

The average copy number variation (CNVA) of chromosome fragments is a potential surrogate marker for TMB in predicting responses to immunotherapy in Non-small cell lung cancer

Yuan yuan Lei

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Guo chao Zhang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital/Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Chao qi Zhang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

li yan xue

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Zhen lin Yang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Zhi liang Lu

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Jian bing Huang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Ruo chuan Zang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Yun Che

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Shuang shuang Mao

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Ling ling Fang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

cheng ming liu

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Xin feng Wang

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Su fei Zheng

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Nan Sun

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Jie He (✉ prof.jiehe@gmail.com)

National Cancer Center/National Clinical Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College

Research

Keywords: CNVA, Tumor mutational burden, Immunotherapy, Non-small cell lung cancer

Posted Date: May 18th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-28650/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Genomic instability plays a large role in the process of cancer. Tumor mutational burden (TMB) is closely related to immunotherapy outcome and is an important manifestation of genomic instability. However, the cost of TMB detection is extremely high, which limits the use of TMB in clinical practice. Another new indicator of genome instability, CNVA (the average copy number variation) which calculates the changes of 0.5 Mb chromosomal fragments, requires extremely low sequencing depth, and is expected to replace TMB as a new marker of immune efficacy.

Methods: A total of 50 samples (23 of which came from patients who received immunotherapy) were subjected to low-depth (10X) chromosome sequencing on the MGI platform. CNVA was calculated by the formula $\text{avg}(\text{abs}(\text{copy number}-2))$. Then, we analyzed the relationship between CNVA and immune infiltration or immunotherapy efficacy. In addition, through the analysis of whole genome sequencing data of 509 lung adenocarcinoma in the TCGA database, we compared CNVA with classic marker TMB to evaluate the value of CNVA as an immune evaluation index.

Results: Compared with the low CNVA group, the high CNVA group had higher expression of PD-L1, CD39 and CD19, and more infiltration of CD8⁺ T cells and CD3⁺ T cells. Among the 23 patients treated with immunotherapy, the average CNVA value of the SD (stable disease)/PR (partial response) group was higher than that of the PD (progressive disease) group ($P < 0.05$). The data of whole genome sequencing data of 509 lung adenocarcinomas from TCGA and real-time quantitative PCR results of 22 frozen specimens found that CNVA was more correlated with CD8 and PD-L1 than TMB. In addition, CNVA showed a specific positive correlation with TMB ($r = 0.2728, p < 0.0001$).

Conclusion: CNVA can be a good indicator of immune infiltration and predicting immunotherapy efficacy. With its low cost and potential clinical application for testing, it is expected to become a substitute for TMB.

Background

The advent of immunotherapy based on checkpoint inhibitors (ICIs) targeting the programmed cell death 1 (PD-1) axis has brought breakthrough progress to the treatment of non-small cell lung cancer. However, due to tumor heterogeneity and the dynamics of PD-L1 expression, only a subset of patients showed a durable clinical benefit from such therapies. The precise application of ICI therapy requires reasonable predictive indicators. According to reports, patients with a PD-L1 tumor proportion score (TPS) $> 50\%$ benefit more than those with a TPS $< 50\%$ [1, 2].

In addition to the TPS, one of the most important predictors of immune efficacy is tumor mutational burden (TMB). High TMB is associated with poor prognosis of various tumors[3–5]. More importantly, TMB is an independent biomarker of ICI response[6], as high TMB is associated with high neoantigen burden and high T-cell infiltration. Initially, the TMB test required whole exome sequencing (WES) of the specimen. It reflects the number of somatic nonsynonymous mutations including coding SNVs (single

nucleotide variation) and indels and is reported in numbers per Mb. However, the detection cost of WES is particularly high and the data are relatively difficult to analyze which is not suitable for large-scale analysis. Recently, some reports have simplified this process to estimate TMB with targeted next-generation sequencing (NGS) panels[7]. The most used are two assays: MSKIMPACT™ [8]and DFCI OncoPanel[9]. Compared that detected by WES, the number of genes detected by targeted NGS sequencing is small, which reduces the cost of testing to a certain extent. However, WES is only concerned with nonsynonymous mutations (leading to changes in the amino acid sequence and numbers of potential neoantigens within a tumor genome), but to make up for the insufficient number of genes, targeted NGS focuses on both synonymous and nonsynonymous mutations, and the depth of sequencing is also far higher than that for WES, making it difficult to analyze the data.

TMB is essentially caused by genomic instability. There are different causes of genomic instability, among which chromosomal instability (CIN), microsatellite instability (MIN) and DNA damage response (DDR) have attracted much attention. The relationship between microsatellite instability and cancer has been confirmed. For example, high MSI is closely related to the efficacy of immunotherapy. Researchers found that the majority (82.1%) of MSI-high tumors were TMB-high; however, only 18.3% of TMB-high tumors were MSI-H[10]. This shows that TMB is not caused only by MSI. In addition, the combination of MSI and TMB detection is better at predicting the efficacy of immunotherapy than TMB alone[10–12]. Specifically in colorectal cancer patients with high MSI, TMB is a better indicator of the efficacy of ICIs[13, 14]. In addition, in small cell lung cancer (SCLC), large number of DDR phenomena have been discovered and the DDR pathway alterations have a positive correlation with high TMB[15]. Since the cost of DDR testing is much lower than that of TMB test, it is expected to become a substitute for TMB[16]. However, the direct relationship between chromosomal instability and cancer is different in different situations. CIN is mainly reflected in the increase or deletion of fragments. The sequencing depth required for detection is also far lower than that of genome TMB detection, which greatly reduces the detection cost and analysis difficulty. The deletion or increase of the entire chromosome will lead to the generation of aneuploidy. Compared with normal diploid cells, most tumor cells are in an aneuploid state[17, 18]. Different types of ploidy have different effects on the occurrence and development of tumors[19]. The overall prognosis of patients with polyploid tumors is poor[20, 21], and polyploid tumors are more malignant than haploid or diploid tumors[22, 23]. However, the analysis of the entire chromosome can be used only for qualitative analysis of tumor specimens. It cannot provide specific values like TMB for quantitative analysis of different specimens, but it can be good for predicting the therapeutic effect.

In this article, to solve this problem, we broke the chromosome into small 0.5 Mb fragments, calculated the copy number changes of all fragments, averaged the absolute value (CNVA) and then evaluated the correlation between CNVA and tumor immune infiltration or immunotherapy effect; in addition, we compared the utility of CNVA with that of TMB. We found that specimens with high CNVA had more chromosomes alterations and, more obvious immune infiltration and were more sensitive to immunotherapy drugs than specimens with low CNVA.

Materials And Methods

Patient samples

A total of 23 patients with non-small cell lung cancer (NSCLC) treated with a PD-1 or PD-L1 inhibitor after surgery or biopsy were retrospectively identified from the Cancer Hospital of the Chinese Academy of Medical Sciences from April 2017 to June 2019. The fresh and frozen specimens used for flow sorting and real-time quantitative PCR came from another 27 surgical patients from the Cancer Hospital of the Chinese Academy of Medical Sciences from March 2019 to February 2020. Patient clinical

characteristics, tumor genomics, and outcome data were collected. Responses to PD-1/L1 targeted therapy were assessed using RECIST 1.1 guidelines and confirmed by the corresponding institute investigator. Ethics approval was granted by the Committee for the Ethics Review of Research Involving Human Subjects of the Cancer Hospital of the Chinese Academy of Medical Sciences.

CNVA analysis

First, the collected paraffin-embedded samples were dissolved, DNA was extracted, and then the DNA library was constructed with the MGI platform. High-quality genomic DNA with $A_{260} / A_{280} = 1.8 \sim 2.0$ was selected for fragmentation. After the fragmentation, the range of DNA distribution was wide, and we used magnetic bead to select specific fragments to control the final library fragment concentration. Those samples with concentrated DNA and high purity after physical fragmentation could be directly subject to end repair. If the DNA concentration range of the sample after physical fragmentation was wide, the fragments were sorted by length. Next, the ligated products were purified and PCR amplified. After constructing the library, we carried out subsequent chromosome low-depth(10X) sequencing. During data analysis, the entire chromosome was cut into t fragments of 0.5 Mb, and each chromosome was cut into more than six thousand fragments. Raw reads were aligned with Burrows-Wheeler Aligner software to determine the chromosomal origin of each sequenced DNA fragment. To calculate the copy number of each position (0.5 Mb), we subtracted 2 from the copy number, then took the absolute value and calculated the average number of all positions. CNVA was calculated with the formula $\text{avg}(\text{abs}(\text{copy number}-2))$.

