Supplemental Experimental procedures

Ovarian cancer cell lines

We used three criteria to find the most representative cell model for detecting the binding profiles of ERα: i) it has been traditionally used as a representative HGSOC cell line, ii) it has the characteristic TP53 mutation and iii) it has a wild type and strong expression of ERα.

The three most commonly used cell lines for HGSOC are SKOV3, A2780 and OVCAR3 (Ordered by the number of citations), with SKOV3 having the strongest expression of ESR1 (RPKM=5.5, 0.004 and 0.1 for SKOV3, A2780 and OVCAR3, respectively, according to CCLE RNAseq data). However, this has been recently questioned by the analysis of genomic profiles [1]. For example, SKOV3 and A2780 were suggested as unlikely HGSOC cell models mainly due to the lack of TP53 mutation [1]. At the moment, CCLE data has been greatly updated by further including whole-genome sequencing (WGS) and RNA sequencing data, as well as WES data released by Sanger institute [2]. Although CCLE WES data fail to reveal a TP53 mutation in SKOV3 [1], three other independent resources, including the CCLE WGS data and RNAseq data, and the Sanger WES data, reveal the same Frame_Shift_Del in TP53, indicating that SKOV3 is a *bona fide* TP53 mutated cell line. SKOV3 was also previously reported to have a 32bp deletion in exon 1 of ESR1 and nonresponsive to estrogen [3]. However, the study didn't verify the identity of the cell line. Our SKOV3 cells was provided by China Infrastructure of cell line Resource (CICLR) and supported by the cell line authentication that matches 100% to ATCC HTB-77 SKOV3 cell line. All the current genomic resources of SKOV3, including WES data of CCLE and Sanger institute, and WGS and RNAseq data of CCLE, support a wild type ERα [2]. Our data in SKOV3 cells further indicate that the DNA binding of ERα shows response to estrogen. Taken together, we selected SKOV3 as the cell model for detecting genome-wide ERα binding profiles. We also used other HGSOC cell lines HO8910 and CAOV3 to validate the findings. Notably, HO8910 shows a strong expression of ERα and a good capacity in forming engrafted tumors (Please refer to Cellosaurus database for more information: https://web.expasy.org/cellosaurus/CVCL_6868), and was used as the cell model over SKOV3 for engrafted tumor given its higher sensitivity to cisplatin.

ChIP-Seq binding enrichment analysis

Sequences generated by the Illumina genome analyzer were aligned against genome version hg19 using bowtie 2 [4]. CtBP bindings were defined as the enriched regions of

the genome that were identified by comparing the CtBP pull-down samples to input samples using MACS2 [5], whereas ERα bindings were defined by comparing E2 treated ChIP samples with Etoh treated samples. Aligned sequences were piled up using MACS2, normalized over input and 10 million reads, and visualized using integrated genome viewer. Heatmaps of binding affinity were generated by R Bioconductor package "ChIPpeakAnno" [6].

Cohort datasets collection

Cohort datasets of ovarian cancer were obtained through database searches of PubMed, ArrayExpress, TCGA and GEO. Each cohort dataset was processed and analyzed independently. All clinical information was extracted either from GEO using R package "GEOquery", or from the publications associated with the data. We required that the datasets should consist of at least 40 patients with clinical information including at least overall survival and vital status. For microarray data, the most sensitive probe set of a gene (highest average signal across all samples) was selected as the expression value of the gene. Meta-analysis for the prognostic value of ERα expression was performed using R package "metafor".

ERα activity signature

A list of 157 ERα co-expressed genes, defined by the overlapped subset of the top 1,000 co-expressed genes (about top ~0.05%) in at least 1/3 of the cohort datasets, was used to construct the signature. Among these signature genes, 124 genes were positively correlated with ERα expression and 33 genes were negatively correlated with ERα expression. An ERα activity α_j of patient j is calculated based on the expression of signature genes:

$$\alpha_j = \frac{\sum_i^n w_i \times x_{ij}}{\sqrt{n}}$$

where x_{ij} is the expression of gene i in patient j, and n is the number of signature genes. The weight w_i of each gene i is determined by:

$$w_i = \frac{\sum_k r_{ki} \times S_k}{\sum_k S_k}$$

where r_{ki} is the Pearson correlation coefficient (PCC) between gene i and ERα in dataset k, and S_k is the number of samples in dataset k.

All gene expression values have been Z-score transformed to ensure that the expression values of different genes are scale free. Therefore, for a given patient j, the ERα activity α_j is a weighted combination of the Z-score of the 157 selected genes. As a validation, we

observed a strong correlation between this EOC-derived signature and ERα in an independent GTEx dataset across normal female tissues (Supplementary Figure 8C).

Differentially expressed genes (DEGs) and differential binding analysis (DBA)

Significant DEGs were identified by the Bioconductor package "edgeR" using read count-based expression data. Significantly differentially bound sites were also identified using "edgeR", where peaks instead of genes were used as units. Peaks that are shared by CtBP and ERα were included in DBA, and the overlapped region was defined as the peak length. An FDR cutoff of 0.05 was used to define significantly differentially bound peaks.

Integration of ChIP-Seq bindings with DEGs

CtBP-specific bindings, ERα-specific bindings, Weaker bindings and No change bindings were annotated separately to genes using Bioconductor package "ChIPpeakAnno". We defined the promoter region as the 10kb window centered on TSSs. Thus, genes that were annotated by "ChIPpeakAnno" have binding events within the 10kb window of its TSS.

Motif analysis of Transcription factors

De novo motif discovery was done using the MEME suite (http://meme-suite.org). The enrichment of known motifs was performed using Bioconductor package "PWMEnrich". 2287 position weight matrixes (PWMs) of USCS hg19 version were analyzed.

PAN-cancer expression of ERα

Pan-cancer ERα expression was downloaded by querying the Cancer Genomic Data Server via the R package "CGDS-R".

