Tumor Mutational Burden by Whole-Genome Sequencing in Resected NSCLC of Never Smokers

Louis-Jacques Ruel[1], Zhonglin Li[1], Nathalie Gaudreault[1], Cyndi Henry[1], Victoria Saavedra Armero[1], Dominique K. Boudreau[1], Tongwu Zhang[2], Maria Teresa Landi[2], Catherine Labbé[1], Christian Couture[1], Patrice Desmeules[1], Philippe Joubert[1], and Yohan Bossé[1,3]

ABSTRACT

Background: Data are scarce about tumor mutational burden (TMB) as a biomarker in never smokers with non–small cell lung cancer (NSCLC).

Methods: TMB was assessed by whole-genome sequencing (WGS) and compared with *in silico* reduced whole-exome sequencing (WES) and targeted commercial next-generation sequencing (NGS) gene panels in 92 paired tumor-normal samples from never smokers who underwent NSCLC resection with curative intent. Analyses were performed to test for association with survival after surgery and to identify the optimal prognostic TMB cutoff.

Results: Tumors of never smokers with NSCLC had low TMB scores (median 1.57 mutations/Mb; range, 0.13–17.94). A TMB cutoff of 1.70 mutations/Mb was associated with a 5-year overall survival of 58% in the high-TMB (42% of cases) compared with 86% in low-TMB patients (Wald $P = 0.0029$). TMB scores from WGS and WES were highly correlated (Spearman $\rho = 0.93$, $P < 2.2e^{-16}$). TMB scores from NGS panels demonstrated high intraindividual fluctuations and identified high-TMB patients with 65% concordance in average compared with WGS.

Conclusions: In resected NSCLC of never smokers, high TMB was associated with worse prognosis. WES provided a good estimate of TMB while targeted NGS panels seem to lack adequate depth and resolution in the setting of low mutation burden.

Impact: TMB is a prognostic indicator of survival in resected NSCLC from individuals who never smoked. In this setting of low mutation counts, TMB can be accurately measured by WGS or WES, but not NGS panels.

Introduction

Tumor mutation burden (TMB) is a biomarker across many tumor types (1). However, its clinical utility in individuals with lung cancer who never smoked is understudied. Herein, we evaluated TMB as a potential prognostic biomarker in resected non—small cell lung cancer (NSCLC) from never smokers. Identifying patients with poor outcomes following surgery is important not only to improve prognostic stratification, but also to eventually guide adjuvant therapies. In fact, NSCLC from never smokers is characterized by a high frequency of targetable drivers (2) and thus, unlikely to respond to immunotherapy (3). In contrast, high TMB in never smokers without NSCLC oncogenic drivers may benefit from immunotherapy. Comprehensive genomic data, which are needed to derive accurate TMB estimation, may also provide useful information about coexisting genomic alterations influencing the depth and duration of response to targeted therapies (4, 5).

There are many technical considerations in TMB analysis including the sequencing panel size, the type of mutations to incorporate, and the cutoffs for the definition of high TMB. As reported previously, whole-genome sequencing (WGS) is the gold standard to estimate TMB (6). However, for logistic reasons (DNA quality and input, cost and ease of analysis), whole-exome sequencing (WES) is the standard for TMB estimation and clinical implementation is presently deemed to rely on targeted next-generation sequencing (NGS) panels below approximately 700 genes (7, 8). The underlying mutations in the calculation of TMB are also a concern. Typical TMB counts considered only mutations causing amino acid substitutions, that is, nonsynonymous mutations with potentially immunogenic aberrations (9). Whether other mutation types, such as small insertion-deletion (indels), frameshifts, splice sites, and synonymous variants, improve the performance of TMB is still under investigation (10). In terms of cutoffs associated with outcomes (survival and response to immunotherapy), it seems that the specific thresholds of high TMB varied markedly across cancer types (1, 11). To the best of our knowledge, no cutoffs have been investigated in NSCLC of never smokers, which are known tumors to be in the low stratum of TMB counts and thus likely associated with decreased immunogenicity (12). Such cutoffs may be useful in the future to identify never smokers with driver-negative NSCLC that may benefit from immunotherapy.

Here, we investigated TMB derived from paired tumor-normal WGS data from 92 never smokers that underwent surgery for NSCLC. The goals of this study were threefold: First, to test whether whole-genome TMB is a prognostic biomarker of survival after surgery in never smokers with NSCLC. Second, to define a clinically relevant TMB cutoff in never smokers with NSCLC. Finally, to evaluate the correlation and agreement of TMB assessed by WGS compared with more practical TMB scores derived from WES and NGS panels. This study addresses important technical considerations of TMB analysis in

[1]Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval, Quebec City, Canada. [2]Division of Cancer Epidemiology and Genetics, NCI, Bethesda, Maryland. [3]Department of Molecular Medicine, Laval University, Quebec City, Canada.

Corresponding Author: Yohan Bossé, Institut universitaire de cardiologie et de pneumologie de Québec, 2725 chemin Sainte-Foy, Québec, Québec G1V 4G5, Canada. Phone: 418-656-8711, ext. 3725; E-mail: yohan.bosse@criucpq.ulaval.ca

Cancer Epidemiol Biomarkers Prev 2022;31:2219–27

doi: 10.1158/1055-9965.EPI-22-0630

never smoker patients with NSCLC and assesses the clinical value of this biomarker following lung cancer resection with curative intent that might help to improve prognosis and guide adjuvant therapies.

Materials and Methods

Patients and tumor samples

All 92 never smoker patients underwent primary lung cancer surgery with curative intent between 2000 and 2018 at the *Institut universitaire de cardiologie et de pneumologie de Québec – Université Laval* (IUCPQ-UL). Last follow-ups took place in summer 2021. Blood samples as well as fresh-frozen tumor specimens were obtained at surgery. None received chemotherapy and/or radiotherapy prior to sample collection. Corresponding clinical variables including demographics, pathology report, and smoking status were collected in our local database. Staging was performed using the 8th edition of the TNM Classification of Malignant Tumours. Patients' medical charts were abstracted for follow-up starting at the time of surgery, vital status, date, and cause of death. Patients were observed until death or last follow-up. Lung tissue samples were obtained in accordance with the Institutional Review Board guidelines, which are consistent with the Declaration of Helsinki. All patients provided written informed consent to provide samples and clinical data to our local biobank, and the ethics committee of the IUCPQ-UL approved the study.

WGS data

The workflow of this study is illustrated in Supplementary Fig. S1 and S2. WGS was performed in tumor and matched "normal" germline (blood) DNA. BAM files were preprocessed following GATK best practices and sequence reads were mapped to GRCh37 using Burrows-Wheeler Aligner. Resulting CRAM files were then used for variant calling using Sentieon's genomics package (bioRxiv 115717v2). The mean sequencing coverage was 32X for normal samples and 88X for tumor samples.

