

## ARTICLE

# Loss of Heterozygosity at the CYP2D6 Locus in Breast Cancer: Implications for Germline Pharmacogenetic Studies

Matthew P. Goetz\*, James X. Sun\*, Vera J. Suman, Grace O. Silva, Charles M. Perou, Yusuke Nakamura, Nancy J. Cox, Philip J. Stephens, Vincent A. Miller, Jeffrey S. Ross, David Chen, Stephanie L. Safgren, Mary J. Kuffel, Matthew M. Ames, Krishna R. Kalari, Henry L. Gomez, Ana M. Gonzalez-Angulo, Octavio Burgues, Hiltrud B. Brauch, James N. Ingle, Mark J. Ratain, Roman Yelensky

Department of Oncology (MPG, MMA, JNI), Department of Health Sciences Research (VJS, KRK), and Department of Molecular Pharmacology and Experimental Therapeutics (MPG, SLS, MK, MMA), Mayo Clinic, Rochester, MN; Department of Genetics, University of North Carolina, Chapel Hill, NC (GOS, CMP); Center for Personalized Therapeutics, University of Chicago, Chicago, IL (YN, NJC, MJR); Foundation Medicine Inc., Cambridge, MA (JXS, PJS, VAM, JSR, RY); Novartis Pharmaceuticals Corporation, East Hanover, NJ (DC); Department of Pathology and Laboratory Medicine, Albany Medical College, Albany, NY (JSR); Division of Hematology/Oncology, Vanderbilt University, Nashville, TN (HG); Department of Breast Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, TX (AMGA, OB); Department Breast Cancer Susceptibility and Pharmacogenomics, Dr Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, University Tuebingen, German Cancer Consortium (DKTK) and German Cancer Research (DKTK), Heidelberg, Germany (HB).

Current affiliation of H. Gomez: Departamento de Medicina Oncologica, Instituto Nacional de Enfermedades Neoplasicas, Lima, Peru.

\*Authors contributed equally to this work.

Correspondence to: Matthew Goetz, MD, Oncology, Mayo Clinic, 200 First Street SW, Rochester, MN 55905 (e-mail: [goetz.matthew@mayo.edu](mailto:goetz.matthew@mayo.edu)).

## Abstract

**Background:** Controversy exists regarding the impact of CYP2D6 genotype on tamoxifen responsiveness. We examined loss of heterozygosity (LOH) at the CYP2D6 locus and determined its impact on genotyping error when tumor tissue is used as a DNA source.

**Methods:** Genomic tumor data from the adjuvant and metastatic settings (The Cancer Genome Atlas [TCGA] and Foundation Medicine [FM]) were analyzed to characterize the impact of CYP2D6 copy number alterations (CNAs) and LOH on Hardy Weinberg equilibrium (HWE). Additionally, we analyzed CYP2D6 \*4 genotype from formalin-fixed paraffin-embedded (FFPE) tumor blocks containing nonmalignant tissue and buccal (germline) samples from

Received: April 1, 2014; Revised: July 25, 2014; Accepted: October 30, 2014

© The Author 2014. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: [journals.permissions@oup.com](mailto:journals.permissions@oup.com).

patients on the North Central Cancer Treatment Group (NCCTG) 89-30-52 tamoxifen trial. All statistical tests were two-sided.

**Results:** In TCGA samples ( $n = 627$ ), the CYP2D6 LOH rate was similar in estrogen receptor (ER)-positive (41.2%) and ER-negative (35.2%) but lower in HER2-positive tumors (15.1%) ( $P < .001$ ). In FM ER+ samples ( $n = 290$ ), similar LOH rates were observed (40.8%). In 190 NCCTG samples, the agreement between CYP2D6 genotypes derived from FFPE tumors and FFPE tumors containing nonmalignant tissue was moderate (weighted Kappa = 0.74; 95% CI = 0.63 to 0.84). Comparing CYP2D6 genotypes derived from buccal cells to FFPE tumor DNA, CYP2D6\*4 genotype was discordant in six of 31 (19.4%). In contrast, there was no disagreement between CYP2D6 genotypes derived from buccal cells with FFPE tumors containing nonmalignant tissue.

**Conclusions:** LOH at the CYP2D6 locus is common in breast cancer, resulting in potential misclassification of germline CYP2D6 genotypes. Tumor DNA should not be used to determine germline CYP2D6 genotype without sensitive techniques to detect low frequency alleles and quality control procedures appropriate for somatic DNA.

The CYP2D6 enzyme metabolizes tamoxifen to its active metabolites (4-hydroxy-tamoxifen and 4-hydroxy-N-desmethyl-tamoxifen [endoxifen]), and numerous studies have demonstrated that CYP2D6 genetic variants are associated with steady state endoxifen concentrations (1–2). However, there is substantial controversy on the validity of CYP2D6 genotype as a predictor of benefit from tamoxifen therapy in the adjuvant setting (reviewed in [3]). Secondary analyses of adjuvant trials administering five years of tamoxifen (the North Central Cancer Treatment Group [NCCTG] 89-30-52 [4], Arimidex, tamoxifen, alone or in combination (ATAC) [5], BIG1-98 [6], and the Austrian Breast and Colorectal Cancer Study Group [ABCSCG] 8 [7] have reached discrepant conclusions). Multiple investigators have voiced concern regarding the unprecedented departure of CYP2D6 allele frequencies from Hardy-Weinberg equilibrium (HWE) in the BIG 1-98 study (8–10). While substantial departure from HWE was not observed in the ABCSCG 8 analysis (7), some departure from HWE was observed with the CYP2D6\*4 allele frequencies reported in the NCCTG 89-39-52 (4) and ATAC (5,9) CYP2D6 analyses. Given previous demonstration of genomic instability at the chromosomal segment where CYP2D6 is located (11–12), it has been hypothesized that when tumor DNA is used for genotyping, the presence of tumor loss of heterozygosity (LOH) at the CYP2D6 locus distorts the frequencies of observed alleles, which could lead to an excessive homozygous assignment of the germline genotype (8–10). To address this question, we undertook a detailed evaluation of whether somatic LOH occurs at the CYP2D6 locus by analyzing genomic tumor data from the adjuvant (The Cancer Genome Atlas [TCGA]) (13) and metastatic settings. Furthermore, we sought to determine whether CYP2D6 LOH could affect the accuracy of calling germline CYP2D6 genotypes. Finally, in a limited number of adjuvant cases in which both formalin-fixed paraffin-embedded (FFPE) tumor blocks and buccal samples were available, we compared CYP2D6 \*4 genotypes obtained from each DNA source.

