

The DNA Damage Response (DDR) landscape of endometrial cancer defines discrete disease subtypes and reveals therapeutic opportunities

Xingyuan Zhang^{1,†}, Sayali Joseph^{1,†}, Di Wu², Jessica L. Bowser^{1,3,*} and Cyrus Vaziri^{1,3,*}

¹Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, School of Medicine, Chapel Hill, NC - 27599, USA

²Department of Biostatistics, University of North Carolina at Chapel Hill, School of Dentistry, Chapel Hill, NC - 27599, USA

³UNC Lineberger Comprehensive Cancer Center, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC - 27599, USA

*To whom correspondence should be addressed. Tel: +1 919 843 9639; Email: cyrus_vaziri@med.unc.edu

Correspondence may also be addressed to Jessica L. Bowser. Tel: +1 919 445 2361; Email: jessica_bowser@med.unc.edu

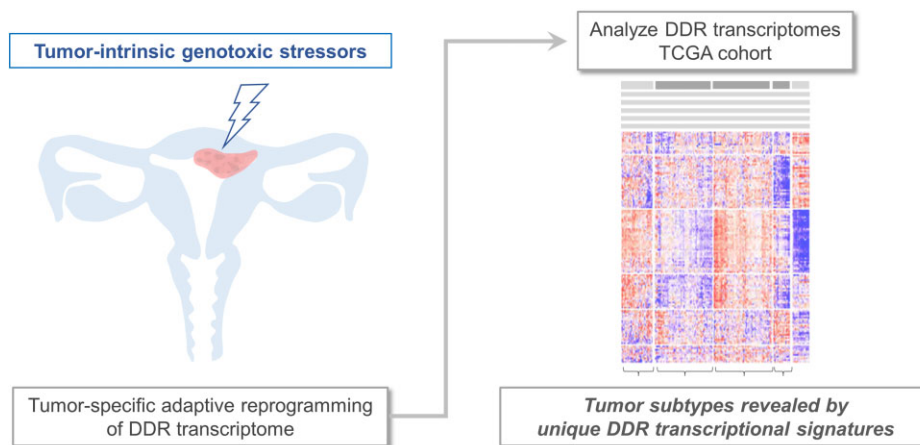
[†]The first two authors should be regarded as Joint First Authors.

[‡]Co-senior authors.

Abstract

Genome maintenance is an enabling characteristic that allows neoplastic cells to tolerate the inherent stresses of tumorigenesis and evade therapy-induced genotoxicity. Neoplastic cells also deploy many mis-expressed germ cell proteins termed Cancer Testes Antigens (CTAs) to promote genome maintenance and survival. Here, we present the first comprehensive characterization of the DNA Damage Response (DDR) and CTA transcriptional landscapes of endometrial cancer in relation to conventional histological and molecular subtypes. We show endometrial serous carcinoma (ESC), an aggressive endometrial cancer subtype, is defined by gene expression signatures comprising members of the Replication Fork Protection Complex (RFPC) and Fanconi Anemia (FA) pathway and CTAs with mitotic functions. DDR and CTA-based profiling also defines a subset of highly aggressive endometrioid endometrial carcinomas (EEC) with poor clinical outcomes that share similar profiles to ESC yet have distinct characteristics based on conventional histological and genomic features. Using an unbiased CRISPR-based genetic screen and a candidate gene approach, we confirm that DDR and CTA genes that constitute the ESC and related EEC gene signatures are required for proliferation and therapy-resistance of cultured endometrial cancer cells. Our study validates the use of DDR and CTA-based tumor classifiers and reveals new vulnerabilities of aggressive endometrial cancer where none currently exist.

Graphical abstract



Introduction

Genome maintenance is an important enabling characteristic of cancer that confers tolerance of intrinsic and therapy-induced DNA damage. Different cancers arise from different precursors, are driven by different oncogenes, and experience different genotoxic exposures. Therefore, each cancer

relies on distinct genome maintenance mechanisms to survive the unique stresses experienced during tumorigenesis. Because genome maintenance mechanisms can be error-free or error-prone, the availability and choice of pathways deployed by cancer cells also shapes the mutational landscape. Elucidating the genome maintenance pathways that sustain each cancer

Received: November 20, 2023. Revised: February 12, 2024. Editorial Decision: March 8, 2024. Accepted: April 3, 2024

© The Author(s) 2024. Published by Oxford University Press on behalf of NAR Cancer.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<http://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

reveals mechanistic underpinnings of tumorigenesis. Moreover, the genome maintenance-dependencies of cancer cells are molecular vulnerabilities and may provide opportunities for therapy.

To define the genome maintenance processes that sustain cancer cells, we recently developed predictive classifiers based upon the relative expression levels of genes encoding core components of the major DNA Damage Response (DDR) and DNA repair pathways (1). In an initial study, our unsupervised analysis of DDR gene expression patterns in breast cancer patients (from both TCGA and North Carolina Breast Cancer cohorts) identified four distinct patient groups and, remarkably, predicted disease subtypes with accuracy (1). Tumor stratification based on DNA repair profiling is a highly innovative approach that reveals underlying mechanisms of carcinogenesis and is highly actionable. The repertoire of DDR pathways available to any tumor has the potential to determine responsiveness to chemotherapy. Therefore, DDR-based classifiers could be used to accurately predict sensitivities to distinct genotoxic therapeutics and inform clinical decisions. By analogy, the PAM50 signature helps guide clinical strategy in breast cancer (2,3).

Here, we have used DDR-based transcriptional profiling to define the genome maintenance landscape of endometrial cancer (EC). Endometrial cancer is the only major solid tumor for which incidence and mortality rates are rising, underscoring the need for improved strategies to predict, prevent, and treat the disease. The DDR landscape of different endometrial cancer subtypes, and the contribution of the DDR to tumorigenic phenotypes in this cancer setting have not been studied.

Endometrial tumors are commonly (~80%) of the endometrioid histotype (designated endometrial endometrioid carcinomas or EEC). There is tremendous heterogeneity in tumors belonging to the EEC subtype. EEC comprise both low grade ('G1' and 'G2'), largely indolent tumors, and high grade ('G3') tumors that are a combination of indolent and highly aggressive tumors. Endometrial Serous Carcinomas (ESC) represent a distinct, highly aggressive, and historically uncommon subtype of endometrial tumors. ESC comprise ~10% of all endometrial tumors yet account for ~40% of all endometrial cancer deaths.(4) The ongoing increase in endometrial cancer mortality is due to a disproportionate rise in ESC (5).

In addition to histology-based subtyping, a comprehensive study of EEC and ESC by The Cancer Genome Atlas (TCGA) revealed four genomic subtypes of endometrial cancer (6). The *POLE* subtype with endometrioid histology is the least common, is 'ultra-mutated', and typically has good prognosis. The Microsatellite Instability (MSI) subtype also contains 'hyper-mutated' tumors of endometrioid histology. The copy-number variation (CNV)-low subtype comprises the majority of low grade endometrioid tumors. The CNV-high subtype comprises the most aggressive endometrioid tumors and all of ESC (6). Subsequent proteogenomic studies have revealed differences in underlying pathways and the immune landscapes of the TCGA subtypes (7).

A defining feature of ESC (and the aggressive CNV-high endometrioid tumors) is their exceptional resistance to DNA-damaging therapies (e.g. carboplatin) (8). Unfortunately, inherent or acquired DNA damage tolerance ('chemoresistance') occurs in 90% of ESC and patients with metastatic or recurrent EEC and is the major cause of death. Current gaps in knowledge of how ESCs and aggressive EECs tolerate chemotherapy limit development of life-saving treat-

ment strategies. To improve patient outcomes, it is critical to elucidate how endometrial cancers tolerate therapy and become chemoresistant. Accordingly, we have sought to identify unique underlying features of aggressive endometrial cancers that account for their chemoresistance.

Our rationale is that a knowledge of the DDR landscape may reveal molecular underpinnings of chemoresistance and provide new opportunities for precision medicine. Our published DDR classifier gene set (1) also included two representative Cancer/Testes Antigens (CTAs) with known roles in DNA repair. CTA genes are normally expressed only in the adult testes. However, many cancer cells aberrantly mis-express CTAs at high levels (9). CTAs often promote DNA repair, thereby conferring cancer cell fitness and tumorigenic phenotypes. For example, in lung cancer cells, Melanoma Antigen A4 (MAGE-A4) associates with the E3 ubiquitin ligase RAD18 and promotes post-replicative DNA repair (10,11). In lung cancer and breast cancer cell lines, *HORMAD1* promotes Homologous Recombination (HR) and radio-resistance (12,13). Interestingly, expression levels of MAGEA4 (10) and *HORMAD1* (12,13) were both highly associated with a specific breast cancer subtype (1). DDR-related functions have also been attributed to many additional CTAs (14). Owing to their absence from all normal somatic cells, CTA that promote DNA repair (and other tumorigenic phenotypes) represent very attractive therapeutic targets. Ablation of CTAs leads to reduced DNA repair capacity and chemo/radio-sensitivity yet is innocuous to non-cancerous cells.

In this study, we have refined our methods for annotating tumors based upon DDR gene expression. Because CTAs provide a critical layer of regulatory input into DNA repair pathways in cancer cells (14), we also interrogated CTA expression profiles in endometrial cancers. Using the TCGA uterine cancer dataset we demonstrate that the different subtypes of endometrial cancer are associated with distinct DDR and CTA expression profiles. Using independent approaches (including unbiased genetic screens) we validate the critical roles of DDR and CTA genes in sustaining growth of endometrial cancer cells. The gene signatures we derived are prognostic and provide new vulnerabilities that can be further harnessed for therapeutic benefit of endometrial cancer patients.

Materials and methods

Statistical analyses and visualization

R (version 4.1.0) was used for visualizing and data analyses. Overall, $P < 0.05$ is considered as statistical significance. Heatmaps were generated using function `Heatmap()` from the R package `ComplexHeatmap` (Version: 2.8.0) (15). Samples were separated as EEC, ESC, and normal endometrium, using the option `column_split`. The standardized \log_2 transformed FPKM-UQ was used as the expression value for genes in the heatmaps. The data in the heatmaps are clustered using hierarchical clustering with Euclidean distance. The DDR pathway heatmaps shown in Figure 2 were derived by calculating the average expression across the genes for each pathways.

Tumor classifiers are used as the labeling in the heatmaps. The clinical information files, somatic variant calling files, and copy number variation files using the GISTIC2 pipeline (citation: <https://doi.org/10.1186/gb-2011-12-4-r41>) were downloaded from the GDC cohort and Xena platform. Histology,

race, height, weight, stage, and tumor grade were obtained directly through clinical data available by the TCGA. Genomic subtypes and genes reported to be frequently mutated or having copy number variations in endometrial cancer were derived through available data from the 2013 TCGA study (6). HRD score was defined as an unweighted sum of three signatures: Number of telomeric Allelic Imbalances (NtAI), Large-scale State Transitions (LST), and loss of heterozygosity (LOH) (16). Replication Stress (RS) was defined as previously described (17). In brief, tumors classified as RS-high included at least one of the following alterations (*CCNE1* amplification, *MYC* and *MYCL1* amplification, *KRAS* amplification, *NF1* mutation, *RB1* two-copy loss, *CDKN2A* two-copy loss, *ERBB2* amplification) whereas tumors without these features were classified as RS-low. CA20 score was defined as the sum of the expression of 20 genes (log₂ median centered). These genes include: *AURKA*, *CCNA2*, *CCND1*, *CCNE2*, *CDK1*, *CEP63*, *CEP152*, *E2F1*, *E2F2*, *LMO4*, *MDM2*, *MYCN*, *NDRG1*, *NEK2*, *PIN1*, *PLK1*, *PLK4*, *SASS6*, *STIL* and *TUBG1* (18).

To find the 6 groups of DDR genes and three groups of CTA genes, the functions of `column_km()` and `row_km()` in the R package `ComplexHeatmap` were used to perform k-means clustering for the DDR genes (indicated as DDR TransSigGroups 1–6; Figure 2A and F, applied in the above standardized log₂ transformed FPKM-UQ) and endometrioid tumors (indicated as Clusters 1–3; Figures 2F and 3A).

Supplementary Figure S2. uses a further step to generate the heatmap. Univariate linear regression is applied to each gene (log₂ transformed FPKM-UQ value as the response variable) with grade as the only covariate to calculate the coefficients of grade for each gene in endometrioid tumors. The residuals are used to draw the heatmap. Samples from the serous tumors and normal endometrial tissue are not included when applying the linear models.

Survival analyses were performed using the following two R packages: `survival` (Version 3.2-13) and `survminer` (Version 0.4.9). Kaplan–Meier curves and log rank test for survival analysis are also performed using R package `survival` and visualized using R package `survminer`.

Cell lines

HEC-1-B, HEC-50, and ARK2 cells were cultured in RPMI-1640 containing 10% fetal bovine serum (Gibco, #26140095) and 1% Penicillin–Streptomycin (Gibco, #15140122). Human Endometrial Epithelial Cells (HEuEC) cells were purchased from (Lifeline Cell Technology, #FC-0078) and cultured in DMEM:F12 medium containing 10% FBS, 1% Penicillin–Streptomycin, and 1% Insulin–transferrin–selenium (Gibco, #41400045). Doxycycline-induced oncogene expressing immortalized-HEuEC (iHEuEC) were cultured in DMEM:F12 media supplemented with 10% tetracycline negative FBS (GeminiBio, #100-800-500) and 1% Penicillin–Streptomycin. All cells were cultured at 37°C in humidified chambers with 5% CO₂.

Immortalization of Human endometrial epithelial cells (HEuEC)

HEuEC were immortalized by overexpressing hTERT upon infection with retrovirus encoding hTERT cDNA. Retrovirus was generated from canine sarcoma D17 packaging cell line stably transduced with the MoMLV-pBABE-hygro-hTERT

vector. Supernatants from this retroviral packaging cell line were collected when cells were 70–80% confluent. Supernatants were centrifuged at 500 × g for 5 min and then filtered through 0.45 μm nylon filter (Corning, #431225). HEuEC cells were seeded in 6-well plates, infected at 50% confluency with 2 ml/well of hTERT-encoding retrovirus containing medium with 8 μg/ml polybrene, and allowed to incubate with the retrovirus for 48 hours. Successfully transduced cells were selected with 100 μg/ml Hygromycin B (Corning, #30240CR) for 14 days. Cells obtained were immortalized and labelled as iHEuEC cells.

Plasmids and RNA interference

siRNA oligonucleotides were used for targeting specific genes transfected into cells using Lipofectamine 2000 as per manufacturer's instructions. The following siRNAs were purchased from Horizon discovery: siKIF2C#1 (#L-004955-00-0005), siNUF2 (#L-006893-01-0005), siCEP55 (#L-005289-00-0005), siKIF2C#2 (#J-004955-06-0005), siKIF2C#3 (#J-004955-08-0005), siRNF168 (#M-007152-03-0050), siMRE11A (#M-009271-01-0050), siUSP7 (#L-006097-00-0005), siPRKDC (#L-005030-00-0005), siControl (#D-001810-10-05). The lentivirus doxycycline-inducible expression system using the pInducer20 vector was purchased from Addgene (#44012). pInducer20 constructs were cloned with *KRASG12V-HA* cDNA insert by HiFi Assembly (NEB, #E2621S). pInducer20-*CCNE1* construct was cloned using Gateway cloning (Invitrogen, #11791020). pInducer20-*cMyc-HA* vector was constructed by VectorBuilder (Vector ID: VB200317-7039asg).

