The 6q22.33 Locus and Breast Cancer Susceptibility

Tomas Kirchhoff,1 Zhang-qun Chen,8 Bert Gold,6 Prodipto Pal,1 Mia M. Gaudet,2 Kristi Kosarin,1 Douglas A. Levine,3 Peter Gregersen,7 Sara Spencer,1 Megan Harlan,1 Mark Robson,1 Robert J. Klein,3 Clifford A. Hudis,5 Larry Norton,7 Michael Dean,6 and Kenneth Offit1

Abstract

Recently, we identified a novel breast cancer susceptibility locus at 6q22.33 following a genome-wide association study in the Ashkenazi Jewish genetic isolate. To replicate these findings, we did a case-control association analysis on 6q22.33 (rs2180341) in an additional 487 Ashkenazi Jewish breast cancer cases and in an independent non-Jewish, predominantly European American, population of 1,466 breast cancer cases and 1,467 controls. We confirmed the 6q22.33 association with breast cancer risk in the replication cohorts [per-allele odds ratio (OR), 1.18; 95% confidence interval (95% CI), 1.04-1.33; P = 0.0083], with the strongest effect in the aggregate meta-analysis (OR, 1.20; 95% CI, 1.04-1.33; P = 2.2 × 10-5) compared with estrogen receptor–negative tumors. Furthermore, this study provides a novel insight into the functional significance of 6q22.33 in breast cancer susceptibility. Due to the stronger association of 6q22.33 with estrogen receptor–positive breast cancer, we examined the effect of candidate genes on estrogen receptor response elements. Upon transfection of overexpressed RNF146 in the MCF-7 breast cancer cell line, we observed diminished expression of an estrogen receptor response element reporter construct. This study confirms the association of 6q22.33 with breast cancer, with slightly stronger effect in estrogen receptor–positive tumors. Further functional studies of candidate genes are in progress, and a large replication analysis is being completed as part of an international consortium.

Introduction

It is estimated that up to 30% of breast cancer cases may be caused by genetic factors (1-4). Family history of breast cancer is responsible for the greatest increase in risk, but the high-penetrant genes that have been identified, such as BRCA1, BRCA2, PTEN, and p53, only explain 20% to 25% of familial breast cancer (5) and 5% of all breast cancers (6). Recently, genome-wide association studies (GWAS) have proven to be useful in identification of additional genetic factors responsible for breast cancer susceptibility (7-10). These large independent studies have reported genetic variation in fibroblast growth factor receptor 2 (FGFR2) conferring 1.2 to 1.4 increased risk of breast cancer in populations with different genetic ancestries. Besides FGFR2, several other loci have been identified with a low penetrant effect on breast cancer risk. However, the replication of these additional candidates has varied among the studies, very likely as a result of statistical power limitations and population stratification. Therefore, large replication studies complemented with functional evidence will be needed to confirm reported associations with breast cancer risk.

In our recent GWAS, conducted in the Ashkenazi Jewish population, we identified a novel candidate region on 6q22.33 associated with ~1.3-fold increased risk of breast cancer. We utilized a two-stage design: the first being an analysis of 250 samples from Ashkenazi Jewish cases with a marked family history of breast cancer compared with 300 Ashkenazi Jewish controls, the second stage being a replication analysis of 384 loci in ~1,000 unselected Ashkenazi Jewish breast cancer cases and 1,000 Ashkenazi Jewish controls (10). We showed the strongest association mapped within a 200-kb region on 6q22.33, where two candidate genes are located: enoyl Coenzyme A hydratase domain containing 1 (ECHDC1) and ring finger protein 146 (RNF146). Prior evidence links both genes to breast cancer tumorigenesis. ECHDC1 has been suggested to play a major role in mitochondrial fatty acid oxidation, and it is well established that endogenous fatty acid synthetic activity is abnormally elevated in a subset of breast carcinomas (11, 12). RNF146, also known as dactylinin, encodes a polypeptide containing an amino-terminal tripeptide. Beside
C3HC4 RING finger domain, characteristic of the ubiquitin proteasome system, which regulates such processes as cell cycle, apoptosis, transcription, protein trafficking, DNA replication and repair, and angiogenesis (13, 14).

In order to validate the association of 6q22.33 with increased breast cancer risk, in this study we carried out a replication analysis on independent cohorts of cases and controls of both Ashkenazi Jewish as well as non-Jewish, predominantly European American, populations. In addition to epidemiologic observations supporting the involvement of 6q22.33 in breast cancer susceptibility, we also provide evidence of a possible functional mechanism accounting for the association of the 6q22.33 locus and breast cancer risk.