## Methods

### Samples

Three previously published data sets were analyzed. The first data set included tumors collected and annotated within The Cancer Genome Atlas breast dataset (13). TCGA collected breast tumors from newly diagnosed patients who underwent surgical resection. Extensive quality control was employed to verify the presence of both tumor DNA and germline DNA. Briefly, each frozen primary tumor specimen had a companion normal tissue DNA specimen that was derived from blood components (including DNA extracted at the tissue source site) ( $n = 684$ ), adjacent

normal tissue taken from greater than 2cm from the tumor ( $n = 76$ ), or both ( $n = 65$ ). Each hematoxylin and eosin (H&E) stained case was reviewed by a board-certified pathologist to confirm that the tumor specimen was histologically consistent with breast adenocarcinoma and the adjacent normal specimen contained no tumor cells. The tumor sections were required to contain an average of 60% of tumor cell nuclei with less than 20% necrosis for inclusion in the study per TCGA protocol requirements. The clinical characteristics of this cohort and the process for informed consent have been previously described (13).

The second set included paraffin-embedded blocks from 360 patients, with relapsed and metastatic ER+ ( $n = 261$ ) or ER- ( $n = 99$ ) breast cancers derived from a subset of patients from the NCT00780676 trial and from pathology departments of several medical centers, as recently described (14). From these samples, CYP2D6 sequencing was performed by Foundation Medicine (FM). In addition, samples were stained for ER, progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER2) and reviewed by a pathologist to confirm ER positivity. All tissue collections were done with the approval of the corresponding institutional review boards, and the process for informed consent was previously published (14).

The third set included specimens from 190 ER-positive breast cancer case patients from the NCCTG 89-30-52 clinical trial (4). In the initial reported CYP2D6 analysis, an H&E section was obtained from FFPE tumors and a board-certified pathologist identified the invasive component and DNA was extracted from a 1cm area of highest tumor cellularity for both DNA (4) and RNA (15) studies. At a later date, the same tissue block was accessed and whole tissue sections containing both invasive and benign tissue were processed for DNA extraction as previously described (16–17). Additionally, germline DNA from a buccal sample was collected and reported initially on 17 patients (4) and an additional 21 patients later provided buccal samples. All tissue collections were done with the approval of the corresponding institutional review boards, and the process for informed consent was previously published (4).

### Genomic Analysis

For the TCGA cohort, DNA copy number at the CYP2D6 locus (Chr. 22: 42522 501 – 42525 911) was determined using the Affymetrix 6.0 single-nucleotide polymorphism (SNP) arrays (13) and copy number segmentation was performed using the Circular Binary Segmentation (CBS) algorithm version 1.12.0, as previously described (13). Copy number segments of interest were identified as regions with intensity values greater than  $|0.3|$ . Frequency landscape plots of these segments were created

using the SWITCHdna R-package plotting function (18). Exome sequencing was performed as previously described (13). Regions of LOH were identified using the Broad Institute's ABSOLUTE method on exome sequencing data and Affymetrix 6.0 SNP arrays (19). LOH landscape frequency plots were created using modifications of SWITCHdna's plotting function. The percentage of overlap between breast TCGA samples analyzed on SNP arrays and those through exome sequencing was 86%.

For the FM cohort, genomic DNA was extracted from 40  $\mu$ m of FFPE tissue and up to 200ng of extracted DNA was sheared by sonication, followed by ligation of Illumina sequencing adaptors. Sequencing libraries were hybridization captured using RNA-based baits (Agilent), targeting a total of 3320 exons of 182 cancer-related genes and 78 polymorphisms in 34 ADME-related genes. Deep (>500x) paired-end sequencing (49 x 49 cycles) was performed using the HiSeq2000 (Illumina). Sequence reads were mapped to the reference human genome (hg19), analyzed for all classes of genomic alterations (substitutions, indels, and copy number alterations), using custom methods optimized for clinical tumor specimens with stromal admixture. Variant calls at the CYP2D6 locus were resolved into genotypes according to the star (\*) allele nomenclature (20). If the minor allele frequency was greater than 5%, the patient was considered to have germline heterozygosity. To determine tumor LOH at CYP2D6, a genome-wide copy number model was fitted to the coverage data at all sequenced exons and more than 1800 SNPs. This profile was segmented and interpreted using allele frequencies of sequenced SNPs to estimate tumor purity and copy number at each segment. Fitting was performed using Gibbs sampling, assigning total copy number and minor allele count to all segments. LOH was called if total copy number at the CYP2D6 locus was 1 (copy loss LOH), or if copy number was 2 or more with a minor allele count of 0 (copy neutral LOH). The distortion of the germline alternate allele frequency from 50% because of LOH is calculated. To assess the impact of LOH, we simulated low-sensitivity genotyping assays by requiring minor allele frequencies to have minimum levels of 10% and 20% before assigning genotypes as heterozygous. The estimate of potential error impact on genotyping methods was then estimated using the HWE test.

