[CANCER RESEARCH 62, 1045–1049, February 15, 2002]

The Androgen Receptor CAG Repeat Polymorphism and Risk of Breast Cancer in the Nurses' Health Study[1]

Christopher A. Haiman,[2] Myles Brown, Susan E. Hankinson, Donna Spiegelman, Graham A. Colditz, Walter C. Willett, Philip W. Kantoff, and David J. Hunter

Departments of Epidemiology [C. A. H., S. E. H., D. S., G. A. C., W. C. W., D. J. H.], Biostatistics [D. S.], and Nutrition [W. C. W., D. J. H.], and the Harvard Center for Cancer Prevention [C. A. H., G. A. C., W. C. W., D. J. H.], Harvard School of Public Health, Boston, Massachusetts 02115; Channing Laboratory [S. E. H., G. A. C., W. C. W., D. J. H.] and Department of Medicine [M. B.], Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115; The Lank Center for Genitourinary Oncology [P. W. K.] and Department of Adult Oncology [M. B., P. W. K.], Dana-Farber Cancer Institute, Boston, Massachusetts 02115; and Department of Preventive Medicine [C. A. H.], USC/Norris Comprehensive Cancer Center, Los Angeles, California 90089

ABSTRACT

Shorter alleles of a polymorphic $[CAG]_n$ repeat in exon 1 of the androgen receptor (*AR*) have been associated with increased risk of prostate cancer and decreased risk of breast cancer. We prospectively assessed the association between the $[CAG]_n$ repeat polymorphism in the androgen receptor and breast cancer risk among Caucasian women in a case-control study nested within the Nurses' Health Study cohort (cases, $n = 727$; controls, $n = 969$). In addition, we assessed whether androgen receptor genotype influences endogenous steroid hormone levels in women and whether the associations between androgen receptor alleles and breast cancer risk differed according to established breast cancer risk factors. Women with one or more long *AR* $[CAG]_n$ repeat alleles (≥22 repeats) were not at increased risk of breast cancer [odds ratio (OR), 1.06; 95% confidence interval (CI), 0.83–1.35]. Significant associations were not observed between *AR* genotypes comprised of two short alleles ($[CAG]_n$ ≤20 *versus* both alleles ≥22: OR, 0.92; 95% CI, 0.62–1.36) or two long alleles ($[CAG]_n$ ≥ 25 *versus* both alleles ≤ 22: OR, 1.42; 95% CI, 0.81–2.50) and breast cancer risk. We also observed no strong overall association between average repeat length and breast cancer risk (OR, 1.04 per CAG repeat; 95% CI, 0.99–1.10) or between average repeat length and plasma hormone levels. We also examined the cross-classification of *AR* genotype and first-degree family history of breast cancer. Compared with women with both alleles <22 and no family history, we observed a significant positive association limited to women with both a first-degree family history of breast cancer and longer alleles (one or two $[CAG]_n$ alleles ≥22; OR, 1.70; 95% CI, 1.20–2.40; *P* for interaction = 0.04). In summary, we observed no overall relation of *AR* genotype with breast cancer risk among mostly postmenopausal Caucasian women. However, these data suggest that longer *AR* $[CAG]_n$ repeat alleles may increase breast cancer risk among women with a first-degree family history of breast cancer.

INTRODUCTION

Estrogen and progesterone regulate multiple normal physiological processes, including the growth and development of the mammary gland (1). Through binding to their respective receptors, these hormones stimulate breast cell proliferation, resulting in greater opportunity for the occurrence and clonal propagation of nucleotide sequence errors caused by carcinogens or errors in DNA replication. The established reproductive breast cancer risk factors, such as early age at menarche and late age at menopause, provide an indirect link between dose and duration of exposure to estrogens and progesterone and breast cancer risk (2). The most direct evidence supporting the role of endogenous steroid hormones in the etiology of breast cancer has come from prospective epidemiological studies in which plasma estrogen levels are higher in women subsequently diagnosed with breast cancer compared with those who remain disease free (3–6).

Among postmenopausal women, androgen levels are also positively associated with breast cancer risk (4–9). The role of endogenous androgens in breast cancer development, however, is less clear, and hypotheses have emerged surrounding their etiological potential (2, 10). Androgens may influence breast cancer risk indirectly through their conversion to estradiol; directly, by binding to the AR[3] and either promoting or opposing breast cell growth, or through binding to ER-α and directly stimulating breast cell proliferation (11). Epidemiological studies support the indirect hypothesis of androgen action; the strength of the association between total testosterone levels and breast cancer risk weakens after controlling for estradiol levels (5–7). Positive associations between adrenal androgens DHEA, DHEAS and androstene-3β, 17β-diol and breast cancer risk have also been observed (4, 6, 8, 9), although whether this effect is mediated by the AR, through their conversion to testosterone, is unclear. Laboratory studies provide support for the indirect and direct effects of adrenal androgens in breast cancer development. Physiologically relevant androgen concentrations have been shown to activate ER-α and stimulate proliferation of breast cancer cell lines (11), and intratumoral aromatase activity has been shown to promote *in situ* conversion of androgens to estrogens and tumor growth in animal models (12).

