

# A Polymorphism in the *CYP17* Gene Increases the Risk of Breast Cancer<sup>1</sup>

Heather Spencer Feigelson, Gerhard A. Coetzee, Laurence N. Kolonel, Ronald K. Ross, and Brian E. Henderson<sup>2</sup>

Departments of Preventive Medicine [H. S. F., R. K. R., B. E. H.] and Urology [G. A. C.], University of Southern California School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, CA 90033-0800, and Cancer Research Center of Hawaii, University of Hawaii at Manoa, Honolulu, Hawaii 96813 [L. N. K.]

## Abstract

We conducted a case-control study to determine whether a polymorphism in the *CYP17* gene was associated with risk of breast cancer. We found an increased risk of advanced breast cancer in women carrying an A2 allele. The odds ratio was 2.5 [95% confidence interval (CI), 1.07–5.94] for regional or metastatic disease. Among controls, the A1/A1 genotype was associated with a later age at menarche. The reduced risk of breast cancer associated with a later age of menarche was largely limited to A1/A1 women: odds ratio, 0.47 (CI, 0.22–0.98) for breast cancer and later age at menarche among A1 homozygotes compared with 0.80 (CI, 0.51–1.27) for A1/A2 and A2/A2 genotypes. These findings suggest that the *CYP17* genotype may be a biomarker for the onset of ovulation and advanced breast cancer risk.

## Introduction

The past 20 years of research have identified numerous risk and protective factors for breast cancer, many of which can be understood as measures of the cumulative exposure of the breast to estrogen and, perhaps, progesterone (1). The two recently identified breast cancer susceptibility genes, *BRCA1* and *BRCA2*, may cause as much as 90% of breast and ovarian cancer in some families, but probably no more than 5–10% of all breast cancer in the United States is attributable to these two loci (2). We and others have hypothesized the existence of other susceptibility genes which carry low absolute risk, but potentially high population-attributable risk, especially when considered in combination (3). One such class of genes is that which codes for enzymes or receptors that control the metabolism and intracellular transport of estrogens.

The *CYP17*<sup>3</sup> gene codes for the cytochrome P450c17 $\alpha$  enzyme. *CYP17* is located on chromosome 10, spans 6569 bp, and is divided into eight exons (4). The cytochrome P450c17 $\alpha$  enzyme mediates both steroid 17 $\alpha$ -hydroxylase and 17,20-lyase activities and functions at key branch points in human steroidogenesis (5). The 5' untranslated region of *CYP17* contains a single base pair polymorphism that creates an Sp1-type (CCACC box) promoter site 34 bp upstream from the initiation of translation, but downstream from the transcription start site (6). This base pair change also creates a recognition site for the *MspAI* restriction enzyme. *MspAI* digestion of a PCR fragment has been used to arbitrarily designate two alleles, A1 (the published sequence) and A2 (6). Since this base pair change creates a CCACC box and it is thought that the number of 5' promoter elements correlates with promoter activity (7), the A2 allele may result in an

increased rate of transcription. This polymorphism has not been examined in any breast cancer studies to date. However, if this polymorphic site affects or is in linkage disequilibrium with other genetic changes that cause variation of transcriptional activity of *CYP17*, it may, in turn, affect estradiol production and in this way may ultimately play a role in the etiology of breast cancer. We have tested this hypothesis preliminarily in a nested case-control study of Asian, African-American, and Latino women from Los Angeles, California and Hawaii.

## Materials and Methods

**Description of Cohort and Identification of Cases and Controls.** In 1993, we initiated a population-based cohort study of individuals ages 45–75 years who were recruited by mail with a projected cohort size in excess of 200,000. The cohort was accessed from the drivers' license files in both Los Angeles and Hawaii. In Los Angeles, primarily African-American and Latino men and women were targeted for enrollment, while in Hawaii, primarily Japanese and white men and women were targeted.

