

Estrogens, Enzyme Variants, and Breast Cancer: A Risk Model

Philip S. Crooke,¹ Marylyn D. Ritchie,² David L. Hachey,³ Sheila Dawling,⁴
Nady Roodi,⁴ and Fritz F. Parl⁴

Departments of ¹Mathematics, ²Molecular Physiology and Biophysics (Center for Human Genetics Research), ³Pharmacology, and ⁴Pathology, Vanderbilt University, Nashville, Tennessee

Abstract

Oxidative metabolites of estrogens have been implicated in the development of breast cancer, yet relatively little is known about the metabolism of estrogens in the normal breast. We developed a mathematical model of mammary estrogen metabolism based on the conversion of 17 β -estradiol (E₂) by the enzymes cytochrome P450 (CYP) 1A1 and CYP1B1, catechol-O-methyltransferase (COMT), and glutathione S-transferase P1 into eight metabolites [i.e., two catechol estrogens, 2-hydroxyestradiol (2-OHE₂) and 4-hydroxyestradiol (4-OHE₂); three methoxyestrogens, 2-methoxyestradiol, 2-hydroxy-3-methoxyestradiol, and 4-methoxyestradiol; and three glutathione (SG)-estrogen conjugates, 2-OHE₂-1-SG, 2-OHE₂-4-SG, and 4-OHE₂-2-SG]. When used with experimentally determined rate constants with purified enzymes, the model provides for a kinetic analysis of the entire metabolic pathway. The predicted concentration of each metabolite during a 30-minute reaction agreed well with the experimentally

derived results. The model also enables simulation for the transient quinones, E₂-2,3-quinone (E₂-2,3-Q) and E₂-3,4-quinone (E₂-3,4-Q), which are not amenable to direct quantitation. Using experimentally derived rate constants for genetic variants of CYP1A1, CYP1B1, and COMT, we used the model to simulate the kinetic effect of enzyme polymorphisms on the pathway and identified those haplotypes generating the largest amounts of catechols and quinones. Application of the model to a breast cancer case-control population identified a subset of women with an increased risk of breast cancer based on their enzyme haplotypes and consequent E₂-3,4-Q production. This *in silico* model integrates both kinetic and genomic data to yield a comprehensive view of estrogen metabolomics in the breast. The model offers the opportunity to combine metabolic, genetic, and lifetime exposure data in assessing estrogens as a breast cancer risk factor. (Cancer Epidemiol Biomarkers Prev 2006;15(9):1620-9)

Introduction

Numerous epidemiologic studies have implicated estrogens in the development of breast cancer (1, 2). The two major estrogens, 17 β -estradiol (E₂) and estrone (E₁), are ligands for the estrogen receptor and substrates for oxidizing phase I enzymes, cytochrome P450 (CYP) 1A1 and CYP1B1. In their dual role of ligand and substrate, estrogens may simultaneously stimulate cell proliferation and gene expression via the estrogen receptor and cause DNA damage via their oxidation products, the catechol estrogens (3, 4). The latter mechanism is based on the unique chemical structure of estrogens. Unlike all other steroid hormones, estrogens have an aromatic A-ring, which yields catechols on oxidation that may be further oxidized to highly reactive semiquinones and quinones (Fig. 1), which in turn can form both oxidative and estrogen DNA adducts. Thus, estrogen quinones seem to share a common feature of many chemical carcinogens (i.e., the ability to covalently modify DNA; refs. 5-8). Support for the carcinogenic activity of estrogens and their oxidative products, the catechol estrogens, comes from experiments in animal models. Treatment with E₂ and the catechol 4-hydroxyestrogen and 2-hydroxyestrogen caused kidney cancer in male Syrian hamsters and endometrial cancer in female CD1 mice (9-11). However, there is no animal model for estrogen-induced breast cancer, and even in the hamster and mouse models, the precise mechanism of DNA damage

is uncertain. Thus, there is a need to understand estrogen metabolism in the human breast to elucidate the role of endogenous and exogenous estrogens in mammary carcinogenesis. To advance this understanding requires not only characterization of the various estrogen metabolites but also equally important a precise definition of the responsible enzymes.

Several investigators have proposed a qualitative model of mammary estrogen metabolism regulated by oxidizing phase I and conjugating phase II enzymes (12, 13). The oxidative estrogen metabolism pathway starts with E₂ and E₁, which are oxidized to the 2-OH and 4-OH catechol estrogens by the phase I enzymes CYP1A1 and CYP1B1 (14, 15). These same enzymes are postulated to further oxidize the catechol estrogens to unstable semiquinones and quinones. Estrogen quinones then form Michael addition products with deoxynucleosides (5, 6, 16). The catechol estrogens and their estrogen quinones/semiquinones also undergo redox cycling, which results in the production of reactive oxygen species capable of causing oxidative DNA damage (17-19). Thus, P450-mediated estrogen metabolism is expected to lead to the formation of both oxidative and estrogen DNA adducts, all of which have been shown to possess mutagenic potential (20-22). It is postulated that the genotoxicity of the oxidative estrogen metabolism pathway is mitigated by alternate reactions of the metabolites with phase II enzymes. Specifically, catechol-O-methyltransferase (COMT) catalyzes the methylation of catechol estrogens to methoxy estrogens, which lowers the catechol estrogens available for conversion to estrogen quinones (23, 24). In turn, the estrogen quinones undergo conjugation with reduced glutathione (GSH) via the catalytic action of glutathione S-transferase (GST) P1 (25, 26). The formation of GSH-estrogen conjugates would reduce the level of estrogen quinones and thereby lower the potential for DNA damage.

Received 3/13/06; revised 5/16/06; accepted 6/22/06.

Grant support: NIH grants 1R01CA/ES83752, 5P30 CA68485, and 5P30 ES00267.

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

Requests for reprints: Fritz F. Parl, Department of Pathology, TVC 4918, Vanderbilt University, Nashville, TN 37232. Phone: 615-343-9117; Fax: 615-343-9563.

E-mail: fritz.parl@vanderbilt.edu

Copyright © 2006 American Association for Cancer Research.

doi:10.1158/1055-9965.EPI-06-0198

The current model of mammary estrogen metabolism has several limitations. First, only single enzymes (e.g., CYP1B1 and COMT) have been analyzed to date with simple substrate-product kinetics, which clearly generates an incomplete picture of the metabolic pathway. Second, although the model incorporates the functional roles of the phase I and II enzymes, it remains uncertain how the enzymes interact quantitatively. Third, each of the phase I and II enzymes contains genetic polymorphisms (15, 23, 27, 28). Studies from several laboratories, including our own, have examined the functional implications of the polymorphisms on estrogen metabolism, again focusing on single enzymes (15, 23, 24, 29, 30). Thus, the multitude of potential kinetic reactions resulting from the complex genetic variations of the phase I and II enzymes is completely outside the scope of the current model of estrogen metabolism. We developed recently an experimental *in vitro* model of mammary estrogen metabolism, in which we used purified, recombinant phase I enzymes CYP1A1 and CYP1B1 with the phase II enzymes COMT and GSTP1 to determine how E_2 is metabolized (31). We used both gas and liquid chromatography with mass spectrometry (GC/MS and LC/MS) to measure the parent hormone E_2 as well as eight metabolites (i.e., the catechol estrogens, methoxyestrogens, and estrogen-GSH conjugates; Fig. 1). In this article, we used these experimental data to develop a multicompartimental kinetic model of the metabolic pathway. Furthermore, we used previously determined rate constants of variant CYP1A1, CYP1B1, and COMT to present an *in silico* kinetic-genomic

model of mammary estrogen metabolism. Finally, we applied the model to a breast cancer case-control population and determined that the combination of enzyme haplotypes with E_2 -3,4-Q production did indeed identify a subset of women with increased breast cancer risk.

Materials and Methods

Mathematical Model. We developed a mathematical model for the estrogen metabolism pathway shown in Fig. 1. We assume that each reaction in the pathway ($A \rightarrow B$, a generic step in the pathway) is an enzyme-catalyzed reaction of the form:

$A + E \xrightleftharpoons[k_2]{k_1} C \xrightarrow{k_3} B + E$, where E denotes the enzyme, C is the enzyme-substrate complex, and k_i , $i = 1, 2$, and 3, are the rate constants of the reaction. For these types of reaction, we approximate the kinetics using the quasi steady-state assumption: $C = \frac{E^*A}{K_m + A}$, $K_m = \frac{k_2 + k_3}{k_1}$, where E^* is the initial enzyme concentration. With this assumption, we have $\frac{dB}{dt} \approx \frac{k_{cat}E^*A}{K_m + A}$, where k_{cat} is a constant. This approach leads to a system of nonlinear, ordinary differential equations for the concentrations of the compounds in the pathway. In each equation, k_{cat_i} and K_{m_i} are constants and E_{enzyme} are the enzyme levels in the respective reactions.

$$\frac{d(E_2)}{dt} = -\frac{k_{cat1} E_{CYP1B1} E_2}{K_{m1} + E_2} - \frac{k_{cat2} E_{CYP1A1} E_2}{K_{m2} + E_2} - \frac{k_{cat3} E_{CYP1B1} E_2}{K_{m3} + E_2} \quad (1)$$

