Significant Prognostic Features and Patterns of Somatic *TP53* Mutations in Human Cancers

Cancer Informatics
1–10

DOI: 10.1177/1176935117691267

Wensheng Zhang[1], Andrea Edwards[1], Erik K Flemington[2] and Kun Zhang[1]

[1]Department of Computer Science and Bioinformatics Facility of Xavier RCMI Center for Cancer Research, Xavier University of Louisiana, New Orleans, LA, USA. [2]Tulane Cancer Center, School of Medicine, Tulane University, New Orleans, LA, USA.

ABSTRACT: *TP53* is the most frequently altered gene in human cancers. Numerous retrospective studies have related its mutation and abnormal *p53* protein expression to poor patient survival. Nonetheless, the clinical significance of *TP53* (*p53*) status has been a controversial issue. In this work, we aimed to characterize *TP53* somatic mutations in tumor cells across multiple cancer types, primarily focusing on several less investigated features of the mutation spectra, and determine their prognostic implications. We performed an integrative study on the clinically annotated genomic data released by The Cancer Genome Atlas. Standard statistical methods, such as the Cox proportional hazards model and logistic regression, were used. This study resulted in several novel findings. They include the following: (1) similar to previously reported cases in breast cancer, the mutations in exons 1 to 4 of *TP53* were more lethal than those in exons 5 to 9 for the patients with lung adenocarcinomas; (2) *TP53* mutants tended to be negatively selected in mammalian evolution, but the evolutionary conservation had various clinical implications for different cancers; (3) conserved correlation patterns (ie, consistent co-occurrence or consistent mutual exclusivity) between *TP53* mutations and the alterations in several other cancer genes (ie, *PIK3CA, PTEN, KRAS, APC, CDKN2A*, and *ATM*) were present in several cancers in which prognosis was associated with *TP53* status and/or the mutational characteristics; (4) among *TP53*-mutated tumors, the total mutation burden in other driver genes was a predictive signature ($P < .05$, false discovery rate <0.11) for better patient survival outcome in several cancer types, including glioblastoma multiforme. Among these findings, the fourth is of special significance as it suggested the potential existence of epistatic interaction effects among the mutations in different cancer driver genes on clinical outcomes.

KEYWORDS: Cancer, *TP53*, somatic mutation, clinical outcomes, prognosis

RECEIVED: October 10, 2016. **ACCEPTED:** January 2, 2017.

PEER REVIEW: Four peer reviewers contributed to the peer review report. Reviewers' reports totaled 1110 words, excluding any confidential comments to the academic editor.

TYPE: Original Research

FUNDING: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This publication was made possible by funding from the NIH NIMHD-RCMI grant 2G12MD007595, the DOD ARO grant W911NF-15-1-0510, and the Louisiana Cancer Research Consortium (LCRC). EKF is funded by NIH R01AI101046 and R01AI106676. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the NIH, DOD, or LCRC.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Kun Zhang, Department of Computer Science and Bioinformatics Facility of Xavier RCMI Center for Cancer Research, Xavier University of Louisiana, 1 Drexel Drive, New Orleans, LA 70125, USA. Email: kzhang@xula.edu

Introduction

The human *TP53* tumor suppressor gene is located on the short arm of chromosome 17, encompassing 11 exons that give a ~2-kb messenger RNA through transcription.[1] The *p53* protein, encoded by *TP53*, responds to diverse cellular stresses to regulate the expression of target genes, thereby inducing cell cycle arrest, apoptosis, senescence, DNA repair, or changes in metabolism.[2] Loss or disruption of *p53* function due to a mutation can lead to uncontrolled cell proliferation and cancer.[3] Some *p53* mutants gain new functions, exhibit oncogenic properties, and exert a dominant negative effect by preventing wild-type (WT) *p53* from binding to the promoter of its target genes.[4] Aberrations in *TP53* could cause a burst of somatic mutations in tumor cells, disrupting the age-related accumulation patterns.[5]

TP53 is the most frequently altered gene in human cancers.[6] Most somatic *TP53* mutations are single-base substitutions distributed throughout exons 5 to 8.[7] Notably, about 20% of these mutations alter 1 of 3 codons (175, 248, or 273) of the 393 amino acids in *p53* protein.[7,8] The clinical significance of *TP53* (*p53*) status for patient outcome has been and continues to be a controversial topic of cancer research. Numerous retrospective studies have associated its mutation and abnormal *p53* protein expression with poor patient survival.[9] Such an association was demonstrated most by the relevant studies in breast, head and neck, hematopoietic, liver, and lymph node cancers. Nevertheless, for other cancer types, it still remains unclear whether *TP53* status and/or its mutational characteristics could serve as a potential biomarker for patient outcomes. For example, the publications disassociating *TP53* mutation from patient survival are largely as many as those relating *TP53* mutation to poor survival.[9]

Beyond the genotypes, more prognostic value of mutations in *TP53* may be hidden in their tumor-specific characteristics as well as the interaction with other genomic aberrations. Skaug et al[10] and Huang et al[11] showed that for patients with non–small cell lung cancer (NSCLC), mutations in exon 8 of the *TP53* gene were more fatal than those in exons 5 and 7. Molina-Vila et al[12] found that nondisruptive *p53* mutations, including in-frame deletions outside of the L2 and L3 loop domains and missense single-base substitutions, were associated with shorter survival in patients with NSCLC. Contrary to this finding in NSCLC, several studies on head and neck squamous cell carcinomas (HNSCs) showed that tumors containing mutations in the DNA-binding regions (L2, L3, and

Table 1. Summary of tumor samples and somatic TP53 mutations of 12 TCGA cancer types.

CANCER	TUMOR SAMPLES[a]	MUTANT RATIO[b]	VARIANT CLASS RATIO[c]			
			INDEL	MISSENSE	NONSENSE	SPLICE
BLCA	233	0.51	0.09	0.72	0.17	0.02
GBM	285	0.29	0.09	0.81	0.04	0.06
HNSC	504	0.71	0.18	0.55	0.16	0.11
LUAD	488	0.62	0.12	0.57	0.19	0.12
LUSC	178	0.79	0.12	0.68	0.14	0.07
BRCA	967	0.32	0.19	0.6	0.14	0.07
OV	370	0.83	0.13	0.66	0.11	0.1
UCEC	248	0.28	0.12	0.77	0.09	0.01
COAD	216	0.56	0.05	0.78	0.12	0.05
ESCA	171	0.74	0	0.73	0.18	0.09
LIHC	197	0.27	0	0.78	0.12	0.1
STAD	288	0.45	0.18	0.64	0.12	0.05

Abbreviations: BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; STAD, stomach adenocarcinoma; UCEC, uterine corpus endometrial carcinoma.
[a]Tumor samples with both clinical and somatic mutation data.
[b]Proportion of tumors with somatic *TP53* mutation burden.
[c]Proportions of different variant categories among the total *TP53* mutations.

loop-sheet-helix domains) of *TP53* led to a significantly worse prognosis and response to radiotherapy than tumors outside those regions.[13,14] A recent publication reported that the combination of *TP53* mutation and loss of chromosome 3p associated with a remarkable decrease in short-period survival rates for patients with HNSC.[15]

In this article, to characterize the *TP53* somatic mutations in tumor cells of multiple cancer types and to determine their prognostic implications, we performed an integrative study on the clinically annotated genomic data published by The Cancer Genome Atlas (TCGA). Our analysis was primarily focused on several less investigated features of the mutation spectra, such as the evolutionary selection of mutant alleles in *TP53* during mammalian evolution, with an extension to the pan-cancer patterns of exclusivity and co-occurrence relationships between *TP53* mutations and the alterations in other cancer driver genes.

