Comprehensive analysis of DNA damage repair deficiency in 10,284 pan-cancer study

Yanni Xiao[1,2#], Di Lu[3#], Mingxing Lei[1], Wenzhuan Xie[4], Yaoxu Chen[4], Yating Zheng[4], Chunli Wang[1], Jing Zhao[4], Zhen Zhu[5], Xiaochen Zhao[4], Mengli Huang[4], Youen Lin[6], Zhongjun Li[2], Li Yang[1]

[1]111 Project Laboratory of Biomechanics and Tissue Repair, College of Bioengineering, Chongqing University, Chongqing, China; [2]Department of Blood Transfusion, Laboratory of Radiation Biology, The Second Affiliated Hospital, Third Military Medical University, Chongqing, China; [3]Department of Thoracic Surgery, Nanfang Hospital, Southern Medical University, Guangzhou, China; [4]The Medical Department, 3D Medicines Inc., Shanghai, China; [5]Department of Thoracic Surgery, Kunshan Hospital of Traditional Chinese Medicine, Kunshan, China; [6]Department of Oncology, Jieyang Yuedong Cancer Hospital, Jieyang, China

Contributions: (I) Conception and design: Y Xiao, D Lu, L Yang; (II) Administrative support: L Yang, Z Li, Y Lin; (III) Provision of study materials or patients: Y Xiao, D Lu, M Lei, W Xie, Y Chen, Y Zheng; (IV) Collection and assembly of data: Y Xiao, D Lu, M Lei, W Xie, Y Chen, Y Zheng, Z Zhu; (V) Data analysis and interpretation: C Wang, W Xie, J Zhao, X Zhao, M Huang, Y Lin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Li Yang, PhD. 111 Project Laboratory of Biomechanics and Tissue Repair, College of Bioengineering, Chongqing University, 83 Shabei Street, Shapingba District, Chongqing 400030, China. Email: yanglibme@cqu.edu.cn; Zhongjun Li, PhD. Department of Blood Transfusion, Laboratory of Radiation Biology, The Second Affiliated Hospital, Third Military Medical University, 83 Xinqiao, Main Street, Shapingba District, Chongqing, China. Email: johnneyusc@gmail.com; Youen Lin, Master. Department of Oncology, Jieyang Yuedong Cancer Hospital, Longshang Community, Yuhu Town, Rongcheng District, Jieyang 522000, China. Email: 82395946@qq.com.

Background: Disruption of the DNA damage repair (DDR) gene is related to cancer progression, treatment selection, and is subjected to radiation and targeted therapies with limited success. This paper conducted a comprehensive analysis to explore the distribution of DDR mutations in Chinese pan-cancer patients.

Methods: A total of 10,284 consecutive cases were analyzed in 24 cancer types [non-small cell lung cancer (NSCLC) 29.0%, liver 12.0%, colorectum 10.7%, etc.]. Tumor tissue samples were subjected to next generation sequencing (NGS) using a 381 gene panel incorporating 100 microsatellite loci. The association of deleterious somatic DDR mutation (del-sDDR^{mut}) with tumor mutational burden (TMB), microsatellite instability (MSI), programmed cell death-ligand 1 (PD-L1) expression of pan-cancers was evaluated. Genomic and clinical data from public cohorts of immunotherapy were analyzed to demonstrate the association between del-sDDR^{mut} and clinical survival.

Results: Del-sDDR^{mut} were found in 802 (7.6%) of all cases, and were most common in cancers of the endometrium, prostate, bladder, etc. cancer with a higher TMB also had a higher prevalence of mutations in DDR pathways. The results of the ridge regression analysis showed that 20 DDR genes were significantly associated with TMB [false discovery rate (FDR) <0.01]. A total of 8,899 patients had both TMB and MSI-data in pan-cancers. Seventy-four percent of patients with MSI-high (MSI-H) were accompanied by del-sDDR^{mut}/TMB-high (TMB-H). The largest proportion of patients with microsatellite stability (MSS) with DDR mutations were classified as TMB-H. The top 6 tumors (NSCLC, melanoma, esophagus, head and neck, thyroid, and mediastinal) had the highest prevalence of PD-L1 ≥1%, and DDR mutations were significantly associated with a higher percent of PD-L1 positive (P<0.05). Furthermore, in the immune cohort analysis of NSCLC, patients with del-sDDR^{mut} significantly improved median progression-free survival (mPFS) and median overall survival (mOS) compared to wild-type DDR patients (P=0.002 and P=0.043), with higher TMB observed (P<0.001).

Conclusions: This study explored the association of DDR mutations with TMB, MSI-H, and PD-L1

expression, and revealed that patients with DDR mutations have a significantly improve prognosis than wild-type patients on immunotherapy.

Keywords: DNA damage repair deficiency (DDR deficiency); tumor mutational burden (TMB); somatic mutation; pan-cancer

Submitted Sep 03, 2021. Accepted for publication Nov 12, 2021.
doi: 10.21037/atm-21-5449
View this article at: https://dx.doi.org/10.21037/atm-21-5449

Introduction

The DNA damage repair (DDR) system plays a key role in maintaining human genomic stability. Around 200 genes have been identified as directly involved in DDR and the resulting DNA damage checkpoint (1). Genomic instability is a crucial feature of cancer. DDR defects, mismatch repair (MMR) defects for instance, will lead to microsatellite instability (MSI), which may result in development of colorectal and endometrial cancer. While, chromosomal instability (CIN) could be observed in most solid tumors. When telomeres of newborn cancer cells become extremely short and lead to chromosome fusion, CIN appears. While activated oncogenes and subsequent DNA replication stress due to DDR defects continue to fuel CIN, and eventually tumors are grew. In addition, hereditary DDR defects are often prone to cancer. Furthermore, DDR defects is also related to therapeutic response, and treatment selections. Several agents, such as PARP inhibitors, immune checkpoint inhibitors (ICIs), have recently received FDA approval in multiple types of solid tumors (2-5).

The therapeutic implications of DDR mutations are becoming better known. Many antitumor compounds that directly target DDR pathways are being evaluated in clinical evaluation, including WEE1 G2 checkpoint kinase (*WEE1*), checkpoint kinase (*CHEK1*/*CHEK2*), ataxia telangiectasia mutated (*ATM*), or ataxia telangiectasia and Rad3-related (*ATR*) inhibitors involved in the calcineurin like EF-hand protein (*CHP*) pathway and polyadenosine diphosphate-ribose polymerase (*PAPR1/2*) inhibitors in the BER pathway (6). Until now, Food and Drug Administration (FDA) has approved four PARP inhibitors (olaparib, rucaparib, niraparib, and talazoparib) have been approved by FDA in at least one type of cancer with somatic or germline *BRCA* mutations or gene mutations in the HRR pathway, including ovarian cancer, breast cancer (7), metastatic castration-resistant prostate cancer (PCa) (8), pancreatic cancer (4), etc. These are limited to patients with breast cancer gene (*BRCA*) mutations, and patients with other DDR mutations, including ATM, ATR, non-BRCA HRR mutations (*RAD51C*, *RAD51D*, *PALB2*, etc.) were found to have a response to PARP inhibitors (9,10).

Additionally, defects in DDR damage have been associated with improved therapeutic sensitivity to chemotherapy or immunotherapy. For chemotherapy, if tumor cells have DDR mutations, genotoxic drugs, including platinum, are susceptible to cause DNA damage that exceeds the repair capacity of DDR systems, which will stop cell replication and induce cell apoptosis or death (11). A recent study has shown that at least one gene alteration among 34 DDR genes was associated with improved clinical outcomes in platinum-treated patients with advanced urothelial carcinoma (12). Furthermore, the DDR mutation was reported to be correlated with a higher tumor mutational burden (TMB) and an improved immune microenvironment, indicating that it may be a predictive biomarker for the application of inhibitors at the immune checkpoint (13). Deficient MMR (dMMR)/MSI-high (MSI-H), mainly caused by mutations in the MMR pathway genes, has been approved by the FDA to guide ICIs in pan-cancer (14,15).

As the DDR system plays a key role in cancer treatment. Previous studies explored the relation between DDR mutation and tumorigenesis, prognosis, drug development. Concerning immunotherapy, there is a study that linked DDR mutation to TMB in gastrointestinal cancer. There is no study that simultaneously explains the relationship between DDR and well-established immunotherapy biomarkers [MSI/TMB/programmed cell death-ligand 1 (PD-L1)] in pan-cancer. Our objective was to perform a comprehensive analysis in 10,284 cancer patients to explore the distribution of DDR deficiency in 24 cancer types, the association of DDR deficiency with TMB, MSI, PD-L1. Our results provide a useful resource to guide the mechanistic, therapeutic, and predictive role of DDR in cancers. We present the following article in accordance with

the MDAR reporting checklist (available at https://dx.doi.org/10.21037/atm-21-5449).

Methods

Clinical cancer specimens

Case information was collected from five hospitals (as follows: The Second Affiliated Hospital of Chongqing, Nanfang Hospital, Kunshan Hospital, Jieyang Yuedong Cancer Hospital, Affiliated Cancer Hospital of Chongqing University), 10,284 patients diagnosed with malignant solid tumors who underwent next generation sequencing (NGS) testing between January 2017 and April 2020 were included in the analysis. Formalin-fixed paraffin-embedded (FFPE) tumor specimens of pan-cancer patients were enrolled in this study. The specimens were confirmed by hematoxylin and eosin (H&E) staining for a pathological diagnosis, and were considered qualified with a size ≥1 mm^3, and the percentage of cancer cells should be over 20%. All procedures performed in this study involving human were in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics board of committee of Jieyang Yuedong Cancer Hospital (No. B2021-1-01). The study was a retrospective study and individual consent for the analysis was waived.

NGS

Library preparation and targeted capture

DNA was cut into 250 bp using S220 focused-ultrasonicator (Covaris, Woburn, MA, USA). The preparation was carried out using the KAPA Hyper Prep kit (Kapa biosystems, Wilmington, USA). The concentration and size distribution of each library were measured by Qubit 3.0 fluorometer (Thermo Fisher Scientific, Shanghai, China) and aLabChip GX Touch HT analyzer (PerkinElmer, Boston, MA, USA). For targeted capture, indexed libraries were subjected to probe-based hybridization with a customized NGS panel targeting 381 cancer-related genes as previously described (16) (Table S1), where the probe baits were individually synthesized 5' biotinylated 120 bp DNA oligonucleotides (IDT). Repetitive elements were filtered from intronic baits according to the annotation by UCSC Genome RepeatMasker. The xGen® Hybridization and Wash Kit (IDT) was used for hybridization enrichment. Briefly, 500 ng of indexed DNA libraries were pooled to obtain a total amount of 2 μg of DNA. The pooled DNA sample was then mixed with human cot DNA and xGen Universal Blockers-TS Mix and dried down in a SpeedVac system. The Hybridization Master Mix was added to the samples and incubated in a thermal cycler at 95 ℃ for 10 min before being mixed and incubated with 4 μL of probes at 65 ℃ overnight. The target regions were captured following the manufacturer's instructions. The final library's concentration and fragment size distribution were determined using a Qubit 3.0 fluorometer (Thermo Fisher Scientific, Shanghai, China) and a LabChip GX Touch HT analyzer (PerkinElmer, Boston, MA, USA), respectively.

