Modification of BRCA1-Associated Breast Cancer Risk by the Polymorphic Androgen-Receptor CAG Repeat


Department of Biostatistics and Epidemiology, and Medicine and Genetics, University of Pennsylvania School of Medicine, and Fox Chase Cancer Center, Philadelphia; Department of Adult Oncology and Lank Center for Genitourinary Cancer, Dana-Farber Cancer Institute, Harvard Medical School, Boston; University of Utah, Salt Lake City; Women’s College Hospital, Toronto; and Department of Preventive Medicine, Creighton University, Omaha

Summary

Compared with the general population, women who have inherited a germline mutation in the BRCA1 gene have a greatly increased risk of developing breast cancer. However, there is also substantial interindividual variability in the occurrence of breast cancer among BRCA1 mutation carriers. We hypothesize that other genes, particularly those involved in endocrine signaling, may modify the BRCA1-associated age-specific breast cancer risk. We studied the effect of the CAG repeat-length polymorphism found in exon 1 of the androgen-receptor (AR) gene (AR-CAG). AR alleles containing longer CAG repeat lengths are associated with a decreased ability to activate androgen-responsive genes. Using a sample of women who inherited germline BRCA1 mutations, we compared AR-CAG repeat length in 165 women with and 139 women without breast cancer. We found that women were at significantly increased risk of breast cancer if they carried at least one AR allele with ≥28 CAG repeats. Women who carried an AR-CAG allele of ≥28, ≥29, or ≥30 repeats were given a diagnosis 0.8, 1.8, or 6.3 years earlier than women who did not carry at least one such allele. All 11 women in our sample who carried at least one AR-CAG allele with ≥29 repeats had breast cancer. Our results support the hypothesis that age at breast cancer diagnosis is earlier among BRCA1 mutation carriers who carry very long AR-CAG repeats. These results suggest that pathways involving androgen signaling may affect the risk of BRCA1-associated breast cancer.

Introduction

Inheritance of a germline mutation in the BRCA1 gene (MIM 113705) is associated with an increased risk of developing breast cancer. However, there is also substantial variability in the ages at which breast cancers are diagnosed in BRCA1 mutation carriers (Easton et al. 1995; Narod et al. 1995; Rebbeck 1999). These observations imply that germline mutations in BRCA1 may be necessary to explain the Mendelian pattern of cancer in some families but may not be sufficient to completely describe the interindividual variability in the age-specific risk of cancer. The ability to effectively apply risk-prediction or cancer-prevention strategies in BRCA1 carriers may therefore depend on knowledge of risk-modifying factors in addition to BRCA1 mutation status.

Steroid hormone pathways regulate BRCA1 expression (Gudas et al. 1995; Marks et al. 1997). Therefore, we hypothesize that allelic variation in genes governing hormonal signaling known to play a role in normal development and cancer risk may be involved in modification of BRCA1-associated cancer risk. For example, the androgen-receptor gene AR (MIM 313700), which functions as a ligand-dependent transcriptional activator in response to androgens, contains a highly polymorphic CAG trinucleotide repeat (AR-CAG) encoding glutamines in its first exon. The length of the AR-CAG polymorphism is inversely associated with the degree of transcriptional activation by the AR (Chamberlain et al. 1994; Kazemi-Esfarjani et al. 1995). Individuals with X-linked spinal and bulbar muscular atrophy (SBMA, Kennedy disease) have 40 or more AR-CAG repeats and manifest clinical androgen insensitivity (LaSpada et al. 1991). AR mediates breast tumor growth and progression (Zhu et al. 1997; Birrell et al. 1998). Increased AR-CAG repeat length has also been associated with decreased prostate cancer risk, presumably because of a decreased ability of androgens to stimulate transcription of genes involved in prostate growth (Hardy et al. 1996; Giovannucci et al. 1997; Ingles et al. 1997). However,
this result has not been replicated in all populations (Eeles et al. 1998). These findings suggest that AR-CAG repeat-length polymorphism may be involved in modifying the development of diseases caused by alterations in endocrine signaling.