Materials and Methods

Data

Clinical and somatic data were downloaded from TCGA database (http://cancergenome.nih.gov/) on April 24, 2015. For each cancer type, TCGA collected multiple (version) somatic mutation data sets. Those data, contributed by different institutes, were generated using various sequencing platforms, somatic mutation–calling algorithms, and computational tools. Except for ovarian serous cystadenocarcinoma (OV), we chose 1 representative data set for each cancer type according to the following criteria. First, the selected data set contains the largest number of tumor samples (or patients). Second, if 2 or more data sets are of the same size, we chose the one in which the mutations were measured by the IlluminaGA DNASeq platform and were called by the latest automated system. Finally, if the decision could not be reached by the previous 2 steps, we selected the data set provided by the UCSC Genome Browser. For OV, we used the data sets from Massachusetts Institute of Technology and Washington University in St Louis. The basic information of the used somatic and clinical data sets is summarized in Supplementary Table 1. Synonymous mutations and those under the categories of "intron" and "rna" were excluded from further analysis.

Basewise conservation

A table containing PhyloP scores (Placental Mammal Basewise Conservation by PhyloP) of *TP53* gene calculated by phylogenetic methods[16] and the PHAST package (http://compgen.cshl.edu/phast/) was retrieved from the UCSC Genome Browser database. In this table, sites predicted to be conserved are assigned positive scores, whereas sites predicted to be fast-evolving are assigned negative scores. The absolute values of the scores represent $-\log_{10}(P \text{ values})$ under a null hypothesis of neutral evolution. For a mutation under several "indel" categories, the score of the base at the start position was used in the comparison analysis.

Statistical analysis

Survival analysis was performed using R package "survival." The Kaplan-Meier survival curves were created by the function "*survfit()*." *P* values for the effect of *TP53* status (or a mutation feature) on patient survival time were calculated by the function "*coxph()*." Benjamini-Hochberg false discovery rate (FDR) was calculated using the function "*p.adjust()*" in the R package "*stats*." The mutational association (or relationship) between *TP53* and another gene was measured by the Yule *phi* coefficient (a Pearson correlation applied to dichotomous data)[17] between the numbered genotypes (1 and 0 were assigned to mutant and WT, respectively). The statistical significance was further evaluated with the *P* value calculated using a logistic regression model, in which *TP53* genotype and the genotype of the paired gene were the independent variable and dependent variable, respectively. A co-occurrence (or mutual exclusivity) relationship was determined by $P < .05$ and $r > 0$ (or $r < 0$). The analysis was performed using the "*corr()*" and "*glm()*" function in the R package "stats." The obtained results (ie, *P* values) are similar to the Fisher test that had been used by Kandoth et al.[18] Hierarchical clustering analysis was conducted using the "*hclust()*" function in the R package "stats." The detailed implementation is described in the "Results" section and the legend of Figure 4.

Results

Among the 33 cancer types with clinically annotated multiomic data available at TCGA database by April 24, 2015, 12 were studied in this work by considering the genetic diversity of patients and the prevalence of somatic *TP53* mutations in these tumors. Each of the selected cancer types had at least 14 patients from a minority population (ie, black American or Asian) besides the dominant white Americans, and the ratio of samples with *TP53* nonsynonymous mutations was more than 25% (Table 1). The studied cancer types included bladder urothelial carcinoma (BLCA), glioblastoma multiforme (GBM), HNSC, lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), breast invasive carcinoma (BRCA), ovarian serous cystadenocarcinoma (OV), uterine corpus endometrial carcinoma (UCEC), colon adenocarcinoma (COAD), esophageal carcinoma, liver hepatocellular carcinoma, and stomach adenocarcinoma (STAD). The sample sizes of those cancer types ranged from 171 to 967.[19]

TP53 status

We performed a series of survival analyses on the 12 cancers using a Cox proportional hazards (PH) regression model. In the modeling, *TP53* (*p53*) status (ie, WT versus mutant) was treated as the stratification factor of primary interest and the patient age at the initial clinical date was included as a covariate. The results demonstrated association between *TP53* mutation and overall poor patient survival in 4 cancer types, namely, HNSC, LUAD, BRCA, and COAD (Supplementary Figure 1). We also noticed that although the *P* value calculated from the hypothesis test was larger than 0.05 for UCEC, the long-term (more than 50 months) survival advantage of *TP53* WT patients was apparent in this cancer type. The mechanism underlying this dilemma could be that the statistical model does not fit the cases with unparallel survival curves. In other words, the primary assumption of a Cox PH model, i.e., the survival curves for 2 strata must have hazard functions that are proportional over time, was apparently not satisfied in the UCEC data. Moreover, it is worth noting that although most of these observed associations were only moderately significant, they were largely consistent with the previously published results (see "Discussion" section) and, therefore, could be regarded as confirmatory findings.

TP53 mutations outside hot-spot regions

We depicted the physical distributions of *TP53* mutations over the coding regions in Supplementary Figure 2. The bar plots clearly showed that the mutations are spread beyond the classical "hot-spot" regions (ie, exons 5-8, which encode DNA-binding domains[20]). In several cancers, including HNSC, LUAD, LUSC, BRCA, and STAD, the presence of nonsynonymous mutations in exon 4 was not rare. Occasionally, mutations also occurred in exons 2, 3, and 9 in these cancer types. The amino acids encoded by exons 2 to 4 constitute the transactivation domains of *p53* and the proline repeat domain that binds directly to the transcriptional coactivator *p300* and allosterically controls DNA-dependent acetylation of *p53*.[21,22] Powell et al[23] found that mutations within exon 4 were particularly associated with poor prognosis in breast cancer. Hereby, we first partitioned the *TP53*-mutated patients within a specific cancer type into 2 groups (ie, E2-4 and E5-9) based on the presence or absence of a mutation burden in exons 2 to 4 and then compared their survival curves with the *TP53* WT counterpart. The result showed that for patients with BRCA and LUAD, the mutations in exons 2 to 4 were more lethal than those in other exons (Figure 1). In BRCA, the *P* values were less than 2×10^{-5} (Benjamini-Hochberg FDR $< 2 \times 10^{-4}$) in the comparison of E2-4 vs WT and larger than 0.05 in the comparison of E5-9 vs WT. In LUAD, no patient with somatic mutations in exons 2 to 4 lived longer than 50 months after the initial clinical date.