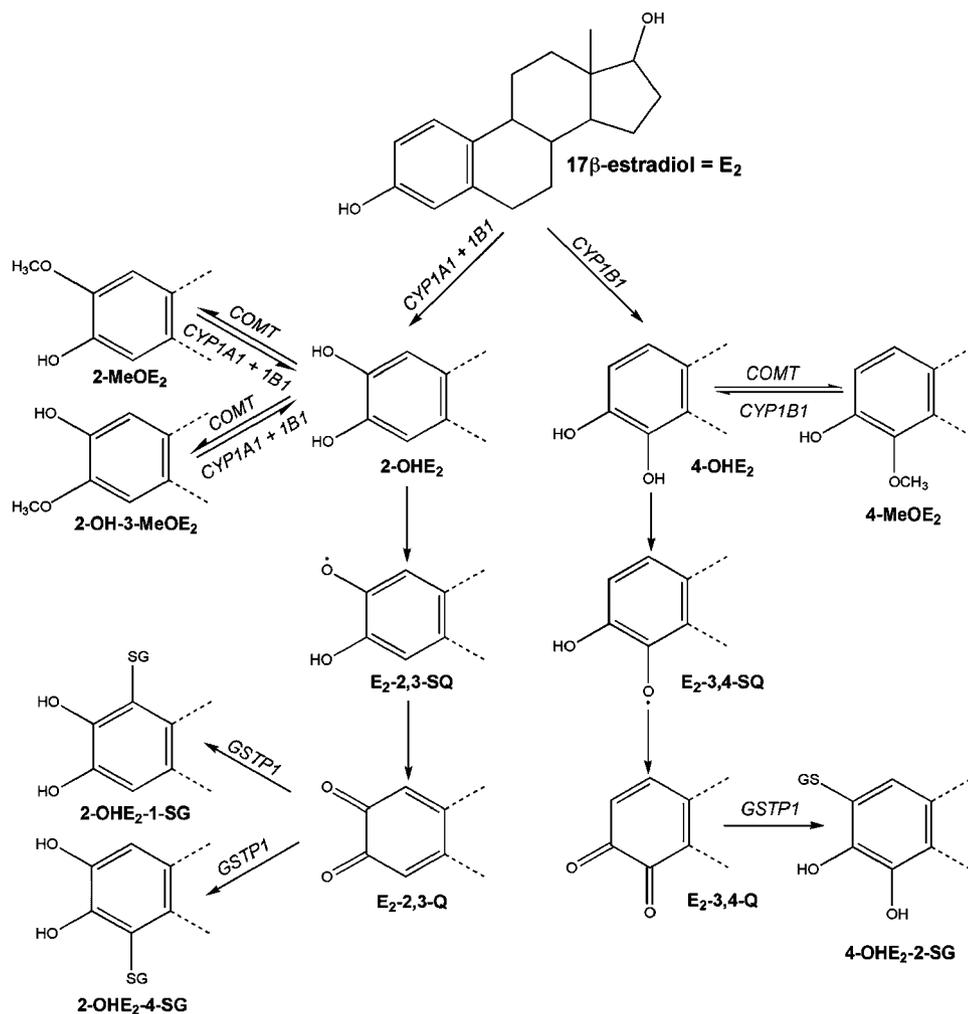


Figure 1. The estrogen metabolism pathway is regulated by oxidizing phase I and conjugating phase II enzymes. CYP1A1 and CYP1B1 catalyze the oxidation of E_2 to catechol estrogens 2-OHE₂ and 4-OHE₂. The catechol estrogens are either methylated by COMT to methoxyestrogens (2-MeOE₂, 2-OH-3-MeOE₂, and 4-MeOE₂) or further oxidized by CYPs to semi-quinones (E₂-2,3-SQ and E₂-3,4-SQ) and quinones (E₂-2,3-Q and E₂-3,4-Q). The estrogen quinones are conjugated by GSTP1 to GSH-conjugates (2-OHE₂-1-SG, 2-OHE₂-4-SG, and 4-OHE₂-2-SG).

$$\begin{aligned} \frac{d(OHE_2^2)}{dt} = & \frac{k_{cat_2} E_{CYP1A1} E_2}{K_{m_2} + E_2} + \frac{k_{cat_3} E_{CYP1B1} E_2}{K_{m_3} + E_2} - \frac{k_{cat_6} E_{COMT} OHE_2^2}{K_{m_6} + OHE_2^2} + \\ & \frac{k_{cat_7} E_{CYP1A1} MeOHE_2^2}{K_{m_7} + MeOHE_2^2} + \frac{k_{cat_8} E_{CYP1B1} MeOHE_2^2}{K_{m_8} + MeOHE_2^2} - \\ & \frac{k_{cat_9} E_{COMT} OHE_2^2}{K_{m_9} + OHE_2^2} + \frac{k_{cat_{10}} E_{CYP1A1} MeOHE_2^{23}}{K_{m_{10}} + MeOHE_2^{23}} + \\ & \frac{k_{cat_{11}} E_{CYP1B1} MeOHE_2^{23}}{K_{m_{11}} + MeOHE_2^{23}} - \frac{V_{max_{Q1}} (OHE_2^2)^{\sigma_{Q1}}}{K_{m_{Q1}} + (OHE_2^2)^{\sigma_{Q1}}} \end{aligned} \quad (2)$$

$$\begin{aligned} \frac{d(OHE_2^4)}{dt} = & \frac{k_{cat_4} E_{CYP1B1} E_2}{K_{m_4} + E_2} - \frac{k_{cat_4} E_{COMT} OHE_2^4}{K_{m_4} + OHE_2^4} + \\ & \frac{k_{cat_5} E_{COMT} MeOHE_2^4}{K_{m_5} + MeOHE_2^4} - \frac{V_{max_{Q2}} (OHE_2^4)^{\sigma_{Q2}}}{K_{m_{Q2}} + (OHE_2^4)^{\sigma_{Q2}}} \end{aligned} \quad (3)$$

$$\frac{d(MeOHE_2^4)}{dt} = \frac{k_{cat_4} E_{COMT} OHE_2^4}{K_{m_4} + OHE_2^4} - \frac{k_{cat_5} E_{CYP1B1} MeOHE_2^4}{K_{m_5} + MeOHE_2^4} \quad (4)$$

$$\begin{aligned} \frac{d(MeOHE_2^2)}{dt} = & \frac{k_{cat_6} E_{COMT} OHE_2^2}{K_{m_6} + OHE_2^2} - \frac{k_{cat_7} E_{CYP1A1} MeOHE_2^2}{K_{m_7} + MeOHE_2^2} - \\ & \frac{k_{cat_8} E_{CYP1B1} MeOHE_2^2}{K_{m_8} + MeOHE_2^2} \end{aligned} \quad (5)$$

$$\begin{aligned} \frac{d(MeOHE_2^{23})}{dt} = & \frac{k_{cat_9} E_{COMT} OHE_2^2}{K_{m_9} + OHE_2^2} - \frac{k_{cat_{10}} E_{CYP1A1} MeOHE_2^{23}}{K_{m_{10}} + MeOHE_2^{23}} - \\ & \frac{k_{cat_{11}} E_{CYP1B1} MeOHE_2^{23}}{K_{m_{11}} + MeOHE_2^{23}} \end{aligned} \quad (6)$$

$$\begin{aligned} \frac{d(EQ_2^{23})}{dt} = & \frac{V_{max_{Q1}} (OHE_2^2)^{\sigma_{Q1}}}{K_{m_{Q1}} + (OHE_2^2)^{\sigma_{Q1}}} - \frac{k_{cat_{13}} E_{GSTP1} EQ_2^{23}}{K_{m_{13}} + EQ_2^{23}} - \\ & \frac{k_{cat_{14}} E_{GSTP1} EQ_2^{23}}{K_{m_{14}} + EQ_2^{23}} - k_1 EQ_2^{23} \end{aligned} \quad (7)$$

$$\frac{d(EQ_2^{34})}{dt} = \frac{V_{max_{Q2}} (OHE_2^2)^{\sigma_{Q2}}}{K_{m_{Q2}} + (OHE_2^2)^{\sigma_{Q2}}} - \frac{k_{cat_{12}} E_{GSTP1} EQ_2^{34}}{K_{m_{12}} + EQ_2^{34}} - k_2 EQ_2^{34} \quad (8)$$

$$\frac{d(OHE_2^{21} SG)}{dt} = \frac{k_{cat_{14}} E_{GSTP1} EQ_2^{23}}{K_{m_{14}} + EQ_2^{23}} \quad (9)$$

$$\frac{d(OHE_2^{24} SG)}{dt} = \frac{k_{cat_{13}} E_{GSTP1} EQ_2^{23}}{K_{m_{13}} + EQ_2^{23}} \quad (10)$$

$$\frac{d(OHE_2^{42} SG)}{dt} = \frac{k_{cat_{12}} E_{GSTP1} EQ_2^{34}}{K_{m_{12}} + EQ_2^{34}} \quad (11)$$

