Somatic mutations in homologous recombination pathway predict favorable prognosis after immunotherapy across multiple cancer types

Zaoqu Liu

Zhengzhou University First Affiliated Hospital

Jing Li

Zhengzhou University First Affiliated Hospital

Chunguang Guo

Zhengzhou University First Affiliated Hospital

Hui Xu

Zhengzhou University First Affiliated Hospital

Taoyuan Lu

Zhengzhou First People Hospital

Libo Wang

Zhengzhou First People Hospital

Long Liu

Zhengzhou First People Hospital

Xinwei Han (✉ fcchanxw@zzu.edu.cn)

Zhengzhou University First Affiliated Hospital <https://orcid.org/0000-0003-4407-4864>

Research

Keywords: homologous recombination, immune checkpoint inhibitor, tumor mutation burden, neoantigen load, microsatellite instability, tumor microenvironment

Posted Date: June 2nd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-556047/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: Homologous recombination (HR) pathway was recently implicated in modifying the tumor immune microenvironment and thus might serve as a candidate biomarker of immune checkpoint inhibitor (ICI) treatment. We aimed to comprehensively explore the clinical and molecular significance of HR mutations in ICI treatment across multiple cancer types.

Methods: We analyzed the clinical and genomic data of 1,752 ICI-treated patients and 4,605 non-ICI-treated patients from cBioPortal, TCGA, and ICGC databases. The impacts of HR mutations on tumor mutations burden (TMB), neoantigen load (NAL), microsatellite instability (MSI), immune molecular characteristics, and multi-omics events were further investigated.

Results: HR mutations correlated with improved prognosis in ICI-treated bladder cancer (BLCA), colorectal cancer (CRC), and non-small-cell lung cancer (NSCLC), while in non-ICI cohorts, HR mutations were not associated with prognosis and even suggested unfavorable prognosis. Multivariable Cox analysis demonstrated HR mutations were an independent predictive factor for ICI treatment. Moreover, HR mutations could accurately predict high level of TMB, NAL, and MSI, and displayed the more prevalent incidence than TMB-high, NAL-high, and microsatellite instability-high (MSI-H), which indicated that HR mutations might be an ideal surrogate for TMB, NAL, and MSI estimation. HR mutations were also proven to correlate with the tumor microenvironment, immunity characteristics, immune checkpoints profiles, and substantial multi-omics alteration events.

Conclusions: HR mutations are predictive of improved clinical outcomes in BLCA, CRC, and NSCLC treated with ICI instead of non-ICI, and might be a promising surrogate for TMB, NAL, and MSI estimation in ICI treatment.

Introduction

Over the past decade, immunotherapy has driven tremendous progress in the treatment of solid tumors, such as non-small-cell lung cancer (NSCLC), colorectal cancer (CRC), and bladder cancer (BLCA), melanoma, and renal cell carcinoma[1-5]. As representatives of immunotherapy, immune checkpoint inhibitors (ICIs) aim to help the immune system recognize and attack cancer cells by acting on the primary targets including programmed death-ligand 1 (*PD-L 1*), programmed death 1 (*PD-1*), and cytotoxic T-lymphocyte-associated protein 4 (*CTLA-4*)[6]. Nevertheless, only a small fraction of patients yield considerable clinical benefit from ICI treatment[7]. Substantial efforts have been put into identifying candidate biomarkers to facilitate the clinical selection of patients for ICI treatment, such as *PD-L 1* expression[1, 8], tumor mutation burden (TMB)[9], neoantigen load (NAL)[4], tumor microenvironment (TME)[10], and mismatch repair deficiency (dMMR)/microsatellite instability-high (MSI-H)[11]. Some of these approaches have received encouraging feedback, but each has limitations in terms of clinical utility[12].

The spatiotemporal heterogeneity, dynamic changes, test platform uniformities, and cutoff threshold were crucial challenges for the clinical application of *PD-L1* expression[12]. The measure of TMB and NAL is expensive and complicated, and as continuous variables, they also lack unified definition and standard for cutoff values[13]. High density of CD8+ tumor-infiltrating lymphocytes (TILs) in the TME indicates is associated with clinical benefit from ICI treatment, but the distribution overlap of baseline CD8+ T cell between responders and non-responders limits the establishment of absolute cutoff values[12, 14]. The dMMR/MSI-H tumors display high level of genome instability and harbor substantial somatic mutations that encode potential tumor neoantigens, which are thus likely to be immunogenic and triggering overexpression of immune checkpoints. Pembrolizumab, an anti-PD-1 monoclonal antibody, is a new standard of care for dMMR/MSI-H tumors[15]. However, the dMMR/MSI-H phenotype only accounts for less than 5% of solid tumors, hindering its clinical utilization[16].

Recently, several studies have demonstrated that homologous recombination deficiency (HRD) could reprogram the TME via elevating the infiltration abundance of CD3+ TILs and CD68+ tumor associated macrophages (TAMs)[14, 17]. In addition, HRD tumors could promote the production of cytosolic DNA and activation of the cGAS-STING signaling axis, which is involved in tumor immune recognition[18]. Thus, HRD might serve as a candidate biomarker for evaluating the response to ICI treatment. HR is an essential and accurate form for repairing DNA double-strand breaks (DSBs) generated from DNA interstrand cross-link resolution, and its deficiency is initially described as heterozygous germline mutations in BRCA1/2[19]. The advent in next-generation sequencing (NGS) has revealed that somatic mutations in HR genes are more prevalent than germline mutations and emerge as a novel biomarker for assessing HRD[20]. However, whether the HR somatic mutations are a promising indicator of ICI treatment remains unclear.

Thus, the current study aims to explore the relationship between HR mutations and clinical outcomes of ICI-treated and non-ICI-treated tumors, and further investigate and discuss their impacts on TMB, NAL, microsatellite instability (MSI), immune molecular characteristics, and multi-omics events, providing a comprehensive and systematic reference for preclinical rationale and clinical translation.

Method And Material

Data sources

A total of 1,752 cancer patients treated with ICI and 4,605 non-ICI-treated patients were enrolled. Additional file 2: Table S1 summarized the data sources and details of this study.

MSK ICI cohort. The MSK ICI cohort contained 1,661 patients in 10 cancer types, who received at least one dose of ICI therapy[9]. The clinical information and DNA sequencing data (MSK-IMPACT sequencing) were retrieved from the cBioPortal (www.cbioportal.org). The following cancer types were included: CRC (n =110), melanoma (n =320), bladder cancer (BLCA, n =215), NSCLC (n =350), renal cell carcinoma (n =151), esophagogastric cancer (n =126), head and neck cancer (HNCA, n =139), breast cancer (n =44),

glioma (n =177), cancer of unknown primary (CUP, n=88). The endpoint event was the overall survival (OS) of patients after receiving ICI treatment.

NSCLC ICI cohort. Data from a total of 91 NSCLC patients was derived from the following studies: Hellmann and colleagues[21] (n =75) and Rizvi and colleagues[22] (n =16). Patients in Hellmann cohort were treated with *PD-1* plus *CTLA-4* blockade, and patients in Rizvi cohort were treated with *PD-1* blockade. All patients were performed with whole-exome sequencing. The endpoint events were the durable benefit and progression-free survival (PFS) of patients after receiving ICI treatment. The durable clinical benefit (DCB) was defined as partial or stable response lasting >6 months.

