The effects of breast cancer therapy on estrogen receptor signaling throughout the body

Droog, M.

Chapter 2

Estrogen Receptor α Wields Treatment-specific Enhancers between Morphologically Similar Endometrial Tumors

Marjolein Droog*, Ekaterina Nevedomskaya*, Gwen M.H.E. Dackus, Renske Fles, Yongsoo Kim, Harry Hollema, Marian Mourits, Petra Nederlof, Hester van Boven, Sabine C. Linn, Flora E. van Leeuwen, Lodewyk F.A. Wessels and Wilbert Zwart

* equal contribution

Proc Natl Acad Sci U S A 114 (2017) E1316-E1325

Abstract

The DNA binding sites of Estrogen Receptor alpha (ERα) show great plasticity under control of hormones and endocrine therapy. Tamoxifen is a widely applied therapy in breast cancer patients that affects ERα interactions with coregulators, and shifts the DNA binding signature of ERα in breast cancer upon prolonged exposure. Although tamoxifen inhibits breast cancer progression, it increases endometrial cancer risk in postmenopausal women. We therefore asked whether the DNA binding signature of ERα differs between endometrial tumors that arise in the presence or absence of tamoxifen, implicating divergent enhancer activity for tumors that develop in a different endocrine milieu.

Using ChIP-seq, we compared ERα profiles between 10 endometrial tumors from tamoxifen-users with 6 endometrial tumors from non-users, which we integrated with transcriptomic data of 47 endometrial tumors from tamoxifen-users and 64 endometrial tumors from non-users. Compared to non-users, tamoxifen-associated endometrial tumors revealed differential ERα binding sites, with distinct underlying DNA sequences, and divergent enhancer activity as marked by H3K27ac. Because tamoxifen acts as an agonist in the postmenopausal endometrium, similar to estrogen in breast, we compared ERα sites in tamoxifen-associated endometrial cancers to publicly available ERα ChIP-seq data in breast tumors, and found striking resemblance of ERα patterns between the two tissue types. Our study highlights the divergence between endometrial tumors that arise in different hormonal conditions, and shows for the first time that ERα enhancer usage in human cancer differs in the presence of nonphysiological endocrine stimuli.

Abbreviations

ChIP-seq, Chromatin Immunoprecipitation coupled with massive parallel sequencing; CPM, counts per million; CTCF, CTCCF-binding factor; ERα, Estrogen Receptor alpha; ESR1, the gene that encodes ERα; GSEA, Gene Set Enrichment Analysis; H/E, Hematoxylin and eosin; IPA, Ingenuity Pathway Analysis; POL2RA, RNA polymerase; RAD21, double strand break repair protein rad21 homolog; SRF, serum response factor; TAF1, transcription initiation factor TFIID subunit 1; TAMARISK, Tamoxifen Associated Malignancies: Aspects of Risk; TCGA, The Cancer Genome Atlas.

Significance

This study exhibits for the first time that the hormonal environment in which a tumor originates may affect enhancer usage of a hormone receptor. We further reveal that enhancer function is less tissue-specific than previously thought. By implementing ChIP-seq in a unique patient cohort, we compared estrogen receptor α profiles between endometrial tumors that developed in different hormonal environments, and integrated this with transcriptomic data. Our data show that tumors that associate with therapeutic intervention, have a distinct ERα DNA binding signature with different regulatory potential that resemble ERα binding patterns in breast cancer. These results highlight the value of cistromic analyses in clinical specimens, which enabled us to distinguish novel subtypes of tumors on the transcriptional regulation-level.

Introduction

Estrogen Receptor (ER)α is a steroid hormone receptor that behaves as a transcription factor by interacting with the DNA. The DNA binding profile (cistrome) of ERα is dependent on context and tissue-type[1]. The hormonal environment of the cell greatly influences this cistrome because estrogen activates ERα by binding its ligand binding domain. Upon activation, ERα's structural conformation changes to interact with cofactors at the DNA[2], and to regulate a transcriptional program that drives cell proliferation[3]. Hence, the hormonal environment modulates the ERα cistrome and thereby rewires downstream effects.

Endocrine therapies manipulate the DNA binding capacities of the steroid hormone receptor ERα as exemplified by tamoxifen. Tamoxifen, a small molecule inhibitor that competes with estrogens to bind ERα, is a major endocrine agent in ERα-positive breast cancer patients. Studies in the breast cancer cell line MCF-7 show that prolonged tamoxifen exposure shifts the ERα cistrome, which consequently changes gene expression[4-6].

Tamoxifen is well-known for its tissue-selective physiological action. Early reports on tamoxifen effects on transplanted MCF-7 cells in athymic mice revealed decreased tumor cell growth, but also an increased uterine weight of the mice in response to drug treatment[7]. Nonetheless, a species-selective action of tamoxifen could not be excluded at this stage. Growth-stimulatory effects of tamoxifen on human endometrial carcinomas were however later shown in a nude mouse model[8], and later reported in breast cancer patients, in whom tamoxifen treatment increased endometrial thickness as well as the risk of endometrial cancer by 2-7 fold in postmenopausal women, depending upon treatment duration[9-15].

Another study directly compared the contrasting actions of tamoxifen in athymic mice, transplanting endometrial EnCa101 tumors and MCF-7 derived tumors within the same mouse[16]. Although tamoxifen blocked tumor growth in the MCF-7 tumor while stimulating growth in the EnCa101 tumor, both tumors had qualitatively very similar patterns of tamoxifen metabolites, ruling out differential tamoxifen metabolism as a potential explanation for the observed tissue selective effects.

Cell line data illustrated that tamoxifen affected gene regulation differently in the endometrium compared to breast[17]. These data showed the agonistic effects of tamoxifen on ERα using the endometrial cancer cell line Ishikawa, but only for a handful of binding sites and related genes[17]. Our previously published data revealed thousands of ERα binding sites in multiple endometrial tumors from tamoxifen-treated patients, which showed remarkable overlap with the ERα cistrome in breast cancer[18], but lacked data on tumors of patients who never received endocrine treatment.

We hypothesize that the ERα cistrome differs between ERα-positive tumors that arise in the presence or absence of tamoxifen, and expect this will have consequences on the tumor's transcriptome. The TAMARISK study, which consists of endometrial tumors from patients who had a history of breast cancer, and half of whom received tamoxifen, provides an opportunity to investigate this hypothesis. We combined chromatin immunoprecipitation, coupled with massive parallel sequencing (ChIP-seq), and gene expression data in endometrial tumors from this cohort, and used bioinformatic analysis to investigate differences between endometrial tumors that originated in different hormonal environments (tamoxifen versus no tamoxifen).

We found that tamoxifen-associated endometrial tumors have a distinct ERα DNA binding signature that differs from endometrial tumors that develop in a hormonal environment without tamoxifen. The differentially enriched ERα sites were associated with gene expression differences, and the enriched ERα sites in tamoxifen-associated endometrial tumors resembled ERα-binding patterns in breast cancer. With this, our data suggest that the hormonal environment in which a tumor arises associates with differential enhancer usage of ERα.

Materials and Methods

Patient material

The design of the TAMARISK (Tamoxifen Associated Malignancies: Aspects of Risk)-study has previously been described as a nation-wide population-based prospective cohort of patients who developed uterine corpus cancer after breast cancer[13,19]. Here we present for the first time

the prospective part of this study, in which samples were obtained from patients who developed uterine corpus cancer between 2003 and 2006 after previously being diagnosed for breast cancer. Residual endometrial samples of anonymized patients, who signed a conformed consent, from the TAMARISK-study were used. This study was performed in accordance with the Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands (http://www.fmwv.nl), and has been approved by the local medical ethics committee of the Netherlands Cancer Institute. Endometrial samples were derived from patients who all had a history of breast cancer. Fresh frozen (frozen within 30 minutes after surgery and stored at -80°C) endometrial tumor specimens were collected and used for ChIP-seq and microarray. Clinicopathological parameters of these endometrial samples can be found in Table 1. Microsatellite instability detection was performed in the retrospective part of the cohort (Supplementary Table S1), described before[13,18].

Hematoxylin and eosin (H/E) and ERα stained sections were previously performed as described[13]. H/E staining of all tumors used (for both ChIP-seq and microarray) were reviewed by multiple gynecologic pathologists in terms of classification and grade (WHO classification 1994). ERα staining was done using mouse monoclonal antibodies (MCA1799, dilution 1:20; Serotec, Oxford, United Kingdom).

Chromatin Immunoprecipitations coupled with high-throughput sequencing (ChIP-seq)

ChIP-seq was performed as previously described[20,21] on sixteen endometrioid adenocarcinomas from the prospective TAMARISK cohort (Table 1). Thirty 30-µm cryosections of fresh-frozen endometrioid adenocarcinomas that contained at least 50% tumor tissue were fixed with 1% formaldehyde for 20 minutes and processed for sonication. For each ChIP 10µg of antibody was used, and 100µl of Protein A magnetic beads (Invitrogen). Antibodies raised to detect ERα (SC-543; Santa Cruz) and H3K27ac (ab4729; Abcam) were used.

High-throughput sequencing and processing

Single-end 51bp (ERα ChIP-seq) and 65bp (H3K27ac ChIP-seq) reads were generated using Illumina HiSeq 2000 Genome Analyzer, and aligned to hg19 human genome using bwa v0.5.9 with default parameters. Reads that were poorly aligned or mapped to multiple locations were filtered out based on the mapping quality: only reads with MAPQ > 20 were retained for further peak calling and analysis. The number of mapped and filtered reads is listed in Supplementary Table S2.

Table 1. Clinicopathological parameters of endometrioid adenocarcinomas.

	ChIP-seq		Gene expression	
	Tam-user	Non-user	Tam-user	Non-user
N	9	6	47	64
Tamoxifen use in years, median (range)	3.2 (1.4-5.6)	N/A	4.35 (2.0-15.0)	N/A
Interval time breast and endometrial cancer in years, median (range)	3.4 (2.0-4.7)	5.6 (0-17.1)	5.0 (2.1-26.5)	6.6 (-0.1-29.9)
Age breast cancer, median (range)	60.5 (51.0-83.2)	55.6 (44.2-67.3)	60.7 (35.9-83.2)	57.4 (35.2-77.3)
Age endometrial cancer, median (range)	64.2 (53.6-87.6)	59.4 (54.3-75.0)	69.5 (53.6-89.0)	68.0 (49.9-84.6)
Menopausal status, n (%)*				
Postmenopausal status	9 (100%)	6 (100%)	43 (91.5%)	60 (93.7%)
Perimenopausal status	0 (0.0%)	0 (0.0%)	0 (0.0%)	1 (1.6%)
Unknown	0 (0.0%)	0 (0.0%)	4 (8.5%)	3 (4.7%)
Recency tamoxifen use as to surgery, n (%)				
Ex-user	0 (0.0%)	0 (0.0%)	9 (19.1%)	0 (0.0%)
Recent user (last 12 months)	0 (0.0%)	0 (0.0%)	11 (23.4%)	0 (0.0%)
Current user	9 (100%)	0 (0.0%)	23 (48.9%)	0 (0.0%)
Non-user	0 (0.0%)	6 (100%)	0 (0.0%)	64 (100%)
Unknown	0 (0.0%)	0 (0.0%)	4 (8.5%)	0 (0.0%)
Histological type, n (%)				
Endometrioid adenocarcinoma and variants	9 (100%)	6 (100%)	47 (100.0%)	64 (100.0%)
Use of other hormonal therapy, n (%)	0 (0.0%)	0 (0.0%)	4 (8.5%)	6 (9.4%)
FIGO stage, n (%)				
I	7 (77.7%)	5 (83.3%)	34 (72.3%)	47 (73.4%)
II	1 (11.1%)	0 (0.0%)	5 (10.6%)	8 (12.5%)
III/IV	1 (11.1%)	1 (16.7%)	4 (8.5%)	6 (9.4%)
Unknown	0 (0.0%)	0 (0.0%)	4 (8.5%)	3 (4.7%)
Chemotherapy, n (%)				
Yes	4 (44.4%)	2 (33.3%)	10 (21.3%)	9 (14.1%)
No	5 (55.6%)	4 (66.7%)	34 (72.3%)	52 (81.3%)
Unknown	0 (0.0%)	0 (0.0%)	3 (6.4%)	3 (4.7%)

*During diagnosis of endometrial cancer.

There are no significant differences between tamoxifen-users and non-users with regard to interval time between breast and endometrial cancer, the age of breast cancer, or the age of endometrial cancer, according to t-test; no difference in FIGO stage or use of chemotherapy, according to Chi-square test.

Peak calling

Two algorithms were used for peak calling of ERα ChIP-seq data: MACS 1.4[22] and DFilter v1[23]. We employed MACS with default parameter settings, except for the p-value cutoff that we set at 10^{-7}. We used DFilter with parameter settings as recommended for transcription factor ChIP-seq peak calling (bs = 50, ks = 30, refine, nonzero). Only peaks called with both peak calling algorithms were considered for further analyses. The number of called peaks is listed in Supplementary Table S2. The sequences reported in this paper have been deposited in the National Center for Biotechnology Information Gene Expression Omnibus database (accession no. GSE94031).

DNA copy number calling

We employed the CopywriteR R package[24] to extract DNA copy number information from off-target (background) reads of ChIP-seq data. We used the package with the default parameters.