On the basis of the hypothesized roles of endogenous androgens in breast cancer development, variation in genes involved in androgen metabolism or hormone signaling may influence susceptibility to breast cancer. Genetic modification of the *AR* has been studied extensively in association with prostate cancer (13–16). An association has been reported between CAG repeat number within exon 1 of the *AR* and prostate cancer risk, with an increase in risk observed for shorter $[CAG]_n$ repeat lengths (14–16). *In vitro* studies also provide evidence of biological function for the repeat polymorphism as longer alleles have decreased AR transactivational activity (17, 18). Genetic variation in *AR* repeat length has also been evaluated in relationship with breast cancer risk (19–24). Among *BRCA1* mutation carriers, Rebbeck *et al.* (19) observed that women with longer *AR* alleles were diagnosed with breast cancer at a significantly younger age, suggesting that androgens may modulate breast cancer penetrance among women predisposed to early-onset breast cancer. However, subsequent studies among women from high-risk breast cancer families (20, 21) and most (23, 24) but not all (22) studies of sporadic breast cases do not support the association between *AR* $[CAG]_n$ repeat length and breast cancer risk.

We conducted a prospective, case-control study nested within the Nurses' Health Study cohort to evaluate *AR* $[CAG]_n$ repeat allele length in relation to breast cancer risk among predominantly post-

Received 9/24/01; accepted 12/12/01.

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 NIH Grants CA40356, CA49449, EH00002, CA65725, CA80111, and CA57374.

[2] To whom requests for reprints should be addressed, at University of Southern California USC/Norris Comprehensive Cancer Center, 1441 Eastlake Avenue, Room 4411, Los Angeles, CA 90089-9175.

[3] The abbreviations used are: AR, androgen receptor; ER, estrogen receptor; BMI, body mass index; DHEA, dehydroepiandrosterone; DHEAS, DHEA sulfate; OR, odds ratio; CI, confidence interval.

menopausal Caucasian women. We also examined whether *AR* repeat alleles modified cancer risk associated with established breast cancer risk factors and whether repeat length is associated with circulating levels of endogenous steroid hormones.

MATERIALS AND METHODS

Study Population. The Nurses' Health Study was initiated in 1976, when 121,700 United States registered nurses between the ages of 30 and 55 returned an initial questionnaire reporting medical histories and baseline health-related exposures. Updated information has been obtained by questionnaire every 2 years, including data on reproductive variables, oral contraceptive and postmenopausal hormone use, cigarette smoking, and since 1980, dietary intake. Incident breast cancers were identified by self-report and confirmed by medical record review. Between 1989 and 1990, blood samples were collected from 32,826 women. Approximately 97% of the blood samples were returned within 26 h of blood draw, immediately centrifuged, aliquoted into plasma, RBCs, and buffy coat fractions, and stored in liquid nitrogen freezers. Follow-up has been >90% in all subsequent questionnaire cycles for this subcohort.

Eligible cases in this study consisted of women with pathologically confirmed incident breast cancer from the subcohort who gave a blood specimen. Cases with a diagnosis any time after blood collection up to June 1, 1996 with no previously diagnosed cancer except for nonmelanoma skin cancer were included. Controls were randomly selected participants who gave a blood sample and were free of diagnosed cancer (except nonmelanoma skin cancer) up to and including the interval in which the case was diagnosed. Controls were matched to cases on year of birth, menopausal status, postmenopausal hormone use, as well as time of day, month, and fasting status at blood draw. Women were defined as postmenopausal at the time of a bilateral oophorectomy or after having no menstrual cycle within the last 12 months prior to blood draw. Women who had a hysterectomy with one or both ovaries left intact were classified as premenopausal until the age at which 10% of the cohort had undergone natural menopause (46 years for smokers and 48 years for nonsmokers) and as postmenopausal at the age at which 90% of the cohort had undergone natural menopause (54 for smokers and 56 for nonsmokers); in the intervening years, these women were classified as being of uncertain menopausal status. For postmenopausal cases not using postmenopausal hormones within 3 months prior to blood draw, we matched a second control to increase statistical power in plasma steroid hormone analyses. The nested case-control study consists of 727 incident breast cancer cases and 969 matched controls. The study sample for the plasma hormone analysis is composed of 474 postmenopausal controls not using hormone replacement therapy within 3 months of blood draw. The protocol was approved by the Committee on Human Subjects at Brigham and Women's Hospital.

Information regarding breast cancer risk factors was obtained from the 1976 baseline questionnaire, subsequent biennial questionnaires, and a questionnaire completed at the time of blood sampling. Menopausal status and use of postmenopausal hormones was assessed at blood draw and updated until date of diagnosis for cases and the equivalent date for matched controls. Histopathological characteristics such as tumor stage and size and estrogen and progesterone receptor status were ascertained from medical records and used in case subgroup analyses when available.

Laboratory Methods. DNA was extracted from buffy coat fractions using the Qiagen QIAamp blood kit (Qiagen, Chatsworth CA). Oligonucleotide primers flanking the $[CAG_n]$ repeat were constructed (5′-TCCAGAATCTGTTCCAGAGCGTGC-3′ and 5′-GCTGTGAAGGTTGCTGTTCCTCAT-3′). Genomic DNA was PCR amplified using fluorescently labeled primers. The length of these fragments varied by the number of CAG repeats. The fragments were run on a 5% denaturing polyacrylamide gel, and amplicon length was determined relative to a Gene Scan size standard (Applied Biosystems) by automated fluorescence detection (Genescan: Applied Biosystems, Foster City, CA) in the Dana-Farber Cancer Institute Molecular Biology Core Facility and the Harvard Center for Cancer Prevention Laboratory. Genotyping was performed by laboratory personnel blinded to case-control status, and blinded quality control samples were inserted to validate genotyping procedures; concordance for the blinded samples was 100%.

