Published in final edited form as:
Horm Cancer. 2012 April ; 3(0): 37–43. doi:10.1007/s12672-011-0086-2.

Functional characterization of a genetic polymorphism in the promoter of the *ESR2* gene.

Santosh Philips[1,4], **Alexandra Richter**[2], **Steffi Oesterreich**[2], **James M Rae**[3], **David A Flockhart**[1], **Narayanan B Perumal**[4,5], and **Todd C Skaar**[1]

[1] Department of Medicine, Division of Clinical Pharmacology, Indiana University School of Medicine, Indianapolis, IN, USA.

[2]University of Pittsburgh Cancer Institute, Department of Pharmacology and Chemical Biology, Pittsburgh, PA, USA.

[3]Department of Internal Medicine, University of Michigan, Ann Arbor, MI,USA.

[4]Indiana University School of Informatics, Indianapolis, IN, USA.

[5]Eli Lilly and Company, Indianapolis, IN.

Abstract

The *ESR2* gene encodes the estrogen receptor beta protein. Several studies have shown that genetic variants in the *ESR2* gene are associated with a variety of clinical phenotypes. However, very little is known about the functional significance of *ESR2* genetic variants. We used a bioinformatics approach to identify regions of the *ESR2* promoter that is evolutionarily conserved across the genomes of several species. We resequenced 1.6 kb of the *ESR2* gene which included 0.8 kb of the promoter, 0.3 kb of exon ON and 0.5kb of the following intron. We identified 5 SNPs in the *ESR2* promoter and 1 SNP in the intron. Phase analysis indicated that the SNPs likely exist in 11 different haplotypes. Three of the SNPs (rs8008187, rs3829768, rs35036378) were predicted to alter transcription factor binding sites in the *ESR2* promoter. All three were detected only in African American subjects. The rs35036378 SNP was in the TATA box and was highly conserved across species. *ESR2* promoter reporter assays in LNCaP and SKBR3 cell lines showed that the variant construct containing the rs35036378 SNP allele had approximately 50% less activity relative to the wild type construct. We conclude that the rs35036378 SNP appears to cause a reduced promoter activity of the *ESR2* gene.

Keywords

ESR2; SNPs; TATA box; promoter polymorphism

INTRODUCTION

The estrogen receptor beta (*ESR2*) is one of the two nuclear receptors that mediate most of the actions of estrogen. *ESR2*, the gene that encodes for the ESR2 protein, has many polymorphisms that have been associated with a wide variety of phenotypes, such as hot flashes [1], serum lipids [2] endometrial cancer [3], cardiovascular diseases [4], Alzheimer's

Corresponding Author: Todd C Skaar, Ph.D., Department of Medicine, Division of Clinical Pharmacology, Indiana University School of Medicine, 1001 W. 10th Street, WD Myers Building W7123, Indianapolis, IN 46202, USA Tel: +1 317 630 2695; fax: + 1 317 630 8185; tskaar@iupui.edu.

Conflict of Interest: We declare no conflict of interest.

disease [5], and breast cancer [6]. Despite the many clinical association studies, it is still not clear which of the identified SNPs are functional. Although the effect size of the *ESR2* SNPs in clinical association studies has been relatively small, this does not necessarily mean that the effects of the *ESR2* genetic variations are small. One of the recent explanations for the relatively small effect size of many of the common SNPs observed in GWAS studies is the phenomenon of "synthetic associations" [7, 8]. This apparent small effect size is observed when multiple rare SNPs with large functional effects load disproportionately onto haplotypes that are imperfectly tagged by the more common SNPs. Since there do not appear to be any common variants in the *ESR2* gene that have obvious recognizable function, we hypothesized that there are multiple rare SNPs in the *ESR2* gene that have larger effects. Since the low minor allele frequency of the rare SNPs would require very large populations to test their associations individually, this is impractical for the phenotypes that we measure because they require intensive prospective phenotype measurements [1, 2]. Therefore, we have taken the approach of functionally characterizing the rare SNPs in the *ESR2* gene so they can be evaluated together in related clinical studies, using an approach similar to the scoring system for polymorphisms in the drug metabolizing enzyme genes [9].

