Mutation burden-orthogonal tumor genomic subtypes delineate responses to immune checkpoint therapy

Shiro Takamatsu ,[1] Junzo Hamanishi ,[1] J B Brown,[2,3] Ken Yamaguchi,[1] Koji Yamanoi,[1] Kosuke Murakami,[4] Osamu Gotoh ,[5] Seiichi Mori,[5] Masaki Mandai,[1] Noriomi Matsumura [4]

Correspondence to
Dr Noriomi Matsumura;
noriomi@med.kindai.ac.jp

ABSTRACT

Background In cancer therapy, higher-resolution tumor-agnostic biomarkers that predict response to immune checkpoint inhibitor (ICI) therapy are needed. Mutation signatures reflect underlying oncogenic processes that can affect tumor immunogenicity, and thus potentially delineate ICI treatment response among tumor types.

Methods Based on mutational signature analysis, we developed a stratification for all solid tumors in The Cancer Genome Atlas (TCGA). Subsequently, we developed a new software (Genomic Subtyping and Predictive Response Analysis for Cancer Tumor ICi Efficacy, GS-PRACTICE) to classify new tumors submitted to whole-exome sequencing. Using existing data from 973 pan-cancer ICI-treated cases with outcomes, we evaluated the subtype-response predictive performance.

Results Systematic analysis on TCGA samples identified eight tumor genomic subtypes, which were characterized by features represented by smoking exposure, ultraviolet light exposure, APOBEC enzyme activity, *POLE* mutation, mismatch repair deficiency, homologous recombination deficiency, genomic stability, and aging. The former five subtypes were presumed to form an immune-responsive group acting as candidates for ICI therapy because of their high expression of immune-related genes and enrichment in cancer types with FDA approval for ICI monotherapy. In the validation cohort, the samples assigned by GS-PRACTICE to the immune-reactive subtypes were significantly associated with ICI response independent of cancer type and TMB high or low status.

Conclusions The new tumor subtyping method can serve as a tumor-agnostic biomarker for ICI response prediction and will improve decision making in cancer treatment.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Mismatch repair deficiency (MMRd) and high tumor mutational burden (TMB-high) are proposed as tumor-agnostic predictive biomarkers for immune checkpoint inhibitors (ICIs), but their frequencies vary among tumor types.

⇒ In a limited number of cancer types, including non-small-cell lung cancer and melanoma, mutagenic processes other than MMRd and the mutation signatures reflecting such processes have been reported to be associated with ICI sensitivity.

WHAT THIS STUDY ADDS

⇒ From the systematic analysis of mutational signatures in all solid tumors of The Cancer Genome Atlas, we developed a new method to classify whole-exome sequenced tumors into eight genomic subtypes with different immunogenicity.

⇒ In validation data including multiple cancer types, the classified tumor subtypes significantly correlated with ICI efficacy, independent of cancer type and TMB status.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ Our method provides a new pan-cancer biomarker for predicting ICI efficacy orthogonal to TMB status.

⇒ The results suggest that the mutational processes underlying carcinogenesis strongly affect tumor immunogenicity, leading to differences in ICI treatment response among tumor types.

INTRODUCTION

The advent of immune checkpoint inhibitors (ICIs) has provided substantial opportunities in cancer treatment. However, the proportion of patients who benefit from ICIs varies widely by cancer type,[1] and tumor-agnostic biomarkers to identify (un)responsive subsets are strongly desired. A recently established predictive biomarker is the loss of mismatch repair protein in immunohistochemistry or microsatellite instability (MSI-high), which indicates mismatch repair deficiency (MMRd) status.[2] MMRd tumors are considered to be highly sensitive to ICI because they carry a large number of tumor-specific neoantigens.[3] Another tumor agnostic biomarker recently approved by the Food and Drug Administration (FDA) is tumor mutational burden (TMB)-high status, where tumors have 10 or more mutations per megabase calculated from the FoundationOne CDx assay.[4] Despite the approval of TMB as a biomarker, there exist a sufficient number of cases that have modest TMB but respond to ICI,[5,6] and more

sophisticated methods for identifying such tumors need to be developed.

Comprehensive gene mutation analysis in cancer enabled by high-throughput next-generation sequencing has revealed that even neutral somatic mutations, previously thought to be 'passenger' mutations, exhibit reproducible patterns of change, or mutational signatures, depending on the underlying endogenous and exogenous mutagenic processes.[7 8] Certain mutational signatures are known to be associated with tumor immunogenicity,[9–11] suggesting that differences in the background mutational processes may play an important role in antitumor immunity.

To advance oncology patient care by leveraging the signature-immunogenicity relationship, we report the development of a computational framework to classify tumors beyond their tissue origin. The tool is subsequently challenged to predict response to ICI independent of cancer type and TMB status using a large external patient dataset, demonstrating its feasibility and position to complement FDA-approved TMB analyses.

MATERIALS AND METHODS
The Cancer Genome Atlas data
Clinical information of all tumors except diffuse large B-cell lymphoma, acute myeloid leukemia, and thymoma in The Cancer Genome Atlas (TCGA) studies was obtained from the cBioPortal (https://www.cbioportal.org/) and the broad GDAC websites (https://gdac.broadinstitute.org/). Among these, 9794 cases, whose somatic mutation profiles analyzed by Mutect2[12] were available on the GDC portal (https://portal.gdc.cancer.gov/), were included in this study. We also obtained the other somatic mutation profiles calculated by the three different variant callers (see the Methods section) and gene expression profiles from a previous report.[13] The annotations of germline mutations and gene promoter methylations were obtained from previous reports.[14 15] The contribution values to COSMIC (V.2) 30 mutational signatures (https://cancer.sanger.ac.uk/signatures/signatures_v2) of each sample were calculated using MutationalPatterns.[16] The annotation of cancer types with FDA approval for ICI monotherapy was based on a previous report.[17] The response rates for ICI monotherapy for each tumor type were obtained from previous reports,[1 18 19]

Validation datasets
Pan-Cancer Analysis of Whole Genomes consortium (PCAWG), Clinical Proteomic Tumor Analysis Consortium (CPTAC), National Bioscience Database Center (NBDC), and cBioPortal datasets were obtained from their databases (online supplemental table S2). For the ICI-treated cohorts, samples collected from metastatic tumors and those with a history of ICI treatment at sample collection were excluded. A total of 973 patients from 13 datasets were included in the analysis (online supplemental table S3 and figure S15).