Steroid hormones estrone sulfate, estrone, estradiol, testosterone, androstenedione, DHEA, and DHEAS were assayed in four separate batches. Estrone sulfate from batches 1 and 2 were assayed in the laboratory of Dr. C. Longcope (University of Massachusetts Medical Center, Worcester, MA). All other analyses were performed by Nichols Institute (San Juan Capistrano, CA). The intra-batch coefficients of variation were ≤13.6%. With the exception of estrone sulfate, estrone, testosterone, and DHEA, there was little or no batch-to-batch variation. For the hormones with batch-to-batch variation, the means for identical quality control samples varied by 23–46% between batches. Hence, to account for this variation, we controlled for batch in all analyses. Methods for plasma hormone assays and information regarding laboratory precision and reproducibility have been published previously (6, 25).

Statistical Analysis. The Wilcoxon signed-rank test was used to compare the distributions of the average repeat length {($[CAG]_n$ of allele 1 + allele 2)/2} and the shorter and the longer of a woman's two *AR* $[CAG]_n$ alleles among cases and controls. A global test to evaluate differences in allele frequencies between cases and controls was performed using a Mantel Haenszel χ^2 test across matched case-control sets. This test was also used to evaluate case-control differences in the frequency of *AR* alleles using the cutpoint designated by Spurdle *et al.* (Ref. 23; $[CAG]_n \geq 22$ *versus* <22), corresponding closely to the average $[CAG]_n$ repeat length in our population.

The relationship between *AR* repeat length and phenotypic variation in waist:hip ratio, a physiological characteristic for which androgens may play a role, was evaluated by age-adjusted Spearman correlation. ORs and CIs were calculated using conditional and unconditional logistic regression. In conditional analyses, 9 controls were excluded due to incomplete matching. In the majority of analyses, we used the cutpoint $[CAG]_n \geq 22$ repeats, which is the average repeat length among our controls. Indicator variables for all three genotypes (0 alleles ≥22, 1 allele ≥22, and 2 alleles ≥22) were created using 0 alleles ≥22 as the reference category in multivariate models. Genotype was evaluated as a dichotomous variable with one or two alleles ≥22 combined and as the linear effect/CAG repeat of the shorter, longer, and average repeat lengths. We also evaluated cutoff points throughout the $[CAG]_n$ repeat (19–29) and average repeat length (19–27) distributions. In addition to the matching variables, we adjusted for the breast cancer risk factors BMI (kg/m^2) at age 18 (continuous), weight gain since age 18 (<5 kg, 5–19.9, ≥20), age of menarche (<12 years, 12, 13, >13), parity/age at first birth (nulliparous, 1–2 children/age at first birth ≤24 years, 1–2 children/age at first birth >24, 3+ children/age at first birth ≤24, 3+ children/age at first birth >24), first-degree family history of breast cancer (yes/no), history of benign breast disease (yes/no), and duration of postmenopausal hormone use (never; past user <5 years, ≥5; current user, <5 years and ≥5). We also adjusted for age at menopause (continuous in years) in analyses limited to postmenopausal women. Fisher's Exact Test was used in case-case analyses to evaluate associations between case characteristics [histological subtype, stage of disease, and ER (ER+/−) and progesterone receptor (+/−) status] and *AR* genotype. Unconditional logistic regression models, including terms for the matching variables and other potential confounders, were used to assess the association between *AR* repeat alleles and breast cancer risk when stratifying by established breast cancer risk factors. Interactions were evaluated by including interaction terms between genotype and risk factor variables in unconditional multivariate logistic regression models. The likelihood ratio test was used to assess the statistical significance of these interactions.

Linear regression models were used to evaluate the association between *AR* repeat length and geometric mean plasma hormone levels among controls, controlling for BMI at blood draw, laboratory batch, and the matching variables. Genotype was evaluated as the linear effect (percent change in hormone level) per average CAG repeat length increment. The natural logarithm of the plasma hormone values was used in the analyses to reduce the skewness of the regression residuals. Hormones were measured in four different batches, and all hormones were not assayed for all women due to insufficient plasma. Subjects with hormone levels below detectable limits ($n = 6$) were assigned the lowest detectable value (estrone, $n = 3$; testosterone, $n = 1$; DHEAS, $n = 2$). Within each batch, hormone values greater than three times the interquartile range above the 75th percentile were treated as outliers and excluded (estrone sulfate, $n = 11$; estrone, $n = 3$; estradiol, $n = 8$; testosterone, $n = 2$; androstenedione, $n = 1$; DHEA, $n = 3$; DHEAS, $n = 1$). We used the Statistical Analysis System (SAS) for all analyses (26).

