ORIGINAL ARTICLE

Cancer Science WILEY

# Characterization of tumors with ultralow tumor mutational burden in Japanese cancer patients

Keiichi Hatakeyama[1] | Takeshi Nagashima[2,3] | Keiichi Ohshima[1] |
Sumiko Ohnami[2] | Shumpei Ohnami[2] | Yuji Shimoda[2,3] | Akane Naruoka[4] |
Koji Maruyama[5] | Akira Iizuka[6] | Tadashi Ashizawa[6] |
Tohru Mochizuki[1] | Kenichi Urakami[2] | Yasuto Akiyama[6] | Ken Yamaguchi[7]

[1]Medical Genetics Division, Shizuoka Cancer
Center Research Institute, Shizuoka, Japan

[2]Cancer Diagnostics Research Division,
Shizuoka Cancer Center Research Institute,
Shizuoka, Japan

[3]SRL Inc, Tokyo, Japan

[4]Drug Discovery and Development Division,
Shizuoka Cancer Center Research Institute,
Shizuoka, Japan

[5]Experimental Animal Facility, Shizuoka
Cancer Center Research Institute, Shizuoka,
Japan

[6]Immunotheraphy Division, Shizuoka Cancer
Center Research Institute, Shizuoka, Japan

[7]Shizuoka Cancer Center, Shizuoka, Japan

Correspondence
Keiichi Hatakeyama, Medical Genetics
Division, Shizuoka Cancer Center Research
Institute, Shizuoka, Japan.
Email: k.hatakeyama@scchr.jp

## Abstract

Tumor mutational burden analysis using whole-exome sequencing highlights features of tumors with various mutations or known driver alterations. Cancers with few changes in the exon regions have unclear characteristics, even though low-mutated tumors are often detected in pan-cancer analysis. In the present study, we analyzed tumors with low tumor mutational burden listed in the Japanese version of The Cancer Genome Atlas, a data set of 5020 primary solid tumors. Our analysis revealed that detection rates of known driver mutations and copy number variation were decreased in samples with tumor mutational burden below 1.0 (ultralow tumor), compared with those in samples with low tumor mutational burden (≤5 mutations/Mb). This trend was also observed in The Cancer Genome Atlas data set. In the ultralow tumor mutational burden tumors, expression analysis showed decreased TP53 inactivation and chromosomal instability. TP53 inactivation frequently correlated with PI3K/mTOR-related gene expression, implying suppression of the PI3K/mTOR pathway in ultralow tumor mutational burden tumors. In common with mutational burden, the T cell-inflamed gene expression profiling signature was a potential marker for prediction of an immune checkpoint inhibitor response, and some ultralow tumor mutational burden tumor populations highly expressed this signature. Our analysis focused on how these tumors could provide insight into tumors with low somatic alteration that are difficult to detect solely using whole-exome sequencing.

### KEYWORDS
gene expression signature, Japanese Cancer Genome Atlas, TMB ultralow, TP53 inactivation, tumor mutational burden

# 1 | INTRODUCTION

TMB, an emerging characteristic of human cancer, was first highlighted by large-scale integrated mutation analysis using next-generation sequencing.[1] A TMB increase can be attributed to both endogenous factors and environmental damage.[2,3] Common TMB is estimated based on the number of somatic mutations in the exon regions and their flanking regions using WES or target panel sequencing. Tumors with high mutation frequency tend to respond to immune checkpoint inhibitors[4], therefore TMB has attracted considerable attention as a biomarker for the treatment of patients with cancer.

Although TMB analyses across the globe are performed using WES, most cancers are determined to be TMB low (TMB ≤ 5 mutation/Mb).[1,5] Driver mutations in exon regions were occasionally found to be absent in diverse tumor types,[6,7] implying the presence of somatic alterations that were undetectable only by WES. Therefore, characterization of TMB-low tumors is valuable for understanding tumorigenesis without known driver alterations.

The availability of data sets from TCGA has expanded the possibilities for mutation analysis and GEP. In analyses using surgical specimens, GEP reflects gene expression in both heterogenous cancer cells and normal tissues, including immune cells, and therefore enabled the extraction of tumorigenesis-related gene set information such as TP53 inactivation,[8] chromosomal instability (CIN),[9-11] methylation,[12] and the PI3K/Akt/mTOR pathway[13-15] as tumor signatures. For instance, TP53 inactivation score derived from the expression of gene sets in multiple cancer types can be used to isolate cell populations with cell cycle progression or apoptosis suppression due to TP53 inactivation.[8] Furthermore, information on the tumor microenvironment could be estimated from a specific gene set.[16-19] Signature analysis based on GEP has the potential to help visualize tumor features that cannot be distinguished by mutations.

We previously investigated TMB, driver mutation, CNV, mutational signature, and gene expression in the tumor microenvironment in Japanese patients with solid tumors[5] and consolidated this data set and pipeline into the JCGA toward the progress of cancer research and clinical sequencing.[7] These studies focused on known somatic alterations in the exon regions and their flanking regions. However, a recent integrated genome analysis revealed that other regions containing intergenic sites harbored important alterations, including genes related to tumorigenesis, despite mutation accumulation in the exon regions.[6] Therefore, in-depth WES and GEP analysis of tumors with low somatic mutations may provide a new perspective prior to whole genome sequencing of samples from Japanese patients with cancer.

In the current study, to characterize low-mutated tumors, we performed WES and GEP of the JCGA data set, which is composed of 5020 primary solid tumors. First, detection rates of known driver mutations and CNV were investigated in low-mutated tumors. Next, to clarify tumor signature based on GEP, expression levels of a tumorigenesis-related gene set were investigated. Finally, a potential expression marker for response prediction to an ICI was compared between low-mutated and high-mutated tumors. Our analysis could identify characteristics of tumors with low somatic alteration that are hardly detected solely by WES.

# 2 | MATERIALS AND METHODS

## 2.1 | Ethical statement

Written informed consent was obtained from all patients, and the institutional review board at the Shizuoka Cancer Center approved all aspects of this study (authorization number 25-33). All experiments using clinical samples were performed in accordance with the approved Japanese ethical guidelines (human genome/gene analysis research, 2017, provided by Ministry of Health, Labor, and Welfare; https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/hokabunya/kenkyujigyou/i-kenkyu/index.html).

