Research Resource: Genome-Wide Profiling of Progesterone Receptor Binding in the Mouse Uterus

Cory A. Rubel, Rainer B. Lanz, Ramakrishna Kommagani, Heather L. Franco, John P. Lydon, and Francesco J. DeMayo

Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030

Progesterone (P4) signaling through its nuclear transcription factor, the progesterone receptor (PR), is essential for normal uterine function. Although deregulation of PR-mediated signaling is known to underscore uterine dysfunction and a number of endometrial pathologies, the early molecular mechanisms of this deregulation are unclear. To address this issue, we have defined the genome-wide PR cistrome in the murine uterus using chromatin immunoprecipitation (ChIP) followed by massively parallel sequencing (ChIP-seq). In uteri of ovariectomized mice, we identified 6367 PR-binding sites in the absence of P4 ligand; however, this number increased at nearly 3-fold (18,432) after acute P4 exposure. Sequence analysis revealed that approximately 73% of these binding sites contain a progesterone response element or a half-site motif recognized by the PR. Many previously identified P4 target genes known to regulate uterine function were found to contain PR-binding sites, confirming the validity of our methodology. Interestingly, when the ChIP-seq data were coupled with our microarray expression data, we identified a novel regulatory role for uterine P4 in circadian rhythm gene expression, thereby uncovering a hitherto unexpected new circadian biology for P4 in this tissue. Further mining of the ChIP-seq data revealed *Sox17* as a direct transcriptional PR target gene in the uterus. As a member of the Sox transcription factor family, *Sox17* represents a potentially novel mediator of PR action in the murine uterus. Collectively, our first line of analysis of the uterine PR cistrome provides the first insights into the early molecular mechanisms that underpin normal uterine responsiveness to acute P4 exposure. Future analysis promises to reveal the PR interactome and, in turn, potential therapeutic targets for the diagnosis and/or treatment of endometrial dysfunction. **(*Molecular Endocrinology* 26: 1428–1442, 2012)**

A well-defined paradigm for the regulation of a healthy endometrium and the maintenance of a successful pregnancy through term is the role of progesterone (P4) signaling through its cognate receptor the progesterone receptor (PR) (1). Historically, focus on the pathology of multiple gynecological diseases has stressed the role of increased estrogen activity, largely ignoring the critical role that dysregulation of P4 signaling can contribute to the etiology of these diseases. Interruption of P4 signaling, occurring from the loss of the PR itself or through the loss of its interacting partners or downstream effectors, lead to a physiological state of P4 resistance. The diseases associated with reproductive health are an important issue affecting a significant population of women in the United States. The latest data from the Centers for Disease Control and Prevention indicate 11.8% of women ages 15–44 are affected with impaired fecundity, whereas the National Cancer Institute identifies more than 43,000 new cases of endometrial cancer per year (2). Current data have highlighted a complex and precisely timed communication occurring between uterine compartments to facilitate successful embryo implantation and initiation of successful pregnancy, which have underscored the identification of several mediators of P4

ISSN Print 0888-8809 ISSN Online 1944-9917
Printed in U.S.A.

doi: 10.1210/me.2011-1355 Received December 12, 2011. Accepted May 8, 2012.
First Published Online May 25, 2012

Abbreviations: BAR, Binary analysis results; BED, Browser Extensible Data; CEAS: *cis*-regulatory element annotation system; ChIP, chromatin immunoprecipitation; ChIP-seq, ChIP followed by massively parallel sequencing; DAVID, database for annotation, visualization, and integrated discovery; GO, gene ontology; HRE, hormone response element; MACS, model-based analysis of ChIP-sequencing; P4, progesterone; PR, progesterone receptor; PRE, progesterone response element; RT-qPCR, real-time quantitative PCR.

signaling (3). However, the mechanism by which P4 elicits these effects has yet to be determined. Understanding the precise mechanism of P4 regulation is of critical importance in our goal to identify the etiology of uterine diseases and to guide our ability to design treatments to minimize or eliminate their symptoms.

The PR is expressed as two major isoforms, PRA and PRB, which are generated from a single gene containing alternate translational start sites. Each isoform contains modular functional domains common to all nuclear receptors (4–6). Canonical PR signaling is activated through binding of its ligand P4. After P4 binding, the PR translocates to the nucleus and binds to the glucocorticoid receptor subgroup of steroid receptor hormone response elements (HRE) (7). This class of HRE typically consists of two 6-bp palindromes separated by a 3-bp spacer with a consensus sequence of 5′-AGAACAnnnTGTTCT-3′ and in addition to the PR, this sequence is also recognized by glucocorticoid receptor, androgen receptor, and mineralcorticoid receptor (8). In addition to the ability of PR to bind chromatin at specific nucleotide sequences, a mechanism exists whereby recruitment of the PR occurs independently of the presence of a functional HRE. Recruitment of the PR to transcriptional complexes of this type occurs through interactions with other transcription factors associating with their specific response elements (9–11). The capacity of the PR to recognize a highly varied and often unpredictable HRE sequence coupled with its ability to bind chromatin independently of any recognizable HRE underscores the diverse nature through which the PR can bind chromatin. This presents a complicated situation in which *in silico* prediction of functional progesterone receptor response elements (PRE) remains difficult and often inaccurate. To overcome the complications in identifying *bona fide* PR-regulatory sites, we must utilize next-generation technologies that will precisely define PR binding throughout the mouse uterine genome.

High-throughput expression arrays have aided in rapidly identifying numerous and previously unknown P4 target genes, the identification of which have given investigators greater insight as to how to select these targets for more in-depth study (12–14). Expression profiling data alone are incomplete however, as they are unable to yield information on the mechanism of PR regulation. It is imperative that we identify those groups of P4 target genes, the induction of which is directly in response to interaction by the PR on the chromatin or through induction by secondary interactions with downstream mediators. The solution to this problem has recently been overcome through the use of chromatin immunoprecipitation followed by deep sequencing (ChIP-seq). We have, for the first time, applied ChIP-seq technology under *in vivo* conditions of acute P4 administration to identify the genome-wide PR-binding regions in the mouse uterus. This has allowed us to correlate genes associated with PR binding with our expression data. In addition, we investigated the role of the hormone response element in directing PR binding to target genes and used functional pathway analysis to identify potential P4 target gene group clustering. Further, this analysis has allowed us to investigate the mechanism of PR regulation of downstream targets and to identify potential new transactivating factors that may cooperate with the PR to regulate uterine function. Novel findings in our study demonstrate a possible role of the PR in regulating circadian rhythm in the mouse uterus and the identification of *Sox17* as a potential new PR-interacting factor. Overall, our study provides an important coupling of expression data and ChIP-seq technologies that will help further define the global characterization of the PR.

Materials and Methods

ChIP

Female C57 mice (6 wk of age) were ovariectomized and left to heal for 2 wk. Mice were then administered vehicle (oil) or 1 mg progesterone (P4) sc followed by extraction of the uterus 1 h post injection. Mouse uteri were fixed with 1% formaldehyde for 15 min and quenched with 0.125 M glycine. Chromatin was isolated by adding lysis buffer, followed by disruption with a hand-held mechanical homogenizer followed by 30 strokes with a Dounce homogenizer. Lysates were sonicated, and the DNA was sheared to an average length of 300–500 bp. Genomic DNA (Input) was prepared by treating aliquots of chromatin with RNase, proteinase K, and heat for de-cross-linking, followed by ethanol precipitation. Pellets were resuspended, and the resulting DNA was quantified on a NanoDrop spectrophotometer. An aliquot of chromatin (30 μg) was precleared with protein A agarose beads (Invitrogen, Carlsbad, CA). ChIP was performed using an antibody against PR (sc-7208, Santa Cruz Biotechnology, Inc., Santa Cruz, CA). After incubation at 4 C overnight, protein A agarose beads were used to isolate the immune complexes. Complexes were washed, eluted from the beads with sodium dodecyl sulfate buffer, and treated with RNase and proteinase K. Cross-links were reversed by incubation overnight at 65 C, and ChIP DNA was purified by phenol-chloroform extraction and ethanol precipitation. For ChIP real-time quantitative PCR (RT-qPCR), primers were designed to amplify regions of PR binding identified through ChIP-seq. RT-qPCR were carried out in triplicate on specific genomic regions using SYBR Green Supermix (Bio-Rad Laboratories, Inc., Hercules, CA). The resulting signals were normalized to Input DNA.

Sox17-binding sites were identified using ChIP-on-Chip. ChIP and Input DNAs were amplified by whole-genome amplification (WGA) using the GenomePlex WGA Kit (Sigma, St. Louis, MO). The resulting amplified DNA were purified, quan-

tified, fragmented, and labeled using the DNA Terminal Labeling Kit from Affymetrix (Santa Clara, CA), and then hybridized to Affymetrix Mouse Tiling 2.0R Array A (containing chromosomes 10, 13, and 14) at 45 C overnight. Arrays were washed and scanned, and the resulting CEL files were analyzed using Affymetrix TAS software.

ChIP-seq analysis

PR and Input ChIP were performed by Active Motif, Inc. (Carlsbad, CA) on mouse uteri treated with vehicle or 1 mg P4 as described above. ChIP and input DNA were amplified using the Illumina ChIP-Seq DNA Sample Prep Kit. Briefly, DNA ends were polished and 59-phosphorylated using T4 DNA polymerase, Klenow polymerase, and T4 polynucleotide kinase. Addition of 3′-adenine to blunt ends using Klenow fragment (3′-5′ exo minus), Illumina genomic adapters were ligated and the sample was size fractionated to 175–225 bp on a 2% agarose gel. After amplification for 18 cycles with Phusion polymerase, the resulting DNA libraries were tested by RT-qPCR at the same specific genomic regions as the original ChIP DNA to assess quality of the amplification reactions. DNA libraries were sent to Illumina Sequencing Services for sequencing on a Genome Analyzer II. Sequences (35 bases) were aligned to the mouse genome (NCBI Build 37, July 2007) using Eland software. Aligns were extended *in silico* at their 39-ends to a length of 110–200 bp and assigned to 32-nucleotide (nt) bins along the genome. The resulting histograms were stored in BAR (binary analysis results) files. Peak locations were determined by applying a threshold of 18 (five consecutive bins containing .18 aligns) and storing the resulting intervals in Browser Extensible Data (BED) files (BED, Affymetrix TAS software). These files were analyzed using Genpathway proprietary software that provides comprehensive information on genomic annotation, peak metrics, and sample comparisons for all peaks (intervals).