ChIP-seq data analysis

DiffBind R package[6] was used to identify genomic regions differentially bound by ERα in two groups of endometrial cancer. Peaks present in at least half of the patients in one of the groups were considered for the analysis. Differential read count analysis was performed without control read subtraction; significance threshold was set at FDR < 0.1. Heatmaps visualizing raw ChIP-seq signal in peaks were built using seqMINER 1.3.3[25]. Snapshots of ChIP-seq signal, as well as average signal profiles in peaks, were generated using the TransView R package[26].

Annotation of ChIP-seq peaks relative to the nearest gene was performed using the CEAS (cis-regulatory element annotation) tool[27] with default settings. Motifs, enriched at ERα binding sites, were identified using SeqPos tools (with default settings) available through Galaxy Cistrome[28]. Genes that have an ERα peak within the gene body or 20kb upstream of the transcription start site were identified as potential targets of the corresponding ERα binding sites. For functional enrichment of the genes we used QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN Redwood City, www.qiagen.com/ingenuity). Gene ontology analysis of the potential target genes was performed using PANTHER gene classification database[29].

Public ChIP-seq data processing and analysis

We previously published on endometrial tumors of tamoxifen-users that we included in this study with accession numbers GSM2144746, GSM2144758, and GSM2144760[18]. We also used publicly available ERα ChIP-seq data from primary breast cancer tissue from two cohorts of patients. The data was obtained from NCBI GEO, accession numbers GSE32222[6] and GSE40867[21]. Raw FASTQ files were aligned to the hg19 genome with bwa. TransView R package was used to generate average signal profiles and to calculate RPKM values in peaks. A list of publicly available cell line ChIP-seq data that was used is presented in Supplementary Table S3. For data provided by the Encode project[30], bed files were downloaded from https://www.encodeproject.org. Intersection of peak lists from two replicates was created, where only peaks shared by the two replicates were used where applicable. For ERα and FOXA1 ChIP-seq in the breast cancer cell line MCF-7, the data from Hurtado et al[4] was used. Raw FASTQ files were downloaded from https://www.ebi.ac.uk/arrayexpress/; alignment and peak calling was performed as described above. Intersection of peak lists from multiple replicates was used.

RNA isolation, cDNA synthesis, and RNA amplification and labeling for microarray

Microarray data was generated early after tissue collection as part of the prospective TAMARISK study. We included endometrial tumors of the endometrioid adenocarcinoma subtype from 47 patients who were on tamoxifen for at least two years[13], and from 64 patients who never used tamoxifen (Table 1). Thirty 30-µm cryosections of fresh-frozen endometrial tumors that were at least 50% tumorigenic were used for RNA isolation using Trizol (#15596-026, Invitrogen) according to manufacturer's instructions. RNA was purified using RNeasy Mini Kit (#74104, Qiagen), and treated with DNAse using RNase-Free DNase Set (#79254, Qiagen). Concentration and purity of the RNA was measured on a nanodrop spectrophotometer (Isogen Life-Science), whereas integrity of the RNA was determined by agarose gel. Next, cDNA was synthesized. First and second strand cDNA synthesis was performed using T7-(dT)24 primer and RT Superscript III (#18064-022, Invitrogen). The cDNA was purified using QIAquick PCR Purification Column (Qiagen). This was checked on a 1% agarose gel. Amplified RNA from the cDNA, was obtained using T7-mRNA amplification Invitrogen Superscript RNA Amplification system (#L1016-001, Invitrogen). Amplified RNA was labelled with Cy5 or Cy3 (#EA-006, Kreatech Biotechnology). The labeled amplified RNA was checked on a nanodrop spectrophotometer (Isogen Life-Science) and pooled with the same amount of reverse color Cy-labeled RNA from the reference. As a reference, a pool of RNAs was made that consisted mostly of endometrial tumors of patients that never

used tamoxifen and RNA of a few patients that had received tamoxifen, and which reflected the ratio of endometrial subtype as it appears within the population. The labelled amplified RNA was then fragmented using RNA fragmentation reagents (#8740, Ambion), and mixed with blocking solution containing Poly d(A) (#27-7836-01, Pharmacia), Cot-1 DNA (#15279-011, Invitrogen), and Yeast t-RNA (#109 495, Roche). Each tumor sample contains a replicate as the tumor samples were profiled once with Cy5 and once with Cy3. Labeled amplified RNAs were kept at 42°C until use and then mixed with 42°C preheated 2x F-hybridization buffer which contained formamide (#F 7503-1000, Sigma Aldrich) and 20X SCC (#19812323, BioSolve BV) at a 1:1 ratio, 0.1%SDS (#51232, BioWhittaker).

Gene expression profiling with microarrays

Spotted oligo microarrays with the Operon V3.0 library, human 35K oligo array (Operon Biotechnologies) were manufactured by the Netherlands Cancer Institute. A hybridization chamber (#10040, Ambion) was used. The microarray was prehybridized at 42°C for one hour using a buffer (5X SSC, 0.1% SDS and 1% BSA), and then washed with distilled water for ten minutes, again for five minutes. Hybridization occurred at 42°C overnight. Washes were performed at 42°C for the following solutions: 5x SSC 0.1%SDS for 30 sec, 2x SSC 0.1%SDS for 30 sec, 1x SSC for 5 sec. Two other washes were performed at room temperature with the solutions: 0.2X SSC for two min, and 0.05X SSC for 20 sec. The hybridized array was scanned on a DNA Microarray Scanner (Model G2505B, Serial number US22502518, Agilent Technologies). The fluorescence intensities were measured using ImaGene software (Biodiscovery).

Gene expression analysis

After background correction, the intensities from the Cy5 and the Cy3 channel were used to calculate log2-transformed ratios. These ratios were then normalized using the LOWESS subarray method[31]. The normalized data were further analyzed in R. To create one dataset, experiments done in dye swap were combined to generate gene expression log-ratios for patients who did not receive tamoxifen or received the drug for over 2 years were used in the analysis (Table 1). Only the probes with gene symbol assigned and statistically significant log-ratios ($p < 0.05$) in at least 40 patients were retained (N = 3734). Differential gene expression between endometrial tumors from tamoxifen-users and non-users was assessed using limma R package[32]. Fold changes from limma analysis were used to rank genes for Gene Set Enrichment Analysis[33] (http://software.broadinstitute.org/gsea/index.jsp). For pathway analyses we used

curated, hallmarks and oncogenic signatures gene sets collections from mSigDB v5.0[33].

Pathway enrichment network was generated using the Enrichment Map[34] app from Cytoscape[35]. To generate gene sets of up- and down-regulated genes in the breast cancer cell line MCF-7, we used publicly available gene expression data from Zwart et al[20]. Gene expression data was processed with beadarray R package[36]. After quantile normalization, differential gene expression between vehicle and estradiol conditions was determined using limma workflow: After fitting gene-wise linear model empirical Bayes statistics was estimated. P-values were adjusted for multiple testing using Benjamini–Hochberg procedure. Genes with adjusted p-value below 0.05 and absolute log-fold change above 1 were considered differentially expressed upon estradiol stimulation. For the endometrial cancer cell line Ishikawa, gene expression data of vehicle and estradiol–stimulated cells was downloaded from Gertz et al[37]. RPKM values were processed using limma package to identify genes differentially expressed between estradiol and vehicle stimulated cells as described above. Genes with adjusted p-value below 0.05 and absolute log-fold change above 1 were used to construct up- and down-regulated gene sets.

TCGA data

The Cancer Genome Atlas (TCGA) pan-cancer processed and normalized gene expression[38] was downloaded from https://tcga-data.nci.nih.gov. We used limma R package to generate fold changes in gene expression between endometrial (endometrioid adenocarcinoma subtype) and breast (ERα-positive subtype) cancers. These fold changes were used to rank genes for Gene Set Enrichment Analysis (GSEA).

Results

ERα binds DNA differentially in endometrial tumors of tamoxifen-users compared to non-users

Tamoxifen is a ligand that binds ERα, and increases the risk for endometrial cancer in postmenopausal women. Tamoxifen-associated endometrioid adenocarcinomas are morphologically indistinguishable from endometrioid adenocarcinomas that arise in a tamoxifen-free environment (Figure 1A). Likewise, they cannot be distinguished based on DNA copy number profile[19]. Because tamoxifen targets ERα, we tested if ERα binding to the DNA differed between endometrial tumors from tamoxifen-users versus non-users.

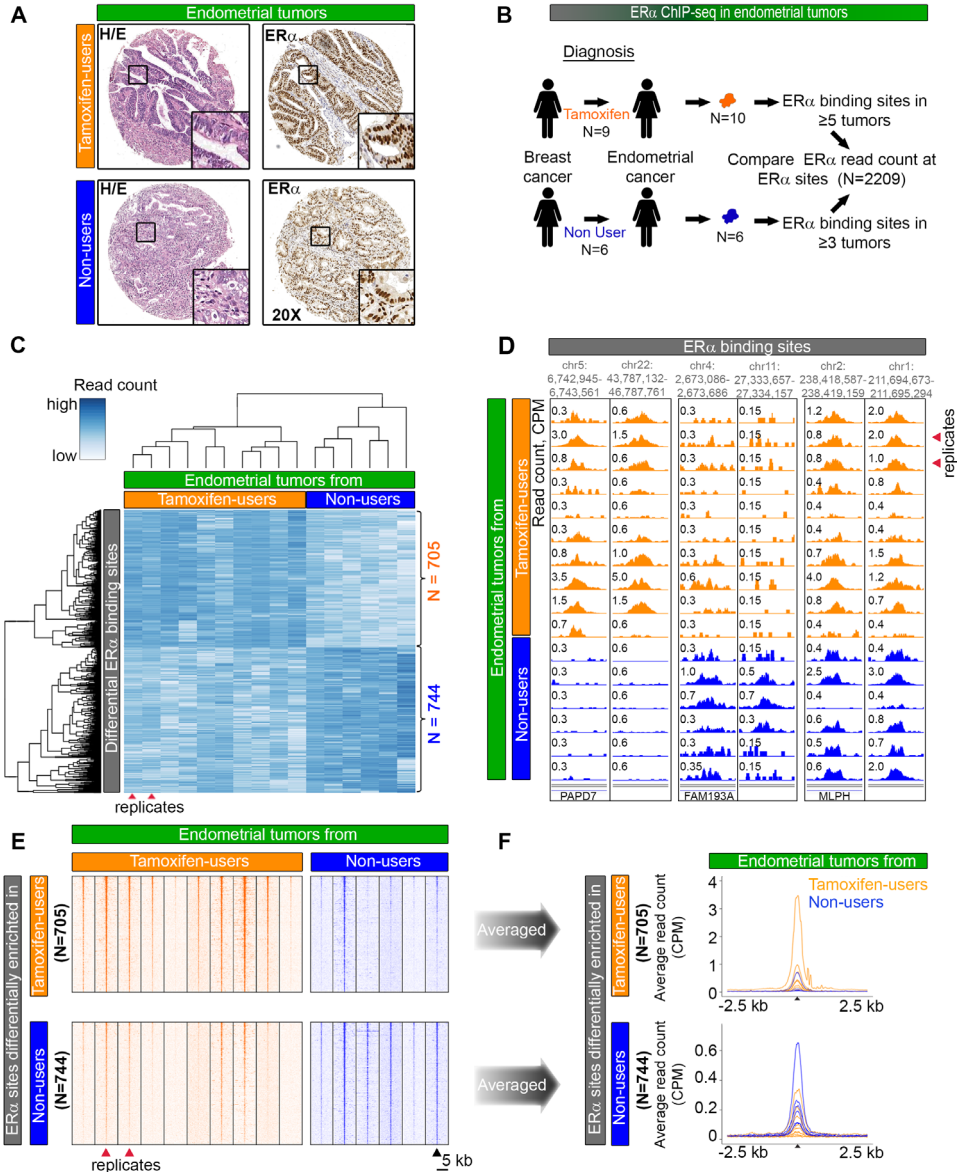
To investigate the ERα cistrome in endometrial tumors from tamoxifen-users and non-users, we used sixteen fresh frozen clinical specimens from the prospective TAMARISK cohort. Patients from the TAMARISK series had breast cancer (half of whom received tamoxifen), and subsequently developed endometrial cancer. To compare the ERα cistrome between endometrial tumors from tamoxifen-users and non-users, we performed ERα ChIP-seq on sixteen endometrial tumors. We previously published on endometrial tumors of tamoxifen-users (and compared biological replicates of ERα ChIP-seq data) that we also included in this study[18]. All tested tumors were of the most common subtype endometrioid adenocarcinoma, as determined by our pathologists (Table 1). Between endometrial tumors from tamoxifen-users and non-users, no differences in clinicopathological parameters including prior chemotherapy usage (Table 1) or microsatellite instability (Supplementary Table S1) were evident.

Of the sixteen endometrial tumors that we used for ChIP-seq, six endometrial cancer samples arose in patients who never used endocrine treatment for their breast cancer ('non-users'). The remaining ten tumors came from nine tamoxifen-users, who up till the day of surgery used tamoxifen (Figure 1B). We included two specimens of the same patient to provide a replicate experiment.