Generating doxycycline inducible oncogene expressing cells

Low-passage HEK 293T cells were seeded in 10 mm dishes to 60% confluency. At 24 h. post seeding, cells were transfected using Lipofectamine 2000 (Invitrogen, #11668019). Each dish was transfected with 6 μg lentiviral packaging plasmids (4.5 μg dNRF and 1.5 μg pMDK64) and 6 μg pInducer20-plasmid encoding oncogene cDNA (*c-Myc-HA*, *KRAS-HA*, *CCNE1*, or pInducer20 backbone). Plasmid DNA mix was diluted in 350 μl serum-free Opti-MEM media (Gibco, #31985070). In a separate reaction mix, 35 μl Lipofectamine 2000 was diluted in 350 μl of Opti-MEM media. Both DNA mix and lipofectamine solution were mixed gently and incubated for 15 mins. at room temperature. The DNA-lipofectamine mix was then added to cultured HEK293T cells which were incubated overnight. Approximately 16 h. post transfection, transfection medium was removed and replenished with 10 ml fresh medium. Lentivirus containing cell culture supernatant was collected at 24 h. and again at 48 h. Collected medium was centrifuged at 500 × g for 5 min to remove cellular debris and the supernatant was filtered through a 0.45 μm nylon syringe filter. Filtered lentivirus containing medium was stored at –80°C. To generate doxycycline-inducible oncogene expressing cell line iHEuEC, cells were seeded in 6-well plates. At 50–60% confluency, cells were incubated with 2 ml lentivirus-containing media mixed with 0.8 μg/ml polybrene for 24 h. After lentiviral infection, cells were selected with Geneticin (Gibco, #10131035) at 1000 μg/ml for 5 days. pInducer20 backbone without oncogene cDNA insert was used for generating empty vector controls.

Immunoblots

Chromatin protein extracts were isolated from cultured cells using CSK lysis buffer (CSK buffer; 10 mM Pipes, pH 6.8, 100 mM NaCl, 300 mM sucrose, 3 mM MgCl₂, 1 mM EGTA, 10 mM NaF, 25 mM β-glycerophosphate and 0.1% Triton X-100) containing fresh protease inhibitor cocktail (Roche, #4693159001) and phosSTOP (Roche, #4906837001). Harvested cell lysates in CSK buffer were incubated on ice for 15 min, followed by centrifugation at 4000 rpm for 4 min. The supernatant was collected as soluble cell fraction and the pellet was washed with 1 ml CSK buffer at 4000 rpm for 4 min. The remaining pellet was re-suspended in minimal volume of CSK buffer which is sonicated, followed by nuclease treatment, and collected as crude chromatin fraction. For whole-cell protein extracts, cultured cells were washed with cold 1× PBS and cold CSK lysis buffer added to cultured monolayers. Cells were collected in lysis buffer by scraping, lysates generated by sonication and then centrifugation at 20 000 × g for 15 mins, and supernatants collected. Proteins extracts were resolved on Tris-glycine gels (Invitrogen, #XP04205BOX) by electrophoresis and transferred to a nitrocellulose membrane. Membranes were incubated for 1 h in 1× Tris-buffered saline containing 0.1% Tween 20 (TBS-T) and 5% non-fat milk (blocking buffer) and then incubated with primary antibodies (dilutions ranging from 1:500 to 1:2000), prepared in blocking buffer, at 4°C overnight. Membranes were washed with 1× TBS-T and then incubated with HRP-conjugated secondary antibody diluted in 1× TBS-T at 1:5000 for 1 h. Chemiluminescent signals were captured using Perkin Elmer Western Lightning Plus ECL reagent and autoradiography film. Primary antibodies used included: γH2AX S139 (Millipore, #05-636); KIF2C (Abcam, #187652); phospho-RPA32-S4/S8 (Bethyl, #A300245A), RPA34 (Millipore Sigma, #NA18) CEP55 (Santa Cruz, #sc-374051); NUF2 (Abcam, #ab180945); HA-tag (Santa Cruz, #sc-7392); c-Myc (Santa Cruz, #sc-40); CCNE1 (sc-248); Vinculin (Sigma, #V9131); GAPDH (Santa Cruz, #32233); phospho-CHEK2-T68 (CST, #2661L); CHK2 (CST, #2662); PCNA (Santa Cruz, #sc-56).

BrdU labeling and flow cytometry

Cultured cells were incubated with 10 mM BrdU for 1 h before trypsinization. Trypsinized cells were washed with 1× PBS and re-suspended in 65% DMEM with 35% ethanol for fixation overnight at 4°C. Fixed cells were denatured using HCl and then neutralized with 0.1 M sodium borate (pH 8.5). Cells were stained with a FITC conjugated anti-BrdU antibody for 1 h. Following BrdU staining, cells were incubated in 1× PBS containing 10 μg/ml of propidium iodide (PI) and 8 μg/ml of RNaseA overnight at 4°C. Stained cells were analyzed for replication dynamics (BrdU staining) and cell cycle phase (PI staining) using an Accuri C6 plus flow cytometer (BD) and analyzed using the FlowJo software.

Clonogenic growth assay

Cells were reverse transfected with siRNA using Lipofectamine 2000 overnight in 6 mm dishes. After transfection, the media on the cells was changed to fresh media. After 6 h, cells were seeded in 6-well plates at a seeding density of 2000 cells per well. Twenty-four hrs. after seeding, cells were either treated with MK1775 or DMSO for 6–7 days of clonogenic growth. Medium changes occurred every 72 h. After 6–7 days,

culture plates were stained with 0.05% crystal violet in 1× PBS containing 1% methanol and 1% formaldehyde. Plates were scanned and colony area measurements calculated using ImageJv2 software.

Immunofluorescence

HEC-1-B cells transfected with siControl or siKIF2C#3 were plated into 6-well plates with glass coverslips. Twenty-four hrs. after seeding, HEC-1-B cells were treated with thymidine 2 mM overnight and then released from the thymidine block into treatment with MK1775 200 nM for 48 h. The cells were then washed with 1× PBS and fixed using 4% PFA for 15 min at room temperature. Fixed cells were incubated in blocking solution (DPBS with 3% BSA and 5% donkey serum) for 1 h followed by incubation with primary antibodies (diluted 1:200 in blocking solution) overnight at 4°C. Cells were washed 3× with 1× DPBS for 5 mins. For detection of the primary antibodies, cells were incubated 1 h with secondary antibodies (Alexa Fluor 488 donkey, and Alexa Fluor 647 Goat anti-mouse, 1:400 dilution). Secondary antibodies were removed by washing 3× with 1× DPBS for 5 min and coverslips mounted to microscope slides using Prolong Gold mounting medium with DAPI (Invitrogen, #P36941). Antibodies: γH2A.X (Millipore, #05-636) and Pericentrin (Abcam, #ab4448).

CRISPR-Cas9 based DNA damage repair (DDR) gene targeting library preparation

The pooled CRISPR-Cas9 library includes 5040 sgRNAs targeting 504 DNA Damage Repair genes (each gene has 10 independent targeting sgRNAs) and 1000 non-targeting control sgRNAs. The oligonucleotide pool encoding sgRNAs in the library was synthesized by CustomArray Inc. The DDR-CRISPR library was PCR amplified with Q5 high-fidelity polymerase. The amplified and PCR purified DDR-CRISPR library was cloned into LentiCRISPRv2-DD-Cas9-BSD by NEBuilder HiFi DNA assembly. Endura Electrocompetent cells (LGC Biosearch Technologies, #602422) were transformed with the DDR-CRISPR library by electroporation. Transformed bacteria were plated on twelve 15 cm LB ampicillin plates and incubated overnight at 37°C. LB media was added to each plate and the transformed bacterial colonies were recovered using sterile scrapers. Pooled bacterial colonies from three plates were transferred to 500 ml LB and grown with shaking at 37°C for 3 h. The DDR-CRISPR library plasmid DNA was purified from transformed bacterial culture using a Qiagen EndoFree plasmid DNA Maxi kit (Qiagen, #12362).

DDR targeting CRISPR-Cas9 screen in HEC-50 cells

HEK293T cells were transfected with DDR-CRISPR-Cas9 pooled library plasmid using Lipofectamine 2000. Lentivirus containing media supernatant was harvested from transfected cells at 24 and 48 h after transfection. Lentivirus containing media was centrifuged at 500 × g for 5 min and filtered through 0.45 μm nylon filters. Aliquots were frozen at –80°C. HEC-50 cells were transduced with DDR-CRISPR-Cas9 library lentivirus at an MOI of ~0.3. Lentivirus infected cells were treated with Blasticidin (10 μg/ml) for 4–5 days to select stably transduced cells. After selection, cells were treated with Shield-1 (1 μM) for 4 days. 3 × 10⁶ cells were harvested at this timepoint (population doubling 0 (PD0)) for isolation of

genomic DNA and for using as library controls after sgRNA enrichment. Another 3×10^6 cells were seeded in three 150 mm cell culture dishes for subsequent passages. Cells were passaged every 3–4 days. At the final passage (population doubling 20 (PD20)), 3×10^6 cells were harvested for genomic DNA isolation and sgRNA enrichment. Assuming a harvest of 6 μ g genomic DNA content from 1×10^6 cells, library preparation was performed with 18 μ g genomic DNA, thereby preserving 500 fold representation of each sgRNA. sgRNA were then PCR amplified in two steps: (i) sgRNA was PCR amplified using NEB Q5 Hot Start polymerase (NEB, #M0494L). The product was purified using AMPure XP beads (Beckman Coulter, Cat.no: A63881). (ii) A PCR step was then used to add indexed sequencing adapters. Each secondary PCR reaction contained 1 ng of purified PCR product from the first PCR step as template in a 100 μ l reaction volume containing indexing primers. After purification using AMPure XP beads, the 320 bp products were quantified using Qubit and diluted to 1 ng/ml and sequenced using HiSeq 4000, PE 2 \times 150 (Novogene). Analysis of sgRNA enrichment was performed using Volundr pipeline.

Results

DDR gene expression profiling reveals distinct endometrial cancer subsets.

We hypothesized that endometrial cancer subtypes each experience unique genotoxic stresses (from different oncogenes, inherent DNA repair defects, supernumerary centrosomes, or other tumor-intrinsic factors), leading to distinct adaptive programs of DDR gene expression. We tested elements of this hypothesis using the experimental work-flow summarized in Figure 1. We curated a list of 131 genes encoding core components of 13 major DNA repair and genome maintenance pathways (Supplemental Table S1). Using the TCGA uterine cancer dataset we generated heatmaps depicting relative expression of these DDR genes in ESC and EEC. We determined relative expression levels of DDR genes in relation to multiple classifiers including demographic features (race and BMI), tumor grade, disease stage, and genome maintenance-related features (e.g. Replication Stress or RS signature; Homologous Recombination-Deficiency or HRD score; Centrosome Amplification or CA20 signature).

Histology, tumor grade, and disease stage are important independent prognostic factors for endometrial cancer (19). ESCs are often metastatic at time of diagnosis and have poor outcomes compared to more indolent endometrial tumors (G1 and G2 EEC). G3 EECs are heterogeneous in that these tumors can be either indolent or aggressive and share similar molecular / genetic features with G1 and G2 EEC or ESC (e.g. *TP53* mutation). Race and BMI are also predictive factors. Lower BMI has been associated with higher disease stage and features of ESC (e.g. high HER2 expression) (20), and non-Hispanic black women are more likely to be diagnosed with ESC and die of the disease compared to white women [<https://seer.cancer.gov/statfacts/html/corp.html>]. Oncogene-induced DNA replication stress is a major source of intrinsic DNA damage in cancer cells (21), which likely necessitates expression of DDR genes for stress tolerance. Different oncogenes induce different types of DNA damage and may create dependencies on distinct DDR pathways for damage tolerance (22–24). Therefore, in our clas-

sifiers we included various individual oncogenes/tumor suppressors that are frequently mutated or amplified in ESC and EEC (Figure 2A). Homologous Recombination Deficiency (HRD) is common in many tumors and creates new dependencies on alternative DDR mechanisms (25). Therefore, we considered that HR status may impact the expression and deployment of other DDR mediators in endometrial cancer. Because aberrant centrosome number is a potential source of mitotic errors and DNA damage in many tumors (26), we also considered the Centrosome Amplification 20 (CA20) gene signature (18,27) as a feature that might influence the DDR.

We performed supervised clustering and examined DDR gene expression patterns in relation to histological features of the tumors using the 2013 TCGA sample set for uterine cancer ($n = 241$) (6). Based on their patterns of DDR gene expression, EEC and ESC tumors were both very distinct from normal tissue (Figure 2A). ESCs are among the most chemoresistant and clinically challenging endometrial tumors. Therefore, it was of particular interest to define DDR features of ESC. Using K-means clustering, we identified 6 Transcriptional Signature Groups ('TransSigGroups') of DDR genes, of which 'DDR TransSigGroup 2' and 'DDR TransSigGroup 5' were expressed at particularly high levels in ESC and a subset of EEC when compared with normal endometrial tissue (Figure 2A). DDR genes and associated pathways represented in DDR TransSigGroup 5 and DDR TransSigGroup 2 are listed in Table 1. We validated the differential expression patterns of DDR TransSigGroup 5 and DDR TransSigGroup 2 genes using a secondary cohort that included all serous and endometrioid tumors in the TCGA database for uterine cancer ($n = 521$) (Supplemental Figure S1). The genes in DDR TransSigGroup 2 and DDR TransSigGroup 5 encode subunits of Ribonucleotide Reductase (RNR) and multiple genes involved in DNA replication fork progression. Most strikingly, DDR TransSigGroup 5 and DDR TransSigGroup 2 gene lists are characterized by transcripts encoding numerous mediators of the Fanconi Anemia (FA) and Homologous Recombination (HR) pathways (Table 1).

For both serous and endometrioid tumors, expression of DDR TransSigGroup 5 genes associates with higher tumor grade (Figure 2B), high HRD and CA20 scores (Figure 2C, D), and high RS (Figure 2E). In contrast with ESC, which showed a homogenous expression pattern for DDR TransSigGroup 5 and TransSigGroup 2 genes, expression of DDR TransSigGroup 5 and TransSigGroup 2 genes was inherently heterogeneous for EEC (Figure 2A and Supplemental Figure S1). Therefore, we used K-means clustering to further subcategorize endometrioid tumors into three distinct 'Clusters' based on their DDR gene expression patterns (Figure 2F). The resulting Cluster 1 and Cluster 2 tumors expressed DDR TransSigGroup 5 genes at low levels and contained mostly CN-low and grade 1 (G1) and grade 2 (G2) tumors (Figure 2F). Tumors in Cluster 3 had DDR gene expression patterns strongly resembling those of ESC (Figure 2F). Of the 37 DDR genes of DDR TransSigGroup 5 that characterized ESC, all were enriched in Cluster 3 endometrioid tumors as validated by Wilcoxon tests. Similar to ESC, high expression of DDR TransSigGroup 5 genes in Cluster 3 endometrioid tumors was not associated with specific oncogenes or tumor suppressors. Also similar to ESC, Cluster 3 endometrioid tumors showed increased enrichment for DDR TransSigGroup 5 and TransSigGroup 2 genes compared to normal endometrium.