Data analysis

The model-based analysis of ChIP-sequencing (MACS) (15) peak finding algorithm was used to normalize ChIP against Input control. Specifically, a P value of 10^{-10} was used with this software to identify ChIP peaks in this work by comparing them with input. Genes associated with intervals were assessed using three increasingly less stringent requirements; if it was within 10 kb, 25 kb, or 50 kb upstream or downstream of a gene it was counted. Sequence conservation, to identify phastCons scores, analysis of enriched motifs, and CEAS were performed using the Cistrome Analysis Pipeline software (http://cistrome.org/ap/) under default settings (16). For gene functional classifications, the public Database for Annotation, Visualization, and Integrated Discovery (DAVID, http://david.abcc.ncifcrf.gov/) running default settings was utilized (17). Manual analysis of PR-binding sequence for the presence of full PRE or half-site variations was performed using the text editing software notepad++, found at http://notepad-plus-plus.org/.

Quantitative real-time PCR analysis

Total RNA was isolated from frozen uterine tissue using the Trizol reagent (Invitrogen, Carlsbad, CA). RNA (1μg) was reverse transcribed into cDNA with Moloney Murine Leukemia Virus (M-MLV, Invitrogen) in a 20-μl volume. Expression levels of mRNA were measured by quantitative real time PCR TaqMan analysis using the ABI Prism 7700 Sequence Detector System according to the manufacturer's instructions (Applied Biosystems, Foster City, CA). Real time probes and primers were purchased from Applied Biosystems; standard curves were generated by serial dilution of a preparation of total cDNA isolated from whole-mouse uterus. All real-time PCR was done by using independent RNA sets. All mRNA quantities were normalized against 18S RNA using ABI rRNA control reagents.

Transfection and luciferase assay

Human endometrial epithelial cells (HEC-1A) were transfected according to manufacturer's instructions using Qiagen SuperFect reagent (QIAGEN, Valencia, CA) with 200 ng of +1.7-kb region of the mouse *Sox17* gene or a −19-kb region of the mouse *Ihh* gene containing a PR interval location fused to the pGL4.23 luciferase vector (Promega Corp., Madison, WI) along with 50 ng of human PRA, human PRB, mouse *Sox17* in pcDNA6 (kindly provided by Dr. James M. Wells, Cincinnati Children's Hospital Medical Center, Cincinnati, OH), or pcDNA3 empty vector control. The cells were treated with 10^{-8} M R5020 (Sigma, St. Louis, MO) or vehicle control (ethanol) in McCoy's 5A media (Invitrogen) for 24 h at 37 C. The cells were harvested and lysed using the Promega Passive lysis buffer (Promega Corp.). Luciferase activity was measured using the Promega luciferase assay system (Promega) in a Centro LB 960 luminometer (Berthold Technologies, Oak Ridge, TN). The luciferase activity was normalized to a *Renilla* luciferase internal control.

Immunofluorescence

Uteri were fixed overnight in 4% paraformaldehyde (vol/vol), thoroughly washed in 70% ethanol, processed, embedded in paraffin, and sectioned. Uterine sections from paraffin-embedded tissue were cut at 5 μm and mounted on silane-coated slides, deparaffinized, and rehydrated in a graded alcohol series. Antigen retrieval was achieved using Vector Antigen Unmasking Solution (Vector Laboratories, Burlingame, CA). After antigen retrieval, sections were washed six times in dH_2O and held in Tris-buffered saline-Tween-20 (TBST) and then incubated with anti-*Sox17* (1:50; Santa Cruz Biotechnology, Inc., Santa Cruz, CA) in 1% BSA in TBST for 1 h. Slides were then washed six times in dH_2O and incubated with antirabbit Alexa 488 (1:200; Invitrogen) for 45 min and then mounted with Vectashield with 4′,6-diamidino-2-phenylindole (Vector Laboratories).

Results

Identification of PR-binding sites in C57 mouse uterine chromatin

Genome-wide PR interaction sites in the mouse uterus were identified using ChIP followed by deep sequencing on uteri collected from ovariectomized C57 mice 1 h after treatment with vehicle or 1 mg P4 by sc injection. In addition to whole uterine analysis, following a protocol of mechanical separation (13), PR binding was identified from isolated uterine epithelium under P4 conditions only. More than 25 million tags of each sample were

mapped to unique locations in the mouse genome. A model-based analysis peak finding algorithm (MACS) was used to normalize immunoprecipitated chromatin against uterine input with a *P* value of 1×10^{-0} (<1% false discovery rate), producing a stringent cutoff for identifying binding sites (intervals). This high-confidence cutoff produced 6367 intervals in the vehicle-treated whole uteri sample, 18,432 intervals in the P4 whole-uteri sample, and 1415 intervals in the P4 epithelial sample; interval locations are provided in BED format in Supplemental Material 1 (SM1) published on The Endocrine Society's Journals online web site at http://mend.endojournals.org and can be accessed on the Gene Expression Omnibus website (GSE34927). Analysis of the vehicle *vs.* P4-treated whole uteri demonstrates that the two datasets share 4235 interval locations, whereas comparison of the P4-treated whole uteri with the epithelial isolated chromatin show an overlap of 1265 intervals (Fig. 1A). The discrepancy between the large numbers of whole-uterus intervals when compared with the epithelial-interval numbers demonstrates the physical limitations in the mechanical method of epithelial isolation crossed with ChIP-seq rather than from a lack of PR binding occurring in the uterine epithelium. This limitation is highlighted in thresholding difficulties observed with peak calling using a *P* value of 1×10^{-7} (estimated 10% false discovery rate). However, of the epithelial intervals that were identified, more than 90% of the PR-binding locations matched the P4 whole-uterus interval set indicating that our ChIP-seq data correspond to actual *in vivo* PR binding. Due to this limitation, our data will focus solely on whole-uterus PR binding, unless otherwise stated.

A comparison of the sequences from our identified PR intervals among various placental mammalian genomes demonstrated a high level of conservation within the regions of PR binding but not in surrounding regions (Fig. 1B). We then mapped the PR intervals to reference genes of the National Center for Biotechnology Information (NCBI) mouse genome database (NCBI37/mm9). Further analysis on the P4 dataset only used the *cis*-regulatory element annotation system (CEAS module) on the web-based application Cistrome (http://cistrome.org/ap/) (16), which provided a visual representation of the distribution of PR binding enrichment on each mouse chromosome (Supplemental Fig. 1 published on The Endocrine Society's Journals Online web site at http://mend.endojournals.org). We then mapped P4-PR interval distribution relative to genomic boundaries (Fig. 1C). Intervals located in the regions of physical gene coordinates showed a majority of PR intervals located within 10 kb of gene boundaries in all ChIP sets (60.9%), with 40.7% residing in introns, 1.1% within exons, and 10.9% within the 10-kb upstream region, consistent with previous data (18–22). When the frequency of P4 PR-binding sites located within the 10-kb upstream region was normalized to the length of this promoter region, as described previously (23), we find enrichment of binding sites at the promoters (Supplemental Fig. 2). Further, the distribution of P4 PR-binding locations in the 10-kb upstream promoter region shows a sharp increase in frequency of binding sites as they approach proximal to the transcriptional start site (Fig. 1D).

Due to the nature of the variable number of PR-binding sites within gene boundaries, ranging from a single interval per gene to several intervals lining entire gene regions, nonredundant gene calling was performed. This analysis allowed for the identified intervals to be reduced to a smaller set of unique genes that contain PR binding. Gene calling analysis was carried out with three different range criteria: an interval was counted for a specific gene if it was located within either ± 10 kb, 25 kb, or 50 kb from the transcriptional start and stop of a called gene. Further analyses performed solely on the ± 25-kb whole-uterus dataset identified 1007 unique genes with PR-binding sites in the vehicle-only condition and 6758 unique sites in the P4 condition only (Fig. 1E). Full gene lists are available in Supplemental Material 2 (SM2). To identify possible *in vivo* direct P4 target genes, it was necessary to combine our ChIP-seq dataset with that of our microarray analysis of uteri acutely treated with P4. This allows for screening of genes that not only contain PR binding but those genes the expression of which is impacted in response to P4 administration. To do this, we next compared our PR interval locations with the acute P4 transcriptional response by performing two independent microarray analyses, each performed on ovariectomized mice treated with vehicle or 1 mg P4 and killed 6 h after injection. Microarray gene lists are available as Supplemental Material 3 (SM3) and can be accessed on the Gene Expression Omnibus website (GSE34902). Each independent microarray yielded 1908 and 2360 unique P4-regulated genes, respectively, under threshold conditions requiring equal to or greater than 1.2-fold expression. In order for an expressed gene to be counted in our analysis, we used two stringency methods. To be counted, a high-stringency method required a gene to be present in both microarrays and have the same direction of expression. The high-stringency method yielded 926 genes (data not shown). In addition, we used a more relaxed stringency method that allowed for a gene to be counted if it was present in only one array. Both the high- and low-stringency methods counted a gene if it contained PR binding within ± 25 kb. This less stringent method yielded a total of 1714 P4-regulated genes with PR binding in either the

