



Integrative analysis of homologous recombination repair patterns unveils prognostic signatures and immunotherapeutic insights in breast cancer

Yan-Shuang Li¹ · Hong-Chuan Jiang¹

Received: 17 January 2024 / Revised: 15 February 2024 / Accepted: 22 February 2024

This is a U.S. Government work and not under copyright protection in the US; foreign copyright protection may apply 2024

Abstract

Globally, breast cancer (BC) is the leading cause of female death and morbidity. Homologous recombination repair (HRR) is critical in BC. However, the prognostic role and immunotherapy response of HRR in BC remains to be clarified. Firstly, we identified HRR types in BC samples from the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) dataset (GSE42568) based on 65 HRR genes (HRRGs). A differentially expressed gene (DEG) list for different HRR types was generated. Then, the influences of gene sets composed of these DEGs on biological pathways and BC prognosis were explored. Next, we identified gene clusters based on gene sets composed of DEGs. Genes associated with prognosis for DEGs were identified using univariate Cox regression. Finally, the HRR score was constructed based on genes associated with prognosis. We analyzed how HRR score correlates with tumor mutation burden (TMB), immune cell infiltration (ICI), and immunotherapy response. Three HRR clusters were discovered. HRR subtype A demonstrated decreased infiltration and a high number of immunosuppressive cells with a poor prognosis. DEGs among various HRR types were predominantly enriched in cell cycle and genomic stability-related pathways. The prognostic model based on sixteen DEGs accurately predicted BC prognosis. The HRRGs were differentially expressed in three DEG clusters. TMB, ICI, and immunotherapy responses differed significantly between the high and low HRR groups (HSG, LSG). The HSG was distinguished by a high degree of ICI and low TMB. LSG had a better response to anti-PD-1 or anti-PD-1 and anti-CTLA4 combination therapy. This work revealed that HRR patterns would contribute to predicting prognosis and immunotherapy response in BC, which may benefit patients.

Keywords Breast cancer · Homologous recombination repair · Immunity · Prognosis · Immunotherapy response

Introduction

Breast cancer (BC) is the leading cause of incidence and mortality among women worldwide (Ahmad 2019). The number of new cases of female BC reported worldwide in 2018 reached approximately 2.1 million, with the incidence and mortality being 24.2% and 15%, respectively (Bray et al. 2018). Although there has been much progress in immunotherapy and targeted therapy for BC, the low overall response continues to be a hurdle, and

the prognosis remains poor for patients with recurrent or metastatic BC (Ahmad 2019; Gerber et al. 2010; Varadé et al. 2021). Additionally, BC is a heterogeneous disease, so the prognosis for the same tumor node metastasis (TNM) stage and immunohistochemical subtype can be quite different (Yersal and Barutca 2014). Thus, exploring the cellular and molecular biological mechanisms of BC is essential for its treatment.

Previous studies have identified different BC subtypes according to immune cell infiltration status, DNA methylation, molecular apocrine, and basal-like-enrichment (Shen et al. 2020a; Jézéquel et al. 2019; Zhang et al. 2018), which confirmed the role of immunity and DNA mutations in BC. The tumor microenvironment (TME) was previously implicated in cancer progression and immunotherapy responses (Mittal et al. 2018; Schlam et al. 2021). The accumulation of mutations is a crucial part of cancer development, and

Communicated by: Ewa Ziętkiewicz

✉ Hong-Chuan Jiang
dr.jianghongchuan@gmail.com

¹ Department of Breast Surgery, Beijing Chaoyang Hospital, Capital Medical University, Beijing 100020, China

repairing DNA damage is crucial to maintaining genomic integrity (Reilly et al. 2019). Breaks of DNA double-strand (DSB) in cells can hinder chromosomal replication, causing a chromosomal deletion. Cell death and tumor transformation will occur if it is not repaired promptly. Homologous recombination repair (HRR) is the most accurate and high-fidelity DNA damage repair approach responsible for DSB repair, involving many genes, including BRCA1 and BRCA2 (Neiger et al. 2021). ADP-ribose polymerase (PARP) is crucial for DNA single-strand break repair, mainly through base excision. Both BRCA1 and BRCA2 dysfunction may cause chromosome instability, cell-cycle arrest, and apoptosis. Dysfunction of either BRCA1 or BRCA2 seems to originate from PARP inhibition and can cause persistent DNA damage, which is usually repaired by HRR. PARP inhibitors (PARPi) selectively kill HRR deficiency tumor cells using the principle of synthetic lethality (Prados-Carvajal et al. 2021).

Our study integrated HRR-related genes (HRRGs) and identified HRR-related subtypes in BC based on public databases. After obtaining differentially expressed genes (DEGs) across HRR subtypes, a prognostic model was constructed based on them. Then, the biological functions of these DEGs were explored. We established the HRR score based on DEGs associated with BC prognosis and investigated the association between the HRR score and tumor mutation burden (TMB), immune cell infiltration (ICI), and therapeutic sensitivity. The research may provide an avenue for prognosis and immunotherapy response prediction in BC.

Materials and methods

Breast cancer data retrieval and homologous recombination repair-related genes

One thousand two hundred twelve BC samples were analyzed from two high-throughput platforms: 1091 from the TCGA and 121 from the GEO dataset (GSE42568). GEO database inclusion criteria are as follows. (1) “Breast cancer,” “Gene expression,” and “Homo sapiens” were used as the search keywords. (2) Entry type was set as “Series.” (3) Experiment type was “Expression profiling by array to facilitate subsequent analysis with R software.” (4) Samples involved in the study contain the gene expression profiles of breast cancer and normal breast biopsies. (5) Prognostic information of the patients was recorded. GEO database exclusion criteria are as follows. (1) The sample number is less than 50 cases. (2) RNA sequencing data and prognosis information of patients were incomplete. TCGA samples were downloaded with gene expression data, copy number variations (CNV), somatic mutation characteristics, clinical information, and survival data. GEO dataset samples were

downloaded with annotated platform data and expression matrices. Probe IDs were converted to gene symbols using Perl. Patients with incomplete RNA-seq data and with insufficient survival data were excluded. TCGA gene expression data expressed as Fragments Per Kilobase Million (FPKMs) was transformed to Transcripts Per Million (TPMs). The “ComBat” function was used in the “SVA” package (version 3.50.0) to combine samples for further analysis (Leek et al. 2012). In total, 65 HRR-related genes (HRRGs) were acquired from the Molecular Signature (MsigDB, <http://software.broadinstitute.org/gsea/msigdb>) (Liberzon et al. 2015). Waterfall plots visualized somatic mutations with the “maftools” package (version 2.18.0) in R software in Mayakonda et al. (2018), and HRRGs with CNV were visualized on 23 chromosomes with the “RCircos” package (version 1.2.2) (Zhang et al. 2013).

Unsupervised clustering using homologous recombination repair genes

To identify various HRR-related patterns mediated by HRRGs, the expression of these 65 HRRGs was extracted from the integrated datasets. Finally, we obtained the expression data of 37 HRRGs, because genes from only one expression cohort were removed during the integration process (Stockwell et al. 2017; Hassannia et al. 2019; Bersuker et al. 2019; Doll et al. 2019; Zhu et al. 2021). Hierarchical agglomerative clustering was performed by the “Consensus-ClusterPlus” package (version 1.66.0) (Wilkerson and Hayes 2010). An unsupervised analysis determined the number of clusters and subtypes. This process was repeated 1000 times to ensure a stable clustering.

Gene set variation analysis

We conducted a gene set variation analysis (GSVA) with “GSVA” package (version 1.50.0) to examine the biological mechanisms across HRR subtypes (Hänzelmann et al. 2013). From MSigDB, gene sets from Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were retrieved for GSVA. The pathways significantly associated with HRR subtypes were shown in heatmaps.

Estimation of immune cell infiltration (ICI)

The ICI in each sample was assessed and quantified using single-sample gene-set enrichment analysis (ssGSEA) provided by the “GSVA” package (Hänzelmann et al. 2013). Charoentong’s study which collected information about immune cell marker gene expression was used to generate an enrichment score. Then, the score was used to reflect each immune cell’s relative abundance of infiltration. The ICI variations among HRR subtypes were subsequently examined.