Noninvasive prenatal testing (Nipt) analysis

The sequencing platform and database building method for Nipt analysis are the same as those for CNVA analysis. Unlike CNVA, which only focuses on copy number changes of 0.5 Mb chromosome fragments, Nipt focuses on changes in the entire chromosome, that is, macroscopic increases or deletions affecting the copy number of the entire chromosome. The specific data analysis method is described in the literature[24].

Analysis of TMB using the TCGA database

TCGA data were downloaded from the Genomic Data Commons Data Portal (GDC Data Portal, <https://portal.gdc.cancer.gov/>). In total, 509 lung adenocarcinoma (LUAD) patients from the TCGA database were analyzed. Somatic mutations were identified using the ecTMB algorithm. The correlation analysis between CNVA of selected samples and whole exome mutation burden was performed using Pearson correlation coefficient analysis.

Immunohistochemistry

The paraffin sections were gradually dewaxed and hydrated with xylene and alcohol. Antigens retrieval was performed in a microwave. After washing 3 times with PBS buffer, goat serum was applied to tissue slides to cover the whole section. The slides were incubated for 30 minutes at room temperature, taking care to prevent the surrounding tissue from drying out. After the goat serum had been used to block the slides, the excess serum was shaken off and diluted primary antibodies against PD-L1 (CST 13683S), CD8 (ab93278) and CD3 (PTG 17617-1-AP) were added to the specimen and incubated at 4 °C overnight. After incubating at room temperature for half an hour the next day, the secondary antibody was added, and the sample was incubated and finally developed by the BCA method.

Flow cytometry

Flow cytometric analysis was undertaken on representative fresh samples of lung cancer obtained at surgical resection. After mechanical disaggregation to obtain monodispersed cell suspensions, the tissue samples were treated with collagenase (Sigma-aldrich, 35285124) and DNase I (0.33 U/ml, Roche) and stained with anti-PD-L1 antibody and DAPI (KeyGEN BioTECH, KGA215-50).

Real-time quantitative PCR

Total RNA was isolated with the standard TRIzol-based protocol (Invitrogen). RT-PCR was performed on an ABI 7900HT Real-Time PCR thermocycler (Life Technologies). After the PCR was completed, the base and threshold of the amplification curve were manually set. Data were analyzed using the $2^{-\Delta\Delta CT}$ method. For the primer sequences used in the article, please refer to the Supplementary Materials.

Statistical analysis

Data analysis was performed using GraphPad Prism 6 (GraphPad Software, Inc.) or SPSS 24 (IBM, USA). Comparisons between two groups were performed using Student's t-test (two-tailed). Pearson correlations were used to evaluate the significance of the association between TMB and immune infiltration or between TMB and CNVA. Differences with p-values < 0.05 were considered significant.

Results

Distribution of chromosomal differences in patients treated with different immunotherapies

In this study, we included a total of 23 patients receiving immunotherapy and 27 non-immunotherapy patients undergoing thoracic surgery. The clinical information is shown in Table 1. According to the method mentioned above, we performed chromosomal variation detection with paraffin-embedded samples of these 50 specimens. Low-depth sequencing (10 X) was performed to analyze the copy numbers of these small fragments (0.5 Mb). The visualization of the fluctuation of these small fragments of chromosomes is shown in Fig. 1A-1C. We observed that chromosomes 1–23 showed different degrees of fluctuation in different specimens, and the copy number of the small fragments (0.5 Mb) of each chromosome fluctuated around the baseline. To quantify this fluctuation state, we proposed a new marker CNVA, which is the average of the copy numbers of all these small 0.5 Mb fragments, and is calculated with formula $\text{avg}(\text{abs}(\text{copy number}-2))$. When the CNVA value was large (0.4032) (Fig. 1A), the fluctuation degree of the chromosome was significantly increased, and the copy value of the 0.5 Mb fragment was randomly distributed around the baseline; however, when the CNVA value was small (0.1471) or moderate (0.238), the fluctuation of chromosomes was low (Fig. 1B and 1C). The above results indicated that the CNVA value was a good indicator of chromosome fluctuations.

Table 1
Clinical information for immunotherapy and non-immunotherapy patients

| Characteristics | Patient (n = 50) | | CNVA value | | P-value |
|-----------------|------------------|-------|------------|------|---------|
| | NO. | % | low | high | |
| All patients | 50 | 100.0 | 24 | 26 | |
| Age | | | | | |
| ≤ 60 | 27 | 54.0 | 15 | 12 | 0.247 |
| > 60 | 23 | 46.0 | 9 | 14 | |
| gender | | | | | |
| male | 29 | 58.0 | 12 | 17 | 0.271 |
| female | 21 | 42.0 | 12 | 9 | |
| Tumor size | | | | | |
| < 4 cm | 36 | 75.0 | 18 | 18 | 1.000 |
| ≥ 4 cm | 12 | 25.0 | 6 | 6 | |
| Tumor stage | | | | | |
| I - II | 32 | 65.3 | 14 | 18 | 0.414 |
| III - IV | 16 | 32.7 | 9 | 7 | |
| LN metastasis | | | | | |
| Absent | 28 | 56.0 | 12 | 16 | 0.412 |
| Present | 22 | 44.0 | 12 | 10 | |

Discussion

Distribution of chromosomal differences in patients treated with different immunotherapies

In this study, we included a total of 23 patients receiving immunotherapy and 27 non-immunotherapy patients undergoing thoracic surgery. The clinical information is shown in Table 1. According to the method mentioned above, we performed chromosomal variation detection with paraffin-embedded samples of these 50 specimens. Low-depth sequencing (10 X) was performed to analyze the copy numbers of these small fragments (0.5 Mb). The visualization of the fluctuation of these small fragments of chromosomes is shown in figure 1A-1C. We observed that chromosomes 1-23 showed different degrees of fluctuation in different specimens, and the copy number of the small fragments (0.5 Mb) of each chromosome fluctuated around the baseline. To quantify this fluctuation state, we proposed a new

marker CNVA, which is the average of the copy numbers of all these small 0.5 Mb fragments, and is calculated with formula $\text{avg}(\text{abs}(\text{copy number}-2))$. When the CNVA value was large (0.4032) (Fig. 1A), the fluctuation degree of the chromosome was significantly increased, and the copy value of the 0.5 Mb fragment was randomly distributed around the baseline; however, when the CNVA value was small (0.1471) or moderate (0.238), the fluctuation of chromosomes was low (Fig. 1B and 1C). The above results indicated that the CNVA value was a good indicator of chromosome fluctuations.

High CNVA value indicates high immune infiltration status

Among these 50 specimens, we found that CNVA values were mainly distributed between 0.1 and 0.5 (Fig. 2A). The CNVA value was not related to age, sex, tumor size and stage (Table 1). To further prove that the CNVA value could indeed reflect the degree of chromosome fluctuations, in addition to focusing on the copy number of 0.5 Mb fragments, we used the Nipt algorithm[24] to analyze amplification or deletion of the entire chromosome. The value returned by the algorithm was defined as the Z value, Z value > 3 indicated a triploid or higher state and Z value < -3 indicated a haploid or lower state. According to the CNVA level, we divided the 50 specimens into high, moderate and low groups. Then we drew a heat map based on the CNVA and Z values. Green represents the group with low CNVA value, orange represents the group with high CNVA value, and blue represents the group with moderate CNVA value. The lightest color in the heat map, that is, the part with the least degree of chromosome aneuploidy, was basically concentrated in the group with low or moderate CNVA value. Conversely, the high CNVA group was particularly dark in color, and the chromosome aneuploidy state was more obvious (Fig. 2B). This proved that CNVA could indeed reflect the aneuploid or stable state of the entire chromosome through the fluctuation of 0.5 Mb fragments.