There are parts of the pathway for which kinetic data are not available. In particular, rate constants cannot be determined experimentally for the reaction sequences 2-hydroxyestradiol ($2-OHE_2 \rightarrow E_2-2,3-SQ \rightarrow 2-OHE_2$ -quinone ($E_2-2,3-Q$) and $4-OHE_2 \rightarrow E_2-3,4-SQ \rightarrow E_2-3,4-Q$ because of the transient nature of the semiquinones (ms half-life; ref. 32). Therefore, we simplified the pathway and collapsed the sequential reactions to single reactions, $2-OHE_2 \rightarrow E_2-2,3-Q$ and $4-OHE_2 \rightarrow E_2-3,4-Q$, respectively. We also assumed that each of these quinone

production reactions ($OHE_2^k \rightarrow EQ_2^j$) satisfies dynamics of the form $\frac{dEQ_2^j}{dt} = \frac{V_{max_{Qj}} (OHE_2^k)^{\sigma_{Qj}}}{K_{m_{Qj}} + (OHE_2^k)^{\sigma_{Qj}}}$, where $V_{max_{Qj}}$, $K_{m_{Qj}}$, and σ_{Qj} are constants. For the mathematical model to be a tractable computational model of the metabolism pathway, it is necessary to have estimates of these unknown constants. We used two types of experimental data to derive the constants $V_{max_{Q1}}$, $V_{max_{Q2}}$, $K_{m_{Q1}}$, $K_{m_{Q2}}$, σ_{Q1} , and σ_{Q2} . First, we used rate constants determined experimentally for individual reactions catalyzed by CYP1A1, CYP1B1, COMT, and GSTP1 (15, 23, 26, 33). Second, we used the concentrations over time determined for every non-quinone compound in the pathway following simultaneous incubation of the parent hormone E_2 with all four enzymes (31). Using the experimental data, a searching algorithm was written in *Mathematica* (Wolfram Research, Inc., Champaign, IL) to find values for $V_{max_{Qj}}$, $K_{m_{Qj}}$, and σ_{Qj} . The derived variables, $V_{max_{Q1}}$, $V_{max_{Q2}}$, $K_{m_{Q1}}$, $K_{m_{Q2}}$, σ_{Q1} , and σ_{Q2} , for the two quinone reactions were chosen to fit the experimental data using numerical solutions of the system of differential equation.

As a measure of the quinone concentrations over the course of time, we introduce the area under the curve (AUC) metric: $AUC_k = \int_0^T EQ_2^k(t) dt$, where $k = 23$ and 34 and $T = 30$ minutes. It is possible to introduce other measures [e.g., $EQ_{2-max}^{ij} = \max_{0 \leq t \leq T} EQ_2^j(t)$], which is the highest concentration achieved during the time interval $(0, T)$. We have chosen the former metric because it incorporates both concentration and time.

CYP1A1 Variants. Wild-type (WT) CYP1A1 cDNA was prepared for expression and purification of recombinant CYP1A1 as described previously (15, 33). Site-directed mutagenesis was done to generate the cDNA variants, which were verified by nucleotide sequence analysis and then similarly expressed and purified: 462Ile \rightarrow Val (m2), 461Thr \rightarrow Asn (m4), and 461Asn-462Val (m2/m4) (27). SDS-PAGE showed >95% protein purity and the reduced-CO difference spectrum revealed the λ_{max} at 450 nm, which allowed quantitation for subsequent enzyme experiments. We used GC/MS (15, 33) to determine the reaction kinetics of E_2 oxidation for WT, m2, m4, and m2/m4 CYP1A1.

Study Population. The hospital-based case-control study group of 221 Caucasian women with primary invasive breast cancer and their age-matched control subjects has been described previously (34-36). Genomic DNA was extracted from tumor tissue or WBCs. The DNA samples of four control subjects had been depleted, leaving 221 cases and 217 controls for the study group.

DNA Analysis. The genotypes of CYP1A1, CYP1B1, and COMT were determined by PCR and restriction endonuclease digestion as described previously (23, 35, 36). Each PCR contained internal controls for the respective gene and random retesting of 50 samples yielded 100% reproducibility. Direct sequencing of five different samples provided further independent genotype validation.

Statistical Analysis. The Wilcoxon rank sum test was used to determine the median difference in age between cases and controls. The χ^2 test was used to compare the distribution of CYP1A1, CYP1B1, and COMT alleles in cases and controls. The χ^2 goodness-of-fit test was used for testing Hardy-Weinberg equilibrium (HWE). Haplotype frequencies were estimated via the expectation-maximization algorithm in Powermarker version 3.23 (37, 38). Haplotype-trait association with breast cancer was tested using a regression approach also in Powermarker version 3.23 (37, 39).

Composite $E_2-3,4-Q$ AUC. There are 4 CYP1A1, 16 CYP1B1, and 2 COMT (23, 35, 36) haplotypes with $4 \times 16 \times 2 = 128$ possible genetic combinations. We calculated an $E_2-3,4-Q$ AUC

for each woman based on her CYP1A1, CYP1B1, and COMT haplotypes, which together we term a composite haplotype. In calculating each AUC, we considered that the *CYP1B1* gene has four polymorphic sites with 16 possible haplotypes. Because any individual can possess only two haplotypes, the certainty of assigning the correct CYP1B1 haplotypes becomes a matter of probability if the individual is heterozygous for more than one polymorphic site. The probability of a particular composite haplotype occurring in our population of 438 women was computed in the following manner. For each individual, a subset of the 128 possible combinations was computed resulting in 438 subsets of haplotype combinations (CYP1A1, CYP1B1, and COMT). Then, a count was done for each of the 128 possible composite haplotypes in the 438 subsets from which a frequency chart was constructed. From this chart, we defined the probabilities of composite haplotypes occurring in the population. Suppose an individual has n possible composite haplotypes of CYP1A1, CYP1B1, and COMT and let $AUC_1, AUC_2, \dots, AUC_n$ denote the AUC values for each haplotype. If the probabilities that these composite haplotypes occur in the population are P_1, P_2, \dots, P_n , where $0 < P_i < 1, i = 1, 2, \dots, n$, then the composite AUC_{comp} is defined as

$$AUC_{\text{comp}} = \frac{\sum_{i=1}^n P_i AUC_i}{\sum_{i=1}^n P_i}$$

Results

Validation of *In silico* Model against Experimental Data.

In a previous study, we determined the metabolism of E_2 , 2-OHE₂, 4-OHE₂, 2-methoxyestradiol (2-MeOE₂), 2-hydroxy-3-methoxyestradiol (2-OH-3-MeOE₂), 4-methoxyestradiol (4-MeOE₂), 2-OHE₂-1-SG, 2-OHE₂-4-SG, and 4-OHE₂-2-SG as a function of time in the presence of CYP1A1 (85 pmol), CYP1B1 (165 pmol), COMT (125 pmol), and GSTP1 (500 pmol). Each experimental reaction contained 10 μmol/L E_2 , 100 μmol/L S-adenosyl methionine, 100 μmol/L glutathione and proceeded for 0, 2, 5, 10, 20, and 30 minutes at 37°C followed by GC/MS and LC/MS analysis (31). Figure 2A shows superimposed the experimental data (dots) and the model simulations (curves) for all nine analytes over the 30-minute reaction time. In the simulations, it was assumed that initially all analyte concentrations are zero, except $E_2(0) = E_2^*$. Enzyme concentrations used in the simulations are consistent with those used in the preceding experimental studies (15, 23, 26, 31, 33). Given the complexity of the pathway, there is excellent agreement between the simulated and experimental results. Of the nine analytes, only two, 2-MeOE₂ and 2-OH-3-MeOE₂, showed a noticeable difference between simulated and measured results. As shown in Fig. 1, the likely reason for this discrepancy lies in the more complex kinetics of 2-MeOE₂ and 2-OH-3-MeOE₂. These methoxyestrogens are the only analytes that are subject to the simultaneous action of three enzymes (i.e., the COMT-mediated production, which is counteracted by CYP1A1- and CYP1B1-mediated demethylation; ref. 33).

Estrogen Quinones. The estrogen quinones are too labile to be reliably quantified in a multienzyme system. However, as outlined in Materials and Methods, we could use the mathematical model to provide functional relations between $E_2(t)$ and the estrogen quinone concentrations: $EQ_2^{23}(t)$ and $EQ_2^{34}(t)$. Figure 2B shows the simulated production and disappearance of the estrogen quinones during the 30-minute reaction with a lower level and faster, nearly complete disappearance of E_2 -2,3-Q compared with the higher, more sustained level of E_2 -3,4-Q.

Enzyme Polymorphisms. Because the model was built on experimentally determined rate constants, we could analyze how variations in these kinetic variables, occurring as the result of enzyme polymorphisms, affect single steps or a combination of steps in the pathway. Table 1 summarizes the kinetic variables determined for variants of CYP1A1 (this study), CYP1B1 (40), and COMT (23). We applied the model to simulate reactions using the rate constants for 4 CYP1A1, 16 CYP1B1, and 2 COMT WT and variant enzymes. For each of the $4 \times 16 \times 2 = 128$ possible genetic combinations, we simulated values for the resulting estrogen metabolites over the 30-minute reaction time and then used interpolatory polynomials for these functions to calculate the respective AUCs. These simulations permitted us to see that the combinations of enzyme variants produced a continuous spectrum of concentrations over time for each of the estrogen metabolites. Accordingly, the model allowed us to identify which variant combinations of CYP1A1, CYP1B1, and COMT produced the highest or lowest estrogen metabolite concentrations over time. Because the catechols and quinones have been shown to cause DNA damage, we focused our attention on these two groups of metabolites (Fig. 3). Of the 128 combinations of CYP1A1, CYP1B1, and COMT variants, we found, for example, that the haplotype combination CYP1A1_{461Asn-462Ile}CYP1B1_{48Arg-119Ser-432Val-453Asn}COMT_{108Met} produced the maximum AUC for both 4-OHE₂ and E_2 -3,4-Q (Fig. 3B and D).