RESULTS

There were 304 premenopausal and 1236 postmenopausal women with mean ages of 48.5 (SD, 3.1) and 61.0 (SD, 5.0), respectively, at blood draw. Compared with controls, cases had the same mean BMI at blood draw (25.5 *versus* 25.6 kg/m^2) and similar ages at menarche (12.5 *versus* 12.6 years), first birth (25.0 *versus* 24.9 years) and menopause (48.0 *versus* 48.2 years). The proportion of women with a first-degree family history of breast cancer was significantly higher among the cases (21% *versus* 15%; $P = 0.001$). The distribution of *AR* alleles for cases and controls was similar (Fig. 1) as was the average repeat lengths (cases, 22.0; 95% CI, 21.8–22.1; controls, 21.8; 95% CI, 21.7–22.0; $P = 0.15$). The mean repeat lengths of the shorter and longer alleles were nearly identical between cases and controls (shorter allele: cases, 20.4; 95% CI, 20.2–20.5; controls, 20.3; 95% CI, 20.1–20.4; $P = 0.35$; longer allele: cases, 23.6; 95% CI, 23.4–23.8; controls, 23.4; 95% CI, 23.3–23.6; $P = 0.18$). We observed no significant difference in allele frequencies between cases and controls dichotomized by $[CAG]_n \geq 22$ ($\chi^2 = 0.007$; df = 1; $P = 0.93$), and the expected genotype distributions based on the allele frequencies among controls were similar to the observed distributions.

The adjusted OR for the effect per average $[CAG]_n$ repeat increment was 1.04 (95% CI, 0.99–1.10). We observed similar results for the shorter (OR, 1.04; 95% CI, 0.99–1.08) and longer (OR, 1.02; 95% CI, 0.98–1.07) repeat alleles. Compared with women with both alleles <22, those with one or two alleles ≥22 did not have a substantially increased risk of breast cancer (Table 1). No material change in relative risk estimates was evident when cutpoints above 22 were selected (women with 1 or 2 alleles ≥23, 25, 27, and 29 compared with those with two short alleles, both alleles <22; Table 1). Among premenopausal and postmenopausal women, no significant association was observed between having one or more $[CAG]_n$ repeat alleles ≥22 and breast cancer risk (OR, 1.38; 95% CI, 0.68–2.82 and OR, 1.01; 95% CI, 0.78–1.31, respectively) or when alternate allele cutpoints were evaluated (Table 1). Risk was not significantly greater for women homozygous for long alleles (*versus* both alleles <22; both alleles ≥25: OR, 1.42; 95% CI, 0.81–2.50). Similarly, an association with breast cancer risk was not evident among women with two short alleles compared with those with both alleles ≥22 (both alleles ≤20: OR, 0.92; 95% CI, 0.62–1.36; ≤19: OR, 1.06; 95% CI, 0.61–1.86). We also did not observe a consistent elevation in breast cancer risk for cutpoints along the average *AR* allele length distribution (Table 1). Using the same cutpoints, no associations were observed for shorter or longer repeat alleles (data not shown).

In analyses by breast cancer histological subtype and clinical parameters, carriers of one or more $[CAG]_n$ alleles ≥22 were not significantly overrepresented among invasive cancers (OR, 1.10; 95% CI, 0.86–1.40). We observed no difference in the frequency of longer alleles, shorter alleles, or the number of women with an average repeat length ≥22 by estrogen or progesterone receptor status, number of involved lymph nodes, tumor size, or degree of differentiation (all *P*s ≥ 0.3). We also did not observe a significant correlation between *AR* repeat length and waist:hip ratio ($r = 0.05$; $P = 0.17$), a physiological factor associated with higher circulating androgen levels.

We next evaluated the relationships between *AR* alleles and the risk of breast cancer according to established breast cancer risk factors. We observed no significant interactions between longer *AR* alleles ($[CAG]_n \geq 22$) and BMI among postmenopausal women ($P = 0.92$, test for interaction), history of benign breast disease ($P = 0.60$, test for interaction), age at menarche ($P = 0.71$, test for interaction), age at first birth ($P = 0.96$, test for interaction), parity ($P = 0.39$, test for interaction), oral contraceptive use ($P = 0.19$, test for interaction), or hormone replacement therapy ($P = 0.10$, test for interaction).

Because the data of Rebbeck *et al.* (19) suggest that *BRCA1*-associated breast cancer risk may be modified by *AR* $[CAG]_n$ repeat length, we examined the potential interaction between *AR* genotype

Table 1 *Association between androgen receptor CAG polymorphism and breast cancer risk, Nurses' Health Study, 1990–1996*

	Cases	Controls	OR (95% CI)	Adjusted OR (95% CI)
All women				
No. of alleles ≥22[a]				
0	179	247	1.00 (ref.)[b]	1.00 (ref.)[c]
1	374	481	1.03 (0.81–1.31)	1.04 (0.81–1.35)
2	174	232	0.99 (0.74–1.32)	1.08 (0.80–1.46)
1 or 2	548	713	1.02 (0.81–1.28)	1.06 (0.83–1.35)
Both alleles <22, 1–2 alleles	179	249	1.00 (ref.)[d]	1.00 (ref.)[e]
≥23	461	610	1.01 (0.80–1.28)	1.06 (1.22–1.34)
≥25	257	304	1.15 (0.89–1.50)	1.22 (0.93–1.60)
≥27	86	97	1.18 (0.82–1.69)	1.25 (0.86–1.81)
≥29	30	35	1.05 (0.60–1.82)	1.09 (0.62–1.94)
Premenopausal women				
Both alleles <22, 1–2 alleles	23	27	1.00 (ref.)[d]	1.00 (ref.)[f]
≥23	62	63	1.17 (0.60–2.29)	1.35 (0.64–2.84)
≥25	32	33	1.19 (0.55–2.56)	1.31 (0.57–3.03)
≥27	11	8	1.86 (0.58–5.96)	2.80 (0.66–11.87)
Postmenopausal women				
Both alleles <22, 1–2 alleles	147	208	1.00 (ref.)[d]	1.00 (ref.)[g]
≥23	356	506	0.96 (0.74–1.24)	0.98 (0.76–1.28)
≥25	196	248	1.09 (0.81–1.46)	1.13 (0.83–1.52)
≥27	63	80	1.05 (0.70–1.57)	1.08 (0.71–1.64)
All women, average repeat number[h]				
<22	340	473	1.00 (ref.)[d]	1.00 (ref.)[e]
≥22	387	496	1.05 (0.87–1.29)	1.09 (0.89–1.33)
≥23	250	310	1.08 (0.87–1.35)	1.14 (0.90–1.43)
≥25	73	76	1.29 (0.90–1.85)	1.40 (0.97–2.02)
≥27	12	17	0.99 (0.46–2.14)	1.07 (0.49–2.36)