For the 50 specimens, there were also 27 nonimmunotherapy frozen tissue. To prove the relationship between CNVA and immune infiltration, we extracted RNA from these 27 frozen tissues specimens. Five of the specimens were of poor quality and were discarded. After that, we detected the expression of immune-related genes via immunohistochemistry in the remaining 22 samples by real-time quantitative PCR. The results showed that compared with the CNVA low group, the high CNVA group had significantly upregulated levels of CD8 ($p=0.0067$), PD-L1 (0.0477), CD39 (0.0337) and the B cell marker gene CD19 (0.0464). The Average expression level of immunosuppressive genes Foxp3 and CTLA-4 were down-regulated in the CNVA high group but there was no statistical difference (Fig. 2C and Fig. 7E). In addition, we also detected the expression of CD3, CD8 and PD-L1 in 50 immunotherapy and nonimmunotherapy samples by immunohistochemistry. The results also showed that, compared with the group with low CNVA, the high CNVA had higher expression of CD3 (0.0370), CD8 (0.0495) and PD-L1 (Fig. 2D-2J). The above results indicate that high CNVA is positively correlated with immune infiltration.

Correlation between the effect of immunotherapy and immune infiltration

We retrospectively collected clinical information on patients receiving immunotherapy at the Cancer Hospital of the Chinese Academy of Medical Sciences. Among them, 21 patients underwent surgery before receiving immunotherapy and had paraffin-embedded samples available. The other two patients

underwent biopsy before immunotherapy, and paraffin samples were also available. The clinical information is shown in Table 2. Of these 23 patients, 14 were lung adenocarcinoma patients and 9 were squamous cell carcinoma patients. All patients received second-, third-, or fourth-line immunotherapy after surgery or biopsy. Based on the evaluation of the efficacy after 12 weeks of immunotherapy, of these 23 patients, 12 were judged as achieving SD or a PR and 11 were judged having PD (efficacy assessment was completed by the clinician). Patients with SD or PR were classified into a single group with good efficacy, and patients with PD was classified as a group with poor efficacy. We performed immunohistochemical analysis on the paraffin samples of these 23 patients. The results showed that compared with the group with poor efficacy, group with good efficacy showed increased expression of the classic immune infiltration marker CD3 (Fig. 3A-3C) ($p < 0.05$), CD8 (Fig. 3D-3F) (There was no statistical difference which might be due to the small sample size) and PD-L1 (Fig. 3G-3I). This result is consistent with literature reports[25].

High CNVA value may predict a superior immunotherapy effect

After proving that CNVA had a specific correlation with immune infiltration, we next compared the relationship between immunotherapy efficacy and CNVA in the 23 patients. The CNVA values of these patients were mainly between 0.1 and 0.4 (Fig. 4A). We divided these 23 patients into two groups according to efficacy, patients with PD as one group and patient with PR/SD as another group. It can be seen from the scatter plot that the average CNVA value of the PD group was lower than that of the PR/SD group with statistically significant difference ($p = 0.0204$) (Fig. 4B). However, we only included a small number of immunotherapy patients and these patients are not strictly first-line immunotherapy patients, as many of them have undergone radiotherapy and chemotherapy or targeted therapy. These treatment methods may affect certain chromosomes fragments. Therefore, we further analyzed the copy number (aneuploidy) of the whole chromosome of chromosomes 1-23 in different therapeutic groups by the Nipt method. The frequency of deletion or amplification of the entire chromosome is very likely to be lower than that of the small 0.5 Mb fragment, and it will be more stable under the influence of various treatment methods. Indeed, we found that the PR/SD group was mainly located in the lighter color area of the heatmap, while the PD group was basically in the darker color area, that is, the area where the chromosome aneuploidy phenomenon was more obvious (Fig. 4C). All the result above confirmed that the degree of chromosome fluctuation was indeed related to the therapeutic effect.

Verification of the correlation between chromosome copy number changes in terms of DNA content and immune infiltration

The most intuitive results from changes in the copy number of chromosomes is changes in the DNA content. According to change in DNA content, it can be roughly speculated whether there is chromosome aneuploid or a major change in chromosome copy number. However, such changes will be noted only when the difference in DNA content caused by the change in the number of chromosomes can be detected. Therefore, a normal DNA content cannot exclude the existence of malignant or chromosomal abnormalities, but an abnormal DNA content always indicates chromosomal abnormalities. In other

words, changes in DNA content will be more specific in reflecting chromosomal instability. Therefore, we can indirectly confirm the relationship between chromosome copy number abnormalities and immune infiltration by analyzing the relationship between DNA content and immune infiltration. We detected the relationship between DNA content (fluorescence value of DAPI) and PD-L1 in a series of lung cancer cell lines and lung cancer tissue samples by flow cytometry. A schematic diagram of the PD-L1 isotype control and positive specimen is shown in figure 5A-5D. We observed that in 15 lung cancer cell lines, the fluorescence intensity of DAPI was basically negatively correlated with the expression of PD-L1 (Fig. 5E). However, in lung cancer tissue specimens, the optimal curve of the trend line was a polynomial of two terms (Fig. 5F). These results showed that when a certain baseline was used as the standard (probably the DNA content of normal diploid cells), the DNA content changed towards one of two extremes,, which could have either very small or very large numbers of chromosome copies (predicting chromosome fluctuations), PD-L1 expression would increase and the result was indeed similar to the CNVA effect (with an emphasis on the absolute value of copy number changes). But the result of the cell lines did not develop in two extreme directions, we speculated that it might be related to the small number of cell lines and the relative concentration of DNA content.

The state of immune infiltration is related to the level of TMB

TMB is a widely studied indicator related to the efficacy of immunotherapy. We further analyzed the correlation between TMB and various infiltrating immune cells in lung adenocarcinoma. We downloaded the information of 509 lung adenocarcinoma patients with both whole genome and transcriptome data from the TCGA database. Whole-genome data were used to calculate the TMB value by ecTMB and the transcriptome data were used to analyze the infiltration of various immune cells by CIBERSORT (<https://cibersort.stanford.edu/>). According to the TMB value, these 509 cases were divided into two groups: TMB high and TMB low. The results of the heat map clustering showed that the group with high TMB (blue) was roughly clustered together and had higher levels of CD8⁺ T cells and M1 type macrophages and lower M2 type macrophages than the low TMB group (Fig. 6A). To observe the difference in infiltrating lymphocytes between the TMB high and TMB low groups more clearly, we subdivided the 509 samples into three groups, TMB high, TMB moderate and TMB low, and plotted histograms. Compared with the group with low TMB, we found that the TMB high group, in addition to having obvious increases in CD8⁺ T cells ($p < 0.0001$) (Fig. 6B and Fig. 7C), also had significant increases in plasma cells and activated CD4⁺ T cells ($p < 0.0001$) (Fig. 6B). In addition, the ratio of M1 to M2 macrophages (which serves as an index of immune activity) in the group with high TMB was much higher than that in the group with low TMB ($p < 0.0001$) (Fig. 6C).

Comparison of the utility of CNVA and TMB in predicting response to immunotherapy

Both CNVA and TMB essentially reflect genome instability. The former is concerned with the copy number changes of chromosome fragments of 0.5 Mb at the macro level, while TMB is only concerned with the number of mutations per 1 Mb DNA. In theory, the two values should be correlated. The cost of TMB testing is extremely high and the analysis is difficult. We wondered whether CNVA could replace TMB as

an indicator of immunotherapy efficacy. First, we compared the changes in the expression levels of two indexes PD-L1 and CD8, which were closely related to immunotherapy efficacy under different TMB or CNVA states. The TMB value was extracted from 509 samples from the TCGA database. The results showed that the expression of PD-L1 and the infiltration of CD8⁺ T cells in the group with high TMB were significantly increased ($p < 0.01$) (Fig. 7A and 7C), and both had a certain correlation with TMB (although the correlation coefficient was not large) (Fig. 7B and 7D). The CNVA data came from the abovementioned 27 nonimmunotherapy specimens, 22 of which had high-quality RNA for real-time quantitative PCR. We also found high expression of PD-L1 and CD8 in samples with high CNVA ($p < 0.05$) and this correlation was stronger than the correlation between high TMB and these two parameters (Fig. 7E and 7G). In addition, the correlation coefficient was also higher than that for TMB, although the P value of PD-L1 is not as significant as TMB due to the lack of samples (Fig. 7F and 7H).