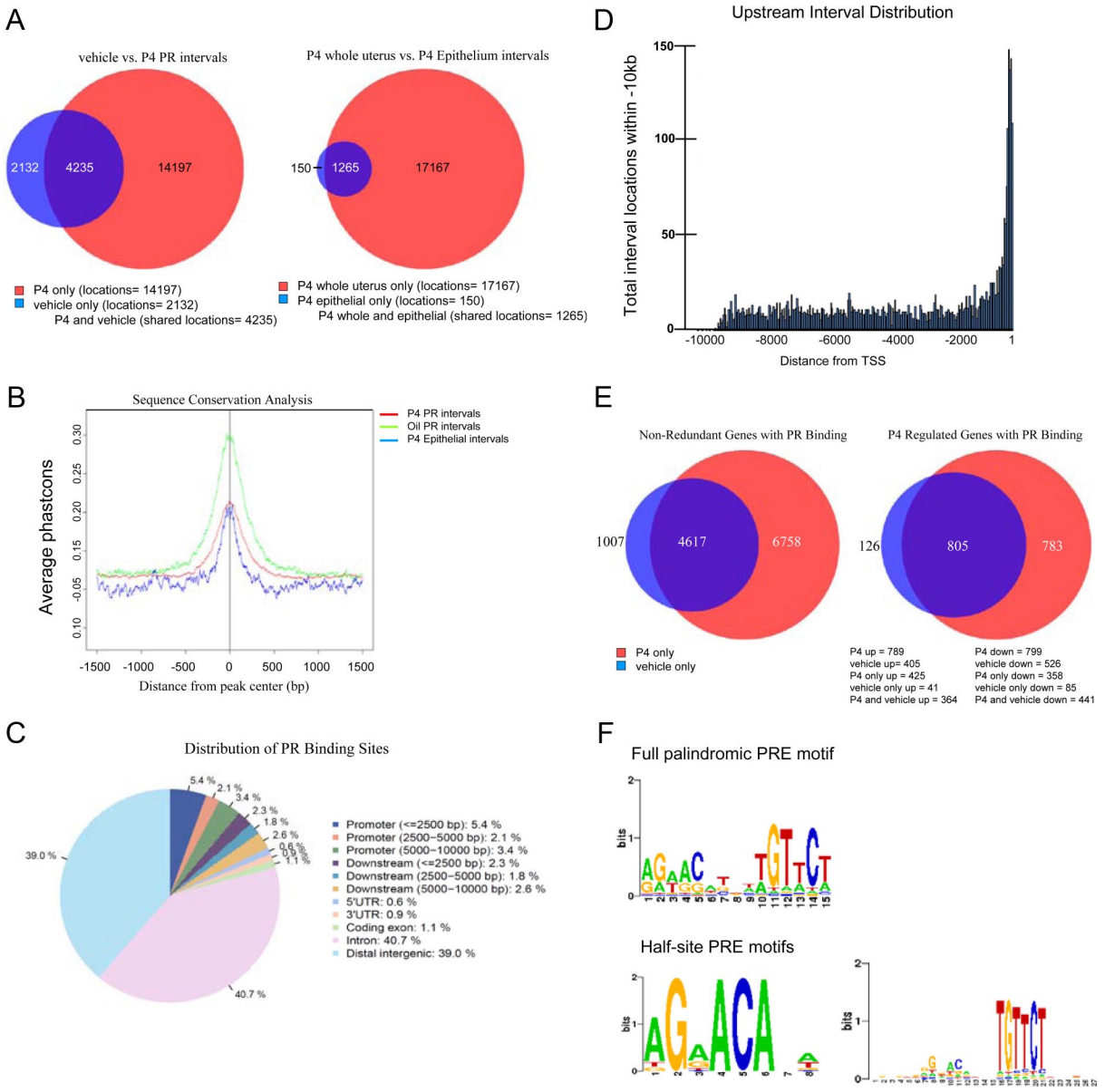

FIG. 1. Identification of PR binding locations with vehicle or P4 on whole-mouse uterine tissue and epithelial-isolated uterine tissue. BED files containing chromosome annotation and start-stop positions of the ChIP-seq sites are provided as Supplemental Material SM1. A, ChIP-seq PR-binding site counts and proximity relationships. Venn diagrams of the intersection of multiple intervals files between vehicle *vs.* P4-treated and whole *vs.* epithelial. B, Sequence conservation analysis. PR ChIP intervals were aligned at their centers and uniformly expanded 1500 bp in each direction, and phastCons scores were retrieved and averaged at each position. C, Distribution of genome-wide PR-binding locations throughout the murine uterus. PR-binding sites were analyzed using the CEAS module at Cistrome. Promoter region was defined in increments of 2500 bp to 10 kb. When the PR-binding site is within a gene, it is further defined as within the 5′-untranslated region (UTR), 3′-UTR, coding exon or intron. Intergenic region is defined as more than 10 kb from gene boundaries. D, In the P4-treated dataset, the frequency of PR-binding sites located within 10 kb upstream of the transcriptional start site were analyzed. E, Gene counts. Venn diagrams for nonredundant genes with PR-binding locations within ± 25 kb and those genes that are regulated by P4 after 6 h treatment by microarray (>1.2-fold) are presented. P4-regulated nonredundant genes are further grouped *underneath the diagram* as a list presenting the gene counts that are either up- or down-regulated in the absence or presence of P4. F, HRE motifs were found enriched in PR-binding sites in the P4-treated group using the SeqPos tool at Cistrome.

vehicle or P4 treatment condition, with 805 of those genes identified as having PR binding in both conditions (Fig. 1E). Additional analysis grouped P4-regulated genes with PR binding sites into datasets based on either up or down-regulation (Fig. 1E). Interestingly, P4-regulated PR binding sites were evenly split between up and down-regulated genes whereas PR binding under vehicle conditions yielded a slight preference for down-regulated genes. When this dataset was divided further into genes containing PR binding only in the presence of P4, there was a slight preference for up-regulated genes, whereas those with PR binding only in the absence of P4 demonstrated a slight preference for down-regulated genes.

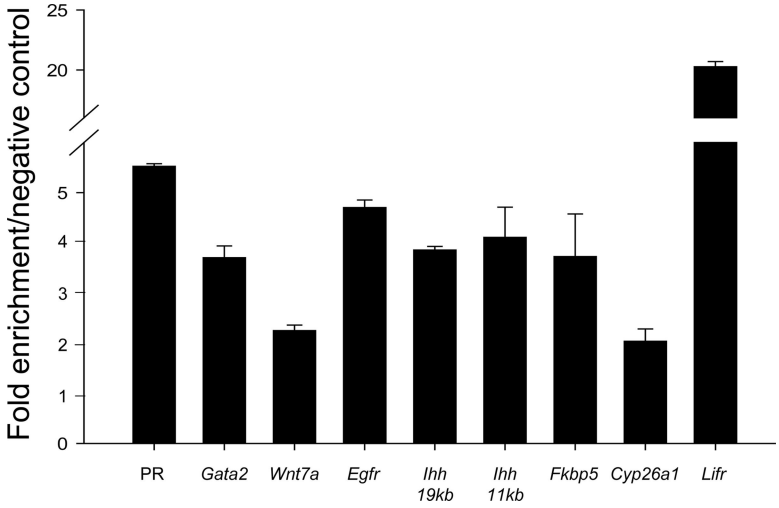

FIG. 2. ChIP validation of P4 target genes. Validation of PR-binding sites by ChIP-RT-qPCR with the PR H-190 antibody (Santa Cruz Biotechnology) on uteri isolated from ovariectomized C57 female mice treated 1 h with P4. Negative control (Untr) is a gene-deficient region. Data are represented as fold enrichment of PR binding of PR-location sites over that of the negative control region.

Evaluation of the interval regions of our vehicle and P4 datasets by the SeqPos tool at Cistrome for enriched motifs identified numerous HRE locations. HRE consistently placed as one of the top most enriched groups, which included full-length palindromic HRE and HRE half-site matrices, in good agreement with previously identified HRE motif searches (Fig. 1F). Manual analysis of each P4-PR interval sequence for the presence of a canonical palindromic HRE, adhering to the consensus AGAACAnnnTGTTCT, resulted in the identification of only 0.36% of the total intervals, which included the known P4 target gene *Fkbp5*. When sequences were searched for palindromes allowing for more variation in the consensus HRE nucleotide sequence, 23.1% of the 18,432 intervals were found to contain HRE; inclusion of canonical HRE half-sites in the analysis increased the percentage of HRE-containing intervals to 73.2%. To validate the results of our ChIP-seq experiment, intervals putatively containing PR binding along with a negative control region were analyzed by ChIP-RT-qPCR using a PR antibody. The determination of PR-binding locations to validate by ChIP-RT-qPCR was chosen to underlie the broad range with which the PR can interact on chromatin. Selected locations demonstrated PR binding on both up-regulated (*Gata2*, *Egfr*, *Ihh*, *Fkbp5*, *Cyp26a1*) and down-regulated (PR, *Wnt7a*, *Lifr*) P4 target genes (Fig. 2). In addition, PR binding was confirmed in the presence of a full-length PRE (PR, *Fkbp5*, *Ihh* 11 kb), with validation on locations with either the presence of PRE half-sites only (*Lifr*, *Cyp26a1*, *Gata2*, *Ihh* 19kb) or no identifiable PRE (*Egfr*, *Wnt7a*).

Uterine circadian biology is regulated by P4 signaling through the PR

Unique genes identified from PR-interval locations were submitted to the DAVID Bioinformatics Database for gene ontology (GO) analysis to identify clusters of gene groups with biologically related themes (17, 24). Full analyses and lists of enriched annotation terms for the gene groups are submitted as Supplemental Material (SM4). Using the gene functional classification tool, we identified terms associated with genes containing PR binding within ± 25 kb irrespective of treatment, terms associated with genes containing PR binding found only in the vehicle treatment, or those terms associated with genes containing PR binding found only in the P4 treatment. These terms were then separated into those gene groups that were demonstrated to associate with either up- (Table 1) or down-regulation (Table 2) of P4. Interestingly, it was noted that circadian rhythm/biological rhythms were found to be significantly enriched. Regulators of circadian rhythm, the endogenous timing system of a biological organism, have classically been shown to originate in the suprachiasmatic nucleus, although recently it has been demonstrated that individual tissues contain their own endogenous clock including those of the reproductive system (25). Our data show that several of the known genes regulating circadian rhythm contain putative PR-binding locations and display altered expression after P4 administration *in vivo*. We analyzed four genes with known circadian regulatory function, *Clock*, *Npas2*, *Cry1*, and *Per1*, and three nuclear receptors, *Nr1d2*, *Ppargc1a*, and *Rorc*, characterized for their negative regulation of the previous group, for further analysis. PR binding in the presence of P4 was confirmed by ChIP-RT-qPCR (Fig. 3A). Of the four circadian genes containing PR-binding locations, only *Npas2* contains a full-length PRE, whereas *Clock*, *Cry1*, and *Per1* lack even the presence of a PRE half-site, indicating that PR may need to cooperate with other factors for stable chromatin binding on these genes. We then assayed for expression of these genes in ovariectomized mice after 6 h administration of P4 (Fig. 3B). *Clock*, *Npas2*, *Cry1*, and *Per1* exhibited increased expression upon P4 administration, whereas *Nr1d2*, *Ppargc1a*, and *Rorc* showed a decrease in expression. Use of PR knockout mice (26) as controls demonstrated that P4 regulation of circadian gene expression requires the PR.