To identify ERα chromatin binding sites that differed between the two tumor groups, we performed differential binding analysis[6]. We included sites that were present in at least half of the tumors per group, and performed the analysis on the union of those sites (N=2209, Figure 1B). In total, we identified 1449 binding sites as significantly different (FDR < 0.1) between the two groups ($p < 0.00013$ based on 8008 available group labels permutations). The ERα read count is higher at 705 sites in tamoxifen-associated endometrial tumors and it is lower at 744 sites compared to endometrial tumors from patients that never used tamoxifen (Figure 1C, Supplementary Tables S4 and Supplementary Table S5). Importantly, two specimens that belonged to the same patient, clustered together. Snapshots of the ERα signal exemplify both differential and non-differential ERα sites (Figure 1D). Analysis of the ERα ChIP-seq data shows the raw (Figure 1E) and average (Figure 1F and Supplementary Figure S1) read count in the ERα peaks that are differentially enriched between the two groups.

Differential ERα binding sites between tumors from tamoxifen-users and non-users have distinct underlying DNA sequences and potential activity

ERα binds the DNA in tamoxifen-associated endometrial tumors differently compared to endometrial tumors from patients who never received


A
Endometrial tumors
Tamoxifen-users
Non-users
H/E
ERα
20X
B
ERα ChIP-seq in endometrial tumors
Diagnosis
Tamoxifen
N=9
N=10
ERα binding sites in ≥5 tumors
Breast cancer
Endometrial cancer
Compare ERα read count at ERα sites (N=2209)
Non User
N=6
N=6
ERα binding sites in ≥3 tumors
C
Read count
high
low
Endometrial tumors from
Tamoxifen-users
Non-users
Differential ERα binding sites
N = 705
N = 744
replicates
D
ERα binding sites
chr5: 6,742,945-6,743,561
chr22: 43,787,132-46,787,761
chr4: 2,673,086-2,673,686
chr11: 27,333,657-27,334,157
chr2: 238,418,587-238,419,159
chr1: 211,694,673-211,695,294
Endometrial tumors from
Tamoxifen-users
Non-users
Read count, CPM
replicates
PAPD7
FAM193A
MLPH
E
Endometrial tumors from
Tamoxifen-users
Non-users
ERα sites differentially enriched in
Tamoxifen-users (N=705)
Non-users (N=744)
replicates
5 kb
Averaged
F
Endometrial tumors from
Tamoxifen-users
Non-users
ERα sites differentially enriched in
Tamoxifen-users (N=705)
Non-users (N=744)
Average read count (CPM)
-2.5 kb
2.5 kb

tamoxifen treatment (Figure 1). To investigate the genomic features of these differential binding sites, we characterized their DNA sequences, their genomic distribution and their regulatory activity (Figure 2).

Sequence motif analysis revealed that the enriched ERα binding sites of tamoxifen-associated endometrial tumors contain different DNA motifs than the ERα binding sites enriched in non-users (Figure 2A). Tamoxifen-associated endometrial tumors exhibited enriched ERα binding sites that contained mostly motifs of estrogen receptor (ESR1), as well as other well-known hormone receptors, such as the Androgen Receptor, Glucocorticoid Receptor and Thyroid Hormone Receptor. In contrast, ERα binding sites enriched in endometrial tumors from non-users included mostly motifs of the forkhead domain family and the high mobility group Box family. This last group contains well-known stem cell markers, such as SOX4 and Nanog, which associate with endometrial cancer[39,40]. Both groups contained motifs for leucine zipper proteins at the differential ERα binding sites.

Differential ERα sites, enriched in endometrial tumors from either tamoxifen-users or non-users, locate mainly at distal intergenic regions and gene introns (Figure 2B). This corresponds to previously described data on distribution of ERα peaks, which is characteristic for enhancer-binding transcription factors[18,41]. We also defined a set of ERα binding sites that are not differential between the two tumor groups (absolute logFC

Figure 1. Comparative analysis of ERα binding in endometrial tumors from tamoxifen-users and non-users.

A) H/E staining and ERα immunohistochemistry staining in endometrial tumors from tamoxifen-users and non-users. Magnification, x20.

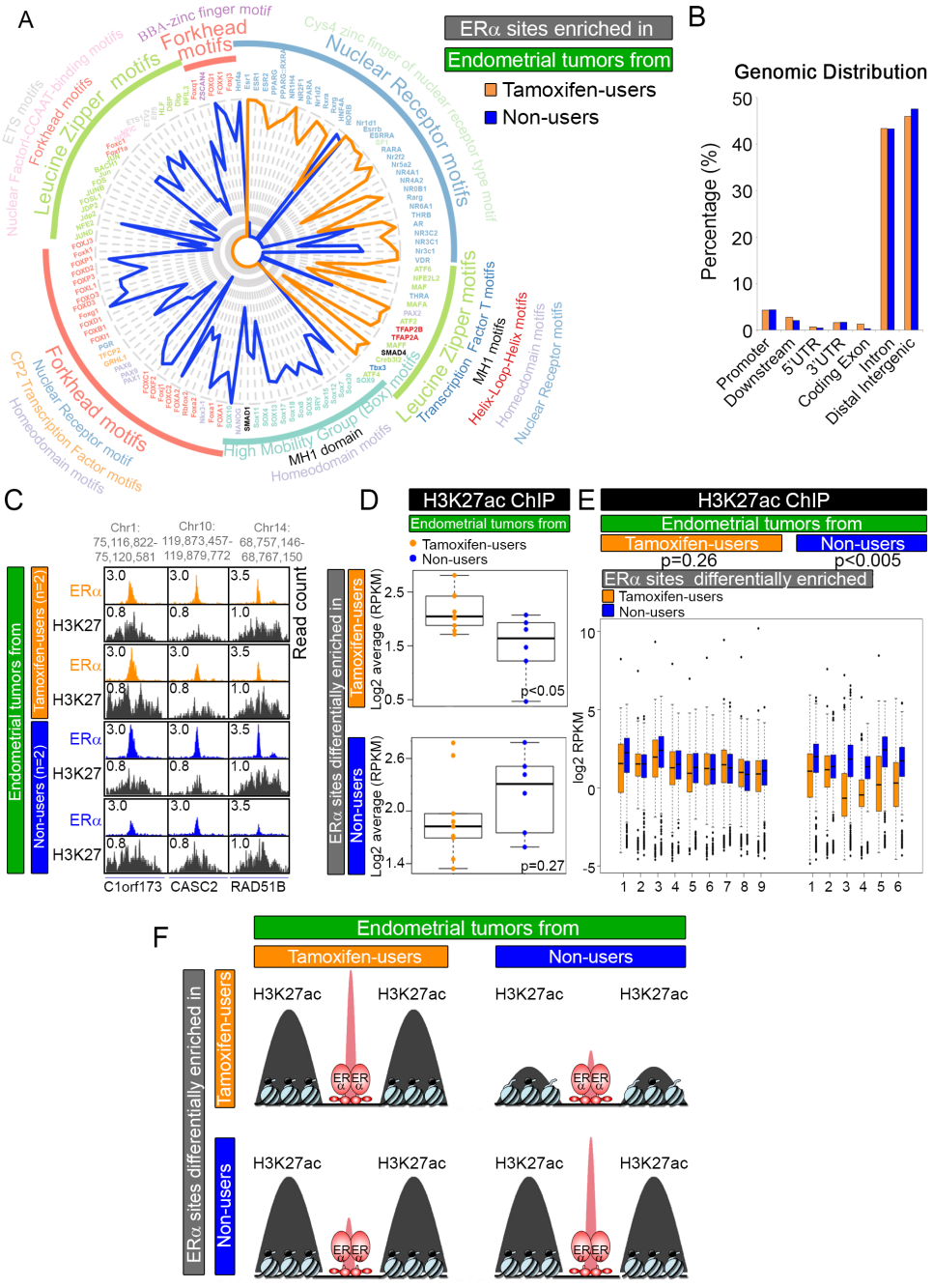
B) Experimental set-up of ChIP-seq analyses in endometrial cancers. The analysis compares ERα binding between ten endometrial tumors from nine tamoxifen-users (orange) and ERα binding in six patients (blue) who never received endocrine therapy for breast cancer treatment. Characteristics are described in Table 1.

C) Hierarchical clustering based on the results of differential binding analysis. Upper part shows 705 ERα binding sites that have a higher read count in tamoxifen-associated endometrial tumors (orange). The lower part shows 744 ERα binding sites that have a higher read count in endometrial tumors from patients that never used tamoxifen (blue). Red arrowheads indicate two tumors that originated from one patient.

D) Snapshots depicting ERα binding sites in sixteen endometrial tumors, at indicated genomic locations. Read counts were normalized (CPM).

E) Heatmap visualizing raw read count intensity of ERα at differential binding sites in tamoxifen-associated endometrial tumors (orange) and endometrial tumors from patients that never used tamoxifen (blue). Upper panel and bottom panel shows differential ERα binding sites as described in Figure C. ChIP-seq signal aligns on the center of the peaks with a window of 5kb.

F) Averaged read counts for ERα ChIP-seq data in tumors from tamoxifen-users (orange) and non-users (blue) at differential ERα binding sites. Data aligns on the center of ERα peaks with a 2.5kb window.


A
ERα sites enriched in
Endometrial tumors from
Tamoxifen-users
Non-users
Forkhead motifs
Leucine Zipper motifs
Nuclear Receptor motifs
High Mobility Group (box) motifs
MH1 domain
Homeodomain motifs
Transcription Factor T motifs
MH1 motifs
Helix-Loop-Helix motifs
B
Genomic Distribution
Percentage (%)
Promoter
Downstream
5'UTR
3'UTR
Coding Exon
Intron
Distal Intergenic
C
Endometrial tumors from
Tamoxifen-users (n=2)
Non-users (n=2)
ERα
H3K27
Read count
C1orf173
CASC2
RAD51B
D
H3K27ac ChIP
Endometrial tumors from
Tamoxifen-users
Non-users
ERα sites differentially enriched in
Log2 average (RPKM)
p<0.05
p=0.27
E
H3K27ac ChIP
Endometrial tumors from
Tamoxifen-users
Non-users
p=0.26
p<0.005
ERα sites differentially enriched
log2 RPKM
F
Endometrial tumors from
Tamoxifen-users
Non-users
ERα sites differentially enriched in
H3K27ac
ERER

below 0.5, N = 423, Supplementary Table S6). These sites show enhancer-like genomic distribution (Supplementary Figure S2) and harbor motifs of estrogen receptor (Supplementary Figure S3).

Chromosomal distribution of the differential ERα binding sites varied between endometrial tumor of tamoxifen-users and non-users (Supplementary Figure S4). We tested if this difference could be caused by distinct chromosomal aberrations present in the two tumor groups. Consistent with previous reports[19], we could not separate tamoxifen-associated endometrial tumors from endometrial tumors that arose in a tamoxifen-free environment, based on their copy number profiles (Supplementary Figure S5). Based on this observation, we conclude that the apparent bias of ERα binding to specific chromosomes is not due to differences in chromosomal copy number.

We next investigated if differential ERα binding sites harbored the H3K27ac histone mark, which would indicate active enhancers. Visual inspection of H3K27ac ChIP-seq data revealed a strong signal at ERα binding sites (Figure 2C, Supplementary Figure S6). We observed that higher binding of ERα in tamoxifen users was accompanied by a more prominent H3K27ac signal at those regions compared to non-users (Figure 2D upper panel). At the sites where ERα binding was enriched in non-users, H3K27ac was equally present in both tamoxifen-users and non-users (Figure 2D lower panel). Similarly, H3K27ac signal was comparable at both tamoxifen-users- and non-users-enriched ERα sites (Figure 2E left). In contrast, in tumors of non-users there is a difference in enhancer activity at the differential sites, with non-users-enriched ERα sites exposing higher H2K27ac (Figure 2E right).

Figure 2. Characterization of ERα sites differentially bound between endometrial tumors from tamoxifen-users and non-users.

A) Radar plot, visualizing DNA motif enrichment at genomic ERα sites differentially enriched in either endometrial tumors from tamoxifen-users (orange) or non-users (blue). Lengths of radii correspond to the fraction of peaks that contain the identified motif. Motif colors correspond to transcription factor families.

B) Genomic distribution of ERα sites that are differentially enriched in endometrial tumors from tamoxifen-users (orange) and non-users (blue), relative to the nearest gene.

C) Snapshots depicting H3K27ac ChIP-seq signal at ERα binding sites in endometrial tumors from tamoxifen-users (orange) and non-users (blue), at indicated genomic locations. Read counts were normalized (CPM).

D) Boxplots visualizing average normalized H3K27ac read count in endometrial tumors from tamoxifen-users (orange dots) and non-users (blue dots) at differential ERα binding sites.

E) Boxplots showing normalized H3K27ac read counts at ERα differential binding sites in endometrial tumors from tamoxifen-users (orange) and non-users (blue). P-values of the paired t-test for each tumor group are shown.

F) Model for the intensity of H3K27ac mark (black) at differential ERα-binding sites (red) in the two tumor groups.

