Published OnlineFirst May 22, 2018; DOI: 10.1158/1055-9965.EPI-17-1174

Research Article

Cancer Epidemiology, Biomarkers & Prevention

Genomic-Epidemiologic Evidence That Estrogens Promote Breast Cancer Development

Fritz F. Parl[1], Philip S. Crooke[2], W. Dale Plummer Jr[3], and William D. Dupont[3]

Abstract

Background: Estrogens are a prime risk factor for breast cancer, yet their causal relation to tumor formation remains uncertain. A recent study of 560 breast cancers identified 82 genes with 916 point mutations as drivers in the genesis of this malignancy. Because estrogens play a major role in breast cancer development and are also known to regulate the expression of numerous genes, we hypothesize that the 82 driver genes are likely to be influenced by estrogens, such as 17ß-estradiol (E2), and the estrogen receptor ESR1 (ERα). Because different types of tumors are characterized by unique sets of cancer driver genes, we also argue that the fraction of driver genes regulated by E2-ESR1 is lower in malignancies not associated with estrogens, e.g., acute myeloid leukemia (AML).

Methods: We performed a literature search of each driver gene to determine its E2-ESR1 regulation.

Results: Fifty-three of the 82 driver genes (64.6%) identified in breast cancers showed evidence of E2-ESR1 regulation. In contrast, only 19 of 54 mutated driver genes (35.2%) identified in AML were linked to E2-ESR1. Among the 916 driver mutations found in breast cancers, 813 (88.8%) were linked to E2-ESR1 compared with 2,046 of 3,833 in AML (53.4%).

Conclusions: Risk assessment revealed that mutations in estrogen-regulated genes are much more likely to be associated with elevated breast cancer risk, while mutations in unregulated genes are more likely to be associated with AML.

Impact: These results increase the plausibility that estrogens promote breast cancer development. *Cancer Epidemiol Biomarkers Prev; 27(8); 899–907. ©2018 AACR.*

Introduction

There is a gap in our understanding of breast cancer between the numerous epidemiologic studies, which have firmly established estrogens as prime breast cancer risk factors, and the molecular studies in animal or cellular models, which have defined in great detail the mechanisms underlying the action of estrogens on individual gene expression (1). What connects the molecular mechanisms to the clinical disease? All cancers carry somatic mutations in their genomes. An in-depth analysis of 560 breast cancers led by the Wellcome Trust Sanger Institute in Cambridge, UK, has produced the most comprehensive portrait to date of the somatic mutations involved in the disease (2). The analysis identified 93 mutated cancer genes as drivers in the genesis of breast cancer. The 93 protein-coding genes contained 1,661 probable driver mutations in the form of point mutations, copy number aberrations, and genomic rearrangements. The most common abnormality were 916 point mutations, which occurred in 82 driver genes as base substitutions. These variants resulted in missense, nonsense, splice site, start–stop mutations, and indel mutations (i.e., small insertions and deletions). Because estrogens play a major role in breast cancer development and are also known to regulate the expression of numerous genes, we hypothesize that the driver genes are likely to be influenced by estrogens, such as 17ß-estradiol (E2), and the estrogen receptor ESR1 (ERα).

Because different types of tumors are characterized by unique sets of cancer driver genes, we also argue that the fraction of driver genes linked to E2-ESR1 should be lower in malignancies not associated with estrogens, e.g., acute myeloid leukemia (AML). A recent in-depth analysis of 1,540 patients with AML, also led by the Wellcome Trust Sanger Institute, has produced the most comprehensive portrait to date of the somatic mutations involved in the disease (3). The analysis identified 5,234 driver mutations across 76 genes or genomic regions. The driver mutations included chromosomal aneuploidies, fusion genes, complex karyotypes, and 54 distinct gene mutations composed of base substitutions and small (<200-bp) insertions or deletions. The 54 individually named genes contained a total of 3,833 driver mutations.

E2 is one of the most potent mitogens in the body and has been implicated in a variety of cancers, such as colon, kidney, and lung cancer, albeit to a lesser extent than in breast cancer. We selected AML in order to compare breast cancer to a malignancy without apparent relation to E2-ESR1. The E2-ESR1 action is mediated via several mechanisms through interaction with estrogen response elements (ERE) to induce or repress the expression of numerous genes (4, 5). These include involvement of coactivators/corepressors and cross-talk with other transcription factors. The classic pathway involves both binding of E2 to ESR1 and ESR1 interaction with the ERE resulting in hormone-induced gene transcription. Subsequent work revealed that neither hormone binding nor receptor–DNA contact is essential and that alternate pathways exist. Thus, ESR1 can interact synergistically with another transcription factor (e.g., AP-1 and Sp1) that is bound to its respective response element. E2 stimulates gene transcription indirectly

[1]Department of Pathology, Microbiology and Immunology, Vanderbilt University, Nashville, Tennessee. [2]Department of Mathematics, Vanderbilt University, Nashville, Tennessee. [3]Department of Health Policy, Vanderbilt University, Nashville, Tennessee.

Corresponding Author: Fritz F. Parl, Vanderbilt University, The Vanderbilt Clinic, Room 4918, Vanderbilt University School of Medicine, Nashville, TN 37232. Phone: 615-482-2446; Fax: 615-343-9563; E-mail: fritz.parl@vanderbilt.edu

doi: 10.1158/1055-9965.EPI-17-1174

Published OnlineFirst May 22, 2018; DOI: 10.1158/1055-9965.EPI-17-1174

Table 1. Relationship of E2-ESR1 to breast cancer driver genes with point mutations