Natural selection of mutant alleles in TP53 during mammalian evolution

p53 is highly conserved from placozoans to man, in the structure, function, and interaction with other proteins.[1,24] Cancer appears to be most common in mammals among vertebrates, and the basic cancer-causing mechanisms are similar in mammalian species.[25-27] Therefore, the *TP53* mutations observed in

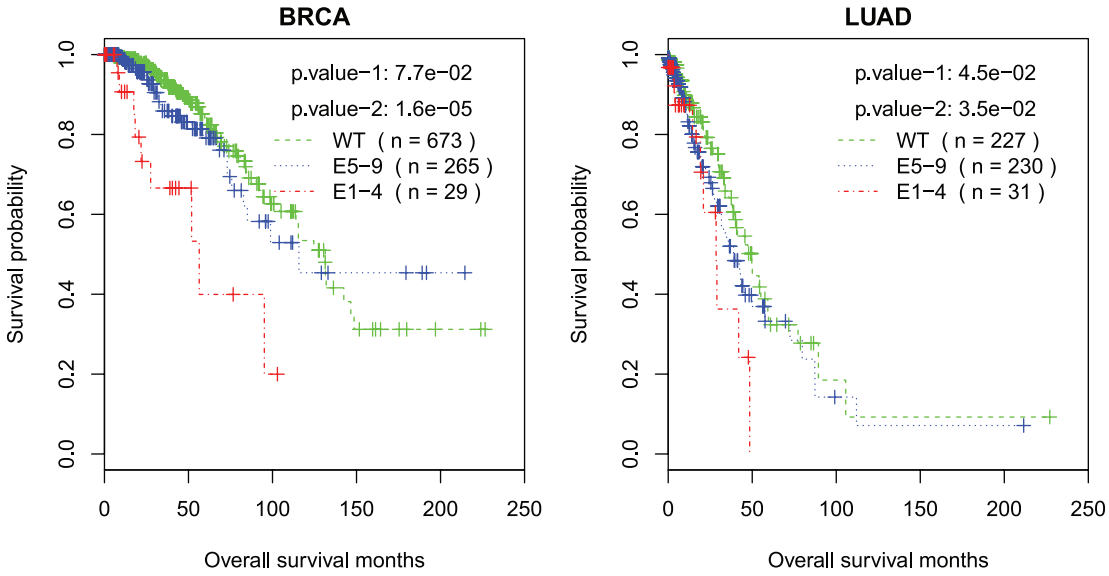

Figure 1. Lethality of *TP53* mutations in exons 2 to 4 for patients with BRCA and LUAD. Green: WT group, tumors without *TP53* mutation burden. Red: E1-4 group, tumors with *TP53* mutation(s) within exons 1 to 4. Blue: E5-9 group, tumors with *TP53* mutation(s) outside exons 1 to 4. *P* value-1 (or *P* value-2) is calculated using Cox proportional hazards model, corresponding to the difference between the groups E5-9 (or E1-4) and WT. BRCA indicates breast invasive carcinoma; LUAD, lung adenocarcinoma; WT, wild type.

human tumors could be subject to the natural selection mechanisms during mammalian evolution.

We evaluated the conservation of human *TP53* gene sequence by the basewise PhyloP scores (see "Materials and Methods" section). The DNA bases where mutants were selected against were determined by the criterion of PhyloP score less than 1.301, which corresponded to $P < .05$ in the test with neutral evolution as the null hypothesis. We found that cancer-related *TP53* variants tended to be negatively selected in the evolution of mammals. This conclusion was drawn from the following observations as shown in Figure 2A and 2B. First, across the *TP53* gene sequence, the peaks of smooth-spline curve of PhyloP scores were positive and overlapped with all the 10 encoding exons (exons 2-11). Second, of the DNA bases in which the somatic mutations arose, those with PhyloP scores more than 1.301 accounted for a large proportion consistently across the 12 cancer types. The ratio was ~40% higher than that of the entire base set of exons 4 to 9, in which more than 99% of the somatic mutations in tumors were located.

We were particularly interested in the potential prognostic implications of the evolutionary conservation of somatic *TP53* mutants. As such, based on the presence or absence of a negatively selected *TP53* variant in mammalian evolution, we partitioned the *TP53*-mutated samples of each cancer type into 2 groups (*Con* and *nCon*) and then performed survival analysis with this classification as the stratification factor of primary interest. The results showed that the class effect on patient outcome was moderately significant ($P < .05$, Benjamini-Hochberg FDR <0.15) only in BLCA and GBM (Figure 3). In the former, the survival curves stratified clearly and *Con* patients had better prognosis. In the latter, the *nCon* patients lived longer.

Specific mutual-exclusivity and co-occurrence patterns

Mutual exclusivity and co-occurrence are 2 important characteristics of somatic mutation spectra in cancers. In this study, we identified 95 (or 97) statistically significant mutual-exclusivity (or co-occurrence) relationships between *TP53* and other 117 cancer driver genes. Those genes were part of the 291 "high-confidence" genes pinpointed by Tamborero et al[19] through the Pearson correlation and logistic regression analysis. The distribution of these identified relationships was skewed across various cancer types, as most of the mutual-exclusivity relationships occurred in COAD, whereas most of the co-occurrence relationships concentrated in LUAD. We further organized these relationships by a hierarchical clustering algorithm and visualized the patterns with a heat map, in which the involved genes were partitioned into 6 groups (Figure 4). The groups highlighted in gray and cyan contained the 6 genes we focused on in the subsequent analysis.

