OXFORD

JNCI J Natl Cancer Inst (2018) 110(3): djx176

doi: 10.1093/jnci/djx176
First published online September 23, 2017
Article

ARTICLE

Genomic Rearrangement Signatures and Clinical Outcomes in High-Grade Serous Ovarian Cancer

R. Tyler Hillman, Gary B. Chisholm, Karen H. Lu, P. Andrew Futreal

Affiliations of authors: Department of Gynecologic Oncology and Reproductive Medicine (RTH, GBC, KHL) and Department of Genomic Medicine (RTH, PAF), The University of Texas M.D. Anderson Cancer Center, Houston, TX

Correspondence to: P. Andrew Futreal, PhD, 1881 East Road, 3SCR6.3434, Unit 1954, Houston, TX 77054 (e-mail: afutreal@mdanderson.org).

Abstract

Background: To identify clinically relevant genomic rearrangement signatures in high-grade serous ovarian cancer (HGSOC), we conducted a retrospective analysis of sequenced HGSOC whole-tumor genomes.
Methods: Clinical data and whole-genome sequencing (WGS) reads were obtained for primary HGSOC tumors sequenced by the Australian Ovarian Cancer Study (AOCS; n = 80). Genomic rearrangements were identified, and non-negative matrix factorization (NMF) was used to extract rearrangement signatures. The cohort was then dichotomized around the median signature contribution, and overall survival (OS) was analyzed. An independent cohort from The Cancer Genome Atlas (TCGA) ovarian cancer study (n = 490) was also examined. The TCGA cohort was dichotomized around the median similarity between tumor copy number profile and a prognostic rearrangement signature, and OS was analyzed. Outcomes were assessed using Kaplan-Meier and multivariable Cox regression methods. All statistical tests were two-sided.
Results: We identified five genomic rearrangement signatures (Ov.RS1-5) in HGSOC. Ov.RS3 exhibited 10 kilobase to 10 megabase deletions and tandem duplications, and patients whose tumors exhibited a high contribution from Ov.RS3 had poor OS. The median OS was 22.7 months (95% confidence interval [CI] = 20.2 to 39.0 months) in the Ov.RS3-high group vs 38.2 months (95% CI = 22.7 to 69.1 months) in the Ov.RS3-low group (hazard ratio [HR] = 1.86, 95% CI = 1.12 to 3.09, $P = .02$). For the independent TCGA cohort, median OS rates were 38.0 months (95% CI = 35.3 to 41.4 months) in the Ov.RS3 high-similarity group vs 48.9 months (95% CI = 44.1 to 57.1 months) in the Ov.RS3 low-similarity group (HR = 1.54, 95% CI = 1.21 to 1.97, $P < .001$).
Conclusion: A novel genomic rearrangement signature is associated with poor prognosis in HGSOC.

High-grade serous ovarian cancer (HGSOC) is a lethal malignancy that is diagnosed in approximately 22 000 women per year in the United States (1). Initial treatment of advanced HGSOC involves a combination of cytoreductive surgery and platinum-based combination chemotherapy. Molecularly targeted therapies such as anti-vascular endothelial growth factor antibodies and poly (ADP ribose) polymerase (PARP) inhibitors are now being used in the treatment of recurrent disease (2). Of these targeted therapies, PARP inhibitors act in synergy with homologous recombination repair pathway deficiency (HRD), a genomic instability phenotype exhibited by a subset of HGSOC (3). Measurement of genomic instability in HGSOC is thus an important clinical biomarker in this disease.

The presence of HRD can be assessed using quantitative measures of genome-wide loss of heterozygosity (LOH), a pattern of genomic instability associated with HRD and originally identified through the study of copy number variants in large HGSOC cohorts (4–7). However, methods based on the detection of LOH fail to capture small-scale rearrangements as well as copy number–neutral events such as inversions and interchromosomal translocations. In contrast, whole-genome sequencing (WGS) provides comprehensive information on genomic

Received: April 17, 2017; **Revised:** June 14, 2017; **Accepted:** August 2, 2017

rearrangements, often with single-nucleotide resolution of rearrangement breakpoints. WGS has now become feasible on a large scale, and many tumor genomes have been sequenced across a variety of cancer types (8–10).

WGS was recently applied to 560 breast cancer samples, producing the largest catalog of somatic genomic rearrangements in any tumor type to date (8). In an effort to identify patterns of genomic rearrangement in these tumors, a well-described mathematical framework known as non-negative matrix factorization (NMF) was used to extract six genomic rearrangement "signatures" from these sequenced genomes (8,11). One of these signatures, characterized by frequent small tandem duplications, was previously found to be associated with germline *BRCA1* loss of function in breast cancer (8,12). We sought to apply this rearrangement signature discovery framework to sequenced HGSOC whole genomes for the purpose of identifying genomic rearrangement patterns with prognostic significance.

Methods

Data Sources

This research was approved by the institutional review board of The University of Texas M.D. Anderson Cancer Center. Aligned paired-end WGS reads produced from primary, chemotherapy-naïve HGSOC fresh frozen tumor and matched normal samples by the Australian Ovarian Cancer Study (AOCS) group (n = 80) were obtained from the European Bioinformatics Institute (EGAD00001000877) (10). In addition, aligned WGS reads for matched recurrent ascites or tumor tissue were obtained for a subset of the AOCS cohort (n = 10). Patients in this cohort provided written informed consent for participation in the original study, and all data were de-identified prior to use in the current study (10). Rearrangement signatures were extracted only from the cohort of primary tumors, but rearrangement signature contribution was determined for both primary and recurrent tumors (Figure 1A). Clinical data, information about homologous recombination loss of function mutations, and *CCNE1* amplification were obtained from tables accompanying the original publication (10). For The Cancer Genome Atlas (TCGA) cohort, clinical data and germline-filtered segmented copy number profiles generated from Affymetrix SNP6.0 microarrays were obtained from the Broad GDAC FireHose (gdac.broadinstitute.org) (5). Grade 1 tumors and early-stage (Fédération Internationale de Gynécologie et d'Obstétrique [FIGO] stage I or II) tumors were removed, resulting in a cohort of 490 patients (Figure 1B). Chemotherapy-"refractory" patients exhibited disease progression by clinical or CA-125 criteria while on primary treatment, and chemotherapy-"resistant" patients exhibited disease progression within six months from the end of primary treatment (10).