For the NCCTG samples, CYP2D6 genotyping (\*3, \*4, \*6, \*10, \*41) was performed at the Mayo Clinic using the Applied Biosystems' Taqman Allelic Discrimination Assay (Foster City, CA), as previously described and reported in the context of a pooled analysis of NCCTG and Stuttgart patients (16) and as submitted to the International Tamoxifen Pharmacogenomics Consortium (21). Analyses were performed irrespective of ethnicity.

## Statistical Methods

Within the TCGA, a Pearson's Chi Square Test was used to determine whether LOH rates differed across intrinsic subtypes. Within the FM cohort, a two-sided Fisher's exact test was used to assess whether copy loss rate differed with respect to ER status. Within the NCCTG cohort, the extent of agreement between CYP2D6 genotypes derived from FFPE tumor and FFPE tumors containing nonmalignant tissue was assessed using weighted Kappa statistics and the corresponding 95% confidence interval. HWE tests were calculated using an exact test (the Simple Hardy-Weinberg Calculator by Michael H Court) (<http://www.tufts.edu/~mcourt01/Documents/Court%20lab%20-%20HW%20calculator.xls>) by comparing the observed and expected genotype frequencies for case patients and control patients. All statistical tests were two-sided, and a P value of less than .05 was considered statistically significant.

## Results

### TCGA Samples

Using SNP array data (n = 728) (13), evaluation of the CYP2D6 locus at chromosome 22 demonstrated copy number alterations (CNA) in 29.0% (n = 211) (Figure 1A). Among the 627 case patients with exome sequencing data, 219 case patients (34.9%) had LOH at the CYP2D6 locus (Figure 1B). While the CNA were higher for the ER-positive (35.0%) (Figure 2A) compared with the ER-negative (12.0%) (Figure 2B), LOH rates were similar comparing ER-positive (41.2%) (Figure 3A) and ER-negative (35.2%) (Figure 3B). Analyzing according to intrinsic subtypes, LOH rates among the ER+ (luminal A [40.3%] luminal B [42.7%]) and basal-like (43.4%) subsets were similar but greater than that in the HER2-enriched subtype (15.1%) (P < .001, Pearson's Chi Square Test). For each of these subtypes, a "zoomed-in plot" of the region containing the CYP2D6 gene is indicated (Supplementary Figures 1 and 2, available online). A further analysis within the clinically defined HER2+ subset demonstrated that LOH rates were lower within the ER-/HER2+ (14.3%) compared with ER+/HER2+ (26.6%).

### Foundation Medicine Samples

The findings among the case patients comprising the FM cohort were similar to those from the TCGA cohort, where 82 of 201 (40.8%) and 23 of 89 (25.8%) of the ER+ and ER- case patients, respectively, had LOH at the CYP2D6 locus (Figure 4). While copy-neutral LOH was similar in both ER+ and ER- (18.9% and 19.1%, respectively), the copy loss rate among ER+ case patients was statistically significantly greater relative to ER- case patients (21.9% vs 6.7%; P = .001, two-sided Fisher's exact test).

Given that standard genotyping assays (eg, Taqman) may not be able to detect an allele that is present at low frequency because of LOH, CYP2D6 genotypes were determined using next generation sequencing (Table 1) and the potential effect of LOH on CYP2D6 genotype was assessed (Table 2). Among the 105 case patients with LOH, a substantial fraction had a low frequency of one of the germline alleles: under 20% (n = 27), under 10% (n = 7). If such samples were assumed to be homozygous, this would result in excessive number of homozygotes and, statistically, departure from HWE (Table 2).

### NCCTG 89-30-52 Samples

The original CYP2D6 \*4 genotyping results were derived from tumor FFPE (FFPE-T) and demonstrated departure from HWE (chi square = 16.1, P  $\leq$  .001) (4). These case patients (n = 190) were reassessed using FFPE sections containing nonmalignant tissue (FFPE-NM) (16). For CYP2D6 \*4, the agreement was moderate comparing CYP2D6 \*4 genotypes derived from FFP-T with FFPE-NM (weighted Kappa 0.74; 95% CI = 0.63 to 0.84), resulting in excess homozygous genotypes and departure from HWE (P < .001). Specifically, 15 original homozygous wild-type (Wt/Wt) cases were reclassified as heterozygous for \*4 (Wt/\*4) and three homozygous variant (\*4/\*4) were reclassified as (Wt/\*4). The \*4 discrepancies among the remaining five cases were likely unexplained by LOH (Table 3). An evaluation for HWE using the genotyping data derived from FFPE-NM demonstrated that CYP2D6\*4 is within HWE (chi square = 1.34, P = .25).

To further investigate the observed discrepancy between these results, the CYP2D6 genotypes derived from FFPE-T tumor (4) and FFPE-NM (16) were compared with CYP2D6\*4 genotype derived from buccal cells (germline). Among the 31 case patients