Literature search and data collection for the prognostic analysis of HRT

We searched the Medline database using designed medical subject heading (MeSH) terms: (("hormone replacement therapy" [MeSH terms] OR "Hormone replacement therapy" [all fields]) OR ("Estrogen replacement therapy" [MeSH terms] OR "Estrogen replacement therapy" [all fields]) OR ("Hormone replacement" [all fields] OR "Hormone therapy" [all fields]) OR "Estrogen Antagonists" [MeSH terms] OR ("Hormone" [MeSH terms] AND ("Tamoxifen" [MeSH terms] OR "Estrogens" [MeSH terms] OR "Estradiol" [MeSH terms]))) AND ("ovarian neoplasms" [MeSH terms] OR "ovarian Carcinoma" [all fields] OR "ovarian cancer"[all fields] OR "ovarian neoplasms" [all fields] OR "ovary cancer"[all fields]). Eligibility criteria of patients were determined in the original

literatures. Meta-analysis of HRT was performed using software RevMan 5.3. Pooled outcomes were determined by the fixed effect model.

Running Enrichment analysis

An enrichment score [7] was used to determine whether samples with CtBP amplifications or BRCA1/2 mutations were enriched among tumors with high mutation burden (SFig.6E), and to determine whether patient samples sensitive to chemotherapy were enriched among tumors with high expression of CtBP (SFig.6H). The enrichment score was defined by the maximum running score across all the samples. Then, we permuted the samples 10,0000 times to generate a background distribution for enrichment score (SFig.6F) to determine the significance.

Annexin V-FITC apoptosis detection

Cells were seeded into 12-well plates. The cells were treated with indicated conditions. Cells were collected and washed with PBS twice, and then stained with fluorescein isothiocyanate (FITC)-conjugated Annexin V and then PI. Cells were quantified by a flow cytometry (BD ACCURI C6).

1. Domcke, S., et al., *Evaluating cell lines as tumour models by comparison of genomic profiles.* Nat Commun, 2013. **4**: p. 2126.
2. Barretina, J., et al., *The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity.* Nature, 2012. **483**(7391): p. 603-7.
3. Lau, K.M., S.C. Mok, and S.M. Ho, *Expression of human estrogen receptor-alpha and -beta, progesterone receptor, and androgen receptor mRNA in normal and malignant ovarian epithelial cells.* Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5722-7.
4. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome Biol, 2009. **10**(3): p. R25.
5. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS).* Genome Biol, 2008. **9**(9): p. R137.
6. Zhu, L.J., et al., *ChIPpeakAnno: a Bioconductor package to annotate ChIP-seq and ChIP-chip data.* BMC Bioinformatics, 2010. **11**: p. 237.
7. Mootha, V.K., et al., *PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.* Nat Genet, 2003. **34**(3): p. 267-73.