		Point mutations				
Genes	**Chromo**	**Substitution**	**Insertion/deletion**	**Total**	**E2-ESR1 regulation**	**References**
TP53	17	170	52	222	Cross-talk	(18)
PIK3CA	3	164	8	172	Cross-talk	(19)
GATA3	10	4	50	54	GATA3 is ESR1 coactivator	(20)
MAP3K1	5	6	43	49	MAP3K1 activates ESR1	(21)
MLL3	7	14	22	36	MLL3 is ESR1 coactivator	(22)
CDH1	16	15	19	34	E2-ESR1 downregulates CDH1	(23)
PTEN	10	15	18	33	E2-ESR1 downregulates PTEN	(24)
RB1	13	12	9	21	RB1 controls ESR1 stability/activity	(25)
ARID1A	1	5	12	17	—	
CBFB	16	11	4	15	RUNX1-CBFB is ESR1 corepressor	(26)
AKT1	14	14	0	14	AKT1 is ESR1 coactivator	(27)
SF3B1	2	13	0	13	—	
SPEN	1	6	7	13	SPEN1 is ESR1 corepressor	(28)
TBX3	12	3	10	13	E2-ESR1 upregulates TBX3	(29)
NCOR1	17	6	6	12	E2-ESR1 downregulates NCOR1	(30)
BRCA2	13	6	4	10	E2-ESR1 upregulates BRCA2	(5)
PIK3R1	5	1	9	10	E2-ESR1 upregulates PIK3R1	(31)
NF1	17	6	3	9	—	
MLLT4	6	2	6	8	—	
ERBB2	17	6	1	7	E2-ESR1 downregulates ERBB2	(4)
FOXA1	14	7	0	7	E2-ESR1 downregulates FOXA1	(5)
MAP2K4	17	4	3	7	E2-ESR1 upregulates MAP2K4	(32)
MED23	6	3	4	7	—	
NOTCH2	1	5	1	6	E2 downregulates NOTCH2	(33)
BRCA1	17	3	2	5	E2-ESR1 upregulates BRCA1	(5)
CTCF	16	5	0	5	Cross-talk	(34)
KDM6A	X	3	2	5	—	
MLL2	12	2	3	5	MLL2 is ESR1 coactivator	(35)
SETD2	3	3	2	5	—	
ARID1B	6	2	2	4	—	
CREBBP	16	2	2	4	CREBBP is ESR1 coactivator	(36)
FOXP1	3	3	1	4	E2-ESR1 upregulates FOXP1	(37)
RUNX1	21	1	3	4	RUNX1 is ESR1 corepressor	(38)
USP9X	X	0	4	4	USP9X binds ESR1 for cofactor recruitment	(39)
XBP1	22	0	4	4	E2-ESR1 upregulates XBP1	(4)
ATR	3	1	2	3	E2 blocks ATR via membrane ESR1	(40)
CDKN2A	9	3	0	3	—	
ERBB3	12	3	0	3	E2-ESR1 downregulates ERBB2	(5)
ESR1	6	3	0	3	E2-ESR1 downregulates ESR1	(4)
FBXW7	4	3	0	3	—	
PALB2	16	1	2	3	—	
RHOA	3	3	0	3	E2-ESR1 upregulates RHOA	(32)
ATM	11	2	0	2	ESR1 downregulates ATM through miRNAs	(7)
ATRX	X	2	0	2	—	
AXIN1	16	1	1	2	E2-ESR1 downregulates Axin1	(41)
BCOR	X	2	0	2	—	
CDKN1B	12	2	0	2	—	
CUX1	7	2	0	2	—	
DNMT3A	2	1	1	2	—	
EGFR	7	2	0	2	Cross-talk	(42)
GNAS	20	2	0	2	—	
MLH1	3	1	1	2	E2-ESR1 downregulates MLH1	(43)
MSH2	2	2	0	2	E2-ESR1 upregulates MSH2	(4)
NOTCH1	9	0	2	2	Cross-talk	(44)
PHF6	X	1	1	2	—	
SMAD4	18	1	1	2	SMAD4 is ESR1 corepressor	(45)
SMARCA4	19	1	1	2	SMARCA4 (BRG1) is ESR1 coactivator	(46)
STAG2	X	1	1	2	—	
ZFP36L1	14	0	2	2	E2-ESR1 downregulates ZFP36L1	(4)
AKT2	19	1	0	1	Cross-talk	(47)
APC	5	0	1	1	E2-ESR1 upregulates APC	(48)
ASXL1	20	0	1	1	ESR1-ASXL1 binding	(49)
BRAF	7	1	0	1	—	
BUB1B	15	0	1	1	E2-ESR1 upregulates BUB1B	(5)
CASP8	2	1	0	1	—	
CBLB	3	0	1	1	E2-ESR1 downregulates CBLB	(5)

(Continued on the following page)

Published OnlineFirst May 22, 2018; DOI: 10.1158/1055-9965.EPI-17-1174

Table 1. Relationship of E2-ESR1 to breast cancer driver genes with point mutations (Cont'd)

		Point mutations				
Genes	**Chromo**	**Substitution**	**Insertion/deletion**	**Total**	**E2-ESR1 regulation**	**References**
CIC	19	1	0	1	—	
CNOT3	19	1	0	1	CNOT3 is ESR1 corepressor	(50)
ECT2L	6	1	0	1	—	
ERCC4	16	1	0	1	—	
FGFR2	10	1	0	1	E2-ESR1 upregulates FGFR2	(51)
HRAS	11	1	0	1	HRAS-EGFR estrogenic effect	(52)
KRAS	12	1	0	1	E2-ESR1 downregulates KRAS	(53)
MEN1	11	0	1	1	MEN1 is ESR1 coactivator	(54)
NF2	22	1	0	1	—	
NRAS	1	1	0	1	E2-ESR1 upregulates NRAS	(4)
PBRM1	3	0	1	1	—	
PMS2	7	1	0	1	—	
PRDM1	6	1	0	1	PRDM1 (BLIMP1) is ESR1 corepressor	(55)
PREX2	8	1	0	1	—	
STK11	19	1	0	1	STK11 (LKB1) is ESR1 coactivator	(56)
TET2	4	1	0	1	—	
		579	337	916		

through protein–protein contact (tethering) of ESR1 with the adjacent transcription factor. This mechanism may explain E2-induced expression of certain genes without identifiable EREs and E2-induced breast cell growth without ESR1-DNA binding (6). In addition, E2 and ESR1 can act via components of the cell membrane (7). For example, E2 can be recognized by G-protein–coupled receptor 30 (GPR30), an integral membrane protein. In response to estrogen, GPR30 initiates nongenomic functions such as stimulation of adenylyl cyclase, synthesis of nuclear phosphatidylinositol (3,4,5)-trisphosphate (PIP3), and activation of the mitogen-activated protein kinase (MAPK) and phosphoinositide 3-kinase (PI3K) signaling pathways. Finally, ESR1 can activate miRNAs to regulate gene expression (8). As complex as the action of estrogens is in the E2-ESR1 system, another layer of complexity is added by the action of nonsteroidal ligands. For example, E2-independent stimulation of ESR1-mediated gene transcription can be achieved by signals from tyrosine kinase-linked cell surface receptors initiated by peptide growth factors (e.g., insulin-like growth factor, IGF-I) and protein kinase activators, which, in turn, can activate ESR1 through phosphorylation. The addition of E2 acts synergistically, providing an example of "cross-talk" between the signal transduction and steroid receptor pathways (9). In the context of this study, we use the term E2-ESR1 regulation to encompass all these processes.

