[CANCER RESEARCH 61, 785–789, January 15, 2001]

Building a Multigenic Model of Breast Cancer Susceptibility: *CYP17* and *HSD17B1* Are Two Important Candidates[1]

Heather Spencer Feigelson,[2] Roberta McKean-Cowdin, Gerhard A. Coetzee, Daniel O. Stram, Laurence N. Kolonel, and Brian E. Henderson

American Cancer Society, National Home Office, Atlanta, Georgia 30329-4251 [H. S. F.]; Departments of Preventive Medicine [H. S. F., R. M-C., G. A. C., D. O. S., B. E. H.] and Urology [G. A. C.], University of Southern California Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, California 90033-0800; and Cancer Etiology Program, Cancer Research Center of Hawaii, University of Hawaii, Honolulu, Hawaii 96813 [L. N. K.]

ABSTRACT

We conducted a nested case-control study to evaluate whether polymorphisms in two genes involved in estrogen metabolism, *CYP17* and *HSD17B1*, were useful in developing a breast cancer risk model that could help discriminate women who are at higher risk of breast cancer. If polymorphisms in these genes affect the level of circulating estrogens, they may directly influence breast cancer risk. The base population for this study is a multiethnic cohort study that includes African-American, Non-Latina White, Japanese, Latina, and Native Hawaiian women. For this analysis, 1508 randomly selected controls and 850 incident breast cancer cases of the first four ethnic groups who agreed to provide a blood specimen were included (76 and 80% response rates, respectively). The *CYP17 A2* allele and the *HSD17B1 A* allele were considered "high-risk" alleles. Subjects were then classified according to number of high-risk alleles. After adjusting for age, weight, and ethnicity, we found that carrying one or more high-risk alleles increases the risk of advanced breast cancer in a dose-response fashion. The risk among women carrying four high-risk alleles was 2.21 [95% confidence interval (CI), 0.98–5.00; *P* for trend = 0.03] compared with those who carried none. This risk was largely limited to women who were not taking hormone replacement therapy (relative risk, 2.60; 95% CI, 0.95–7.14) and was most pronounced among those weighing 170 pounds or less (RR, 3.05; 95% CI, 1.29–7.25). These findings suggest that breast cancer risk has a strong genetic component and supports the theory that the underlying mechanism of "complex traits" can be understood using a multigenic model of candidate genes.

INTRODUCTION

Lander and Schork (1) defined "complex trait" as any phenotype that does not exhibit classic Mendelian inheritance attributable to a single gene locus. Such traits include susceptibilities to heart disease, hypertension, and cancer. Given the large and compelling body of epidemiological and experimental evidence that implicates estrogens in the etiology of human breast cancer, we have proposed a multigenic model of breast cancer predisposition that included genes involved in estrogen biosynthesis and intracellular binding (2). We hypothesized that functionally relevant polymorphisms in such genes would exhibit small, but additive, effects on individual susceptibility to breast cancer, and that specific combinations could result in a high-risk profile by influencing lifetime levels of estrogen.

One such gene, *CYP17*, encodes the cytochrome p450c17α enzyme, which mediates both steroid 17α-hydroxylase and 17,20-lyase activities, and functions at key branch points in human steroidogenesis (3). A single-bp polymorphism (T27C) in the 5′ untranslated region of *CYP17* (34 bp upstream from the initiation of translation and 27 bp downstream from the transcription start site) creates a recognition site for the *Msp*AI restriction enzyme and has been used to designate two alleles, *A1* (the published sequence) and *A2*. We and others have found that endogenous hormone levels are associated with this polymorphism (4, 5). Furthermore, several studies have examined the association with *CYP17* and breast cancer with mixed results (5–12). Most recently, we have shown that women who carry the *CYP17 A2/A2* genotype were about half as likely as women with the *A1/A1* genotype to be current HRT[3] users (13).

We have now examined the importance of a second candidate gene in this polygenic model: the 17β-hydroxysteroid dehydrogenase 1 (*HSD17B1*) gene. *HSD17B1* encodes the 17HSD type 1 enzyme that catalyzes the final step of estradiol biosynthesis, i.e., the conversion of estrone to the more biologically active estradiol. 17HSD type 1 is expressed in both normal and malignant breast epithelium (14). Several polymorphisms have been identified in *HSD17B1* including a common polymorphism in exon 6 that results in an amino acid change from serine (allele *A*) to glycine (allele *G*) at position 312 (14, 15). Although current evidence indicates that this amino acid change may not affect the catalytic or immunological properties of the enzyme (16), an early report suggested that individuals who were homozygous for serine were at marginally significantly increased risk for breast cancer (14).

