Racial Disparity in Breast Cancer and Functional Germ Line Mutation in Galectin-3 (rs4644): A Pilot Study

Vitaly Balan,1 Pratima Nangia-Makker,1 Ann G. Schwartz,2 Young Suk Jung,1 Larry Tait,3 Victor Hogan,1 Tirza Raz,1 Yi Wang,2 Zeng Quan Yang,2 Gen Sheng Wu,2 Yongjun Guo,6 Huixiang Li,2 Judith Abrams,3 Fergus J. Couch8 Wilma L. Lingle,8 Ricardo V. Lloyd,8 Stephen P. Ethier,3 Michael A. Tainsky,4 and Avraham Raz1

1Tumor Progression and Metastasis, 2Population Studies and Prevention, 3Molecular Biology and Genetics, School of Medicine, 4Biostatistics Core, Karmanos Cancer Institute, The Department of Pathology, School of Medicine, Wayne State University, Detroit, Michigan; 5Tankenau Institute for Medical Research, Wynnewood, Pennsylvania; 6The First Affiliated Hospital of Zhengzhou University, Zhengzhou, China; and 7Department of Laboratory Medicine and Pathology, Mayo Clinic College of Medicine, Rochester, Minnesota

Abstract

For reasons largely unknown, Caucasian women are at a significantly higher risk of developing breast cancer than Asian women. Over a decade ago, mutations in BRCA1/2 were identified as genetic risk factors; however, the discovery of additional breast cancer genes and genes contributing to racial disparities are lacking. We report a functional germline mutation (polymorphism) in the galectin-3 gene at position 191 (rs4644) substituting proline with histidine (P64H), which results in susceptibility to matrix metalloproteinase cleavage and acquisition of resistance to drug-induced apoptosis. This substitution correlates with incidence of breast cancer and racial disparity. Genotype analysis of 338 Caucasian (194 disease free and 144 breast cancer patients) and 140 Asian women (79 disease free and 61 breast cancer patients) showed that the allele homozygous for H64 exists in disease free Caucasian and Asian women at a frequency of 12% and 5%, respectively, versus 37% and 82% in breast cancer patients. The data indicate that H/H allele is associated with increased breast cancer risk in both races. The data implicate galectin-3 H64 in breast cancer and explain, in part, the noted racial disparity, thus providing a novel target for diagnosis and treatment. [Cancer Res 2008;68(24):10045–50]

Introduction

Breast cancer is the second leading cause of cancer deaths among women. It is estimated that in 2008, ~182,460 new cases of invasive breast cancer will be diagnosed among women in the United States (1). Although many epidemiologic and etiologic risk factors have been identified, the cause of any individual breast cancer is often unknown. Breast cancer incidence and mortality can vary tremendously with race. Population genetics studies revealed genetic variation within racial or ethnic subpopulations, but to what degree such genetic variability accounts for the differences in disease outcomes is still unknown. Investigation of the risk factors associated with racial and ethnic populations may highlight how they might affect prevalence, severity, and responses to treatment.