Materials and Methods

Subjects. Cases and controls for association analyses were identified from two populations: Ashkenazi Jewish and non-Jewish populations of predominantly European ancestry. For replication analysis on the Ashkenazi Jewish population, we used 487 breast cancer patients, ascertained by the Clinical Genetics Service at Memorial Sloan-Kettering Cancer Center (MSKCC); a substantial proportion of these cases were described in the context of prior epidemiologic studies (15). All these cases tested negative for Ashkenazi Jewish founder mutations in the BRCA1 and BRCA2 genes. Cases were compared with 1,149 healthy Ashkenazi Jewish controls used in our previous study (10). Replication population of non-Jewish cases included: (a) 171 familial breast cancer cases of European ancestry ascertained from clinical protocols at MSKCC, with eligible cases having ≥3 individuals with breast cancer present in a single lineage; (b) 751 non-Jewish sporadic breast cancer cases unselected for a family history of the disease and collected as a part of a separate protocol at MSKCC; and (c) 544 non-Jewish sporadic breast cancer cases unselected for a family history of the disease, which were ascertained from anonymized protocols at MSKCC. Overall, as illustrated in Table 1A, although all non-Jewish cases were predominantly European American (n = 1,604), other ancestries were present in this ascertainment, such as African American (n = 167), Hispanic (n = 117), Asian (n = 58), and other populations (n = 7).

For non-Jewish controls, we ascertained 630 cancer-free women who participated in the New York Cancer Project, an ongoing cohort study (16), all of whom were of European ancestry. A second group of non-Jewish controls comprised 837 women who were either participating in cancer screening and were cancer free or were spouses of patients with prostate cancer, and who did not have a personal or family history of breast cancer. The population structure of non-Jewish controls was similar to the non-Jewish sporadic breast cancer group, as detailed in Table 1A: European ancestry (n = 1,337), African Americans (n = 51), Hispanics (n = 58), Asians (n = 18), and other ancestries (n = 3). Previous published association data from our GWAS (phase I and phase II) were included in a final aggregate meta-analysis only, with a detailed description of breast cancer cohorts and controls published in that prior study (10).

Genotyping. Genomic DNA was prepared using the Gentra Autopure system, according to the manufacturer's protocol (Qiagen). Other DNA extraction procedures were done as previously described (17). Genotyping of rs2180341, rs6569479, rs6569480, and rs7776136 was done by the TaqMan allelic discrimination procedure under standard conditions (Applied Biosystems). In order to avoid potential bias by inclusion of data from samples previously genotyped by other methods (Affymetrix 500 K and Illumina GoldenGate assay), we regenotyped all published sample populations by conventional TaqMan allelic discrimination. All genotypes showed 100% concordance. The clustering of genotype calls was done by SDS 2.1 software (Applied Biosystems).

Statistical Methods. Deviations of the genotype frequencies in the controls from those expected under Hardy-Weinberg equilibrium were evaluated by $\chi^2$ tests (1 degree of freedom). Breast cancer risk associated with rs2180341 was estimated as odds ratios (OR) and 95% confidence intervals (95% CI) using unconditional logistic regression with multiple genetic models including the genotype model (separate indicators for heterozygotes
and rare homozygotes), dominant model (indicator for heterozygotes and rare homozygotes combined), recessive model (indicator for homozygotes), and log additive (per-allele) model (each copy of rare allele) with the common homozygote as the reference category. For cases with no apparent trend in ORs, in all analysis we also used a 2-degrees-of-freedom test (genotype test). All models were adjusted for continuous age at diagnosis (cases) or at the time of inclusion in the study (controls) and ethnicity (European American, African American, Hispanic, etc.).

**Sequencing.** Coding regions of ECHDC1 and RNFI46 were sequenced by ABI3700 capillary sequencing. The primers were designed to cover entire transcribed regions of both genes and to capture ~50 bp of sequence on both sides of each exon (primer sequences available upon request). Sequencing was done from both directions and sequence data were analyzed by both Sequencher (Genes Codes) and Mutation surveyor (Softgenetics) software packages.

**Cell Culture.** MCF-7 cells (American Type Culture Collection) were grown in normal DMEM supplemented with 10% (v/v) FCS, 0.01 mg/mL bovine insulin, and antibiotics at 37°C in a humidified atmosphere of 95% air and 5% CO2.