The initial mailing consisted of a 26-page questionnaire which collected information on diet, physical activity, prior history of specific medical illnesses, use of vitamins and selected drugs, reproductive history (from women), and family history of cancer. All respondents are being followed for incident cancers by matching with Surveillance, Epidemiology, and End Results (SEER) registries in the two locations as well as by active follow-up.

In both locations, biological samples (blood and urine) are being collected from cases of incident cancer of selected sites and from an approximately 3% random sample of healthy cohort members to serve as controls. Eligible controls for the present analysis were those female cohort members who had reported no history of cancer. Participation rates for the sample collection component in all populations has exceeded 70% thus far.

**Laboratory Methods.** As samples are collected, blood components are separated and stored in 0.5-ml aliquots at  $-80^{\circ}\text{C}$ . DNA was purified from buffy coats of peripheral blood samples for all cases and controls using a rapid DNA preparation method (8).

The *CYP17* assay has been described previously (6). Briefly, a PCR fragment containing the base pair change was generated using the following primers: CYP-1, 5'-CATTTCGCACCTCTGGAGTC-3', and CYP-2, 5'-GGCTCTGGGGTACTTG-3'. PCR reactions were carried out in 25- $\mu\text{l}$  aliquots containing about 50 ng of genomic DNA, 50 pmol of each primer, 1 $\times$  reaction buffer, 100  $\mu\text{M}$  deoxynucleotide triphosphates, and 1 unit of *Taq* polymerase. The amplification was for 30 cycles with denaturation at 94 $^{\circ}\text{C}$  for 1 min, annealing at 57 $^{\circ}\text{C}$  for 1 min, and extension at 72 $^{\circ}\text{C}$  for 1 min. An initial denaturation step of 5 min at 94 $^{\circ}\text{C}$  and a final extension at 72 $^{\circ}\text{C}$  for 5 min were used. The PCR products were digested for 3 h at 37 $^{\circ}\text{C}$  using *MspAI* and separated by agarose gel electrophoresis and staining with ethidium bromide to identify the base pair change.

**Data Analysis.** Stratified analysis and logistic regression models were used to determine the association between *CYP17* alleles and incident breast cancer while controlling for age and ethnicity, and to assess other potential confounders. Possible interaction between *CYP17* and reproductive characteristics was also evaluated. Where appropriate,  $\chi^2$  tests and ANOVA were used and their associated *P* values are shown.

Received 12/6/96; accepted 2/1/97.

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.

<sup>1</sup> This work was supported by State of California Grants 11B-0084 and 2FB-0212 and National Cancer Institute Grant CA-54281.

<sup>2</sup> To whom requests for reprints should be addressed, at Department of Preventive Medicine, University of Southern California/Norris Comprehensive Cancer Center, 1441 Eastlake Avenue, MS 44, Los Angeles, CA 90033-0800. Phone: (213) 764-0419; Fax: (213) 764-0127.

<sup>3</sup> The abbreviations used are: *CYP17*, cytochrome P450c17 $\alpha$  gene; OR, odds ratio; CI, 95% confidence interval; FFTP, first full-term pregnancy.

Table 1 *CYP17 genotype frequencies among Asian, African-American and Latino controls*

Ethnic group	Genotype frequency <sup>a</sup> (%)			Total
	A1/A1	A1/A2	A2/A2	
Asian	12 (26.1)	26 (56.5)	8 (17.4)	46
African-American	53 (38.7)	66 (48.2)	18 (13.1)	137
Latino	31 (30.4)	53 (52.0)	18 (17.6)	102

<sup>a</sup>  $\chi^2 = 3.547$ ,  $P = 0.471$ .

## Results

This study included 174 cases of incident breast cancer and 285 female cohort controls. The mean age of the cases was 63.0 years ( $\pm 8.4$  SD); the mean age of the control group was 61.4 years ( $\pm 8.3$  SD). The *CYP17* allelic distribution among controls is shown in Table 1. The allelic distributions in each ethnic group are in Hardy-Weinberg equilibrium (data not shown), and there were no statistically significant differences in the allelic frequency between the three ethnic groups ( $\chi^2 = 3.547$ ,  $P = 0.471$ ). Thus, the three groups were combined and analyzed together.