DNA sequencing, data processing, and variant calling

The libraries were then loaded into a NovaSeq6000 platform (Illumina) and subjected to a sequencing depth of 1,000×. The data from the samples were then mapped to the reference human genome hg19 by burrows Wheeler comparator (v0.7.12) (17). PCR data was collected using SAMtools (v1.1.19) and Picard (v1.130). A variant detection model was then developed to detect Somatic single nucleotide variants (SNVs) based on the Binomial test. PCR data was also analyzed using a modified R package. A local realignment procedure was performed to detect indels. The filtered variables were then filtered according to their unique base quality and supporting read depth (16). A filter was then used to filter out variants, and ensure sensitivity and specificity at allele frequency (AF) ≥5%. The dbSNP (v138), 1000Genome and ESP6500 (population frequency >0.015) databases were annotated with single nucleotide polymorphisms (SNPs) and indels by ANNOVAR. Only missense, stop gain, frame shifted and non-frameshift indel mutations were retained.

Somatic and germline alterations were identified. Germline variation was screened by comparing each tumor tissue with an adjacent normal samples or blood controls. Pathogenic and possible pathogenic mutations were explained by bioinformatics experts according to the consensus of previous reports and the recommendations of the American College of Medical Genetics and Genomics and the Association of Molecular Pathology (18).

TMB, MSI, and PD-L1 testing

TMB was defined as the number of mutated bases per million bases tested. Among them, missense, silent, stop gain, stop loss, in-frame and frameshift mutation types were included. TMB-high (TMB-H) was defined as greater than the median value.

One hundred microsatellite loci were selected to

determine the MSI, and for each assay, the top 30 loci with the best coverage were included for the final calculation of the MSI score. An internally developed R script was used to evaluate the distribution of reading counts among various repeat lengths for each microsatellite locus of each sample. Any sample with an MSI score of ≥0.4 was classified as MSI-H and otherwise microsatellite stability (MSS).

FFPE tissue sections were subjected to assessment of PD-L1 expression using the PD-L1 immunohistochemistry (IHC) 22C3 pharmDx assay (Agilent Technologies, Shanghai, China) or PD-L1 IHC SP263 (Roche Diagnostics GmbH, Shanghai, China). Staining for 22C3 was performed on the Dako Link-48 autostainer system at Teddy Clinical Research lab, while staining for SP263 was performed on the Roche BenchMark Ultra platform at QIAGEN Suzhou Clinical Lab. PD-L1 expression was determined using the tumor proportion score (TPS), the proportion of viable tumor cells showing partial or complete membrane PD-L1 staining at any intensity. TPS ≥1% was considered PD-L1 positive.

Determination of a deleterious DDR mutation

Thirty-one genes that belong to seven different pathways, including base excision repair (BER), nucleotide excision repair (NER), MMR, Fanconi anemia (FA), homology-dependent recombination (HR), non-homologous end joining (NHEJ), and damage sensor (DS), were identified as DDR pathway genes based on searches of the PubMed, NCBI Gene, and Biosystems Databases (Table S1). All pathogenic variants or likely pathogenic variants in DDR genes were considered deleterious, including TRUNC (Frame_Shift_Del, Frame_Shift_Ins, Nonsense, Nonstop, Splice_Site, Translation_Start_Site), INFRAME (In_Frame_Del and In_Frame_Ins), and MISSENSE mutations were considered. Patients harboring one or more deleterious DDR mutations were defined as DDR mutations, while patients without deleterious DDR mutations were defined as DDR wild-type subgroup. A mutation in the DDR pathway was determined by at least one alteration of the DDR gene in the corresponding pathway.

Ridge regression analysis

Linear ridge regression was performed on 31 DDR genes using the alteration status, MSI-H status, and encoding tumor type as 24 additional binary variables. The coefficients were determined by the method developed by Cule and De Iorio (19). Regression was performed in R-3.6.0 using the Ridge package.

Immune cohort analysis

Genomic and clinic data from public cohorts involving immunotherapeutic patients [Rizvi 2015 (20); Rizvi 2018 (21); Miao 2018 (22); Hellmann 2018 (23); Samstein 2019 (24)] were analyzed. Overall survival (OS)/progression-free survival (PFS) were analyzed in R-3.6.0 using the Survival package. Meta-analysis was performed in R-3.6.0 using the Meta package.

Statistical analysis

For normally distributed continuous variables, the Student's *t*-test or the Mann-Whitney U test was used to determine the differences between the two groups. Fisher's exact test or the Chi-square test was used to identify the association of two categorical variables. P values of all reported were two-tailed, and $P<0.05$ was considered was considered statistically significant. All analyses and graphs in the present study were performed by R 3.6.0.

Results

Patients' characteristics and prevalent DDR mutations across cancer types

A total of 10,284 patients with pan-cancers and successful tumor NGS between January 2017 and August 2019 were identified [2,876 non-small cell lung cancer (NSCLC), 1,237 liver, 1,097 colorectum, 850 biliary tracts, 638 stomach, etc.]; the median age of this cohort was 58 (IQR, 49–66) years, most tumors demonstrated male (61%). Positive expression of PD-L1 expression (PD-L1 ≥1%) was found in 41%, while TMB-H (higher than the median TMB in each tumor) was found in 48%. The median TMB in the entire cohort was 6 mutations/Mb (mut/Mb, range, 4–10); the baseline clinical of the 10,284 patients is detailed in Table S2.

Tumors from 1,218 patients (11.8%) were defined as DDR mutations, while the remaining 9,066 (88.2%) were defined as DDR wide-type. Among DDR-mutations pan-cancer, 720 patients (59.1%) were identified as only having deleterious somatic DDR mutation (del-sDDR^{mut}), and 424 (34.8%) patients were identified as having germline DDR mutations (gDDR^{mut}) alone, while the remaining 74 (6.1%) patients have both del-sDDR^{mut} and gDDR^{mut}. The most commonly mutated DDR genes were *ATM* (19.13%),

BRCA2 (17.16%), *BRCA1* (10.92%), *RAD50* (8.92%) and *ATR* (7.8%; *Figure 1*).

The del-sDDR^{mut} was 7.4% (213/2,876) in NSCLC, while the most frequent in the endometrium (26/106, 24.5%), followed by the bladder/urinary tract (31/226, 13.7%), the intestinal tract (15/119, 12.6%), colorectum (124/1,097, 11.3%), and the cervix (19/169, 11.2%; *Figure 2A*; *Table 1*). Cancer types with a higher TMB also had a higher prevalence of mutations in DDR pathways (*Figure 2A,2B*). Various mutations in the DDR pathways accumulated in different types of cancer (*Figure 2B*). In endometrium cancer, the mutation frequency of BER, NHEJ, and NER was more than 2.8%, 4.7%, and 5.7%, respectively. The MMR pathway was also recurrently altered in cancers of the endometrium, cervix, colorectum, and intestine (*Figure 2B*). HR and FA alteration were more common in the endometrium, bladder/urinary tract, intestine, prostate, colorectum, and ovary cancer. NHEJ had relatively lower mutation frequencies, mainly in the liver, lung, pancreas, sarcoma, kidney, etc. (*Figure 2B*). The tile plots showed that the alterations in NER and MMR, HR and MMR, or HR and NER were exclusive (*Figure 2C*). Compared to del-sDDR^{mut}, patients with ovary (15.7%) and prostate (13.2%) have the highest frequencies in gDDR^{mut} than other cancers, and the frequency of BCRA2 mutation in the prostate has more than 5.9%, much higher than other DDR genes (Figure S1).

The association of DDR mutation and TMB, MSI-H, and PD-L1

Among 95.2% (756/794) del-sDDR^{mut} cancer samples with TMB and MSI levels, we then explored the association of del-sDDR^{mut} with TMB by using ridge regression. The volcano plot showed that 21 DDR genes were significantly associated with TMB [false discovery rate (FDR) ≤0.01, *Figure 3A*]. Furthermore, polymerase (DNA) epsilon (*POLE*), polymerase (DNA) delta 1 (*POLD1*), *RAD51*, BRCA1-associated RING domain gene 1 (*BARD1*), Fanconi anemia complementation group C (*FANCC*), Fanconi anemia complementation group G (*FANCG*), MutS homolog 2 (*MSH2*), and Fanconi anemia complementation group D2 (*FANCD2*) contribute to TMB more than MSI-H (*Figure 3A*).

Of the del-sDDR^{mut} cases in prevalent cancers of MSI-H (endometrium, prostate, colorectum, stomach, intestinal, cervix), 43.27% (90/208) were MSS and TMB-low (TMB <83.9 mut/Mb). Of the DDR-altered/TMB-H cases, 74.42% (32/43) were MSI-H (*Figure 3B*). Of DDR-altered cases, 34.62% (72/208) had one or more gene alterations in the MMR pathway [MutL homolog 1 (*MLH1*), *MSH2*, MutS homolog 6 (*MSH6*), or postmeiotic segregation increased 2 (*PMS2*)]. Of MMR-altered cases, 83.33% (60/72) were MSI-H, and 37.50% (27/72) were TMB-H (TMB ≥83.9 mut/Mb) (*Figure 3B*). For certain genes, patients with *POLE*, *CHEK2*, *BRAD1*, *FANCC*, *BRCA1*, and *FANCG* had the highest proportion of TMB-H/MSS (*Figure 3C*).

