Chromatin topology defines estradiol-primed Progesterone Receptor and PAX2 binding in endometrial cancer cells

Alejandro La Greca¹, Nicolás Bellora^{3,6†}, Francois Le Dily^{2,5†}, Rodrigo Jara¹, Ana Silvina Nacht^{2,5}, Javier Quilez Oliete², José Luis Villanueva², Enrique Vidal², Gabriela Merino⁴, Cristóbal Fresno⁴, Inti Tarifa Rieschle¹, Griselda Vallejo^{1,6}, Guillermo Vicent², Elmer Fernández^{4,6}, Miguel Beato^{2,5}, Patricia Saragüeta^{1,6*}

*For correspondence:

patriciasaragüeta2@gmail.com (PS)

†These authors contributed equally to this work

¹ Biology and Experimental Medicine Institute, IBYME-CONICET, Buenos Aires, Argentina; ² Centre for Genomic Regulation (CRG), Barcelona Institute for Science and Technology, Barcelona, Spain; ³ Institute of Nuclear Technologies for Health, INTECNUS-CONICET, Bariloche, Argentina.; ⁴ Bioscience Data Mining Group, Córdoba University, Córdoba, Argentina; ⁵ Universitat Pompeu Fabra (UPF), Barcelona, Spain; ⁶ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina

Abstract Estrogen (E2) and Progesterone (Pg), via their specific receptors (ERalpha and PR), are major determinants in the development and progression of endometrial carcinomas. However, their precise mechanism of action and the role of other transcription factors involved are not entirely clear. Using Ishikawa endometrial cancer cells, we report that E2 treatment exposes a set of progestin-dependent PR binding sites which include both E2 and progestin target genes. ChIP-seq results from hormone-treated cells revealed a non-random distribution of PAX2 binding in the vicinity of these estrogen-promoted PR sites. Altered expression of hormone regulated genes in PAX2 knockdown cells suggests a role for PAX2 in fine-tuning ERalpha and PR interplay in transcriptional regulation. Analysis of long-range interactions by Hi-C coupled with ATAC-seq data showed that these regions, that we call "progestin control regions" (PgCRs), exhibited an open chromatin state even before hormone exposure and were non-randomly associated with regulated genes. Nearly 20% of genes potentially influenced by PgCRs were found to be altered during progression of endometrial cancer. Our findings suggest that endometrial response to progestins in differentiated endometrial tumor cells results in part from binding of PR together with PAX2 to accessible chromatin regions. What maintains these regions open remains to be studied.

Introduction

Progesterone (Pg) is a key regulator in the female reproductive tract, including uterine and mammary gland development (Lydon *et al.*, 1995). Endometrial and breast tissues exhibit significantly different responses to hormones, resulting in very distinctive morphologies and functions. During pregnancy, Pg prepares the uterine epithelium to receive the embryo and initiates the process of differentiation of stromal cells towards their decidual phenotype. In the mammary gland and in co-

ordination with prolactin, Pg stimulates epithelial proliferation and differentiation of alveolar lobes in the mammary gland (*Mulac-Jericevic and Conneely, 2004*). Unlike Pg, estradiol (E2) is the main proliferative signal in the uterine epithelium and exerts its function through activating estrogen receptor (ER) alpha and beta (ERalpha and beta, respectively) (*Ishiwata et al., 1997; Kayisli et al., 2004*).

The physiological role of Pg is mediated by the interaction and consequent activation of isoforms A (PRA) and B (PRB) of the progesterone receptor (PR), which are transcribed from alternate promoters of the gene (*Hovland et al., 1998*). While PRA is more abundant in stromal endometrial cells, PRB is the most representative isoform in epithelial cells of endometrium. Steroid hormones exert their transcriptional effects through binding of the steroid receptors (SR) to specific DNA sequences in the promoters or enhancers of target genes known as "hormone response elements" (HRE). Estradiol exposure triggers ER binding to estrogen response elements (ERE) regulating target genes such as *PGR*. Previous work showed E2-dependent upregulation of PR in many different target cells, species and pathological conditions (*Graham et al., 1995; Kraus and Katzenellenbogen, 1993*). Exposure to progestins triggers binding of PR to PRE. Once bound to their HREs the hormone receptors interact with other transcription factors, co-regulators (*Beato et al., 1995*), such as the p160 family of co-activators of steroid receptors SRC-1-3, and chromatin remodelling enzymes. This evidence favors tissue specific roles of PR isoforms and their co-regulators orientated towards differential transactivation of target genes.

High levels of PRA and PRB have been described in endometrial hyperplasia (*Miyamoto et al., 2004*) while low and high-grade endometrial cancers reveal reduced or absent expression of one or both isoforms in epithelia or stroma (*Shao, 2013*). This PR decrease is often associated with shorter progression-free survival and overall survival rates (*Leslie et al., 1997; Miyamoto et al., 2004; Sakaguchi et al., 2004; Jongen et al., 2009; Kreizman-Shefer et al., 2014*). The absence of PR gene expression may be attributed to hypermethylation of CpG islands within the promoter or first exon regions of the PR gene or to the presence of associated deacetylated histones. These modifications were reported for endometrial cancer cell lines as well as tumor samples and may be exclusive to PRB (*Sasaki et al., 2001; Xiong et al., 2005; Ren et al., 2007*). Treatment of such cells with DNA methyltransferase or histone deacetylase inhibitors can restore both PRB expression and its regulation of target genes such as *FOXO1*, p21 (*CDKN1A*), p27 (*CDKN1B*), and cyclin D1 (*CCND1*) (*Xiong et al., 2005; Yang et al., 2014*). Down-regulation of PR by post-transcriptional mechanisms and through post-translational modifications of PR may contribute to progesterone resistance in endometrial cancer but have not been extensively explored in the context of endometrial cancer. It is known that oncogenic activation of KRAS, PI3K or AKT and/or loss of functional tumor suppressors such as *PTEN* are common genetic alterations (*Hecht and Mutter, 2006*), together with *ARID1A* (*Liang et al., 2012*), all of them observed in endometrial cancer. Although there are numerous reports of hormonally regulated enhancers and super-enhancers in mammary cancer cells (see in dbsuper-enhancer, <http://bioinfo.au.tsinghua.edu.cn/dbsuper/>) (*Khan and Zhang, 2016; Hnisz et al., 2015*), there is a void of information about their presence in endometrial cells.

To better understand the response to progestin in endometrial cancer cells, we have studied the genomic binding of ER and PR, the global gene expression changes and the state of chromatin by ATACseq as well as the genomic interactions by Hi-C in Ishikawa cells exposed to progestin or estrogen, and also in cells exposed to progestin after a period of estradiol pretreatment. Inside TADs with progestin regulated genes, we identified regions that we named "progestin control regions" (PgCRs) that correlate with the open chromatin compartment independently of hormonal stimuli and include binding sites for the partner transcription factor PAX2.

Results

Ishikawa endometrial epithelial cells respond to R5020 through activation of PR, whose levels increase upon exposure to E2

Endometrial epithelial cells respond to ovarian steroid hormones -progesterone (Pg) and estradiol (E2)-, E2 being the main proliferative stimulus and Pg its antagonist. After treating Ishikawa cells with E2 10 nM for 48h we observed an increment in number of cells compared to vehicle (OH) (FC 1.78 ± 0.08 v. OH) that was suppressed by addition of R5020 10 nM (FC 1.15 ± 0.08 v. OH) (**Figure 1A**). Treatment with R5020 10 nM alone did not induce proliferation on Ishikawa cells (FC 0.77 ± 0.08 v. OH) (**Figure 1A**; **Figure 1-source data 1**). E2-induced cell proliferation was also abrogated by pre-incubation with estrogen receptor (ER) antagonist ICI182780 1 μ M (ICI 10×10^{-6} M) (FC 1.05 ± 0.05 v. OH) (**Figure 1-Figure Supplement 1A**), but not by pre-incubation with PR antagonist RU486 1 μ M (RU486 10×10^{-6} M) (FC 1.42 ± 0.07 v. OH) (**Figure 1-Figure Supplement 1B**), proving that ER but not PR was directly involved in the proliferative response to E2. Suppression of E2-induced cell proliferation by R5020 was inhibited by pre-incubation with RU486 (FC 1.50 ± 0.06 v. OH), indicating that R5020 effect was mediated by PR in Ishikawa cells (**Figure 1-Figure Supplement 1B**). The effects of E2 and R5020 on proliferation were corroborated by BrdU incorporation and cell cycle phase analysis 18h after hormone exposure (**Figure 1-Figure Supplement 1C** and **D**). E2 increased the number of BrdU positive cells and percentage of cells in S phase compared to untreated control cells and to cell exposed to the vehicle (OH), and these increments were inhibited by R5020. Treatment with the histone deacetylase inhibitor Trichostatin A 250 nM (TSA 250 nM) was used as negative control for BrdU incorporation and cell cycle progression (**Figure 1-Figure Supplement 1C**).

Ishikawa cells contain isoforms A and B of PR (PRA and PRB), both of which increased their steady state levels by treating cells with E2 10 nM for 12h (**Figure 1-Figure Supplement 1E** and **F**). Pretreating cells with E2 for 12h (preE2) followed by R5020 had little effect on the proliferative response (**Figure 1A**), while E2 pre-treatment for 48h significantly increased the proliferative effect of E2 exposure compared to non-pretreated cells (FC 1.47 ± 0.08 v. no preE2). The percentage of cells exhibiting nuclear localization of PR increased upon E2 pretreatment prior to R5020 exposure (T0). Upon exposure to R5020 for 60min the percentage of cells exhibiting nuclear PR was not affected by E2 pretreatment, though the intensity of the fluorescence signal increased in E2-pretreated cells (**Figure 1B**; **Figure 1-source data 2**). Ishikawa cells express considerably higher levels of ERalpha than of ERbeta (**Figure 1-Figure Supplement 1G**), suggesting that the proliferative effect of E2 was mediated by ERalpha. R5020 increased nuclear ERalpha, suggesting a functional PR-ER crosstalk in response to hormonal stimuli (**Figure 1-Figure Supplement 1H**). Such interactions have already been proven in breast cancer T47D cells (**Ballaré et al., 2003**) and in U118 rat endometrial stromal cells (**Vallejo et al., 2005**), though in the latter PR remains strictly cytoplasmic.

Treatment with hormones during 12h produced transcriptomic changes consistent with the physiological stages of normal cycling endometrial tissue (**Chi et al., 2020**). RNAseq results from Ishikawa cells (**Figure 1-source data 3**) exposed to E2 10 nM for 12h showed a significant resemblance to proliferative endometrium (**Figure 1C**), while 12h treatments with R5020 10 nM regulated a gene expression profile similar to a mid-secretory phase (**Figure 1D**). In line with these findings, among the top overrepresented biological processes for E2-treated Ishikawa cells showed angiogenesis and positive regulation of smooth muscle cell proliferation and for R5020-treated cells processes like protein targeting to Golgi and SRP-dependent cotranslational protein targeting to membrane were found (**Figure 1-Figure Supplement 2A**). In addition, the majority of regulated genes (81% of R5020 and 63% of E2) were not shared by both hormones (**Figure 1-Figure Supplement 2B**). Genes like *PGR* (progesterone receptor) and cell-cycle regulator *CCND2* (cyclin d2) were upregulated by E2 but not by R5020, while *TGFA* (transforming growth factor alpha) was upregulated by both hormones under these single-hormone treatments (**Figure 1-Figure Supplement 2B** and **C**).

Compared to T47D cells, PR protein levels in Ishikawa cells were significantly lower (**Figure 1-**

Figure 1. R5020 inhibits E2-induced Ishikawa cell proliferation through an active PR that is capable of transactivating an exogenous MMTV promoter sequence and an endogenous enhancer sequence located 50kb upstream of *EGFR* gene. (A) Proliferation of Ishikawa cells either pretreated with E2 10nM for 12h (preE2) or not (no preE2) and later treated with vehicle (OH), E2 10 nM (E2), R5020 10 nM (R5020), E2 combined with R5020 (E2+R5020) and FBS (10%FBS), expressed as mean number of cells ± SE of three independent experiments. (***) p < 0.001. (B) Immunofluorescence of PR in untreated (T0; top left), 60min R5020-treated (R5020; bottom left), 12h E2-pretreated (top right) and 12h E2-pretreated 60min R5020-treated (bottom right) Ishikawa cells. Scale bar is equivalent to 30 μm. Mean nuclear signal of PR for every cell in all images was determined and shown to the right of the images as arbitrary units (log a.u.). Horizontal dashed lines in boxplots indicate background signal for secondary antibody. (***) p < 0.001. (C and D) Gene set enrichment analysis (GSEA) results using R5020- and E2-treated Ishikawa expression profiles as phenotypes for classification of normal endometrium (proliferative and secretory) samples. Enrichment profile (green) shows correlation of normal samples at the top or bottom of a ranked gene list (phenotypes). Normalized enrichment scores (NES) and nominal p values (nom.p) are shown in the graphs. (E) Ishikawa cells transfected with an MMTV-Luciferase reporter gene and treated with vehicle (OH) and R5020 10 nM (R5020) for 18h. Diagram at the top depicts MMTV LTR promoter features, including several hormone response elements (HRE) and a nuclear factor 1 (NF1) binding site within nucleosome B (dark grey circle and magnification). Numbers in the diagram indicate base pair position relative to transcription start site (TSS). Results are expressed as relative units (r.u.) of Luciferase activity. (F) Representation of *EGFR* TSS and the enhancer sequence located 50kb upstream used to evaluate PR recruitment. Black arrows indicate position of qPCR primers employed on samples treated or not (0) with R5020 for 5, 30 and 60 min. Unspecific immunoprecipitation of chromatin was performed in parallel with normal rabbit IgG (IgG). Results are expressed as %input DNA and bars represent mean fold change in PR enrichment relative to time 0 (untreated cells) ± SE of two independent experiments. (*) p < 0.05.

Figure 1-Figure supplement 1. E2-dependent cell cycle progression and proliferation of Ishikawa cells is mediated by ERα. R5020 antagonism on E2 is mediated by PR.

Figure 1-Figure supplement 2. RNAseq results from hormone-treated Ishikawa cells.

Figure 1-Figure supplement 3. Ishikawa cells express six times less PR than T47D cells.

Figure 1-source data 1. Proliferation assays in treated and untreated Ishikawa cells. Measure represents number of cells x10⁵ in the table.

Figure 1-source data 2. Ishikawa mean nuclear signal intensity.

Figure 1-source data 3. Ishikawa treated and untreated normalized expression dataset.

Figure 1-source data 4. PR binding Ct values in *EGFR* enhancer sequence.

135 **Figure Supplement 3A** and B). Copy number variation analysis revealed that Ishikawa cells carry
 136 neither chromosomal alterations nor multiple copies of the PGR locus as was shown in T47D cells,
 137 possibly accounting for differences in PR content between both cell lines (**Figure 1–Figure Supple-**
 138 **ment 3C**; (Le Dily et al., 2014)). To support the functional activity of PR in Ishikawa cells, a reporter
 139 gene downstream of the MMTV promoter containing several Pg response elements (PREs) (**Figure**
 140 **1E**) was transiently transfected and the cells were exposed to either R5020 10nM or vehicle (OH) for
 141 20 hours. A 2.8 fold increment in Luciferase activity was found in cells exposed to R5020 relative
 142 to OH controls, supporting the capacity of ligand-activated PR to activate PREs (**Figure 1E**). Bind-
 143 ing of PR to cellular chromatin targets was confirmed by ChIP analysis of the PRE at the enhancer
 144 sequence located 50kb upstream from the transcription start site (TSS) of the *EGFR* (epidermal
 145 growth factor receptor) gene (**Figure 1F**; **Figure 1–source data 4**). Together, these observations
 146 demonstrate that PR is responsive to progestins and functional in Ishikawa cells under our cul-
 147 ture conditions and that these cells resemble E2 and Pg hormone regulation to proliferative and
 148 secretory normal endometrium respectively.

149 **Binding of PR and ERalpha to the Ishikawa endometrial cancer genome**

150 To explore the genome-wide distribution of PR and ERalpha binding (PRbs and ERbs respectively)
 151 in Ishikawa cells, ChIPseq was performed in different conditions (**Figure 2A** and **Figure 2–Figure**
 152 **Supplement 1**). First, we analyzed untreated cells (T0) and cells exposed for 5, 30, and 60min
 153 to 10nM R5020 using a specific antibody to PR that detects both isoforms, PRA and PRB. Results
 154 showed robust PR binding after 30 min of R5020 treatment (R5020 30min) with 1,446 sites, of
 155 which 322 sites (22%) were present in untreated cells (PRbs at time zero, T0=331). After 60min of
 156 treatment with R5020 (R5020 60min), the majority of sites identified at 30min were still evident
 157 (78%) (**Figure 2A** and **Figure 2–Figure Supplement 1A**). The representation of PREs in 22% of the
 158 PR binding sites that were lost between 30 and 60 min of R5020 treatment was analyzed taking
 159 into account common, and unique 30min or 60min PRbs. De novo motif discovery, analysis of
 160 information content and quantification occurrences of PRE motifs in such regions did not show
 161 differences in the information content (the strength of PRE motif), nor new motif different from
 162 PRE, but revealed a higher abundance of PREs in common and unique 60min datasets, yielding
 163 1.72 fold and 1.78 relative unique sites in 30min respectively. qPCR performed on six regions in
 164 the vicinity of hormone regulated genes and occupied by PR at 30 and 60min of R5020 exposure
 165 validated ChIPseq results (**Figure 2–Figure Supplement 1B**). These regions were selected according
 166 to differentially expressed genes from RNAseq data and top-ranked by peak signal. These results
 167 indicate that hormone-dependent PR occupancy increased 5-fold by 30min and stabilized between
 168 30 and 60min of treatment, in accordance with qPCR results (**Figure 2–Figure Supplement 1C**).