Statistical analyses
Statistical analyses were mainly performed in Python (V.3.7.4); the Mann-Whitney U test, χ^2 test, and Spearman's rank correlation coefficient test were performed using SciPy (V.1.6.1), survival analyses including Kaplan-Meier curve, log-rank test, and Cox proportional hazard regression using Lifelines (V.0.25.10) and StatsModels (V.0.12.2), machine learning analyses using Scikit-learn (V.0.24.1). The Venn diagram, the Jonckheere-Terpstra test, and the Passing-Bablok regression analysis were performed using 'VennDiagram' (V.1.6.20), 'clinfun' (V.1.0.15), and 'mcr' (V.1.2.2) packages in R. We considered a p<0.05 as being statistically significant.

Details are provided in online supplemental data.

RESULTS
Identification of eight genomic subtypes based on mutational signature analysis
Based on Mutect2-derived[12] mutation annotations from whole-exome sequencing (WES) data, score profiles of COSMIC (V.2) mutational signatures were derived for each solid tumor in TCGA (n=9794). Eight tumor groups were obtained after clustering logarithm-transformed profiles (figure 1A). Based on the enrichment of signatures with proposed etiologies (online supplemental table S1), seven of these subtypes were labeled as groups associated with smoking (SMK), ultraviolet light (UVL), APOBEC (APB), DNA polymerase epsilon deficiency (POL), mismatch repair deficiency (MRD), homologous recombination deficiency (HRD), and aging (AGE). The remaining group that showed no specific accumulation of mutation signatures and the lowest number of mutations was assigned the genomic stability (GNS) subtype.

In terms of clinical information, age, gender, stage, and mortality differed considerably among the subtypes (figure 1A, online supplemental figure S1A). The proportion of patients with SMK history was highest in the SMK group. Molecularly characterized groups also contained enriched annotations, including high *POLE* mutations in the POL group, as well as MMR mutations, *MLH1* methylation, and MSI high status in the MRD group. The HRD group contained characteristic *BRCA* alterations.[20] The distribution of genomic subtypes differed among tumor types. (figure 1B, online supplemental figure S1B,C). Extensive analytics of each subtype are provided in online supplemental figures S1D and S2–S8.

Transcriptomes of genes associated with tumor immune response were assessed. Genes representing the infiltration of cytotoxic CD8 +T cells (*CD8A, GZMB,* and *IFNG*) and genes related to ICI response (*CXCL9* and *CXCL13*)[21] were upregulated in the five subtypes (SMK, UVL, APB, POL, MRD) relative to the others (HRD, GNS, AGE). The CYT score[22] and GEP score[23] related to ICI response were also higher in the same five subtypes. Postsubtyping also demonstrates that the five subtypes were more frequently of tumor origin with FDA approval for ICI monotherapy (figure 1A). Further, when the proportion of samples
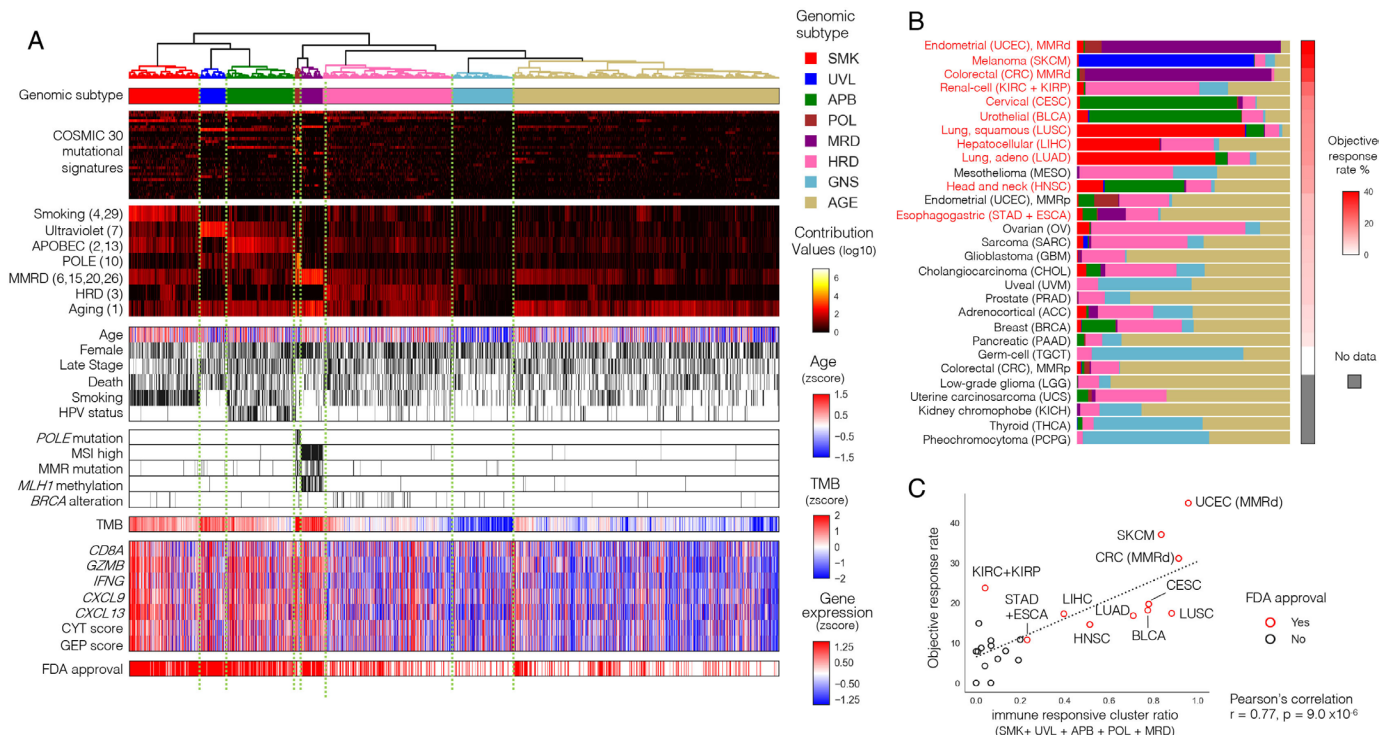
Figure 1 Identification of tumor genomic subtypes associated with immune responses. (A) Unsupervised hierarchical clustering based on mutational signatures divided TCGA solid tumors (N=9794) into eight distinctive subtypes, five of which showed strong tumor immune responses. The first panel shows the genomic subtypes in color. The second panel shows COSMIC 30 mutational signature contribution values, and the third panel highlights the seven known etiology-related ones, with COSMIC mutational signature numbers in parentheses. The fourth panel shows clinical information (age, sex, stage, death, smoking habits, HPV infection), and the fifth panel shows DNA repair-related gene alterations, including *POLE* mutations, MSI high, MMR mutations, *MLH1* methylation, and *BRCA* alterations are shown in this order. The sixth panel shows TMB, and the seventh panel shows immune-related gene expression (*CD8A, GZMB, IFNG, CXCL9,* and *CXCL13*) and scores (CYT and GEP). The bottom panel shows the cancer types for which ICI therapy as a single agent is approved by the FDA. These results are summarized in online supplemental figure S1A,B. (B) The distribution of the subtypes (shown with the same colors as (A)) and reported objective response rate to ICI monotherapy per tumor type. Tumor types are arranged in descending order of the response rate, with red letters indicating those with FDA approval for ICI monotherapy. (C) The proportion of the five subtypes with high immune-responsive scores (SMK, UVL, APB, POL, and MRD) per tumor type showed a significant positive correlation with the reported objective response rate to ICI monotherapy for that type. ICI, immune checkpoint inhibitor; MMR, mismatch repair; MMRd, MMR deficiency; MMRp, MMR proficiency; MSI, microsatellite instability; TCGA, The Cancer Genome Atlas.

