rather relatively small proportion.¹ PSC, a poorly differentiated tumor containing

both cancerous and sarcoma or sarcoma-like components, is generally resistant to conventional chemotherapy.²⁻⁴ Recently, mutations found in *TP53*, *MET*, *EGFR* and *KRAS* as well as high PD-L1 expressions in PSC have rendered targeted therapy and immunotherapies preferable choices for PSC patients.⁵⁻⁷ However, genomic analyses of Chinese PSC patients are relatively few. Therefore, a comprehensive analysis of the genomic profile of Chinese PSC patients is of vital importance, which may be helpful in providing instructive treatment strategies.

Recent studies have found that tumor mutation burden (TMB), PD-L1 expression, tumor-infiltrating lymphocyte (TIL) infiltration are associated with clinical outcomes in response to immunotherapies.⁸⁻¹¹ It has been commonly recognized that tumors with high expressions of TMB, PD-L1 and TIL are associated with good response to immunotherapies.¹²⁻¹⁴ Tumors have generally been categorized into four different types according to CD8⁺ TIL and PD-L1 expression. Patients with high PD-L1 and CD8⁺ TIL expression have been associated with benefit from immunotherapies.¹⁵⁻¹⁷ In addition, previous studies have found that genomic status could influence the immunotherapy efficacy by harnessing the potent cytotoxic T lymphocyte. However, the association between genomic mutation and immune-associated markers has not been fully defined in PSC.

To address the limited knowledge, we performed this study in patients with surgically resected PSC to evaluate the landscape of genomic variations and the profile of immune-related factors. We analyzed association between the genomic alterations and TMB, PD-L1 and CD8⁺ TIL, as well as the clinical relevance of genomic variations, TMB, PD-L1 and CD8⁺ TIL, overall survival including the distinct subgroup of PSC, and other clinicopathologic features. These results might provide important insights into the genomic features and biology and a roadmap to inform us of genome and immuno-guided personalized treatment for Chinese PSC patients.

Methods

Sample Collection

We retrospectively identified patients who were pathologically confirmed with PSC from January 2014 to April 2019. These patients were recruited from Shandong Cancer Hospital and Institute, Shandong Provincial Hospital, Liaocheng People's Hospital, Linyi People's Hospital. Their pathological types were classified into two types: either pure sarcomatoid component or mix type. All specimens of eligible case must have a confirmed diagnosis of PSC and had at least 80% sarcomatoid cellularity. Both tumor and non-tumor tissues of each patient were collected. All diagnoses were further independently confirmed by two experienced pathologists. The histologic subtypes of pure PSC were confirmed by immunohistochemistry (IHC), in which adenocarcinoma or squamous carcinoma components were not found. Non-pure SPC refers to the subtype in which adenocarcinoma or squamous carcinoma components were found. As you have kindly suggested, we have defined pure SPC and non-pure SPC in the methods section, as indicated in our revised manuscript. A total of 43 PSC patients were collected, leaving a final cohort of 38 patients after exclusion of two polluted samples and three non-tumor tissues without sufficient DNA. This study was approved by the ethics committee of Shandong Cancer Hospital and Institute. All included patients in this study offered written informed consent. Overall survival (OS) was defined as the interval between diagnosis and death or between diagnosis and the last observation point. For surviving patients, the data were censored at the last follow-up.

Sample Processing and Genomic DNA Extraction

After surgical resection, tumor tissues were fixed with formalin, subsequently embedded in paraffin (FFPE). Corresponding non-tumor tissues adjacent to tumors were set as controls. Genomic DNA was extracted from each FFPE sample using the GeneRead DNA FFPE Kit (Qiagen, USA) and those extracted from non-tumor tissues were also served as controls.

DNA Library Construction and Whole Exome Sequencing (WES)

Genomic DNA was digested into fragments with the size of 200 bp using enzymatic method (5X WGS Fragmentation Mix, Qiagen, USA). T-adapters were added to both ends after end repairing and A tailing. For the library construction, the purified DNA was amplified by ligation-mediated PCR, and subsequently the final sequencing libraries were

generated using the 96 rxn xGen Exome Research Panel v1.0 (Integrated DNA Technologies, USA) according to the manufacturer's instructions. Paired-end multiplex samples were sequenced with the NovaSeq 6000 System. Sequencing depth was 200× per tissue sample and 100× per control.

Sequence Alignment and Variants Detecting

The raw data were preprocessed by FASTP to trim adaptor sequences,¹⁸ and subsequently, the clean reads in Fast Q format were aligned to the reference human genome (hg19/GRCh37) by Burrows-Wheeler Aligner (BWA, v0.7.15).¹⁹ SAM tools and Picard (2.12.1) (<http://picard.sourceforge.net/>) were used to sort the mapped BAM files and process PCR duplicates. To compute the sequencing coverage and depth, the final BAM files were generated by GATK (the Genome Analysis Toolkit 3.8) for local realignment and base quality recalibration.²⁰ The somatic single nucleotide variations (SNVs) were identified by MuTect and somatic small insertions and deletions (InDels) were detected by GATK Somatic Indel Detector. ANNOVAR software was performed for variant annotation based on multiple databases,²¹ including variant description (HGVS), population frequency databases (1000 Genomes Project, dbSNP, ExAC), variant functional prediction databases (PolyPhen-2 and SIFT), and phenotype or disease databases (OMIM, COSMIC, ClinVar). After annotation, the retained non-synonymous SNVs were screened with variant allele frequency (VAF) (cutoff $\geq 3\%$) or with VAF (cutoff $\geq 1\%$) for cancer hotspots from disease databases for the further analysis. Tumor mutation burden (TMB) was calculated with the total number of non-synonymous SNVs and indel variants per megabase of coding regions. Significant driver genes were identified by combining two methods of MutsigCV and dNdScv as described in the previous studies,^{22,23} with a false discovery rate (FDR, cutoff $< 10\%$).

Copy number variations (CNVs) were identified using the Genome Identification of Significant Targets in Cancer (GISTIC) 2.0 algorithm.²⁴ At chromosomal arm-level, significant amplification or deletion were screened with FDR (cutoff $< 10\%$) for further analyses. At focal-level, significant amplification was screened with FDR (cutoff $< 5\%$) and G-score (cutoff > 0.3), and significant deletion was screened with FDR (cutoff $< 5\%$) and G-score (cutoff < -0.2) for further analyses.

Mutational signatures were analyzed by R Foundation for Somatic Signatures.²⁵ We further compared mutational signatures with the Sanger signatures in the Catalogue of Somatic Mutations in Cancer (COSMIC) database.