To date, numerous attempts have been made at identifying potential breast cancer risk alleles, genetic signatures, and their distribution across populations. However, because the identification of the two major breast cancer hereditary susceptibility genes BRCA1 and BRCA2 that are responsible for ~5% of cases, no other obvious candidate genes contributing to a significant breast cancer risk have been identified (2). Of note, within BRCA1 and BRCA2 carriers, broad variation in breast cancer risk exists with additional factors contributing to breast cancer development (3). Mutations in p53, ATM, CHEK2, PIK3CA, and overexpression of genes such as HER-2, CCND1, and others by gene amplification have been identified in subfractions of breast cancer cases (4, 5). A number of recent reports have focused on single nucleotide polymorphisms (SNP) with the notion that such a genetic variations can/may contribute to breast cancer development as well as other cancers (3, 6). Genomic-wide association studies (GWAS) and intensive bioinformatic searches of existing databases showed that non-synonymous (ns) SNPs of a few candidate genes exhibit limited association with breast cancer (3, 4, 6). Although a powerful experimental tool, GWAS is not without challenges; critical to success is the development of robust study designs, sufficient sample sizes, rigorous phenotypes, and comprehensive maps. Candidate gene studies conducted to date have focused mainly on established oncogenes or genes of perceived interest, whereas many other genes of potential significance have been overlooked. For example, the commercially available Affymetrix Genome-Wide Human SNP Array 6.0 features 1.8 million genetic markers, includes >906,600 SNPs, although still missing numerous others. The Human Hap550 chip from Illumina suffers from a similar shortcoming. One such missing msSNP that has not been evaluated to date is the rs4644 in the galectin-3 gene. Galectin-3 is a member of the galectin gene family expressing binding/specificity to β-galactoside residues through evolutionarily conserved sequence elements of a carbohydrate recognition/binding domain and a collagen-α-like motif cleavable by matrix metalloproteases (MMP; ref. 7). Clinical investigations have shown a correlation between expression of galectin-3 and the malignant properties of several types of cancer (8); consequently, galectin-3 is thought to be a cancer-associated protein. An allelic variation (C to A) in the DNA sequence of galectin-3 at position 191 (rs4644), leading to the expression of either proline (P) or histidine (H) at position 64, was previously noted and regarded as a nontumor specific variation (9). Recently, we have addressed the functional significance of this variation.
Amino acids Ala62-Tyr63 of galectin-3 harbor the cleavage site for MMPs (7) and substitution of the subsequent amino acid H64 with P resulted in resistance to cleavage by MMPs (10). Introducing galectin-3 H64 into galectin-3 null cells resulted in acquisition of tumorigenicity and metastatic phenotype in experimental models (11). In sharp contrast, cells transfected with galectin-3 P64 result in reduced tumorigenicity associated with abrogated angiogenesis and increased apoptosis, suggesting that cleavage of galectin-3 H64 by MMPs plays a role during tumor development/progression (10). In this article, we have performed the genotype analysis of the nsSNP (rs4644) in cancer-free and breast cancer patients, and correlated it with the incidence of breast cancer in Caucasian and Asian populations.

Materials and Methods

Control and patient population. The racial distribution of SNP rs4644 in the normal population was retrieved from Ensembl release 59 (50 Caucasian and 59 Asians women). DNA samples were randomly selected from women of the Southeast Michigan Tri-county area (139 Caucasian), and an additional 25 disease free volunteers (5 Caucasian and 20 Asian). DNA samples were isolated from the frozen tissues of 43 Caucasian and 45 Asian (Asterand) and 71 Caucasian and 1 Asian women (Southeast Michigan Tri-county area). Matched DNA (from blood, frozen normal, and cancer) and normal breast DNA were obtained from Asterand. Matched DNA from frozen cancer, normal tissue and blood, and DNA from cancer and normal tissue from 17 Caucasian patients were obtained from the Biospecimens Core of the Mayo Breast Cancer Specialized Programs of Research Excellence (SPORE). The mean age of the combined breast cancer group was 51.7 y. In all cases, a pathologist independently confirmed the diagnosis. DNAs were isolated using AccuPrep Genomic DNA extraction kit (Bioneer).

PCR, sequencing, and RFLP. The following primer pair was designed to amplify exon 3 of human galectin-3 gene Sense 5’TCCATATGGCCGTAT- CTGGGTTCTGG 3’ Anti-sense 3’ CCTATTGGGCCCCTGCTTTGCCACTG 5’. The target sequence was amplified using Eppendorf Thermal Cycler by Qiagen Fast Cycling PCR. The resulting 324 bp DNA fragment was purified using QIAquick PCR Purification kit (Qiagen), visualized on 1% agarose gel containing Syto 60 stain, scanned by Odyssey IR imaging system (LI-COR Biosciences), and sequenced using the sense primer at agarose gel containing Syto 60 stain, scanned by Odyssey IR imaging system (LI-COR Biosciences), and sequenced using the sense primer at an OLYMPUS BX40 microscope supporting a Sony DXC-979MD 3CCD video camera and stored with the M5+ microcomputer imaging device (Interfocus).

