

OPEN

# Homologous recombination deficiency status-based classification of high-grade serous ovarian carcinoma

Hisamitsu Takaya<sup>1</sup>, Hidekatsu Nakai<sup>1</sup>, Shiro Takamatsu<sup>2</sup>, Masaki Mandai<sup>2</sup> & Noriomi Matsumura<sup>1\*</sup>

Homologous recombination repair (HRR) pathway deficiency (HRD) is involved in the tumorigenesis and progression of high-grade serous ovarian carcinoma (HGSOC) as well as in the sensitivity to platinum chemotherapy drugs. In this study, we obtained data from The Cancer Genome Atlas (TCGA) on HGSOC and identified scores for the loss of heterozygosity, telomeric allelic imbalance, and large-scale state transitions, and calculated the HRD score. We then investigated the relationships among the score, genetic/epigenetic alterations in HRR-related genes, and the clinical data. We found that *BRCA1/2* mutations were enriched in the group with HRD scores  $\geq 63$ . Compared with the groups with scores  $\leq 62$ , this group had a good prognosis; we thus considered HRD scores  $\geq 63$  to be the best cutoff point for identifying HRD cases in HGSOC. Classification of HGSOC cases by the HRD status revealed a better prognosis for HRD cases caused by genetic alterations (genetic HRD) than those caused by epigenetic changes and those caused by undetermined reasons ( $p = 0.0002$ ). Among cases without macroscopic residual tumors after primary debulking surgery, 11 of 12 genetic HRD cases survived after the median observation period of 6.6 years, showing remarkably high survival rates ( $p = 0.0059$ ). In conclusion, HGSOC can be classified into subtypes with different prognoses according to HRD status. This classification could be useful for personalized HGSOC treatment.

Ovarian cancer has the worst prognosis of all gynecologic malignancies<sup>1</sup>. Specifically, high-grade serous ovarian cancer (HGSOC), which is the most common type of ovarian cancer, is mostly advanced stage III–IV disease at the time of diagnosis<sup>2</sup>. The primary treatment approach for this cancer is a combination of debulking surgery and a platinum-chemotherapy. However, the majority of ovarian cancer patients experience tumor recurrence; they develop chemotherapy resistance that ultimately becomes lethal<sup>3</sup>.

HGSOC is characterized by chromosomal instability due to homologous recombination repair (HRR) pathway deficiency (HRD)<sup>4</sup>. Germline *BRCA1/2* mutations, somatic *BRCA1/2* mutations, and *BRCA* gene promoter methylations are well-known causes of HRD, but other genetic abnormalities of the HRR pathway could also cause HRD<sup>4,5</sup>, although no consensus has been reached. The presence of HRD results in irreparable DNA damage from platinum-containing drugs, which leads to cell death. Moreover, an underlying HRD in tumor cells makes the cells sensitive to PARP inhibitors. PARP inhibitors bind to and trap PARP1 and PARP2 on DNA at the sites of single-strand breaks, which results in the generation of double-strand breaks. In cancer cells with HRD, double-strand DNA breaks are repaired by error-prone pathways (i.e. nonhomologous end joining), ultimately leading to cell death<sup>6</sup>.

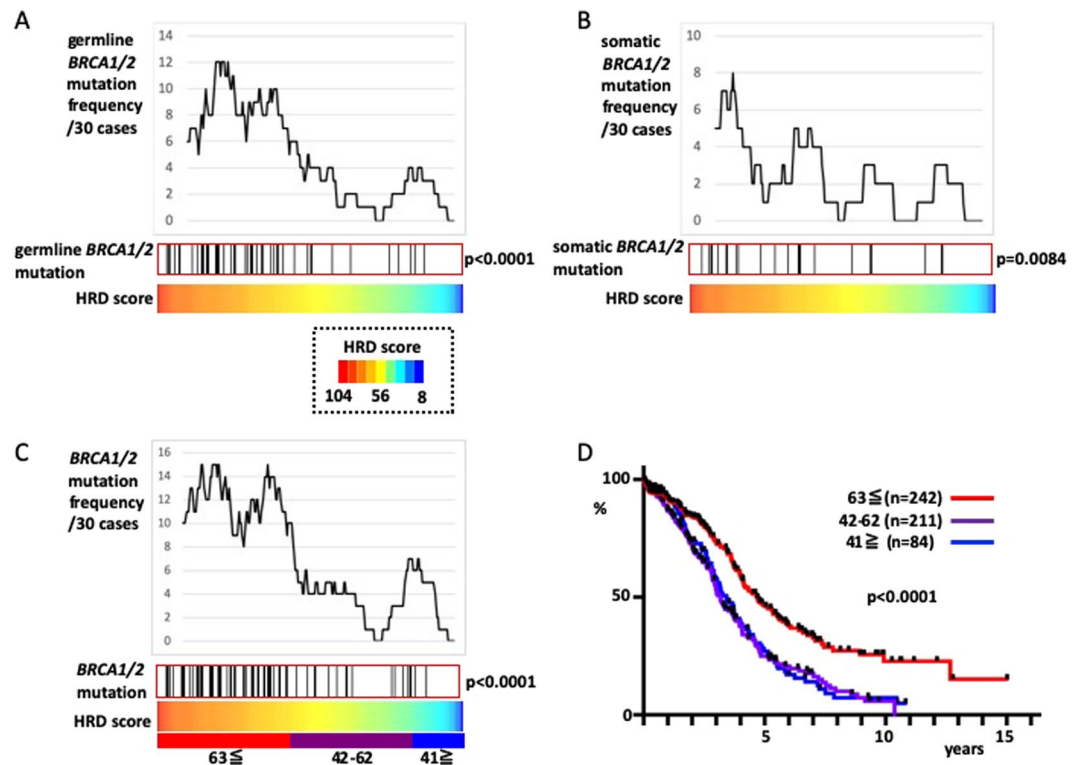
HRD causes characteristic genomic scar signatures, namely, the loss of heterozygosity (LOH)<sup>7</sup>, telomeric allelic imbalance (TAI)<sup>8</sup>, and large-scale state transitions (LST)<sup>9</sup>. The HRD score is the sum of these scar signature scores<sup>10</sup>. The HRD score correlates with sensitivity to niraparib, which is a PARP inhibitor<sup>11</sup>.

