Concordance of metabolic enzyme genotypes assayed from paraffin-embedded, formalin-fixed breast tumors and normal lymphatic tissue

Thomas P Ahern1
Mariann Christensen2
Deirdre P Cronin-Fenton3
Kathryn L Lunetta4
Carol L Rosenberg5
Henrik Toft Sørensen3
Timothy L Lash3
Stephen Hamilton-Dutoit2

1Department of Epidemiology, Boston University School of Public Health, Boston, Massachusetts; 2Institute of Pathology; 3Department of Clinical Epidemiology, Aarhus University Hospital, Aarhus, Denmark; 4Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts; 5Hematology/Oncology Section, Department of Medicine, Boston University School of Medicine, Boston, Massachusetts, USA

Objectives: Translational epidemiology studies often use archived tumor specimens to evaluate genetic hypotheses involving cancer outcomes. When the exposure of interest is a germline polymorphism, a key concern is whether the genotype assayed from tumor-derived DNA is representative of the germline. We evaluated the concordance between breast tumor-derived and normal lymph node-derived genotypes for three polymorphic tamoxifen-metabolizing enzymes.

Methods: We assayed paired DNA samples extracted from archived tumor and normal lymph node tissues from 106 breast cancer patients. We used TaqMan assays to determine the genotypes of three enzyme variants hypothesized to modify tamoxifen effectiveness, ie, CYP2D6*4, UGT2B15*2, and UGT1A8*2. We assessed genotype agreement between the two DNA sources by calculating the percent agreement and the weighted kappa statistic.

Results: We successfully obtained genotypes for CYP2D6*4, UGT2B15*2, and UGT1A8*2 in 99%, 100%, and 84% of the paired samples, respectively. Genotype concordance was perfect for the CYP2D6*4 and UGT1A8*2 variants (weighted kappa for both = 1.00; 95% confidence interval [CI] 1.00, 1.00). For UGT2B15*2, one pair out of 106 gave a discordant result that persisted over several assay repeats.

Conclusions: We observed strong agreement between DNA from breast tumors and normal lymphatic tissue in the genotyping of polymorphisms in three tamoxifen-metabolizing enzymes. Genotyping DNA extracted from tumor tissue avoids the time-consuming practice of microdissecting adjacent normal tissue when other normal tissue sources are not available. Therefore, the demonstrated reliability of tumor-derived DNA allows resources to be spent instead on increasing sample size or the number of polymorphisms examined.

Keywords: molecular epidemiology, breast neoplasms, cytochrome P450 CYP2D6, glucuronosyltransferase

Introduction

Translational clinical epidemiology studies frequently use tissue archives as a resource for bioassays of cellular and molecular features, such as germline genotype or protein expression in a tumor. The utility of archived tissues is tempered by inherent weaknesses of stored biologic material, for example, degradation of macromolecules over time, or the questionable representativeness of a tissue sample to a whole tumor or to a person.1–4 A specific example of the latter limitation is the use of archived breast tumors to assay germline genotypes of drug-metabolizing enzymes, which may modify the effectiveness of adjuvant therapies in terms of disease recurrence and survival.5 In such studies, it is uncommon to have available samples from more reliable sources of germline DNA, such as buccal cells or peripheral blood, and it is...
therefore necessary to assume that a tumor-derived genotype is representative of a patient’s germline. However, there is cause for skepticism over this assumption, owing to a variety of potential DNA alterations in tumors that may interfere with accurate genotyping.

Earlier studies have explored genotype concordance between reliable sources of germline DNA and breast tumor or tumor-adjacent tissue. Schneider et al compared genotype agreement for a panel of angiogenesis gene polymorphisms in 17 breast tumor samples with paired lymph node samples. Xie et al compared genotype agreement for five genes at loci with high rates of loss of heterozygosity in paired DNA samples from peripheral blood and microdissected breast tumor-adjacent normal tissue in 106 subjects. While both of these studies demonstrated 100% concordance between the different DNA sources, the use of tumor-adjacent normal tissue by Xie et al and the small sample size enrolled by Schneider et al leave open the question of whether DNA derived from formalin-fixed, paraffin-embedded (FFPE) breast tumor sections provides a reliable medium for the assay of germline genotype. Here we report the results from a large study of the concordance between genotypes of three tamoxifen-metabolizing enzymes assayed from archived FFPE breast carcinoma and normal lymph node tissues.