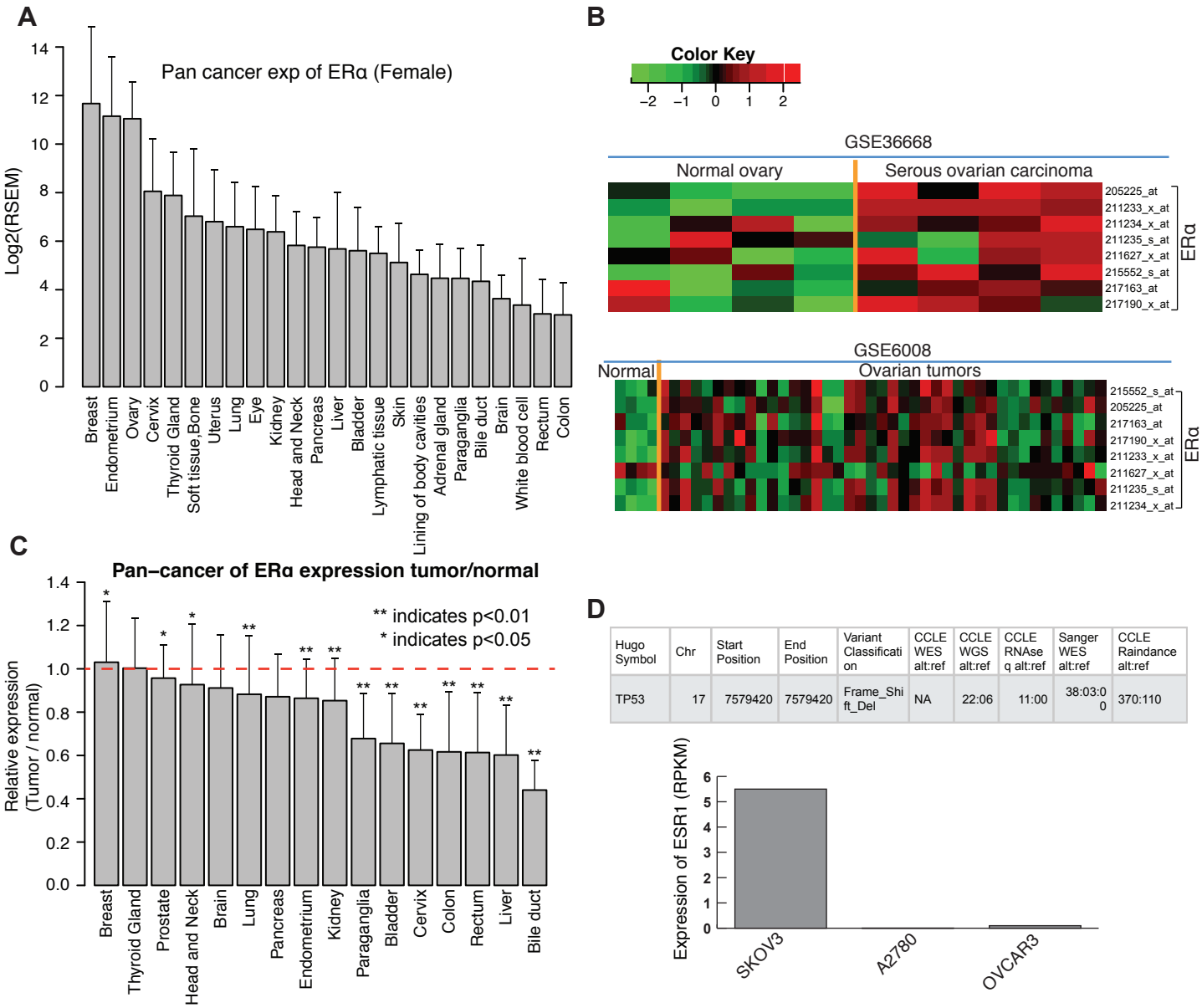

Figure S1. Characterization of ESR1 Expression in human cancer. **A**. Cancer-specific mRNA expression of ER α in TCGA RNA-seq datasets. RNA-seq data were obtained from cBioportal for each of the cancer types shown by their corresponding tissue name. Error bar shows the standard deviation of ESR1 expression across tumors. **B**. Heatmaps showing the expression of ESR1 in ovarian cancers and normal ovary epithelia of different GEO datasets. Rows of heatmaps correspond to the values of the probe sets used in the microarray platform for detecting ESR1 expression. To get a statistial result, student T-test is used to compare the average expression of different probes between normal samples and tumors ($P < 0.05$ for both datasets). **C**. Pan-cancer analysis of ESR1 relative expression compared to normal control. Only cancer types where the RNA-seq data of both cancer and normal control samples are available from TCGA were analyzed. Error bar shows the standard deviation of the relative expression. **D**. CCLE records about the TP53 mutation of SKOV3 and the expression of ESR1 in SKOV3, A2780 and OVCAR3. Although TP53 mutation is not detected by CCLE WES, four independent resources reported the same frame shift deletion of TP53, suggesting SKOV3 is a *bona fide* TP53 mutated cell line.

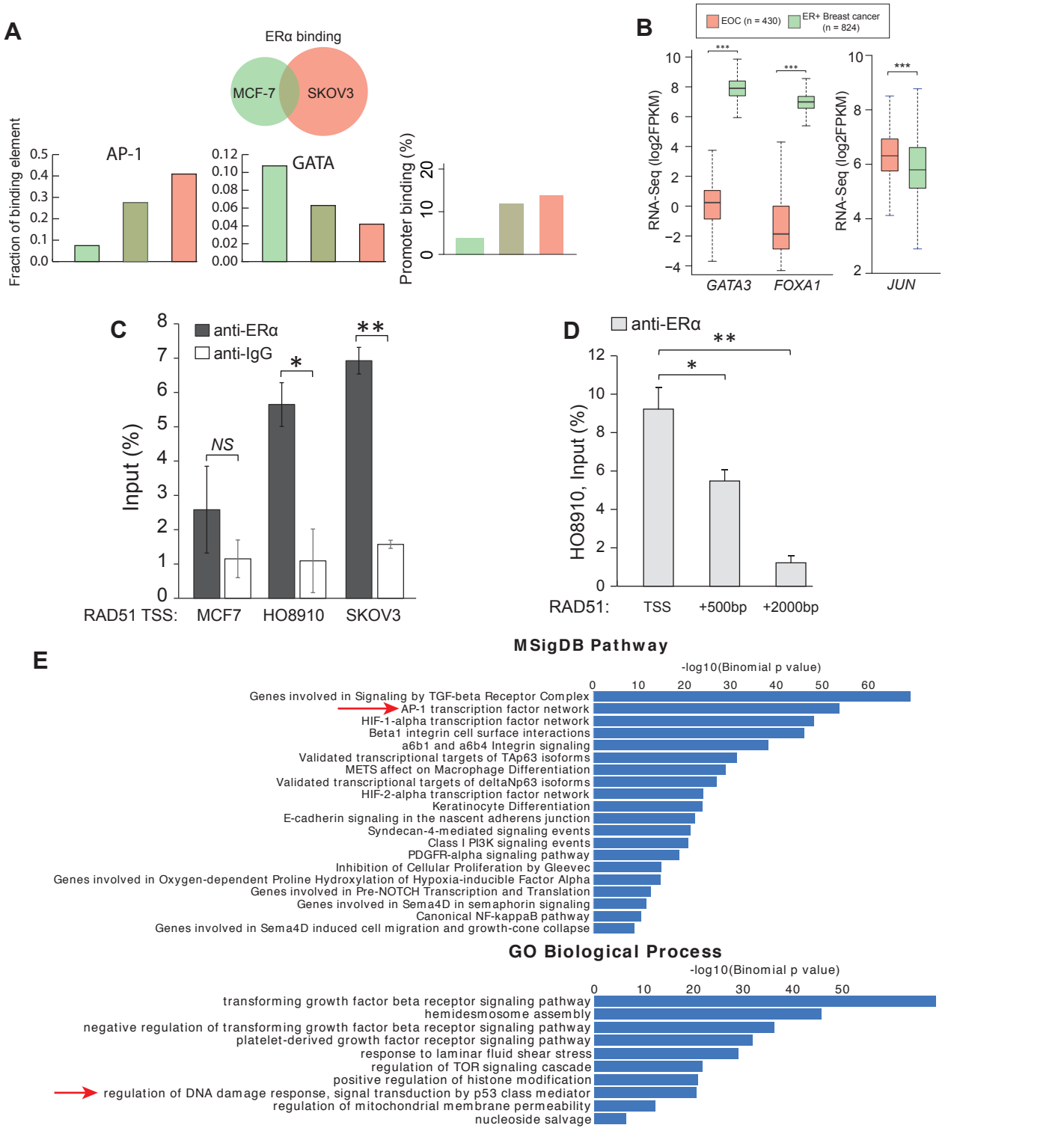

Figure S2. Characteristics of ERα binding profile. **A**. The fraction of each category of bindings that are matched with an AP-1 motif, a GATA motif or a gene promoter (+2kb to -500bp). **B.** Validation for the expression difference of GATA3, FOXA1 and c-Jun between TCGA breast cancers and EOCs using RNAseq data. Significance is determined by Wilcoxon rank sum test. **C.** ChIP-qPCR results testing the binding at the TSS of RAD51 in SKOV3, HO8910 and MCF7 cells. The data represent the mean and standard error of the three independent results, normalized to the corresponding input DNA. **D.** ChIP-qPCR results testing the binding at the TSS of RAD51, 500bp upstream of the TSS and 2000bp upstream of TSS in HO8910 cells. **E**. Enrichment analysis of SKOV3 ERα bindings by genomic regions enrichment of annotations tool (GREAT). The single nearest gene of bindings is used for function enrichment.