Cell viability and apoptosis. BT-549, BT-549 Gal-3 H64, and BT-549 Gal-3 P64 cells were seeded at a density of 2,500 cells per well in triplicates onto 24-well culture dishes. In one set of wells, 25 μmol/L cisplatin (Sigma-Aldrich, Inc.) was added after 24 h of seeding, the other nontreated set served as control. Cell viability was determined after 72 h by the mitochondrial-dependent reduction of 3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide (MTT; Sigma-Aldrich, Inc.) to formazan.

To determine if cell death is due to apoptosis, 1×106 cells were seeded in 100-mm Petri dishes and treated similarly with cisplatin. After 24, 48, and 72 h, the cells were trypsinized, lysed, and equivalent number of cells (1×107) were subjected to SDS-PAGE and Western blot analysis with a 1:200 dilution of PARP mouse monoclonal antibody (BioMol). Blots were also immunoreacted with a 1:5,000 dilution of anti–β-actin mouse monoclonal antibody to normalize for variation in protein loading. After washes, the blots were reacted with secondary antibody mix containing 1:5,000 dilutions of IR Dye 680 or IR Dye 800 conjugated corresponding antibody (Molecular Probes) and scanned by Odyssey IR imaging system to identify the respective proteins.

Immunohistochemical analysis. A breast cancer progression tissue array (BR480) containing 48 tissue cores of 2-mm size was purchased from U.S. Biomax. Four-micrometer serial sections were deparafinized, rehydrated, blocked, and incubated with primary antibodies (anti–galectin-3 monoclonal, anti–galectin-3 polyclonal) at 4°C overnight at the suitable dilution and linked with the appropriate host secondary antibodies (Vector Laboratories) tagged with Avidin biotinylated horseradish peroxidase, colorized with 3'-diaminobenzidine, and counterstained with hematoxylin. Visualization and documentation were accomplished with an OLYMPUS BX40 microscope supporting a Sony DIX-C979MD 3CCCD video camera and stored with the M5+ microcomputer imaging device (Interfocus).

Statistical analysis. The assumption of Hardy-Weinberg equilibrium in the control groups was evaluated using Fisher exact tests. Odds ratios (OR) were used to estimate the magnitude of the association of breast cancer with the alleles HH and HP. Cornfield’s method was used to estimate 95% confidence intervals (CI). If there was no evidence of heterogeneity in the ORs based on Tarone’s test, the Mantel-Haenszel combined OR was calculated to estimate the overall OR across ethnic groups. Two-sided P values are reported. Calculations were made with Stata SE 10.2.

The assumption of Hardy-Weinberg equilibrium was found to be tenable in both the Asian control group (P = 0.12) and the Caucasian control group (P = 0.4). The Biostatistics Core of the Karmanos Cancer Institute performed statistical analysis. Microsoft Excel software was used for calculating statistical significance of apoptosis assay by two sample t test: using unequal variance. P values of <0.05 were considered statistically significant.

Results and Discussion

To examine the possibility that the P64H substitution contributes to breast cancer, we initially screened a panel of commonly used human breast cancer cell lines, analyzed their galectin-3 genomic sequence, and compared it with their reported experimental tumorigenicity in nude mice. The normal human breast epithelial cell line (MCF10A) was found to be homozygous for the P allele similar to normal human fibroblast cells (HS68 and IM90).
Two breast carcinoma cell lines BT-549 and SK-BR-3 (11) that are galectin-3 null and exhibited none to very low tumorigenicity in experimental models were homozygous for P allele at DNA level, whereas the tumorigenic human breast carcinoma cell lines, MDA-MB-231, MDA-MB-435, SUM 149, SUM 1315, and SUM 102 were homozygous for the H allele and the MDA-MB-468 (12, 13) was heterozygous carrying both alleles (Fig. 1A). Thus, we hypothesized that in these cell lines galectin-3 P64 and galectin-3 H64 could be associated with normal and tumor phenotypes, respectively. Next, we questioned whether this genotype variation is human specific or is evolutionarily conserved. Homology search of galectin-3 from chicken to mammals revealed that appearance of H64 variant is relatively late occurrence in evolution and found only in humans (Fig. 1B).