We evaluated whether polymorphisms in these two steroid biosynthesis genes were useful in developing a breast cancer risk model that could help discriminate women who are at higher risk of breast cancer. If polymorphisms in these genes affect the level of circulating estrogens, they may directly influence breast cancer risk. We hypothesized that the effect of these polymorphisms would be most pronounced in women without other sources of estrogen, *i.e.*, lean women in whom peripheral conversion of androgens in the adipose tissue would be minimal and women who are not currently receiving HRT. We further evaluated whether these genetic components showed evidence of increased penetrance by stage at diagnosis, age at onset, or family history of breast cancer.

MATERIALS AND METHODS

Study Population. This nested case-control study is part of a large, ongoing, multiethnic cohort study in Hawaii and Los Angeles, California with an emphasis on diet and other lifestyle characteristics in the etiology of cancer. The cohort totals 215,251 men and women, ages 45–75 years at baseline, and includes African-Americans, Japanese, Native Hawaiians, Latinos, Non-Latino Whites, and small numbers of other racial/ethnic groups. The assembly of the cohort began in Spring 1993 and was completed in 1996. Drivers' license files were used in both Hawaii and Los Angeles County to establish a cohort that would was both ethnically and socioeconomically diverse. Two additional sources of participants were ultimately needed to reach our enrollment goals: in Hawaii, the voters' registration file was used to identify additional older Japanese women; in California, the Health Care Financing Administration was used to identify African-Americans of ages 65 and older. All cohort members

Received 5/15/00; accepted 11/7/00.

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.

[1] Supported by Grants 2FB-0212 and 4KB-0147 from the California Breast Cancer Research Program of the University of California and by National Cancer Institute Grants CA63464 and CA54281.

[2] To whom requests for reprints should be addressed, at American Cancer Society, National Home Office, 1599 Clifton Road, NE, Atlanta, GA 30329-4251. Phone: (404) 929-6815; Fax: (404) 327-6450; E-mail: hfeigels@cancer.org.

[3] The abbreviations used are: HRT, hormone replacement therapy; RR, relative risk; CI, confidence interval; lbs., pounds.

completed a mailed 26-page questionnaire at baseline that included information regarding medical history, family cancer history, diet, medication use, physical activity, and reproductive history, including the use of hormones. Further details of the cohort study are provided elsewhere (17).

We identified incident cancer cases through the population-based tumor registries in Los Angeles and Hawaii (both of which are members of the National Cancer Institute's Surveillance, Epidemiology, and End Results program) and the California State Tumor Registry. As of July 1, 1999, 1320 cases of incident breast cancer among the four larger ethnic groups (*i.e.*, excluding Native Hawaiians) have been identified from the cohort of whom 80% ($n = 1056$) agreed to provide a blood specimen. For this analysis, we included 850 women who were diagnosed with incident breast cancer of stage 1 or greater. Of these, 235 cases had stage 2 or higher tumors (regional and metastatic disease) and are classified here as "advanced" disease. We excluded 200 cases of breast carcinoma *in situ* and 6 cases of unknown stage. Cohort members who did not give a blood sample were similar with respect to age, ethnicity, and education level to those who did provide a sample, and participation rates were similar for stage 1 and advanced stage cases.

A random sample of men and women was generated to provide potential cohort controls. Controls were contacted by phone and asked to provide a blood specimen. As of July 1, 1999, 1984 female cohort members had been asked to donate a blood specimen. This study includes 1508 (76%) cohort controls who agreed to provide a blood specimen and reported no history of breast cancer. At the time of the blood draw, informed consent forms were completed by all participants. Controls who reported a personal history of breast cancer on the baseline questionnaire were excluded from this analysis.

Statistical Analysis. Data were analyzed using logistic regression methods to estimate RRs and 95% CIs. Age and ethnicity were included in the statistical models to adjust for possible differences in allele distribution. Weight was also included in the models based on the *a priori* assumption that small differences in serum hormone levels that can be attributed to genetic polymorphisms may be masked by the peripheral production of estrogens in the adipose tissue of postmenopausal women. The *CYP17 A2* allele and the *HSD17B1 A* allele were considered the "high-risk" alleles. Subjects were then classified according to number of high-risk alleles. For example, a woman whose genotype was *A1/A1* for *CYP17* and *G/G* for *HSD17B1* would be scored as zero high-risk alleles, a women with *A1/A2* and *A/G* would be scored as 2, and so on to the highest risk category: *A2/A2* and *A/A*, which would be scored as 4. This method of classification allows the computation of a test for trend. Although the high-risk alleles for *CYP17 (A2)* and *HSD17B1 (A)* may not equally affect risk of breast cancer, they were considered exchangeable in the allele counting based statistical model. Dummy variables were also created for each number of high-risk alleles and were entered into the logistic regression model to obtain risk estimates for each number of high-risk alleles.