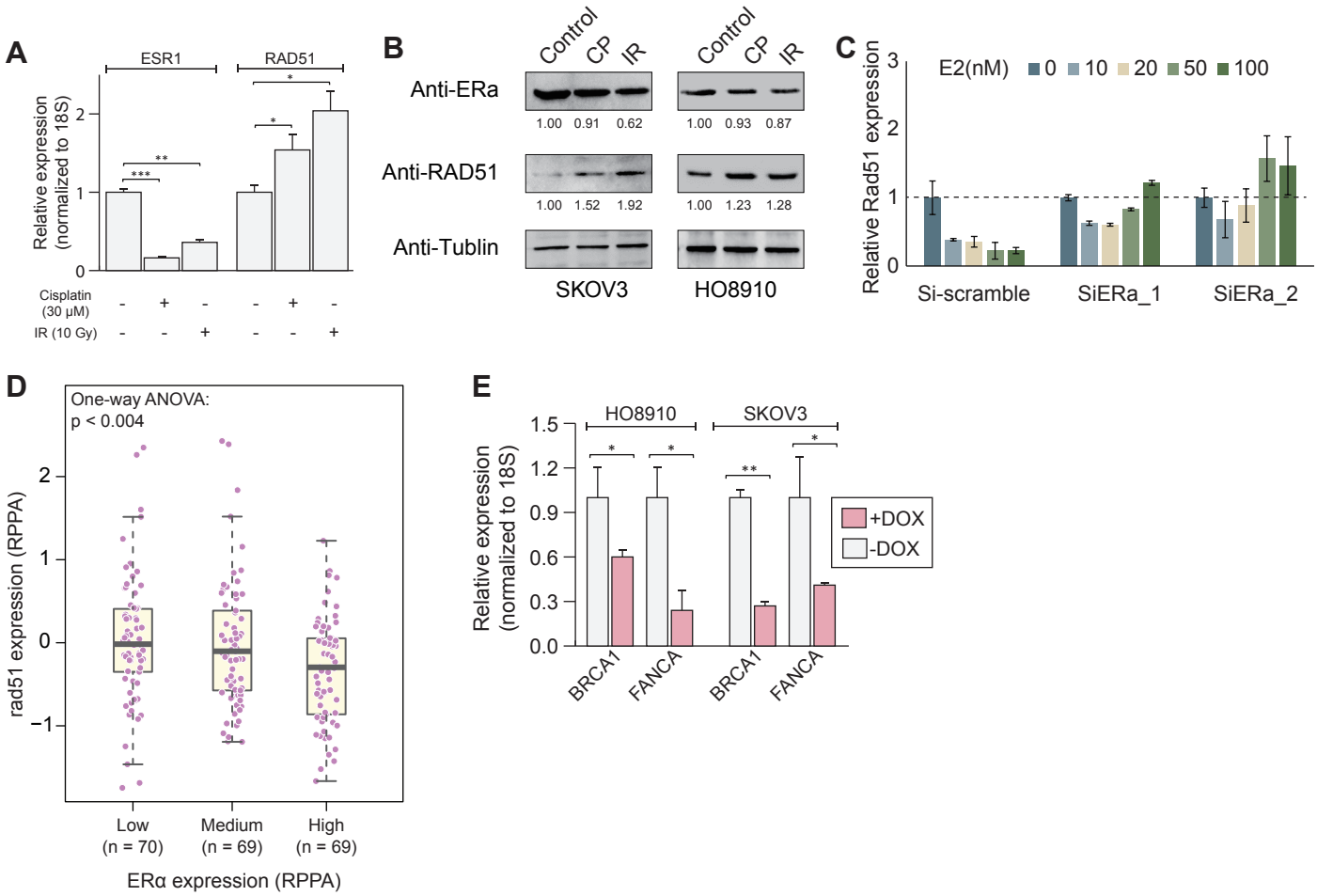

Figure S3. Correlation between ERα and HRR genes. **A**. Expression of *ESR1* and *RAD51* under the treatment of cisplatin or irradiation (IR) in HO8910 cells. **B**. Western blot for protein expression change of ERα and RAD51 under the treatment of cisplatin or irradiation (IR). **C**. Relative expression of RAD51 in SKOV3 cells exposed to gradient estrogen (0nM, 10nM, 20nM, 50nM, and 100nM)-containing medium when ERα is knocked down. Values are normalized to E2 (0 nm) in each group. **D.** Negative correlation between ERα and RAD51 (Spearman rho = -0.25, $p < 10^{-6}$). TCGA HGSOCs are divided into three equal groups by the expression value of ERα. Protein expression detected by reverse phase protein array (RPPA) of TCGA EOC dataset is used. One-way ANOVA is used to determine the significance of the difference of rad51 expression among different groups. **E.** qPCR validation of BRCA1 and FANCA expression comparing between ERα overexpressing SKOV3 cells (DOX+) with control cells (DOX-). Relative expression compared to -DOX and the Std. of 3 replicates are shown.