TABLE 1. GO analysis of groups of genes separated based on the presence of PR binding within 25 kb of RefSeq gene boundaries and shown by microarray analysis to be up-regulated by P4

25-kb up	Term	*P* value
GOTERM_CC_FAT	GO:0043232~intracellular non-membrane-bounded organelle	8.00E-23
UP_SEQ_FEATURE	Transit peptide:mitochondrion	2.19E-05
GOTERM_BP_FAT	GO:0006396~RNA processing	1.47E-27
GOTERM_MF_FAT	GO:0005524~ATP binding	1.50E-05
GOTERM_MF_FAT	GO:0008565~protein transporter activity	1.01E-27
GOTERM_BP_FAT	GO:0016481~negative regulation of transcription	2.10E-16
GOTERM_BP_FAT	GO:0045941~positive regulation of transcription	5.02E-30
KEGG_PATHWAY	mmu04710:circadian rhythm	0.0135227
SP_PIR_KEYWORDS	Biological rhythms	4.72E-05
Oil only up		
GOTERM_MF_FAT	GO:0005524~ATP binding	0.0012782
P4 only up		
GOTERM_CC_FAT	GO:0043232~intracellular non-membrane-bounded organelle	5.35E-14
GOTERM_MF_FAT	GO:0008565~protein transporter activity	8.29E-14
GOTERM_MF_FAT	GO:0005524~ATP binding	8.49E-29
INTERPRO	IPR000886:endoplasmic reticulum, targeting sequence	1.32E-08
GOTERM_BP_FAT	GO:0045941~positive regulation of transcription	6.11E-11
GOTERM_MF_FAT	GO:0005525~GTP binding	4.95E-07

Groups were further separated into those up-regulated genes that were found to contain PR binding in vehicle-only or P4-only conditions.

Identification of novel cooperative factors of PR biology

After GO analysis, interval sequences were evaluated further for motifs in addition to those belonging to the PR. Using the SeqPos tool at Cistrome, we searched for motifs that were found to be enriched in the vehicle and P4 ChIP datasets. The results of this analysis are represented with top significant motifs shown in the vehicle (Table 3) and P4 (Table 4) groups. Several significant factors were found with enrichment of their motifs located near PR-binding sites. Examples of identified motifs were those for ERE, *Foxo*, *Mcm*, ETS, retinoid X receptor-α/vitamin D receptor, *Sp1*, *Msx*, Jun-AP1, *Klfs*, *Gata*, and *Sox* transcription factors. Of particular interest were the latter two, with *Gata* binding of interest because previous expression data from our laboratory has identified *Gata2* as a novel P4 target gene (27). Additionally, it was observed that the motif analysis produced an abundance of *Sox* (sex-determining region on the Y chromosome-related high mobility group box) motifs, including those for *Sox17*. This factor was of interest because it has previously been reported as a possible P4 target gene and regulator of P4 gene expression in the rabbit endometrium (28). An examination of putative *Sox* motif locations throughout the mouse uterine genome demonstrated significant overlap with PR-binding sites located

TABLE 2. GO analysis of groups of genes separated based on the presence of PR binding within 25 kb of RefSeq gene boundaries and shown by microarray analysis to be down-regulated by P4

25 kb down	Term	*P* value
INTERPRO	IPR001849:Pleckstrin homology	1.37E-13
GOTERM_MF_FAT	GO:0030695~GTPase regulator activity	7.91E-28
GOTERM_MF_FAT	GO:0030528~transcription regulator activity	1.25E-21
GOTERM_BP_FAT	GO:0010628~positive regulation of gene expression	3.11E-26
GOTERM_BP_FAT	GO:0016481~negative regulation of transcription	5.58E-26
INTERPRO	IPR000504:RNA recognition motif, RNP-1	2.31E-10
GOTERM_MF_FAT	GO:0004725~protein tyrosine phosphatase activity	4.22E-14
GOTERM_CC_FAT	GO:0005925~focal adhesion	8.98E-08
GOTERM_BP_FAT	GO:0006917~induction of apoptosis	1.82E-06
Oil only down		
GOTERM_CC_FAT	GO:0016021~integral to membrane	0.004128
P4 only down		
INTERPRO	IPR001849:Pleckstrin homology	3.14E-06
GOTERM_BP_FAT	GO:0017148~negative regulation of translation	7.50E-06
GOTERM_MF_FAT	GO:0016887~ATPase activity	5.83E-10
GOTERM_MF_FAT	GO:0005099~Ras GTPase activator activity	1.12E-07
INTERPRO	IPR001128:cytochrome P450	1.13E-07

Groups were further separated into those down-regulated genes that were found to contain PR binding in vehicle-only or P4-only conditions.

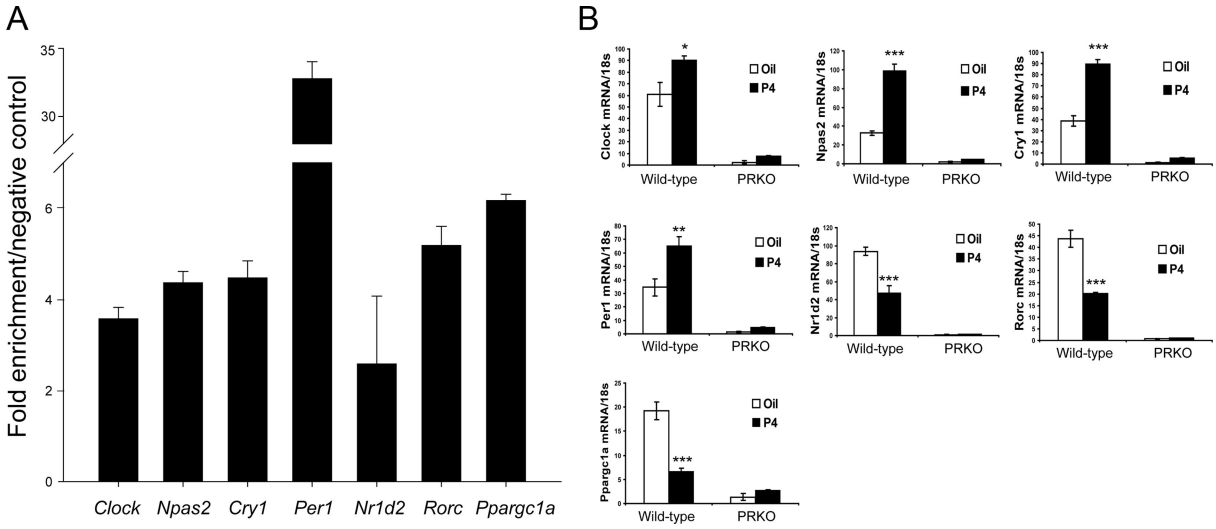

FIG. 3. P4 regulates the expression of several genes involved in circadian rhythm regulation in the uterus. A, ChIP-RT-qPCR validation of PR binding near genes involved in regulation of circadian rhythm (*Clock*, *Npas2*, *Cry1* and *Per1*, *Nr1d2*, *Rorc*, *Ppargc1a*) with the PR H-190 antibody (Santa Cruz Biotechnology) on uteri isolated from ovariectomized C57 female mice treated 1 h with P4. Data are represented as fold enrichment of PR binding of PR-location sites over that of the negative control region. B, Analysis of several regulators of circadian rhythm. Ovariectomized mice were treated with vehicle or P4 for 6 h, and uterine gene expression was analyzed by RT-qPCR. All circadian genes confirmed to contain PR binding demonstrate P4 regulation. Expression data for each gene were normalized to that of 18s rRNA. Data are based on three independent experiments. *Error bars* represent SEM. *, $P < 0.05$, **, $P < 0.01$; ***, $P < 0.001$.

within known P4 target genes, such as *Fkbp5*, PR, *Ihh*, *Areg*, and *ERα* (Fig. 4A). Analysis of other motifs associated around putative *Sox*-binding sites showed enrichment of HRE along with an abundance of potential transcription factors in common with the PR-interval dataset. GO analysis provided enrichment terms matching those seen in our P4-PR interval analysis, including that of regulation of transcription, ATP binding, and regulation of GTPase activity. Several GO terms identified were related to reproductive development, which prompted us to investigate the role, if any, of *Sox* factors in PR biology. ChIP-seq analysis predicted PR binding for two locations within 25 kb of the *Sox17* gene, and PR binding was confirmed by ChIP-RT-qPCR at these locations (Fig. 4B). Our microarray identified *Sox17* as a potential P4 up-regulated gene, which was confirmed in the mouse uterus through *in vivo* administration of P4 for 6 h by RT-qPCR analysis (Fig. 4C). Additionally, up-regulation of SOX17 protein was demonstrated in the epithelium by immunofluorescence on mouse uterine sections after P4 administration (Fig. 4D). To further demonstrate the ability of our ChIP-seq data to accurately predict PR binding locations that are sites of *bona fide* P4-PR-regulatory sequences, we placed the genomic *Sox17* PR binding location in front of a minimal promoter. We then cotransfected with both PR isoforms and treated with the progestin R5020 to determine the regulation, if any, of *Sox17* by the PR. Consistent with literature identifying the PRA isoform as the critical PR in the mouse uterus regulating P4 signaling (29), induction of gene expression was achieved only with the PRA isoform (Fig. 4E).