Because our hypothesis is that these polymorphisms act by influencing lifetime levels of endogenous estrogens (from ovarian synthesis), we were interested in statistically testing whether these high-risk alleles could predict circulating estrogen levels (where estrogen is estrone + estradiol). Using plasma estrogen measurements from postmenopausal control women (who were not taking estrogen or progestogen in the 2 weeks prior to blood draw), we estimated the relationship between log estrogen and *CYP17* and *HSD17B1* as: log estrogen = constant + $\beta 1 \cdot$ (number of high-risk *CYP17* alleles) + $\beta 2 \cdot$ (number of high risk *HSD17B1* alleles). This is equivalent to log estrogen = constant + (number of high-risk *CYP17* alleles) + $(\beta 2/\beta 1) \cdot$ (number of *HSD17B1* high-risk alleles). Our fitted results gave an estimated $\beta 2/\beta 1 = 1.41$. We then used this fitted equation in a logistic regression model to estimate the risk of breast cancer per unit change in log estrogen as predicted by the genotypes of these two genes. Finally, we compared the likelihoods of the two models (allele counting *versus* predicted hormone equations) to determine whether the two models fit the data equivalently. If the likelihoods are similar, this suggests that the allele counting model is consistent with the apparent effect of the alleles on plasma estrogens.

We used the results of earlier studies (13, 18) to guide our stratified and subset analyses. Because a significant amount of circulating estrogen results from aromatization of androstenedione to estrone in peripheral tissues, we stratified the data by weight. An *a priori* cutpoint of 170 pounds (~80 kg) was chosen to reflect the nonlinearity of the body weight-endogenous hormone relationship (18). It has been shown that <170 pounds, there is no statistically significant association with body weight and endogenous hormone levels. Above 170 pounds, levels of endogenous estrogens rise in response to aromatase activity in adipose tissue.[4] This 170 pound cutpoint also corresponds with the 75th percentile of the weight distribution among our control women. We have reported that women with *CYP17 A2* alleles were less likely to use HRT (13), thus we stratified on HRT use. We also used stratified analysis to examine genotype by stage, age of onset, and family history, because these are often factors that can reflect gene expression. We examined the effect of *CYP17* and *HSD17B1* among women <55 years of age; 324 controls and 40 advanced cases were <55 years at the time of diagnosis (or time of blood draw among controls). Women were considered to have a positive family history of breast cancer if they reported having a mother or sister(s) with breast cancer on the baseline questionnaire. Thirty-nine women with advanced breast cancer and 162 controls reported a positive family history. Finally, to help assess the consistency of the association, the data were stratified by ethnicity.

Genotyping. DNA was purified from buffy coats of peripheral blood samples. The *CYP17* assay has been described previously (19). A PCR fragment containing the bp change was generated using the following primers: CYP-1, 5-CATTCGCACTCTGGAGTC-3; and CYP-2, 5-AGGCTCTTGGGGTACTTG-3. PCR reactions were carried out in 25-μl aliquots containing about 50 ng of genomic DNA, 50 pmol of each primer, 1× reaction buffer, 100 μM deoxynucleotide triphosphates, and 1 unit of Taq polymerase (Pharmacia). The amplification was for 30 cycles with denaturation at 94°C for 1 min, annealing at 57°C for 1 min, and extension at 72°C for 1 min. An initial denaturation step of 5 min at 94°C and a final extension at 72°C for 5 min were used. The PCR products were digested for 3 h at 37°C using *Msp*AI, separated by agarose gel electrophoresis, and stained with ethidium bromide to identify the bp change.