In this study, we analyzed the driver genes with point mutations in breast cancer and AML and determined the fraction of driver genes linked to E2-ESR1. We found a significantly higher fraction of driver genes and mutations linked to E2-ESR1 in breast cancer. Risk assessment revealed that mutations in estrogen-regulated genes are much more likely to be associated with elevated breast cancer risk, while mutations in unregulated genes are more likely to be associated with AML. The data provide evidence that estrogens promote breast cancer development.

Materials and Methods

The comprehensive lists of driver genes published in the breast cancer and AML studies, respectively, form the basis of the present investigation (2, 3). To determine if there is a relationship between a driver gene and E2-ESR1, we performed a literature search of each gene using two key terms: the HGNC gene symbol (or alias) and estrogen (or estradiol). Only publications containing experimental studies of E2-ESR1 interactions with driver genes were accepted as evidence. Clinical correlation of the expression of a gene with the ESR1 status ($ER^{+/-}$) of breast cancers was considered supportive evidence when accompanied by experimental studies of the interaction mechanism. Clinical correlation in the absence of experimental studies was deemed insufficient evidence.

Data published online (2, 3) were reformatted with one row for each patient in either study. For each AML or breast cancer driver gene, we recorded the number of point mutations in each patient. This dataset also specified whether each patient had breast cancer or AML, and whether each of the 111 driver genes for these cancers was, or was not, regulated by E2-ESR1. Copy number aberrations and rearrangements were not considered in this paper because the involved genes were not recorded in the AML study. Patients with no point mutation in any of the driver genes for these cancers were deleted from the analyses. For each gene, the data were then dichotomized to indicate whether the patient did, or did not, have at least one point mutation. Maximum likelihood ratio χ^2 tests together with 95% confidence intervals (CI) were performed to assess the association between the patients' type of cancer and the E2-ESR1 regulation status of each gene (10). A similar approach was taken to assess whether any point mutation in any of the regulated or unregulated genes was associated with cancer type. Logistic regression was used to model the joint effects of regulated and unregulated genes using an additive model (11). Wald tests were used to compare the effects of mutations in regulated genes with those in unregulated genes. In secondary analyses, we regressed cancer type against the number of point mutations of each gene. These results were consistent with those based on the presence or absence of a point mutation. The significance of an association was indicated by $-\log_{10} P$ values, which are not adjusted for multiple comparisons. Because we assessed 111 driver genes, the Bonferroni cutoff for statistical significance was 4.5×10^{-4}. Mutational spectrum analyses were assessed by 2×6 χ^2 contingency table tests.

Results

Interaction of E2-ESR1 with breast cancer driver genes containing point mutations

Table 1 lists the 82 driver genes carrying 916 point mutations in form of substitutions and small insertions/deletions and summarizes the molecular mechanisms of E2-ESR1 regulation of these

Published OnlineFirst May 22, 2018; DOI: 10.1158/1055-9965.EPI-17-1174

genes. There is evidence of E2-ESR1 regulation for 53 of the 82 driver genes (64.6%) and the percentage is even higher when the total number of point mutations is considered, i.e., 813 of 916 (88.8%) point mutations show evidence of E2-ESR1 regulation.

Interaction of E2-ESR1 with AML driver genes

AML is characterized by 54 driver genes (Table 2; ref. 3). Nineteen of the 54 driver genes (35.2%) show evidence of E2-ESR1 regulation. The total number of point mutations observed in the AML study was 3,833, of which 2,046 (53.4%) could be linked to E2-ESR1. Twenty-nine of the 54 driver genes (53.7%) are unique to AML, whereas 25 (46.3%) are shared with breast cancer.

Effect of E2-ESR1 regulation on elevated risks of breast cancer and AML

Altogether, there are 111 driver genes with point mutations identified in both studies (2, 3). These mutations were observed in 1,403 AML patients and 478 breast cancer patients. Figure 1 shows

Table 2. Relationship of E2-ESR1 to AML driver genes

		Point mutations				
Genes	**Chromo**	**Substitution**	**Indel <200 bp**	**Total**	**E2-ESR1 regulation**	**References**
FLT3	13	222	339	561	E2-ESR1 upregulates FLT3	(57)
NPM1	5	0	432	432	E2-ESR1 upregulates NPM1	(58)
DNMT3A[a]	2	346	29	375	—	
NRAS	1	289	0	289	E2-ESR1 upregulates NRAS	(4)
TET2	4	121	81	202	—	
IDH2	15	150	0	150	—	
RUNX1	21	78	70	148	RUNX1 is ESR1 corepressor	(38)
CEBPA	19	27	103	130	—	
PTPN11 (SHP2)	12	129	0	129	E2-ESR1 upregulates SHP2	(59)
TP53	17	103	6	109	Cross-talk	(18)
IDH1	2	104	0	104	—	
SRSF2	17	69	22	91	—	
WT1	11	56	24	80	WT1 binds to ESR1	(60)
KRAS	12	79	0	79	E2-ESR1 downregulates KRAS	(53)
MLL (KMT2A)	11	1	76	77	—	
ASXL1	20	36	34	70	ESR1-ASXL1 binding	(49)
KIT	4	69	0	69	—	
STAG2	23	43	24	67	—	
RAD21	8	31	23	54	—	
EZH2	7	41	6	47	E2-ESR1 upregulates EZH2	(61)
PHF6	23	36	10	46	—	
CBL	11	42	0	42	—	
NF1	17	19	21	40	—	
SF3B1	2	39	0	39	—	
U2AF1	21	37	0	37	—	
GATA2	3	37	0	37	—	
BCOR	23	17	18	35	—	
MYC	8	28	0	28	E2-ESR1 upregulates MYC	(4)
EP300	22	19	4	23	EP300 is ESR1 coactivator	(62)
ETV6	12	12	10	22	—	
KDM5A	12	12	1	13	—	
MLL2 (KMT2D)	12	10	3	13	MLL2 is ESR1 coactivator	(35)
ZRSR2	23	8	5	13	—	
KDM6A	23	11	1	12	—	
MLL3 (KMT2C)	7	5	5	10	MLL3 is ESR1 coactivator	(22)
JAK2	9	10	0	10	E2-ESR1 upregulates JAK2	(63)
FBXW7	4	9	0	9	—	
BRAF	7	9	0	9	—	
CREBBP	16	9	0	9	CREBBP is ESR1 coactivator	(36)
ATRX	23	6	0	6	—	
CUX1	7	3	2	5	—	
RB1	13	5	0	5	RB1 controls ESR1 activity	(25)
MPL	1	3	0	3	—	
MLL5 (KMT2E)	7	2	1	3	—	
PTEN	10	3	0	3	E2-ESR1 downregulates PTEN	(24)
PRPF40B	12	3	0	3	—	
CDKN2A	9	2	0	2	—	
SF1	11	2	0	2	—	
U2AF2	19	2	0	2	—	
GNAS	20	2	0	2	—	
CBLB	3	1	0	1	E2-ESR1 downregulates CBLB	(5)
IKZF1	7	1	0	1	—	
SH2B3	12	1	0	1	—	
SF3A1	22	1	0	1	—	
		2400	1350	3833		