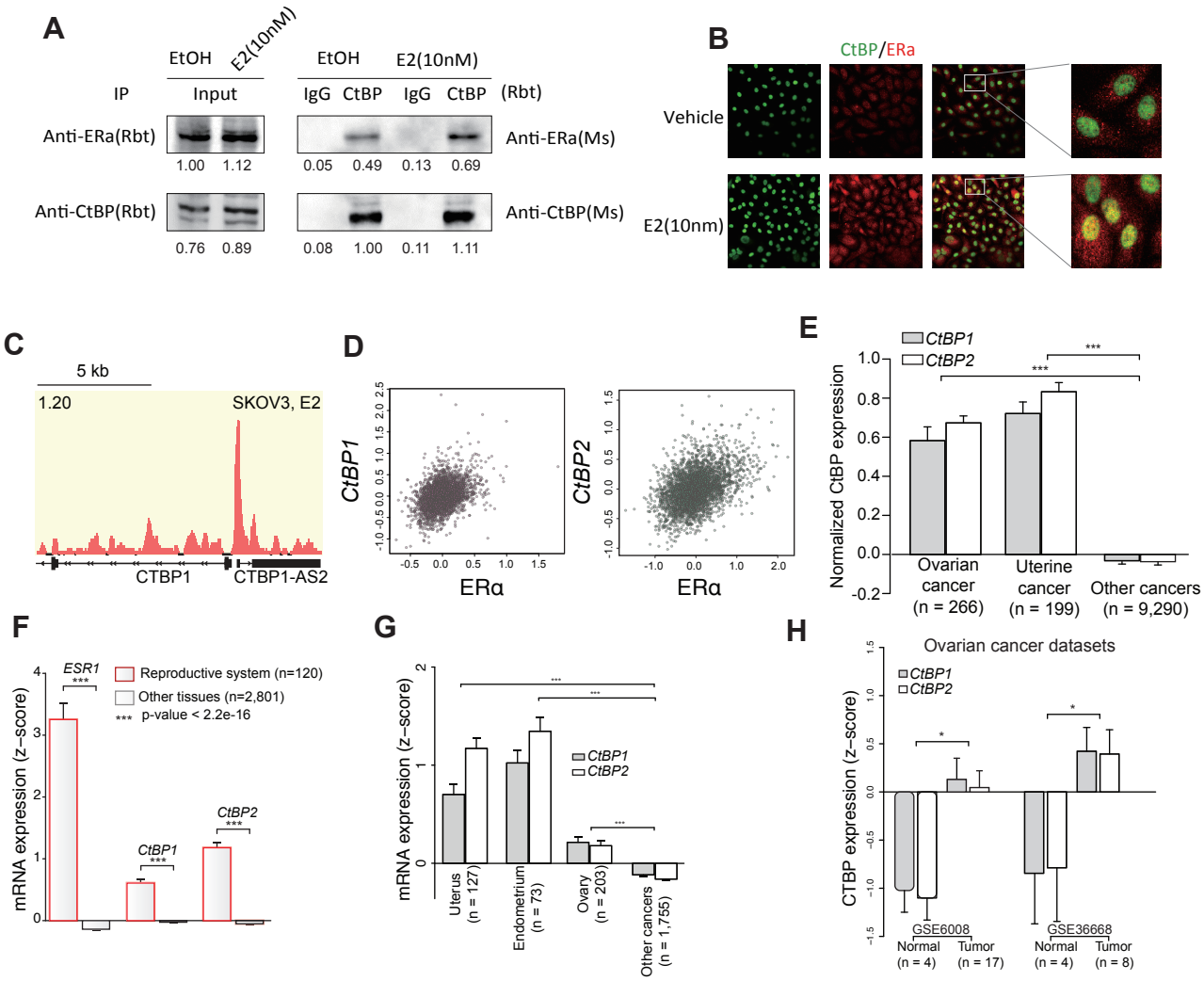

Figure S4. A). Results from co-IP experiments in SKOV3 cells. Nuclear extractions are from E2- and EtOH-treated SKOV3 cells. Immunoprecipitation is performed with a mouse derived-IgG or anti-CtBP antibody, followed by western blotting detection using rabbit-derived ERa and CtBP antibodies. **B)**. Co-localization of ERα and CtBP. SKOV3 cells were seeded in chamber and cultured in phenol red-free DMEM medium supplemented with 10% charcoal-treated FBS for 48h. Cells were then treated with vehicle (veh) or E2 (10 nM) for 24h. Then cells were fixed for IF staining. Images were taken on a confocal microscopy. **C).** ERα binding at the promoter of CtBP1 locus in E2-treated SKOV3 cells. **D).** Coexperssion between CTBP and ESR1 in integrated gene co-expression database COXPRESdb. E). Normalized expression of CtBP of TCGA pan-cancer datasets. Expression values of CtBP has been Z-score transformed. **F).** Expression of ESR1 and CtBP in tissues of women reproductive system vs. other tissues, according to the data of GTEx project. **G).** Same as Figure 3B but using data from ExpO consortium. H). Upregulation of CtBP in EOC compared to normal controls in multiple EOC datasets.