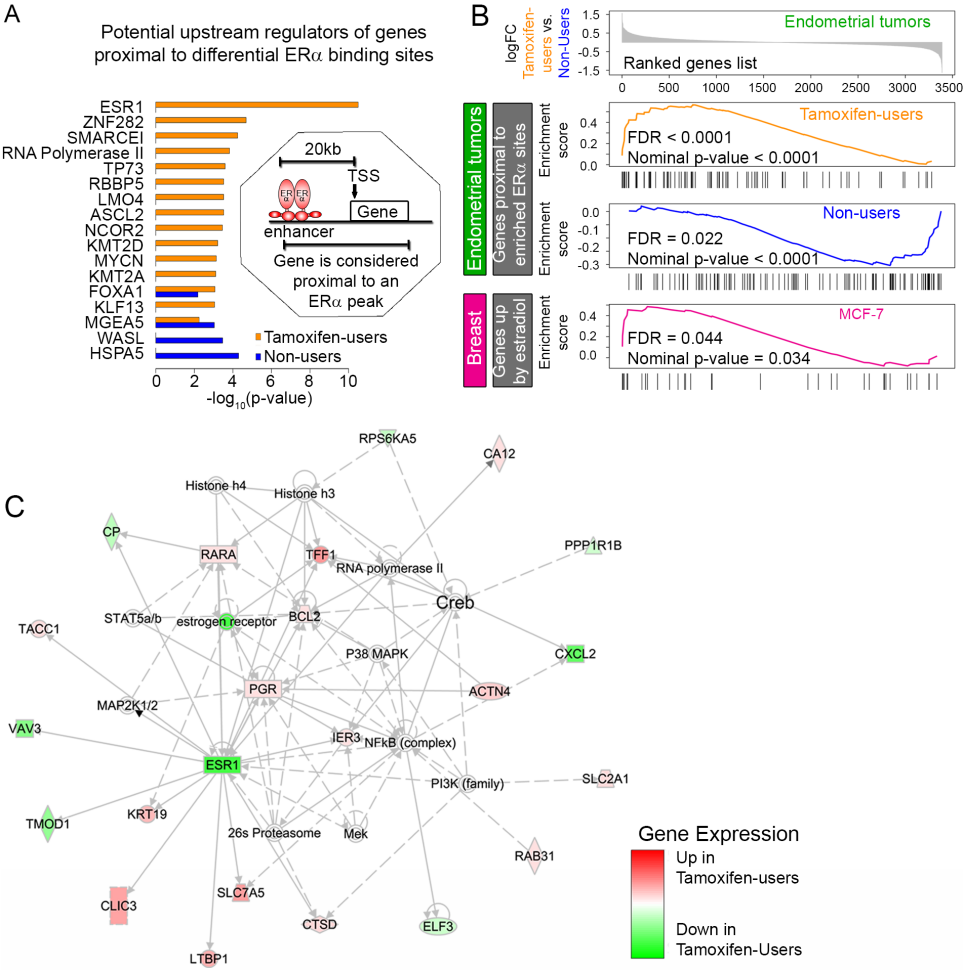


Figure 3. ERα-mediated gene regulation in endometrial tumors of tamoxifen-users versus non-users.

A) Barplot visualizes potential upstream regulators of genes proximal to differential ERα binding sites according to Ingenuity Pathway Analysis. The inlet shows how potential target genes are defined. TSS stands for transcriptional start site.

B) Gene Set Enrichment Analysis based on differential gene expression between endometrial tumors of tamoxifen-users and non-users from the TAMARISK cohort. Top panel shows ranked log-fold change of gene expression between the two cancer groups. Lower panels show enrichment scores versus gene rank in three significantly enriched gene sets: genes proximal to the binding sites enriched in tamoxifen-associated tumors, genes proximal to the binding sites enriched in tumors of non-users and genes upregulated by estradiol in MCF-7 breast cancer cell line. Patients characteristics are described in Table 1.

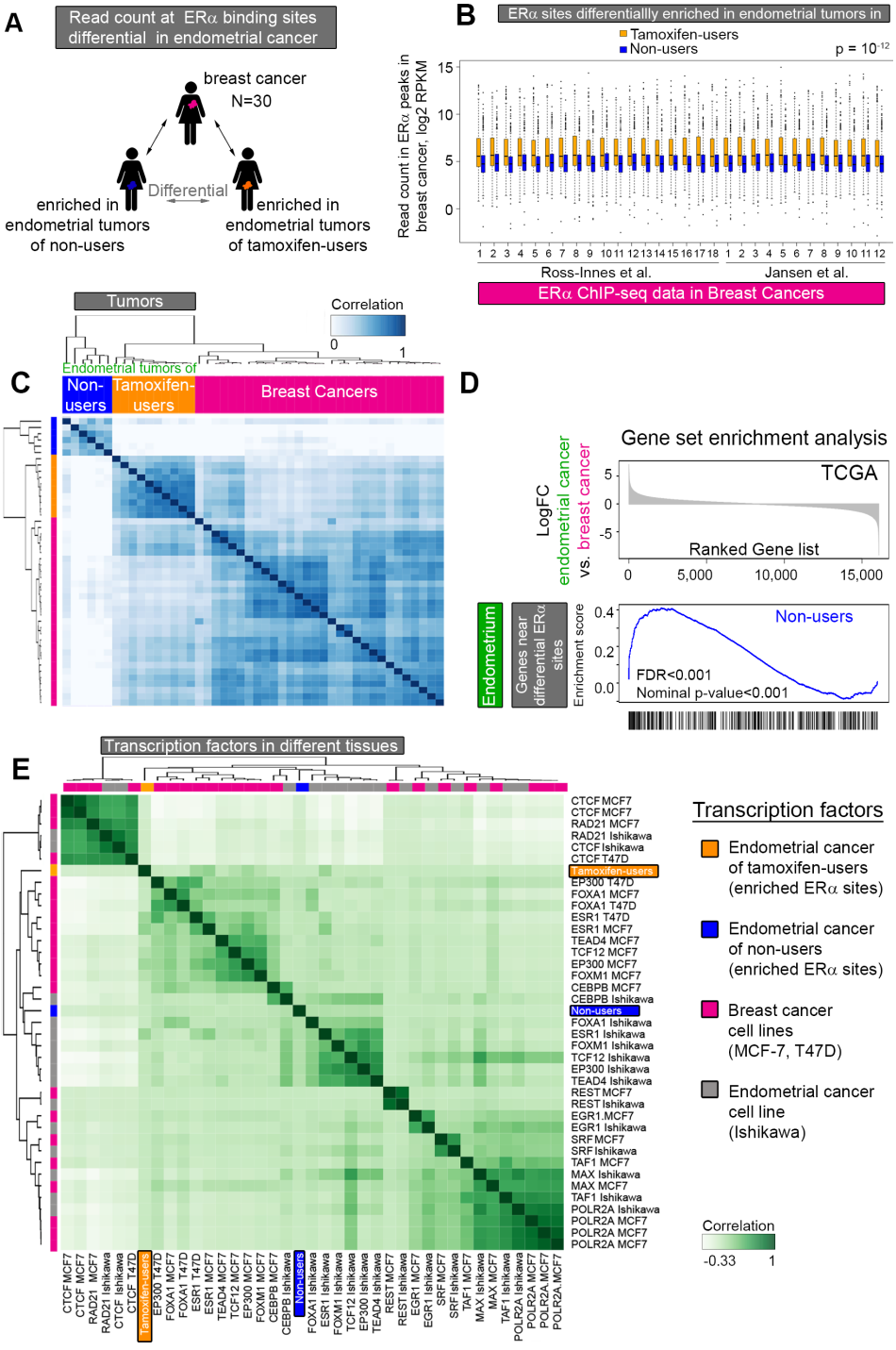
C) Top network from Ingenuity Pathway Analysis based on genes identified as potential targets of ERα by combined analysis of gene expression and ChIP-seq data.

Taken together, these data reveal that the differential ERα binding sites between endometrial tumors from tamoxifen-users and non-users are enriched for different DNA motifs, distributed mostly to active enhancers. Activity of ERα sites enriched in non-users, based on H3K27Ac, does not differ between the two groups of endometrial tumors. In contrast, H3K27Ac levels at tamoxifen-associated ERα sites are higher in the group of tamoxifen-users (Figure 2F).

Differential ERα binding sites between tamoxifen-users and non-users affect gene regulation differently

Differences in ERα profiles and H3K27ac signals in endometrial tumors between tamoxifen-users and non-users suggest deviations in corresponding gene expression (Figure 2). To link tamoxifen treatment with gene activity, we generated microarray data from endometrial tumors (endometrioid adenocarcinoma subtype, largely ER-positive[42]) of 47 tamoxifen-users and 64 non-users of the TAMARISK series (Table 1). Pathway analysis revealed a number of biological processes associated with the genes differentially expressed between the two tumor groups (Supplementary Tables S7, S8). Network representation of these pathways shows that, besides ERα targets, genes related to extracellular matrix and mesenchymal transition, as well as genes downregulated by RB1 and TP53 and interferon targets are upregulated in the group of tamoxifen-users (Supplementary Figure S7).

To understand possible functions in gene expression of the identified differential ERα binding sites, we first characterized their potential target genes (388 genes for tamoxifen-associated endometrial tumors, and 402 genes for endometrial tumors of non-users, Supplementary Tables S9, S10). We considered a gene to be a potential target if an ERα peak was positioned within the gene body, or within 20 kb upstream of the transcriptional start site (TSS), as we performed before[5]. Using Ingenuity Pathway Analysis, we found that estrogen receptor (ESR1) was a potential upstream regulator of genes proximal to ERα binding sites differentially enriched in endometrial tumors of tamoxifen-users, but not non-users. Instead the top potential upstream regulator for genes proximal to ERα binding sites in endometrial tumors of non-users was HSPA5, a heat shock protein (Figure 3A). These data suggest that between differential ERα sites of both groups, only the binding sites enriched in tamoxifen-associated endometrial tumors regulate known target genes of ERα. Gene ontology analysis revealed a number of biological processes that were specific for the genes potentially targeted by ERα peaks enriched in non-users, including negative regulation of collagen biosynthesis and metabolism, negative regulation of multicellular organismal metabolic process and phosphorus metabolic process (Supplementary Tables S11, S12).


A
Read count at ERα binding sites differential in endometrial cancer
breast cancer N=30
Differential
enriched in endometrial tumors of non-users
enriched in endometrial tumors of tamoxifen-users
B
ERα sites differentiallly enriched in endometrial tumors in
Tamoxifen-users
Non-users
p = 10-12
Read count in ERα peaks in breast cancer, log2 RPKM
Ross-Innes et al.
Jansen et al.
ERα ChIP-seq data in Breast Cancers
C
Tumors
Correlation
Endometrial tumors of
Non-users
Tamoxifen-users
Breast Cancers
D
Gene set enrichment analysis
LogFC endometrial cancer vs. breast cancer
TCGA
Ranked Gene list
Endometrium
Genes near differential ERα sites
Enrichment score
Non-users
FDR<0.001
Nominal p-value<0.001
E
Transcription factors in different tissues
Transcription factors
Endometrial cancer of tamoxifen-users (enriched ERα sites)
Endometrial cancer of non-users (enriched ERα sites)
Breast cancer cell lines (MCF-7, T47D)
Endometrial cancer cell line (Ishikawa)
Correlation
-0.33
1

We further investigated the regulatory link between differential ERα binding sites and gene expression by ranking genes from the microarray data according to the difference in expression between endometrial tumors from tamoxifen-users and non-users, and used six gene sets for gene set enrichment analysis (GSEA): (1) genes proximal to ERα binding sites enriched in tamoxifen-associated endometrial cancer; (2) genes proximal to ERα binding sites enriched in endometrial tumors of non-users; (3/4) genes upregulated/downregulated by estradiol in MCF-7 breast cancer cell line; (5/6) and genes upregulated/downregulated by estradiol in endometrial cancer cell line Ishikawa. RNA expression levels of genes proximal to enriched ERα sites are higher in the corresponding tumor group compared with the tumor group in which the ChIP-seq signal at these ERα sites is less pronounced. Genes upregulated by estradiol in breast cancer cell line MCF-7 are among the genes that are higher expressed in endometrial tumors from tamoxifen-users (Figure 3B).

In order to further focus on transcriptional effects of the differential ERα binding sites, we narrowed the list of potential target genes by combining gene expression and ChIP-seq data by means of gene set enrichment analysis: only the genes that contributed to the leading edge (core enrichment) in the GSEA analysis were taken (Supplementary Table S13, S14). Ingenuity Pathway Analysis revealed a strong enrichment of ESR1-regulated genes (15 out of 70 target genes are described as ESR1-regulated, p = 4.7e-13) and constructed a functional network that is centered around ERα and includes well known targets such as PGR, RARA, TFF1, VAV3 and others (Figure 3C).

Figure 4. Comparative analysis of ERα binding sites in endometrial and breast tumor tissue.

A) Analysis set-up. ERα binding in breast cancer at the sites differential between the two endometrial cancer groups was evaluated.

B) Boxplot visualizing normalized ERα ChIP-seq read count in breast cancers at differential ERα binding sites enriched in endometrial tumors from tamoxifen-users (orange boxplots) and non-users (blue boxplots). The p-value of the paired t-test is $p = 10^{-12}$.

C) Heatmap visualization of the correlation matrix based on ERα ChIP-seq read count at differential ERα binding sites in endometrial tumors of tamoxifen-users (orange), non-users (blue), and in breast tumors (pink).

D) Gene Set Enrichment Analysis based on differential gene expression between endometrial and breast cancers from TCGA pan-cancer project. Top panel shows ranked log-fold change of gene expression between endometrial cancer and breast cancer. The lower panel shows enrichment scores versus gene rank in the significantly enriched gene set: genes proximal to the binding sites enriched in tumors of non-users.