Clinical Application of Kinetic-Genomic Model. We applied the model to a hospital-based breast cancer case-control population that has been analyzed previously (34-36). The two principal differences to the preceding studies are (a) the evaluation of haplotypes rather than genotypes and (b) the integrated examination of CYP1A1, CYP1B1, and COMT instead of as independent entities. Table 2 summarizes the allele and haplotype data of the case-control population. Only 3 of the 4 possible CYP1A1 haplotypes and 12 of 16 possible CYP1B1 haplotypes were observed in the study group. Among the 12 CYP1B1 haplotypes, 8 were present in both cases and controls. Three of the uncommon haplotypes were seen only in cases and one rare haplotype was found only in controls. The overall P for the CYP1B1 haplotype distribution among cases and controls was 0.63. In the simulations, we focused our analysis on the E_2 -3,4-Q AUC because E_2 -3,4-Q has been identified as the principal estrogen metabolite causing DNA damage (8-10, 41). We calculated a composite E_2 -3,4-Q AUC for each woman based on her CYP1A1, CYP1B1, and COMT haplotypes as outlined in Materials and Methods. This information was then used to rank every woman in the entire study population based on her individual E_2 -3,4-Q AUC (Fig. 4).

A major weakness of genetic studies is the neglect of phenotypic factors. This is particularly true for polymorphic enzymes whose activity levels can vary considerably more in response to inducing agents than as a result of a single inherited amino acid substitution. For this reason, we considered the effect of changing the concentration of the phase I enzymes, which play the principal role in the metabolic pathway. As shown in Fig. 4, we varied the CYP1B1/CYP1A1 ratio from 2 to 5 in the model to reflect reported 4-OHE₂/2-OHE₂ ratios in breast tissue (42-44). In these simulations, we changed the concentrations of CYP1B1 while keeping CYP1A1 constant. The concentrations of COMT and GSTP1 remained unchanged. When we ranked the E_2 -3,4-Q AUCs for the entire study population at different CYP1B1/CYP1A1 ratios, we observed an increase in median AUCs for cases and controls with increasing CYP ratio (Fig. 4A). There were no significant differences between case and control E_2 -3,4-Q AUCs at any CYP1B1/CYP1A1 ratio. However, cases predominated in the top tier of the population as shown for the top 8 percentile

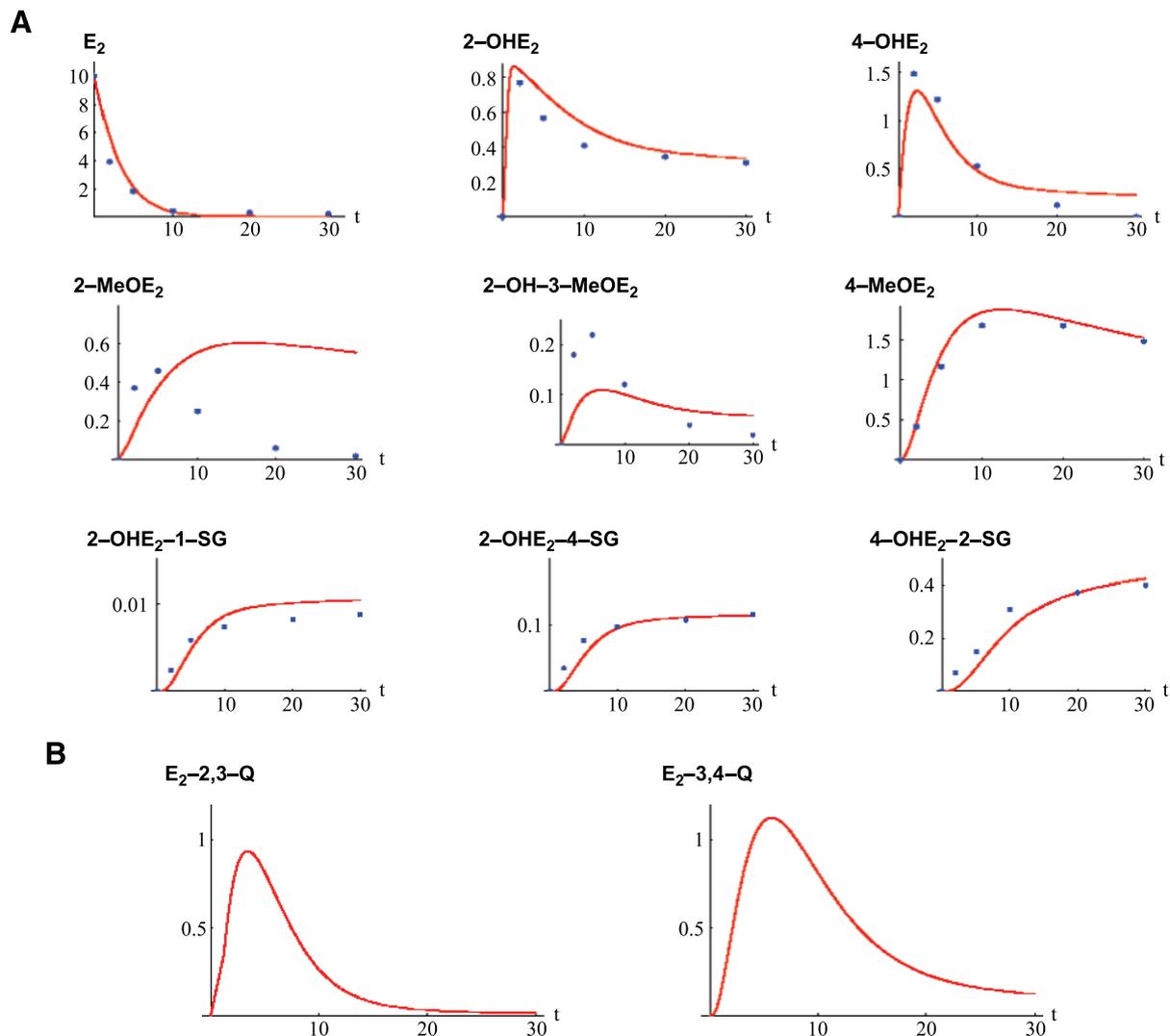


Figure 2. A. Comparison of mathematical model with experimental data. The metabolism of E_2 , 2-OHE₂, 4-OHE₂, 2-MeOE₂, 2-OH-3-MeOE₂, 4-MeOE₂, 2-OHE₂-1-SG, 2-OHE₂-4-SG, and 4-OHE₂-2-SG is shown as a function of time. Concentration is expressed in $\mu\text{mol/L}$. Blue dots, experimental data (31); red curves, derived from the mathematical model. B. Simulated kinetics of estrogen quinones E_2 -2,3-Q and E_2 -3,4-Q.

(35 subjects) of the study group (Fig. 4B). At CYP1B1/CYP1A1 = 5, the model identified 23 cases and 12 controls ($P = 0.06$). The discriminating ability was even more pronounced in the top 2 percentile (10 subjects) of the study population (Fig. 4C). At CYP1B1/CYP1A1 = 5, there were nine cases and 1 control ($P = 0.01$). Table 3 summarizes the composite CYP1A1, CYP1B1, and COMT haplotypes together with the E_2 -3,4-Q AUCs for the top 10 women.

Discussion

The complexity of mammary estrogen metabolism was recognized several years ago and outlined in a qualitative model (12, 13). Although this model defined the role of specific components (e.g., the oxidizing phase I and conjugating phase II enzymes), the quantitative effect of these enzymes in the overall pathway could not be assessed. The experimental analysis of single enzymes with simple substrate-product kinetics offered an incomplete picture of the pathway limited to the enzyme examined. Here, we present a novel approach that incorporates experimental data obtained previously with individual enzymes into a mathematical model of the estrogen metabolism pathway. Instead of simply doing a parametric

fitting exercise, we used actual experimental rate constants to develop the model, which consists of 11 differential equations that permit us to simulate the kinetics of E_2 and 8 metabolites in the multienzyme pathway. The model simulations were validated against experimental results obtained previously by incubating E_2 with the combined enzymes CYP1A1, CYP1B1, COMT, and GSTP1 (31) and showed excellent concordance of simulated and measured results. Of the nine analytes, only 2-MeOE₂ and 2-OH-3-MeOE₂ showed a noticeable deviation of the simulated from the measured results, most likely due to their more complex kinetics resulting from the simultaneous involvement of three enzymes, COMT, CYP1A1, and CYP1B1. It is noteworthy that the deviation of the two methoxyestrogens did not affect the simulation of more distal metabolites, such as the GSH-estrogen conjugates, which showed excellent agreement (Fig. 2A).