In The Cancer Genome Atlas (TCGA) project, about half of HGSOC cases are reported to have HRD due to an HRR pathway abnormality<sup>4</sup>. However, the relationship between the HRD score and HRR pathway gene abnormalities other than those in *BRCA1/2* has not been thoroughly investigated. We thus investigated the connection

<sup>1</sup>Kindai University Faculty of Medicine, Department of Obstetrics and Gynecology, Osakasayama, 589-8511, Japan.

<sup>2</sup>Kyoto University Graduate School of Medicine, Department of Gynecology and Obstetrics, Kyoto, 606-8507, Japan.

\*email: [noriomi@med.kindai.ac.jp](mailto:noriomi@med.kindai.ac.jp)



**Figure 1.** Definition of HRD cases according to HRD scores. The 296 cases for which both exome sequence and SNP array data were available were sorted by HRD score. (A) Distribution of cases with germline *BRCA1/2* mutations. The upper graph shows the frequency of germline *BRCA1/2* mutation cases by sliding the windows of 30 cases. The black bars in the central box represent germline *BRCA1/2* mutations cases. (B) Distribution of cases with somatic *BRCA1/2* mutations. (C) germline *BRCA1/2* mutations and somatic *BRCA1/2* mutations are presented together as *BRCA1/2* mutations; these cases are concentrated in HRD scores  $\geq 63$ . The cutoff HRD score  $\geq 42$ , which is generally used, is also shown. (D) Comparison of overall survival rate by HRD scores; 537 cases with known prognosis and SNP array data were divided according to HRD scores  $\geq 63$ , 42–62, and  $\geq 41$ .

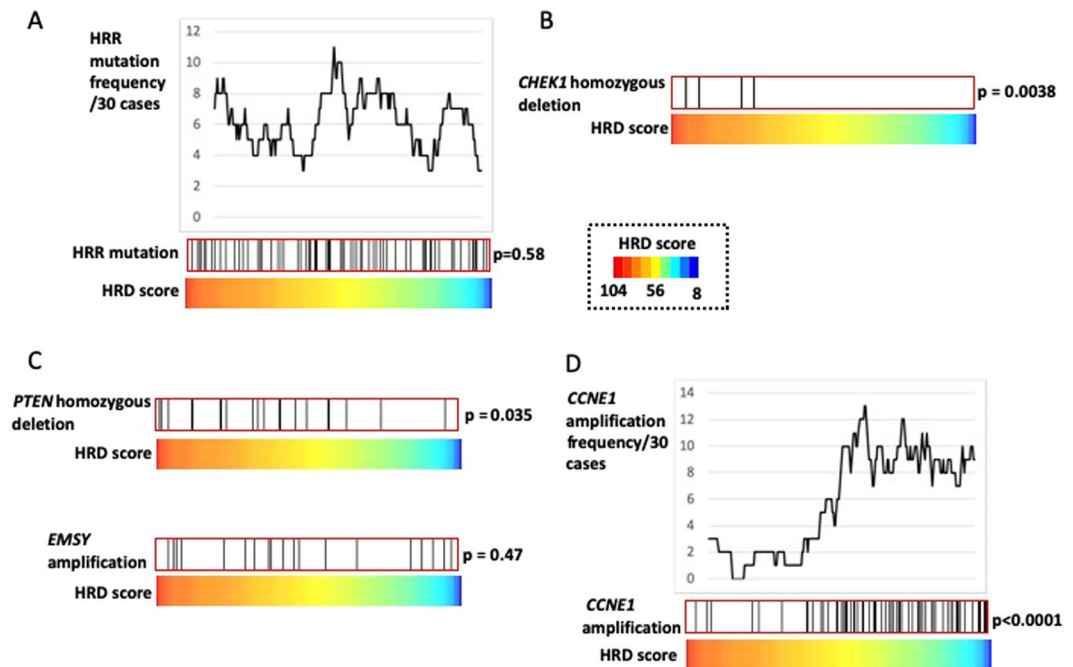
between HRD status and HRR pathway gene abnormalities in HGSOc data in the TCGA, and we show that HGSOc can be classified according to HRD status into subtypes with different prognoses.

## Results

First, we defined HRD cases by investigating the distribution of cases that harbored germline *BRCA1/2* or somatic *BRCA1/2* mutations. Compared with the cases without germline *BRCA1/2* mutations, those cases with germline *BRCA1/2* mutations had higher scores for LOH<sup>7</sup> ( $p = 0.0018$ ), TAI<sup>8</sup> ( $p < 0.0001$ ), and LST<sup>9</sup> ( $p < 0.0001$ ). Cases with somatic *BRCA1/2* mutations scored higher for LOH compared with those without somatic *BRCA1/2* mutations ( $p = 0.0018$ ;  $p = 0.28$  and  $0.06$  for TAI and LST scores, respectively). HRD scores (TAI + LST + LOH) were higher in those cases with germline *BRCA1/2* mutations or somatic *BRCA1/2* mutations, compared with those without (Fig. 1A,B;  $p = 0.0001$  and  $0.0084$  for germline *BRCA1/2* mutations and somatic *BRCA1/2* mutations, respectively). When germline *BRCA1/2* mutations and somatic *BRCA1/2* mutations were analyzed together as *BRCA* mutations, it was found that *BRCA* mutation cases had high HRD scores ( $p < 0.0001$ , Fig. 1C).