## 2.2 | Clinical samples

Tumors and their surrounding tissue (≥0.1 g) were dissected from surgical specimens immediately after resection of the lesion at the Shizuoka Cancer Center Hospital. The tumor samples were visually assessed by a clinical pathologist in our hospital when tumor content was ≥50%, and were not filtered further by pathophysiological features or cancer type. Tumor-adjacent tissue specimens were used as the normal control for microarrays. In addition, peripheral blood was collected as a pair control for excluding germline mutation. Details of experimental protocols have been described previously.[5,7,20-23]

## 2.3 | Data sets for analysis of somatic alteration

To analyze both sequencing and microarray data derived from the freshly frozen clinical samples, we used our own pipeline detailed in a previous report.[7] In brief, the exome library for WES was constructed using an Ion Torrent AmpliSeq RDY Exome Kit (Thermo Fisher Scientific). The exome library supplied 292 903 amplicons covering 57.7 Mb of the human genome, comprising 34.8 Mb of exonic sequences from 18 835 genes registered in RefSeq. In the pipeline, blood and tumors harboring enormous differences in single nucleotide polymorphisms (SNP) derived from the same patient were reconfirmed and retested to eliminate sampling error as much as possible. Furthermore, to avoid sequencer- and amplicon-derived errors, arbitrary somatic mutations were manually inspected using the Integrative Genomics Viewer (IGV), and somatic mutation candidates containing multiple nucleotide variations (~1000 sites) were validated using Sanger sequencing. This data set comprised 5521 tumor specimens (5020 primary tumors and 501 metastatic tumors) derived from 5143 patients, and all samples were analyzed using WES and target panel sequencing. These data were then submitted to the National Bioscience Database

Center (NBDC) Human Database as 'Controlled-Access Data' (Research ID, hum0127.v1; https://humandbs.biosciencedbc.jp/en/). All data sets are scheduled to be released in January 2021. For this study, part of the data set was extracted for analysis of somatic alterations including TMB, CNV, tumor cellularity, gene expression signature, and fusion genes (Table S1). Tumor cellularity was estimated by averaging 2 individual algorithms (FACETS[24] and Sequenza[25]) as described in a previous study.[7] Somatic genomic alterations contributing to tumorigenesis were extracted from multiple databases, and then curated in-house as driver mutations.[7] To analyze known fusion transcripts, a 491 fusion-gene panel for constructing fusion-specific amplicons was constructed.[22] Somatic CNVs were detected using saasCNV.[26] This method accounted for both the read-depth ratio and B allele frequency and achieved the best performance among 6 CNV detection tools.[27] To compare between our platform and others, mutation and expression profiles were extracted from the public database in the TCGA project.[28-33] The list used in this study is provided in Table S2.

## 2.4 | Gene expression signature analysis

For RNA analysis, tumor and normal tissues were immediately immersed in RNAlater solution (Thermo Fisher Scientific). Purified total RNA for GEP was amplified and fluorescently labeled using a One-Color Low Input Quick Amp Labeling Kit (Agilent Technologies). Cy3-labeled cRNAs were hybridized to a SurePrint G3 Human Gene Expression 8 × 60K v2 Microarray (Agilent Technologies), which contained 50 599 probes representing 29 833 genes registered in the Entrez Gene Database. Signature analysis based on gene expression was performed using the expression ratio of the tumor and the corresponding normal tissue (T/N). The expression signature/score was calculated from the average of genes in unique gene sets corresponding to each individual signature (TP53 inactivation,[8] CIN,[9-11] methylation,[12] PI3K/mTOR CMAP UP,[13-15] and T cell-inflamed GEP[17-19]). To suppress influence of gene expression derived from normal tissue, the signature analysis was performed using T/N expression ratio.[5]

## 2.5 | Statistical analysis

To compare CNV size, protein length, and age, Wilcoxon signed-rank test was performed. Comprehensive data, such as those obtained from expression analysis using microarray or RNA-seq, were compared using the Welch $t$ test based on a standard protocol. To graphically assume distribution as a standard normal, a quantile-quantile (Q-Q) plot was employed. The log-rank test was performed for overall survival and progression-free survival curves. To control the FDR, the Benjamini-Hochberg procedure ($q < .01$) was carried out, and results were considered significant at $P$-values $< .01$. The CNV of each individual gene on each chromosome was considered significant at $P$-values $< 1 \times 10^{-6}$.

## 2.6 | Data accessibility statement

The authors declare that all the other data supporting the findings of this study are available within the article and its supporting information files and from the corresponding author upon reasonable request.

# 3 | RESULTS

## 3.1 | Tissue distribution and tumor cellularity in low-TMB tumors

We extracted 5020 primary tumors from the JCGA data set, of which 4070 samples (81.1%) had TMB ≤ 5 mutations/Mb (Figure 1A), defined as TMB low. Tissue distribution of TMB-low tumors was similar to that in the data set without TMB filtering. TMB and total CNV size of these samples tended to decrease with low tumor cellularity (Figure 1B). To avoid any influence of low tumor content on the detection of somatic alterations, 1394 samples with low tumor cellularity (<0.3) were excluded from our main analysis (samples meeting the above criteria are listed in Table S1), and then tumors with <1.0 mutation/Mb were defined as TMB ultralow (n = 430). Tissue distribution for sufficient tumor cellularity (≥0.3) was similar to that in the data set without TMB filtering (Figure S1A). TMB of these populations was estimated as abnormal distribution from the histogram and Q-Q plot (Figure S1B,C). The CNV size was significantly low ($P < 2.2 \times 10^{-16}$) in TMB-ultralow tumors, implying that these ultralow-mutated tumors were less prone to somatic chromosome alterations (Figure S1D).

## 3.2 | Driver mutations in TMB-ultralow tumor

The frequency of TMB classes (low and ultralow) was compared across tumor types (Figure 1C). Benign tumors contained the highest TMB-ultralow frequency, exceeding 50% in gastrointestinal stromal tumor (GIST), thymus, and pancreas. The detection rate of driver-mutated samples was next investigated for 6 tumor types; this included more than 20 samples for both TMB classes (Figure 1D). The driver mutations were annotated as Tier 1 or 2 according to a previous report.[7] The frequency of driver mutations was decreased in TMB-ultralow tumors compared with that in TMB-low tumors. We observed this tendency in the TCGA data set extracted by our analysis pipeline (Figure S2A). Furthermore, we also checked known fusion genes and confirmed that the number of detected fusion genes was decreased in TMB-ultralow tumors (Figure S2B).

