Phenol Sulfotransferases: Hormonal Regulation, Polymorphism, and Age of Onset of Breast Cancer

Pankaj Seth, Kathryn L. Lunetta, Daphne W. Bell, Heather Gray, Selim M. Nasser, Esther Rhei, Carolyn M. Kaelin, Dirk J. Iglehart, Jeffrey R. Marks, Judy E. Garber, Daniel A. Haber, and Kornelia Polyak

Departments of Adult Oncology [P. S., H. G., J. E. G., K. P.]; and Biostatistics [K. L. L.], Dana-Farber Cancer Institute, Boston, Massachusetts 02115; Brigham and Women’s Hospital [E. R., C. M. K., D. J. I.], Boston, Massachusetts 02115; Center for Cancer Risk Analysis, Massachusetts General Hospital, Charlestown, Massachusetts 02129 [D. W. B., D. A. H.;]. Department of Pathology, Massachusetts General Hospital, Boston, Massachusetts 02114 [S. M. N.]; Harvard Medical School, Boston, Massachusetts 02115 [P. S., D. W. B., H. G., S. M. N., E. R., C. M. K., D. J. I., J. E. G., D. A. H., K. P.]; Harvard School of Public Health, Boston, Massachusetts 02115 [K. L. L.]; and Department of Surgery, Duke University Medical Center, Durham, North Carolina 27710 [J. R. M.]

Abstract

In recent years, significant effort has been made to identify genes that influence breast cancer risk. Because the high-penetrance breast cancer susceptibility genes BRCA1 and 2 play a role only in a small fraction of breast cancer cases, understanding the genetic risk of the majority of breast cancers will require the identification and analysis of several lower penetrance genes. The estrogen-signaling pathway plays a crucial role in the pathophysiology of breast cancer; therefore, polymorphism in genes involved in this pathway is likely to influence breast cancer risk. Our detailed analysis of gene expression profiles of estrogen- and 4-OH-tamoxifen-treated ZR75-1 breast cancer cells identified members of the sulfotransferase family as downstream targets of tamoxifen. On the basis of the inactivation of SULT1A by 4-OH-tamoxifen and the known inherited variability in SULT1A enzymatic activity, we hypothesized that polymorphism in sulfotransferases might influence the risk of breast cancer. Using an RFLP that distinguishes an arginine to histidine change in exon 7 of the SULT1A1 gene, we characterized SULT1A1 genotypes in relation to breast cancer risk. An analysis of 444 breast cancer patients and 227 controls revealed no effect of SULT1A1 genotype on the risk of breast cancer (P = 0.69); however, it did appear to influence the age of onset among early-onset affected patients (P = 0.04). Moreover, individuals with the higher activity SULT1A1*1 allele were more likely to have other tumors in addition to breast cancer (P = 0.004; odds ratio, 3.02; 95% confidence interval, 1.32, 8.09). The large number of environmental mutagens and carcinogens activated by sulfotransferases and the high frequency of the SULT1A1*1 allele in human populations warrants additional studies to address the role of SULT1A genes in human cancer.

Introduction

Breast cancer is a leading cause of cancer death in women of the Western world. Despite advances in early detection and treatment, breast cancer mortality rates have not decreased significantly over the past few decades. Therefore, major emphasis has been placed on the identification of risk factors and potential targets for chemoprevention with the aim of decreasing the incidence of the disease. Estrogen is a major risk factor for breast cancer and, as demonstrated by recent clinical trials (1, 2), is an excellent target for chemopreventive therapies. Despite the significance of this issue, our understanding of the mechanism by which estrogen promotes and inhibits the growth of normal and cancerous mammary epithelial cells is relatively limited. Both estrogen and tamoxifen initiate cellular responses by binding to the ER, a ligand-dependent transcription factor; therefore, it is reasonable to believe that some genes induced by these ligands might directly mediate their growth regulatory effects (3, 4). By analyzing the global transcriptional response to estrogen or tamoxifen in human breast cancer cells using SAGE (5), we have identified members of the highly homologous SULT1A phenol sulfotransferase gene family as the only transcript(s) among over 8000 transcripts analyzed that is significantly up-regulated by 4-OH-tamoxifen.