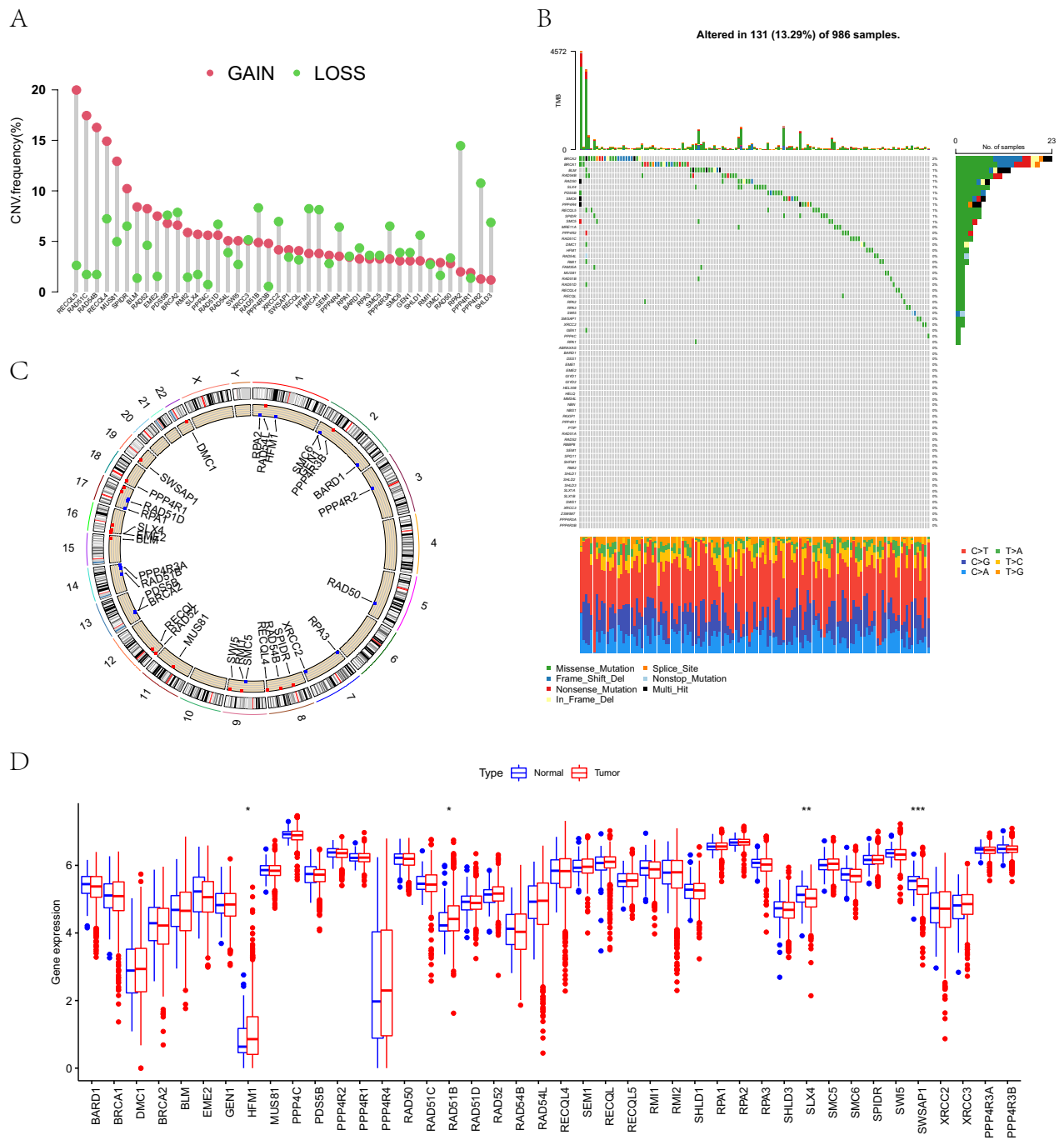


Fig. 1 The CNV frequency of HRRGs and expression differences of HRRGs in normal and tumor tissues in the TCGA-BRCA cohort. **A** The CNV frequency of HRRGs in the TCGA cohort. The points represent the CNV frequency of HRRGs in the TCGA cohort. The height of the column indicates the frequency of CNV change. The green dots indicate the frequency of deletions, and the red ones indicate the amplification frequency. **B** Frequency of HRRG mutations in

1090 patients from the TCGA-BRCA cohort. The columns represent patients, and the different colors at the bottom of the figure represent the proportion of different types of mutations. **C** The position of HRRGs with CNV on 23 chromosomes. **D** The expression differences of HRRGs in normal tissues and tumor tissues. CNV, copy number variation; HRRGs, homologous recombination repair genes; BRCA, breast cancer

Identification of differentially expressed genes among homologous recombination repair subtypes in breast cancer

Unsupervised clustering results were used to separate BC samples into different subtypes based on HRR expression patterns. We screened differentially expressed

genes (DEGs) across subtypes using “Limma” package (version 3.46.0) in R software (Ritchie et al. 2015). An adjusted false discovery rate (FDR) of 0.05 was used as a threshold, as well as the absolute value of LogFC was greater than 1. The common DEGs up-regulated or downregulated of distinct subtypes were found via drawing Venn diagrams. The common DEGs from different

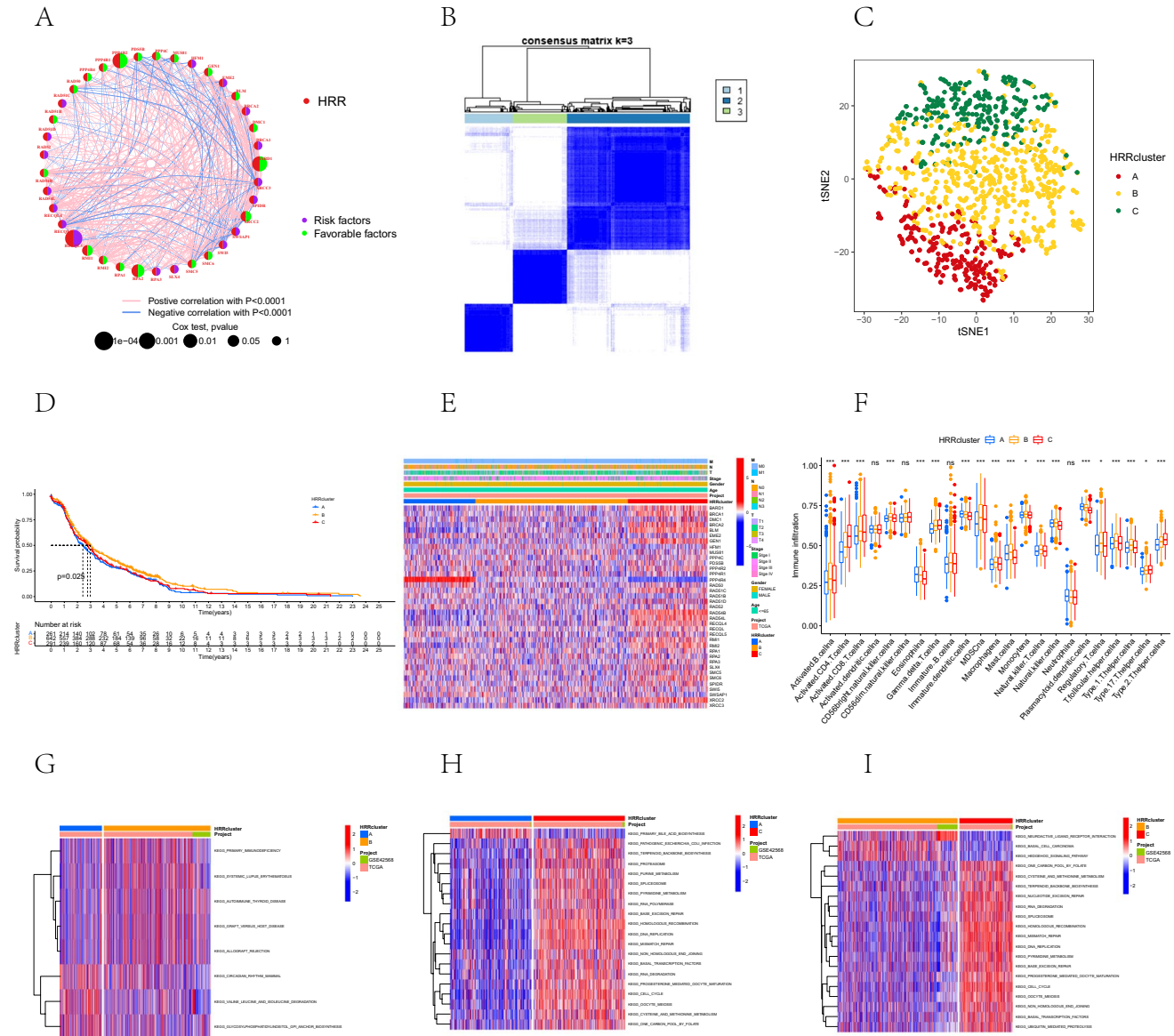


Fig. 2 Identification of HRR clusters and ICI features and biological characteristics in different HRR clusters. **A** Interaction between HRRGs. The size of the circle indicates the range of prognostic significance of each HRRG. P -values were calculated by the log-rank test. Blue dots represent protective factors for prognosis. Red dots represent risk factors for prognosis. The connected lines represent their correlation. The thickness of line indicates the correlation strength. Gray and yellow lines show negative and positive correlations, respectively. **B** The consensus matrix of all samples. All samples were clustered into the appropriate number of subtypes ($k=3$). **C** The distribution of the three HRR subtypes is shown by tSNE dimension reduction. **D** Kaplan-

Meier curves for the difference in overall survival between HRR subtypes A, B, and C ($P=0.025$). **E** The heatmap of the expression of HRRGs in different HRR clusters. The color of the heatmap indicates the relative expression of HRRGs. **F** Abundance of each ICI in HRR clusters A, B, and C. **G** The heatmap of top 20 biological pathways in HRR clusters A and B through GSVA. Red and blue represent activated and inhibited pathways, respectively. **H** The heatmap of top 20 biological pathways in HRR clusters B and C through GSVA. **I** The heatmap of the top 20 biological pathways in HRR clusters B and C through GSVA. HRRGs, homologous recombination repair genes; ICI, immune cell infiltration; GSVA, gene set variation analysis

HRR subtypes were submitted for GO and KEGG analysis to investigate their molecular biology using “clusterProfiler” package (version 4.10.0) (Yu et al. 2012). A *P*-value less than 0.05 was considered a statistically significant difference.

Construction and validation of prognostic model of breast cancer

We used the Cox regression model to investigate the influence of multiple factors on survival. The integrated dataset