To address this hypothesis, we have resequenced a section of the *ESR2* promoter and identified rare SNPs that are predicted to impact promoter function. We have also validated the bioinformatic prediction using *in vitro* functional promoter assays. Our studies have identified a SNP in the *ESR2* promoter TATA box that substantially reduces the promoter activity.

MATERIAL AND METHODS

DNA sequencing

We resequenced 1.6 kb of the *ESR2* gene. The target region included exon ON and the flanking regions (nucleotide 45760236 to 45761880 from accession # NT_026437). The DNA was amplified using three PCR reactions. The primers were designed using primer3 software and obtained from Integrated DNA Technologies, Inc (Coralville, Iowa, USA). The sequences of the primers are listed in the Supplementary Table 1. The target region was amplified in 50 African American DNA samples from the Coriell diversity panel and in 50 samples from the COBRA tamoxifen breast cancer clinical trial (ClinicalTrials.gov Identifier #NCT00228930). The subjects in this tamoxifen trial were mostly Caucasians. The PCR reaction consisted of 50 ng of genomic DNA, 200 nM of primers, 5% DMSO and the final volume was made up to 50 μl using Accuprime Pfx Supermix (Invitrogen Corporation, Carlsbad, CA). The DNA was amplified using GeneAmp PCR System 9700 (Applied Biosystems, Foster city, CA) using the following cycle: (a) an initial denaturation step of 95°C for 5 min; (b) 40 cycles of 95°C for 15 sec, 52°C (Amplicon 2 & 3) / 57°C (Amplicon 1) for 30 sec, 68°C for 1 min and (c) a final primer extension step of 68°C for 5 min. The size and purity of the amplicons were confirmed by agarose gel electrophoresis. The amplicons were purified using Microcon centrifugal filter devices (Millipore Corporation, Billerica, MA) and sequenced at the Biochemistry Biotechnology Facility (Indiana University, Indianapolis, IN) using the Big Dye® Terminator v3.1 cycle sequencing kit (Applied Biosystems, Inc, Foster city, CA) on the ABI 3100 Genetic Analyzer (Applied Biosystems, Foster city, CA). The ABI trace files from sequencing containing the sequences were assembled and analyzed using the VectorNTI software (Invitrogen Corporation, Carlsbad, CA). The chromatogram for each subject was visually scanned for the presence of SNPs. The haplotypes from the DNA sequencing data were estimated using the PHASE software [10, 11].

Bioinformatic analysis of the polymorphic regions of the *ESR2* gene

Two bioinformatic analyses of the polymorphic regions of the *ESR2* gene were conducted to prioritize the SNPs for functional laboratory studies. The first was a determination of species conservation of the regions containing the SNP. The human (NT_026437), mouse (NW_047761), rat (NT_039551) and chimpanzee (Chimp Nov. 2003 Assembly, UCSC) sequences were aligned and visualized using ClustalW [12] and VISTA [13]. The level of conservation of the predicted TATA box among the species was observed using ClustalW. It was further confirmed using the rVISTA tool[14]; the cut-offs for the similarity score and core similarity were left at the default of 0.7 and 0.75 respectively. The second bioinformatic analysis was to determine if the SNPs caused any predicted changes in any transcription factor binding sites. Transcription factor binding sites were identified using TFSEARCH against the TRANSFAC database using the default threshold score of 85. For SNPs that are located in transcription factor binding sites, the variant sequences were also analyzed to determine if the SNP caused a change in the score of the predicted binding site.

TATA Box SNP genotyping

We designed a Taqman™ assay to genotype the TATA SNP (rs35036378) in samples from additional populations. The sequences of the primers and probes are listed in the Supplementary Table 1. The Taqman assay was used to genotype an additional 50 African American, 10 Chinese, 10 Japanese and 9 Indo-Pakistanis from the Coriell repository and 221 Caucasians from the COBRA tamoxifen breast cancer clinical trial (ClinicalTrials.gov Identifier #NCT00228930). This was a prospective observational clinical trial of women 18 years or older who were treated with tamoxifen in the adjuvant setting or for the prevention of breast cancer[15]. The use of the samples from the COBRA tamoxifen clinical trial was approved by the IRBs from Georgetown University Medical Center, the University of Michigan and Indiana University Medical Center.