To confirm whether the decrease in the number of driver mutations, including fusion genes, accompanied the decline in TMB, the frequency of driver mutations was reanalyzed using equally segmented sample populations (Figure S3A). In TMB-ultralow tumors, the detection rate of Tier 1/2 driver mutations dropped below 50%, whereas these mutations were observed in over 80% of samples for
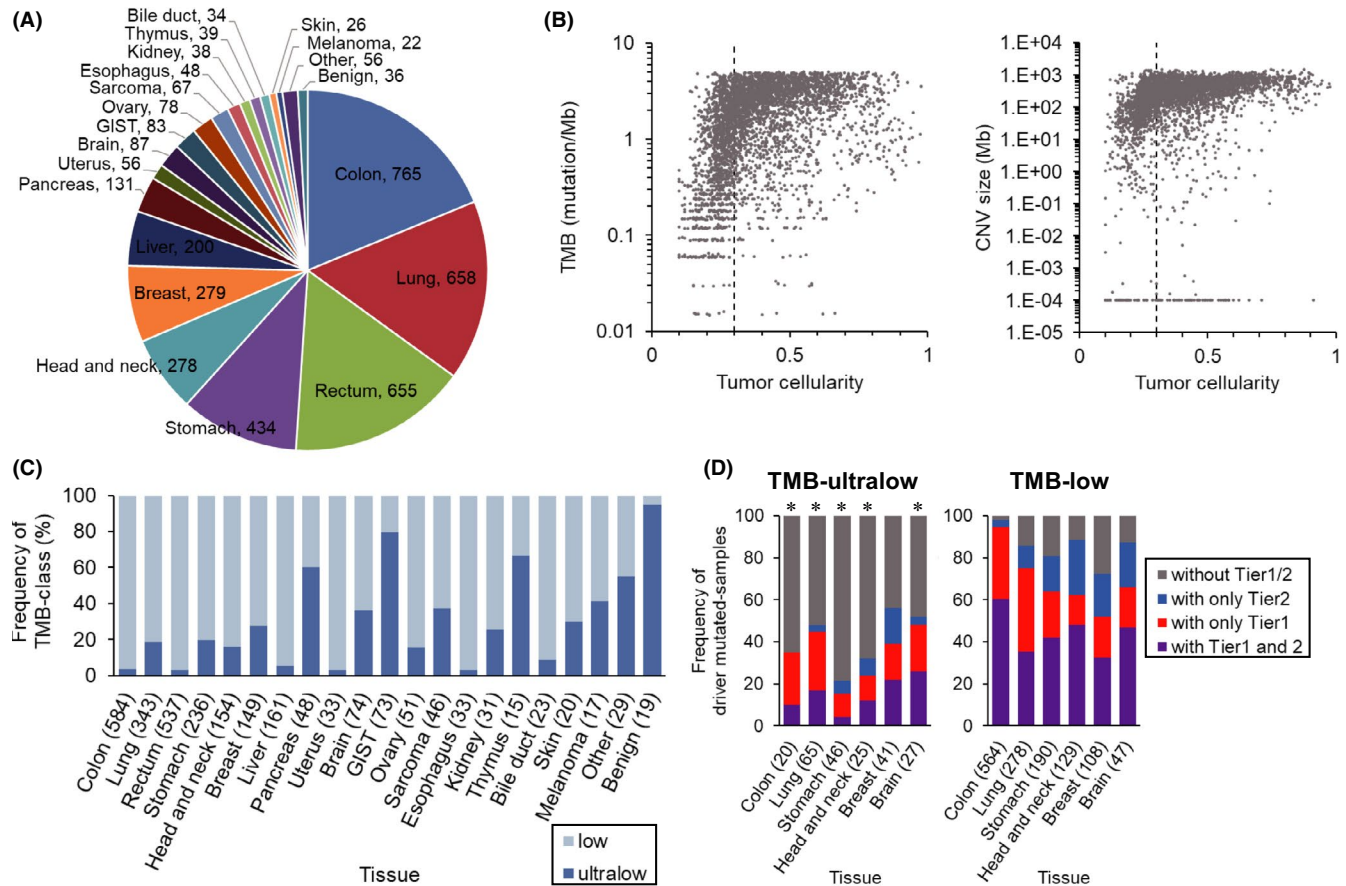
**FIGURE 1** Sample profile in tumor mutational burden (TMB)-ultralow and TMB-low tumors. A, Distribution of tumor types in 4070 samples. The 'other' group contains multiple tumor types that comprise <20 samples. B, Influence of tumor cellularity on TMB and copy number variation (CNV) in whole-exome sequencing (WES). The samples were sorted in descending order by CNV size, calculated from the sum of loss (CNV ≤ 1.5) and gain (CNV ≥ 2.5) in the genome estimated by WES. C, Frequency of TMB-ultralow and TMB-low samples in each tumor type. These samples met the criteria of tumor cellularity (≥0.3). D, Detection rate of driver-mutated samples. The tumor types that include more than 20 samples in both TMB classes (ultralow and low) were selected. Driver mutation was defined as Tier 1 or 2 according to a previous report.[7] Integer in parentheses represents number of samples. *Significant differences in the number of driver mutations detected between TMB-ultralow and TMB-low tumors (Fisher exact test, $P < .01$)

all intervals of TMB-low tumors. To further evaluate the impact of high tumor cellularity on the frequency of driver-mutated samples, we divided samples into intermediate and high (≥0.65) tumor cellularity (Figure S3B). No significant difference in frequency of each TMB class was observed between different degrees of cellularity, indicating that high tumor cellularity has an insignificant effect on the detection of driver mutations.