Catechol estrogens and estrogen quinones occupy pivotal positions in the oxidative estrogen metabolism pathway (Fig. 1). Using GC/MS, we could follow the production and disappearance of the catechol estrogens (Fig. 2A). Ideally, measurements of the estrogen quinones should be made, but they are highly reactive with short half-lives (seconds to minutes) due to the strained 1,2-diketone functionality inherent in *o*-quinones (45). Although estrogen quinones are

too labile to be reliably quantified in a multienzyme system, the model allowed us to simulate their production and disappearance during the 30-minute reaction (Fig. 2B). The disappearance of the quinones is due to two factors, the conversion of the quinones into stable GSH-estrogen conjugates and the irreversible loss from the system, most likely due to binding of the reactive quinones to protein (46). The more rapid disappearance of E₂-2,3-Q compared with E₂-3,4-Q is consistent with the shorter half-life of 42 seconds for E₁-2,3-Q compared with 12 minutes for E₁-3,4-Q (47). Overall, the model captures the joint action of the phase I and II enzymes rather well, allowing the simulation of the pathway from the parent hormone E₂ through several enzymatic steps to the most distal metabolites, the GSH-estrogen conjugates. Because these conjugates are produced via the quinones, the excellent agreement between simulated and measured GSH-estrogen conjugate levels provides further assurance about the validity of modeling the estrogen quinones.

Although other phase I enzymes, such as CYP1A2 and CYP3A4, are involved in hepatic and extrahepatic estrogen oxidation, CYP1A1 and CYP1B1 display the highest levels of expression in breast tissue and therefore are considered the principal oxidizing enzymes in mammary estrogen metabolism (48, 49). COMT shows ubiquitous expression in all tissues, including breast (50). Although COMT is the sole methylating enzyme, there are potentially three GSH-conjugating enzymes active in the pathway. Based on protein levels in breast tissue, GSTP1 is the predominant member of the GST family with

GSTM1 and GSTA1 present at much lower levels (51-53). GSTs are known to have selective as well as overlapping substrate specificities and it is presently not known whether GSTM1 and GSTA1 share with GSTP1 the ability to conjugate estrogen quinones (26). To determine the potential roles of GSTM1 and GSTA1 in estrogen metabolism, we plan to prepare each as purified, recombinant enzyme followed by kinetic studies to define their respective rate constants. Besides COMT, there are two other classes of phase II enzymes capable of conjugating catechol estrogens (i.e., the sulfotransferases and UDP-glucuronosyltransferases). It seems that the catechol estrogens are converted predominantly to methyl conjugates and to a lesser extent to sulfate and glucuronide conjugates (54). In future experiments, we will assess the role of sulfotransferases and UDP-glucuronosyltransferases. The present mathematical model only incorporates the key phase I enzymes CYP1A1 and CYP1B1 and the phase II enzymes COMT and GSTP1. However, the model can readily accommodate additional enzymes and allow inclusion of other GST members as well as sulfotransferases and UDP-glucuronosyltransferases in the same manner as we currently do for CYP1A1 and CYP1B1. In contrast to the complex kinetics of the methoxyestrogens, the sulfate and glucuronide conjugation reactions follow simple substrate-product kinetics like the GSTP1-mediated GSH conjugation. Therefore, we anticipate straightforward modeling with good agreement of simulated and experimental data.

Each of the phase I and II enzymes involved in estrogen metabolism possesses genetic variants that (*a*) are associated

Table 1. Kinetic variables for WT and variant CYP1A1, CYP1B1, and COMT

Reaction	Enzyme	Allele	<i>k</i> _{cat} or <i>V</i> _{max}	<i>K</i> _m
E ₂ → 2-OHE ₂	CYP1A1	461Thr-462Ile (WT)	1.50	17
		461Asn-462Ile	1.10	23
		461Thr-462Val	3.60	18
		461Asn-462Val	1.70	23
E ₂ → 2-OHE ₂	CYP1B1	48Arg-119Ala-432Val-453Asn (WT)	0.36	24
		48Gly-119Ala-432Val-453Asn	0.17	12
		48Arg-119Ser-432Val-453Asn	0.29	19
		48Arg-119Ala-432Leu-453Asn	0.35	17
		48Arg-119Ala-432Val-453Ser	0.91	49
		48Gly-119Ser-432Leu-453Ser	0.50	33
		48Arg-119Ala-432Leu-453Ser	0.56	30
		48Gly-119Ser-432Leu-453Asn	0.28	17
		48Gly-119Ser-432Val-453Asn	0.19	13
		48Gly-119Ser-432Val-453Ser	0.64	34
		48Arg-119Ser-432Leu-453Asn	0.56	33
		48Arg-119Ser-432Leu-453Ser	0.47	24
		48Gly-119Ala-432Leu-453Ser	0.47	31
		48Gly-119Ala-432Leu-453Asn	0.19	9.1
		48Arg-119Ser-432Val-453Ser	0.65	36
		48Gly-119Ala-432Val-453Ser	0.23	75
E ₂ → 4-OHE ₂	CYP1B1	48Arg-119Ala-432Val-453Asn (WT)	2.10	14
		48Gly-119Ala-432Val-453Asn	0.80	6.6
		48Arg-119Ser-432Val-453Asn	1.70	9.3
		48Arg-119Ala-432Leu-453Asn	1.40	9.6
		48Arg-119Ala-432Val-453Ser	3.10	21
		48Gly-119Ser-432Leu-453Ser	1.70	15
		48Arg-119Ala-432Leu-453Ser	2.20	13
		48Gly-119Ser-432Leu-453Asn	0.71	5.8
		48Gly-119Ser-432Val-453Asn	1.10	5.5
		48Gly-119Ser-432Val-453Ser	2.20	15
		48Arg-119Ser-432Leu-453Asn	1.90	15
		48Arg-119Ser-432Leu-453Ser	1.90	13
		48Gly-119Ala-432Leu-453Ser	1.80	12
		48Gly-119Ala-432Leu-453Asn	0.73	7.2
		48Arg-119Ser-432Val-453Ser	2.70	17
		48Gly-119Ala-432Val-453Ser	0.81	28
2-OHE ₂ → 2-MeOE ₂	COMT	108Val (WT)	6.80	117
		108Met	2.72	99
2-OHE ₂ → 2-OH-3-MeOE ₂	COMT	108Val (WT)	1.50	51
		108Met	0.62	58
4-OHE ₂ → 4-MeOE ₂	COMT	108Val (WT)	3.40	24
		108Met	1.94	28

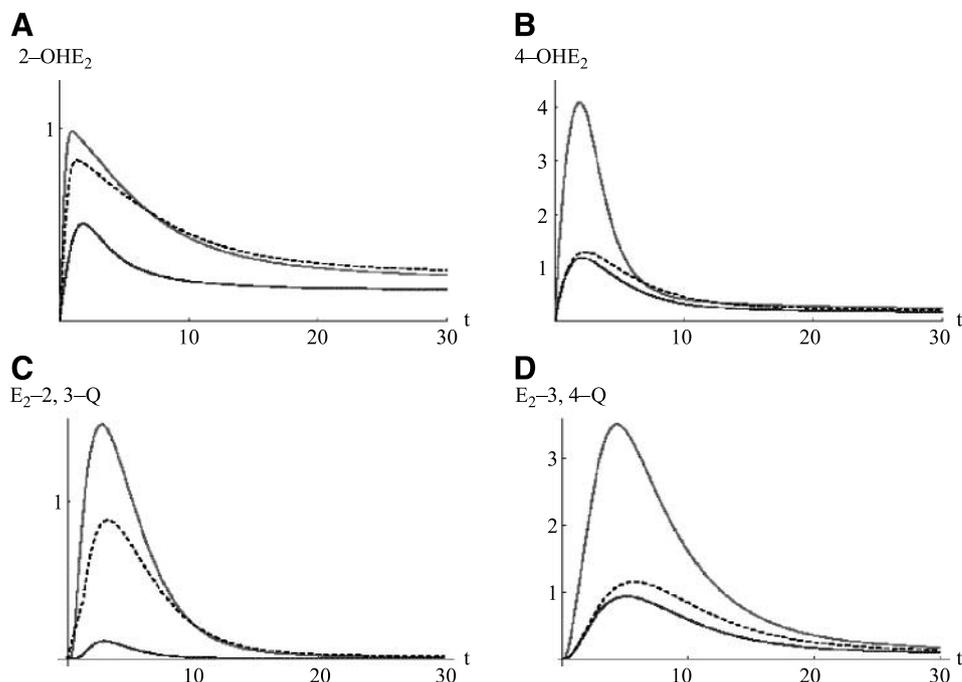


Figure 3. Kinetic-genomic modeling of catechol estrogens, 2-OHE₂ (A) 4-OHE₂ (B), and estrogen quinones, E₂-2,3-Q (C) and E₂-3,4-Q (D), using rate constants for WT and variant CYP1A1, CYP1B1, and COMT. Only the highest, lowest, and WT (dotted line) AUCs are shown.

with altered enzyme function and (b) occur in a sizeable portion of the population (55, 56). Because our experimental analysis used only WT recombinant enzymes, the results provide a limited view of estrogen metabolism. To obtain a more realistic and inclusive view of estrogen metabolism in the female population, we used the mathematical model to simulate how variations in the kinetic variables resulting from polymorphisms of the enzymes affect the metabolite concentrations. We examined 4 CYP1A1, 16 CYP1B1, and 2 COMT alleles. GSTP1 also has two polymorphisms (i.e., 104Ile → Val and 113Ala → Val; refs. 28, 57), but it is unknown whether they affect GSH-estrogen conjugation. Thus, our simulations are based on the examination of $4 \times 16 \times 2$ genetic combinations to show the utility of the model. Although each of the metabolites can be modeled, we concentrated our analysis on the catechols and quinones because of their documented carcinogenic activity (5-7, 16). As shown in Fig. 3, modeling of the 128 haplotype combinations produced a continuous spectrum of catechol and quinone concentrations over time, as expressed by a range of AUCs. The simulations identified the haplotype combinations producing the highest and lowest AUCs. For example, the maximum AUCs for 4-OHE₂ and E₂-3,4-Q were produced by the haplotype CYP1A1_{461Asn-462Ile}CYP1B1_{48Arg-119Ser-432Val-453Asn-COMT_{108Met}}, which were 2.6- and 4.6-fold higher, respectively, than the minimum AUCs produced by haplotype CYP1A1_{461Thr-462Val}CYP1B1_{48Gly-119Ala-432Val-453Ser}COMT_{108Val}. Although these differences may not appear large, it is important to consider that they affect on lifetime exposure, which is consistent with the hormonal risk model presented by Pike et al. (58).