Immunohistochemical Staining

Immunohistochemical staining was conducted using Enhance Labelled Polymer System (ELPS). Briefly, the PSC tissues were stained with anti-PD-L1 (CST, 13684, 1:100), anti-CD8+ (CST, 85336, 1:100) at 4°C overnight, then washed with PBS for three times, with 5 minutes per time. Corresponding secondary antibody was used to stain the tissues at 37°C for 30 minutes, after which they were then washed with PBS for three times with 5 minutes per time. Further staining with 3,3-diaminobenzidine (DAB), after which washed with the distilled water. Next, hematoxylin staining was conducted, followed by dehydration, clearing and mounting with neutral gums. Images of the stained tissues were captured by Digital Pathology Slide Scanner (KF-PRO-120, KF-BIO).

The Calculation of PD-L1

For the evaluation of PD-L1 expression, tumor proportion score (TPS) was defined as the number of PD-L1-staining tumor cells divided by the total number of viable tumor cells multiplied by 100; combined positive score (CPS) was defined as the number of PD-L1-staining cells divided by the total number of viable tumor cells multiplied by 100. PD-L1 staining of the tonsil had been adopted to ensure eligibility of the enrolled specimens. The qualified staining should be strong positivity for PD-L1 in intratonsillar cleft epithelia and negative staining for PD-L1 in lymphocytes (mantle zone and germinal center B cells) and superficial epithelial cells. The cutoff TPS/CPS value was 1.0% for positivity of PD-L1.

The Calculation of CD8⁺ T Cell

We observed CD8⁺ T cell distribution in the tumor stroma at a 200-fold magnification. If they were equally distributed, they were calculated in three randomly chosen areas (0.1 mm² per area). If unequally distributed, corresponding areas would be selected according to the percentage of CD8⁺ T cells in areas of various densities (0.1 mm² per area). CD8⁺ T

cell infiltration was calculated as $CD8^+$ T cell count/ $0.1 \text{ mm}^2 \times 10$, namely $CD8^+$ T cell count/ mm^2 . For each sample, TMB was defined as the total number of nonsynonymous SNVs/Indels of coding area of a tumor genome based on whole-exome sequencing (WES). TNB was determined as the number of all putative neoantigens per megabase of genome. No staining of $CD8^+$ T cell infiltration was defined as negative and otherwise positive.

Statistical Analysis

R Foundation for Statistics Computing, Package (R package, version 3.3.3) was used to perform the statistics analysis. Fisher's exact test and Wilcoxon test were used to analyze the relationship between TMB and clinical indexes. Kaplan-Meier method was used to estimate overall survival. Hazard ratios (HR) was analyzed by Cox regression analysis. $p < 0.05$ was defined as statistically significant.

Results

Patient Characteristics

Thirty-eight surgically resected PSC samples were collected and successfully sequenced. Baseline characteristics of the enrolled patients with PSC are listed in Table 1. The median age was 61 years (range 39–82 years). Thirty patients (78.9%) were male. The percentage of patients with AJCC stage I–II (63.2%) was higher than that with AJCC stage III (36.8%). The percentage of smokers (60.5%) was higher than that of the non-smokers (34.2%), and two patients were unknown in smoking status. We classified the 38 PSC patients into the following categories: spindle-like (84.2%) and non-spindle-like (15.8%) as defined by morphology. The proportion of pure and non-pure PSC patients was 57.9% and 42.1%, respectively. The threshold of PD-L1 and $CD8^+$ T cell was set to 1% and 5%, respectively, according to previous

Table 1 Baseline Characteristics of Included Patients (N = 38)

Variables	N=38
Age	
Median	61
Range	39–82
Gender	
Male	30 (78.9%)
Female	8 (21.1%)
AJCC stage	
I–II	24 (63.2)
III	14 (36.8)
Smoking history	
Smokers	23 (60.5%)
Non-smokers	13 (34.2%)
Not in detail	2 (5.3%)
Morphology	
Spindle-like	32 (84.2%)
Non spindle-like	6 (15.8%)
Pathological type	
Pure sarcomatoid	22 (57.9%)
Non pure sarcomatoid	16 (42.1%)
PD-L1 expression	
≥1%	21 (55.3%)
<1%	17 (44.7%)
CD8+ T cell	
≥5%	28 (73.7%)
<5%	10 (26.3%)

studies.^{26–28} The number of patients with PD-L1 higher than 1% and lower than 1% was 21 and 17. And those whose CD8⁺ T higher than 5% accounted for 73.7% whereas those with CD8⁺ T lower than 5% accounted for 26.3%.

Genomic Landscape of PSC Patients

The somatic mutations were detected by WES, the top ten of frequently mutated gene which were *TP53*, *TTN*, *ZFH4*, *USH2A*, *FAT3*, *PKHD1L1*, *MUC16*, *CSMD2*, *MDN1* and *PTPRD* (Figure 1A and Supplementary Figure 1). Two software packages, MutSigCV and dndscv, identified mutated genes with high frequency. *TP53* was the most frequently mutated gene in our PSC cohort. A total of 52.63% (20/38) of the patients harbored different forms of mutations in *TP53*. These mutations were demonstrated in the forms of missense, splice-site, stop-loss, stop-gain, in-frame indel, frame shift indel.

At chromosomal arm-level, amplification of chromosomal 1p, 19p, 19q, 20p, 20q, 22q and deletion of chromosomal 4p, 4q, 5q, 6q, 13q, 18p, 18q and 21q could be detected in our PSC cohort, as demonstrated in Figure 1B. Similarly, at focal-level, amplification of *FBLN2*, *TERT*, *CARD11*, *ZNF479*, *CUX1*, *LZTR1*, *DGCR8* and deletion in *PIK3R1*, *ARHGAP26*, *HNRNBA2B1*, *HOXA11*, *HOXA13* and *HOXA9* were revealed in our PSC cohort, as shown in Figure 1C.

Mutational Features of PSC Patients

We further analyzed the mutational features of each PSC patient in our cohort using the nonnegative matrix factorization (NMF) algorithm, which identified four independent mutational signatures (S1–S4) that matched four of the Sanger signatures in the Catalogue of Somatic Mutations in Cancer (COSMIC) database, including deficient mismatch repair (dMMR) (signature 6), dMMR (signature 15), smoking and an unknown feature (Figure 2A and B). The type of base-pair substitutions of each sample and mutational spectrum are demonstrated (Figure 2A). Each mutation status of the 38 PSC patients as defined by the four mutational signatures, sex, morphology, stage, survival, smoker, age, pathological subtype was shown (Figure 2C). There was a significant disparity between dMMR (signature 6) and dMMR (signature 15) ($p < 0.0001$), between dMMR (signature 6) and smoking ($p = 0.00051$), between dMMR (signature 15) and smoking ($p < 0.0001$) (Figure 2D).