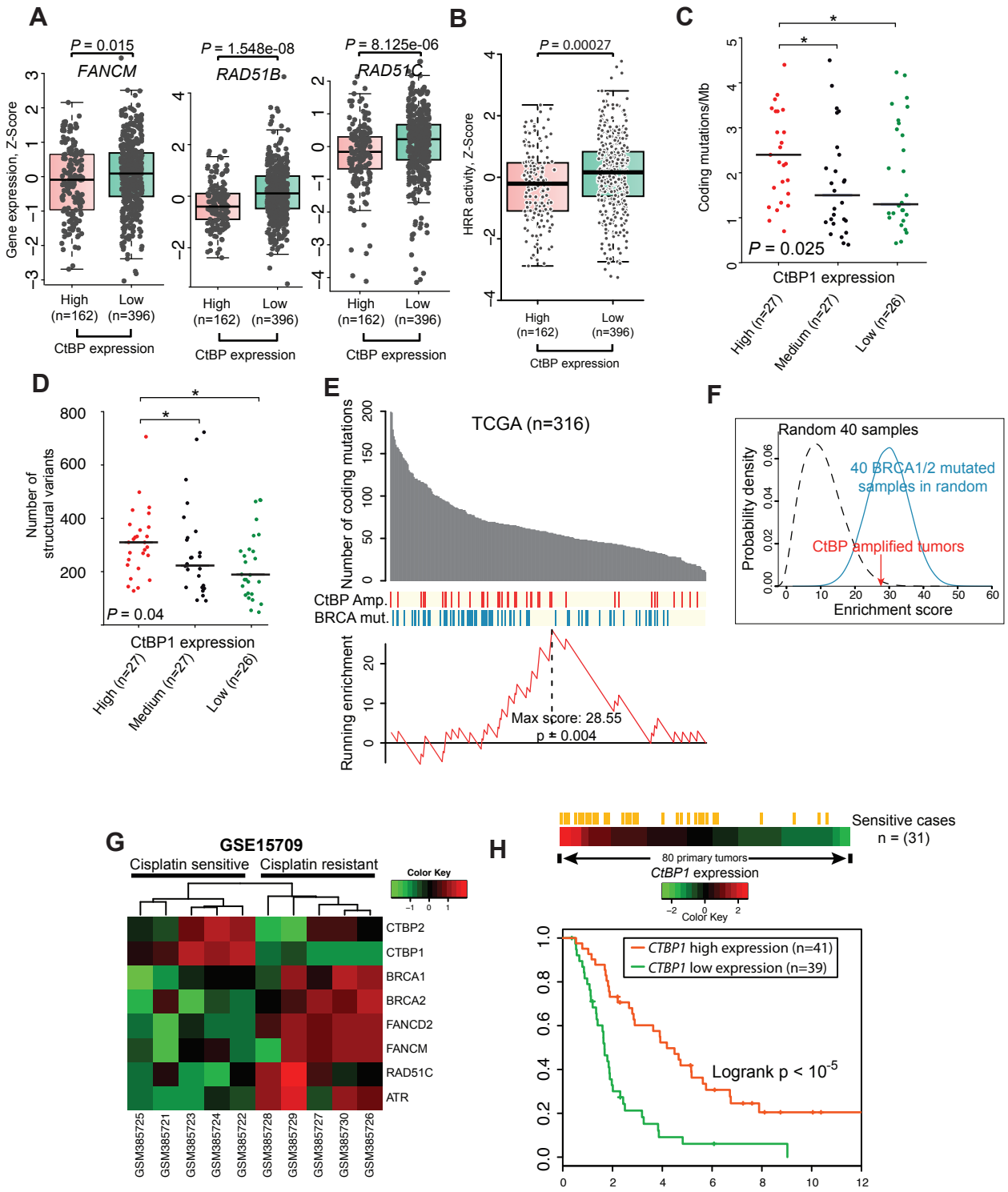

Figure S5. A). Examples for expression of HRR genes grouped by CtBP gene expression (CtBP1 high & CtBP2 high versus other cases). U133 array data of 568 TCGA EOCs were analyzed. Each dot represents the normalized expression of an HRR gene in a tumor. P values are given by Mann-Whitney test. **B**). HRR activity is a combined z-score derived from the expression of core HRR genes, inluding *BRCA1, BRCA2, FANCD2, FANCM, RAD51C, RAD51B, PTEN, ATR*. Alteration of these genes represents most known HRR defects of EOC. **C**). Coding mutations per megabase for samples grouped by CtBP1 gene expression level. The Solid lines indicate the median. Kruskal-Wallis test P value is shown in the figure. (* denotes P <0.05, by Mann-Whitney test). Only mutations that are covered by at least 30 reads and at least 30% of these reads contain the mutant allele are included in this analysis. **D**) Same as C) but for the number of structural variants per tumor. **E**) Enrichment score of hypermutated samples in CtBP amplificated tumors (n=40). The top panel shows the total number of coding mutations of 316 TCGA EOCs in a decreasing order. The middle panel indicates the EOCs with CtBP amplification or BRCA1/2 mutation. The bottom panel shows the running enrichment score. The empirical p value is determined by comparing to random 40 samples from the pool of 316 sequenced EOCs. **F**) Distributions of enrichment score of 40 random samples and 40 BRCA1/2 mutated samples in random. A pool of 70 samples with germline or somatic mutations of BRCA1/2 are used. The enrichment score of CtBP amplification is indicated by the arrow. **G**) Heatmap showing the expression of CtBP and its targeting HRR genes in cisplatin-sensitive and cisplatin-resistant A7880 cells in a public dataset. The cisplatin-resistant cells were clonally derived from the cisplatin-sensitive A2780 cells after rounds of drug treatment. Five biological replicates for each of the two cell types were automatically clustered into a cisplatin-sensitive group and a cisplatin-resistant group. **H**) The association of CTBP1 with the response to chemotherapy and clinical outcome in 80 EOC patients (OV-AU dataset). The same dataset in A) and B) is analyzed here. The association between the expression of CTBP1 and sensitivity to platinum-based chemotherapy is significant (p<0.0001), as empirically determined by comparing the enrichment score with random perturbations of samples. No significant correlations are observed for CTBP2 in this dataset.