Because *HSD17B1* has an adjacent pseudo-gene, the PCR amplification was nested to insure that the pseudo-gene was not coamplified (14). The first PCR fragment was generated using the following primers: HSD1-F, 5′-CGGGAGCCGCTCTGGGGCGATCT-3′ (forward); and HSD1-R, 5′-GGTGCCACTGTGCTGATTTTTAAATTTTCT-3′ (reverse). The primers for the second PCR reaction were: HSD2-F, 5′-AAGCCGACCCTGCGCTACTTCAC-3′ (forward); and HSD2-R, 5′-TCTATCTTAATTAGCCACCCACAGC-3′ (reverse). The PCR reactions were carried out in 25-μl aliquots containing 50 pmol of each primer, 1× reaction buffer, 100 μM deoxynucleotide triphosphates, and 1 unit of Taq polymerase (Pharmacia) + DMSO (5% final). In the first PCR, ~50 ng of genomic DNA were used as a template; in the second PCR, 1 μl of the first PCR reaction mix was used as a template. The amplification was for 30 cycles for the first PCR and 19 cycles for the second PCR with denaturation at 94°C for 1 min, annealing at 60°C for 1 min, and extension at 72°C for 1.5 min for the first PCR and 1 min for the second PCR. An initial denaturation step of 3 min at 94°C and a final extension at 72°C for 5 min were used in both cases. The PCR products were digested for 3 h at 60°C using *Bst*UI, separated by agarose gel electrophoresis and stained with ethidium bromide to identify the bp change.

All samples were run in batches that contained positive controls (of each genotype) and negative controls (samples with no DNA added). All batches included both breast cancer cases and non-cases. In addition, 5% of the samples were repeated (blind) in subsequent batches, and results of both batches were accepted only if the duplicates were identical. Gels were read blind to case/control status.

RESULTS

The study included 536 African-Americans, 400 Japanese, 481 Latinas, and 402 non-Latina Whites. The mean age of controls (62.8) was similar to that of stage 1 and advanced stage cases (65.7 and 63.5, respectively). Controls were slightly (but not statistically significantly) heavier than controls. The mean weight for controls was 155 lbs., and 151 lbs. for both stage 1 and advanced stage cases. Thirty % of controls compared with 36% of cases were using HRT at baseline.

The data shown in Table 1 suggest that *CYP17* and *HSD17B1* both contribute to a modest increased risk of advanced breast cancer after adjusting for each other and for age, weight, and ethnicity, although

[4] N. Probst-Hensch, personal communication.

Table 1 *Independent age-, weight-, and ethnicity-adjusted RRs*[a] and 95% CIs for breast cancer by CYP17 and HSD17B1 genotypes

	Stage 1				Advanced		
	Controls	Cases	RR	95% CI	Cases	RR	95% CI
CYP17							
A1A1	542	222	1.00		70	1.00	
A1A2	739	287	0.95	0.77–1.18	122	1.28	0.93–1.75
A2A2	227	106	1.08	0.81–1.44	43	1.45	0.96–2.20
						P for trend = 0.05	
HSD17B1							
GG	369	154	1.00		45	1.00	
GA	739	297	1.00	0.79–1.27	126	1.39	0.96–2.00
AA	400	164	0.99	0.76–1.29	64	1.28	0.85–1.93
						P for trend = 0.28	
Total	1508	615			235		

[a] Logistic regression model includes *CYP17, HSD17B1,* age, weight, and ethnicity.

the CIs do not differ from 1.0. The test for linear trend is of borderline statistical significance for the *CYP17 A2* allele ($P = 0.05$) but not for the *HSD17B1 A* allele, although the magnitude of risk for the high-risk alleles are similar. Neither gene was associated with stage 1 breast cancer. In the logistic regression model, the risk factors (including *CYP17* and *HSD17B1*) are assumed to act in a multiplicative manner, and there was no statistical evidence that this assumption did not hold.

Table 2 shows that when the genes are considered together, there is evidence of a linear trend of increasing risk for advanced breast cancer as the number of high-risk alleles increases ($P = 0.03$). Women who carry four high-risk alleles (*i.e.,* those whose genotype is *CYP17 A2/A2* and *HSD17B1 A/A*) have a RR of 2.21 (95% CI, 0.98–5.00) for advanced breast cancer. Regardless of the number of high-risk alleles, there is no elevation of risk in women with stage 1 breast cancer.

To test our hypothesis that the effect of these polymorphisms would be most pronounced in women without other sources of estrogen, we stratified these data on HRT use. As shown in Table 2, the increasing risk of advanced breast cancer associated with the *CYP17* and *HSD17B1* high-risk alleles was largely limited to women who reported that they had never used HRT or were former HRT users (P for trend = 0.09).