The TMB was significantly higher among MSS patients in patients with a DDR alteration than in DDR-WT cases (median, 9.68 *vs.* 6.45 mut/MB; P<0.0001; *Figure 3D*). Except for RAD51, Fanconi anemia complementation group F (*FANCF*), protein kinase DNA-activated catalytic subunit (*PRKDC*), *PALB2*, *FANCG*, *FANCC*, *POLD1*, the other 22 mutations of DDR genes were enriched with TMB-H. Within MSI-H patients, although 20% of the patients were TMB-H, the TMB in patients with DDR alterations was higher than DDR-WT (median, 59.79 *vs.* 16.11 mut/Mb; P<0.0001; *Figure 3E*). The specific DDR gene with significantly higher TMB included *FANCE*, *PALB2*, *FANCD2*, breast cancer 1 interacting protein 1 (*BRIP1*), *BRCA2*, bloom syndrome, RecQ like helicase (*BLM*), *ATR*, *ATM*, *MSH6*, *MSH2*, *RAD50*, meiotic recombination 11 homolog A (*MRE11A*) (*Figure 3E*).

The top 6 tumors with the highest prevalence of PD-L1 ≥1% were NSCLC (56.93%, 1,125/1,976), melanoma (60.82%, 59/97), esophagus (62.30%, 76/122), head and neck (64.19%, 95/148), thyroid (64.29%, 9/14) and mediastinal (85.71%, 18/21) (*Figure 4A*). Among 31 DDR genes, alteration of *ATR*, *BLM*, *CHEK2*, *MSH6*, *PMS2* was significantly associated with a higher percent of PD-L1 positive (P<0.05; *Figure 4B*; Figure S2).

The association of del-sDDR^{mut} and immunotherapy in NSCLC

Five independent cohorts (Rizvi 2015, Rizvi 2018, Miao 2018, Samstein 2019, and Hellmann 2018; study cohort) (20-24) with data from patients with advanced NSCLC were used to analyze the correlation of del-sDDR^{mut} and immunotherapy. The meta-analysis showed that patients with del-sDDR^{mut} exhibited a significantly reduced risk of death compared to the wild-type DDR group (hazard ratio =0.71; 95% CI: 0.56–0.90, fixed effect model; *Figure 5A*). The del-sDDR^{mut} patients significantly improved median PFS (mPFS) and median OS (mOS) than the DDR wild-type patients (mPFS: 5.4 *vs.* 3.5 months, hazard ratio =0.68,

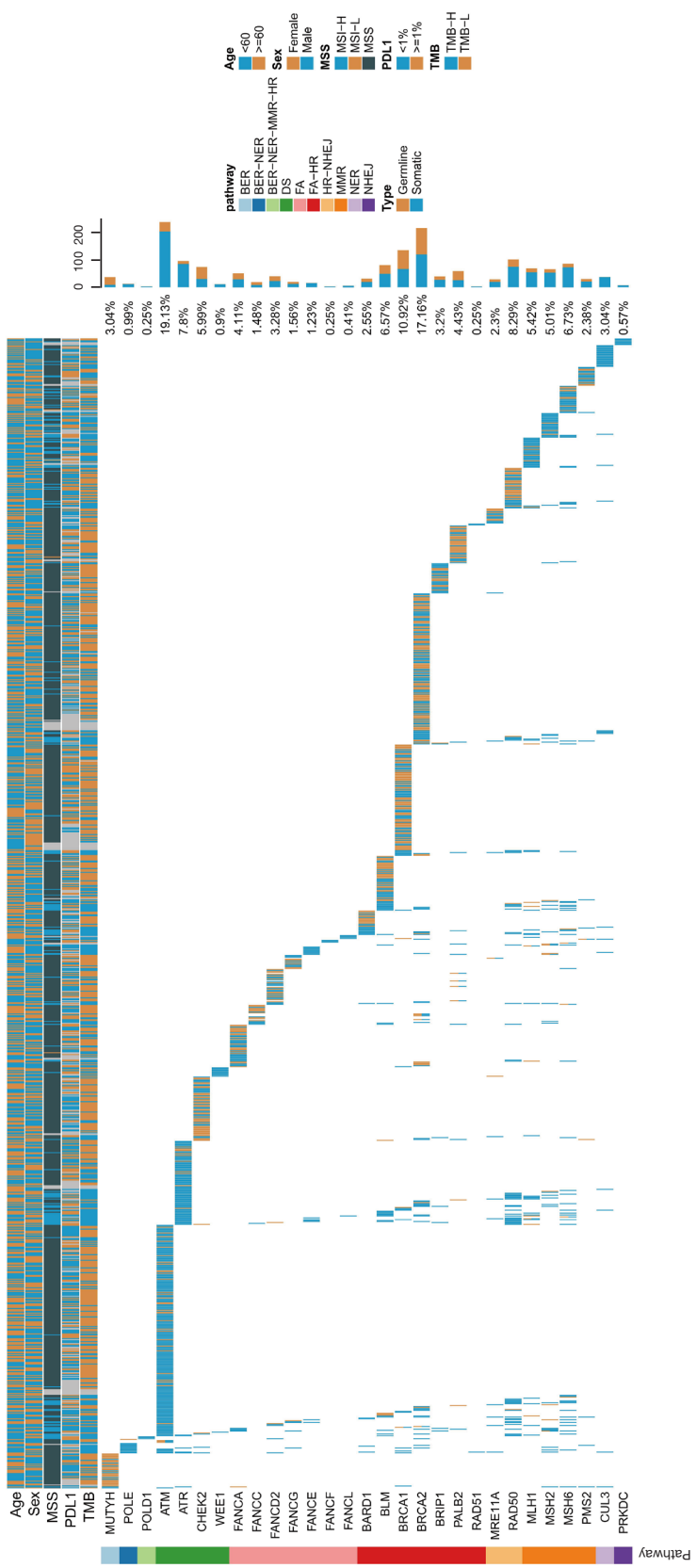

Figure 1 Mutual exclusion of sDDR^{mut} and gDDR^{mut}. sDDR^{mut}, somatic DDR mutation; gDDR^{mut}, germline DDR mutations; DDR, DNA damage repair; BER, base excision repair; NER, nucleotide excision repair; MMR, mismatch repair; DS, damage sensor; FA, Fanconi anemia; HR, homology-dependent recombination; NHEJ, non-homologous end joining; TMB, tumor mutational burden; TMB-H, TMB-high; TMB-L, TMB-low.

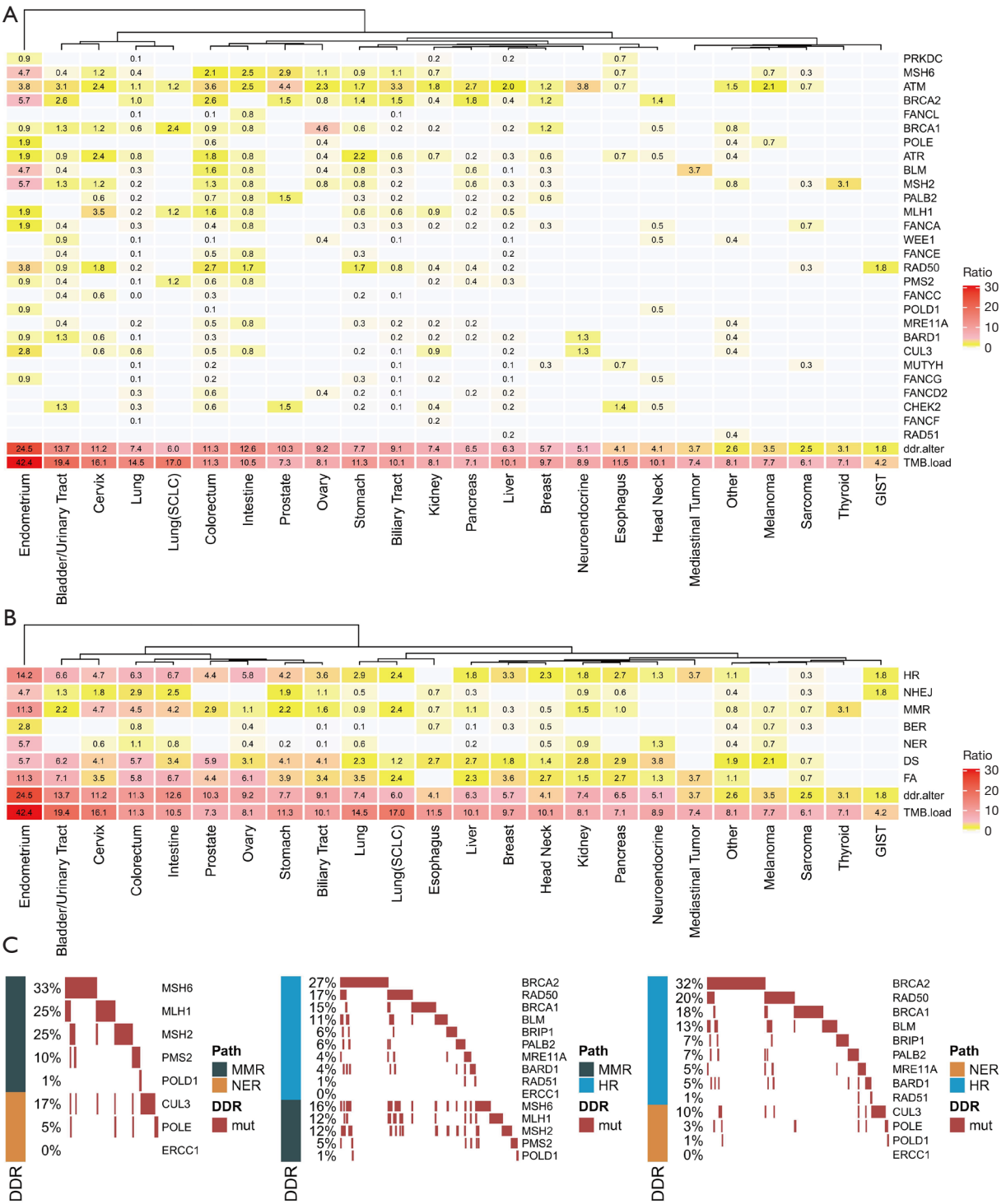

Figure 2 Analysis of somatic DDR gene alterations in pan cancer. (A) Somatic DDR gene mutations are frequently and unevenly distributed types and frequencies of cancer types. (B) Somatic DDR pathway mutations are frequently and unevenly distributed types and frequencies of cancer types. (C) Mutual exclusion of somatic mutations in different DDR pathways. DDR, DNA damage repair; SCLC, small cell lung cancer; GIST, gastrointestinal stromal tumor; BER, base excision repair; NER, nucleotide excision repair; MMR, mismatch repair; DS, damage sensor; FA, Fanconi anemia; HR, homology-dependent recombination; NHEJ, non-homologous end joining.