Immunological Landscape of PSC Patients

Previous reports have shown that the high degree of TMB, PD-L1 expressions and CD8⁺ T infiltration are correlated with the immunogenic features of the tumor and reliably predict a good response to immunotherapy. High TMB, related to tumor neoepitope burden, indicating a T-cell-inflamed tumor microenvironment (TME), act as predictive biomarkers for immunotherapy. In the present study, we assessed TMB, PD-L1 and CD8⁺ T infiltration comprising the immunological

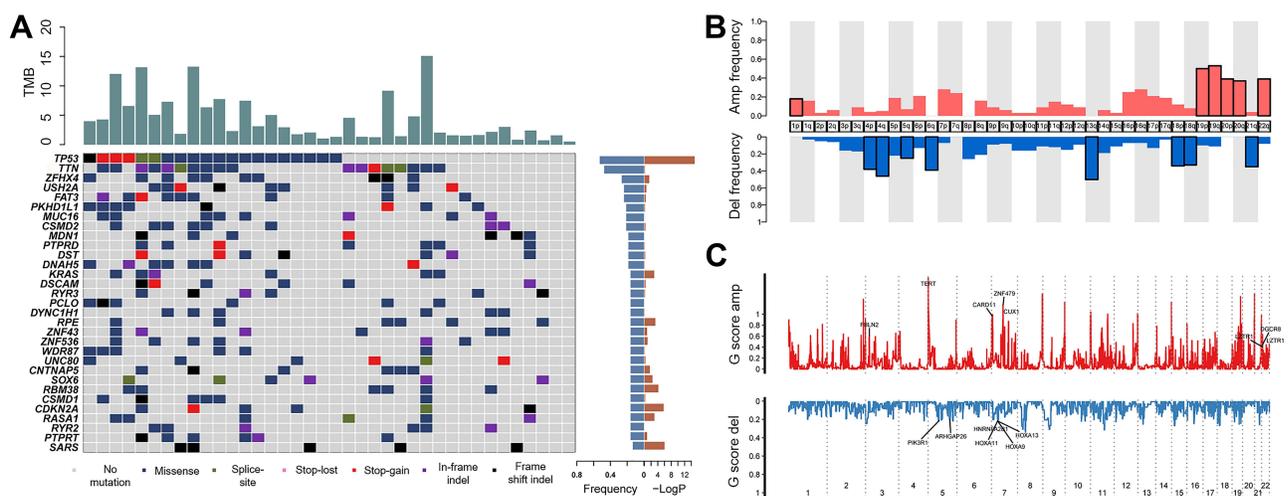


Figure 1 Genomic alterations in pulmonary sarcomatoid carcinoma. (A) Spectrum of the key molecular alterations in PSC. Tumor mutation burden are listed at the bottom according to the samples. Frequency of each mutation has been calculated. (B) Amplification and deletion frequency of copy number variation (CNV) on chromosome levels. (C) Zoom in the significant amplification and deletion region.

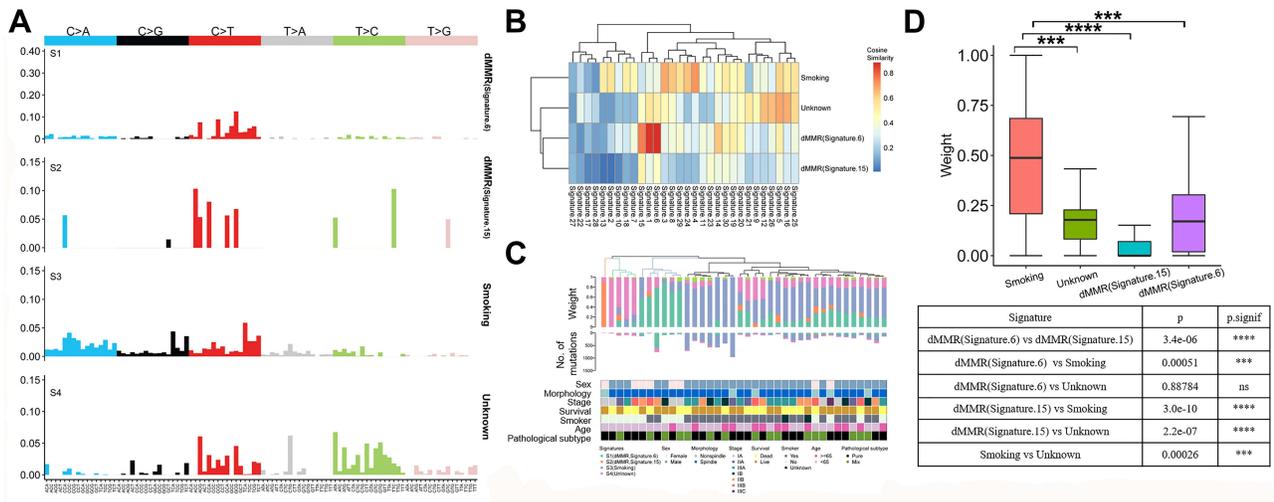


Figure 2 Somatic mutations and copy number alterations in PSC. (A) Signatures are displayed according to the 96-substitution classification, with x-axis showed mutation types and y-axis showed the estimated mutations of each mutation type, which are identified by a Bayesian nonnegative matrix factorization (NMF) algorithm. (B) Mutational signatures in our cohort (n=30). Samples are featured by the following: smoking, unknown, deficient mismatch repair (dMMR) (Signature 6), dMMR (Signature 15). (C) Mutational burden and weight stratified by dominant mutational signature and clinical variables. (D) Mutational weight among PSC patients with smoking, unknown feature, dMMR (Signature 6) and dMMR (Signature 15). ***: $p < 0.001$; ****: $p < 0.0001$.

milieu of TME among PSC patients. We analyzed the association between TMB, PD-L1 expression and CD8⁺ T cell infiltration. As demonstrated in Figure 3A, there was a positive correlation between CD8⁺ T cell and PD-L1 expression ($p < 0.01$). However, there was no significant association between TMB versus CD8⁺ T cell, TMB versus PD-L1. Among the high frequently mutated genes, we detected higher TMB in patients with mutations in *ZFH4*, *TTN*, *TP53*, *KRAS*, *FAT3*, *PKHD11*, *CDKN2A*, *CSMD2*, *DNAH5* and *DST* as compared with their corresponding wild type ($p < 0.01$ for