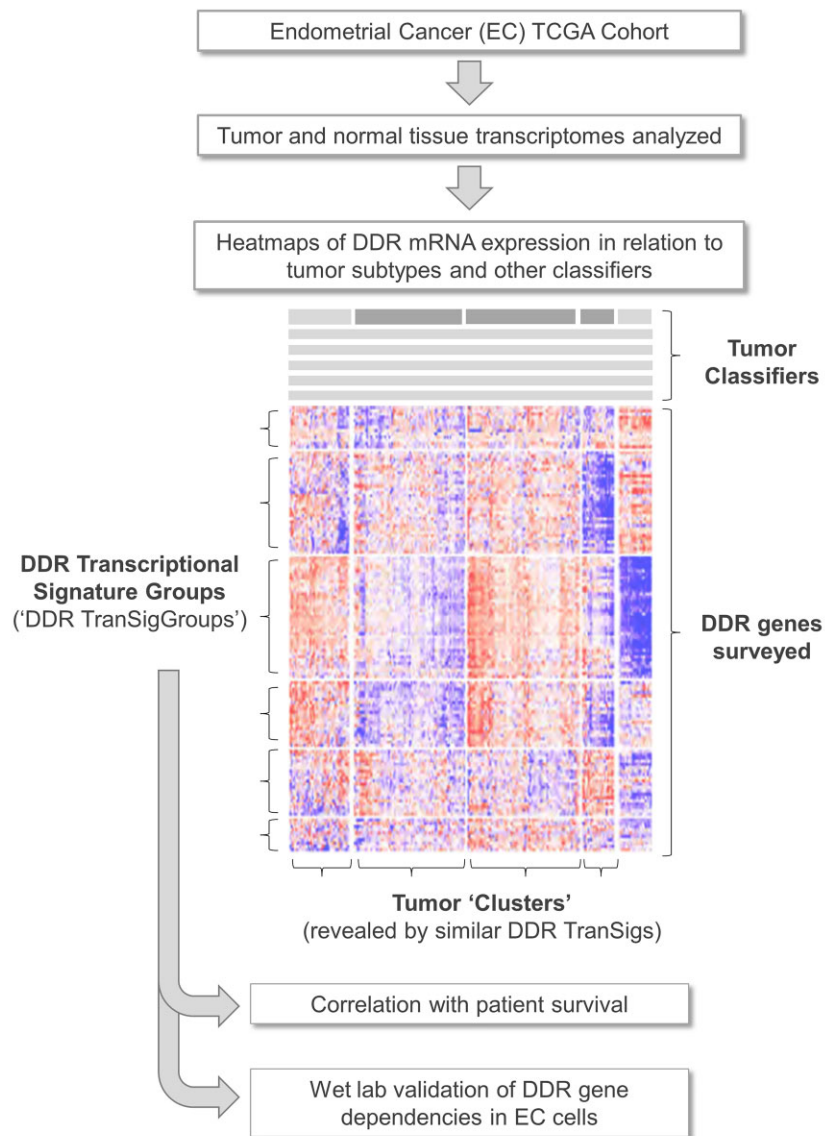


Figure 1. Summary of experimental design for establishing DDR gene expression patterns in different endometrial cancer subtypes. Using the TCGA uterine cancer dataset, heatmaps were generated depicting relative expression levels of 131 DDR genes (listed in [Supplemental Table S1](#), encoding core components of 13 major DNA repair pathways) in ESC, EEC and normal tissues. Relative expression levels of DDR genes were examined in relation to multiple classifiers including demographic features (race and BMI), tumor grade, disease stage, oncogenes and tumor suppressors that are frequently mutated or amplified in ESC and EEC, and genome maintenance-related features (including RS signature; HRD score; and CA20 signature). K-means clustering was used to identify DDR Transcriptional Signature Groups (TransSigGroups) associated with different tumor clusters. Expression levels of individual DDR genes and DDR TransSigGroups were correlated with clinical outcomes. Independent wet-lab studies were performed to establish dependencies of cultured endometrial cancer cells on individual DDR genes.

Distinct endometrial cancer subsets based on DDR gene expression profiles are not driven solely by tumor grade effects

We considered the possibility that the high expression of DDR TransSigGroup 5 and TransSigGroup 2 genes associated with Cluster 3 endometrioid tumors was driven by tumor grade. However, when we removed the grade effect for each gene using linear models and re-analyzed DDR gene expression patterns, we still detected 3 main Clusters of endometrioid tumors ([Supplementary Figure S2](#)). DDR TransSigGroup 5 and TransSigGroup 2 genes remained enriched in Cluster 3 tumors when tumor grade effect was removed (Figure 2G). Approximately, 78% of Cluster 3 endometrioid tumors identified in Figure 2F (without removal of tumor grade effect) remained

in Cluster 3 after we removed the tumor grade effect (Figure 2G and [Supplemental Figure S2](#)). Similarly, most of the tumors remained in Clusters 1 and 2 after removing the tumor grade effect (Figure 2G). Therefore, tumor grade alone cannot explain the separation of endometrioid tumors into three distinct Clusters based on DDR gene expression profiles. Expression of DDR TransSigGroup 5 and TransSigGroup 2 genes did not correlate with any particular oncogene or tumor suppressor mutations in EEC. HRD scores were not significantly different between the three endometrioid tumor Clusters (data not shown). However, Cluster 3 tumors showed higher RS and CA20 scores compared to Cluster 1 and Cluster 2 (RS score: Pearson's chi-square statistic X is 4.49 for 1 degree of freedom, and P -value = 0.03; CA20 score: Figure 2H). Therefore,

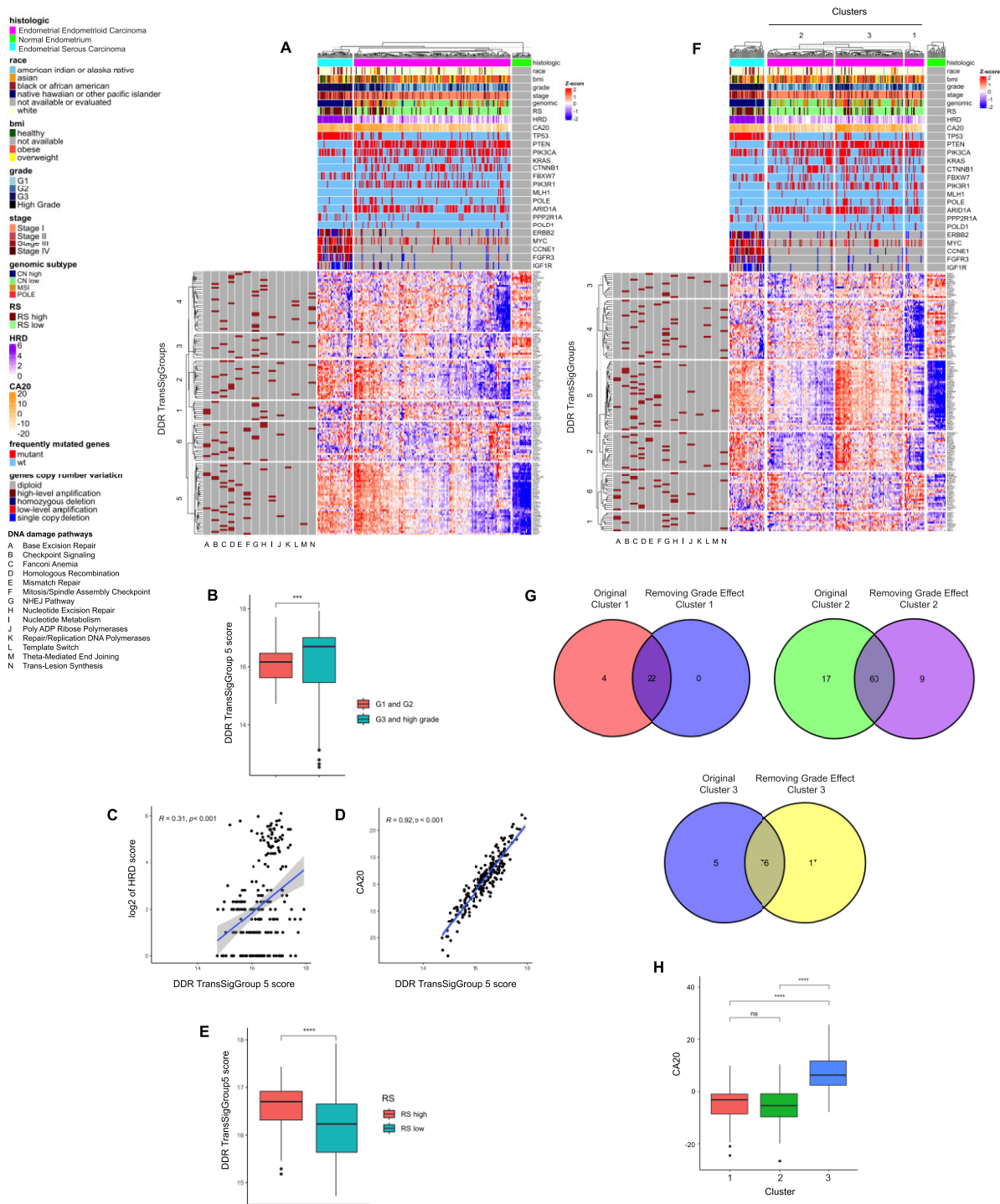


Figure 2. Expression of DDR genes in endometrial cancer with supervised clustering based on histological subtypes. **(A)** Genomic and clinical data from the 2013 TCGA sample set for uterine cancer were downloaded from TCGA portal and Xena Browser. Samples were clustered by endometrial tumor histology (serous or endometrioid) or as normal endometrium. Heatmaps are depicting relative mRNA expression of DDR genes assembled from a curated list of 131 genes that encode core components of 13 major DDR and genome maintenance pathways. Based on DDR gene expression patterns, K-means clustering ($K = 6$) identified 6 transcript gene signature groups ('DDR TransSigGroups 1–6'). DDR TransSigGroup 5 is highly expressed in serous endometrial tumors compared to normal endometrial tissue and contains 37 genes of different DDR pathways. DDR pathways and related genes are highlighted on the far left. Clinical, molecular, and genomic classifiers relevant to endometrial cancer and genome maintenance features (HRD and CA20 scores) are depicted. Correlation of DDR TransSigGroup 5 genes of endometrioid serous tumors to tumor grade (B) and HRD (C) and CA20 (D) scores and replication stress (RS). DDR TransSigGroup 5 gene score is the average expression of the genes in the cluster. **(B)** Wilcoxon's rank sum test was applied to compare DDR TransSigGroup 5 score with tumor grade ('G1', 'G2' as low grade, 'G3' and high grade). **(C–E)** Spearman's rank correlation. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. **(F)** K-means clustering separating endometrioid tumors into three sub-clusters based on DDR gene expression profiles. **(G)** Venn diagrams depicting similar composition of the three sub-clusters for endometrioid tumors before (A) and after (Supplemental Figure S2) removing tumor grade effect. **(H)** CA20 scores between the 3 sub-clusters for endometrioid tumors. (H) Wilcoxon's rank sum test, *** $P < 0.0001$.

Table 1. Group 5 and Group 2 DDR genes

Group 5 genes	DDR pathway(s)
BLM	HR
BRCA1	HR
BRCA2	HR
BRIP1	HR, FA
BUB1	Mitosis/SAC
BUB1B	Mitosis/SAC
BUB3	Mitosis/SAC
CDH1	Mitosis/SAC
CHEK1	CS
CHEK2	CS
CLSPN	CS
DNA2	HR
ESPL1	Mitosis/SAC
EXO1	MMR
FANCA	FA
FANCB	FA
FANCC	FA
FANCD2	FA
FANCG	FA
FANCI	FA
FEN1	BER
LIG1	NER
MAD2L1	Mitosis/SAC
NEIL3	BER
PARPBP	Poly ADP ribose polymerases
PCNA	BER
POLE	Repair/replication DNA polymerases
POLQ	TMEJ
PTTG1	Mitosis/SA
RAD51	HR
RAD54L	HR
RRM2	NM
TIMELESS	CS
TIPIN	CS
TK1	NM
UBE2T	FA
XRCC2	HR
Group 2 genes	DDR pathway(s)
BARD1	HR
FANCE	FA
FBXO5	HR
HLTF	TS
MDC1	CS
MSH2	MMR
MSH6	MMR
PALB2	FA
PARP2	Poly ADP ribose polymerases
POLH	TL
PRKDC	NHEJ
RAD1	CS
RAD18	TL
RAD51C	FA
RNASEH1	NM
RNF168	NHEJ
RRM1	NM
TOPBP1	HR
USP1	FA
XRCC5	NHEJ

BER, Base Excision Repair; CS, Checkpoint Signaling; FA, Fanconi Anemia; HR, Homologous Recombination; MMR, Mismatch Repair, Mitosis/SAC, Mitosis/Spindle Assembly Checkpoint; NHEJ, Non-Homologous End Joining; NER, Nucleotide Excision Repair; NM, Nucleotide Metabolism; TS, Template Switch; TMEJ, Theta-Mediated End Joining; TL, Trans-Lesion Synthesis.

DDR TransSigGroup 5 and TransSigGroup 2 genes enriched in endometrioid tumors (Cluster 3) might allow tolerance of genotoxicities resulting from DNA replication stress and supernumerary centrosomes.

DDR gene expression profiles reveal distinct endometrial cancer subsets unrelated to molecular subtypes

Next, we asked whether DDR TransSigGroup 5 and TransSigGroup 2 genes showed any relationship with the endometrial cancer molecular subtypes identified by proteogenomic characterization (7). Therefore, we performed supervised clustering of both ESC and EEC based on the four TCGA genomic subtypes (namely CN-high, CN-low, MSI, and POLE-mutated) instead of histological subtypes. As shown in [Supplemental Figure S3](#), DDR TransSigGroup 5 genes corresponded to the CNV-high subtype (a category which contains nearly all ESC). A subset of the POLE-mutated tumors and a subset of MSI tumors (two categories which contain endometrioid tumors) also were enriched for DDR TransSigGroup 5 genes ([Supplemental Figure S3](#)). Of the various classifiers we used, high tumor grade and CA20 scores were most closely associated with expression of DDR TransSigGroup 5 genes in CNV-high, MSI, and POLE. The estimated Spearman Rho for CA20 score is 0.38 with a P -value < 0.001 . Thus, genomic subtypes, even those related to DNA damage (MSI and POLE-mutated), cannot explain the DDR gene expression profiles of endometrioid tumors. We conclude that DDR gene expression profiles reveal distinct endometrial cancer subsets.

ESC and Cluster 3 endometrioid tumors are defined by multiple DNA repair pathways and high CA20 scores

Most DNA repair processes rely on the coordinated actions of multiple DNA repair proteins that function in a common pathway. To complement our analysis of individual DNA repair genes in ESC and EEC, we performed pathway-based analyses to determine the overall status of different dedicated DNA repair processes. As shown in [Figure 3A](#) our pathway-based analysis of DDR gene expression also revealed marked differences between ESC, EEC Clusters 1, 2 and 3, and normal endometrial tissues. In ESC and Cluster 3 EEC, we observed increases in expression of key DNA repair genes corresponding to multiple pathways. Thus, TMEJ, checkpoint signaling, FA, HR, BER and MMR pathways were generally elevated in ESC and Cluster 3 EEC when compared with normal endometrium or Clusters 1 and 2 EEC ([Figure 3A](#)).