Table 1 The sample size across 24 tumors

Tumor	N	Freq (%)
Lung (NSCLC)	2,935	27.655
Liver	1,272	11.985
Colorectum	1,100	10.365
Biliary tract	881	8.301
Stomach	664	6.256
Pancreas	517	4.871
Kidney	482	4.542
Breast	337	3.175
Sarcoma	334	3.147
Ovary	263	2.478
Head neck	230	2.167
Bladder/urinary tract	229	2.158
Cervix	181	1.705
Melanoma	151	1.423
Esophagus	146	1.376
Intestine	121	1.14
Endometrium	107	1.008
Lung (SCLC)	84	0.791
Neuroendocrine	83	0.782
Prostate	79	0.744
GIST	60	0.565
Mediastinal tumor	34	0.32
Thyroid	32	0.302
Others	291	2.742

NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer; GIST, gastrointestinal stromal tumor.

95% CI: 0.54–0.87, P=0.002, *Figure 5B-a*; mOS: 16 *vs.* 11 months, hazard ratio =0.75, 95% CI: 0.57–0.99, P=0.043, *Figure 5B-b*). The del-sDDR^{mut} was associated with a higher TMB than the wild-type DDR wild-type (P<0.001; *Figure 5C-a*). The objective response rate (ORR) to immunotherapy was 30.56% for patients with del-sDDR^{mut} and 20.43% for the wild-type subgroup (P=0.035; *Figure 5C-b*). However, the del-sDDR^{mut} groups were significantly associated with reduced colorectal and melanoma cancer risk in analyzing immunotherapy cohorts in pan-cancer (Figure S3).

Discussion

This study identifies deleterious somatic alterations of the DDR genes in 7.6% and germline alterations in 4.7% of patients in 10,284 patients representing 24 different tumors and an association with TMB MSI-H, and PD-L1 expression.

DDR gene mutations were immanent: about 25% of cancer types showed enrichment of DDR gene somatic mutations. The potential functional consequences of these alterations were readily inferred. For example, the DDR pathway was altered in almost 26% of endometrial carcinoma patients, and the HR and MMR pathway mutation represented 15% and 11.5%, respectively, less than the previous study, which showed that nearly 17% of endometrial carcinoma has dMMR (14). *ATM* and *MSH6* were the top 2 DDR gene somatic alterations in a patient with PCa, while in germline alterations, *BRCA2*, *BARD1*, and *ATM* were the top 3 DDR genes. A retrospective case-case study that included 799 patients showed that inherited mutations in *BRCA1/2* and *ATM* distinguish the risk of lethal and localized PCa and are associated with earlier death age and shorter survival time (25). In this study, we included 5 MMR pathway genes, and *MSH6*, *MSH2*, and *MLH1* alterations were detected in cervix patients, of which more than 4.2% of patients have *MLH1* alterations. Early reports have noted that dysfunctional *MLH1* was associated with chemoresistance and did not prolong survival after neoadjuvant chemotherapy in gastric cancer (26), but has not been reported in cervical cancer. In particular, looking at the genomic signatures of the DDR genes, taking the HR, MMR, BER pathway, for example, we found that the somatic alternation patterns of the DDR gene from different pathways showed a clear mutually exclusive signature across all types of tumors, which was consistent with a previous study (1).

Besides *POLE/POLD1* and three MMR pathway genes, we found that other 15 genes deficiency alterations were also significantly associated with TMB, including *BARD1*, *FANCD2*, *RAD51*, etc. Furthermore, in patients with altered DDR, 5.3% were TMB-H/MSS. This subtype showed that DDR deficiency could cause a hypermutated state in addition to dMMR. In terms of specific genes, the alteration of *ATM*, *ATR*, *BRCA1*, *CHEK2*, *BARD1*, *FANCC*, and *FANCG* was present in most TMB-H/MSS patients. This may be the role of *ATR* in protecting genomic integrity. When *ATR* is damaged, cells cannot cope with genome breakage or mutation, and the normal response of ATR

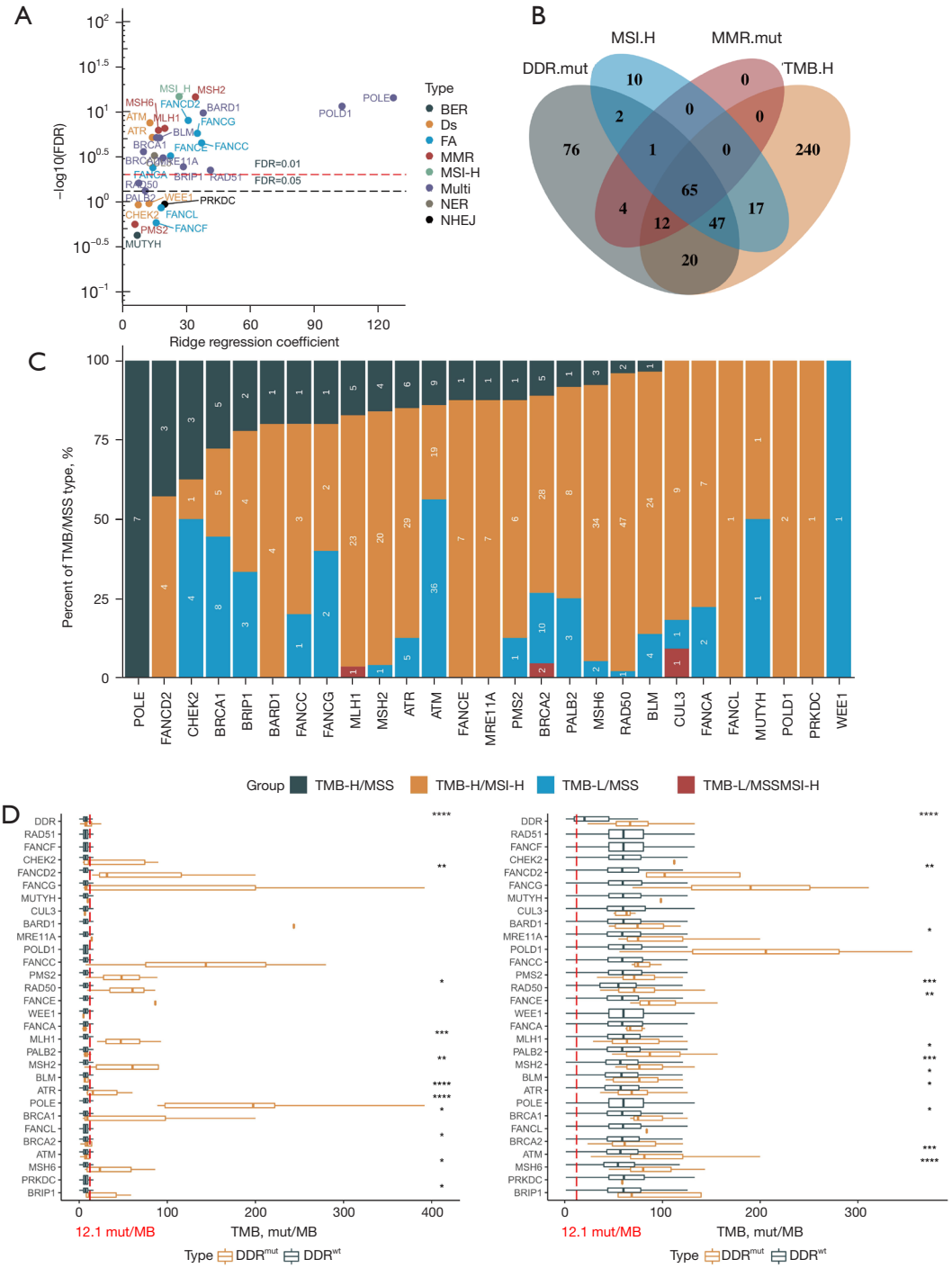

Figure 3 The association of DDR alterations with MSI status and TMB. (A) Magnitude and meaningful volcano plot of the DDR gene ridge. (B) The Venn diagram shows overlapping cases defined by the indicated molecular characteristics. (C) Characteristics of TMB and MSI in cases with GA. (D) MSS samples only: TMB-H prevalence and TMB score box plot for the subgroup with GA in each gene or feature. (E) MSI-H samples only: TMB-H prevalence and TMB score box plot for the subgroup with GA in each gene or feature. * indicates P<0.05, ** indicates P<0.01, *** indicates P<0.001, **** indicates P<0.0001. DDR, DNA damage repair; TMB, tumor mutational burden; MSI, microsatellite instability; MSS, microsatellite stability; GA, genetic algorithm; MSI-H, MSI-high; TMB-H, TMB-high; BER, base excision repair; NER, nucleotide excision repair; MMR, mismatch repair; DS, damage sensor; FA, Fanconi anemia; HR, homology-dependent recombination; NHEJ, non-homologous end joining.