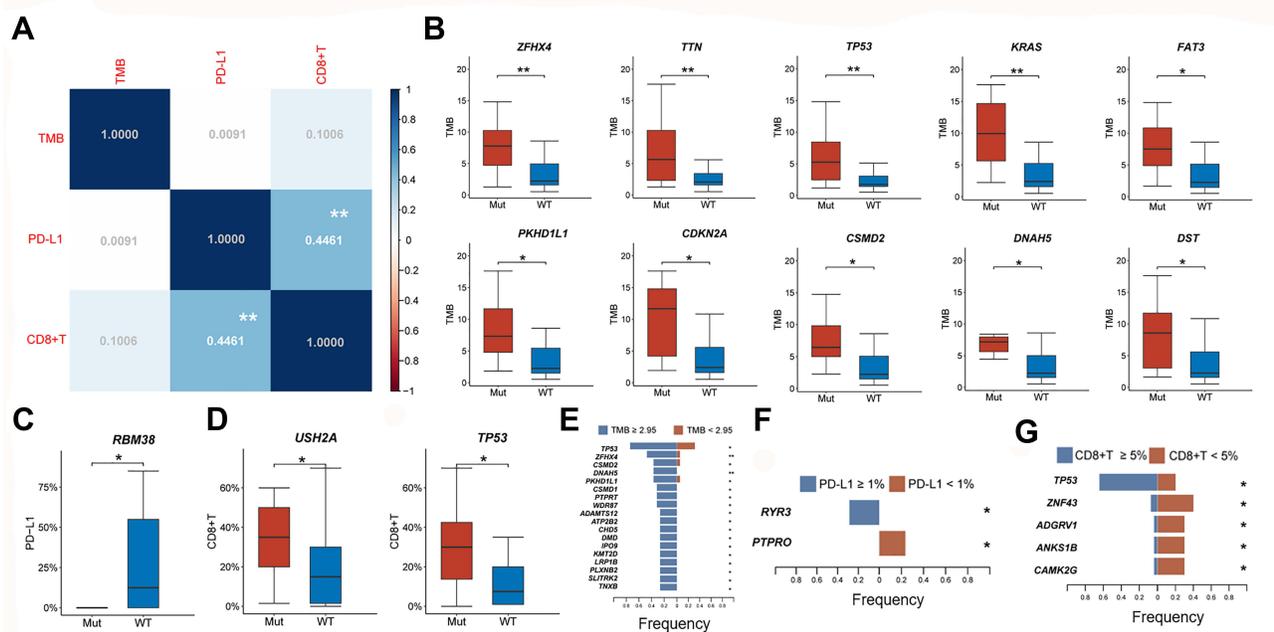


Figure 3 Differential expression of TMB, PD-L1 expression, CD8⁺ T infiltration based on gene mutation status. (A) The association among TMB, PD-L1 expression and CD8⁺ T infiltration. The numbers are shown as correlation coefficient between each of them. (B) TMB expression in PSC patients with wild type and high frequency mutations in *ZFH4*, *TTN*, *TP53*, *KRAS*, *FAT3*, *PKHD11*, *CDKN2A*, *CSMD2*, *DNAH5*, *DST*. (C) PD-L1 expression in PSC patients with wildtype and mutant *RBM38*. (D) CD8⁺ T cell infiltration in PSC patients with wildtype and mutant *USH2A*, *TP53*. (E) Plots showing the contribution of gene signatures in TMB high and TMB low. (F) Plots showing the contribution of gene signatures in PD-L1 high and PD-L1 low. (G) Plots showing the contribution of gene signatures in CD8⁺ T cell high and CD8⁺ T cell low. *: $p < 0.05$; **: $p < 0.01$.

ZFH4, *TTN*, *TP53*, *KRAS*; $p < 0.05$ for *FAT3*, *PKHD1L1*, *CDKN2A*, *CSMD2*, *DNAH5*, *DST*) (Figure 3B). As for PD-L1 expression, we observed lower expression of PD-L1 in PSC with mutation in *RBM38* than those with wild type ($p < 0.05$) (Figure 3C). High level of CD8⁺ T cell infiltration was found to be in mutations in *USH2A* and *TP53* genes than PSC patients with wild-type *USH2A* and *TP53* ($p < 0.05$) (Figure 3D). To the contrary, we shifted out genes whose mutational status may affect TMB, PD-L1 and CD8⁺ T cell expression. Results have shown that among the 38 PSC patients, those with higher TMB tend to associate with a series of genes mutations such as *TP53*, *ZFH4*, *CSMD2*, *DNAH5*, *PKHD1L1*, *CSMD1*, *PTPRT*, *WDR87*, *ADAMTS12*, *ATP2B2*, *CHD5*, *DMD*, *IPO9*, *KMT2D*, *LRP1B*, *PLXNB2*, *SLITRK2*, *TNXB* ($p < 0.05$ for all) (Figure 3E). Mutations vary between those with high PD-L1 and lower PD-L1 expression. For samples whose PD-L1 expressions higher than 1%, *RYR3* is shown to be reported whereas *PTPRO* tends to be mutated in samples with relatively lower PD-L1 expressions (<1%) ($p < 0.05$) (Figure 3F). For specimen with CD8⁺ T cell infiltration more than 5%, *TP53* mutation was detected. However, *ZNF43*, *ADGRV1*, *ANKS1B* and *CAMK2G* were proven to be mutated in samples with CD8⁺ T cell infiltration less than 5% ($p < 0.05$) (Figure 3G).

Association of Genomic Variations and Immunological Factors with Clinicopathological Features

Considering the possible impact of clinical variables such as sex, age, tumor stage, smoking status, morphology, pathological type and living status on genetic mutations, we further analyzed the association between gene mutation and clinical variables. As shown in Figure 4A, higher frequencies of *AASDH* and *PTPRK* mutations were found in female PSC patients compared with male patients ($p < 0.05$). *NLRP7*, *AKAP2*, *ALMS1*, *PALM2*, *SOGA3*, *SPTBN1* mutations were found to be more in PSC patients over 65-year-old compared with those in younger patients (Figure 4B). PSC patients with stage III were more likely to harbor *PCDH15*, *STON1*, *CHD8*, *COL12A1*, *DRC1*, *HLTF*, *PPFIA2*, *SEMA6D*, *TMEM260* mutations compared with those with stage I/II ($p < 0.05$) (Figure 4C). There were more smokers in PSC patients with mutations in *TTN*, *ZFH4*, *DNAH5*, *PKHD1L1* compared with nonsmokers. And *EGFR*, *AASDH*, *HLTF*, *PTPRK* and *ZFAND4* mutations were more likely to be found in nonsmoker PSC patients compared with their smoker counterpart (Figure 4D). Non-spindle cell PSC showed a significantly higher frequency of *VPS13B*, *ANAPC2*, *BMP7*, *C9orf3*, *CCDC70*, *CHD3*, *DHX32*, *DSG2*, *FRMPD4*, *LRRC26*, *LRRC8A*, *LZTS1*, *NDUFV3*, *OSBP2*, *PIAS3*, *PKP1*, *SEC24A*, *SFMBT2*, *ZHX3*, *ZNF148* mutations than spindle cell PSC (Figure 4E). Mix PSC demonstrated higher rate of