Sulfotransferases are involved in the metabolism of xenobiotics and endogenous chemicals (steroids, catecholamines, and iodothyronines; Refs. 6, 7). Sulfonation is generally associated with inactivation of xenobiotics, although a large number of carcinogens and mutagens are activated by this process. In humans, there are three members of the phenol sulfotransferase family (SULT1A1, SULT1A2, and SULT1A3). DNA sequences and the structure of these three enzymes are highly homologous, and all three genes are localized tightly linked on chromosome 16p12.1-p11.2 (8). Similar to other drug-metabolizing enzymes (glutathione transferases, N-acetyltransferases, and others), inherited differences in the enzymatic activity of sulfotransferases are likely to influence the individual’s risk for cancer (9), but this connection has not yet been investigated. Interestingly, exon 7 of the SULT1A1 gene contains a recently described, functionally relevant polymorphism that significantly influences its enzymatic activity (10). Cell extracts from individuals homozygous for the low activity allele (SULT1A*2) have ~10-fold lower phenol sulfotransferase activity compared with that of individuals with the high activity (SULT1A*1) allele (11). Similarly, there are several functionally relevant polymorphisms described in the highly related (96% identity at the amino acid level) but much lower activity SULT1A2 gene; however, thus far no high throughput assay has been developed for their analysis (11–13).

To determine whether SULT1A1 polymorphism and the resulting variability in enzymatic activity would influence breast cancer risk, we analyzed SULT1A1 genotypes in relation to breast cancer risk. To our knowledge, this is the first report investigating the association between polymorphism in sulfotransferases and cancer risk in humans.

Materials and Methods

Cell Culture and Analysis of Gene Expression Profiles Using SAGE

The ZR75-1 cell line was obtained from the American Type Culture Collection and cultured in RPMI medium supplemented with 10% fetal bovine serum and...
10 μg/ml insulin. For the generation of SAGE libraries, ZR75-1 cells were cultured for 7 days in phenol red-free RPMI medium supplemented with 5% charcoal-treated fetal bovine serum, after which one plate received fresh medium (untreated controls), one plate received fresh medium containing 10 nM estradiol (estrogen-treated cells), and the third plate received fresh medium containing 10 μM 4-OH-tamoxifen (tamoxifen-treated cells). Cells were collected after 16 h and used for the generation of SAGE libraries essentially as described (14).

**Northern and Immunoblot Analysis.** For Northern blot analysis, RNA from ZR75-1 cells was probed with a 78 bp PCR-derived probe corresponding to the 1B alternatively spliced exon that contained the 5'UTR of the SULT1A cDNA. poly(A) RNA was prepared using μMACs (Miltenyi Biotech) kit following the manufacturer’s instructions. RNA electrophoresis and hybridization were performed as described (14). Immunoblot analysis of ZR75-1 cell extracts was performed with antihuman SULT1A3 polyclonal antibody (Oxford Biomed.).

**Gene-specific RT-PCR Reactions.** To determine which SULT1A genes are expressed in ZR75-1 cells, the following oligonucleotides were used for RT-PCR analysis. The forward primers were SULT-F2 (5'-agggccagtgccagcaggggtg-3'), SULT-F3 (5'-agtcctccacaaacaccc-3'), and SULT-FC (coding region), 5'-agtcctccacaaacaccc-3'. The reverse primers were SULT1A1R (specific for SULT1A1), 5'-ccctcaatttcatcatatcgtgtg-3', SULT1A2R1 (specific for SULT1A2), 5'-gacatcctccacaaacaccc-3', SULT1A2R2 (specific for SULT1A2), 5'-gacatcctccacaaacaccc-3', and SULT1A2R3 (specific for SULT1A2), 5'-acctccagtgccagcaggggtg-3'. RT-PCR was performed essentially as described (14). The identity of the PCR fragments was confirmed by cycle sequencing (Thermosequenase, USB).