Germline and somatic variant calling

Somatic variant calling was carried out using TNsnv, TNhaplotyper2 (bioRxiv 115717) and TNscope (bioRxiv 250647) for single-nucleotide variation (SNV) and TNhaplotyper2, TNscope and Strelka (13) for indels. Each tool produced its own variant calling format (VCF) files with corresponding SNVs and/or indels. We used GRCh37 human reference genome for variant mapping.

Variant filtering and annotation

Filters were applied to VCF files as described previously (14). Briefly, only variants that passed the default filters implemented in variant calling tools were retained. In addition, variant calling was considered only at genomic positions with read depth >12 in tumor samples and >6 in normal samples. Finally, the variant read count had to be >5 in tumor samples and with variant allele frequency <0.02 in normal samples. Using ANNOVAR (15), variants were retained if present in 1000 genome phase III v5, ExAC v0.3.1 and gnomAD v2.1.1 databases with a minor allele frequency (MAF) <0.001. The MAF value was chosen according to variant calling best practices (16). For both SNVs and indels, filtered variants shared by at least two mutation caller tools were used for TMB calculations to minimize type I and type II errors. Variants were then annotated using ANNOVAR with dbSNP151, 1000 genome phase III v5, ExAC v0.3.1, and gnomAD v2.1.1 databases to distinguish coding and noncoding variants. Homozygous and heterozygous genotypes to normal and tumor samples, respectively, were added to Strelka output files before using ANNOVAR. Unmapped contigs (starting with GL000) acquired from variant calling were excluded from analyses.

WGS TMB

TMB assessed by WGS (TMB_{WGS}) was determined by the total number of filtered and annotated coding and noncoding variants. We considered coding variants as nonsynonymous, frameshifts, and splicing variants. Other variant types from ANNOVAR annotation files were considered as noncoding variants. TMB_{WGS} was also converted in variants per megabase unit (mutations/Mb) by dividing the total number of variants by the size of the human reference genome. Here we applied the size of reference genome GRCh37 (17), which is estimated at 3,000 Mb.

In silico WES and NGS panels

We achieved *in silico* WES (isWES) by downsampling WGS to coding sequence regions comprised in GRCh37 reference exome available from the GATK public server. Thus, WES was computationally generated from WGS data. The resulting length of the exome was 32.8 Mb. Only coding variants that passed filters and annotation were retained. WGS was also downsampled to gene list of commercial NGS panels: FoundationOne CDx (F1CDx), Illumina TruSight Oncology 500 (TSO500), Memorial Sloan Kettering-Integrated Mutation Profiling of Actionable Cancer Targets (MSK-IMPACT), NEOplus RUO assay, Oncomine Tumor Mutation Load Assay (Oncomine TML) and QIAseq TMB (QIAseq) (Supplementary Table S1). TMB calculations for NGS panels factored in the total size of the gene coding regions instead of the total panel size, as reported in the study performed by Merino and colleagues (8).

Statistical analysis

Overall survival (OS) was calculated as the interval starting at the date of surgery to the date of death of any cause or last follow-up. The time interval for relapse-free survival (RFS) was from the day of surgery to the date of first relapse or the last follow-up. For this endpoint, patients had to come for at least one postsurgery visit to be included in the analyses. The optimal TMB cutoff was obtained using the bootstrapped Youden index with R library *cutpointr* (18). The agreement in TMB between WGS and isWES was compared with Pearson and Spearman correlations. These analyses were carried out with RStudio statistical software version 4.1.1. Kaplan–Meier curves and Cox proportional hazards analyses were performed to test the association between TMB and survival using SAS 3.8 with a macro by Jeff Meyers (19). Cox analyses were performed with and without adjusting for pathologic stage. Other plots were produced with R library *ggplot* (20). For NGS panels, the median TMB was used as the cutoff to assess their specificity, sensitivity, positive predictive value (PPV), and negative predictive value in relation with TMB_{WGS}.

Mutational signature analysis

Mutational signatures were identified using default parameters in SigProfilerExtractor (ref. 21; bioRxiv 2020.12.13.422570) v1.1.8 with GRCh37 reference genome. This derives single-base substitutions (SBS) signatures reported in the Catalogue Of Somatic Mutations In Cancer.

Driver mutation analysis

Driver mutations were annotated using OpenCRAVAT (22) with ClinVar v2021.10.01, dbSNP v154.0.2, gnomAD v2.2.0 and Cancer Gene Census v85.0.12 databases. The driver genes considered were *EGFR*, *ERBB2*, *KRAS*, *BRAF*, *NRAS*, *MAP2K1*, *MAP2K4*, *TP53*,

PIK3CA, *MET*, *ALK*, *EML4*, *ROS1*, and *RET*. Each annotated variant was manually verified to confirm their pathogenicity according to OncoKB database (23), in which case they were confirmed positive. Only drivers found in at least 1 patient were reported.

Clonal hematopoiesis variants assessment

To evaluate the presence of clonal hematopoiesis variants in tumor samples that were potentially incorporated into the TMB value, we annotated variants that passed filters using OpenCRAVAT with ClinVar v2021.10.01, dbSNP v154.0.2, gnomAD v2.2.0, and Cancer Gene Census v85.0.12 databases and searched for mutations in the following genes commonly altered by clonal hematopoiesis (24): *ASXL1*, *BRCC3*, *CBL*, *CREBBP*, *DNMT3A*, *GNAS*, *GNB1*, *JAK2*, *NRAS*, *PPM1D*, *RAD21*, *SETD2*, *SETDB1*, *SF3B1*, *SRSF2*, *TET2*, and *U2AF1*.

Data availability

The data generated in this study are available upon request from the corresponding author.

Results

Never smoker patients with lung cancer

The clinical characteristics of the 92 never smoker patients are shown in **Table 1**. All patients were self-reported White French Canadian (European ancestry), which was confirmed using WGS data. The mean age at surgery is 66 ± 10 years and 82% of patients are female. Histologic types are adenocarcinoma (n = 82; 89%), sarcomatoid carcinoma (n = 6; 7%), squamous cell carcinoma (n = 2; 2%), and adenosquamous carcinoma (n = 2; 2%). A positive history of passive tobacco smoking is self-reported in 20% of cases. The average duration of follow-up is 87 ± 55 months and death rate at 5 years is 25%. As expected, higher tumor stages are negatively associated with survival (Supplementary Fig. S3).

Somatic variant calling and TMB

Coding and noncoding variant count for all samples and by histological subtypes are shown in **Fig. 1A**. Coding variants consist

Table 1. Demographic and clinical characteristics of never smoker patients with lung cancer.