Materials and methods

Study population

This study was approved by the Boston University Medical Campus Institutional Review Board and the Regional Committee on Biomedical Research Ethics of Aarhus County, Denmark. We conducted the study among women aged 35–70 years who were diagnosed with incident invasive breast carcinoma, ie, Union for International Cancer Control (UICC) Stage I, II, or III, while living in one of seven counties in western Denmark between 1985 and 2001. The study sample was identified using the Danish Breast Cancer Cooperative Group registry to study associations between polymorphisms in genes encoding drug-metabolizing enzymes and tamoxifen resistance. Cases were defined as women with local or distant breast cancer recurrence. One recurrence-free control was incidence-density sampled and matched to each case for estrogen receptor expression, receipt of tamoxifen therapy, menopausal status at diagnosis (pre- or postmenopausal), date of breast cancer surgery (caliper-matched ± 12 months), county of residence at diagnosis, and UICC stage at diagnosis (Stage I, II, or III). For each case and control, FFPE primary breast tumors were requested from the pathology departments at the admitting hospitals. Cancer-free FFPE lymph node tissue resected during primary surgical treatment was also requested for cases and controls when such tissue was available. For the present study, we aimed to enroll approximately 100 subjects for whom both tumor and normal tissues were received from pathology departments.

Tissue processing

Hematoxylin and eosin stained sections or written pathologic descriptions of the tumor and lymph node tissue blocks were reviewed by a pathologist to identify blocks for processing. Tissue blocks were processed in a laminar flow hood that had been sterilized with at least one hour of ultraviolet light exposure before use. Contaminating nucleases were removed by wiping all work surfaces and instruments with an aqueous solution of 1% (w/v) sodium dodecyl sulfate, followed by 99.9% ethanol. Several sections were cut from each tissue block and discarded to ensure a clean starting surface for sample procurement. For DNA extractions, three to six 10 µm pieces were cut from each lymph node and tumor tissue block and placed in a 1.5 mL microtube. The cutting knife was wiped clean with the sodium dodecyl sulfate/ethanol series after each block, and the blade was replaced after every two blocks. Each blade was long enough to permit the cutting of two blocks without cross-contamination. Laboratory personnel changed gloves between blocks, and cut a pure paraffin control block after every 10 tissue blocks to serve as a contamination checkpoint.

DNA extraction

Before DNA extraction, all tissue samples were deparaffinized by treatment with xylene; 1 mL of xylene was added to each microtube containing three to six slices of paraffin-embedded tissue and the microtubes were placed in a shaking incubator at 45°C and 800 rpm for 20 minutes. Samples were then centrifuged at 12,200 rpm for five minutes, after which the supernatant was removed and discarded. The xylene treatment steps were repeated once for all samples. After deparaffinization, samples were incubated with 1 mL of 99% ethanol in a shaking incubator at 60°C and 800 rpm for 20 minutes, centrifuged at 12,200 rpm for five minutes, with subsequent removal of supernatant; ethanol treatment was repeated once for all samples. One drop of acetone was added to each tube, and tissue samples were evaporated to dryness at 60°C. Tissues were then allowed to dissolve overnight in a shaking incubator at 55°C and 1000 rpm after addition of 150 µL proteinase K (10 mg/mL) solution. The following morning, proteinase K was inactivated by a 20-minute treatment with 1 mL of 4 M guanidine thiocyanate and 0.4 M calcium chloride.
incubation at 98°C. DNA was then extracted from the tissue samples using a robotic magnet-assisted nucleic acid isolation instrument (MagnaPure, Roche Applied Science).

Genotyping
From each tissue sample, 50 ng of extracted DNA were amplified in 25 µL polymerase chain reactions with 50 denaturation cycles at 92°C for 15 seconds, followed by annealing and extension at 60°C for 90 seconds, using primers and reagents supplied with TaqMan genotyping kits (Applied BioSystems, Foster City, CA). We compared normal lymph node tissue with tumor tissue on genotypes of three metabolic enzyme single-nucleotide polymorphisms (SNPs), ie, cytochrome P450 2D6 (CYP2D6)*4 (rs3892097, ABI kit: C-27102431-D0); UDP-glucuronosyltransferase 2B15 (UGT2B15)*2 (rs1902023, ABI kit: C-27028164-10); and UDP-glucuronosyltransferase 1A8 (UGT1A8)*2 (rs1042597, ABI kit: C-11742072-10). Additional data on these SNPs are shown in Table 1. These genes were selected for analysis because each plays a key role in the phase I metabolism of tamoxifen, and it is hypothesized that functional polymorphisms at these loci may modify the effectiveness of adjuvant tamoxifen therapy.5,13–17 All sample pairs were assayed in duplicate using the MX3000P real-time polymerase chain reaction instrument (MagnaPure, Roche Applied Science).