Results were similar for the A1/A2 and A2/A2 genotypes compared with the A1/A1 group: the age- and ethnicity-adjusted OR and 95% CIs for breast cancer were 1.33 (CI, 0.87–2.06) and 1.26 (CI, 0.70–2.28), respectively, for the A1/A2 and A2/A2 genotypes. Thus, the A1/A2 and A2/A2 genotypes have been combined in all analyses to increase the precision of the estimates presented. Table 2 shows the results for the association between *CYP17* genotype and breast cancer. Women who carry at least one A2 allele are at increased risk for breast cancer, although the result is not statistically significant (OR, 1.32; CI, 0.87–2.00). However, *CYP17* genotype is strongly associated with the risk of advanced breast cancer. Homozygous or heterozygous A2 women are two and one-half times more likely to have advanced breast cancer, defined by having either regional or metastatic disease (OR, 2.52; CI, 1.07–5.94).

Because of the role of *CYP17* in estrogen biosynthesis, we next determined whether the *CYP17* genotype was associated with age at menarche, age at menopause, or age at FFTP. Among controls, a higher proportion of A1/A1 women than A2-carrying women had a later age of menarche (38.3% versus 28.5%, respectively,  $\chi^2 = 3.06$ ,  $P = 0.080$ ), and the mean age of menarche among A1/A1 women was statistically significantly higher than in women carrying an A2 allele (13.4 versus 13.0,  $P = 0.047$ ). Women with menarche at age 13 years or older had a reduced, but not statistically significant, risk of breast cancer compared with women with earlier menarche (OR, 0.69; CI, 0.47–1.02; Table 3). Age at FFTP, but not age at menopause, was associated with the *CYP17* genotype. Thus, age at FFTP was treated as a confounder in the multivariate analysis. When we examined the possible interaction between age at menarche and *CYP17* genotype with risk of breast cancer, we found that the protective effect of older age at menarche was largely limited to women with the A1/A1 genotype. As shown in Table 3, A1/A1 women with an age of menarche 13 years or older had less than one-half the risk of breast cancer compared with women with an earlier age at menarche (OR, 0.47; CI, 0.22–0.98). Women with the A1/A2 or A2/A2 genotype, on the other hand, showed a much smaller and not statistically significant protective effect from later age at menarche (OR, 0.80; CI, 0.51–

1.27). Neither age at FFTP nor age at menopause showed evidence of interaction with the *CYP17* genotype.

## Discussion

This is the first study to our knowledge to investigate the possible role of *CYP17* in the etiology of breast cancer and represents the first substantial evidence of the role of genetic susceptibility to breast cancer through estradiol biosynthesis. We found a markedly increased risk of advanced breast cancer in women who carry at least one A2 allele (OR, 2.52; CI, 1.07–5.94).

Carey *et al.* (6) first suggested an association between this *CYP17* polymorphism and polycystic ovaries and male pattern baldness, conditions associated with elevated serum androgens. Because circulating estrogens result from conversion of androstenedione and testosterone by the aromatase enzyme, one would expect that increased production of such androgens would result in increased levels of serum estrogens.

Estradiol biosynthesis occurs in mammals in the adrenal glands, gonads, and placenta via metabolic pathways that involve several cytochrome P450 enzymes (4). Cholesterol may ultimately be converted by a variety of pathways to progestins, mineralocorticoids, glucocorticoids, androgens, and estrogens. The *CYP17* gene product is the key enzyme that determines the choice of these pathways. The 17 $\alpha$ -hydroxylase activity converts steroids to precursors of the glucocorticoid cortisol, and 17,20-lyase activity yields precursors to estradiol and testosterone (4). It is conceivable that changes in expression or activity of this enzyme may have significant effects on steroidogenesis.