We also used a PCR restriction fragment length polymorphism (RFLP) assay to verify a subset of the sequencing and genotyping results for rs35036378. We used the amplicon 4 primers and digested the resultant amplicon with the DraI restriction enzyme (New England Biolabs, Ipswich, MA), which cuts the wild-type (T) but not the variant (G).

ESR2-TATA box luciferase reporter assays

A 410 bp sequence of the *ESR2* promoter containing the TATA box (nucleotide 45761496 to 45761087 from accession # NT_026437) was cloned into HindIII/XhoI site of the pGL3 Basic promoterless vector (Promega Corporation, Madison, WI). The plasmid with the variant TATA box was created using the Quick Change® Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). The sequences of the wild type and variant inserts were confirmed by DNA sequencing. The wild type and variant plasmids were transfected into the LNCaP and SKBR3 cells.

LNCaP cells were plated in a 24-well plate at a cell density of 7×10^4 cells/well and were grown overnight in DMEM with 10% FBS to approximately 80% confluency. The DNA mix for transfection was prepared in Opti-MEM I Reduced Serum Medium (Invitrogen Corporation, Carlsbad, CA) and consisted of 0.5 μg of the test plasmid (wild type or variant) with 0.05 μg of the phRGTK renilla control vector (Promega, Madison, WI) that served as an internal control to normalize luciferase activity. The transfection was carried out using Lipofectamine LTX (Invitrogen Corporation, Carlsbad, CA).

SKBR3 cells were plated in a 12-well plate at a cell density of 4×10^5 cells/well and were grown overnight in DMEM with 10% FBS to approximately 80% confluency. The DNA mix for transfection was prepared in Opti-MEM I Reduced Serum Medium (Invitrogen

Corporation, Carlsbad, CA) and consisted of 1.6 μg of the test plasmid (wild type or variant) with 0.30 μg of the phRGTK renilla control vector (Promega, Madison, WI) that served as an internal control to normalize luciferase activity. The transfection was carried out using Lipofectamine 2000(Invitrogen Corporation, Carlsbad, CA).

Cells were harvested after 24 hr and the luciferase activity was determined using the Dual Luciferase assay kit (Promega Corporation, Madison, WI) in a SIRIUS luminometer (Berthold Technologies, Oak Ridge, TN). The experiments were repeated on three different days.

To assure reproducibility of the reporter assay results, the experiments with the LNCaP cells were repeated on different days, in 2 different laboratories (TS and SO), using different batches of DNA.

RESULTS

ESR2 resequencing

Re-sequencing of a 1.6kb *ESR2* region in 100 samples led to the identification of five novel SNPs in the promoter and one in intron 1of the *ESR2* gene (Table 1; Figure 1). These results have been submitted to the PharmGKB (PS203888, PS203889) and dbSNP (rs#'s are in Table 1) databases. The five promoter SNPs were observed in the African American population, but not in the Caucasian samples, whereas the intronic SNP was observed in both populations. No SNPs were identified in the exon ON. The PHASE software was utilized to estimate the haplotypes from the DNA sequencing data (Table 2). The program predicted 11 different haplotypes with wild type being the most frequent.

The TFsearch software was used to predict the functional significance of promoter SNPs. Based on the difference in TFsearch TFBS scores, three of the promoter SNPs were predicted to alter a transcription factor binding site(Table 1). We focused additional studies on confirming the rs35036378 (-43T>G) SNP, which was predicted to reduced the function of the *ESR2* TATA box.

Functional variant in TATA box

To confirm the presence of this TATA box SNP (rs35036378) in the Coriell DNA samples, we used a restriction fragment length polymorphism (RFLP) assay. The results were concordant with the sequencing results (data not shown). Using a Taqman assay, we genotyped additional samples for a total of 100 African American, 271 Caucasian, 10 Chinese, 10 Japanese, and 9 Indo-Pakistani samples. We did not identify this SNP in any of the Caucasian, Chinese, Japanese or Indo-Pakistani samples, but only in the African American samples. In total, there were four samples (all African Americans) that were heterozygous for the TATA box SNP (NA17110, NA17131, NA17170 and NA17196).