Until now, the proposed cutoff value for the HRD score has been  $\geq 42$ <sup>10</sup>. In the report where this cutoff was proposed, the ovarian cancer and breast cancer datasets were analyzed together. Therefore, we also analyzed the breast cancer TCGA dataset<sup>12</sup>, and found that generally, breast cancer had lower HRD scores than HGSOc (Supplementary Fig. 1A,  $p < 0.0001$ ), but higher HRD scores were still associated with *BRCA* mutation cases ( $p < 0.0001$ , Supplementary Fig. 1B). Combined analysis of ovarian and breast cancer datasets from 1257 cases showed that the distribution of *BRCA* mutation cases was bimodal. Indeed, there also seemed to be a cutoff at HRD scores  $\geq 42$  (Supplementary Fig. 1C). However, in the analysis of ovarian cancer alone, *BRCA* mutation cases were more likely to have HRD scores  $\geq 63$  (Fig. 1C). *BRCA* mutations represented 38% of the cases with HRD scores  $\geq 63$  (49/128), 10% with HRD scores between 42 and 62 (12/118), and 10% with HRD scores 41 and below (5/50); no enrichment of *BRCA* mutation cases with HRD scores from 42 to 62 was observed compared with those with scores  $\leq 41$ . An analysis of the prognosis revealed overlap in the survival curves of patients with HRD scores of 42–62 ( $n = 211$ ) and  $\leq 41$  ( $n = 84$ ), but the survival rate of patients with HRD scores  $\geq 63$  ( $n = 242$ ) was clearly better than that of the other two groups (Fig. 1D,  $p < 0.0001$ ). Thus, we designated an HRD score  $\geq 63$  as the cutoff.

Subsequently, we investigated the connection between the HRD score and overall HRR pathway gene mutations other than *BRCA* mutations but found none (Fig. 2A,  $p = 0.58$ ). Moreover, no difference was observed in

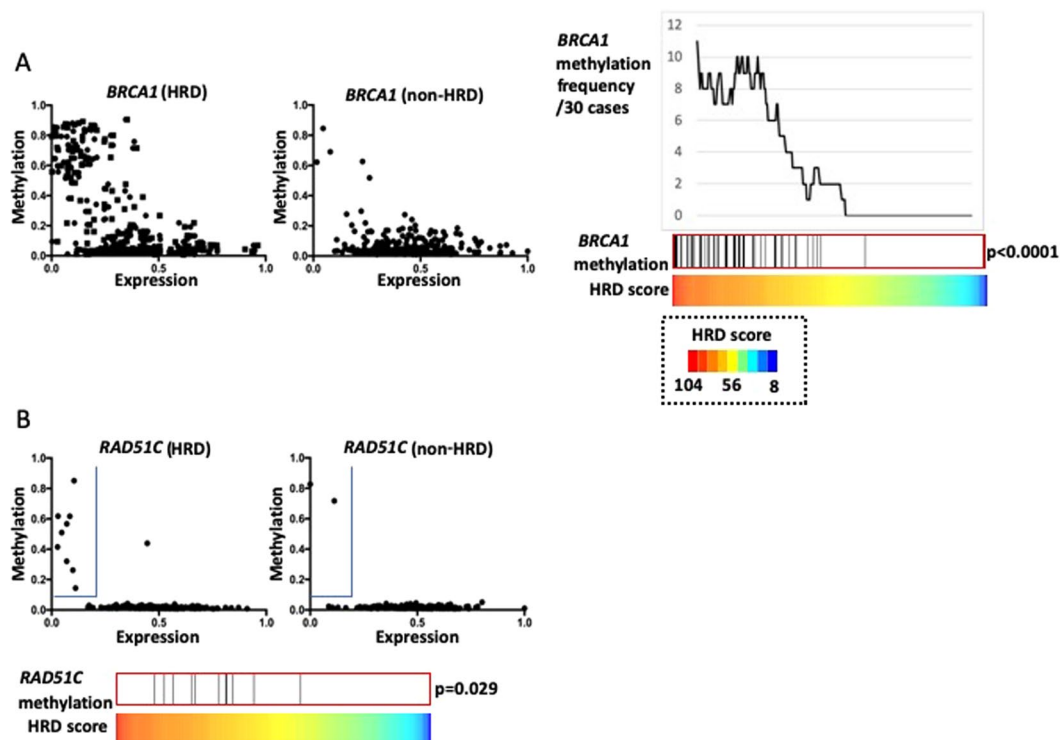


**Figure 2.** Relationship between HRD scores and mutations in HRR-related genes other than *BRCA1/2*. Dividing 296 cases according to HRD scores shows the distribution of cases with mutations. (A) Distribution of cases with HRR pathway gene mutations other than those in *BRCA1/2*, including homozygous deletions. (B) *CHEK1* homozygous deletions had high HRD scores among the HRR pathway gene mutations other than those in *BRCA1/2*. (C) *PTEN* homozygous deletions and *EMSY* amplifications are two gene mutations that may cause HRD. (D) Relationship between *CCNE1* amplifications and HRD score; *CCNE1* amplification cases were enriched in low HRD scores.