VAF is known experimentally to be influenced by tumor cellularity. We checked the correlation of tumor cellularity with VAF of driver mutations (Figure S3C,D). The VAF score was moderately correlated with tumor cellularity, and was decreased in TMB-ultralow tumors, raising the possibility that detection sensitivity for driver mutations in TMB-ultralow tumors may be reduced due to a few mutation reads in exome sequencing. To clarify the influence of VAF on detection sensitivity, driver mutations detected using WES were compared with those detected by target panel sequencing, which has a higher read-depth than WES (Figure S3E). Of the driver mutations, 86.3% were identified by both platforms,

indicating that the detection of driver mutations in samples extracted based on the criterion (tumor cellularity > 0.3) was impervious to VAF. Furthermore, in TMB-ultralow tumors with or without driver mutations, no significant differences in CNV size and tumor cellularity were observed (Figure S3F), suggesting that detection of driver mutations is not influenced by tumor cellularity. These results indicated that TMB-ultralow tumors are less prone to accumulating known somatic alterations related to tumorigenesis in gene coding regions.

## 3.3 | Impact of increasing age and malignancy on mutation accumulation

The number of driver mutations was positively correlated with increasing age.[34,35] We investigated the relationship between age and accumulation of known driver mutations, including fusion genes (Figure S4). Increasing patient age correlated with a higher number

of driver mutations in 5 tumor types (colon, lung, head and neck, breast, and brain tumors) except stomach tumors. However, no significant difference in age at operation was observed between TMB-ultralow and TMB-low tumors (Table 1).

Benign tumors were mostly classified as TMB-ultralow (see Figure 1C). This finding raises the possibility that the number of mutations reflects malignancy in TMB-low tumors. Using TCGA data sets, we checked overall and progression-free survival in TMB-ultralow and TMB-low tumors (Figure S5). No significant differences in Kaplan-Meier survival curves were observed between TMB classes. These results suggested that biases of increasing age and malignancy were insignificant for our TMB classification.

## 3.4 | Specific somatic mutations in TMB-ultralow and TMB-low tumors

The detection rate of driver mutations, including fusion genes, differed between TMB-ultralow and TMB-low tumors. To further investigate other somatic mutations, the accumulation of mutations in all genes targeted by WES was measured in 6 tumor types (colon, lung, stomach, head and neck, breast, and brain tumors) without known driver alterations, including fusion genes (Figure 2). Among the genes with the top 10 mutations in TMB-ultralow and TMB-low tumors, *MUC17* was consistently mutated in different tumor types. *TP53* alterations except known driver mutations (Tier1/2) were detected only in TMB-low tumors. Furthermore, 1988 cancer-related genes listed in the previous report[7] were undetectable in TMB-ultralow tumors. The mutated genes shown in Figure 2 were confirmed using IGV to avoid sequencer- and amplicon-derived errors (Figure S6), resulting in the exclusion of *ZNF714* mutations due to potential sequencer-derived errors. Protein length after translating these genes in TMB-ultralow tumors was shorter than that in TMB-low tumors (Figure S7), implying that the propensity of mutations may be dependent on the length of exons in the gene. These results suggested that TMB-ultralow tumors were less likely to accumulate known driver mutations and alterations in exon regions in cancer-related genes.

**TABLE 1** Age at operation in TMB-ultralow and TMB-low tumors

| Tissue | Average of age at operation (Ave. ± SD) | | P-value[a] |
| | TMB-ultralow | TMB-low | |
|---|---|---|---|
| Colon | 70 ± 12 | 68 ± 11 | .148 |
| Lung | 68 ± 10 | 69 ± 10 | .470 |
| Stomach | 69 ± 10 | 69 ± 9 | .744 |
| Head and neck | 58 ± 19 | 66 ± 12 | .0553 |
| Breast | 54 ± 15 | 59 ± 13 | .107 |
| Brain | 54 ± 18 | 57 ± 15 | .970 |

[a]Wilcoxon signed-rank test.

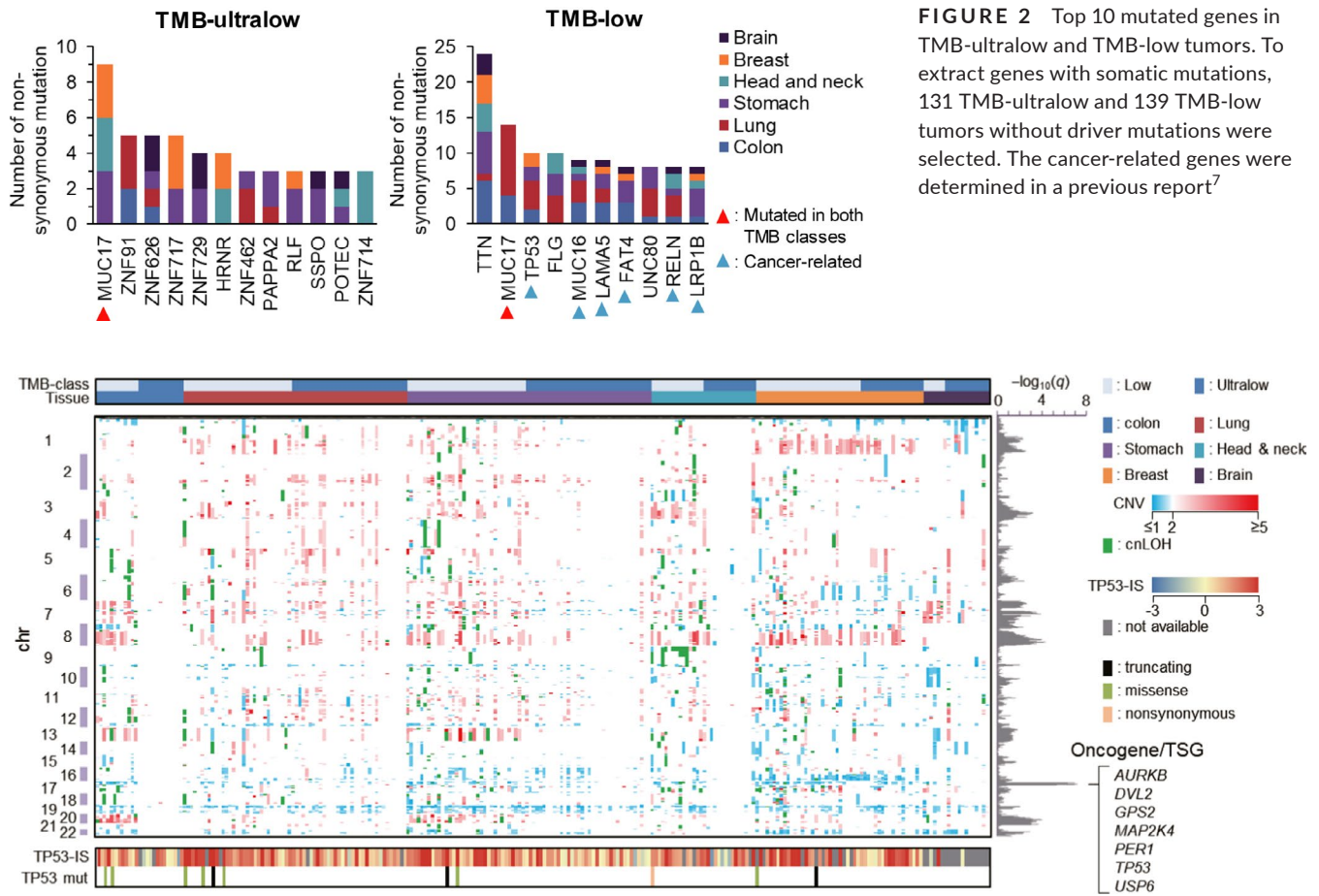## 3.5 | CNV and TP53 inactivation in TMB-ultralow tumors