We prioritized the SNPs for functional studies in part based on the documented conservation of the SNP regions. Using ClustalW and rVISTA, we determined the inter-species conservation of the region of the *ESR2* gene that was sequenced. The wild type TATA box sequence, which includes rs35036378, and several neighboring nucleotides were 100% conserved across all four species (Figure 2). In contrast, the regions around the other SNPs were not highly conserved (data not shown).

Functional Studies

We used the pGL3 promoter luciferase plasmid to determine the functional significance of the TATA box SNP (rs35036378). We cloned 410 bp (nucleotide 45761087 to 45761496

from accession # NT_026437) of the wild type *ESR2* promoter into the pGL3 basic vector. The rs35036378 SNP was incorporated using site-directed mutagenesis. These plasmids were then transfected into the LNCaP and SKBR3 cells. These cell lines were chosen for this study because they are known to express *ESR2*, suggesting that the cloned *ESR2* promoter would be active in these cells. Compared to the wild type plasmid, the luciferase activity from the variant plasmid was reduced by ~50% in both cell lines (Figure 3). These results suggest that rs35036378 in the *ESR2* TATA box is a functional SNP with significant effect on *ESR2* gene expression.

DISCUSSION

In this study, we identified six SNPs in the promoter of the *ESR2* gene. Three of these are in predicted transcription factor binding sites. They all reduced the TRANSFAC transcription factor binding site scores for their respective binding sites. We focused our laboratory validation on rs35036378, since our bioinformatic analyses predicted that this SNP would reduce the functionality of the *ESR2* promoter TATA box. Others have previously shown that this TATA box is functionally important in the proximal promoter of the *ESR2* gene [16]. Previous studies have shown that two specific isoforms of *ESR2* originated from two distinct untranslated first exons namely exon OK and exon ON [17] and that these regulatory mechanisms can influence the deregulation of *ESR2* expression in cancer [18]. Studies by Ashworth et al have shown an association between venous ulceration and SNPs in the region spanning exon ON and promoter [19]. Various other studies have shown that methylation of the *ESR2* promoter ON is associated with reduced expression of *ESR2* isoforms in breast, prostate and ovarian cancer tissue and cell lines [20-22].

The rs35036378 SNP found in the untranslated region 5' of exon ON reduced the *ESR2* promoter luciferase reporter activity by approximately 50%. The luciferase assays were performed in the SKBR3 breast cancer and LNCaP prostate cancer cell lines because they are known to express *ESR2* mRNA. In the context of cell types *in vivo*, the SNP could have more or less effects. Since ESR2 appears to have roles in multiple tissues, such as the brain [23], ovary [24], prostate [25], this SNP could have effects on a variety of functions. The altered phenotypes observed in the *ESR2* knockout mice (βERKO) [26] and in mice treated with ESR2 selective ligands [27] should be useful in guiding which phenotypes may be affected by this SNP in humans. These would likely include effects on neurologic and reproductive function.

Since the *ESR2* promoter SNPs are relatively rare, the large numbers of subjects that would be required to obtain sufficient power to detect an association would likely make a traditional genotypephenotype study impractical. Consequently, deriving an algorithm for combining multiple rare SNPs will be necessary. We have used such a scoring system for the drug metabolizing enzyme genes for which the functions of multiple rare SNPs are well characterized [9]. In order to establish these algorithms, the function of each SNP must be known. The studies described here provide a functional characterization of the rs35036378 SNP that would be the beginning of such an effort. In addition, a better understanding of the function of the individual *ESR2* SNPs may also help understand the mechanisms behind the SNPs that have been associated with clinical phenotypes [1-6].