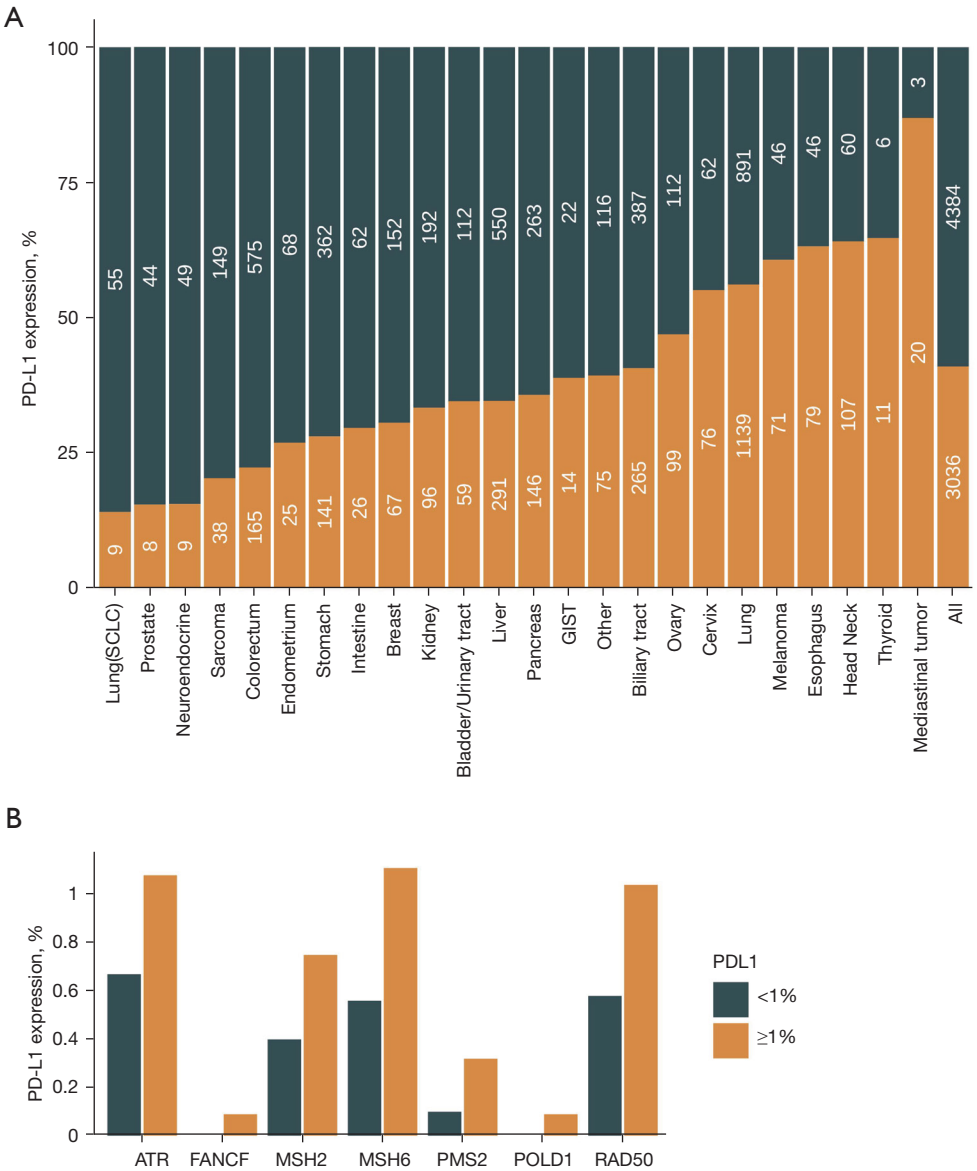

Figure 4 The association of DDR alterations and PD-L1 expression. (A) The prevalence of positive PD-L1 TPS across 24 tumors. (B) The association of positive PD-L1 TPS with DDR genes or pathways alteration. DDR, DNA damage repair; PD-L1, programmed cell death-ligand 1; TPS, tumor proportion score.

to stress phosphorylates transducers, including CHK (27). This indicating MSS patients with DDRmut potentially had higher TMB. The optimal genomic environment for targeting *ATR* is unclear. We can assume that the *ATR* mutant/*TP53* mutant with elevated TMB may be the best candidate gene for the combination of *ATR* inhibitor and immunotherapy (28).

We found that the dysfunction of *ATR*, *BLM*, *CHEK2*, *MSH6*, and *PMS2* was associated with PD-L1 expression. ICIs therapy has revolutionized cancer treatment, resulting in significant and long-lasting clinical benefits, despite a small number of patients. Apart from MSI status and TMB, PD-L1 expression by IHC has been related to responses to ICI, and the pembrolizumab monotherapy has been approved for patients with PD-L1 expression ≥1% (29).

Besides *BRCA1/2*, we also found 19 DDR genes with recurrent germline alteration. Few studies have been reported evaluating the correlation between germline

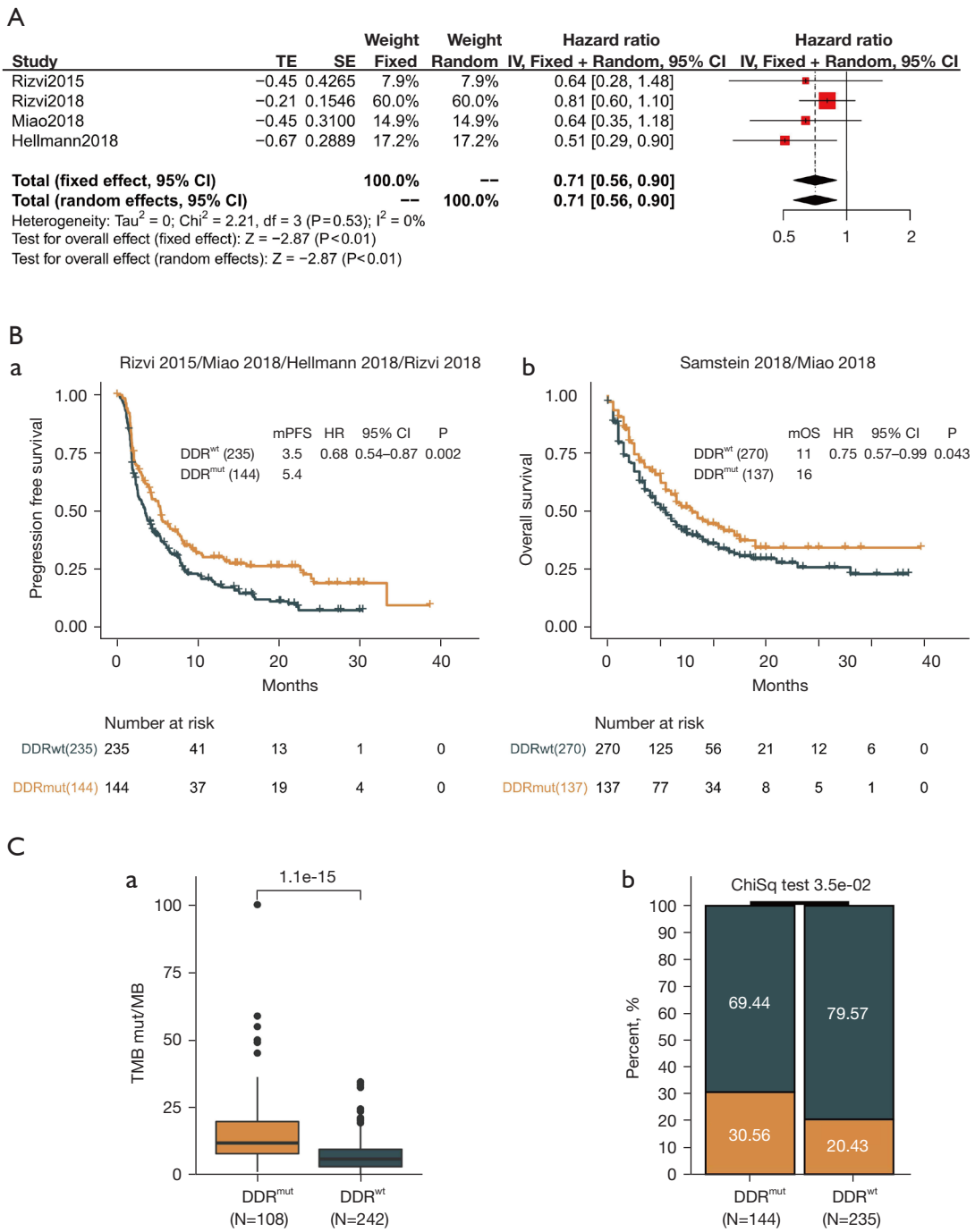

Figure 5 The association of del-sDDR^{mut} and immunotherapy in NSCLC. (A) Meta-analysis of survival for del-sDDR^{mut} in NSCLC among five independent immunotherapy cohorts. (B-a) Correlation between del-sDDR^{mut} and PFS in immunotherapy cohorts. (B-b) Correlation between del-sDDR^{mut} and OS in immunotherapy cohorts. (C-a) Correlation between del-sDDR^{mut} and TMB in different groups. (C-b) The ORR of immunotherapy in different groups. del-sDDR^{mut}, deleterious somatic DDR mutation; DDR, DNA damage repair; NSCLC, non-small cell lung cancer; PFS, progression-free survival; OS, overall survival; mPFS, median PFS; mOS, median OS; TMB, tumor mutational burden; ORR, objective response rate.

mutations and different cancers, especially for DDR genes. In a study investigating the germline mutation status of DDR genes (gDDRm) in 98 patients with bladder cancer (BCa) by Na *et al.* (30), germline mutations in DDR genes were associated with BCa risk and poor prognosis. Similarly, gDDRm, especially *BRCA1/2* and *ATM* alterations, have been reported to be associated with aggressive disease and poor survival of PCa (25,31). Furthermore, it has shown an increased risk of progression to castration resistance in patients with *de novo* metastatic and castration sensitive PCa (mCSPC), who may earlier benefit from the treatment of poly (ADP-ribose) polymerase (PARP) inhibitors or platinum-based chemotherapy if the status of gDDRm was identified at the diagnosis (9,32,33). Furthermore, germline mutations in DDR genes cause sensitivity to PARP inhibitors in advanced solid tumors, including ovarian, pancreatic, and breast cancer (34,35).

DDR genes play a crucial role in maintaining genome stability (36). TMB, which is a predictive biomarker for response to PD-(L)1 inhibitors, may reflect the level of genomic instability at the nucleotide level. In this study, we observed a trend of greater TMB in tumors harboring DDR somatic mutations than germline mutations in pan-cancer, which was consistent with a previous study (37). In our study, the median TMB for each tumor type with somatic mutations ranged from 5.65 mut/Mb in sarcoma cancer to 51.61 mut/Mb in endometrium cancer, while the range for germline mutations ranged from 2.82 mut/Mb in melanoma to 10.48 mut/Mb in small cell lung cancer (SCLC). It should be noted that, except for one elderly patient with pancreatic cancer with relatively high TMB, whose somatic alternation of *POLE* was known as variants of unknown significance (VOUS), patients with higher levels of TMB in the colorectum, endometrium, and ovarian cancer all have somatic *POLE* alternations defined as pathogenic or very similar pathogenic mutations (MUT/VLM). Further, these patients with relatively high TMB in each tumor type, which suggested that these patients in each tumor type might respond well to immunotherapy.

There is growing evidence indicating that DDR may potentially predict the clinical benefits of immunotherapy. DDR consists of 7 pathways, including MMR, BER, HR, NER, FA, DS, and NHEJ (38). The widely known evidence for a relationship between DDR deficiency and ICB treatment response involves the MMR pathway. dMMR/MSI-H is the first FDA-approved pan-cancer biomarker (39). *POLE/POLD1*, in the BER pathway, are potential biomarkers of genomic instability and response to immunotherapy among different types of cancer (40). However, the administration of ICB in HR-deficient cancers has shown contradictory results. In melanoma patients treated with PD-1 inhibitors, *BRCA2* mutations were accumulated in responders (41). However, a phase 1b study showed that *BRCA* status was not associated with the efficacy of anti-PD-L1 therapy in patients with recurrent or refractory ovarian cancer (42). Regarding the FA, DS, NER, and NHEJ pathways, no clinical trials of ICB have been reported in cancer with NER/NHEJ/FA/DS deficiency.