assigned to the five subtypes was scored per tumor type, the score was strongly correlated with the previously reported objective response rate to ICI monotherapy for that tumor type (figure 1C).[1 18 19] The SMK/UVL/APB/POL/MRD subtypes thus serve to prognosticate positive response to ICI administration, and are hereafter termed immuno-responsive genomic subtypes (irGS).

Development of Genomic Subtyping and Predictive Response Analysis for Cancer Tumor ICi Efficacy

A software tool embedding machine learning was developed to stratify newly sequenced tumors into the eight genomic subtypes derived above (figure 2A). First, hierarchical clusters were again derived using each of three alternative variant calling schemes (online supplemental figure S9A, see the Methods section). High concurrence with analyses based on Mutect2 was observed (online supplemental figure S9A,B). To extract samples typical for each subtype as a training dataset, samples with matching classification results in at least three of the four methods,

including concomitant classification with Mutect2, were selected and used for subsequent analysis (online supplemental figure S9B,C). The resulting 7181 samples and their 30 COSMIC signature scores were used as features to construct k-nearest neighbor, support vector machine, random forest, and logistic regression classifiers with optimized hyperparameters (see the Methods and online supplemental figure S10A). All classifiers showed more than 95% subset accuracy (exact match ratio) in multi-label classification (online supplemental figure S10B), yielding a robust eight-class ensemble-based stratification tool.

For new query inputs of somatic mutation profile scores derived from tumor sequencing, each of the four classifiers is executed, and predictions are deemed consistent when the three or four resultant classifications concur; otherwise a classification of undeterminable (UND) is assigned to the sample. Further, when a majority prediction is one of the SMK, UVL, APB, POL, or MRD subtypes,
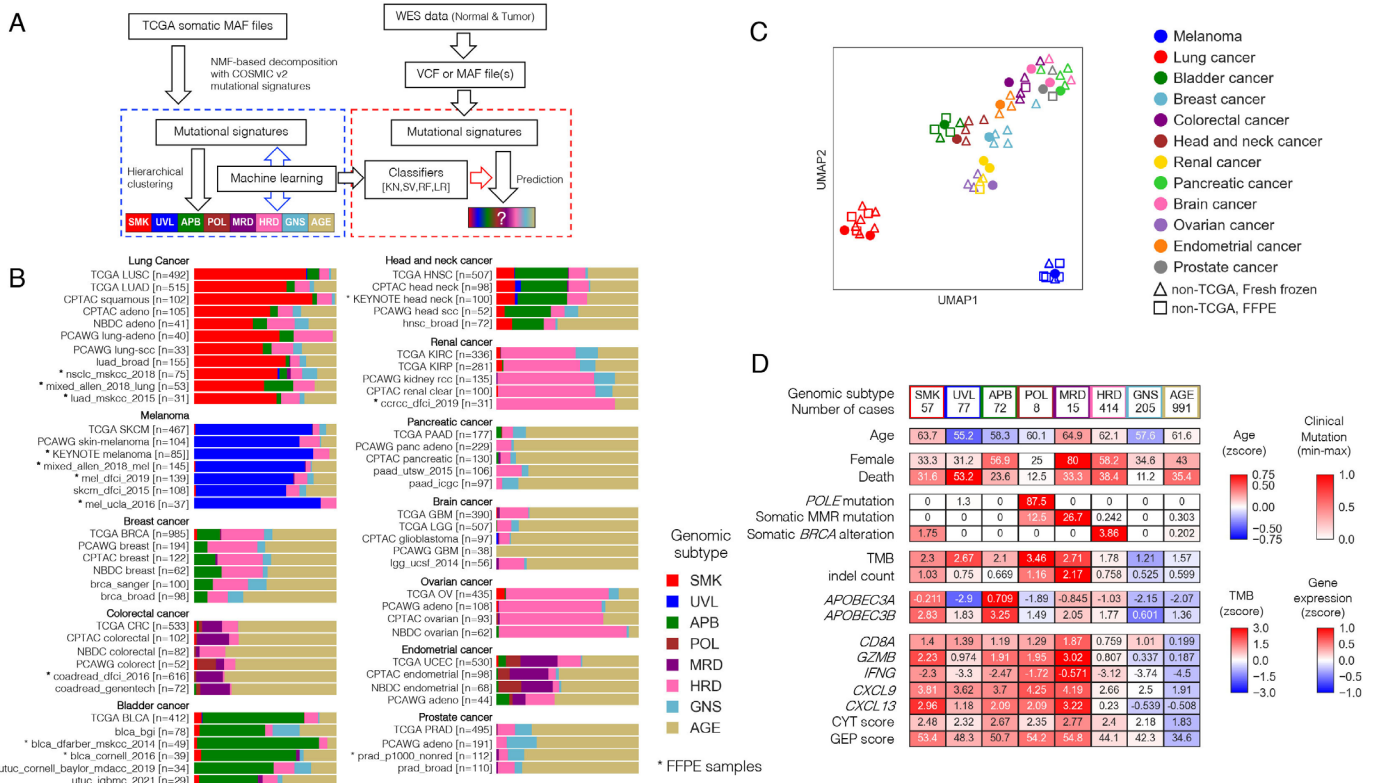
Figure 2 Development of GS-PRACTICE. (A) Overview of the program. Using the TCGA dataset, four different classifiers were built from four different algorithms, namely k-nearest neighbor (KN), support vector machine (SV), random forest (RF), and logistic regression (LR). Using external somatic mutation profiles from WES data as input, the four classifiers output classification results. (B) Subtyping results by GS-PRACTICE for each cancer type in the publicly available data (details in online supplemental table S2). Asterisks indicate data obtained from FFPE samples, which are similar to data obtained from frozen samples. Note that for NBDC colorectal cancer, the percentage of MMRd tumors has been reported to be low in Japanese.[50] (C) UMAP plot using the proportion of assigned subtypes as feature values leading to spatial projection. Marker color indicates the derived organ. Dot markers indicate TCGA data, triangles indicate non-TCGA data from frozen samples, and squares indicate non-TCGA data from FFPE samples. Datasets with the same cancer type are adjacent to each other, indicating a similar distribution of genomic subtypes across differing data sources. (D) Comparison between the genomic subtypes in PCAWG datasets with multiple cancer types (n=1916). Immune-related gene expression and scores were higher in irGS. The distribution of genomic subtypes in individual cancer types are indicated in figure 2B and online supplemental figure S11B. FFPE, formalin-fixed paraffin-embedded; CPTAC, Clinical Proteomic Tumor Analysis Consortium; GS-PRACTICE, Genomic Subtyping and Predictive Response Analysis for Cancer Tumor ICi Efficacy; irGS, immune-reactive genomic subtype; MMRd; mismatch repair deficiency; NBDC, National Bioscience Database Center; PCAWG, Pan-Cancer Analysis of Whole Genomes consortium; TCGA, The Cancer Genome Atlas.