TMB-low tumors harbored more driver or cancer-related mutations than TMB-ultralow tumors. To find other alterations superseding driver mutations, we performed CNV analysis of each chromosome (Figure 3). In TMB-ultralow tumors, accumulation of CNV was lower for almost all chromosomes. Especially, the loss of chromosome 17 encoding *TP53* was observed independently of tumor type, implying that normal TP53 function was maintained in TMB-ultralow tumors. Defective TP53 function is known to contribute to CIN progress. Based on previous expression analysis,[8] we calculated the TP53 inactivation score (Figures 3 and S8A). This inactivation score was significantly decreased in TMB-ultralow tumors, and gene set enrichment analysis (GSEA)[36,37] revealed that cell cycle pathways related to TP53 inactivation were enriched, this finding is consistent with that in a previous report[8] (Figure S8B). Furthermore, the TP53 inactivation score was highly correlated with CIN signature[9-11] (Figure S8C). Conversely, the inactivation score was not associated with methylation signature estimated from GEP[12] (Figure S8D). These results suggested that both CNV and TP53 inactivation related to CIN were suppressed simultaneously in TMB-ultralow tumors.

## 3.6 | Influence of TP53 inactivation on the PI3K/Akt/mTOR pathway

To find variations in the expression of other genes related to tumorigenesis, we performed a volcano plot analysis. Seven genes were differentially expressed in TMB-ultralow tumors (Figure 4A), of which 3 (*ECRG4*, *OGN*, and *ATP1A2*) were likewise upregulated in the TCGA data set (Figure S9) and had moderate negative correlation with TP53 inactivation (Figure S10). To elucidate the relationship between TP53 inactivation and the PI3K/Akt/mTOR pathway, correlation analysis was performed on 134 genes upregulated during PI3K/mTOR activation.[13,15] One-third of these genes had moderate positive correlation with TP53 inactivation, whereas TMB-ultralow tumors showed downregulation of these genes, accompanied by upregulation of *ECRG4*, *OGN*, and *ATP1A2* (Figure 4B).

## 3.7 | Response prediction of ICI in TMB-ultralow tumors

A combination of TMB and GEP related to T cell activation can improve the prediction of responses to ICI.[17-19] To explore a candidate responder to ICI therapy and further characterize TMB-ultralow tumors, the distribution of the T cell-inflamed GEP signature was investigated. In our analysis, no significant correlation between this signature and TMB was observed, except for breast cancer, which showed a weak correlation (Figure 5A). TMB-ultralow samples accounted for 10.3% of all Q3 (≥75-percentile) and this rate was similar to that for TMB-high samples (8.75%; Figure 5B). TMB-ultralow

**FIGURE 2** Top 10 mutated genes in TMB-ultralow and TMB-low tumors. To extract genes with somatic mutations, 131 TMB-ultralow and 139 TMB-low tumors without driver mutations were selected. The cancer-related genes were determined in a previous report[7]



**FIGURE 3** Copy number variation (CNV) analysis in TMB-ultralow and TMB-low tumors. Samples were classified using TMB in each tissue. Gene symbols of oncogenes or tumor suppressor genes (TSG) with significant difference ($q < 1 \times 10^{-6}$) between TMB-ultralow and TMB-low tumors are described. TP53 inactivation score (TP53-IS) was calculated from the expression of 4 genes (*CDC20*, *PLK1*, *CENPA*, and *KIF2C*).[8] cnLOH, copy neutral loss of heterozygosity

samples in Q3 included breast and head and neck tumors that were undetectable in TMB-high tumors (Figure 5C). These results indicated that TMB-ultralow tumors contained T cell-inflamed GEP signatures, similar to that found in TMB-high tumors, indicating that this signature was independent of TMB.

## 4 | DISCUSSION

Integrative analysis of 2658 whole-cancer genomes has recently revealed the presence of various somatic alterations that were difficult to detect using WES.[6] The features of low-mutated tumors, which comprise the majority of solid tumors, are still poorly understood. We focused on TMB-ultralow tumors and attempted to characterize these samples using WES and GEP. To reduce the number of false-negative mutations in WES, mutation analysis is necessary to mitigate the concern for low tumor content.[38] In our analysis, low tumor cellularity influenced somatic mutations and CNV, whereas the frequency of mutation detection was almost constant for high tumor contents. Therefore, samples with low tumor cellularity (<0.3) were excluded from the mutation analysis.

Although no significant differences related to increasing age were observed between TMB-low and TMB-ultralow tumors, aging often coincided with accumulation of driver mutations (see Figure S4). The JCGA (not TCGA) data set was not sufficient to clarify the influence of mutation accumulation on malignancy based on a survival curve, therefore its ability to completely deny this relationship is limited. The impact of aging in TMB-ultralow tumors should be re-evaluated in further studies that include survival outcome.