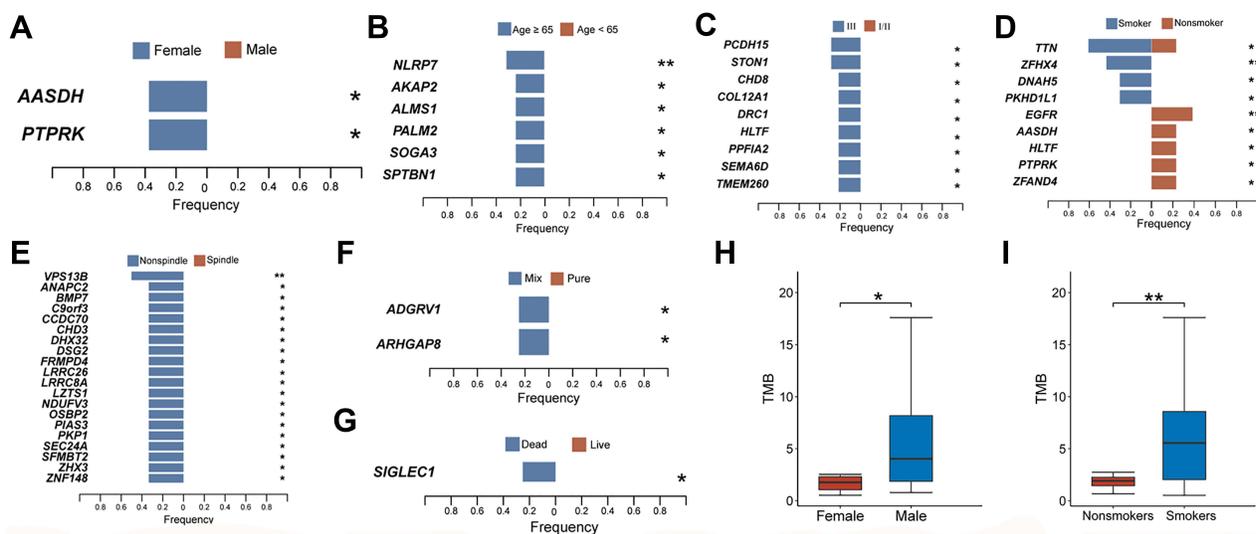


Figure 4 The association between clinical variables and gene mutation status. (A) *AASDH* and *PTPRK* mutation differentiated between female and male. (B) *NLRP3*, *AKAP2*, *ALMS1*, *PALM2*, *SOGA3*, *SPTBN1* mutation were frequently seen in PSC patients older than 65-year-old. (C) Gene mutation that are more frequently seen in PSC patients with stage III. (D) Gene mutations that are differentially expressed between smokers and non-smokers among PSC patients. (E) Genes that are mutated more frequently in non-spindle PSC patients. (F) Mutated genes that differentially expressed between mix and pure PSC. (G) *SIGLEC1* mutation is frequently seen in dead PSC patients. (H) Male PSC patients harbor higher TMB than their female counterpart. (I) Smokers harbor higher TMB than non-smoker among PSC patients. * $p < 0.05$; ** $p < 0.01$.

frequency of *ADGRV1* and *ARHGAP8* mutations compared with pure PSC ($p < 0.05$) (Figure 4F). Besides, higher frequency of *SIGLEC1* mutations tend to result in death in our cohort ($p < 0.05$) (Figure 4G). Further analyses revealed that the level of TMB was higher in male compared with female among PSC patients ($p < 0.05$) (Figure 4H). And smokers also boosted higher TMB value than non-smokers among PSC patients ($p < 0.01$) (Figure 4I).

To ascertain the association between morphology and gene mutation, we next analyzed the gene mutation in spindle cell and non-spindle cell PSC. Non-spindle PSC patients were found to be mutated at a higher frequency in *VPS13B*, *ANAPC2*, *BMP7*, *C9orf3*, *CCDC70*, *CHD3*, *DHX32*, *DSG2*, *FRMPD4*, *LRRC26*, *LRRC8A*, *LZST1*, *NDUFV3*, *OSBP2*, *PIAS3*, *PKPI*, *SEC24A*, *SFMBT2*, *ZHX3* and *ZNF148* compared with spindle subgroup. Similarly, PSC of mix pathological type had significantly higher frequency of *ADGRV1* and *ARHGAP8* mutations than that of pure PSC (Figure 5). Considering the association with morphology and pathological type with gene mutation among PSC patients, we further analyzed OS and immunological landscape of PSC patients with different morphology and pathological subtype. No significant difference in OS as well as immunological landscape as featured by TMB, PD-L1 and CD8⁺ T cell infiltration has been found between spindle-cell and non-spindle cell carcinoma. Likewise, PSC patients with mix and pure pathological subtype also witnessed no significant difference in OS, TMB and PD-L1, as demonstrated in [Supplementary Figure 2](#).

Survival Curves for PSC Patients with Different Mutation Status and Immunological Landscape

We further explore the impact of different gene mutations on OS. It is observed that *SARS* status has significant influence on OS. The presence of *SARS* mutation seems to be associated with significantly worse OS, suggesting the prognostic effect of *SARS* in the prediction of poor survival (log-rank $p = 0.04$) (Figure 6). As shown in [Supplementary Figure 3A](#), although OS was improved by a minimal degree, we did not observe significantly longer OS in PSC patients with higher TMB, PD-L1 expression, CD8⁺ T infiltration (log-rank $p = 0.6431$, log-rank $p = 0.1794$, log-rank $p = 0.06168$). Furthermore, we found that the combination of TMB, PD-L1 and CD8⁺ T cell could not significantly discriminate the population with different OS. In the present study, the threshold for TMB, PD-L1 and CD8⁺ T cell infiltration has been opted for 2.95, 1% and 5%, respectively. For PD-L1, those above the threshold have been deemed positive whereas those below the threshold have been deemed negative. For TMB and CD8⁺ T cell, those above the threshold have been deemed