The increase risk associated with *CYP17* and *HSD17B1* high-risk alleles is most pronounced among women weighing <170 lbs. (Table 3). Compared with having no high-risk alleles, women who are homozygous for the high-risk alleles of both genes have a >3-fold increased risk for developing advanced breast cancer (OR, 3.05; 95% CI, 1.29–7.25; P for trend = 0.02). In heavier women, the effect of genotype is still present but may be diluted by the contribution of estrogen production in the adipose tissue (data not shown). Using body mass index instead of weight did not affect the results.

This trend of increasing breast cancer risk with number of high-risk alleles was also evident in the model estimating risk of breast cancer/unit change in log estrogen as predicted by genotype. Compared with the *A1/A1, GG* genotype, we found that RR = 1.13 for the *A1/A1, GA* genotype; RR = 1.18 for the *A1/A2, GG* genotype; RR = 1.27 for the *A1/A1, AA* genotype; RR = 1.29 for the *A2/A2, GG* genotype; RR = 1.32 for the *A1/A2, GA* genotype; RR = 1.49 for the *A1/A2, AA* genotype; RR = 1.57 for the *A2/A2, GA* genotype; and RR = 1.76 for the *A2/A2, AA* genotype (P for trend = 0.038). The allele counting model gave a fit consistent with this model, where the *CYP17* and *HSD17B1* alleles are the independent variables in the prediction of plasma estrogens.

Ethnic-specific relative risks for advanced breast cancer by number of high-risk alleles are presented in Table 4 to illustrate that the genotype effects are consistent in each ethnic group included in the

Table 2 *Combined effect of CYP17 and HSD17B1 high-risk alleles on breast cancer risk*[a] *startified by stage and by HRT status*

	Number of high-risk alleles[b]					*P*
	0	1	2	3	4	(trend)
Controls ($n = 508$)	139	435	558	323	53	
Stage 1 cases ($n = 615$)	52	176	233	134	20	
RR[c]	1.00	1.12	1.13	1.11	.97	
95% CI		(0.77–1.63)	(0.79–1.63)	(0.75–1.64)	(0.52–1.81)	0.89
Advanced cases ($n = 235$)	15	62	87	58	13	
RR[d]	1.00	1.33	1.42	1.67	2.21	
95% CI		(0.73–2.41)	(0.79–2.54)	(0.91–3.05)	(0.98–5.00)	0.03
Former and never HRT users						
Controls ($n = 1025$)	90	298	379	222	36	
Advanced cases ($n = 147$)	9	40	55	34	9	
RR[d]	1.00	1.36	1.47	1.58	2.60	
95% CI		(0.63–2.92)	(0.70–3.09)	(0.72–3.43)	(0.95–7.14)	0.09
Current HRT users						
Controls ($n = 434$)	45	114	166	92	17	
Advanced cases ($n = 82$)	6	22	31	21	2	
RR[d]	1.00	1.48	1.43	1.77	0.81	
95% CI		(0.56–3.93)	(0.56–3.69)	(0.66–4.75)	(0.15–4.46)	0.58

[a] Age, weight, and ethnicity adjusted.
[b] High risk alleles: *CYP17, A2; HSD17B1, A.*
[c] Relative risk for stage 1 breast cancer.
[d] Relative risk for advanced breast cancer.

Table 3 *Combined effect of CYP17 and HSD17B1 high-risk alleles on risk*[a] *of advanced breast cancer among women weighing 170 lbs. or less*

	No. of high-risk alleles[b]					*P*
	0	1	2	3	4	(trend)
Subjects weighing ≤170 lbs.						
Controls ($n = 1116$)	109	319	406	247	35	
Advanced cases ($n = 178$)	13	44	64	44	13	
RR[c]	1.00	1.17	1.29	1.48	3.05	
95% CI		(0.61–2.26)	(0.68–2.43)	(0.76–2.86)	(1.29–7.25)	0.02

[a] Age, weight, and ethnicity adjusted.
[b] High risk alleles: *CYP17, A2; HSD17B1, A.*
[c] Relative risk for advanced breast cancer.

Table 4 *Relative risks for advanced breast cancer by number of CYP17 and HSD17B1 high-risk alleles by ethnicity*[a]

Sample size and no. of high-risk alleles	African-American	Latina	White	Japanese
Controls (n = 1508)	459	485	277	287
Stage 1 cases (n = 615)	135	138	170	172
Stage 2 cases (n = 235)	68	51	64	52
0[b]	1.00 (4)[c]	1.00 (2)	1.00 (2)	1.00 (7)
1	1.49 (19)	3.14 (20)	1.95 (13)	0.53 (10)
2	1.85 (27)	2.00 (16)	2.95 (29)	0.55 (15)
3	1.98 (15)	1.91 (10)	3.60 (16)	1.02 (17)
4	2.42 (3)	3.38 (3)	3.29 (4)	1.70 (3)

[a] Age and weight adjusted.
[b] High risk alleles: *CYP17, A2; HSD17B1, A.*
[c] Number of advanced stage cases by number of high-risk alleles.

study. For women carrying four high-risk alleles, the RR for advanced breast cancer ranges from 1.70 among Japanese women to 3.38 among Latina women.