Many forms of DNA damage can be repaired by error-free or error-prone pathways. Therefore, DNA repair pathway choice and availability can impact genome stability. We tested correlations between DDR gene expression patterns at the pathway level and two specific forms of genomic instability: copy number alterations (CNA) and single-nucleotide variants (SNV). As shown in [Figure 3B](#) and [C](#), distinct DDR pathways were associated with SNV and CNA in ESC and EEC. For example, EEC with higher CNA generally had higher expression of DDR genes involved in TMEJ, nucleotide metabolism, and checkpoint signaling. SNV-high EEC showed enrichment of DDR genes that mediate BER, checkpoint signaling, the FA pathway, HR, mitosis/spindle assembly checkpoint regulators, NER, nucleotide metabolism, poly ADP ribose polymerases, and TMEJ. Conversely, ESC with CNA-high status

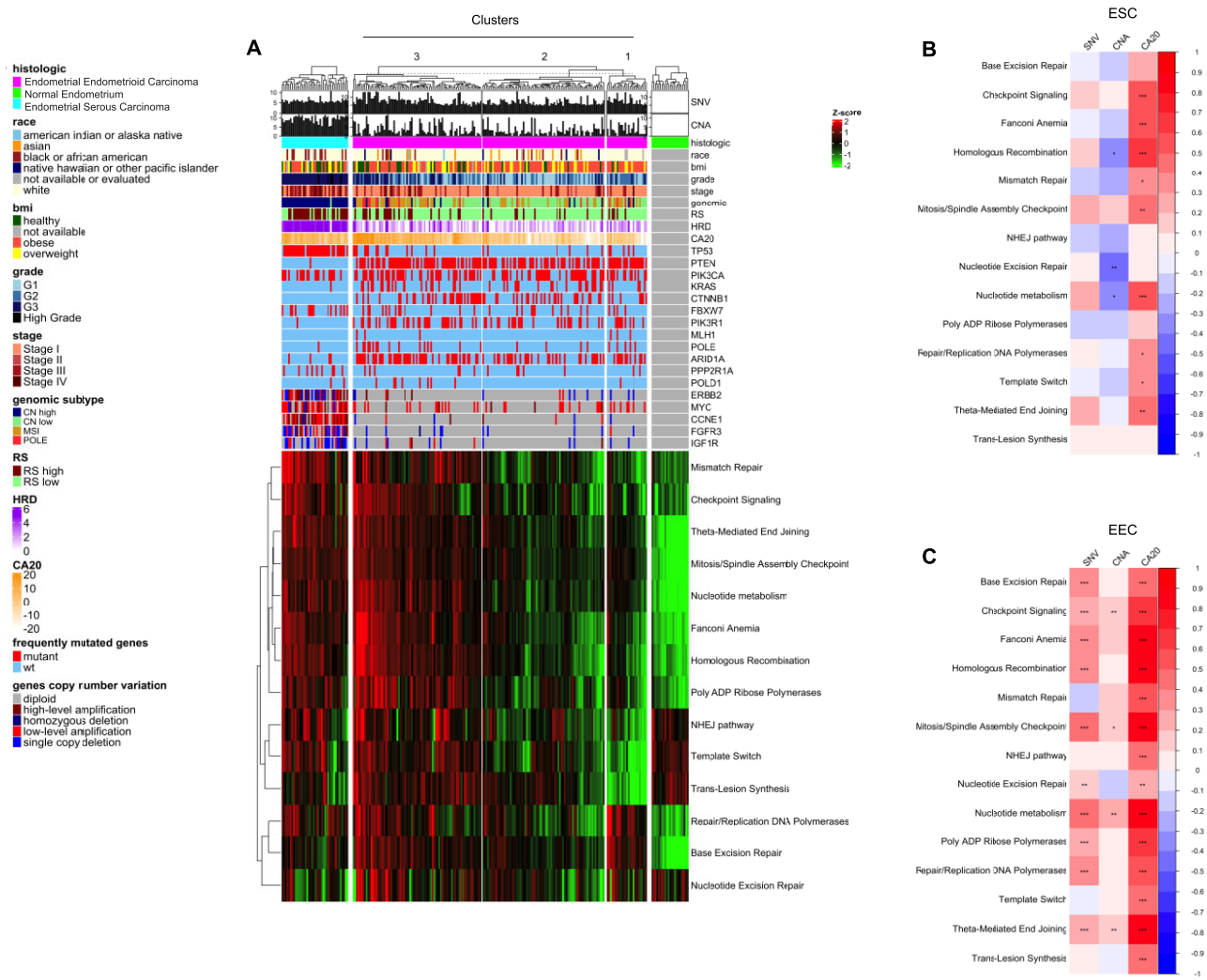


Figure 3. Comparison of serous and endometrioid endometrial tumors based on DDR pathway level analysis and SNV and CNA aberrations. (A) Supervised clustering of DDR pathway Z-scores by histology. K-means clustering of endometrioid tumors described in Figure 2 was applied to endometrioid tumors after removing tumor grade effect. Spearman correlation matrix comparing DDR pathway Z-scores to CA20 score, total SNV counts, or CNV counts in (B) ESC and (C) EEC. (B) Spearman's rank correlation. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

showed reduced expression of genes involved in HR, NER, and nucleotide metabolism. We observed no correlations between DDR pathway gene expression and SNV-high ESC. For ESC and EEC, CA20 scores were positively correlated with expression of genes corresponding to multiple DDR pathways (Figure 3B-C). Notably, none of the 37 DDR TransSigGroup 5 and 20 DDR TransSigGroup 2 genes overlap with CA20 signature genes. Thus, ESC and Cluster 3 endometrioid tumors may have a high dependency on these DDR pathways.

DDR TransSigGroup 5 and 2 gene signatures predict poor patient outcomes in endometrioid endometrial carcinomas

We asked whether our DDR TransSigGroup 5 or TransSigGroup 2 gene signatures were predictive of patient survival and/or disease recurrence. Because ESC and EEC are clinically different uterine tumors, we assessed overall survival and progression-free survival for each histological subtype. For ESC, high expression of DDR TransSigGroup 5 and TransSigGroup 2 genes were not predictive of overall survival or progression-free survival (Figure 4A and C). These

data are fully expected because ESC tumors are homogeneous and all express similarly high levels of DDR TransSigGroup 5 and TransSigGroup 2 genes (Figure 2A). In contrast with ESC, EEC are heterogeneous with respect to expression of DDR TransSigGroup 5 and TransSigGroup 2 genes (Figure 2A, Supplementary Figure S1). Given that DDR TransSigGroup 5 and TransSigGroup 2 genes were enriched in approximately half of *POLE*-mutant EEC (Supplemental Figure S3), *POLE*-mutants were removed in our Kaplan–Meier analyses for EEC. *POLE* mutations occur ~9% of endometrial tumors and have excellent clinical outcomes even when tumors have high-risk features (e.g. high grade, lymphovascular space invasion, are an aggressive non-EEC histotype) (28,29). Recent evidence suggests that *POLE* mutations override the biological relevance of other genetic/molecular alterations (e.g. p53) of the tumors (30,31). Interestingly, high expression of DDR TransSigGroup 5 and TransSigGroup 2 gene signatures in *POLE*-proficient EEC were predictive of poor patient outcomes for EEC. High expression of DDR TransSigGroup 5 or TransSigGroup 2 genes correlated with worse overall survival and progression-free survival of EEC patients (Figure 4B and C). Expression of several individual genes from DDR

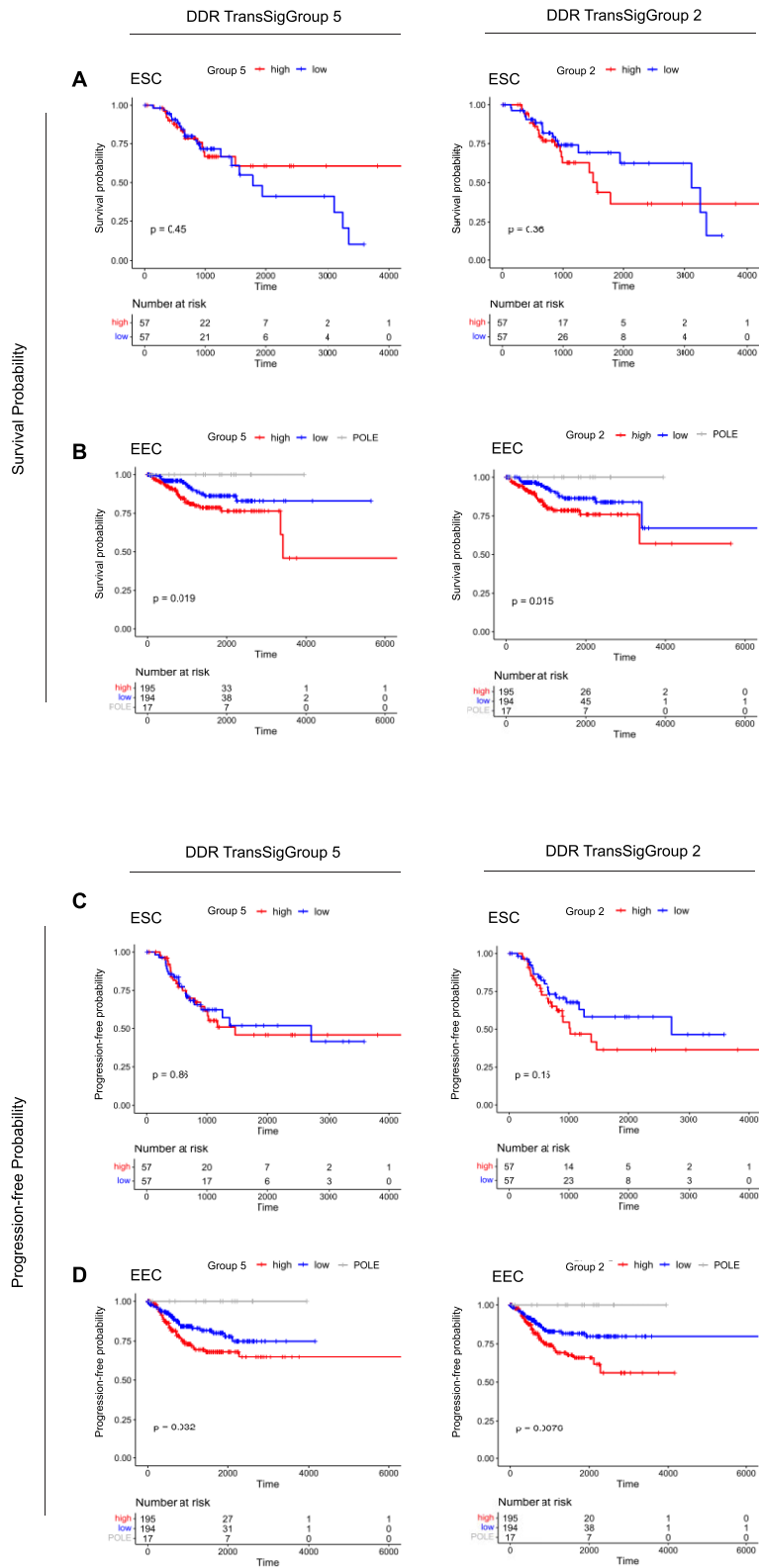


Figure 4. DDR TransSigGroup 5 and TransSigGroup 2 predict poor outcomes in endometrioid endometrial carcinomas. Kaplan-Meier curves for overall survival (A, B) and progression free survival (C, D) of DDR TransSigGroup 5 and DDR TransSigGroup 2 in ESC or EEC. Patient data used are from the entire uterine cancer cohort from the TCGA (similar to Figure 2B). Patient cases ESC and non-POLE mutant EEC were divided into two groups ('high' and 'low') based on median DDR TransSigGroup 5 or DDR TransSigGroup 2 DDR gene expression. Log-rank tests were used to infer statistical significance between high and low groups.

TransSigGroup 5 and TransSigGroup 2 was strongly associated with poor overall survival (18 genes) and/or worse progression-free survival (15 genes) (Supplemental Figures S4 and S5). These results demonstrate that DDR TransSigGroup 5 and TransSigGroup 2 identify aggressive endometrial cancers, regardless of being CN high, CN low, or MSI subtypes.

Endometrial cancer-defined DDR gene signatures and POLQ association are not generalizable to all HRD-high tumors

HR pathway deficiency is associated with several cancer types, including high grade ovarian serous carcinoma (HGOSC) and pancreatic adenocarcinoma (PDAC). Figure 2A shows that HRD scores are uniformly high in ESC compared to low grade EEC. Based on histologic and genetic/molecular similarities and clinical behavior ESC is considered to be clinically equivalent to HGOSC (6,32). Therefore, we determined whether DDR TransSigGroup 5 and TransSigGroup 2 gene signatures associated with ESC were also hallmarks of HGOSC and other HRD-high tumors (such as PDAC). Interestingly, ESC and a large subset of HGOSC shared similar enrichment Z-scores for DDR TransSigGroup 5 genes (Supplementary Figure S6A) and several DNA repair pathways (Supplementary Figure S6B). In contrast, DDR TransSigGroup 5 genes were expressed at low levels in PDAC (Supplementary Figure S6A). Furthermore, expression patterns of all DDR genes in PDAC were distinct from those in ESC and HGOSC (Supplementary Figure S6B). We conclude that expression of DDR TransSigGroup 5 and TransSigGroup 2 genes is not a byproduct of HR-deficiency, but instead is a specific characteristic of ESC and HGOSC.

Reportedly, in a subset of HGOSC with HRD features there is compensatory upregulation of DNA Polymerase theta (POLQ), the rate-limiting enzyme in the Theta-Mediated End-Joining (TMEJ) DSB repair pathway (33). Because POLQ is error-prone, TMEJ generates a microhomology deletion (MHD) mutational scar that is represented in COSMIC mutational signature 3 (34). Given the clinical resemblance of ESC and HGOSC, we investigated links between HRD status, POLQ expression, and mutation signature 3 in endometrial cancer. We used the method of Marquard *et al.* (35) to assess HRD status in high and low grade EEC, ESC, HGOSC and PDAC. As shown in Supplementary Figure S6C, ESC, HGOSC, and PDAC had relatively high HRD scores when compared with high and low grade EEC. In contrast with previous reports, we did not observe a correlation between POLQ expression and HRD status in HGOSC (Supplementary Figure S6D). Similarly, POLQ expression did not correlate with HRD status in ESC (Supplementary Figure S6D). We only observed positive associations between POLQ expression and HRD in EEC and PDAC (Supplementary Figure S6D). We conclude that the expression of POLQ previously described in HGOSC (33) cannot be attributed solely to HR-deficiency. COSMIC mutation signature 3 (whose etiology is attributed to HR-deficiency and may involve POLQ) was present in 45.2% (19/42) of ESC and 87.2% (198/227) of HGOSC. We found no correlation between COSMIC mutation signature 3 and POLQ expression with ESC or HGOSC (Supplementary Figure S6E). Cosmic signature 3 was not detectable in most EEC or PDAC (data not shown) and therefore we did not test correlations with POLQ expression in those tumor types. Taken together, we show that the relationship

between HRD status, POLQ expression and COSMIC signature 3 is more similar in clinically related gynecological cancers (ESC and HGOSC) than in the clinically distinct PDAC and EEC.