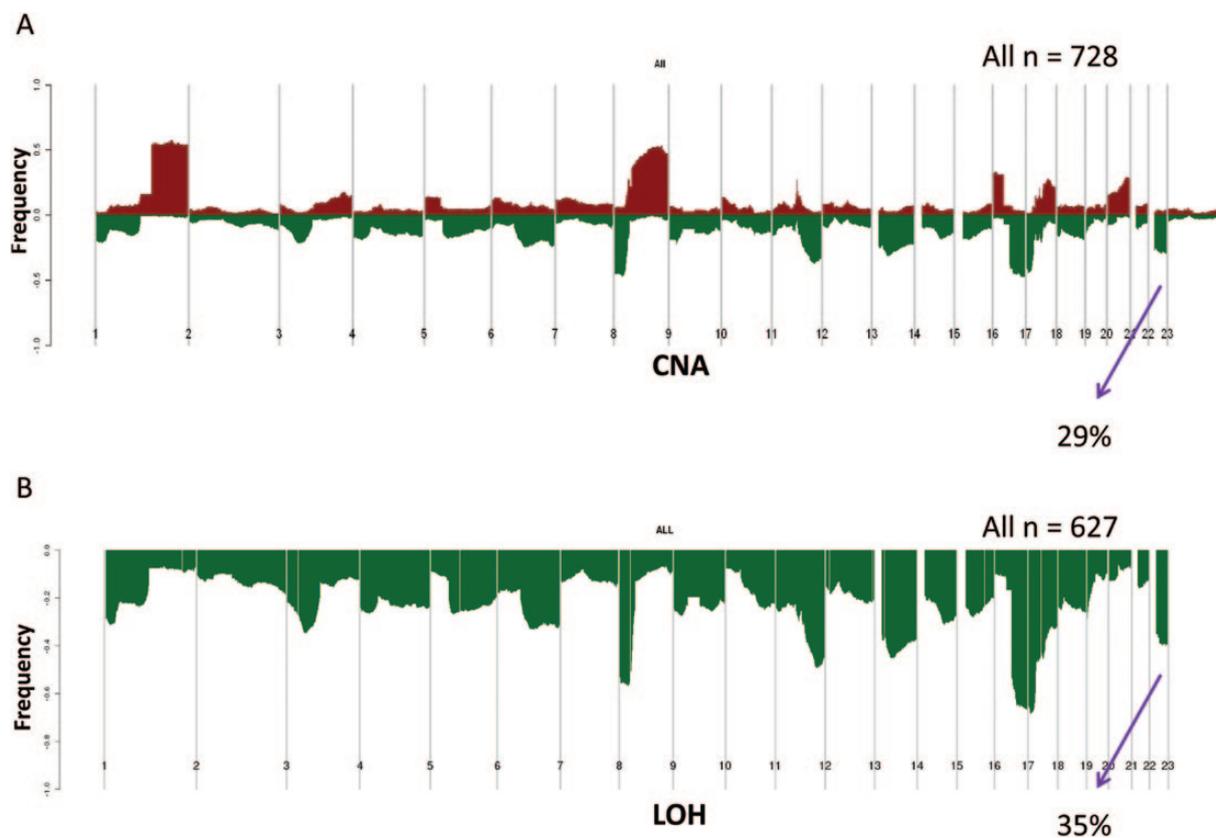


Figure 1. Cytochrome P450 2D6 (CYP2D6) copy number alterations (A) and loss of heterozygosity (B) within the entire Cancer Genome Atlas cohort. CNA = copy number alteration; LOH = loss of heterozygosity.

with both FFPE-T and buccal cells available for CYP2D6\*4 genotyping, there were six (19.4%) cases of disagreement. In four of these six case patients, CYP2D6 \*4 genotypes classified as homozygous wild-type using FFPE-T were determined to be heterozygous for \*4 (Wt/\*4) using DNA derived from buccal cells, and, in another case, a homozygous variant (\*4/\*4) based on FFPE-T was classified as (Wt/\*4) using DNA from buccal cells. One of the errors appeared to be unrelated to LOH, as the tumor-derived genotype of \*4/\*4 was classified as Wt/Wt using buccal cells. In contrast, among the 35 case patients with DNA from both FFPE-NM and buccal cells, there was 100% agreement comparing CYP2D6 \*4 genotypes from each source.

## Discussion

Using two large breast cancer datasets, we have demonstrated the presence of extensive LOH at the CYP2D6 locus in breast cancer. Furthermore, our data demonstrate that determination of germline CYP2D6 genotype using cancer tissue can result in substantial departure from HWE, as was seen in the original NCCTG CYP2D6 analysis (4), ATAC (5), and BIG 1-98 (6) studies. In the cohorts examined in this study, CYP2D6 genotyping using DNA extracted from FFPE-T blocks resulted in erroneous classification of up to 40% of CYP2D6\*4 heterozygotes (intermediate metabolizers) as either extensive metabolizers or poor metabolizers.

Recently, Rae et al., in a cohort of 122 patients, extracted DNA from three 0.6-mm diameter cores obtained from FFPE breast tumor blocks as well as DNA derived from either normal lymph nodes or leukocytes (22). Rae et al. used DNA from these sources to genotype for CYP2D6 and demonstrated a concordance rate of

over 94% between these different sources, concluding that this modest quality control study was sufficient to support the use of breast cancer tissue for germline genotyping of CYP2D6 (22). The results of our studies in this report clearly refute the conclusions of Rae and colleagues and provide further confirmation of the concerns raised by multiple authors (8–10) regarding the fidelity of the CYP2D6 genotyping performed in the context of the BIG 1-98 study (6).

Quality control procedures are critical for accurate genotyping. This includes a requirement to develop assays for all relevant variants, particularly for a locus as complex as CYP2D6 (23). An additional critical aspect of quality control relates to the source of DNA used for germline genotyping. In ATAC (5), FFPE tumor blocks from the trans-ATAC tumor collection were used for DNA extraction. In BIG 1-98 (6), DNA was extracted from one or two 1 mm cores that were punched into an area of the FFPE block most representative of the invasive tumor component.

Given our observation of LOH at the chromosomal locus containing CYP2D6, it was critical to understand whether the use of tumor DNA could contribute to the observed departures from HWE. In the FM cohort, nearly one-third of the tumors with LOH had a frequency of the germline allele under 20%, suggesting that use of a low-sensitivity polymerase chain reaction (PCR) assay could result in misclassification of heterozygous CYP2D6 genotypes as homozygous. Therefore, we directly compared CYP2D6 genotyping results from different laboratories using DNA from the same patients. In the original publication of the NCCTG 89-30-52 clinical trial, CYP2D6 genotyping (using DNA extracted from tumors) was performed in the laboratory of Rae et al. at the University of Michigan (4). When CYP2D6 genotyping was repeated at the Mayo Clinic using DNA derived from the same