E) Hierarchical clustering of the correlation between transcription factor genomic occupancy (peaks) from publicly available ChIP-seq data in the breast cancer cell lines MCF-7 and T47D (pink), the endometrial cancer cell line Ishikawa (grey), and the ERα sites enriched in endometrial tumors from tamoxifen-users (orange) and non-users (blue).

Taken together, these data reveal that the differential ERα binding sites between endometrial tumors from tamoxifen-users and non-users regulate gene expression differently. Gene expression of ERα targets in tamoxifen-associated endometrial tumors resemble estradiol-responsive genes in MCF-7, suggesting an ERα cistrome potentially similar to breast tumors.

ERα binding to DNA in tamoxifen-associated endometrial tumors resembles ERα chromatin binding in breast cancer

In contrast to endometrial tumors of non-users, ERα binding sites enriched in tamoxifen-associated endometrial tumors are proximal to known targets of estrogen receptor in breast cancer (Figure 3). Because tamoxifen has been reported to stimulate cell growth in the endometrium in postmenopausal patients, similar to estrogen in breast, we compared our findings to publicly available ChIP-seq data on ERα and other transcription factors in 30 primary breast tumors (Figure 4A).

To group the tumors according to similarity in ERα ChIP-seq signal, we performed hierarchical clustering of the two tumor types (breast and endometrial). We first analyzed global ERα ChIP-seq signal in the two endometrial groups (tamoxifen-users and non-users) and breast cancer at ERα binding sites present in at least five out of 46 tumors analyzed (N = 16516). Based on the ERα ChIP-seq read count, the tumors clustered on tumor type (Supplementary Figure S8).

Next, we focused on the sites that are differentially bound by ERα in endometrial tumors of tamoxifen-users compared to non-users. We analyzed the ERα ChIP-seq data of the 30 breast tumors at genomic sites that are differentially bound by ERα in tamoxifen-associated endometrial cancer. ERα ChIP-seq signal in breast tumors is significantly higher at sites that are enriched in tamoxifen-associated endometrial tumors compared to binding sites enriched in endometrial tumors from non-users (Figure 4B). In addition, unsupervised hierarchical clustering showed that the ERα read count at differential ERα sites (between endometrial tumors of tamoxifen-users and non-users) correlated most between breast tumors and tamoxifen-associated endometrial tumors (correlation heatmap: Figure 4C, readcount at differential sites: Supplementary Figure S9).

To investigate whether differential ERα binding sites might regulate gene expression in a variable manner between endometrial cancer and breast cancer, we used gene expression data from the TCGA pan-cancer project[38] for gene set enrichment analysis. We ranked genes from the TCGA data according to expression fold change between endometrioid adenocarcinoma and estrogen receptor-positive breast cancer, and used 2 gene sets for the analysis: (1) genes proximal to ERα sites enriched in endometrial

tumors from tamoxifen-users (2) genes proximal to ERα sites enriched in endometrial tumors from non-users. We found one gene set (genes proximal to ERα sites enriched in non-users) enriched among the genes that are higher expressed in endometrial tumors compared to breast tumors (Figure 4D). This analysis suggests that the ERα sites enriched in tumors from non-users indeed are involved in execution of transcriptional programs specific for endometrial cancer.

To investigate the transcription factor network in which the differential ERα binding sites function, we used public cell line data. We correlated the differentially enriched ERα binding sites of endometrial tumors from tamoxifen-users and non-users with binding sites of several transcription factors from the endometrial cancer cell line Ishikawa, and the breast cancer cell lines MCF-7 and T47D (Figure 4E). Transcription factors involved in DNA looping (CTCCF-binding factor (CTCF) and double strand break repair protein rad21 homolog (RAD21)) but also transcription factors that bind promoters (RNA polymerase POL2RA, serum response factor SRF, and transcription initiation factor TFIID subunit 1 TAF1) cluster together irrespective of cell line or tissue type. In contrast, enhancer-binding transcription factors (ESR1, FOXA1, EP300) cluster according to tissue type. In accordance with Figure 2B, the differential ERα binding sites of endometrial tumors cluster with enhancer-binding transcription factors rather than promoter-binding factors. ERα binding sites enriched in tumors from non-users clustered with transcription factor binding sites in Ishikawa, whereas ERα binding sites enriched in tamoxifen-associated endometrial tumors clustered with transcription factor binding sites in breast cancer cell lines. Taken together, these data illustrate a resemblance between breast cancer and tamoxifen-associated endometrial cancer at sites that are enriched for ERα in endometrial tumors of tamoxifen-users compared to non-users. In contrast, genomic regions enriched for ERα in endometrial tumors from non-users correspond to ERα enhancers in the endometrial cancer cell line Ishikawa and potentially regulate expression of genes more specific for endometrial tumors.

Discussion

We found that even though on a morphological level endometrioid adenocarcinomas of tamoxifen-users and non-users are indistinguishable, a large part of the ERα cistrome — and its downstream transcriptional programs — differ. The differential ERα binding sites have distinct underlying DNA sequences and potential regulatory function. Interestingly, ERα binding to the DNA in tamoxifen-associated endometrial tumors resembles ERα chromatin binding in breast cancer, highlighting a conserved ERα pathway between the two tumor types from different organs despite different ligands.

Studies in the breast cancer cell line MCF-7 show that prolonged tamoxifen exposure shifts the ERα cistrome[4-6], which consequently changes gene expression[4], possibly by changing its interactome[2,43-45]. These data are hard to translate between tissues because there are far less models for endometrial cancer to study ERα. Thus far, a tamoxifen-associated endometrial cancer model is nonexistent; the only model for the effects of tamoxifen in endometrial tissue is the endometrial cancer cell line Ishikawa, which is derived from a tumor of a non-user[46]. The effects of tamoxifen on the ERα-cistrome in this model lack genome-wide data[37,47]. Our previous study was the first to show genome-wide ERα binding sites in tamoxifen-associated endometrial cancer, but it lacked data to identify differential ERα sites compared to endometrial tumors of non-users[18].

Using patient samples from the unique TAMARISK-study (Table 1), we now reveal that the ERα cistrome differs between endometrial tumors that originated from different hormonal backgrounds (tamoxifen-rich vs. tamoxifen-free), illustrating that these tumors are epigenetically distinguishable. Prior chemotherapy usage for the treatment of breast cancer did not differ between the two patient groups (Table 1), precluding differences in systemic therapy beyond tamoxifen usage as a potential confounder. Furthermore, we excluded a genetic predisposition in the form of Lynch Syndrome in either patient group by showing that microsatellite instability between tamoxifen-users and non-users was comparable (Supplementary Table S1)[48]. Yet unknown genetic predispositions however, cannot be excluded at this point.

Although several studies report on the effects of ligands, including tamoxifen, on the conformation of ERα, other determinants of the ERα cistrome in endometrial tissue remain obscure. The motifs we found hint at proteins involved at ERα/chromatin interactions at differential ERα binding sites between endometrial tumors of tamoxifen-users and non-users. These motifs indicate a role for stem cell markers, such as SOX4[39] and Nanog[40], in non-users, and for members of the nuclear receptor family in tamoxifen-users, including the androgen receptor, glucocorticoid, and thyroid hormone receptor.

Compared to endometrial tumors of non-users, tamoxifen-associated endometrial tumors showed upregulation of genes involved in pathways that contribute to cancer progression such as EMT, RB, TP53 and interferon targets. These data suggest that endometrial tumors that originate in presence of tamoxifen may expose intrinsic different tumor biology, resulting in different tumor-drivers in this setting. Our previous immunohistochemical studies[13] have shown that longer tamoxifen exposure relates with worse survival, higher TP53 expression and lower ESR1 expression[13], which is in line with our current results.

Previous studies have illustrated that enhancer activity differs per tissue[49]. Our data reveal that ERα profiles in tamoxifen-associated endometrial tumors resemble those found in breast tumors, suggesting that endocrine stimuli reprogram this pathway in endometrial tissue.

To conclude, our study sheds new light on the ERα cistrome and gene expression regulation in endometrial tumors, and implicate that the two kinds of endometrioid adenocarcinomas that we investigated in this report are clearly distinguishable on a cistromic and transcriptional level, albeit morphologically identical. Our results pave the way for new discoveries in endometrial cancer, and further highlight the added value of cistromic analyses in clinical specimens, especially in settings were model systems are not available. By functionally distinguishing tumors on the transcriptional regulation-level, novel subtypes may be revealed with further clinical and prognostic implications.

Acknowledgments

We thank Ron Kerkhoven, Shan Baban and Marja Nieuwland from the NKI genomics facility for sample processing and Arno Velds for bioinformatics support. We thank Lisette Hoogendoorn for clinical data collection. We thank Koen van de Vijver for help with pathological analyses. We would like to acknowledge the NKI- AVL Core Facility Molecular Pathology& Biobanking (CFMPB) for supplying NKI-AVL Biobank material and lab support. This work was supported by grants from the Dutch Cancer Society KWF (NKI 2002-586) and Pink Ribbon.

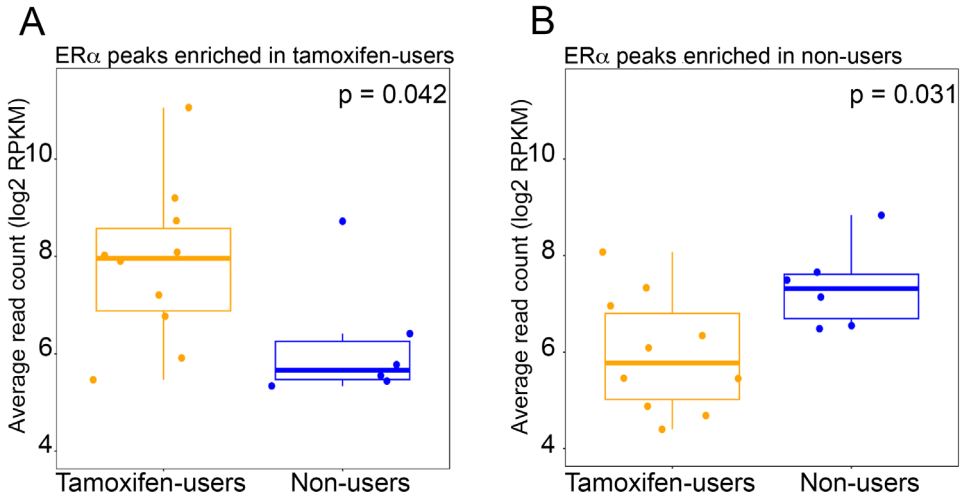
References

1. Droog M, Mensink M, & Zwart W (2016) The Estrogen Receptor alpha-Cistrome Beyond Breast Cancer. Mol Endocrinol 30(10):1046-1058.
2. Shiau AK, et al. (1998) The structural basis of estrogen receptor/coactivator recognition and the antagonism of this interaction by tamoxifen. Cell 95(7):927-937.
3. Flach KD & Zwart W (2016) The first decade of estrogen receptor cistromics in breast cancer. J Endocrinol 229(2):R43-56.
4. Hurtado A, Holmes KA, Ross-Innes CS, Schmidt D, & Carroll JS (2011) FOXA1 is a key determinant of estrogen receptor function and endocrine response. Nature genetics 43(1):27-33.
5. Severson TM, et al. (2016) Neoadjuvant tamoxifen synchronizes ERalpha binding and gene expression profiles related to outcome and proliferation. Oncotarget.
6. Ross-Innes CS, et al. (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481(7381):389-393.