Finally, to further verify that there was a correlation between CNVA and TMB, we reanalyzed the whole genome data of 509 samples from TCGA. Because the whole genome sequencing process does not include break the chromosome into small fragments of 0.5 Mb, and the data present only changes in mutations or copy number throughout the genome, we analyzed the changes in the copy number of all exons and introns, and the average value was calculated and used to replace the CNVA value. Pearson correlation showed that there was indeed a clear correlation between TMB and CNVA values ($r = 0.2728$, $p < 0.0001$) (Fig. 7I). These results indicated that CNVA had the potential to replace TMB.

Conclusions

CNVA is an excellent marker to indicate immune infiltration and predict immune efficacy, and is expected to become a potential substitute for TMB.

Declarations

Ethics approval and consent to participate

Ethics approval was granted by the Committee for the Ethics Review of Research Involving Human Subjects of the Cancer Hospital of the Chinese Academy of Medical Sciences.

Consent for publication

Not applicable.

Availability of data and materials

The analyzed data sets generated during the study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Natural Science Foundation of China, CAMS Innovation Fund for Medical Sciences (2017-I2M-1-005, 2016- I2M-1-001), the National Key Research and Development Program of China (grant numbers 2016YFC1303201, 2016YFC0901400), the National Natural Science Foundation of China(81802299, 81502514),the Fundamental Research Funds for the Central Universities (3332018070), and the National Key Basic Research Development Plan (2018YFC1312105).

Author information

Affiliations

National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical School, Beijing

Yuanyuan Lei, Guochao zhang, Chaoqi Zhang, Zhenlin Yang, Zhiliang Lu, Jianbing Huang, Ruochuan Zang, Yun Che, Shuangshuang Mao, Lingling Fang, chengming Liu, Xinfeng Wang, Sufei Zheng, Nan Sun, Jie He

Acknowledgments

We thank Min Wu for the excellent technical assistance.

Abbreviations

CNVA The average copy number variation SD stable disease; PD: progressive disease; PR: partial response

References

1. Gainor, J.F., et al., *Clinical activity of programmed cell death 1 (PD-1) blockade in never, light, and heavy smokers with non-small-cell lung cancer and PD-L1 expression ≥ 50* . *Ann Oncol*, 2020. **31**(3): p. 404-411.
2. Reck, M., et al., *Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer*. *N Engl J Med*, 2016. **375**(19): p. 1823-1833.
3. Samstein, R.M., et al., *Tumor mutational load predicts survival after immunotherapy across multiple cancer types*. *Nat Genet*, 2019. **51**(2): p. 202-206.
4. Lee, D.W., et al., *Tumor Mutation Burden and Prognosis in Patients with Colorectal Cancer Treated with Adjuvant Fluoropyrimidine and Oxaliplatin*. *Clin Cancer Res*, 2019. **25**(20): p. 6141-6147.

5. Wang, F., et al., *Safety, efficacy and tumor mutational burden as a biomarker of overall survival benefit in chemo-refractory gastric cancer treated with toripalimab, a PD-1 antibody in phase Ib/II clinical trial NCT02915432*. Ann Oncol, 2019. **30**(9): p. 1479-1486.
6. Ready, N., et al., *First-Line Nivolumab Plus Ipilimumab in Advanced Non-Small-Cell Lung Cancer (CheckMate 568): Outcomes by Programmed Death Ligand 1 and Tumor Mutational Burden as Biomarkers*. J Clin Oncol, 2019. **37**(12): p. 992-1000.
7. Budczies, J., et al., *Optimizing panel-based tumor mutational burden (TMB) measurement*. Ann Oncol, 2019. **30**(9): p. 1496-1506.
8. Rizvi, H., et al., *Molecular Determinants of Response to Anti-Programmed Cell Death (PD)-1 and Anti-Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non-Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing*. J Clin Oncol, 2018. **36**(7): p. 633-641.
9. Ricciuti, B., et al., *Use of targeted next generation sequencing to characterize tumor mutational burden and efficacy of immune checkpoint inhibition in small cell lung cancer*. J Immunother Cancer, 2019. **7**(1): p. 87.
10. Goodman, A.M., et al., *Microsatellite-Stable Tumors with High Mutational Burden Benefit from Immunotherapy*. Cancer Immunol Res, 2019. **7**(10): p. 1570-1573.
11. Georgiadis, A., et al., *Noninvasive Detection of Microsatellite Instability and High Tumor Mutation Burden in Cancer Patients Treated with PD-1 Blockade*. Clin Cancer Res, 2019. **25**(23): p. 7024-7034.
12. Luchini, C., et al., *ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach*. Ann Oncol, 2019. **30**(8): p. 1232-1243.
13. Schrock, A.B., et al., *Tumor mutational burden is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer*. Ann Oncol, 2019. **30**(7): p. 1096-1103.
14. Innocenti, F., et al., *Mutational Analysis of Patients With Colorectal Cancer in CALGB/SWOG 80405 Identifies New Roles of Microsatellite Instability and Tumor Mutational Burden for Patient Outcome*. J Clin Oncol, 2019. **37**(14): p. 1217-1227.
15. Park, S., et al., *DNA Damage Response and Repair Pathway Alteration and Its Association With Tumor Mutation Burden and Platinum-Based Chemotherapy in SCLC*. J Thorac Oncol, 2019. **14**(9): p. 1640-1650.
16. Arulananda, S., P. Mitchell, and T. John, *DDR Alterations as a Surrogate Marker for TMB in SCLC - Use it or Lose it?* J Thorac Oncol, 2019. **14**(9): p. 1498-1500.
17. Pinto, A.E., et al., *Frequent 7q gains in flow cytometric multiploid/hypertetraploid breast carcinomas: a study of chromosome imbalances by comparative genomic hybridisation*. J Clin Pathol, 2006. **59**(4): p. 367-72.
18. Sugai, T., et al., *Analysis of genetic alterations associated with DNA diploidy, aneuploidy and multiploidy in gastric cancers*. Oncology, 2005. **68**(4-6): p. 548-57.
19. Sugai, T., et al., *Analysis of genetic alterations, classified according to their DNA ploidy pattern, in the progression of colorectal adenomas and early colorectal carcinomas*. J Pathol, 2003. **200**(2): p. 168-

76.

20. Michels, J.J., et al., *Flow cytometry in primary breast carcinomas: prognostic impact of multiploidy and hypoploidy*. Cytometry B Clin Cytom, 2003. **55**(1): p. 37-45.
21. Michels, J.J., et al., *Flow cytometry in primary breast carcinomas: prognostic impact of S-phase fraction according to different analysis patterns*. Cytometry B Clin Cytom, 2004. **59**(1): p. 32-9.
22. Carloni, S., et al., *DNA ploidy and S-phase fraction analysis in peritoneal carcinomatosis from ovarian cancer: correlation with clinical pathological factors and response to chemotherapy*. Onco Targets Ther, 2017. **10**: p. 4657-4664.
23. Dayal, J.H., et al., *Multiparameter DNA content analysis identifies distinct groups in primary breast cancer*. Br J Cancer, 2013. **108**(4): p. 873-80.
24. Chiu, R.W., et al., *Noninvasive prenatal diagnosis of fetal chromosomal aneuploidy by massively parallel genomic sequencing of DNA in maternal plasma*. Proc Natl Acad Sci U S A, 2008. **105**(51): p. 20458-63.
25. Jing, C.Y., et al., *HHLA2 in intrahepatic cholangiocarcinoma: an immune checkpoint with prognostic significance and wider expression compared with PD-L1*. J Immunother Cancer, 2019. **7**(1): p. 77.
26. Monteiro, L.S., et al., *DNA content in malignant salivary gland tumours*. Oral Dis, 2009. **15**(4): p. 295-301.