PIK3CA, PTEN, KRAS, APC, CDKN2A, and *ATM* are among the most common cancer driver genes.[19,28] We found that there existed a "conserved" pattern (Figure 5) in the mutual-exclusivity and co-occurrence relationships between *TP53* mutations and the alterations of these genes. That is, the mutational associations for a specific gene pair were always in the same category, co-occurrence or mutual-exclusivity, across the cancer types. For example, consistently significant mutual-exclusivity relationship between PIK3CA and *TP53* was shown in 4 cancers (BRCA, STAD, COAD, and HNSC), but none of those cancer types held a significant co-occurrence relationship between these 2 genes. We

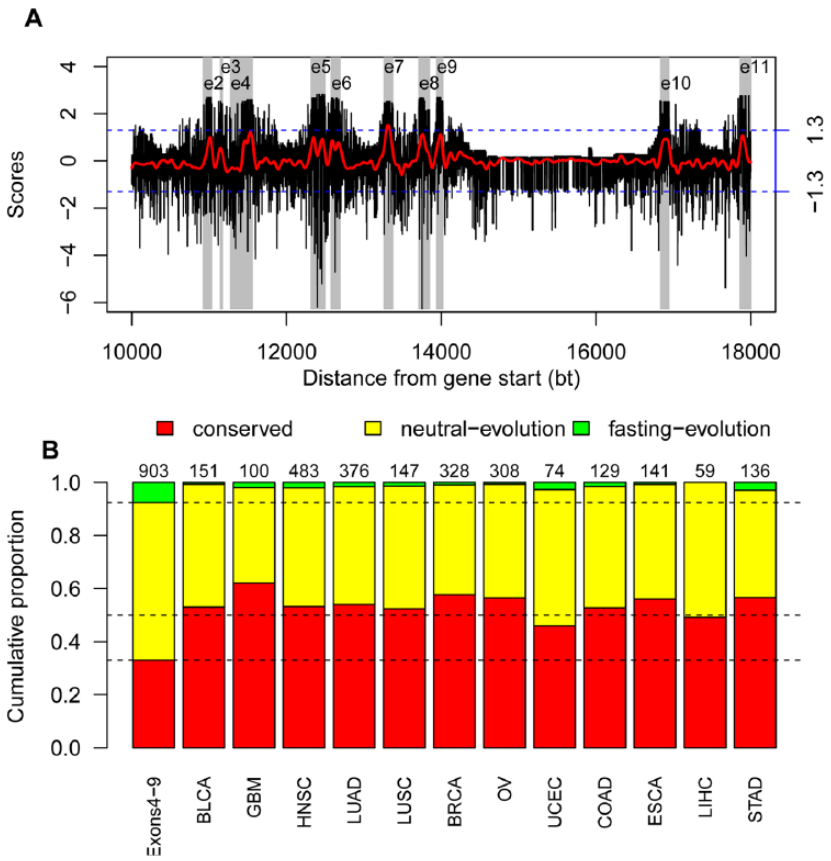

Figure 2. The evolutionary conservation of somatic *TP53* mutation sites. (A) The distribution of PhyloP scores (Placental Mammal Basewise Conservation by PhyloP) over the *TP53* gene sequence. The exon regions (e2, e3, . . ., e11) are marked by a gray background. The red smooth curve was generated by the "smooth.spline()" function in R package "stats." (B) The proportions of neutral-evolution, conserved, and fasting-evolution sites (bases) in the entire sequences of exons 4 to 9 and in the sites where somatic mutations occur over 12 cancer types. The total base number of exons 4 to 9 of *TP53* genes is printed at the top of the first column. The total number of nonsynonymous *TP53* mutations in the cancer samples of each cohort is printed at the top of the corresponding cancer-type column. Additional notes: (1) in the context of this study, "evolutionary conservation" of a DNA base carries the same meaning as "negative selection to a variant" in the site, and we use these 2 terms interchangeably, and (2) we use the PhyloP scores to divide the sequence bases into 3 groups (ie, "conserved," "neutral-evolution," and "fasting-evolution") rather than to identify the sites that significantly deviated from neutral evolution. As a result, two less strict cutoffs (ie, 1.301 and –1.301, corresponding to $P < .05$) are adopted. BLCA indicates bladder urothelial carcinoma; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; OV, ovarian serous cystadenocarcinoma; STAD, stomach adenocarcinoma; UCEC, uterine corpus endometrial carcinoma.

further noted that these significant relationships were primarily present in several cancers where patient outcome was associated with *TP53* status and/or the positional and evolutionary characteristics of the mutations, as shown in Supplementary Figure 1, Figure 1, and Figure 2. It is worth noting that these results obtained by a cherry-picking way will be helpful for a more accurate understanding of mutual exclusivity and co-occurrence of genomic alterations in cancer (see "Discussion" section).

A general co-occurrence pattern

It is well known that most cancers, including those with mutated *TP53*, are driven by multiple genetic mutations. Therefore, a general TP53-involved co-occurrence mutation pattern (or model) can be expressed by $[M_{p53}, M_{other}]$, where M_{p53} is the genotype (ie, mutant and WT) of *TP53* gene in a tumor and M_{other} is the number of mutations on other cancer driver genes. We found an interesting prognostic implication hidden in this simple mutation model. That is, M_{other} could serve as a predictive signature for favorable clinical outcome of patients ($P < .05$; Benjamini-Hochberg FDR <0.11) in the *TP53*-mutated groups of 4 cancers (ie, BLCA, GBM, LUAD, and OV). However, such an observation did not exist in any of the corresponding *TP53* WT groups. Table 2 presents the results of the survival analysis. In the used Cox PH model, a negative regression coefficient indicates that the hazard function decreases (or equivalently, survival time increases) as the quantity of the predictive variable increases.[29] Notably, this finding indicates that *TP53* mutation is also related to the prognosis of OV tumors, whereas the relevance cannot be detected by directly analyzing the association between *TP53* genotypes (and other mutational features) and the survival times of patients.

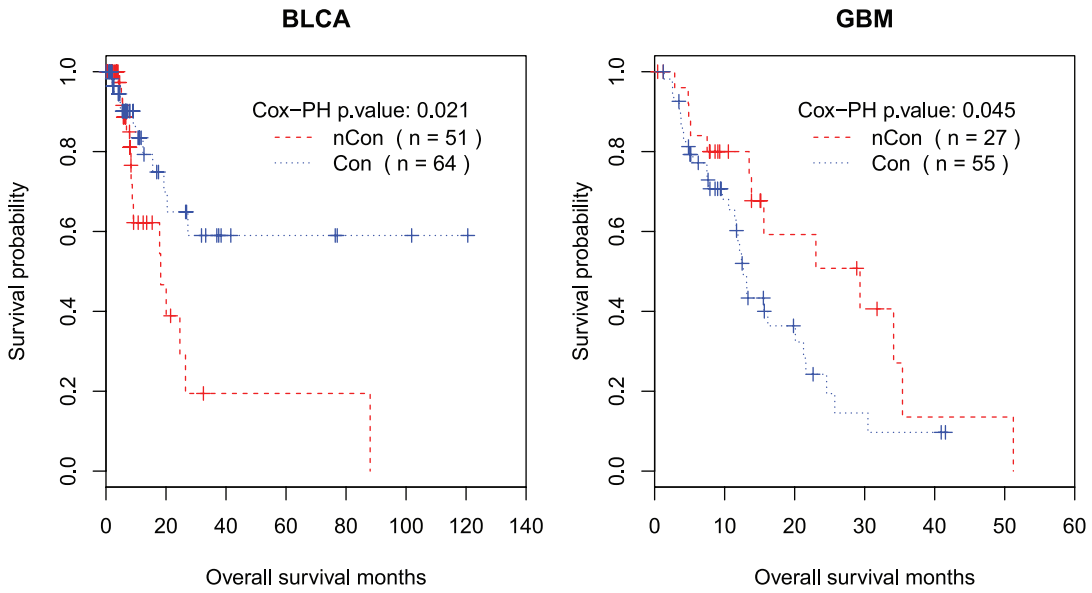

Figure 3. The prognostic implications of the evolutionary conservation of somatic *TP53* mutants in patients with BLCA and GBM. Blue: "*Con*" group in which tumors with a burden of *TP53* mutation(s) that aroused in the evolutionarily conserved sites (DNA bases). Red: "*nCon*" group in which tumors with a burden of *TP53* mutation(s) that aroused in evolutionarily neutral or fast-evolving sites (DNA bases) but not in evolutionarily conserved sites. BLCA indicates bladder urothelial carcinoma; GBM, glioblastoma multiforme.