Non-ICI cohorts for BLCA, CRC, and NSCLC. We enrolled a total of 11 non-ICI cohorts, as follows, three BLCA cohorts including TCGA-BLCA (n =404), Hikmat-BLCA[23] (n =30), and Sfakianos-BLCA[24] (n =77), three CRC cohorts including TCGA-CRC (n =501), ICGC-COCA-CN (n =309), and Yaeger-CRC[25] (n =1134), and five NSCLC cohorts including TCGA-NSCLC (TCGA-LUAD and TCGA-LUSC, n =971), MSKCC-2020-LUAD (n =604), ICGC-LUSC-KR (n =170), Chen-LUAD[26] (n =305), and PDX-NSCLC (n =100). All patients did not receive ICI treatment and have both clinical information and DNA sequencing data. The endpoint event was the OS.

Multi-omics data for BLCA, CRC, and NSCLC. The multi-omics data for BLCA, CRC, and NSCLC, including HumanMethylation450 array, copy number alteration (CNA) data, whole-exome sequencing data, and transcriptome profiling data, were derived from TCGA portal (<https://portal.gdc.cancer.gov>).

Assessments of TMB, NAL, and MSI

TMB was defined as the total non-silent somatic mutation counts in coding regions, encompassing missense, nonsense, frame shift insertion, frame shift deletion, in-frame insertion, in-frame deletion, and splice site mutation. Data in the MSK ICI cohort was directly retrieved from cBioPortal online, which was processed via MSK-IMPACT. The TCGA pan-cancer cohort contained 10,114 patients across 32 cancer types, and the whole-exome sequencing data was obtained from TCGA portal. The “maftools” R package was utilized to process mutation data and calculate the TMB of each patient. Neoantigens of 5,935 solid tumors measured by TCGA official were available in TCIA database (<https://tcia.at/neoantigens>)[27]. TMB or NAL was categorized into high and low groups according to the top quartile[28]. The MSI of 5,930 solid tumors across 18 tumor types examined by Hause and colleagues was also collected[29].

Mutational status of HR genes

The “maftools” software was employed to process the DNA sequencing data. Based on the recommend of Memorial Sloan Kettering Cancer Center (MSKCC), HR mutations were defined as any non-silent mutations in 17 recommended genes, including *ATM*, *BAP1*, *BARD1*, *BLM*, *BRCA1*, *BRCA2*, *BRIP1*, *CHEK2*, *ABRAXAS1*, *FANCA*, *FANCC*, *NBN*, *PALB2*, *RAD50*, *RAD51*, *RAD51C*, and *RTEL1*[30, 31]. According to the presence or absence of mutations in those genes, patients were segmented into HR mutation (HR-Mut) and HR wild-type (HR-WT) subgroups.

Gene set enrichment analysis

To reveal the potential molecular mechanisms underlying the HR phenotypes, the gene set enrichment analysis (GSEA) algorithm implemented in “clusterProfiler” R package was performed to identify dramatically enriched terms related to cancer Hallmark pathways (h.all.v7.2.symbols.gmt). Permutations were set to 1000 to obtain a normalized enrichment score (NES). Gene sets with false discovery rate (FDR) <0.05 were considered to be significantly enriched.

TME characterization analysis

Gene expression profiles were utilized to decipher the TME characterization of tumor samples with multiple bioinformatics tools. The “ESTIMATE” R package was utilized to infer the fraction of stromal and immune fraction in solid tumor samples, and generated immune, stromal scores, and tumor purity. Based on the rationale that immunity within tumors is a dynamic process, Karasaki and colleagues[32] have proposed an immunogram for the cancer-immunity cycle (CIC) depicted by eight axes of the immunogram score (IGS): IGS1, T cell immunity; IGS2, tumor antigenicity; IGS3, priming and activation (activated dendritic cell enrichment); IGS4, trafficking and infiltration; IGS5, recognition of tumor cells; IGS6, inhibitory cells; IGS7, checkpoint expression; and IGS8, inhibitory molecules. The gene sets of IGS1-IGS8 were retrieved from a previous study[32]. The single-sample gene-set enrichment analysis (ssGSEA) approach was leveraged to measure the IGS, and the immunogram radar displayed the mean IGS value of two HR phenotypes. To further evaluate the infiltration abundance of immune cell populations in tumor tissues, we applied two different tools, including TIMER and CIBERSORT algorithms.

Delineate the immune checkpoints profiles

To obtain the normalized gene expression value, the “DESeq2” R package was employed to performed variance stabilized transformations on the raw read counts (HTseq) from TCGA database. The “ComBat” algorithm was used to reduce the likelihood of batch effects from non-biological technical biases (Additional file 1: Fig. S1). Next, we compared the expression differences of 27 immune checkpoint members, including B7-CD28 family (*PD-L1*, *PD-L2*, *PD-1*, *CDLA4*, *CD276*, *HHLA2*, *ICOS*, *ICOSLG*, *TMIGD2*, and *VTCN1*)[33], the TNF superfamily (*BTLA*, *CD27*, *CD40*, *CD40LG*, *CD70*, *TNFRSF18*, *TNFRSF4*, *TNFRSF9*, and *TNFSF14*)[34], and several other molecules (*ENTPD1*, *FGL1*, *HAVCR2*, *IDO1*, *LAG3*, *NCR3*, *NT5E*, and *SIGLEC15*)[35, 36].

Characteristics of multi-omics alterations

Additional file 1: Fig. S2 displays the detailed information regarding the data processing in this study.

For mutation analysis, using the VarScan pipeline and MutSigCV algorithm, we retained genes with mutation frequency >5% and MutSigCV q-value <0.05, and performed differential analysis with the threshold of FDR <0.05 and |log2 Fold change| >1.

For CNA analysis, the DNACopy pipeline used Affymetrix SNP 6.0 array data to identify genomic regions that are repeated and infer the copy number of these repeats. The copy number values were further transformed into segment mean values, which were represented as $\log_2(\text{copy-number}/2)$. With -0.3 and 0.3 as cut-off points, genes were marked as deletion (<-0.3), neutral ($-0.3 \sim 0.3$), and amplification (>0.3). GISTIC 2.0 software was applied to define the recurrently amplified and deleted regions. Subsequently, we retained genes with CNA frequency $>5\%$ and GISTIC q-value <0.05 , and performed differential analysis with the threshold of FDR <0.05 .

For methylation profile analysis, using the ChAMP pipeline and MethylMix algorithm, we identified a total of 833 DNA methylation-driven genes and performed differential analysis with the threshold of FDR <0.05 .

For transcriptome analysis, the RNA raw read count was putted into the DESeq2 pipeline. We retained genes whose counts ≥ 1 in $\geq 50\%$ of the samples and performed differential analysis with the threshold of FDR <0.05 and $|\log_2 \text{Fold change}| > 1$.

Identification of key multi-omics events regarding the HR phenotypes

To further identify the key multi-omics events associated with the HR phenotypes, a least absolute shrinkage and selection operator (Lasso) logistic approach was employed to determine the essential genes using “glmnet” R package. To select the optimal value of lambda, we used 10-fold cross-validations with the 1-SE (standard error) criteria. The optimal lambda is the largest value for which the binomial deviance is within one SE of the smallest value of binomial deviance. This process was repeated 1,000 times, and we selected the model with the most occurrences as the final solution. With the key factors determined by the final Lasso model, we further employed the random forest algorithm to ensure the reliability and accuracy of the results. Subsequently, the protein-coding genes within the key factors were further submitted into STRING database (<https://string-db.org>) to construct protein-protein intersection (PPI) network. The “MCODE” plugin implemented in Cytoscape software was utilized to extract the key module from PPI network.