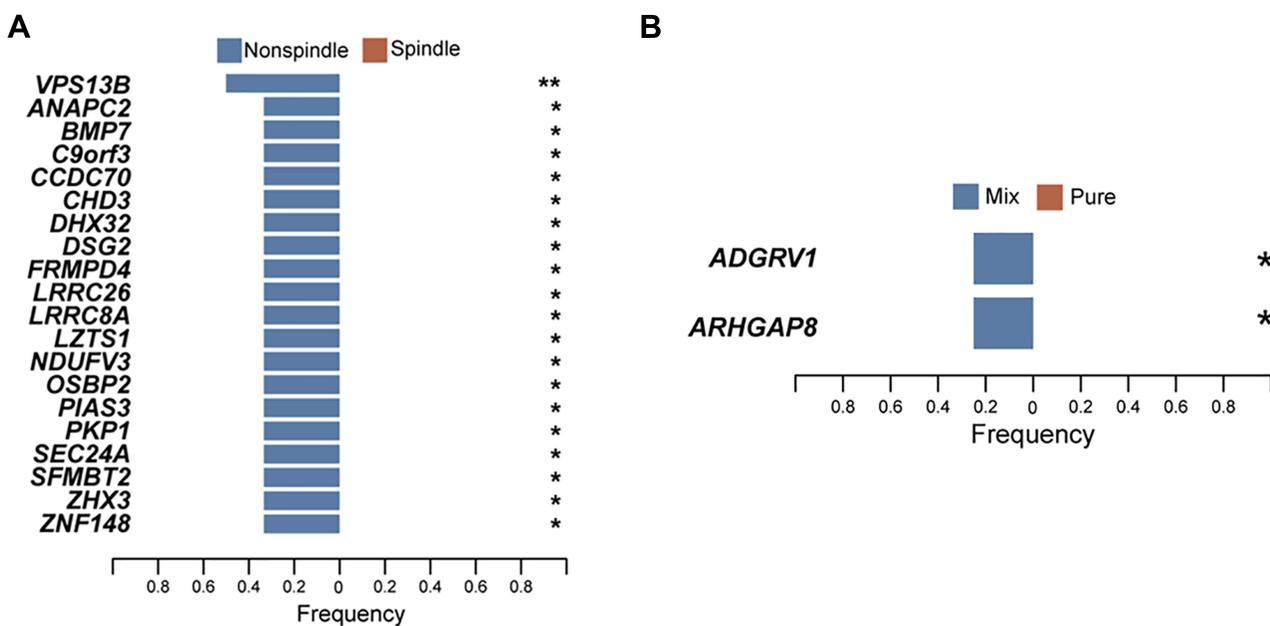


Figure 5 Gene mutations that are differentiated between non-spindle and spindle PSC, between mix and pure PSC. **(A)** Gene mutations that are differentiated between non-spindle and spindle PSC. **(B)** Gene mutations that are differentiated between mix and pure PSC. *: $p < 0.05$; **: $p < 0.01$.

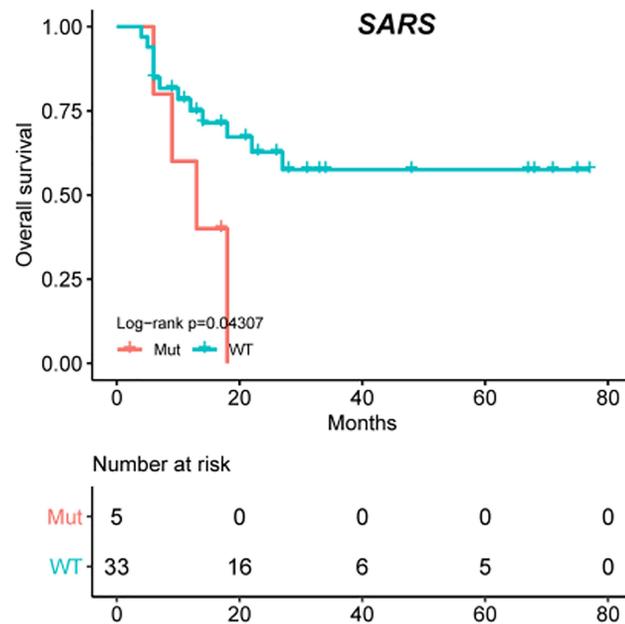


Figure 6 PSC patients harboring SARS mutations have worsened overall survival (log-rank p value = 0.04307).

higher whereas those below the threshold have been deemed lower. Prolonged OS was not observed in patients with $TMB \geq 2.95$ plus $PD-L1 \geq 1\%$ compared with others ($p=0.3065$). Similarly, OS was not significantly longer for those with $TMB \geq 2.95$ plus $CD8^+$ T cell infiltration $\geq 5\%$ compared with other groups ($p=0.4763$). And for those with $PD-L1 \geq 1\%$ plus $CD8^+$ T cell infiltration $\geq 5\%$, they also did not boost significantly prolonged OS in comparison with other groups ($p=0.06192$) ([Supplementary Figure 3B](#)). For PSC patients with relatively higher TMB, PD-L1 and $CD8^+$ T cell infiltration, OS was also not significantly elevated compared with other groups ($p=0.3562$) ([Supplementary Figure 3B](#)). Hazard ratio was numerically elevated in subgroups of TMB+ and PDL1+, TMB- and PDL1+, TMB+ and CD8-, TMB- and CD8+, PDL1+ and CD8+, PDL1+ and CD8- as compared with other groups ([Supplementary Figure 3C](#)). We next analyzed OS of patients with frequent mutation and found no significant difference in OS between patients with mutation in *TP53*, *TTN*, *ZFH4*, *USH2A*, *FAT3*, *MUC16*, *PKHD11*, *CSMD2*, *DST*, *DNAH5*, *CDKN2A*, *KRAS*, *RBM38*, *MDN1* and their wildtype counterpart ([Supplementary Figure 3D](#)).

Discussion

PSC is a rare subtype of lung cancer with poor prognosis and generally resistant to traditional chemotherapy. The present study summarized the landscape of genomic variations by whole exome sequencing (WES) and immune profiling by immunohistochemistry.

In the present study, the most frequent mutations in PSC are *TP53* (52.6%), *TTN* (47.4%), *ZFH4* (26.3%), *USH2A* (23.7%) and *FAT3* (23.7%). It is not surprising that mutation in *TP53* is the most frequently mutated gene in PSC due to its ubiquitous mutation detected in all types of cancers. The previous study reported the most frequent mutations in PSC were *TP53*, *CDKN2A* and *KRAS*, accounting for 73.6%, 37.6% and 34.4% using comprehensive genomic profiling (CGP). It has to be noted that CGP only covers genetic mutation that has clinical values.²⁹ In Yang's study, the top five mutated genes were *TP53*, *CDKN2A*, *MYC*, *MET* and *CCND1* in PSC.³⁰

Though we are not the first to conduct genomic analysis among PSC patients, our WES analysis of PSC patients boosts distinctive merits as a validation of previous findings on the mutational landscape of PSC. More importantly, we exhibited genomic and immune profiling in PSC patients of distinct morphology and pathological type, which has not been previously reported in other studies. It has to be noted that although no difference in TMB, PD-L1, $CD8^+$ T infiltration in PSC patients stratified by the pathological type was found. Mix type PSC demonstrated higher *ADGRV1* and *ARHGAP8* mutations compared with pure type. *ARHGAP8*, which is located within a critical region of loss-of-

heterozygosity on chromosome 22q13.31, has found to be overexpressed in breast and colorectal carcinomas.^{31,32} *ADGRV1* has also been reported in a series of biological processes such as myoclonic epilepsy and usher syndrome type II disease.^{33,34} Therefore, deeper probing into these two genes may help us better understand the genomic disparity between mix and pure PSC, thus offering optimized treatments.