Discussion

It has been proposed that a pivotal step toward personalized cancer medicine is the determination of molecular tumor subtypes (groups), especially those with potential applicability in clinical routines.[30,31] In this study, we found that the mutations in exons 2 to 4 of *TP53* gene defined a poor prognosis group in BRCA and LUAD. Similar results have been reported for BRCA[23] but not for LUAD, according to our knowledge. Meanwhile, we noticed that prevalence of mutations in these exon regions was an important characteristic of several cancers, including HNSC, LUAD, LUSC, BRCA, and STAD, where drinking and smoking are 2 common risk factors.[32–38] Due to this finding, an interesting question worth further study is whether alcohol and/or tobacco is a specific mutagen responsible for the alterations in exons 2 to 4 of *TP53* gene.

Our analysis shows that *TP53* variants tend to be negatively selected in the evolution of mammals. While being consistent with the existing knowledge that recurrent mutations in tumors can be differentiated from single mutations by the evolutionary conservation-based functional impact score,[39,40] this observation cannot be sufficiently explained by the survival disadvantage that is exerted on the carrier of the mutations (in a specific species) by the increased risk of cancer. The reason is that most cancer incidences could be postponed until the postreproductive portion of life spans across mammalian taxa.[41] As a result, the survival disadvantage of the individuals with a germline mutation burden in *TP53* gene is not equivalent to a lower fitness in evolution. In this regard, the negative selection of *TP53* mutants may involve reproduction-related mechanisms that could be interrupted by the loss or disruption of *p53* function in DNA repair. This hypothesis is supported by the varied clinical implications of a *TP53* mutation occurred in evolutionarily conserved DNA bases for different cancers (Figure 3). In particular, the favorable prognosis of the BLCA patients with mutations in conserved sites of *TP53* sequence suggests that the lower evolutionary fitness of *TP53* mutants cannot be simply attributed to the cancer-caused death.

Mutual exclusivity and co-occurrence of genomic alterations have been heavily studied in the past years.[18,42-46]. A proposed naive rule is that mutations in genes functioning in different pathways can occur in the same cancer, whereas those in genes functioning in the same pathway are rarely mutated in the same sample.[42,47] Nevertheless, previous studies also showed that certain combinatorial mutational patterns were inconsistent with this rule and demonstrated tissue-specific variations.[42] Our work showed that the rule was systematically violated when the analysis was focused on the specific relationships between *TP53* mutations and genetic alterations occurred in several other major cancer driver genes (ie, *PIK3CA, PTEN, KRAS, APC, CDKN2A*, and *ATM*). For example, the CDKN2A gene encodes the p16 and $p14^{ARF}$ proteins that inhibit the activity of *CDK4* and *CDK6* complex, thus blocking the transition between G and S phases in cell cycle.[8,48] Activation of *p53* results in the transcriptional upregulation of *CDKN1A* and increases the expression of *p21*, a universal inhibitor of cyclin-dependent kinases (*CDK*s).[8,49] However, as shown in Figure 5, this common involvement of *TP53* and *CDKN2A* in the CDK pathway, a critical component of cancer pathways, did not lead to the mutual exclusivity of their mutations

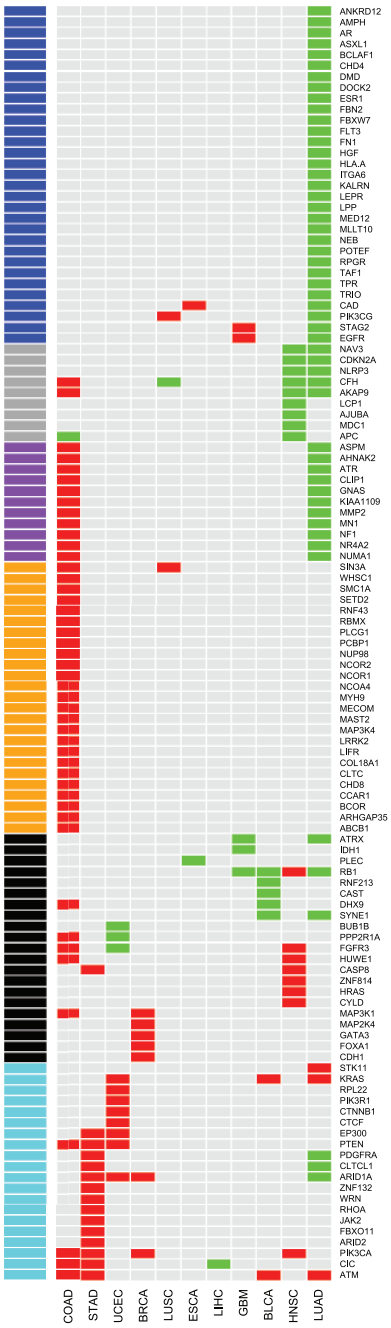

Figure 4. The co-occurrence or mutual-exclusivity relationships between somatic *TP53* mutations and the alterations in other 117 cancer driver genes over 11 cancer types. Ovarian serous cystadenocarcinoma (OV) was not included in this figure due to the lack of significant mutational relationships. The cell color, that is, green, red, or gray, indicates co-occurrence, mutual exclusivity, or the lack of association, respectively. The gene clusters, indicated by the color bar on the left side of the figure, are determined by a hierarchical clustering analysis (Manhattan distance and Ward method) with a sparse matrix (ie, M) as the input. In the matrix, rows and columns represent genes and cancer types, respectively. When the *i*th gene has a significant co-occurrence (or mutual exclusivity) relationship ($P < .05$) with *TP53* in somatic mutation for *j*th cancer, the element m_{ij} of M is 1 (or –1); otherwise, it is 0. BLCA indicates bladder urothelial carcinoma; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; STAD, stomach adenocarcinoma; UCEC, uterine corpus endometrial carcinoma.

in any cancer. Oppositely, the measure of their mutational co-occurrence was significant in HNSC and LUAD. Another example, *PIK3CA* gene plays roles in the PI3K/ART pathway,[50] which regulates cell proliferation and apoptosis in a different manner from the *p53* pathway, but, our results showed that the alterations in *PIK3CA* and *TP53* tended to occur in a mutually exclusive way within 4 cancers (ie, BRCA, STAT, COAD, and HNSC). A clinical characteristic shared by these cancers was that patient survival was associated with *TP53* status and/or the mutation features. Underlying this prognostic stratification are the intrinsic subtypes (within the same cancer) that are approximately determined by the mutations in *TP53* and the other driver genes such as *PIK3CA*.