Statistical analysis

All data processing, statistical analysis, and plotting were performed in R 4.0.5 software. The Kaplan-Meier method and the log-rank test were utilized to estimate the different survival between two groups. The “survival” R package was utilized to perform Cox regression analysis. Fisher’s exact test or Pearson’s chi-squared test was applied to compare categorical variables. Continuous variables were compared between two groups through the Wilcoxon rank-sum test or T test. Correlations between two continuous variables were assessed via Pearson’s correlation coefficients. The receiver operating characteristic curve (ROC) was implemented using the “pROC” R package. All statistical tests were two-sided. $P < 0.05$ was regarded as statistically significant.

Results

HR mutations were an independent prognostic factor in ICI-treated BLCA, CRC, and NSCLC

In this study, we characterized the mutational landscape of HR genes in MSK ICI cohort. As displayed in Additional file 1: Fig. S3, the mutation frequency of HR genes varied widely. *ATM* and *BRCA2* mutated more frequently (6.0%) than other HR genes, followed by *BAP1* (4%), *BRCA1* (3.0%), and *RTEL1* (3.0%). The mutation frequency of *FANCC* (0.9%), *ABRAXAS1* (0.7%), *RAD51C* (0.7%), and *RAD51* (0.6%) was obviously lower than other HR genes. According to the presence or absence of mutations in HR genes, patients were divided into HR-Mut and HR-WT subgroups. The proportion of HR-Mut cases in different cancer types was summarized in Fig. 1a, and CRC took the first place (40.0%), followed by melanoma (35.0%), BLCA (29.8%), and NSCLC (25.1%).

To further investigate the impacts of HR mutations on ICI-treated patients, we performed the Kaplan-Meier survival analysis. In MSK ICI pan-cancer cohort, patients with the HR-Mut phenotype displayed significantly prolonged OS compared with the HR-WT phenotype ($P < 0.0001$, Fig. 1b). In the micro-view, there was no survival difference between two phenotypes in melanoma, renal cell carcinoma, esophagogastric cancer, HNSC, breast cancer, glioma, and CUP ($P > 0.05$, Fig. 1c-i). On the other hand, the HR-Mut phenotype showed dramatically improved OS relative to the HR-WT phenotype in BLCA ($P = 0.0058$, Fig. 1j), CRC ($P = 0.0011$, Fig. 1k), and NSCLC ($P = 0.0340$, Fig. 1l). We further enrolled two NSCLC ICI cohorts and observed the same results, patients with HR-Mut have better PFS in Hellmann-NSCLC ($P = 0.0061$, Fig. 1m) and Rizvi-NSCLC ($P = 0.061$, Fig. 1n). The HR-Mut phenotype also exhibited higher fraction of DCB compared with the HR-WT phenotype in Hellmann-HSCLC (71% vs. 41%, $P = 0.022$, Additional file 1: Fig. S4A) and Rizvi-NSCLC (67% vs. 30%, $P = 0.302$, Additional file 1: Fig. S4B). The lack of statistical significance in the Rizvi-NSCLC cohort may be due to the small sample size. In multivariate Cox analysis, the HR phenotypes remained an independent protective factor for ICI-treated patients in BLCA (Hazard ratio (*HR*) = 0.506 [0.288-0.891], $P = 0.018$, Fig. 1o), CRC ($HR = 0.254$ [0.083-0.775], $P = 0.016$, Fig. 1p), and NSCLC ($HR = 0.677$ [0.469-0.977], $P = 0.037$, Fig. 1q), after adjusting for confounding variables (including age, sex, tumor type, and TMB). These results suggested that after ICI treatment, the impacts of HR mutations in different cancer types were significantly discrepant, and HR mutations could indicate the clinical outcomes of ICI-treated patients in BLCA, CRC, and NSCLC.

Impacts of HR mutations on non-ICI treated BLCA, CRC, and NSCLC

To examine the impacts of HR mutations on patients with traditional treatments in BLCA, CRC, and NSCLC, we enrolled other 11 non-ICI cohorts to investigate the association between HR mutations and patients prognosis. In BLCA, there was no survival difference between the HR-Mut and HR-WT phenotypes in TCGA-BLCA ($n = 404$, $P = 0.300$, Fig. 2a), Hitmat-BLCA ($n = 30$, $P = 0.350$, Fig. 2b), and Sfakianos-BLCA ($n = 77$, $P = 0.570$, Fig. 2c). In CRC, the median OS of the HR-Mut phenotype was also similar to that of the HR-WT phenotype in TCGA-CRC ($n = 501$, $P = 0.570$, Fig. 2d), ICGC-COCA-CN ($n = 309$, $P = 0.640$, Fig. 2e), and Yaeger-CRC ($n = 1134$, $P = 0.650$, Fig. 2f). Likewise, there was no survival difference between two phenotypes in TCGA-NSCLC ($n = 971$, $P = 0.710$, Fig. 2g), MSKCC-2020-LUAD ($n = 604$, $P = 0.560$, Fig. 2h), and ICGC-LUSC-KR ($n = 170$, $P = 0.620$, Fig. 2i). Patients with the HR-Mut phenotype had relatively lower

but not significant OS in Chen-LUAD ($n=305$, $P=0.063$, Fig. 2j), and significantly lower in PDX-NSCLC ($n=100$, $P=0.035$, Fig. 2k), which were opposite to the above results in ICI cohorts. Overall, HR mutations are indicative of favorable prognosis in BLCA, CRC, and NSCLC treated with ICI instead of non-ICI.

HR mutations were favorable surrogates for TMB and NAL estimation

By analysis the DNA sequencing data from MSK and TCGA pan-cancer cohorts, we found the HR-Mut phenotype had predominantly more single-nucleotide polymorphisms (SNP), insertion or deletion (INDEL), and TMB than the HR-WT phenotype in pan-cancer, BLCA, CRC, and NSCLC cohorts (Fig. 3a-b). The proportion of TMB-high in the HR-Mut phenotype was also substantially higher than the HR-WT phenotype in MSK-pan-cancer (56% vs. 11%, $P<0.0001$, Fig. 3c), MSK-BLCA (59% vs. 10%, $P<0.0001$, Fig. 3d), MSK-CRC (59% vs. 2%, $P<0.0001$, Fig. 3e), MSK-NSCLC (62% vs. 12%, $P<0.0001$, Fig. 3f). These results were further verified in TCGA dataset (TCGA-pan-cancer: 71% vs. 16%, $P<0.0001$; TCGA-BLCA: 48% vs. 11%, $P<0.0001$; TCGA-CRC: 65% vs. 11%, $P<0.0001$; TCGA-NSCLC: 52% vs. 16%, $P<0.0001$) (Fig. 3g-j). The ROC curve that using HR mutations to predict TMB-high also presented relatively high AUC in MSK-pan-cancer (0.746), MSK-BLCA (0.776), MSK-CRC (0.872), MSK-NSCLC (0.760) (Fig. 3k-n). In addition, similar results were validated in TCGA TMB data, with the AUC of 0.706 in TCGA-pan-cancer, 0.728 in TCGA-BLCA, 0.779 in TCGA-CRC, and 0.685 in TCGA-NSCLC (Additional file 1: Fig. S5A-D).