Meanwhile, DDR deficiency has also been studied as an integral biomarker for the application of ICB. A recent study suggested that del-sDDR^{mut} was associated with longer progression-free and survival of ICB in metastatic urothelial carcinoma (43). Co-mutation in the DDR gene (HRR-MMR or HRR-BER) was found to be a promising predictor of ICB response for future clinical application (13). In our study, in five public NSCLC cohorts, we observed that the del-sDDR^{mut} group had better clinical results than the del-sDDR^{wt} group. This phenomenon can also be observed in colorectal cancer and melanoma.

This is the first and largest study to systematically analyze the alteration of DDR genes in pan-cancers in the Chinese population. Our study has two limitations. First, although the 31 DDR genes in our study contain most core genes in the DDR pathway, it is worth expanding the list of DDR genes. Second, the most detailed pathological information was unavailable, so we could only classify tumors according to location.

Conclusions

We analyzed the DDR pathway in 10,284 samples of 24 tumor types in the Chinese population, and identified deficiency somatic and germline alteration of 31 DDR genes based on an NGS 381 genes panel; explored the relationship between DDR changes and TMB, MSI-H and PD-L1; and also provided insights into the potential application of DDR defects in cancer risk and drug development, and exploration of biomarkers for immunotherapy.

Acknowledgments

Funding: This study was supported by National Natural Science Foundation of China (82003384), Fundamental Research Funds for the Central Universities (2020CDJYGSG003), Chongqing Talents Program (CQYC2020058022), Scientific Research Foundation from

Chongqing University (02210011044110), and Innovation and Attracting Talents Program for College and University ("111" Project, B06023), China.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-5449

Data Sharing Statement: Available at https://dx.doi.org/10.21037/atm-21-5449

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-5449). WX, YC, YZ, JZ, XZ, MH were employed by the 3D Medicines Inc., Shanghai, China. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures performed in this study involving human were in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the ethics board of committee of Jieyang Yuedong Cancer Hospital (No. B2021-1-01). The study was a retrospective study and individual consent for the analysis was waived.

References

1. 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.e6.
2. Swisher EM, Lin KK, Oza AM, et al. Rucaparib in relapsed, platinum-sensitive high-grade ovarian carcinoma (ARIEL2 Part 1): an international, multicentre, open-label, phase 2 trial. Lancet Oncol 2017;18:75-87.
3. Litton JK, Rugo HS, Ettl J, et al. Talazoparib in Patients with Advanced Breast Cancer and a Germline BRCA Mutation. N Engl J Med 2018;379:753-63.
4. Mukherjee S, Asaithamby A. Pancreatic cancer patients with germline BRCA mutations can benefit from olaparib treatment. Transl Cancer Res 2020;9:2154-6.
5. Mouw KW, Goldberg MS, Konstantinopoulos PA, et al. DNA Damage and Repair Biomarkers of Immunotherapy Response. Cancer Discov 2017;7:675-93.
6. Gourley C, Balmaña J, Ledermann JA, et al. Moving From Poly (ADP-Ribose) Polymerase Inhibition to Targeting DNA Repair and DNA Damage Response in Cancer Therapy. J Clin Oncol 2019;37:2257-69.
7. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. Science 2017;355:1152-8.
8. Abida W, Patnaik A, Campbell D, et al. Rucaparib in Men With Metastatic Castration-Resistant Prostate Cancer Harboring a BRCA1 or BRCA2 Gene Alteration. J Clin Oncol 2020;38:3763-72.
9. Mateo J, Carreira S, Sandhu S, et al. DNA-Repair Defects and Olaparib in Metastatic Prostate Cancer. N Engl J Med 2015;373:1697-708.
10. Norquist BM, Brady MF, Harrell MI, et al. Mutations in Homologous Recombination Genes and Outcomes in Ovarian Carcinoma Patients in GOG 218: An NRG Oncology/Gynecologic Oncology Group Study. Clin Cancer Res 2018;24:777-83.
11. Goldstein M, Kastan MB. The DNA damage response: implications for tumor responses to radiation and chemotherapy. Annu Rev Med 2015;66:129-43.
12. Teo MY, Bambury RM, Zabor EC, et al. DNA Damage Response and Repair Gene Alterations Are Associated with Improved Survival in Patients with Platinum-Treated Advanced Urothelial Carcinoma. Clin Cancer Res 2017;23:3610-8.
13. Wang Z, Zhao J, Wang G, et al. Comutations in DNA Damage Response Pathways Serve as Potential Biomarkers for Immune Checkpoint Blockade. Cancer Res 2018;78:6486-96.
14. 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.
15. Overman MJ, McDermott R, Leach JL, et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2

study. Lancet Oncol 2017;18:1182-91.
16. Su D, Zhang D, Chen K, et al. High performance of targeted next generation sequencing on variance detection in clinical tumor specimens in comparison with current conventional methods. J Exp Clin Cancer Res 2017;36:121.
17. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 2009;25:1754-60.
18. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med 2015;17:405-24.
19. Siegel RL, Miller KD, Fuchs HE, et al. Cancer Statistics, 2021. CA Cancer J Clin 2021;71:7-33.
20. 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.
21. Rizvi H, Sanchez-Vega F, La K, et al. Molecular Determinants of Response to Anti-Programmed Cell Death (PD)-1 and Anti-Programmed Death-Ligand 1 (PD-L1) Blockade in Patients With Non-Small-Cell Lung Cancer Profiled With Targeted Next-Generation Sequencing. J Clin Oncol 2018;36:633-41.
22. Miao D, Margolis CA, Vokes NI, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. Nat Genet 2018;50:1271-81.
23. Hellmann MD, Nathanson T, Rizvi H, et al. Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. Cancer Cell 2018;33:843-852.e4.
24. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. Nat Genet 2019;51:202-6.
25. Na R, Zheng SL, Han M, et al. Germline Mutations in ATM and BRCA1/2 Distinguish Risk for Lethal and Indolent Prostate Cancer and are Associated with Early Age at Death. Eur Urol 2017;71:740-7.
26. Hashimoto T, Kurokawa Y, Takahashi T, et al. Predictive value of MLH1 and PD-L1 expression for prognosis and response to preoperative chemotherapy in gastric cancer. Gastric Cancer 2019;22:785-92.
27. Yazinski SA, Zou L. Functions, Regulation, and Therapeutic Implications of the ATR Checkpoint Pathway. Annu Rev Genet 2016;50:155-73.
28. Minchom A, Aversa C, Lopez J. Dancing with the DNA damage response: next-generation anti-cancer therapeutic strategies. Ther Adv Med Oncol 2018;10:1758835918786658.
29. Mok TSK, Wu YL, Kudaba I, et al. Pembrolizumab versus chemotherapy for previously untreated, PD-L1-expressing, locally advanced or metastatic non-small-cell lung cancer (KEYNOTE-042): a randomised, open-label, controlled, phase 3 trial. Lancet 2019;393:1819-30.
30. Na R, Wu Y, Jiang G, et al. Germline mutations in DNA repair genes are associated with bladder cancer risk and unfavourable prognosis. BJU Int 2018;122:808-13.
31. Castro E, Goh C, Olmos D, et al. Germline BRCA mutations are associated with higher risk of nodal involvement, distant metastasis, and poor survival outcomes in prostate cancer. J Clin Oncol 2013;31:1748-57.
32. Wei Y, Wu J, Gu W, et al. Prognostic Value of Germline DNA Repair Gene Mutations in De Novo Metastatic and Castration-Sensitive Prostate Cancer. Oncologist 2020;25:e1042-50.
33. Pomerantz MM, Spisák S, Jia L, et al. The association between germline BRCA2 variants and sensitivity to platinum-based chemotherapy among men with metastatic prostate cancer. Cancer 2017;123:3532-9.
34. Kaufman B, Shapira-Frommer R, Schmutzler RK, et al. Olaparib monotherapy in patients with advanced cancer and a germline BRCA1/2 mutation. J Clin Oncol 2015;33:244-50.
35. Fong PC, Boss DS, Yap TA, et al. Inhibition of poly(ADP-ribose) polymerase in tumors from BRCA mutation carriers. N Engl J Med 2009;361:123-34.
36. Tubbs A, Nussenzweig A. Endogenous DNA Damage as a Source of Genomic Instability in Cancer. Cell 2017;168:644-56.
37. Ileana Dumbrava EE, Brusco L, Daniels MS, et al. Pathogenic variants in DNA damage response (DDR) genes in patients with advanced solid tumors. J Clin Oncol 2017;35:abstr 11567.
38. Lord CJ, Ashworth A. The DNA damage response and cancer therapy. Nature 2012;481:287-94.
39. Zhao P, Li L, Jiang X, et al. Mismatch repair deficiency/microsatellite instability-high as a predictor for anti-PD-1/PD-L1 immunotherapy efficacy. J Hematol Oncol 2019;12:54.
40. Chen K, Zhao H, Shi Y, et al. Perioperative Dynamic Changes in Circulating Tumor DNA in Patients with Lung Cancer (DYNAMIC). Clin Cancer Res 2019;25:7058-67.
41. Hugo W, Zaretsky JM, Sun L, et al. Genomic and Transcriptomic Features of Response to Anti-PD-1

Therapy in Metastatic Melanoma. Cell 2016;165:35-44.
42. Disis ML, Taylor MH, Kelly K, et al. Efficacy and Safety of Avelumab for Patients With Recurrent or Refractory Ovarian Cancer: Phase 1b Results From the JAVELIN Solid Tumor Trial. JAMA Oncol 2019;5:393-401.
43. Teo MY, Seier K, Ostrovnaya I, et al. Alterations in DNA Damage Response and Repair Genes as Potential Marker of Clinical Benefit From PD-1/PD-L1 Blockade in Advanced Urothelial Cancers. J Clin Oncol 2018;36:1685-94.