Genomic DNAs from healthy women were isolated from two control studies: one, random digit dialing, the other, volunteers. The data independently obtained from sequencing and RFLP analysis revealed that 64% of the cancer-free Caucasian women were either homozygous (12%) or heterozygous (52%) for H64 and the remaining 36% were homozygous for P64. In contrast, 70% of the cancer-free control Asian women were homozygous for P64, the remaining 30% carried at least one H64 allele, 5% of which were homozygous and 25% were heterozygous (Fig. 2A). This racial difference of genotype distribution is in accordance with the data retrieved from available databases (Ensembl release 50; Fig. 2B).

The National Cancer Institute Surveillance, Epidemiology, and End Results (SEER) data indicated that Caucasian women have ~1.5 times higher propensity to develop breast cancer compared with Asian women (Fig. 2B, inset). Genotype distribution of rs4644 in Caucasian and Asian breast cancer patients showed that 70% and 97% of the tumor samples from Caucasians and Asians, respectively, harbor the H64 allele, of which 37% of the Caucasians and 82% of the Asians were homozygous (Supplementary Data Table S1; Fig. 3A).

Kienan and colleagues (14) have recently reported that Asian and European ancestors shared the same population bottleneck expanding out of Africa, but that both also experienced a more recent genetic drift, which was greater in Asians. We presume that this genetic drift led in part to an enrichment in the Asian population of P genotype due to evolutionary selection. Breast cancer is a complex disease, consisting of diverse structures, genetic, epigenetic variations, and clinical outcomes. Genetic linkage studies have not identified a significant number of breast

Figure 1. A, distribution of nsSNP rs4644 among cell lines and their tumorigenicity. B, multiple sequence alignment for galectin-3 of chicken and a few mammalian species. The dots in alignment represent gaps. Gray boxes, the identity between the proteins. X, either H or P. The numbers represent the position of amino acid in human galectin-3.

Figure 2. Genotype distribution of nsSNP rs4644 in normal population by race. A, genotype distribution of rs4644 from 144 cancer-free Caucasian and 20 Asian volunteer women. B, data from 50 Caucasian and 59 Asian unrelated women were retrieved from Ensembl release 46 (from the database we have excluded family related members). Black bar, H-homozygous alleles; lined bar, H/P- heterozygous alleles; open bar, phomozygous alleles. Inset, breast cancer incidence by race. Rates are expressed as cases per 100,000. Statistics were generated from malignant cases only. SEER Program<sup>10</sup> SEER Stat Database: Incidence-SEER 17 Regs Limited-Use, Nov 2006 Sub (1995–2004), National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch, released April 2007, based on the November 2006 submission.

<sup>10</sup> http://www.seer.cancer.gov
cancer-causing genes other than BRCA1/2 (15). This lack of identifiable susceptibility genes suggests that breast cancer is most probably conferred by a large number of loci, each contributing in a small additive manner to the overall breast cancer risk (4). In addition, variation in risk for breast cancer has been attributed in part to diet, local environmental factors, and/or yet to be identified genetic predisposing factors. Thus, to date, the majority of clinical cases lack identifiable risk factor(s) other than gender, age, race, and family history. The differences in the propensity to develop breast cancer are likely attributable to multiple genes/factors, and the data presented here suggest that the reported H64 and P64 allelic variation may contribute, in part, to the racial disparity observed between Caucasian and Asian women in the incidence of breast cancer. There are statistically significant differences in the odds of breast cancer for the H/H genotype compared with the P/P genotype for Caucasian and Asian patients were 2.7; 95% CI, 1.4–4.9; \( P = 0.001 \) and 94.6; 95% CI, 10.0–892.4; \( P < 0.001 \), respectively.