Tumor neoantigens, also termed tumor-specific antigens, derived from the expression of non-silent mutations[27]. Using the neoantigens data from 5,935 patients in TCGA dataset, we further explored the association between HR mutations and NAL. As showed in Additional file 1: Fig. S6, patients with HR Mut possessed dramatically more NAL than those with HR-WT ($P<0.001$). The proportion of NAL-high in the HR-Mut phenotype was also remarkably higher than the HR-WT phenotype in TCGA-pan-cancer (61% vs. 16%, $P<0.0001$, Additional file 1: Fig. S7A), TCGA-BLCA (40% vs. 14%, $P<0.0001$, Additional file 1: Fig. S7B), TCGA-CRC (57% vs. 13%, $P<0.0001$, Additional file 1: Fig. S7C), TCGA-NSCLC (45% vs. 16%, $P<0.0001$, Additional file 1: Fig. S7D). Using the HR phenotypes to assess tumor NAL, the AUCs were 0.683, 0.668, 0.737, and 0.653 in TCGA-pan-cancer, TCGA-BLCA, TCGA-CRC, and TCGA-NSCLC, respectively (Additional file 1: Fig. S7E-H). Taken together, these results confirmed our hypothesis that HR mutations could predict higher TMB and NAL, and displayed accurate performance in BLCA, CRC, and NSCLC. In addition, we exhibited novel evidence that HR mutations may facilitate identify more potential responders to ICI treatment than TMB-high and NAL-high (20% of tumors), which indicated that HR mutations were a favorable surrogate for TMB and NAL estimation.

Association between HR mutations and MSI

Given the latent interactions between HR pathway and mismatch repair (MMR) pathway[28], we revealed the association between two DNA damage repair pathway in different cancer types. As showed in Fig. 4a, there were significant HR/MMR comutation patterns in tumors, such as BLCA, CRC, etc. Therefore, we speculated that HR mutations may have a potential connection with MSI. Using the MSI data from 5,930 patients in TCGA dataset, we observed that nearly all of MSI-H tumors belonged to the HR-Mut phenotype (HR-Mut vs. HR-WT: 19% vs. 1% in TCGA-pan-cancer, 1% vs. 0% in TCGA-BLCA, 50% vs. 3% in TCGA-CRC,

and 2% vs. 0% in TCGA-NSCLC) (Fig. 4b-e). Further ROC analysis revealed that the HR mutations could accurately predict the MSI-H phenotype in TCGA-pan-cancer (AUC =0.805, Fig. 4f), TCGA-BLCA (AUC =0.825, Fig. 4g), TCGA-CRC (AUC =0.855, Fig. 4h), and TCGA-NSCLC (AUC =0.876, Fig. 4i). Previous study had demonstrated that not all microsatellite stable (MSS) tumors respond poorly to ICI treatment[37], moreover, MSI-H tumors only account for less than 5% of solid tumors[16], limiting the clinical utilization. In this study, most MSI-H tumors were HR-Mut, and HR mutations also identified a subset of MSS tumors that may respond to immunotherapy, and importantly, the HR-Mut phenotype covered higher fraction of cancers. These findings suggested the HR phenotypes may be a better biomarker than MSI for clinical translation.

Tumor microenvironment characterization of the HR phenotypes

Since the two HR phenotypes presented significantly different clinical outcomes for ICI-treated BLCA, CRC, and NSCLC, we speculated that HR mutations was closely linked with immune characterizations in the TME. GSEA was applied to decipher the underlying mechanisms in terms of cancer Hallmarks, and 22 pathways dramatically enriched (FDR <0.05, Fig. 5a). We found that the HR-Mut subtype was mainly correlated with cell cycle and interferon response, such as “E2F targets”, “G2M checkpoint”, “Interferon gamma response”, and “Interferon alpha response”. Meanwhile, the HR-WT subtype enriched plenty of stromal activation and tumor metastasis-related terms, such as “Epithelial mesenchymal transition”, “WNT beta catenin signaling”, “TGF beta signaling”, and “Angiogenesis” (Fig. 5a). To further explore the TME, ESTIMATE software was employed to infer the fraction of immune and stromal components. We found the HR-Mut phenotype scored better in the immune score and also displayed lower stromal score and tumor purity, which was in line with the GSEA results (Additional file 1: Fig. S8A). There was no significant difference in the correlations between immune and stromal scores in the two HR phenotypes (HR-Mut: Pearson’s $r=0.726$; HR-WT: Pearson’s $r=0.746$), and the tumor purity was lower as the immune and stromal scores were increasing in both subgroups (Fig. 5b).

Cancer-immunity cycle and immune checkpoints profiles in the HR phenotypes

In cancer-immunity cycle, we found four key immune-active steps including T cell immunity, tumor antigenicity, trafficking and infiltration, and recognition of tumor cells were significantly enhanced in the HR-Mut phenotype versus the HR-WT phenotype. Meanwhile, the HR-Mut phenotype also displayed the substantial enhancement of checkpoint expression, implying an immune-hot but suppressive microenvironment (Fig. 5c and Additional file 1: Fig. S8B). To gain more detailed insights into this issue, we applied the TIMER approach to quantify the infiltration abundance of different immune cell populations. The infiltration abundance of CD4+ and CD8+ T cells was higher in HR-Mut tumors than HR-WT tumors, which further demonstrated T cell enrichment was a dominant factor in the TME of HR-Mut tumors (Fig. 5d). Meanwhile, the deconvolution algorithm CIBERSORT validated these results, and also presented that the HR-Mut phenotype had superior infiltration of activated Nature Killer cells, while the HR-WT phenotype harbored more resting cells such as resting dendritic cells and resting CD4+ T memory cells (Fig. 5e and Additional file 1: Fig. S8C). Notably, the M2/M1 ratio was also higher in HR-WT tumors,

which was a feature of immune-suppression and cancer development (Fig. 5e and Additional file 1: Fig. S8C). In addition, the analysis was extended on the expression profiles of 27 immune checkpoint molecules. We observed the expression level of *CD274*, *PDCD1*, and *CTLA4* was superior in the HR-Mut subtype, which may explain its better response to ICI treatment (Fig. 5f and Additional file 1: Fig. S8B). There were also some other molecules that were significantly upregulated in HR-Mut tumors, such as *FGL1*, *IDO1*, *LAG3*, and *TNFRSF18*. Meanwhile, *CD40LG*, *HHLA2*, and *NT5E* were highly expressed in HR-WT tumors, which may be the potential targets to improve the prognosis of patients with the HR-WT phenotype.

Multi-omics events in the HR phenotypes

To explore the multi-omics characteristics between the HR phenotypes, multiple bioinformatics tools were utilized to analysis genomics, epigenetics, and transcriptomics data in BLCA, CRC, and NSCLC. The processing pipeline was showed in Additional file 1: Fig. S2. As expected, HR-Mut tumors displayed superior genomic alterations than HR-WT tumors. In total, we identified 380 genes with frequent hypermutations in HR-Mut tumors, such as *ATM*, *MYCBP2*, *DNAH17*, *BRCA2*, *ITPR2*, etc (Fig. 6a and Additional file 2: Table S2). Meanwhile, patients with the HR-Mut phenotype showed more amplifications of 11q13.3 (*FADD*), 8q22.3 (*GRHL2*), and 1q23.3 (*NOS1AP*, *NIT1*, *DEDD*, *PFDN2*, *KLHDC9*, and *ARHGAP30*), as well as deletions of 4q22.1 (*CCSER1*) and 16q23.1 (*WWOX*) (Fig. 6b and Additional file 2: Table S2). The two subgroups also presented significant discrepancy in methylation profiles, with 177 dramatically different DNA methylation-driven genes (Additional file 2: Table S2). Moreover, a total of 90 mRNAs, 14 lncRNAs, and 10 miRNAs were differentially expressed between the two HR phenotypes (Additional file 1: Fig. S9A-C and Additional file 2: Table S2).