We also examined the possible effects of family history and early age of onset. The data suggested that the *CYP17 A2* allele, but not the *HSD17B1 A* allele, may have a stronger effect among women with a family history compared with women with no affected first-degree relatives (data not shown). Risk of breast cancer by *CYP17* and *HSD17B1* genotype both appeared to be more pronounced among women who were <55 years at diagnosis (data not shown). However, for both family history and younger age at onset, we have an insufficient sample size to evaluate these data with certainty.

DISCUSSION

These results demonstrate the considerable promise that genetic polymorphisms along the endocrine pathway can be used, separately, but more importantly, in combination, to discriminate between women who have biologically different risks of more aggressive breast cancer. We found that, especially among women who are of average weight and among women who are not using HRT, *CYP17* and *HSD17B1* polymorphisms can be used to identify women at increased risk for advanced breast cancer. Huang *et al.* (10) have presented a similar model that includes polymorphisms in three genes: *CYP17, CYP1A1,* and *COMT*. Certainly, the complete model would include a number of such genes involved in hormone synthesis and degradation. The risk attributed to *CYP17* alone (RR = 1.45 for *A2/A2*) was more modest than in our initial publication (6) and more consistent with subsequent reports (8, 9, 11). Because our original study of *CYP17* was published, at least seven other studies have reported on *CYP17* and breast cancer (5, 7–12). The results of these studies are largely negative and suggest heterogeneity by ethnicity. However, there may be several reasons for the discrepant results. The study by Dunning *et al.* (7) is the largest to date but did not examine the possible confounding effects of HRT, which may mask the *CYP17*-breast cancer association. The results from Kristensen *et al.* (11) suggested that the effect of *CYP17* may be limited to older cases (*i.e.,* >55 years at diagnosis), whereas Bergman-Jungestrom *et al.* (12) found a strong association between *CYP17* and breast cancer in very young women. Three other smaller studies (8–10) found a modest elevation in breast cancer risk with the *CYP17 A2* allele in some subgroups, which is consistent with the *CYP17* association that we report here. However, in the only study to date that is both significantly larger than our original study and gives adequate consideration to potential confounding, Haiman *et al.* (5) did not find an association between *CYP17* and breast cancer. Additional studies must be of sufficient size and quality to detect the modest risk predicted for *CYP17 A2* in advanced breast cancer while examining the influence of HRT use and other potentially important confounders and effect modifiers.

The *HSD17B1* gene originally received much attention as a promising candidate gene for *BRCA1,* given its function and chromosomal location (14, 15). However, after it was eliminated as a candidate for *BRCA1,* it has received little attention, despite a rather compelling body of biochemical evidence for its important role in the synthesis of estradiol. To our knowledge, only one other epidemiological study of breast cancer has evaluated the role of this *HSD17B1* polymorphism (14). Our results warrant further exploration of *HSD17B1*.

Although the ethnic-specific risk estimates are imprecise, the data in Table 4 suggest that these findings are consistent across different ethnic groups. Because these ethnic groups traditionally differ in overall rates of breast cancer incidence, mortality, and factors such as socioeconomic status, access to medical care, parity, and HRT use, the similarities of these risk estimates are compelling. These loci, as well as others involved in the synthesis and metabolism of steroid hormones, may help explain the increasing risk of breast cancer in countries such as Japan, who are generally leaner than United States and European women.

The primary limitation of this study is its size. Although it includes over 850 cases of incident breast cancer, only 28% (235) of them are advanced cases. Further stratification by either weight or HRT use creates strata with small numbers, especially among advanced cases who are homozygous for both *CYP17* and *HSD17B1*. However, when we used this model to also estimate risk of breast cancer per unit change of log estrogen, the risk estimates are remarkably similar, adding confidence to the validity of the model.