[a] Nine controls were excluded because of incomplete matching.
[b] Conditional logistic regression adjusted for the matching variables: age, menopausal status, postmenopausal hormone use, date of blood draw, time of blood draw, and fasting status.
[c] Conditional logistic regression adjusted for the matching variables and age at menarche, parity, age at first birth, BMI at 18 years of age, weight gain since age 18, benign breast disease, first-degree family history of breast cancer, and duration of postmenopausal hormone use.
[d] Unconditional logistic regression adjusted for the matching variables.
[e] Unconditional logistic regression adjusted for the matching variables and age at menarche, parity, age at first birth, BMI at 18 years of age, weight gain since age 18, benign breast disease, first-degree family history of breast cancer, and duration of postmenopausal hormone use.
[f] Unconditional logistic regression adjusted for the matching variables and age at menarche, parity, age at first birth, BMI at 18 years of age, weight gain since age 18, benign breast disease, and first-degree family history of breast cancer.
[g] Unconditional logistic regression adjusted for the matching variables and age at menarche, parity, age at first birth, BMI at 18 years of age, weight gain since age 18, benign breast disease, first-degree family history of breast cancer, duration of postmenopausal hormone use, and age at menopause.
[h] Average repeat number = ($[CAG]_n$ of allele 1 + allele 2)/2.

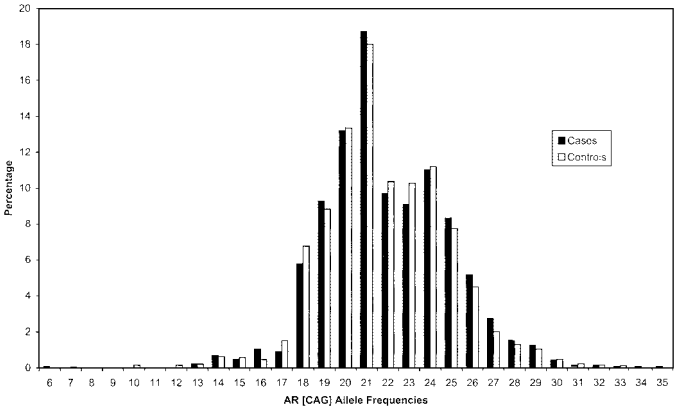

Fig. 1. The distribution of androgen receptor alleles among breast cancer cases and controls.

and first-degree family history of breast cancer. The test for interaction between first-degree family history of breast cancer and *AR* genotype (cutpoint, 22 repeats) was statistically significant ($P = 0.04$; Table 2). In stratified analyses, among women without a first-degree family history of breast cancer, the OR for carriers of 1 or 2 alleles ≥22 (*versus* noncarriers) was 0.92 (95% CI, 0.71–1.17). A nonsignificant positive association was observed among women with a first-degree family history of breast cancer (1 or 2 alleles ≥22: OR, 1.64; 95% CI, 0.93–2.89). When evaluating longer repeat cutpoints (23, 25, and 27), positive associations were limited to the family history-positive strata. We also examined the cross-classification of *AR* genotype and first-degree family history of breast cancer (Table 2). Compared with women with both alleles <22 and no family history, we observed a significant positive association limited to women with both a first-degree family history of breast cancer and longer alleles (Table 2). Point estimates did not change much when adjusting for all established risk factors in multivariate models.

We also evaluated the relationship of *AR* repeat length with circulating steroid hormone levels among postmenopausal controls not currently using postmenopausal hormones (Table 3). The *AR* gene is located on the X chromosome, and because one X chromosome is randomly inactivated in women, we analyzed the average of both *AR* repeat alleles to evaluate the systemic effects of *AR* genotype on endogenous steroid hormone levels. With increasing repeat number, we observed a nonsignificant inverse linear association with plasma testosterone levels and a nonsignificant-positive association with DHEAS levels. Relationships between the hormones androstenedione, DHEA, estradiol, estrone, and estrone sulfate and average repeat length were all nonsignificant.