169 Next, we explored the recruitment of ERalpha to chromatin of Ishikawa cells exposed to E2
 170 (10nM) for 5, 60 and 360min. Poor ERalpha binding was detected at T0 (25 sites), of which 90%
 171 remained occupied throughout all times of treatment with E2. Exposure to E2 resulted in the de-
 172 tection of 178 ERalpha binding sites (ERbs) at 5min, 1,591 at 60min and 1,973 at 360min (**Figure 2A**
 173 and **Figure 2–Figure Supplement 1D**). The majority (85%) of ERbs found at 60min was also identified
 174 at 360min (**Figure 2–Figure Supplement 1D**). ERalpha binding at 0, 60 and 360min of E2 treatment
 175 was confirmed by qPCR on four of the sites identified (**Figure 2–Figure Supplement 1E**). ChIPseq
 176 results point to a clear and sustained E2-dependent enhancement of ERalpha binding (**Figure 2–**
 177 **Figure Supplement 1F**).

178 De novo motif discovery confirmed that PR binding occurred mostly through PREs exhibiting
 179 the complete palindromic response elements (**Figure 2–Figure Supplement 1G**), while ER binding
 180 sites were enriched in half-palindromic ERE motifs (**Figure 2–Figure Supplement 1H**). Comparison
 181 with previous findings in T47D cells (Ballaré et al., 2013) enabled clustering of both PRbs and ERbs
 182 into two classes (**Figure 2–Figure Supplement 1G** and **H**, respectively): sites specific for Ishikawa
 183 cells (group I; 595 PRbs, group III: 1101 ERbs) and sites present in both Ishikawa and T47D cell lines
 184 (group II: 896 PRbs; group IV: 490 ERbs). Classification revealed that PR binds through complete

Figure 2. Estradiol induces R5020-dependent PR binding to specific regions in chromatin. (A) Upper table shows total number of PRbs obtained by ChIPseq for untreated (0min) and R5020-treated (5, 30 and 60min) endometrial Ishikawa cells under three different conditions: non-pretreated with E2 (PR), pretreated with E2 for 12h (prePR) and exogenous expression of PR (FPR). Lower table shows number of ERbs using anti-ERalpha antibody, untreated (0min) and E2-treated (5, 60 and 360min) Ishikawa cells. Venn Diagram shows shared binding sites among PRbs (red), prePRbs (blue) and ERbs (green) at 60min. (B) Venn Diagram shows intersection between ERbs (green), FPRbs (dark grey) and prePRbs (blue) at 60min. (C) Fraction of peaks in FPR and prePR after subtraction of shared PRbs (FPR_s and prePR_s, respectively) that are not shared with each other (unique), that are common to each other (common) and that are common with ER (ERbs). (D) Normalized coverage of PR and ERalpha binding in untreated (T0) and 60min hormone-treated (R60 and E60) Ishikawa cells and PR binding in proliferative (GSE1327133) and secretory (GSE1327134) endometrium. Black arrow indicates peak of interest. R60: 60min R5020 10 nM; E60: 60min E2 10 nM. The three regions displayed include *TGFA*, *PGR* and *CCND2* genes (indicated at the top). An estrogen response element (ERE) and a half ERE are indicated below the peaks.

Figure 2-Figure supplement 1. Genome-wide analysis of PR and ERalpha binding in Ishikawa cells.

Figure 2-Figure supplement 2. PR overexpression on Ishikawa cells.

PREs regardless of cell line identity, but in Ishikawa cells ERalpha binds mostly sites with only half of the characteristic palindrome.

Estrogenic environment defines the landscape for PR binding to the endometrial genome

Shifts in the synthesis and secretion of the ovarian steroids (estrogen and progesterone) during the menstrual cycle serve as the principal hormonal drivers for endometrial changes. Rising circulating estradiol during the mid-to-late follicular phase of the cycle promotes the proliferation of the functional endometrium, and higher E2 levels upregulate *PGR* gene expression (Graham et al., 1995; Kraus and Katzenellenbogen, 1993). A similar result was reported in Ishikawa cells treated with E2 (Diep et al., 2016). To explore the effect of E2 on PR binding to DNA we performed PR ChIPseq analyses on Ishikawa cells exposed to E2 10 nM for 12h (preE2) before treatment with R5020 for 30 and 60min. Pretreatment with E2 significantly increased the number of R5020-dependent PRbs (prePRbs), which included most of PRbs already identified in non-pretreated Ishikawa cells (Figure 2A, Table and Venn Diagram). Quantitative real-time PCR validations performed on 6 sites occupied by PR confirmed positioning of the receptor in both non-preE2 (non E2 pre-treatment) and preE2 conditions (Figure 2–Figure Supplement 2A). It also showed that E2 pretreatment augments both number of PRbs and occupancy of the receptor (signal) (Figure 2–Figure Supplement 2B). Contrary to PRbs in non-pretreated cells, the number of PRbs doubled between 30 and 60min of R5020 in preE2 cells, reaching 5,701 sites (Figure 2A). All conditions replicated their binding profiles in an independent experimental setting (Figure 2–Figure Supplement 2C).

Sequencing experiments performed on T47D cells exposed to 10 nM R5020 revealed over 25,000 PRbs (Ballaré et al., 2013; Nacht et al., 2016), likely reflecting the high content of PR in these cells. However, a large proportion of these PRbs was considered functionally irrelevant as indicated by the lack of nucleosome remodelling in response to hormone treatment (Ballaré et al., 2013). More recent experiments in T47D exposed to subnanomolar R5020 revealed that around 2,000 PRbs are sufficient to evoke a functional response (Zaurin et al, 2020, personal communication). Hence, the number of PRbs found in Ishikawa cells probably reflects the low concentration of PR, which is compatible with a functional response to progestins. To test this possibility we increased the levels of PR in Ishikawa cells by expressing a recombinant FLAG-PR vector. These cells, FPR Ishikawa (FPR), expressed levels of PR comparable to T47D cells (Figure 2–Figure Supplement 2D) and showed no impairment in hallmark phosphorylation of serine 294 in PR (Figure 2–Figure Supplement 2E), indicating that FPR cells were capable of responding to hormone. Upon hormone exposure, FPR cells exhibited rapid binding of PR to the *EGFR* enhancer sequence (Figure 2–Figure Supplement 2F). ChIPseq experiments after R5020 exposure showed twice the number of PRbs in FPR cells compared to parental Ishikawa cells. The majority of PRbs identified in Ishikawa cells (>90%) were also detected in FPR cells (Figure 2–Figure Supplement 2G), meaning that PR overexpression reflected mostly on an increase in number of binding sites.

Upon hormone induction, sites engaged by PR in Ishikawa cells were also occupied in FPR and pretreated cells, denoting a strong similarity between them. Although a small number of binding sites was shared between ERalpha and PR in all three conditions, PR binding in pretreated cells exhibited a higher degree of similarity to ERalpha binding than FPRbs (Figure 2B). Moreover, subtracting PRbs from FPRbs (FPR_s) and prePRbs (prePR_s) heightens this difference, with a much larger fraction of binding sites shared with ERalpha in the case of prePRbs (Figure 2C). Among these sites, one located close to the promoter of *TGFA* gene, identified as an ERbs, showed significant PR binding only in preE2 Ishikawa cells, but not in FPR (Figure 2D, left panel). ERE-containing ERbs, such as the ones found in the transcription termination site of *PGR* gene and immediately upstream of *CCND2* promoter, were occupied by R5020-bound PR in preE2 Ishikawa cells (Figure 2D, middle and right panels). These three genes were upregulated by E2 treatment in RNAseq experiments performed on Ishikawa cells (Figure 1–Figure Supplement 2C). Importantly, PRbs identified in non-pretreated and E2-pretreated Ishikawa cells resembled the PR binding profile found in samples

from normal proliferating and mid-secretory endometrium (*Chi et al., 2020*) (FIGSUPP[fig2]sf5H).

The distribution of PRbs and ERbs in non-pretreated Ishikawa cells, in FPR cells and cells pretreated with E2 (prePRbs) relative to TSS of regulated genes was consistent with previous reports in other cell lines (*Ballaré et al., 2013; Need et al., 2015*), in that they were enriched in intronic and distal intergenic regions (*Figure 3A*). Nearly 50% of binding sites localized to distal regions (>50Kb) and approximately 30% to introns other than the first intron, indicating that regulation of gene expression by the steroid receptors PR and ERalpha is not mediated through proximal promoters but mostly by distal enhancer/silencer sequences. We corroborated these results employing another strategy based on binding site-gene association using the GREAT web tool (see Methods for further details (*McLean et al., 2010*)). First, we defined a set of genes associated to binding sites with a basal plus extension rule (extended up to 100kb away) and then we intersected this group of genes with R5020- or E2-regulated genes. Of the 1,886 genes regulated by R5020, only 224 (12%) were potentially associated to PRbs, while only 199 of the 950 genes regulated by E2 (21%) proved to be associated to ERbs (*Figure 3B*).

As expected, from the sequences contained in 10kb windows centered in peak summits of PRbs, FPRbs and prePRbs, the PRE emerged as the most representative binding motif (*Figure 3C*), including sites uniquely found in FPR (group c: 633) or preE2 (group d: 3,162) cells. While comparison between ERalpha and PR ChIPseq results showed few similarities regarding identity of binding sites, with a set of 216 shared by both hormone receptors, pretreatment with E2 added nearly twice as many binding sites to the pool shared with ERalpha (from *Figure 2A*, Venn Diagram). The most representative motif discovered in these sites -only shared by ERalpha and prePR- was a half ERE (*Figure 3D*, group h: 329) that was highly similar to the motif observed in sites uniquely found in Ishikawa ERbs (from *Figure 2-Figure Supplement 1H*, group III). Sites shared by ERalpha and PR in all three conditions resulted in an unclear combination of PRE and ERE motifs (*Figure 3D*, group e-g). Degenerated motif logo in group g showed no association to any known motif, probably due to a corrupt analysis performed on insufficient data, and the partially degenerated motif logo in group e showed limited association to both PRE and ERE (PRE/ERE).

Taken together, this evidence suggests that, provided there is an estrogenic background, activated PR could regulate estrogen-dependent Ishikawa-specific transcriptome by binding sites already or formerly bound by ERalpha.

PAX2 binds chromatin in close proximity to ERalpha and PR binding sites in Ishikawa cells

Evidence described so far partially explains cell type specific hormone-dependent gene regulation, though it is not sufficient to understand the mechanisms underlying differential binding of hormone receptors to chromatin. Initially, we addressed this by contrasting the sequences of ERbs and PRbs from groups I-IV, i.e. hormone regulated Ishikawa specific, (from *Figure 2-Figure Supplement 1G* and *H*) with an array of 1,395 known TF binding motifs (see Methods). Results revealed an enrichment ($p\text{-value} < 1e^{-4}$) of multiple members of the PAX family -including variants 2, 5, 6 and 9- in groups I and III, i.e. In Ishikawa specific PRbs and ERbs (*Figure 4-Figure Supplement 1A* and *B*, respectively), suggesting that members of the Pax family may be involved in PR and ERalpha action in Ishikawa cells. Unbiased comparison (all sites) of enrichment in TF binding motifs between Ishikawa and T47D PRbs showed similar results for PRbs, although the enrichment was less significant (*Figure 4-Figure Supplement 1C*). Moreover, while enrichment of PAX motifs was also observed around ERbs in Ishikawa cells (*Figure 4-Figure Supplement 1D*), this was not the case with T47D cells, in which examples like the well-known breast-related pioneer transcription factor FOXA1, were found instead (*Figure 4-Figure Supplement 1E*).

Enrichment of NR3C1-4 (mineralocorticoid, glucocorticoid, progesterone and androgen receptors) and ESR1 motifs included into the 1,395 known motifs corroborated de novo discovery performed with MEME in both Ishikawa and T47D cells. Stronger enrichment of PAX motifs was observed in prePRbs compared to PRbs (*Figure 4A*), indicating that PR binding to regions poten-

Figure 3. A fraction of E2-induced PRBs localize on ERbs and contain half ERE motifs. (A) Classification of steroid receptor binding relative to genomic features expressed as percentage (%) of peaks after 60min of hormone treatment inside each feature. Legend at the top right corner indicates the color key for ERbs (green dots) and three conditions of PR binding: non-pretreated with E2 (PRbs, red dots), pretreated with E2 for 12h (prePRbs, blue dots) and exogenous expression of PR (FPRbs, grey dots). The table below shows percentages represented in the plot. (B) To the left: Representation of GREAT tool association rules adapted with modifications (McLean *et al.*, 2010). To the right: Venn diagrams show intersection between PRbs-associated genes and R5020-regulated genes (top), and ERbs-associated genes and E2-regulated genes (bottom). (C) Peak signals in PRbs, FPRbs and prePRbs from 60min R5020-treated Ishikawa cells were plotted as heatmaps. Regions were defined inside a window of 10kb centered in peak summit (± 5 kb) and intensity of the signal correspond to number of reads in each region. Heatmap is subdivided into 4 mutually exclusive groups depending on shared/partly shared/non-shared binding sites: a (n= 1342), sites shared by all three conditions of PR binding; b (n=1072), sites uniquely found in FPR and prePR; c (n=633), sites found only in FPR; and d (n=3162), sites found only in prePR. De novo motif discovery (MEME) was performed on all groups and results are indicated as sequence logos to the right of the map, including the name of the most related known motif. PRE: progesterone response element. (D) Peak signals in PRbs, FPRbs and prePRbs as in (C), and ERbs from 60min E2-treated Ishikawa cells. Heatmap was subdivided into 5 mutually exclusive groups: e (n=198), sites shared by all three conditions of PR binding and ER binding; f (n=112), sites shared by FPRbs, prePRbs and ERbs; g (n=24), sites shared by FPRbs and ERbs; h (n=329), sites shared by prePRbs and ERbs; and i (n=918), sites uniquely found in ERbs. Motif discovery was performed as in A for all groups and results are shown to the right of the map, including the most related known motif. ERE: estrogen response element; 1/2 ERE: half ERE.

Figure 4. Putative PAX2 binding sites are associated with PR and ERalpha binding and hormone-regulated genes in Ishikawa cells. (A) Fold enrichment values ($\log_2$ FE) of 1.395 known TF binding motifs on prePRbs and PRbs. Combined p-values for enrichment analyses are indicated through the color key displayed at the lower right corner of the plot. Relevant motifs pointed on the plot correspond to NR3C1-4, members of the PAX family (1, 2, 5 and 9) and SOX9. (B) Comparison as in (A) between prePRbs and ERbs. Relevant motifs pointed on the plot correspond to NR3C1-4, members of the PAX family (1, 2, 5 and 9), SOX9, ESR1 and estrogen related (ESRr) and retinoic acid receptor (RARr). (C) Predicted UCSC Transcription Factor (TFBS) binding on genes regulated by 12h treatments with R5020 10 nM and E2 10 nM in Ishikawa cells were analysed using DAVID web-based functional enrichment tool. Heatmap shows the top 20 TFBS predicted ($p < 0.05$) for R5020- and E2-regulated genes from RNAseq results expressed as $-\log(p\text{-value})$. Arrows indicate position of PAX2, GR (PR-like binding motif) and ER.

Figure 4–Figure supplement 1. PAX2 binding to chromatin is hormone dependent in Ishikawa cells while it is not localized to nuclei of T47D cells.

285 tially bound by PAX is favored after E2 pretreatment. Coherently, while equivalent fold enrich-
286 ment values were detected when comparing prePRbs to ERbs (Figure 4B), comparison between
287 prePRbs and FPRbs showed that increased PR levels alone were not sufficient for a greater asso-
288 ciation to PAX binding motifs (Figure 4–Figure Supplement 1F). Consistently, RNAseq experiments
289 on Ishikawa cells treated either with R5020 10 nM or E2 10 nM for 12h showed putative PAX2 binding
290 sites among the top 20 significantly enriched TFs (DAVID web-based tool (Huang et al., 2009)) on
291 differentially regulated genes (Figure 4C). PAX2 and the ubiquitous AP1 were the only TFs (includ-
292 ing other PAX family members) predicted to bind both E2 and R5020-regulated genes. ER was also
293 predicted to bind on E2-responsive genes, while glucocorticoid receptor (GR) motif (PR-like motif)
294 was detected on R5020-responsive genes.

295 Next, we evaluated the expression of PAX2 and other family members in Ishikawa RNAseq sam-
296 ples (Figure 5–Figure Supplement 1A). Although none of the detected PAX genes seemed to be
297 strongly regulated by R5020 or E2, expression of PAX2, PAX3 and PAX5 was found to be lower and
298 more variable (comparing experimental replicates) than that of PAX6, PAX8 and PAX9. Despite ex-
299 pression levels of the latter were markedly higher, robust evidences indicate that reduced levels
300 or absence of PAX2 were strongly associated to early onset of neoplastic processes in endometrial
301 tissue and precancerous lesions (Monte et al., 2010; Raffone et al., 2019). We corroborated those
302 findings using publicly available data comprising of several RNAseq samples from normal and can-
303 cerous endometrial tissue (Dou et al., 2020), in which expression of PAX2 was significantly reduced
304 in tumor samples (Endometrioid and Serous adenocarcinomas) compared to normal tissue (Fig-

ure 5A).