In most cancer types, an average tumor sample contains at least half of hundreds of nonsynonymous somatic alterations, including a few cancer driver mutations that are fixed by conferring the recipient cells' fitness advantage and numerous passenger mutations that are fixed by the Muller ratchet and hitchhiking.[40,51] By definition, a passenger mutation has minor or ignorable impact on cell growth and proliferation. However, a recently emerging theory proposes that some passenger mutations could be deleterious to the host cells and their accumulation may build strength to alter cancer progression.[40,52] Our previous analyses of TCGA data suggested that some passenger mutations would exert significant impacts on the resistance of cancer cells to the treatments in ovarian carcinomas.[53] For example, the patients with 2 or multiple somatic mutations in the genes encoding lysosomal membrane proteins had significantly poorer prognosis in terms of survival months after the initial clinical dates. Significant interactive effects of cancer driver mutations and passenger mutations on the clinical outcome of patients was implied by another recent publication, which showed that high mutation number forecasted a remarkably favorable outcome in ovarian patients carrying mutations in *BRCA1* and *BRCA2* genes.[54] In this study, we found that among *TP53*-mutated tumors, the total mutation burden in other driver genes was a predictive signature for good patient survival in several cancers. This result is of special significance. First, it holds potential applicability in clinical routines in the relevant cancer, that is, BLCA, GBM, LUAD, and OV. Second, for the first time (to our knowledge), it suggested the potential existence of epistatic effects between the mutations in different cancer driver genes on clinical outcomes. Third, it suggests that *TP53* status may influence the sensitivity of tumors to the treatment therapies that kill cancer cells by inducing new somatic mutations.[52,55,56]

In summary, through an integrative analysis of the genomic and clinical data of 12 cancers generated by TCGA, we pinpointed a set of significant prognostic features and patterns of somatic *TP53* mutations. We further scrutinized the biological implications of our findings.

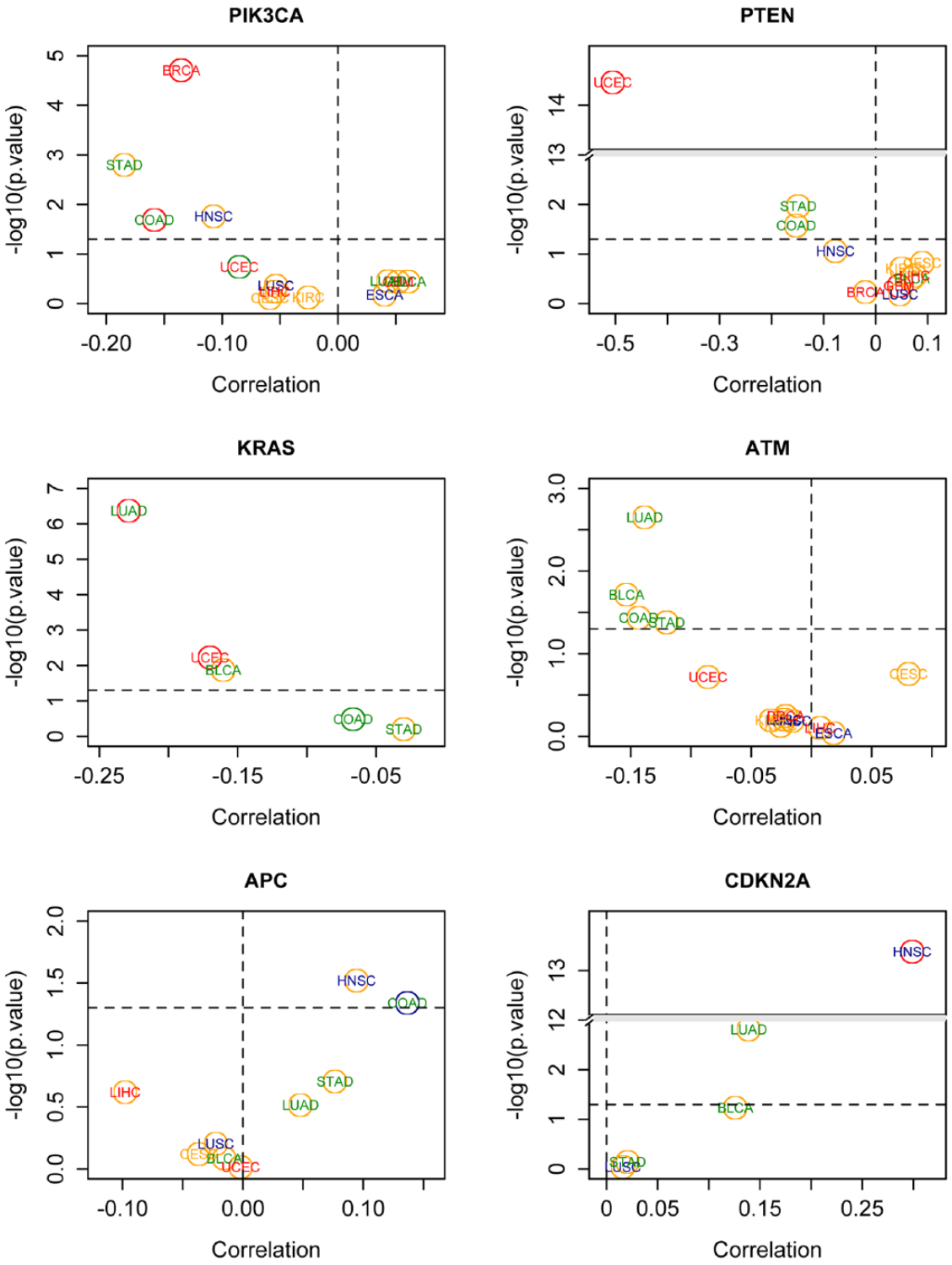

Figure 5. The conserved correlation (ie, co-occurrence or mutual exclusivity) patterns between *TP53* mutations and the alterations in the partner genes, including *PIK3CA, PTEN, KRAS, APC, CDKN2A*, and *ATM*. Circle label: the cancer type whose data were used to evaluate the mutational relationships. Color of circle label (or circle) indicates the proportion (x) of tumors with a mutation burden on *TP53* (or the partner gene): orange: $0 < x < 0.2$; red: $0.2 \leqslant x < 0.4$; green: $0.4 \leqslant x < 0.6$; and blue: $x \geqslant 0.6$. BLCA indicates bladder urothelial carcinoma; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; STAD, stomach adenocarcinoma; UCEC, uterine corpus endometrial carcinoma.