the subtype is additionally classified as irGS; otherwise it is labeled as non-irGS. The prediction system, GS-PRAC-TICE (acronym for 'Genomic Subtyping and Predictive Response Analysis for Cancer Tumor ICi Efficacy'), is publicly available in the GitHub page (https://github.com/shirotak/GS-PRACTICE).

GS-PRACTICE was tested for its ability to stratify a diverse collection of samples from various sources into genomic subtypes (online supplemental table S2). Collectively, 96%–98% of samples were successfully assigned a subtype, indicating consensus in the ensemble's individual classifiers. The classifier concordance rate was also consistent across different data sources, and was consistent irrespective of whether samples were of formalin-fixed paraffin-embedded (FFPE) or frozen tissue origin (online supplemental figure S10C). Reanalyses restricted

to individual cancer types yielded identical conclusions with respect to tissue origin and source (figure 2B,C).

We applied GS-PRACTICE to 1916 samples of the PCAWG datasets[24] using somatic mutation profiles in coding regions obtained from the UCSC Xena[25] as input (online supplemental figure S11A,B). Results paralleled that of TCGA data (figure 2D). The differences in age, sex, and mortality among the subtypes were similar to those of TCGA. Somatic *POLE* mutations were common in the POL group, somatic MMR mutations in the MRD group, and somatic *BRCA* alterations (*BRCA1/2* mutations with LOH) in the HRD group. APOBEC3 family gene expression was elevated in the APB group. The five irGS subgroups demonstrated increased gene expression and biomarker scores associated with infiltration
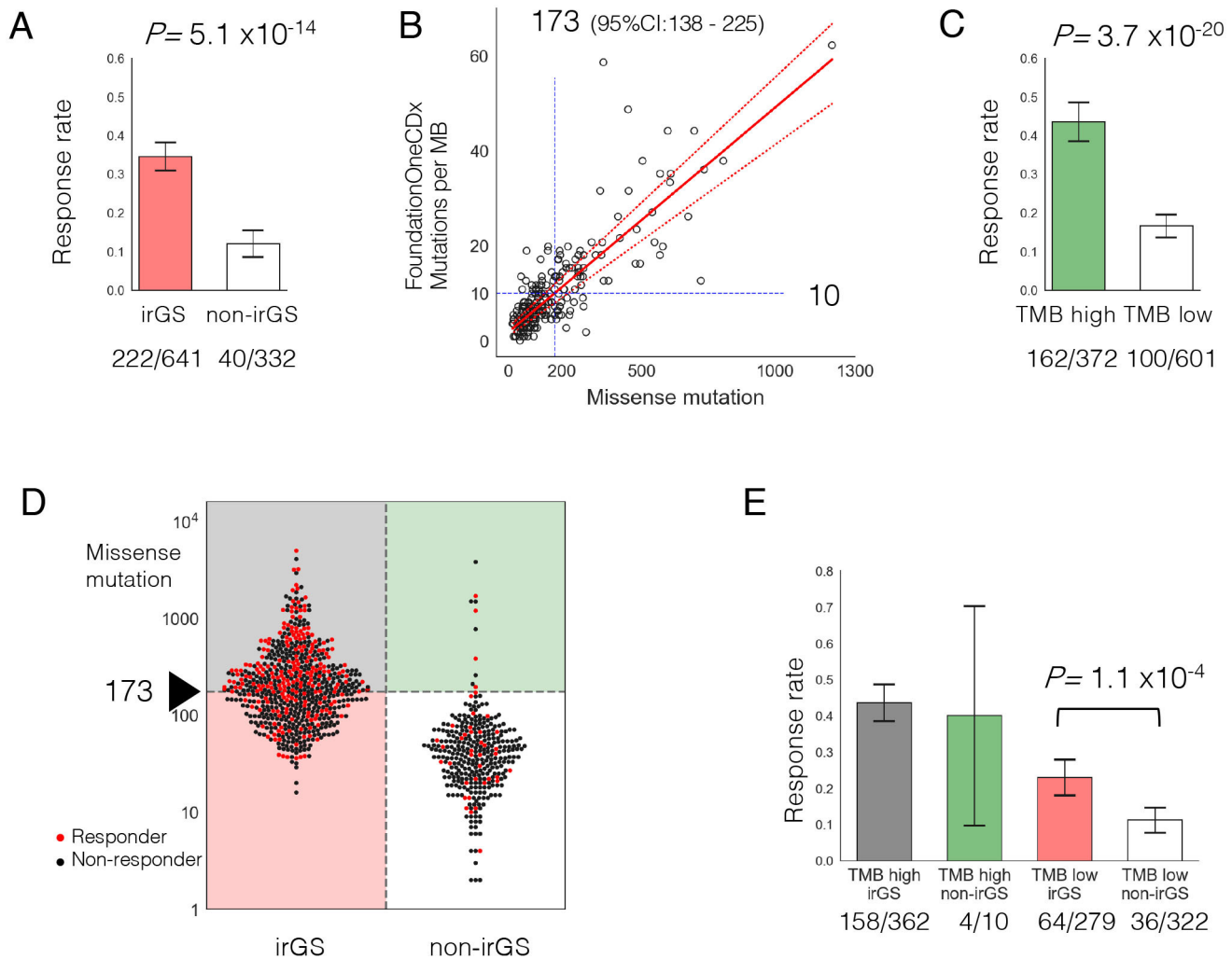
Figure 3 Prediction of ICI response by GS-PRACTICE and TMB. (A) ICI response rate was significantly higher in irGS tumors than non-irGS. (B) Determination of TMB cut-off. The bladder cancer dataset (n=218) from Mariathasan et al[26] was examined. The number of missense mutations in the whole-exome sequencing and the number of mutations per megabase from the FoundationOneCDx panel assay were plotted. From the regression equation, 10 mutations per megabase in the panel corresponded to 173 missense mutations (95% CI 138 to 225). (C) ICI response rate was significantly higher in TMB high tumors than TMB low tumors. (D) Association between distribution of TMB and ICI response per sample divided by irGS status. Red dots indicate responders, and black dots indicate non-responders. (E) Comparison of ICI response rates in the four groups stratified by irGS and TMB status. irGS tumors had a significantly higher response rate than non-irGS within the samples classified as TMB low. GS-PRACTICE, Genomic Subtyping and Predictive Response Analysis for Cancer Tumor ICi Efficacy; ICI, immune checkpoint inhibitor; irGS, immune-reactive genomic subtype; TMB, tumor mutational burden.