Identification of Genomic Rearrangement Signatures

For each primary (n = 80) and recurrent (n = 10) tumor sample in the AOCS cohort, ascatNgs was used to estimate purity and ploidy, and structural rearrangements were identified using BReakpoint AnalySiS (BRASS) (8,13). Small insertions and deletions (<1 kilobase) and those rearrangements with a low BRASS assembly score (less than 90) were removed. We used the R statistical software package to implement the rearrangement signature framework described by Nik-Zainal et al. (8,14). Briefly, rearrangement breakpoints were first categorized as nonclustered or clustered by applying a piecewise constant fitting method ("copynumber" R package) to the inter-rearrangement distance, which was taken as the length in base pairs from one rearrangement breakpoint to the next along the chromosome (8). Rearrangements were then categorized by length (1 kilobase to 10 kilobase, 10 kilobase to 100 kilobase, 100 kilobase to 1 megabase, 1 megabase to 10 megabase, and more than 10 megabase) and type (deletion, duplication, inversion, interchromosomal translocation). This method resulted in 32 categories of structural rearrangements (8).

The "NMF" R package was used to extract rearrangement signatures from the primary tumor samples (n = 80), implementing the procedure described by Brunet et al. (Figure 1C) (15). In order to determine the optimal number of rearrangement signatures present in the WGS cohort, we used published silhouette width–based methods and determined that five rearrangement signatures resulted in the highest reproducibility (Supplementary Figure 1, available online). Comparisons between vector representations of signatures were performed using the cosine similarity metric. Rearrangement contributions in each tumor were calculated using a version of the "deconstructSigs" R package with minor adaptations for use with genomic rearrangement signatures (16).

Extrapolation of Rearrangement Signatures to an Independent TCGA Cohort

In order to extrapolate WGS rearrangement signatures to TCGA copy number profile data, the copy number–neutral components (inversions, interchromosomal translocations) were excluded and the signature was rescaled to produce signatures containing only rearrangements resulting in a copy number change (Figure 1C). Copy number profile segments for each sample in the TCGA cohort were filtered to remove segments supported by less than two microarray probes and also to remove low-amplitude events, defined as a log ratio greater than –0.3 and less than 0.3. Copy number segments were categorized by size and type (deletion, duplication). For each sample in the TCGA cohort (n = 490), a vector representation of the copy number profile was then compared with Ov.RS3-CN using the cosine similarity metric. Tumors in the TCGA cohort were then dichotomized around the median signature similarity into "high"– and "low"–Ov.RS3-CN similarity groups. TCGA copy number profiles were similarly compared with Ov.RS2-CN using the cosine similarity metric for analysis by germline *BRCA1* mutation status.

Statistical Analyses

For the five WGS rearrangement signatures, Kaplan-Meier methodology was applied to estimate the median overall survival of patients with tumors in the "high"–signature contribution group compared with patients with tumors in the "low"–signature contribution group (17). A two-sided log rank test was used for comparison between groups. The hazard ratio (HR) was estimated using a Cox regression model for overall survival. Univariate comparisons between continuous variables were performed using a two-sided Wilcoxon rank-sum test, and either a Fisher's exact test or X^2 test was used for categorical data as appropriate. A two-sided *P* value of less than .05 was considered statistically significant. Time-dependent area under the receiver operating characteristic curve was computed using the R package "timeROC" (18). Calculation of Spearman's rank correlation coefficient and tests of association were performed using R.