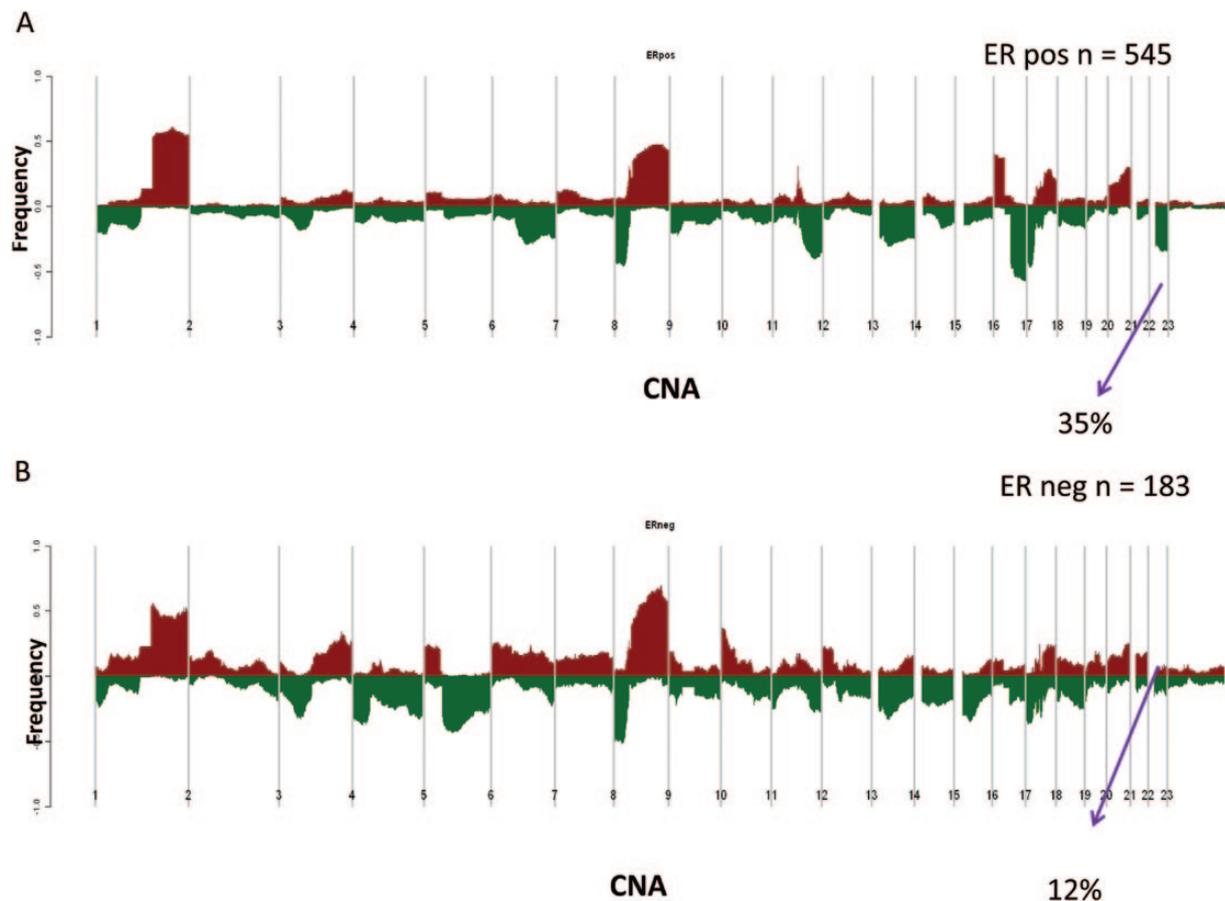


Figure 2. Cytochrome P450 2D6 (*CYP2D6*) copy number alterations within The Cancer Genome Atlas estrogen receptor (ER)-positive (A) and ER-negative (B) cohorts. CNA = copy number alteration; ER = estrogen receptor.

FFPE blocks but using whole tissue sections containing benign tissue, genotyping errors were identified, which appeared to be partially related to the lack of detection of low-frequency alleles in the 2005 analysis; however, additional discrepancies were observed that appear to be unrelated to LOH (Table 3). A full reanalysis of the NCCTG data set demonstrated that *CYP2D6* genotypes met HWE, with complete agreement (35/35) between the updated genotype results with the germline (buccal) cells in those patients that provided a buccal sample. Furthermore, as previously reported, *CYP2D6* genotype was statistically significantly associated with the risk of recurrence (16,21).

In ATAC (5), the departure from HWE with regard to *CYP2D6* \*4 was similar in magnitude as observed in the original NCCTG *CYP2D6* analysis (HWE  $\chi^2 = 18.1$ ,  $P = .000021$ ). While we are confident in our conclusions that LOH at the *CYP2D6* locus is common in breast cancer and that the use of tumor DNA for *CYP2D6* genotype results in misclassification of germline *CYP2D6* genotype, we were unable to reproduce the extreme departure from HWE observed in BIG 1-98 ( $P = 10^{-92}$ ) (6). Stanton noted that if LOH was the sole cause of deviation from HWE in BIG 1-98, the distorted genotype frequencies could be normalized by adjusting for LOH (9). Therefore, we agree with Stanton that the extreme departure from HWE in BIG 1-98 may be related to other factors, such as the use of nonstandard PCR techniques (use of upwards of 60 PCR cycles) (6).

Following the simultaneous publication of the *CYP2D6* analyses of the ATAC and BIG 1-98 data sets, the authors of these studies argued that testing for *CYP2D6* has no value in clinical practice, and an accompanying editorial concluded that this

matter can be likely laid to rest (24). However, our findings have validated the initial concerns raised by multiple investigators regarding genotyping error (8–10) and the conclusions that were generated based on these erroneous data. It is now clear that data from ongoing prospective clinical trials will be necessary to settle the debate on whether or not *CYP2D6* genotyping can identify patients in whom tamoxifen would be an inferior therapy. However, until such data are available, clinicians and patients should be aware of the data generated from secondary analyses of prospective clinical trials that support the importance of both *CYP2D6* genotype (7,16) and endoxifen concentrations (25) and that these data fulfill the basic criteria of Simon et al. for a “prospective-retrospective” design in which the biomarker test is analytically and preanalytically validated for use with archived tissue (26).