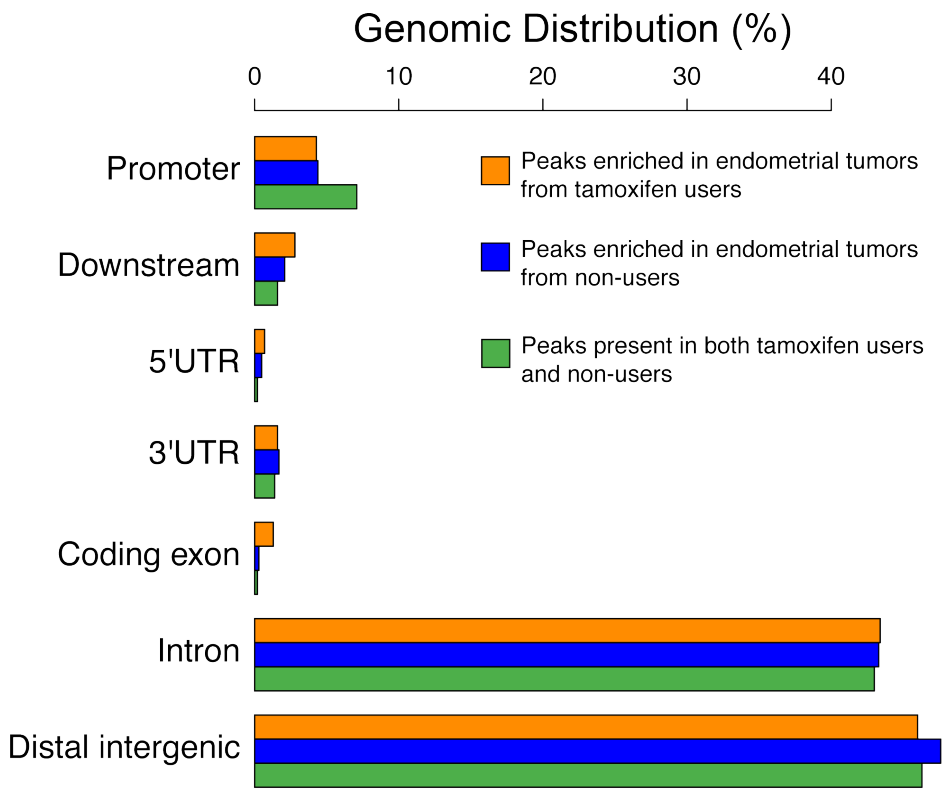
7. Osborne CK, Hobbs K, & Clark GM (1985) Effect of estrogens and antiestrogens on growth of human breast cancer cells in athymic nude mice. Cancer research 45(2):584-590.
8. Satyaswaroop PG, Zaino RJ, & Mortel R (1984) Estrogen-like effects of tamoxifen on human endometrial carcinoma transplanted into nude mice. Cancer research 44(9):4006-4010.
9. van Leeuwen FE, et al. (1994) Risk of endometrial cancer after tamoxifen treatment of breast cancer. Lancet 343(8895):448-452.
10. Fisher B, et al. (1994) Endometrial cancer in tamoxifen-treated breast cancer patients: findings from the National Surgical Adjuvant Breast and Bowel Project (NSABP) B-14. J Natl Cancer Inst 86(7):527-537.
11. Lahti E, et al. (1993) Endometrial changes in postmenopausal breast cancer patients receiving tamoxifen. Obstet Gynecol 81(5 (Pt 1)):660-664.
12. Bergman L, et al. (2000) Risk and prognosis of endometrial cancer after tamoxifen for breast cancer. Comprehensive Cancer Centres' ALERT Group. Assessment of Liver and Endometrial cancer Risk following Tamoxifen. Lancet 356(9233):881-887.
13. Hoogendoorn WE, et al. (2008) Prognosis of uterine corpus cancer after tamoxifen treatment for breast cancer. Breast cancer research and treatment 112(1):99-108.
14. Jones ME, et al. (2012) Endometrial cancer survival after breast cancer in relation to tamoxifen treatment: pooled results from three countries. Breast cancer research : BCR 14(3):R91.
15. Swerdlow AJ, Jones ME, & British Tamoxifen Second Cancer Study G (2005) Tamoxifen treatment for breast cancer and risk of endometrial cancer: a case-control study. J Natl Cancer Inst 97(5):375-384.
16. Gottardis MM, Robinson SP, Satyaswaroop PG, & Jordan VC (1988) Contrasting actions of tamoxifen on endometrial and breast tumor growth in the athymic mouse. Cancer research 48(4):812-815.
17. Shang Y & Brown M (2002) Molecular determinants for the tissue specificity of SERMs. Science 295(5564):2465-2468.
18. Droog M, et al. (2016) Comparative Cistromics Reveals Genomic Crosstalk between FOXA1 and ERalpha in Tamoxifen-Associated Endometrial Carcinomas. Cancer research 76(13):3773-3784.
19. Fles R, et al. (2010) Genomic profile of endometrial tumors depends on morphological subtype, not on tamoxifen exposure. Genes, chromosomes & cancer 49(8):699-710.
20. Zwart W, et al. (2011) Oestrogen receptor-co-factor-chromatin specificity in the transcriptional regulation of breast cancer. The EMBO journal 30(23):4764-4776.
21. Jansen MP, et al. (2013) Hallmarks of aromatase inhibitor drug resistance revealed by epigenetic profiling in breast cancer. Cancer research 73(22):6632-6641.
22. Zhang Y, et al. (2008) Model-based analysis of ChIP-Seq (MACS). Genome biology 9(9):R137.

23. Kumar V, et al. (2013) Uniform, optimal signal processing of mapped deep-sequencing data. Nature biotechnology 31(7):615-622.
24. Kuilman T, et al. (2015) CopywriteR: DNA copy number detection from off-target sequence data. Genome biology 16:49.
25. Ye T, et al. (2011) seqMINER: an integrated ChIP-seq data interpretation platform. Nucleic acids research 39(6):e35.
26. Muller J (2016) TransView: Read density map construction and accession. Visualization of ChIPSeq and RNASeq data sets. R package version 1.16.0.
27. Ji X, Li W, Song J, Wei L, & Liu XS (2006) CEAS: cis-regulatory element annotation system. Nucleic acids research 34(Web Server issue):W551-554.
28. Liu T, et al. (2011) Cistrome: an integrative platform for transcriptional regulation studies. Genome biology 12(8):R83.
29. Mi H, Muruganujan A, Casagrande JT, & Thomas PD (2013) Large-scale gene function analysis with the PANTHER classification system. Nature protocols 8(8):1551-1566.
30. Consortium EP (2012) An integrated encyclopedia of DNA elements in the human genome. Nature 489(7414):57-74.
31. Yang YH, et al. (2002) Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic acids research 30(4):e15.
32. Ritchie ME, et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic acids research 43(7):e47.
33. Subramanian A, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 102(43):15545-15550.
34. Isserlin R, Merico D, Voisin V, & Bader GD (2014) Enrichment Map - a Cytoscape app to visualize and explore OMICs pathway enrichment results. F1000Research 3:141.
35. Shannon P, et al. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 13(11):2498-2504.
36. Dunning MJ, Smith ML, Ritchie ME, & Tavare S (2007) beadarray: R classes and methods for Illumina bead-based data. Bioinformatics 23(16):2183-2184.
37. Gertz J, Reddy TE, Varley KE, Garabedian MJ, & Myers RM (2012) Genistein and bisphenol A exposure cause estrogen receptor 1 to bind thousands of sites in a cell type-specific manner. Genome research 22(11):2153-2162.
38. Cancer Genome Atlas Research N, et al. (2013) The Cancer Genome Atlas Pan-Cancer analysis project. Nature genetics 45(10):1113-1120.
39. Huang YW, et al. (2009) Epigenetic repression of microRNA-129-2 leads to overexpression of SOX4 oncogene in endometrial cancer. Cancer research 69(23):9038-9046.
40. Hubbard SA, et al. (2009) Evidence for cancer stem cells in human endometrial carcinoma. Cancer research 69(21):8241-8248.

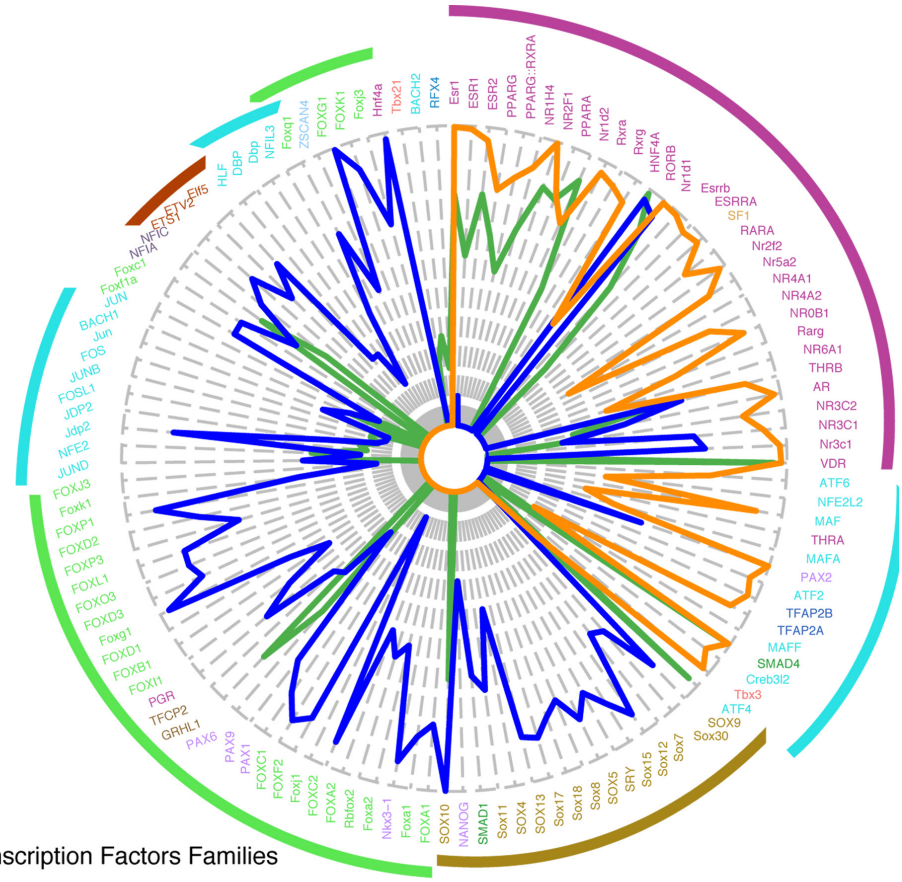
41. Carroll JS, et al. (2005) Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. Cell 122(1):33-43.
42. Nyholm HC, Nielsen AL, Lyndrup J, Norup P, & Thorpe SM (1992) Biochemical and immunohistochemical estrogen and progesterone receptors in adenomatous hyperplasia and endometrial carcinoma: correlations with stage and other clinicopathologic features. American journal of obstetrics and gynecology 167(5):1334-1342.
43. Mohammed H, et al. (2016) Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. Nature protocols 11(2):316-326.
44. Brzozowski AM, et al. (1997) Molecular basis of agonism and antagonism in the oestrogen receptor. Nature 389(6652):753-758.
45. Mohammed H, et al. (2013) Endogenous purification reveals GREB1 as a key estrogen receptor regulatory factor. Cell reports 3(2):342-349.
46. Nishida M, Kasahara K, Kaneko M, Iwasaki H, & Hayashi K (1985) [Establishment of a new human endometrial adenocarcinoma cell line, Ishikawa cells, containing estrogen and progesterone receptors]. Nihon Sanka Fujinka Gakkai zasshi 37(7):1103-1111.
47. Gertz J, et al. (2013) Distinct properties of cell-type-specific and shared transcription factor binding sites. Mol Cell 52(1):25-36.
48. Risinger JI, et al. (1993) Genetic instability of microsatellites in endometrial carcinoma. Cancer research 53(21):5100-5103.
49. Andersson R, et al. (2014) An atlas of active enhancers across human cell types and tissues. Nature 507(7493):455-461.



Supplementary Figure S1. Boxplots visualizing the normalized average read count of differentially enriched ERα sites in endometrial tumors of tamoxifen users (A) and non-users (B). P-values are according to Mann-Whitney test.



Supplementary Figure S2. Genomic distribution of ERα sites that are either shared (green) or differentially enriched in endometrial tumors from tamoxifen-users (orange) and non-users (blue), relative to the nearest gene.

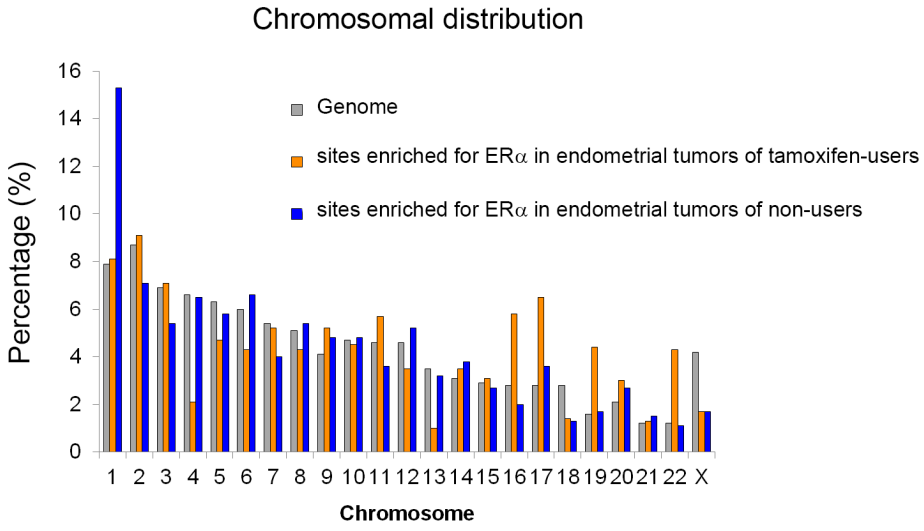


Transcription Factors Families

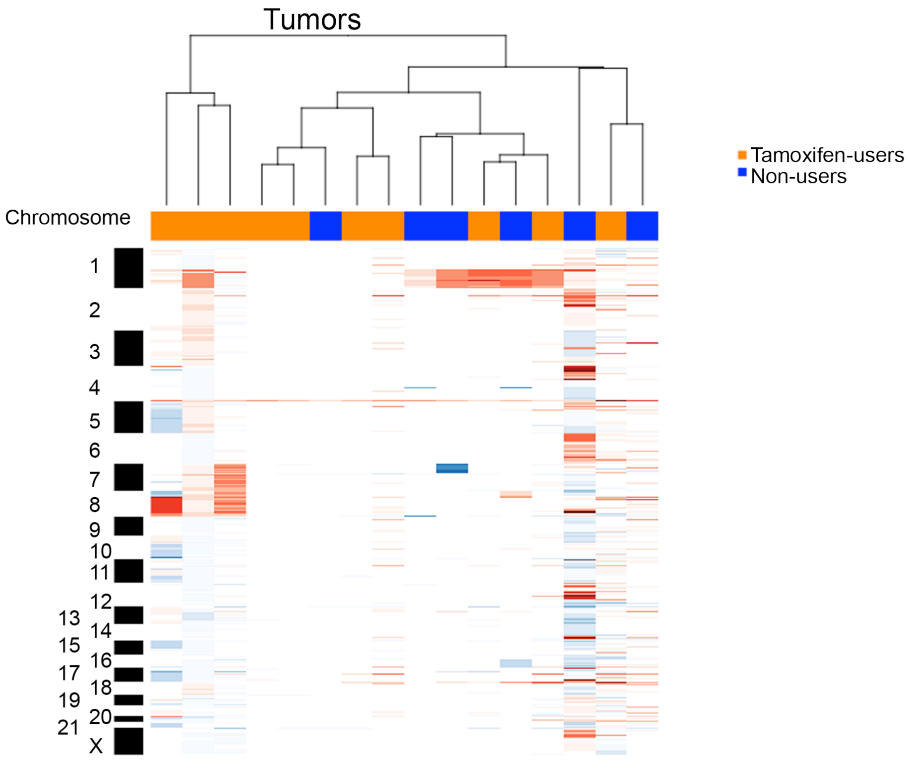
High Mobility Group (Box) Family
MH1 Domain Family
Homeodomain Family
Forkhead Domain Family
Hormone–nuclear Receptor Family
CP2 Transcription Factor Domain Family
Leucine Zipper Family
NFI–CTF–binding Transcription Factor Family
Ets Domain Family
Beta Beta Alpha–zincfinger Family
Cys4 zincfinger of nuclear receptor type
Helix–Loop–Helix Family
Transcription Factor T–Domain
RFX Domain Family