CRISPR screening defines DDR dependencies of endometrial cancer

Our results suggest that profiles of DDR gene expression in endometrial cancer reflect the genome maintenance requirements for tolerating intrinsic genotoxic stresses of tumorigenesis. To formally identify the DDR genes pathways that sustain high grade endometrial cancer cells, we devised an unbiased genetic screening approach. We engineered a lentiviral vector for co-expressing sgRNAs and a Shield-1-inducible CAS9 (DD-CAS9). The resulting vector was used to construct a library of sgRNAs targeting 504 DDR genes (10 unique sgRNAs per gene) spanning every major DNA repair pathway and many stress response pathways. Using our library, we performed CRISPR dropout screens to identify DDR gene dependencies of HEC-50 (also known as HEC-50B) endometrial cancer cells. HEC-50 cells were originally derived from ascites of a recurrent stage 3, high grade endometrioid tumor (36). HEC-50 cells have many oncogenic features that are shared by other HGEC (37). In mice, HEC-50 cells form tumors with serous-like histology (36). Thus, HEC-50 cells model both ESC and high grade EEC. The design of our CRISPR screen is illustrated in Figure 5A. Our CRISPR screen identified 28 DDR genes that were necessary for the growth of HEC-50 cells (Table 2). Based on stringent thresholds (\log_2FC) for sgRNA dropout, HEC-50 cells showed greatest dependency on *ATM*, *USP7*, *FANCA*, *RNF168*, *RAD51C* and *MRE11A* genes (Figure 5B). We selected some of these genes for independent validation experiments. The clonogenic survival assays in Figure 5C confirm that *USP7*, *RNF168*, and *MRE11* are required for HEC-50 cell growth. Therefore, our unbiased survey of DDR dependencies validates the importance of DDR TransSigGroup genes for HGEC viability.

Next, we examined expression patterns in patient tumors and associations with clinical outcomes for the 28 DDR genes identified by our screen as HEC-50 dependencies (Figure 5D).

As shown in Figure 5E, five of the DDR genes that were required for HEC-50 growth overlapped with the DDR TransSigGroup 5 genes from Figure 2A (*BLM*, *RRM2*, *FANCA*, *TIPIN*, *XRCC2*). Six of the DDR genes required for HEC-50 growth were present in DDR TransSigGroup 2 (*RAD51C*, *RRM1*, *RNF168*, *MDC1*, *BARD1*) Figure 5E. Notably, half of these genes (*BARD1*, *RNF168*, *RRM2*, *FANCA*) correlate with low survival probability and/or worse progression-free survival in EEC patients (Supplemental Figures S4 and S5), which is consistent with pro-tumorigenic roles for these DDR genes. Taken together, these results show that DDR genes expressed at high levels in ESC and a subset of EEC are associated with poor clinical outcomes and are necessary to sustain proliferation of endometrial cancer cells.

ESC and Cluster 3 EEC share distinct CT antigen expression patterns which are associated with genome maintenance and poor overall survival

Many CT Antigens (CTAs; germ cell proteins that are mis-expressed in cancers) pathologically activate DNA repair and confer DNA damage tolerance in tumors (10,12,14). The ways in which CTAs impact genome maintenance (and

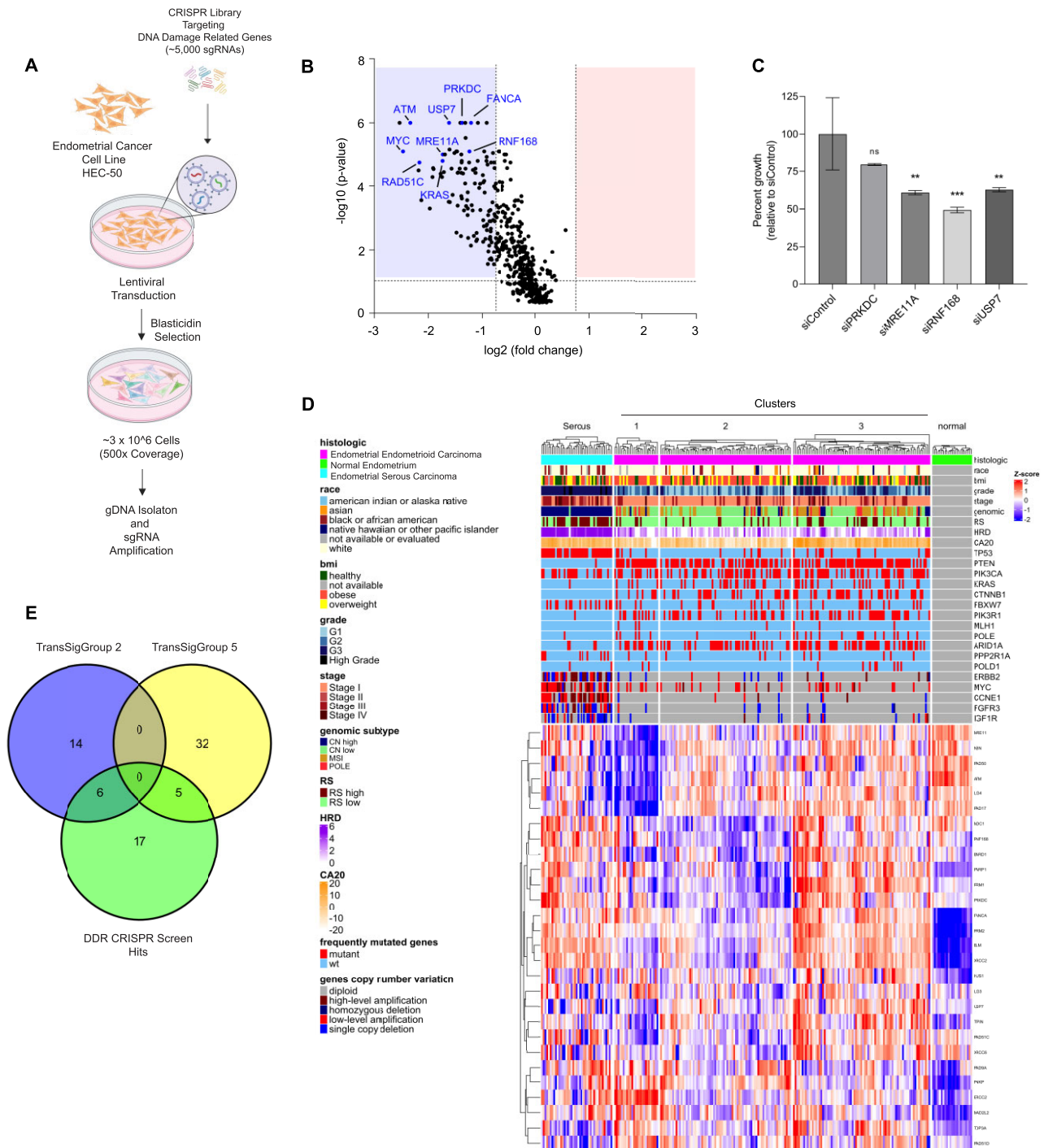


Figure 5. DDR CRISPR screen in high grade endometrial cancer cells reveals dependencies identified with TCGA transcriptome signatures. **(A)** Schematic illustration of CRISPR-Cas9 screen targeting DDR genes in HEC-50 cells, created using BioRender.com. **(B)** Fold depletion in sgRNA targeting DDR genes analyzed by VOLUNDR pipeline. **(C)** Percent growth in HEC-50 cells after siRNA-mediated knockdown of indicated target genes, analyzed using clonogenic survival assay. **(D)** Expression patterns in patient tumors for the 28 DDR genes that dropped out ('hits') with CRISPR screens in HEC-50 cells. **(E)** Venn diagrams depicting DDR CRISPR screen 'hits' in HEC-50 cells compared to DDR genes in DDR TransSigGroup 5 and DDR TransSigGroup 2. **(C)** One-way ANOVA, Dunnett post-test, mean \pm SD. * $P < 0.05$, ** $P < 0.005$, *** $P < 0.0005$, ns = non-significant.

Table 2. DDR CRISPR screen 'hits'

Genes	DDR pathway (s)
ATM	CS
BARD1	HR
BLM	HR
ERCC2	NER
FANCA	FA
HUS1	CS
LIG3	NER
LIG4	NHEJ
MAD2L2	TLS
MDC1	CS
MRE11	NHEJ
NBN	NHEJ
PARP1	Poly ADP ribose polymerases
PNKP	BER
PRKDC	NHEJ
RAD17	CS
RAD50	NHEJ
RAD51C	HR
RAD51D	HR
RAD9A	CS
RNF168	NHEJ
RRM1	NM
RRM2	NM
TIPIN	CS
TOP3A	HR
USP7	TLS
XRCC2	HR
XRCC6	NHEJ

BER, Base Excision Repair; CS, Checkpoint Signaling; FA, Fanconi Anemia; HR, Homologous Recombination; NHEJ, Non-Homologous End Joining; NER, Nucleotide Excision Repair; NM, Nucleotide Metabolism; TLS, Trans-Lesion Synthesis

other tumorigenic characteristics) of endometrial cancers have not been defined. To determine potential roles of CTAs in genome maintenance and the pathobiology of endometrial cancer, we curated a list of 219 genes encoding known CTAs (Supplemental Table S2) and determined their expression patterns in endometrial tumors, using an experimental strategy similar to the workflow described in Figure 1 for DDR genes. The heatmap in Figure 6A shows CTA expression patterns in relation to ESC and EEC Clusters 1, 2 and 3 (as defined by DDR gene expression K means clustering in Figure 2F). Unexpectedly, several CTAs were expressed in normal endometrial tissues, indicating that some genes previously designated as CTAs may also have roles in non-malignant tissues.

Our supervised clustering revealed a distinct set of 16 CTA genes that had very low expression in normal endometrial tissue but were highly expressed in ESC and EEC Cluster 3 (Figure 6A; CTA TransSigGroup 1). Those 'CTA transcriptome gene signature Group 1 (CTA TransSigGroup 1)' are listed in Table 3. The elevated expression of CTA TransSigGroup 1 genes in EEC cluster 3 remained when tumor grade effect was removed (Supplemental Figure S7). CTA TransSigGroup 1 genes significantly correlated with HRD and CA20 scores (Figure 6B, C) and high RS signature (Figure 6D), potentially consistent with genome maintenance-related roles for these CTA.

CTA TransSigGroup 1 gene expression did not correlate with any particular oncogene or tumor suppressor mutations (Figure 6A and Supplemental Figure S7) or molecular sub-

type (Supplemental Figure S8A). Similar to the DDR TransSigGroup 5 and TransSigGroup 2 gene signatures, high expression of CTA TransSigGroup 1 genes was collectively predictive of poor overall survival in EEC and not ESC (Figure 6E, F). Therefore, expression of these CTAs defines distinct disease subsets that cannot be captured by tumor grading, molecular subtyping, or histopathology. These data underscore the potential contribution of CTAs to tumorigenic properties of endometrial tumors.

Given the correlation of CTA TransSigGroup 1 genes with high HRD scores, we asked whether these CTAs share a similar expression pattern in ESC and HGOSC (two clinically similar HRD-high tumors) and PDAC (which are HRD-high yet clinically dissimilar to ESC). Supplemental Figure S8B shows that CTA TransSigGroup 1 genes were expressed at similarly high levels in ESC and HGOSC, but were present at very low levels in PDAC. Therefore, although expression of CTA TransSigGroup 1 genes is associated with HR-deficiency in endometrial cancers, HRD status alone is insufficient to drive expression of CTA TransSigGroup 1 genes in all cancers.

CTA TransSigGroup 1 genes promote tumorigenic phenotypes in endometrial cancer cell lines

We tested the contribution of CTA TransSigGroup 1 genes to tumorigenic behavior of endometrial cancer cells. First, we confirmed that several of the CTA TransSigGroup 1 genes (*KIF2C*, *NUF2*, *CEP55*) were expressed at high levels in endometrial cancer cell lines when compared with normal endometrial epithelial cells (iHEuEC) (Figure 7A). Knockdown of *KIF2C*, *CEP55* and *NUF2* using siRNA reduced clonogenic survival of endometrial cancer cells (Figure 7B), thus implicating a role for these CTAs in promoting a key tumorigenic characteristic. To test the relationship between CTAs and oncogene signaling we conditionally expressed endometrial cancer-relevant oncogenes (*CCNE1*, *KRAS*^{G12V}, *MYC*) in normal endometrial epithelial cells (iHEuEC). As shown in Figure 7C, induced expression of *CCNE1* and *KRAS* oncogenes led to increased expression of *KIF2C*, *CEP55* and *NUF2*. Moreover, clonogenic survival of *KRAS*^{G12V}-expressing cells (but not parental cells lacking *KRAS*^{G12V}) was attenuated by treatment with *NUF2* siRNA (Figure 7D). We conclude that CTA expression can be an adaptive response that confers tolerance of oncogene-associated mitogenic stresses.

KIF2C allows endometrial cancer cells to tolerate DNA replication-stress

One of the most highly expressed CTA TransSigGroup 1 genes was *KIF2C*. This CTA encodes a kinesin-like motor that regulates the dynamic association of microtubules with kinetochores (38). The significance of high-level *KIF2C* expression in neoplastic cells, and the potential contribution of *KIF2C* to tumorigenic and/or chemoresistant phenotypes have not been addressed (for any malignancy). We asked whether *KIF2C* affects the responses of endometrial cancer cells to DNA-damaging events. In cells treated with Hydroxyurea (HU, an agent which depletes nucleotide pools and causes DNA replication stalling), *KIF2C* knockdown led to increased accumulation of pRPA and γ H2AX (Figure 8A). However, in cells treated with neocarzinostatin (NCS) or ultraviolet (UV) radiation (agents which induce DNA DSB and bulky DNA