of cytotoxic CD8 +T cells and ICI response (figure 2D, online supplemental figures S1A and S4A).

GS-PRACTICE as a tumor agnostic predictive biomarker for ICI response

973 cases, most of whom have metastatic lesions (online supplemental table S3), with information on objective response to ICI treatment were used to challenge and assess the subtyping and (non-)irGS assignment from GS-PRAC-TICE (online supplemental table S3). Taken in total, ICI response rate was significantly higher in irGS than non-irGS (34.6% vs 12.0%, p=5.1×10^{-14}, figure 3A). When analyzed by the eight subtypes, the five subtypes belonging to irGS tended to have a higher response rate than the three non-irGS subtypes (online supplemental figure S12).

Next, to determine a cut-off for assignment of TMB-high, we compared the number of mutations detected in our WES pipeline with those in FoundationOne CDx using a bladder cancer dataset.[26] Based on Passing-Bablok regression analysis, the cut-off of 10 mutations per megabase in the CDx panel corresponds to 173 missense mutations in a WES sample (95% CI 138 to 225) (figure 3B). Using this value as the cut-off for TMB-high, tumors categorized as TMB-high showed higher ICI response rate than those as TMB-low (43.5% vs 16.6%, p=3.7×10^{-20}, figure 3C).

When we divided the tumors into four groups according to the pairwise stratifications of (non-)irGS and TMB-low/high, 97.2% of TMB-high tumors belonged to irGS and 96.9% of non-irGS tumors belonged to TMB-low

All cases (N=973) Variables		No.	Response rate (%)	Univariate Odds Ratio	95%CI	p-value	Multivariate Odds Ratio	95%CI	p-value
irGS	No	332	12.0						
	Yes	641	34.6	3.87	(2.68 - 5.59)	6.1×10^{-13}	2.18	(1.40 - 3.40)	5.6×10^{-4}
TMB	low	601	16.6						
	high	372	43.5	3.87	(2.87 - 5.20)	4.4×10^{-19}	2.65	(1.87 - 3.77)	5.0×10^{-8}
Cancer Type	Melanoma	335	32.2						
	Head neck	118	18.6	0.48	(0.29 - 0.81)	5.6×10^{-3}	0.83	(0.48 - 1.44)	0.51
	Lung	162	30.2	0.91	(0.61 - 1.37)	0.65	1.19	(0.77 - 1.84)	0.43
	Bladder	263	26.6	0.76	(0.53 - 1.09)	0.14	0.88	(0.60 - 1.28)	0.50
	Others	95	13.7	0.33	(0.18 - 0.62)	6.1×10^{-4}	0.87	(0.44 - 1.74)	0.70
ICI agent	Double	77	31.2						
	Single	896	26.6	0.80	(0.48 - 1.32)	0.38			
Sex	Female	342	24.3						
	Male	631	28.4	1.24	(0.91 - 1.67)	0.17			
Age	< 62	334	25.4						
	>= 62	339	26.8	1.08	(0.76 - 1.52)	0.68			

0   15   30   45

Figure 4 Univariate and multivariate logistic regression analysis for ICI response in the validation cohort. Multivariate logistic regression analysis showed that irGS was significantly associated with the ICI response after adjusting by TMB status and cancer type. Single, either anti-PD-1, anti-PD-L1, or anti-CTLA4 antibody; double, combination of anti-CTLA4 and anti-PD-1 or anti-PD-L1 antibodies. ICI, immune checkpoint inhibitor; irGS, immune-reactive genomic subtype; TMB, tumor mutational burden.

(figure 3D). Response rate to ICI was highest in the TMB-high irGS group (43.6%). Critically, within TMB-low tumors, irGS tumors had a significantly higher response rate than non-irGS (22.9% vs 11.2%, p=1.1×10^{-4}, figure 3E). Additionally, in a multivariate logistic regression analysis, irGS status was significantly associated with the objective response to ICI after adjustment for TMB-high status and cancer type (adjusted OR, 2.18; 95% CI, 1.40 to 3.40; p=5.6×10^{-4}, figure 4). The trends were similar when examined separately by anti-PD-1 antibody or anti-PD-L1 antibody therapy, as well as by anti-CTLA-4 monotherapy and anti-CTLA-4/anti-PD-1 combination therapy (online supplemental figure S13). These results were also significant when limited to data from the KEYNOTE clinical trials (n=311), a prospective cohort of patients treated solely with anti-PD-1 antibody, pembrolizumab (online supplemental figure S14). Although the KEYNOTE trials excluded patients with clinically diagnosed MMRd tumors at enrolment, two tumors from the cohort (one each with gastric cancer and biliary tract cancer) were classified into the MRD subtype, and both of them responded to ICI. Furthermore, the results were similar when using the cohort's optimal TMB cutoff determined by the ROC curve and the Youden index or using log(10)-transformed TMB as a continuous variable (online supplemental figure S15). Even when the recently reported score for estimating T-cell infiltration in tumors from WES data[27] was added as a covariate, there remained a significant correlation between irGS and ICI response (online supplemental figure S16). The definition of objective response was different in the data of

Anagnostou et al compared with other data (see online supplemental methods), but results were similar even after excluding such data (online supplemental figure S17). Genome subtyping and ICI response analysis by GS-PRACTICE for each of the 13 individual ICI studies comprizing the combined 973 patients are described in online supplemental figure S18 and table S4.