PAX2 association to PR and ERalpha action was also evaluated by immunofluorescence against PAX2. Nuclear localization of PAX2 was observed predominantly after 60min of R5020 in pretreated and non-pretreated PR+ cells (**Figure 5B**), indicating that hormonal treatment promotes co-localization of PAX2 and PR in nuclei of Ishikawa cells. Similar results in PAX2 localization were obtained after treating Ishikawa cells with E2 for 60min (**Figure 5–Figure Supplement 1B**). The increase in nuclear PAX2 signal is not due to changes in protein levels, which were not affected by treatment with either R5020 or E2 (**Figure 5–Figure Supplement 1D**). In accordance to motif analysis results, PAX2 was not detected in nuclei of T47D cells after hormonal treatments (**Figure 5–Figure Supplement 1D**).

To extend these findings, we performed PAX2 ChIPseq experiments on untreated cells and in cells exposed for 60min to either R5020 or E2. The results confirmed PAX2 binding to chromatin following hormonal treatment (**Figure 5C**). Even though identified PAXbs were few (T0: 43, R60: 202 and E60: 209), most of PAX2 binding occurred after R5020 and E2 treatments. Moreover, PAX2 binding was not stochastically distributed in the genome of Ishikawa cells but rather partially associated to ERbs and PRbs (**Figure 5D**). This association was stronger for PR binding in cells pretreated with E2 than in non-pretreated cells or in cells overexpressing recombinant PR (**Figure 5E** and **F**). Similar results were observed for ERbs in response to E2 (**Figure 5E**). In agreement with the coverage plots, the association between PAX2 and binding of the receptors was statistically significant (two-tailed Fisher's exact test; p-value<0.001). On the contrary, we found no significant overlap between randomly rearranged PAXbs (shuffled PAXbs) and binding of receptors (two-tailed Fisher's exact test; p-value>0.1). These results indicate that PAX2, and possibly other members of the PAX family may co-operate with PR and ERalpha for binding to chromatin in Ishikawa cells but not in T47D cells, in which neither enrichment for PAX binding motif nor nuclear localization of PAX2 was detected.

In line with evidence on the role of PAX2 in endometrial cancer, we studied the effect of down-regulating expression levels of PAX2 in Ishikawa cells using a mix of specific siRNAs that produced a marked reduction of its transcript and protein (**Figure 5G**). Consistently with results described above, pretreatment with E2 for 12h did not alter levels of PAX2 or interfered with its knock down. Interestingly, knock down of PAX2 under estrogenic conditions led to a significant R5020-dependent increased binding of PR to regulatory regions upstream of *TGFA* and in *TIPARP* promoter (**Figure 5H**, **Figure 5–source data 1**), which suggest that lack of PAX2 is probably exacerbating E2 stimulation on PR recruitment. No increment in PR binding was detected in the promoter of *TGFA* or in an unrelated genomic region (negative).

This effect did not seem to alter the expression of hormone responsive genes in a similar fashion. In the absence of E2 pretreatment, R5020-dependent expression of *ALPP* was abolished by PAX2 knock down, while *TIPARP* regulation was practically unaffected (**Figure 5I**, **Figure 5–source data 2**). This is consistent with the fact that *TIPARP*, but not *ALPP*, had its promoter regulatory region occupied by both PR and ERalpha upon hormone stimulation (**Figure 5–Figure Supplement 2A**). Thus suggesting that PAX2 could modulate PR binding in regions potentially co-occupied by ERalpha. This effect was corroborated under estrogenic conditions and though reduced levels of PAX2 clearly affected the gene regulatory profiles of both genes, differences were not statistically significant probably due to the variable gene expression background in Ishikawa cells and residual levels of PAX2 among replicates. As was observed in the control condition (non-target siRNA), E2-pretreatment increased *ALPP* transcript levels after PAX2 knock down (pretreated vs. non-pretreated cells), which were then further increased by R5020 treatment. However, as in non-pretreated cells, *ALPP* transcript levels were lower compared to siRNA control conditions. In the case of *TIPARP*, we found that in agreement with previous reports on breast cancer cells (**Rasmussen et al., 2021**), expression levels were increased by E2 (pretreated vs. non-pretreated cells) and apparently by R5020 as well (treated vs. untreated conditions). In line with our hypothesis, supported by PR recruitment to *TIPARP* promoter region, R5020 treatment increased *TIPARP* levels

Figure 5. PAX2 co-localizes with PR and ERalpha in the nuclei of Ishikawa cells and it is positioned in the vicinity of PR and ERalpha binding sites. (A) PAX2 expression in Non Tumor (normal endometrium) and two histologically distinct endometrial cancer samples (Dou et al., 2020). Data stored at the National Cancer Institute's CPTAC program was accessed through cptac python package developed by Sam Payne lab. (B) Immunofluorescent detection of PR (green) and PAX2 (red) in untreated (T0) and 60min R5020-treated (R5020) Ishikawa cells which were pretreated or not with E2 for 12h (non-pretreated, E2-pretreated). Images were merged for co-localization analysis (merge). Scale bar is shown in the panels and is equivalent to 30µm. (C) PAX2 binding profile inside a region of 70kb of chromosome 14. Number of PAX2 binding sites for untreated Ishikawa cells and treated with R5020 for 60min or E2 for 60min is shown to the right of the profiles as well as the intersection between these last two groups (venn diagram). Tracks for PRbs, prePRbs, FPRbs and ERbs are displayed below the profiles for the same region. (D) Pairwise intersections of the fraction of overlap between PRbs, prePRbs, ERbs and PAXbs. Color key is indicated below the plot and the size of each set is shown as a bar plot to the right side. Intervene software was employed in the analysis (Khan and Mathelier, 2017).

Figure 5 (previous page). (E) Binding profiles of ER (green), PR (red), FPR (black) and prePR (blue) on PAX2 binding sites of 60min R5020-treated Ishikawa cells. PAX2 binding after 60min E2 treatment was included (purple). Inset shows signal profiles centered on shuffled R5020-dependent PAX2 binding sites. P-value for fisher exact test is reported in the plots. (F) Binding profiles as in (E) on PAX2 binding sites of 60min E2-treated Ishikawa cells. PAX2 binding after 60min R5020 treatment was included (orange). As in (E), inset shows signal profiles centered on shuffled E2-dependent PAX2 binding sites. P-value for fisher exact test is reported in the plots. (G) PAX2 knockdown evaluated by qPCR (left panel) and Western blot (right panel) after treating Ishikawa cells with a control siRNA (siCont) or a mix of specific siRNAs (pool). PCR data from three replicates were normalized by *GAPDH* expression and shown in the plot together with the mean $\pm$ SEM. PR was used as control for western blot. (H) Binding of PR in cells treated with siRNA control or siPAX2 followed by 12h pretreatment with E2 and treatment with R5020 for 60min. PCR signal was corrected by beta-actin and relativized to untreated control cells. Bars show the mean $\pm$ SEM of the three replicates displayed as dots. (I) Gene expression levels of hormone responsive genes before and after siPAX2 treatment. Expression was normalized to *GAPDH* and relativized to T0 using $\Delta\Delta C_t$ method and expressed as mean log2 fold change $\pm$ SEM of the three replicates displayed as dots. For all items, statistically significant comparisons resulting from ANOVA followed by Tukey's HSD are denoted by the adjusted p-value.

Figure 5–Figure supplement 1. Expression analysis of PAX2 and other PAX family members in Ishikawa cells.

Figure 5–Figure supplement 2. Expression analysis of hormone regulated genes after PAX2 knock down.

Figure 5–source data 1. PR recruitment measured by qPCR in PAX2-knocked down Ishikawa cells.

Figure 5–source data 2. Gene expression measured by qPCR in PAX2-knocked down Ishikawa cells.

in PAX2-knocked down cells.

These results suggest that the lack of PAX2 resembles some characteristics of an unopposed estrogen action, promoting reduced levels of *ALPP* and increased levels of *TIPARP*. Moreover, PAX2 knockdown reduced *ESR1* and *PGR* expression (**Figure 5–Figure Supplement 2B**) pointing to a pivotal role of PAX2 in sustaining PR and ERalpha target gene regulation. However, it did not affect *TGFA* levels (**Figure 5–Figure Supplement 2B**) that were substantially upregulated by E2 compared to non-pretreated cells, but downregulated by R5020 under estrogenic conditions.

Under estrogenic conditions, PR and PAX2 conform endometrial regulatory domains in open chromatin compartments

Nuclear architecture is a major determinant of hormonal gene regulatory patterns (*Le Dily et al., 2014*). Therefore, we used in nucleo Hi-C technology to study the folding of chromatin across the genome of Ishikawa cells by generating genome-wide contact datasets of cells untreated (T0) or pretreated with E2 for 12h, and exposed to R5020 or E2 for 60min. A comparison of contact matrices at 20 kb resolution of untreated Ishikawa cells to T47D cells confirmed the high degree of conservation on the borders of topologically associating domains (TADs) (**Figure 6–Figure Supplement 1A**). TADs are grouped into two chromatin compartments A and B, which represent the active open chromatin (A) and the closed inactive chromatin (B) respectively. Analysis of such compartments showed a cell type-specific patterning (**Figure 6–Figure Supplement 1B**), in which Ishikawa samples from two independent experiments were more closely related to each other than any of them to a T47D sample (**Figure 6–Figure Supplement 1C** and **D**). However, A/B profile distribution in Ishikawa cells was independent from hormonal treatments (**Figure 6–Figure Supplement 1B** and **E**), meaning that chromatin was in a primed state that conditioned hormone-dependent regulation of gene expression. Detailed analysis revealed that 7% of A domains in Ishikawa cells were B in T47D cells, and 12% of B domains in Ishikawa cells were A in T47D cells (**Figure 6–Figure Supplement 1F**). A total of 861 genes encompassed in the A compartment in Ishikawa cells belong in the B compartment in T47D cells, and 1,438 genes in B compartments in Ishikawa cells belong in A in T47D cells (12%), suggesting that distribution of A and B compartments could in part explain cell type specific gene expression profiles.

To evaluate whether chromatin states are related to gene expression through differential binding of hormone receptors to DNA, we intersected PR and ERalpha ChIPseq results with the A/B compartment coordinates. Both transcription factors, PR and ERalpha, bound A compartments more frequently than B, meaning that open genomic regions in Ishikawa showed preferential binding of the hormone receptors (**Figure 6–Figure Supplement 1G**). Neither pre-treatment with E2 nor expression of recombinant PR modified the preferential binding of PR to the A compartments.

As mentioned above, PAX2 binding occurs mostly in close proximity to PR and ERalpha binding sites. In fact, distances between PAXbs and PRbs were remarkably shorter in E2 pretreated cells than in any other condition (**Figure 6–Figure Supplement 1H**). This raised the question of whether recruitment of PR together with PAX2 to open chromatin compartments facilitates regulation of gene expression. To study this notion, we defined putative endometrial regulatory domains that we named "Progestin Control Regions" (PgCR) with the capacity to potentially regulate nearby genes. The restrictions for being a regulatory domain, which consisted in containing at least two PRbs separated by a maximum distance of 25kb and a PAXbs (represented in **Figure 6A: PgCRs Definition**), were met mostly under E2 pretreated conditions. This outcome was due to the strong association between prePRbs and PAXbs, though it may have been aided by the increased PR protein levels. However, the sole increment in PR protein levels was not enough to force an association to PAXbs, given that FPR cells did not show similar results (**Figure 6A: PgCRs Definition**).

Considering that TAD borders may act as regulatory barriers (**Figure 6–source data 1**), we removed from further analysis any region that, in spite of satisfying the rules for being a PgCR, was localized across a barrier as well. In agreement with this restriction, the sizes of PgCR -with an average of 25kb- were smaller than TADs -with an average of 1000kb- (**Figure 6–Figure Supplement 1I**). In addition, the majority of the 121 identified PgCRs (coordinates in hg38 **Figure 6–source data 2**) were not located near the TAD borders, but in the TAD center (**Figure 6–Figure Supplement 1J**), where most non-housekeeping genes are found (*Le Dily et al., 2019*). Moreover, PgCRs seem to be located in A compartments in the vicinity of hormone-regulated genes like *PGR* and *ALPP* (**Figure 6A**). Regulation in the expression of these genes was validated by RNAseq and qPCR of total RNA samples of Ishikawa cells exposed to R5020, E2 and E2 followed by R5020 for 12h, which showed that *ALPP* is induced by both hormones and *PGR* is regulated by R5020 only after treatment with E2 (**Figure 5I, Figure 5–Figure Supplement 2B and Figure 6–Figure Supplement 1K**).

As was mentioned before, the Hi-C matrices were used to determine the spatial segregation of chromatin in both open and closed chromatin compartments (A/B), and the A:B ratio was independent of hormone treatment. Consistent with these results, ATACseq signal on PgCRs remained unchanged upon hormone exposure (**Figure 6B, top panels**). We evaluated the possibility that ATAC-seq signal on PgCRs was simply reproducing any potential region on the genome by relocating the reference regions ("shuffled regions") preserving their size and the chromosome they were originally in. Coverage signal was lost upon shuffling the reference regions indicating that chromatin was readily and non-randomly accessible to TFs in these locations. Although ATACseq peaks were also detected on promoters of hormone-regulated genes, the signal did not differ after hormone exposure (**Figure 6B, bottom panels**), implying that treatments were not responsible for opening the chromatin in these regions. In addition, both R5020- and E2-regulated genes with highest FC values (v. T0) were concentrated under 1Mb ("close") away from PgCRs (**Figure 6C**), though the comparison between FC values of "close" and "far" (over 1Mb) regulated genes was significant only in the case of R5020 ($p=4.4 \times 10^{-3}$; **Figure 6C, inset**).

Further analysis on Hi-C contact matrices revealed that PgCRs preferentially interact with promoters of hormone-regulated genes (**Figure 6D**). Employing a "one-to-one" virtual 4C approach we found that *TGFA* promoter was in close contact with a nearby PgCR (enhancer sequences validated by qPCR in **Figure 2–Figure Supplement 2A and Figure 5H**) possibly regulating the cyclical E2-dependent increment of the transcript and the R5020-dependent downregulation that followed. Although PgCR-promoter interactions were non-random and mostly intra-TAD, we found no difference in contact enrichment between treated and untreated cells (**Figure 6E**). These results are consistent with ATACseq profiles and imply that chromatin would be pre-assembled into regulatory loops -involving PgCRs and promoters- which are transcriptionally inactive until hormone-dependent binding of steroid receptors and PAX2 triggers PolII activation (**Figure 6F**).

These results suggest that specific binding of PR, PAX2 and ERalpha to chromatin occurs in compartments that are present in a permissive (open) or restrictive (closed) status depending on the cell line, and are not modified by short term hormone exposure (**Figure 6F**). However, it is

not yet clear the role of PAX2 in PR binding to PgCRs. Summing up, PR and ER bind mostly to non-common sites that exhibit the corresponding consensus sequences, and are adjacent to PAX2 binding. Therefore, the endometrial specific hormone response results in part from specific chromatin compartments, unique receptor binding sites and selective TFs binding partners to regulate gene expression.

Genes contained in TADs with PgCRs are associated to endometrial tumor progression

To explore the possibility that alterations in the expression profile of genes under the influence of PgCRs were related to disease progression such as endometrial cancer, we examined 497 RNAseq samples from a cohort of patients diagnosed with endometrial cancer (The Cancer Genome Atlas, TCGA, Project TCGA-UCEC). Processing and exploratory analysis of raw count data are described in Methods and Materials section (**Figure 7–Figure Supplement 1A**). Samples were classified according to available clinical metadata, including the FIGO system (International Federation of Gynecology and Obstetrics), resulting in 488 samples of which 267 are Stage I, 53 are Stage II, 137 are Stage III and 31 are Stage IV. Principal component analysis (PCA) on these samples revealed a clear bias in the scattered distribution, mostly explained by the first component (PC1: 18%). Tumors in advanced stages (Stage II to IV) showed a tendency to group together and they were characterized by a marked reduction in *PGR* and *ESR1* levels (**Figure 7A**). However, just as it was observed for hormone-treated Ishikawa cells, no change in *PAX2* expression was detected throughout stages (**Figure 7–Figure Supplement 1B**). Expression profiles of selected genes previously implicated in tumor progression (causally or not) (*Liu et al., 2020; Dou et al., 2020*) were evaluated to validate our analysis (**Figure 7B**).

Progression of the malignant tissue is driven by important alterations in gene expression. Shifts in transcript abundance towards Stage IV were evaluated by differential expression analysis in DESeq2 package ($-1 > \log_2\text{FoldChange} > 1$ and $p\text{-value} < 0.05$), producing a set of 2,842 altered genes (**Figure 7C**). Among these differentially expressed genes (DEGs), we found that 102 genes (**Figure 7–source data 1**) were also located in TADs with PgCRs including both upregulated and downregulated genes (**Figure 7D** and **Figure 7–Figure Supplement 1C** and **D**). These 102 genes constitute 20% of the 522 protein coding genes in TADs with PgCRs (PgCR-genes; **Figure 7–source data 2**). To test if any randomly picked set of genes could produce the same number of intersected genes as the DEGs, we ran 10,000 iterations of an adapted bootstrapped model in which each randomly generated (without replacement) new set of 2,842 genes was contrasted with PgCR-genes. Each intersection was then used to build a distribution ($\mu=71.58$) and a 95% confidence interval ($2.5\%=57$; $97.5\%=87$) that demonstrated that the 102 intersected genes were in fact not part of the normal distribution (**Figure 7E**).