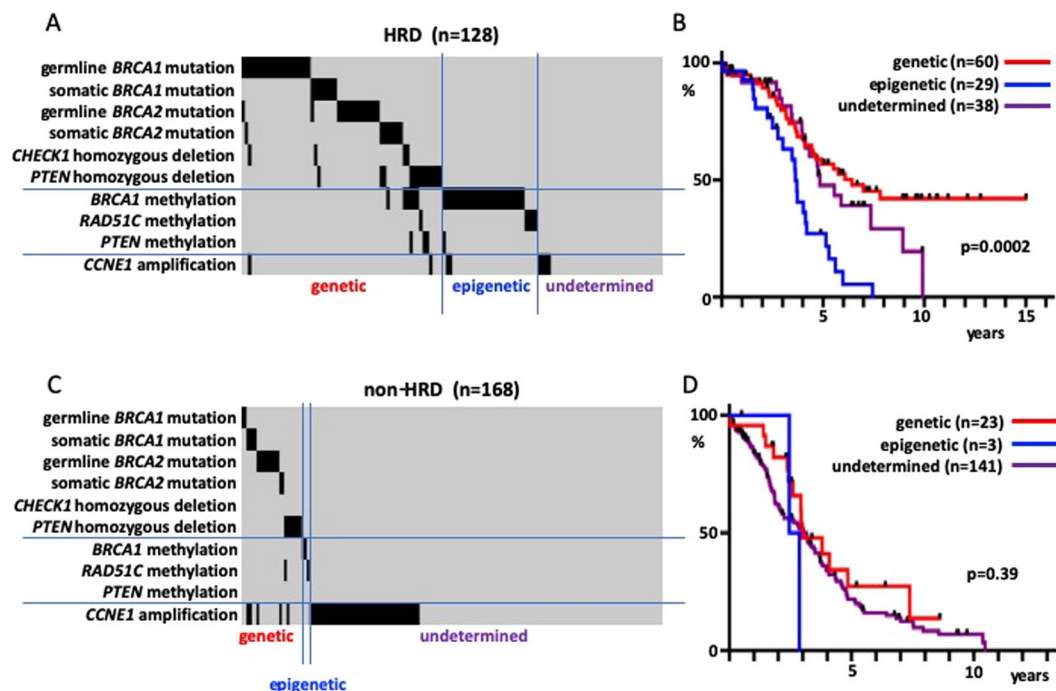
survival rates between patients with and without these HRR mutations ( $p = 0.78$ , data not shown). However, individual analyses of these HRR mutations revealed that the *CHEK1* homozygous deletions was associated with high HRD scores (Fig. 2B, Supplementary Fig. 2,  $p = 0.0038$ ). The *PTEN* homozygous deletions and *EMSY* amplifications are candidate mutations for causing HRD<sup>5</sup>. Of these, *PTEN* homozygous deletions were positively correlated with HRD scores ( $p = 0.035$ ), whereas *EMSY* amplifications were not (Fig. 2C, Supplementary Fig. 2). *CCNE1* amplifications, a genetic abnormality that characterizes HR-proficient HGSOCS<sup>5</sup>, were clearly associated with low HRD scores (Fig. 2D;  $p < 0.0001$ ).

Next, we analyzed DNA methylation in the promoter region of HRR-related genes. First, we investigated *BRCA1* methylation. Of the nine *BRCA1* methylation probes, the  $\beta$  value of cg04658354 was most negatively correlated with the *BRCA1* mRNA expression level ( $r = -0.49$ ). When *BRCA1* mRNA expression and cg04658354  $\beta$  values were plotted, markedly more methylated cases with a high  $\beta$  value and low mRNA expression were observed in the HRD group than in the non-HRD group (Fig. 3A). The 56 *BRCA1* methylated cases described in the TCGA paper<sup>4</sup> revealed higher HRD scores than non-methylated cases (Fig. 3A,  $p < 0.0001$ ). Similarly, *RAD51C* methylation probe cg14837411 showed strong negative correlation with the mRNA expressions ( $r = -0.40$ ). Cases with methylation, which were defined by a  $\beta$  value  $> 0.1$  and a normalized expression  $< 0.2$ , were concentrated in the high HRD cases ( $p = 0.029$ , Fig. 3B). Among the other HRR pathway-related genes, the only one with a strong negative correlation between the methylation probe  $\beta$  value and gene expression was *PTEN* cg21573601 ( $r = -0.40$ ). However, the maximum  $\beta$  value of this probe was low (0.12), and it is thus unclear whether this gene was actually methylated. Candidate *PTEN* methylation cases selected using a cutoff  $\beta$  value  $> 0.03$  and a normalized expression  $< 0.4$  had high HRD scores ( $p = 0.0022$ , Supplementary Fig. 3).

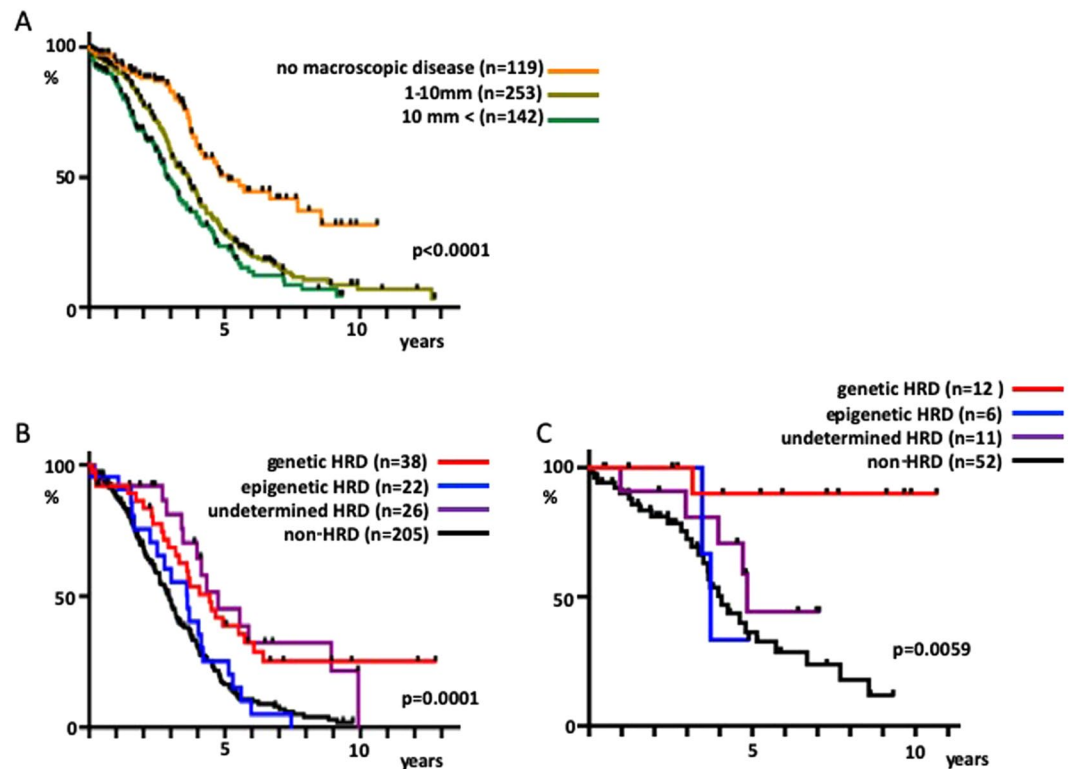
Considering the above findings, the HRD cases for which both SNP array and exome sequence data were available ( $n = 128$ ) were classified according to the underlying molecular mechanism. Those cases that had one of the six factors that could induce HRD genetically (germline *BRCA1* mutations, somatic *BRCA1* mutations, germline *BRCA2* mutations, somatic *BRCA2* mutations, *CHEK1* homozygous deletions, and *PTEN* homozygous deletions) were classified as genetic HRD. Other cases that had *BRCA1* methylations or *RAD51C* methylations were classified as epigenetic HRD, while the remainder were classified as undetermined HRD. There was no case with *PTEN* candidate methylation alone (Fig. 4A). The survival rates of these three groups (comprising 127 patients for which the overall survival (OS) data were available) revealed a good prognosis for genetic HRD and a poor prognosis for epigenetic HRD (Fig. 4B,  $p = 0.0002$ ). No difference was found in the prognosis between genetic HRD cases with *BRCA1/2* mutations and those with *CHEK1* homozygous deletions or *PTEN* homozygous deletions. Since there were only a few *RAD51C* methylation cases among the epigenetic HRD cases, it was difficult to compare their prognoses with those of the *BRCA1* methylation cases (Supplementary Fig. 4A). A similar classification of the non-HRD cases ( $n = 168$ ;  $n = 167$  for those with OS information; Fig. 4C) revealed