ARTICLE

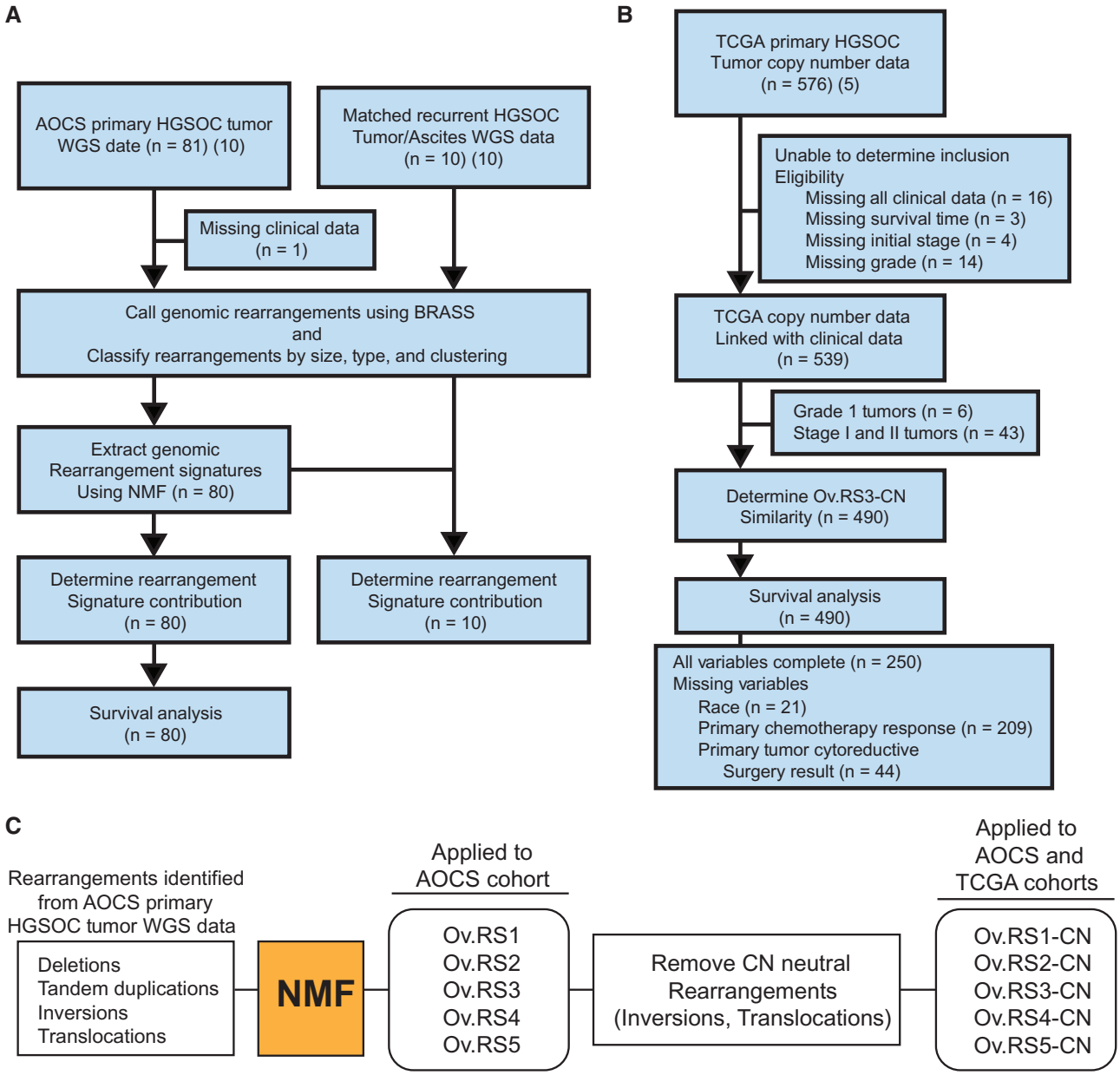

Figure 1. Diagram of experimental procedures. **A)** Australian Ovarian Cancer Study (AOCS) whole-genome sequencing cohort included 81 primary and 10 matched recurrent HGSOC tumor/normal pairs. One sample from this cohort was excluded because of the lack of available clinical information. Genomic rearrangement signatures were extracted from primary tumors in this cohort using non-negative matrix factorization. Survival analyses were performed using the same samples (n = 80) (10). **B)** The independent The Cancer Genome Atlas cohort was comprised of clinical information and copy number profiles generated from Affymetrix SNP6.0 microarrays (5). Note that for survival analysis, several records were missing data for multiple variables and thus the sum of complete records and records with missing variables does not equal 490. **C)** All rearrangements identified from the AOCS cohort were used for extraction of rearrangement signatures. Copy number–neutral events were then removed from these signatures, resulting in modified copy number rearrangement signatures. AOCS = Australian Ovarian Cancer Study; BRASS = BReakpoint AnalySiS; CN = copy number; HGSOC = high-grade serous ovarian cancer; NMF = non-negative matrix factorization; TCGA = The Cancer Genome Atlas; WGS = whole genome sequencing.

For the TCGA cohort (n = 490), complete clinical data were present for only 250 patients (Figure 1B). We therefore constructed Cox regression models using "unknown" categories to represent missing data. Variables were included in the multivariable model if they demonstrated a statistically significant univariate association with overall survival at a statistical significance threshold of a *P* value of less than .1. Variables examined included age at diagnosis (<65 vs ≥ 65 years), race (white vs black vs other), tumor grade (grade 2 vs grade 3 vs grade 4), FIGO stage at diagnosis (III vs IV) (19), Ov.RS3-CN similarity (below median vs above median), primary chemotherapy response (resistant/refractory vs sensitive), and primary tumor-reductive surgery result (no residual disease vs 1–10 mm residual disease vs > 10 mm residual disease). Variables were retained in the multivariable Cox regression model for overall survival at a statistical significance threshold of a *P* value of less than .05.

Plotting Schoenfeld residuals for covariates included in the Cox model suggested that age at diagnosis, chemotherapy response, and primary tumor-reductive surgery result violated the proportional hazards assumption (20). After stratification for these variables, the HR for the primary covariate of interest (Ov.RS3-CN similarity) remained statistically significant and relatively unperturbed (HR = 1.65, 95% confidence interval [CI] = 1.27 to 2.14, $P < .001$). Thus we chose to present the full Cox model, acknowledging that the HR estimates for age at diagnosis, chemotherapy response, and primary tumor-reductive surgery may reflect the averaging of nonproportional hazards over time. To determine the effect of missing data points on conclusions regarding the TCGA cohort, we used the covariates from the original Cox regression model to construct an additional model using listwise deletion of records with missing values (n = 250 remaining). For the analysis of the listwise deleted data, the assignment of the Ov.RS3-CN similarity category was recomputed for the reduced data set.

Results

Patient Characteristics

A total of 80 patients with WGS data from primary HGSOC tumors were included in the rearrangement signature discovery cohort (Figure 1A) (10). All patients in this cohort had advanced-stage disease (FIGO stage III or IV) and underwent primary cytoreductive surgery prior to receiving chemotherapy (Table 1). The TCGA cohort included 490 patients with advanced primary HGSOC (Table 1). Of patients in the TCGA cohort with known chemotherapy response, fewer patients had tumors that were resistant/refractory to primary chemotherapy compared with the AOCS cohort (26.3% vs 61.2%, $P < .001$), an observation that reflects the intention of the original AOCS WGS study to examine chemoresistance in HGSOC.

Table 1. Comparison of clinical and pathologic variables between AOCS and TCGA cohorts

Variables	AOCS cohort (n = 80) No. (%)	TCGA cohort (n = 490) No. (%)	P*
Age at diagnosis, y			.54
< 65	55 (68.8)	316 (64.5)	
≥65	25 (31.2)	174 (35.5)	
Primary chemotherapy response			<.001†
Refractory or resistant	49 (61.2)	74 (15.1)	
Sensitive	31 (38.8)	207 (42.2)	
Unknown	0 (0)	209 (42.7)	
Primary tumor-reductive surgery result			.005†
No residual disease	4 (5.0)	90 (18.4)	
1–10 mm residual disease	47 (58.8)	226 (46.1)	
>10 mm residual disease	29 (36.2)	130 (26.5)	
Unknown	0 (0)	44 (9.0)	
FIGO stage at diagnosis			.97
III	68 (85.0)	412 (84.1)	
IV	12 (15.0)	78 (15.9)	
Neo-adjuvant chemotherapy			—
No	80 (100)	492 (100)	
Yes	0 (0)	0 (0)	

*Two-sided P value from Fisher's exact test for age, primary chemotherapy response, and FIGO stage at diagnosis or from χ^2 test for primary tumor-reductive surgery result. AOCS = Australian Ovarian Cancer Study; FIGO = Fédération Internationale de Gynécologie et d'Obstétrique; TCGA = The Cancer Genome Atlas.

†Unknown data were excluded from the TCGA data set for the purpose of statistical comparison between cohorts.