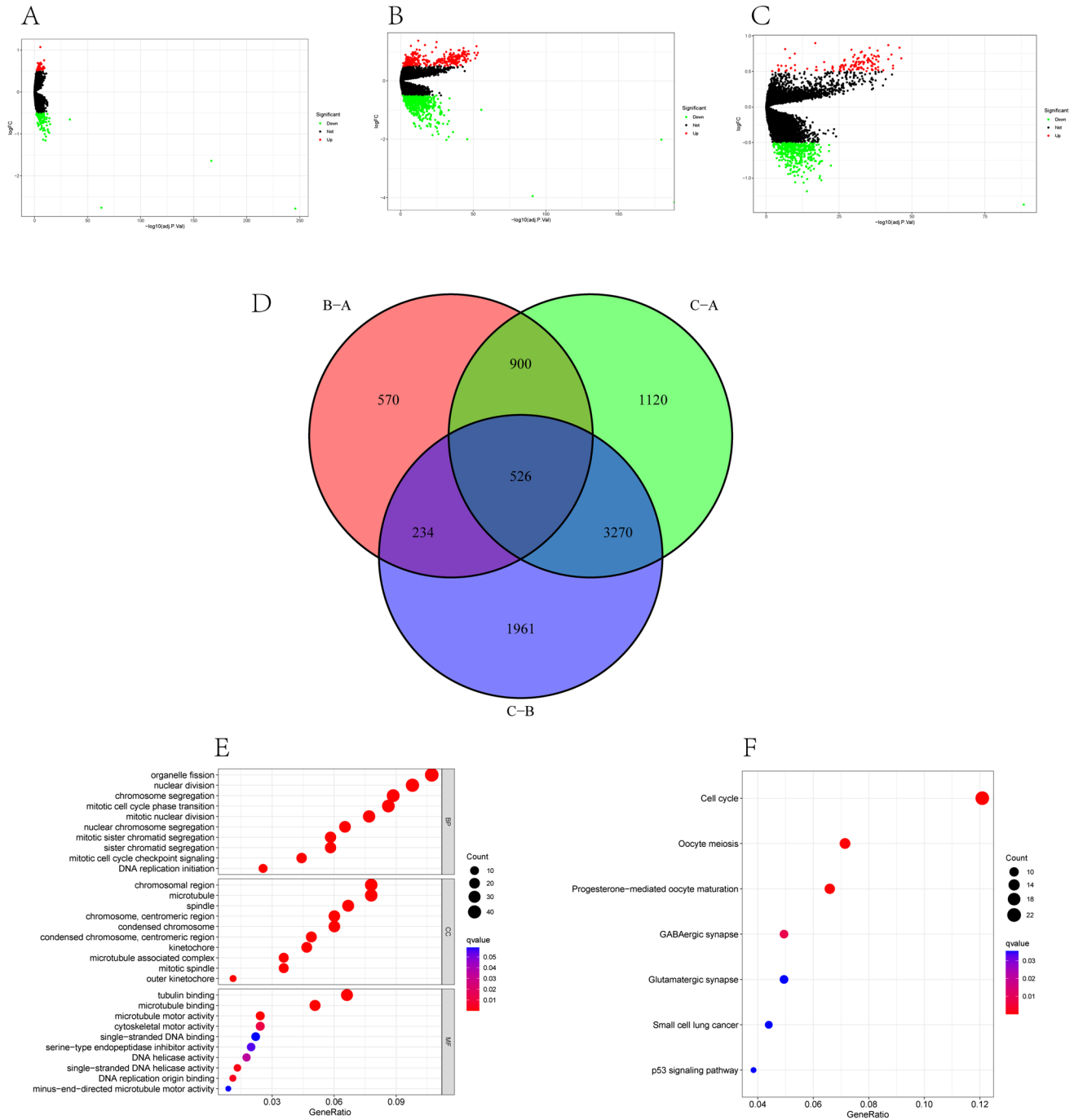
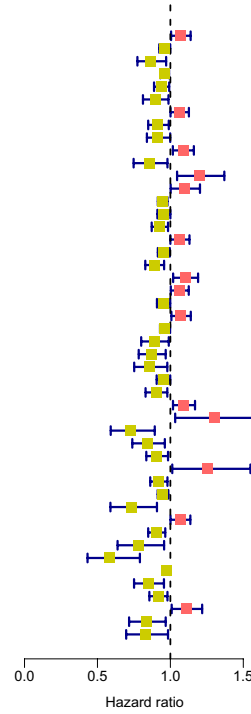


Fig. 3 Screening of DEGs among HRR clusters and functional enrichment analysis of DEGs. **A** Volcano map of group A and B DEGs. **B** Volcano map of group A–C DEGs. **C** Volcano map of group B and C DEGs; **D** Venn diagram of DEGs in three groups (A,

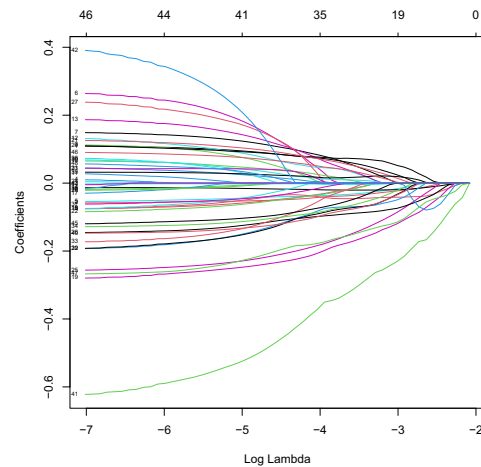
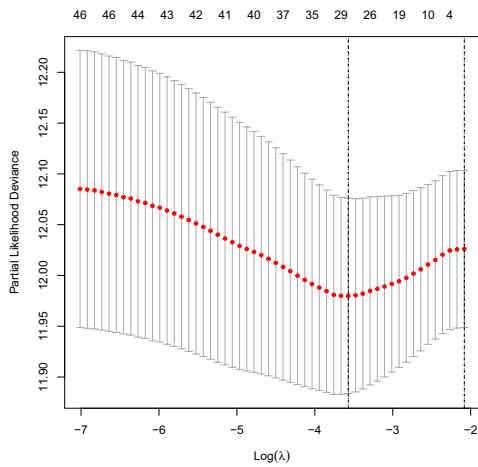
B, and C). **E** GO enrichment analysis of DEGs. **F** KEGG enrichment analysis of DEGs. DEGs, differentially expressed genes; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes

A

	pvalue	Hazard ratio
KRTAP3-3	0.033	1.070(1.005-1.139)
ERICH3	0.043	0.961(0.924-0.999)
SLC4A7	0.014	0.867(0.774-0.971)
PGR	0.011	0.961(0.931-0.991)
RTN1	0.026	0.939(0.888-0.992)
IL6ST	0.027	0.895(0.812-0.987)
GLRB	0.046	1.062(1.001-1.127)
FGD3	0.025	0.915(0.848-0.989)
MEX3A	0.047	0.915(0.839-0.999)
SEMA3B	0.013	1.087(1.017-1.161)
SLC5A6	0.025	0.856(0.748-0.980)
DKK1	0.009	1.137(1.046-1.370)
SH3D21	0.046	1.098(1.002-1.203)
ARHGAP36	0.002	0.946(0.914-0.980)
EGOT	0.049	0.954(0.911-1.000)
CCDC158	0.012	0.926(0.872-0.984)
PCSK4	0.049	1.064(1.000-1.131)
NAT1	0.034	0.954(0.914-0.996)
SUSD3	0.002	0.890(0.828-0.957)
IZUMO4	0.015	1.101(1.019-1.190)
NHLRC4	0.039	1.063(1.003-1.126)
LINC00472	0.047	0.952(0.908-0.999)
FUT8-AS1	0.029	1.072(1.007-1.140)
NPY1R	0.036	0.962(0.929-0.998)
SH3BGR	0.030	0.889(0.800-0.989)
SMAGP	0.011	1.037(1.0.783-0.968)
DEF4	0.023	0.858(0.752-0.979)
DNAH5	0.046	0.952(0.906-0.999)
XRCC2	0.014	0.902(0.830-0.979)
LMNTD2	0.014	1.091(1.018-1.169)
CD99L2	0.026	1.305(1.033-1.649)
SLC25A12	0.002	0.726(0.591-0.893)
ZC3H6	0.012	0.843(0.739-0.962)
NR4A2	0.019	0.906(0.835-0.984)
TCOF1	0.038	1.252(1.013-1.549)
VASH2	0.010	0.919(0.862-0.980)
GRK3	0.007	0.949(0.913-0.986)
SEPHS1	0.005	0.731(0.589-0.908)
AC068831.3	0.047	1.067(1.001-1.137)
STEAP2	0.002	0.905(0.848-0.965)
STMN1	0.017	0.761(0.638-0.957)
CASP2	<0.001	0.584(0.432-0.791)
LINC00993	0.021	0.973(0.950-0.996)
EIF4E3	0.007	0.848(0.751-0.956)
FBXL16	0.011	0.916(0.856-0.980)
ERICH6	0.032	1.108(1.009-1.218)
MARCKSL1	0.017	0.833(0.717-0.968)
B4GALT5	0.032	0.828(0.697-0.984)



B



C

	pvalue	Hazard ratio
'KRTAP3-3'	0.034	1.115(1.009-1.232)
ERICH3	0.065	0.936(0.872-1.004)
GLRB	0.018	1.124(1.020-1.239)
EGOT	0.143	0.942(0.870-1.020)
SUSD3	0.003	0.795(0.683-0.926)
IZUMO4	0.069	1.140(0.990-1.312)
NHLRC4	0.020	1.135(1.020-1.263)
LINC00472	0.083	0.931(0.859-1.009)
SH3BGR	0.007	0.806(0.690-0.942)
SMAGP	0.091	0.867(0.734-1.023)
NR4A2	0.020	0.852(0.745-0.975)
STEAP2	0.017	0.864(0.767-0.975)
STMN1	<0.001	0.500(0.339-0.736)
FBXL16	0.013	0.871(0.782-0.971)
MARCKSL1	0.128	0.844(0.678-1.050)
SEMA3B	0.157	1.101(0.964-1.258)

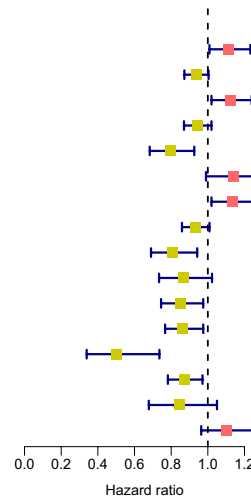


Fig. 4 Results of univariate Cox regression, LASSO regression, and multivariate analysis. **A** Forest plot of univariate Cox regression analysis. **B** LASSO regression for univariate results. **C** Forest plot of multivariate Cox regression analysis

was divided into training and validation sets at random. The model was built on the training set and then evaluated on the validation. To select genes associated with survival, we used $P=0.05$ as a filter for univariate Cox. “glmnet” package (version 4.1–8) was used for LASSO regression analysis and determining the K value through minimal lambda to avoid significant variances (Engelbrecht and Bohlin 2019).