These studies are the beginning of what will be required to thoroughly understand the functional genetic variability in the *ESR2* gene. Within our study, there are two additional SNPs that are predicted to affect 3 transcription factor binding sites. Each of these will require careful testing. Individually, each of the three SNPs that were predicted to be functional (rs8008187, rs3829768, rs35036378) are relatively rare. However, if the bioinformatic functional predictions are correct, then the combined frequency of the

haplotypes containing one or more functionally affected alleles would total nearly 10% in the African American population. It may then be possible to conduct genotype-phenotype association studies with all of the reduced function alleles combined into one group, similar to the way that we combine *CYP2D6* poor metabolizer alleles, so the frequency would be approximately 10%, which should be sufficient to detect genotype-phenotype associations in reasonable sized studies. This illustrates the importance of further *in vitro* studies to test each of the promoter SNPs. There are also many SNPs throughout the rest of the *ESR2* gene that may also have effects on gene function through mechanisms such as alternative splicing. Since methylation appears to be part of the regulatory mechanisms of *ESR2* expression [28], SNPs that affect *ESR2* methylation may also be of substantial importance. These SNPs will all require similar testing to obtain a comprehensive understanding of the influence of genetic variability on the function of the *ESR2* gene.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This research was support by NIH-PGRN (5U01GM061373, D.A.F.), NIH-NIGMS (U01GM061373, 5K24RR020815), NIH-NCRR M01-RR000042 (University of Michigan), M01-RR020359 (Georgetown University), and M01-RR00750 (Indiana University). We would like to acknowledge the additional COBRA investigators, Daniel Hayes, N. Lynn Henry, Vered Stearns, and Anna Maria Storniolo, who were the clinical investigators on the tamoxifen trial from which some of the samples were used in this study.

References

1. Jin Y, Hayes DF, Li L, Robarge JD, Skaar TC, Philips S, Nguyen A, Schott A, Hayden J, Lemler S, Storniolo AM, Flockhart DA, Stearns V. Estrogen receptor genotypes influence hot flash prevalence and composite score before and after tamoxifen therapy. J Clin Oncol. 2008; 26:5849–5854. [PubMed: 19018086]
2. Ntukidem NI, Nguyen AT, Stearns V, Rehman M, Schott A, Skaar T, Jin Y, Blanche P, Li L, Lemler S, Hayden J, Krauss RM, Desta Z, Flockhart DA, Hayes DF. Estrogen receptor genotypes, menopausal status, and the lipid effects of tamoxifen. Clin Pharmacol Ther. 2008; 83:702–710. [PubMed: 17713466]
3. Ashton KA, Proietto A, Otton G, Symonds I, McEvoy M, Attia J, Gilbert M, Hamann U, Scott RJ. Estrogen receptor polymorphisms and the risk of endometrial cancer. BJOG. 2009; 116:1053–1061. [PubMed: 19438492]
4. Domingues-Montanari S, Subirana I, Tomas M, Marrugat J, Senti M. Association between ESR2 genetic variants and risk of myocardial infarction. Clin Chem. 2008; 54:1183–1189. [PubMed: 18487282]
5. Pirskanen M, Hiltunen M, Mannermaa A, Helisalmi S, Lehtovirta M, Hanninen T, Soininen H. Estrogen receptor beta gene variants are associated with increased risk of Alzheimer's disease in women. Eur J Hum Genet. 2005; 13:1000–1006. [PubMed: 15944651]
6. Yu KD, Rao NY, Chen AX, Fan L, Yang C, Shao ZM. A systematic review of the relationship between polymorphic sites in the estrogen receptor-beta (ESR2) gene and breast cancer risk. Breast Cancer Res Treat. 2010
7. Wang K, Dickson SP, Stolle CA, Krantz ID, Goldstein DB, Hakonarson H. Interpretation of association signals and identification of causal variants from genome-wide association studies. Am J Hum Genet. 2010; 86:730–742. [PubMed: 20434130]
8. Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. PLoS Biol. 2010; 8:e1000294. [PubMed: 20126254]
9. Borges S, Desta Z, Jin Y, Faouzi A, Robarge JD, Philip S, Nguyen A, Stearns V, Hayes D, Rae JM, Skaar TC, Flockhart DA, Li L. Composite functional genetic and comedication CYP2D6 activity

score in predicting tamoxifen drug exposure among breast cancer patients. J Clin Pharmacol. 2010; 50:450–458. [PubMed: 20081063]