Our kinetic-genomic model is pertinent to the numerous epidemiologic studies that have examined the association of genetic variants of enzymes involved in estrogen metabolism with breast cancer risk (56, 59). These studies have been handicapped by investigating only one or two enzymes, but even those examining all enzymes have been fundamentally limited by not being able to assess the underlying metabolic interactions (60). Our model attempts to fill this gap and we applied it to a hospital-based case-control population that has been analyzed previously with respect to CYP1A1, CYP1B1, and COMT genotypes (34-36). Here, we went beyond genotypes and used the model to determine for each woman the effect of her composite CYP1A1, CYP1B1, and COMT haplotypes on estrogen metabolite production. Inherited

variations in enzyme genotype persist throughout life and can therefore be regarded as constants for each individual. However, the very same genes are also subject to induction and levels of enzyme expression may vary considerably as a result of the high degree of inducibility by a variety of agents. For example, CYP enzymes are induced by hundreds of compounds (dietary and environmental chemicals and drugs) and human exposure to such xenobiotics is unavoidable (61). Intraindividual and interindividual variation in xenobiotic exposure has several consequences: (a) the P450 activity in an individual may change over time, (b) the P450 activity may differ between individuals of the same genotype, and (c) the phenotypic variability in P450 activity may be greater than the effect of genetic polymorphisms due to the strong inducing power of certain xenobiotics. Thus, although each individual has a unique composite E₂-3,4-Q AUC based on her subset of 128 genetic combinations, the AUC value can vary with the phenotype. In the model, we attempted to incorporate both the certainty of the enzyme genotype and the ambiguity of the phenotype, the latter indicated by the changing ratio of phase I enzymes CYP1B1/CYP1A1 (Fig. 4). Although we have to accept imprecise information about P450 activity in breast tissue, we can assume that the concentration of CYP1B1 is greater than that of CYP1A1 based on mRNA expression levels, higher levels of 4-OHE₂ than 2-OHE₂, and the observation that 2-OHE₂ is produced by both CYP isoforms, whereas 4-OHE₂ is formed only by CYP1B1 (14, 15, 48, 49). Because the 4-OHE₂/2-OHE₂ ratio can be ~3 and reach as high as ~5 (42-44), we varied the CYP1B1/CYP1A1 ratio in the model from 2 to 5. For CYP ratios >2, the model identified a top tier of E₂-3,4-Q AUCs with significantly increased numbers of breast cancer cases in the top percentiles (Fig. 4B and C; Table 3), suggesting that E₂-3,4-Q AUC may be an indicator of breast cancer risk. The ranking order of E₂-3,4-Q AUCs is primarily determined by the enzyme genotype (i.e., the composite CYP1A1-CYP1B1-COMT haplotype of a subject). However, the AUC ranking is also affected by the enzyme phenotype and a change in CYP1B1/CYP1A1 ratio may lead to a different ranking of a subject in the population (Fig. 4B and C). This is due to the fact that CYP1B1 and CYP1A1 catalyze different reactions in the metabolic pathway. Changing their ratio will have different results on the E₂-3,4-Q AUC for subjects with different composite haplotypes. Estrogens

have long been recognized as prime risk factor for the development of breast cancer, but their assessment has not progressed beyond traditional exposure data, such as parity, age at menarche and menopause, etc. Here, we present a novel approach that is based on the molecular analysis of mammary estrogen metabolism. The E₂-3,4-Q AUC is a plausible metabolic risk factor for breast cancer because E₂-3,4-Q has been identified as principal estrogen metabolite causing DNA adduct formation in experimental animals and E₂-3,4-Q-derived DNA adducts have been detected in human breast cancer tissues (8-10, 13, 41, 44, 62). Whether the E₂-3,4-Q AUC is an independent risk factor, as suggested by the present analysis, will need to be confirmed by a larger separate study. The value of E₂-3,4-Q AUC as a new metabolic-genetic risk factor may yet be in its combination with traditional measures of endogenous and exogenous estrogen exposure.

Table 2. CYP1A1, CYP1B1, and COMT allele and haplotype frequencies of age-matched study population

	Cases	Controls	P
No.	221	217	
Age (y)			
Mean	57.4	57.3	0.99
Median	56	57	
Allele frequency			
CYP1A1			
Codon 461			
Thr	0.955	0.956	
Asn	0.045	0.044	0.916
HWE	1.000	1.000	
Codon 462			
Ile	0.950	0.963	
Val	0.050	0.037	0.348
HWE	0.011	0.245	
CYP1B1			
Codon 48			
Arg	0.661	0.684	
Gly	0.339	0.316	0.455
HWE	1.000	0.871	
Codon 119			
Ala	0.649	0.682	
Ser	0.351	0.318	0.305
HWE	0.463	0.633	
Codon 432			
Val	0.423	0.433	
Leu	0.577	0.567	0.763
HWE	0.205	0.078	
Codon 453			
Asn	0.824	0.823	
Ser	0.176	0.177	0.971
HWE	0.362	0.657	
COMT			
Codon 108			
Val	0.516	0.505	
Met	0.484	0.495	0.740
HWE	0.358	0.889	
Haplotype frequency			
CYP1A1			
461Thr-462Ile	0.905	0.919	
461Thr-462Val	0.045	0.037	0.411
461Asn-462Ile	0.050	0.044	
CYP1B1			
48Arg-119Ala-432Val-453Asn	0.379	0.380	
48Gly-119Ser-432Leu-453Asn	0.302	0.261	0.630
48Arg-119Ala-432Leu-453Ser	0.164	0.151	
48Arg-119Ala-432Leu-453Asn	0.104	0.139	
48Gly-119Ser-432Val-453Asn	0.029	0.040	
48Gly-119Ser-432Leu-453Ser	0.004	0.011	
48Arg-119Ala-432Val-453Ser	0.003	0.011	
48Arg-119Ser-432Leu-453Ser	0.001	0.005	
48Arg-119Ser-432Val-453Asn	0.007	NO	
48Gly-119Ser-432Val-453Ser	0.004	NO	
48Arg-119Ser-432Leu-453Asn	0.003	NO	
48Gly-119Ala-432Val-453Asn	NO	0.002	

Abbreviation: NO, not observed.