Subjects and Methods

Subjects

A sample of 304 women who carry disease-associated germline BRCA1 mutations was ascertained through families with a history of breast and/or ovarian cancer at Creighton University, the Dana Farber Cancer Institute, The University of Michigan, Fox Chase Cancer Center, The University of Pennsylvania, The University of Utah, and Women’s College Hospital (Toronto) between 1978 and 1997. Women were self- or physician-referred because of a strong family history of breast and/or ovarian cancer. These women provided written informed consent for research under protocols approved by the institutional review boards at each institution. Of these 304 women, 165 (54%) were affected (mean age 44.5 years, range 19–89 years) and 139 (46%) were unaffected by breast cancer (mean age 44.5 years, range 19–89 years).

Genotype Analysis

All study participants provided peripheral blood samples from which genomic DNA was extracted according to standard protocols. We used PCR to amplify the AR-CAG trinucleotide repeat found in exon 1 of AR, as described elsewhere (Giovannucci et al. 1997). The AR-CAG repeat-length polymorphism was modeled in three ways to reflect two alternative hypotheses about the activity of AR in breast carcinogenesis. The first two models reflect the activity of specific alleles acting at the level of the breast epithelial cell to modulate androgen signaling. First, the shorter of the two repeat-length alleles for each subject was considered in a survival analysis with \( \log \text{rank} \) statistic estimated from Kaplan-Meier models. Risk (hazard) ratios were estimated by using Cox proportional hazards models. These analyses involved dichotomizing the total sample by using cutpoints along the AR-CAG repeat-length distribution to compare women whose AR-CAG repeat allele was less than, greater than, or equal to the specified number of repeats. Cutpoints were made within the range of observed AR-CAG repeat lengths. First, the effect of having at least one very short allele was evaluated by comparing groups divided at repeat lengths \(<15\) through \(<25\) through \(\geq 25\), where the repeat-length cutpoint was determined by the shorter of a woman’s two AR alleles. Analyses were undertaken in this range because few alleles with \(<14\) AR-CAG repeats were observed. The 25-repeat allele was used as the upper cutpoint bound, because few shorter alleles with \(>25\) repeats were observed. Second, the effect of having at least one very long allele was evaluated by comparing groups divided at allele lengths \(<20\) through \(<30\) through \(\geq 30\), where the repeat-length cutpoint was determined by the longer of a woman’s two AR alleles. The 30-repeat AR-CAG allele was used as the upper bound, because few 31- or 32-repeat alleles were observed in this sample (fig. 1). This

Statistical Methods

Cox proportional hazards models were used to evaluate the difference in breast cancer penetrance across AR-CAG repeat lengths. To correct for nonindependence of observations among participants drawn from the same families, the robust variance-covariance estimation approach of Lin and Wei (1989) was used, as implemented in STATA (StataCorp., release 5). Participants were followed up (retrospectively) from birth until one of several events occurred. The primary event of interest was the first diagnosis of a primary invasive breast cancer \((n = 165, 54\%)\). Participants with no prior breast cancer diagnosis were censored when they developed ovarian cancer \((n = 40, 13\%)\), had a prophylactic mastectomy or oophorectomy \((n = 46, 15\%)\) or died \((n = 15, 5\%)\)—or when none of these events had occurred by the end of the observation period \((n = 38, 13\%)\). All Cox proportional hazards analyses were undertaken with and without adjustment for three hormone-related risk factors: age at menarche, age at first live birth, and total number of full-term pregnancies (parity). Parity and age at menarche are the only factors on this list that have been previously suggested as modifiers of breast cancer risk in BRCA1 carriers (Narod et al. 1995).

AR-CAG repeat length is known to be correlated with endocrine signaling. However, the AR-CAG repeat is continuously distributed, and there is no a priori point at which cutoffs may be applied to identify allele-specific risk groups. Therefore, sensitivity analyses were undertaken to compare AR-CAG genotype classes by use of the log rank statistic estimated from Kaplan-Meier models. Risk (hazard) ratios were estimated by using Cox proportional hazards models. These analyses involved dichotomizing the total sample by using cutpoints along the AR-CAG repeat-length distribution to compare women whose AR-CAG repeat allele was less than, greater than, or equal to the specified number of repeats. Cutpoints were made within the range of observed AR-CAG repeat lengths. First, the effect of having at least one very short allele was evaluated by comparing groups divided at repeat lengths \(<15\) through \(<25\) through \(\geq 25\), where the repeat-length cutpoint was determined by the shorter of a woman’s two AR alleles. Analyses were undertaken in this range because few alleles with \(<14\) AR-CAG repeats were observed. The 25-repeat allele was used as the upper cutpoint bound, because few shorter alleles with \(>25\) repeats were observed. Second, the effect of having at least one very long allele was evaluated by comparing groups divided at allele lengths \(<20\) through \(<30\) through \(\geq 30\), where the repeat-length cutpoint was determined by the longer of a woman’s two AR alleles. The 30-repeat AR-CAG allele was used as the upper bound, because few 31- or 32-repeat alleles were observed in this sample (fig. 1). This
analysis allowed us to identify critical cutpoints along the continuous allele distribution for which breast cancer penetrance may be modified.