An important finding within the TCGA *CYP2D6* analysis was the observation of a substantially higher rate of LOH within the luminal A (40%), luminal B (43%), and basal-like subsets (40%), compared with the HER2-enriched (15%) and normal-like (8%) subtypes. Within the clinically defined HER2+ subset, LOH rates were lower within the ER-/HER2+ (14%) compared with ER+/HER2+ (27%). Within the FM cohort, the *CYP2D6* loss rate among ER+ case patients was statistically significantly greater relative to ER- case patients. While the biological relevance of these findings is unknown, the demonstration of substantial LOH at chromosome 22q13, the cytogenetic segment which contains the *CYP2D6* gene, has been implicated in breast (11), colon (11,27), and insulinomas (28), suggesting that a putative tumor suppressor gene in this region may be important in the pathogenesis of

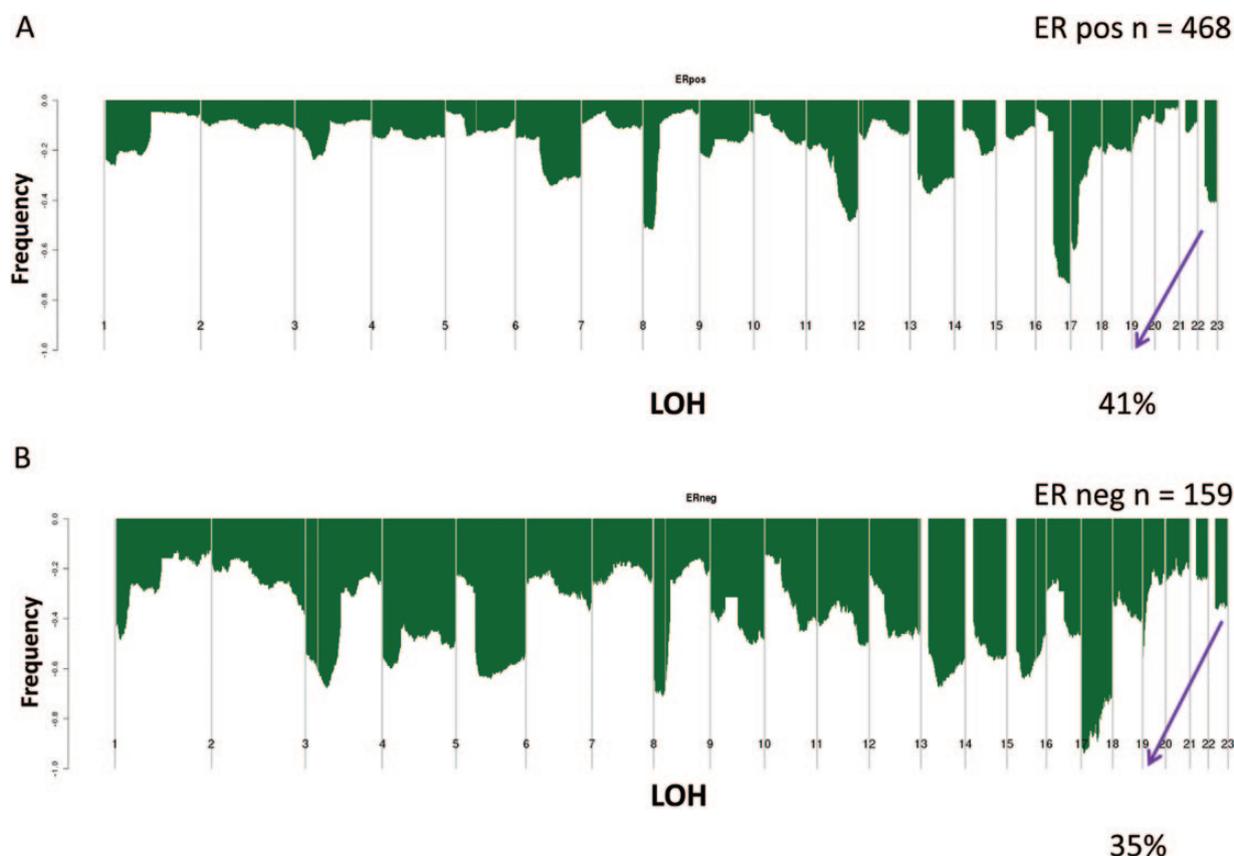


Figure 3. Cytochrome P450 2D6 (*CYP2D6*) loss of heterozygosity within The Cancer Genome Atlas estrogen receptor (ER)-positive (A) and ER-negative (B) cohorts. ER = estrogen receptor; LOH = loss of heterozygosity.

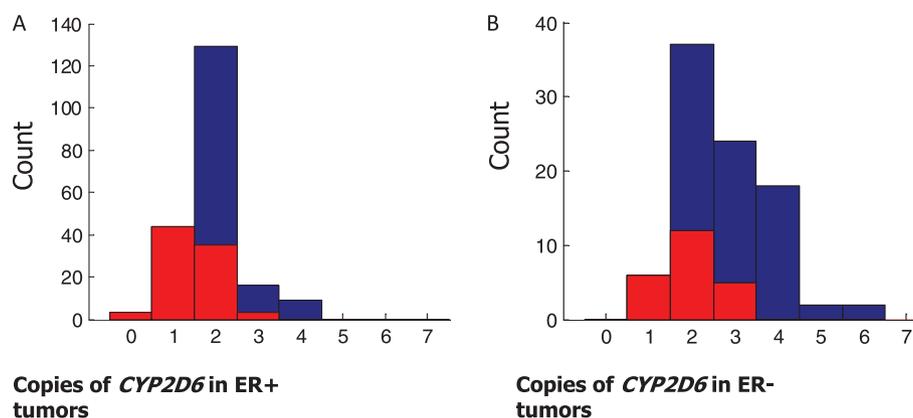


Figure 4. Frequency of Cytochrome P450 2D6 (*CYP2D6*) loss of heterozygosity within the Foundation Medicine estrogen receptor (ER)-positive (A) and ER-negative (B) cohorts. Tumor LOH is denoted in red. ER = estrogen receptor.

cancer, and particularly in the luminal and basal-like subtypes of breast cancer.