Subsequently, risk genes were identified using multivariate Cox regression, and a prognostic risk model was developed. Gene expression and Cox regression coefficients were used to calculate the risk score. Patients were grouped into high-risk and low-risk groups based on their median risk scores in training, validation, and total sets. To determine whether high-risk and low-risk groups had similar survival rates, we used Kaplan–Meier (K-M) analysis, to draw survival curves (Wu et al. 2021). The time-dependent receiver operating curves (ROC) were drawn using the R package “Survival ROC” (version 1.0.3.1) to evaluate the model’s predictive ability and stability over 1-year, 3-year, and 5-year periods

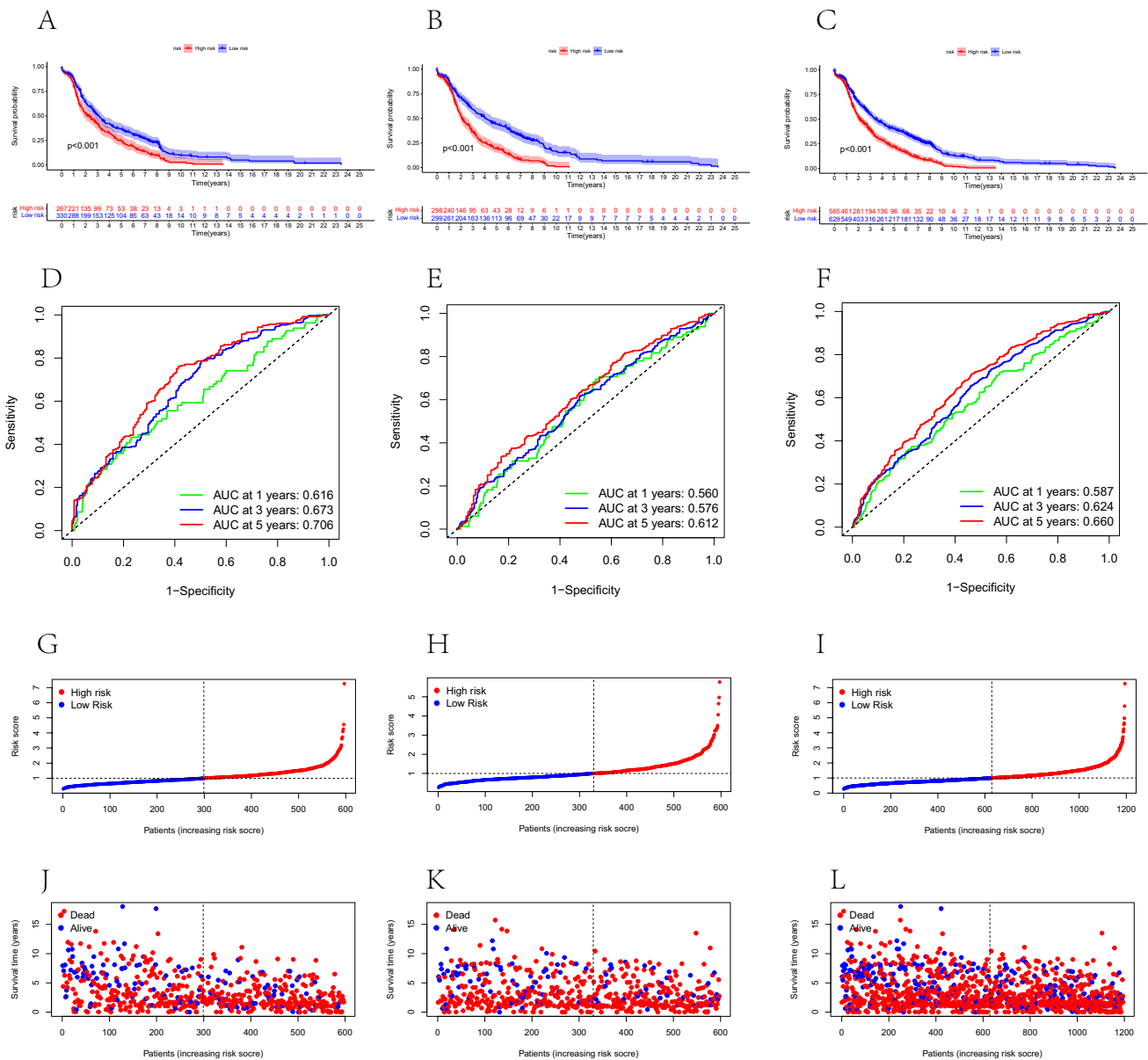


Fig. 5 Kaplan–Meier survival curves, time-dependent ROC curves, and survival status maps of the total data set, training set, and validation set. **A–C** Kaplan–Meier survival curves and risk curves of the total data set, training

set, and validation set. **D–F** Time-dependent ROC curves of the total data set, training set, and validation set. **G–L** Survival status maps of the total data set, training set, and validation set. ROC, receiver operating characteristic

(Liu et al. 2021a). The area under the curve (AUC) > 0.5 and closer to 1 indicates a better prognosis. Additionally, clinical information (age, sex, and stage) was unitized in this prognostic risk model for clinical application.

Construction of homologous recombination repair score

The HRR score was developed to measure the features of HRR in BC. Genes with significant prognosis differences were found through the above methods. Unsupervised clustering was used to group patients for an in-depth examination. Then, these DEGs' primary components were extracted using principal component analysis (PCA), and the HRRG signatures were created. Signature scores were chosen from principal components 1 and 2 (Ringnér 2008). In order to calculate the HRR score of each patient, the following equation was used: $HRR\ score = \sum PCA1i + \sum PCA2i$ (i is the expression of prognostic DEGs). With the "maxstat" package (version 0.7–25), we obtained the best cutoff value for separating High and Low HRR score groups (HSG and LSG) linked to prognosis (Ogluszka et al. 2019).

Prediction of immunotherapy sensitivities

To compare HSG and LSG expression differences, we downloaded the immunophenoscore (IPS) data from the Cancer Immune Atlas (TCIA) database (<http://tcia.at/>) (Guo et al. 2022). Then, differences in anti-cytotoxic T-lymphocyte antigen 4 (CTLA-4) and anti-programmed cell death protein 1 (PD-1) antibody responses were compared between different HRR score groups to predict immunotherapy sensitivity.

Statistical analysis

Statistical analysis in this study was performed using R software (version 4.0.5). A comparison of numerical type variables between two or more groups was conducted using Wilcoxon and Kruskal–Wallis tests. Kaplan–Meier curves were used for drawing prognosis and survival curves, and the chi-square test was used for categorized variable comparisons between HSG and LSG groups. The correlation coefficient was calculated using Spearman's test. A P -value < 0.05 was considered a statistically significant difference.

Result

Genetic variations in homologous recombination repair genes in breast cancer

There were 65 HRRGs in this study. First, we estimated and displayed that HRRGs had frequent CNV changes in BC

and that most HRRGs were concentrated on CNV amplifications, while some HRRGs had a higher frequency of CNV deletions (Fig. 1A). RECQL5, RAD51C, and RAD54B were the genes with the highest copy number amplification frequency. RPA2, PPP4R2, and RAD51B were among the genes with the highest prevalence of deletions. HRRG mutations were found in 131 of 986 samples, suggesting a 13.29% mutation frequency. Among 42 HRRGs carrying mutations, a 2% mutation frequency was found in BRCA2, followed by BRCA1 (Fig. 1B). The findings revealed that CNV was expected, and deletions or amplifications of different HRRGs had distinct features. The location of the mutated genes on 23 pairs of chromosomes and the prevalence of CNVs on HRRGs are shown in Fig. 1C. HRRG expression levels differed between tumor and normal samples, as shown in Fig. 1D. Some HRRGs, such as HFM1 and PPP4R4, were higher expressed significantly in tumors, while some genes, such as EME2 and MUS81, were lower expressed in tumors, which reminded us that BC may be triggered and progressed by differences in HRRG expression.

Identification of homologous recombination repair phenotypes based on gene expression patterns

The interaction relationships and prognostic effects of HRRGs were further explored with multivariate COX and correlation analyses. Thirty-seven HRRGs had effects on the prognosis of BC patients, with some HRRGs having a significant positive correlation with prognoses, such as SLX4, RAD54L, RECQL4, and RECQ, and some HRRGs, including RAD51B, RAD50, and PD55B, were negatively correlated with prognosis (Fig. 2A). Additionally, the favorable prognostic HRRGs were negatively correlated with the unfavorable HRRGs. For example, there was a negative association between SLX4 and RAD51B, RAD50, and PD55B. The complex interplay between the HRRGs may play a crucial role in patient prognosis and ICI characteristics. Subsequently, we performed a cluster analysis based on differential expression patterns of HRRGs using the R package "ConsensusClusterPlus" (version 1.66.0). Three HRR-related phenotypes were identified and labeled as HRR clusters A, B, and C (Fig. 2B). The clustering effect was good in three groups (A, B, and C) according to T-SNE analysis (Fig. 2C). A cluster with a worse prognosis than clusters B and C was revealed by the KM curves ($P=0.025$) (Fig. 2D). Heatmaps showed HRRG expression patterns, and we can conclude that HRRGs were expressed differently in each cluster (Fig. 2E).

Immune cell infiltration features and biological characteristics in different homologous recombination repair phenotypes

Activated dendritic cells, CD56 bright natural killer (NK) cells, neutrophils, and immature B cells did not differ

significantly, as ICI analysis showed. Most of the immune cells in the subtype A cluster had lower infiltration. It

was flooded with immune-suppressive cells, including eosinophils, myeloid-derived suppressor cells (MDSCs),