For these different multi-omics events, we further used the Lasso algorithm to extract key drivers in the HR phenotypes. With 10-fold cross-validation, we totally performed 1,000 iterations and included 8 Lasso models for further screening. As illustrated in Fig. 6c, the 45-gene model had the highest frequency of 357 compared with other models. These 45 variables with beta coefficients that were not zero were considered to be key multi-omics events, including 38 mutant genes, 4 DNA methylation-driven genes, 2 mRNAs and 1 lncRNA with dysregulated expression (Fig. 6d). This suggested that the difference between the two phenotypes was mainly concentrated in mutational events. Subsequently, we performed another algorithm (random forest) to fit a novel model using these key factors. The results showed a perfectly correct prediction with an AUC of 1, which further validated the reliability and accuracy of the Lasso results (Additional file 1: Fig. S9D-E). The random forest also assessed the importance of each variable, and as expected, *ATM* and *BRCA2*, which determined the HR phenotypes and possessed the highest mutation frequency, were the most critical (Additional file 1: Fig. S9F). In addition, we constructed the PPI network using 44 altered protein-coding genes, further revealing the interactive relationships among the key genes in the whole network (Fig. 6e). The MCODE plugin (degree algorithm) implemented in Cytoscape software was utilized to extract the most significant module from PPI network (Fig. 7f). Taken together, these key genes constituted the multi-omics characteristics of the HR phenotypes, and might be the potential targets for improving clinical outcomes.

Discussion

HR is a DNA repair pathway of clinical interest owing to tumor cells with HRD was sensitive to poly-(ADP)-ribose polymerase inhibitor (PARPi)[19, 38]. Similar to immunotherapy, PARPi has also made groundbreaking success over the past decade. Monotherapy with PARPi represents a therapeutic strategy based on the mechanism of synthetic lethality that displays significant clinical benefit in ever-growing list of malignancies[38]. However, a considerable number of patients receiving either ICI or PARPi treatment alone provide dismal clinical outcomes and cannot benefit, and thus a rationale to combine these therapy strategies has emerged[39]. An important question is whether PARPi-sensitive HRD tumors also benefit from ICI treatment, rather than generating unmeaningful immune-related adverse events (irAEs). Despite several studies have reported that HRD could enhance the tumor immunogenicity and reprogram the TME, lack of data demonstrating the predictive value of HR mutations in ICI treatment may restrict its clinical application[14, 17, 18, 39]. Our study systematically explored the relationship between HR mutations and clinical outcomes of ICI-treated and non-ICI-treated tumors, and further demonstrated their impacts on TMB, NAL, MSI, immune molecular characteristics, and multi-omics events, providing a full-scale reference for preclinical rationale and clinical translation.

According to the MSK-IMPACT sequencing data, we noticed that the impacts of HR mutations in different cancer types were significantly discrepant. There was no survival difference between two phenotypes in melanoma, renal cell carcinoma, esophagogastric cancer, HNSC, breast cancer, glioma, and CUP, which indicated HR mutations were not indicative of prognosis in these cancers treated with ICI. Conversely, after receiving ICI treatment, HR mutations suggested favorable OS in BLCA, CRC, and NSCLC. Two other independent cohorts including Hellmann-NSCLC and Rizvi-NSCLC also demonstrated better PFS and DCB in HR-Mut tumors. More importantly, we proved that the HR phenotypes was an independent protective factor in these three cancers with ICI treatment. In addition, using data from 11 non-ICI cohorts, we observed that there was no survival difference in most cohorts, and even some cohorts such as Chen-LUAD and PDX-NSCLC showed opposite results. Therefore, HR mutations are predictive of favorable clinical outcomes in BLCA, CRC, and NSCLC treated with ICI instead of non-ICI.

HRD in tumor cells give rise to the high level of genome instability and substantial non-silent mutations that encode potential tumor neoantigens. Thus, HRD tumors tend to drive the heightened immunogenicity and trigger the enrichment of lymphocytes[14, 17, 18, 39], which might explain its better response to ICI treatment. By analyzing the DNA sequencing data from MSK-pan-cancer and verifying the findings in TCGA-pan-cancer, we observed that HR-Mut tumors suggested increased numbers of SNP, INDEL, and TMB, thereby generated more NAL. Tumor neoantigens derived from a minority of somatic mutations can be processed by the antigen presentation system and appear on the cell surface, further activating T cells[13, 27]. Previous studies have demonstrated that high TMB and NAL are associated with improved survival in patients receiving ICI treatment across multiple solid cancers[9, 12, 13, 27]. The HR-Mut phenotype accounted for the majority of TMB-high and NAL-high tumors in pan-cancer, BLCA, CRC, and NSCLC. Moreover, we found there were significant HR/MMR comutation patterns across multiple cancer types, and the proportion of MSI-H tumors were almost entirely distributed in the HR-Mut phenotype,

particularly in CRC. Several clinical trials have demonstrated that dMMR/MSI-H is dramatically associated with long-term immunotherapeutic response and favorable prognosis in CRC and noncolorectal carcinomas treated with ICI[11]. Although our data showed that fewer MSI-H tumors were present in BLCA and NSCLC, all MSI-H tumors belonged to the HR-Mut phenotype. ROC analysis further proved HR mutations can accurately predict TMB-high, NAL-high, and MSI-H phenotypes. As well-studied biomarkers, TMB, NAL, and MSI have exhibited profound clinical value in assessing the benefit from ICI treatment. However, the unified cutoff value or small percentage may be the major stumbling block in clinical utilization. The HR phenotypes is a binary variable and does not require a cutoff value. Additionally, the HR phenotypes may facilitate identify more potential responders to ICI treatment due to HR-Mut cover higher fraction of tumors than TMB-high, NAL-high and MSI-H. Hence, these findings suggested the HR phenotypes may be a promising surrogate for TMB, NAL, and MSI estimation.

Of note, HR mutations also displayed profound impacts on the TME and immunity. The HR-Mut phenotype was mainly correlated with interferon response and abundant immune milieu, while the HR-WT phenotype was characterized by stromal activation. In 2017, Karasaki and colleagues[32] have proposed a cancer immunogram, a theoretical framework that integrates multiple parameters to tailor ICI treatment. The immunogram is established on the hypothesis that T cell immunity is the ultimate effector mechanism that can be regulated by seven other immune parameters[32]. The five key steps including T cell immunity, tumor antigenicity, trafficking and infiltration, recognition of tumor cells, and checkpoint expression were significantly enhanced in the HR-Mut phenotype versus the HR-WT phenotype. In anti-tumor immunity, the status and abundance of TILs, particularly CD4⁺/CD8⁺ T cells, is considered to be the core determinant of immunotherapeutic efficacy[40, 41]. Hence, the enrichment of CD4⁺/CD8⁺ T cell infiltration in HR-Mut tumors further demonstrated that HR mutations was strongly associated with enhanced immune activity. Furthermore, we observed the expression level of *CD274*, *PDCD1*, and *CTLA4* was superior in the HR-Mut subtype, which was in line with its better response to ICI treatment. There were also some other molecules that were significantly upregulated in HR-Mut tumors, such as *FGL1*, *IDO1*, *LAG3*, and *TNFRSF18*. Meanwhile, *CD40LG*, *HHLA2*, and *NT5E* were highly expressed in HR-WT tumors. Focus on the superior immune checkpoint molecules in two HR phenotypes, the development of new immunotherapeutic strategies could tailor individual treatment and provide backup resources for immunotherapy of two phenotypes. Overall, it is of great interest to clarify the underlying mechanisms of the interaction between HR mutations and microenvironment remodeling.