Table 2. The effects of mutation burden (ie, the total number of somatic mutations) in cancer driver genes on patient overall survival time.[a]

TYPES	TP53-MUTATED GROUPS		TP53 WILD-TYPE GROUPS	
	B[b]	*P* VALUE (FDR)[c]	B	*P* VALUE
BLCA	−0.824	**.039 (0.11)**	−0.217	.431
GBM	−0.954	**.026 (0.11)**	0.232	.127
HNSC	0.020	.900	0.408	.055
LUAD	−0.356	**.037 (0.11)**	0.058	.785
LUSC	−0.209	.367	−0.191	.613
BRCA	−0.291	.399	−0.315	.126
OV	−0.571	**.010 (0.11)**	0.183	.579
UCEC	−0.069	.864	−0.868	.078
COAD	0.083	.778	−0.513	.298
ESCA	0.545	.190	−0.022	.950
LIHC	−0.504	.276	0.293	.181
STAD	0.224	.527	−0.173	.476

Abbreviations: BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; COAD, colon adenocarcinoma; ESCA, esophageal carcinoma; FDR, false discovery rate; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; STAD, stomach adenocarcinoma; UCEC, uterine corpus endometrial carcinoma.

[a]*TP53* was excluded in calculating the mutation burden.

[b]B represents the regression coefficient estimated by the Cox proportional hazards model. A negative (or positive) B indicates that overall survival time increases (or decreases) as mutation burdens increase.

[c]P Values less than 0.05 is marked in bold fonts, indicating the corresponding regression coefficient (B) is significantly different from zero. FDR was estimated using the Benjamini-Hochberg procedure.

Acknowledgements

The analyses presented here are based on the data published by The Cancer Genome Atlas (TCGA). The authors downloaded the data sets via TCGA Data Portal, which were managed by the NCI and NHGRI. At present, all TCGA data reside at the Genomic Data Commons (https://gdc-portal.nci.nih.gov/legacy-archive/search/f). They thank the 4 reviewers for their constructive comments.

Author Contributions

WZ and KZ conceived and designed the experiments. WZ performed the experiments. WZ and KZ analyzed the data. WZ, AE, EKF, and KZ wrote the paper. EKF and AE helped with experiment design. All authors read and approved the final manuscript.

REFERENCES

1. Lamb P, Crawford L. Characterization of the human p53 gene. *Mol Cell Biol*. 1986;6:1379–1385.
2. Liu J, Zhang C, Feng Z. Tumor suppressor p53 and its gain-of-function mutants in cancer. *Acta Biochim Biophys Sin (Shanghai)*. 2014;46:170–179.
3. George P. p53 how crucial is its role in cancer. *Int J Curr Pharm Res*. 2011;3:19–25.
4. Willis A, Jung EJ, Wakefield T, Chen X. Mutant p53 exerts a dominant negative effect by preventing wild-type p53 from binding to the promoter of its target genes. *Oncogene*. 2004;23:2330–2338.
5. Zhang W, Flemington E, Zhang K. Mutant TP53 disrupts age-related accumulation patterns of somatic mutations in multiple cancer types. *Cancer Genet*. 2016;209:376–380.
6. Vogelstein B, Sur S, Prives C. p53: the most frequently altered gene in human cancers. *Nat Educ*. 2010;3:6.
7. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. *Cold Spring Harb Perspect Biol*. 2010;2:a001008.
8. Bunz F. *Principles of Cancer Genetics*. Dordrecht, The Netherlands: Springer; 2008.
9. Robles AI, Harris CC. Clinical outcomes and correlates of TP53 mutations and cancer. *Cold Spring Harb Perspect Biol*. 2010;2:a001016.
10. Skaug V, Ryberg D, Kure EH, et al. p53 mutations in defined structural and functional domains are related to poor clinical outcome in non-small cell lung cancer patients. *Clin Cancer Res*. 2000;6:1031–1037.
11. Huang C, Taki T, Adachi M, Konishi T, Higashiyama M, Miyake M. Mutations in exon 7 and 8 of p53 as poor prognostic factors in patients with non-small cell lung cancer. *Oncogene*. 1998;16:2469–2477.
12. Molina-Vila MA, Bertran-Alamillo J, Gasco A, et al. Nondisruptive p53 mutations are associated with shorter survival in patients with advanced non-small cell lung cancer. *Clin Cancer Res*. 2014;20:4647–4659.
13. Poeta ML, Manola J, Goldwasser MA, et al. TP53 mutations and survival in squamous-cell carcinoma of the head and neck. *N Engl J Med*. 2007;357:2552–2561.
14. Peltonen JK, Vahakangas KH, Helppi HM, Bloigu R, Paakko P, Turpeenniemi-Hujanen T. Specific TP53 mutations predict aggressive phenotype in head and neck squamous cell carcinoma: a retrospective archival study. *Head Neck Oncol*. 2011;3:20.
15. Gross AM, Orosco RK, Shen JP, et al. Multi-tiered genomic analysis of head and neck cancer ties TP53 mutation to 3p loss. *Nat Genet*. 2014;46:939–943.
16. Pollard KS, Hubisz MJ, Rosenbloom KR, Siepel A. Detection of nonneutral substitution rates on mammalian phylogenies. *Genome Res*. 2010;20:110–121.
17. Yule GU. On the methods of measuring the association between two attributes. *J Royal Stat Soc*. 1912;75:579–652.
18. Kandoth C, McLellan MD, Vandin F, et al. Mutational landscape and significance across 12 major cancer types. *Nature*. 2013;502:333–339.
19. Tamborero D, Gonzalez-Perez A, Perez-Llamas C, et al. Comprehensive identification of mutational cancer driver genes across 12 tumor types. *Sci Rep*. 2013;3:2650.
20. Courtois S, Caron de Fromentel C, Hainaut P. p53 protein variants: structural and functional similarities with p63 and p73 isoforms. *Oncogene*. 2004;23:631–638.
21. Edwards SJ, Hananeia L, Eccles MR, Zhang YF, Braithwaite AW. The proline-rich region of mouse p53 influences transactivation and apoptosis but is largely dispensable for these functions. *Oncogene*. 2003;22:4517–4523.
22. Dornan D, Shimizu H, Burch L, Smith AJ, Hupp TR. The proline repeat domain of p53 binds directly to the transcriptional coactivator p300 and allosterically controls DNA-dependent acetylation of p53. *Mol Cell Biol*. 2003;23:8846–8861.
23. Powell B, Soong R, Iacopetta B, Seshadri R, Smith DR. Prognostic significance of mutations to different structural and functional regions of the p53 gene in breast cancer. *Clin Cancer Res*. 2000;6:443–451.
24. Lane DP, Cheok CF, Brown C, Madhumalar A, Ghadessy FJ, Verma C. Mdm2 and p53 are highly conserved from placozoans to man. *Cell Cycle*. 2010;9:540–547.
25. Nagy JD, Victor EM, Cropper JH. Why don't all whales have cancer? a novel hypothesis resolving Peto's paradox. *Integr Comp Biol*. 2007;47:317–328.
26. Galis F, Metz JA. Anti-cancer selection as a source of developmental and evolutionary constraints. *Bioessays*. 2003;25:1035–1039.
27. Leroi AM, Koufopanou V, Burt A. Cancer selection. *Nat Rev Cancer*. 2003;3:226–231.
28. Ellrott K. *MutSig.pancan12.sig_genes.v2*. Seattle, WA: Sage Bionetworks; 2013.
29. Korosteleva O. *Clinical Statistics: Introducing Clinical Trials, Survival Analysis, and Longitudinal Data Analysis*. Sudbury, MA: Jones & Bartlett Learning; 2009.
30. Santos C, Sanz-Pamplona R, Nadal E, et al. Intrinsic cancer subtypes—next steps into personalized medicine. *Cell Oncol (Dordr)*. 2015;38:3–16.
31. Ogino S, Fuchs CS, Giovannucci E. How many molecular subtypes? implications of the unique tumor principle in personalized medicine. *Expert Rev Mol Diagn*. 2012;12:621–628.