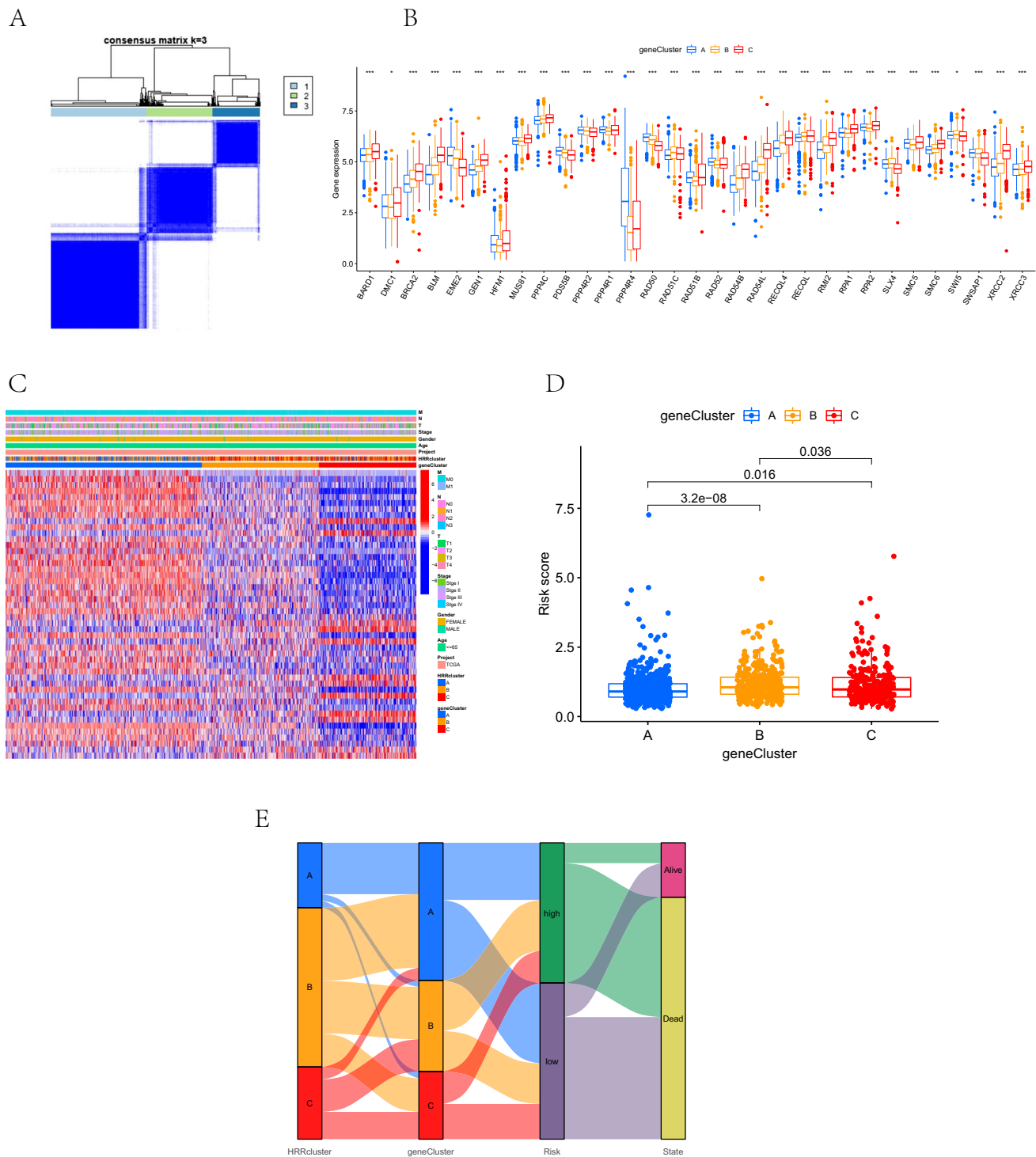


Fig. 6 Construction of gene clusters based on HRR cluster-related DEGs. **A** Consensus matrix based on the TCGA-BRCA cohort in three HRR clusters. TCGA samples were clustered into the appropriate number of subtypes ($k=3$). **B** The differential expression of HRRGs in different gene clusters. **C** Heatmap of the HRR-related DEGs in different HRR clusters and gene clusters. **D** Differences in HRR scores among

the three gene clusters. The statistical differences among the three gene clusters were compared by the Wilcoxon test ($P < 0.05$). **E** Sankey diagram of the distribution of patients with HRR cluster, gene cluster, risk group, and survival status. Positive correlation is marked in red and negative correlation in blue. HRR, homologous recombination repair; DEGs, differentially expressed genes; BRCA, breast cancer

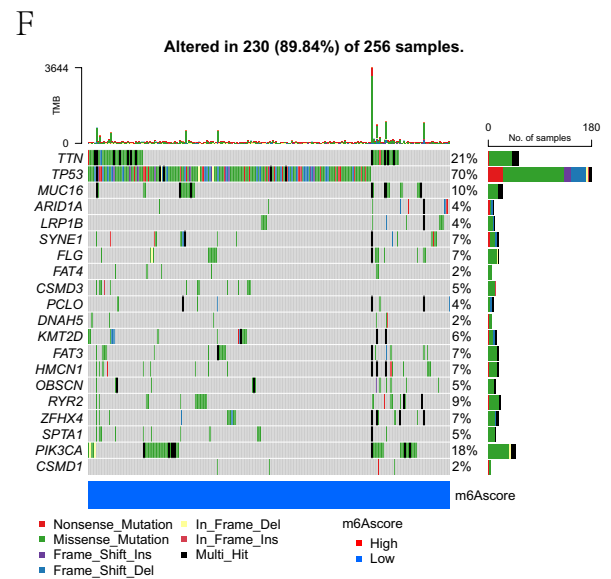
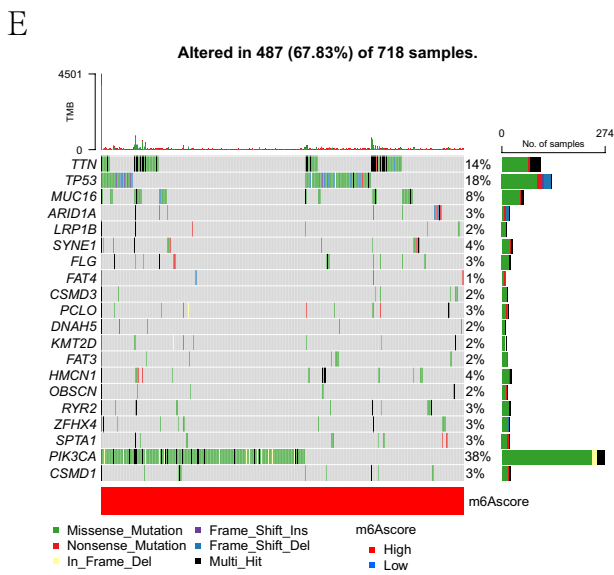
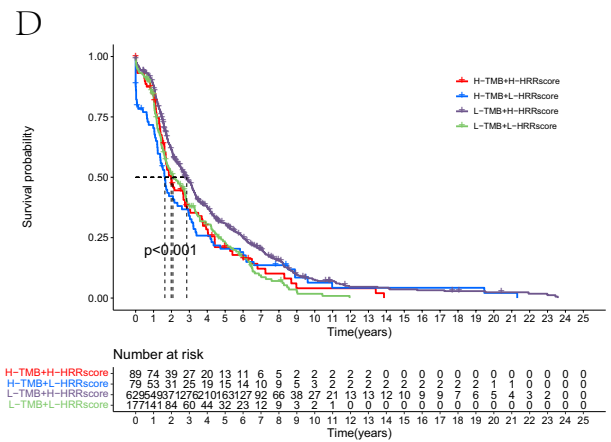
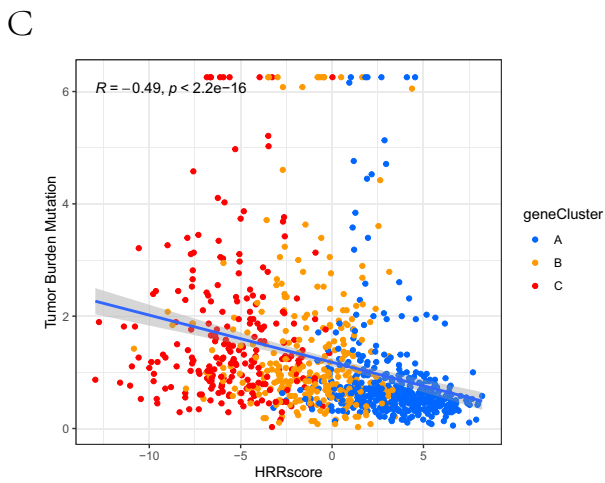
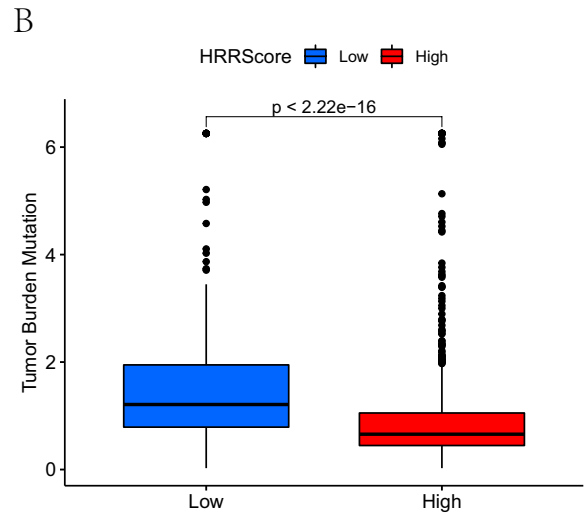
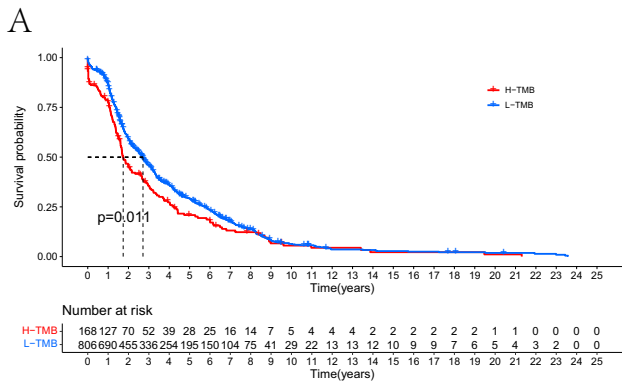


Fig. 7 The impact of TMB on prognosis and its association with HRR score. **A** Kaplan–Meier curves of the overall survival difference between high- and low-TMB groups ($P=0.011$). The statistical differences were compared by the Wilcoxon test. **B** The TMB levels between the HSG and LSG groups. **C** Correlation between HRR score and TMB (Spearman coefficient). **D** Kaplan–Meier curves for the differences in overall survival stratified by TMB and HRR score ($P<0.001$). **E, F** Waterfall plots for top 20 driver genes with the highest mutation frequency in HSG (**E**) and LSG (**F**). TMB, tumor mutation burden; HRR, homologous recombination repair; HSG, high HRR score group; LSG, low HRR score group

macrophages, mast cells, and Tregs, which may be related to different prognoses (Fig. 2F). Thus, we speculated that HRR might cause immunosuppression and accelerate tumor progression by suppressing immunity.

The biological behaviors among three HRR-related subtypes were explored through GSVA analysis (Fig. 2G–I). In cluster A, valine, leucine, and isoleucine degradations; primary immunodeficiency; and other related signals were significantly involved. Signals, including DNA replication, mismatch repair, and homologous recombination, were mainly enriched in cluster C. On the contrary, cluster B was primarily enriched in complement, coagulation cascades, and ECM receptor interactions.