Characteristics	All patients n = 92	TMB high[a] n = 39	TMB low[a] n = 53
Age (years)	66 ± 10	65 ± 11	66 ± 9
Sex			
Female	75 (82)	27 (69)	48 (91)
Male	17 (18)	12 (31)	5 (9)
Body mass index (kg/m2)	25.7 ± 4.4	24.6 ± 3.7	26.4 ± 4.7
Passive tobacco smoking	18 (20)	7 (18)	11 (21)
Histology			
Adenocarcinoma	82 (89)	32 (82)	50 (94)
Squamous cell carcinoma	2 (2)	2 (5)	0
Adenosquamous carcinoma	2 (2)	1 (3)	1 (2)
Sarcomatoid carcinoma	6 (7)	4 (10)	2 (4)
Pathologic stage			
IA	43 (47)	13 (33)	30 (57)
IB	15 (16)	8 (21)	7 (13)
IIA	1 (1)	0	1 (2)
IIB	17 (18)	8 (21)	9 (17)
III	15 (16)	9 (23)	6 (11)
IV[b]	1 (1)	1 (3)	0
Type of surgery			
Lobectomy	69 (75)	29 (74)	40 (75)
Bilobectomy	5 (5)	3 (8)	2 (4)
Pneumonectomy	6 (7)	4 (10)	2 (4)
Wedge resection	4 (4)	2 (5)	2 (4)
Segmentectomy	8 (9)	1 (3)	7 (13)
Tumor size (cm)			
≤3	57 (62)	20 (51)	37 (70)
>3–≤5	19 (20)	7 (18)	12 (23)
>5–≤7	12 (13)	8 (21)	4 (8)
>7	4 (4)	4 (10)	0
Deaths at 5 years	23 (25)	16 (41)	7 (13)
Follow-up censored at 5 years	16 (17)	4 (10)	12 (23)
Comorbidities			
Hypertension	37 (40)	15 (38)	22 (42)
Diabetes	8 (9)	2 (5)	6 (11)
COPD	1 (1)	1 (3)	0
Asthma	6 (7)	3 (8)	3 (6)
Emphysema	1 (1)	1 (3)	0

Note: Continuous variables are presented as mean ± SD. Discrete variables are presented as n (%).
[a]TMB high and low are defined as above or below 1.70 mutations/Mb, respectively.
[b]This patient was originally classified as stage III with the 6th edition of AJCC Cancer Staging Manual, then later updated as stage IV with the 8th edition.

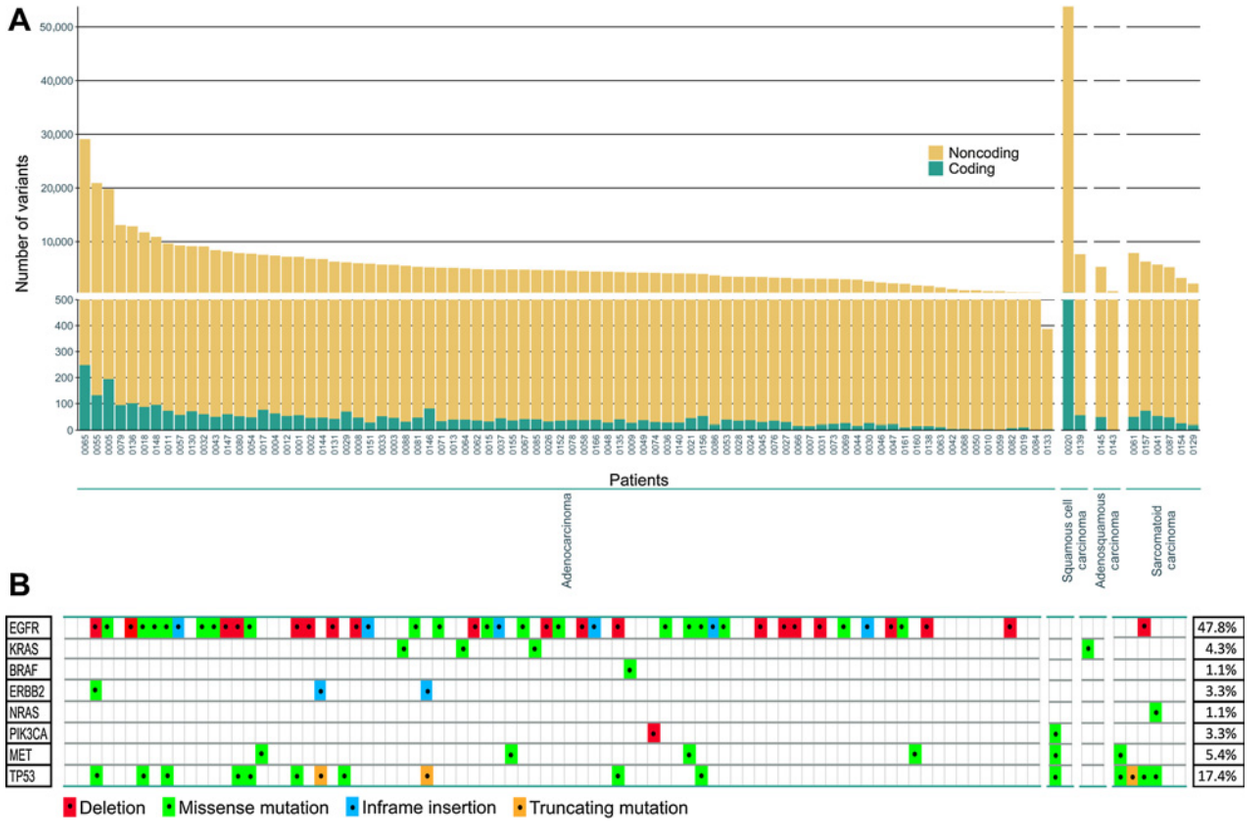

Figure 1.

A, All samples sorted by the number of variants and histologic subtypes. Bars are stacked for noncoding (yellow) and coding (green) variants. A *y*-axis break separates above and below 500 variants. Histologic types are adenocarcinoma ($n = 82$), squamous cell carcinoma ($n = 2$), adenosquamous carcinoma ($n = 2$), and sarcomatoid carcinoma ($n = 6$). **B,** Most frequent and potentially oncogenic driver mutations.

of 0.13% to 1.75% of all variants. The median TMB_{WGS} was 1.57 mutations/Mb with a range of 0.13–17.94. Supplementary Figure S4 shows the median number of SNVs and indels identified by the different variant callers. A total of 17 potential clonal hematopoiesis variants with VAF < 0.02 in normal samples were found in 15 tumor samples (16.3% of patients) used for TMB assessment. However, these mutations were not found in normal samples (VAF = 0) and their VAF in tumor samples ranged from 0.093 to 0.344, suggesting their tumor somatic origin.