Ovarian Cancer Genomic Rearrangement Signatures

We identified five HGSOC rearrangement signatures (Ov.RS1-5) that were distinct from one another (mean pairwise cosine similarity = 0.28) and notably did not contain substantial contributions from clustered rearrangements (mean signature contribution from clustered rearrangements < 0.01%) (Figure 2A). Comparison of these five HGSOC signatures with the six breast cancer rearrangement signatures identified by Nik-Zainal et al. indicated the presence of both shared and unique rearrangement signatures across tumor types (8). Specifically, Ov.RS1, Ov.RS2, Ov.RS4, and Ov.RS5 were very similar to signatures previously identified in breast cancer (cosine similarity > 0.90) (Figure 2B). Two breast cancer rearrangement signatures characterized by high contributions from complex clustered rearrangements were not identified in HGSOC (Br.RS4, Br.RS6) (8). Ov.RS3, characterized by medium to large (10 kilobase to 10 megabase) duplications and deletions, did not have high similarity with breast cancer rearrangement signatures (maximum cosine similarity = 0.70) (Figure 2B). The signature contribution from copy number–neutral events (inversions, translocations) was less than 50% for Ov.RS1, Ov.RS2, Ov.RS3, and Ov.RS5, but was 72.8% for Ov.RS4 (Figure 2C).

We next examined ovarian cancer rearrangement signature contributions in the 10 matched primary and recurrent tumors from the AOCS cohort (Figure 2D). We found an increase in the number of genomic rearrangements attributable to Ov.RS2 (median increase of 27.0 rearrangements per tumor, 95% CI = 13.5 to 41.0, $P = .004$) and Ov.RS3 (median increase of 22.5 rearrangements per tumor, 95% CI = 10.0 to 47.0, $P = .006$) when recurrent tumors were compared with matched primary tumors.

Genomic Rearrangement Signatures and BRCA1 Loss of Function

A previously identified rearrangement signature in breast cancer (Br.RS3) was associated with germline *BRCA1* loss of function (8,12). Br.RS3 is characterized by small tandem duplications, a phenotype reminiscent of the previously described tandem duplicator phenotype associated with *BRCA1* loss of function (21). We identified a genomic rearrangement signature (Ov.RS2) in HGSOC (Figure 3A) that demonstrated a high similarity with Br.RS3 (cosine similarity > 0.98).

Among ovarian cancers, we found that median Ov.RS2 rearrangement contribution showed a statistically significant difference between tumors with germline or somatic *BRCA1* mutations (median = 56.9%, interquartile range [IQR] = 36.5%–64.3%) and tumors with no homologous recombination pathway lesion (median = 4.6%, IQR = 0.4%–13.1%, $P < .001$) or tumors with non-*BRCA1* homologous recombination pathway mutations (median = 1.9%, IQR = 0.0%–2.2%, $P < .001$), which in this cohort included tumors with *BRCA2*, *BRIP1*, *CHEK2*, *FANCI*, *CDK12*, *PTEN*, or *RAD51C* mutations (Figure 3B). Median Ov.RS2 rearrangement contribution showed a statistically significant difference between tumors with germline *BRCA1* mutations (median = 62.5%, IQR = 54.1%–65.0%) compared with non-*BRCA1* mutated tumors (median = 3.4%, IQR = 0.0%–10.5%, $P < .001$). Similarly, median