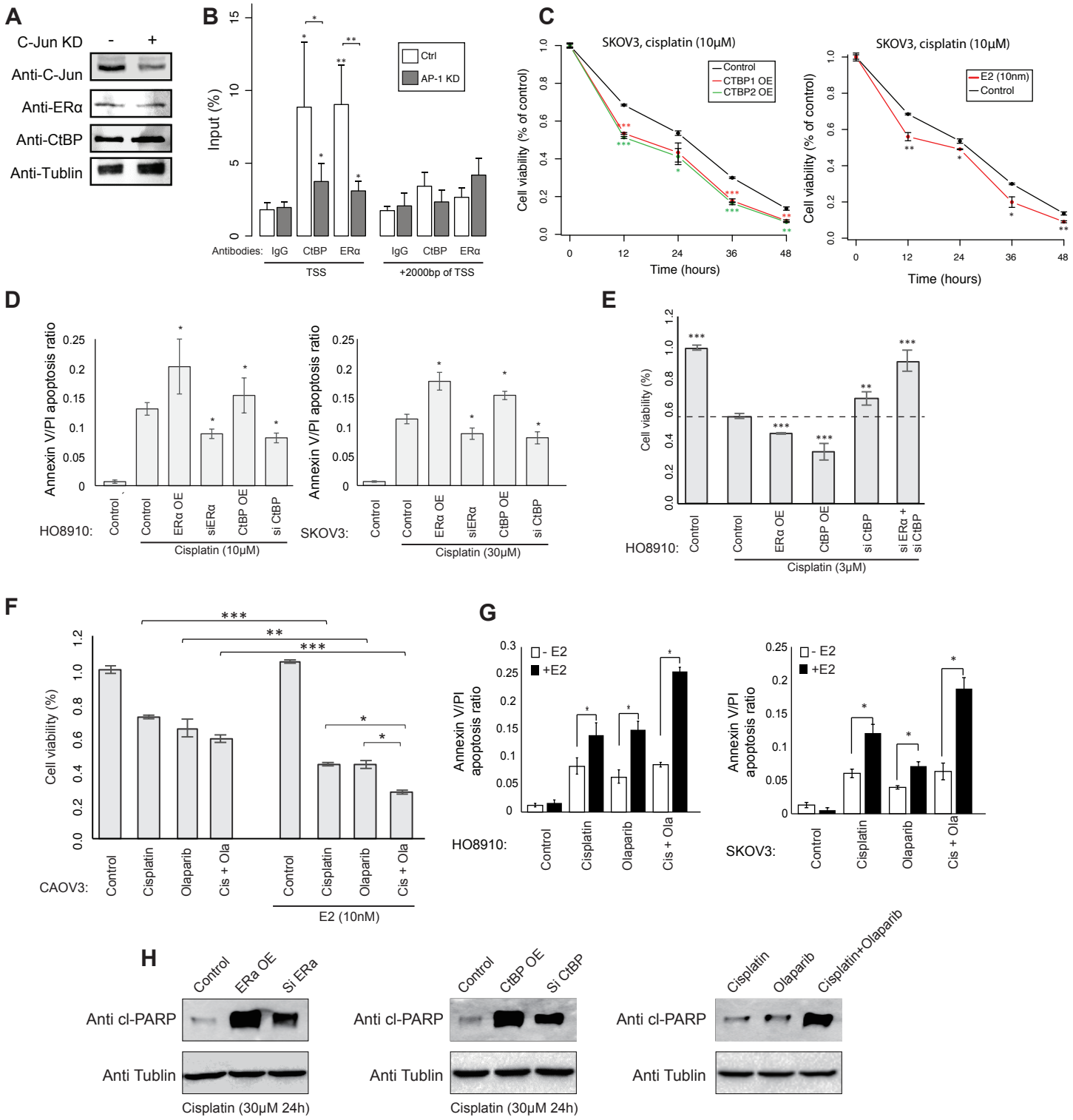

Figure S6. CtBP- and ERα-mediated response to chemotherapy drugs. **A**. Western blot showing the expression of CtBP, ERα, C-Jun and tublin with knockdown of C-Jun by lentivirirus delivered RNAi in SKOV3 cells. **B.** ChIP-qPCR valiation for AP-1-mediated CtBP and ERα binding at the TSS of RAD51 in CAOV3 cells. Unspecified arsterisks indicate the significance comparing with corresponding IgG. **C.** Cisplatin-dependent survival of SKOV3 cells, CtBP-overexpressed SKOV3 cells and SKOV3 cells treated with E2 (10nM). The measurement is the cell viability using the WST assay. The time indicates the time point when cisplatin was used to treat the cells after plating. **D**. Apoptosis Assay using Flow Cytometry after Annexin V-FITC/Propidium Iodide (PI) staining cells treated with indicated conditions (HO8910: CP 10μM HO8910 for 24h; SKOV3: 30μM for 24h). **E.** Cell viability assay for HO8910cells transfected with indicated vector or siRNA, after the treatment of cisplatin (3μM) for 71h. **F.** Cell viability assay for CAOV3 cells after different treatments as indicated for 24h (cisplatin: 3μM; olaparib: 10μM; E2: 10nM). **G**. Apoptosis Assay using Flow Cytometry after Annexin V-FITC/Propidium Iodide (PI) staining cells treated with indicated conditions (HO8910: CP 10 μM HO8910 for 24h; SKOV3: 30μM for 24h). **H**. Westernblot showing the cleaved PARP protein level when SKOV3 cells treated with indicated conditions for 24h (Cisplatin: 30μM; Olaparib: 10μM).