A cohort of candidate genes also demonstrated enrichment for mutation in our PSC cohort. For instance, somatic mutations in *ZFHX4* gene are associated with poor overall survival of Chinese esophageal squamous cell carcinoma patients. Aberrations in *MUC16* have been described in ovarian and pancreatic cancers. *CSMD2*, has been reported to be implicated in the central nervous system of primary lymphoma and systemic melanoma metastasis. Our current findings have emphasized and implicated the potential role of these mutant genes involved in PSC development. Additionally, our results revealed that *SARS* mutation was indicative of worsened survival among PSC patients. *SARS*, also known as seryl tRNA synthetase (SerRS), was reported to affect glucose-induced lipid biosynthesis in breast cancer.³⁵ We therefore assume that the worse OS incurred by *SARS* mutation in PSC could also be possibly attributed to the dysregulated pathways that may be potentially activated.

Amounts of studies have shown that TMB, PD-L1 and immune cell infiltration are intimately correlated with the efficacy of immunotherapy.^{36–38} It is still not clear what molecular factors may affect patient responses to immunotherapies and whether PSC of different mutational status exhibit differential immune signatures. To this end, we measured immunological landscape of different mutational background and depicted their differences. We observed that patients with the significantly mutated genes *ZFHX4*, *TTN*, *TP53*, *KRAS*, *FAT3*, *PKHD1L1*, *CDKN2A*, *CSMD2*, *DNAH5*, *DST* showed higher TMB compared with the wild type. It is possible that neoantigens, resulted from the somatic nonsynonymous mutations, may account for higher TMB. The identification of these mutations among PSC patients opens the way for both targeted and immunotherapies.

Our results showed that PSC patients boosted with higher TMB, PD-L1 and CD8⁺ T cell infiltration had longer OS, despite being statistically insignificant. In addition, a number of studies have confirmed the association between TMB, PD-L1, CD8⁺ T infiltration and immunotherapy response, showing their higher intensity may render tumors higher immunogenicity, leading to improved clinical response to immunotherapy.^{39,40} Undeniably, there are several limitations in our study. First, the information on the number of patients should be provided who showed recurrence, who received cytotoxic chemotherapy, tyrosine kinase inhibitors, or immune checkpoint inhibitors were not collected. Moreover, it is a pity that the association between these markers and immune response was not explored in the whole PSC population, which warrants further study.

Conclusion

In the present study, first of all, we identified a cohort of gene and chromosomal mutation in PSC. We also analyzed the association between gene mutation at higher frequency and PD-L1, TMB and CD8⁺ T cell infiltration. Next, by the investigation of prognostic impact of gene mutation and immune landscape on survival, we revealed that a worsened survival could be predicted in PSC patients presenting with *SARS* mutation. Notably, mutational and immune landscape of PSC based on the morphology and pathological type have been depicted. In an analogous meaning, it has some clinical utility for the morphological and pathological type could be largely speculated according to the genomic mutation and immune landscape of PSC patients. We have demonstrated the genomic alterations of PSC patients. The description of genetic driver mutations has provided a new hope for personalized medicine going forward, bringing PSC into the era of targeted therapy and immunotherapy. However, more clinical trials are needed to further delineate possible genomic status of PSC and to predict clinical benefit from targeted and immunotherapies.

Data Sharing Statement

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics Approval and Consent to Participate

This study was approved by the ethics committee of Shandong Cancer Hospital and Institute.

Consent for Publication

This manuscript has been read and approved by all the authors to publish and is not submitted or under consideration for publication elsewhere.

Acknowledgments

We thank Dianbin Mu of Shandong cancer hospital and institute for his technical support for this research.

Funding

This study was supported jointly by Special Funds for Taishan Scholars Project (Grant No. tsqn201812149), Academic promotion program of Shandong First Medical University (2019RC004).

Disclosure

All authors declare that they have no competing interests.

References

1. Yendamuri S, Caty L, Pine M, et al. Outcomes of sarcomatoid carcinoma of the lung: a surveillance, epidemiology, and end results database analysis. *Surgery*. 2012;152(3):397–402. doi:10.1016/j.surg.2012.05.007
2. Vieira T, Girard N, Ung M, et al. Efficacy of first-line chemotherapy in patients with advanced lung sarcomatoid carcinoma. *J Thorac Oncol*. 2013;8(12):1574–1577. doi:10.1097/01.JTO.0000437008.00554.90
3. Liu X, Jia Y, Stoopler MB, et al. Next-generation sequencing of pulmonary sarcomatoid carcinoma reveals high frequency of actionable MET gene mutations. *J Clin Oncol*. 2016;34(8):794–802. doi:10.1200/JCO.2015.62.0674
4. Ung M, Rouquette I, Filleron T, et al. Characteristics and clinical outcomes of sarcomatoid carcinoma of the lung. *Clin Lung Cancer*. 2016;17(5):391–397. doi:10.1016/j.clcc.2016.03.001
5. Italiano A, Cortot AB, Ilie M, et al. EGFR and KRAS status of primary sarcomatoid carcinomas of the lung: implications for anti-EGFR treatment of a rare lung malignancy. *Int J Cancer*. 2009;125(10):2479–2482. doi:10.1002/ijc.24610
6. Leone A, Graziano P, Gasbarra R, et al. Identification of EGFR mutations in lung sarcomatoid carcinoma. *Int J Cancer*. 2011;128(3):732–735. doi:10.1002/ijc.25335
7. Ushiki A, Koizumi T, Kobayashi N, et al. Genetic heterogeneity of EGFR mutation in pleomorphic carcinoma of the lung: response to gefitinib and clinical outcome. *Jpn J Clin Oncol*. 2009;39(4):267–270. doi:10.1093/jjco/hyn155
8. Yu Y, Zeng D, Ou Q, et al. Association of survival and immune-related biomarkers with immunotherapy in patients with non-small cell lung cancer: a meta-analysis and individual patient-level analysis. *JAMA Netw Open*. 2019;2(7):e196879. doi:10.1001/jamanetworkopen.2019.6879
9. Jardim DL, Goodman A, de Melo Gagliato D, Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell*. 2021;39(2):154–173. doi:10.1016/j.ccell.2020.10.001
10. Bao R, Stapor D, Luke JJ. Molecular correlates and therapeutic targets in T cell-inflamed versus non-T cell-inflamed tumors across cancer types. *Genome Med*. 2020;12(1):90. doi:10.1186/s13073-020-00787-6
11. Reuss JE, Anagnostou V, Cottrell TR, et al. Neoadjuvant nivolumab plus ipilimumab in resectable non-small cell lung cancer. *J Immunother Cancer*. 2020;8(2):e001282. doi:10.1136/jitc-2020-001282
12. Acharya N, Madi A, Zhang H, et al. Endogenous glucocorticoid signaling regulates CD8(+) T cell differentiation and development of dysfunction in the tumor microenvironment. *Immunity*. 2020;53(3):658–671 e656. doi:10.1016/j.immuni.2020.08.005
13. Sanchez-Magrner L, Miles J, Baker CL, et al. High PD-1/PD-L1 checkpoint interaction infers tumor selection and therapeutic sensitivity to anti-PD-1/PD-L1 treatment. *Cancer Res*. 2020;80(19):4244–4257. doi:10.1158/0008-5472.CAN-20-1117
14. Marabelle A, Fakih M, Lopez J, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, Phase 2 KEYNOTE-158 study. *Lancet Oncol*. 2020;21(10):1353–1365. doi:10.1016/S1470-2045(20)30445-9
15. Ngiow SF, Young A, Jacquilot N, et al. A threshold level of intratumor CD8+ T-cell PD1 expression dictates therapeutic response to anti-PD1. *Cancer Res*. 2015;75(18):3800–3811. doi:10.1158/0008-5472.CAN-15-1082
16. Taube JM, Klein A, Brahmer JR, et al. Association of PD-1, PD-1 ligands, and other features of the tumor immune microenvironment with response to anti-PD-1 therapy. *Clin Cancer Res*. 2014;20(19):5064–5074. doi:10.1158/1078-0432.CCR-13-3271
17. Wouters MCA, Nelson BH. Prognostic significance of tumor-infiltrating B cells and plasma cells in human cancer. *Clin Cancer Res*. 2018;24(24):6125–6135. doi:10.1158/1078-0432.CCR-18-1481
18. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*. 2018;34(17):i884–i890. doi:10.1093/bioinformatics/bty560
19. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25(14):1754–1760. doi:10.1093/bioinformatics/btp324
20. McKenna A, Hanna M, Banks E, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res*. 2010;20(9):1297–1303. doi:10.1101/gr.107524.110
21. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res*. 2020;38(16):e164. doi:10.1093/nar/gkq603
22. Lawrence MS, Stojanov P, Polak P, et al. Mutational heterogeneity in cancer and the search for new cancer-associated genes. *Nature*. 2013;499(7457):214–218. doi:10.1038/nature12213