**Subjects and Sample Preparation.** Patients were selected based on early onset of breast cancer (under the age of 40 in the MGH cohort and under the age of 65 (majority of the patients under the age 60) in the DFCI cohort). Both these women and the healthy blood donors provided written informed consent for research under protocols approved by the institutional review boards at each institution. The majority (79-99%) of the subjects were Caucasians, and most affected patients did not have a family history of breast cancer. Genomic DNA was extracted using standard protocols.

**PCR-RFLP Assay for SULT1A1 Genotype Analysis.** PCR amplifications and RFLP analysis were performed essentially as described with minor modifications (10). A detailed protocol is available from the authors upon request.

**Statistical Analysis.** Pearson χ² test was used to test for independence of alleles (Hardy-Weinberg Equilibrium, HWE) within each sample. Fisher’s exact test was used to test for differences in genotype and allele frequencies among the three samples and between patients and controls. Within each sample, Fisher's exact test was used to test for relationships between genotype and various available predictors, such as race, ER status, and other tumor present. We used ANOVA to determine whether the genotypes that coded four different ways (three genotypes, dominant IA1*1, IA1*2, and the additive IA1*1 allele) were significantly related to age of onset in each sample. Because the data may not conform well to the assumptions of ANOVA, we also performed the nonparametric K-W test in each case.

**Results**

**SULT1A mRNA Levels Are Induced after Tamoxifen Treatment.** To determine the global transcriptional response of breast cancer cells to estrogen or tamoxifen, we generated SAGE libraries from an estrogen-dependent human breast cancer cell line, ZR75-1, before and after estrogen or 4-OH-tamoxifen treatment.2 Of over 8,000 transcripts analyzed, only one SAGE tag (GCTCGGGGACT) was found to be markedly (10-fold; \( P = 0.00229 \)) increased by 4-OH-tamoxifen but not by estrogen treatment (Fig. 1A). This SAGE tag can correspond to SULT1A1, SULT1A2, or SULT1A3 (phenol-prefering sulfotransferase 1, 2, and 3) because of their high degree of similarity; therefore, we arbitrarily refer to this transcript as SULT1A. To confirm the induction of SULT1A by tamoxifen, we performed Northern blot analysis using poly(A) RNA and a 78 bp probe that corresponded to an alternatively spliced exon 1B that contained the 5'UTR (Fig. 1B). Because of the high similarity of the SULT1A cDNAs, we were unable to design a gene-specific probe for Northern blot analysis. Sequence analysis of PCR fragments determined that both SULT1A1 and SULT1A2 mRNAs contain this alternatively spliced exon. As Fig. 1 shows, the mRNA levels of SULT1A were induced by 24 h with an even more significant increase detected at 72 h after 4-OH-tamoxifen treatment. However, even at this time point the increase in the mRNA levels is modest (2–3-fold). Surprisingly, estrogen treatment also increased SULT1A mRNA levels at later time points (Fig. 1B), which were not detected by SAGE because this analysis was performed at earlier (16 h) time points. To determine which one of the three SULT1A genes is affected by tamoxifen treatment, we performed RT-PCR and immunoblot analysis using gene-specific primers and antibodies (when available). In these experiments, we found that, although SULT1A3 was not affected, we were not able to determine conclusively if SULT1A1, SULT1A2, or both mRNA levels are affected by hormonal treatment in ZR75-1 cells (data not shown). Because of the lack of commercially available antibodies, we were unable to analyze the protein levels of SULT1A1 and SULT1A2.

**SULT1A Genotype and Breast Cancer Risk.** The involvement of SULT1A in the metabolism of carcinogens and endogenous hormones and its induction by estrogenic compounds prompted us to investigate the relationship between a functionally relevant polymorphism in the SULT1A gene and breast cancer risk. We used a PCR-RFLP approach developed for detecting a G to A transition at nucleotide 638 in exon 7 of the SULT1A gene. Because the high similarity of the SULT1A cDNAs, we were unable to design a gene-specific probe for Northern blot analysis. Sequence analysis of PCR fragments determined that both SULT1A1 and SULT1A2 mRNAs contain this alternatively spliced exon. As Fig. 1 shows, the mRNA levels of SULT1A were induced by 24 h with an even more significant increase detected at 72 h after 4-OH-tamoxifen treatment. However, even at this time point the increase in the mRNA levels is modest (2–3-fold). Surprisingly, estrogen treatment also increased SULT1A mRNA levels at later time points (Fig. 1B), which were not detected by SAGE because this analysis was performed at earlier (16 h) time points. To determine which one of the three SULT1A genes is affected by tamoxifen treatment, we performed RT-PCR and immunoblot analysis using gene-specific primers and antibodies (when available). In these experiments, we found that, although SULT1A3 was not affected, we were not able to determine conclusively if SULT1A1, SULT1A2, or both mRNA levels are affected by hormonal treatment in ZR75-1 cells (data not shown). Because of the lack of commercially available antibodies, we were unable to analyze the protein levels of SULT1A1 and SULT1A2.