TMB_{WGS} and clinicopathologic characteristics

TMB_{WGS} was not associated with histologic types, age, and sex ($P > 0.05$; Supplementary Fig. S5A and S5B). Conversely, TMB_{WGS} increases with tumor stages and sizes ($P < 0.001$; Supplementary Fig. S5C and S5D). The average TMB_{WGS} of positive and negative passive tobacco smoking history are 2.31 ± 0.50 and 1.91 ± 0.26, respectively, and show no significant difference between the two groups (Wilcoxon rank-sum *P* value = 0.315). The most extreme TMB_{WGS} outlier, patient 20, was characterized by APOBEC mutational signatures (SBS2 and SBS13), see Supplementary Fig. S6. In this cohort of never smokers, 47.8% and 4.3% of tumors were *EGFR* and *KRAS* positive, respectively (**Fig. 1B**). In the TMB_{WGS}-high group, 51% ($n = 20/39$) were *EGFR* positive compared with 45% ($n = 24/53$) in the TMB_{WGS}-low group (Wilcoxon rank-sum *P* value < 0.001). Overall, 77% ($n = 30/39$) TMB_{WGS}-high patients and 57% ($n = 30/53$) TMB_{WGS}-low patients featured a driver mutation (Wilcoxon rank-sum *P* value < 0.001).

Association between TMB_{WGS} and patient outcomes

The optimal TMB_{WGS} cutoff value using the Youden index was 1.70 ± 0.11 mutations/Mb. This cutoff delineated two groups with distinct survival (**Fig. 2**). The 5-year OS was 58% in high-TMB_{WGS} compared with 86% in low-TMB_{WGS} (Wald $P = 0.0029$). The median survival rates were 93 and 216 months in high- and low-TMB_{WGS}, respectively (**Fig. 2A**). The association between TMB and OS remained statistically significant ($P < 0.05$) following adjustment for the presence of *EGFR* mutations or other oncogenic drivers. Similarly for RFS, Kaplan–Meier curves were different for patients with high versus low TMB (Supplementary Fig. S7).

To investigate the potential confounding effect of histology, the analyses were repeated keeping only patients with adenocarcinoma ($n = 82$). In this histologic type, the Youden cutoff was also 1.70 mutations/Mb. The 5-year OS was 64% in high-TMB_{WGS} compared with 87% in low-TMB_{WGS} (Wald $P = 0.0170$; **Fig. 2B**). The median survival rate was 167 months for high-TMB_{WGS} and unreached for low-TMB_{WGS} (**Fig. 2B**).

The effect of downsampling to isWES and NGS panels

Figure 3 compares TMB measured by WGS, isWES, and NGS panels. TMB assessed by isWES is generally lower than TMB_{WGS}.

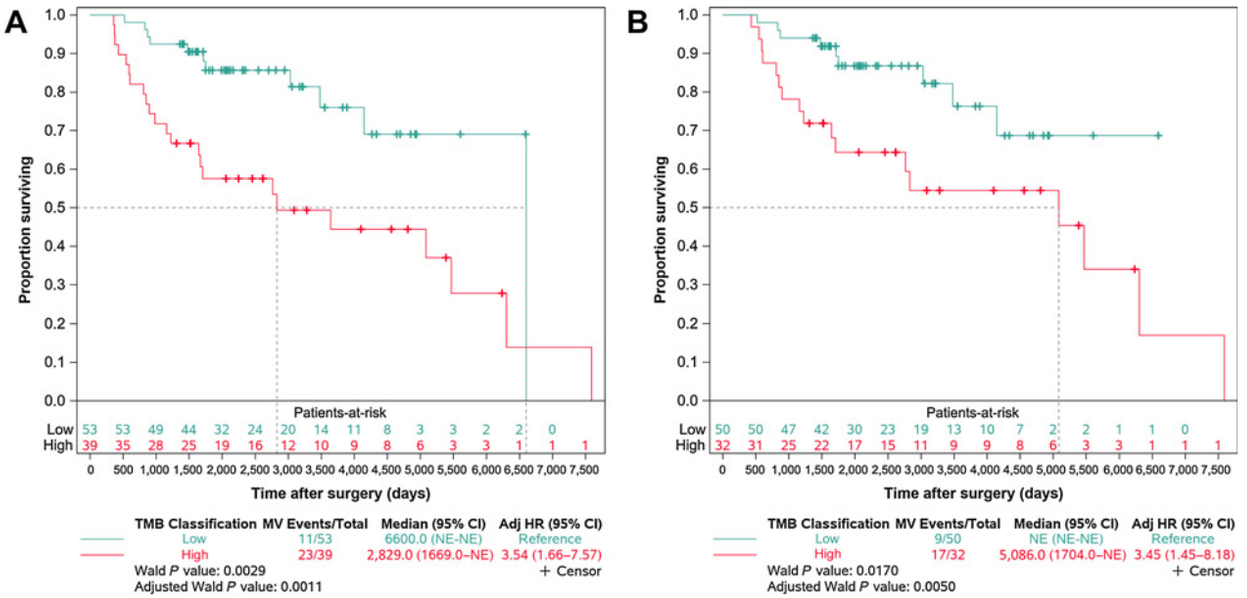

Figure 2.

A, Kaplan–Meier plot of OS for all patients; high TMB HR is 3.54, and significance is measured by Wald *P* values of 0.0011 and 0.0029 with and without adjusting for pathologic stage, respectively. **B,** Kaplan–Meier plot of OS for patients with adenocarcinoma histology; high TMB HR is 3.45, and significance is measured by Wald *P* values of 0.0050 and 0.0170 with and without adjusting for pathologic stage, respectively. High versus low TMB_{WGS} is defined as above or below 1.70 mutations/Mb. CI, confidence interval; HR, hazard ratio; OS, overall survival.

Both measures (TMB_{WGS} and TMB_{isWES}) are strongly correlated (Spearman coefficient of 0.93 and Pearson coefficient of 0.98, both with a *P* value <2.2e-16; **Fig. 4A**). This remains true after removing extreme values (**Fig. 4B**). TMB_{WGS} is on average 1.36-fold higher than TMB_{isWES}, as shown in **Fig. 3**. **Figure 4C** shows that 86 of 92 patients have a lower TMB_{isWES} than TMB_{WGS}. On the basis of regression formula presented in **Fig. 4A**, the optimal TMB cutoff by isWES corresponds to 1.20 mutations/Mb (**Fig. 5**).