Table 2 *ORs and 95% CIs[a] for breast cancer risk by AR genotype and first-degree family history of breast cancer*

First-degree family history	Genotype[b] $[CAG]_n$	Cases *n*	Controls *n*	OR (95% CI)[c]
no	<22	149	205	1.00
no	≥22	428	620	0.91 (0.71–1.17)
yes	<22	30	44	1.01 (0.60–1.70)
yes	≥22	120	100	1.70 (1.20–2.40)
				LRT[d] $P = 0.04$
no	<22	149	205	1.00
no	≥23	356	524	0.90 (0.70–1.16)
yes	<22	30	44	1.02 (0.60–1.72)
yes	≥23	105	86	1.76 (1.22–2.53)
				LRT[d] $P = 0.13$
no	<22	149	205	1.00
no	≥25	195	256	1.02 (0.76–1.36)
yes	<22	30	44	1.06 (0.62–1.79)
yes	≥25	62	48	2.01 (1.28–3.14)
				LRT[d] $P = 0.31$
no	<22	149	205	1.00
no	≥27	65	77	1.10 (0.73–1.65)
yes	<22	30	44	1.09 (0.64–1.86)
yes	≥27	21	20	1.58 (0.80–3.12)
				LRT[d] $P = 0.40$

[a] Unconditional logistic regression adjusted for the matching variables: age, menopausal status, postmenopausal hormone use, date of blood draw, time of blood draw, and fasting status.

[b] AR genotype defined as having one or more long alleles ≥22, 23, 25 or 27 *versus* both alleles <22.

[c] Reference groups are those with both alleles <22 and no first-degree family history of breast cancer.

[d] LRT is the likelihood ratio test for the interaction between *AR* genotype and first-degree family history of breast cancer.

Table 3 *Percent differences in hormone levels[a] per AR average repeat length among postmenopausal controls not using postmenopausal hormones*

Hormone (*n*)[b]	% change per six CAG repeat increment	(95% CI)
Estrone sulfate (442)	3.0%	(−11.9% to 20.5%)
Estrone (456)	2.5%	(−7.2% to 13.2%)
Estradiol (464)	−3.3%	(−13.7% to 8.3%)
Testosterone (461)	−6.0%	(−18.1% to 7.9%)
Androstenedione (459)	4.5%	(−10.3% to 20.4%)
DHEA (435)	0.1%	(−16.4% to 19.8%)
DHEAS (463)	9.6%	(−9.2% to 32.3%)

[a] Controlling for BMI, age, date of blood draw, time of blood draw, fasting status, and laboratory batch.

[b] Numbers vary due to insufficient plasma for some women and exclusion of outliers.

DISCUSSION

If androgens act directly on breast epithelium via the AR to inhibit tumor growth, then women with decreased AR transactivation (carriers of longer repeat alleles) may have an elevated risk of breast cancer. Conversely, through a physiological feedback mechanism, women with less functional receptors may have elevated androgen concentrations and be at increased risk of breast cancer as a result of increased estrogen levels and an increase in the fraction of bioavailable estradiol unbound to sex hormone-binding globulin due to a greater affinity of sex hormone-binding globulin for testosterone (27). Our data do not directly support either of these hypotheses, as we observed no association between $[CAG]_n$ repeat length in exon 1 of the *AR* gene and risk of breast cancer, nor did we observe *AR* allele length to be significantly associated with endogenous steroid hormone levels.

The association between *AR* $[CAG]_n$ alleles and breast cancer risk has been evaluated in three case-control studies (22–24). In a study conducted among French-Canadians (cases, $n = 255$; controls, $n = 461$), Giguére *et al.* (22) observed a significant inverse association between *AR* repeat allele length and breast cancer risk (2 alleles ≤20: OR, 0.50; 95% CI, 0.27–0.82). A null association was reported between *AR* genotype and breast cancer risk among 368 cases and 284 age-matched controls <40 from an Australian population (1 allele ≥22: OR, 1.40; 95% CI, 0.92–2.15; 2 alleles ≥22: OR, 1.40; 95% CI, 0.87–2.26; Ref. 23). A British study consisting of 508 cases and 426 controls >50 years of age (24) also provided no evidence of an association between *AR* repeat length and breast cancer risk (*versus* alleles <22; 1 allele >23: OR, 0.82; 95% CI, 0.62–1.09; 2 alleles >23; OR, 1.31; 95% CI, 0.87–1.97). Unlike the results of Giguére *et al.* (22), our data from a prospective study among mainly Caucasian women predominantly >50 years of age support these findings. Small sample size and the analysis of multiple potential *AR* CAG cutpoints could potentially explain the inconsistent findings between studies. In a recent case series of 133 Italian women with primary breast cancer, shorter *AR* $[CAG]_n$ repeat alleles were observed to be overrepresented among more aggressive forms of breast cancer (28). Our data also do not support these findings, as we observed no evidence of a relationship among cases between *AR* repeat length and breast cancer stage or grade.

In a study among 304 *BRCA1* mutation carriers, Rebbeck *et al.* (19) reported that women with at least one long *AR* allele (≥28) were diagnosed with breast cancer at a significantly younger age. Recent studies among *BRCA1/2* mutation carriers and women from high-risk breast/ovary cancer families have not confirmed these findings (20, 21). We did not observe a strong association between *AR* genotype and breast cancer risk; however, phenotypic and etiological differences between sporadic and *BRCA1*-associated breast cancers may be attributable to physiological variation in the effects of hormonal factors, such as *AR* genotype, on breast cancer development. We examined the hypothesis that having a family history of breast cancer may act as a modifier of breast cancer risk associated with *AR* genotype. Longer *AR* repeat alleles were overrepresented among cases with a first-degree family history of breast cancer; women in this dual category had a significantly elevated risk of breast cancer (OR, 1.70;

95% CI, 1.20–2.40). Although few, if any, of the women in this study are *BRCA1* mutation carriers, our data are consistent with the observations of Rebbeck *et al.* (19) suggesting that *AR* genotype may influence family history-associated breast cancer risk.