Screening of homologous recombination repair phenotype-related differentially expressed genes and functional enrichment analysis

DEGs among these three subtypes have been determined, and the results were visualized using volcano plots (Fig. 3A–C). There were 2230 DEGs between subtypes A and B, 5991 DEGs between subtypes B and C, and 5816 DEGs between subtypes A and C. Five hundred twenty-six common DEGs were identified by intersecting the DEGs of the three groups (Fig. 3D). We used these DEGs for functional enrichment analysis. The critical GO biological processes were organelle fission, chromosomal region, and tubulin binding (Fig. 3E). Analysis of KEGG showed that the genes were primarily involved in the cell cycle, oocyte meiosis, and progesterone-mediated maturation of oocytes (Fig. 3F).

Construction and validation of the prognostic risk model

The prognostic model was constructed using 1194 samples with complete gene expression and survival data. Samples were split into training and validation sets randomly. Forty-eight DEGs were related to prognostic survival in the univariate Cox regression ($P<0.05$) (Fig. 4A). Lasso regression was performed to prevent overfitting, and 28 genes related to prognosis were obtained (Fig. 4B). A risk model was constructed using 16 DEGs through multivariate COX

analysis, and a risk score was calculated to assess the prognosis (Fig. 4C). The 16 genes included were KRTAP3-3, ERICH3, GLRB, EGOT, SUS3, IZUMO4, NHLRC4, LINC00472, SH3BGR, SMAGP, NR4A2, STEAP2, STMN1, FBXL16, MARCKSL1, and SEMA3B. Calculations of risk scores were performed as follows: Risk score = $(0.108 \times \text{KRTAP3-3}) + (-0.06 \times \text{ERICH3}) + (0.117 \times \text{GLRB}) + (0.05 \times \text{EGOT}) + (-0.229 \times \text{SUS3}) + (0.13 \times \text{IZUMO4}) + (0.126 \times \text{NHLRC4}) + (0.07 \times \text{LINC00472}) + (0.215 \times \text{SH3BGR}) + (-0.143 \times \text{SMAGP}) + (-0.16 \times \text{NR4A2}) + (-0.145 \times \text{STEAP2}) + (-0.693 \times \text{STMN1}) + (-0.137 \times \text{FBXL16}) + (-0.169 \times \text{MARCKSL1}) + (0.096 \times \text{SEMA3B})$. The median risk scores of BC patients in total, training, and validation sets were used to categorize them into high- and low-risk groups. There was a worse survival rate among all three sets of patients who had high-risk scores ($P<0.001$) (Fig. 5A–C). In the total set, AUC at 1 year was 0.616; at 3 years, 0.673; and at 5 years, 0.706 (Fig. 5D). The AUC for the training set was 0.560 after 1 year, 0.576 after 3 years, and 0.612 after 5 years (Fig. 5E). Additionally, the AUC was 0.587 after 1 year, 0.624 after 3 years, and 0.660 after 5 years in the validation set (Fig. 5F). This model predicted BC prognosis well, with good sensitivity and specificity. The visualized risk scores showed increased mortality and decreased survival time, along with the increase in risk scores in the total set, training set, and validation set, respectively (Fig. 5G–L).

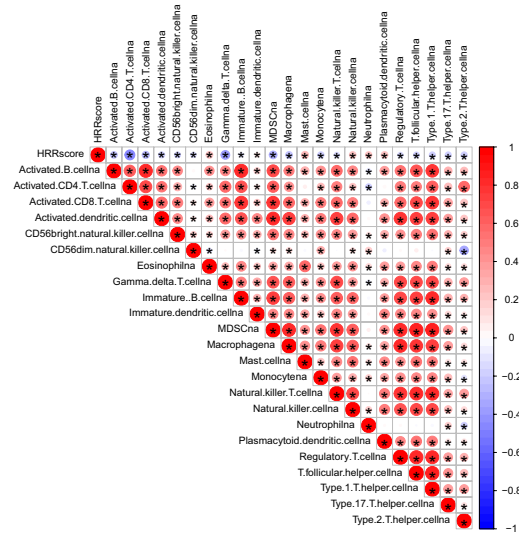
Identification of prognostic-related subtypes based on secondary clustering using differentially expressed genes

To determine the biological characteristics of HRR subtypes, we conducted an unsupervised cluster analysis of 526 overlapping DEGs. We identified three DEG clusters (gene clusters A–C) associated with the HRR phenotype (Fig. 6A). There were significant differences in the expressions of HRRG gene clusters (Fig. 6B). The heatmap illustrated HRRG expression patterns and showed differential expression of HRRGs in three gene clusters (Fig. 6C). There was a lower risk score of cluster A than that of clusters B and C (Fig. 6D). The Sankey diagram visualized each patient's HRR clusters, gene clusters, risk group, and survival status (Fig. 6E). High-risk patients were most likely to have poor prognoses in gene cluster A. In gene cluster B, the majority of patients having poor prognoses were at high risk.

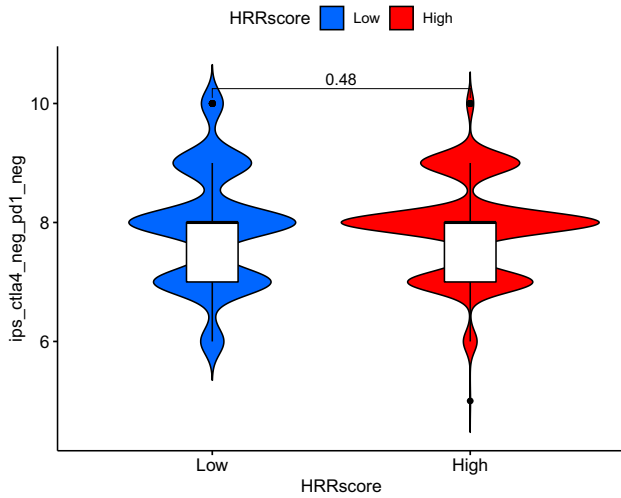
Correlation between homologous recombination repair score and tumor mutation burden

A series of somatic mutation analyses were made on some genome stability pathways involved in the GSVA analysis. The impact of TMB on prognosis was explored. As shown

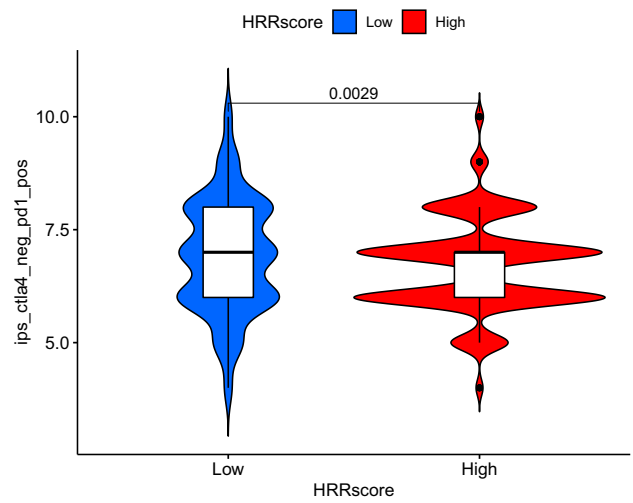
A



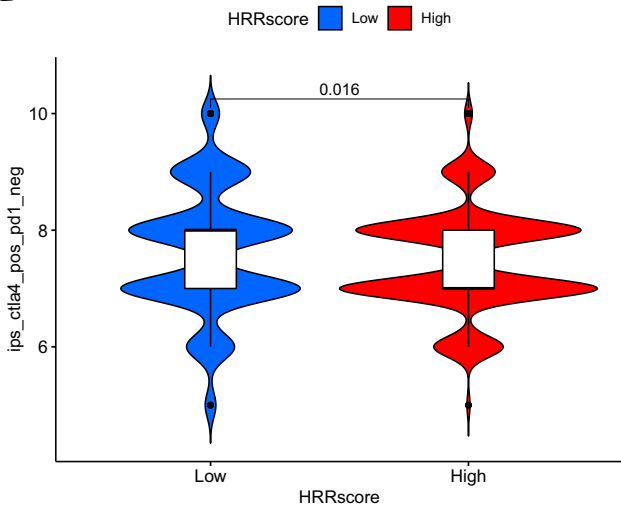
B



C



D



E

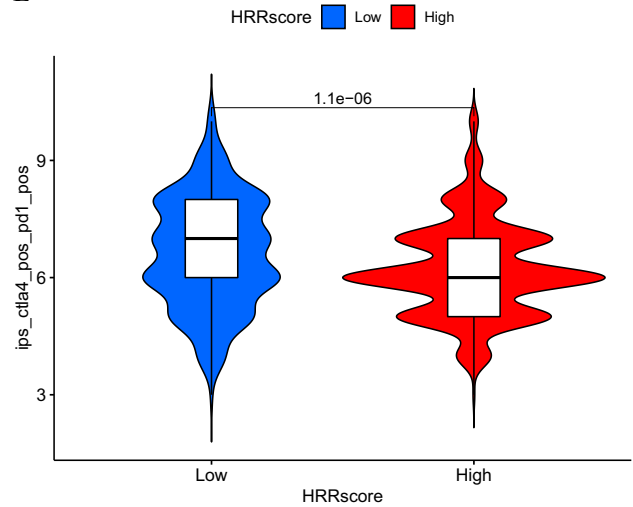


Fig. 8 The association between HRR score and ICI and differences of IPS between HSG and LSG. **A** Spearman analysis of the correlation between HRR score and ICI abundance in the TCGA cohort. **B** In PD-1_NEG+CTLA4_neg, there was no significant difference between the IPS of the HSG and the LSG. **C–E** In PD-1_NEg+CTLA4_pos, CTLA4_pos+PD-1_POS, and PD-1_POS+CTLA4_neg, IPS increased in the LSG. HRR, homologous recombination repair; ICI, immune cell infiltration; IPS, immunophenotypes; HSG, high HRR score group; LSG, low HRR score group

in Fig. 7A, the higher TMB group had a worse prognosis. As shown in Fig. 7B, compared to LSG, HSG had a lower TMB level ($P < 0.001$). TMB and HRR scores correlated negatively (Spearman coefficient: $r = 0.49$, $P < 0.001$, Fig. 7C). We performed a stratified prognostic analysis based on the synergism of the TMB and HRR scores. High HRR and low TMB scores were associated with a significant survival advantage for patients (Fig. 7D), which revealed that the HRR score could indicate prognosis effectively. After this, we analyzed the differences between HSG and LSG in terms of the distribution of somatic mutations. The frequency of mutations of the top 20 genes with the most mutations was generally lower in HSG than in LSG (Fig. 7E, F).