Table 1 Summary data for the SNPs analyzed for tumor/normal tissue genotype concordance

<table>
<thead>
<tr>
<th>Gene/variant</th>
<th>Reference SNP</th>
<th>Location</th>
<th>Sequence [SNP]</th>
<th>Expected allele frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cytchrome P450 2D6/ CYP2D6*4</td>
<td>rs3892097</td>
<td>22q13.1</td>
<td>CCCCTTACCCGCATCTCCCACCCCCCA [A/G]</td>
<td>A: 0.763</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GACGCCCTTTTGGCCCACACGGTCT</td>
<td>G: 0.237*</td>
</tr>
<tr>
<td>UDP-glucuronosyltransferase 2B15/ UGT2B15*2</td>
<td>rs1902023</td>
<td>4q13</td>
<td>TTATCCTACATCTTTAATGAAAAAT [G/T]</td>
<td>G: 0.560</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A T G G A G A T T C T CT G A A T A T</td>
<td>T: 0.440b</td>
</tr>
<tr>
<td>UDP-glucuronosyltransferase 1A8/ UGT1A8*4</td>
<td>rs1042597</td>
<td>2q37</td>
<td>CTCGTTGGTCTTGGCCAGGGGGAATAG [C/G]</td>
<td>C: 0.714</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>T T G C C A C T AT T G A A G A G T G C A</td>
<td>G: 0.286†</td>
</tr>
</tbody>
</table>


Results
There were 106 paired normal lymph node and breast tumor tissue samples from the individual patients included in this study. The distribution of the paired samples according to key clinical characteristics is reported in Table 2.
We successfully assayed \( CYP2D6^*4 \) genotype in 105 pairs (99%), \( UGT1A8^*2 \) genotype in 89 pairs (84%), and \( UGT2B15^*2 \) genotype in 106 pairs (100%). There was no difference between samples that genotyped successfully and those that failed to genotype successfully with regard to DNA purity (\( A_{260}/A_{280} \) measurement), DNA concentration, length of amplified sequence, or age of the tissue block at the time of DNA extraction. Our main analyses are based on sample pairs with successfully assayed genotypes from both tissue sources.

Table 3 reports the crosstabulation of genotype frequencies according to lymph node- and tumor-derived DNA, the kappa statistic for agreement between the two tissue types, and results from Chi-square tests for Hardy–Weinberg equilibrium based on genotype frequencies from the normal tissue samples. For all three genes, allele frequencies observed in our study sample agreed with values reported in similar populations (Tables 1 and 3). Wild-type and variant allele frequencies for all three genes did not deviate from Hardy–Weinberg equilibrium \( (P \geq 0.11 \text{ for all Chi-square tests}) \). Genotyping concordance was perfect for the \( CYP2D6^*4 \) and \( UGT1A8^*2 \) variants (for both, weighted kappa = 1.00; 95% CI 1.00, 1.00). Among the 106 pairs with a successfully assayed \( UGT2B15^*2 \) genotype, one pair (0.9%) gave a discordant result. The normal lymph tissue indicated a homozygous variant genotype and the tumor tissue indicated a homozygous wild-type genotype (weighted kappa = 0.97; 95% CI 0.91, 1.00). This outcome was consistent across four repeated assays on the discordant pair, even after re-extraction of DNA from the source tissue blocks.

### Discussion

Our study demonstrated high concordance between breast tumor- and normal lymph node-derived genotypes for key SNPs in drug-metabolizing enzymes, and agrees with observations from similar earlier studies. Previous work by Rae et al showed 100% concordance between \( CYP2D6 \) genotypes from 10 FFPE archival breast tumors and matched peripheral blood samples.\(^6\) Schneider et al demonstrated 100% concordance between 17 breast tumor and lymph node samples for polymorphisms in two angiogenesis genes.\(^11\) Xie et al reported 100% concordance for five genes from different high loss of heterozygosity sites in 106 paired samples of peripheral blood and microdissected normal tissue adjacent to breast tumor tissue.\(^12\) Our study provides evidence that breast tumor-derived genotypes are an acceptable proxy for germline genotypes when more suitable DNA sources are not available. This finding agrees with conclusions from two earlier reviews on this topic,\(^21,22\) and with results from other concordance studies using colorectal\(^23\) and non-small cell lung\(^24\) tumor DNA.

One of our target genes, \( CYP2D6 \), is located in a chromosomal region (22q13.1) that is commonly deleted in breast cancer cells.\(^8\) Thus, SNPs in \( CYP2D6 \) assayed from tumor-derived DNA are at risk of misclassification due to loss of heterozygosity, yielding the expectation of imperfect genotype concordance between normal and tumor-derived DNA. The other SNPs lie in chromosomal regions that experience little or no deletion, so there is less expectation of imperfect concordance between normal and tumor-derived DNA for them. The perfect concordance we observed between \( CYP2D6 \) genotypes from the two different tissue types indicates either that loss of heterozygosity does not meaningfully distort \( CYP2D6 \) genotype classification in breast tumor cells, or that despite substantial loss of heterozygosity, tumor sections contain sufficient stromal or adjacent normal tissue to report an accurate germline genotype. The perfect concordance observed by Xie et al in five genes from high loss of heterozygosity sites also supports these notions, providing reassurance that loss of heterozygosity is not a major threat.
Breast cancer genotype concordance to genotype misclassification when breast tumor-derived DNA must be relied upon.