Results

We evaluated the effect of AR-CAG repeat length on breast-cancer penetrance in 304 women who carry germ-line mutations in the BRCA1 gene. Mutations in BRCA1 spanned the majority of the gene’s coding region (ranging from Met1Ile to 5439delA). Mutations included 198 (65%) deletions (including large genomic deletions), 55 (18%) nonsense mutations, 31 (10%) insertions, and 20 (7%) disease-associated missense mutations. The three most commonly identified mutations were 185delAG (n = 49), Q1313ter (n = 33), and 5382insC (n = 26).

A frequency histogram depicting the observed distribution of AR-CAG repeat lengths is presented in figure 1. Although the distributions among cancer patients and cancer-free women are generally similar, we observed a skew in the distribution of breast cancer cases toward inheritance of at least one longer-repeat allele compared with cancer-free women. The median repeat length was 22 (range 8–32). The overall allele repeat distribution in our sample (including mean, median, and range) is similar to that of control populations reported elsewhere (Giovannucci et al. 1997; Hakimi et al. 1997). The median repeat length of the shorter allele carried by each individual was 21 (range 8–30). The median repeat length of the longer allele carried by each individual was 24 (range 17–32).

We found no simple association between breast cancer penetrance and continuous CAG repeat variables coded as mean repeat length (RR = 1.03, 95% CI 0.96–1.10), shorter repeat length (RR = 1.01, 95% CI 0.95–1.07), or longer repeat length (RR = 1.04, 95% CI 0.98–1.10). To identify points in the continuous AR-CAG repeat-length distribution associated with modified breast cancer penetrance, we performed sensitivity analyses using the shorter or longer of a woman’s two alleles. No effect

**Figure 1** AR-CAG repeat-length distribution by breast cancer status
of a woman’s shorter AR allele was observed (fig. 2). In contrast, women who carried at least one long AR-CAG repeat allele (≥28 repeats) had a significantly earlier age at breast cancer diagnosis than women without one long-repeat allele (fig. 3). In the group of women who carried at least one allele with ≥28 repeats, we observed eight breast cancers (mean age at diagnosis 41.0 years) and four unaffected women (mean age at diagnosis 37.5 years) with 28-repeat alleles, four breast cancers (mean age at diagnosis 40.8 years) in women with 29-repeat alleles, and seven breast cancers (mean age at diagnosis 34.0 years) in women with ≥30-repeat alleles. There were no women in this study who carried two alleles with ≥28 repeats.

Estimates of hazard ratios (HRs) from Cox proportional hazards models indicated an increasing breast cancer penetrance as the repeat-length cutpoint increased (fig. 3). Women with ≥28 AR-CAG repeats (n = 19) developed breast cancer 0.8 years earlier than women who had only shorter alleles (n = 146; HR = 1.81, 95% CI 1.06–3.08). Similarly, women with ≥29 AR-CAG repeats (n = 11) developed breast cancer 1.8 years earlier than women who had only shorter alleles (n = 154; HR = 2.66, 95% CI 1.51–4.69), and women with ≥30 AR-CAG repeats (n = 7) developed breast cancer 6.3 years earlier than women who had only shorter alleles (n = 158; HR = 4.45, 95% CI 1.31–15.16). As indicated in figure 1, all women in our sample who carried at least one allele with ≥28 repeats were affected. All estimates presented here were made without adjustment for other potential confounding variables. However, inferences from survival analyses comparing AR-CAG repeats on breast cancer penetrance were identical in analyses that were unadjusted and adjusted for parity, age at menarche, age at first live birth, or ascertainment site. Note that the 24 individuals who carried at least one AR-CAG allele of ≥28 repeats included 20 unrelated women and two pairs of relatives (i.e., two relatives in each of two families, each pair carrying the same mutation). Five of the unrelated women in this group carried the same BRCA1 mutation (185delAG).