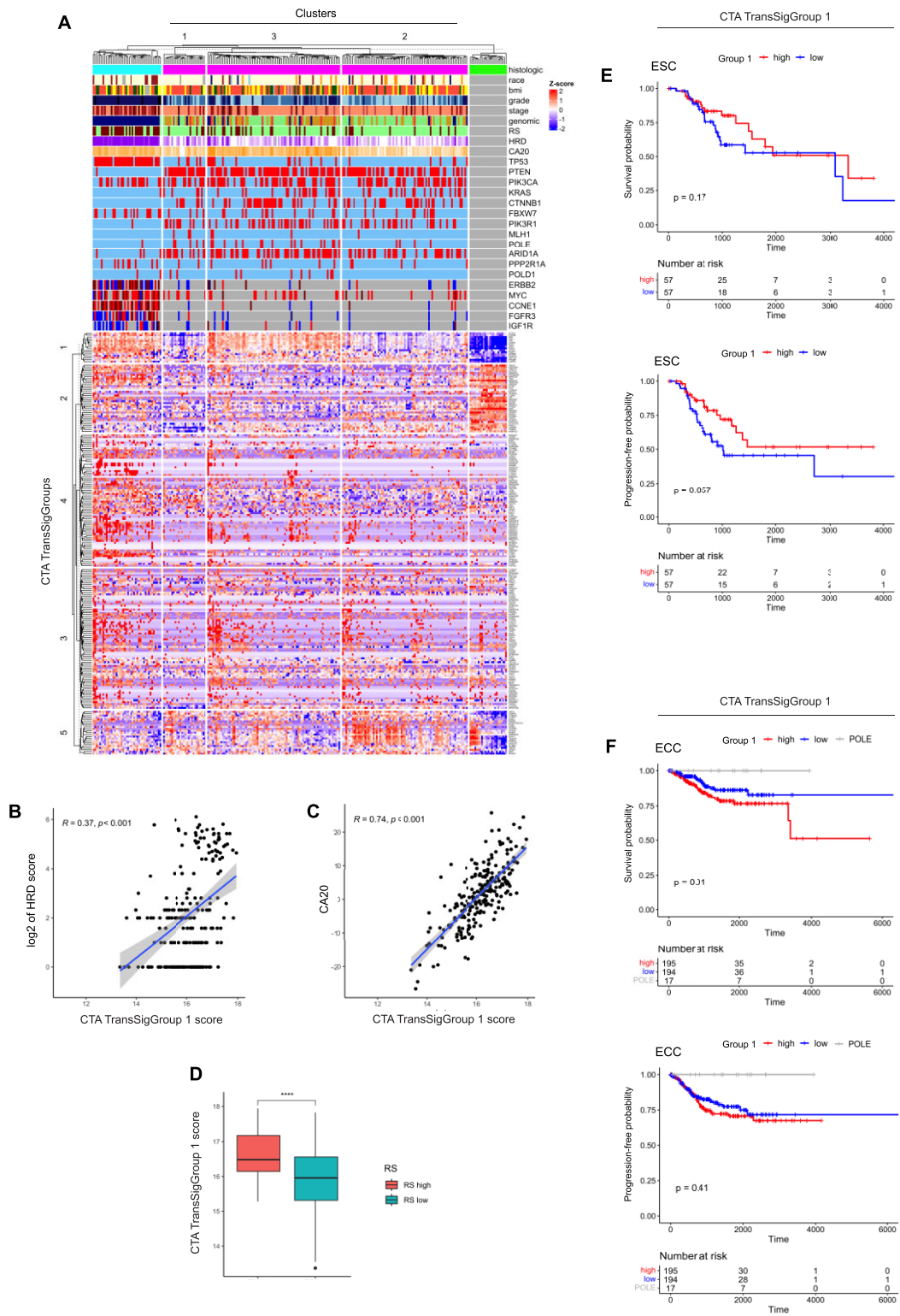


Figure 6. Expression of CT Antigen genes in endometrial cancer with supervised clustering based on histological subtypes. **(A)** Samples from the 2013 TCGA sample set for uterine cancer clustered by endometrial tumor histology (serous or endometrioid), normal endometrium, and K-means clustering of endometrial endometrioid carcinomas (defined in Figure 2A). Heatmaps are depicting relative mRNA expression of CTA genes assembled from a curated list of 248 genes. 219 genes are displayed in the heatmap as 29 genes had no expression in normal endometrium or endometrial tumors. CTA transcriptome gene signatures ('CTA TransSigGroups 1–6'), identified by K-means clustering of DDR gene expression in Figure 1A, are depicted on the heatmap. Spearman's rank correlation of genes of CTA TransSigGroup 1 of all endometrial tumors vs. HRD **(B)** or CA20 **(C)** scores or by RS high or low status **(D)**. CTA TransSigGroup 1 gene score is the average expression of the genes in the grouping. Kaplan-Meier curves for overall survival and progression free survival in ESC **(E)** and non-*POLE* mutant ECC **(F)**. Patient data is from the entire uterine cancer cohort from the TCGA. Data were divided into two groups ('high' and 'low') based on median CTA TransSigGroup 1 gene expression. Log-rank tests were used to infer statistical significance between high and low groups.

Table 3. Group 1 CTA genes

Group 1 genes	CTA role(s)
ACTL8	Actin architecture
ATAD2	Transcriptional regulator
CEP55	Mitosis; centromeric protein
CNOT9	Regulates mRNA degradation
DCAF12	Component of E3 ubiquitin-protein ligase complex
DMRT1	Transcriptional factor
KDM5B	Histone demethylase
KIF20B	MT motor enzyme; MT attachment and chromosome segregation
KIF2C	MT motor enzyme; MT attachment and chromosome segregation
KNL1	Kinetochores-MT attachment and chromosome segregation
NUF2	Mitosis; centromeric proteins
OIP5	Mitosis; centromeric proteins
PBK	Mitotic protein kinase
PRAME	Component of E3 ubiquitin-protein ligase complex
TTK	Protein kinase; chromosome alignment and chromosome segregation
ZNF165	Transcriptional regulator

MT, microtubules.

photo-adducts respectively) KIF2C-depletion did not cause persistent pRPA and γ H2AX (Figure 8A).

We also asked whether KIF2C protects against therapy-induced DNA damage. Inhibitors of the WEE1 protein kinase have been used as experimental therapeutic agents in patients with ESC (39,40). Mechanistically, WEE1 inhibition can lead to DNA replication stress, bypass of G2 checkpoints, and mitotic catastrophe (22). In HEC-1-B cells, KIF2C was recruited to chromatin after WEE1 inhibitor-treatment (MK1775; Figure 8A). MK1775-treatment led to more rapid and high-level accumulation of DNA damage markers pRPA and γ H2AX in KIF2C-ablated HEC-1-B cells when compared with control (KIF2C-replete) cells (Figure 8A).

HEC-1-B cells were also sensitized to MK1775-treatment following KIF2C depletion (Figure 8B; Supplementary Figure S9). In flow cytometry experiments, cells lacking KIF2C failed to maintain normal rates of ongoing DNA synthesis in the presence of WEE1 inhibitor (Figure 8C and D; Supplemental Figure S9F and G). In immunoblotting experiments, the increased levels of pCHK2, pRPA and γ H2AX expression in KIF2C-depleted MK1775-treated HEC-1-B and ARK2 cells are indicative of a role of KIF2C in remediating DNA replication stress (Figure 8E). Immunofluorescence microscopy revealed pan-nuclear γ H2AX in KIF2C-depleted and WEE1 inhibitor-treated cultures (Figure 8F). The γ H2AX present at sites of DNA DSB repair typically shows a pattern that is focal, not pan-nuclear. Therefore, the γ H2AX staining pattern shown in Figure 8F is indicative of DNA replication fork collapse, fully consistent with the reduced DNA synthesis rates of KIF2C-depleted cells shown in Figure 8C and D. Taken together, we conclude that KIF2C is important for tolerance of a subset of endogenously arising as well as pharmacologically induced species of DNA lesions that arise in endometrial cancer cells during S-phase. Moreover, our results indicate that KIF2C, CEP55, NUF2, and likely other CTA confer key tumorigenic properties and contribute to the pathobiology of endometrial cancers.

Discussion

Here, we present the first comprehensive characterization of the DDR and CTA transcriptional landscape of endometrial cancer in relation to conventional histological and molecular subtypes. We show that endometrial tumors have distinctive DDR and CTA expression profiles when compared with other tumor types. Critically, DDR and CTA expression patterns of endometrial cancers associate strongly with clinical outcomes. In fact, DDR-based profiling is prognostic of endometrial cancer patient outcomes in ways that other classification methods fail to capture. Historically, it has been observed that a subset of EEC behave very similarly to ESC (clinical aggressiveness and poor outcomes), despite having different histological and molecular characteristics. Defining these aggressive EEC tumors has been challenging and even high tumor grade (grade 3) has not always been a reliable indicator of EEC patients with aggressive disease. Remarkably, our studies reveal that aggressive *POLE*-proficient endometrioid tumors with poor outcomes can be clearly defined (and shown to resemble the ESC subtype) based on DDR and CTA expression profiling. Thus, DDR and CTA expression-based classifiers (specifically DDR TransSigGroup 5 and DDR TransSigGroup 2) identify aggressive, endometrioid tumors very consistently and independently of histology and molecular subtype and these signatures are not driven solely by tumor grade effects.

Why then do DDR and CTA expression profiles represent such robust and reliable biomarkers for tumor classification? Neoplastic cells exist in various stressful environments and must tolerate unique genotoxic stresses, while acquiring specific genomic changes that propel tumorigenesis. The DDR allows cells to withstand genotoxic stresses. Moreover, deployment of error-prone DNA repair pathways can promote genomic instability. Our work suggests that the unique stresses experienced by different endometrial cancer subtypes may necessitate specific DDR dependencies that are accurately reflected by their DDR transcriptomes.

The two DDR mRNA gene signatures that are strongly associated with ESC and the aggressive EEC subtypes (DDR TransSigGroup 2 and TransSigGroup 5, comprising 20 and 35 DDR genes respectively) have some interesting and noteworthy characteristics: DDR TransSigGroup 2 and 5 contain the *RRM1* and *RRM2* transcripts which encode subunits of Ribonucleotide Reductase (RNR), an enzyme essentially required for the conversion of rNTPs into dNTPs that sustain DNA synthesis and repair (41). DDR TransSigGroup 5 contains mRNAs encoding all three members of the Replication Fork Protection Complex (TIM, TIPIN, CLASPIN) which stabilizes DNA replication forks and promotes S-phase checkpoint signaling (42). DDR TransSigGroup 2 and TransSigGroup 5 also contain genes encoding the S-phase Checkpoint kinase CHEK1 and its activators TOPBP1 and RAD1 (43). Genes encoding other important mediators of DNA replication fork progression are also present in DDR TransSigGroup 2 and TransSigGroup 5 including the Trans-Lesion Synthesis (TLS) factors RAD18 and POLH, and the Template Switch (TS) factor HLTf which cooperates with RAD18 (44). In DDR TransSigGroup 5 and TransSigGroup 2 gene lists there is very strong representation of mRNAs encoding mediators of the Fanconi Anemia (FA) and Homologous Recombination (HR) pathways (Table 1) which also facilitate replication fork progression in times of stress.

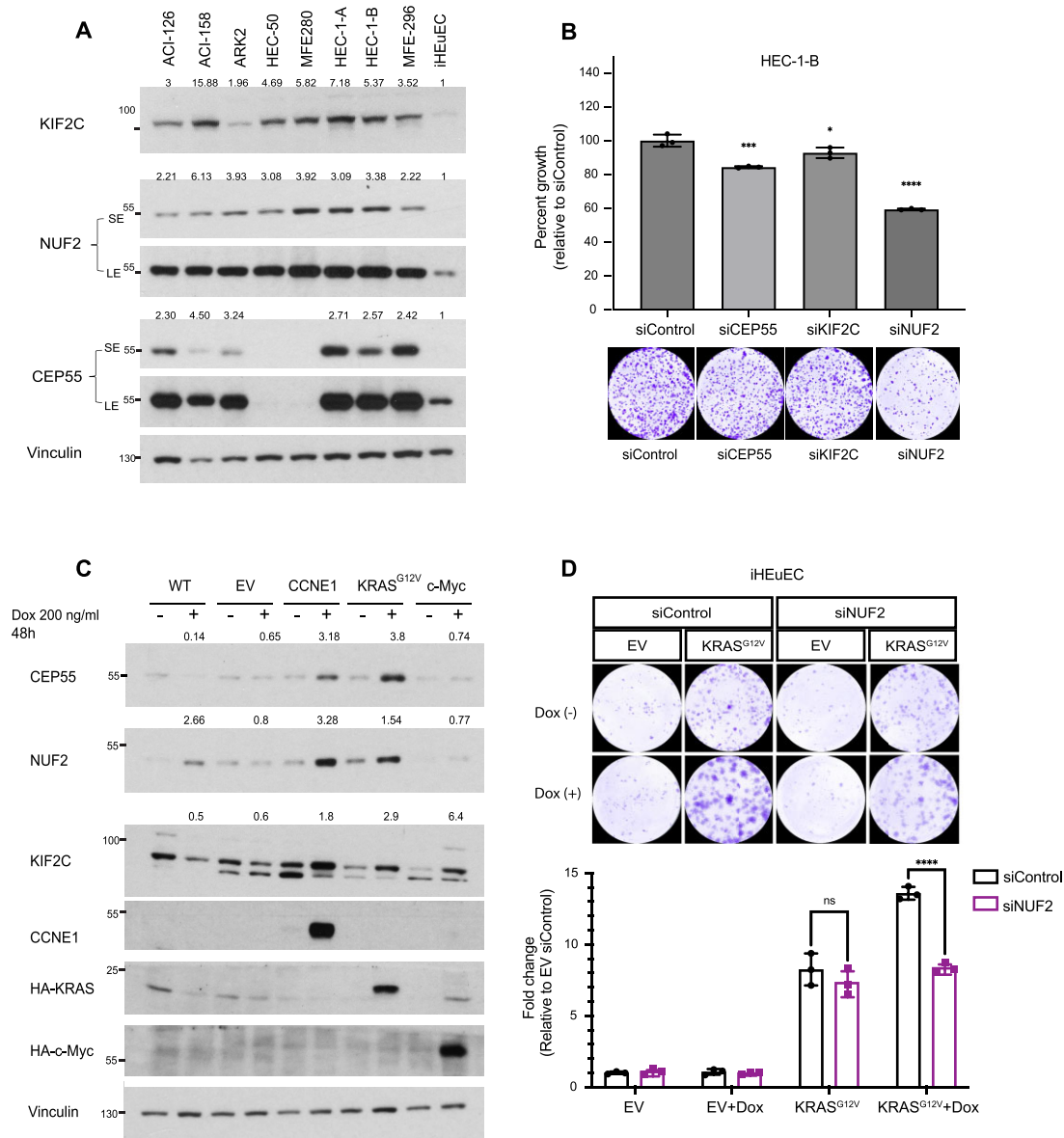


Figure 7. Endometrial cancer cell growth is dependent on CTA TransSigGroup 1 genes. **(A)** Immunoblots showing protein expression of CTAs KIF2C, NUF2, and CEP55 in immortalized human endometrial epithelial cells (iHEuEC) and various endometrial cancer cell lines. Numerical values in each lane indicate fold change in expression of indicated protein relative to its expression in iHEuEC cells. **(B)** Relative growth of HEC-1-B cells after siRNA-mediated knockdown of indicated CTA genes (KIF2C, CEP55 and NUF2) by clonogenic survival assay, quantified relative to control siRNA transfected cells. **(C)** Immunoblots showing protein levels of KIF2C, NUF2 and CEP55 in iHEuEC cells stably expressing Dox-inducible oncogenes *CCNE1*, *KRAS^{G12V}* and *c-MYC*, relative to WT and iHEuEC empty vector (EV) transformed cells. iHEuEC cells were treated with Doxycycline (Dox) 200 ng/ml for 48h. Numerical values in each lane indicate fold change in expression relative to protein levels in their respective no Dox controls. **(D)** Relative growth of iHEuEC EV and iHEuEC expressing mutant *KRAS^{G12V}* after transfection with either non-targeting Control (siControl) or siRNA targeting NUF2 (siNUF2). Relative growths were measured with or without Dox treatment (200 ng/ml) by clonogenic survival assay. Quantification of growth is relative to iHEuEC EV siControl cells (-Dox). (B) One-way ANOVA, Dunnett post-test, mean \pm SD. (C) Two-way ANOVA, Dunnett post-test, mean \pm SD. * $P < 0.05$, *** $P < 0.0005$, **** $P < 0.00005$.