Next, we employed GSEA algorithm to test whether the expression profile of these 102 genes in advanced tumor samples was different to the one observed in Ishikawa cells treated with R5020 or E2. Genes modulated by hormones in Ishikawa cells were mainly regulated in the opposite direction during tumor progression, though most of the genes were not regulated in the conditions of our experiments (**Figure 7F**). Although GSEA results were not statistically robust, genes downregulated in tumor progression were mostly clustered together among genes upregulated by E2 (NES=1.25; FDR=0.1669) or R5020 (NES=0.99; FDR=0.9547) in Ishikawa cells. On the other hand, genes upregulated in tumor progression were more evenly distributed across treated and untreated Ishikawa samples.

These results indicate that hormones via their receptors maintain a fine balance over the expression levels of these genes, which become disrupted upon tumor progression and that PgCRs regulation involves genes that participate directly or indirectly in key steps of tumor onset and progression.

Figure 6. Convergence of PR and PAX2 binding in TADs with regulated genes defines potential endometrial regulatory domains. (A) Upper panel shows the contact matrices at a resolution of 20kb obtained by In Nucleo Hi-C in *PGR* and *ALPP* loci. Middle panel shows the spatial segregation of chromatin as open or closed compartments inside TADs (green bars: A compartment; white bars: B compartment - see methods section). The bottom panels show ChIP-seq signal distribution of PR, FPR, prePR, PAX2 and ERα as well as the location of PgCRs and genes over the region. The dashed rectangle restricts the TAD of interest and the vertical arrow marks the TSS of *PGR* and *ALPP*. Definition of PgCR: Coverage profiles of PR (red), FPR (black) and prePR (blue) binding on Progesterone Control Regions (PgCRs) delimited by the start and end

Figure 6 (previous page). labels, and flanked upstream and downstream by 1.5Kb regions. Input sample (grey) was included in the plot. Rules for qualifying as a control region are depicted on top of the profile plot. Magnified images over Control Regions are shown to the right (zoom on PgCRs). (B) ATACseq peaks from cells untreated (T0), treated with R5020 for 60min, 12h E2-pretreated (preE2 T0) and E2-pretreated followed by 60min treatment with R5020. Signal was plotted over Control Regions, shuffled Control Regions (Shuffled Regions), promoters of all annotated genes from GENCODE database (GENCODEv38) and promoters of genes regulated by 12h treatments with R5020 or E2. (C) Plot shows fold change values of genes regulated by R5020 and E2 (v. untreated cells) relative to Control Regions. Genes located upstream of PgCRs are represented with negative distance values. Dashed horizontal lines mark fold change cut-off points ($|\log_2FC|=0.8$) and vertical lines are placed at position -1 and 1Mb. Insets depict comparison of fold change values (absolute values) between genes located beneath (close) and over (far) a 1Mb distance from PgCRs. Statistical significance for this comparison was determined with Welch Two Sample t-test and is represented by a p value on the plot. (D) Top panel: Hi-C contact map at 5kb resolution of chr2 (70,200,000-71,200,000) obtained in Ishikawa cells and showing the organization around TGFA gene locus. Middle panel: Virtual 4C profile at 5kb resolution (expressed as normalized counts per thousands within the region depicted above) using the TGFA promoter as bait and showing the contacts engaged between TGFA promoter and the PgCR detected in this region (highlighted in green). Arrow on top panel highlights the position of the loop in the map. Bottom panel shows the positions of genes in the region depicted. (E) Distributions of observed versus expected interactions established between promoters (red - left), between PgCRs (blue - middle) and between Promoters and PgCRs (purple - right) located within a same TAD in Ishikawa cells treated as indicated below. (F) Representation of a chromatin loop involving a PgCR and the promoter of a regulated gene. Initially, the gene is transcriptionally inactive even though the loop is already formed. After hormone induction (E2 pretreatment followed by R5020), PR, PAX2 and in some cases ERalpha occupy open chromatin compartments in contact with promoters resulting in transcriptional activation.

Figure 6–Figure supplement 1. Chromosome compartments analysis and definition of Progesterone Control Regions.

Figure 6–source data 1. TAD coordinates (hg38).

Figure 6–source data 2. PgCR coordinates (hg38).

Discussion

There seems to be consensus that the way in which combinations of TFs assemble their binding sites contributes to the folding of the genome in cell type specific patterns that orchestrate the physiological coordination of gene expression programs required for the proper development and function of complex organisms (Lambert *et al.*, 2018; Stadhouders *et al.*, 2019). There is evidence that the same TF can regulate different gene sets in different cell types (Gertz *et al.*, 2012), but the mechanisms through which hormone receptors regulate endometrial specific gene networks had not been previously deciphered. Here, we describe ERalpha and PR binding to the genome of endometrial cancer cells and analyse their specific chromatin context. In this genomic study we used Ishikawa cells, given that they are a good model of Type I epithelial endometrial cancer (Nishida, 2002) containing ERalpha and PR.

It should be noted that progesterin regulation of gene expression in Ishikawa cells is decidedly variable. Then, the effects of perturbations to the model (i.e. treatments, siRNA) are tested in a highly variable gene expression background. We would like to remark that the contribution of our results to the progesterone receptor binding and its interaction with their neighbouring genomic regions in endometrial cancer cells is novel and more consistent than their contribution to productive gene expression. Therefore, it was relevant to our model to validate the results by contrasting them with previously published transcriptomic and PR binding data of normal endometrium and endometrial cancer.

It was reported that in Pgr Knockout (PRKO) mice the absence of PR results in unopposed estrogen-induced endometrial hyperplasia (Lydon *et al.*, 1995). As for the two isoforms of PR, the PRB isoform is considered a strong transcriptional activator while PRA can function as a transcriptional inhibitor of PRB activity (Mulac-Jericevic *et al.*, 2000). Selective ablation of PRA in mice results in a PRB dependent gain of function, with enhanced estradiol-induced endometrial proliferation (Conneely *et al.*, 2003). Ishikawa cells express more PRB than PRA, coherent with PRB dominance in glandular epithelial cells (Mote *et al.*, 1999). To explore the mechanism underlying the endometrial specific response to ovarian steroids hormones, we studied the genomic binding of ERalpha and PR by ChIPseq in hormone untreated Ishikawa cells and in cells exposed to hormone for different time periods. We discovered that the majority (67%) of PRBs after estradiol pretreatment were new sites not present in untreated cells and different as well from ERBs occupied after estradiol

519 treatment. Just 639 PR binding sites (11% of all PRbs) were the same for both PR and ERalpha. This
 520 indicates that contrary to what was described in breast cancer cells (*Mohammed et al., 2015; Sing-*
 521 *hal et al., 2016*), in endometrial cells PR binding has little influence on ERalpha binding. In Ishikawa
 522 cells, binding of ER and PR occurs mainly at ERE and PRE sequences, respectively, in regions that
 523 are also enriched in PAX response elements. Ishikawa cells are rich in PAX TF and PAX ChIPseq
 524 shows a similar overlapping with ERbs and PRbs.

525 When we analyzed chromatin topology of Ishikawa cells using Hi-C we found that PRbs and ERbs
 526 are enriched in Topologically Associating Domains (TADs) containing hormone regulated genes.
 527 These TADs were predominantly part of the open (A) chromosome compartment, even in cells
 528 not exposed to hormone. This was confirmed by ATACseq results showing that the sites where
 529 the hormone receptors will bind were already accessible, suggesting that hormone independent
 530 mechanisms were responsible for the generation and maintenance of the hormone responsive
 531 TADs. In that respect, it is interesting that we found an enrichment of PAXbs near PRbs in these
 532 TADs containing progesterone regulated genes, suggesting that PAX2 could generate the open
 533 chromatin conformation that enables PR binding and facilitates the interacting loops detected in
 534 Hi-C experiments. Loss of PAX2 expression has been implicated in the development of endometrial
 535 intraepithelial neoplasia (EIN) (*Sanderson et al., 2017*) and PAX2 is potentially useful in the diagnos-
 536 tic of difficult EIN cases (e.g. where there is no "normal" tissue available to act as an internal control
 537 when assessing nuclear morphology) (*Quick et al., 2012*). The fact that PAX2 expression is reduced
 538 upon tamoxifen treatment or through neoplastic processes (*Monte et al., 2010*) in which altered
 539 hormone receptor pathways are hallmark features is indicative of a central role of the estrogen
 540 receptor in its mechanism of action. In line with this notion, we propose that E2-induced ERal-
 541 pha acts together with PAX2 to define key PR regulatory regions that modulate gene expression.
 542 Our results connect PR response elements with PAX2 and 3D chromatin conformation, which is
 543 consistent with the preservation of progestin regulation in differentiated cancer cells expressing
 544 hormone receptors. Undifferentiated tumor cells, which do not express hormone receptors, lose
 545 this regulatory capabilities leading to more aggressive manifestations of the malignancy. We hy-
 546 pothesize that PR-PAX-PR binding sites containing regulatory domains that we name PgCRs could
 547 reflect PR shadow enhancers (*Cannavò et al., 2016*) in endometrial cells.

548 The redundancy of PRbs associated to endometrial specific gene expression may reinforce a
 549 genetic mechanism to ensure progestin regulation in tissue under hormonal influence, in periods
 550 in which there is low or no circulating hormone. Notably, the only described super-enhancer in
 551 endometrial carcinomas is the Myc super-enhancer and is not hormonally regulated (*Zhang et al.,*
 552 *2016*). We postulate the existence of a novel subset of 121 strategic endometrial regulatory do-
 553 mains in this hormonally responsive endometrial cancer cell line. Among them the *TGFA* gene
 554 presents one of PgCR-promoter interaction that could explain hormone regulation previously re-
 555 ported in this cells (*Hata et al., 1993*). This concept could be exploited to guide treatments oriented
 556 to recover progestin regulation over estrogen proliferative effects in endometrial malignancy.

557 Previous results in T47D mammary cancer cells have shown Hormone Control Regions, which
 558 include ERbs and PRbs acting in conjunction with FOXA1 and C/EBPa (*Nacht et al., 2019*) interact
 559 with promoters of hormone regulated genes in hormone responsive TADs and organize the high
 560 level folding of the genome (*Le Dily et al., 2019*). Although the analysis of interaction between PgCR
 561 and different ERalpha enriched binding regions in endometrial cells remains to be performed, our
 562 present study proposes that PR binding sites originated under estrogenic conditions and acting in
 563 conjunction with PAX2, fulfil a similar function in differentiated hormone-responsive endometrial
 564 cancer cells. Thus combinations of the same hormone receptors and different transcription factors
 565 account for cell type specific expression of different gene regulatory networks in part by generating
 566 and maintaining different genome topologies.

567 *Droog et al. (2017)* highlights that "the divergence between endometrial tumors that arise in
 568 different hormonal conditions and shows that ERalpha enhancer use in human cancer differs in
 569 the presence of nonphysiological endocrine stimuli". They reported that ERalpha-binding sites in

Figure 7. Altered expression of PgCR-genes correlates with drivers of endometrial tumor progression. (A) Scatter plot of PCA results showing scores of components 1 and 2 (PC1 and PC2) using transcriptomic data of endometrial cancer samples (protein coding genes) obtained from TCGA-UCEC database (n=488). Samples are classified by FIGO stage (I to IV) and identified by color. Marginal density plots represent distribution of scores for each stage. Normalized expression values of PGR and ESR1 genes in these samples is shown to the right of the scatter plot. (B) Expression of genes positively (top row) and negatively (bottom row) regulated during endometrial cancer progression (FIGO stage) from TCGA-UCEC samples. (C) Heatmap shows scaled normalized counts of 2,842 differentially expressed genes (DEGs) between Stage IV and the other three stages. (D) Intersection between DEGs presented in C and protein coding genes contained in TADs with PgCRs (522) resulted in 102 gene identities detailed in the box to the left of the Venn diagram. Genes were arranged as Upregulated or Downregulated with respect to Stage IV samples. (E) Bootstrapping statistical approach to test whether intersection described in D could result from randomly picking any set of 2,842 genes among all known protein coding genes (GENCODE v38). Vertical red lines mark the 2.5 and 97.5 percentile, which denote the 95% confidence interval of the distribution. Intersection from D is indicated as a vertical black dashed line (102). (F) GSEA plot results using the 102 genes from intersection in D as gene set to match against an expression dataset from Ishikawa cells treated or not with R5020 or E2 for 12h. Normalized enrichment scores and FDR values are reported on the plots.

Figure 7-Figure supplement 1. Exploratory analysis and processing of TCGA-UCEC endometrial cancer samples.

Figure 7-source data 1. Intersection between Tumor DEGs and PgCR-genes.

Figure 7-source data 2. Protein coding genes in TADs with PgCRs (PgCR-genes).

tamoxifen-associated endometrial tumors are different from those in the tumors from nonusers. It has yet to be explored whether the response to progesterone and synthetic progestins, used in treatments of hormone-dependent endometrial cancers, is affected by the changes resulting from the use of tamoxifen.

On the other hand, ERalpha and glucocorticoid receptor (GR) are expressed in the uterus and have differential effects on growth (*Vahrenkamp et al., 2018*). Expression of both receptors was associated with poor outcome in endometrial cancer and the simultaneous induction of ERalpha and GR leads to molecular interplay between the receptors (*Vahrenkamp et al., 2018*). In our conditions, R5020 induces genes with GR/PR putative binding sites, enabling regulation that could result in a similar ERalpha-GR pathological outcome.

In the last ten years an enormous effort has been placed in gathering massive amounts of good quality omic and clinical data from healthy women and endometrial cancer patients in different stages of this often sub-diagnosed disease (*Cancer Genome Atlas Research Network et al., 2017; Dou et al., 2020*). Besides the underlying mutational triggers, the mechanisms by which normal endometrial tissue progresses into aggressive and potentially lethal malignancies are being extensively studied (*Ma et al., 2018; Bai et al., 2019; Monsivais et al., 2019; Rodriguez et al., 2020*). Employing this data, we verified that PAX2, PR and ERalpha levels are reduced in tumor samples (versus normal tissue) and they decreased even further throughout all stages of the disease. In conclusion, it is our belief that loss of PAX2 forces aberrant PR and ERalpha signaling in endometrial cells that may lead to the altered expression of both receptors and consequently PgCR-genes, facilitating the progression of the disease.

591 **Methods and Materials****Table 1.** Key Resources Table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
cell line (<i>Homo sapiens</i>)	Endometrial adenocarcinoma (Epithelial) Ishikawa	Dr. Rochefort Unité Hormones and Cancer INSERM, France		Female
cell line (<i>Homo sapiens</i>)	FPR Ishikawa	This paper		flag-tagged PR overexpression; Female
cell line (<i>Homo sapiens</i>)	Breast cancer (Epithelial) T47D	Dr. Beato (Center for Genomic Regulation)		female
recombinant DNA reagent	p3xFLAG-CMV-14 (plasmid)	Dr. Beato (Center for Genomic Regulation)		flag-tagged human PR
sequence-based reagent	human PAX2 siRNA	Dharmacon		pool of 4 on-target oligos
sequence-based reagent	human control siRNA	QIAGEN 1027310		scramble non-target oligo
antibody	anti-PR (H-190 Rabbit polyclonal)	Santa Cruz Bio. sc-7208	RRID:AB_2164331	ChIP:30 μ l xIP Western (1:200)
antibody	anti-ERalpha (HC-20 Rabbit polyclonal)	Santa Cruz Bio. sc-543	RRID:AB_631471	Western (1:200)
antibody	anti-ERalpha (HC-20X Rabbit polyclonal)	Santa Cruz Bio. sc-543X		ChIP (25 μ l xIP)
antibody	anti-ERalpha (H-184X Rabbit polyclonal)	Santa Cruz Bio. sc-7204		ChIP (25 μ l xIP)
antibody	anti-phosphoserine 294 PR (Rabbit polyclonal)	Cell Signaling 13736	RRID:AB_2798307	IF (1:100)
antibody	anti-PAX2 (Rabbit polyclonal)	BioLegend PRB-276P (Covance)	RRID:AB_291611	ChIP:6 μ l xIP Western (1:200)
antibody	normal rabbit IgG	Santa Cruz Bio. sc-2027	RRID:AB_737197	ChIP (12 μ l xIP)
antibody	anti- α -tubulin (Mouse monoclonal)	Merck T5168 (Sigma-Aldrich)	RRID:AB_477579	Western (1:1000)

592 **Cell culture and hormonal treatments**

593 Endometrial adenocarcinoma Ishikawa cells and FPR Ishikawa cells were cultured in phenol red
594 DMEM/F12 medium (GIBCO, Thermo Fisher Scientific) supplemented with 10% FCS (GreinerBioOne)
595 and gentamycin (Thermo Fisher Scientific) at 37 °C and 5% carbon dioxide to maintain cell line stock.
596 Ishikawa cells were obtained from Dr. Rochefort at the Unité Hormones and Cancer, INSERM U148
597 et VAC 59 CNRS, 34090 Montpellier, France. The cell line was authenticated by karyotyping (G band)
598 a sample in comparison with the original published karyotype (*Nishida, 2002*) and has been tested
599 negative for mycoplasma contamination. Before each experiment, cells were plated in phenol red-
600 free DMEM/F12 medium supplemented with 5% dextran-coated charcoal-treated (DCC)-FCS and

gentamycin for 48h. Then, the medium was replaced by serum-free DMEM/F12 and kept in it for 18h (overnight). Treatments were performed with R5020 and E2 to a final concentration of 10 nM and ethanol (vehicle) for the times indicated for each experiment. When indicated, pretreatment with E2 consisted of a single administration of E2 to a final concentration of 10 nM 12h before hormonal treatments. T47D cells were cultured in RPMI 1640 medium as previously described (*Nacht et al., 2016*).