A limitation of our study is that our standard for germline genotype was DNA extracted from FFPE normal lymphatic tissue instead of from fresh nonmalignant tissue. Earlier work by Rae et al showed 100% genotype concordance between DNA from freshly harvested tumor cell cultures and DNA from tumor cell pellets that were formalin-fixed and embedded in paraffin before DNA extraction.6 Therefore, we consider our FFPE normal lymphatic tissue to be an accurate standard for the germline genotype that would be observed using fresh tissue.

In summary, we observed excellent agreement between archived breast tumor- and normal lymphatic tissue-derived DNA in classifying the germline genotype of three drug-metabolizing enzyme variants (CYP2D6*4, UGT1A8*2, and UGT2B15*2) in sample sizes ranging from 89 to 106. Only one of the assayed gene variants (UGT2B15*2) gave a discrepant result, although in only one of 106 tested pairs. The discrepant result was a shift from homozygous wild-type in normal lymphatic tissue to a homozygous variant in the breast tumor tissue, in a chromosomal region not typically deleted in breast cancer.8 Therefore, while it is possible that the discrepancy arose from a tumor genome alteration, it is more plausible that a clerical error during either tissue archiving or processing resulted in the pairing of tumor and lymphatic tissues from separate patients.

Together, our observations indicate that FFPE archived breast tumors provide a reliable resource for the determination of germline genotypes in CYP2D6, UGT2B15, UGT1A8, and likely other drug-metabolizing enzymes. It remains possible that mutations at other loci on the tumor genome could yield poorer concordance proportions than those observed for the three metabolic enzyme variants we studied. Our finding is important because genotyping DNA extracted from tumor tissue avoids the time-consuming practice of microdissection of adjacent normal tissue when other sources of normal tissue (eg, peripheral blood or lymph node) are not available. Study resources can then be devoted more cost-efficiently to, for example, increase sample size or genotyping capacity.

Acknowledgments
The authors are most grateful to Kristina L Lauridsen, Jesper Bertelsen, and Inge K Krohn for their outstanding technical assistance. This research was supported by grants from the US National Cancer Institute (R01 CA118708), the Danish Cancer Society (DP06117), and the Karen Elise Jensen Foundation. Mr Ahern’s effort on this study was supported by a CDMRP predoctoral training award (BC073012).

Disclosure
The Department of Clinical Epidemiology at Aarhus University Hospital is involved in studies with funding from various companies as research grants to, and administered by, Aarhus University. None of these studies is related to the present study or supported any of the work reported herein. No author has any other potential competing interest.

Table 3 Cross-tabulation of genotypes determined using DNA extracted from formalin-fixed, paraffin-embedded tumor and normal tissue from Danish breast cancer patients

<table>
<thead>
<tr>
<th>Gene variant (Hardy–Weinberg P value)</th>
<th>Tumor tissue genotypes</th>
<th>Normal tissue genotypes</th>
<th>Concordance Weighted $\kappa$ (95% CI)</th>
<th>Agreement % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>wt/wt</td>
<td>wt/var</td>
<td>var/var</td>
<td></td>
</tr>
<tr>
<td>CYP2D6*4 ($P = 0.46$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>wt/wt</td>
<td>66</td>
<td>0</td>
<td>1.00 (1.00, 1.00)</td>
</tr>
<tr>
<td></td>
<td>wt/var</td>
<td>0</td>
<td>36</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>var/var</td>
<td>0</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal tissue allele frequencies</td>
<td>G: 0.80</td>
<td>A: 0.20</td>
<td></td>
</tr>
<tr>
<td>UGT2B15*2 ($P = 0.11$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>wt/wt</td>
<td>25</td>
<td>0</td>
<td>1.00 (0.91, 1.00)</td>
</tr>
<tr>
<td></td>
<td>wt/var</td>
<td>0</td>
<td>61</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>var/var</td>
<td>0</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal tissue allele frequencies</td>
<td>G: 0.52</td>
<td>T: 0.48</td>
<td></td>
</tr>
<tr>
<td>UGT1A8*2 ($P = 0.22$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>wt/wt</td>
<td>47</td>
<td>0</td>
<td>1.00 (1.00, 1.00)</td>
</tr>
<tr>
<td></td>
<td>wt/var</td>
<td>0</td>
<td>32</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>var/var</td>
<td>0</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Normal tissue allele frequencies</td>
<td>G: 0.71</td>
<td>A: 0.29</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: wt, wild-type; var, variant; CI, confidence interval.
References