**Figure 3.** Relationship between HRD status and DNA promoter methylation of HRR-related genes. (A) *BRCA1* methylation. Left: HRD cases; Center: non-HRD cases. X-axis: Average values of two *BRCA1* mRNA expression probes normalized to 0–1. Y-axis:  $\beta$  value of cg04658354. Right: Distribution of *BRCA1* methylated cases. (B) *RAD51C* methylation. X-axis: Average values of two *RAD51C* mRNA expression probes normalized to 0–1. Y-axis:  $\beta$  value of cg14837411. *RAD51C* methylation cases are surrounded by a blue line in the upper-left region. The frequency of *RAD51C* methylation differs greatly between HRD cases (left) and non-HRD cases (right).



**Figure 4.** Relationship between the molecular mechanism of HRD and prognosis; sorting of cases by mechanism into the genetic mutation (genetic), DNA methylation (epigenetic), and unknown mechanism (undetermined) groups. (A) Classification of HRD cases. (B) Survival rate of HRD cases. (C) Classification of non-HRD cases. (D) Survival rate of non-HRD cases.



**Figure 5.** Relationship among residual tumor at the time of surgery, HRD status, and prognosis. **(A)** The 514 cases were sorted by size of the residual tumor if data on residual tumor at the time of surgery were available (no macroscopic disease, 1–10 mm and >10 mm) and their survival rates were compared. **(B)** Cases with residual tumor  $\geq 1$  mm and available SNP array data were further divided into the HRD and non-HRD groups. HRD cases with exome sequencing data were further sorted by molecular mechanism into genetic HRD, epigenetic HRD, and undetermined HRD to compare the survival rate. **(C)** Similarly, cases with no macroscopic disease were classified by the HRD status for survival rate comparison.

no difference in survival rates (Fig. 4D,  $p = 0.39$ ). Additionally, a survival rate comparison between HRD and non-HRD cases with the aforementioned six factors causing genetic HRD revealed a better prognosis for HRD cases (Supplementary Fig. 4B,  $p = 0.019$ ).

Finally, in the analysis, we included data on residual tumor at the time of surgery. The 514 cases were sorted by tumor size if data on residual tumor at the time of surgery were available, and a poor long-term prognosis for those groups with residual tumors  $\geq 1$  mm was revealed regardless of size (Fig. 5A,  $p < 0.0001$ ). Among the cases with residual tumor  $\geq 1$  mm, approximately 25% of genetic HRD and undetermined HRD patients had a long-term survival  $>7$  years, whereas non-HRD and epigenetic HRD patients had low long-term survival rates (Fig. 5B,  $p = 0.0001$ ). Intriguingly, among patients with no macroscopic diseases, genetic HRD patients had a very high survival rate, and 11 of 12 patients survived after a median observation period of 6.6 years (Fig. 5C,  $p = 0.0059$ ).

## Discussion

HGSOC data from the TCGA became publicly available in 2011 and have been analyzed in numerous reports since then. However, in this study, we still found some important observations from these data. Our contributions are as follows: (1) establishing a cutoff value to identify HRD from the analysis of HGSOC data alone (2); a comprehensive analysis of HRR pathway genes (3); proposal of a classification method for HRD cases according to cause; and (4) integration of residual tumor data.