Beyond the apparent internal consistency of these data, there is evidence that stage at diagnosis is dependent on the same combination of low- or high-risk alleles. Only 22% of cases who are double homozygotes for the low-risk alleles present with advanced disease, *versus* 39% of those homozygous for the high-risk alleles. The fact that increased risk from the high-risk alleles in limited to advanced cases of breast cancer may be explained at least two ways: (*a*) these advanced cases may have a different etiology; (*b*) these tumors may be more aggressive and progress to an advanced stage more rapidly. We are currently investigating whether survival or histopathology differ by (germ-line) *CYP17* and *HSD17B1* status. Although these genes do not appear to play an important role in localized disease, it is perhaps more relevant to look for characteristics to discriminate women who are at risk for advanced breast cancer, because these are the women who may most benefit from early interventions or be candidates for chemopreventive therapies.

It is yet to be determined whether these polymorphisms are in themselves functional or are linked to a variation elsewhere in the gene or other nearby locus. It was shown recently (11) that the T27C polymorphism in *CYP17* (converting the sequence CACT into CACC) does not influence Sp-1 binding in *in vitro* assays, as had been suggested based on its similarity to other known Sp-1 binding sequences. Additional functional studies are needed to determine whether the *A2* allele confers specifically a higher expression level of *CYP17* (11). Early work (16) on site-directed mutagenesis of *HSD17B1* failed to demonstrate changes in the catalytic or immunological properties of the enzyme resulting from this Ser→Gly change. However, one would not expect standard assays to necessarily detect the relatively small differences in activity predicted from the epidemiological data. For example, the model of breast tissue age by Pike *et al.* (20) demonstrates that a relative risk of 2.0 reflects only a 20% difference in levels of circulating estrogen. This 20% difference is generally consistent with what has been reported in the studies that have looked at the association between endogenous hormone levels and *CYP17* genotype (4, 5).

There is compelling evidence from this work and others that breast cancer risk has a strong genetic component. Some lifestyle factors, such as exercise or severe dietary change, may influence risk, either directly or by interrupting ovulatory function, and thus diminish the expression of this underlying genetic determination. In the same way, the use of exogenous hormones, or obesity, may add additional sources of estrogen, which would tend to augment this underlying susceptibility. Nevertheless, the further characterization of germ-line and somatic sequence variants in relevant genes holds promise for individualizing diagnosis, screening, and therapeutic intervention. We present this two-gene model as an example of how a multigene model can contribute to our understanding of, and ultimately the prevention of, diseases such as breast cancer.

ACKNOWLEDGMENTS

We thank Dr. Malcolm C. Pike for expert advice and assistance with the statistical analysis of these data.