Based on WES and GEP, TMB-low tumors without known driver alterations harbored *TP53* mutations (not driver) and exhibited TP53 inactivation correlated with CIN, a tendency that was not observed in TMB-ultralow tumors. Studies on cell culture, experimental animal models, and human cancers have shown that defective TP53 activation is associated with enhanced CIN.[8,39-42] Recent pan-cancer analysis suggested that copy number changes attributed to CIN are delayed events.[43] Therefore, TMB-ultralow tumors may be less prone to accumulate somatic mutations and CNV due to retained TP53 activation in early tumorigenesis.

TMB-ultralow tumors usually had less variation in all chromosomes including TP53-encoded chromosome 17 than did TMB-low tumors, whereas a few cases harbored a change in CNV gain and/
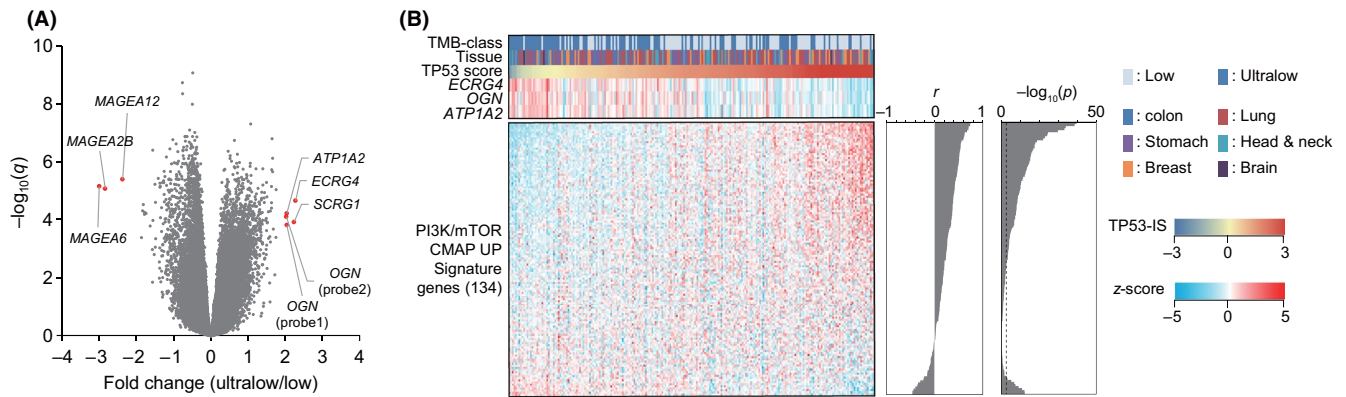
**FIGURE 4** Gene expression profiling (GEP) for extraction of TMB-ultralow-specific alteration in mRNA expression. A, Volcano plot showing the results of microarray analysis in TMB-ultralow and TMB-low tumors. Red dots represent significant differential genes ($q < .01$, fold change ≥ 2.0). B, Heatmap of PI3K/mTOR pathway-related genes, sorted by correlation profile with TP53 inactivation score (TP53-IS). PI3K/mTOR CMAP UP signature genes were previously defined as gene transcription signature of PI3K/Akt/mTOR in pan-cancer.[13-15] *ECRG4*, *OGN*, and *ATP1A2* genes extracted in Figure 4A were significantly downregulated in TCGA data set (see Figure S8). *OGN* expression was calculated from the average of *OGN* probes 1 and 2. The dashed line represents $P = .01$
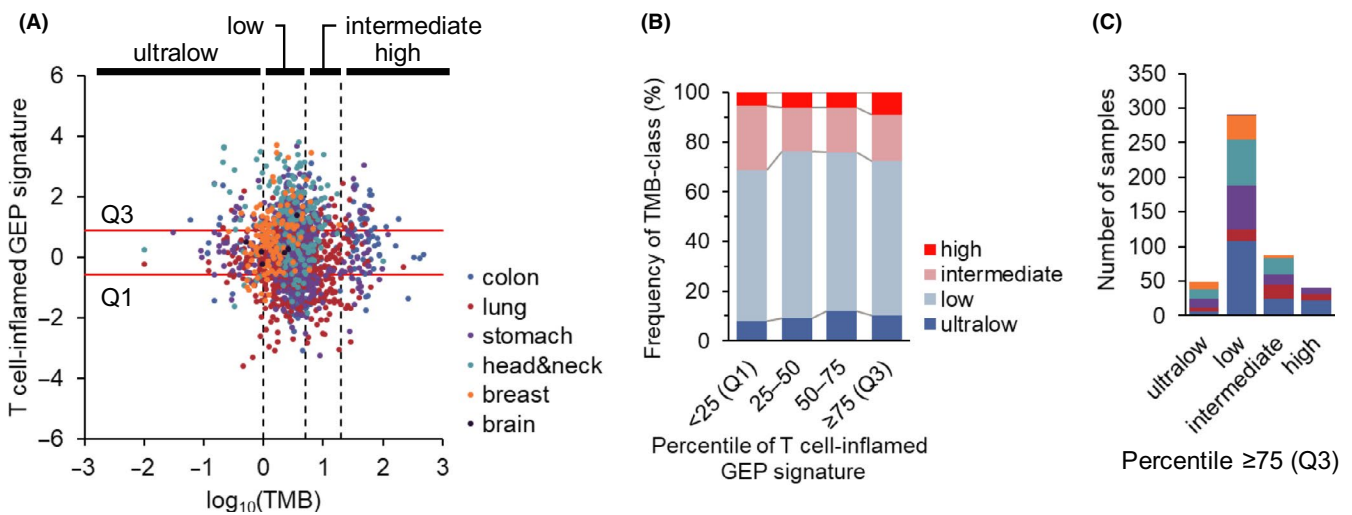


**FIGURE 5** Relationship between T cell-inflamed gene expression profile (GEP) signature and tumor mutational burden (TMB). A, Scatter plot of T cell-inflamed GEP signature in each tissue. This signature was calculated using mRNA expression of 18 genes that were determined in previous reports.[17-19] The signature in breast cancer showed weak correlation with TMB ($r = 0.26$, $P = 3.91 \times 10^{-3}$). In this analysis, represented data include TMB-intermediate and TMB-high tumors that met the criterion (tumor cellularity ≥ 0.3). B, Frequency of TMB class in each percentile. C, Distribution of tumor type in Q3 interval (≥75 percentile). Color coding of tumor type corresponds to Figure 5A

or loss (see Figure 3 and Figure S1D). Recent pan-cancer analysis revealed that chromothripsis-(like) events often occur in tumors with low driver mutations.[44] The large chromosomal aberration in TMB-ultralow tumors without driver alterations may be due to this event. Conversely, latest mutational analyses showed that composite mutations led to tumorigenesis.[45,46] TMB-ultralow tumors may harbor multiple alterations with weak pathogenicity as somatic or germline mutations, resulting in the development of tumors without driver mutations. These findings implied that TMB-ultralow tumors may include not just populations with suppressed TP53 inactivation but also other tumors via another tumorigenic process. To further understand TMB-ultralow tumors, integration of our analysis and whole genome sequencing, including germline mutation analysis, is an issue for the future.