**Western Blot Analysis.** MCF7 cells were grown in a 35-mm dish to about 80% confluence. Four micrograms pCMV6-XL5/RNF146 or pCMV6-XL5 (Origin Technologies Inc.) with 8 μL FuGene HD transfection reagent (Roche Applied Science) were mixed and added into the cultured cells according to Roche’s protocol. After 48 or 72 h transfection, the cells were washed with 1X PBS and lysed with 0.5 mL lysis buffer (Pierce Biotechnology) containing a mammalian protease inhibitor mixture (Sigma-Aldrich). The cell lysates were centrifuged at 13,200 rpm in eppendorf-centrifuge, 4°C for 10 min. The supernatants were used for Western blot analysis as described by Chen et al. (18). RNFI46 antibody dilution was as recommended by the Vander (Santa Cruz Biotechnology, Inc.). Horseradish peroxidase–conjugated secondary antibodies and an enhanced chemiluminescence kit were purchased from Amersham Biosciences (GE Healthcare).

**Luciferase Assay.** MCF-7 cells were grown in 24-well plates to near confluence. Two hundred nanograms of DLC1 promoter-Luciferase construct (19) or pCMV3XERE-Luciferase plasmid (superArray Bioscience) and 1 μg pCMV6-XL5/RNF146 or pCMV6-XL5 with 3 μL FuGene HD were mixed and added to plate wells for transfection. After 24 h incubation, estradiol (Sigma-Aldrich) was added to 10 nmol/L final concentration. Luciferase activities of cell lysates were measured according to the manufacturer’s instructions for the Dual Luciferase Assay System (Promega) in a Moonlight 2010 luminometer (Analytical Luminescence laboratory; Turner Designs). The experiments were carried out in triplicate and the statistical analysis was done on the mean values from three independent measurements, given as the mean ± SD.

**Results**

**Reproduction of 6q22.33 Association with Breast Cancer Risk.** In this study, we tested the association of a locus at 6q22.33 with breast cancer in independent populations of breast cancer cases and controls. We had previously shown that in the Ashkenazi Jewish population there is a strong linkage disequilibrium (LD) among four single nucleotide polymorphisms (SNP; rs2180341, rs6569479, rs6569480, rs7776136), corresponding to a breast cancer association signal on 6q22.33 (10). Because strong LD among these four SNPs was also confirmed in other reference populations, such as the population of northern European ancestry (CEU population) in International HapMap project (HapMap), for all replication screens in this study we used rs2180341, which tags to a LD block including the associated locus (data not shown). First, we expanded the initial Ashkenazi Jewish set used in the previous study by an additional 487 clinically ascertained Ashkenazi Jewish breast cancer cases that previously tested negative for BRCA1 and BRCA2 mutations (Table 2A, left panel). These were compared with 1,149 Ashkenazi healthy controls used in our previous study (10). The analysis confirmed a significant increase of breast cancer risk assuming a dominant genetic model (age-adjusted OR, 1.28; 95% CI, 1.03–1.59; P = 0.024), but the association was statistically marginal by per-allele test (age-adjusted OR, 1.18; 95% CI, 0.99–1.41; P = 0.066), possibly due to the small sample size of this replication case group (n = 487). However, the aggregate Ashkenazi Jewish analysis consisting of a total of 1,565 cases and 1,149 controls from present and prior studies confirmed the strong association of 6q22.33 with breast cancer risk (per-allele OR, 1.32; 95% CI, 1.15–1.50; P = 0.000057; Table 2A, right panel).

Second, we did a replication analysis in three sets of non-Jewish breast cancer cases compared with 1,467 non-Jewish controls from the same geographic region and including predominantly European Americans (Table 2B). To replicate the 6q22.33 association from phase I of our previous study, in which we did a GWAS on 250 affected Ashkenazi Jewish probands from families ≥3 breast cancer cases, we genotyped rs2180341 in the cohort of 171 non-Jewish cases enriched for family history of breast cancer (Table 2B, left panel). Although statistically insignificant (per-allele P = 0.24), the genotype frequencies of G/G and A/G (5.3% and 41.5% in cases versus 4.8% and 37% in controls) showed a minor trend toward association that is also reflected in per-allele OR (OR, 1.17, 95% CI, 0.9-1.52). To increase the statistical power, we genotyped additional populations of 1,295 non-Jewish consecutive breast cancer cases consisting of nonoverlapping groups of 751 and 544 non-Jewish breast cancer patients collected at the same cancer center. As seen in the right panel of Table 2B, after adjustment for age and ethnicity, the aggregated analysis of replication non-Jewish breast cancer cases (n = 1,466) and controls (n = 1,467) showed a significant association in all modes of analysis, with the strongest association signal assuming the recessive model (OR, 1.57, 95% CI, 1.15-2.15; P = 0.0039). Although the non-Jewish cohorts in this study were derived predominantly from a European background, a small fraction of the study population represented other ancestries (Table 1A and B). With the exception of the African American population, where we noted differences for rs2180341, the allele and genotype frequencies did not significantly differ in other ancestries (Table 1B). Nevertheless, to correct for population stratification the statistical analyses of non-Jewish cohorts were adjusted for age and ethnicity in multivariate models.