Tables

Table 1 Clinical information for immunotherapy and non-immunotherapy patients

| Characteristics | Patient (n=50) | | CNVA value | | P-value |
|-----------------|----------------|-------|------------|------|---------|
| | NO. | % | low | high | |
| All patients | 50 | 100.0 | 24 | 26 | |
| Age | | | | | |
| ≤ 60 | 27 | 54.0 | 15 | 12 | 0.247 |
| > 60 | 23 | 46.0 | 9 | 14 | |
| gender | | | | | |
| male | 29 | 58.0 | 12 | 17 | 0.271 |
| female | 21 | 42.0 | 12 | 9 | |
| Tumor size | | | | | |
| < 4 cm | 36 | 75.0 | 18 | 18 | 1.000 |
| ≥4cm | 12 | 25.0 | 6 | 6 | |
| Tumor stage | | | | | |
| I - II | 32 | 65.3 | 14 | 18 | 0.414 |
| III - IV | 16 | 32.7 | 9 | 7 | |
| LN metastasis | | | | | |
| Absent | 28 | 56.0 | 12 | 16 | 0.412 |
| Present | 22 | 44.0 | 12 | 10 | |

Table 2 Specific treatment information for immunotherapy patients

| Patient ID | TNM stage | pathology | Specimen category | Immune-therapy | treatment effect |
|------------|-----------|-------------------------|-------------------|----------------|------------------|
| 1 | T2N2M0 | Adenocarcinoma | surgery | Third line | PD |
| 2 | T2aN0M0 | Adenocarcinoma | surgery | Second line | PR |
| 3 | T2N0M0 | Adenocarcinoma | surgery | Second line | SD |
| 4 | T2aN0M0 | Adenocarcinoma | surgery | Third line | PD |
| 5 | T2N2M0 | Adenocarcinoma | surgery | Third line | PD |
| 6 | T2aN0M0 | Squamous cell carcinoma | surgery | Second line | SD |
| 7 | T2bN0M0 | Adenocarcinoma | surgery | Second line | SD |
| 8 | T1aN1M0 | Squamous cell carcinoma | surgery | Third line | SD |
| 9 | T2N0M0 | Squamous cell carcinoma | surgery | Second line | PR-SD |
| 10 | T1aN2M0 | Adenocarcinoma | surgery | Third line | PD |
| 11 | T2N1M0 | Squamous cell carcinoma | surgery | Third line | PD |
| 12 | T3N2M0 | Adenocarcinoma | surgery | Third line | PD |
| 13 | T2N0M0 | Adenocarcinoma | surgery | Third line | PR |
| 14 | T3N0M0 | Squamous cell carcinoma | surgery | Third line | PD |
| 15 | T2aN1M0 | Squamous cell carcinoma | surgery | Third line | PR |
| 16 | T2N2M0 | Adenocarcinoma | surgery | Fourth line | PD |
| 17 | T2N2M1 | Squamous cell carcinoma | surgery | Fourth line | PD |
| 18 | T1bN0M0 | Adenocarcinoma | surgery | Fourth line | SD |
| 19 | T2N0M0 | Squamous cell carcinoma | surgery | Third line | SD |
| 20 | ypT1N2M0 | Squamous cell carcinoma | surgery | Third line | PD |
| 21 | T3N2M0 | Adenocarcinoma | surgery | Third line | PD |
| 22 | / | Adenocarcinoma | biopsy | Fourth line | SD |
| 23 | / | Adenocarcinoma | biopsy | Third line | SD |

Abbreviations: *PD* progressive disease, *SD* stable disease, *PR* partial response

Figures

figure 1

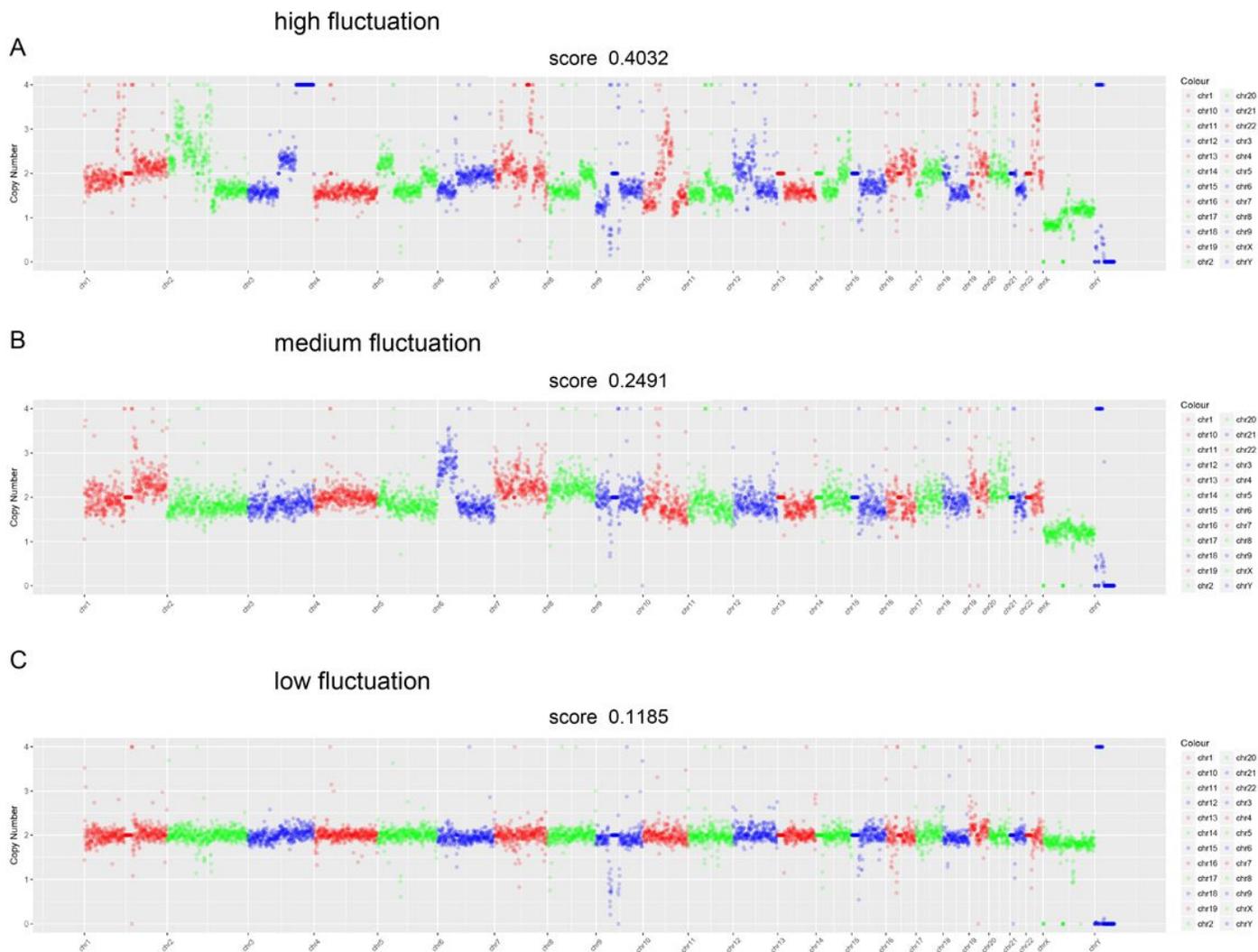


Figure 1

Fluctuation status of chromosomes in different specimens (A). When the CNVA score was 0.4032, the chromosome shows a high fluctuation state. (B). When the CNVA score was 0.4032, the chromosome shows a moderate fluctuation state. (C). When the CNVA score was 0.1185, the chromosome shows a low degree of fluctuation.