A cutoff HRD score  $\geq 42$  was previously proposed in a study that combined ovarian and breast cancer SNP array data from the TCGA and two other datasets from custom panels<sup>10</sup>. After running a similar analysis that combined the breast cancer and HGSOC TCGA datasets, *BRCA1/2* mutations seemed to be enriched at HRD scores  $\geq 42$  (Supplementary Fig. 1C). However, a comparison of the HGSOC and breast cancer data revealed a large gap in HRD scores (Supplementary Fig. 1A). In another study, a pan-cancer analysis using the TCGA data revealed that the TAI, LST, and LOH scores of HGSOC, with a median of 24, 20, and 15, respectively, were higher than those of cancers derived from other organs; the corresponding median scores for breast cancer were 12, 8, and 8, respectively, for TAI, LST, and LOH, which were approximately half those of HGSOC<sup>13</sup>. Another study calculated the HRD scores of formalin-fixed paraffin-embedded samples using an OncoScan SNP array and found that HGSOC had far higher HRD scores than ovarian clear cell carcinoma, as over 80% of cases scored  $\geq 42$ <sup>14</sup>. HRD scores thus vary greatly between organ of origin and tissue type. As the aim of our study was to classify HGSOC by HRD status, we based our analyses on HGSOC data only and adopted an HRD score  $\geq 63$  as the cutoff. As a result, a little less than half of HGSOC patients were classified in the HRD group (Fig. 1C). In other

studies, authors performed SNP arrays with a custom panel, classified a little less than half of their samples as HRD with a cutoff HRD score  $\geq 42$ , and found a high frequency of *BRCA* mutations<sup>15,16</sup>. The ratio of HRD is consistent with findings from our study even though the cutoff point differs. In order to clarify why such a difference occurred, it would be beneficial if the raw SNP array data that were used to generate the HRD scores for research purposes would be open to the public.

The question of whether HRR pathway gene mutations other than *BRCA1/2* actually cause HRD has not been thoroughly investigated. Our TCGA data analyses indicated that these overall mutations are unrelated to the HRD score (Fig. 2A) and prognosis and that they are not involved in HRD or platinum sensitivity. Among these overall genes, only those cases with *CHEK1* homozygous deletions displayed high HRD scores (Fig. 2B). *CHEK1* is required for commencing HRR in cases of DNA double-strand breaks<sup>17</sup>, which means that *CHEK1* homozygous deletions are likely to cause HRD. Moreover, *PTEN* homozygous deletions correlated with high HRD scores (Fig. 2C). Therefore, we suggest that the only genetic changes that are involved in HRD in HGSOC besides the germline/somatic mutations in *BRCA1/2* are *CHEK1* and *PTEN* homozygous deletions. In support of our results, the HRD-associated LOH score is high in ovarian cancer with *PTEN* homozygous deletions, but ovarian cancer cases with *ATM*, *ATR*, *FANCA*, *FANCD2*, *FANCM*, and *PALB2* mutations do not have high LOH scores<sup>7</sup>.

It is well known that HRD is caused by promoter methylation of *BRCA1*<sup>5</sup>; indeed, HRD scores were high in cases in which *BRCA1* was methylated (Fig. 3A). The TCGA database only lists 168 genes including *BRCA1* because of the strict standards for  $\beta$  value and mRNA expression for the selection of methylated genes<sup>4</sup>. In addition, tumors with methylated *RAD51C* also transform to HRD<sup>5</sup> and have high LOH scores<sup>7</sup>. We focused on the negative correlation between the *RAD51C*  $\beta$  value and gene expression and found, as in previous reports, that *RAD51C* is often methylated in HRD cases (Fig. 3B).

We classified the HRD cases and found that epigenetic HRD cases had a poor prognosis, just like non-HRD cases (Fig. 4B). Previous studies have reported the correlation between *BRCA1* and *RAD51C* methylation and a good prognosis in ovarian cancer<sup>18,19</sup>. However, many recent studies do not support these findings<sup>20–23</sup>. In some cases, methylated *BRCA1* in tumors was found to be demethylated after chemotherapy<sup>24,25</sup>. The presence of *BRCA1* methylation before treatment was associated with post relapse platinum resistance in a study of triple-negative breast cancer<sup>26</sup>. These results suggest that demethylation of methylated *BRCA1* occurs more readily than genetic changes such as the reversion mutation in *BRCA1/2* mutation cases<sup>24,27</sup>. Our findings for the elevated HRD scores in the tumors with *BRCA1* or *RAD51C* methylations (Fig. 3) indicate that there is a poor prognostic subtype in the high-HRD score cases and warrants HRD-positive cases should be classified into epigenetic and non-epigenetic cases when conducting clinical trials.

Compared with genetic HRD cases, non-HRD cases had a poorer prognosis despite the presence of HRD-inducing genetic changes (Supplementary Fig. 3B). Presumably, tumors that managed to regain DNA repair functions through some mechanism became dominant in such cases, even though the DNA repair defects is initially important to carcinogenesis. In HGSOC with *BRCA* abnormalities, tumor infiltration by lymphocytes was augmented<sup>28,29</sup>, which was likely due to tumor immunity triggered by neo-antigen production as a result of DNA repair disorders. In the course of tumor progression, some tumors that regain DNA repair function may become dominant as the tumors interact with immune cells.

Finally, we performed analyses that integrated residual tumors after surgery. The HGSOC TCGA dataset comprises patients who underwent primary debulking surgery. We found that over 90% of genetic HRD patients who had no residual tumor are long-term survivors (Fig. 5C). Consistent with our data, stage III HGSOC cases who survived more than ten years were associated with germline or somatic *BRCA1/2* mutations and no residual disease after primary debulking surgeries<sup>30</sup>. This finding indicates that genetic HRD patients may be those who should undergo the most extensive surgery in order to extend their survival. Presumably, it is possible that tumor cells which harbor genetic alterations like reversion mutations<sup>24,27</sup> could be minimized in genetic HRD cases by complete resection. Whether or not a similarly high survival rate could be obtained in genetic HRD patients with complete surgery by interval debulking surgery following neoadjuvant chemotherapy remains to be elucidated. Several recent studies have evaluated interval debulking surgery<sup>31,32</sup>; genomic analyses of such clinical study samples would provide the answer to this question.

In conclusion, we classified the HGSOC cases into genetic HRD, epigenetic HRD, undetermined HRD, and non-HRD, each of which had a different prognosis. We found that in the genetic HRD fraction, a high survival rate could be increased by complete resection. Our data could be the foundational basis for future clinical studies in HGSOC.