We then followed this analysis with a preliminary ChIP-on-chip on chromosomes 10, 13, and 14 for *Sox17* binding on 1 h P4-treated uteri. A total of 1095 binding locations were identified, provided in Supplemental Material 5 (SM5). When we compared these sites for PR binding on the same chromosomes, we found 2685 P4 and 1034 vehicle PR intervals, respectively, with 87 intervals directly overlapping P4 PR-binding sites and 61 overlapping vehicle PR-binding sites (Fig. 5A). Interval distribution analysis was performed using the CEAS module, as previously described, which highlighted the abundance of *Sox17* binding within the upstream 10-kb promoter region of murine genes (Fig. 5, B and C). After interval analysis, we then analyzed these regions for unique genes containing *Sox17*-binding sites located within ± 10 kb of gene boundaries. This identified 813 and 659 genes for *Sox17* and PR, respectively. A total of 314 genes containing *Sox17* binding also were found to contain PR binding (Fig. 5D). From this analysis, we confirmed *ERα* as a P4 target gene that contains overlapping *Sox17* and PR-binding sites (Fig. 5E), as predicted by our earlier motif analysis on P4-PR interval sequences and thus validating our Chip-seq as accurately predicting novel uterine transcription factors and binding locations. Considering the possibility that the 1-h time point for *Sox17* ChIP is too early for complete induction of this factor and enrichment of chromatin binding, and there-

TABLE 3. Listing of top significantly enriched sequence motifs identified using the SeqPos tool on the web-based application Cistrome for the dataset containing PR binding in the absence of ligand

Oil-PR intervals			
ID	**Factors**	**DNA-binding domain**	***p* value**
M00957	PGR, Pgr	Hormone-nuclear receptor family	1.00E-30
M01016	SOX17, Sox17	HMG	1.00E-30
MA0258	ESR2	Hormone-nuclear receptor	1.00E-30
M01801	ERalpha, Esr1	Cys4 zinc finger of nuclear receptor type	1.00E-30
M01270	PPARG	Cys4 zinc finger of nuclear receptor type	1.00E-30
M01282	PPARA, Ppara	Hormone-nuclear receptor family	1.00E-30
hPDI022	RXRA	zf-C4	1.00E-30
M01590	SMAD1, Smad1	MH1 Domain Family	1.00E-30
M00511	ESRRA	Cys4 zinc finger of nuclear receptor type	7.44E-22
M00939	E2F-1, E2F1	Fork head/winged helix	2.17E-19
MA0071	RORA_1	Hormone-nuclear receptor	4.22E-17
UP00009	Nr2f2	ZnF_C4	8.48E-15
UP00048	Rara	ZnF_C4	4.18E-14
M00926	AP-1, FOS, FOSB, JUN, JUNB, JUND	Leucine zipper factors (bZIP)	2.05E-12
hPDI019	KLF3	Zf-C2H2	1.55E-11
M01393	Msx-2, MSX2	Homeodomain	7.68E-11
MA0300	GAT1	GATA	1.92E-10
M00971	Ets	None	1.38E-09
M01588	KLF4, Klf4	Cys2His2 zinc finger domain	3.07E-07
M00348	GATA2, Gata2	GATA domain family	2.2E-06
UP00093	Klf7	zf-C2H2	2.4E-06
M01652	p53, TP53	p53	7.1E-06
M00223	STATx	STAT	1.3E-05
M00117	CEBPB, Cebpb	Leucine zipper factors (bZIP)	2.3E-05
M00733	SMAD4, Smad4	MH1 domain family	2.8E-05
M00473	FOXO1, Foxo1	Forkhead domain family	3E-05
MA0148	FOXA1	Forkhead	0.00011
M00932	Sp1, DAND5, SP1, PSG1	Cys2His2 zinc finger domain	0.00012

Chromosome location coordinates were uploaded into SeqPos and analyzed for motifs most enriched near the center of PR-binding locations.

fore representing a non-biologically-relevant time point, we followed up this preliminary ChIP-on-chip with ChIP RT-qPCR validation using uteri of mice given P4 for 6 h, a period known to give maximal P4 induction. We chose five genes identified from the ChIP-on-chip to contain *Sox17* binding sites, *Sfrs12*, *Gtf3c6*, *Ptch1,* and the two PR overlapping locations within *ERα* (Fig. 5F). In addition, based on our initial motif analysis, which was predictive for *Sox17* binding on the *Ihh* −19-kb PR interval (Fig. 4A), we hypothesized that in uteri treated with P4 for 6 h we would induce binding at this location. As expected, at 6 h of P4 treatment we validated *Sox17* binding on the *Ihh* −19-kb PR interval (Fig. 5F). Finally, we sought to determine the potential cooperative transactivating properties of *Sox17* with the PR. To do this, we placed the −19-kb genomic PR binding location identified for *Ihh*, a known up-regulated P4 target gene, in front of a luciferase containing a minimal promoter. Hec-1a cells were cotransfected with PRA, PRB, or *Sox17* alone or in combination and treated with R5020. Interestingly, we found that induction of the luciferase shows specificity with the PRB isoform, but not with either the PRA or *Sox17* expression vectors when each is transfected alone (Fig. 5G). However, when PRA is cotransfected with *Sox17*, there is a cooperative increase in induction that matches the level of induction seen with the PRB alone. Importantly, no additional induction is seen in the PRB-*Sox17* cotransfection, indicating that *Sox17* cooperatively induces expression of *Ihh* specifically through the PRA isoform.

Taken together with our ChIP-seq data, validation of PR binding on the *Sox17* gene, confirmation of *Sox17* as a *bona fide* P4 target gene through both *in vitro* and *in vivo* experiments, and ChIP-on-chip, we hypothesize that *Sox17* may act as cooperative transactivating factor in P4 target gene induction. Further investigation of *Sox17* will determine the effect, if any, of its role in female fertility and endometrial health.

Discussion

To fully understand the wide reaching effects of the PR, identification of its target genes, *cis*-regulatory elements, and genome-wide binding is necessary. The introduction of ChIP-seq technology coupled with microarray expression data has provided the ability to produce an unbiased

TABLE 4. Listing of top significantly enriched sequence motifs identified using the SeqPos tool on the web-based application Cistrome for the dataset containing PR binding in the presence of ligand

P4-PR intervals			
ID	**Factors**	**DNA-binding domain**	***P* value**
M00957	PGR, Pgr	Hormone-nuclear receptor family	1.00E-30
M01016	SOX17, Sox17	HMG	1.00E-30
MA0258	ESR2	Hormone-nuclear Receptor	1.00E-30
M01801	ERα, Esr1	Cys4 zinc finger of nuclear receptor type	1.00E-30
M01590	SMAD1, Smad1	MH1 domain family	1.00E-30
M00032	Ets1	None	1.00E-30
MA0074	RXRA, VDR	Hormone-nuclear Receptor	1.00E-30
M00926	AP-1, FOS, FOSB, JUN, JUNB, JUND	Leucine zipper factors (bZIP)	1.00E-30
M00007	Elk-1, ELK1	None	1.00E-30
M01595	STAT3	STAT	1.00E-30
hPDI019	KLF3	Zf-C2H2	1.00E-30
M01282	PPARA, Ppara	Hormone-nuclear Receptor Family	1.58E-22
M01145	c-Myc, MYC	Helix-loop-helix/leucine zipper factors (bHLH-ZIP)	9.47E-22
MA0066	PPARG	Hormone-nuclear receptor	7.83E-21
MA0039	Klf4	ββα-zinc finger	2.42E-20
M01652	p53, TP53	p53	4.33E-20
MA0300	GAT1	GATA	1.67E-17
M01393	Msx-2, MSX2	Homeodomain	3.65E-17
M01656	Trp63	None	3.40E-12
M00254	CCAAT box	None	3.88E-08
UP00093	Klf7	zf-C2H2	5.83E-07
M00468	KLF12, Klf12	ββα-zinc finger Family	1.6E-06
MA0157	FOXO3	Forkhead	2E-06
M00932	Sp1, DAND5, SP1, PSG1	Cys2His2 zinc finger domain	2.7E-06
MA0331	MCM1	MADS	0.00012
M00348	GATA2, Gata2	GATA domain family	0.00024
M00156	RORA	None	0.00026
MA0148	FOXA1	Forkhead	0.00026
MA0133	BRCA1	Other	0.0005
M00472	FOXO4, Foxo4	Forkhead domain family	0.00076
M01116	CLOCK::ARNTL, Clock::Arntl	None	0.0006

Chromosome location coordinates were uploaded into SeqPos and analyzed for motifs most enriched near the center of PR-binding locations.

method for elucidating the mechanism of PR action in the presence or absence of P4 in the mouse uterus *in vivo*. Several previous investigations have been published yielding analysis of the cistromes of multiple nuclear receptors (18, 21, 30, 31); however, none have thus far focused on PR binding in the mouse uterus. Our study has combined enriched functional annotation and motif analysis to further characterize PR transcriptional regulation. Interestingly, we have identified a subset of PR binding that occurs solely in the absence of P4 (2132 intervals). However, a clear majority of PR binding was found to occur after acute treatment with P4 (14,197 intervals) consistent with the ligand-dependent pathway of transcriptional regulation of P4 signaling. The binding of PR in the absence of ligand needs further investigation into whether a potential ligand-independent mechanism of regulation of a subset of genes by the PR in the uterus is physiologically relevant. The strength of this method is demonstrated with the confirmation of PR binding near or within known P4-regulated genes. Analysis of PR-interval sequences demonstrated 67% contain a minimum of a PRE half-site, with 23% possessing a full palindromic PRE. PR-binding sites that occur in the absence of either a PRE half-site or full-length PRE may require PR to function at transcriptional complexes independent of direct contact with chromatin, presumably through protein-protein interactions with other transcription factors binding to their own response elements. The identification of these PR-recruiting factors and the mechanism of protein-protein interaction will be of importance in future investigations. As has been noted in genome-wide mapping of other nuclear receptors, a large number of PR-binding sites occur in distal intergenic regions whereas those located within gene coordinates are found predominantly within introns (Fig. 1C). However, of those binding sites located in designated gene promoter regions, PR binding was highly distributed in close proximity to the transcriptional start sites. It will be of interest to determine what functional role, if any, those PR-binding locations found in intergenic sites play in P4 regulation of