ARTICLE

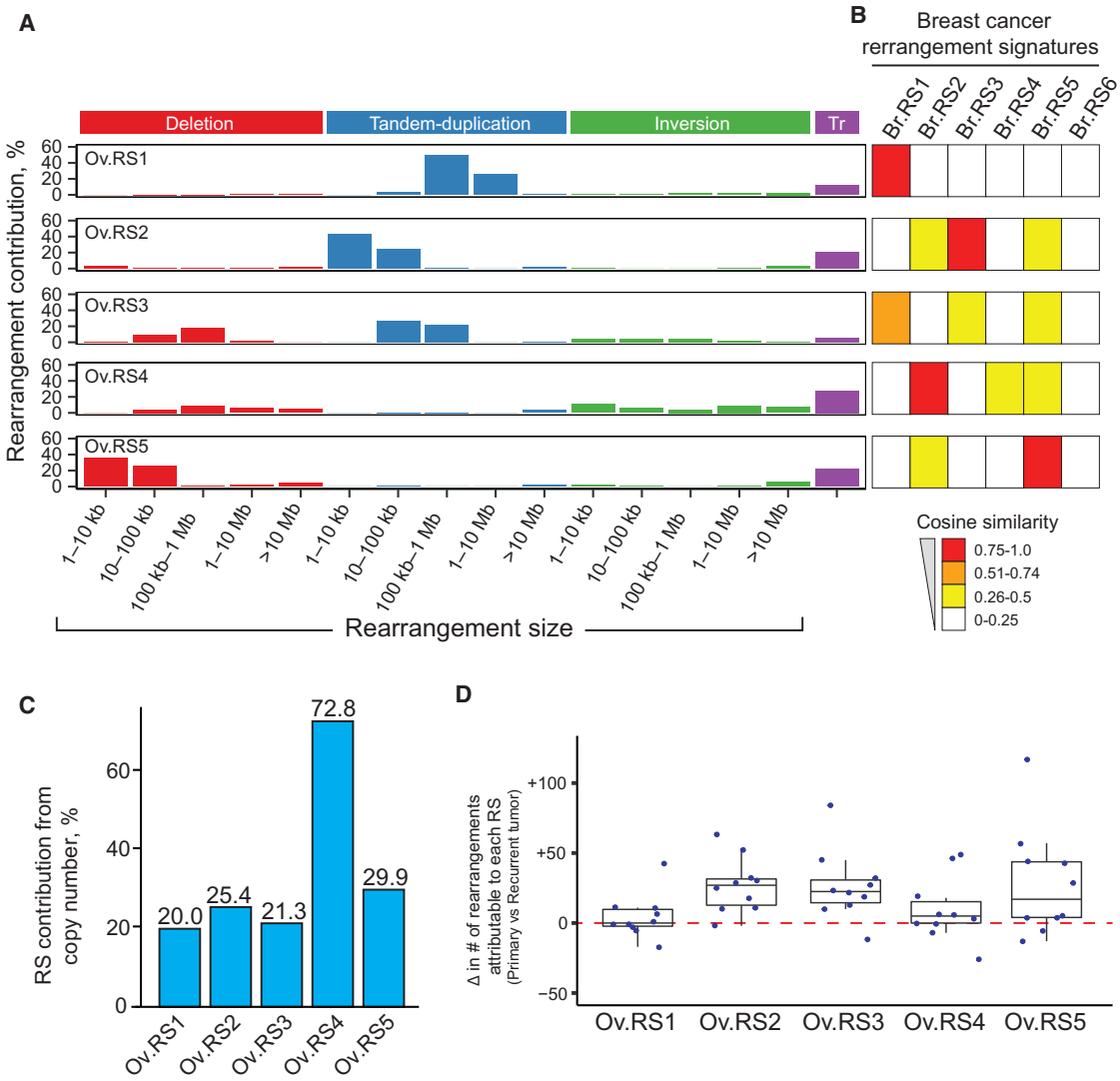

Figure 2. Five genomic rearrangement signatures identified from high-grade serous ovarian cancer (HGSOC) tumor whole-genome sequencing data. **A)** The five rearrangement signatures include distinct contributions from rearrangement classes and are dissimilar from one another (mean pairwise cosine similarity = 0.28). No signature had a substantial contribution from clustered rearrangements (mean signature contribution from clustered rearrangements < 0.01%), and thus for clarity clustered rearrangement contribution is not shown. **B)** The HGSOC rearrangement signatures were compared with previously identified breast cancer rearrangement signatures using cosine similarity, as represented by heatmap values (8). Ov.RS1, Ov.RS2, Ov.RS4, and Ov.RS5 were similar to breast cancer rearrangement signatures (cosine similarity > 0.90), but in contrast Ov.RS3 did not have high similarity with any breast cancer rearrangement signature (maximum cosine similarity = 0.70). **C)** The percent contribution from copy number–neutral events for each of the five HGSOC rearrangement signatures. **D)** Change in the number of rearrangements attributable to each rearrangement signature, as compared with 10 matched primary and recurrent tumor pairs from the same patients. RS = rearrangement signature; Tr = inter-chromosomal translocation.

Ov.RS2 rearrangement contribution showed a statistically significant difference between tumors with somatic *BRCA1* mutations (median = 27.1%, IQR = 20.6%–50.3%) compared with non-*BRCA1* mutated tumors (median = 3.4%, IQR = 0.0%–10.5%, $P = .01$) (Figure 3C). For the TCGA cohort, the Ov.RS2-CN copy number profile similarity showed a statistically significant difference between germline/somatic *BRCA1* mutated tumors (median = .20, IQR = .17–.27) compared with non-*BRCA1* mutated tumors (median = .15, IQR = .12–.21, $P < .001$) (Figure 3D).

Genomic Rearrangement Signatures and Overall Survival

Ov.RS1, Ov.RS2, Ov.RS4, and Ov.RS5 did not show a statistically significant association with overall survival (Ov.RS1: HR = 1.08, 95% CI = 0.66 to 1.77, $P = .76$; Ov.RS2: HR = 0.75, 95% CI = 0.46 to 1.23, $P = .25$; Ov.RS4: HR = 1.05, 95% CI = 0.64 to 1.71, $P = .86$; Ov.RS5: HR = 1.16, 95% CI = 0.53 to 1.40, $P = .55$; data not shown). In contrast, the median OS was 22.7 months (95% CI = 20.2 to 39.0 months) in the Ov.RS3-high group vs 38.2 months (95% CI = 22.7 to 69.1 months) in the Ov.RS3-low group (HR = 1.86, 95% CI = 1.12 to 3.09, $P = .02$) (Figure 4A).

Ov.RS3 is comprised of relatively equal contributions from medium to large nonclustered tandem duplications and deletions, with a small contribution from inversions (Figure 2A). No significant associations were identified between the Ov.RS3 contribution group and patient age at diagnosis ($P = .63$), primary chemotherapy response ($P = .17$), primary tumor-cytoreductive surgery result ($P = .54$), or FIGO stage at diagnosis ($P = .75$; data not shown). The per-tumor contribution of Ov.RS3