10. Stephens M, Donnelly P. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am J Hum Genet. 2003; 73:1162–1169. [PubMed: 14574645]
11. Stephens M, Smith NJ, Donnelly P. A new statistical method for haplotype reconstruction from population data. Am J Hum Genet. 2001; 68:978–989. [PubMed: 11254454]
12. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994; 22:4673–4680. [PubMed: 7984417]
13. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I. VISTA: computational tools for comparative genomics. Nucleic Acids Res. 2004; 32:W273–279. [PubMed: 15215394]
14. Loots GG, Ovcharenko I, Pachter L, Dubchak I, Rubin EM. rVista for comparative sequence-based discovery of functional transcription factor binding sites. Genome Res. 2002; 12:832–839. [PubMed: 11997350]
15. Jin Y, Desta Z, Stearns V, Ward B, Ho H, Lee KH, Skaar T, Storniolo AM, Li L, Araba A, Blanchard R, Nguyen A, Ullmer L, Hayden J, Lemler S, Weinshilboum RM, Rae JM, Hayes DF, Flockhart DA. CYP2D6 genotype, antidepressant use, and tamoxifen metabolism during adjuvant breast cancer treatment. J Natl Cancer Inst. 2005; 97:30–39. [PubMed: 15632378]
16. Li LC, Yeh CC, Nojima D, Dahiya R. Cloning and characterization of human estrogen receptor beta promoter. Biochem Biophys Res Commun. 2000; 275:682–689. [PubMed: 10964723]
17. Hirata S, Shoda T, Kato J, Hoshi K. The multiple untranslated first exons system of the human estrogen receptor beta (ER beta) gene. J Steroid Biochem Mol Biol. 2001; 78:33–40. [PubMed: 11530282]
18. Smith L, Coleman LJ, Cummings M, Satheesha S, Shaw SO, Speirs V, Hughes TA. Expression of oestrogen receptor beta isoforms is regulated by transcriptional and post-transcriptional mechanisms. Biochem J. 2010; 429:283–290. [PubMed: 20462399]
19. Ashworth JJ, Smyth JV, Pendleton N, Horan M, Payton A, Worthington J, Ollier WE, Ashcroft GS. Polymorphisms spanning the 0N exon and promoter of the estrogen receptor-beta (ERbeta) gene ESR2 are associated with venous ulceration. Clin Genet. 2008; 73:55–61. [PubMed: 18070128]
20. Zhu X, Leav I, Leung YK, Wu M, Liu Q, Gao Y, McNeal JE, Ho SM. Dynamic regulation of estrogen receptor-beta expression by DNA methylation during prostate cancer development and metastasis. Am J Pathol. 2004; 164:2003–2012. [PubMed: 15161636]
21. Suzuki F, Akahira J, Miura I, Suzuki T, Ito K, Hayashi S, Sasano H, Yaegashi N. Loss of estrogen receptor beta isoform expression and its correlation with aberrant DNA methylation of the 5'-untranslated region in human epithelial ovarian carcinoma. Cancer Sci. 2008; 99:2365–2372. [PubMed: 19032364]
22. Zhao C, Lam EW, Sunters A, Enmark E, De Bella MT, Coombes RC, Gustafsson JA, Dahlman-Wright K. Expression of estrogen receptor beta isoforms in normal breast epithelial cells and breast cancer: regulation by methylation. Oncogene. 2003; 22:7600–7606. [PubMed: 14576822]
23. Fan X, Xu H, Warner M, Gustafsson JA. ERbeta in CNS: new roles in development and function. Prog Brain Res. 2010; 181:233–250. [PubMed: 20478441]
24. Drummond AE, Fuller PJ. The importance of ERbeta signalling in the ovary. J Endocrinol. 2010; 205:15–23. [PubMed: 20019181]
25. Ellem SJ, Risbridger GP. The dual, opposing roles of estrogen in the prostate. Ann N Y Acad Sci. 2009; 1155:174–186. [PubMed: 19250203]
26. Couse JF, Korach KS. Estrogen receptor null mice: what have we learned and where will they lead us? Endocr Rev. 1999; 20:358–417. [PubMed: 10368776]
27. Minutolo F, Macchia M, Katzenellenbogen BS, Katzenellenbogen JA. Estrogen receptor beta ligands: Recent advances and biomedical applications. Med Res Rev. 2009
28. Xue Q, Lin Z, Cheng YH, Huang CC, Marsh E, Yin P, Milad MP, Confino E, Reierstad S, Innes J, Bulun SE. Promoter methylation regulates estrogen receptor 2 in human endometrium and endometriosis. Biol Reprod. 2007; 77:681–687. [PubMed: 17625110]