In contrast to isWES, TMB scores derived from *in silico* NGS panels fluctuate considerably (**Fig. 3**). Supplementary Figure S8

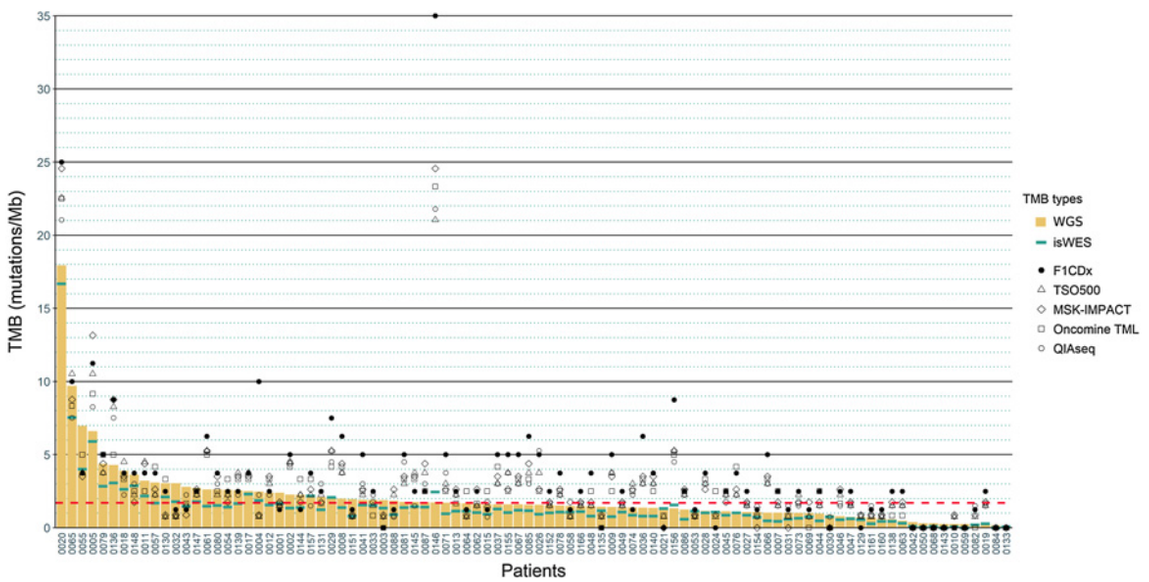

Figure 3.

isWES-based and NGS panel–based TMB estimates by tumors compared with WGS. isWES is represented by green lines and panels are represented by shapes. Patients are sorted by TMB_{WGS}. The WGS cutoff of 1.70 mutations/Mb is represented by a dashed red line. NEOplus RUO panel was not plotted on purpose as it is similar to the MSK-IMPACT Panel (Supplementary Fig. S8).

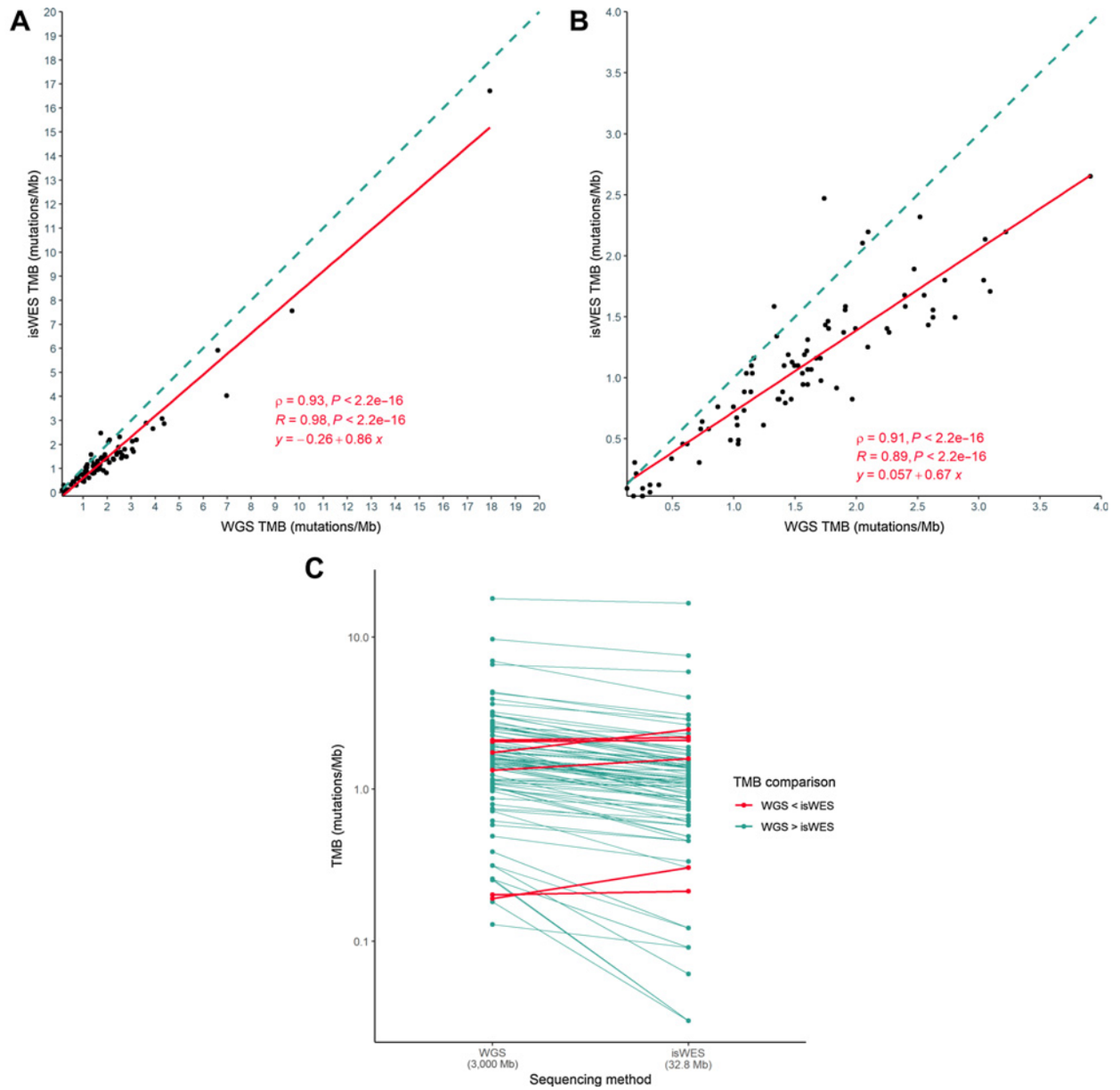

Figure 4.
Comparison between TMB measured by WGS and isWES. All patients (**A**) and after excluding extreme values (**B**). The identity line is represented by a green dashed line. **C,** Intraindividual differences in TMB measurement (base-10 log scale) between WGS and isWES.

shows TMB values by WGS compared with TMB estimated by targeted sequencing panels. All panels are characterized by larger interquartile range. The concordance to discriminate between high and low TMB was then evaluated between NGS panels compared with TMB_{WGS}. The overall sensitivity of NGS panels was relatively good (average of 80%), but specificity (average of 54%) and PPV (average of 57%) are low (Supplementary Table S2). The percentage of misclassified patients were 39% for F1CDx, 28% for TSO500, 42% for MSK-IMPACT, 40% for NEOplus, 35% for Oncomine, and 26% for QIAseq.