figure 2

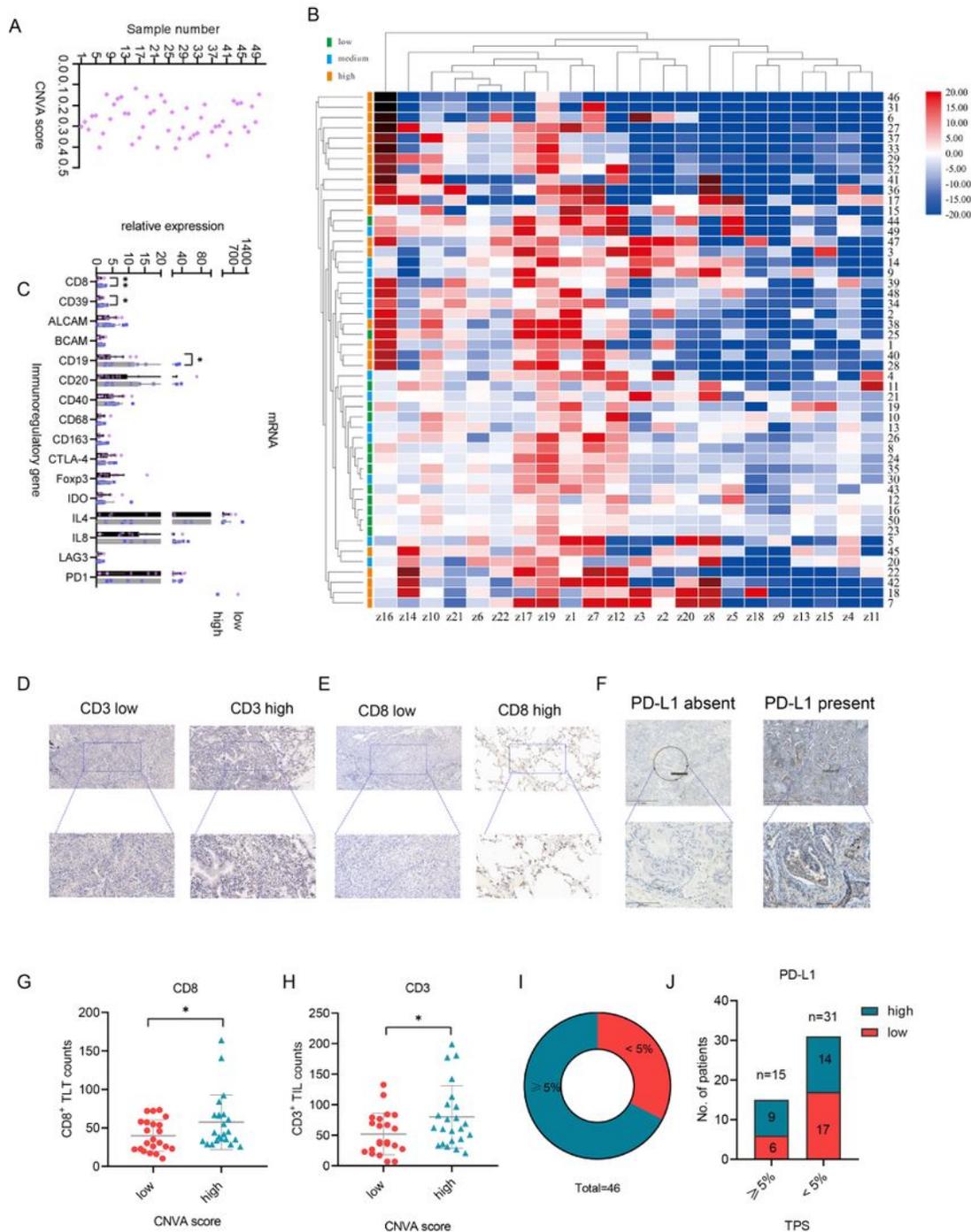


Figure 2

High CNVA value indicated high immune infiltration status (A). Scatterplot depicted the distribution of CNVA score in 50 specimens. (B). The entire deletion or increase of chromosomes 1-23 in the three groups of CNVA score high, moderate and low. Z value > 3 meant chromosome is triploid or more, Z value < -3 meant chromosome is haploid or less. (C). The histogram showed the mRNA expression of different immunoblot-related genes in 22 frozen specimens detected by q-PCR. (D-F). Representative

immunohistochemical pictures of CD8, CD3 and PD-L1. (G-H). Scatterplot showed immunohistochemical scores of CD8 and CD3 in high CNVA group and low CNVA group. (I). Proportion of 46 specimens with PD-L1 TPS $\geq 5\%$ and $< 5\%$. (J). The histogram showed the specific distribution of CNVA score in PD-L1 $\geq 5\%$ and $< 5\%$ groups. Data are presented as the mean \pm SD, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ns means no significance.



Figure 3

Relationship between immune efficacy and immune infiltration (A). Scatterplot depicted the distribution of CNVA score in 50 specimens. (B). The entire deletion or increase of chromosomes 1-23 in the three groups of CNVA score high, moderate and low. Z value > 3 meant chromosome is triploid or more, Z value < -3 meant chromosome is haploid or less. (C). The histogram showed the mRNA expression of different immunoblot-related genes in 22 frozen specimens detected by q-PCR. (D-F). Representative immunohistochemical pictures of CD8, CD3 and PD-L1. (G-H). Scatterplot showed immunohistochemical scores of CD8 and CD3 in high CNVA group and low CNVA group. (I). Proportion of 46 specimens with PD-L1 TPS $\geq 5\%$ and $< 5\%$. (J). The histogram showed the specific distribution of CNVA score in PD-L1 $\geq 5\%$ and $< 5\%$ groups. Data are presented as the mean \pm SD, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. ns means no significance.