## Materials and Methods

**Datasets.** Raw data from the Affymetrix Genome-Wide Human SNP Array 6.0 and the HT Human Genome U133 Array (Affymetrix) as well as clinical data on HGSOC<sup>4</sup> and breast invasive carcinoma<sup>12</sup> were obtained from the GDC Data Portal (<https://portal.gdc.cancer.gov>). In addition, processed data from Infinium HumanMethylation27 BeadChip (Illumina) was obtained from the GDAC Firehose (<http://gdac.broadinstitute.org>).

**Data processing.** SNP array data of cases for which tumor samples could be matched with a normal sample were extracted. For both tumor and normal sample data, the log R ratio and B-allele frequency of each probe was calculated using Affymetrix Power Tools and PennCNV<sup>33</sup>, and segmented copy number data were generated using allele-specific copy number analysis of tumors<sup>34</sup>. Gene expression data were normalized by the Robust Multi-array Average method using the “affy” package of R.

**HRD score analysis.** Genomic scar signatures were analyzed using an R program, as previously described<sup>13</sup>, and TAI, LST, and LOH scores were calculated. Briefly, TAI was defined as the number of regions of allelic imbalance that extended to the telomeres after deciding upon the major copy number state to be used to adjust biases in

samples with an uneven chromosome count. LOH was defined as the number of chromosomal LOH regions longer than 15 Mb. LST was defined as the number of break points between regions longer than 10 Mb after filtering out regions shorter than 3 Mb. The HRD score was calculated as the sum of the TAI, LST, and LOH scores. The calculated scores are shown in Supplementary Table 1 (ovarian cancer) and in Supplementary Table 2 (breast cancer).

**Homologous recombination repair (HRR)-related gene analysis.** According to the ARIEL3 study and other reports<sup>35,36</sup>, *BRCA1*, *BRCA2*, *ATM*, *ATR*, *BARD1*, *BLM*, *BRIP1*, *CDK12*, *CHEK1*, *CHEK2*, *FANCA*, *FANCC*, *FANCD2*, *FANCE*, *FANCF*, *FANCI*, *FANCL*, *FANCM*, *MRE11*, *NBN*, *PALB2*, *RAD50*, *RAD51*, *RAD51B*, *RAD51C*, *RAD51D*, *RAD52*, *RAD54L*, and *RPA1* were identified as HRR pathway genes. Cases of germline mutations or somatic mutations including homozygous deletions were identified. Moreover, cases of *EMSY* amplification and *PTEN* homozygous deletions, which may cause HRD, and the *CCNE1* amplifications<sup>5</sup>, which are associated with HR proficiency, were identified. These cases were identified using the cBioPortal for Cancer Genomics (<http://www.cbioportal.org>)<sup>37,38</sup>.

The correlation analysis of HRR pathway genes was performed using promoter methylation DNA microarray data and mRNA expression microarray data. When multiple mRNA expression probes existed, average of these probes was calculated for each case and then normalized from 0 to 1. Subsequently, we selected methylation probes which  $\beta$  values showed strong negative correlation ( $r < -0.30$ ) with the mRNA expressions. If multiple methylation probes were selected for one gene, the most negatively correlated probe was used.

**Heatmap.** A heatmap of the HRD score was created using R 3.2.2 and the MATLAB RGB color specifier. Other heatmaps were drawn using Java TreeView version 1.1.6r4.

**Statistical analysis.** We determined the presence of genetic abnormalities, and using the Mann–Whitney U test, we compared the TAI, LST, LOH, and HRD scores between tumors with genetic abnormalities and those without them. We also compared the HRD scores between HGSOC and breast cancer using the Mann–Whitney U test. Correlation analyses were performed using the Pearson correlation coefficient. Survival function was estimated using a Kaplan–Meier analysis, and survival curves were compared between groups using the log-rank test. All statistical analyses were performed using GraphPad Prism 6 (GraphPad Software, San Diego) with an  $\alpha$  of 0.05.

Received: 25 July 2019; Accepted: 3 February 2020;