Finally, survival analysis was performed using data from the above ICI-treated cohorts (n=606, see the Methods section) to investigate whether irGS assignment by GS-PRACTICE was associated with overall survival. In univariate analysis, both irGS and TMB-high status were associated with favorable outcomes (log-rank test p=5.8×10^{-9}, 1.5×10^{-9}, figure 5A). Stratification analysis by the two statuses showed that the TMB-low non-irGS group had the worst overall survival (log-rank test p=9.0×10^{-11}, figure 5B). This trend was similarly observed when analyzed per cancer type (figure 5C). Furthermore, Cox proportional hazard model analysis adjusted for irGS, TMB status (binary or continuous), and cancer type showed that both irGS and TMB status were independent favorable prognostic factors (figure 5D).

DISCUSSION

The relationship between mutational signatures and ICI response has been previously reported for several specific types of cancer. For example, mutational signatures in melanoma[10] and non-small-cell lung cancer[11] correlate with response to ICI, and these data are explained by the idea that the process of carcinogenesis by exogenous
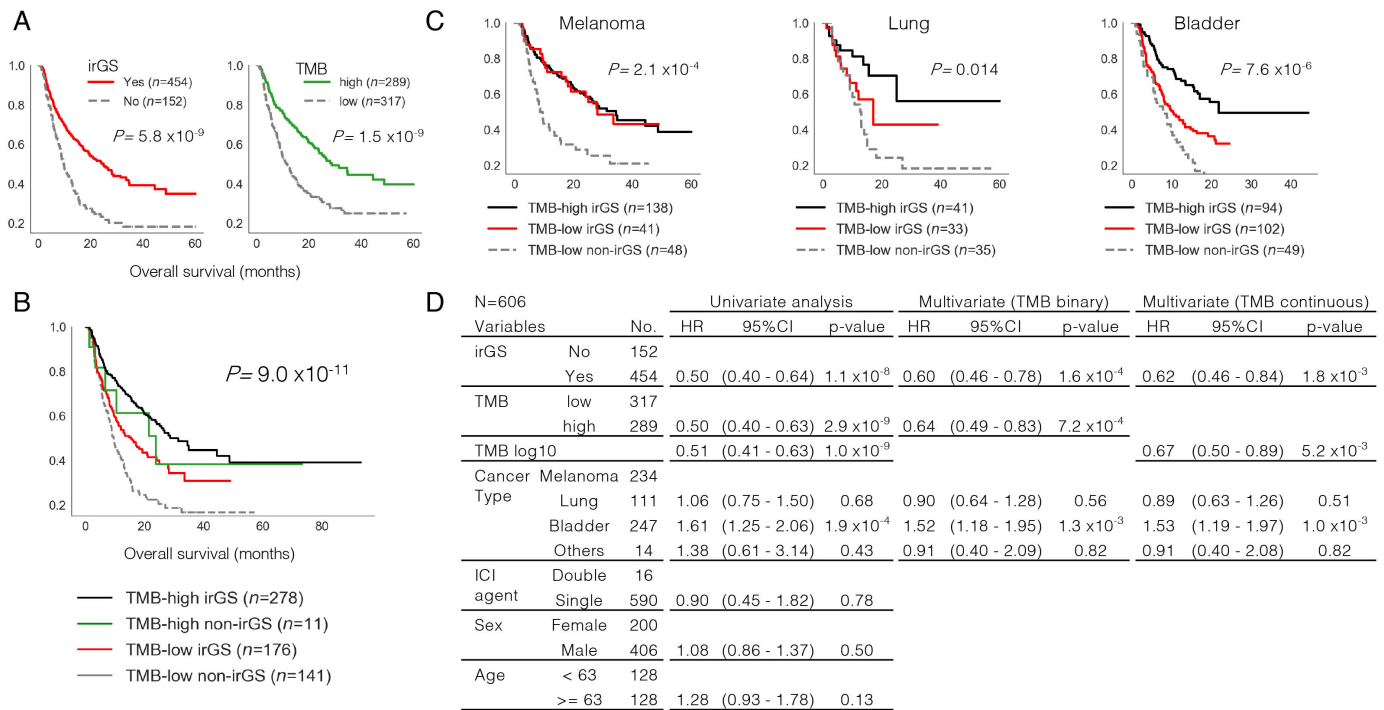
Figure 5 Overall survival analysis in ICI-treated cases. (A) Kaplan-Meier curve analysis. Both irGS (left) and TMB high (right) correlated with better survival outcome. (B) Stratification analysis by pairwise irGS and TMB status. The TMB-low non-irGS group showed the worst survival outcome. (C) Stratification analysis by irGS, TMB, and cancer type. The TMB-low non-irGS groups showed unfavorable outcomes when analyzed per cancer type. The TMB-high non-irGS groups are not shown due to insufficient subgroup size. (D) Cox hazard regression model analysis. Both irGS and TMB status (binary or continuous) were shown to be independent favorable prognostic factors even after adjusting for cancer type. ICI, immune checkpoint inhibitor; irGS, immune-reactive genomic subtype; TMB, tumor mutational burden.

mutagens (UV, tobacco) results in highly immunogenic tumor antigens.[28] In addition, APOBEC-related mutational signatures are associated with viral infections and a specific mutational pattern called kataegis, which also produces highly immunogenic antigens[29 30] and is associated with ICI response in non-small-cell lung cancer.[31 32] On the other hand, it has been reported that high copy number, aneuploidy, and HRD-associated scores inversely correlate with tumor immune response,[33–35] and negative results in a recent clinical trial in ovarian cancer where half of the tumors showed HRD[36 37] suggest that HRD-related signatures are unlikely to be associated with high sensitivity to ICIs. Aging-related (clock-like) mutational signatures are reported to be associated with lower immune activity in melanoma and non-small-cell lung cancer treated with ICI.[9 38] Since many age-related gene mutations also occur in non-tumor cells,[39] they may be related to immune tolerance. Our categorization of irGS and non-irGS in this study is supported by previous reports on the relationship between specific mutation signatures and tumor immunogenicity, and provides a cross-organ assessment of this relationship.