[a]Driver genes listed in bold are shared with breast cancer.

Published OnlineFirst May 22, 2018; DOI: 10.1158/1055-9965.EPI-17-1174

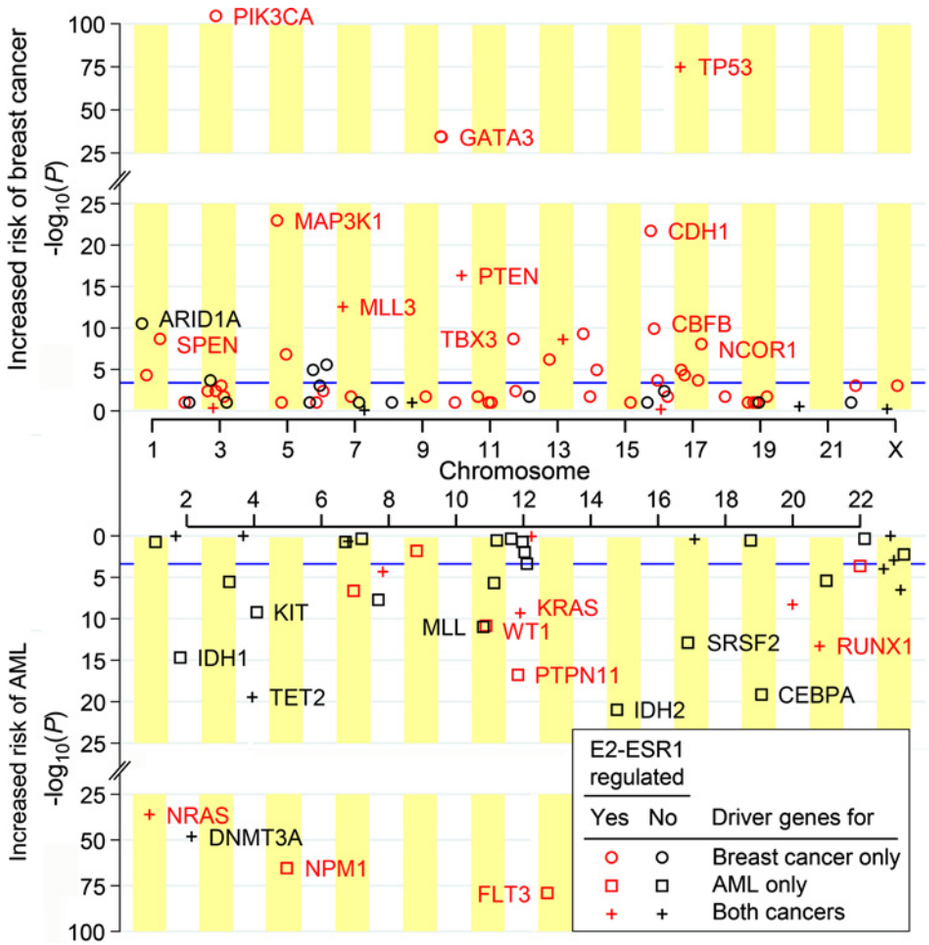

Figure 1.
Manhattan plot of *P* values for the association between breast cancer or AML and point mutations in driver genes for these two cancers. In this cross-sectional study, all subjects had either breast cancer or AML. The *P* values on the top indicate an increased risk of breast cancer given that the patient has one or the other of these cancers. *P* values on the bottom are similarly defined but indicate an increased risk of AML. Each *P* value is color coded by whether or not the associated gene is E2-ESR1 regulated, and the marker shapes indicate whether the associated driver gene is for breast cancer only, AML only, or for both cancers. The Bonferroni-corrected significance level of 4.5×10^{-4} is indicated by blue dashed lines.

a Manhattan plot of increased risks of breast cancer or AML among patients, all of whom have either one or the other of these cancers. For each driver gene there are –log10 *P* values indicating the association between having at least one point mutation and breast cancer or AML. Each *P* value is color coded by whether or not the associated gene is E2-ESR1 regulated, and the marker shapes indicate whether the associated driver gene is for breast cancer only, AML only, or for both cancers. There were 111 driver genes for these cancers that had at least one point mutation. The Bonferroni-corrected 0.05 significance level is thus $0.05/111 = 4.5 \times 10^{-4}$. The blue lines in Fig. 1 mark this threshold of significance. Twenty-one of 25 mutations (84.0%) that were associated with elevated breast cancer risk at this level of significance were E2-ESR1 regulated. In contrast, only 11 of 26 mutations (42.3%) that were associated with elevated AML risk at this level of significance were E2-ESR1 regulated. Thus, while there are mutations in unregulated genes that are associated with elevated breast cancer risk, and mutations in regulated genes that are associated with elevated AML risk, mutations in regulated genes are much more likely to be associated with breast cancer, while mutations in unregulated genes are more likely to be associated with AML. Table 3 gives odds ratios for breast cancer among study subjects. The odds of breast cancer in cancer patients with at least one mutation in a regulated gene are 11.4 times that of patients with no mutations in these genes ($P = 2.9 \times 10^{-22}$; 95% CI, 6.0–25). In contrast, the odds of breast cancer in a cancer patient with at least one mutation in an unregulated gene were 0.191 times that of patients with no mutations in these genes ($P = 3.4 \times 10^{-51}$; 95% CI, 0.15–0.24). The breast cancer odds among patients with any mutation in a regulated gene were 22.4 times that of patients with at least one mutation in an unregulated gene ($P = 2.4 \times 10^{-17}$; 95% CI, 11–46).