32. Marur S, Forastiere AA. Head and neck cancer: changing epidemiology, diagnosis, and treatment. *Mayo Clin Proc*. 2008;83:489–501.
33. Potter JD, McMichael AJ. Alcohol, beer and lung cancer—a meaningful relationship? *Int J Epidemiol*. 1984;13:240–242.
34. Cancer IAoRo. IARC Monographs programme on the evaluation of the carcinogenic risk of chemicals to humans. Preamble. *IARC Monogr Eval Carcinog Risk Chem Hum*. 1986;39:13–32.
35. Hamajima N, Hirose K, Tajima K, et al. Alcohol, tobacco and breast cancer—collaborative reanalysis of individual data from 53 epidemiological studies, including 58,515 women with breast cancer and 95,067 women without the disease. *Br J Cancer*. 2002;87:1234–1245.
36. Moy KA, Fan Y, Wang R, Gao YT, Yu MC, Yuan JM. Alcohol and tobacco use in relation to gastric cancer: a prospective study of men in Shanghai, China. *Cancer Epidemiol Biomarkers Prev*. 2010;19:2287–2297.
37. Nishino Y, Inoue M, Tsuji I, et al. Tobacco smoking and gastric cancer risk: an evaluation based on a systematic review of epidemiologic evidence among the Japanese population. *Jpn J Clin Oncol*. 2006;36:800–807.
38. Zaridze D, Borisova E, Maximovitch D, Chkhikvadze V. Alcohol consumption, smoking and risk of gastric cancer: case-control study from Moscow, Russia. *Cancer Causes Control*. 2000;11:363–371.
39. Reva B, Antipin Y, Sander C. Predicting the functional impact of protein mutations: application to cancer genomics. *Nucleic Acids Res*. 2011;39:e118.
40. McFarland CD, Korolev KS, Kryukov GV, Sunyaev SR, Mirny LA. Impact of deleterious passenger mutations on cancer progression. *Proc Natl Acad Sci U S A*. 2013;110:2910–2915.
41. Rozhok AI, DeGregori J. Toward an evolutionary model of cancer: considering the mechanisms that govern the fate of somatic mutations. *Proc Natl Acad Sci U S A*. 2015;112:8914–8921.
42. Yeang CH, McCormick F, Levine A. Combinatorial patterns of somatic gene mutations in cancer. *FASEB J*. 2008;22:2605–2622.
43. Constantinescu S, Szczurek E, Mohammadi P, Rahnenfuhrer J, Beerenwinkel N. TiMEx: a waiting time model for mutually exclusive cancer alterations. *Bioinformatics*. 2015;32:968–975.
44. Szczurek E, Beerenwinkel N. Modeling mutual exclusivity of cancer mutations. *PLoS Comput Biol*. 2014;10:e1003503.
45. Babur O, Gonen M, Aksoy BA, et al. Systematic identification of cancer driving signaling pathways based on mutual exclusivity of genomic alterations. *Genome Biol*. 2015;16:45.
46. Thomas RK, Baker AC, Debiasi RM, et al. High-throughput oncogene mutation profiling in human cancer. *Nat Genet*. 2007;39:347–351.
47. Zhang J, Wu LY, Zhang XS, Zhang S. Discovery of co-occurring driver pathways in cancer. *BMC Bioinformatics*. 2014;15:271.
48. Agarwal P, Kabir FML, DeInnocentes P, Bird RC. Tumor suppressor gene p16/INK4A/CDKN2A and its role in cell cycle exit, differentiation, and determination of cell fate, tumor suppressor genes. In: Cheng Y, ed. *Tumor Suppressor Gene*. Rijeka, Croatia: InTech; 2012; 1–34.
49. Wu M, Bellas RE, Shen J, Sonenshein GE. Roles of the tumor suppressor p53 and the cyclin-dependent kinase inhibitor p21WAF1/CIP1 in receptor-mediated apoptosis of WEHI 231 B lymphoma cells. *J Exp Med*. 1998;187:1671–1679.
50. Majchrzak A, Witkowska M, Smolewski P. Inhibition of the PI3K/Akt/mTOR signaling pathway in diffuse large B-cell lymphoma: current knowledge and clinical significance. *Molecules*. 2014;19:14304–14315.
51. Bozic I, Antal T, Ohtsuki H, et al. Accumulation of driver and passenger mutations during tumor progression. *Proc Natl Acad Sci U S A*. 2010;107:18545–18550.
52. McFarland CD, Mirny LA, Korolev KS. Tug-of-war between driver and passenger mutations in cancer and other adaptive processes. *Proc Natl Acad Sci U S A*. 2014;111:15138–15143.
53. Zhang W, Edwards A, Flemington E, Zhang K. Somatic mutations favorable to patient survival are predominant in ovarian carcinomas. *PLoS ONE*. 2014;9:e112561.
54. Birkbak NJ, Kochupurakkal B, Izarzugaza JM, et al. Tumor mutation burden forecasts outcome in ovarian cancer with BRCA1 or BRCA2 mutations. *PLoS ONE*. 2013;8:e80023.
55. Prindle MJ, Fox EJ, Loeb LA. The mutator phenotype in cancer: molecular mechanisms and targeting strategies. *Curr Drug Targets*. 2010;11:1296–1303.
56. Fox EJ, Loeb LA. Lethal mutagenesis: targeting the mutator phenotype in cancer. *Semin Cancer Biol*. 2010;20:353–359.