ARTICLE

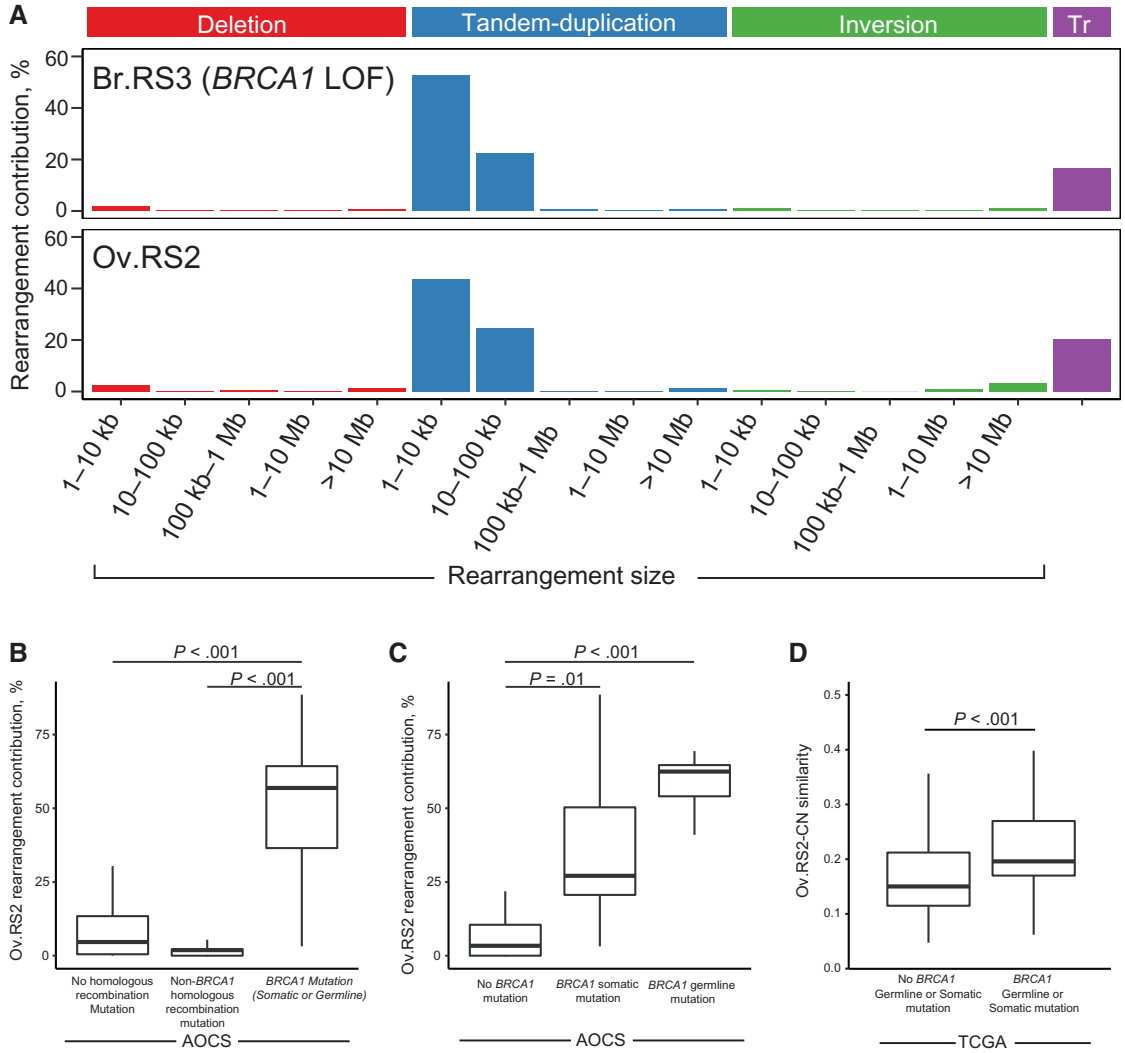

Figure 3. Ov.RS2 and *BRCA1* loss of function in high-grade serous ovarian cancer (HGSOC). **A)** Comparison between Ov.RS2 and the previously identified Br.RS3 rearrangement signature (8). The contribution of clustered rearrangements was omitted for clarity. **B)** Ov.RS2 rearrangement signature contribution in HGSOC tumors with germline or somatic *BRCA1* mutations compared with tumors with no homologous recombination pathway lesion. **C)** Ov.RS2 rearrangement signature contribution in HGSOC tumors with germline *BRCA1* mutations compared with tumors with somatic *BRCA1* mutations and non-*BRCA1* mutated tumors. **D)** Ov.RS2-CN rearrangement signature similarity in HGSOC tumors with germline or somatic *BRCA1* mutations compared with non-*BRCA1* mutated tumors in The Cancer Genome Atlas cohort. All comparisons were performed using a two-sided Wilcoxon rank-sum test. AOCS = Australian Ovarian Cancer Study; LOF = loss of function; TCGA = The Cancer Genome Atlas; Tr = inter-chromosomal translocation.

ARTICLE

showed a weak negative correlation with the number of genomic rearrangements (Spearman rank correlation coefficient = –0.29, two-sided P = .009), but was not correlated with tumor ploidy (Spearman rank correlation coefficient = 0.04, two-sided P = .75). In total, 16 of 80 samples (20.0%) demonstrated *CCNE1* amplification in this cohort. The per-tumor contribution of Ov.RS3 showed an association with *CCNE1* amplification that was close to achieving statistical significance (P = .06). In contrast, we found a statistically significant difference in per-tumor contribution of Ov.RS2 between tumors with *CCNE1* amplification (median = 0.8%, IQR = 0.0%–4.4%) and tumors without *CCNE1* amplification (median = 11.4%, IQR = 2.5%–46.4%, P < .001) (5,10).

Extrapolation of Prognostic Rearrangement Signature to TCGA Cohort

Among the TCGA cohort (n = 490), the median OS was 38.0 months (95% CI = 35.3 to 41.4 months) in the Ov.RS3-CN high-similarity group vs 48.9 months (95% CI = 44.1 to 57.1 months) in the Ov.RS3-CN low-similarity group (HR = 1.54, 95% CI = 1.21 to 1.97, P < .001) (Figure 4B). No statistically significant associations were identified between the Ov.RS3-CN similarity group and chemotherapy response (P = .46), primary tumor-cytoreductive surgery result (P = .51), or FIGO stage at diagnosis (P = .26). A statistically significant univariate association was identified between the Ov.RS3-CN similarity group and patient age at diagnosis (P = .006). In a multivariable Cox regression analysis of the TCGA cohort (n = 490), Ov.RS3-CN high-similarity patients had a higher risk of death after adjusting for covariates (HR = 1.59, 95% CI = 1.24 to 2.03, P < .001) (Table 2). A similar adjusted HR ratio for death among Ov.RS3-CN high-similarity patients was obtained by multivariable analysis of this data set with listwise deletion of records with missing values (n = 250; HR = 1.46, 95% CI = 1.08 to 1.98, P = .01) (Supplementary Table 1, available online).