Mutational spectrum analysis

Mutagens that generate somatic mutations imprint particular patterns of mutations on cancer genomes. We analyzed the mutational spectra of substitutions in the driver genes of breast cancer and AML patients in an attempt to discern any potential effect of estrogens in the mutational process. We assigned each mutation to one of the six classes of base substitution, conventionally expressed as the pyrimidine of a mutated Watson–Crick base pair: C>A, C>G, C>T, T>A, T>C, and T>G (Fig. 2). Overall, there was a preponderance of C>T mutations in both breast cancer and AML, in agreement with genomic studies showing C>T substitutions as predominant lesion in multiple types of malignancies (12). A more detailed analysis of our data revealed that the mutational spectra of E2-ESR1–regulated and unregulated driver genes in breast cancer were similar (Fig. 2A), consistent with the exposure of all driver genes to the same mutagens. In contrast, the mutational spectra of driver genes shared by breast cancer and

Published OnlineFirst May 22, 2018; DOI: 10.1158/1055-9965.EPI-17-1174

Table 3. Odds ratios for breast cancer among cancer patients who either have AML or breast cancer by the presence or absence of point mutations in driver genes that either are, or are not, regulated by E2-ESR1

	AML Number (%)	Breast cancer Number (%)	OR	*P*	95% CI
No mutations in regulated genes[a]	228 (16.3)	8 (1.7)	1.0[c]		
At least one mutation in regulated genes[a]	1,175 (83.7)	470 (98.3)	11.4	2.9×10^{-22}	6.0–25
Total	1,403 (100)	478 (100)			
No mutations in unregulated genes[b]	419 (29.9)	330 (69.0)	1.0[c]		
At least one mutation in unregulated genes[b]	984 (70.1)	148 (31.0)	0.191	3.4×10^{-51}	0.15–0.24
Total	1,403 (100)	478 (100)			
At least one mutation in unregulated genes[b]	984	148	1.0[c]		
At least one mutation in regulated genes[a]	1,175	470	22.4	2.4×10^{-17}	11–46

[a]Genes regulated by E2-ESR1.
[b]Genes not regulated by E2-ESR1.
[c]Denominator of following odds ratio.

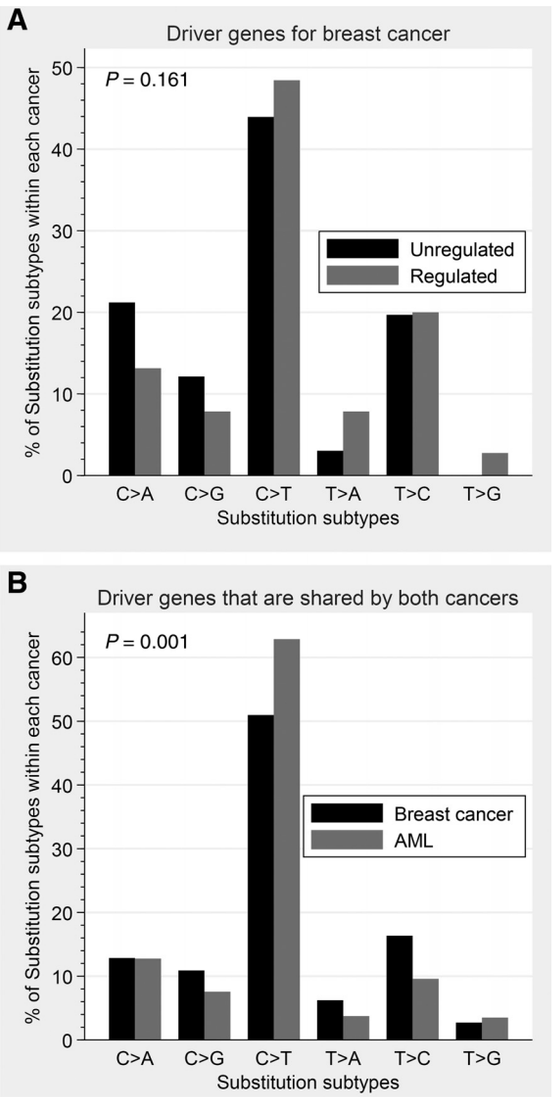

Figure 2.
Bar graphs showing mutational spectra of driver gene substitutions in (**A**) E2-ESR1-regulated and unregulated breast cancer driver genes and (**B**) shared driver genes of breast cancer and AML patients.

AML differed significantly (Fig. 2B), reflecting different mutagens in the two malignancies.

Discussion

Life events and exposures related to estrogens are prime risk factors for breast cancer, yet the causal relation between estrogens and tumor formation in women remains uncertain. We reasoned that the effect of estrogens in tumor formation might leave an imprint in the molecular changes found in breast cancer. The complex landscapes of somatic mutations observed in tumors are typically the result of a relatively small number of functional oncogenic alterations (called driver events), which are outnumbered by nonfunctional alterations (passenger events). The driver mutations confer clonal selective advantage on cancer cells and are causally implicated in oncogenesis, whereas the passenger mutations do not substantially contribute to oncogenesis and progression. The analysis of 560 breast cancers led by the Wellcome Trust Sanger Institute in Cambridge, UK, has produced the most comprehensive portrait to date of the somatic mutations involved in the disease (2). The analysis identified 93 mutated cancer genes as drivers in the genesis of breast cancer, of which 82 carried point mutations with evidence of E2-ESR1 regulation in 53 (64.6%; Table 1). In contrast, only 19 of the 54 (35.2%) driver genes with point mutations in AML patients showed evidence of E2-ESR1 regulation (Table 2).