Transfection with flag-tagged PR (FPR Ishikawa cells)

Plasmid p3xFLAG-CMV-14 carrying the complete sequence for progesterone receptor gene (HindIII-924 - 938EcoRI) was introduced in Ishikawa cells using Lipofectamine 2000 (Thermo Fisher Scientific) following manufacturer recommendations. After 24h of transfection, cells were exposed to 0.6mg/mL G418 for selection. Then on, every two passages, FRP cells were exposed to a reduced concentration of G418 (0.4mg/mL), except during hormonal treatments.

Transfection with small interference RNA (siRNA)

Ishikawa cells were seeded in 6 well plates (3×10^5 per well, Western blot and qRT-PCR) and p100 plates (1.5×10^6 per plate, ChIP) for transfection with pool of 4 on-target siRNAs (Dharmacon; #1: GAAGUCAAGUCGAGUCUAU, #2: CGACAGAACCCGACUAUGU, #3: GGACAAGAUUGCUGAAUAC, #4: CAUCAGAGCACAUCAAAUC) or scramble non-target (QIAGEN; #1: AATTCTCCGACGTGTACAGT). The following day, cells were transfected using Lipofectamine 3000 (Thermo Fisher Scientific) following instructions from transfection reagent manufacturer. Briefly, mix A containing OptiMEM and siRNAs and mix B containing OptiMEM and Lipo3000 were incubated together for 20min at room temperature and added to the cultured in a drop by drop fashion. During incubation, media was replaced by phenol red-free DMEM/F12 5% DCC-FCS without antibiotics. 36h later, cells were treated with E2 10 nM for 12h (pretreatment) and then 60min with R5020 10 nM. Cells were seeded in three replicates and statistical differences among conditions were evaluated by ANOVA followed by Tukey's HSD and pairwise t test with correction for multiple comparison to produce adjusted p-values (Benjamini-Hochberg method).

Proliferation assay

Ishikawa cells were seeded at 5×10^4 cells/plate density in 35mm dish plates. After 48h in 5% DCC-FCS, the medium was replaced for 1% DCC-FCS for 18h. Treatments were performed for 48h and cells were then collected using trypsin (0.25%). Antagonists for ER and PR, ICI182780 and RU486 1 μ M respectively, were added for 60min and removed before hormonal treatments. The number of live cells was determined using trypan blue (0.1%) in Neubauer chamber, repeating the procedure sixteen times for each sample and performing three independent experiments.

BrdU incorporation assay and cell cycle analysis

Ishikawa cells were seeded and prepared for hormonal treatments as described for Proliferation assay. Treatments were carried out for 15h, the last two hours of which includes incubation with BrdU. Cells were treated with cell cycle inhibitor TSA A 250 μ M as negative control of BrdU incorporation. After collecting cells in trypsin and washing them with PBS, ethanol 70% was added to fix and permeabilize them. DNA denaturation was achieved with 0.5% BSA and 2 M HCl after which cells were incubated in 1:2000 solution of anti-BrdU (BD Pharmingen) for 1h at RT. FITC secondary antibody (Dako) was incubated for 1h in obscurity at RT followed by propidium iodide for 5min. BrdU incorporation and cell cycle phases were evaluated by flow cytometry (BD FACS Canto II) in three replicates.

Western blot

Cell extracts were collected at the times indicated by the experiment with 1% SDS, 25 mM Tris-HCl pH 7.8, 1 mM EDTA, 1 mM EGTA and protease and phosphatase inhibitors. Total protein extracts

were loaded in 8% SDS-PAGE and incubated with the following antibodies: PR (H190, Santa Cruz Bio.), ERalpha (HC-20, Santa Cruz Bio.) and alpha-tubulin (Sigma Aldrich). Quantification of gel images was performed with ImageJ software and expressed as abundance in relative units to alpha-tubulin.

Immunofluorescence

Cells were seeded onto coverslips in six-well plates in a density of 10×10^3 cells/150 μ l using the protocol described in Cell culture and hormonal treatments and either pretreated or not with E2 10 nM during the last 12h of serum-free culture. After hormonal treatments cells were washed with ice cold PBS followed by fixation and permeabilization by incubation in 70% ethanol for 12h at -20°C . After rinsing three times for 5min in 0.1% Tween-PBS, the coverslips were incubated for 2h with 10% BSA in 0.1% Tween-PBS to reduce nonspecific staining. To detect PR (H-190 Santa Cruz Bio.), phosphoserine 294 PR (S294 Cell Signaling), ERalpha (HC-20 Santa Cruz Bio.) and PAX2 (Biolegend) cells were incubated with corresponding antibodies diluted in 10% BSA 0.1% Tween-PBS at 4°C overnight. After several washes in Tween-PBS, coverslips were exposed to secondary antibodies Alexa 488 and Alexa 555 (Thermo Fisher Scientific, Thermo Fisher Scientific) diluted 1:1000 in 10% BSA 0.1% Tween-PBS for 1h at room temperature using DAPI to reveal nuclei. Coverslips were mounted on slides with Mowiol mounting medium (Sigma Aldrich) and analyzed in TIRF Olympus DSU IX83 (Olympus Life Sciences Solutions). Quantification of nuclear fluorescence was done with ImageJ software after generating a binary mask in dapi images.

qRT-PCR

After 12h of treatment with R5020 and E2, cell extracts were collected in denaturing solution (4 M Guanidine thiocyanate, 25 mM Sodium citrate pH 7, 0.1 M 2-Mercaptoethanol, 0.5% Sarkosyl) and total RNA was prepared following phenol-chloroform protocol (*Chomczynski and Sacchi, 1987*). Integrity-checked RNA was used to synthesize cDNA with oligodT (Biodynamics) and MMLV reverse transcriptase (Thermo Fisher Scientific). Quantification of candidate gene products was assessed by real-time PCR. Expression values were corrected by GAPDH and expressed as mRNA levels over time zero (T0). Primer sequences are: *PGR* (FW:5'-CCCACAGGAGTTGTCAAGC-3' RV:5'-TAACTTCAGACATCATTTC-3'), *ESR1* (FW:5'-TCTATCCGAGTATGATCCTACCA-3' RV:5'-CAGACGAGACC-AATCATCAG-3'), *ESR2* (FW:5'-GTCCTGCTGTGATGAACTAC-3' RV:5'-CCCTCTTTCGCTTTGGACTA-3'), *ALPP* (FW:5'-CAACCTGAGCTGCCTTTCTC-3' RV:5'-GAACTGTGTCCCGGCTTCT-3'), *TGFA* (FW:5'-CTTCAAGCCA-GGTTTTTCGAG-3' RV:5'-GGCAGGTTGGAAGAGATCAA-3'), *TIPARP* (FW:5'-CACACCAGCTCACTTCCA GA-3' RV:5'-CAGCTCAAACAGGAGGTCAA-3'), *GAPDH* (FW:5'-GAGTCAACGGATTGGTCGT-3' RV:5'-TTGATTTT-GGAGGGATCTCG-3'), *PAX2* (FW:5'-CCAATGGTGAGAAGAGGAAA-3' RV:5'-CTCAAAGACCCGATCCAAAG-3').

Luciferase reporter assay

Ishikawa cells were seeded and prepared for hormonal treatments as described for Proliferation assay without addition of gentamycin. Cells were co-transfected with MMTV LTR-Firefly Luciferase (pAGMMTVLu, gift from Laboratory of Patricia Elizalde) and CMV-Renilla luciferase (pRL-CMV, Promega) plasmids using lipofectamine plus 2000 (Thermo Fisher Scientific). After 5h, media were renewed with the addition of antibiotics and 12h later cells were treated with vehicle (ethanol) and R5020 for 20h. Firefly and Renilla activities (arbitrary units) were determined with Dual-Luciferase Reporter assay system (Promega) and expressed as Firefly units relative to internal control Renilla for each sample (Firefly $\times 10^4$ /Renilla).

RNAseq

Total RNA was collected from untreated (T0) and 12h R5020- and E2-treated Ishikawa cells using RNeasy Plus Mini Kit (QIAGEN) and subjected to high-throughput sequencing in Illumina HiSeq 2000 and 2500. Poly-A-enriched RNA was used to prepare libraries with TruSeq RNA Sample Preparation

kit v2 y v4 (ref. RS-122-2001/2, Illumina) according to instructions from manufacturer followed by single-end (run1) and paired-end (run2) sequencing. Good quality 50bp reads were aligned to the reference human genome (hg19, UCSC) using Tophat software (*Trapnell et al., 2009*) keeping those that mapped uniquely to the reference with up to two mismatches and transcript assembly, abundance quantification and differential expression analyses were performed with the Cufflinks tool (*Trapnell et al., 2010*). Genes under 200bp in length or with FPKM values below 0.1 were excluded from downstream analyses. Genes were classified into induced, repressed or non-regulated depending on log2FC value relative to untreated cells (T0). Threshold value was arbitrarily set at log2FC $\geq \pm 0.8$ and $q < 0.05$ (FDR). Enriched terms and TFBS were determined through RDAVIDWeb-service (*Fresno and Fernández, 2013*) and DAVID web-based tool (*Huang et al., 2009*) under standard parameter settings for each tool.

Gene Set Enrichment Analysis (GSEA)

GSEA tool was implemented following instructions from developers under default parameters (*Subramanian et al., 2005*). The expression dataset was created using Ishikawa RNAseq results, labelling samples as "R5020" and "E2" for categorical classification (phenotypes). Gene sets were constructed from proliferative and mid-secretory normal endometrial RNAseq samples (*Chi et al., 2020*). Differential expression analysis to extract genes representative of each stage was performed with DESeq2 package ($|\log_2FC| > 2$, $p < 0.05$) (*Love et al., 2014*).

Chromatin immunoprecipitation (ChIP)

ChIP experiments were performed as described in (*Strutt and Paro, 1999*) and (*Vicent et al., 2011*). Antibodies used for immunoprecipitation were PR (H-190, Santa Cruz Bio.), ERalpha (HC-20X and H-184X, Santa Cruz Bio.), PAX2 (PRB-276P, BioLegend) and normal rabbit IgG (sc-2027, Santa Cruz Bio.). Enrichment to DNA was expressed as percentage of input (non-immunoprecipitated chromatin, 1%) relative to untreated Ishikawa cells (T0) using the comparative Ct method. Ct values were acquired with BioRad CFX Manager software. Primers sequences are: EGFR UPS (FW:5'-GCGTGAGACACAAACATTCCAAAGTGA-3' RV:5'-GTTCAAGCAATGGGATCGAGTTGT-3'), ALPP UPS (FW:5'-AACTGTTCCAGCTGCGTTTT-3' RV:5'-AGAACACGGTCACTTCCTTGA-3'), ALPP PROM (FW:5'-TGACAGGGTGTCTTGTTCCA-3' RV:5'-GGGTGCGGTATTGAGTACAGA-3'), TGFA UPS (FW:5'-CACATCCGGAGTTCAGACAA-3' RV:5'-CACCTGGGAGCAGGTTACTC-3'), TGFA MID (FW:5'-GGCATTGGAGGGTGTCTAA-3' RV:5'-GAGCAGAGGGGTCACTGAAG-3'), TGFA P1 (FW:5'-CTCTCACACCAGACGAAGCACA-3' RV:5'-CAGTGACCCCTGAGTTGGAGACT-3'), TGFA PROM (FW:5'-GGGAAAAAGACGCAGACTAGG-3' RV:5'-GGTAGCCGCCTTCCTATTCTC-3'), TIPARP PROM (FW:5'-GAGGCTGGAGGCGTCTGGGGAGTAGG-3' RV:5'-CTGCGGACAGATGGAGGGTCACTTTG-3'), SERPINA3 (FW:5'-GCATCATCAAAGTGAAGCCTGAGAA-3' RV:5'-CAGTAGAAAAGCCTCTTTGTTACTCCCA-3'), RPS6KA1 (FW:5'-GGTACTGTTGTCTGGTCCCCCT-3' RV:5'-CTCCAGTGAGAACAGCCCAACT-3').

ChIPseq

After minor modifications to the ChIP protocol described in (*Vicent et al., 2011*), purified ChIP-DNA was submitted to deep sequencing using Illumina HiSeq-2000. Libraries were prepared by the Genomics unit of the CRG Core Facility (Centre for Genomic Regulation, Barcelona, Spain) with NEBNext ChIPseq Library Prep Reagent Set (ref. E6200S, Illumina) and 50bp sequencing reads were trimmed to remove Illumina adapters and low-quality ends using Trimmomatic (*Bolger et al., 2014*) version 0.33 in single-end mode. Good quality reads were aligned to the reference human genome (hg19 or hg38, UCSC) with BWA (*Li and Durbin, 2009*) v0.7.12 (BWA_MEM algorithm with default parameters) keeping alignments that mapped uniquely to the genome sequence (Samtools version 1.2, (*Li and Durbin, 2009*)). Overlapping reads were clustered and significant signal enrichments (peaks) were identified by MACS2 v2.1.0 (*Zhang et al., 2008*) using input as background signal. FDR value during initial peak calling steps was set to 0.05 (q), though downstream analyses included only those with $q < 10^{-5}$. Replication of binding sites was evaluated among treatments (time of

exposure to hormone) and conditions (no pretreated, pretreated and FPR) using scatter plots, venn diagrams and heatmaps. Selected sites were validated by qPCR. When necessary peak files were converted to hg38 coordinates using the batch conversion tool from UCSC. Statistical significance of the association between peak files (bed) was evaluated using bedtools (v.2.28.0 (Quinlan, 2014)) shuffle (parameters for randomization: -chrom -noOverlapping) and fisher (pairwise comparisons) modules.

Heatmaps, Scatterplots and Motif analysis

Overlap of ChIPseq peak regions defined by upstream peak calling procedures (MACS2) were determined using intersectBed program from the bedTools suite (Quinlan, 2014). An overlap of at least one bp was considered positive. De novo motif discovery (MEME software) performed on sequences contained in 10kb windows centered in peak summits. Graphs, correlation tests, non-linear regression and statistical analyses in general were performed for common peaks between ChIPseq samples using R (R Development Core Team). Heatmaps were plotted using the summit of the peaks as a reference central position. Reference positions were taken from common and exclusive peaks within experiments and were sorted by height of the peak. Genome aligned reads occurring between -5000 and +5000 bp from reference sites were mapped using count_occurrences program (Kremisky et al., 2015) and the number of reads per bins of 200bp was used for the color intensity of heatmap cells with R. For Motif discovery, genomic regions of top 500 peaks ranked by their height were extracted from each set and regions that overlap with repeats, low complexity regions or transposable elements (extracted from the UCSC genome browser, hg19 human assembly), were removed from the analysis. Motif discovery was performed using MEME program suite executed with the following parameters: -maxsize 250000 -revcomp -dna -nmotifs 3 -mod oops (Bailey et al., 2015). Motif enrichments were evaluated with the procedure and statistics described in (Agirre et al., 2015). Additionally, the analysis utilized a 5mers collection of 1,395 human position frequency matrices modelling transcription factors binding sites (Weirauch et al., 2014), which were scanned ($p\text{-value} < 1 \times 10^{-4}$) and their enrichment evaluated in regions of 200bp centered in the summits of whole peaks sets. To uncover motif profiles, discovered and library motifs were whole-genome scanned ($p\text{-value} < 1 \times 10^{-4}$). Their occurrences around the sets of summits were obtained with count_occurrences ($\pm 2000\text{bp}$, bin size=200bp) and the profiles showing the proportion of regions per bin having at least one match were plotted using R.

Binding site-gene association

Genomic coordinates of PR and ERalpha binding sites (hg38) were fed to GREAT web tool (McLean et al., 2010) to identify potential cis-regulatory interactions. Association was determined in a "basal plus extension" process using a proximal regulatory domain of 5kb upstream and 1kb downstream from each TSS (GRCh38, UCSC hg38) and an extension of 100kb in both directions. The group of genes associated with PRbs or ERbs were respectively intersected to R5020 and E2 RNAseq results, employing simple python scripting.

ATACseq

ATACseq was performed as previously described (Buenrostro et al., 2013). Briefly, 50,000 cells were lysed with 50 μL cold lysis buffer (Tris-Cl pH 7.4 10mM; NaCl 10mM; MgCl₂ 3mM; NP-40 0.1% v/v) and centrifuged at 500xg for 10min at 4 °C. Nuclei were resuspended in TD Buffer with 1.5 μL Tn5 Transposase (Nextera, Illumina) and incubated 15 minutes at 37 °C. DNA was isolated using Qia-gen MinElute column and submitted to 10 cycles of PCR amplification using NEBNext High-Fidelity 2X PCR Master Mix (Univ. primer: AATGATACGGCGACCAACGAGATCTACACTCGTCGGCAGCGTCAGATGTG ; Indexed primers: CAAGCAGAAGACGGCATACGAGATNNNNNNNNGTCTCGTGGGCTCGGAGATGT). Library were size selected using AMPure XP beads and sequenced on a NextSeq 500 instrument (2x75nt).