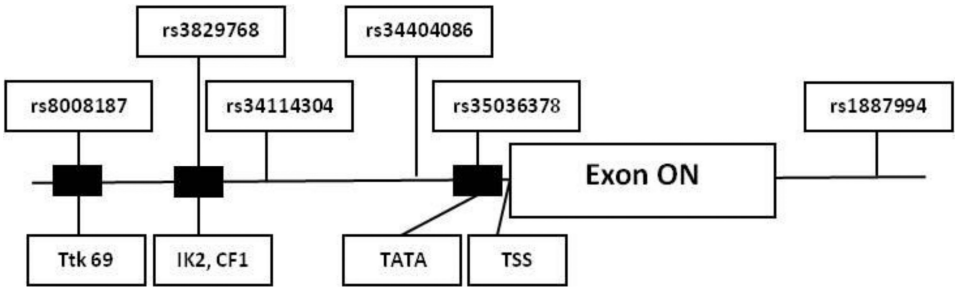

Fig 1.

Schematic representation of the positions of the SNPs in the *ESR2* gene. The designated position of each SNP is relative to the first nucleotide of exonON, which is nucleotide# 45761078 of accession# NT_026437. The black boxes represent the transcription factor binding sites in which the SNP exists. [TSS : transcription start site]

Fig 2.

Conservation of the TATA box region of the *ESR2* gene. The alignment was constructed using ClustalW. The * below the sequence indicates conservation across all 4 species. The bold nucleotides indicate the TATA box and the arrow above indicates the position of the SNP.