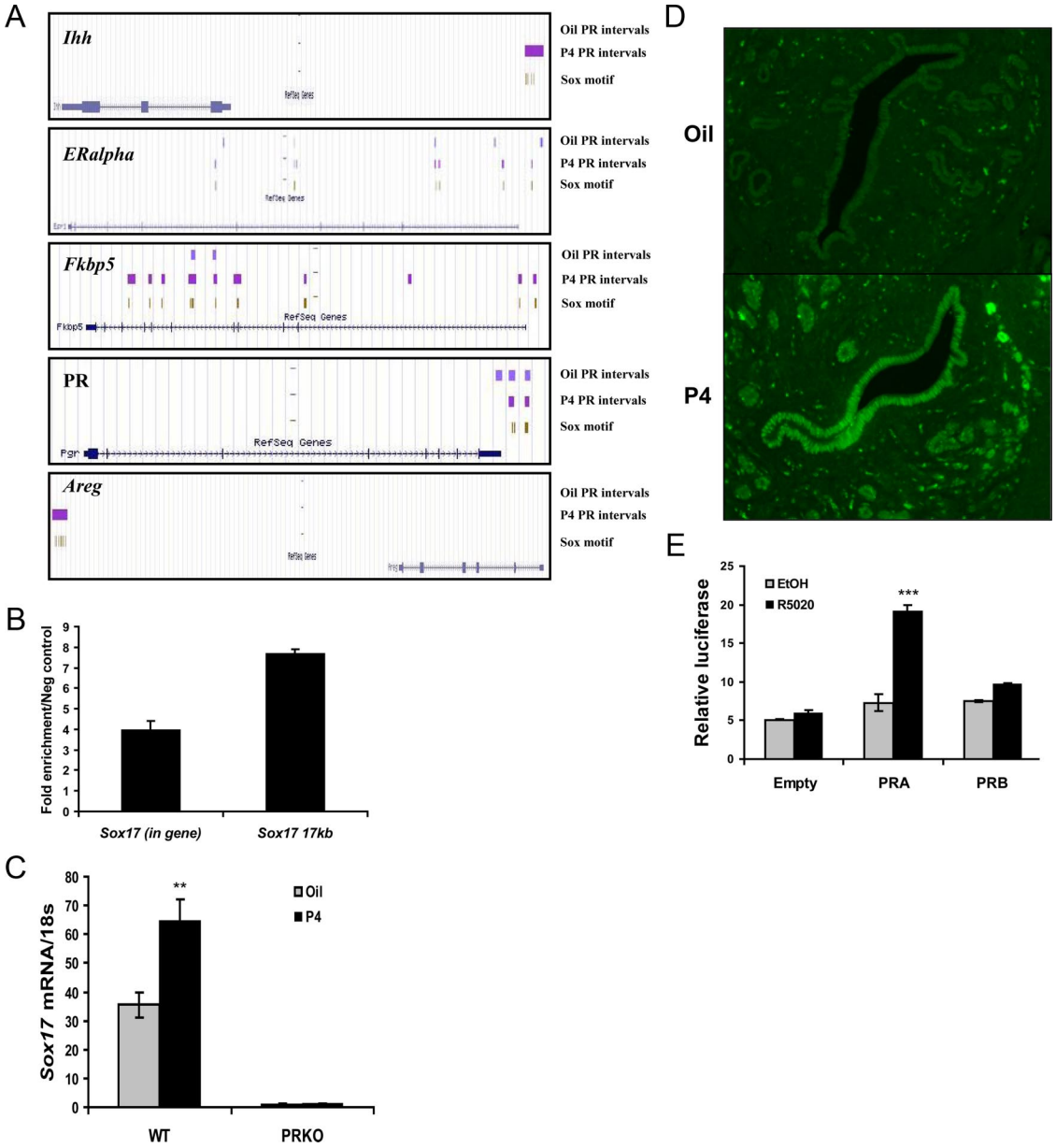

FIG. 4. Identification of a potential novel PR transcriptional interaction with *Sox17*. A, UCSC Genome Browser illustrations of select genes showing distributions of PR binding in vehicle and P4 treatments and potential Sox-binding locations based on motif analysis. Sequence positions and other generic UCSC annotations were removed for clarity. *Light blue* indicates PR binding in vehicle condition, *purple* indicates PR binding in P4 condition and *yellow* indicates location of putative Sox motifs. B, Validation of PR binding on the *Sox17* gene by ChIP-RT-qPCR. Two putative locations, within the *Sox17* gene and one approximately 17 kb upstream, were validated with the PR H-190 antibody (Santa Cruz Biotechnology) on uteri isolated from ovariectomized C57 female mice treated 1 h with P4. Data are represented as fold enrichment of PR binding of PR-location sites over that of the negative control region. C, Analysis of *Sox17* mRNA expression. Analysis performed on RNA isolated from uteri of ovariectomized mice treated with vehicle or P4 for 6 h. Expression data for each gene were normalized to that of 18s rRNA. Data based on three independent experiments. D, Immunofluorescence of SOX17 protein on uterine sections of ovariectomized mice treated with vehicle or P4 for 6 h with SOX17 Ab (AF1924; R&D Systems). E, *In vitro* analysis of PR-interval region within the *Sox17* gene cloned into the pGL4.24 minimal promoter construct cotransfected with PRA and PRB expression vectors (pcDNA3) treated with R5020 for 24 h in the Hec-1a cell line. Results obtained from triplicate experiments. *Error bars* represent SEM. ***, $P < 0.001$.

gene transcription, possibly through the mechanism of DNA looping. This model is consistent with the findings of other groups that distal enhancer sites are found to interact with transcriptional complexes located near gene boundaries through verified DNA-looping structures (22, 32) and reviewed here (33).

In our analysis of PR-interval locations, we noticed several PR-binding sites were located within or near the gene boundaries of many genes (*Clock*, *Npas2*, *Cry1*, *Per1*, *Ppargc1a*, *Rorc*, *Nr1d2*) associated with regulation of mammalian circadian rhythm (34, 35). Microarray expression confirmed by RT-qPCR analysis confirms these