Correlation between HRR score and immune cell infiltration

Based on DEG expression patterns, we developed an algorithm using PCA and defined the result as an HRR score. We determined the optimal cutoff value using the “maxstat” package. Due to the differences in ICI characteristics among HRR subtypes, we also examined the association between HRR score and ICI (Fig. 8A). A positive correlation was demonstrated between HRR scores and most infiltrating immune cells, following the results of HRR subtypes.

Correlation between HRR score and immunotherapy response

A comparison of the IPS difference between HSG and LSG in four types (CTLA4_negative + PD-1_negative, CTLA4_positive + PD-1_negative, CTLA4_negative + PD-1_positive, CTLA4_positive + PD-1_positive) was conducted in order to determine if HRR score would predict the clinical immunotherapy effects of BC, visualized by an immunophenotype map (Fig. 8 B–E). The result indicated that for PD-1_NEG + CTLA4_neg types, no significant difference existed between the IPS of the HSG and the LSG (Fig. 8B). However, for PD-1_NEg + CTLA4_pos, CTLA4_pos + PD-1_POS, and PD-1_POS + CTLA4_neg types, IPS increased in the LSG (Fig. 8C–E). Results showed that patients in the LSG responded better to anti-PD-1 therapy or anti-PD-1 and anti-CTLA4 therapy.

Discussion

In this study, we identified three HRR clusters based on HRRGs. The three HRR clusters showed different ICI characteristics and HRRG expression patterns. DEGs identified in HRR clusters were primarily involved in pathways related to cell cycle and genome stability. The prognostic model based on 16 DEGs accurately predicted BC prognosis. We also identified three gene clusters based on DEGs, which could identify prognostic subtypes more stably. High ICI, low TMB, and worse immunotherapy response were observed in HSG. The results may help predict prognosis and potential therapeutic benefits.

We mapped three clusters of HRR-related DEGs and built a prognostic nomogram for BC. The ROC curves showed that this model was good at predicting BC prognosis. The nomogram contained 16 DEGs. The majority of these genes play a role in the development of BC and other types of cancers. Researchers have found that ERICH1-AS1 expression can predict non-small cell lung cancer prognosis (Tang et al. 2015). BC cells can be triggered to proliferate and migrate by SUSD3 (Moy et al. 2015). EGOT has also been used to construct a risk prediction model related to BC (Lv et al. 2021). By inactivating the Hedgehog signaling pathway in BC, overexpression of EGOT reduces the viability and migration of the cells (Qiu et al. 2020). Cells can become more sensitive to paclitaxel toxicity by increasing the levels of EGOT, and regulating EGOT may be a new way to boost paclitaxel toxicity (Xu et al. 2019). Up-regulation of LINC00472 inhibited the viability, invasion, migration, and EMT of lung cancer cells but increased the apoptosis rate of lung cancer cells (Mao et al. 2019). Cell growth in vitro and xenograft tumor growth in vivo in BC tumors were inhibited where SH3BGRL2 was downregulated (Li et al. 2020). SMAGP knock-down can inactivate the PI3K/Akt pathway, thereby inhibiting the malignant phenotypes of glioblastoma cells (Ni et al. 2020). Overall survival and recurrence-free survival were positively associated with NR4A1-NR4A3 expression. Additionally, NR4A family genes regulate oxidative phosphorylation in BC (Yousefi et al. 2022). Down-regulation of STEAP2 expression is associated with a poor prognosis for BC. STEAP2 acts as an anti-oncogene during BC development. Through PI3K/AKT/mTOR signaling, STEAP2 downregulation can promote tumor cell proliferation and metastasis (Yang et al. 2020). MiR-770 inhibits BC metastasis by targeting STMN1 directly (Li et al. 2018). The high level of STMN1 protein in BC tissues was related to poor prognosis (Tang et al. 2020). Angiogenesis and epithelial-mesenchymal transition in BC can be blocked by FBXL16. BC with FBXL16 downregulation has a higher node and high-grade tumors and poor survival (Kim et al. 2021). In lung adenocarcinoma cells, MARCKSL1 promotes

proliferation, migration, and invasion (Liang et al. 2020). SEMA3B-AS1 could target the miR-3940/KLLN axis to inhibit BC progression (Hu et al. 2022). This proved that the 16 DEGs in nomograms were mostly able to play a role in the development and prognosis of BC and other types of cancers.

In HRR cluster A, most immune cells had lower infiltration, including immunosuppressive cells, such as eosinophils, MDSCs, macrophages, mast cells, and Tregs. It has been shown in previous studies that patients with a low eosinophil count have a higher recurrence rate and a worse prognosis (Ownby et al. 1983; Onesti et al. 2020). Innate and adaptive immunity are inhibited by MDSCs, which are heterogeneous groups of immature myeloid cells. MDSC expansion triggers the pre-cancerous immune microenvironment, thus accelerating BC progression (Liu et al. 2021b). The mononuclear phagocytic system includes macrophages. They are usually broadly classified into M1 and M2. M1 macrophages might help metastatic BC cell dormancy, while M2 macrophages might promote tumor outgrowth (Lin et al. 2019). The increased mast cell density and distribution are a worse prognostic factor for BC (Carpenco et al. 2019). Treg cells are known to play a crucial role in peripheral tolerance and can suppress effector T cells to prevent unwanted immune responses. Treg cells have an invasive appearance and are associated with a reduced chance of relapse-free survival and overall survival in BC biopsies (Mittal et al. 2018; Togashi et al. 2019). A decrease in Tregs and naive CD4+ T cells in TME might suppress BC metastasis (Li et al. 2021). These results suggested that HRR subtypes were associated with ICI in TME. So, we analyzed how the HRR score correlated with ICI and found a positive correlation between HRR scores and most infiltrating immune cells.

A positive link between HRR score and immune cells infiltrating indicated the possible association between HRR and immunity. The HRRGs had frequent CNV amplifications and deletions, especially in LSG. HRR-related DEGs tended to be enriched in cell cycle and genome stability. HRR is the most accurate and high-fidelity way to repair DNA damage. Inherited mutations in HRRGs can increase the risk of BC development (Mersch et al. 2015). TMB elevation is caused by mutations in the DNA damage response genes, predicting a poor prognosis (Klempner et al. 2020). We found that TMB was negatively related to survival and HRR scores. Combining TMB and HRR scores can further improve prognosis prediction. It has been reported that immunotherapy might benefit patients with poor prognoses because their TMB is higher and their genes are more mutated (Barroso-Sousa et al. 2020; Liu et al. 2020). In line with previous research, anti-PD-1 therapy or anti-PD-1 and anti-CTLA4 therapy worked better for patients in the LSG. Significant breakthroughs have been made in some advanced therapies, including PARP-targeted therapy. There

are several PARPi approved to treat BC, including taraparib and olaparib (Lyons 2019; Shen et al. 2020b). PARPi can further hinder DNA repair in tumor cells in patients with HRR deficiency, accelerating their death and achieving precise targeting (Chopra et al. 2020).

In this study, we identified three HRR subtypes and comprehensively explored the association between HRR and ICI. Our gene signature model predicts BC prognosis with high accuracy based on HRR-related DEGs, which may help develop new treatment options. The HRR score could predict TMB and immunotherapy responses, which may provide references for clinical treatment.

Author contribution Hong-Chuan Jiang designed the study. Yan-Shuang Li wrote the original draft. Yan-Shuang Li collected raw data. Yan-Shuang Li performed statistical and bioinformatics analyses. Hong-Chuan Jiang supervised the study. All authors read and approved the final version of the manuscript.

Data availability The data are available at the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>) (GSE42568), TCGA database (<http://cancergenome.nih.gov/>), and TCIA database (<https://tcia.at/>).

Declarations

Ethical approval N/A

Informed consent N/A

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahmad A (2019) Breast cancer statistics: recent trends. *Adv Exp Med Biol* 1152:1–7
- Barroso-Sousa R, Jain E, Cohen O, Kim D, Buendia-Buendia J, Winer E et al (2020) Prevalence and mutational determinants of high tumor mutation burden in breast cancer. *Ann Oncol : Off J Eur Soc Med Oncol* 31(3):387–394
- Bersuker K, Hendricks JM, Li Z, Magtanong L, Ford B, Tang PH et al (2019) The CoQ oxidoreductase FSP1 acts parallel to GPX4 to inhibit ferroptosis. *Nature* 575(7784):688–692
- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of