The A2 allele corresponds to an additional Sp-1 type promoter site in the 5' untranslated region of *CYP17*. Since it is thought that the number of 5' promoter elements correlates with promoter activity, it might be expected that an additional CCACC site may influence promoter activity, thereby up-regulating transcription (6). Thus, the A2 allele may reflect an increased rate of gene transcription and ultimately higher estradiol levels as a result of greater enzyme activity. At this time, there are no functional data available as to whether this CCACC site in *CYP17* affects endocrine status. Further laboratory investigation is required to determine whether the A1/A2 polymorphic site in *CYP17* is, in fact, the causative mutation or is simply a marker for another mutation that affects enzymatic activity. Nevertheless, this polymorphism appears to be a marker of enzyme activity that can be easily determined using standard PCR methodology.

Our findings suggest that the *CYP17* genotype may be an important biomarker for the onset of ovulation (*i.e.*, menarche) in adolescents or the initiation of regular ovulatory cycles (possibly the most critical reproductive determinant on breast cancer risk). Previous studies (9) have shown that women with early menarche and rapid establishment of regular

Table 2 *ORs and 95% CIs for the association between breast cancer and CYP17 genotype for all cases and by disease stage*

Group	A1/A2 + A2/A2	A1/A1	OR <sup>a</sup>	95% CI	P
Controls	189	96	1.00		
Cases	125	49	1.32	0.87–2.00	0.193
Cases by stage					
Carcinoma <i>in situ</i> /localized	92	42	1.12	0.72–1.75	0.616
Regional/metastatic	33	7	2.52	1.07–5.94	0.034

<sup>a</sup> Adjusted for age and ethnicity.

Table 3 ORs and 95% CIs for the risk of breast cancer and age at menarche for all women combined and then stratified by CYP17 genotype

	All data (n = 459)					
	Cases	Controls	Unadjusted		Adjusted <sup>a</sup>	
			OR	95% CI	OR	95% CI
Menarche						
<13 yr	94	130	1.00		1.00	
≥13 yr	80	155	0.71	0.49–1.04	0.69	0.47–1.02
A1/A1 genotype (n = 145)						
	Cases	Controls	Unadjusted		Adjusted <sup>a</sup>	
			OR	95% CI	OR	95% CI
Menarche						
<13 yr	26	37	1.00		1.00	
≥13 yr	23	59	0.55	0.26–1.18	0.47	0.22–0.98
A1/A2 + A2/A2 genotype (n = 314)						
	Cases	Controls	Unadjusted		Adjusted <sup>a</sup>	
			OR	95% CI	OR	95% CI
Menarche						
<13 yr	68	93	1.00		1.00	
≥13 yr	57	96	0.81	0.50–1.32	0.80	0.51–1.27

<sup>a</sup> Adjusted for age, ethnicity, and age at FFTP.

menstrual cycles had an almost 4-fold increased risk for breast cancer when compared with women with later menarche and long duration of irregular cycles. Apter and Vihko (10), in a longitudinal study of 200 schoolgirls, found that girls with early menarche establish regular ovulatory cycles more quickly, and have higher serum estradiol levels, than girls with a later age at menarche. Ovarian steroids, most importantly estradiol and progesterone, can modulate the release of pituitary gonadotropins by negative and positive feedback effects exerted at the level of the anterior pituitary and the hypothalamus (11). This, then, is one possible way that *CYP17* may influence the timing of menarche or the establishment of regular ovulatory cycles.

Our finding that a later age at menarche may only be protective in women with the A1/A1 genotype has important implications for cancer prevention. It is possible that the *CYP17* A2 allele results in a large enough lifetime increase in estrogen exposure that it overrides the protective benefit of late menarche. Thus, its effects on breast cancer risk would be 2-fold. First, higher estradiol levels would lead to an earlier age at menarche. Second, in women who carry the A2 allele, even when menarche is delayed as a result of physical activity or other physiological mechanisms, higher lifetime estradiol levels may negate, or substantially diminish, the otherwise protective benefit of late menarche. If this is the case, women with an A2 allele may be good candidates for early intervention or possibly chemoprevention efforts.