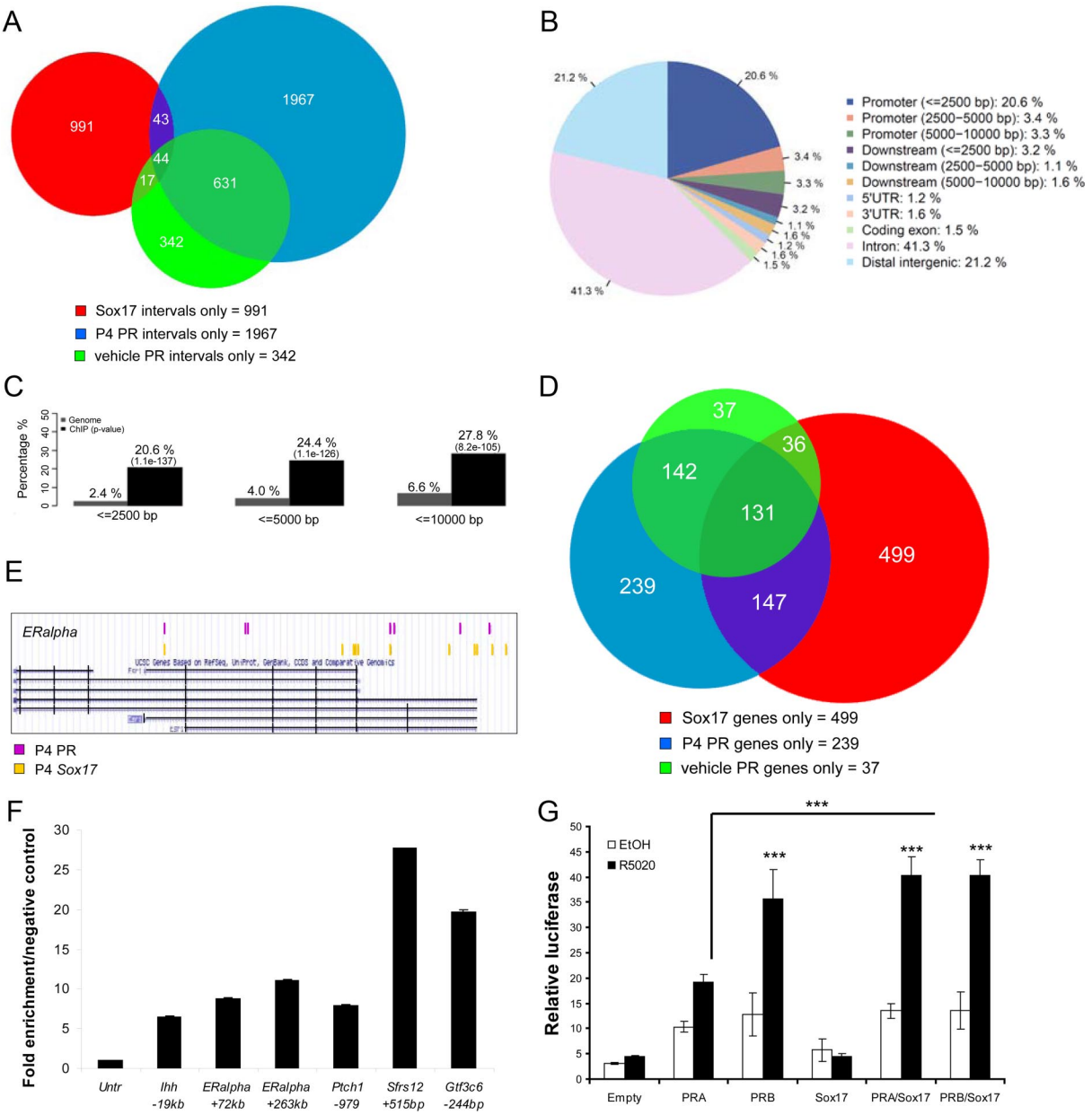

FIG. 5. Analysis of *in vivo Sox17* binding by ChIP-on-chip. A, ChIP-on-chip *Sox17*-binding site counts and proximity relationships. Venn diagrams of the intersection of multiple interval files from vehicle and P4-treated PR-binding sites, and P4-treated *Sox17* sites. All locations present on chromosomes 10, 13, and 14. B, Distribution of *Sox17*-binding locations on chromosomes 10, 13, and 14. Binding locations were analyzed using the CEAS module at Cistrome. Promoter region was defined in increments of 2500 bp up to 10 kb. When the PR-binding site is within a gene, it is further defined as within the 5′-untranslated region (UTR), 3′-UTR, coding exon, or intron. Intergenic region is defined as more than 10 kb from gene boundaries. C, *Sox17* binding is enriched along the 10-kb promoter region. The percentage of ChIP regions that reside in a 10-kb region upstream of gene boundaries was calculated and compared with the genome background percentages of the same region. Detailed analysis of calculation can be found in Ref. 21. D, Gene counts. Venn diagrams for nonredundant genes with PR (vehicle and P4) and *Sox17* (P4) binding locations within ± 10 kb. E, UCSC Genome Browser illustration of the *ERα* gene showing distributions of PR and *Sox17* binding in P4-treated uteri. Binding of the two factors demonstrate direct overlap in two interval locations. Sequence positions and other generic UCSC annotations were removed for clarity. *Purple* indicates PR binding and *yellow* indicates *Sox17* binding. F, Validation of *Sox17* binding sites by ChIP-RT-qPCR on uteri isolated from ovariectomized C57 female mice treated 6 h with P4. Negative control (Untr) is a gene-deficient region. Data are represented as fold enrichment of *Sox17*-binding sites over that of the negative control region. G, Cotransfection of 19-kb enhancer region of *Ihh* placed into the pGL4.24 minimal promoter construct with PRA, PRB expression vectors (pcDNA3) and a mouse *Sox17* expression vector (pcDNA6) treated with R5020 for 24 h in the Hec-1a cell line. Results are obtained from triplicate experiments. *Error bars* represent SEM. ***, $P < 0.001$.

genes are regulated by uterine P4. Circadian control in mammals is located within the hypothalamus in a compartment termed the "suprachiasmatic nucleus"; however, it has also been demonstrated that timely expression of circadian genes is present in other tissues, including the uterus, which may be regulated by self-sustained circa-

dian oscillations independent of regulation by the suprachiasmatic nucleus (36–38). Expression of these circadian genes has been shown to be rhythmically expressed in the uterus with evidence indicating that this expression may be regulated by the ovarian steroid hormones, consistent with our data (39–43). The effects on female fertility are strictly dependent on the specificity of circadian gene disruption (44–46). There is a need for more investigation into the physiological significance of the role of circadian genes in uterine regulation. It is intriguing to speculate that P4 regulation of the window of receptivity for embryo implantation may involve coordination of the circadian clock in the receptive uterus.

A much more provocative finding was that of the abundance of *Sox* motifs present throughout our PR interval location data. Interestingly, it was the presence of *Sox17*, a member of the *Sox* (sex-determining region on the Y chromosome-related high mobility group box) group F, which includes *Sox7*, *Sox17*, and *Sox18* (47, 48) that stood out. This factor has been characterized for its role as an early inducer of endodermal tissue, cardiac muscle development, fetal hematopoiesis, and spermatogenesis (49–53). This highlights *Sox17* as an important factor in the developmental regulation of many diverse tissues and, in the context of our findings, may indicate that *Sox17* expression in the uterus is of importance for the regulation of normal endometrial function and maintenance of pregnancy. We show that the location of *Sox* motifs in our data highly overlap areas of PR binding on known P4-regulated genes in the mouse uterus. Further evidence of the potential for a regulatory role of *Sox17* in mediating P4 signaling in the uterus was demonstrated by its confirmation as a P4 target gene that is required for the induction of the uteroglobin gene in rabbit endometrium (28). Additionally, *Sox17* expression has been confirmed in human endometrium (54), providing an important context as a potentially conserved role throughout mammalian species as a mediator of P4 signaling. We confirmed the status of *Sox17* as a P4 target gene through demonstration of up-regulation in both mRNA and protein after acute P4 treatment. From our motif analysis, which first brought to our attention the abundance of *Sox* motifs, we were able to subsequently identify those genes containing *Sox* motifs. An analysis of GO using this set of genes was analyzed, producing data that repeatedly identified terms involving reproductive processes. This data coupled with the confirmation of *Sox17* expression in the uterus provides strong evidence that *Sox17* may participate as a mediator of P4 signaling, perhaps acting in a transcriptional complex with the PR.

Pursuing this hypothesis, we performed ChIP-on-chip for *Sox17* on chromosomes 10, 13, and 14 and identified numerous P4-regulated genes that contain both PR and *Sox17* binding, including the P4-regulated gene, *ERα*. It has been noted that the 1-h time point may not be the optimal condition for *Sox17* binding, because our data show full induction of *Sox17* after 6 h of P4 treatment. It will be necessary to identify the most optimal conditions for further ChIP-seq analysis, which may demonstrate an increase in the occurrence of binding site overlap between the two factors. In addition to this idea, ChIP RT-qPCR validation on uteri treated for 6 h with P4 demonstrated strong induction of *Sox17* binding on six potential target genes, implicating this as the biologically relevant time point at which to perform future ChIP-seq analysis. Most notably, however, was our transfection analysis using a region of an *Ihh* enhancer site cotransfected with either PR isoform or *Sox17*. When each factor was present alone, only PRB demonstrated *Ihh* induction. Remarkably, when *Sox17* was present with the PRA isoform, induction of *Ihh* mirrored the level seen with PRB alone, whereas no further increase in expression was seen with *Sox17* and PRB. The relevance of this last finding has not escaped our notice. It has long been traditionally reported that the PRB isoform acts as a strong inducer of gene activation *in vitro*, with PRA acting as either a weak inducer of activation or a repressor of nuclear receptor activity (55). These data have been in conflict with *in vivo* data that demonstrably show the PRA isoform to the critical PR in murine uterine biology (29). Our finding provides evidence of the limitations of the *in vitro* data and insight into how the PRA interacts with chromatin of P4 target genes. Alone, the PRB acts as a strong activator of gene expression whereas PRA alone lacks the ability to fully induce gene expression. However, *in vivo*, the PRA may require a cooperative trans-activating partner to fully bind chromatin and elicit a transcriptional response that is equal to or greater than that seen with the PRB. It will be of interest to investigate the role of *Sox17* in uterine biology further, using genetically modified mice coupled with high-throughput technologies. Confirmation of its potential role as a PR cooperative factor may aid in identifying the underlying mechanistic origin of several endometrial diseases, because the presence of pioneering factors working in concert with other nuclear receptors have shown to be associated with prognosis of disease (56, 57).

Our study provides the first genome-wide identification of the PR cistrome in uterine tissue, providing an important dataset to be used for future investigation into the action of P4 signaling through its receptor. Additionally, these data will be useful in the identification of novel cooperating transactivating factors that participate in normal uterine regulation. Dysregulation of P4 signaling

in the endometrium contributes to a number of serious gynecological diseases that can affect both the physical and mental health of those afflicted. The identification of novel genes targeted by P4, in addition to those transcriptional regulators that aid uterine PR function, will better characterize new pathways regulated through P4 signaling and potentially lead to development of therapeutic targets to reverse endometrial disease.

Acknowledgments

We thank Dr. Paul Labhart of Active Motif for Performing the ChIP-seq.

Address all correspondence and requests for reprints to: Francesco J. DeMayo, Ph.D., Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030. E-mail: fdemayo@bcm.edu.

This work was supported by National Institutes of Health Grants R01HD042311 (to F.J.D.) and R01CA77530 (to J.P.L.); and *Eunice Kennedy Shriver National Institute* of Child Health and Human Development/National Institutes of Health Grants U54HD0077495 (to F.J.D.) and U54HD055787 (to F.J.D.).

Disclosure Summary: The authors have nothing to disclose.