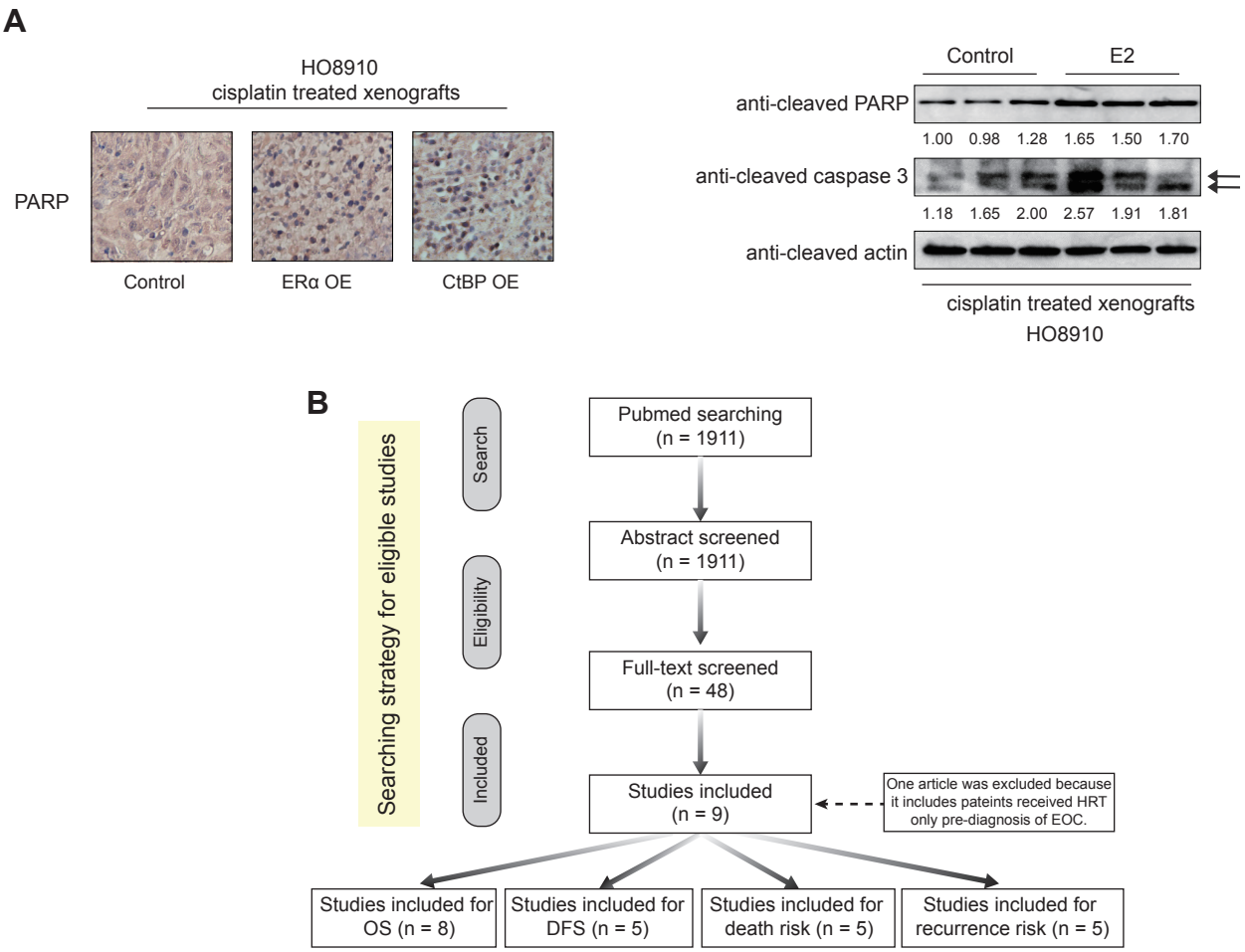

Figure S7. Effects of HRT in xenografts and EOC patients receiving chemotherapy. **A**). Left, representative images of IHC staining of PARP of xenografted tumors of HO8910 cells. Right, measurement of cleaved PARP and Caspase 3 in xenografted tumors of HO8910 cells by western blot. **B**). The flowchart of the searching strategy for eligible studies. **C** and **D**). Forest plots for meta-analysis of death risk (**C**) and recurrence risk (**D**). The analysis was performed using software RevMan 5.3. CIs are set at 95% and shown as horizontal lines. Solid vertical lines indicate the no-difference point between HRT and control group. Pooled odds ratios are determined by the Fixed effect model and shown by diamonds.

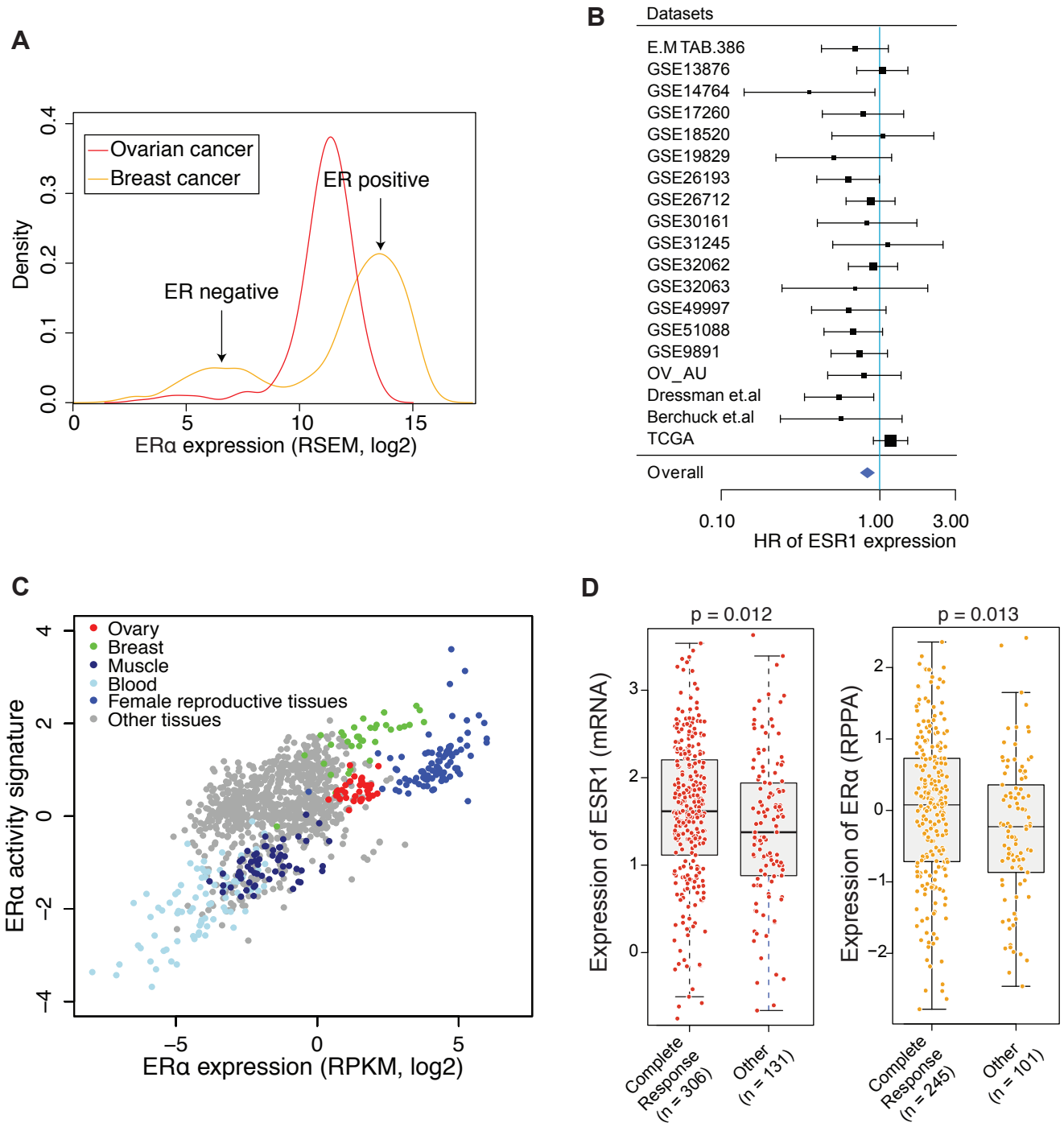

Figure S8. The prognostic value of ERα in EOC. **A**. The distribution of ESR1 expression for EOC and breast cancer of TCGA datasets. **B.** Forest plot visualing the hazard ratios (HRs) of ESR1 expression in each dataset evaluated by the multivariate Cox proportional hazards model. Squares show HR estimates of gene expression. The sizes of square are determined by the weights in meta-analysis summaries. Segments show 95% CIs, and the blue diamonds show the fixed-effects meta-analysis summaries of HRs over all the datasets. **C.** Scatter plot showing the correlation between ERα expression and ERα activity across human tissues (female), using the data from GTEx project. **D**. The expression of ESR1 and ERα in TCGA datasets. EOCs are divided into the group having complete response to chemotherapy and others (patial response and progressive or stable disease). Agilent microarray data are used for the expression of ESR1 to increase the number of samples. Significance is determined by the Wilcox rank sum test.