**SULT1A Genotype and Breast Cancer Risk.** The involvement of SULT1A in the metabolism of carcinogens and endogenous hormones and its induction by estrogenic compounds prompted us to investigate the relationship between a functionally relevant polymorphism in the SULT1A gene and breast cancer risk. We used a PCR-RFLP approach developed for detecting a G to A transition at nucleotide 638 in exon 7 of the SULT1A gene. This transition leads to an arginine to histidine change and significantly reduced enzymatic activity. A representative gel with an enlarged area demonstrating fragment sizes that correspond to the three genotypes is shown in Fig. 2. We tested a total of 444 breast cancer patients from three different cohorts and 227 controls (healthy blood donors, male and female) free of malignancy. Two of the cohorts consisted of 378 early-onset breast cancer patients whose data were collected at MGH (280 cases; <40 years of age at diagnosis) and DFCI (98 cases; <57 years of age at diagnosis), whereas the third cohort included 66 sporadic breast cancer pa-

---

1 Seth et al. manuscript in preparation.
patients whose data were collected from consecutive surgeries at Brigham and Women’s Hospital, MGH, and Duke University Medical Center. Age and ethnicity information was not available for 98 controls, whereas 129 controls were ethnically and age matched to the MGH patient set.

Table 1 shows the SULT1A1 genotype and allele frequency distribution among the three patient samples and in controls. No evidence for differences in allele or genotype frequencies was found between each of the three patient samples and controls (Fisher’s exact test; \( P = 0.75, 0.55, 0.47 \) for the MGH, DFCI, and Other sample set, respectively). Similarly, no evidence for the association of genotype or allele frequency with various other parameters (ER status and tumor stage at diagnosis) was found in any of the three samples. Previous data indicated that SULT1A1 allele frequency may be influenced by ethnicity (10). In accordance with this, in the MGH sample set we found some evidence of allele frequency differences between Caucasian and non-Caucasian (11) allele frequency (0.70 and 0.93, respectively; \( P = 0.08 \)), but this did not reach statistical significance and is based on only seven non-Caucasians in the sample versus 273 Caucasians.

Interestingly, in the MGH patient set, there were 27 patients who had a history of other tumors in addition to breast cancer at the time of diagnosis. These tumors included cervical (\( n = 1 \)), vulval (\( n = 1 \)), ovarian (\( n = 6 \)), colon (\( n = 1 \)), thyroid (\( n = 1 \)), and basal cell (\( n = 8 \)) carcinomas; osteosarcomas (\( n = 2 \)); melanomas (\( n = 4 \)); and Hodgkin’s lymphomas (\( n = 3 \)). None of the breast/osteosarcoma patients had Li-Fraumeni syndrome, and none of the breast/thyroid cancer patients had Cowden syndrome. One of the breast/ovarian cancer patients was found to have a BRCA1 mutation, and another patient had a strong family history of breast cancer that indicated the potential involvement of a high-penetration breast cancer susceptibility gene. Although high-penetration genes might confound the weaker effect of the SULT1A1 polymorphism, removal of these two (or even all six) cases from the analysis would not influence the overall result. Statistical analysis revealed that patients are more likely to have SULT1A1*1/1A*1 (20 cases) or SULT1A1*1/1A*2 (7 cases) genotype if they have other tumors in addition to breast cancer than if they do not (\( P = 0.022 \)). Using allele data, the 1A*1 allele is associated with having other tumors present (\( P = 0.004; \) odds ratio = 3.02; 95% confidence interval, 1.32, 8.09). This may indicate that in certain genetic background or environmental conditions, high sulfotransferase enzymatic activity may increase cancer risk in general, or it may modify the penetrance of mutations in certain tumor suppressor gene(s). This observation is not unexpected because several animal and in vitro studies have found an association between sulfonation activity and incidence of chemically induced cancers (15).