References

1. **DeMayo FJ, Zhao B, Takamoto N, Tsai SY** 2002 Mechanisms of action of estrogen and progesterone. Ann NY Acad Sci 955:48–59; discussion 86–88, 396–406
2. **Jemal A, Siegel R, Xu J, Ward E** 2010 Cancer statistics, 2010. CA Cancer J Clin 60:277–300
3. **Rubel CA, Jeong JW, Tsai SY, Lydon JP, Demayo FJ** 2010 Epithelial-stromal interaction and progesterone receptors in the mouse uterus. Semin Reprod Med 28:27–35
4. **Conneely OM, Maxwell BL, Toft DO, Schrader WT, O'Malley BW** 1987 The A and B forms of the chicken progesterone receptor arise by alternate initiation of translation of a unique mRNA. Biochem Biophys Res Commun 149:493–501
5. **Tsai MJ, O'Malley BW** 1994 Molecular mechanisms of action of steroid/thyroid receptor superfamily members. Annu Rev Biochem 63:451–486
6. **Ribeiro RC, Kushner PJ, Baxter JD** 1995 The nuclear hormone receptor gene superfamily. Annu Rev Med 46:443–453
7. **Li X, O'Malley BW** 2003 Unfolding the action of progesterone receptors. J Biol Chem 278:39261–39264
8. **Umesono K, Evans RM** 1989 Determinants of target gene specificity for steroid/thyroid hormone receptors. Cell 57:1139–1146
9. **Owen GI, Richer JK, Tung L, Takimoto G, Horwitz KB** 1998 Progesterone regulates transcription of the p21(WAF1) cyclin-dependent kinase inhibitor gene through Sp1 and CBP/p300. J Biol Chem 273:10696–10701
10. **Gao J, Mazella J, Seppala M, Tseng L** 2001 Ligand activated hPR modulates the glycodelin promoter activity through the Sp1 sites in human endometrial adenocarcinoma cells. Mol Cell Endocrinol 176:97–102
11. **Gizard F, Robillard R, Gervois P, Faucompré A, Révillion F, Peyrat JP, Hum WD, Staels B** 2005 Progesterone inhibits human breast cancer cell growth through transcriptional upregulation of the cyclin-dependent kinase inhibitor p27Kip1 gene. FEBS Lett 579:5535–5541
12. **Cheon YP, Li Q, Xu X, DeMayo FJ, Bagchi IC, Bagchi MK** 2002 A genomic approach to identify novel progesterone receptor regulated pathways in the uterus during implantation. Mol Endocrinol 16:2853–2871
13. **Pan H, Zhu L, Deng Y, Pollard JW** 2006 Microarray analysis of uterine epithelial gene expression during the implantation window in the mouse. Endocrinology 147:4904–4916
14. **Jeong JW, Lee KY, Kwak I, White LD, Hilsenbeck SG, Lydon JP, DeMayo FJ** 2005 Identification of murine uterine genes regulated in a ligand-dependent manner by the progesterone receptor. Endocrinology 146:3490–3505
15. **Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS** 2008 Model-based analysis of ChIP-Seq (MACS). Genome Biol 9:R137
16. **Liu T, Ortiz JA, Taing L, Meyer CA, Lee B, Zhang Y, Shin H, Wong SS, Ma J, Lei Y, Pape UJ, Poidinger M, Chen Y, Yeung K, Brown M, Turpaz Y, Liu XS** 2011 Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol 12:R83
17. **Huang da W, Sherman BT, Lempicki RA** 2009 Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4:44–57
18. **Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, Brodsky AS, Keeton EK, Fertuck KC, Hall GF, Wang Q, Bekiranov S, Sementchenko V, Fox EA, Silver PA, Gingeras TR, Liu XS, Brown M** 2006 Genome-wide analysis of estrogen receptor binding sites. Nat Genet 38:1289–1297
19. **Lin CY, Vega VB, Thomsen JS, Zhang T, Kong SL, Xie M, Chiu KP, Lipovich L, Barnett DH, Stossi F, Yeo A, George J, Kuznetsov VA, Lee YK, Charn TH, Palanisamy N, Miller LD, Cheung E, Katzenellenbogen BS, Ruan Y, Bourque G, Wei CL, Liu ET** 2007 Whole-genome cartography of estrogen receptor α binding sites. PLoS Genet 3:e87
20. **Lanz RB, Bulynko Y, Malovannaya A, Labhart P, Wang L, Li W, Qin J, Harper M, O'Malley BW** 2010 Global characterization of transcriptional impact of the SRC-3 coregulator. Mol Endocrinol 24:859–872
21. **Wyce A, Bai Y, Nagpal S, Thompson CC** 2010 Research Resource: the androgen receptor modulates expression of genes with critical roles in muscle development and function. Mol Endocrinol 24:1665–1674
22. **Carroll JS, Liu XS, Brodsky AS, Li W, Meyer CA, Szary AJ, Eeckhoute J, Shao W, Hestermann EV, Geistlinger TR, Fox EA, Silver PA, Brown M** 2005 Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell 122:33–43
23. **Shin H, Liu T, Manrai AK, Liu XS** 2009 CEAS: cis-regulatory element annotation system. Bioinformatics 25:2605–2606
24. **Huang da W, Sherman BT, Lempicki RA** 2009 Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13
25. **Boden MJ, Kennaway DJ** 2006 Circadian rhythms and reproduction. Reproduction 132:379–392
26. **Lydon JP, DeMayo FJ, Funk CR, Mani SK, Hughes AR, Montgomery Jr CA, Shyamala G, Conneely OM, O'Malley BW** 1995 Mice lacking progesterone receptor exhibit pleiotropic reproductive abnormalities. Genes Dev 9:2266–2278
27. **Rubel CA, Franco HL, Jeong JW, Lydon JP, Demayo FJ** 2012 GATA2 is expressed at critical times in the mouse uterus during pregnancy. Gene Expr Patterns 12:196–203
28. **Garcia C, Calvo E, Nieto A** 2007 The transcription factor SOX17 is involved in the transcriptional control of the uteroglobin gene in rabbit endometrium. J Cell Biochem 102:665–679
29. **Mulac-Jericevic B, Mullinax RA, DeMayo FJ, Lydon JP, Conneely OM** 2000 Subgroup of reproductive functions of progesterone mediated by progesterone receptor-B isoform. Science 289:1751–1754

30. **Lin B, Wang J, Hong X, Yan X, Hwang D, Cho JH, Yi D, Utleg AG, Fang X, Schones DE, Zhao K, Omenn GS, Hood L** 2009 Integrated expression profiling and ChIP-seq analyses of the growth inhibition response program of the androgen receptor. PLoS One 4:e6589
31. **Welboren WJ, van Driel MA, Janssen-Megens EM, van Heeringen SJ, Sweep FC, Span PN, Stunnenberg HG** 2009 ChIP-Seq of ERα and RNA polymerase II defines genes differentially responding to ligands. EMBO J 28:1418–1428
32. **Wang Q, Li W, Liu XS, Carroll JS, Jänne OA, Keeton EK, Chinnaiyan AM, Pienta KJ, Brown M** 2007 A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. Mol Cell 27:380–392
33. **Bulger M, Groudine M** 2011 Functional and mechanistic diversity of distal transcription enhancers. Cell 144:327–339
34. **Maywood ES, O'Brien JA, Hastings MH** 2003 Expression of mCLOCK and other circadian clock-relevant proteins in the mouse suprachiasmatic nuclei. J Neuroendocrinol 15:329–334
35. **Reick M, Garcia JA, Dudley C, McKnight SL** 2001 NPAS2: an analog of clock operative in the mammalian forebrain. Science 293: 506–509
36. **Yamamoto T, Nakahata Y, Soma H, Akashi M, Mamine T, Takumi T** 2004 Transcriptional oscillation of canonical clock genes in mouse peripheral tissues. BMC Mol Biol 5:18
37. **Liu S, Cai Y, Sothern RB, Guan Y, Chan P** 2007 Chronobiological analysis of circadian patterns in transcription of seven key clock genes in six peripheral tissues in mice. Chronobiol Int 24:793–820
38. **Johnson MH, Lim A, Fernando D, Day ML** 2002 Circadian clockwork genes are expressed in the reproductive tract and conceptus of the early pregnant mouse. Reprod Biomed Online 4:140–145
39. **Nakamura TJ, Sellix MT, Menaker M, Block GD** 2008 Estrogen directly modulates circadian rhythms of PER2 expression in the uterus. Am J Physiol Endocrinol Metab 295:E1025–E1031
40. **Horard B, Rayet B, Triqueneaux G, Laudet V, Delaunay F, Vanacker JM** 2004 Expression of the orphan nuclear receptor ER-Ralpha is under circadian regulation in estrogen-responsive tissues. J Mol Endocrinol 33:87–97
41. **He PJ, Hirata M, Yamauchi N, Hattori MA** 2007 Up-regulation of Per1 expression by estradiol and progesterone in the rat uterus. J Endocrinol 194:511–519
42. **Nakamura TJ, Moriya T, Inoue S, Shimazoe T, Watanabe S, Ebihara S, Shinohara K** 2005 Estrogen differentially regulates expression of Per1 and Per2 genes between central and peripheral clocks and between reproductive and nonreproductive tissues in female rats. J Neurosci Res 82:622–630
43. **Nakamura TJ, Sellix MT, Kudo T, Nakao N, Yoshimura T, Ebihara S, Colwell CS, Block GD** 2010 Influence of the estrous cycle on clock gene expression in reproductive tissues: effects of fluctuating ovarian steroid hormone levels. Steroids 75:203–212
44. **Kennaway DJ, Boden MJ, Voultsios A** 2004 Reproductive performance in female Clock δ19 mutant mice. Reprod Fertil Dev 16: 801–810
45. **Alvarez JD, Hansen A, Ord T, Bebas P, Chappell PE, Giebultowicz JM, Williams C, Moss S, Sehgal A** 2008 The circadian clock protein BMAL1 is necessary for fertility and proper testosterone production in mice. J Biol Rhythms 23:26–36
46. **Pilorz V, Steinlechner S** 2008 Low reproductive success in Per1 and Per2 mutant mouse females due to accelerated ageing? Reproduction 135:559–568
47. **Bowles J, Schepers G, Koopman P** 2000 Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev Biol 227:239–255
48. **Wegner M** 1999 From head to toes: the multiple facets of Sox proteins. Nucleic Acids Res 27:1409–1420
49. **Hudson C, Clements D, Friday RV, Stott D, Woodland HR** 1997 Xsox17α and -β mediate endoderm formation in *Xenopus*. Cell 91:397–405
50. **Liu Y, Asakura M, Inoue H, Nakamura T, Sano M, Niu Z, Chen M, Schwartz RJ, Schneider MD** 2007 Sox17 is essential for the specification of cardiac mesoderm in embryonic stem cells. Proc Natl Acad Sci USA 104:3859–3864
51. **Sakamoto Y, Hara K, Kanai-Azuma M, Matsui T, Miura Y, Tsunekawa N, Kurohmaru M, Saijoh Y, Koopman P, Kanai Y** 2007 Redundant roles of Sox17 and Sox18 in early cardiovascular development of mouse embryos. Biochem Biophys Res Commun 360: 539–544
52. **Jang YY, Sharkis SJ** 2007 Fetal to adult stem cell transition: knocking Sox17 off. Cell 130:403–404
53. **Kanai Y, Kanai-Azuma M, Noce T, Saido TC, Shiroishi T, Hayashi Y, Yazaki K** 1996 Identification of two Sox17 messenger RNA isoforms, with and without the high mobility group box region, and their differential expression in mouse spermatogenesis. J Cell Biol 133:667–681
54. **Van Vaerenbergh I, Van Lommel L, Ghislain V, In't Veld P, Schuit F, Fatemi HM, Devroey P, Bourgain C** 2009 In GnRH antagonist/rec-FSH stimulated cycles, advanced endometrial maturation on the day of oocyte retrieval correlates with altered gene expression. Hum Reprod 24:1085–1091
55. **Giangrande PH, McDonnell DP** 1999 The A and B isoforms of the human progesterone receptor: two functionally different transcription factors encoded by a single gene. Recent Prog Horm Res 54: 291–313; discussion 313–314
56. **Nakshatri H, Badve S** 2009 FOXA1 in breast cancer. Expert Rev Mol Med 11:e8
57. **Chou J, Provot S, Werb Z** 2010 GATA3 in development and cancer differentiation: cells GATA have it! J Cell Physiol 222:42–49