(English Language Editor: J. Chapnick)

Cite this article as: Xiao Y, Lu D, Lei M, Xie W, Chen Y, Zheng Y, Wang C, Zhao J, Zhu Z, Zhao X, Huang M, Lin Y, Li Z, Yang L. Comprehensive analysis of DNA damage repair deficiency in 10,284 pan-cancer study. Ann Transl Med 2021;9(22):1661. doi: 10.21037/atm-21-5449

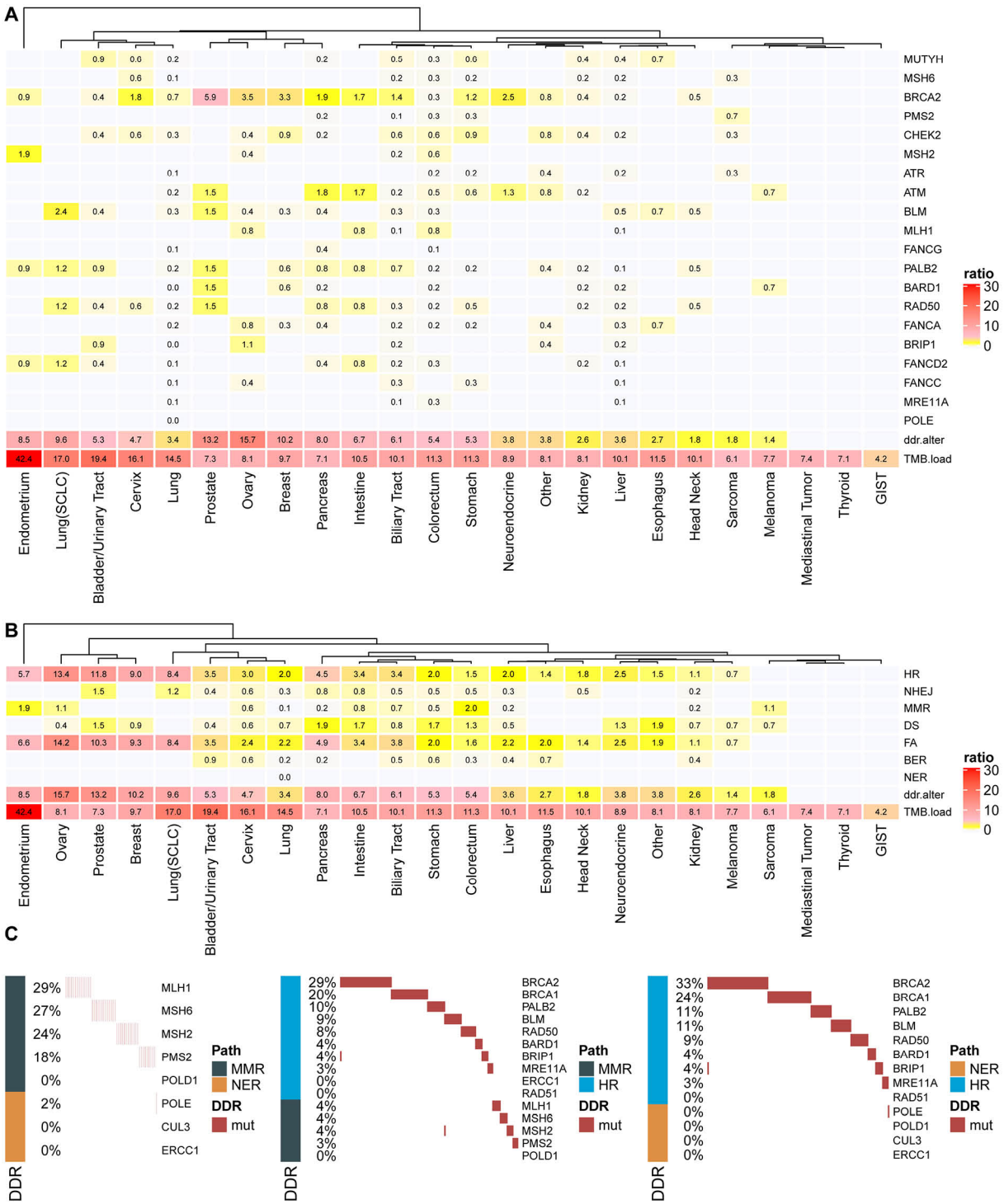

Figure S1 Analysis of germline DDR gene alterations in pan cancer. (A) Germline DDR gene alterations are frequent and non-uniformly distributed by type and frequency across cancer types. (B) Germline DDR pathway alterations are frequent and non-uniformly distributed by type and frequency across cancer types. (C) Mutual exclusion of germline mutations in different DDR pathways. DDR, DNA damage repair; BER, base excision repair; NER, nucleotide excision repair; MMR, mismatch repair; DS, damage sensor; FA, Fanconi anemia; HR, homology-dependent recombination; NHEJ, non-homologous end joining.

https://dx.doi.org/10.21037/atm-21-5449

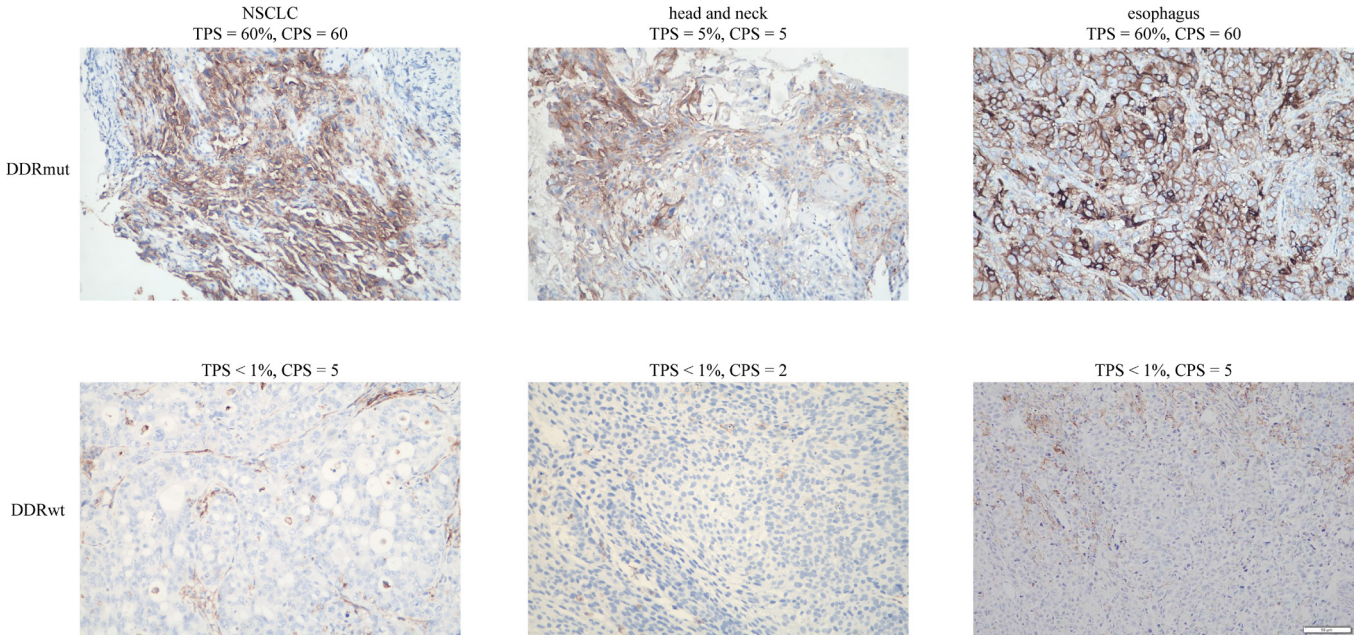

Figure S2 The representative images of PD-L1 IHC in different cancers with or without DDR mutation. PD-L1, programmed cell death-ligand 1; IHC, immunohistochemistry; DDR, DNA damage repair; NSCLC, non-small cell lung cancer; TPS, tumor proportion score; CPS, combined positive score.

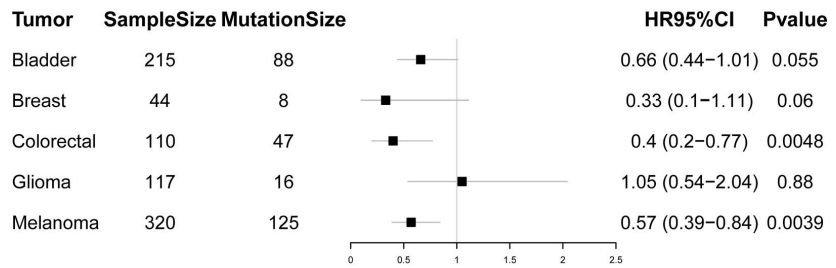

Figure S3 The result of survival analysis for del-sDDRmut in pan-cancer immunotherapy cohorts. del-sDDRmut, deleterious somatic DDR mutation; DDR, DNA damage repair.