It was not possible to conduct a complete evaluation of the effect of BRCA1 mutation type given the extreme heterogeneity in the type and location of BRCA1 mutations. However, we repeated our analyses including only women with at least one AR-CAG allele of ≥28 repeats by using a sample of women with unique BRCA1 mutations. Our inferences were similar to those presented in figure 3: having at least one allele with 28 or more AR-CAG repeats (HR = 3.7, 95% CI: 1.6–8.4), 29 or more AR-CAG repeats (HR = 3.6, 95% CI: 1.3–9.8), or 30 or more AR-CAG repeats (HR = 4.5, 95% CI 1.1–18.1) remained significantly associated with breast cancer risk.

**Discussion**

We report that AR-CAG allele size is associated with breast cancer penetrance in BRCA1 mutation carriers. Although the present association study is primarily hy-

![Figure 2](https://example.com/figure2.png)

**Figure 2** HRs associated with shorter AR-CAG repeat length among 304 BRCA1 mutation carriers
Rebbeck et al.: Androgen Receptor in BRCA1 Breast Cancer

Hypothesis generating, there is a strong biological rationale for our findings. It is well known that AR-CAG repeat length modulates the transactivational activity of AR in vitro and is inversely associated with androgen sensitivity (LaSpada et al. 1991; Chamberlain et al. 1994; Kazemi-Esfarjani et al. 1995; Tut et al. 1997). The AR is expressed in normal breast epithelial cells and in some breast tumors, and it may be coexpressed with the estrogen and progesterone receptors in breast tumors (Hackenberg and Schulz 1996). Androgens are known to inhibit the growth of some breast cancer cell lines (Birrell et al. 1995). Ectopic AR expression leads to the inhibition of breast tumor–cell proliferation in response to androgens (Szelei et al. 1997). Although side effects have limited their use, androgens are effective in the treatment of women with metastatic breast cancer (Goldenberg et al. 1973). This information supports the hypothesis that decreased androgenic activity associated with an increased number of AR-CAG repeats may result in increased breast cancer risk.

Figure 3  HRs associated with longer AR-CAG repeat length among 304 BRCA1 mutation carriers

One means by which AR may act in breast tumorigenesis is by androgen signaling acting on the level of the mammary epithelial cell. Since AR maps to the X chromosome, breast epithelial cells in women express only one of the two AR alleles a woman has inherited. Thus, each cell is under the influence of only a single AR allele. Acting on knowledge of AR activity in breast cell proliferation, we hypothesize that decreased androgenic activity in breast cells expressing a very long AR-CAG repeat allele may result in increased breast epithelial cell proliferation. This increased proliferation may in turn affect the penetrance of breast cancer in BRCA1 mutation carriers. In support of this hypothesis, Elhaji et al. (1997) have reported somatic mutations leading to a significant lengthening of the AR-CAG repeats in breast tumors from postmenopausal women. Thus, our finding of a relationship between the longer AR-CAG allele and breast cancer risk in BRCA1 mutation carriers supports a model in which the effect is mediated at the level of the mammary epithelium in a cell-autonomous fashion.

An alternative to this hypothesis is an endocrine or paracrine mechanism for the action of AR in breast tumorigenesis. Androgens may modulate BRCA1 risk via an endocrine mechanism by altering the levels of circulating hormones, such as estradiol, or via a paracrine mechanism involving effects mediated by the mammary stroma. For example, a recent epidemiologic study of the relationship between testosterone and breast cancer risk reported that decreased serum testosterone levels may have an indirect effect by influencing the bioavailability of estrogen (Zeleniuch-Jacquotte et al. 1997). Alternatively, androgens might act on the mammary stroma to indirectly affect the growth of the mammary epithelium. If either of these indirect mechanisms is mediating the effect of the AR polymorphism on BRCA1 risk, the effect would be expected to be the result of the action of the AR in a number of cells, rather than in a cell-autonomous fashion, and thus would reflect the activity of both AR alleles. We modeled this effect of both...
AR alleles as the average of the signal from the two alleles, and we found no support for this hypothesis. However, our ability to reject these hypotheses was limited, and we cannot rule out the possibility that endocrine or paracrine effects of androgens may affect breast cancer risk in BRCA1 mutation carriers.