Due to slightly different allele frequencies in African Americans, and in order to verify whether rs2180341 can
Table 2. Replication of 6q22.33 association with breast cancer risk: odds ratios and 95% confidence intervals of rs2180341 and breast cancer risk in:

| A. Ashkenazi Jewish controls versus Ashkenazi Jewish BRCA1/2 negative breast cancer cases (left panel) and aggregate AJ cases (right panel) |
|---|---|---|---|---|
| | AJ controls* | AJ (BRCA1/2 negative) cases | OR (95% CI) | P |
| | n = 1,149 | n = 487 | | |
| A/A | 710 (61.8%) | 271 (55.6%) | 1 | |
| A/G | 382 (33.2%) | 192 (39.4%) | 1.31 (1.05-1.64) | 0.10 |
| G/G | 57 (5%) | 24 (4.9%) | 1.09 (0.66-1.79) | 0.006 |

2 df test
Dominant | 1.28 (1.03-1.59) | 0.024 | 1.42 (1.21-1.67) | 0.000043 |
Recessive | 0.98 (0.60-1.60) | 0.93 | 1.27 (0.90-1.80) | 0.17 |
Per allele | 1.18 (0.99-1.41) | 0.066 | 1.32 (1.15-1.50) | 0.000057 |

B. Non-Jewish controls of mixed ethnicities versus non-Jewish familial cases (left panel) and aggregate of all non-Jewish breast cancer cases of mixed ethnicities (right panel)

| | All non-AJ controls | Non-AJ familial cases | OR (95% CI) | P |
| | | | | |
| | n = 1,467 | n = 171 | | |
| A/A | 854 (58.2%) | 91 (53.2%) | 1 | |
| A/G | 542 (37%) | 71 (41.5%) | 1.31 (0.94-1.83) | 0.63 |
| G/G | 71 (4.8%) | 9 (5.3%) | 1.07 (0.51-2.25) | 0.3 |

2 df test
Dominant | 1.28 (0.92-1.76) | 0.14 | 1.15 (0.99-1.34) | 0.06 |
Recessive | 0.96 (0.67-1.39) | 0.92 | 1.27 (0.90-1.80) | 0.0062 |
Per allele | 1.17 (0.90-1.52) | 0.24 | 1.18 (0.94-1.53) | 0.0081 |

C. Breast cancer cases and controls of European ancestry that are a subset of non-Jewish populations

| European controls | European cases | OR (95% CI) | P |
| | | | |
| | n = 1337 | n = 1117 | | |
| A/A | 783 (58.6%) | 612 (54.8%) | 1 | |
| A/G | 492 (36.8%) | 425 (38%) | 1.1 (0.94-1.31) | 0.016 |
| G/G | 62 (4.6%) | 80 (7.2%) | 1.63 (1.15-2.32) | 0.055 |

2 df test
Dominant | 1.28 (0.92-1.76) | 0.14 | 1.15 (0.99-1.34) | 0.06 |
Recessive | 0.96 (0.67-1.39) | 0.92 | 1.27 (0.90-1.80) | 0.0062 |
Per allele | 1.17 (0.90-1.52) | 0.24 | 1.18 (0.94-1.53) | 0.0081 |

D. Aggregate of replication population of Ashkenazi Jewish and non-Jewish breast cancer cases and controls, excluding previously published data

| All replication controls | All replication cases | OR (95% CI) | P |
| | | | |
| | n = 1,467 | n = 1,953 | | |
| A/A | 854 (58.2%) | 1,060 (54.3%) | 1 | |
| A/G | 542 (37%) | 753 (38.6%) | 1.09 (0.94-1.31) | 0.1 |
| G/G | 71 (4.8%) | 140 (7.2%) | 1.62 (1.18-2.23) | 0.01 |