789 Hi-C

790 High-throughput chromosome conformation capture assays were performed as previously de-
 791 scribed (*Lieberman-Aiden et al., 2009; Rao et al., 2014*). Adherent cells were directly cross-linked
 792 on the plates with 1% formaldehyde for 10min at room temperature. After addition of glycine
 793 (125mM final) to stop the reaction, cells were washed with PBS and recovered by scrapping. Cross-
 794 linked cells were incubated 30min on ice in 3C lysis Buffer (10 mM Tris-HCl pH=8, 10 mM NaCl, 0.2%
 795 NP40, 1X anti-protease cocktail), centrifuged 5min at 3,000 rpm and resuspended in 190 μ L of NEB-
 796 uffer2 1X (New England Biolabs - NEB). 10 μ L of 10% SDS were added and cells were incubated for
 797 10min at 65 °C. After addition of Triton X-100 and 15min incubation at 37 °C, nuclei were centrifuged
 798 5min at 3,000 rpm and resuspended in 300 μ L of NEBuffer2 1X. Digestion was performed overnight
 799 using 400U Mbol restriction enzyme (NEB). To fill-in the generated ends with biotinylated-dATP, nu-
 800 clei were pelleted and resuspended in fresh repair buffer 1x (1.5 μ L of 10 mM dCTP; 1.5 μ L of 10 mM
 801 dGTP; 1.5 μ L of 10 mM dTTP; 37.5 μ L of 0.4 mM Biotin-dATP; 50U of DNA Polymerase I Large (Klenow)
 802 fragment in 300 μ L NEBuffer2 1X). After 45min incubation at 37 °C, nuclei were centrifuged 5min
 803 at 3,000 rpm and ligation was performed 4h at 16 °C using 10,000 cohesive end units of T4 DNA
 804 ligase (NEB) in 1.2 mL of ligation buffer (120 μ L of 10X T4 DNA Ligase Buffer; 100 μ L of 10% Triton
 805 X-100; 12 μ L of 10mg/ml BSA; 963 μ L of H₂O). After reversion of the cross-link, DNA was purified
 806 by phenol extraction and EtOH precipitation. Purified DNA was sonicated to obtain fragments of
 807 an average size of 300-400bp using a Bioruptor Pico (Diagenode; 8 cycles; 20s on and 60s off).
 808 3 μ g of sonicated DNA was used for library preparation. Briefly, biotinylated DNA was pulled down
 809 using 20 μ L of Dynabeads Myone T1 streptavidine beads in Binding Buffer (5 mM Tris-HCl pH7.5;
 810 0.5 mM EDTA; 1 M NaCl). End-repair and A-tailing were performed on beads using NEBnext library
 811 preparation end-repair and A-tailing modules (NEB). Illumina adaptors were ligated and libraries
 812 were amplified by 8 cycles of PCR. Resulting Hi-C libraries were first controlled for quality by low se-
 813 quencing depth on a NextSeq500 prior to higher sequencing depth on HiSeq2000. Hi-C data were
 814 processed using an in-house pipeline based on TADbit (*Serra et al., 2017*). Reads were mapped ac-
 815 cording to a fragment-based strategy: each side of the sequenced read was mapped in full length
 816 to the reference genome Human Dec. 2013 (GRCh38/hg38). In the case reads were not mapped
 817 when intra-read ligation sites were found, they were split. Individual split read fragments were then
 818 mapped independently. We used the TADbit filtering module to remove non-informative contacts
 819 and to create contact matrices as previously described (*Serra et al., 2017*) PCR duplicates were
 820 removed and the Hi-C filters applied corresponded to potential non-digested fragments (extra-
 821 dandling ends), non-ligated fragments (dandling-ends), self-circles and random breaks.

822 CNV

823 The copy number variation (CNV) analysis was estimated comparing the coverage obtained in the
 824 Hi-C datasets with the expected coverage for a diploid genome based on the density of restriction
 825 sites and genomic biases (*Vidal et al., 2018*). Indeed, the linear correlation between number of
 826 Hi-C contacts and number of restriction sites is lost in case of altered copy number allowing the
 827 estimation of a relative number of copy as compared to diploid chromosomes in each dataset.
 828 Such estimations are consistent with other analyses and with karyotyping (*Le Dily et al., 2014*).

829 Virtual 4C

830 Hi-C matrices were normalized for sequencing depth and genomic biases using OneD (*Vidal et al.,*
 831 *2018*) and further smoothed using a focal average. Virtual 4C plots were generated from the ma-
 832 trices locally normalized and expressed as normalized counts per thousands within the region.

833 Intra-TAD interactions between specific loci

834 Each bin of a TAD was labeled as part of a PgCR or TSS (or "others" if they did not belong to the
 835 previous types). We collected the observed contacts between the different types of bins and com-
 836 puted the expected contacts frequencies based on the genomic distance that separate each pair.

In the figure, results are expressed as Log2 of the ratio of observed contacts between the different types of pairs above the intra-TAD background.

Endometrial cancer samples (TCGA)

Raw count data from human endometrial cancer RNAseq samples (n=575) were downloaded from The Cancer Genome Atlas (TCGA), project TCGA-UCEC. Each sample was matched to its corresponding clinical metadata including FIGO stage and histologic type and only protein coding genes above arbitrary threshold (mean expression among samples > 100 counts) were kept for further analyses. Raw counts were normalized and transformed (variance stabilization) in DESeq2 package and later used for filtering Stage I samples by *ESR* and *PGR* expression levels. Considering the inherently heterogeneous nature of tissue samples, only samples clustered together above threshold were kept for further analysis (see **Figure 7–Figure Supplement 1A**). After filtering, the raw counts of 488 samples were used to find differential expressed genes (DEG) between stages, downstream heatmaps (pheatmap R package (*Kolde and Kolde, 2015*)) and Principal Component Analysis in DESeq2 and PCAtools packages (*Blighe et al., 2019*). To construct the normal distribution of intersected genes we randomly sampled (without replacement) a group of the same size of DEGs among all possible protein coding genes (GENCODE v38) 10,000 times using custom python scripting (code is publicly available at https://github.com/alelagreca/ishikawa_scripting) (*La Greca, 2021*), then intersected each one with PgCR-genes to produce the curve. Statistical significance of our non-random intersection was determined by calculating the 2.5 and 97.5 percentiles in the distribution (95% confidence interval).

Competing interests

The authors declare no competing interests.

Acknowledgments

We are grateful to members of the Beato and Saragüeta laboratories for help and suggestions.

Data Availability

All raw and processed sequencing data generated in this study have been submitted to the NCBI Gene Expression Omnibus (GEO) under accession number GSE139398. T47D ChIPseq data is available under GEO accession number GSE41466 (*Ballaré et al., 2013*) and Hi-C data in GEO accession GSE53463 (*Le Dily et al., 2014*). RNAseq datasets from proliferative (GSM3890623, GSM3890624, GSM3890625 and GSM3890626) and mid-secretory (GSM3890627, GSM3890628, GSM3890629, GSM3890630 and GSM3890631) human endometrium were obtained from GEO accession GSE132711 (SuperSeries GSE132713) (*Chi et al., 2020*). ChIPseq coverage data of proliferative and secretory normal endometrium were downloaded from GEO accession GSE132712 (SuperSeries GSE132713) (*Chi et al., 2020*). Human endometrial cancer RNAseq samples (n=575) were downloaded from The Cancer Genome Atlas (TCGA), project TCGA-UCEC. Additional normal and endometrial cancer samples (n=109) were accessed through CPTAC program in the National Cancer Institute using cptac platform installed with python (*Dou et al., 2020*).

References

- Agirre E**, Bellora N, Alló M, Pagès A, Bertucci P, Kornblihtt AR, Eyra E. A chromatin code for alternative splicing involving a putative association between CTCF and HP1 α proteins. *BMC Biol.* 2015 May; 13:31. doi: 10.1186/s12915-015-0141-5.
- Bai R**, Cui Z, Ma Y, Wu Y, Wang N, Huang L, Yao Q, Sun J. The NF- κ B-modulated miR-19a-3p enhances malignancy of human ovarian cancer cells through inhibition of IGFBP-3 expression. *Mol Carcinog.* 2019 12; 58(12):2254–2265. doi: 10.1002/mc.23113.

- 881 **Bailey TL**, Johnson J, Grant CE, Noble WS. The MEME Suite. *Nucleic Acids Res.* 2015 Jul; 43(W1):W39–49. doi:
882 10.1093/nar/gkv416.
- 883 **Ballaré C**, Castellano G, Gaveglia L, Althammer S, González-Vallinas J, Eyraas E, Le Dily F, Zaurin R, Soronellas D,
884 Vicent GP, Beato M. Nucleosome-driven transcription factor binding and gene regulation. *Mol Cell.* 2013 Jan;
885 49(1):67–79. doi: 10.1016/j.molcel.2012.10.019.
- 886 **Ballaré C**, Uhrig M, Bechtold T, Sancho E, Di Domenico M, Migliaccio A, Auricchio F, Beato M. Two domains of
887 the progesterone receptor interact with the estrogen receptor and are required for progesterone activation
888 of the c-Src/Erk pathway in mammalian cells. *Mol Cell Biol.* 2003 Mar; 23(6):1994–2008.
- 889 **Beato M**, Herrlich P, Schütz G. Steroid hormone receptors: many actors in search of a plot. *Cell.* 1995 Dec;
890 83(6):851–7.
- 891 **Blighe K**, Lewis M, Lun A, Blighe MK. Package ‘PCAtools’. . 2019; .
- 892 **Bolger AM**, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.*
893 2014 Aug; 30(15):2114–20. doi: 10.1093/bioinformatics/btu170.
- 894 **Buenrostro JD**, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native chromatin for fast and sen-
895 sitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods.*
896 2013 Dec; 10(12):1213–8. doi: 10.1038/nmeth.2688.
- 897 **Cancer Genome Atlas Research Network**, Albert Einstein College of Medicine, Analytical Biological Services,
898 Barretos Cancer Hospital, Baylor College of Medicine, Beckman Research Institute of City of Hope, Buck In-
899 stitute for Research on Aging, Canada's Michael Smith Genome Sciences Centre, Harvard Medical School,
900 Helen F Graham Cancer Center & Research Institute at Christiana Care Health Services, HudsonAlpha Insti-
901 tute for Biotechnology, ILSbio, LLC, Indiana University School of Medicine, Institute of Human Virology, In-
902 stitute for Systems Biology, International Genomics Consortium, Leidos Biomedical, Massachusetts General
903 Hospital, McDonnell Genome Institute at Washington University, Medical College of Wisconsin, et al. Inte-
904 grated genomic and molecular characterization of cervical cancer. *Nature.* 2017 03; 543(7645):378–384. doi:
905 10.1038/nature21386.
- 906 **Cannavò E**, Khoueiry P, Garfield DA, Geeleher P, Zichner T, Gustafson EH, Ciglar L, Korbel JO, Furlong EEM.
907 Shadow Enhancers Are Pervasive Features of Developmental Regulatory Networks. *Curr Biol.* 2016 Jan;
908 26(1):38–51. doi: 10.1016/j.cub.2015.11.034.
- 909 **Chi RPA**, Wang T, Adams N, Wu SP, Young SL, Spencer TE, DeMayo F. Human Endometrial Transcriptome and
910 Progesterone Receptor Cistrome Reveal Important Pathways and Epithelial Regulators. *J Clin Endocrinol*
911 *Metab.* 2020 Apr; 105(4). doi: 10.1210/clinem/dgz117.
- 912 **Chomczynski P**, Sacchi N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-
913 chloroform extraction. *Anal Biochem.* 1987 Apr; 162(1):156–9. doi: 10.1006/abio.1987.9999.
- 914 **Conneely OM**, Mulac-Jericevic B, Lydon JP. Progesterone-dependent regulation of female reproductive activity
915 by two distinct progesterone receptor isoforms. *Steroids.* 2003 Nov; 68(10-13):771–8.
- 916 **Diep CH**, Ahrendt H, Lange CA. Progesterone induces progesterone receptor gene (PGR) expression via
917 rapid activation of protein kinase pathways required for cooperative estrogen receptor alpha (ER) and
918 progesterone receptor (PR) genomic action at ER/PR target genes. *Steroids.* 2016 10; 114:48–58. doi:
919 10.1016/j.steroids.2016.09.004.
- 920 **Dou Y**, Kawaler EA, Cui Zhou D, Gritsenko MA, Huang C, Blumenberg L, Karpova A, Petyuk VA, Savage
921 SR, Satpathy S, Liu W, Wu Y, Tsai CF, Wen B, Li Z, Cao S, Moon J, Shi Z, Cornwell M, Wyczalkowski MA,
922 et al. Proteogenomic Characterization of Endometrial Carcinoma. *Cell.* 2020 02; 180(4):729–748.e26. doi:
923 10.1016/j.cell.2020.01.026.
- 924 **Droog M**, Nevedomskaya E, Dackus GM, Fles R, Kim Y, Hollema H, Mourits MJ, Nederlof PM, van Boven HH,
925 Linn SC, van Leeuwen FE, Wessels LFA, Zwart W. Estrogen receptor α yields treatment-specific enhancers
926 between morphologically similar endometrial tumors. *Proc Natl Acad Sci U S A.* 2017 02; 114(8):E1316–E1325.
927 doi: 10.1073/pnas.1615233114.
- 928 **Fresno C**, Fernández EA. RDAVIDWebService: a versatile R interface to DAVID. *Bioinformatics.* 2013 Nov;
929 29(21):2810–1. doi: 10.1093/bioinformatics/btt487.

- 930 **Gertz J**, Reddy TE, Varley KE, Garabedian MJ, Myers RM. Genistein and bisphenol A exposure cause estrogen
931 receptor 1 to bind thousands of sites in a cell type-specific manner. *Genome Res.* 2012 Nov; 22(11):2153–62.
932 doi: [10.1101/gr.135681.111](https://doi.org/10.1101/gr.135681.111).
- 933 **Graham JD**, Roman SD, McGowan E, Sutherland RL, Clarke CL. Preferential stimulation of human progesterone
934 receptor B expression by estrogen in T-47D human breast cancer cells. *J Biol Chem.* 1995 Dec; 270(51):30693–
935 700.
- 936 **Hata H**, Hamano M, Tsujii A, Togasaki Y, Kuramoto H, Nishida M, Anzai Y, Gurpide E. [In vitro study for hormones
937 and growth factors dependent cell proliferation of endometrial adenocarcinoma cells]. *Hum Cell.* 1993 Sep;
938 6(3):182–7.
- 939 **Hecht JL**, Mutter GL. Molecular and pathologic aspects of endometrial carcinogenesis. *J Clin Oncol.* 2006 Oct;
940 24(29):4783–91. doi: [10.1200/JCO.2006.06.7173](https://doi.org/10.1200/JCO.2006.06.7173).
- 941 **Hnisz D**, Schuijers J, Lin CY, Weintraub AS, Abraham BJ, Lee TI, Bradner JE, Young RA. Convergence of develop-
942 mental and oncogenic signaling pathways at transcriptional super-enhancers. *Mol Cell.* 2015 Apr; 58(2):362–
943 70. doi: [10.1016/j.molcel.2015.02.014](https://doi.org/10.1016/j.molcel.2015.02.014).
- 944 **Hovland AR**, Powell RL, Takimoto GS, Tung L, Horwitz KB. An N-terminal inhibitory function, IF, suppresses
945 transcription by the A-isoform but not the B-isoform of human progesterone receptors. *J Biol Chem.* 1998
946 Mar; 273(10):5455–60. doi: [10.1074/jbc.273.10.5455](https://doi.org/10.1074/jbc.273.10.5455).
- 947 **Huang DW**, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID
948 bioinformatics resources. *Nat Protoc.* 2009; 4(1):44–57. doi: [10.1038/nprot.2008.211](https://doi.org/10.1038/nprot.2008.211).
- 949 **Ishiwata I**, Kiguchi K, Ishiwata C, Soma M, Nakaguchi T, Ono I, Tachibana T, Hashimoto H, Ishikawa H, Nozawa
950 S. Histogenesis of hollow cell ball structure of ovarian and endometrial adenocarcinoma cells in vivo and in
951 vitro. *Hum Cell.* 1997 Sep; 10(3):209–16.
- 952 **Jongen V**, Briët J, de Jong R, ten Hoor K, Boezen M, van der Zee A, Nijman H, Hollema H. Expression of estrogen
953 receptor- α and - β and progesterone receptor-A and -B in a large cohort of patients with endometrioid
954 endometrial cancer. *Gynecol Oncol.* 2009 Mar; 112(3):537–42. doi: [10.1016/j.ygyno.2008.10.032](https://doi.org/10.1016/j.ygyno.2008.10.032).
- 955 **Kayisli UA**, Luk J, Guzeloglu-Kayisli O, Seval Y, Demir R, Arici A. Regulation of angiogenic activity of human
956 endometrial endothelial cells in culture by ovarian steroids. *J Clin Endocrinol Metab.* 2004 Nov; 89(11):5794–
957 802. doi: [10.1210/jc.2003-030820](https://doi.org/10.1210/jc.2003-030820).
- 958 **Khan A**, Mathelier A. Intervene: a tool for intersection and visualization of multiple gene or genomic region
959 sets. *BMC Bioinformatics.* 2017 May; 18(1):287. doi: [10.1186/s12859-017-1708-7](https://doi.org/10.1186/s12859-017-1708-7).
- 960 **Khan A**, Zhang X. dbSUPER: a database of super-enhancers in mouse and human genome. *Nucleic Acids Res.*
961 2016 Jan; 44(D1):D164–71. doi: [10.1093/nar/gkv1002](https://doi.org/10.1093/nar/gkv1002).
- 962 **Kolde R**, Kolde MR. Package 'pheatmap'. *R Package.* 2015; 1(7).
- 963 **Kraus WL**, Katzenellenbogen BS. Regulation of progesterone receptor gene expression and growth in the rat
964 uterus: modulation of estrogen actions by progesterone and sex steroid hormone antagonists. *Endocrinol-*
965 *ogy.* 1993 Jun; 132(6):2371–9. doi: [10.1210/endo.132.6.8504742](https://doi.org/10.1210/endo.132.6.8504742).
- 966 **Kreizman-Shefer H**, Pricop J, Goldman S, Elmalah I, Shalev E. Distribution of estrogen and progesterone recep-
967 tors isoforms in endometrial cancer. *Diagn Pathol.* 2014 Mar; 9:77. doi: [10.1186/1746-1596-9-77](https://doi.org/10.1186/1746-1596-9-77).
- 968 **Kremisky I**, Bellora N, Eyra E. A Quantitative Profiling Tool for Diverse Genomic Data Types Reveals Poten-
969 tial Associations between Chromatin and Pre-mRNA Processing. *PLoS One.* 2015; 10(7):e0132448. doi:
970 [10.1371/journal.pone.0132448](https://doi.org/10.1371/journal.pone.0132448).
- 971 **La Greca AD**, Ishikawa scripts. 25a4757.; 2021. https://github.com/alelagreca/ishikawa_scripting.
- 972 **Lambert SA**, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT. The
973 Human Transcription Factors. *Cell.* 2018 02; 172(4):650–665. doi: [10.1016/j.cell.2018.01.029](https://doi.org/10.1016/j.cell.2018.01.029).
- 974 **Le Dily F**, Baù D, Pohl A, Vicent GP, Serra F, Soronellas D, Castellano G, Wright RHG, Ballare C, Filion
975 G, Marti-Renom MA, Beato M. Distinct structural transitions of chromatin topological domains corre-
976 late with coordinated hormone-induced gene regulation. *Genes Dev.* 2014 Oct; 28(19):2151–62. doi:
977 [10.1101/gad.241422.114](https://doi.org/10.1101/gad.241422.114).