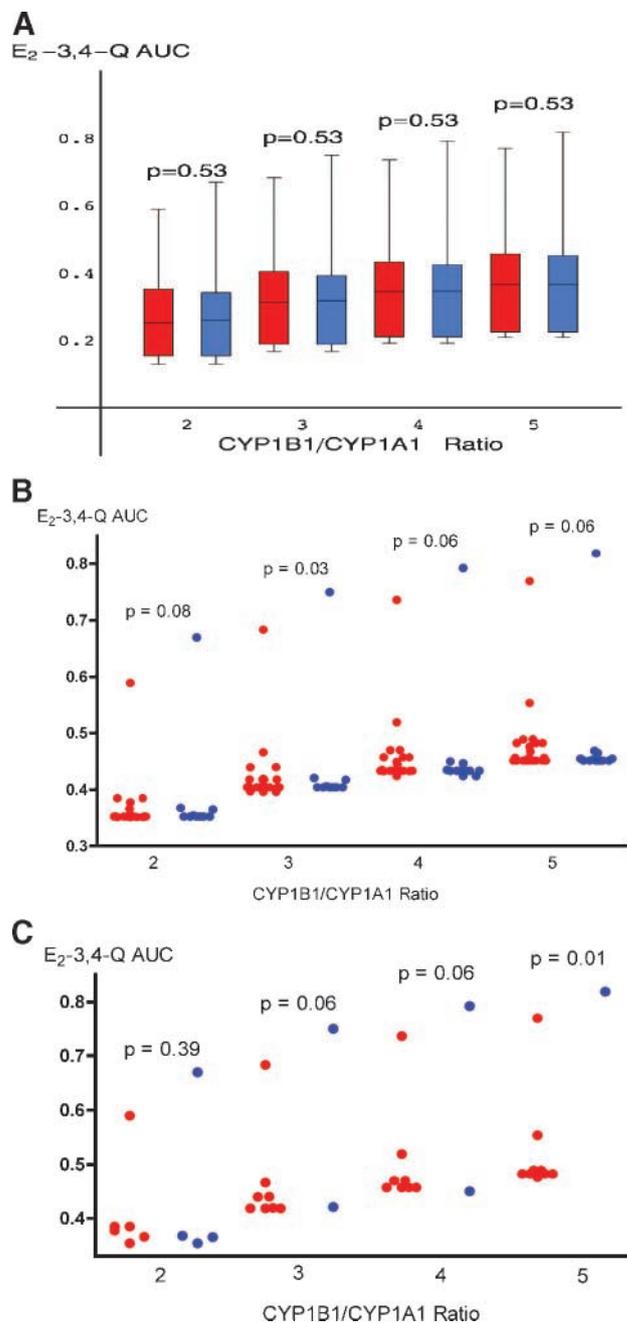


Figure 4. Correlation of E₂-3,4-Q AUC with CYP1B1/CYP1A1 ratio for cases and controls. **A.** Box and whisker graph of E₂-3,4-Q AUCs for entire population of 221 cases (red) and 217 controls (blue). Box, includes 84% of the respective group; whiskers, the top and bottom 8 percentiles. As indicated by the medians (center line in each box), the AUCs for cases and controls increase with increasing CYP ratio. However, there are no significant differences between case and control medians at any CYP ratio tested (see *P*s). **B.** Column scatter graph of E₂-3,4-Q AUCs for top 8 percentile (35 subjects) of entire study population. Dot, an individual case (red) or control (blue). Subjects with the same composite CYP1A1-CYP1B1-COMT enzyme haplotype have the same E₂-3,4-Q AUC. As the CYP ratio increases, their E₂-3,4-Q AUC changes in the same manner. However, subjects with different composite enzyme haplotypes may yield different E₂-3,4-Q AUC values, resulting in a change in their ranking with increasing CYP ratio. **C.** Column scatter graph of E₂-3,4-Q AUCs for top 2 percentile (10 subjects) of entire study population. There are significantly more cases (red) than controls (blue). *P* = 0.01 at CYP1B1/CYP1A1 = 5.

Table 3. Composite CYP1A1, CYP1B1, and COMT haplotypes of women with top 10 E₂-3,4-Q AUC values at CYP1B1/CYP1A1 ratio = 5

Rank	AUC	Subject	Age	CYP1A1	CYP1B1	COMT
x1	0.8069	Control	81	Asn-Ile	Arg-Ala-Val-Asn	Met
2	0.7542	Case	86	Asn-Ile	Arg-Ala-Leu-Asn	Met
				Asn-Ile	Arg-Ala-Leu-Asn	Val
				Asn-Ile	Arg-Ala-Leu-Ser	Val
				Asn-Ile	Arg-Ala-Leu-Asn	Met
3	0.5381	Case	68	Asn-Ile	Arg-Ala-Leu-Ser	Met
				Asn-Ile	Arg-Ala-Leu-Asn	Val
				Asn-Ile	Gly-Ser-Leu-Asn	Val
				Asn-Ile	Arg-Ser-Leu-Asn	Val
				Asn-Ile	Gly-Ala-Leu-Asn	Val
4	0.4803	Case	76	Thr-Ile	Arg-Ala-Leu-Asn	Met
				Thr-Val	Arg-Ala-Leu-Asn	Met
5	0.4800	Case	47	Thr-Ile	Arg-Ala-Val-Asn	Met
				Asn-Ile	Arg-Ala-Val-Asn	Met
6	0.4709	Case	59	Thr-Ile	Arg-Ala-Leu-Ser	Met
				Thr-Ile	Arg-Ala-Leu-Ser	Val
7	0.4709	Case	45	Thr-Ile	Arg-Ala-Leu-Ser	Met
				Thr-Ile	Arg-Ala-Leu-Ser	Val
8	0.4709	Case	53	Thr-Ile	Arg-Ala-Leu-Ser	Met
				Thr-Ile	Arg-Ala-Leu-Ser	Val
9	0.4709	Case	48	Thr-Ile	Arg-Ala-Leu-Ser	Met
				Thr-Ile	Arg-Ala-Leu-Ser	Val
10	0.4617	Case	56	Thr-Ile	Arg-Ala-Leu-Ser	Val
				Asn-Ile	Arg-Ala-Leu-Ser	Val

For example, one could estimate the overall exposure of a woman to E₂-3,4-Q AUC by taking into account (a) total years of menstruation or menopause age, (b) total pregnancy time, (c) years of menstruation before first full-term pregnancy, (d) body mass index, (e) dosage and duration of oral contraceptives, and (f) dosage and duration of hormone replacement therapy. Altogether, one could derive an individualized risk factor of estrogen exposure for each woman that combines her reproductive life history with her unique genetic and metabolic traits. Data on traditional variables related to estrogen exposure were unfortunately not obtained for all subjects of the present study population, such as the control subject who had the highest E₂-3,4-Q AUC in the entire population (Table 3).

In summary, using experimentally determined rate constants, we developed a mathematical model of mammary estrogen metabolism that allowed the kinetic simulation of E₂ and eight metabolites. The simulations showed excellent agreement with experimental results and provided a quantitative assessment of the metabolic interactions. The model permits the simulation of the carcinogenic estrogen quinones, whose transient nature prevents their direct quantitation. Using rate constants of genetic variants of CYP1A1, CYP1B1, and COMT, the model allows examination of the kinetic effect of enzyme polymorphisms on the entire pathway, including the identification of those haplotypes producing the largest amounts of catechols and quinones. We conceptually addressed the ambiguity of phenotypic information about enzyme concentration by varying the CYP1B1/CYP1A1 ratio. Application of the model to a breast cancer case-control population defined E₂-3,4-Q AUC as a potential risk factor. The model identified a subset of women with an increased risk of breast cancer based on their enzyme haplotype and consequent E₂-3,4-Q production. The model offers for the first time the opportunity to combine genetic, metabolic, and lifetime exposure data in assessing estrogens as breast cancer risk factor.

References

- Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. *Lancet* 1996;347:1713-27.
- The Endogenous Hormones and Breast Cancer Collaborative Group. Endogenous sex hormones and breast cancer in postmenopausal women: reanalysis of nine prospective studies. *J Natl Cancer Inst* 2002;94:606-16.
- Parl FF. Estrogens, estrogen receptor, and breast cancer. Amsterdam: IOS Press; 2000.
- Yager JD, Davidson NE. Estrogen carcinogenesis in breast cancer. *N Engl J Med* 2006;354:270-82.
- Stack DE, Byun J, Gross ML, Rogan EG, Cavalieri EL. Molecular characteristics of catechol estrogen quinones in reactions with deoxyribonucleosides. *Chem Res Toxicol* 1996;9:851-9.
- Akanni A, Abul-Hajj YJ. Estrogen-nucleic acid adducts: reaction of 3,4-estrone-*o*-quinone radical anion with deoxyribonucleosides. *Chem Res Toxicol* 1997;10:760-6.
- Cavalieri EL, Li KM, Balu N, et al. Catechol ortho-quinones: the electrophilic compounds that form depurinating DNA adducts and could initiate cancer and other diseases. *Carcinogenesis* 2002;23:1071-7.
- Li KM, Todorovic R, Devanesan P, et al. Metabolism and DNA binding studies of 4-hydroxyestradiol and estradiol-3,4-quinone *in vitro* and in female ACI rat mammary gland *in vivo*. *Carcinogenesis* 2004;25:289-97.
- Liehr JG, Fang WF, Sirbasku DA, Ari-Ulubelen A. Carcinogenicity of catechol estrogens in Syrian hamsters. *J Steroid Biochem* 1986;24:353-6.
- Li JJ, Li SA. Estrogen carcinogenesis in Syrian hamster tissues: role of metabolism. *Fed Proc* 1987;46:1858-63.
- Newbold RR, Liehr JG. Induction of uterine adenocarcinoma in CD-1 mice by catechol estrogens. *Cancer Res* 2000;60:235-7.
- Yager JD, Liehr JG. Molecular mechanisms of estrogen carcinogenesis. *Annu Rev Pharmacol Toxicol* 1996;36:203-32.
- Cavalieri EL, Stack DE, Devanesan PD, et al. Molecular origin of cancer: catechol estrogen-3,4-quinones as endogenous tumor initiators. *Proc Natl Acad Sci U S A* 1997;94:10937-42.
- Hayes CL, Spink DC, Spink BC, Cao JQ, Walker NJ, Sutter TR. 17 β -estradiol hydroxylation catalyzed by human cytochrome P450 1B1. *Proc Natl Acad Sci U S A* 1996;93:9776-81.
- Hanna IH, Dawling S, Roodi N, Guengerich FP, Parl FF. Cytochrome P450 1B1 (CYP1B1) pharmacogenetics: association of polymorphisms with functional differences in estrogen hydroxylation activity. *Cancer Res* 2000;60:3440-4.
- Bolton JL, Pisha E, Zhang F, Qiu S. Role of quinoids in estrogen carcinogenesis. *Chem Res Toxicol* 1998;11:1113-27.
- Liehr JG, Roy D. Free radical generation by redox cycling of estrogens. *Free Radic Biol Med* 1990;8:415-23.
- Nutter LM, Wu YY, Ngo EO, Sierra EE, Gutierrez PL, Abul-Hajj YJ. An *o*-quinone form of estrogen produces free radicals in human breast cancer cells: correlation with DNA damage. *Chem Res Toxicol* 1994;7:23-8.
- Han X, Liehr JG. Microsome-mediated 8-hydroxylation of guanine bases of DNA by steroid estrogens: correlation of DNA damage by free radicals with metabolic activation to quinones. *Carcinogenesis* 1995;16:2571-4.
- Shibutani S, Takeshita M, Grollman A. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. *Nature* 1991;349:431-4.
- Chakravarti D, Mailander PC, Li KM, et al. Evidence that a burst of DNA depurination in SENCAR mouse skin induces error-prone repair and forms mutations in the *H-ras* gene. *Oncogene* 2001;20:7945-53.
- Terashima I, Suzuki N, Shibutani S. Mutagenic properties of estrogen-quinone derived DNA adducts in Simian kidney cells. *Biochemistry* 2001;40:166-72.
- Dawling S, Roodi N, Mernaugh RL, Wang XY, Parl FF. Catechol-*O*-methyltransferase (COMT)-mediated metabolism of catechol estrogens: comparison of wild-type and variant COMT isoforms. *Cancer Res* 2001;61:6716-22.
- Goodman JE, Jensen LT, He P, Yager JD. Characterization of human soluble high and low activity catechol-*O*-methyltransferase catalyzed catechol estrogen methylation. *Pharmacogenetics* 2002;12:517-28.
- Strange RC, Spiteri MA, Ramachandran S, Fryer A. Glutathione-S-transferase family of enzymes. *Mutat Res* 2001;482:21-6.
- Hachey DL, Dawling S, Roodi N, Parl FF. Sequential action of phase I and II enzymes cytochrome P450 1B1 and glutathione S-transferase P1 in mammary estrogen metabolism. *Cancer Res* 2003;63:8492-9.
- Casorbi I, Brockmoller J, Roots I. A C4887A polymorphism in exon 7 of human CYP1A1: population frequency, mutation linkages, and impact on lung cancer susceptibility. *Cancer Res* 1996;56:4965-9.
- Ali-Osman F, Akande O, Antoun G, Mao J, Buolamwini J. Molecular cloning, characterization, and expression in *Escherichia coli* of full-length cDNAs of three human glutathione S-transferase Pi gene variants. *J Biol Chem* 1997;15:10004-12.
- Li DN, Seidel A, Pritchard MP, Wolf CR, Friedberg T. Polymorphisms in P450 CYP1B1 affect the conversion of estradiol to the potentially carcinogenic metabolite 4-hydroxyestradiol. *Pharmacogenetics* 2000;10:343-53.
- Shimada T, Watanabe J, Inoue K, Guengerich FP, Gillam EMJ. Specificity of 17 β -oestradiol and benzo[*a*]pyrene oxidation by polymorphic human cytochrome P4501B1 variants substituted at residues 48, 119, and 432. *Xenobiotica* 2001;31:163-76.
- Dawling S, Hachey DL, Roodi N, Parl FF. *In vitro* model of mammary estrogen metabolism: structural and kinetic differences in mammary metabolism of catechol estrogens 2- and 4-hydroxyestradiol. *Chem Res Toxicol* 2004;17:1258-64.