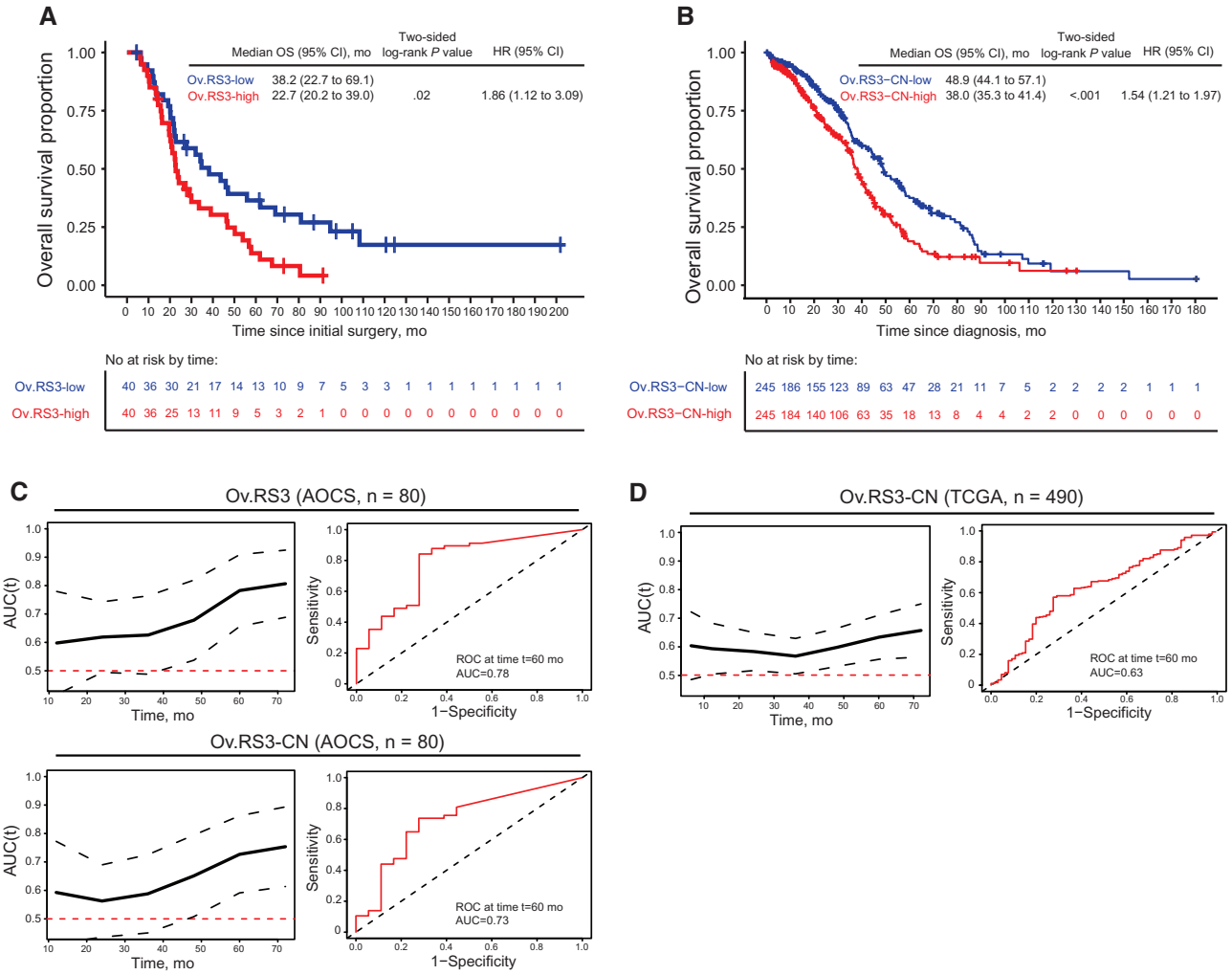

Figure 4. Overall survival and Ov.RS3. **A)** Kaplan-Meier analysis of overall survival in the Australian Ovarian Cancer Study (AOCS) cohort stratified by Ov.RS3 contribution relative to the median. **B)** Kaplan-Meier analysis of overall survival in the TCGA cohort stratified by Ov.RS3-CN similarity relative to the median. **C)** Time-dependent area under the receiver operating characteristic (ROC) curve for Ov.RS3 and Ov.RS3-CN applied to the AOCS cohort, including representative ROC curve for overall survival at 60 months. **D)** Time-dependent area under the ROC curve for Ov.RS3-CN applied to the TCGA cohort, including representative ROC curve for overall survival at 60 months. Comparisons were performed using a two-sided log-rank test, and HRs were estimated using univariate Cox regression. **Dotted lines** on the time-dependent area under the ROC curve plots represent 95% confidence interval for estimates. AOCS = Australian Ovarian Cancer Study; AUC = area under the receiver operating characteristic curve; CI = confidence interval; HR = hazard ratio; OS = overall survival; ROC = receiver operating characteristic; TCGA = The Cancer Genome Atlas.

Examination of the time-dependent area under the receiver operating characteristic curve (AUC(t)) for prediction of poor overall survival revealed diagnostic utility for Ov.RS3 applied to the AOCS cohort (Figure 4C), with an AUC of 0.78 at 60 months. When Ov.RS3-CN was applied to this cohort, the AUC was 0.73 at 60 months. A similar analysis of AUC(t) for Ov.RS3-CN applied to the TCGA cohort copy number profiles revealed somewhat lower diagnostic utility, with an AUC of 0.63 at 60 months (Figure 4D).

Discussion

In this study, we used a mathematical framework for signal decomposition known as NMF to identify distinct rearrangement signatures in sequenced HGSOC whole genomes. Of the five rearrangement signatures we identified, four are shared between HGSOC and breast cancer, including a signature associated with germline *BRCA1* loss of function (Ov.RS2) (8,12). We also identified Ov.RS3, a novel rearrangement signature comprised of medium to large deletions and duplications. HGSOC tumors with high contributions of genomic rearrangements from this signature demonstrated a poor clinical prognosis, and recurrent tumors exhibit a higher contribution from Ov.RS3 relative to matched primary tumors from the same patient. By extrapolating Ov.RS3 to the analysis of copy number profiles, we were able to demonstrate a similar prognostic association in an independent TCGA cohort. To our knowledge, this is the first description of a non-HRD-related genomic rearrangement phenotype associated with prognosis in HGSOC.

It is notable that the contribution of a rearrangement signature such as Ov.RS3 represents a qualitative description of rearrangement patterns and is independent of the total number of rearrangements present in a tumor genome. Just as Ov.RS2/Br.RS3 is associated with *BRCA1* loss of function, it is possible that Ov.RS3 results from the dysregulation of a different DNA repair pathway. It is interesting to note that high Ov.RS3-CN similarity was associated with increased age at the time of diagnosis among TCGA patients. The association between Ov.RS3-CN similarity and increasing age suggests the possibility

Table 2. Multivariable Cox regression model adjusted for all covariates estimating overall survival among patients from the TCGA cohort, stratified by Ov.RS3-CN similarity relative to the median (n = 490)

Variable	HR (95% CI)	P*
Age at diagnosis, y		
<65	1.00 (reference)	
≥65	1.54 (1.20 to 2.00)	<.001
FIGO stage at diagnosis		
III	1.00 (reference)	
IV	1.40 (1.01 to 1.93)	.04
Ov.RS3-CN similarity		
Below median	1.00 (reference)	
Above median	1.59 (1.24 to 2.03)	<.001
Primary chemotherapy response		
Refractory or resistant	1.00 (reference)	
Sensitive	0.39 (0.28 to 0.54)	<.001
Unknown	0.47 (0.32 to 0.68)	<.001
Primary tumor-reductive surgery result		
No residual disease	1.00 (reference)	
1–10 mm residual disease	1.73 (1.14 to 2.61)	.01
>10 mm residual disease	1.69 (1.10 to 2.62)	.02
Unknown	0.72 (0.40 to 1.31)	.28

*Multivariable analyses were performed using the Cox model for overall survival, and the statistical test was two-sided. CI = confidence interval; FIGO = Fédération Internationale de Gynécologie et d'Obstétrique; HR = hazard ratio; TCGA = The Cancer Genome Atlas.

that the mutational process underlying Ov.RS3 may be active in cells prior to the development of clinically detectable disease. If this is indeed the case, it is possible to speculate that rearrangement patterns detected in circulating cell-free DNA may allow for the early detection of ovarian cancer in the future.