One limitation of the present study is that the participants carried a variety of BRCA1 mutations, and we could not evaluate the effect of BRCA1 mutation type or location on the present results. However, it is unlikely that the heterogeneity of BRCA1 mutations affected the inferences of this study. Previous reports suggest that the location of the BRCA1 mutation may, in part, determine breast versus ovarian cancer risk (Gayther et al. 1995). The data available from large consortia indicate that no differences exist in breast cancer risk by mutation location or type (D. Easton, personal communication). The common Ashkenazi Jewish mutations in BRCA1 (i.e., 185delAG and 5382insC) confer approximately the same lifetime breast cancer risk (Struewing et al. 1997). In addition, the effect of AR-CAG repeat length among women who carried at least one longer allele of ≥28 repeats remained, even after limiting the sample to women who had unique BRCA1 mutations. This suggests that BRCA1 mutation type did not artificially induce the association of AR-CAG alleles and breast cancer risk. Given the extreme heterogeneity of mutations in BRCA1, it is unlikely that a comprehensive analysis of the effect of mutation location or type could be done. Furthermore, because the analyses were not limited to a particular class of mutations, the present results may be applicable to the general population of BRCA1 mutation carriers from high-risk families. An additional limitation is that some individuals in the families studied here may have been excluded because they had died or were otherwise unable to participate in this research. As a result, the present results do not allow us to determine whether this effect implicates AR as an independent breast cancer risk factor or as a modifier of BRCA1-associated breast carcinogenesis.

Our results suggest that female BRCA1 mutation carriers who have inherited at least one very long AR-CAG repeat may be diagnosed with breast cancer at a significantly earlier age than women who do not carry a very long AR-CAG repeat. Because the frequency of long AR-CAG repeats is rare, the AR-CAG repeat polymorphism may be relevant only to some BRCA1 mutation carriers. However, the large magnitude of effect suggests that the personal impact on breast cancer risk to women who carry the very long repeats could be substantial. We conclude that the length of the AR-CAG repeat may affect the timing of breast cancer diagnosis in BRCA1 mutation carriers, possibly through modulation of hormonal responses of individual mammary epithelial cells. However, these results are preliminary, and it is premature to consider knowledge about AR-CAG genotype in making clinical decisions about breast cancer risk, surveillance, or prevention among BRCA1 mutation carriers.

Acknowledgments

The authors wish to thank Andre Rogatko for his helpful discussions of the statistical analyses. This study was supported by grants from the Public Health Service (ES08030 and CA737370 [to T.R.R.], CA57601 [to B.L.W.], CA70328 [to A.K.G.], and CA55914 and CA74415 [to S.L.N.]), the University of Pennsylvania Cancer Center (to T.R.R. and B.L.W.), The Breast Cancer Research Foundation (to B.L.W.), the Dana-Farber Women’s Cancers Program (to J.E.G. and M.B.), the Dana-Farber Prostate Cancer Research Fund (to P.W.K. and K.K.), the Department of Defense (DAMD-17-96-I-6088 [to A.K.G.], DAMD-17-94-J-4425 [to M.D.B.], DAMD17-94-J-4260 [to S.L.N.], and DAMD-17-94-J-4340 and DAMD-17-97-I-7112 [to H.T.L.]), The Utah Cancer Registry (funded by Public Health Service Grant NO1-CN-6700) and the Utah State Department of Health, and the Nebraska State Cancer and Smoking-Related Diseases Research Program (LB595 [to H.T.L.]).

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

Breast Cancer Information Core, http://www.nhgri.nih.gov/Intramural_research/Lab_transfer/Bic/index.html (for BRCA1)
Genome Database, http://www.gdb.org/ (for BRCA1 and AR)
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nlm.nih.gov/Omim/ (for BRCA1 [MIM 113705] and AR [MIM 313700])

References

repeat of the androgen receptor and female breast cancer. Am J Hum Genet 61:A64