- 978 **Le Dily F**, Vidal E, Cuartero Y, Quilez J, Nacht AS, Vicent GP, Carbonell-Caballero J, Sharma P, Villanueva-Cañas
979 JL, Ferrari R, De Llobet LI, Verde G, Wright RHG, Beato M. Hormone-control regions mediate steroid receptor-
980 dependent genome organization. *Genome Res.* 2019 01; 29(1):29–39. doi: [10.1101/gr.243824.118](https://doi.org/10.1101/gr.243824.118).
- 981 **Leslie KK**, Kumar NS, Richer J, Owen G, Takimoto G, Horwitz KB, Lange C. Differential expression of the A
982 and B isoforms of progesterone receptor in human endometrial cancer cells. Only progesterone receptor B
983 is induced by estrogen and associated with strong transcriptional activation. *Ann N Y Acad Sci.* 1997 Sep;
984 828:17–26. doi: [10.1111/j.1749-6632.1997.tb48520.x](https://doi.org/10.1111/j.1749-6632.1997.tb48520.x).
- 985 **Li H**, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics.* 2009
986 Jul; 25(14):1754–60. doi: [10.1093/bioinformatics/btp324](https://doi.org/10.1093/bioinformatics/btp324).
- 987 **Liang H**, Cheung LWT, Li J, Ju Z, Yu S, Stemke-Hale K, Dogruluk T, Lu Y, Liu X, Gu C, Guo W, Scherer SE, Carter H,
988 Westin SN, Dyer MD, Verhaak RGW, Zhang F, Karchin R, Liu CG, Lu KH, et al. Whole-exome sequencing com-
989 bined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer. *Genome*
990 *Res.* 2012 Nov; 22(11):2120–9. doi: [10.1101/gr.137596.112](https://doi.org/10.1101/gr.137596.112).
- 991 **Lieberman-Aiden E**, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ,
992 Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny
993 LA, Lander ES, Dekker J. Comprehensive mapping of long-range interactions reveals folding principles of the
994 human genome. *Science.* 2009 Oct; 326(5950):289–93. doi: [10.1126/science.1181369](https://doi.org/10.1126/science.1181369).
- 995 **Liu J**, Feng M, Li S, Nie S, Wang H, Wu S, Qiu J, Zhang J, Cheng W. Identification of molecular markers associated
996 with the progression and prognosis of endometrial cancer: a bioinformatic study. *Cancer Cell Int.* 2020;
997 20:59. doi: [10.1186/s12935-020-1140-3](https://doi.org/10.1186/s12935-020-1140-3).
- 998 **Love MI**, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with
999 DESeq2. *Genome Biol.* 2014; 15(12):550. doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8).
- 1000 **Lydon JP**, DeMayo FJ, Funk CR, Mani SK, Hughes AR, Montgomery CA Jr, Shyamala G, Conneely OM, O'Malley
1001 BW. Mice lacking progesterone receptor exhibit pleiotropic reproductive abnormalities. *Genes Dev.* 1995
1002 Sep; 9(18):2266–78.
- 1003 **Ma J**, Li D, Kong FF, Yang D, Yang H, Ma XX. miR-302a-5p/367-3p-HMGA2 axis regulates malignant processes
1004 during endometrial cancer development. *J Exp Clin Cancer Res.* 2018 Feb; 37(1):19. doi: [10.1186/s13046-](https://doi.org/10.1186/s13046-018-0686-6)
1005 [018-0686-6](https://doi.org/10.1186/s13046-018-0686-6).
- 1006 **McLean CY**, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, Wenger AM, Bejerano G. GREAT im-
1007 proves functional interpretation of cis-regulatory regions. *Nat Biotechnol.* 2010 May; 28(5):495–501. doi:
1008 [10.1038/nbt.1630](https://doi.org/10.1038/nbt.1630).
- 1009 **Miyamoto T**, Watanabe J, Hata H, Jobo T, Kawaguchi M, Hattori M, Saito M, Kuramoto H. Significance of pro-
1010 gesterone receptor-A and -B expressions in endometrial adenocarcinoma. *J Steroid Biochem Mol Biol.* 2004
1011 Oct; 92(3):111–8. doi: [10.1016/j.jsbmb.2004.07.007](https://doi.org/10.1016/j.jsbmb.2004.07.007).
- 1012 **Mohammed H**, Russell IA, Stark R, Rueda OM, Hickey TE, Tarulli GA, Serandour AA, Serandour AAA, Birrell SN,
1013 Bruna A, Saadi A, Menon S, Hadfield J, Pugh M, Raj GV, Brown GD, D'Santos C, Robinson JLL, Silva G, Launch-
1014 bury R, et al. Progesterone receptor modulates ER α action in breast cancer. *Nature.* 2015 Jul; 523(7560):313–7.
1015 doi: [10.1038/nature14583](https://doi.org/10.1038/nature14583).
- 1016 **Monsivais D**, Peng J, Kang Y, Matzuk MM. Activin-like kinase 5 (ALK5) inactivation in the mouse uterus re-
1017 sults in metastatic endometrial carcinoma. *Proc Natl Acad Sci U S A.* 2019 02; 116(9):3883–3892. doi:
1018 [10.1073/pnas.1806838116](https://doi.org/10.1073/pnas.1806838116).
- 1019 **Monte NM**, Webster KA, Neuberg D, Dressler GR, Mutter GL. Joint loss of PAX2 and PTEN expression in endome-
1020 trial precancers and cancer. *Cancer Res.* 2010 Aug; 70(15):6225–32. doi: [10.1158/0008-5472.CAN-10-0149](https://doi.org/10.1158/0008-5472.CAN-10-0149).
- 1021 **Mote PA**, Balleine RL, McGowan EM, Clarke CL. Colocalization of progesterone receptors A and B by dual
1022 immunofluorescent histochemistry in human endometrium during the menstrual cycle. *J Clin Endocrinol*
1023 *Metab.* 1999 Aug; 84(8):2963–71. doi: [10.1210/jcem.84.8.5928](https://doi.org/10.1210/jcem.84.8.5928).
- 1024 **Mulac-Jericevic B**, Mullinax RA, DeMayo FJ, Lydon JP, Conneely OM. Subgroup of reproductive functions of
1025 progesterone mediated by progesterone receptor-B isoform. *Science.* 2000 Sep; 289(5485):1751–4. doi:
1026 [10.1126/science.289.5485.1751](https://doi.org/10.1126/science.289.5485.1751).
- 1027 **Mulac-Jericevic B**, Conneely OM. Reproductive tissue selective actions of progesterone receptors. *Reproduc-*
1028 *tion.* 2004 Aug; 128(2):139–46. doi: [10.1530/rep.1.00189](https://doi.org/10.1530/rep.1.00189).

- 1029 **Nacht AS**, Ferrari R, Zaurin R, Scabia V, Carbonell-Caballero J, Le Dily F, Quilez J, Leopoldi A, Brisken C, Beato M,
1030 Vicent GP. C/EBP α mediates the growth inhibitory effect of progestins on breast cancer cells. *EMBO J*. 2019
1031 09; 38(18):e101426. doi: [10.15252/embj.2018101426](https://doi.org/10.15252/embj.2018101426).
- 1032 **Nacht AS**, Pohl A, Zaurin R, Soronellas D, Quilez J, Sharma P, Wright RH, Beato M, Vicent GP. Hormone-
1033 induced repression of genes requires BRG1-mediated H1.2 deposition at target promoters. *EMBO J*. 2016
1034 08; 35(16):1822–43. doi: [10.15252/embj.201593260](https://doi.org/10.15252/embj.201593260).
- 1035 **Need EF**, Selth LA, Trotta AP, Leach DA, Giorgio L, O'Loughlin MA, Smith E, Gill PG, Ingman WV, Graham JD,
1036 Buchanan G. The unique transcriptional response produced by concurrent estrogen and progesterone treat-
1037 ment in breast cancer cells results in upregulation of growth factor pathways and switching from a Luminal
1038 A to a Basal-like subtype. *BMC Cancer*. 2015 Oct; 15:791. doi: [10.1186/s12885-015-1819-3](https://doi.org/10.1186/s12885-015-1819-3).
- 1039 **Nishida M**. The Ishikawa cells from birth to the present. *Hum Cell*. 2002 Sep; 15(3):104–17.
- 1040 **Quick CM**, Laury AR, Monte NM, Mutter GL. Utility of PAX2 as a marker for diagnosis of endometrial intraep-
1041 ithelial neoplasia. *Am J Clin Pathol*. 2012 Nov; 138(5):678–84. doi: [10.1309/AJCP80MLT7KDWLMF](https://doi.org/10.1309/AJCP80MLT7KDWLMF).
- 1042 **Quinlan AR**. BEDTools: The Swiss-Army Tool for Genome Feature Analysis. *Curr Protoc Bioinformatics*. 2014
1043 Sep; 47:11.12.1–34. doi: [10.1002/0471250953.bi1112s47](https://doi.org/10.1002/0471250953.bi1112s47).
- 1044 **Raffone A**, Travaglino A, Saccone G, Mascolo M, Insabato L, Mollo A, De Placido G, Zullo F. PAX2 in en-
1045 dometrial carcinogenesis and in differential diagnosis of endometrial hyperplasia: A systematic review
1046 and meta-analysis of diagnostic accuracy. *Acta Obstet Gynecol Scand*. 2019 03; 98(3):287–299. doi:
1047 [10.1111/aogs.13512](https://doi.org/10.1111/aogs.13512).
- 1048 **Rao SSP**, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD,
1049 Lander ES, Aiden EL. A 3D map of the human genome at kilobase resolution reveals principles of chromatin
1050 looping. *Cell*. 2014 Dec; 159(7):1665–80. doi: [10.1016/j.cell.2014.11.021](https://doi.org/10.1016/j.cell.2014.11.021).
- 1051 **Rasmussen M**, Tan S, Somisetty VS, Hutin D, Olafsen NE, Moen A, Anonsen JH, Grant DM, Matthews J. PARP7 and
1052 Mono-ADP-Ribosylation Negatively Regulate Estrogen Receptor α Signaling in Human Breast Cancer Cells.
1053 *Cells*. 2021 03; 10(3). doi: [10.3390/cells10030623](https://doi.org/10.3390/cells10030623).
- 1054 **Ren Y**, Liu X, Ma D, Feng Y, Zhong N. Down-regulation of the progesterone receptor by the methylation of
1055 progesterone receptor gene in endometrial cancer cells. *Cancer Genet Cytogenet*. 2007 Jun; 175(2):107–16.
1056 doi: [10.1016/j.cancergencyto.2007.02.002](https://doi.org/10.1016/j.cancergencyto.2007.02.002).
- 1057 **Rodriguez AC**, Vahrenkamp JM, Berrett KC, Clark KA, Guillen KP, Scherer SD, Yang CH, Welm BE, Janát-Amsbury
1058 MM, Graves BJ, Gertz J. ETV4 Is Necessary for Estrogen Signaling and Growth in Endometrial Cancer Cells.
1059 *Cancer Res*. 2020 Mar; 80(6):1234–1245. doi: [10.1158/0008-5472.CAN-19-1382](https://doi.org/10.1158/0008-5472.CAN-19-1382).
- 1060 **Sakaguchi H**, Fujimoto J, Hong BL, Nakagawa Y, Tamaya T. Drastic decrease of progesterone receptor form B
1061 but not A mRNA reflects poor patient prognosis in endometrial cancers. *Gynecol Oncol*. 2004 May; 93(2):394–
1062 9. doi: [10.1016/j.ygyno.2004.01.042](https://doi.org/10.1016/j.ygyno.2004.01.042).
- 1063 **Sanderson PA**, Critchley HOD, Williams ARW, Arends MJ, Saunders PTK. New concepts for an old prob-
1064 lem: the diagnosis of endometrial hyperplasia. *Hum Reprod Update*. 2017 03; 23(2):232–254. doi:
1065 [10.1093/humupd/dmw042](https://doi.org/10.1093/humupd/dmw042).
- 1066 **Sasaki M**, Dharia A, Oh BR, Tanaka Y, Fujimoto S, Dahiya R. Progesterone receptor B gene inactivation and CpG
1067 hypermethylation in human uterine endometrial cancer. *Cancer Res*. 2001 Jan; 61(1):97–102.
- 1068 **Serra F**, Baù D, Goodstadt M, Castillo D, Filion GJ, Marti-Renom MA. Automatic analysis and 3D-modelling of
1069 Hi-C data using TADbit reveals structural features of the fly chromatin colors. *PLoS Comput Biol*. 2017 Jul;
1070 13(7):e1005665. doi: [10.1371/journal.pcbi.1005665](https://doi.org/10.1371/journal.pcbi.1005665).
- 1071 **Shao R**. Progesterone receptor isoforms A and B: new insights into the mechanism of progesterone resistance
1072 for the treatment of endometrial carcinoma. *Ecancermedicalscience*. 2013 Dec; 7:381. doi: [10.3332/ecan-
1073 cer.2013.381](https://doi.org/10.3332/ecancer.2013.381).
- 1074 **Singhal H**, Greene ME, Tarulli G, Zarnke AL, Bourgo RJ, Laine M, Chang YF, Ma S, Dembo AG, Raj GV, Hickey TE,
1075 Tilley WD, Greene GL. Genomic agonism and phenotypic antagonism between estrogen and progesterone
1076 receptors in breast cancer. *Sci Adv*. 2016 06; 2(6):e1501924. doi: [10.1126/sciadv.1501924](https://doi.org/10.1126/sciadv.1501924).
- 1077 **Stadhouders R**, Filion GJ, Graf T. Transcription factors and 3D genome conformation in cell-fate decisions.
1078 *Nature*. 2019 May; 569(7756):345–354. doi: [10.1038/s41586-019-1182-7](https://doi.org/10.1038/s41586-019-1182-7).