To establish whether this tumor-specific genotype reflects a somatic or germline change and/or loss of heterozygosity, genotyping was performed in RNA and/or DNA isolated from cancer tissue, surrounding normal tissue and blood of Caucasian and Asian breast cancer patients (Supplementary Table S2). The genotype distribution of rs4644 from these samples was analyzed Figure 3. Genotype distribution of rsSNP rs4644 in breast cancer patients by race. \( A \), genotype distribution of rs4644 from 141 Caucasian and 61 Asian breast cancer patients was determined using independent sequencing and RFLP analyses. \( B \), RFLP and sequence analysis of the C191A site of the galectin-3 gene. 324-bp amplified DNA was digested with NcoI and the resulting fragments were separated by electrophoresis on a 2% agarose gel at 100 V for 30 min. PCR product containing the H allele can be digested by NcoI creating two fragments of 171 and 153 bp length (lanes H/H and H/P). Sequencing of the galectin-3 gene around the C191A site was performed in Applied Genomics Technology Center using the sense primer (inset, DNA sequence chromatograms indicating allelic variation). The OR of breast cancer for the H/H genotype compared with the P/P genotype for Caucasian and Asian patients were 2.7; 95% CI, 1.4–4.9; \( P = 0.001 \) and 94.6; 95% CI, 10.0–892.4; \( P < 0.001 \), respectively.

Figure 4. Cisplatin induced apoptosis of BT-549 cells and cloned variants. \( I \), percent apoptosis in cells treated with 25 \( \mu \)mol/L cisplatin for 72 h at 37 °C and analyzed after MTT assay. Absorbances in the untreated control cells were given a value of 100 and relative percent apoptosis was calculated as 100%, and the values are the average of triplicate experiments. Columns, mean; bars, SD. *, \( P > 0.001 \). \( II \), Western blot analysis of the expression of PARP in cells treated with 50 \( \mu \)mol/L cisplatin after 24, 48, and 72 h. \( A \), BT-549; \( B \), BT-549 Gal-3 H64; \( C \), BT-549 Gal-3 P64. j-Actin was used as loading control. 1, untreated; 2, 24 h; 3, 48 h; 4, 72 h. Data shown is representative of three clones for each variant. \( III \), immunohistochemical analysis of intact and cleaved galectin-3 in breast cancer progression. \( A \) to \( C \), intact protein was visualized using monoclonal anti–galectin-3 antibody TIB166; \( A' \) to \( C' \), full-length plus cleaved galectin-3 using polyclonal galectin-3 antibody hL31. For detailed description see 10. Arrows in \( B' \) and \( C' \), the cleaved protein. Representative pictures are presented from a study of 5 normal breast tissues, 10 lobular hyperplasia, and 8 infiltrating lobular carcinoma cases. Magnification, \( \times 200 \).
Galectin-3 is implicated with the progression and metastasis of several types of cancers by affecting cell growth, differentiation, transformation, angiogenesis, immune response, and apoptosis (16). Although most of these functions are positively influenced by galectin-3, it may act as a double-edged sword either protecting against or stimulating cell death. Overexpression of the galectin-3 H64 variant resulted in resistance to cisplatin-induced apoptosis in transfected BT-549 breast cancer cells (17). Of note, the secreted form could lead to apoptosis of tumor infiltrating T cells leading to the evasion of cancer cells from the host immune system, which is regulated in part by a functional crosstalk between intracellular and extracellular galectins (18). To investigate the possible functional significance of this allelic variation, we used BT-549, a galectin-3 null cell line, to generate cell clones stably overexpressing either galectin-3 P64 or H64 variant allele (10). BT-549 cell clones harboring either H64 or P64 allele were treated with cisplatin. A dose- and time-dependent decrease in the number of viable cells was observed. Cells expressing the H64 galectin-3 variant were resistant to apoptosis with only 36.2% ± 2.2% apoptotic cells, whereas the parental and P64 cell clones were highly susceptible to drug induced apoptosis with 76.5% ± 1.34% and 71.5% ± 3.02%, respectively (Fig. 4f). That the cell death was due to apoptosis was confirmed by the cleavage of PARP. Figure 3B indicates that parental and galectin-3 P64 expressing cells showed the 85 kDa cleaved polypeptide of PARP, whereas galectin-3 H64 cells were refractive to PARP cleavage.