In addition, based on multiple bioinformatics tools and machine learning algorithms, we identified a total of 45 key multi-omics characteristics derived from HR mutations, including 38 mutant genes, 4 DNA methylation-driven genes, 2 mRNAs and 1 lncRNA with dysregulated expression. This suggested that the multi-omics difference between the two HR phenotypes was mainly concentrated in mutational events. *ATM* and *BRCA2* mutations served as the core variables in random forest, and also accounted for the most prevalent mutations in HR pathway of cancers. The PPI network further revealing the interactive relationships among the key genes in the whole network. The most significant module consisting of *ATM2*, *BRCA2*, *ARID1A*, *RELN*, *EP300*, *SMARCA4*, *EPHA6*, *KDR*, and *PLCB1*, was extracted from PPI network. Mutations in these genes have been all reported to be associated with DNA damage repair and

tumor immune microenvironment remodeling[17, 30, 42-47]. For example, *ARID1A* mutant tumors was characterized by high TMB, MSI-H, and high level of *PD-L1* expression and TILs infiltration, and notably, *ARID1A* mutations could strongly regulated DNA repair pathways[42]. A recent study has demonstrated that *EP300* mutation is dramatically associated with high TMB and advances antitumor immunity in BLCA[44]. Schoenfeld and colleague has reported that *SMARCA4* mutant NSCLC is enriched in patients with *KRAS*, *STK1*, and *KEAP1* mutations, and display long-term outcomes and increased sensitivity to immunotherapy[45]. Moreover, HR mutations were also associated with several epigenetics and transcriptomics driven events, including *SSX1*, *KCNE3*, *KLHDC7A*, *CDR2L* methylations and *F11*, *PSG4*, *AC129926.1* abnormal expression. Briefly, *SSX1* is an epigenetic target that give rise to the disturbances of interacting multiprotein complexes such as the SWItch/Sucrose Non-Fermentable (SWI/SNF) complex in synovial sarcomas[48]; the hypermethylation of *KCNE3* promoter is reported to be a diagnostic and prognostic biomarker in prostate cancer[49]; *CDR2* is the main antigen for the onconeural antibody Yo, which is strongly associated with ovarian cancer[50]. It is notable that the role of *KLHDC7A*, *F11*, *PSG4*, and *AC129926.1* have not previously been reported in cancers, and thus requires further exploration. Taken together, these key genes constituted the multi-omics characteristics of the HR phenotypes, and might be the potential targets for improving clinical outcomes for ICI treatment.

Although the role of HR mutations in ICI treatment is promising, some limitations should be acknowledged. First, all the samples from this study were retrospective. However, MSK cancer center and TCGA are the most reliable public datasets, and the findings were mutually verified in both databases. Second, this study was limited to focus on 9 cancer types, due to the lack of immunotherapeutic information of other cancers in MSK ICI cohort. Further exploration of HR mutations in other solid tumors should be conducted in the future.

Conclusion

In conclusion, HR mutations are predictive of improved clinical outcomes in BLCA, CRC, and NSCLC treated with ICI instead of non-ICI. HR mutations could accurately predict TMB-high, NAL-high, and MSI-H, suggesting that HR mutations may be a promising surrogate for TMB, NAL, and MSI estimation. HR mutations are associated with the TME, immunity characteristics, immune checkpoints profiles, and substantial multi-omics alteration events. Our findings lay a foundation for the clinical application of HR mutations in cancer immunotherapy.

Abbreviations

HR: homologous recombination; ICI: immune checkpoint inhibitor; TMB: tumor mutations burden; NAL: neoantigen load; MSI: microsatellite instability; BLCA: bladder cancer; CRC: colorectal cancer; NSCLC: non-small-cell lung cancer; PD-1: programmed death-ligand 1; PD-1: programmed death 1; CTLA-4: cytotoxic T-lymphocyte-associated protein 4; TME: tumor microenvironment; HRD: homologous recombination deficiency; TAMs: tumor associated macrophages; OS: overall survival; PFS: progression-free survival; DCB: durable clinical benefit; CNA: copy number alteration; GSEA: gene set enrichment analysis; NES:

normalized enrichment score; FDR; false discovery rate; CIC: cancer-immunity cycle; IGS: immunogram score; ssGSEA: single-sample gene-set enrichment analysis; Lasso: least absolute shrinkage and selection operator; PPI; protein-protein intersection; ROC: operating characteristic curve; SNP: single-nucleotide polymorphisms; INDEL: insertion or deletion; MSS: microsatellite stable; PARPi: poly-(ADP)-ribose polymerase inhibitor; irAEs: immune-related adverse events.

Declarations

Ethics approval and consent to participate

Not available.

Consent for publication

The authors confirm that they have obtained written consent from each patient to publish the manuscript.

Availability of data and materials

Public data used in this work can be acquired from the cBioPortal database (<https://www.cbioportal.org>), The Cancer Genome Atlas (TCGA, <http://portal.gdc.cancer.gov>), the International Cancer Genome Consortium (ICGC, <http://dcc.icgc.org>).

Competing interests

The authors declare no competing interests.

Funding

Not applicable.

Authors' contributions

ZQL and XWH designed this work. ZQL integrated and analyzed the data. ZQL wrote this manuscript. ZQL, JL, CGG, HX, TYL, LBW, LL, and XWH edited and revised the manuscript. All authors approved this manuscript.

Acknowledgements

Not applicable.

Author details

¹Department of Interventional Radiology, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China. ²Interventional Institute of Zhengzhou University, Zhengzhou, Henan 450052, China. ³Interventional Treatment and Clinical Research Center of Henan Province, Zhengzhou,

Henan 450052, China. ⁴Department of Endovascular Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China. ⁵Department of Cerebrovascular Disease, Zhengzhou University People's Hospital, Zhengzhou, Henan 450003, China. ⁶Department of Hepatobiliary and Pancreatic Surgery, The First Affiliated Hospital of Zhengzhou University, Zhengzhou, Henan 450052, China.

References

1. Reck M, Rodriguez-Abreu D, Robinson AG, Hui R, Csoszi T, Fulop A, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N Engl J Med*. 2016; 375:1823-1833.
2. Ganesh K, Stadler ZK, Cercek A, Mendelsohn RB, Shia J, Segal NH, et al. Immunotherapy in colorectal cancer: rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol*. 2019; 16:361-375.
3. Bellmunt J, Powles T, Vogelzang NJ. A review on the evolution of PD-1/PD-L1 immunotherapy for bladder cancer: The future is now. *Cancer Treat Rev*. 2017; 54:58-67.
4. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, et al. Genomic and Transcriptomic Features of Response to Anti-PD-1 Therapy in Metastatic Melanoma. *Cell*. 2016; 165:35-44.
5. Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med*. 2015; 373:1803-1813.
6. Mahoney KM, Rennert PD, Freeman GJ. Combination cancer immunotherapy and new immunomodulatory targets. *Nat Rev Drug Discov*. 2015; 14:561-584.
7. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, Adaptive, and Acquired Resistance to Cancer Immunotherapy. *Cell*. 2017; 168:707-723.
8. Herbst RS, Baas P, Kim DW, Felip E, Perez-Gracia JL, Han JY, et al. Pembrolizumab versus docetaxel for previously treated, PD-L1-positive, advanced non-small-cell lung cancer (KEYNOTE-010): a randomised controlled trial. *Lancet*. 2016; 387:1540-1550.
9. Samstein RM, Lee CH, Shoushtari AN, Hellmann MD, Shen R, Janjigian YY, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet*. 2019; 51:202-206.
10. Liu Z, Zhang Y, Shi C, Zhou X, Xu K, Jiao D, et al. A novel immune classification reveals distinct immune escape mechanism and genomic alterations: implications for immunotherapy in hepatocellular carcinoma. *J Transl Med*. 2021; 19:5.
11. Zhao P, Li L, Jiang X, Li Q. Mismatch repair deficiency/microsatellite instability-high as a predictor for anti-PD-1/PD-L1 immunotherapy efficacy. *J Hematol Oncol*. 2019; 12:54.
12. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol*. 2016; 17:e542-e551.
13. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol*. 2019; 30:44-56.