Published online: 17 February 2020

## References

1. Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2019. *CA. Cancer J. Clin.* **69**, 7–34 (2019).
2. Seidman, J. D. *et al.* The histologic type and stage distribution of ovarian carcinomas of surface epithelial origin. *Int. J. Gynecol. Pathol.* **23**, 41–44 (2004).
3. Miller, D. S. *et al.* Phase II evaluation of pemetrexed in the treatment of recurrent or persistent platinum-resistant ovarian or primary peritoneal carcinoma: a study of the Gynecologic Oncology Group. *J. Clin. Oncol.* **27**, 2686–2691 (2009).
4. Cancer Genome Atlas Research Network. Integrated genomic analyses of ovarian carcinoma. *Nature*. **474**, 609–615 (2011).
5. Konstantinopoulos, P. A., Ceccaldi, R., Shapiro, G. I. & D'Andrea, A. D. Homologous Recombination Deficiency: Exploiting the Fundamental Vulnerability of Ovarian Cancer. *Cancer Discov.* **5**, 1137–1154 (2015).
6. Ledermann, J. A. & Pujade-Lauraine, E. Olaparib as maintenance treatment for patients with platinum-sensitive relapsed ovarian cancer. *Ther. Adv. Med. Oncol.* **11**, 1758835919849753 (2019).
7. Abkevich, V. *et al.* Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br. J. Cancer.* **107**, 1776–1782 (2012).
8. Birkbak, N. J. *et al.* Telomeric allelic imbalance indicates defective DNA repair and sensitivity to DNA-damaging agents. *Cancer Discov.* **2**, 366–375 (2012).
9. Popova, T. *et al.* Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res.* **72**, 5454–5462 (2012).
10. Telli, M. L. *et al.* Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. *Clin. Cancer Res.* **22**, 3764–3773 (2016).
11. Mirza, M. R. *et al.* Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. *N. Engl. J. Med.* **375**, 2154–2164 (2016).
12. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature*. **490**, 61–70 (2012).
13. Marquard, A. M. *et al.* Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomark. Res.* **3**, 9 (2015).
14. Chao, A. *et al.* Genomic scar signatures associated with homologous recombination deficiency predict adverse clinical outcomes in patients with ovarian clear cell carcinoma. *J. Mol. Med. (Berl)*. **96**, 527–536 (2018).
15. Stronach, E. A. *et al.* Biomarker Assessment of HR Deficiency, Tumor BRCA1/2 Mutations, and CCNE1 Copy Number in Ovarian Cancer: Associations with Clinical Outcome Following Platinum Monotherapy. *Mol. Cancer Res.* **16**, 1103–1111 (2018).
16. Patel, J. N. *et al.* Characterisation of homologous recombination deficiency in paired primary and recurrent high-grade serous ovarian cancer. *Br. J. Cancer.* **119**, 1060–1066 (2018).
17. Smith, J., Tho, L. M., Xu, N. & Gillespie, D. A. The ATM-Chk2 and ATR-Chk1 pathways in DNA damage signaling and cancer. *Adv. Cancer Res.* **108**, 73–112 (2010).
18. Ignatov, T. *et al.* BRCA1 promoter methylation is a marker of better response to platinum-taxane-based therapy in sporadic epithelial ovarian cancer. *J. Cancer Res. Clin. Oncol.* **140**, 1457–1463 (2014).
19. Bai, X. *et al.* BRCA1 promoter hypermethylation in sporadic epithelial ovarian carcinoma: Association with low expression of BRCA1, improved survival and co-expression of DNA methyltransferases. *Oncol. Lett.* **7**, 1088–1096 (2014).
20. Ruscito, I. *et al.* BRCA1 gene promoter methylation status in high-grade serous ovarian cancer patients—a study of the tumour Bank ovarian cancer (TOC) and ovarian cancer diagnosis consortium (OVCAD). *Eur. J. Cancer.* **50**, 2090–2098 (2014).
21. Sun, T. *et al.* Genetic Versus Epigenetic BRCA1 Silencing Pathways: Clinical Effects in Primary Ovarian Cancer Patients: A Study of the Tumor Bank Ovarian Cancer Consortium. *Int. J. Gynecol. Cancer.* **27**, 1658–1665 (2017).
22. Bernards, S. S. *et al.* Clinical characteristics and outcomes of patients with BRCA1 or RAD51C methylated versus mutated ovarian carcinoma. *Gynecol. Oncol.* **148**, 281–285 (2018).

23. Zhu, X., Zhao, L. & Lang, J. The BRCA1 Methylation and PD-L1 Expression in Sporadic Ovarian Cancer. *Int. J. Gynecol. Cancer*. **28**, 1514–1519 (2018).
24. Patch, A. M. *et al.* Whole-genome characterization of chemoresistant ovarian cancer. *Nature*. **521**, 489–494 (2015).
25. Prieske, K. *et al.* Loss of BRCA1 promotor hypermethylation in recurrent high-grade ovarian cancer. *Oncotarget*. **8**, 83063–83074 (2017).
26. Tutt, A. *et al.* Carboplatin in BRCA1/2-mutated and triple-negative breast cancer BRCAness subgroups: the TNT Trial. *Nat. Med.* **24**, 628–637 (2018).
27. Sakai, W. *et al.* Secondary mutations as a mechanism of cisplatin resistance in BRCA2-mutated cancers. *Nature*. **451**, 1116–1120 (2008).
28. McAlpine, J. N. *et al.* BRCA1 and BRCA2 mutations correlate with TP53 abnormalities and presence of immune cell infiltrates in ovarian high-grade serous carcinoma. *Mod. Pathol.* **25**, 740–750 (2012).
29. George, J. *et al.* Nonequivalent gene expression and copy number alterations in high-grade serous ovarian cancers with BRCA1 and BRCA2 mutations. *Clin. Cancer Res.* **19**, 3474–3484 (2013).
30. Yang, S. Y. C. *et al.* Landscape of genomic alterations in high-grade serous ovarian cancer from exceptional long- and short-term survivors. *Genome Med.* **10**, 81 (2018).
31. Vergote, I. *et al.* Neoadjuvant chemotherapy or primary surgery in stage IIIC or IV ovarian cancer. *N. Engl. J. Med.* **363**, 943–953 (2010).
32. Kehoe, S. *et al.* Primary chemotherapy versus primary surgery for newly diagnosed advanced ovarian cancer (CHORUS): an open-label, randomised, controlled, non-inferiority trial. *Lancet*. **386**, 249–257 (2015).
33. Wang, K. *et al.* PennCNV: an integrated hidden Markov model designed for high-resolution copy number variation detection in whole-genome SNP genotyping data. *Genome Res.* **17**, 1665–1674 (2007).
34. Van Loo, P. *et al.* Allele-specific copy number analysis of tumors. *Proc. Natl. Acad. Sci. USA* **107**, 16910–16915 (2010).
35. Coleman, R. L. *et al.* Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. *Lancet*. **390**, 1949–1961 (2017).
36. Bajrami, I. *et al.* Genome-wide profiling of genetic synthetic lethality identifies CDK12 as a novel determinant of PARP1/2 inhibitor sensitivity. *Cancer Res.* **74**, 287–297 (2014).
37. Gao, J. *et al.* Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, 11 (2013).
38. Cerami, E. *et al.* The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).

## Acknowledgements

Noriomi Matsumura reports that this work was supported by JSPS KAKENHI Grant Number 18H02947.

## Author contributions

H.T. designed the study and prepared the manuscript draft. N.M. and H.N. contributed in analyzing the data as well as in drafting the manuscript. S.T. and M.M. collected and interpreted the data and conducted a comprehensive literature review related to this project. All authors thoroughly reviewed the manuscript. All authors read and approved the final version of this manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-59671-3>.

**Correspondence** and requests for materials should be addressed to N.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020