The genetic alterations in malignant tumors consist of three main categories: point mutations, copy number aberrations, and rearrangements, all of which are found in breast cancer and AML. Whereas point mutations occur in specific genes, copy number aberrations and rearrangements frequently extend beyond individual genes to chromosomal regions. While the breast cancer study (2) identified driver genes affected by point mutations, copy number aberrations, and rearrangements, the AML study (3) only reported individual driver genes afflicted by point mutations. Because we are examining the effect of estrogens on individual genes, the study is limited to point mutations in both types of malignancy. It is important to realize that this is not a conventional case–control study. Rather, it is a concurrent study of patients who all have either breast cancer or AML and whose somatic cancer driver mutations are known. This greatly affects how our results should be interpreted. For example, in Fig. 1, the *P* value associated with point mutations in the PIK3CA gene is 3.0×10^{-105}, and the 95% CI for its breast cancer odds ratio is 353 – ∞. What this means is that, among patients who have either

breast cancer or AML, the odds of them having breast cancer if they also have a point driver mutation in PIK3CA is extremely high. It does not mean that women with one of these mutations have a risk of breast cancer relative to the general population that is anywhere near this magnitude. The fact that we are studying known driver mutations for breast cancer or AML in patients who have one or the other of these cancers means that the magnitude of the *P* values reported in this paper is not surprising. What is new and important is the extent to which driver mutations that are associated with elevated breast cancer risk are found in genes that are regulated by E2-ESR1. Table 3 shows that mutations in E2-ESR1–regulated genes are much more likely to be associated with elevated breast cancer risk, while mutations in unregulated genes are more likely to be associated with AML. This increases the evidence that the well-known epidemiologic association between estrogens and breast cancer is causally mediated via E2-ESR1–regulated driver genes.

Tumor development or carcinogenesis is usually viewed as a stepwise process beginning with genotoxic effects (initiation) followed by enhanced cell proliferation (promotion; ref. 13). Our analysis of the mutational spectrum of substitutions in breast cancer driver genes shows similar spectra for E2-ESR1–regulated and unregulated driver genes (Fig. 2A), indicating that the mutations in both classes of genes have the same etiology among breast cancer patients. The nature of the initiating mutagens remains uncertain; candidates include catechol estrogens, which are oxidative metabolites of estrogens with known mutagenic activity (14, 15). While our findings do not add to our understanding of tumor initiation, the results shown in Fig. 1 demonstrate that the carcinogenic action of estrogens in breast cancer is mediated by promotion of cell proliferation via estrogen-regulated driver genes. A core group of genes such as DNMT3A, PTEN, RB1, RUNX1, and TP53 is mutated regularly in many tumors, evidence for their key roles in the malignant process. In the present study, 25 of the driver genes were mutated in both breast cancer and AML. Although being common to both diseases, the shared genes displayed significantly different mutational spectra (Fig. 2B), consistent with the notion that breast cancer and AML are caused by different, disease-specific mutagens.

Most malignant tumors carry multiple driver and passenger mutations, e.g., an average of 32 were observed in breast cancer and 26 in AML (16). Even if we focus only on driver genes and use the present studies as examples, each breast cancer contains approximately 3 (1,661 mutations in 560 tumors) and each AML an estimated 3.4 driver mutations (5,234 mutations in 1,540 leukemias). We are simplifying the analysis by focusing on single driver gene mutations rather than the combined impact of three mutations in each tumor. We can only speculate how mutations in three genes would interact, but given the apparent influence of E2-ESR1 on single mutational driver events, estrogens are likely to exert a compounding effect. Breast cancer is a heterogeneous disease. A recent analysis of nearly 2,000 breast cancers defined 10 subtypes by correlating the genomic and transcriptomic data with long-term clinical follow-up (17). We viewed breast cancer as a single entity while it is likely that any E2-ESR1 influence would vary among the 10 subtypes. Future genomic-epidemiologic studies with sufficiently large numbers of cancers will be required to determine the extent of estrogen carcinogenicity in individual subtypes.

In summary, we observed a significantly higher fraction of driver genes and point mutations linked to E2-ESR1 in breast cancer than in AML. Risk assessment revealed that mutations in E2-ESR1–regulated genes are much more likely to be associated with elevated breast cancer risk, while mutations in unregulated genes are more likely to be associated with AML. Together, these results increase the plausibility that estrogens promote breast cancer development.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: F.F. Parl, P.S. Crooke, W.D. Dupont
Development of methodology: F.F. Parl, P.S. Crooke
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): W.D. Plummer Jr, W.D. Dupont
Writing, review, and/or revision of the manuscript: F.F. Parl, P.S. Crooke, W.D. Dupont
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): F.F. Parl, P.S. Crooke, W.D. Plummer Jr
Study supervision: F.F. Parl, P.S. Crooke

Acknowledgments

W.D. Dupont was supported by NCI grants R01 CA050468 and P30 CA068485 and NCATS7NIH grant UL1 TR000445.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received December 14, 2017; revised February 23, 2018; accepted May 14, 2018; published first May 22, 2018.

References

1. Parl FF. The etiology of breast cancer: endogenous and exogenous causes. 2014. Amazon ISBN 978-0615993737; DOI 10.13140/2.1.2321.0568.
2. Nik-Zainal S, Davies H, Staaf J, Ramakrishna M, Glodzik D, Zou X, et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. Nature 2016;534:47–54.
3. Papaemmanuil E, Gerstung M, Bullinger L, Gaidzik VI, Paschka P, Roberts ND, et al. Genomic classification and prognosis in acute myeloid leukemia. N Engl J Med 2016;374:2209–21.
4. Carroll JS, Meyer CA, Song J, Li W, Geistlinger TR, Eeckhoute J, et al. Genome-wide analysis of estrogen receptor binding sites. Nat Genet 2006; 38:1289–97.
5. Frasor J, Danes JM, Komm B, Chang KC, Lyttle CR, Katzenellenbogen BS. Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype. Endocrinology 2003;144:4562–74.
6. DeNardo DG, Cuba VL, Kim H, Wu K, Lee AV, Brown PH. Estrogen receptor DNA binding is not required for estrogen-induced breast cell growth. Mol Cell Endocrinol 2007;277:13–25.
7. Pietras RJ, Marquez-Garban DC. Membrane-associated estrogen receptor signaling pathways in human cancers. Clin Cancer Res 2007;13:4672–6.
8. Guo X, Yang C, Qian X, Lei T, Li Y, Shen H, et al. Estrogen receptor alpha regulates ATM Expression through miRNAs in breast cancer. Clin Cancer Res 2013;19:4994–5002.
9. Levin ER. Bidirectional signaling between the estrogen receptor and the epidermal growth factor receptor. Mol Endocrinol 2003;17:309–17.
10. Choi L, Blume JD, Dupont WD. Elucidating the foundations of statistical inference with 2 x 2 tables. PLoS One 2015;10:e0121263.