14. Hamid O, Schmidt H, Nissan A, Ridolfi L, Aamdal S, Hansson J, et al. A prospective phase II trial exploring the association between tumor microenvironment biomarkers and clinical activity of ipilimumab in advanced melanoma. *J Transl Med.* 2011; 9:204.
15. Andre T, Shiu KK, Kim TW, Jensen BV, Jensen LH, Punt C, et al. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. *N Engl J Med.* 2020; 383:2207-2218.
16. Cortes-Ciriano I, Lee S, Park WY, Kim TM, Park PJ. A molecular portrait of microsatellite instability across multiple cancers. *Nat Commun.* 2017; 8:15180.
17. Morse CB, Toukatly MN, Kilgore MR, Agnew KJ, Bernards SS, Norquist BM, et al. Tumor infiltrating lymphocytes and homologous recombination deficiency are independently associated with improved survival in ovarian carcinoma. *Gynecol Oncol.* 2019; 153:217-222.
18. Sen T, Rodriguez BL, Chen L, Corte CMD, Morikawa N, Fujimoto J, et al. Targeting DNA Damage Response Promotes Antitumor Immunity through STING-Mediated T-cell Activation in Small Cell Lung Cancer. *Cancer Discov.* 2019; 9:646-661.
19. Venkitaraman AR. Linking the cellular functions of BRCA genes to cancer pathogenesis and treatment. *Annu Rev Pathol.* 2009; 4:461-487.
20. Hoppe MM, Sundar R, Tan DSP, Jeyasekharan AD. Biomarkers for Homologous Recombination Deficiency in Cancer. *J Natl Cancer Inst.* 2018; 110:704-713.
21. Hellmann MD, Nathanson T, Rizvi H, Creelan BC, Sanchez-Vega F, Ahuja A, et al. Genomic Features of Response to Combination Immunotherapy in Patients with Advanced Non-Small-Cell Lung Cancer. *Cancer Cell.* 2018; 33:843-852 e844.
22. Rizvi NA, Hellmann MD, Snyder A, Kvistborg P, Makarov V, Havel JJ, et al. Cancer immunology. Mutational landscape determines sensitivity to PD-1 blockade in non-small cell lung cancer. *Science.* 2015; 348:124-128.
23. Al-Ahmadie HA, Iyer G, Lee BH, Scott SN, Mehra R, Bagrodia A, et al. Frequent somatic CDH1 loss-of-function mutations in plasmacytoid variant bladder cancer. *Nat Genet.* 2016; 48:356-358.
24. Kim PH, Cha EK, Sfakianos JP, Iyer G, Zabor EC, Scott SN, et al. Genomic predictors of survival in patients with high-grade urothelial carcinoma of the bladder. *Eur Urol.* 2015; 67:198-201.
25. Yaeger R, Chatila WK, Lipsyc MD, Hechtman JF, Cercek A, Sanchez-Vega F, et al. Clinical Sequencing Defines the Genomic Landscape of Metastatic Colorectal Cancer. *Cancer Cell.* 2018; 33:125-136 e123.
26. Chen J, Yang H, Teo ASM, Amer LB, Sherbaf FG, Tan CQ, et al. Genomic landscape of lung adenocarcinoma in East Asians. *Nat Genet.* 2020; 52:177-186.
27. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep.* 2017; 18:248-262.
28. Wang Z, Zhao J, Wang G, Zhang F, Zhang Z, Zhang F, et al. Comutations in DNA Damage Response Pathways Serve as Potential Biomarkers for Immune Checkpoint Blockade. *Cancer Res.* 2018; 78:6486-6496.

29. Hause RJ, Pritchard CC, Shendure J, Salipante SJ. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med*. 2016; 22:1342-1350.
30. Park W, Chen J, Chou JF, Varghese AM, Yu KH, Wong W, et al. Genomic Methods Identify Homologous Recombination Deficiency in Pancreas Adenocarcinoma and Optimize Treatment Selection. *Clin Cancer Res*. 2020; 26:3239-3247.
31. Lord CJ, Ashworth A. BRCAness revisited. *Nat Rev Cancer*. 2016; 16:110-120.
32. Karasaki T, Nagayama K, Kuwano H, Nitadori JI, Sato M, Anraku M, et al. An Immunogram for the Cancer-Immunity Cycle: Towards Personalized Immunotherapy of Lung Cancer. *J Thorac Oncol*. 2017; 12:791-803.
33. Janakiram M, Chinai JM, Zhao A, Sparano JA, Zang X. HHLA2 and TMIGD2: new immunotherapeutic targets of the B7 and CD28 families. *Oncoimmunology*. 2015; 4:e1026534.
34. Ward-Kavanagh LK, Lin WW, Sedy JR, Ware CF. The TNF Receptor Superfamily in Co-stimulating and Co-inhibitory Responses. *Immunity*. 2016; 44:1005-1019.
35. Chretien S, Zerdes I, Bergh J, Matikas A, Foukakis T. Beyond PD-1/PD-L1 Inhibition: What the Future Holds for Breast Cancer Immunotherapy. *Cancers (Basel)*. 2019; 11.
36. Wang J, Sanmamed MF, Datar I, Su TT, Ji L, Sun J, et al. Fibrinogen-like Protein 1 Is a Major Immune Inhibitory Ligand of LAG-3. *Cell*. 2019; 176:334-347 e312.
37. Chen EX, Jonker DJ, Loree JM, Kennecke HF, Berry SR, Couture F, et al. Effect of Combined Immune Checkpoint Inhibition vs Best Supportive Care Alone in Patients With Advanced Colorectal Cancer: The Canadian Cancer Trials Group CO.26 Study. *JAMA Oncol*. 2020; 6:831-838.
38. Pilie PG, Gay CM, Byers LA, O'Connor MJ, Yap TA. PARP Inhibitors: Extending Benefit Beyond BRCA-Mutant Cancers. *Clin Cancer Res*. 2019; 25:3759-3771.
39. Peyraud F, Italiano A. Combined PARP Inhibition and Immune Checkpoint Therapy in Solid Tumors. *Cancers (Basel)*. 2020; 12.
40. Kluger HM, Zito CR, Barr ML, Baine MK, Chiang VL, Sznol M, et al. Characterization of PD-L1 Expression and Associated T-cell Infiltrates in Metastatic Melanoma Samples from Variable Anatomic Sites. *Clin Cancer Res*. 2015; 21:3052-3060.
41. Borst J, Ahrends T, Babala N, Melief CJM, Kastenmuller W. CD4(+) T cell help in cancer immunology and immunotherapy. *Nat Rev Immunol*. 2018; 18:635-647.
42. Tokunaga R, Xiu J, Goldberg RM, Philip PA, Seeber A, Battaglin F, et al. The impact of ARID1A mutation on molecular characteristics in colorectal cancer. *Eur J Cancer*. 2020; 140:119-129.
43. Govindan R, Ding L, Griffith M, Subramanian J, Dees ND, Kanchi KL, et al. Genomic landscape of non-small cell lung cancer in smokers and never-smokers. *Cell*. 2012; 150:1121-1134.
44. Zhu G, Pei L, Li Y, Gou X. EP300 mutation is associated with tumor mutation burden and promotes antitumor immunity in bladder cancer patients. *Aging (Albany NY)*. 2020; 12:2132-2141.
45. Schoenfeld AJ, Bandlamudi C, Lavery JA, Montecalvo J, Namakydoust A, Rizvi H, et al. The Genomic Landscape of SMARCA4 Alterations and Associations with Outcomes in Patients with Lung Cancer.