In June 2020, the FDA approved pembrolizumab for the treatment of tumors diagnosed 10 mutations per megabase or greater by FoundationOne CDx.[4] This cut-off corresponded to 173 missense mutations in our WES analysis (figure 3B) and was close to the optimal cut-off value of 165 calculated by the ROC curve based on the collated cohort we assembled (online supplemental figure S14). However, TMB quantification based on a panel assay is still subject to fluctuation (figure 3B).[40] In particular, tumor-only gene panel testing, including FoundationOne CDx, may overestimate TMB in non-Caucasians due to the paucity of public databases for germline variant filtering.[41 42] As sequencing now practically impacts clinical decisions, comprehensive sequencing methods including WES are optimal for reproducible and reliable measurements.[43] Furthermore, in currently available gene panel test data,[44] less than 0.5% of tumor samples have more than 100 gene mutations, even including synonymous ones, making it difficult to apply GS-PRACTICE to data from such panel assays. As the cost of WES decreases and efforts toward the implementation of WES as a routine cancer treatment continue to advance,[45 46] the combination of precision-improved TMB calculation and the orthogonal GS-PRACTICE method will usher in precise patient selection for ICI treatment.

There have been some criticisms that there is no logical basis for setting a universal TMB threshold for all solid tumors, since such an index is a continuous value that varies considerably among cancer types.[7 47 48] Our analysis showed that almost all non-irGS tumors belonged to TMB-low (figure 3D, online supplemental figures S13–S15), indicating that the current TMB cut-off has the consequence to exclude non-irGS tumors, which have little or no immunogenic background mutational processes. In

other words, our method may add biological rationales to the empirically determined TMB cut-off. Additionally, the previous report that the optimal cutoffs for TMB-high differed among cancer types[5] may be explained by the different distribution of genomic subtypes per tumor origin (figure 1B).

While GS-PRACTICE represents an advance in cancer diagnostics and clinical decision making, some limitations of this work must be made transparent. First, due to lack of data from randomized controlled trials, it cannot be concluded that the differences in response rate and prolonged survival observed in this study are fully attributable to the ICI efficacy. To elucidate this, design and logistics of appropriate randomized control trials using ICI are needed. Second, the clinical cohorts validated by GS-PRACTICE were mostly Caucasian patients, so future validation is needed to determine whether the program is applicable to non-Caucasian patients. Third, accurate subtyping may not be possible for tumors with a small number of mutations due to computational reasons. The clustering results using the four variant callers showed relatively low concordance rates for the HRD, AGE, and GNS subtypes (online supplemental figure S9B). Renal cancers had a moderately low number of mutations and were mostly classified as HRD, but their HRD scores and indel signature six ratios were low (online supplemental figure S7B), indicating that they are unlikely to have HRD properties. It is known that the response to ICI in renal cancer is not associated with TMB,[49] and the present analysis also did not identify any characteristic mutation patterns associated with ICI response. One method to improve on the state of the art would be to apply GS-PRACTICE to whole genome sequencing, which can detect dozens of times more mutations than WES.[8] This may allow for higher resolution mutation signature analysis and more sophisticated tumor genome subtyping even in tumors with a small number of coding mutations.

GS-PRACTICE represents a pan-cancer advancement in both solid tumor diagnostics and precision medicine, as it subtypes tumors by leveraging mutational signatures with defined etiologies, and the subtypes were shown to be indicative of ICI response. The method can be reproducibly applied to WES data derived from FFPE specimens, and thus immediately provide a predictive biomarker for ICI treatment in clinical practice. Future analyses of randomized control trials and whole genome sequencing will spur improved dataset generation for model building, which will subsequently strengthen the clinical utility of the protocol developed herein.

Author affiliations
[1]Department of Gynecology and Obstetrics, Kyoto University Graduate School of Medicine, Kyoto, Japan
[2]Life Science Informatics Research Unit, Department of Molecular Biosciences, Kyoto University Graduate School of Medicine, Kyoto, Japan
[3]Center for Cancer Immunotherapy and Immunobiology, Kyoto University Graduate School of Medicine, Kyoto, Japan
[4]Department of Obstetrics and Gynecology, Kindai University Faculty of Medicine, Osaka, Japan
[5]Project for Development of Innovative Research on Cancer Therapeutics, Cancer Precision Medicine Center, Japanese Foundation for Cancer Research, Tokyo, Japan

ORCID iDs
Shiro Takamatsu http://orcid.org/0000-0002-1569-6421
Junzo Hamanishi http://orcid.org/0000-0002-7750-0623
Osamu Gotoh http://orcid.org/0000-0002-9363-4584
Noriomi Matsumura http://orcid.org/0000-0002-4512-7975

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 Lemery S, Keegan P, Pazdur R. First FDA approval agnostic of cancer site - when a biomarker defines the indication. *N Engl J Med* 2017;377:1409–12.

3 Le DT, Durham JN, Smith KN, *et al*. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science* 2017;357:409–13.

4 Subbiah V, Solit DB, Chan TA, *et al*. The FDA approval of pembrolizumab for adult and pediatric patients with tumor mutational burden (TMB) ≥10: a decision centered on empowering patients and their physicians. *Ann Oncol* 2020;31:1115–8.

5 Valero C, Lee M, Hoen D, *et al*. Response rates to anti-PD-1 immunotherapy in microsatellite-stable solid tumors with 10 or more mutations per Megabase. *JAMA Oncol* 2021;7:739–43.

6 McGrail DJ, Pilié PG, Rashid NU, *et al*. High tumor mutation burden fails to predict immune checkpoint blockade response across all cancer types. *Ann Oncol* 2021;32:661–72.

7 Alexandrov LB, Nik-Zainal S, Wedge DC, *et al*. Signatures of mutational processes in human cancer. *Nature* 2013;500:415–21.

8 Alexandrov LB, Kim J, Haradhvala NJ, *et al*. The repertoire of mutational signatures in human cancer. *Nature* 2020;578:94–101.

9 Budczies J, Seidel A, Christopoulos P, *et al*. Integrated analysis of the immunological and genetic status in and across cancer types: impact of mutational signatures beyond tumor mutational burden. *Oncoimmunology* 2018;7:e1526613.

10 Pham TV, Boichard A, Goodman A, *et al*. Role of ultraviolet mutational signature versus tumor mutation burden in predicting response to immunotherapy. *Mol Oncol* 2020;14:1680–94.

11 Rizvi NA, Hellmann MD, Snyder A, *et al*. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science* 2015;348:124–8.

12 Cibulskis K, Lawrence MS, Carter SL, *et al*. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* 2013;31:213–9.

13 Thorsson V, Gibbs DL, Brown SD, *et al*. The immune landscape of cancer. *Immunity* 2018;48:812–30.

14 Huang K-L, Mashl RJ, Wu Y, *et al*. Pathogenic germline variants in 10,389 adult cancers. *Cell* 2018;173:355–70.

15 Knijnenburg TA, Wang L, Zimmermann MT, *et al*. Genomic and molecular landscape of DNA damage repair deficiency across the cancer genome atlas. *Cell Rep* 2018;23:239–54.