figure 4

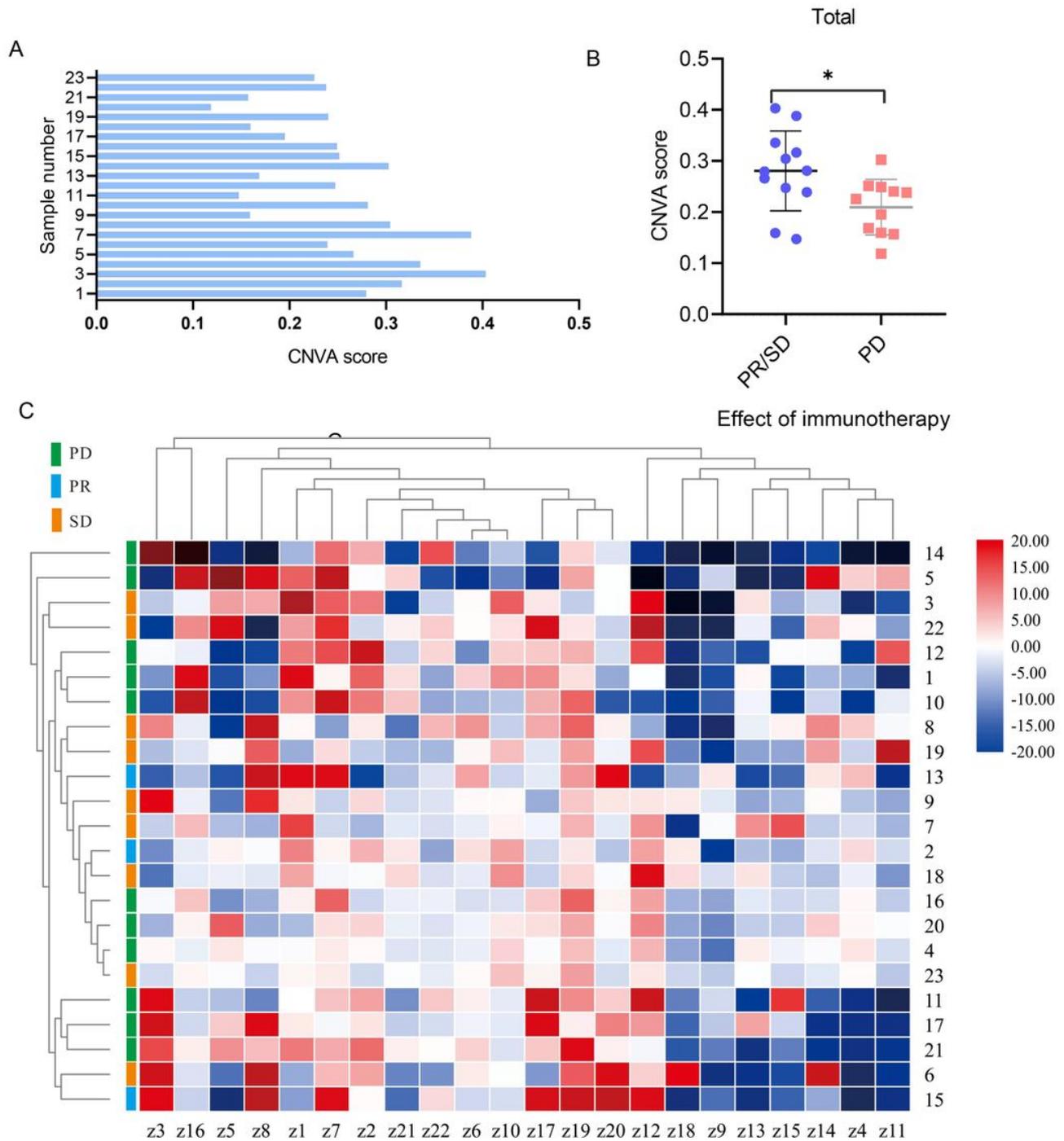


Figure 4

Correlation between CNVA score and immune efficacy (A). Histogram showed the distribution of CNVA values in 23 immunotherapy specimens. (B). The scatterplot showed the distribution of CNVA score in the PR/SD group and PD group. (C). Heat map displayed complete deletion or increase of chromosomes 1-23 in PR group, SD group and PD group. Z value > 3 meant chromosome is triploid or more, Z value < -3 meant chromosome is haploid or less. Data are presented as the mean \pm SD, * p < 0.05, ** p < 0.01, *** p < 0.001. ns means no significance.

figure 5

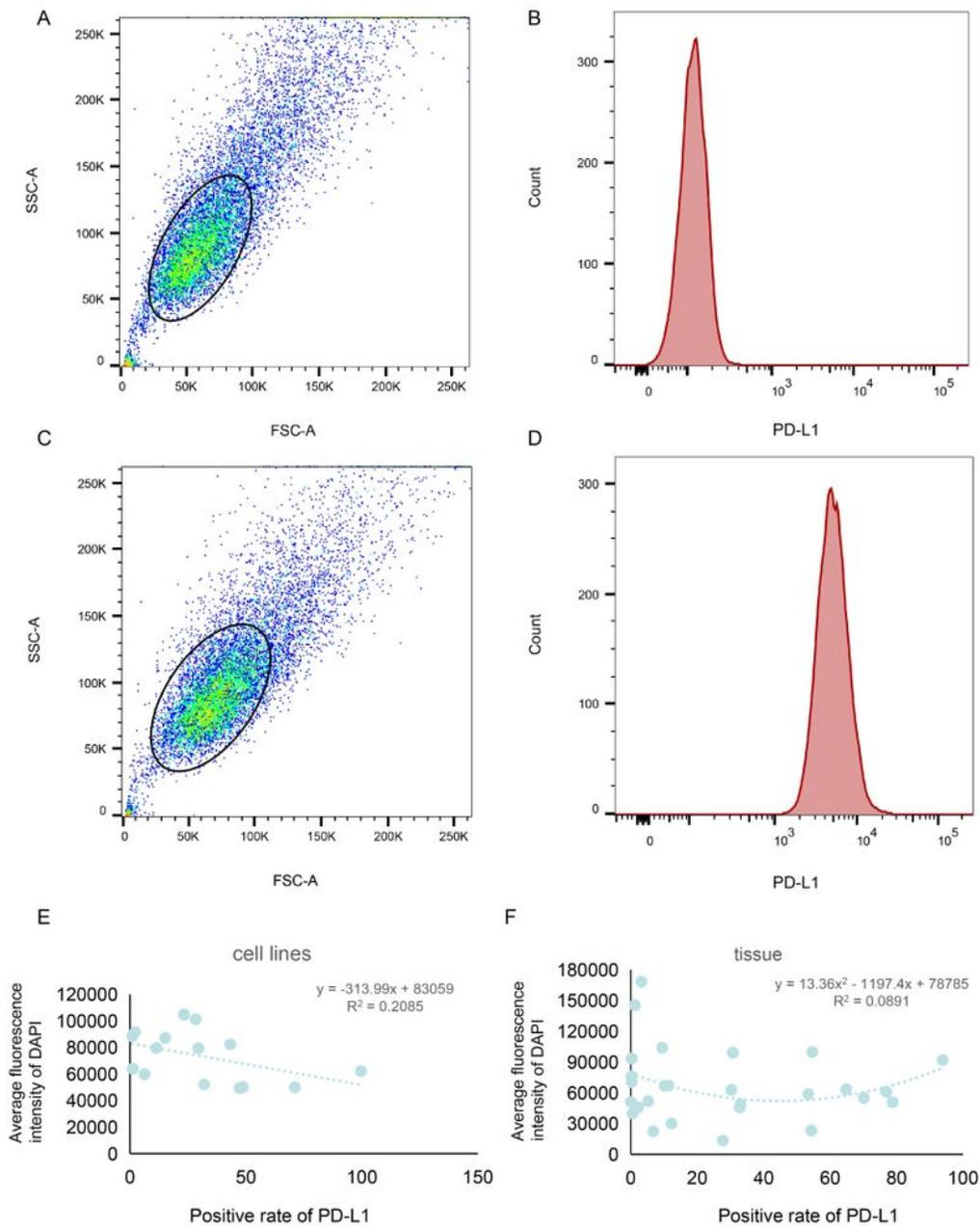


Figure 5

Verify the correlation between chromosome copy number changes and immune infiltration at the DNA content level (A-B). Representative pictures of the expression of PD-L1 in the isotype control by flow cytometry. (C-D). Representative pictures of the expression of PD-L1 in the experimental group detected by flow cytometry. (E). Scatterplot showed the relationship between DNA content and PD-L1 expression in 15 lung cancer cell lines (PC9, H157, H358, A549, H1299, H460, H1781, H1915, H2030, H827, SK-LU-1,

H226, H322, H441, H1650). The vertical axis represented the average fluorescence value of DAPI, and the horizontal axis represented the percentage of PD-L1 positive cells. (E). Scatterplot showed the relationship between DNA content and PD-L1 expression in Lung cancer tissue. The vertical axis represented the average fluorescence value of DAPI, and the horizontal axis represented the percentage of PD-L1 positive cells. Data are presented as the mean \pm SD, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ ns means no significance.