Discussion

The number of lung cancer surgeries performed in individuals who never smoked is likely to increase in the coming years (25). In this study, we demonstrated that TMB is a prognostic indicator of OS in this population. Gold standard measurement of TMB using WGS confirmed a relatively low TMB score in never smokers, but with substantial interindividual variability. The optimal prognostic threshold to define high- compared with low-TMB_{WGS} was found at 1.70 mutations/Mb and strongly separated two groups of never smokers with distinct OS rate at 5-year after surgery (58% vs. 86%,

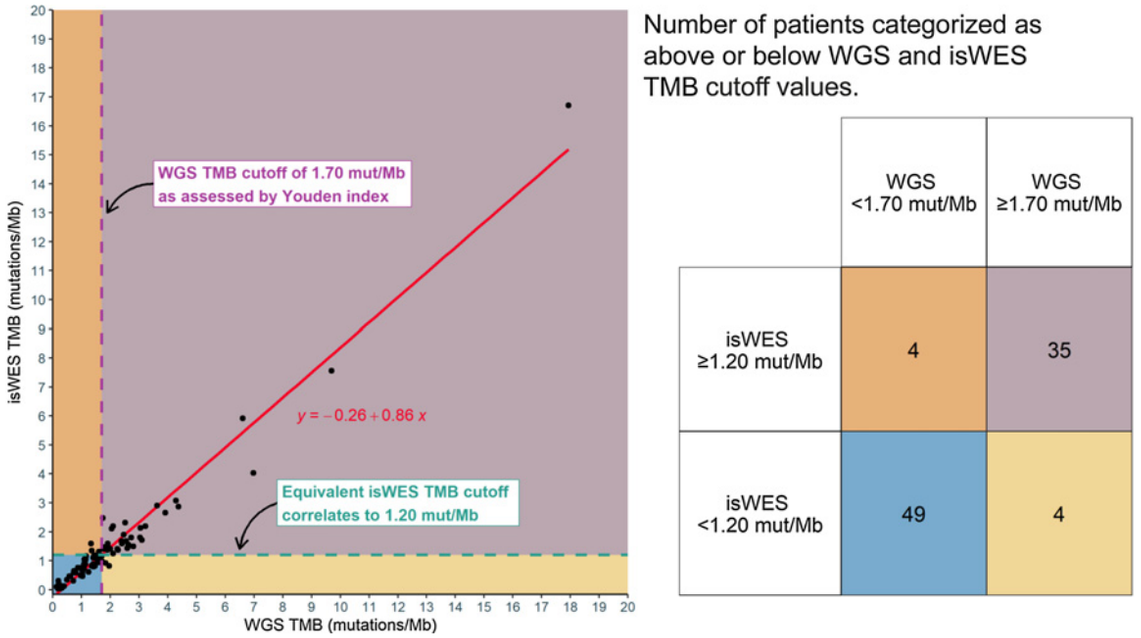

Figure 5.

Agreement in detecting never smokers with high TMB. A simple linear regression ($y \sim x$) is shown in red. TMB_{WGS} (1.70 mutations/Mb) and TMB_{isWES} (1.20 mutations/Mb) cutoffs are shown in purple and green, respectively. The TMB_{isWES} cutoff corresponds to the intersection between the linear regression and the TMB_{WGS} cutoff.

respectively). Importantly, in this setting of low mutation counts, we demonstrated the transferability of using WES, reported as isWES, to obtain TMB approximation comparable with the ground truth detected by WGS. In general, TMB by isWES was underestimated and we thus estimated that the corresponding cutoff to distinguish patients with high and low TMB using WES should be 1.20 mutations/Mb. In contrast, NGS panels were not reliable for TMB estimation in never smokers. The misclassification of low versus high TMB obtained using NGS panels were too important to recommend clinical usage, even for NGS panel spanning more than 400 genes or with sequencing panel size >1 Mb.

TMB is a biomarker of immunotherapy but is not recognized as a prognostic biomarker. In pan-cancer analysis of 3,014 immune checkpoint inhibitor (ICI)-naïve patients where TMB was estimated from a 468 gene panel (MSK-IMPACT), high TMB defined by the highest 20% in each histology was not associated with improved OS (log-rank $P = 0.11$; ref. 11). In the subset of non–ICI-treated NSCLC, the top 20% of TMB ($n = 125$) had a survival curve that overlapped with the bottom 80% of TMB ($n = 498$; log-rank $P = 0.5$). This suggests that TMB offered no prognostic value for all comers with NSCLC. In contrast, in this lower spectrum of TMB values of never smokers, we found a strong association of high TMB with worse prognosis. From a biological standpoint, more DNA damage caused by mutations is expected to decrease survival. On the other hand, mutated proteins may provide additional opportunities for the immune system to recognize and kill tumor cells. The ultimate survival outcome is thus highly dependent on the interplay between immune cells and tumor, and it may thus be difficult to predict prognosis using TMB alone. In this perspective, TMB as a complementary biomarker to PD-L1 testing is a promising avenue.

In driver-negative NSCLC, there is a correlation between high TMB and clinical benefit in response to ICI (9). In advanced NSCLC with a positive smoking history in more than 90% of patients, high TMB was established at 10 mutations/Mb (26). In the current study, only 1 of 92 (1%) patients had TMB above this threshold. This compares with breast cancer where such hypermutated tumors occur in 5% of cases (27). In fact, median TMB, the initially proposed method of cancer mutational burden classification (9), varies widely across cancer types (1) and the median TMB in this study of lung cancer in never smokers (1.57 mutations/MB) is similar to the range observed in other solid tumors such as breast (2.6 mut/Mb; ref. 27), prostate (76 mutations with WES, ∼2 mut/Mb; ref. 28), and renal cell carcinomas (1.76 mut/Mb; ref. 29).

We demonstrated that in this lower spectrum of TMB values, a much lower threshold can improve prognostic stratification following curative intent resection. This raises the possibility that the 10 mutations/Mb cutoff used for ICI efficacy may need to be adapted in this setting. The optimal cutoff associated with improved survival after ICI treatment is also known to vary substantially across tumor types (11), but to the best of our knowledge a threshold as low as 1 mutation/Mb has never been reported. Obviously, other studies are needed to demonstrate whether TMB is a biomarker of immunotherapy response in never smokers as well as to evaluate whether the cutoff associated with prognosis is the same or different than the cutoff potentially associated with ICI efficacy. Nevertheless, TMB in our study identified a group of never smokers that may benefit from adjuvant therapy. As TMB is a biomarker of immunotherapy, this class of drugs seems like of reasonable choice of treatment in this subgroup of patients if no targetable mutation is identified. Our study is thus opening the door for considering ICI clinical trials in never smokers with high TMB defined by a cutoff point adapted for this cancer type.