2 df test
Dominant | 1.28 (0.92-1.76) | 0.14 | 1.15 (0.99-1.34) | 0.06 |
Recessive | 0.96 (0.67-1.39) | 0.92 | 1.27 (0.90-1.80) | 0.0062 |
Per allele | 1.17 (0.90-1.52) | 0.24 | 1.18 (0.94-1.53) | 0.0081 |

E. Aggregate of all case/control ascertainment

| Aggregate controls* | Aggregate cases* | OR (95% CI) | P |
| | | | |
| | n = 2,616 | n = 3,031 | | |
| A/A | 1,564 (59.8%) | 1,626 (53.6%) | 1 | |
| A/G | 924 (35.4%) | 1,192 (39.3%) | 1.25 (1.11-1.40) | 3.65E-07 |
| G/G | 128 (4.9%) | 213 (7%) | 1.51 (1.19-1.92) | 3.30E-06 |

2 df test
Dominant | 1.28 (1.15-1.43) | 0.01 | 1.28 (1.15-1.43) | 0.0065 |
Recessive | 1.38 (1.09-1.75) | 0.000057 | 1.38 (1.09-1.75) | 0.0083 |
Per allele | 1.24 (1.13-1.36) | 0.0083 | 1.24 (1.13-1.36) | 0.0083 |

Abbreviations: AJ, Ashkenazi Jewish; non-AJ, non-Jewish; 2 df, two degrees of freedom.
*Indicates inclusion of Ashkenazi Jewish datasets from a previous study (ref. 10; n = 1,149 AJ controls and n = 1,078 AJ cases). All statistical tests were adjusted for age and ethnicity as detailed in Materials and Methods.
be used to tag the association signal in the non-Jewish population, we examined the LD structure of 6q22.33 in non-Jewish by genotyping all four SNPs from our prior study in a subgroup of 847 non-Jewish controls and confirmed strong LD (D' = 0.98) across all four subpopulations in this set, which was comparable with observations in Ashkenazi Jewish (data not shown). Because of the population substructure of the non-Jewish cases and controls used in this study, we carried out a separate analysis on the population of European ancestry only, and the association remained statistically significant (per-allele OR, 1.19; 95% CI, 1.04-1.35; \(P = 0.01\); Table 2C). The replication of association of 6q22.33 with breast cancer risk was further shown by pooled analysis of all replication ascertainment, excluding the data from the previous Ashkenazi Jewish association study (Table 2D). As shown, the significant associations were observed under all models of analysis with per-allele OR of 1.18 (95% CI, 1.04-1.33; \(P = 0.008\)).

Finally, we did an aggregate meta-analysis, including all aforementioned cases (\(n = 3,031\)) and controls (\(n = 2,616\)) from this and the prior GWAS. As shown in Table 2E, the association, adjusted for age and ethnicity, was consistent with previous findings and was strongly significant (per-allele OR, 1.24; 95% CI, 1.13-1.36; \(P = 2.2 \times 10^{-7}\)).

**Sequencing Analysis of Candidate Genes in 6q22.33.** We previously showed that rs2180341 tags a LD block made up of four SNPs that were equally associated with breast cancer risk covering a region of \(\sim 200\) kb. Two candidate genes map within this region: ECHDC1 and RNF146. To further investigate the functional consequences of 6q22.33 association, we sequenced the transcribed regions of both genes. Besides several rare polymorphisms, for which no significant differences were observed after testing them independently in a subset of case-control population, we did not find any potentially pathogenic mutation that would explain the association signal (Supplementary Table S1).

**Epidemiological and Functional Evaluation of 6q22.33 Association with Estrogen Receptor–Positive Tumors.** To evaluate whether rs2180341 is more strongly associated with estrogen receptor–positive compared with estrogen receptor–negative tumors, we stratified the model by estrogen receptor status (Table 3). We were able to retrieve estrogen receptor status data on 1,658 breast cancer cases that were genotyped in this or the previous Ashkenazi Jewish study. These included 248 estrogen receptor–negative and 979 estrogen receptor–positive tumors from Ashkenazi Jewish patients and 100 estrogen receptor–negative and 331 estrogen receptor–positive tumors from non-Jewish cases. The data were compared with 1,779 European American controls representing the studies for which the estrogen receptor data on breast cancer cases were available, excluding a subset of non-Jewish, ethnically mixed controls (\(n = 837\) with no annotations for estrogen receptor status. Each copy of the rare allele of rs2180341 was associated with slightly stronger risk of estrogen receptor–positive tumors (per-allele OR, 1.35; 95% CI, 1.20-1.51; \(P = 2.2 \times 10^{-7}\)) compared with estrogen receptor–negative tumors (per-allele OR, 1.19; 95% CI 0.97-1.47; \(P = 0.1\); Table 3A). To eliminate the possibility of potential sampling bias in these estimates as suggested in other studies (20), we compared ORs of the cases for which estrogen receptor status data were available to overall ORs of all breast cancer cases used for the study and did not note any significant difference (data not shown).