Our CRISPR screen validated many DDR TransSigGroup 2 and TransSigGroup 5 genes (and their associated pathways) as important dependencies of cultured endometrial cancer HEC-50 cells. Genes identified by our CRISPR screen include, the RNR subunits RRM1 and RRM2, the RFC component TIPIN, and RAD9A and HUS1 (whose encoded proteins associate with RAD1 to form the 9–1–1 complex that activates CHEK1). CRISPR screening also confirmed requirements for FA and HR pathway genes including BARD1, FANCA, XRCC2 (FANCU), RAD51C and RAD51D. RNF168 (iden-

tified by our CRISPR screen) is not a designated FA gene, but regulates localization of BARD1 which in turn recruits the BRCA1-PALB2 complex to DNA damage (45). Taken together, our DDR expression profiling and unbiased CRISPR screens suggest that high-grade endometrial cancer cells are particularly dependent on a subset of DDR factors involved in dNTP production, replication fork protection, the FA pathway and HR. The high degree of overlap between the DDR genes overexpressed in TCGA samples and those identified by our CRISPR screen is remarkable considering that cancer

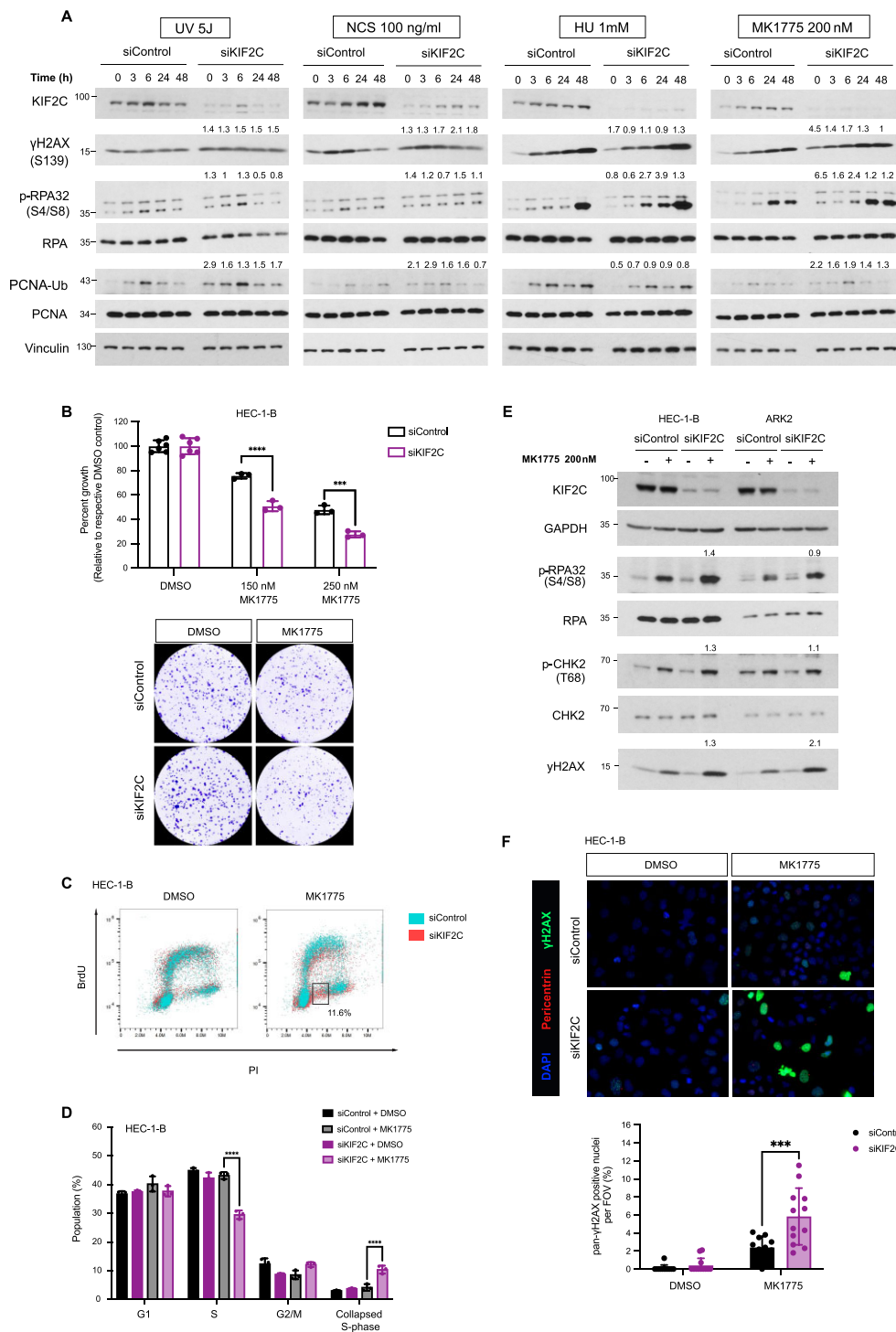


Figure 8. Loss of KIF2C sensitizes endometrial cancer cells to Wee-1-inhibition-induced replication stress. **(A)** Immunoblots showing expression of DNA damage markers in crude chromatin lysates from HEC-1-B cells transfected with siControl or siKIF2C, treated with indicated genotoxic agents at different timepoints. Numerical values in siKIF2C lanes indicate fold change in expression, relative to respective siControl treatment group. **(B)** Relative growth of HEC-1-B cells after transfection with either non-targeting Control (siControl) or siRNA-targeting KIF2C (siKIF2C). Cells were treated with either DMSO or MK1775 (150 nM) for 72 h. Growth was analyzed by a clonogenic survival assay. **(C)** Overlay of BrdU-PI cell cycle profile analyzed by flow cytometry of HEC-1-B cells. Cells were transfected with either siControl or siKIF2C, treated with either DMSO or MK1775 (200 nM) for 48 h. **(D)** Quantification of percentage of cells in different phases of cell cycle (G1, S, and G2/M) along with S-phase defects that can be visualized as collapsed $<4n$ populations. **(E)** Immunoblots showing expression of DNA damage markers in whole cell lysates from HEC-1-B and ARK2 cell lines transfected with siControl or siKIF2C, treated with DMSO or MK1775 (200 nM) for 48 h. Numerical values in siKIF2C lanes indicate fold change in expression, relative to expression in respective siControl treatment group (after normalization to loading control or respective total protein). **(F)** Immunofluorescence images of HEC-1-B cells transfected with siControl or siKIF2C \pm MK1775 (200 nM) treatment for 48h. Transfected cells were synchronized with single thymidine block (2 mM) treatment. After thymidine release, cells are treated with MK1775 for 48 h and then fixed for detection of immunodetection of γ H2AX, pericentrin and DAPI. Quantification of pan- γ H2AX positive nuclei in immunofluorescent images indicative of replicative stress in indicated groups. (B, D and F) Two-way ANOVA, Tukey post-test, mean \pm SD. *** $P < 0.0005$, **** $P < 0.00005$.

cell growth in culture may create very different dependencies from those that arise in an *in vivo* pathological environment. Many of the FA pathway genes that were overexpressed in patient tumors were not required for growth in cell culture. In humans the FA pathway is particularly important for tolerance of acetaldehyde genotoxicity (46). We speculate that metabolic differences between tumors growing *in vivo* and cells in culture might explain the lack of complete overlap between DDR TransSigGroup 2 and 5 genes and dependencies identified by CRISPR screens.

Oncogene-induced DNA replication stress represents a major source of intrinsic DNA damage in cancer cells (11,21,44,47–49). Diverse mechanisms have been proposed to explain how oncogenes induce DNA replication stress (including re-replication, hyper-transcription and R-loop collisions, reactive oxygen species, and altered replication licensing). Moreover, different oncogenes may induce different species of DNA damage. Therefore, we predicted that DDR gene expression patterns might correlate with specific RS-inducing oncogenic drivers and/or with a composite ‘RS’ signature. Consistent with this prediction, the more aggressive ESC subtype (which expressed high levels of DDR TransSigGroup 2 and TransSigGroup 5 genes) did generally have stronger RS signatures when compared with all other endometrial patient tumors. However, in the aggressive Cluster 3 EEC subset (which similar to ESC expressed DDR TransSigGroup 2 and TransSigGroup 5 genes), DDR gene expression patterns did not correlate with alterations in any of the *KRAS*, *MYC*, *CCNE1*, *PTEN*, *TP53*, *PIK3CA*, *FBXW7*, *CTTNB1* or with the RS signature. However, DDR TransSigGroup 5 and TransSigGroup 2 genes (expressed in ESC and aggressive EEC subsets) correlated strongly with the CA20 centrosome amplification signature.

Centrosome amplification occurs in many cancers and contributes to genomic instability and tumor initiation (26,50–53). The formation of a normal bipolar mitotic spindle strictly requires the presence of two centrosomes at mitosis. Excessive numbers of centrosomes can cause spindle multipolarity, and mitotic delay leading to DNA damage through direct and indirect mechanisms (54–59). Therefore, cancer cells with supernumerary centrosomes must adapt to withstand the stresses and DNA damage caused by multipolar spindles. Altered transcriptional programs such as upregulation of DDR TransSigGroup 2 and TransSigGroup 5 genes is a plausible explanation for how EEC tolerate DNA damage caused by supernumerary centrosomes in CA20-high tumors.

In both ESC and EEC, high CA20 scores also correlated strongly with expression of a subset of 10 CTA genes (which we designated ‘CTA TransSigGroup 1’). Genes represented in the CTA TransSigGroup 1 list are *CNOT9*, *ATAD2*, *PBK*, *KNL1*, *CEP55*, *KIF20B*, *TTK*, *NUF2C*, *KIF2C* and *OIP5*. While two of the CTA TransSigGroup 1 gene products have roles in gene expression (*CNOT9* regulates mRNA degradation (60) and *ATAD2* is a transcriptional regulator (61)), most have roles in mitosis: *NUF2* and *OIP5* are centromeric proteins with roles in chromosome segregation (62,63). *PBK* is a mitotic protein kinase that mediates DNA damage-induced G2/M checkpoint control (64). *TTK* (*MPS1*) encodes a dual specificity protein kinase that is essential for chromosome alignment and segregation during mitosis (65). *CEP55* (Centrosomal Protein 55) regulates mitotic cytokinesis (66). *KNL1* is a *TTK*/*MPS1* substrate required for creation of kinetochore-microtubule (MT) attachments and chromosome

segregation (67). *KIF20B* and *KIF2C* are MT motor enzymes required for cytokinesis (68). Remarkably, we show that several Group 1 CTA are oncogene-inducible and required for growth of endometrial cancer cell lines.

We hypothesize that there is a pathologically meaningful relationship between oncogenes, DNA replication stress, HRD, and mitotic CTAs. It is well known that oncogenes cause DNA replication stress (44). We and others have demonstrated that HR allows cells to tolerate oncogene-induced DNA replication stress (22). Moreover, when DNA replication stress is not remediated in S-phase backup genome maintenance processes become more important later in the cell cycle (22,44,69). This likely explains why mitotic regulators such as *PLK1* are important for viability of oncogene-expressing cells (70). Taken together, the most reasonable explanation of our results is that elevated expression of mitotic CTAs is an adaptive response to oncogene-induced DNA replication stress. This interpretation is fully supported both by our past results (22) and by results of this study showing: (i) that CTAs are induced when we express oncogenes in normal endometrial epithelial cells and (ii) that CTAs are required for viability of normal epithelial cells only when oncogenes are expressed. There is no paradigm for why any of the CTA genes studied here are aberrantly expressed in any cancer. There is also no existing paradigm for CTA expression as an enabling feature of oncogene-expressing cells (for any oncogene or any cancer). Taken together our results suggest that CTA TransSigGroup1 are very important for regulating chromosome segregation and remediating mitotic stresses experienced by endometrial cancer cells with supernumerary centrosomes. Interestingly, other CTAs have also been implicated as important requirements for mitotic progression of lung cancer cells (71).

Because CTA TransSigGroup 1 genes are absent from normal somatic cells, these proteins are attractive as therapeutic targets for sensitizing cancer cells to mitotic stress. Importantly, CTA-directed therapies would be innocuous to normal healthy cells (which lack CTA expression). Two of the CTA TransSigGroup 1 genes identified here (*PBK* and *TTK*) are protein kinases and are druggable (65). The other genes in CTA TransSigGroup 1 have not been evaluated for chemical tractability. However, the MT motors *KIF2C* and *KIF1B* use ATP as a substrate and therefore might well be druggable based on ATP-binding site competitive inhibitors. Mitotic stress can result from intrinsic defects as well as microtubule poisons such as taxanes (72). Microtubule inhibitors are commonly used in endometrial cancer chemotherapy regimens. The mechanism of anti-mitotic agents such as paclitaxel is attributed to induction of spindle multi-polarity (72). Therefore, therapies targeting CTA TransSigGroup 1 genes might also be useful for sensitizing endometrial cancers to paclitaxel and other microtubule inhibitors. Taken together, the DDR and CTA landscape revealed by our study provides new opportunities for therapies that target vulnerabilities of aggressive endometrial cancer subtypes. Our study fully validates the use of DDR and CTA-based tumor classifiers and justifies extension of this classification method to other cancer types.

Data availability

All data are incorporated into the article and its online supplementary material. Source code is available on FigShare under DOI: 10.6084/m9.figshare.24582798. Any more data required are available upon request.

Supplementary data

Supplementary Data are available at NAR Cancer Online.

Funding

National Institutes of Health (R01 ES009558, CA215347, CA229530 to C.V.); University of North Carolina Lineberger Developmental Funding Program Stimulus Award (to J.L.B. and C.V.); University of the North Carolina Center for Environmental Health and Susceptibility 2023–2024 Pilot Project Awards (to J.L.B.).

Conflict of interest statement

None declared.