Previous studies have compared TMB values across tumor profiling approaches ranging from small NGS panels to WES. Using The Cancer Genome Atlas (TCGA) dataset, Endris and colleagues (30) compared the performance of three gene panels (Oncomine Comprehensive v3, TruSight Tumor 170, Oncomine Tumor Mutational Load) and observed a strong correlation between panel-based TMB estimation compared with WES, especially for sequencing panel size of >1 Mb. Using the same pan-cancer dataset, Buchhalter and colleagues (31) concluded that panels of size between 1.5 and 3 Mb are optimal to estimate TMB. Heeke and colleagues (32) compared TMB obtained from three targeted sequencing panels (FoundationOne, Oncomine, QIAseq) in 30 largely ever smoker patients with NSCLC. Strong correlations in TMB values were observed between pairs of panels ($R2 > 0.79$). However, this was substantially reduced in subsets of tumors with TMB values ranging from 5 to 25 mutations/Mb ($R2 < 0.25$), which is within the range used to separate low and high TMB. Poor agreement between pairs of panels was also observed using the Bland–Altman method. The latter study highlighted discordance across panels, especially for low TMB values, and that simple correlation is a poor metric to compare panels. Similarly, Wu and colleagues (33) demonstrated in TCGA that accuracy, defined by correctly identifying either high or low TMB, outperformed correlation in assessing the performance of panel-based TMB estimation. The same study also showed that NGS panels TMB estimations are less likely to be reliable in cancer types with intermediate to low TMB levels. A study estimating TMB in circulating tumor DNA (blood TMB) has suggested that a panel of 150 genes was sufficient in patients with NSCLC (34). However, the minimum number of genes needed to estimate TMB is known to vary by cancer types and is negatively correlated with the median TMB, that is, cancer types with lower TMB levels (e.g., lung cancer in never smokers) require panels with more genes (33).

Our study is the first to use WGS-based TMB to assess the performance of TMB approximation derived from WES and NGS panels. Like ongoing TMB harmonization efforts (7, 8), we used tumor sequencing data and downsampled the mutation calls to the exome and gene list of commercial NGS panels to compare the tumor profiling approaches. We confirmed strong correlation and accuracy of WES compared with WGS, suggesting that WES is sufficient in clinical practice for NSCLC in never smokers. In contrast, and as expected for tumors at the lower spectrum of TMB counts, NGS panels were not reliable to estimate TMB levels. In addition, the proportion of patients misclassified as high and low TMB vary substantially across NGS panels. Taken together, we propose WES for accurate TMB estimation of lung cancer in never smokers, but not NGS panels. Further research is needed to identify a panel size threshold for accurate classification of high and low TMB in NSCLC of never smokers.

This study has limitations. First, it was conducted in a research setting with unfixed fresh-frozen tissues available and where TMB measurements were successful in all patients. This may not reflect samples encountered in clinical practice where a TMB failure rate is expected and where WGS/WES testing is usually not available. Second, WES and NGS panels were based on *in silico* and not experimental data. Although we used high-coverage WGS in this study, WES and NGS assays offer superior coverage and sensitivity for mutation calling in targeted genes. Third, the main purpose of TMB is to predict response to ICI. As standard of care, we were unable to predict the response to ICI in our patients. Fourth, the performance of TMB as a prognostic biomarker was not compared with PD-L1 protein expression by IHC or other genomic biomarkers. We have not investigated blood TMB (bTMB) in the same patients, but this will be explored as part of our research plan. In advanced NSCLC, bTMB is emerging as a biomarker of response to immunotherapy (34, 35). In early-stage NSCLC, with lower ctDNA shedding into the bloodstream, bTMB estimations were found less consistent compared with TMB from tissue (36). On the other hand, for patients with complete surgical resection, as patients treated in the current study, ctDNA analysis has been shown to detect molecular residual disease and predict recurrence (37). Accordingly, with further methodology improvements to enhance sensitivity of ctDNA analysis, bTMB may become an attractive alternative in postsurgery setting, especially for longitudinal disease monitoring. Finally, there are still a lot of methodology issues to overcome to generate a reproducible TMB score as well as uncertainties in its clinical value in lung cancer management (6, 7).

In conclusion, we demonstrated a prognostic impact of TMB in never smokers that underwent lung cancer surgery. High TMB set at a cutoff of 1.70 mutations/Mb in this population was associated with worse survival. TMB estimations from WES, but not NGS panels, provided good concordance with WGS. Whether TMB can be a predictive biomarker of response to checkpoint blockade in never smokers deserves further investigations.

Authors' Disclosures

C. Labbé reports personal fees from Amgen, AstraZeneca, Bristol-Myers Squibb, Jazz Pharmaceuticals, LEO Pharma, Merck, Pfizer, Roche, and Sanofi Genzyme outside the submitted work. P. Desmeules reports grants from AstraZeneca, Pfizer, Roche, Eli-Lilly, EMD Serrono, Novartis, and Amgen outside the submitted work. No disclosures were reported by the other authors.

Authors' Contributions

L.-J. Ruel: Formal analysis, visualization, methodology, writing–original draft. **Z. Li:** Formal analysis. **N. Gaudreault:** Data curation, project administration, writing–review and editing. **C. Henry:** Data curation, project administration, writing–review and editing. **V. Saavedra Armero:** Data curation, project administration, writing–review and editing. **D.K. Boudreau:** Data curation, project administration, writing–review and editing. **T. Zhang:** Resources, data curation. **M.T. Landi:** Resources, data curation. **C. Labbé:** Data curation, writing–review and editing. **C. Couture:** Data curation, writing–review and editing. **P. Desmeules:** Data curation, writing–review and editing. **P. Joubert:** Data curation, writing–review and editing. **Y. Bossé:** Conceptualization, supervision, funding acquisition, methodology, writing–original draft.

Acknowledgments

This work was supported by the Fondation de l'Institut universitaire de cardiologie et de pneumologie de Québec (FIUCPQ) to Y. Bossé and P. Joubert owing to a generous donation from Mr. Normand Lord, President & CEO, St-Augustin Canada Electric. L-J. Ruel is the recipient of a master's degree scholarship from the *Fonds de recherche Québec - Santé* (FRQS). P. Joubert is the recipient of a Junior 1 Clinical Research Scholar award from the *Fonds de recherche Québec - Santé* (FRQS). Y. Bossé holds a Canada Research Chair in Genomics of Heart and Lung Diseases.

The publication costs of this article were defrayed in part by the payment of publication fees. Therefore, and solely to indicate this fact, this article is hereby marked "advertisement" in accordance with 18 USC section 1734.

Note

Supplementary data for this article are available at Cancer Epidemiology, Biomarkers & Prevention Online (http://cebp.aacrjournals.org/).

Received May 30, 2022; revised August 9, 2022; accepted September 19, 2022; published first September 20, 2022.