Although the per-allele ORs for estrogen receptor–positive tumors versus estrogen receptor–negative tumors were not statistically different (per-allele \(P = 0.19\);
Table 3B), a trend resulted from a slightly elevated genotype frequency of G/G (7.1% estrogen receptor positive versus 5.8% estrogen receptor negative) and A/G (41% estrogen receptor positive versus 39.1% estrogen receptor negative) in estrogen receptor-positive tumors, corresponding to increased ORs (per-allele OR, 1.14; 95% CI, 0.94-1.38). Therefore, we further investigated the functional consequences of estrogen receptor-positive tumor association with breast cancer risk in RNF146. We found that overexpression of RNF146 under control of a dynein light chain 1 (DLC1) promoter with a luciferase reporter in MCF-7 cells (Fig. 1A) resulted in remarkably decreased luciferase activity (Fig. 1B). The DLC1 promoter contains half of an estrogen receptor binding site (18). This implies that RNF146 plays a role in estrogen-mediated cellular activities. To verify this possibility, we transfected RNF146 and pCMV-3XERE-Luciferase-luciferase constructs into MCF-7 cells. Similarly in this experiment the luciferase activity was significantly inhibited (Fig. 1C). The pCMV-3XERE-luciferase contains only a minimal cytomegalovirus (CMV) promoter, three copies of an estrogen response element (ERE), and a luciferase reporter. In summary, these experiments show diminished luciferase activity from the ERE reporter vector in the presence of RNF146 overexpression, suggesting that RNF146 plays a role in down-regulation of estrogen response elements.

Discussion

This study represents a replication of the association of the 6q22.33 locus with increased risk of breast cancer. We confirmed a significant association utilizing an expanded set of additional Ashkenazi Jewish cases as well as a pooled analysis of Ashkenazi Jewish case-control data from current and previous studies (Table 2A). The strong association observed in Ashkenazi Jewish was driven by frequency differences in distribution of G/G (62.2% in cases and 5% in controls) and A/G (40.3% in cases and 33.2% controls) genotypes. When we compared the observed genotype frequencies in Ashkenazi Jewish with reported data on 60 CEU individuals from HapMap (G/G 6.7%, A/G 41.7%), we suspected a potential sampling bias due to lower allele frequencies of rs2180341 in the Ashkenazi Jewish control population used in our previous study that might have contributed to false association discovery. The replication analysis in non-Jewish populations, however, showed that the allele and genotype frequencies do not deviate from those observed in Ashkenazi Jewish control cohorts, which was consistent throughout the study in both non-Jewish control ascertainment.

The significant association observed in the aggregate analysis of non-Jewish cases and controls provides the first evidence that genetic variation in 6q22.33 may contribute to breast cancer risk in populations other than Ashkenazi Jewish (Table 2B). Analysis of a small subset of non-Jewish cases derived from families with a strong history of breast cancer did not reveal a significantly stronger association with the 6q22.33 locus than that observed for sporadic breast cancer cases unselected for a family history of the disease, consistent with our previous observations in the Ashkenazi Jewish population (10). Although there was a trend for a slight increase of minor allele homozygotes and heterozygotes for the non-Jewish

Figure 1. RNF146 inhibition of Estrogen Response Element (ERE) reporter gene expression. A. Western blot analysis of RNF146 in MCF7 cells. MCF7 cells were transfected with pCMV6-XL5/RNF146 and pCMV6-XL5 plasmids by FuGene HD. About 0.5 mL of the cell lysates were subjected to immunoprecipitation and Western blot analysis with RNF146 antibody. RNF146 antibody was diluted 500×. Lane 1, pCMV6XL5/RNF146 transfected; lane 2, pCMV6-XL5 transfected. B. RNF146 suppresses ERE-Luciferase activity. After 48 h transfection, the MCF7 were washed with 1XPBS buffer and lysed with 100 μL lysate buffer. The cell lysates were centrifuged at 12,000 rpm at 4°C for 10 min. A 10-μL supernatant was used for Luciferase activity measurement using DLC1 promoter-Luciferase with RNF146 construct assay. C. pCMV-3XERE-Luciferase assay. pCMV-3XERE contains three copies of an ERE and a luciferase reporter.
familial cases, as seen in Table 2B, the addition of this “enriched” group did not offer an obvious advantage in powering this study.