11. Dupont WD. Statistical modeling for biomedical researchers: a simple introduction to the analysis of complex data. Cambridge: Cambridge University Press; 2009.
12. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, et al. Signatures of mutational processes in human cancer. Nature 2013; 500:415–421.
13. Moolgavkar SH, Day NE, Stevens RG. Two-stage model for carcinogenesis: epidemiology of breast cancer in females. J Natl Cancer Inst 1980;65: 559–69.
14. Yager JD, Davidson NE. Estrogen carcinogenesis in breast cancer. New Engl J Med 2006;354:270–82.
15. Parl FF, Dawling S, Roodi N, Crooke PS. Estrogen metabolism and breast cancer: a risk model. Ann N Y Acad Sci 2009;1155:68–75.
16. Lawrence MS, Stojanov P, Mermel CH, Robinson JT, Garraway LA, Golub TR, et al. Discovery and saturation analysis of cancer genes across 21 tumour types. Nature 2014;505:495–501.
17. Curtis C, Shah SP, Chin SF, Turashvili G, Rueda OM, Dunning MJ, et al. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. Nature 2012;486:346–52.
18. Berger C, Qian Y, Chen X. The p53-estrogen receptor loop in cancer. Curr Mol Med 2013;13:1229–40.
19. Bosch A, Li Z, Bergamaschi A, Ellis H, Toska E, Prat A, et al. PI3K inhibition results in enhanced estrogen receptor function and dependence in hormone receptor-positive breast cancer. Sci Transl Med 2015;7:283ra51.
20. Theodorou V, Stark R, Menon S, Carroll JS. GATA3 acts upstream of FOXA1 in mediating ESR1 binding by shaping enhancer accessibility. Genome Res 2013;23:12–22.
21. Pham TT, Angus SP, Johnson GL. MAP3K1: genomic alterations in cancer and function in promoting cell survival or apoptosis. Genes Cancer 2013;4:419–26.
22. Ansari KI, Shrestha B, Hussain I, Kasiri S, Mandal SS. Histone methylases MLL1 and MLL3 coordinate with estrogen receptors in estrogen-mediated HOXB9 expression. Biochemistry 2011;50:3517–27.
23. Cardamone MD, Bardella C, Gutierrez A, Di Croce L, Rosenfeld MG, Di Renzo MF, et al. ERalpha as ligand-independent activator of CDH-1 regulates determination and maintenance of epithelial morphology in breast cancer cells. Proc Natl Acad Sci USA 2009;106:7420–5.
24. Noh EM, Lee YR, Chay KO, Chung EY, Jung SH, Kim JS, et al. Estrogen receptor alpha induces down-regulation of PTEN through PI3-kinase activation in breast cancer cells. Mol Med Rep 2011;4:215–9.
25. Caligiuri I, Toffoli G, Giordano A, Rizzolio F. pRb controls estrogen receptor alpha protein stability and activity. Oncotarget 2013;4:875–83.
26. van Bragt MP, Hu X, Xie Y, Li Z. RUNX1, a transcription factor mutated in breast cancer, controls the fate of ER-positive mammary luminal cells. Elife 2014;3:e03881.
27. Bhat-Nakshatri P, Wang G, Appaiah H, Luktuke N, Carroll JS, Geistlinger TR, et al. AKT alters genome-wide estrogen receptor alpha binding and impacts estrogen signaling in breast cancer. Mol Cell Biol 2008;28: 7487–503.
28. Legare S, Cavallone L, Mamo A, Chabot C, Sirois I, Magliocco A, et al. The estrogen receptor cofactor SPEN functions as a tumor suppressor and candidate biomarker of drug responsiveness in hormone-dependent breast cancers. Cancer Res 2015;75:4351–63.
29. Fillmore CM, Gupta PB, Rudnick JA, Caballero S, Keller PJ, Lander ES, et al. Estrogen expands breast cancer stem-like cells through paracrine FGF/Tbx3 signaling. Proc Natl Acad Sci USA 2010;107:21737–42.
30. Frasor J, Danes JM, Funk CC, Katzenellenbogen BS. Estrogen down-regulation of the corepressor N-CoR: mechanism and implications for estrogen derepression of N-CoR-regulated genes. Proc Natl Acad Sci USA 2005;102:13153–7.
31. Harvell DM, Richer JK, Allred DC, Sartorius CA, Horwitz KB. Estradiol regulates different genes in human breast tumor xenografts compared with the identical cells in culture. Endocrinology 2006;147:700–13.
32. Zheng S, Huang J, Zhou K, Zhang C, Xiang Q, Tan Z, et al. 17beta-Estradiol enhances breast cancer cell motility and invasion via extra-nuclear activation of actin-binding protein ezrin. PLoS One 2011;6:e22439.
33. Raafat A, Goldhar AS, Klauzinska M, Xu K, Amirjazil I, McCurdy D, et al. Expression of Notch receptors, ligands, and target genes during development of the mouse mammary gland. J Cell Physiol 2011;226:1940–52.
34. Fiorito E, Sharma Y, Gilfillan S, Wang S, Singh SK, Satheesh SV, et al. CTCF modulates Estrogen Receptor function through specific chromatin and nuclear matrix interactions. Nucleic Acids Res 2016;44:10588–602.
35. Mo R, Rao SM, Zhu YJ. Identification of the MLL2 complex as a coactivator for estrogen receptor alpha. J Biol Chem 2006;281:15714–20.
36. Jaber BM, Mukopadhyay R, Smith CL. Estrogen receptor-alpha interaction with the CREB binding protein coactivator is regulated by the cellular environment. J Mol Endocrinol 2004;32:307–23.
37. Shigekawa T, Ijichi N, Ikeda K, Horie-Inoue K, Shimizu C, Saji S, et al. FOXP1, an estrogen-inducible transcription factor, modulates cell proliferation in breast cancer cells and 5-year recurrence-free survival of patients with tamoxifen-treated breast cancer. Horm Cancer 2011;2:286–97.
38. Chimge NO, Frenkel B. The RUNX family in breast cancer: relationships with estrogen signaling. Oncogene 2013;32:2121–30.
39. Oosterkamp HM, Hijmans EM, Brummelkamp TR, Canisius S, Wessels LF, Zwart W, et al. USP9X downregulation renders breast cancer cells resistant to tamoxifen. Cancer Res 2014;74:3810–20.
40. Pedram A, Razandi M, Evinger AJ, Lee E, Levin ER. Estrogen inhibits ATR signaling to cell cycle checkpoints and DNA repair. Mol Biol Cell 2009;20:3374–89.
41. Chimge NO, Little GH, Baniwal SK, Adisetiyo H, Xie Y, Zhang T, et al. RUNX1 prevents oestrogen-mediated AXIN1 suppression and beta-catenin activation in ER-positive breast cancer. Nat Commun 2016;7:10751.
42. Skandalis SS, Afratis N, Smirlaki G, Nikitovic D, Theocharis AD, Tzanakakis GN, et al. Cross-talk between estradiol receptor and EGFR/IGF-IR signaling pathways in estrogen-responsive breast cancers: focus on the role and impact of proteoglycans. Matrix Biol 2014;35:182–93.
43. Singh KP, Treas J, Tyagi T, Gao W. DNA demethylation by 5-aza-2-deoxycytidine treatment abrogates 17 beta-estradiol-induced cell growth and restores expression of DNA repair genes in human breast cancer cells. Cancer Lett 2012;316:62–9.
44. Hao L, Rizzo P, Osipo C, Pannuti A, Wyatt D, Cheung LW, et al. Notch-1 activates estrogen receptor-alpha-dependent transcription via IKKalpha in breast cancer cells. Oncogene 2010;29:201–13.
45. Wu L, Wu Y, Gathings B, Wan M, Li X, Grizzle W, et al. Smad4 as a transcription corepressor for estrogen receptor alpha. J Biol Chem 2003; 278:15192–200.
46. DiRenzo J, Shang Y, Phelan M, Sif S, Myers M, Kingston R, et al. BRG-1 is recruited to estrogen-responsive promoters and cooperates with factors involved in histone acetylation. Mol Cell Biol 2000;20:7541–9.
47. Sun M, Paciga JE, Feldman RI, Yuan Z, Coppola D, Lu YY, et al. Phosphatidylinositol-3-OH Kinase (PI3K)/AKT2, activated in breast cancer, regulates and is induced by estrogen receptor alpha (ERalpha) via interaction between ERalpha and PI3K. Cancer Res 2001;61:5985–91.
48. Prosperi JR, Becher KR, Willson TA, Collins MH, Witte DP, Goss KH. The APC tumor suppressor is required for epithelial integrity in the mouse mammary gland. J Cell Physiol 2009;220:319–31.
49. Katoh M. Functional proteomics of the epigenetic regulators ASXL1, ASXL2 and ASXL3: a convergence of proteomics and epigenetics for translational medicine. Expert Rev Proteomics 2015;12:317–28.
50. Winkler GS, Mulder KW, Bardwell VJ, Kalkhoven E, Timmers HT. Human Ccr4-Not complex is a ligand-dependent repressor of nuclear receptor-mediated transcription. EMBO J 2006;25:3089–99.
51. Fletcher MN, Castro MA, Wang X, de Santiago I, O'Reilly M, Chin SF, et al. Master regulators of FGFR2 signalling and breast cancer risk. Nat Commun 2013;4:2464.
52. Treeck O, Weber A, Boester M, Porz S, Frey N, Diedrich K, et al. H-ras dependent estrogenic effects of epidermal growth factor in the estrogen-independent breast cancer cell line MDA-MB-231. Breast Cancer Res Treat 2003;80:155–62.
53. Buterin T, Koch C, Naegeli H. Convergent transcriptional profiles induced by endogenous estrogen and distinct xenoestrogens in breast cancer cells. Carcinogenesis 2006;27:1567–78.
54. Dreijerink KM, Mulder KW, Winkler GS, Hoppener JW, Lips CJ, Timmers HT. Menin links estrogen receptor activation to histone H3K4 trimethylation. Cancer Res 2006;66:4929–35.
55. Wang X, Belguise K, O'Neill CF, Sanchez-Morgan N, Romagnoli M, Eddy SF, et al. RelB NF-kappaB represses estrogen receptor alpha expression via induction of the zinc finger protein Blimp1. Mol Cell Biol 2009;29: 3832–44.