16 Blokzijl F, Janssen R, van Boxtel R, *et al*. MutationalPatterns: comprehensive genome-wide analysis of mutational processes. *Genome Med* 2018;10:33.

17 Vaddepally RK, Kharel P, Pandey R, *et al*. Review of indications of FDA-approved immune checkpoint inhibitors per NCCN guidelines with the level of evidence. *Cancers* 2020;12:738.

18 Yarchoan M, Albacker LA, Hopkins AC, *et al*. PD-L1 expression and tumor mutational burden are independent biomarkers in most cancers. *JCI Insight* 2019;4:e126908.

19 Gómez-Raposo C, Merino Salvador M, Aguayo Zamora C, *et al*. Immune checkpoint inhibitors in endometrial cancer. *Crit Rev Oncol Hematol* 2021;161:103306.

20 Takamatsu S, Brown JB, Yamaguchi K, *et al*. Utility of homologous recombination deficiency biomarkers across cancer types. *JCO Precis Oncol* 2022;6:e2200085.

21 Litchfield K, Reading JL, Puttick C, *et al*. Meta-analysis of tumor- and T cell-intrinsic mechanisms of sensitization to checkpoint inhibition. *Cell* 2021;184:596–614.

22 Rooney MS, Shukla SA, Wu CJ, *et al*. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* 2015;160:48–61.

23 Cristescu R, Mogg R, Ayers M, *et al*. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science* 2018;362:eaar3593.

24 ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium. Pan-cancer analysis of whole genomes. *Nature* 2020;578:82–93.

25 Goldman MJ, Craft B, Hastie M, *et al*. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol* 2020;38:675–8.

26 Mariathasan S, Turley SJ, Nickles D, *et al*. TGFβ attenuates tumour response to PD-L1 blockade by contributing to exclusion of T cells. *Nature* 2018;554:544–8.

27 Bentham R, Litchfield K, Watkins TBK, *et al*. Using DNA sequencing data to quantify T cell fraction and therapy response. *Nature* 2021;597:555–60.

28 Jardim DL, Goodman A, de Melo Gagliato D, *et al*. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell* 2021;39:154–73.

29 Boichard A, Tsigelny IF, Kurzrock R. High expression of PD-1 ligands is associated with *kataegis* mutational signature and APOBEC3 alterations. *Oncoimmunology* 2017;6:e1284719.

30 Boichard A, Pham TV, Yeerna H, *et al*. APOBEC-related mutagenesis and neo-peptide hydrophobicity: implications for response to immunotherapy. *Oncoimmunology* 2018;8:1550341.

31 Wang S, Jia M, He Z, *et al*. APOBEC3B and APOBEC mutational signature as potential predictive markers for immunotherapy response in non-small cell lung cancer. *Oncogene* 2018;37:3924–36.

32 Chen H, Chong W, Teng C, *et al*. The immune response-related mutational signatures and driver genes in non-small-cell lung cancer. *Cancer Sci* 2019;110:2348–56.

33 Davoli T, Uno H, Wooten EC, *et al*. Tumor aneuploidy correlates with markers of immune evasion and with reduced response to immunotherapy. *Science* 2017;355:eaaf8399.

34 Kraya AA, Maxwell KN, Wubbenhorst B, *et al*. Genomic signatures predict the immunogenicity of BRCA-Deficient breast cancer. *Clin Cancer Res* 2019;25:4363–74.

35 van Wilpe S, Tolmeijer SH, Koornstra RHT, *et al*. Homologous recombination repair deficiency and implications for tumor immunogenicity. *Cancers* 2021;13:2249.

36 Monk BJ, Colombo N, Oza AM, *et al*. Chemotherapy with or without avelumab followed by avelumab maintenance versus chemotherapy alone in patients with previously untreated epithelial ovarian cancer (JAVELIN Ovarian 100): an open-label, randomised, phase 3 trial. *Lancet Oncol* 2021;22:1275–89.

37 Moore KN, Bookman M, Sehouli J, *et al*. Atezolizumab, bevacizumab, and chemotherapy for newly diagnosed stage III or IV ovarian cancer: placebo-controlled randomized phase III trial (IMagyn050/GOG 3015/ENGOT-OV39). *J Clin Oncol* 2021;39:1842–55.

38 Chong W, Wang Z, Shang L, *et al*. Association of clock-like mutational signature with immune checkpoint inhibitor outcome in patients with melanoma and NSCLC. *Mol Ther Nucleic Acids* 2021;23:89–100.

39 Kakiuchi N, Ogawa S. Clonal expansion in non-cancer tissues. *Nat Rev Cancer* 2021;21:239–56.

40 Merino DM, McShane LM, Fabrizio D, *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.

41 Parikh K, Huether R, White K, *et al*. Tumor mutational burden from tumor-only sequencing compared with germline subtraction from paired tumor and normal specimens. *JAMA Netw Open* 2020;3:e200202.

42 Asmann YW, Parikh K, Bergsagel PL, *et al*. Inflation of tumor mutation burden by tumor-only sequencing in under-represented groups. *NPJ Precis Oncol* 2021;5:22.

43 Mankor JM, Paats MS, Groenendijk FH, *et al*. Impact of panel design and cut-off on tumour mutational burden assessment in metastatic solid tumour samples. *Br J Cancer* 2020;122:953–6.

44 AACR Project GENIE Consortium. AACR project GENIE: powering precision medicine through an international Consortium. *Cancer Discov* 2017;7:818–31.

45 Suwinski P, Ong C, Ling MHT, *et al*. Advancing personalized medicine through the application of whole exome sequencing and big data analytics. *Front Genet* 2019;10:49.

46 Bayle A, Droin N, Besse B, *et al*. Whole exome sequencing in molecular diagnostics of cancer decreases over time: evidence from a cost analysis in the French setting. *Eur J Health Econ* 2021;22:855–64.

47 Rousseau B, Foote MB, Maron SB, *et al*. The spectrum of benefit from checkpoint blockade in hypermutated tumors. *N Engl J Med* 2021;384:1168–70.

48 Strickler JH, Hanks BA, Khasraw M. Tumor mutational burden as a predictor of immunotherapy response: is more always better? *Clin Cancer Res* 2021;27:1236–41.

49 Braun DA, Hou Y, Bakouny Z, *et al*. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. *Nat Med* 2020;26:909–18.

50 Akagi K, Oki E, Taniguchi H, *et al*. Nationwide large-scale investigation of microsatellite instability status in more than 18,000 patients with various advanced solid cancers. *J Clin Oncol* 2020;38:803.