https://dx.doi.org/10.21037/atm-21-5449

Table S1 Annotation of 31 DDR genes to specific DDR pathways

Gene description									DDR pathway membership						
Entrez gene ID	Gene symbol	Gene description	Alias (selected)	Additional comments	Approved sy	Approved name	HGNC ID	Location	BER	NER	MMR	FA	HR	NHEJ	DS
4361	*MRE11A*	MRE11 homolog A, double strand break repair nuclease			*MRE11*	MRE11 homolog, double strand break repair nuclease	HGNC:723 0	11q21					MRE11A	MRE11A	
5591	*PRKDC*	Protein kinase, DNA-activated, catalytic polypeptid			*PRKDC*	protein kinase, DNA-activated, catalytic polypeptide	HGNC:941 3	8q11.21						PRKDC	
10111	*RAD50*	RAD50 double strand break repair protein			*RAD50*	RAD50 double strand break repair protein	HGNC:981 6	5q31.1					RAD50	RAD50	
4292	*MLH1*	mutL homolog 1			*MLH1*	mutL homolog 1	HGNC:712 7	3p22.2			MLH1				
4436	*MSH2*	mutS homolog 2			*MSH2*	mutS homolog 2	HGNC:732 5	2p21-p16.3			MSH2				
2956	*MSH6*	mutS homolog 6			*MSH6*	mutS homolog 6	HGNC:732 9	2p16.3			MSH6				
5395	*PMS2*	PMS1 homolog 2, mismatch repair system component			*PMS2*	PMS1 homolog 2, mismatch repair system component	HGNC:912 2	7p22.1			PMS2				
5424	*POLD1*	Polymerase (DNA directed), delta 1, catalytic subunit		Replication	*POLD1*	DNA polymerase delta 1, catalytic subunit	HGNC:917 5	19q13.3	POLD1	POLD1	POLD1		POLD1		
472	*ATM*	ATM serine/threonine kinase			*ATM*	ATM serine/threonine kinase	HGNC:795	11q22.3							ATM
545	*ATR*	ATR serine/threonine kinase			*ATR*	ATR serine/threonine kinase	HGNC:882	3q23							ATR
580	*BARD1*	BRCA1 associated RING domain 1			*BARD1*	BRCA1 associated RING domain 1	HGNC:952	2q35				BARD1	BARD1		
641	*BLM*	Bloom syndrome, RecQ helicase-like			*BLM*	Bloom syndrome RecQ like helicase	HGNC:105 8	15q26.1				BLM	BLM		
672	*BRCA1*	Breast cancer 1, early onset			*BRCA1*	BRCA1, DNA repair associated	HGNC:110 0	17q21.31				BRCA1	BRCA1		
675	*BRCA2*	Breast cancer 2, early onset			*BRCA2*	BRCA2, DNA repair associated	HGNC:1101	13q13.1				BRCA2	BRCA2		
83990	*BRIP1*	BRCA1 interacting protein C-terminal helicase 1	BACH1, FANCJ, OF		*BRIP1*	BRCA1 interacting protein C-terminal helicase 1	HGNC:204 73	17q23.2				BRIP1	BRIP1		
1111	*CHEK1*	Checkpoint kinase 1			*CHEK1*	Checkpoint kinase 1	HGNC:192 5	11q24.2							CHEK1
11200	*CHEK2*	Checkpoint kinase 2			*CHEK2*	Checkpoint kinase 2	HGNC:166 27	22q12.1							CHEK2
8452	*CUL3*	Cullin 3			*CUL3*	Cullin 3	HGNC:255 3	2q36.2		CUL3					
2067	*ERCC1*	Excision repair cross-complementation group 1			*ERCC1*	ERCC excision repair 1, endonuclease non-catalytic subunit	HGNC:343 3	19q13.32		ERCC1		ERCC1	ERCC1		
2175	*FANCA*	Fanconi anemia, complementation group A			*FANCA*	Fanconi anemia complementation group A	HGNC:358 2	16q24.3				FANCA			
2176	*FANCC*	Fanconi anemia, complementation group C			*FANCC*	Fanconi anemia complementation group C	HGNC:358 4	9q22.32				FANCC			
2177	*FANCD2*	Fanconi anemia, complementation group D2			*FANCD2*	Fanconi anemia complementation group D2	HGNC:358 5	3p25.3				FANCD2			
2178	*FANCE*	Fanconi anemia, complementation group E			*FANCE*	Fanconi anemia complementation group E	HGNC:358 6	6p21.31				FANCE			
2188	*FANCF*	Fanconi anemia, complementation group F			*FANCF*	Fanconi anemia complementation group F	HGNC:358 7	11p14.3				FANCF			
2189	*FANCG*	Fanconi anemia, complementation group G			*FANCG*	Fanconi anemia complementation group G	HGNC:358 8	9p13.3				FANCG			
55120	*FANCL*	Fanconi anemia, complementation group L			*FANCL*	Fanconi anemia complementation group L	HGNC:207 48	2p16.1				FANCL			
4595	*MUTYH*	mutY DNA glycosylase			*MUTYH*	mutY DNA glycosylase	HGNC:752 7	1p34.1	MUTYH						
79728	*PALB2*	Partner and localizer of BRCA2	FANCN		*PALB2*	Partner and localizer of BRCA2	HGNC:261 44	16p12.2				PALB2	PALB2		
5426	*POLE*	Polymerase (DNA directed), epsilon, catalytic subunit		Replication	*POLE*	DNA polymerase epsilon, catalytic subunit	HGNC:917 7	12q24.33	POLE	POLE					
5888	*RAD51*	RAD51 recombinase	FANCR		*RAD51*	RAD51 recombinase	HGNC:981 7	15q15.1				RAD51	RAD51		
7465	*WEE1*	WEE1 G2 checkpoint kinase			*WEE1*	WEE1 G2 checkpoint kinase	HGNC:127 61	11p15.4							WEE1

DDR, DNA damage repair; BER, base excision repair; NER, nucleotide excision repair; MMR, mismatch repair; DS, damage sensor; FA, Fanconi anemia; HR, homology-dependent recombination; NHEJ, non-homologous end joining.

https://dx.doi.org/10.21037/atm-21-5449

Table S2 The baseline clinical of the 10,284 patients

Characteristics	Overall, n=10,2841	Biliary tract, n=8,501	Bladder/urinary tract, n=2,261	Breast, n=3,321	Cervix, n=1,691	Colorectum, n=1,0971	Endometrium, n=1,061	Esophagus, n=1,461	GIST, n=571	Head neck, n=2,201	Intestine, n=1,191	Kidney, n=4,581	Liver, n=1,2371	Lung, n=2,8761	Lung (SCLC), n=831	Mediastinal tumor, n=271	Melanoma, n=1,431	Neuroendocrine, n=791	Other, n=2,641	Ovary, n=2,611	Pancreas, n=5,121	Prostate, n=681	Sarcoma, n=2,841	Stomach, n=6,381	Thyroid, n=321	DDR^{wt}, n=9,0661	DDR^{mut}, n=1,2181	P value
Age, n [%]																												0.7
<60	5,407 [53]	399 [47]	67 [30]	268 [81]	130 [77]	588 [54]	55 [52]	64 [44]	36 [63]	125 [57]	66 [55]	287 [63]	862 [70]	1,167 [41]	25 [30]	18 [67]	78 [55]	52 [66]	171 [65]	162 [62]	250 [49]	14 [21]	196 [69]	305 [48]	22 [69]	4,774 [53]	633 [52]	
≥60	4,877 [47]	451 [53]	159 [70]	64 [19]	39 [23]	509 [46]	51 [48]	82 [56]	21 [37]	95 [43]	53 [45]	171 [37]	375 [30]	1,709 [59]	58 [70]	9 [33]	65 [45]	27 [34]	93 [35]	99 [38]	262 [51]	54 [79]	88 [31]	333 [52]	10 [31]	4,292 [47]	585 [48]	
Sex, n [%]																												0.066
Female	4,062 [39]	388 [46]	61 [27]	328 [99]	169 [100]	463 [42]	106 [100]	26 [18]	21 [37]	53 [24]	48 [40]	129 [28]	157 [13]	1,026 [36]	16 [19]	12 [44]	69 [48]	26 [33]	109 [41]	260 [100]	196 [38]	0 [0]	147 [52]	236 [37]	16 [50]	3,551 [39]	511 [42]	
Male	6,222 [61]	462 [54]	165 [73]	4 [1.2]	0 [0]	634 [58]	0 [0]	120 [82]	36 [63]	167 [76]	71 [60]	329 [72]	1,080 [87]	1,850 [64]	67 [81]	15 [56]	74 [52]	53 [67]	155 [59]	1 [0.4]	316 [62]	68 [100]	137 [48]	402 [63]	16 [50]	5,515 [61]	707 [58]	
MSS type, n [%]																												<0.001
MSI-H	220 [2.2]	19 [2.4]	5 [2.3]	0 [0]	7 [4.4]	71 [6.8]	24 [24]	0 [0]	1 [1.8]	2 [1.0]	7 [6.3]	5 [1.1]	9 [0.8]	19 [0.7]	0 [0]	0 [0]	1 [0.7]	1 [1.4]	2 [0.8]	6 [2.5]	5 [1.0]	3 [4.6]	0 [0]	33 [5.4]	0 [0]	55 [0.6]	165 [14]	
MSI-L	20 [0.2]	1 [0.1]	1 [0.5]	2 [0.7]	0 [0]	1 [<0.1]	0 [0]	0 [0]	0 [0]	1 [0.5]	0 [0]	0 [0]	1 [<0.1]	9 [0.3]	0 [0]	0 [0]	1 [0.7]	1 [1.4]	0 [0]	0 [0]	0 [0]	0 [0]	0 [0]	1 [0.2]	1 [3.2]	15 [0.2]	5 [0.4]	
MSS	9,555 [98]	760 [97]	211 [97]	300 [99]	151 [96]	979 [93]	78 [76]	138 [100]	54 [98]	204 [99]	104 [94]	441 [99]	1,136 [99]	2,794 [99]	80 [100]	25 [100]	138 [99]	70 [97]	251 [99]	234 [98]	477 [99]	62 [95]	259 [100]	579 [94]	30 [97]	8,558 [99]	997 [85]	
PD-L1, n [%]																												0.039
<1%	4,384 [59]	387 [59]	112 [65]	152 [69]	62 [45]	575 [78]	68 [73]	46 [37]	22 [61]	60 [36]	62 [70]	192 [67]	550 [65]	891 [44]	55 [86]	3 [13]	46 [39]	49 [84]	116 [61]	112 [53]	263 [64]	44 [85]	149 [80]	362 [72]	6 [35]	3,868 [60]	516 [56]	
≥1%	3,036 [41]	265 [41]	59 [35]	67 [31]	76 [55]	165 [22]	25 [27]	79 [63]	14 [39]	107 [64]	26 [30]	96 [33]	291 [35]	1,139 [56]	9 [14]	20 [87]	71 [61]	9 [16]	75 [39]	99 [47]	146 [36]	8 [15]	38 [20]	141 [28]	11 [65]	2,629 [40]	407 [44]	
TMB (mut/Mb), median [IQR]	6 [4, 10]	6 [4, 9]	10 [6, 17]	6 [4, 9]	6 [4, 14]	8 [6, 10]	8 [6, 22]	8 [6, 11]	3 [2, 4]	6 [4, 9]	6 [4, 9]	5 [3, 7]	6 [5, 9]	7 [4, 13]	11 [8, 16]	5 [2, 6]	4 [3, 6]	5 [3, 7]	4 [2, 7]	6 [3, 7]	5 [3, 6]	5 [2, 6]	3 [2, 5]	6 [4, 10]	2 [2, 6]	6 [4, 9]	9 [6, 21]	<0.001
TMB group, n [%]																												<0.001
< Median	5,027 [52]	415 [54]	107 [50]	159 [53]	78 [50]	561 [54]	57 [56]	72 [53]	28 [51]	111 [54]	61 [55]	232 [52]	576 [51]	1,388 [50]	40 [51]	14 [56]	67 [50]	36 [51]	128 [51]	132 [56]	256 [54]	40 [62]	137 [53]	316 [52]	16 [52]	4,637 [54]	390 [34]	
≥ Median	4,650 [48]	359 [46]	107 [50]	140 [47]	78 [50]	483 [46]	45 [44]	63 [47]	27 [49]	96 [46]	50 [45]	213 [48]	564 [49]	1,375 [50]	39 [49]	11 [44]	66 [50]	35 [49]	123 [49]	104 [44]	221 [46]	25 [38]	120 [47]	291 [48]	15 [48]	3,880 [46]	770 [66]	

GIST, gastrointestinal stromal tumor; SCLC, small cell lung cancer; DDR, DNA damage repair; MSS, microsatellite stability; MSI, microsatellite instability; MSI-H, MSI-high; MSI-L, MSI-low; PD-L1, programmed cell death-ligand 1; TMB, tumor mutational burden; mut/Mb, mutations/Mb.

https://dx.doi.org/10.21037/atm-21-5449