23. Martincorena I, Raine KM, Gerstung M, et al. Universal patterns of selection in cancer and somatic tissues. *Cell*. 2017;171(5):1029–1041 e1021. doi:10.1016/j.cell.2017.09.042
24. Mermel CH, Schumacher SE, Hill B, Meyerson ML, Beroukhi R, Getz G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011;12(4):R41–R41. doi:10.1186/gb-2011-12-4-r41
25. Gehring JS, Fischer B, Lawrence M, Huber W. SomaticSignatures. inferring mutational signatures from single-nucleotide variants. *Bioinformatics*. 2015;31(22):3673–3675. doi:10.1093/bioinformatics/btv408
26. 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
27. Bai Y, Chen X, Hou L, et al. PD-L1 expression and its effect on clinical outcomes of EGFR-mutant NSCLC patients treated with EGFR-TKIs. *Cancer Biol Med*. 2018;15(4):434–442. doi:10.20892/j.issn.2095-3941.2018.0223
28. Tokito T, Azuma K, Kawahara A, et al. Predictive relevance of PD-L1 expression combined with CD8+ TIL density in stage III non-small cell lung cancer patients receiving concurrent chemoradiotherapy. *Eur J Cancer*. 2016;55:7–14. doi:10.1016/j.ejca.2015.11.020
29. Schrock AB, Li SD, Frampton GM, et al. Pulmonary sarcomatoid carcinomas commonly harbor either potentially targetable genomic alterations or high tumor mutational burden as observed by comprehensive genomic profiling. *J Thorac Oncol*. 2017;12(6):932–942. doi:10.1016/j.jtho.2017.03.005
30. Yang Z, Xu J, Li L, et al. Integrated molecular characterization reveals potential therapeutic strategies for pulmonary sarcomatoid carcinoma. *Nat Commun*. 2020;11(1):4878. doi:10.1038/s41467-020-18702-3
31. Fritz G, Brachetti C, Bahlmann F, Schmidt M, Kaina B. Rho GTPases in human breast tumours: expression and mutation analyses and correlation with clinical parameters. *Br J Cancer*. 2002;87(6):635–644. doi:10.1038/sj.bjc.6600510
32. Johnstone CN, Castellvi-Bel S, Chang LM, et al. ARHGAP8 is a novel member of the RHOGAP family related to ARHGAP1/CDC42GAP/p50RHOGAP: mutation and expression analyses in colorectal and breast cancers. *Gene*. 2004;336(1):59–71. doi:10.1016/j.gene.2004.01.025
33. Myers KA, Nasioulas S, Boys A, et al. ADGRV1 is implicated in myoclonic epilepsy. *Epilepsia*. 2018;59(2):381–388. doi:10.1111/epi.13980
34. Jouret G, Poirsier C, Spodenkiewicz M, et al. Genetics of usher syndrome: new insights from a meta-analysis. *Otol Neurotol*. 2019;40(1):121–129. doi:10.1097/MAO.0000000000002054
35. Zhao J, Bai H, Li X, et al. Glucose-sensitive acetylation of Seryl tRNA synthetase regulates lipid synthesis in breast cancer. *Signal Transduct Target Ther*. 2021;6(1):303. doi:10.1038/s41392-021-00714-0
36. Somasundaram A, Burns TF. The next generation of immunotherapy: keeping lung cancer in check. *J Hematol Oncol*. 2017;10(1):87. doi:10.1186/s13045-017-0456-5
37. Liu L, Bai X, Wang J, et al. Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. *Clin Cancer Res*. 2019;25(24):7413–7423. doi:10.1158/1078-0432.CCR-19-0558
38. Reading JL, Galvez-Cancino F, Swanton C, et al. The function and dysfunction of memory CD8(+) T cells in tumor immunity. *Immunol Rev*. 2018;283(1):194–212. doi:10.1111/imr.12657
39. Goodman AM, Kato S, Bazhenova L, et al. Tumor mutational burden as an independent predictor of response to immunotherapy in diverse cancers. *Mol Cancer Ther*. 2017;16(11):2598–2608. doi:10.1158/1535-7163.MCT-17-0386
40. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol*. 2016;17(12):e542–e551. doi:10.1016/S1470-2045(16)30406-5

International Journal of General Medicine

Dovepress

Publish your work in this journal

The International Journal of General Medicine is an international, peer-reviewed open-access journal that focuses on general and internal medicine, pathogenesis, epidemiology, diagnosis, monitoring and treatment protocols. The journal is characterized by the rapid reporting of reviews, original research and clinical studies across all disease areas. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.

Submit your manuscript here: <https://www.dovepress.com/international-journal-of-general-medicine-journal>