In our study, *AR* repeat length was not a strong predictor of plasma steroid hormone levels. In a prospective study among males 40–70 years of age ($n = 882$; Ref. 29), the age-related decrease in serum testosterone was significantly attenuated among men with longer *AR* alleles (% change in free testosterone: +2.8% per three CAG repeats; 95% CI, 1.0–4.7). More recently, in a small cross-sectional study among premenopausal women ($n = 239$), Westberg *et al.* (30) observed a significant association between short *AR* repeat alleles and higher follicular phase testosterone levels. Our data do not support an association in postmenopausal women. In addition, the modest inverse association between testosterone levels and increasing $[CAG]_n$ repeats that we observed, as well as the findings of Westberg *et al.* (30), do not support the hypothesis in women that longer, less functional *AR* alleles have increased plasma androgen levels. In addition, we did not observe even modest associations between *AR* repeat length and other hormones.

In summary, we did not observe a direct association between *AR* repeat length and breast cancer risk or evidence of a relationship between *AR* genotype and endogenous testosterone levels. However, our data are consistent with the hypothesis that longer *AR* repeat alleles may be involved in modifying family history-associated breast cancer risk.

ACKNOWLEDGMENTS

We thank Jeanne Sparrow, Lisa Li, and Carsta Cieluch for technical and laboratory assistance. We are also indebted to the participants in the Nurses' Health Study for their dedication and commitment.