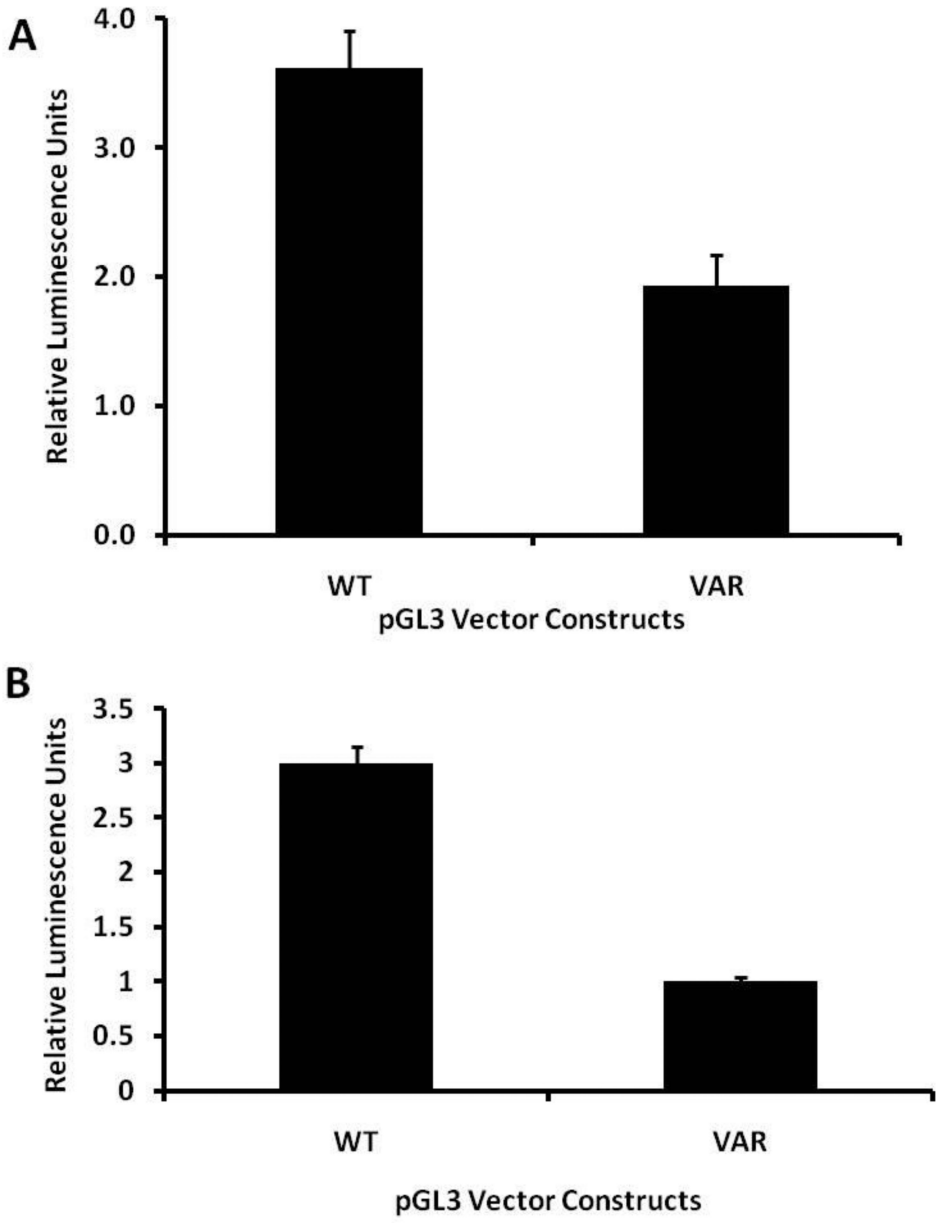

Fig 3.

Effect of the TATA box SNP on *ESR2* promoter activity. Wild-type (WT) and variant (VAR; -43(T/G)) promoters were cloned into the pGL3 plasmid and expressed in LNCaP (A) and SKBR3 (B) cells. Values are means and standard errors of the firefly luciferase activity normalized to the renilla luciferase activity of 3 experiments performed on three different days (p-value: 0.009(LNCaP), 0.0002(SKBR3)).

Table 1

SNPs in the *ESR2* gene.

dbSNP rsID	Minor Allelic Freq.		SNP Position	Domain	Transcription factor binding site consensus sequence	Functional TFBS	TFBS Scores	
	CA	AA					Wild Type	Variant
rs8008187	0.00	0.01	-727 C>G	Promoter	tgggGCTCCTG**C**tgg	Ttk 69	93.3	86.7
rs3829768	0.00	0.07	-516 T>C	Promoter	aagGCCTTCCCAG**T**Gacc	IK-2	89.5	87.7
					tcccAG**T**GACCTCttga	CF1	87.7	77.6
rs34114304	0.00	0.01	-440 G>A	Promoter	-	-	-	-
rs34404086	0.00	0.01	-154 G>A	Promoter	-	-	-	-
rs35036378	0.00	0.02	-43 T>G	Promoter	cggtc**T**TTAAAAggaag	TATA Box	94	84.4
rs1887994	0.06	0.04	+468 G>T	Intron	-	-	-	-

The position is relative to the first nucleotide of exon1. The bold nucleotide in the consensus sequence is the polymorphic nucleotide. The transcription factor binding sites (TFBS) were determined using the TFsearch. The (-) means there were no TFBS predicted to be affected by the SNP. The minor allele frequencies are from 50 African Americans (AA) and 50 Caucasian (CA).

Table 2

The 11 haplotypes that were identified in the 50 African American and 50 Caucasians along with their frequency of occurrence.

Index	Haplotype	Frequency(AA)	Frequency(CA)
1	000000	0.854	0.930
2	000001	0.040	0.070
3	000010	0.020	0.000
4	000100	0.002	0.000
5	001000	0.010	0.000
6	010000	0.063	0.000
7	010100	0.002	0.000
8	100000	0.002	0.000
9	100100	0.003	0.000
10	110000	0.002	0.000
11	110010	0.004	0.000

The SNP at each position in the haplotype from left to right correspond with the SNPs in Table 1 from top to bottom.