There is a possibility that, similar to other metabolic enzymes, SULT1A1 polymorphism may not influence breast cancer risk, but it may influence the age of onset in affected patients (16). To determine this, we analyzed whether there is a relationship between genotype and mean age of onset using parametric and nonparametric statistical methods (ANOVA and K-W test) for each sample set (Tables 2 and 3). By this approach in the MGH sample set, there is some borderline significant evidence that genotype affects age of onset, using a dominant SULT1A1 allele model (ANOVA, \( P = 0.05; \) K-W, \( P = 0.06 \)). In the DFCI sample set, there is some evidence that genotype affects age of onset, using a recessive SULT1A1*1 allele model (ANOVA and K-W, \( P = 0.03 \)). We found no evidence of a genotype effect on age of onset using any genotype coding in the sporadic patients’ sample set. Because prior analysis indicated that the SULT1A1 genotype and allele frequency is not statistically different in the different patient sample sets, we combined the early onset MGH and DFCI sample sets and performed a two-way ANOVA that accounted for genotype and sample. The three-genotype ANOVA model has \( P = 0.04 \) for genotype. Hence, although the evidence for dominance/recessiveness of the SULT1A1*1 allele in the two sample sets is different, the two do not cancel each other out. Thus, the combined sample does provide evidence that genotype has an effect.

Table 1 SULT1A1 genotype and allele distribution

<table>
<thead>
<tr>
<th>Sample</th>
<th>IA<em>1/IA</em>1</th>
<th>IA<em>1/IA</em>2</th>
<th>IA<em>2/IA</em>2</th>
<th>IA*1+ allele frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MGH</td>
<td>139 (49.6%)</td>
<td>118 (42.1%)</td>
<td>23 (8.2%)</td>
<td>0.707</td>
</tr>
<tr>
<td>DFCI</td>
<td>54 (55.1%)</td>
<td>36 (36.7%)</td>
<td>8 (8.2%)</td>
<td>0.735</td>
</tr>
<tr>
<td>Other</td>
<td>36 (54.5%)</td>
<td>22 (33.3%)</td>
<td>8 (12.1%)</td>
<td>0.712</td>
</tr>
<tr>
<td>Controls</td>
<td>110 (48.5%)</td>
<td>94 (41.4%)</td>
<td>23 (10.1%)</td>
<td>0.692</td>
</tr>
</tbody>
</table>

Table 2 SULT1A1 genotype and age of onset of breast cancer: mean age of onset by genotype (IA*1/IA*1, IA*1/IA*2, and IA*2/IA*2) and sample (MGH, DFCI, and Other) n indicates the number of patients. Mean indicates average age.