REFERENCES

1. Plu-Bureau, G., Touraine, P., and Mauvais-Jarvis, P. Interactions between estradiol and progesterone in normal breast. *In:* A. Manni (ed.), Endocrinology of Breast Cancer, pp. 21–37. Totowa, NJ: Humana Press, 1999.
2. Bernstein, L., and Ross, R. K. Endogenous hormones and breast cancer risk. Epidemiol. Rev., *15:* 48–65, 1993.
3. Toniolo, P. G., Levitz, M., Zeleniuch-Jacquotte, A., Banerjee, S., Koenig, K. L., Shore, R. E., Strax, P., and Pasternack, B. S. A prospective study of endogenous estrogens and breast cancer in postmenopausal women. J. Natl. Cancer Inst., *87:* 190–197, 1995.
4. Dorgan, J. F., Longcope, C., Stephensen, H. E., Falk, R. T., Miller, R., Franz, C., Kahle, L., Campbell, W. S., Tangrea, J. A., and Schatzkin, A. Relation of prediagnostic serum estrogen and androgen levels to breast cancer risk. Cancer Epidemiol. Biomark. Prev., *5:* 533–539, 1996.
5. Thomas, H. V., Key, T. J., Allen, D. S., Moor, J. W., Dowsett, M., Fentiman, I. S., and Wang, D. Y. A prospective study of endogenous serum hormone concentrations and breast cancer risk in postmenopausal women on the island of Guernsey. Br. J. Cancer, *76:* 401–405, 1997.
6. Hankinson, S. E., Willett, W. C., Manson, J. A. E., Colditz, G. A., Hunter, D. J., Spiegelman, D., Barbieri, R. L., and Speizer, F. E. Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. J. Natl. Cancer Inst., *90:* 1292–1299, 1998.
7. Zeleniuch-Jacquotte, A., Bruning, P. F., Bonfrer, J. M. G., Koenig, K. L., Shore, R. E., Kim, M. Y., Pasternack, B. S., and Toniolo, P. Relation of serum levels of testosterone and dehydroepiandrosterone sulfate to risk of breast cancer in postmenopausal women. Am. J. Epidemiol., *145:* 1030–1038, 1997.
8. Dorgan, J. F., Stanczyk, F. Z., Longcope, C., Stephenson, H. E., Chang, L., Miller, R. E., Franz, C., Falk, R. T., and Kahle, L. Relationship of serum DHEA, DHEAS and 5-androstene-3β, 17β-diol to risk of breast cancer in postmenopausal women. Cancer Epidemiol. Biomark. Prev., *6:* 177–181, 1997.
9. Cauley, J. A., Lucas, F. L., Kuller, L. H., Stone, K., Browner, W., and Cummings, S. R., for the Study of Osteoporotic Fractures Research Group. Elevated serum estradiol and testosterone concentrations are associated with a high risk for breast cancer. Ann. Intern. Med., *130:* 270–277, 1999.
10. Adams, J. B. Adrenal androgens and human breast cancer: A new appraisal. Breast Cancer Res. Treat., *51:* 183–188, 1998.
11. Maggiolini, M., Donzé, O., Jeannin, E., Andò, S., and Picard, D. Adrenal androgens stimulate the proliferation of breast cancer cells as direct activators of estrogen receptor α. Cancer Res., *59:* 4864–4869, 1999.
12. Yue, W., Wang, J. P., Hamilton, C. J., Demers, L. M., and Santen, R. J. *In situ* aromatization enhances breast tumor estradiol levels and cellular proliferation. Cancer Res., *58:* 927–932, 1999.
13. Ross, R. K., Pike, M. C., Coetzee, G. A., Reichart, J. K. V., Yu, M. C., Feigelson, H., Stanczyk, F. Z., Kolonel, L. N., and Henderson, B. E. Androgen metabolism and prostate cancer: establishing a model of genetic susceptibility. Cancer Res., *58:* 4497–4504, 1998.
14. Giovannucci, E., Stampfer, M. J., Krithivas, K., Brown, M., Brufsky, A., Talcott, J., Hennekens, C. H., and Kantoff, P. W. The CAG repeat within the androgen receptor gene and its relationship to prostate cancer. Proc. Natl. Acad. Sci. USA, *94:* 3320–3323, 1997.
15. Ingles, S. A., Ross, R. K., Yu, M. C., Irvine, R. A., La Pera, G., Haile, R. W., and Coetzee, G. A. Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. J. Natl. Cancer Inst., *89:* 166–170, 1997.
16. Stanford, J. L., Just, J. J., Gibbs, M., Wicklund, K. G., Neal, C. L., Blumenstein, B. A., and Ostrander, E. A. Polymorphic repeats in the androgen receptor gene: molecular markers of prostate cancer risk. Cancer Res., *57:* 1194–1198, 1997.
17. Chamberlain, N. L., Driver, E. D., and Miesfeld, R. L. The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. Nucleic Acids Res., *22:* 3181–3186, 1994.
18. Tut, T. G., Ghadessy, F. J., Trifiro, M. A., Pinsky, L., and Yong, E. L. Long polyglutamine tracts in the androgen receptor are associated with reduced *trans*-activation, impaired sperm production, and male infertility. J. Clin. Endocrinol. Metab., *82:* 3777–3782, 1997.
19. Rebbeck, T. R., Kantoff, P. W., Krithivas, K., Neuhausen, S., Blackwood, M. A., Godwin, A. K., Daly, M. B., Narod, S. A., Garber, J. E., Lynch, H. T., Weber, B. L., and Brown, M. Modification of *BRCA1*-associated breast cancer risk by the polymorphic androgen-receptor CAG repeat. Am. J. Hum. Genet., *64:* 1371–1377, 1999.
20. Kadouri, L., Easton, D. F., Edwards, S., Hubert, A., Kote-Jarai, Z., Glaser, B., Durocher, F., Abeliovich, D., Peretz, T., and Eeles, R. A. CAG and GGC repeat polymorphisms in the androgen receptor gene and breast cancer susceptibility in *BRCA1/2* carriers and non-carriers. Br. J. Cancer, *85:* 36–40, 2001.
21. Menin, C., Banna, G. L., De Salvo, G., Lazzarotto, V., De Nicolo, A., Agata, S., Montagna, M., Sordi, G., Nicoletto, O., Chieco-Bianchi, L., and D'Andrea, E. Lack of an association between androgen receptor CAG polymorphism and familial breast/ovarian cancer. Cancer Lett., *168:* 31–36, 2001.
22. Giguére, Y., Dewailly, E., Brisson, J., Ayotte, P., Laflamme, N., Demers, A., Forest, V-I., Dodin, S., Robert, J., and Rousseau, F. Short polyglutamine tracts in the androgen receptor are protective against breast cancer in the general population. Cancer Res., *61:* 5869–5874, 2001.
23. Spurdle, A. B., Dite, G. S., Chen, X., Mayne, C. J., Southey, M. C., Batten, L. E., Chy, H., Trute, L., McCredie, M. R. E., Giles, G. G., Armes, J., Venter, D. J., Hopper, J. L., and Chenevix-Trench, G. Androgen receptor exon 1 CAG repeat length and breast cancer in women before age forty years. J. Natl. Cancer Inst., *91:* 961–966, 1999.
24. Dunning, A. M., McBride, S., Gregory, J., Durocher, F., Foster, N. A., Healey, C. S., Smith, N., Pharoah, P. D. P., Luben, R. N., Easton, D. F., and Ponder, B. A. J. No association between androgen receptor or vitamin D receptor gene polymorphism and risk of breast cancer. Carcinogenesis (Lond.), *20:* 2131–2135, 1999.
25. Hankinson, S. E., Manson, J. E., Spiegelman, D., Willet, W. C., Longcope, C., and Speizer, F. E. Reproducibility of plasma hormone levels in postmenopausal women over a 2–3-year period. Cancer Epidemiol. Biomark. Prev., *4:* 649–654, 1995.
26. SAS Institute Inc. SAS/STAT User's Guide, Version 6, Ed. 4. Cary, NC: SAS Institute, Inc., 1989.
27. Siiteri, P. K., Hammond, G. L., and Nisker, J. A. Increased availability of serum estrogens in breast cancer: a new hypothesis. *In:* M. C. Pike, P. K. Siiteri, and C. W. Welsch (eds.), Hormones and Breast Cancer, pp. 87–101. Banbury Report No. 8. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory, 1981.
28. Yu, H., Bharaj, B., Vassilikos, E. J. K., Giai, M., and Diamandis, E. P. Shorter CAG repeat length in the androgen receptor gene is associated with more aggressive forms of breast cancer. Breast Cancer Res. Treat., *59:* 153–161, 2000.
29. Krithivas, K., Yurgalevitch, S. M., Mohr, B. A., Wilcox, C. J., Batter, S. J., Brown, M., Longcope, C., McKinlay, J. B., and Kantoff, P. W. Evidence that the CAG repeat in the androgen receptor gene is associated with the age-related decline in serum androgen levels in men. J. Endocrinol., *162:* 137–142, 1999.
30. Westberg, L., Baghaei, F., Rosmand, R., Hellstrand, M., Landén, M., Jansson, M., Holm, G., Björntorp, P., and Eriksson, E. Polymorphisms of the androgen receptor gene and the estrogen receptor β gene are associated with androgen levels in women. J. Clin. Endocrinol. Metab., *86:* 2562–2568, 2001.