The design of this study, a case-control study nested in a large cohort, minimizes many possible biases and issues of comparability of the case and control groups that are of concern in other study designs. It also allows us to expand the study as cases accrue so that we may address several issues that are still of concern. We plan an expanded study that will include a fourth ethnic group, non-Hispanic whites, and will have a sufficient sample size to conduct ethnicity-specific analysis to determine whether any differences exist in the association between *CYP17* and breast cancer that were not apparent in these preliminary analyses. This larger study will also allow us to analyze the A1/A2 and A2/A2 groups separately and determine whether risk varies based on carrying one or two A2 alleles and to evaluate additional clinical parameters and tumor characteristics such as histological grade, receptor status, and her-2/neu

expression. This study provides an important early step in defining a model of breast cancer that can explain individual susceptibility in terms of both underlying genetic susceptibility and endogenous estrogen exposure. Ultimately, such a model could help us better identify women who may be at increased risk of breast cancer.

#### Acknowledgments

We thank Leslie Kuang, Michael Henderson, David Richardson, and Wu Zhang for laboratory assistance, Kristine R. Monroe for data management, and Enrique Zelaya and Rose Rahman for sample collection.

#### References

- Feigelson, H. S., and Henderson, B. E. Estrogens and breast cancer. *Carcinogenesis* (Lond.), 17: 2279–2284, 1996.
- Castilla, L. H., Couch, F. J., Erdos, M. R., Hoskins, K. F., Calzone, K., Garber, J., Boyd, J., Lubin, M. B., Deshane, M. L., Brody, L. C., Collins, F. S., and Weber, B. Mutations in the *BRCA1* gene in families with early-onset breast and ovarian cancer. *Nat. Genet.*, 8: 387–391, 1994.
- Feigelson, H. S., Ross, R. K., Yu, M. C., Coetzee, G. A., Reichardt, J. K. V., and Henderson, B. E. Genetic susceptibility to cancer from exogenous and endogenous exposures. *J. Cell Biochem.*, in press, 1997.
- Picado-Leonard, J., and Miller, W. L. Cloning and sequence of the human gene for P450c17 (steroid 17 $\alpha$ -hydroxylase/17,20 lyase): similarity with the gene for P450c21. *DNA*, 6: 439–448, 1987.
- Brentano, S. T., Picado-Leonard, J., Mellon, S. H., Moore, C. C. D., 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.
- 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.
- Kadonaga, J. T., Jones, K. A., and Tjian, R. Promoter-specific activation of RNA polymerase II transcription by Sp1. *Trends Biochem. Sci.*, 11: 20–23, 1986.
- Talmud, P., Tybjaerg-Hansen, A., Bhatnagar, D., Mbewa, A., Miller, J. P., Durrington, P., and Humphries, S. Rapid screening for specific mutations in patients with a clinical diagnosis of familial hypercholesterolaemia. *Atherosclerosis*, 89: 137–141, 1991.
- Henderson, B. E., and Bernstein, L. Endogenous and exogenous hormonal factors. *In: J. R. Harris, M. E. Lippmann, M. Morrow, and S. Helman* (eds), *Diseases of the Breast*, pp. 185–200. Philadelphia: Lippincott-Raven, 1996.
- Apter, D., and Vihko, R. Early menarche, a risk factor for breast cancer, indicates early onset of ovulatory cycles. *J. Clin. Endocrinol. Metab.*, 57: 82–86, 1983.
- Bardin, C. W., and Thau, R. B. Hormonal regulation of the ovary. *In: J. B. West* (ed.), *Physiologic Basis of Medical Practice*, pp 921–933. Baltimore: Williams & Wilkins, 1985.