This study is limited by its retrospective design, utilizing previously published HGSOC genomic data sets (5,10). It will be essential to validate the association between Ov.RS3 and clinical prognosis in prospective studies of HGSOC patients. One further limitation is the lack of a second large WGS cohort in which to validate the association between Ov.RS3 and prognosis. To circumvent this limitation, we assessed the similarity between an extrapolated version of this signature (Ov.RS3-CN) and copy number profiles in an independent TCGA cohort. However, WGS and copy number profile data are fundamentally different genomic analysis platforms, and it is possible that measurement of Ov.RS3-CN similarity does not fully capture the phenotype represented by the whole-genome Ov.RS3 rearrangement signature. A separate limitation of applying rearrangement signatures as a percent contribution per tumor is that an increasing contribution from one signature will necessarily decrease the contribution from the remaining signatures. For example, it seems likely that the strong association between *CCNE1* amplification and Ov.RS2 contribution in the AOCS cohort is responsible for the near statistical significance of the association between Ov.RS3 contribution and *CCNE1* amplification.

Prospective validation of the prognostic value of Ov.RS3 would result in a clinically applicable biomarker of HGSOC behavior. Such a biomarker could be used to identify patients with a poor prognosis who may be candidates for clinical trials in the adjuvant setting. The falling cost of next-generation sequencing makes it possible to foresee a future where the selective use of WGS in a clinical setting becomes a reality, especially in the management of tumor types such as HGSOC with frequent genomic rearrangements (22). Deriving clinical utility from WGS data will depend upon the concomitant application of analytical frameworks such as NMF, which are capable of distilling high-resolution genomic rearrangement data into relevant biomarkers and tumor subtypes.

Funding

RTH is supported by National Institutes of Health grant No. T32CA101642-11.

Notes

The funder had no role in design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

An oral abstract presentation related to this work was given at the American Society of Clinical Oncology annual meeting in Chicago, IL June 2–6, 2017.

References

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics. *CA Cancer J Clin*. 2014;64(1):9–29.
2. Yap TA, Carden CP, Kaye SB. Beyond chemotherapy: Targeted therapies in ovarian cancer. *Nat Rev Cancer*. 2009;9(3):167–181.
3. Ledermann JA, Drew Y, Kristeleit RS. Homologous recombination deficiency and ovarian cancer. *Eur J Cancer*. 2016;60:49–58.
4. Abkevich V, Timms KM, Hennessy BT, et al. Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br J Cancer*. 2012;107(10):1776–1782.
5. The Cancer Genome Atlas. Integrated genomic analyses of ovarian carcinoma. *Nature*. 2011;474(7353):609–615.
6. Watkins JA, Irshad S, Grigoriadis A, Tutt AN. Genomic scars as biomarkers of homologous recombination deficiency and drug response in breast and ovarian cancers. *Breast Cancer Res*. 2014;16(3):211.
7. Telli ML, Timms KM, Reid J, et al. homologous recombination deficiency (HRD) score predicts response to platinum-containing neoadjuvant chemotherapy in patients with triple-negative breast cancer. *Clin Cancer Res*. 2016;22(15):3764–3773.
8. Nik-Zainal S, Davies H, Staaf J, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature*. 2016;534(7605):1–20.
9. Bailey P, Chang DK, Nones K, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature*. 2016;531(7592):47–52.
10. Patch A-M, Christie EL, Etemadmoghadam D, et al. Whole-genome characterization of chemoresistant ovarian cancer. *Nature*. 2015;521(7553):489–494.
11. Alexandrov LB, Nik-Zainal S, Wedge DC, et al. Signatures of mutational processes in human cancer. *Nature*. 2013;500(7463):415–421.
12. Glodzik D, Morganella S, Davies H, et al. A somatic-mutational process recurrently duplicates germline susceptibility loci and tissue-specific super-enhancers in breast cancers. *Nat Genet*. 2017;49(3):341–348.
13. Raine KM, Van Loo P, Wedge DC, et al. ascatNgs: Identifying somatically acquired copy-number alterations from whole-genome sequencing data. *Curr Protoc Bioinforma*. 2016;56:15.9.1–15.9.17.
14. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. 2016. https://www.R-project.org/. Accessed July 24, 2017.
15. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A*. 2004;101(12):4164–4169.
16. Rosenthal R, McGranahan N, Herrero J, Taylor BS, Swanton C. deconstructSigs: Delineating mutational processes in single tumors distinguishes DNA repair deficiencies and patterns of carcinoma evolution. *Genome Biol*. 2016;17(1):31.
17. Kaplan EL, Meier P. Nonparametric estimation from incomplete observation. *J Am Stat Assoc*. 1958;53(282):457–481.
18. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med*. 2013;32(30):5381–5397.
19. Prat J, FIGO Committee on Gynecologic Oncology. Staging classification for cancer of the ovary, fallopian tube and peritoneum. *Obstet Gynecol*. 2015;126(1):171–174.
20. Schoenfeld D. Partial residuals for the proportional hazards regression model. *Biometrika*. 1982;69(1):239–241.
21. Menghi F, Inaki K, Woo X, et al. The tandem duplicator phenotype as a distinct genomic configuration in cancer. *Proc Natl Acad Sci U S A*. 2016;113(17):E2373–E2382.
22. National Human Genome Research Institute. The cost of sequencing a human genome. 2017. https://www.genome.gov/sequencingcosts/. Accessed July 24, 2017.

ARTICLE