IJC
International Journal of Cancer

Association between breast cancer genetic susceptibility variants and terminal duct lobular unit involution of the breast

Clara Bodelon[1], Hannah Oh[1], Nilanjan Chatterjee[1], Montserrat Garcia-Closas[1], Maya Palakal[1], Mark E. Sherman[1,2], Ruth M. Pfeiffer[1], Berta M. Geller[3], Pamela M. Vacek[3], Donald L. Weaver[3], Rachael E. Chicoine[3], Daphne Papathomas[1], Jackie Xiang[1], Deesha A. Patel[1], Zeina G. Khodr[1], Laura Linville[1], Susan E. Clare[4], Daniel W. Visscher[5], Carolyn Mies[6], Stephen M. Hewitt[7], Louise A. Brinton[1], Anna Maria Storniolo[8], Chunyan He[9], Stephen J. Chanock[1], Gretchen L. Gierach[1] and Jonine D. Figueroa[1,10]

[1] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, MD
[2] Division of Cancer Prevention, National Cancer Institute, Bethesda, MD
[3] Department of Biostatistics, University of Vermont College of Medicine and Vermont Cancer Center, Burlington, VT, USA
[4] Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL
[5] Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, MN
[6] Department of Pathology and Laboratory Medicine, University of Pennsylvania, Philadelphia, PA
[7] Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, Bethesda, MD
[8] Susan G. Komen Tissue Bank at the Indiana University Simon Cancer Center, Indianapolis, IN
[9] Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, IN
[10] Usher Institute of Population Health Sciences and Informatics and Edinburgh Cancer Research Centre, University of Edinburgh, United Kingdom

Terminal duct lobular units (TDLUs) are the predominant source of future breast cancers, and lack of TDLU involution (higher TDLU counts, higher acini count per TDLU and the product of the two) is a breast cancer risk factor. Numerous breast cancer susceptibility single nucleotide polymorphisms (SNPs) have been identified, but whether they are associated with TDLU involution is unknown. In a pooled analysis of 872 women from two studies, we investigated 62 established breast cancer SNPs and relationships with TDLU involution. Poisson regression models with robust variance were used to calculate adjusted per-allele relative risks (with the non-breast cancer risk allele as the referent) and 95% confidence intervals between TDLU measures and each SNP. All statistical tests were two-sided; P < 0.05 was considered statistically significant. Overall, 36 SNPs (58.1%) were related to higher TDLU counts although this was not statistically significant ($p = 0.25$). Six of the 62 SNPs (9.7%) were nominally associated with at least one TDLU measure: rs616488 (*PEX14*), rs11242675 (*FOXQ1*) and rs6001930 (*MKL1*) were associated with higher TDLU count ($p = 0.047$, 0.045 and 0.031, respectively); rs1353747 (*PDE4D*) and rs6472903 (*8q21.11*) were associated with higher acini count per TDLU ($p = 0.007$ and 0.027, respectively); and rs1353747 (*PDE4D*) and rs204247 (*RANBP9*) were associated with the product of TDLU and acini counts ($p = 0.024$ and 0.017, respectively). Our findings suggest breast cancer SNPs may not strongly influence TDLU involution. Agnostic genome-wide association studies of TDLU involution may provide new insights on its biologic underpinnings and breast cancer susceptibility.

Key words: genetic susceptibility, terminal duct lobular unit, involution, breast cancer
Additional Supporting Information may be found in the online version of this article.
G.L.G. and J.D.F. contributed equally to this work.
Disclaimers: The authors have no conflicts of interest.
Grant sponsor: Intramural Research Program of the National Cancer Institute
DOI: 10.1002/ijc.30512
History: Received 25 Apr 2016; Accepted 5 Oct 2016; Online 9 Nov 2016
Correspondence to: Clara Bodelon, PhD, Division of Cancer Epidemiology and Genetics, 9609 Medical Center Dr., Rm 7-E236, Bethesda, MD 20892, USA, Tel.: [2402767327], Fax: +[240-276-7838], E-mail: clara.bodelon@nih.gov

Terminal duct lobular units (TDLUs) are the anatomical structures of the breast that produce milk and represent the predominant histological source of breast cancers.[1] During normal aging the number and size of TDLUs and acini (epithelial substructures within TDLUs) are reduced through a process known as TDLU involution. Reduced levels of TDLU involution has been found to be associated with increased breast cancer risk,[2,3] yet whether a shared genetic basis exists is unknown.

Genome-wide association studies (GWAS) have identified numerous common breast cancer susceptibility loci[4–24] but it is unclear if they exert their effects through TDLU involution. Understanding whether there are associations between breast cancer susceptibility loci and TDLU involution could provide important insights into the role of these susceptibility loci in breast microanatomy, which in turn is linked to breast cancer risk. We have recently developed quantitative measures of

What's new?

As women age, the ducts in their breasts that produce milk, called TDLUs, begin to shrink and disappear. However, when they disappear more slowly than normal, it may signal an increased risk of breast cancer. These authors wanted to know whether genetic loci linked to breast cancer also relate to TDLU shrinkage. They pooled data from two studies to investigate whether any of 62 breast cancer susceptibility loci were associated with TDLU shrinkage measures. Six SNPs showed a nominal association, but it was not evident that breast cancer genes did influence TDLU appearance.

TDLU involution, including TDLU count standardized to tissue area and median acini count per TDLU, and found multiple associations with breast cancer risk factors including parity,[25] menopausal status[25] and circulating hormones[26] among women who donated normal breast tissues. In women with benign breast disease, we have also found these TDLU measures to be related to breast cancer risk factors, including circulating insulin-like growth factors,[27] estrogens[28] and mammographic density (MD).[29] This supports the notion that some breast cancer risk factors may influence their risk in part through delaying or inhibiting TDLU involution. Furthermore, studies demonstrating lesser TDLU involution among women with a positive family history of breast cancer[2,25] suggest that this trait may have a heritable component.