We identified 3 upregulated genes (*ECRG4*, *OGN*, and *ATP1A2*) that were inversely correlated with TP53 inactivation in TMB-ultralow tumors, and their expression levels in TCGA showed the same trend. Upregulation of *ECRG4* in esophageal carcinoma increased TP53 protein expression, resulting in inhibition of cancer cell migration and invasion[47], therefore *ECRG4* may suppress TP53 inactivation in TMB-ultralow tumors. OGN inhibited the PI3K/Akt/mTOR pathway in breast and colorectal cancers as a tumor suppressor gene,[48,49] and the genes related to this pathway were highly expressed in *OGN*-upregulated tumors with ultralow TMB. Although the molecular function of *ATP1A2* in cancer remains unclear, this gene is affected in mouse muscle lacking mTOR activity,[50] implying an association with the PI3K/Akt/mTOR pathway. Therefore, we concluded that the PI3K/Akt/mTOR pathway

was suppressed, along with TP53 inactivation, in TMB-ultralow tumors.

The prediction of responses to ICI using gene expression focuses on tumors harboring some mutations (TMB ≥ 10 mutation/Mb),[17-19] although most solid tumors were categorized as TMB low. Our analysis revealed that the expression of the T cell-inflamed GEP signature was increased in low-mutated tumors. This score was independent of TP53 inactivation, CIN, and the PI3K/Akt/mTOR pathway (data not shown). Notably, populations showing high signatures (≥75 percentile) were TMB-low and TMB-ultralow including breast and head and neck tumors that show the high T cell-inflamed phenotype without high TMB.[51] Patients with low-mutated tumors of head and neck cancers included a few responders to ICI.[52] The population with high expression of the T cell-inflamed GEP signature may include candidate ICI responders, despite low mutation numbers.

In conclusion, the present study characterized TMB-ultralow and TMB-low tumors in the JCGA data set, which is composed of 5020 primary solid tumors. Detection rates of driver mutations and CNV were deceased in TMB-ultralow tumors, a trend that was observed in TCGA data set. Furthermore, several tumorigenesis-related pathways were not featured compared with those for TMB-low tumors. GEP suggested that low mutators include candidate ICI responders. Our analysis focusing on TMB-ultralow tumors can provide insight into tumors with low numbers of somatic alterations that are hardly detected solely by WES.

## CONFLICT OF INTEREST
The authors declare no conflicts of interest for this article.

## ORCID
*Keiichi Hatakeyama* https://orcid.org/0000-0001-6000-5899
*Takeshi Nagashima* https://orcid.org/0000-0002-7261-9908
*Keiichi Ohshima* https://orcid.org/0000-0002-2882-7566
*Akane Naruoka* https://orcid.org/0000-0002-6401-0318
*Yasuto Akiyama* https://orcid.org/0000-0003-3813-7726

## REFERENCES
1. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500:415-421.
2. Roberts SA, Gordenin DA. Hypermutation in human cancer genomes: footprints and mechanisms. *Nat Rev Cancer*. 2014;14:786-800.
3. Alexandrov LB, Kim J, Haradhvala NJ, et al. The repertoire of mutational signatures in human cancer. *Nature*. 2020;578:94-101.
4. Yarchoan M, Hopkins A, Jaffee EM. Tumor mutational burden and response rate to PD-1 inhibition. *N Engl J Med*. 2017;377:2500-2501.
5. Hatakeyama K, Nagashima T, Ohshima K, et al. Mutational burden and signatures in 4000 Japanese cancers provide insights into tumorigenesis and response to therapy. *Cancer Sci*. 2019;110:2620-2628.
6. Consortium ITP-CAoWG. Pan-cancer analysis of whole genomes. *Nature*. 2020;578:82-93.
7. Nagashima T, Yamaguchi K, Urakami K, et al. Japanese version of The Cancer Genome Atlas, JCGA, established using fresh frozen tumors obtained from 5143 cancer patients. *Cancer Sci*. 2020;111:687-699.
8. Donehower LA, Soussi T, Korkut A, et al. Integrated analysis of TP53 gene and pathway alterations in the cancer genome atlas. *Cell Rep*. 2019;28(5):1370-1384.e5.
9. Carter SL, Eklund AC, Kohane IS, Harris LN, Szallasi Z. A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nat Genet*. 2006;38:1043-1048.
10. Lee K, Kim JH, Kwon H. The actin-related protein BAF53 is essential for chromosomal subdomain integrity. *Mol Cells*. 2015;38:789-795.
11. Teixeira VH, Pipinikas CP, Pennycuick A, et al. Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nat Med*. 2019;25:517-525.
12. Saghafinia S, Mina M, Riggi N, Hanahan D, Ciriello G. Pan-cancer landscape of aberrant DNA methylation across human tumors. *Cell Rep*. 2018;25(4):1066-1080.e8.
13. Creighton CJ, Fu X, Hennessy BT, et al. Proteomic and transcriptomic profiling reveals a link between the PI3K pathway and lower estrogen-receptor (ER) levels and activity in ER+ breast cancer. *Breast Cancer Res*. 2010;12:R40.
14. Duan Q, Flynn C, Niepel M, et al. LINCS canvas browser: interactive web app to query, browse and interrogate LINCS L1000 gene expression signatures. *Nucleic Acids Res*. 2014;42:W449-W460.
15. Zhang Y, Kwok-Shing Ng P, Kucherlapati M, et al. A Pan-cancer proteogenomic atlas of PI3K/AKT/mTOR pathway alterations. *Cancer Cell*. 2017;31(6):820-832.e3.
16. Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity*. 2013;39:782-795.
17. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-gamma-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest*. 2017;127:2930-2940.
18. Cristescu R, Mogg R, Ayers M, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science*. 2018;362(6411):eaar3593.
19. Ott PA, Bang YJ, Piha-Paul SA, et al. T-cell-inflamed gene-expression profile, programmed death ligand 1 expression, and tumor mutational burden predict efficacy in patients treated with pembrolizumab across 20 cancers: KEYNOTE-028. *J Clin Oncol*. 2019;37:318-327.
20. Nagashima T, Shimoda Y, Tanabe T, et al. Optimizing an ion semiconductor sequencing data analysis method to identify somatic mutations in the genomes of cancer cells in clinical tissue samples. *Biomed Res*. 2016;37:359-366.
21. Shimoda Y, Nagashima T, Urakami K, et al. Integrated next-generation sequencing analysis of whole exome and 409 cancer-related genes. *Biomed Res*. 2016;37:367-379.
22. Urakami K, Shimoda Y, Ohshima K, et al. Next generation sequencing approach for detecting 491 fusion genes from human cancer. *Biomed Res*. 2016;37:51-62.
23. Ohshima K, Hatakeyama K, Nagashima T, et al. Integrated analysis of gene expression and copy number identified potential cancer driver genes with amplification-dependent overexpression in 1,454 solid tumors. *Sci Rep*. 2017;7:641.
24. Shen R, Seshan VE. FACETS: allele-specific copy number and clonal heterogeneity analysis tool for high-throughput DNA sequencing. *Nucleic Acids Res*. 2016;44:e131.