- incidence and mortality worldwide for 36 cancers in 185 countries. *CA: Cancer J Clin* 68(6):394–424
- Carpenco E, Ceaușu RA, Cimpean AM, Gaje PN, Șaptefrați L, Fulga V et al (2019) Mast cells as an indicator and prognostic marker in molecular subtypes of breast cancer. *In Vivo (Athens, Greece)* 33(3):743–748
- Chopra N, Tovey H, Pearson A, Cutts R, Toms C, Proszek P et al (2020) Homologous recombination DNA repair deficiency and PARP inhibition activity in primary triple negative breast cancer. *Nat Commun* 11(1):2662
- Doll S, Freitas FP, Shah R, Aldrovandi M, da Silva MC, Ingold I et al (2019) FSP1 is a glutathione-independent ferroptosis suppressor. *Nature* 575(7784):693–698
- Engelbreten S, Bohlin J (2019) Statistical predictions with glmnet. *Clin Epigenetics* 11(1):123
- Gerber B, Freund M, Reimer T (2010) Recurrent breast cancer: treatment strategies for maintaining and prolonging good quality of life. *Deutsches Arzteblatt Int* 107(6):85–91
- Guo JN, Chen D, Deng SH, Huang JR, Song JX, Li XY et al (2022) Identification and quantification of immune infiltration landscape on therapy and prognosis in left- and right-sided colon cancer. *Cancer Immunol Immunother : CII* 71(6):1313–1330
- Hänzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinforma* 14:7
- Hassannia B, Vandenabeele P, Vanden BT (2019) Targeting ferroptosis to iron out cancer. *Cancer Cell* 35(6):830–849
- Hu J, Huang H, Xi Z, Ma S, Ming J, Dong F et al (2022) LncRNA SEMA3B-AS1 inhibits breast cancer progression by targeting miR-3940/KLLN axis. *Cell Death Dis* 13(9):800
- Jézéquel P, Kerdraon O, Hondermarck H, Guérin-Charbonnel C, Lasla H, Gouraud W et al (2019) Identification of three subtypes of triple-negative breast cancer with potential therapeutic implications. *Breast Cancer Res : BCR* 21(1):65
- Kim YJ, Zhao Y, Myung JK, Yi JM, Kim MJ, Lee SJ (2021) Suppression of breast cancer progression by FBXL16 via oxygen-independent regulation of HIF1 α stability. *Cell Rep* 37(8):109996
- Klempner SJ, Fabrizio D, Bane S, Reinhart M, Peoples T, Ali SM et al (2020) Tumor mutational burden as a predictive biomarker for response to immune checkpoint inhibitors: a review of current evidence. *Oncologist* 25(1):e147–e159
- Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD (2012) The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics (Oxford, England)* 28(6):882–883
- Li Y, Liang Y, Sang Y, Song X, Zhang H, Liu Y et al (2018) MiR-770 suppresses the chemo-resistance and metastasis of triple negative breast cancer via direct targeting of STMN1. *Cell Death Dis* 9(1):14
- Li DD, Deng L, Hu SY, Zhang FL, Li DQ (2020) SH3BGR2 exerts a dual function in breast cancer growth and metastasis and is regulated by TGF- β 1. *Am J Cancer Res* 10(4):1238–1254
- Li J, Wang S, Wang N, Zheng Y, Yang B, Wang X et al (2021) Aiduqing formula inhibits breast cancer metastasis by suppressing TAM/CXCL1-induced Treg differentiation and infiltration. *Cell Commun Signal* 19(1):89
- Liang W, Gao R, Yang M, Wang X, Cheng K, Shi X et al (2020) MARCKSL1 promotes the proliferation, migration and invasion of lung adenocarcinoma cells. *Oncol Lett* 19(3):2272–2280
- Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P (2015) The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell Syst* 1(6):417–425
- Lin Y, Xu J, Lan H (2019) Tumor-associated macrophages in tumor metastasis: biological roles and clinical therapeutic applications. *J Hematol Oncol* 12:76
- Liu Z, Mi M, Li X, Zheng X, Wu G, Zhang L (2020) A lncRNA prognostic signature associated with immune infiltration and tumour mutation burden in breast cancer. *J Cell Mol Med* 24(21):12444–12456
- Liu TT, Li R, Huo C, Li JP, Yao J, Ji XL et al (2021a) Identification of CDK2-related immune forecast model and ceRNA in lung adenocarcinoma, a pan-cancer analysis. *Front Cell Dev Biol* 9:682002
- Liu C, Qiang J, Deng Q, Xia J, Deng L, Zhou L et al (2021b) ALDH1A1 activity in tumor-initiating cells remodels myeloid-derived suppressor cells to promote breast cancer progression. *Can Res* 81(23):5919–5934
- Lv W, Wang Y, Zhao C, Tan Y, Xiong M, Yi Y et al (2021) Identification and validation of m6A-related lncRNA signature as potential predictive biomarkers in breast cancer. *Front Oncol* 11:745719
- Lyons TG (2019) Targeted therapies for triple-negative breast cancer. *Curr Treat Options Oncol* 20(11):82
- Mao X, Zhou X, Liu J, Mao Y, Zhou H (2022) Retracted: Up-regulated Linc00472 suppresses development of lung cancer cell via inhibition of MiR-196b-5p. *Biosci Biotechnol Biochem* 86(8):e1–e13
- Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP (2018) Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 28(11):1747–1756
- Mersch J, Jackson MA, Park M, Nebgen D, Peterson SK, Singletary C et al (2015) Cancers associated with BRCA1 and BRCA2 mutations other than breast and ovarian. *Cancer* 121(2):269–275
- Mittal S, Brown NJ, Holen I (2018) The breast tumor microenvironment: role in cancer development, progression and response to therapy. *Expert Rev Mol Diagn* 18(3):227–243
- Moy I, Todorović V, Dubash AD, Coon JS, Parker JB, Buranapramest M et al (2015) Estrogen-dependent sushi domain containing 3 regulates cytoskeleton organization and migration in breast cancer cells. *Oncogene* 34(3):323–333
- Neiger HE, Siegler EL, Shi Y (2021) Breast cancer predisposition genes and synthetic lethality. *Int J Mol Sci* 22(11):5614
- Ni H, Ji D, Huang Z, Li J (2020) SMAGP knockdown inhibits the malignant phenotypes of glioblastoma cells by inactivating the PI3K/Akt pathway. *Arch Biochem Biophys* 695:108628
- Ogłuszka M, Orzechowska M, Jędraszka D, Witas P, Bednarek AK (2019) Evaluate cutpoints: adaptable continuous data distribution system for determining survival in Kaplan-Meier estimator. *Comput Methods Programs Biomed* 177:133–139
- Onesti CE, Josse C, Boulet D, Thiry J, Beaumecker B, Bours V et al (2020) Blood eosinophilic relative count is prognostic for breast cancer and associated with the presence of tumor at diagnosis and at time of relapse. *Oncoimmunology* 9(1):1761176
- Ownby HE, Roi LD, Isenberg RR, Brennan MJ (1983) Peripheral lymphocyte and eosinophil counts as indicators of prognosis in primary breast cancer. *Cancer* 52(1):126–130
- Prados-Carvajal R, Irving E, Lukashchuk N, Forment JV (2021) Preventing and overcoming resistance to PARP inhibitors: a focus on the clinical landscape. *Cancers* 14(1), 44
- Qiu S, Chen G, Peng J, Liu J, Chen J, Wang J et al (2020) LncRNA EGOT decreases breast cancer cell viability and migration via inactivation of the Hedgehog pathway. *FEBS Open Bio* 10(5):817–826
- Reilly NM, Yard BD, Pittman DL (2019) Homologous recombination-mediated DNA repair and implications for clinical treatment of repair defective cancers. *Methods Mol Biol (Clifton, NJ)* 1999:3–29
- Ringnér M (2008) What is principal component analysis? *Nat Biotechnol* 26(3):303–304
- Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W et al (2015) limma powers differential expression analyses for RNA-seq and microarray studies. *Nucleic Acids Res* 43(7):e47
- Schlam I, Church SE, Hether TD, Chaldeckas K, Hudson BM, White AM et al (2021) The tumor immune microenvironment of

- primary and metastatic HER2- positive breast cancers utilizing gene expression and spatial proteomic profiling. *J Transl Med* 19(1):480
- Shen Y, Peng X, Shen C (2020a) Identification and validation of immune-related lncRNA prognostic signature for breast cancer. *Genomics* 112(3):2640–2646
- Shen M, Pan H, Chen Y, Xu YH, Yang W, Wu Z (2020b) A review of current progress in triple-negative breast cancer therapy. *Open Med (Warsaw, Poland)* 15(1):1143–1149
- Stockwell BR, Friedmann Angeli JP, Bayir H, Bush AI, Conrad M, Dixon SJ et al (2017) Ferroptosis: a regulated cell death nexus linking metabolism, Redox Biology, and Disease. *Cell* 171(2):273–285
- Tang Q, Ni Z, Cheng Z, Xu J, Yu H, Yin P (2015) Three circulating long non-coding RNAs act as biomarkers for predicting NSCLC. *Cell Physiol Biochem : Int J Exp Cell Physiol Biochem Pharmacol* 37(3):1002–1009
- Tang J, Luo Y, Wu G (2020) A glycolysis-related gene expression signature in predicting recurrence of breast cancer. *Aging* 12(24):24983–24994
- Togashi Y, Shitara K, Nishikawa H (2019) Regulatory T cells in cancer immunosuppression — implications for anticancer therapy. *Nat Rev Clin Oncol* 16:356–371
- Varadé J, Magadán S, González-Fernández Á (2021) Human immunology and immunotherapy: main achievements and challenges. *Cell Mol Immunol* 18(4):805–828
- Wilkerson MD, Hayes DN (2010) ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics (Oxford, England)* 26(12):1572–1573
- Wu Z, Lu Z, Li L, Ma M, Long F, Wu R et al (2021) Identification and validation of ferroptosis-related lncRNA signatures as a novel prognostic model for colon cancer. *Front Immunol* 12:783362
- Xu S, Wang P, Zhang J, Wu H, Sui S, Zhang J et al (2019) Ai-lncRNA EGOT enhancing autophagy sensitizes paclitaxel cytotoxicity via upregulation of ITPR1 expression by RNA-RNA and RNA-protein interactions in human cancer. *Mol Cancer* 18(1):89
- Yang Q, Ji G, Li J (2020) STEAP2 is down-regulated in breast cancer tissue and suppresses PI3K/AKT signaling and breast cancer cell invasion in vitro and in vivo. *Cancer Biol Ther* 21(3):278–291
- Yersal O, Barutca S (2014) Biological subtypes of breast cancer: prognostic and therapeutic implications. *World J Clin Oncol* 5(3):412–424
- Yousefi H, Fong J, Alahari SK (2022) NR4A Family genes: a review of comprehensive prognostic and gene expression profile analysis in breast cancer. *Front Oncol* 12:777824
- Yu G, Wang LG, Han Y, He QY (2012) clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: J Integr Biol* 16(5):284–287
- Zhang H, Meltzer P, Davis S (2013) RCircos: an R package for Circos 2D track plots. *BMC Bioinforma* 14:244
- Zhang S, Wang Y, Gu Y, Zhu J, Ci C, Guo Z et al (2018) Specific breast cancer prognosis-subtype distinctions based on DNA methylation patterns. *Mol Oncol* 12(7):1047–1060
- Zhu L, Yang F, Wang L, Dong L, Huang Z, Wang G et al (2021) Identification the ferroptosis-related gene signature in patients with esophageal adenocarcinoma. *Cancer Cell Int* 21(1):124

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.