Peaks enriched in tamoxifen users
Peaks enriched in non-users
Peaks present in both

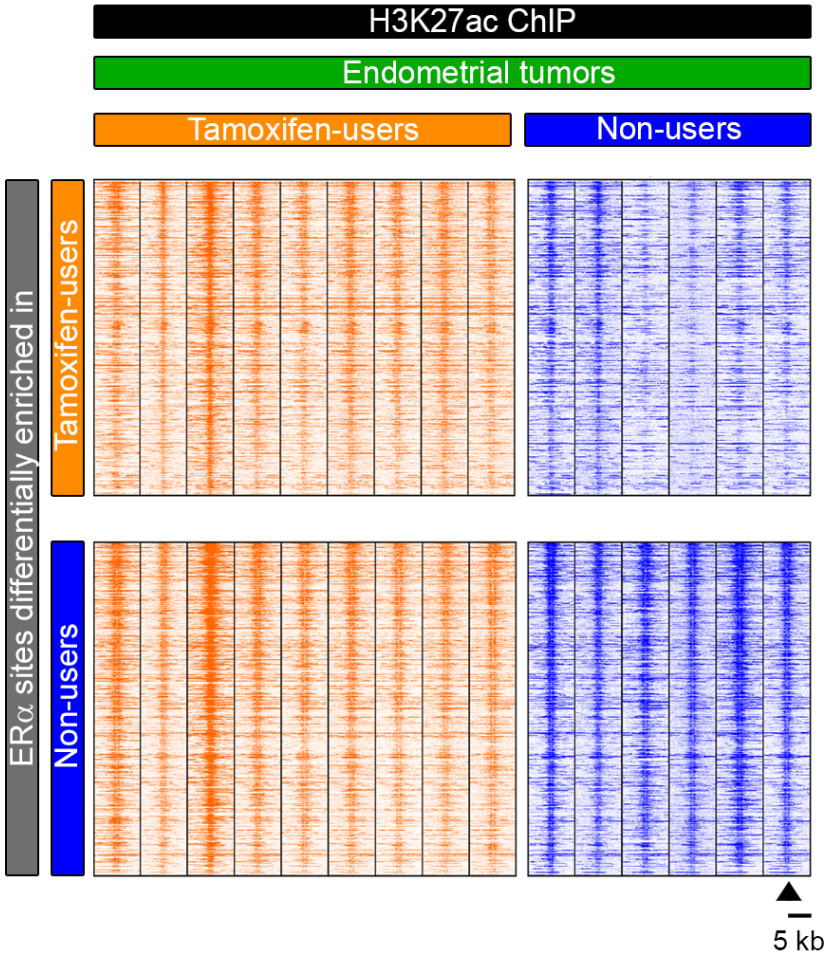
Supplementary Figure S3. Radar plot, visualizing DNA motif enrichment at genomic ERα sites that are either shared (green) or differentially enriched in endometrial tumors from tamoxifen-users (orange), or non-users (blue). Lengths of radii correspond to the fraction of peaks that contain the identified motif. Motif colors correspond to transcription factor families.



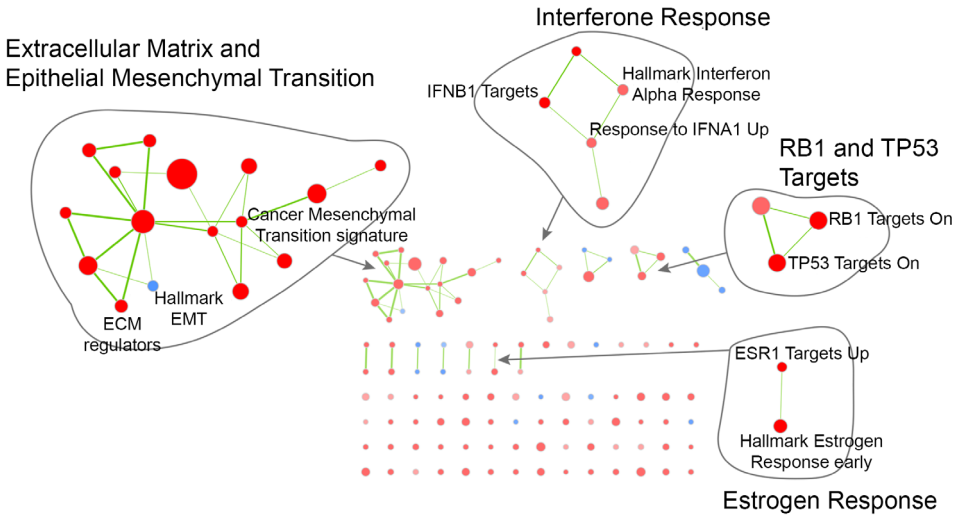
Supplementary Figure S4. Chromosomal distribution of ERα binding sites differentially enriched in endometrial tumors from tamoxifen-users (orange) or non-users (blue).



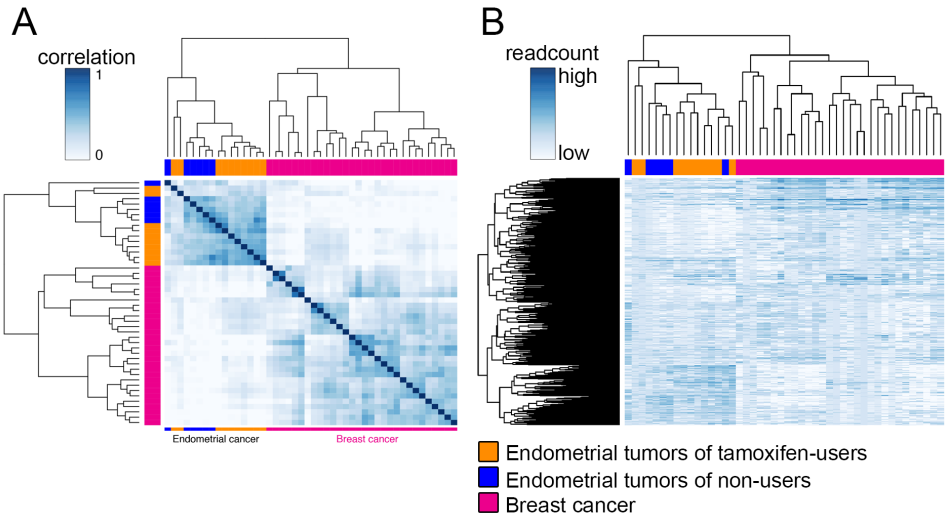
Supplementary Figure S5. Copy Number Variations analysis of endometrial cancers used for ChIP-seq (red indicates genomic gains, blue indicates loss). Copy number profiles are clustered based on correlation; endometrial tumors from tamoxifen-users and non-users do not cluster according to tumor group.



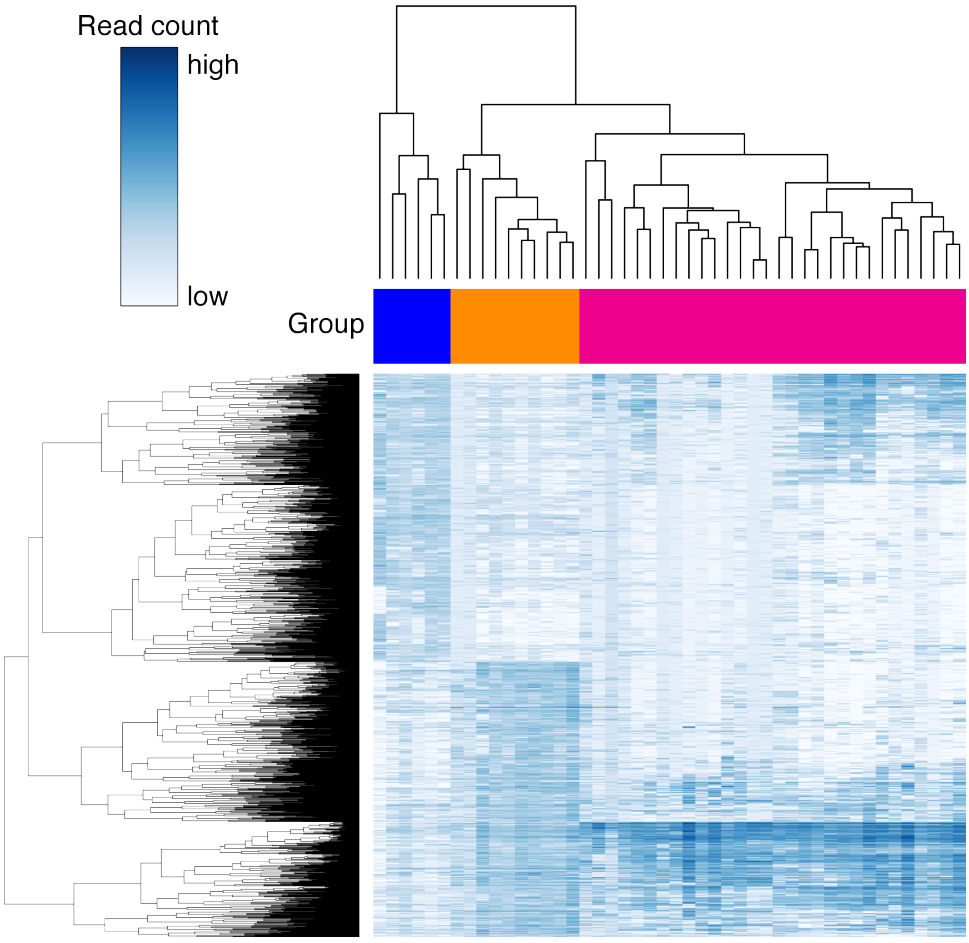
Supplementary Figure S6. Heatmap visualizing raw read count intensity of H3K27ac in endometrial tumors from tamoxifen-users (orange) and non-users (blue) at DNA sites that are differentially bound by ERα per group.



Supplementary Figure S7. The networks illustrate the results of the pathway enrichment analysis (with GSEA) of differential gene expression between tamoxifen users and non-users. Nodes represent pathways that link overlapping genes (overlap coefficient cutoff 0.5). Red nodes illustrate upregulated pathways in tamoxifen users while blue nodes represent upregulated pathways in non-users.



Supplementary Figure S8. Heatmap visualization of the correlation matrix based on ERα ChIP-seq read count at ERα binding sites in endometrial tumors of tamoxifen-users (orange), non-users (blue), and in breast tumors (pink). Only peaks found in at least five tumors were included.
Hierarchical clustering of ERα ChIP-seq signal in breast tumors (pink) and endometrial tumors (tamoxifen users are orange, non-users blue) in peaks present in at least five tumors.



Supplementary Figure S9. Hierarchical clustering of ERα ChIP-seq signal in endometrial tumors from tamoxifen-users (orange), non-users (blue), and breast tumors (pink) in peaks that are differential between endometrial tumors of tamoxifen users and non-users.

Supplementary Table S1. Clinicopathological parameters of the endometrioid adenocarcinomas from the retrospective part of the TAMARISK study.