25. Favero F, Joshi T, Marquard AM, et al. Sequenza: allele-specific copy number and mutation profiles from tumor sequencing data. *Ann Oncol*. 2015;26:64-70.

26. Zhang Z, Hao K. SAAS-CNV: a joint segmentation approach on aggregated and allele specific signals for the identification of somatic copy number alterations with next-generation sequencing data. *PLoS Comput Biol*. 2015;11:e1004618.

27. Kim HY, Choi JW, Lee JY, Kong G. Gene-based comparative analysis of tools for estimating copy number alterations using whole-exome sequencing data. *Oncotarget*. 2017;8:27277-27285.

28. Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487:330-337.

29. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490:61-70.

30. Cancer Genome Atlas Network. Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature*. 2015;517:576-582.

31. Cancer Genome Atlas Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature*. 2012;489:519-525.

32. Cancer Genome Atlas Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*. 2014;513:202-209.

33. Cancer Genome Atlas Network. Comprehensive molecular profiling of lung adenocarcinoma. *Nature*. 2014;511:543-550.

34. Bozic I, Antal T, Ohtsuki H, et al. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci USA*. 2010;107:18545-18550.

35. Iranzo J, Martincorena I, Koonin EV. Cancer-mutation network and the number and specificity of driver mutations. *Proc Natl Acad Sci USA*. 2018;115:E6010-E6019.

36. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet*. 2003;34:267-273.

37. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102:15545-15550.

38. Hatakeyama K, Nagashima T, Urakami K, et al. Tumor mutational burden analysis of 2,000 Japanese cancer genomes using whole exome and targeted gene panel sequencing. *Biomed Res*. 2018;39:159-167.

39. Smith ML, Fornace AJ Jr. Genomic instability and the role of p53 mutations in cancer cells. *Curr Opin Oncol*. 1995;7:69-75.

40. Tainsky MA, Bischoff FZ, Strong LC. Genomic instability due to germline p53 mutations drives preneoplastic progression toward cancer in human cells. *Cancer Metastasis Rev*. 1995;14:43-48.

41. Lu X, Nguyen TA, Moon SH, Darlington Y, Sommer M, Donehower LA. The type 2C phosphatase Wip1: an oncogenic regulator of tumor suppressor and DNA damage response pathways. *Cancer Metastasis Rev*. 2008;27:123-135.

42. Wasylishen AR, Lozano G. Attenuating the p53 pathway in human cancers: many means to the same end. *Cold Spring Harb Perspect Med*. 2016;6(8):a026211.

43. Gerstung M, Jolly C, Leshchiner I, et al. The evolutionary history of 2,658 cancers. *Nature*. 2020;578:122-128.

44. Cortes-Ciriano I, Lee JJ, Xi R, et al. Comprehensive analysis of chromothripsis in 2,658 human cancers using whole-genome sequencing. *Nat Genet*. 2020;52:331-341.

45. Gorelick AN, Sanchez-Rivera FJ, Cai Y, et al. Phase and context shape the function of composite oncogenic mutations. *Nature*. 2020;582:100-103.

46. Saito Y, Koya J, Araki M, et al. Landscape and function of multiple mutations within individual oncogenes. *Nature*. 2020;582:95-99.

47. Li L, Zhang C, Li X, Lu S, Zhou Y. The candidate tumor suppressor gene ECRG4 inhibits cancer cells migration and invasion in esophageal carcinoma. *J Exp Clin Cancer Res*. 2010;29:133.

48. Hu X, Li YQ, Li QG, Ma YL, Peng JJ, Cai SJ. Osteoglycin-induced VEGF inhibition enhances T lymphocytes infiltrating in colorectal cancer. *EBioMedicine*. 2018;34:35-45.

49. Xu T, Zhang R, Dong M, et al. Osteoglycin (OGN) inhibits cell proliferation and invasiveness in breast cancer via PI3K/Akt/mTOR signaling pathway. *Onco Targets Ther*. 2019;12:10639-10650.

50. Kleinert M, Parker BL, Fritzen AM, et al. Mammalian target of rapamycin complex 2 regulates muscle glucose uptake during exercise in mice. *J Physiol*. 2017;595:4845-4855.

51. Hegde PS, Chen DS. Top 10 Challenges in cancer immunotherapy. *Immunity*. 2020;52:17-35.

52. Hanna GJ, Lizotte P, Cavanaugh M, et al. Frameshift events predict anti-PD-1/L1 response in head and neck cancer. *JCI Insight*. 2018;3(4):e98811.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.