There are some limitations to our study. While we have demonstrated that the use of tumor-derived DNA contributes to *CYP2D6* genotyping error (analytical validity), this is unlikely to be the only factor contributing to the heterogeneity in the tamoxifen *CYP2D6* literature. In addition to “analytical validity,” Simon et al. pointed out that an “adequate number of patients with archived tissue must be present,” and suggested that the correlative study “include at least two-thirds of the total accrued patients” (26). It should be noted that in the ATAC study, less than 19% of the patients receiving tamoxifen were analyzed

with regard to *CYP2D6* genotype. Lastly, Simon et al. pointed out the critical nature of “clinical validity” (26). Here, it should be noted that the tamoxifen *CYP2D6* literature contains variability in tamoxifen dosing (20-40mg/day), duration of therapy (one to 10 years), ER status of the primary tumor, use of *CYP2D6* inhibiting medications, and, finally, lack of control for drugs that alter the hazard for recurrence (chemotherapy and aromatase inhibitors) (21). Therefore, we recommend careful control for each of these factors when analyzing and interpreting the tamoxifen *CYP2D6* literature.

In summary, we have provided definitive data from independent data sets that over 40% of primary and metastatic

**Table 1.** CYP2D6 allele frequencies determined by NGS in the Foundation Medicine cohort\*

CYP2D6 allele	Enzyme activity	Count in FM cohort	Frequency in FM cohort, %	Expected frequency, % (20) (for Europeans)
*1 or *2	Normal (wild-type)	461	64.0	63.1
*4	None	120	16.7	17.2
*41	Reduced	65	9.0	7.0
*9	Reduced	15	2.1	2.5
*10	Reduced	14	1.9	2.9
*5 (deletion of allele)	None	11	1.5	3.2
*6	None	7	1.0	0.6
*29	Reduced	7	1.0	0
*3	None	6	0.8	0.3
*17	Reduced	4	0.6	0
Other rare alleles	Various	10	1.4	0.4
Tandem duplications	Increased	Not assessed	Not assessed	2.8

\* CYP2D6 = cytochrome P450 2D6; FM = Foundation Medicine; NGS = next generation sequencing.

**Table 2.** The potential effects of CYP2D6 tumor LOH on Hardy Weinberg equilibrium

	Clinical subtype	HWE test: CYP2D6 *4	HWE test: CYP2D6 *41
NGS-based calls	ER+	0.005	0.19
NGS-based calls	ER-	0.95	0.75
NGS-based calls	All	0.02	0.11
*Germline allele <10%	ER+	$7.2 \times 10^{-4}$	0.16
*Germline allele <10%	ER-	0.78	0.75
*Germline allele <10%	All	$3.2 \times 10^{-4}$	0.09
*Germline allele <20%	ER+	$1.2 \times 10^{-6}$	0.04
*Germline allele <20%	ER-	0.018	0.75
*Germline allele <20%	All	$8.3 \times 10^{-8}$	0.02

\* Hardy Weinberg equilibrium calculation assuming a low-sensitivity genotyping assay would misclassify the low frequency allele as homozygous. CYP2D6 = cytochrome P450 2D6; ER = estrogen receptor; HWE = Hardy Weinberg equilibrium; LOH = loss of heterozygosity; NGS = next generation sequencing.

**Table 3.** CYP2D6\*4 genotypes obtained from FFPE blocks enriched for tumor or benign tissues

CYP2D6 *4 genotype using tumor-enriched DNA(4)	CYP2D6 *4 genotype using DNA from tumors containing benign tissues (16)			No call on updated analysis	Total
	Wt/Wt	Wt/*4	*4/*4		
Wt/Wt	121	15	1	0	137
Wt/*4	2	34	0	4	40
*4/*4	2	3	8	0	13
Total	125	52	9	4	190

\* CYP2D6 = Cytochrome P450 2D6; FFPE = formalin-fixed paraffin-embedded; Wt = wild-type.

breast tumors exhibit tumor LOH at the CYP2D6 locus and that the use of standard PCR (eg, Taqman) genotyping techniques applied to purified tumor DNA to detect germline CYP2D6 variation results in genotyping error because of an excess number of homozygotes and departure from HWE. Based on these results, we recommend that CYP2D6 genotyping be repeated in those studies in which the use of tumor DNA to derive germline CYP2D6 genotype resulted in substantial departure from HWE. Furthermore, recommendations and/or guidelines for the use of CYP2D6 genotyping should not be derived from studies with evidence for genotyping error.

## Funding

Supported in part by 1R01CA133049-01 (Goetz, Suman, Ames) the Mayo Clinic Breast Cancer Specialized Program of Research Excellence (SPORE) grant CA116201 (Ingle, Goetz) and CA058223 (Perou), the Breast Cancer Research Foundation, and U01GM61393 (Ratain).

## Notes

The study funders had no role in the design of the study, the collection, analysis, or interpretation of the data, the writing of the manuscript, nor the decision to submit the manuscript for publication.