- 1079 **Strutt H**, Paro R. Mapping DNA target sites of chromatin proteins in vivo by formaldehyde crosslinking. *Methods Mol Biol.* 1999; 119:455–67. doi: [10.1385/1-59259-681-9:455](https://doi.org/10.1385/1-59259-681-9:455).
1080
- 1081 **Subramanian A**, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub
1082 TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpret-
1083 ing genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005 Oct; 102(43):15545–50. doi:
1084 [10.1073/pnas.0506580102](https://doi.org/10.1073/pnas.0506580102).
- 1085 **Trapnell C**, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics.* 2009
1086 May; 25(9):1105–11. doi: [10.1093/bioinformatics/btp120](https://doi.org/10.1093/bioinformatics/btp120).
- 1087 **Trapnell C**, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. Transcript
1088 assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell
1089 differentiation. *Nat Biotechnol.* 2010 May; 28(5):511–5. doi: [10.1038/nbt.1621](https://doi.org/10.1038/nbt.1621).
- 1090 **Vahrenkamp JM**, Yang CH, Rodriguez AC, Almomen A, Berrett KC, Trujillo AN, Guillen KP, Welm BE, Jarboe EA,
1091 Janat-Amsbury MM, Gertz J. Clinical and Genomic Crosstalk between Glucocorticoid Receptor and Estrogen
1092 Receptor α In Endometrial Cancer. *Cell Rep.* 2018 03; 22(11):2995–3005. doi: [10.1016/j.celrep.2018.02.076](https://doi.org/10.1016/j.celrep.2018.02.076).
- 1093 **Vallejo G**, Ballaré C, Barañao JL, Beato M, Saragüeta P. Progesterin activation of nongenomic pathways via cross
1094 talk of progesterone receptor with estrogen receptor beta induces proliferation of endometrial stromal cells.
1095 *Mol Endocrinol.* 2005 Dec; 19(12):3023–37. doi: [10.1210/me.2005-0016](https://doi.org/10.1210/me.2005-0016).
- 1096 **Vicent GP**, Nacht AS, Font-Mateu J, Castellano G, Gaveglia L, Ballaré C, Beato M. Four enzymes cooperate to
1097 displace histone H1 during the first minute of hormonal gene activation. *Genes Dev.* 2011 Apr; 25(8):845–62.
1098 doi: [10.1101/gad.621811](https://doi.org/10.1101/gad.621811).
- 1099 **Vidal E**, le Dily F, Quilez J, Stadhouders R, Cuartero Y, Graf T, Marti-Renom MA, Beato M, Filion GJ. OneD:
1100 increasing reproducibility of Hi-C samples with abnormal karyotypes. *Nucleic Acids Res.* 2018 05; 46(8):e49.
1101 doi: [10.1093/nar/gky064](https://doi.org/10.1093/nar/gky064).
- 1102 **Weirauch MT**, Yang A, Albu M, Cote AG, Montenegro-Montero A, Drewe P, Najafabadi HS, Lambert SA, Mann I,
1103 Cook K, Zheng H, Goity A, van Bakel H, Lozano JC, Galli M, Lewsey MG, Huang E, Mukherjee T, Chen X, Reece-
1104 Hoyer JS, et al. Determination and inference of eukaryotic transcription factor sequence specificity. *Cell.*
1105 2014 Sep; 158(6):1431–1443. doi: [10.1016/j.cell.2014.08.009](https://doi.org/10.1016/j.cell.2014.08.009).
- 1106 **Xiong Y**, Dowdy SC, Gonzalez Bosquet J, Zhao Y, Eberhardt NL, Podratz KC, Jiang SW. Epigenetic-mediated
1107 upregulation of progesterone receptor B gene in endometrial cancer cell lines. *Gynecol Oncol.* 2005 Oct;
1108 99(1):135–41. doi: [10.1016/j.ygyno.2005.05.035](https://doi.org/10.1016/j.ygyno.2005.05.035).
- 1109 **Yang S**, Xiao X, Jia Y, Liu X, Zhang Y, Wang X, Winters CJ, Devor EJ, Meng X, Thiel KW, Leslie KK. Epigenetic modifi-
1110 cation restores functional PR expression in endometrial cancer cells. *Curr Pharm Des.* 2014; 20(11):1874–80.
- 1111 **Zhang X**, Choi PS, Francis JM, Imielinski M, Watanabe H, Cherniack AD, Meyerson M. Identification of focally
1112 amplified lineage-specific super-enhancers in human epithelial cancers. *Nat Genet.* 2016 Feb; 48(2):176–82.
1113 doi: [10.1038/ng.3470](https://doi.org/10.1038/ng.3470).
- 1114 **Zhang Y**, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS.
1115 Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* 2008; 9(9):R137. doi: [10.1186/gb-2008-9-9-r137](https://doi.org/10.1186/gb-2008-9-9-r137).

Figure 1-Figure supplement 1. (A) Cells were pretreated with ICI182780 1 μ M (ICI 10×10^{-6} M) and then treated with vehicle (OH) and E2 10 nM (E2) for 48h. Number of cells was determined and presented as in Figure 1A. (***) $p < 0.001$. (B) Cells were pretreated with RU486 1 μ M (RU 486 10×10^{-6} M) and later treated with vehicle (OH), E2 10 nM (E2) and E2 and R5020 (E2+R5020) for 48h. Number of cells was determined and expressed as in A. (**) $p < 0.01$; (***) $p < 0.001$; ns: not significant. (C) BrdU incorporation was measured in untreated cells (control) and treated with vehicle (OH), E2 10 nM (E2), R5020 10 nM (R5020), E2 and R5020 (E2+R5020) and TSA 250 nM (TSA) for 15-18h. Mean fold change in BrdU positive cells for each treatment over Control $\pm$ SE of three independent experiments is shown. (D) Percentage of BrdU positive cells in each phase of the cell cycle for treatments shown in D: %G1 (white bars), %S (grey bars) and %G2-M (black bars). (E) PR (PRB and PRA) and ERα (ERa) protein levels of serum-deprived Ishikawa cells were assayed by western blot (Ishikawa) in two different conditions: phenol red-positive medium (phenol red +) and phenol red-negative medium (phenol red -) with 5% dextran-charcoal-treated serum (DCC-FBS). Alfa tubulin was used as loading control (tub-a). Western blot image at the bottom of the figure shows PR protein levels in non-pretreated (no preE2) and 12h E2-pretreated (preE2) Ishikawa cells using alfa tubulin as loading control (tub-a). (F) Western blot image showing PR protein levels in Ishikawa cells pretreated with E2 for 12h (preE2) and non-pretreated (no preE2). Alfa tubulin was used as loading control (tub-a). (G) mRNA levels of ERα, ERβ and actin in serum-deprived Ishikawa cells. (***) $p < 0.001$. (H) Representative image of anti-ERα detection in Ishikawa cells treated or not with R5020 for 60min. Quantification was performed as in Figure 1. Nuclei were revealed with DAPI (blue).

Figure 1-Figure supplement 2. (A) Overrepresented Biological Processes in RNAseq results from Ishikawa cells treated with R5020 and E2 for 12h analysed using DAVID web-based tool. Only the top 20 terms with $p < 0.05$ were selected for each hormone. (B) Venn Diagram represents intersection between differentially expressed genes (v.T0; $\log_2FC = \pm 0.8$, $q < 0.05$) in Ishikawa cells treated with R5020 and E2 for 12h. Upregulated and downregulated genes are indicated by Up and Dw, respectively. (C) Normalized expression profiles of genes located in the proximity of steroid receptor binding. RNAseq results expressed in RPKM for genes *TGFA*, *PGR* and *CCND2* displayed in IGV-based environment. Tracks for untreated (T0) and R5020 10nM (R5020) and E2 10nM (E2) 12h treatments were represented in the same vertical viewing range value.

Figure 1-Figure supplement 3. (A) Representative image of western blot assays using protein extracts from Ishikawa and T47D cells. 30 μ g (1) of T47D extracts were seeded followed by 10 μ g (1/3), 5 μ g (1/6), 3.3 μ g (1/9) and 2 μ g (1/15) of the same extract. 30 μ g of protein extract from Ishikawa cells (Ishi) were loaded on the same gel. PR isoforms (PRA and PRB) levels were evaluated and tubulin (tub-a) as control. Image was processed to remove lanes that did not contribute to the present analysis. (B) Quantification of PR protein levels in Ishikawa and T47D cells expressed as mean units of PRA+PRB relative to tubulin protein levels (tub-a) of three independent experiments. Numbers above bars indicate value of quantification relative to tubulin determined by densitometry of western blot assays. (*) $p < 0.05$. (C) Copy number variation (CNV) analysis of 100kb regions distributed over all chromosomes (hg19 UCSC) in T47D (upper panel) and Ishikawa (lower panel) cell lines. Black dotted lines limit diploid (2) and tetraploid (4) numbers. Location of estrogen receptor 1 (ESR1) and progesterone receptor (PGR) loci are indicated with red arrows. Enlarged image of regions covering chromosome 6 (ESR1) and chromosome 11 (PGR) is shown below graph.

Figure 2-Figure supplement 1. (A) ChIPseq results using anti-PR antibody in Ishikawa cells. Cells were treated or not (T0) for 5min (R5), 30min (R30) and 60min (R60) with R5020 10nM. The number of PRbs for T0, R30 and R60 is indicated in the Venn Diagram. (B) Boxplot shows validated PR binding sites expressed as percentage of input (%input) for 0, 30 and 60min R5020-treated Ishikawa cells. Sites included in this analysis are located in: promoter of *ALPP* gene, intergenic sequence upstream of *ALPP* gene, intergenic sequence upstream of *EGFR* gene, distal promoter of *SERPINA3* gene, promoter of *RPS6KA1* and intragenic sequence in *TGFA* gene. Significance was evaluated by ANOVA followed by Tukey's multiple contrasts and significant results are indicated for each individual comparison. (ns) not significant; (*) $p < 0.05$. (C) Scatterplots depict correlation in peak signal intensity among samples. Number of sites in common (n) as well as correlation coefficient (r) and p-value are shown over each plot. Non-linear regression curves are indicated as red lines. (D) ChIPseq results using anti-ERα antibody in Ishikawa cells treated or not (T0) with E2 10nM for 5 (E5), 60 (E60) and 360min (E360). In the Venn Diagram are shown untreated (T0), E60 and E360 as well as the number of ER binding sites (ERbs) for each group. (E) Boxplots shows validated ERα binding sites expressed as percentage of input (%input) for 0, 60 and 360min E2-treated Ishikawa cells. Sites included in this analysis are located in: intergenic region upstream of *ALPP* gene, intergenic region upstream of *TGFA* gene, intragenic region upstream of *TGFA* gene and promoter of *TIPARP* gene. Significance was evaluated by ANOVA followed by Tukey multiple contrast tests and significant results are indicated for each individual comparison. (ns) not significant; (*) $p < 0.05$. (F) Scatterplots as in C depict correlation in peak signal intensity among samples immunoprecipitated with anti-ERα. (G) Heatmaps of PR ChIPseq data from untreated (-) and 60min R5020-treated (+) Ishikawa and T47D cells. Regions were defined inside a window of -5kb to +5kb centered in peak summit. Intensity of the signal corresponds to the number of reads in the region. ChIPseq data are organized according to the Venn diagram scheme on the top right side of the figure as follows: I, PRbs solely found in Ishikawa cells; II, PRbs shared by both cell lines. De novo discovered motifs in I and II are indicated as sequence logos to the right of the map. PRE: progesterone response element. (H) Heat map constructed as in G of ERα ChIPseq data from untreated (-) and 60min E2-treated (+) Ishikawa and T47D cells. ChIPseq data are organized according to the Venn diagram scheme on the top right side of the figure as follows: III, ERbs solely found in Ishikawa cells; IV, ERbs shared by both cell lines. Discovered motifs in III and IV indicated presented as sequence logos are on the right site of the map. ERE: estrogen response element.

Figure 2-Figure supplement 2. (A) Validation of R5020-dependent PR binding in pretreated (+) or not (-) Ishikawa cells expressed as percentage of input (%input). Regions included are described in **Figure 2-Figure Supplement 1B**. Significance was evaluated by ANOVA followed by Tukey's multiple contrast tests. (ns) not significant; (*) $p < 0.05$; (**) $p < 0.01$. (B) Signal strength for PR and prePR binding of untreated (T0) and R5020-treated cells. Intensity was defined as reads per million relative to distance to site center in kb inside a window of ± 1 kb. (C) PR and ERalpha ChIPseq replicates from two independent experiments are displayed as a heatmap in a 10 kb window centered in peaks from pretreated cells. Labels on the left indicate common peaks between PR and ER (601), unique ER (924) and unique PR (5,064) binding sites. (D) (left side panels) Western blot of PR protein levels on Ishikawa and FPR Ishikawa (exogenous expression of triple flag-PR) cells. (right side panel) Representative image of western blot assays using protein extracts from FPR Ishikawa and T47D cells. 30 μ g (1) of T47D extract were seeded followed by 10 μ g (1/3), 5 μ g (1/6), 3.3 μ g (1/9) and 2 μ g (1/15) of the same extract. 30 μ g of extract from FPR Ishikawa cells (FPR) were loaded on the same gel. Anti-PR antibody was used to reveal PR isoforms (PRA and PRB) and tubulin (tub-a) as control. (E) Immunofluorescence of serine 294 phosphorylated PR (pS294) (green). Percentage of positive nuclei was calculated as green dots $\times$ 100/blue dots. (F) Recruitment of PR to *EGFR* enhancer sequence -described in **Figure 1F**- in FPR cells treated with R5020 10 nM and unspecific IgG control. Results are expressed as in **Figure 1F** (two independent experiments). (***) $p < 0.001$. (G) Shared binding sites among 30min- and 60min-treated FPR and parental Ishikawa cells. (H) Binding profiles (coverage signal) of prePRbs, PRbs, ERbs and FRPbs on PR peaks from normal proliferative and mid-secretory endometrium. Normal endometrium data was obtained from (*Chi et al., 2020*).

Figure 4-Figure supplement 1. Scatter plots showing fold enrichment values ($\log_2FE$) of 1,395 known TF binding motifs on (A) PRbs in group I and group II from Supp Fig S4G, (B) ERbs in group III and group IV from Figure S4H, (C) PRbs from Ishikawa and T47D cells, (D) ERbs and PRbs from Ishikawa cells, (E) ERbs and PRbs from T47D cells and (F) FPRbs and prePRbs. Combined p-values (product of x and y p-value multiplication) for enrichment analyses are indicated through the color key displayed at the lower right corner of the plots. Relevant motifs pointed on the plots correspond to NR3C1-4, ESR1, rESR, members of the Pax family (1, 2, 5, 6 and 9), GRHL1 (TF involved in epithelial development), SOX9 and FOXA1.

Figure 5—Figure supplement 1. (A) Normalized expression of detected PAX family members in untreated and hormone-treated Ishikawa RNAseq samples. Expression is shown for two independent experimental replicates. (B) Representative images of PAX2 signal (red) in cells treated or not with E2 for 60min and dapi staining (blue). Scale bar is shown in the figure and equals to 30 μ m. Distribution of intensities in nuclear signal for both conditions is shown to the right of the images. (C) Western blot image of PAX2 protein levels in untreated (T0) Ishikawa cells and treated with R5020 for 60min (R60) or E2 for 60min (E60). GAPDH was used as normalizing control. (D) PR (red) and PAX2 (green) staining in T47D cells treated with R5020 for 60min. Dapi was used for revealing nuclei.

Figure 5-Figure supplement 2. (A) IGV-based visualization of treated Ishikawa ChIPseq data. PR, ER α , prePR and PAX2 (R5020 and E2 sites) binding in regulatory regions of *ALPP* and *TIPARP* genes is shown in different tracks. (B) Ishikawa cells were incubated with a pool of siRNAs to reduce *PAX2* levels followed by 12h of E2 (pretreatment) and then treated with R5020 for 12h. RNA was collected and gene expression was analyzed by RT-qPCR. Results are shown as fold change over untreated cells (T0) after normalizing by *GAPDH* expression using $\Delta\Delta C_t$ method and expressed as mean fold change $\pm$ SD of three replicates.

Figure 6-Figure supplement 1. (A) Contact matrices of chr12 region at 100kb resolution obtained by In Nucleo Hi-C performed in two replicates of Ishikawa (Ishikawa_rep1 and Ishikawa_rep2) and T47D cells. (B) Hi-C datasets of untreated (T0) Ishikawa cells or treated with R5020 (R60), E2 (E60) and both hormones combined (RE60) were used to determine the spatial segregation of chromatin in open (A: green) or closed (B: purple) chromatin compartments compared to T47D cells. A fraction of chr14 is shown. (C) Relationship among samples was evaluated after clustering Hi-C datasets. Plot reflects differences as a function of distance. (D and E) Correlation analysis between Ishikawa replicates and T47D samples using eigen values determined from Hi-C datasets. Correlation coefficients are indicated in the figures. (F) Genome-wide proportions of 100kb windows of the genome found commonly in A or B in the two cell lines or showing divergent chromatin states (light green and light purple). (G) Histograms showing the average density of PR and ER binding sites (per 100kb) in the different types of regions mentioned in F. Below are the proportions of total PR and ERalpha binding sites found in those distinct regions. (H) Relative distance in bp from PRbs, FPRbs and prePRbs to PAXbs (both E2 and R5020 originated sites). Red dashed line marks overlapping events (distance = 0). Peaks located upstream of PAXbs are noted as negative distance values, while downstream peaks are positive. (I) Size distribution of PgCRs and TADs, expressed as $\log_{10}$. (J) Distribution of PgCRs in harbouring TADs divided into center (center), upstream end (up) and downstream end (down). (K) RNAseq results for genes *ALPP* and *PGR* displayed in a genome browser-based environment. Tracks for untreated (T0) and R5020 10nM (R5020) and E2 10nM (E2) 12h treatments were represented in the same vertical viewing range value (vvr). A different vvr was used for each gene in display.

Figure 7–Figure supplement 1. (A) Scatter plot of *ESR1* and *PGR* expression levels for Stage I endometrial cancer samples. Orthogonal black dashed lines denote the quadrants for sample selection (colored quadrant). (B) Normalized expression levels of *PAX* genes detected in TCGA-UCEC samples classified by FIGO stage (I to IV). (C) Normalized expression of genes obtained from intersection of endometrial cancer DEGs and PgCR-genes. Samples were arranged by FIGO stage (I to IV). (D) List of genes from heatmap in C classified as up and downregulated relative to Stage IV.