REFERENCES

1. Lander, E. S., and Schork, N. J. Genetic dissection of complex traits. Science (Washington DC), *265:* 2037–2048, 1994.
2. Feigelson, H. S., Ross, R. K., Yu, M. C., Coetzee, G. A., Reichardt, J. K., and Henderson, B. E. Genetic susceptibility to cancer from exogenous and endogenous exposures. J. Cell. Biochem. Suppl., *25:* 15–22, 1996.
3. Brentano, S. T., Picado-Leonard, J., Mellon, S. H., Moore, C. C., and Miller, W. L. Tissue-specific, cyclic adenosine 3′,5′-monophosphate-induced, and phorbol ester-repressed transcription from the human P450c17 promoter in mouse cells. Mol. Endocrinol., *4:* 1972–1979, 1990.
4. Feigelson, H. S., Shames, L. S., Pike, M. C., Coetzee, G. A., Stanczyk, F. Z., and Henderson, B. E. Cytochrome P450c17α gene (*CYP17*) polymorphism is associated with serum estrogen and progesterone concentrations. Cancer Res., *58:* 585–587, 1998.
5. Haiman, C. A., Hankinson, S. E., Spiegelman, D., Colditz, G. A., Willett, W. C., Speizer, F. E., Kelsey, K. T., and Hunter, D. J. The relationship between a polymorphism in *CYP17* with plasma hormone levels and breast cancer. Cancer Res., *59:* 1015–1020, 1999.
6. Feigelson, H. S., Coetzee, G. A., Kolonel, L. N., Ross, R. K., and Henderson, B. E. A polymorphism in the *CYP17* gene increases the risk of breast cancer. Cancer Res., *57:* 1063–1065, 1997.
7. Dunning, A. M., Healey, C. S., Pharoah, P. D., Foster, N. A., Lipscombe, J. M., Redman, K. L., Easton, D. F., Day, N. E., and Ponder, B. A. No association between a polymorphism in the steroid metabolism gene *CYP17* and risk of breast cancer. Br. J. Cancer, *77:* 2045–2047, 1998.
8. Helzlsouer, K. J., Huang, H. Y., Strickland, P. T., Hoffman, S., Alberg, A. J., Comstock, G. W., and Bell, D. A. Association between *CYP17* polymorphisms and the development of breast cancer. Cancer Epidemiol. Biomark. Prev., *7:* 945–949, 1998.
9. Weston, A., Pan, C. F., Bleiweiss, I. J., Ksieski, H. B., Roy, N., Maloney, N., and Wolff, M. S. *CYP17* genotype and breast cancer risk. Cancer Epidemiol. Biomark. Prev., *7:* 941–944, 1998.
10. Huang, C-S., Chern, H-D., Chang, K-J., Cheng, C-W., Hsu, S-M., and Shen, C-Y. Breast cancer risk associated with genotype polymorphisms of the estrogen-metabolizing genes *CYP17*, *CYP1A1*, and *COMT*: a multigenic study on cancer susceptibility. Cancer Res., *59:* 4870–4875, 1999.
11. Kristensen, V. N., Haraldsen, E. K., Anderson, K. B., Lonning, P. E., Erikstein, B., Karesen, R., Gabrielsen, O. S., and Borresen-Dale, A-L. *CYP17* and breast cancer risk: the polymorphism in the 5′ flanking area of the gene does not influence binding to Sp-1. Cancer Res., *59:* 2825–2828, 1999.
12. Bergman-Jungestrom, M., Gentile, M., Lundin, A-C., Group, S-E. B. C., and Wingren, S. Association between *CYP17* gene polymorphism and risk of breast cancer in young women. Int. J. Cancer, *84:* 350–353, 1999.
13. Feigelson, H. S., McKean-Cowdin, R., Pike, M. C., Coetzee, G. A., Kolonel, L. N., Nomura, A. M., Le Marchand, L., and Henderson, B. E. Cytochrome P450c17α gene (*CYP17*) polymorphism predicts use of hormone replacement therapy. Cancer Res., *59:* 3908–3910, 1999.
14. Mannermaa, A., Peltoketo, H., Winqvist, R., Ponder, B., Kiviniemi, H., Easton, D., Poutanen, M., Isomaa, V., and Vihko, R. Human familial and sporadic breast cancer: analysis of the coding regions of the *17β-hydroxysteroid dehydrogenase 2* gene (*EDH17B2*) using a single-strand conformation polymorphism assay. Hum. Genet., *93:* 319–324, 1994.
15. Normand, T., Narod, S., Labrie, F., and Simard, J. Detection of polymorphisms in the estradiol *17β-hydroxysteroid dehydrogenase 2* gene at the *EDH17B2* locus on 17q11–q21. Hum. Mol. Genet., *2:* 479–483, 1993.
16. Puranen, T., Poutanen, M., Peltoketo, H., Vihko, P., and Vihko, R. Site-directed mutagenesis of the putative active site of human 17β-hydroxysteroid dehydrogenase type 1. Biochem. J., *304:* 289–293, 1994.
17. Kolonel, L., Henderson, B., Hankin, J., Nomura, A., Wilkens, L., Pike, M., Stram, D., Monroe, K., Earle, M., and Nagamine, F. A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am. J. Epidemiol., *151:* 346–357, 2000.
18. Probst-Hensch, N. M., Pike, M. C., McKean-Cowdin, R., Stanczyk, F. Z., Kolonel, L. N., and Henderson, B. E. Ethnic differences in postmenopausal plasma estrogen levels: high estrone levels in Japanese-American women despite low weight. Br. J. Cancer, *11:* 1867–1870, 2000.
19. Carey, A. H., Waterworth, D., Patel, K., White, D., Little, J., Novelli, P., Franks, S., and Williamson, R. Polycystic ovaries and premature male pattern baldness are associated with one allele of the steroid metabolism gene *CYP17*. Hum. Mol. Genet., *3:* 1873–1876, 1994.
20. Pike, M. C., Krailo, M. D., Henderson, B. E., Casagrande, J. T., and Hoel, D. G. Hormonal risk factors, breast-tissue age, and age-incidence of breast cancer. Nature (Lond.), *330:* 767–770, 1983.

Building a Multigenic Model of Breast Cancer Susceptibility: *CYP17* and *HSD17B1* Are Two Important Candidates

Heather Spencer Feigelson, Roberta McKean-Cowdin, Gerhard A. Coetzee, et al.

Cancer Res 2001;61:785-789.