References

1. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. N Engl J Med 2017;377:2500–1.
2. Devarakonda S, Li Y, Martins Rodrigues F, Sankararaman S, Kadara H, Goparaju C, et al. Genomic profiling of lung adenocarcinoma in never-smokers. J Clin Oncol 2021;39:3747–58.
3. Mazieres J, Drilon A, Lusque A, Mhanna L, Cortot AB, Mezquita L, et al. Immune checkpoint inhibitors for patients with advanced lung cancer and oncogenic driver alterations: results from the IMMUNOTARGET registry. Ann Oncol 2019;30:1321–8.
4. Planchard D, Besse B, Groen HJM, Hashemi SMS, Mazieres J, Kim TM, et al. Phase 2 study of dabrafenib plus trametinib in patients with BRAF V600E-mutant metastatic NSCLC: updated 5-year survival rates and genomic analysis. J Thorac Oncol 2022;17:103–15.
5. Skoulidis F, Heymach JV. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. Nat Rev Cancer 2019;19:495–509.
6. Sholl LM, Hirsch FR, Hwang D, Botling J, Lopez-Rios F, Bubendorf L, et al. The promises and challenges of tumor mutation burden as an immunotherapy biomarker: a perspective from the international association for the study of lung cancer pathology committee. J Thorac Oncol 2020;15:1409–24.
7. Stenzinger A, Endris V, Budczies J, Merkelbach-Bruse S, Kazdal D, Dietmaier W, et al. Harmonization and standardization of panel-based tumor mutational burden measurement: real-world results and recommendations of the quality in pathology study. J Thorac Oncol 2020;15:1177–89.
8. Merino DM, McShane LM, Fabrizio D, Funari V, Chen SJ, White JR, et al. Establishing guidelines to harmonize tumor mutational burden (TMB): in silico assessment of variation in TMB quantification across diagnostic platforms: phase I of the friends of cancer research TMB harmonization project. J Immunother Cancer 2020;8:e000147.
9. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. Science 2015;348:124–8.
10. Hellmann MD, Nathanson T, Rizvi H, Creelan BC, Sanchez-Vega F, Ahuja A, et al. Genomic features of response to combination immunotherapy in patients with advanced non-small-cell lung cancer. Cancer Cell 2018;33:843–52.
11. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019;51:202–6.
12. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. Cell 2012;150:1121–34.
13. Kim S, Scheffler K, Halpern AL, Bekritsky MA, Noh E, Källberg M, et al. Strelka2: fast and accurate calling of germline and somatic variants. Nat Methods 2018;15:591–4.
14. Zhang T, Joubert P, Ansari-Pour N, Zhao W, Hoang PH, Lokanga R, et al. Genomic and evolutionary classification of lung cancer in never smokers. Nat Genet 2021;53:1348–59.
15. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.
16. Koboldt DC. Best practices for variant calling in clinical sequencing. Genome Med 2020;12:91.
17. Church DM, Schneider VA, Graves T, Auger K, Cunningham F, Bouk N, et al. Modernizing reference genome assemblies. PLoS Biol 2011;9:e1001091.
18. Thiele C, Hirschfeld G. cutpointr: improved estimation and validation of optimal cutpoints in R. J Stat Softw 2021;98:1–27.
19. Meyers J. Kaplan-Meier Survival Plotting Macro %Newsurv. SAS Support Communities; 2021. Available from: https://communities.sas.com/t5/SAS-Communities-Library/Kaplan-Meier-Survival-Plotting-Macro-NEWSURV/tap/479747.
20. Wickham H. ggplot2: elegant graphics for data analysis. Springer International Publishing; 2016.
21. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013;500:415–21.
22. Pagel KA, Kim R, Moad K, Busby B, Zheng L, Tokheim C, et al. Integrated informatics analysis of cancer-related variants. JCO Clin Cancer Inform 2020;4:310–7.
23. Chakravarty D, Gao J, Phillips S, Kundra R, Zhang H, Wang J, et al. OncoKB: a precision oncology knowledge base. JCO Precis Oncol 2017;2017:PO.17.00011.
24. Steensma DP. Clinical consequences of clonal hematopoiesis of indeterminate potential. Blood Adv 2018;2:3404–10.
25. Jeon J, Holford TR, Levy DT, Feuer EJ, Cao P, Tam J, et al. Smoking and lung cancer mortality in the United States from 2015 to 2065: a comparative modeling approach. Ann Intern Med 2018;169:684–93.
26. Hellmann MD, Ciuleanu TE, Pluzanski A, Lee JS, Otterson GA, Audigier-Valette C, et al. Nivolumab plus ipilimumab in lung cancer with a high tumor mutational burden. N Engl J Med 2018;378:2093–104.
27. Barroso-Sousa R, Jain E, Cohen O, Kim D, Buendia-Buendia J, Winer E, et al. Prevalence and mutational determinants of high tumor mutation burden in breast cancer. Ann Oncol 2020;31:387–94.
28. Subudhi SK, Vence L, Zhao H, Blando J, Yadav SS, Xiong Q, et al. Neoantigen responses, immune correlates, and favorable outcomes after ipilimumab treatment of patients with prostate cancer. Sci Transl Med 2020;12:eaaz3577.
29. Lee M, Samstein RM, Valero C, Chan TA, Morris LGT. Tumor mutational burden as a predictive biomarker for checkpoint inhibitor immunotherapy. Hum Vaccin Immunother 2020;16:112–5.
30. Endris V, Buchhalter I, Allgäuer M, Rempel E, Lier A, Volckmar AL, et al. Measurement of tumor mutational burden (TMB) in routine molecular diagnostics: in silico and real-life analysis of three larger gene panels. Int J Cancer 2019;144:2303–12.
31. Buchhalter I, Rempel E, Endris V, Allgäuer M, Neumann O, Volckmar AL, et al. Size matters: dissecting key parameters for panel-based tumor mutational burden analysis. Int J Cancer 2019;144:848–58.
32. Heeke S, Benzaquen J, Hofman V, Long-Mira E, Lespinet V, Bordone O, et al. Comparison of three sequencing panels used for the assessment of tumor mutational burden in NSCLC reveals low comparability. J Thorac Oncol 2020;15:1535–40.
33. Wu HX, Wang ZX, Zhao Q, Wang F, Xu RH. Designing gene panels for tumor mutational burden estimation: the need to shift from 'correlation' to 'accuracy'. J Immunother Cancer 2019;7:206.
34. Wang Z, Duan J, Cai S, Han M, Dong H, Zhao J, et al. Assessment of blood tumor mutational burden as a potential biomarker for immunotherapy in patients with non-small cell lung cancer with use of a next-generation sequencing cancer gene panel. JAMA Oncol 2019;5:696–702.
35. Gandara DR, Paul SM, Kowanetz M, Schleifman E, Zou W, Li Y, et al. Blood-based tumor mutational burden as a predictor of clinical benefit in non-small-cell lung cancer patients treated with atezolizumab. Nat Med 2018;24:1441–8.
36. Zhang Y, Chang L, Yang Y, Fang W, Guan Y, Wu A, et al. The correlations of tumor mutational burden among single-region tissue, multi-region tissues and blood in non-small cell lung cancer. J Immunother Cancer 2019;7:98.
37. Chaudhuri AA, Chabon JJ, Lovejoy AF, Newman AM, Stehr H, Azad TD, et al. Early detection of molecular residual disease in localized lung cancer by circulating tumor DNA profiling. Cancer Discov 2017;7:1394–403.