## References

1. Stearns V, Johnson MD, Rae JM, et al. Active tamoxifen metabolite plasma concentrations after coadministration of tamoxifen and the selective serotonin reuptake inhibitor paroxetine. *J Natl Cancer Inst.* 2003;95(23):1758–1764.
2. Murdter TE, Schroth W, Bacchus-Gerybadze L, et al. Activity levels of tamoxifen metabolites at the estrogen receptor and the impact of genetic polymorphisms of phase I and II enzymes on their concentration levels in plasma. *Clin Pharmacol Ther.* 2011;89(5):708–717.
3. Brauch H, Schwab M. Prediction of tamoxifen outcome by genetic variation of CYP2D6 in post-menopausal women with early breast cancer. *Br J Clin Pharmacol.* 2014;77(4):695–703.

4. Goetz MP, Rae JM, Suman VJ, et al. Pharmacogenetics of tamoxifen biotransformation is associated with clinical outcomes of efficacy and hot flashes. *J Clin Oncol*. 2005;23(36):9312–9318.
5. Rae JM, Drury S, Hayes DF, et al. CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. *J Natl Cancer Inst*. 2012;104(6):452–460.
6. Regan MM, Leyland-Jones B, Bouzyk M, et al. CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the breast international group 1-98 trial. *J Natl Cancer Inst*. 2012;104(6):441–451.
7. Goetz MP, Suman VJ, Hoskin TL, et al. CYP2D6 metabolism and patient outcome in the Austrian Breast and Colorectal Cancer Study Group trial (ABCSCG) 8. *Clin Cancer Res*. 2013;19(2):500–507.
8. Nakamura Y, Ratain MJ, Cox NJ, et al. Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial. *J Natl Cancer Inst*. 2012;104(16):1264; author reply 1266–1268.
9. Stanton V Jr. Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial. *J Natl Cancer Inst*. 2012;104(16):1265–1266; author reply 1266–1268.
10. Pharoah PD, Abraham J, Caldas C. Re: CYP2D6 genotype and tamoxifen response in postmenopausal women with endocrine-responsive breast cancer: the Breast International Group 1-98 trial and Re: CYP2D6 and UGT2B7 genotype and risk of recurrence in tamoxifen-treated breast cancer patients. *J Natl Cancer Inst*. 2012;104(16):1263–1264; author reply 1266–1268.
11. Castells A, Gusella JF, Ramesh V, et al. A region of deletion on chromosome 22q13 is common to human breast and colorectal cancers. *Cancer Res*. 2000;60(11):2836–2839.
12. Hirano A, Emi M, Tsuneizumi M, et al. Allelic losses of loci at 3p25.1, 8p22, 13q12, 17p13.3, and 22q13 correlate with postoperative recurrence in breast cancer. *Clin Cancer Res*. 2001;7(4):876–882.
13. Comprehensive molecular portraits of human breast tumours. *Nature*. 2012;490(7418):61–70.
14. Jeselsohn R, Yelensky R, Buchwalter G, et al. Emergence of constitutively active estrogen receptor- $\alpha$  mutations in pretreated advanced estrogen receptor-positive breast cancer. *Clin Cancer Res*. 2014;20(7):1757–1567.
15. Goetz MP, Suman VJ, Ingle JN, et al. A two-gene expression ratio of homeobox 13 and interleukin-17B receptor for prediction of recurrence and survival in women receiving adjuvant tamoxifen. *Clin Cancer Res*. 2006;12(7)(Pt 1):2080–2087.
16. Schroth W, Goetz MP, Hamann U, et al. Association between CYP2D6 polymorphisms and outcomes among women with early stage breast cancer treated with tamoxifen. *JAMA*. 2009;302(13):1429–1436.
17. Brauch H, Schroth W, Goetz MP, et al. Tamoxifen use in postmenopausal breast cancer: CYP2D6 matters. *J Clin Oncol*. 2013;31(2):176–180.
18. Weigman VJ, Chao HH, Shabalin AA, et al. Basal-like Breast cancer DNA copy number losses identify genes involved in genomic instability, response to therapy, and patient survival. *Breast Cancer Res Treat*. 2012;133(3):865–880.
19. Pinto D, Darvishi K, Shi X, et al. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nature Biotech*. 2011;29(6):512–520.
20. Sistonen J, Sajantila A, Lao O, et al. CYP2D6 worldwide genetic variation shows high frequency of altered activity variants and no continental structure. *Pharmacogenet Genomics*. 2007;17(2):93–101.
21. Province MA, Goetz MP, Brauch H, et al. CYP2D6 Genotype and Adjuvant Tamoxifen: Meta-analysis of Heterogeneous Study Populations. *Clin Pharmacol Ther*. 2013;95(2):216–227.
22. Rae JM, Regan MM, Thibert JN, et al. Concordance Between CYP2D6 Genotypes Obtained From Tumor-Derived and Germline DNA. *J Natl Cancer Inst*. 2013;105(17):1332–1334.
23. Committee THCPAN. The human cytochrome P450 (CYP) allele nomenclature database. Available at: <http://www.cyp-alleles.ki.se/>. Accessed November 21, 2014.
24. Kelly CM, Pritchard KI. CYP2D6 genotype as a marker for benefit of adjuvant tamoxifen in postmenopausal women: lessons learned. *J Natl Cancer Inst*. 2012;104(6):427–428.
25. Madlensky L, Natarajan L, Tchu S, et al. Tamoxifen metabolite concentrations, CYP2D6 genotype, and breast cancer outcomes. *Clin Pharmacol Ther*. 2011;89(5):718–725.
26. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst*. 2009;101(21):1446–1452.
27. Zheng HT, Peng ZH, Zhou CZ, et al. Detailed deletion mapping of loss of heterozygosity on 22q13 in sporadic colorectal cancer. *World J Gastroenterol*. 2005;11(11):1668–1672.
28. Jonkers YM, Claessen SM, Feuth T, et al. Novel candidate tumour suppressor gene loci on chromosomes 11q23-24 and 22q13 involved in human insulinoma tumourigenesis. *J Pathol*. 2006;210(4):450–458.