The biological functions regulated by galectin-3 include mRNA splicing, cell growth, cell cycle, and apoptosis resistance, whereas secreted galectin-3 modulates cellular adhesion and signaling, immune response, angiogenesis, and tumorigenesis by binding to cell surface glycoconjugates such as laminin, fibronectin, and collagen IV (16). The functions regulated by secreted galectin-3 such as angiogenesis and tumor growth were inhibited only in BT-549 cell clones harboring the MMP cleavage-resistant galectin-3 P64 variant (10).

An important role for galectin-3 in breast cancer progression has been suggested because it was found to be significantly elevated in the sera of breast cancer patients compared with healthy controls (19). Furthermore, galectin-3 localizes in cells proximal to the stroma in comedo-ductal carcinoma in situ, which could lead to increased invasive potential by inducing interactions with the stromal counterparts (20). Secreted galectin-3 H64 binds to the endothelial cells of the microvessels and induces migration, morphogenesis, and angiogenesis (21). The cleaved galectin-3 H64 bound to the endothelial cell surface with higher affinity (~20 times) than the full-length protein in a carbohydrate-dependent manner, and the MMPs cleavage resistant galectin-3 P64 did not mimic galectin-3 H64 functions (10). Of note, in vivo cleavage of galectin-3 was found to be a surrogate marker for the activity of MMPs in the growing breast cancer (10). Thus, we examined here, whether and at what progressive stage this cleavage of galectin-3 could be detected in breast cancer. Immunohistochemical analysis of a breast cancer progression tissue array with specific antibodies to detect intact (Fig. 4II A–C) and cleaved galectin-3 (Fig. 4II A–C’) revealed that in the normal breast tissue (A–A’), no cleavage of galectin-3 could be detected. In ductal hyperplasia, the intact protein could be detected on the tubular side of the ducts, whereas the cleaved protein could be detected adjacent to the stroma (B–B’). In the infiltrating lobular carcinoma, most of the invasive cell clusters were positive for the cleaved product (C–C’). These results emphasize the significance of P6H1 mutation in breast cancer progression and highlight the role of MMPs in this process.

Because galectin-3 is not a classic oncogene but rather a cancer helper gene product, we suggest that it most probably exerts its function during the progression stage, i.e., after cancer initiation and promotion by other genes/factors. Future larger scale studies will be required to determine which allele is dominant in an H/P heterozygote and whether this affects breast cancer development.

In conclusion, we provide novel data implicating a galectin-3 germine nonsynonymous polymorphism in breast cancer and note that variation in allelic frequencies for this nsSNP between Caucasian and Asian women in the general population may explain, in part, the evolutionary differential incidence of breast cancer observed in the two races.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Acknowledgments

Received 8/19/2008; revised 9/25/2008; accepted 10/18/2008.

Grant support: NIH R01CA6120-19 (A. Raz) and RO1CA87895 (A.G. Schwartiz), NIH Mayo Breast Cancer SPORE Grant CA116201 (F.J. Couch and W.L. Lingle), and the Karmanos Cancer Institute’s Strategic Research Grant (A. Raz and P.N. Makker) is thankfully acknowledged.

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 Des, PVM Shetkar and George Brush for critical evaluation of the manuscript, Nancy K. Levin for management and maintenance cancer samples database, and all the volunteers who participated in this study. The samples were obtained with approval of Institutional Review Boards HIC#: 017208M1X and HIC#: 10330M1E.

References

Racial Disparity in Breast Cancer and Functional Germ Line Mutation in Galectin-3 (rs4644): A Pilot Study