	Tam-user	Non-user
N	67	129
Tamoxifen use in years, median (range)	3.8 (2.0-17.9)	N/A
Interval time breast and endometrial cancer in years, median (range)	7.2 (2.1-18.7)	5.8 (1.0-21.5)
Age breast cancer, median (range)	64.0 (34.9-79.5)	59.9 (37.4-87.9)
Age endometrial cancer, median (range)	70.8 (52.1-86.6)	68.0 (44.1-93.5)
Menopausal status, n (%)		
Postmenopausal status*	66 (98.5%)	120 (93.0%)
Perimenopausal status*	1 (1.5%)	5 (3.9%)
Premenopausal status*	0 (0.0%)	3 (2.3%)
Unknown*	0 (0.0%)	1 (0.8%)
Recency tamoxifen use as to surgery, n (%)		
Ex-user	20 (29.9%)	0 (0.0%)
Recent user (last 12 months)	47 (70.1%)	0 (0.0%)
Current user	0 (0.0%)	0 (0.0%)
Non-user	0 (0.0%)	129 (100.0%)
Unknown	0 (0.0%)	0 (0.0%)
Histological type, n (%)		
Endometrioid adenocarcinoma and variants	67 (100.0%)	129 (100.0%)
Use of other hormonal therapy, n (%)**	6 (9.0%)	9 (7.0%)
FIGO stage, n (%)		
I	43 (64.2%)	99 (76.7%)
II	3 (4.5%)	7 (5.4%)
III/IV	1 (1.5%)	0 (0.0%)
Unknown	19 (28.4%)	21 (0.0%)
MSI status, n (%)		
Negative or low	42 (62.7%)	74 (57.4%)
High	9 (13.4%)	25 (19.4%)
Unknown	16 (23.9%)	30 (23.3%)
Chemotherapy, n (%)		
Yes	7 (10.4%)	12 (9.3%)
No	60 (89.6%)	115 (89/1%)
Unknown	0 (0.0%)	2 (1.6%)

*During diagnosis of endometrial cancer.
There are no significant differences between tamoxifen-users and non-users with regard to interval time between breast and endometrial cancer, the age of breast cancer, or the age of endometrial cancer, according to t-test; no difference in FIGO stage, use of chemotherapy or MSI status, according to Chi-square test.

Supplementary Table S2. Number of sequencing reads, mapped reads and called peaks.

SampleID	Factor	Group	Total reads	Mapped reads	% Mapped reads	Filtered reads	% Filtered reads	No peaks
wz154	ERα	tamoxifen-user	15331544	14207826	92.67	12342539	86.87	1589
wz483	ERα	tamoxifen-user	18754169	17283692	92.16	15005839	86.82	11560
wz484	ERα	tamoxifen-user	19287976	18138678	94.04	15509489	85.51	2355
wz489	ERα	tamoxifen-user	18569672	16528291	89.01	14129722	85.49	714
wz487	ERα	tamoxifen-user	23924152	15740024	65.79	13543920	86.05	100
WZ485	ERα	tamoxifen-user	32296216	26465862	81.95	22901235	86.53	504
wz439	ERα	tamoxifen-user	19603948	17976811	91.70	15752509	87.63	21800
wz152	ERα	tamoxifen-user	14930170	12101826	81.06	10116835	83.6	27915
wz170	ERα	tamoxifen-user	20680117	19906669	96.26	17433136	87.57	1631
wz438	ERα	tamoxifen-user	15611205	14725046	94.32	12681132	86.12	123
wz490	ERα	non-user	20899419	18917776	90.52	16402313	86.70	455
wz488	ERα	non-user	19444847	18274374	93.98	16019957	87.66	26714
wz486	ERα	non-user	30403020	18373438	60.43	15788112	85.93	19055
wz470	ERα	non-user	32936222	29363575	89.15	24987722	85.10	1312
wz503	ERα	non-user	14880365	13780675	92.61	11885851	86.25	391
wz504	ERα	non-user	22411738	21007108	93.73	18174472	86.52	3223
wz984	H3K27ac	tamoxifen-user	23175814	22016196	95.00	20519524	93.20	
wz985	H3K27ac	tamoxifen-user	22900854	22227909	97.06	20621089	92.77	
wz565	H3K27ac	tamoxifen-user	41850397	40395208	96.52	36913144	91.38	
wz990	H3K27ac	tamoxifen-user	29613393	28327999	95.66	26416294	93.25	
wz991	H3K27ac	tamoxifen-user	24977421	23739843	95.05	21921727	92.34	
wz992	H3K27ac	tamoxifen-user	31443859	29791112	94.74	27645455	92.80	
wz1368	H3K27ac	tamoxifen-user	20886714	19953687	95.53	18656222	93.50	
wz1045	H3K27ac	tamoxifen-user	25954861	24892026	95.91	23468485	94.28	
wz981	H3K27ac	tamoxifen-user	26601717	25544357	96.03	23749000	92.97	
Tam_5	H3K27ac	non-user	27210981	26584367	97.70	25114577	94.47	
Tam_33	H3K27ac	non-user	28796446	27937355	97.02	26006528	93.09	
Tam_38	H3K27ac	non-user	23348596	22763584	97.49	21272824	93.45	
Tam_55	H3K27ac	non-user	29667180	28638563	96.53	26233736	91.60	
Tam_102	H3K27ac	non-user	22491570	21890598	97.33	20508226	93.69	
Tam_119	H3K27ac	non-user	22791103	22188485	97.36	20885046	94.13	

Supplementary Table S3. Public cell lines ChIP-seq data.

Cell line	Factor	Source	Accession
MCF7	CTCF	Encode	ENCSR000DMS
MCF7	CTCF	Encode	ENCSR000DMR
Ishikawa	CTCF	Encode	ENCSR000BQE
T47D	CTCF	Encode	ENCSR000BNO
MCF7	TCF12	Encode	ENCSR000BUN
Ishikawa	TCF12	Encode	ENCSR000BUV
MCF7	FOXM1	Encode	ENCSR000BUJ
Ishikawa	FOXM1	Encode	ENCSR000BUS
MCF7	REST	Encode	ENCSR000BSP
Ishikawa	REST	Encode	ENCSR000BUU
MCF7	EGR1	Encode	ENCSR000BUX
Ishikawa	EGR1	Encode	ENCSR000BSQ
MCF7	MAX	Encode	ENCSR000BUL
Ishikawa	MAX	Encode	ENCSR000BTY
MCF7	EP300	Encode	ENCSR000BTR
Ishikawa	EP300	Encode	ENCSR000BUE
T47D	EP300	Encode	ENCSR000BLM
MCF7	RAD21	Encode	ENCSR000BTQ
Ishikawa	RAD21	Encode	ENCSR000BTU
MCF7	CEBPB	Encode	ENCSR000BSR
Ishikawa	CEBPB	Encode	ENCSR000BTT
MCF7	TEAD4	Encode	ENCSR000BUO
Ishikawa	TEAD4	Encode	ENCSR000BSW
MCF7	SRF	Encode	ENCSR000BVA
Ishikawa	SRF	Encode	ENCSR000BTD
MCF7	TAF1	Encode	ENCSR000AHF
Ishikawa	TAF1	Encode	ENCSR000BTO
MCF7	POLR2	Encode	ENCSR000DMN
MCF7	POLR2	Encode	ENCSR000DMT
MCF7	POLR2	Encode	ENCSR000DMK
Ishikawa	POLR2	Encode	ENCSR000BKI
Ishikawa	ESR1	Encode	ENCSR000BIY
T47D	ESR1	Encode	ENCSR000BKN
Ishikawa	FOXA1	Encode	ENCSR000BKW

T47D	FOXA1	Encode	ENCSR000BKY
MCF7	FOXA1	ArrayExpress	ERR022028/ERR022029
MCF7	ESR1	ArrayExpress	ERR022052/ERR022053/ERR022048/ERR022049/ ERR022026

Supplementary Tables S4 to **S10** are too long to be printed. Digital copies are available on request:

Supplementary Table S4. Genomic coordinates of the ERα binding sites enriched in tamoxifen-users.

Supplementary Table S5. Genomic coordinates of the ERα binding sites enriched in non-users.

Supplementary Table S6. Genomic coordinates of the peaks that are shared (non-differential) between tamoxifen-users and non-users. Supplementary Table S7. Pathways enriched among the genes upregulated in tamoxifen users.

Supplementary Table S7. Pathways enriched among the genes upregulated in tamoxifen users.

Supplementary Table S8. Pathways enriched among the genes upregulated in non-users.

Supplementary Table S9. Genes potentially targeted by ERα peaks enriched in tamoxifen users as defined by peak present in 20kb upstream of the TSS of the gene or in the gene body.

Supplementary Table S10. Genes potentially targeted by ERα peaks enriched in non-users as defined by peak present in 20kb upstream of the TSS of the gene or in the gene body.

Supplementary Table S11. Gene Ontology biological processes enriched among the genes potentially targeted by ERα peaks enriched in tamoxifen users.

GO biological process	Number of genes in the reference set	Number of genes in the uploaded list	Expected number of genes	Enrichment (over/under)	Fold Enrichment	Enrichment p-value
system development (GO:0048731)	4042	111	69.38	+	1.60	9.17E-04
negative regulation of biological process (GO:0048519)	4622	122	79.34	+	1.54	1.41E-03
single-multicellular organism process (GO:0044707)	5417	137	92.99	+	1.47	1.98E-03
negative regulation of cellular process (GO:0048523)	4281	114	73.49	+	1.55	2.90E-03
single-organism process (GO:0044699)	12544	260	215.33	+	1.21	5.02E-03
animal organ development (GO:0048513)	2888	84	49.57	+	1.69	6.01E-03
developmental process (GO:0032502)	5333	133	91.54	+	1.45	7.69E-03
multicellular organism development (GO:0007275)	4640	119	79.65	+	1.49	1.02E-02
single-organism cellular process (GO:0044763)	11217	237	192.55	+	1.23	1.12E-02
biological_process (GO:0008150)	17072	325	293.05	+	1.11	1.95E-02
single-organism developmental process (GO:0044767)	5239	129	89.93	+	1.43	2.55E-02
anatomical structure development (GO:0048856)	4986	124	85.59	+	1.45	2.71E-02

Supplementary Table S12. Gene Ontology biological processes enriched among the genes potentially targeted by ERα peaks enriched in non-users.

GO biological process complete	Number of genes in the reference set	Number of genes in the uploaded list	Expected number of genes	Enrichment (over/under)	Fold Enrichment	Enrichment p-value
single-multicellular organism process (GO:0044707)	5417	159	99.19	+	1.60	1.88E-07
multicellular organism development (GO:0007275)	4640	142	84.96	+	1.67	2.47E-07
anatomical structure development (GO:0048856)	4986	146	91.29	+	1.60	2.87E-06
single-organism process (GO:0044699)	12544	286	229.68	+	1.25	9.51E-06
single-organism developmental process (GO:0044767)	5239	148	95.93	+	1.54	2.66E-05
anatomical structure morphogenesis (GO:0009653)	1956	73	35.81	+	2.04	3.59E-05
developmental process (GO:0032502)	5333	149	97.65	+	1.53	4.93E-05
system development (GO:0048731)	4042	120	74.01	+	1.62	1.27E-04
multicellular organismal process (GO:0032501)	6482	168	118.69	+	1.42	6.79E-04
single-organism cellular process (GO:0044763)	11217	255	205.38	+	1.24	1.55E-03
animal organ development (GO:0048513)	2888	90	52.88	+	1.70	2.03E-03
negative regulation of collagen biosynthetic process (GO:0032966)	10	5	0.18	+	27.31	1.20E-02
biological_process (GO:0008150)	17072	346	312.59	+	1.11	1.51E-02
negative regulation of collagen metabolic process (GO:0010713)	11	5	0.2	+	24.82	1.90E-02
phosphorus metabolic process (GO:0006793)	2192	70	40.14	+	1.74	2.69E-02
negative regulation of multicellular organismal metabolic process (GO:0044252)	12	5	0.22	+	22.76	2.89E-02

Supplementary Table S13. Potential target genes of ERα peaks enriched in tamoxifen users as identified by combined analysis of gene expression and ChIP-seq. Genes in the list contain ERα peak 20kb upstream of TSS or in the gene body and strongly contribute to the enrichment among upregulated genes in gene set enrichment analysis (GSEA).

Gene Symbol	Core Enrichment
ACTN4	Yes
ANO1	Yes
B4GALT1	Yes
BCL2	Yes
C16orf45	Yes
CA12	Yes
CGNL1	Yes
CLIC3	Yes
CTSD	Yes
DCXR	Yes
DOCK1	Yes
HS6ST1	Yes
IER3	Yes
IGFBP4	Yes
KLC1	Yes
KRT19	Yes
LAMC1	Yes
LTBP1	Yes
MPST	Yes
NBN	Yes
NOSIP	Yes
PGR	Yes
RAB31	Yes
RAB5B	Yes
RARA	Yes
RASGRF1	Yes
SETD7	Yes
SLC2A1	Yes
SLC47A1	Yes
SLC7A5	Yes
SLCO2A1	Yes
SMOC2	Yes

SPINT2	Yes
TACC1	Yes
TFF1	Yes
TST	Yes
ZFHX3	Yes
ZNF185	Yes

Supplementary Table S14. Potential target genes of ERα peaks enriched in tamoxifen users as identified by combined analysis of gene expression and ChIP-seq. Genes in the list contain ERα peak 20kb upstream of TSS or in the gene body and strongly contribute to the enrichment among downregulated genes in gene set enrichment analysis (GSEA).

Gene Symbol	Core Enrichment
AKAP7	Yes
AQP9	Yes
AUNIP	Yes
C10orf11	Yes
CADPS2	Yes
CP	Yes
CRYL1	Yes
CXCL2	Yes
DLGAP5	Yes
ELF3	Yes
ESR1	Yes
EYA2	Yes
FAM107B	Yes
FAM129A	Yes
FGF18	Yes
GCNT2	Yes
GRAMD3	Yes
HGD	Yes
HIGD1A	Yes
HOMER2	Yes
MGP	Yes
PDE4DIP	Yes
PIGR	Yes
PLLP	Yes
PPP1R1B	Yes
RBM47	Yes
RPS6KA5	Yes
TMEM101	Yes
TMOD1	Yes
UGDH	Yes
VAV3	Yes
WWTR1	Yes