Clin Cancer Res. 2020; 26:5701-5708.

46. Li S, Zhu A, Ren K, Li S, Chen L. IFN β -induced exosomal linc-EPHA6-1 promotes cytotoxicity of NK cells by acting as a ceRNA for hsa-miR-4485-5p to up-regulate NKp46 expression. Life Sci. 2020; 257:118064.
47. Banerjee S, Mohammed A, Wong HR, Palaniyar N, Kamaleswaran R. Machine Learning Identifies Complicated Sepsis Course and Subsequent Mortality Based on 20 Genes in Peripheral Blood Immune Cells at 24 H Post-ICU Admission. Front Immunol. 2021; 12:592303.
48. Hale R, Sandakly S, Shipley J, Walters Z. Epigenetic Targets in Synovial Sarcoma: A Mini-Review. Front Oncol. 2019; 9:1078.
49. Li Y, Meng L, Shi T, Ren J, Deng Q. Diagnosis and prognosis potential of four gene promoter hypermethylation in prostate cancer. Cell Biol Int. 2021; 45:117-126.
50. Raspotnig M, Haugen M, Thorsteinsdottir M, Stefansson I, Salvesen HB, Storstein A, et al. Cerebellar degeneration-related proteins 2 and 2-like are present in ovarian cancer in patients with and without Yo antibodies. Cancer Immunol Immunother. 2017; 66:1463-1471.

Figures

Figure 1

The incidence and survival analysis of HR mutations in ICI cohorts across multiple cancer types. a Frequency of HR mutations in patients with different cancer types. b-l Kaplan–Meier survival curves of OS comparing the HR-Mut and HR-WT phenotypes in ICI-treated pan-cancer (b), melanoma (c), renal cell carcinoma (d), esophagogastric cancer (e), HNCA (f), breast cancer (g), glioma (h), CUP (i), BLCA (j), CRC (k), and NSCLC (l) from MSK ICI cohort. m-n Kaplan–Meier survival curves of PFS comparing the HR-Mut and HR-WT phenotypes in Hellmann-HSCLC (m) and Rizvi-NSCLC (n) cohorts. o-q. Multivariable Cox analysis of the HR phenotypes in MSK-BLCA (o), MSK-CRC (p), and MSK-NSCLC (q).

Figure 2

Survival analysis of HR mutations in non-ICI treated BLCA, CRC, and NSCLC. A-k Kaplan-Meier survival curves of OS comparing the HR-Mut and HR-WT phenotypes in TCGA-BLCA (a), Hikmat-BLCA (b), Sfakianos-BLCA (c), TCGA-CRC (d), ICGC-COCA-CN (e), Yaeger-CRC (f), TCGA-NSCLC (g), MSKCC-2020-LUAD (h), ICGC-LUSC-KR (i), Chen-LUAD (j), PDX-NSCLC (k).

Figure 3

The association between HR mutations and TMB. a Comparison of SNP, INDEL, and TMB between the HR-Mut and HR-WT phenotypes in MSK-pan-cancer, MSK-BLCA, MSK-CRC, and MSK-NSCLC. b Comparison of SNP, INDEL, and TMB between the HR-Mut and HR-WT phenotypes in TCGA-pan-cancer, TCGA-BLCA, TCGA-CRC, and TCGA-NSCLC. c-j. The proportion of TMB-high and TMB-low in the HR-Mut and HR-WT phenotypes in MSK-pan-cancer (c), MSK-BLCA (d), MSK-CRC (e), MSK-NSCLC (f), TCGA-pan-cancer (g), TCGA-BLCA (h), TCGA-CRC (i), and TCGA-NSCLC (j). k-n. ROC curves of HR mutations to

predict higher TMB in MSK-pan-cancer (k), MSK-BLCA (l), MSK-CRC (m), and MSK-NSCLC (n). ****P <0.0001.

Figure 4

The association between HR mutations and MSI. a Fraction of HR/MMR commutations in multiple cancer types. b-e. The proportion of MSI-H and MSS in the HR-Mut and HR-WT phenotypes in TCGA-pancancer (b), TCGA-BLCA (c), TCGA-CRC (d), and TCGA-NSCLC (e). f-i. ROC curves of HR mutations to

predict MSI-H in TCGA-pan-cancer (f), TCGA-BLCA (g), TCGA-CRC (h), and TCGA-NSCLC (i). nsP >0.05, ***P <0.001, ****P <0.0001.

Figure 5

TME and immunity characterization of the HR phenotypes. a GSEA analysis of cancer Hallmarks between the HR-Mut and HR-WT phenotypes. Green points represent pathways significantly related to HR-Mut tumors, orange points represent pathways significantly related to HR-WT tumors, and grey points represent pathways with no significance between the two phenotypes. b Scatterplots between stromal and immune scores with tumor purity gradient were shown, and its correlation coefficient was indicated by each phenotype. The color grading corresponds to the tumor purity, indexed as shown on the color bar

at the top right of the panel. c Radar plots showed that the immunogram patterns of the two phenotypes. The axes of the radar chart were generated with the median IGS for the HR-Mut and HR-WT phenotypes, respectively. d The TIMER algorithm showed the infiltration difference of 6 immune cell populations including B, CD4+ T, CD8+ T, neutrophil, macrophage, and myeloid dendritic cells between two phenotypes. e The CIBORSORT algorithm showed the proportion of 22 immune cells in the HR-Mut and HR-WT phenotypes. f Differences in the distribution of 27 immune checkpoint molecules between two phenotypes. nsP >0.05, *P <0.05, **P <0.01, ***P <0.001, ****P <0.0001.

Figure 6

Multi-omics events in the HR phenotypes. a Significantly different mutated genes between the two phenotypes. Waterfall plot display the top 20 mutations with the highest mutational frequency. b Significantly different amplified and deleted regions between the two phenotypes. c Generation of the eight Lasso model after 1000 iteration. The Lasso model with 45 variables were selected due to its highest frequency of 357 compared to other seven Lasso model. d Based on the seed of 45-gene model,

ten-fold cross-validations to tune the parameter selection in the LASSO model. The two dotted vertical lines are drawn at the optimal values by minimum criteria (left) and 1-SE (standard error) criteria (right). We selected the 1-SE criteria to determine the final model. e the protein-coding genes within the key factors were further submitted into STRING database to construct the PPI network. f The MCODE plugin implemented in Cytoscape software was utilized to extract the key module from PPI network.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure.docx](#)
- [SupplementaryTable.xlsx](#)