figure 6

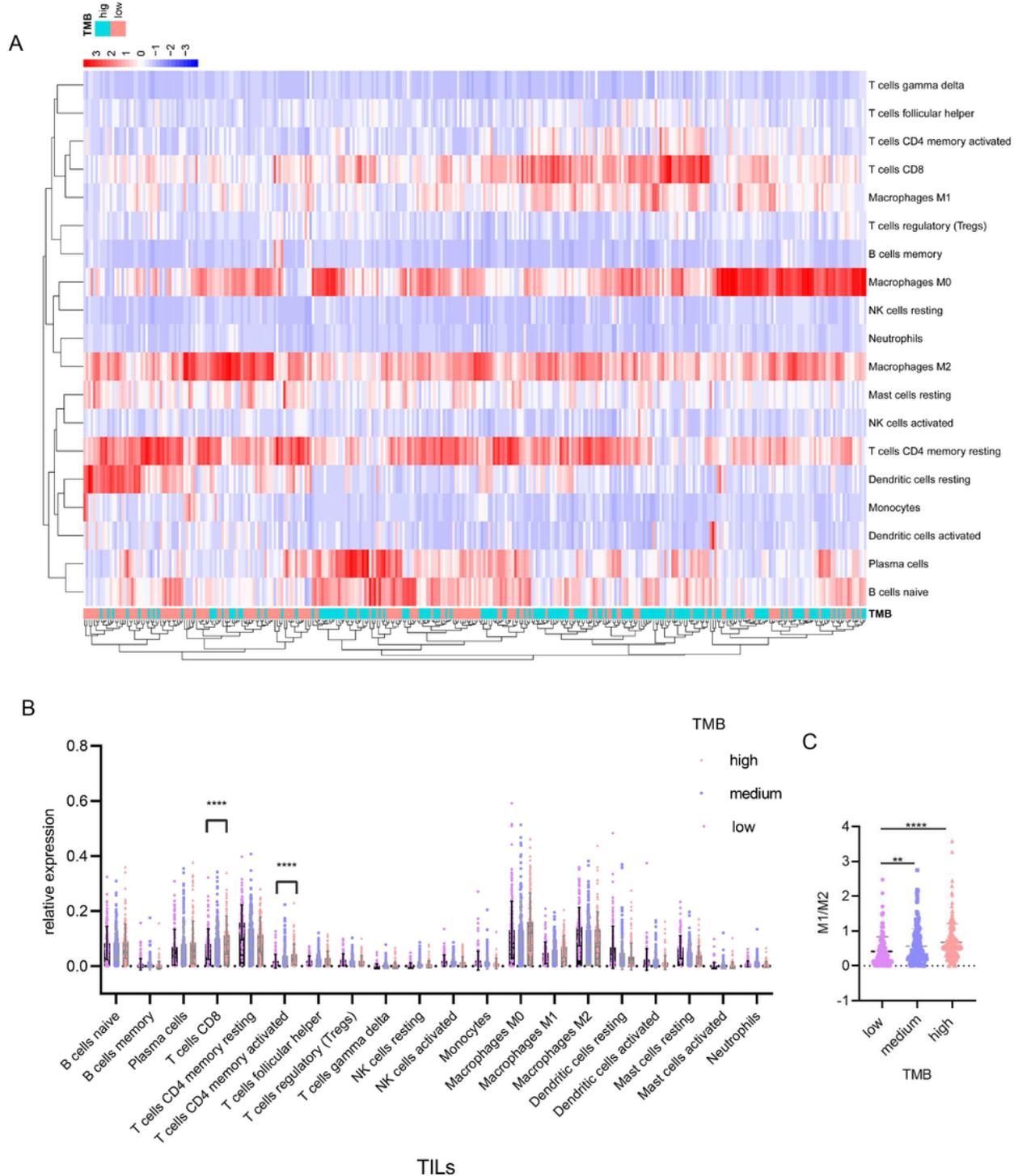


Figure 6

Relationship between TMB and immune infiltration (A). Heat map showed the infiltration of 19 immune cell subclasses in the 509 TCGA specimens in the group with high TMB and low TMB. (B). The histogram displayed the infiltration of 19 immune cells after further subdividing TMB into high, moderate and low groups. (C). Scatterplot showed the ratio of M1 and M2 macrophages at different TMB levels in the 509 specimens. Data are presented as the mean \pm SD, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. ns means no significance.

figure 7

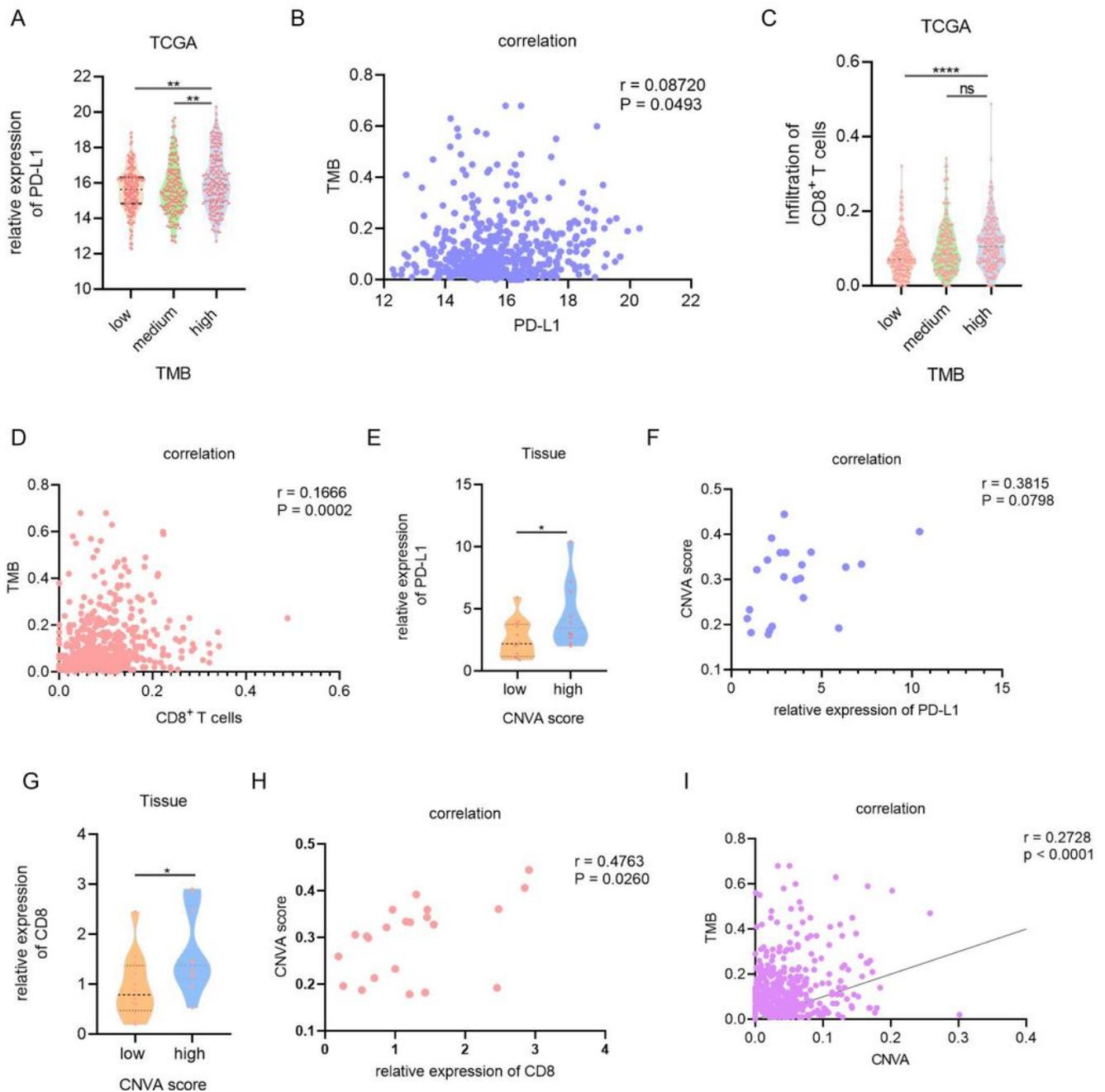


Figure 7

Compare the role of CNVA and TMB in the indication of immune efficacy and the relationship between the two (A). Violin diagram showed the expression of PD-L1 in groups of 509 TCGA specimens with different levels of CNVA. (B). Scatterplot analysis of the correlation between PD-L1 and TMB. (C). Violin diagram showed the expression of CD8 in groups of 509 TCGA specimens with different levels of CNVA. (D). Scatterplot analysis of the correlation between CD8 and TMB. (E). The expression level of PD-L1 mRNA in 22 frozen tissue samples with different CNVA levels. (F). Scatterplot showing the relationship between PD-L1 and CNVA values in 22 frozen specimens. (G). The expression level of CD8 mRNA in 22 frozen tissue samples with different CNVA levels. (H). Scatterplot showing the relationship between CD8 and CNVA values in 22 frozen specimens. (I). Correlation between the average copy number change of all genes (including introns and exons) in 507 TCGA specimens and TMB. Data are presented as the mean \pm SD, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. ns means no significance.