Because we used populations of mixed ethnicities in the replication analysis, there is the potential for population admixture, or confounding by ancestry. As we have detailed in Table 1A, besides being predominantly of European ancestry, a small fraction of our sample collections represents other ancestries. This may, in turn, affect overall association findings if the frequencies of rs2180341 differ among these populations or if LD structure of the region varies substantially. After quantification of admixture in non-Jewish case/control populations based on allele and genotype frequencies (Table 1B) we did not note significant differences among European, Asian, or Hispanic populations. We noted, however, a slightly elevated minor allele frequency of this SNP in a subset of African American ancestry (37% versus 26% in Europeans). These findings prompted us to control all statistical analyses for ethnicity. The different allele frequencies in AAs were, however, present equally in controls as well as cases, which predicts no dramatic effect on overall association statistics.

The conclusion that population stratification in this study does not affect association findings was further supported when the analysis was restricted to European ancestry only (Table 2C). The significant breast cancer risk effect associated with 6q22.33 in the European American subset was no different compared with the overall association statistics in the entire ethnically mixed non-Jewish set (Table 2B, left panel) as well as pooled analysis of all replication case-controls population (Table 2D). This suggests that population substructure does not significantly confound association findings in this study and that adjustment for ethnicity provided sufficient correction for population stratification. Nevertheless, a preliminary analysis of a small African American subset showed a trend towards the association of 6q22.33 with breast cancer risk in AA, with higher minor allele frequencies for both cases and controls compared with samples of European ancestry (OR, 1.16; 95% CI, 0.72-1.85; P = 0.53; complete data not shown). However, these data were limited by very small sample size, and a separate association study in this population is currently underway.

The population differences in prior GWAS to date (7, 9) may explain the failure to detect the association with 6q22.33, as well as the inconsistent replication of other loci, for example the chromosome 2q35 detected in the Iceland cohorts but not in the initial United Kingdom-led consortium (8). Besides global population stratification, other factors may account for these differences among studies. These include sample size differences, geographical differences within the same population, and possible sample ascertainment bias that may result in different pools of cancer predisposing alleles in different studies. Although findings from both Ashkenazi Jewish and non-Jewish populations in our study replicate the association of 6q22.33 with breast cancer risk when analyzed separately as well as in aggregate (Table 2E), we observed a significantly stronger effect in the Ashkenazi Jewish population as compared with non-Jewish, predominantly European, case-control ascertainment. This may indicate population differences in LD structure of the region, although the LD block of 200 kb defined by perfectly correlated SNPs (rs2180341, rs6569479, rs6569480, rs7776136) is comparable in both population ascertainties. In non-Jewish, however, there may be differences in the extent of LD beyond the conserved 200-kb locus. Other variants near this region will need to be genotyped in order to better define the association signal outside Ashkenazi Jewish ancestry.

As the case for other loci mapped by GWAS, sequencing of coding regions of two candidate genes in the implicated region failed to identify any mutation or variant linked to the associated allele. It is possible that the breast cancer–causative variant in 6q22.33 maps outside the 200-kb core region. If the extent of LD is reduced in the non-Jewish population, this will also cause a reduction of the association signal. A more distant variant linked to the 200-kb core region on 6q22.33 could be involved in distant transcription regulation of either of the two candidate genes in this locus; further fine mapping will be needed to address these points. Such long-range effects could, in turn, distally affect the expression of two candidate genes mapping in this region, ECHDC1 and RNF146. A detailed expression analysis in a large collection of pairs of clinically well-annotated normal tissue specimens as well as corresponding breast tumor counterparts is currently underway to correlate gene expression of both candidate genes with the presence of high risk allele at the 6q22.33 locus.