56. Nath-Sain S, Marignani PA. LKB1 catalytic activity contributes to estrogen receptor alpha signaling. Mol Biol Cell 2009;20:2785–95.
57. Seillet C, Rouquie N, Foulon E, Douin-Echinard V, Krust A, Chambon P, et al. Estradiol promotes functional responses in inflammatory and steady-state dendritic cells through differential requirement for activation function-1 of estrogen receptor alpha. J Immunol 2013;190: 5459–70.
58. Zhou Y, Shen J, Xia L, Wang Y. Estrogen mediated expression of nucleophosmin 1 in human endometrial carcinoma clinical stages through estrogen receptor-alpha signaling. Cancer Cell Int 2014;14:540.
59. Li J, Kang Y, Wei L, Liu W, Tian Y, Chen B, et al. Tyrosine phosphatase Shp2 mediates the estrogen biological action in breast cancer via interaction with the estrogen extranuclear receptor. PLoS One 2014; 9:e102847.
60. Reizner N, Maor S, Sarfstein R, Abramovitch S, Welshons WV, Curran EM, et al. The WT1 Wilms' tumor suppressor gene product interacts with estrogen receptor-alpha and regulates IGF-I receptor gene transcription in breast cancer cells. J Mol Endocrinol 2005;35:135–44.
61. Bhan A, Hussain I, Ansari KI, Bobzean SA, Perrotti LI, Mandal SS. Histone methyltransferase EZH2 is transcriptionally induced by estradiol as well as estrogenic endocrine disruptors bisphenol-A and diethylstilbestrol. J Mol Biol 2014;426:3426–41.
62. Hanstein B, Eckner R, DiRenzo J, Halachm S, Liu H, Searcy B, et al. p300 is a component of an estrogen receptor coactivator complex. Proc Natl Acad Sci USA 1996;93:11540–5.
63. Coughlan N, Thillainadesan G, Andrews J, Isovic M, Torchia J. beta-Estradiol-dependent activation of the JAK/STAT pathway requires p/CIP and CARM1. Biochim Biophys Acta 2013;1833:1463–75.

Published OnlineFirst May 22, 2018; DOI: 10.1158/1055-9965.EPI-17-1174

Cancer Epidemiology, Biomarkers & Prevention

AACR American Association for Cancer Research

Genomic-Epidemiologic Evidence That Estrogens Promote Breast Cancer Development

Fritz F. Parl, Philip S. Crooke, W. Dale Plummer, Jr, et al.

Cancer Epidemiol Biomarkers Prev 2018;27:899-907. Published OnlineFirst May 22, 2018.