References

- Walens,A., Van Alsten,S.C., Olsson,L.T., Smith,M.A., Lockhart,A., Gao,X., Hamilton,A.M., Kirk,E.L., Love,M.I., Gupta,G.P., *et al.* (2022) RNA-based classification of homologous recombination deficiency in racially diverse patients with breast cancer. *Cancer Epidemiol. Biomarkers Prev.*, **31**, 2136–2147.
- Parker,J.S., Mullins,M., Cheang,M.C., Leung,S., Voduc,D., Vickery,T., Davies,S., Fauron,C., He,X., Hu,Z., *et al.* (2009) Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.*, **27**, 1160–1167.
- Prat,A., Parker,J.S., Fan,C. and Perou,C.M. (2012) PAM50 assay and the three-gene model for identifying the major and clinically relevant molecular subtypes of breast cancer. *Breast Cancer Res. Treat.*, **135**, 301–306.
- Hamilton,C.A., Cheung,M.K., Osann,K., Chen,L., Teng,N.N., Longacre,T.A., Powell,M.A., Hendrickson,M.R., Kapp,D.S. and Chan,J.K. (2006) Uterine papillary serous and clear cell carcinomas predict for poorer survival compared to grade 3 endometrioid corpus cancers. *Br. J. Cancer*, **94**, 642–646.
- Clarke,M.A., Devesa,S.S., Harvey,S.V. and Wentzensen,N. (2019) Hysterectomy-corrected uterine corpus cancer incidence trends and differences in relative survival reveal racial disparities and rising rates of nonendometrioid cancers. *J. Clin. Oncol.*, **37**, 1895–1908.
- Cancer Genome Atlas Research, N., Kandoth,C., Schultz,N., Cherniack,A.D., Akbani,R., Liu,Y., Shen,H., Robertson,A.G., Pashtan,I., Shen,R., *et al.* (2013) Integrated genomic characterization of endometrial carcinoma. *Nature*, **497**, 67–73.
- Dou,Y., Kawaler,E.A., Cui Zhou,D., Gritsenko,M.A., Huang,C., Blumenberg,L., Karpova,A., Petyuk,V.A., Savage,S.R., Satpathy,S., *et al.* (2020) Proteogenomic characterization of endometrial carcinoma. *Cell*, **180**, 729–748.
- Moxley,K.M. and McMeekin,D.S. (2010) Endometrial carcinoma: a review of chemotherapy, drug resistance, and the search for new agents. *Oncologist*, **15**, 1026–1033.
- Whitehurst,A.W. (2014) Cause and consequence of cancer/testis antigen activation in cancer. *Annu. Rev. Pharmacol. Toxicol.*, **54**, 251–272.
- Gao,Y., Mutter-Rottmayer,E., Greenwalt,A.M., Goldfarb,D., Yan,F., Yang,Y., Martinez-Chacin,R.C., Pearce,K.H., Tateishi,S., Major,M.B., *et al.* (2016) A neomorphic cancer cell-specific role of MAGE-A4 in trans-lesion synthesis. *Nat. Commun.*, **7**, 12105.
- Gao,Y., Tateishi,S. and Vaziri,C. (2016) Pathological trans-lesion synthesis in cancer. *Cell Cycle*, **15**, 3005–3006.
- Gao,Y., Kardos,J., Yang,Y., Tamir,T.Y., Mutter-Rottmayer,E., Weissman,B., Major,M.B., Kim,W.Y. and Vaziri,C. (2018) The cancer/testes (CT) antigen HORMAD1 promotes homologous recombinational DNA repair and radioresistance in lung adenocarcinoma cells. *Sci. Rep.*, **8**, 15304.
- Nichols,B.A., Oswald,N.W., McMillan,E.A., McGlynn,K., Yan,J., Kim,M.S., Saha,J., Mallipeddi,P.L., LaDuke,S.A., Villalobos,P.A., *et al.* (2018) HORMAD1 Is a negative prognostic indicator in lung adenocarcinoma and specifies resistance to oxidative and genotoxic stress. *Cancer Res.*, **78**, 6196–6208.
- Jay,A., Reitz,D., Namekawa,S.H. and Heyer,W.D. (2021) Cancer testis antigens and genomic instability: more than immunology. *DNA Repair (Amst.)*, **108**, 103214.
- Gu,Z., Eils,R. and Schlesner,M. (2016) Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics*, **32**, 2847–2849.
- Knijnenburg,T.A., Wang,L., Zimmermann,M.T., Chambwe,N., Gao,G.F., Cherniack,A.D., Fan,H., Shen,H., Way,G.P., Greene,C.S., *et al.* (2018) Genomic and molecular landscape of DNA damage repair deficiency across the cancer genome atlas. *Cell Rep.*, **23**, 239–254.
- Konstantinopoulos,P.A., da Costa,A., Gulhan,D., Lee,E.K., Cheng,S.C., Hendrickson,A.E.W., Kochupurakkal,B., Kolin,D.L., Kohn,E.C., Liu,J.F., *et al.* (2021) A replication stress biomarker is associated with response to gemcitabine versus combined gemcitabine and ATR inhibitor therapy in ovarian cancer. *Nat. Commun.*, **12**, 5574.
- de Almeida,B.P., Vieira,A.F., Paredes,J., Bettencourt-Dias,M. and Barbosa-Morais,N.L. (2019) Pan-cancer association of a centrosome amplification gene expression signature with genomic alterations and clinical outcome. *PLoS Comput. Biol.*, **15**, e1006832.
- Coll-de la Rubia,E., Martinez-Garcia,E., Dittmar,G., Gil-Moreno,A., Cabrera,S. and Colas,E. (2020) Prognostic biomarkers in endometrial cancer: a systematic review and meta-analysis. *J. Clin. Med.*, **9**, 1900.
- Gates,E.J., Hirschfield,L., Matthews,R.P. and Yap,O.W. (2006) Body mass index as a prognostic factor in endometrioid adenocarcinoma of the endometrium. *J. Natl. Med. Assoc.*, **98**, 1814–1822.
- Bowry,A., Kelly,R.D.W. and Petermann,E. (2021) Hypertranscription and replication stress in cancer. *Trends Cancer*, **7**, 863–877.
- Yang,Y., Gao,Y., Mutter-Rottmayer,L., Zlatanou,A., Durando,M., Ding,W., Wyatt,D., Ramsden,D., Tanoue,Y., Tateishi,S., *et al.* (2017) DNA repair factor RAD18 and DNA polymerase polkappa confer tolerance of oncogenic DNA replication stress. *J. Cell Biol.*, **216**, 3097–3115.
- Yang,Y., Gao,Y., Zlatanou,A., Tateishi,S., Yurchenko,V., Rogozin,I.B. and Vaziri,C. (2018) Diverse roles of RAD18 and Y-family DNA polymerases in tumorigenesis. *Cell Cycle*, **17**, 833–843.
- Vaziri,C., Rogozin,I.B., Gu,Q., Wu,D. and Day,T.A. (2021) Unravelling roles of error-prone DNA polymerases in shaping cancer genomes. *Oncogene*, **40**, 6549–6565.
- Cong,K. and Cantor,S.B. (2022) Exploiting replication gaps for cancer therapy. *Mol. Cell*, **82**, 2363–2369.
- Godinho,S.A. and Pellman,D. (2014) Causes and consequences of centrosome abnormalities in cancer. *Philos. Trans. Roy. Soc. Lond.*, **369**, 20130467.
- Ogden,A., Rida,P.C. and Aneja,R. (2017) Prognostic value of CA20, a breast based on centrosome amplification-associated genes. *Breast Tumors. Sci. Rep.*, **7**, 262.
- McConechy,M.K., Talhouk,A., Leung,S., Chiu,D., Yang,W., Senz,J., Reha-Krantz,L.J., Lee,C.H., Huntsman,D.G., Gilks,C.B., *et al.* (2016) Endometrial carcinomas with POLE exonuclease domain mutations have a favorable prognosis. *Clin. Cancer Res.*, **22**, 2865–2873.
- Billingsley,C.C., Cohn,D.E., Mutch,D.G., Hade,E.M. and Goodfellow,P.J. (2016) Prognostic significance of POLE exonuclease domain mutations in high-grade endometrioid endometrial cancer on survival and recurrence: a subanalysis. *Int. J. Gynecol. Cancer*, **26**, 933–938.

30. Leon-Castillo,A., Gilvazquez,E., Nout,R., Smit,V.T., McAlpine,J.N., McConechy,M., Kommoss,S., Brucker,S.Y., Carlson,J.W., Epstein,E., *et al.* (2020) Clinicopathological and molecular characterisation of ‘multiple-classifier’ endometrial carcinomas. *J. Pathol.*, **250**, 312–322.
31. Das Ghosh,D., Roy Chowdhury,R., Dutta,R., Mukhopadhyay,I., Mukhopadhyay,A. and Roychoudhury,S. (2023) In-silico analysis of TCGA data showing multiple POLE-like favourable subgroups overlapping with TP53 mutated endometrial cancer: implications for clinical practice in low and middle-income countries. *Gynecol Oncol. Rep.*, **47**, 101209.
32. Rashid,S., Arafah,M.A. and Akhtar,M. (2022) The many faces of serous neoplasms and related lesions of the female pelvis: a review. *Adv. Anat. Pathol.*, **29**, 154–167.
33. Ceccaldi,R., Liu,J.C., Amunugama,R., Hajdu,I., Primack,B., Petalcorin,M.I., O’Connor,K.W., Konstantinopoulos,P.A., Elledge,S.J., Boulton,S.J., *et al.* (2015) Homologous-recombination-deficient tumours are dependent on Poltheta-mediated repair. *Nature*, **518**, 258–262.
34. van Schendel,R., van Heteren,J., Welten,R. and Tijsterman,M. (2016) Genomic scars generated by Polymerase Theta reveal the versatile mechanism of alternative end-joining. *PLoS Genet.*, **12**, e1006368.
35. Marquard,A.M., Eklund,A.C., Joshi,T., Krzystanek,M., Favero,F., Wang,Z.C., Richardson,A.L., Silver,D.P., Szallasi,Z. and Birkbak,N.J. (2015) Pan-cancer analysis of genomic scar signatures associated with homologous recombination deficiency suggests novel indications for existing cancer drugs. *Biomark. Res.*, **3**, 9.
36. Kuramoto,H., Nishida,M., Morisawa,T., Hamano,M., Hata,H., Kato,Y., Ohno,E. and Iida,T. (1991) Establishment and characterization of human endometrial cancer cell lines. *Ann. N.Y. Acad. Sci.*, **622**, 402–421.
37. Devor,E.J., Gonzalez-Bosquet,J., Thiel,K.W. and Leslie,K.K. (2020) Genomic characterization of five commonly used endometrial cancer cell lines. *Int. J. Oncol.*, **57**, 1348–1357.
38. Bakhom,S.F., Thompson,S.L., Manning,A.L. and Compton,D.A. (2009) Genome stability is ensured by temporal control of kinetochore-microtubule dynamics. *Nat. Cell Biol.*, **11**, 27–35.
39. Liu,J.F., Xiong,N., Campos,S.M., Wright,A.A., Krasner,C., Schumer,S., Horowitz,N., Veneris,J., Tayob,N., Morrissey,S., *et al.* (2021) Phase II study of the WEE1 inhibitor Adavosertib in recurrent uterine serous carcinoma. *J. Clin. Oncol.*, **39**, 1531–1539.
40. Madariaga,A. and Oza,A.M. (2021) Wee1 Inhibition in recurrent serous uterine cancer: science paving the way in a challenging disease. *J. Clin. Oncol.*, **39**, 1513–1517.
41. Greene,B.L., Kang,G., Cui,C., Bennati,M., Nocera,D.G., Drennan,C.L. and Stubbe,J. (2020) Ribonucleotide reductases: structure, chemistry, and metabolism suggest new therapeutic targets. *Annu. Rev. Biochem.*, **89**, 45–75.
42. Leman,A.R. and Noguchi,E. (2012) Local and global functions of timeless and tipin in replication fork protection. *Cell Cycle*, **11**, 3945–3955.
43. Reinhardt,H.C. and Yaffe,M.B. (2009) Kinases that control the cell cycle in response to DNA damage: chk1, Chk2, and MK2. *Curr. Opin. Cell Biol.*, **21**, 245–255.
44. Anand,J., Chiou,L., Sciandra,C., Zhang,X., Hong,J., Wu,D., Zhou,P. and Vaziri,C. (2023) Roles of trans-lesion synthesis (TLS) DNA polymerases in tumorigenesis and cancer therapy. *NAR Cancer*, **5**, zcad005.
45. Kraiss,J.J., Wang,Y., Patel,P., Basu,J., Bernhardt,A.J. and Johnson,N. (2021) RNF168-mediated localization of BARD1 recruits the BRCA1-PALB2 complex to DNA damage. *Nat. Commun.*, **12**, 5016.
46. Garaycochea,J.I. and Patel,K.J. (2014) Why does the bone marrow fail in Fanconi anemia? *Blood*, **123**, 26–34.
47. Kotsantis,P., Petermann,E. and Boulton,S.J. (2018) Mechanisms of oncogene-induced replication stress: jigsaw falling into place. *Cancer Discov.*, **8**, 537–555.
48. Petropoulos,M., Champeris Tsaniras,S., Taraviras,S. and Lygerou,Z. (2019) Replication licensing aberrations, Replication stress, and genomic instability. *Trends Biochem. Sci.*, **44**, 752–764.
49. Bartkova,J., Rezaei,N., Liontos,M., Karakaidos,P., Kletsas,D., Issaeva,N., Vassiliou,L.V., Kolettas,E., Niforou,K., Zoumpourlis,V.C., *et al.* (2006) Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature*, **444**, 633–637.
50. Godinho,S.A., Kwon,M. and Pellman,D. (2009) Centrosomes and cancer: how cancer cells divide with too many centrosomes. *Cancer Metastasis Rev.*, **28**, 85–98.
51. Levine,M.S., Bakker,B., Moyett,J., Lu,J., Vitre,B., Spierings,D.C., Lansdorp,P.M., Cleveland,D.W., Lambrechts,D., *et al.* (2017) Centrosome amplification is sufficient to promote spontaneous tumorigenesis in mammals. *Dev. Cell*, **40**, 313–322.
52. Fukasawa,K. (2005) Centrosome amplification, chromosome instability and cancer development. *Cancer Lett.*, **230**, 6–19.
53. Godinho,S.A., Picone,R., Burute,M., Dagher,R., Su,Y., Leung,C.T., Polyak,K., Brugge,J.S., Thery,M. and Pellman,D. (2014) Oncogene-like induction of cellular invasion from centrosome amplification. *Nature*, **510**, 167–171.
54. Passerini,V., Ozeri-Galai,E., de Pagter,M.S., Donnelly,N., Schmalbrock,S., Kloosterman,W.P., Kerem,B. and Storchova,Z. (2016) The presence of extra chromosomes leads to genomic instability. *Nat. Commun.*, **7**, 10754.
55. Sharma,S.S., Ma,L., Bagui,T.K., Forinash,K.D. and Pledger,W.J. (2012) A p27Kip1 mutant that does not inhibit CDK activity promotes centrosome amplification and micronucleation. *Oncogene*, **31**, 3989–3998.
56. Loffler,H., Lukas,J., Bartek,J. and Kramer,A. (2006) Structure meets function—centrosomes, genome maintenance and the DNA damage response. *Exp. Cell Res.*, **312**, 2633–2640.
57. Ganem,N.J. and Pellman,D. (2012) Linking abnormal mitosis to the acquisition of DNA damage. *J. Cell Biol.*, **199**, 871–881.
58. Yang,Z., Loncarek,J., Khodjakov,A. and Rieder,C.L. (2008) Extra centrosomes and/or chromosomes prolong mitosis in human cells. *Nat. Cell Biol.*, **10**, 748–751.
59. Hayashi,M.T. and Karlseder,J. (2013) DNA damage associated with mitosis and cytokinesis failure. *Oncogene*, **32**, 4593–4601.
60. Pavanello,L., Hall,B., Airhihen,B. and Winkler,G.S. (2018) The central region of CNOT1 and CNOT9 stimulates deadenylation by the Ccr4-not nuclease module. *Biochem. J.*, **475**, 3437–3450.
61. Boussouar,F., Jamshidikia,M., Morozumi,Y., Rousseaux,S. and Khochbin,S. (2013) Malignant genome reprogramming by ATAD2. *Biochim. Biophys. Acta*, **1829**, 1010–1014.
62. Salmon,E.D., Cimini,D., Cameron,L.A. and DeLuca,J.G. (2005) Merotelic kinetochores in mammalian tissue cells. *Philos. Trans. Roy. Soc. Lond.*, **360**, 553–568.
63. Fujita,Y., Hayashi,T., Kiyomitsu,T., Toyoda,Y., Kokubu,A., Obuse,C. and Yanagida,M. (2007) Priming of centromere for CENP-A recruitment by human hMis18alpha, hMis18beta, and M18BP1. *Dev. Cell*, **12**, 17–30.
64. Gaudet,S., Branton,D. and Lue,R.A. (2000) Characterization of PDZ-binding kinase, a mitotic kinase. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 5167–5172.
65. Lan,W. and Cleveland,D.W. (2010) A chemical tool box defines mitotic and interphase roles for Mps1 kinase. *J. Cell Biol.*, **190**, 21–24.
66. Sinha,D., Duijf,P.H.G. and Khanna,K.K. (2019) Mitotic slippage: an old tale with a new twist. *Cell Cycle*, **18**, 7–15.
67. Funabiki,H. and Wynne,D.J. (2013) Making an effective switch at the kinetochore by phosphorylation and dephosphorylation. *Chromosoma*, **122**, 135–158.
68. Walczak,C.E. (2003) The kin I kinesins are microtubule end-stimulated ATPases. *Mol. Cell*, **11**, 286–288.
69. Tirman,S., Quinet,A., Wood,M., Meroni,A., Cybulla,E., Jackson,J., Pegoraro,S., Simoneau,A., Zou,L. and Vindigni,A. (2021) Temporally distinct post-replicative repair mechanisms fill

- PRIMPOL-dependent ssDNA gaps in human cells. *Mol. Cell*, **81**, 4026–4040.
70. Luo,J., Emanuele,M.J., Li,D., Creighton,C.J., Schlabach,M.R., Westbrook,T.F., Wong,K.K. and Elledge,S.J. (2009) A genome-wide RNAi screen identifies multiple synthetic lethal interactions with the Ras oncogene. *Cell*, **137**, 835–848.
71. Cappell,K.M., Sinnott,R., Taus,P., Maxfield,K., Scarbrough,M. and Whitehurst,A.W. (2012) Multiple cancer testis antigens function to support tumor cell mitotic fidelity. *Mol. Cell. Biol.*, **32**, 4131–4140.
72. Weaver,B.A. (2014) How Taxol/paclitaxel kills cancer cells. *Mol. Biol. Cell*, **25**, 2677–2681.