There was also evidence for a possible mechanism of the putative association of the 6q22.33 locus with breast cancer risk suggested by analysis of the estrogen receptor status in our case populations. These findings were similar to other GWAS reports in which novel breast cancer susceptibility loci were slightly strongly associated with estrogen receptor–positive than with estrogen receptor–negative tumors (21, 22). In the case of 6q22.33, we found a statistically significant association with risk of estrogen receptor–positive tumors (Table 3A), but the differences in the associations for risk of estrogen receptor–positive tumors did not differ from estrogen receptor–negative tumors in our study (Table 3B). A likely explanation was the limited power of this comparison due to the small sample size for estrogen receptor–negative cancers in this and other analyses. The association of 6q22.33 with estrogen receptor–positive tumors would also predict an association with greater age at onset in cases; however, no such difference was noted (data not shown).

Although a larger series will be needed to confirm the 6q22.33 association with estrogen receptor–positive tumors, this finding provides an avenue for further biological investigation. Among the two candidate genes at 6q22.33, RNF146 was found to down-regulate an estrogen response driven reporter gene in vitro in MCF-7 breast cancer cells (Fig. 1). We interpret this to mean that modest up-regulation of RNF146 expression associated with high-risk allele could interfere with estrogen signaling. Although the current experimental evidence from the literature describing the function of RNF146 is limited, the protein dactylin has been shown to possess protein ubiquitin (E3 ubiquitin) ligase activity. These enzymes were observed to be highly overexpressed in many cancers and their overexpression was often associated with poor prognosis (23-25). Moreover, many E3 ligases are targets for small molecules in anticancer therapies (26). Several recent studies have shown that nuclear E3 ubiquitin ligases may possess other essential functions besides their role in ubiquitination, i.e., in modulating the estrogen receptor pathway (27, 28). Our data suggest that RNF146 may be another negative regulator of estrogen response in breast cancer.
tissue. We hypothesize that 6q22.33 high-risk allele harbors a genetic variation that may impact gene expression of RNF146. Thus, overexpressed RNF146 subsequently diminishes the estrogen receptor response upon activation by estrogen, perhaps due to competition with estrogen via ERα regulatory elements. This observation is in accordance with the slightly stronger association of 6q22.33 in estrogen receptor–positive tumors. A minor allele at this locus, associated with breast cancer risk in estrogen receptor–positive tumors, may also be associated with elevated expression of RNF146. Testing this hypothesis will require additional investigation using approaches such as full sequencing of the region to identify putative functional genetic variation associated with 6q22BC risk alleles, as well as correlated genotype/expression profiling in tumors and normal tissues.

Overall, the findings presented here add further support for the 6q22.33 locus as a candidate for further investigation in breast cancer genetic epidemiologic investigations. The 30% increased risk associated with this allele in final aggregate analysis (Table 2E) is comparable with confirmed breast cancer loci identified in recent whole genome scans. As a single marker, rs2180341 is of limited clinical or predictive utility. Also, as shown recently, multiplicative models including the several breast cancer risk variants discovered to date (9) do not yield a significant improvement in individual predictive efficacy, although public health uses for genetic panels of such predictive markers may emerge in the future (29). Eventually, a large-scale analysis of all significant “hits” from these studies, including modeling of potential gene–gene and gene–environment interactions, may provide more clinically relevant risk prediction models. Additional approaches, such as epigenome mapping and tumor expression profiling, may also help account for the remainder of the hereditary fraction of breast cancer risk. For the 6q22.33 and other loci there also remains the need for additional epidemiologic and clinical information to adjust the associations for other potential confounding variables, i.e., traditional breast cancer risk factors. Because of the small breast cancer risk effect predicted for this locus, this would require a large association analysis of well-annotated case/control collections. For that purpose, to precisely estimate the association of 6q22.33 and breast cancer risk and to evaluate the possible effects of population stratification, a large international pooled analysis of case-control studies is underway in the Breast Cancer Association Consortium. This study involves the analysis of 6q22.33 genotypes in >50,000 breast cancer cases and controls.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

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.

We thank Mathew Danzig for review of charts for estrogen receptor data. Control samples were provided by the New York Cancer Project supported by the Academic Medicine Development Company of New York.

References

The 6q22.33 Locus and Breast Cancer Susceptibility


Cancer Epidemiol Biomarkers Prev  Published OnlineFirst August 18, 2009.

Updated version  Access the most recent version of this article at:
doi:10.1158/1055-9965.EPI-09-0151

Supplementary Material  Access the most recent supplemental material at:
http://cebp.aacrjournals.org/content/suppl/2009/08/17/1055-9965.EPI-09-0151.DC1

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, use this link http://cebp.aacrjournals.org/content/early/2009/08/14/1055-9965.EPI-09-0151. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.