<table>
<thead>
<tr>
<th>Sample</th>
<th>IA<em>1+IA</em>1</th>
<th>IA<em>1+IA</em>2</th>
<th>IA<em>2+IA</em>2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>MGH</td>
<td>139</td>
<td>35.79</td>
<td>3.77</td>
</tr>
<tr>
<td>DFCI</td>
<td>54</td>
<td>40.30</td>
<td>9.95</td>
</tr>
<tr>
<td>Other</td>
<td>36</td>
<td>55.50</td>
<td>15.44</td>
</tr>
</tbody>
</table>
of the close proximity of the two genes (we were unable to determine conclusively if SULT1A1, SULT1A2, or Because of the high similarity of the relationship between SULT1A1 polymorphism and breast cancer risk. ZR75-1 breast cancer cells, together with the known functionally recent analysis of gene expression profiles of tamoxifen-treated as tamoxifen, thereby affecting their activity (18, 19). The identification of SULT1A1 as a potential downstream target of tamoxifen in our recent analysis of gene expression profiles of tamoxifen-treated ZR75-1 breast cancer cells, together with the known functionally relevant polymorphism in SULT1A1, prompted us to investigate the relationship between SULT1A1 polymorphism and breast cancer risk. Because of the high similarity of the SULT1A1 and SULT1A2 gene, we were unable to determine conclusively if SULT1A1, SULT1A2, or both mRNAs are affected by hormonal treatment. However, because of the close proximity of the two genes (−45 kb), alleles for SULT1A1 are in linkage disequilibrium with alleles for SULT1A2 (8, 13). Moreover, high activity alleles for SULT1A1 are linked to high activity alleles for SULT1A2; the same is true for the low activity alleles (13). In addition, the enzymatic activity of SULT1A2 is much lower than that of SULT1A1 (13). On the basis of all of this data, if SULT1A1 enzymatic activity influences breast cancer risk, analysis of SULT1A1 polymorphism in relation to breast cancer is likely to be informative and is not likely to be confounded by the potential effect of the SULT1A2 gene, although this possibility cannot be completely excluded.

Our analysis of 444 breast cancer patients and 227 controls revealed no evidence of increased risk of breast cancer associated with any SULT1A1 genotype. The control group consisted of 129 healthy blood donors, ethnically and age matched to the MGH patient set. An additional 98 controls were healthy blood donors with no data available on age and ethnicity. However, on the basis of the average patient population of the institutions where the samples were collected, we have no reason to believe that the patient and control groups would be significantly different.

Interestingly, we did find evidence that SULT1A1 genotype influences the age of onset of the disease in patients preselected for early-onset breast cancer (MGH and DFCI) but not in unselected (Other) cases. Because these early-onset breast cancer patients may have a genetic predisposition to breast cancer attributable to a high-penetrance gene or to other low-penetrance genes, the influence of the SULT1A1 polymorphism on the age of onset may indicate a genetic interaction between SULT1A1 genotype and other breast cancer susceptibility genes. Similarly, the intriguing association between SULT1A1 genotype and the presence of multiple different tumors in the same patient may suggest such an interaction between SULT1A1 and a higher penetrance cancer susceptibility gene(s). Because these other cancer types were of diverse origin, SULT1A1 may modify the effect of several different cancer-predisposing genes, an observation that is likely to stimulate additional studies.

This type of genetic interaction could explain the differing effect the SULT1A1*1 allele has on the age of onset in the MGH and DFCI sample sets. In the case of the MGH patients (age of onset <40), the SULT1A1*1 allele behaves as a dominant allele with both SULT1A1*1 homozygotes and heterozygotes having an earlier onset. Conversely, in the DFCI patients (<65 years of age), it behaves as a recessive allele with only the SULT1A1*1 homozygotes demonstrating an earlier onset. Although this difference between the two samples can be attributable to various exogenous (exposure to carcinogens) or endogenous (estrogen levels and others) factors as well, at this point we have no data available to differentiate between the above-mentioned possibilities.

Currently, we do not know the mechanism through which SULT1A1 activity influences the development of breast cancer and other cancer types. One hypothesis is that certain environmental agents (alkylphenols, octylephenol, bisphenol A, and others), either by interfering with the metabolism of endogenous steroid hormones or by becoming mutagenic upon sulfonation, increase the probability of acquiring a mutation in oncogenes or tumor suppressor genes. It is notable in this respect that phenolic compounds have been shown to influence sexual development and reproductive function in lower vertebrates. In addition, several animal and in vitro studies have found an association between high sulfotransferase activity and the risk of developing chemically induced cancers (15).

Because of the high frequency of the high activity SULT1A1 allele in Caucasian populations and the large number of xenobiotics sulfonated by this enzyme, SULT1A1 may turn out to be an important low-penetrance cancer-predisposing gene. The possible biological interaction of breast cancer preventive agents such as tamoxifen with SULT1A1 genotype merits additional attention because of their widespread use in high-risk patients.

Acknowledgments

We thank Bert Vogelstein, William Sellers, and Ian Krop for their critical review of the manuscript, and we thank the patients who participated in this study.

References