32. Kalyanaraman B, Sealy RC, Sivarajah K. An electron spin resonance study of osemiquinones formed during the enzymatic and autoxidation of catechol estrogens. *J Biol Chem* 1984;259:14018–22.
33. Dawling S, Roodi N, Parl FF. Methoxyestrogens exert feedback inhibition on cytochrome P450 1A1 and 1B1. *Cancer Res* 2003;63:3127–32.
34. Ritchie MD, Hahn LW, Roodi N, et al. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. *Am J Hum Genet* 2001;69:138–47.
35. Bailey LR, Roodi N, Verrier CS, Yee CJ, Dupont WD, Parl FF. Breast cancer and CYP1A1, GSTM1, and GSTT1 polymorphisms: evidence of a lack of association in Caucasians and African Americans. *Cancer Res* 1998;58:65–70.
36. Bailey LR, Roodi N, Dupont WD, Parl FF. Association of cytochrome P450 1B1 (CYP1B1) polymorphism with steroid receptor status in breast cancer. *Cancer Res* 1998;58:5038–41. [Erratum: *Cancer Res* 1999;59:1388].
37. Liu K, Muse S. PowerMarker: new genetic data analysis software version 3.0. Available from: <http://www.powermarker.net>.
38. Excoffier L, Slatkin M. Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol Biol Evol* 1995;12:921–7.
39. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. *Hum Hered* 2002;53:79–91.
40. Lewis DFV, Gillam EMJ, Everett SA, Shimada T. Molecular modelling of human CYP1B1 substrate interactions and investigation of allelic variant effects on metabolism. *Chem Biol Interact* 2003;145:281–95.
41. Embrechts J, Lemièrè F, Van Dongen W, et al. Detection of estrogen DNA-adducts in human breast tumor tissue and healthy tissue by combined nano LC-nano ES tandem mass spectrometry. *J Am Soc Mass Spectrom* 2003;14:482–91.
42. Liehr JG, Ricci MJ. 4-Hydroxylation of estrogens as marker of human mammary tumors. *Proc Natl Acad Sci U S A* 1996;93:3294–6.
43. Castagnetta LAM, Granata OM, Traina A, et al. Tissue content of hydroxyestrogens in relation to survival of breast cancer patients. *Clin Cancer Res* 2002;8:3146–55.
44. Rogan EG, Badawi AF, Devanesan PD, et al. Relative imbalances in estrogen metabolism and conjugation in breast tissue of women with carcinoma: potential biomarkers of susceptibility to cancer. *Carcinogenesis* 2003;24:697–702.
45. Tabakovic K, Gleason WB, Ojala WH, Abul-Hajj YJ. Oxidative transformation of 2-hydroxyestrone. Stability and reactivity of 2,3-estrone quinone and its relationship to estrogen carcinogenicity. *Chem Res Toxicol* 1996;9:860–5.
46. Tabakovic K, Abul-Hajj YJ. Reaction of lysine with estrone 3,4-*o*-quinone. *Chem Res Toxicol* 1994;7:696–701.
47. Iverson SL, Shen L, Anlar N, Bolton JL. Bioactivation of estrone and its catechol metabolites to quinoid-glutathione conjugates in rat liver microsomes. *Chem Res Toxicol* 1996;9:492–9.
48. Huang Z, Fasco MJ, Figge HL, Keyomarsi K, Kaminsky LS. Expression of cytochromes P450 in human breast tissue and tumors. *Drug Metab Dispos* 1996;24:899–905.
49. Spink DC, Spink BC, Cao JQ, et al. Differential expression of CYP1A1 and CYP1B1 in human breast epithelial cells and breast tumor cells. *Carcinogenesis* 1998;19:291–8.
50. Weisz J, Fritz-Wolz G, Gestl S, et al. Nuclear localization of catechol-*O*-methyltransferase in neoplastic and nonneoplastic mammary epithelial cells. *Am J Pathol* 2000;156:1841–8.
51. Cairns J, Wright C, Cattan AR, et al. Immunohistochemical demonstration of glutathione *S*-transferases in primary human breast carcinomas. *J Pathol* 1992;166:19–25.
52. Kelley MK, Engqvist-Goldstein A, Montali JA, Wheatley JB, Schmidt DE, Jr., Kauvar LM. Variability of glutathione *S*-transferase isoenzyme patterns in matched normal and cancer human breast tissue. *Biochem J* 1994;304:843–8.
53. Alpert LC, Schecter RL, Berry DA, et al. Relation of glutathione *S*-transferase α and μ isoforms to response to therapy in human breast cancer. *Clin Cancer Res* 1997;3:661–7.
54. Raftogianis R, Creveling C, Weinshilboum R, Weisz J. Estrogen metabolism by conjugation. *J Natl Cancer Inst Monogr* 2000;27:113–24.
55. Garte S, Gaspari L, Alexandrie AK, et al. Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev* 2001;10:1239–48.
56. Mitrunen K, Hirvonen A. Molecular epidemiology of sporadic breast cancer. The role of polymorphic genes involved in oestrogen biosynthesis and metabolism. *Mutat Res* 2003;544:9–41.
57. Zimniak P, Nanduri B, Pikula S, et al. Naturally occurring human glutathione *S*-transferase GSTP1-1 isoforms with isoleucine and valine in position 104 differ in enzymic properties. *Eur J Biochem* 1994;224:893–9.
58. Pike MC, Krailo MD, Henderson BE, Casagrande JT, Hoel DG. "Hormonal" risk factors, "breast tissue age," and the age-incidence of breast cancer. *Nature* 1983;303:767–70.
59. Dunning AM, Healey CS, Pharoah PDP, Teare MD, Ponder BAJ, Easton DF. A systematic review of genetic polymorphisms and breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 1999;8:843–54.
60. Thomas DC. The need for a systematic approach to complex pathways in molecular epidemiology. *Cancer Epidemiol Biomarkers Prev* 2005;14:557–9.
61. Conney AH. Introduction to drug-metabolizing enzymes: a path to the discovery of multiple cytochromes P450. *Annu Rev Pharmacol Toxicol* 2003;43:1–30.
62. Markushin Y, Zhong W, Cavalieri EL, et al. Spectral characterization of catechol estrogen quinone (CEQ)-derived DNA adducts and their identification in human breast tissue extract. *Chem Res Toxicol* 2003;16:1107–17.