We investigated the relationships between 62 well-established breast cancer susceptibility loci and standardized measures of TDLU involution by pooling data from two studies, the Susan G. Komen Tissue Bank (KTB) at the Indiana University Simon Cancer Center[30] and the NCI Breast Radiology Evaluation and Study of Tissues (BREAST) Stamp Project.[31]

Material and Methods

Study population

The Susan G. Komen tissue bank. The KTB is a repository of normal breast tissues that has been recruiting women volunteers, aged 18–91, since 2007. This analysis targeted a subset of 2,321 participants recruited from January 10, 2009 through September 14, 2012 and aged 18–84 years that have been previously analyzed for TDLU involution.[25] Details of this study population and subject ascertainment are described elsewhere (http://komentissuebank.iu.edu/).[30] All volunteers provided informed consent for the use of their donated specimens and questionnaire data for breast cancer research. Briefly, paired normal breast tissues and blood samples were collected along with a self-administered questionnaire (including demographic, lifestyle, reproductive and cancer related data). Tissue cores were removed from the upper outer quadrant of the breast (left or right) using a standardized technique with a 10-gauge needle. One sample was fixed in 10% buffered formalin, routinely processed to prepare paraffin embedded blocks, sectioned at 5 μm, and stained with hematoxylin and eosin (H&E).

Of the 2,321 women, we excluded 801 subjects who were missing genotyping data for the single nucleotide polymorphisms (SNPs) of interest, previously diagnosed with cancer ($n = 177$), pregnant ($n = 21$), non-white ($n = 555$) or currently taking hormone therapy ($n = 95$). Repeated samples from the same donors were excluded ($n = 124$) and one subject with uninterpretable tissue morphology was also excluded. The final analytic population was 548 women. All KTB data were collected with the approval of the Indiana University Institutional Review Board (IRB) and the National Institutes of Health Office of Human Subjects Research (NIH OHSR #4508).

The BREAST stamp project. Women diagnosed with benign breast disease enrolled in the BREAST Stamp Project, a cross-sectional molecular epidemiologic study of MD, were also included in the study population.[31] A total of 465 women, 40–65 years of age, were clinically referred to undergo image-guided breast biopsy following an abnormal breast imaging exam at the University of Vermont College of Medicine and its affiliated academic medical center, the UVM Medical Center (formerly Fletcher Allen Health Care) and were enrolled from 2007 through 2010. Women who were diagnosed with *in situ* or invasive breast cancer ($n = 78$), missing SNP ($n = 3$) or tissue morphology data ($n = 14$), non-white ($n = 23$) and/or currently taking hormone therapy ($n = 23$) were excluded, leaving an analytic population of 324 BREAST Stamp participants in our analytic population.

Demographic and breast cancer risk factor information were collected *via* a self-administered questionnaire and a Supporting Information telephone interview. Participants underwent clinically indicated ultrasound-guided (14-gauge needle) or vacuum-assisted (9-gauge needle) breast biopsies, which were processed as formalin-fixed paraffin-embedded blocks, sectioned at 5 μm and H&E stained.

Digital raw mammographic images were transferred to the University of California at San Francisco for quantitative density assessment.[29,31] Area measures of density were estimated using computer-assisted thresholding software. One trained experienced reader measured absolute dense area (cm^2) by setting a pixel threshold for dense tissue. Nondense area was defined as the difference between the total breast area and the dense area. The percentage of dense area was calculated by dividing absolute dense breast area by total breast area (*i.e.*, absolute dense area + absolute nondense area) and multiplying by 100. Participants provided written informed consent and the study was approved by the IRBs at the University of Vermont and the NCI.

KTB sample collection and DNA extraction

Details of the KTB samples collection are described in (http://komentissuebank.iu.edu/wp-content/uploads/downloads/2012/10/SOP-002V3.0-Acquisition-of-Whole-Blood.pdf). Briefly, whole blood samples were collected using Vacuette® EDTA tubes. DNA was extracted from blood cells at the Indiana CTSI Specimen Storage Facility (ICTSI-SSF) lab using an AutogenFlex Star (SN 401033) instrument and the Flexigene AGF3000 blood kit for DNA extractions from whole blood specimens following manufacturer's specifications. For this study, a 50 μL aliquot of sample was stored using Biomatrica® DNAstable® Handbook. Samples were reconstituted at the Cancer Genomics Research laboratory (Leidos Biomedical Research, Frederick, MD) for genotyping.

BREAST stamp sample collection and DNA extraction

Whole blood samples were collected pre-biopsy, allowed to clot for 30 min and processed at the University of Vermont General Clinical Research Center using standard techniques. Mouthwash samples were collected as previously described.[32] Blood samples were centrifuged at 3,000 rpm for 15 min, and the serum and clot fractions were frozen at −80°C. Mouthwash samples were centrifuged at 1500*g* for 15 min and buccal cell pellet was resuspended with 3.0 mL TE buffer. The buccal cells were frozen at −80°C. Frozen samples were shipped to SeraCare Life Sciences (Gaithersburg, MD), where they were stored in liquid nitrogen. At SeraCare, leukocyte DNA was isolated from blood clots using phenol chloroform extraction methods, and DNA was isolated from buccal cells using Puregene methods (Gentra Puregene Buccal Cell Kits, Qiagen). DNA was quantified at the Cancer Genomics Research Laboratory with the QuantiFluor® dsDNA System (Promega) according to the manufacturer's instructions.

Breast cancer susceptibility SNPs and genotyping

Sixty-two breast cancer susceptibility SNPs reported in GWAS identified as of 2013 and for which Taqman assays were available and validated were included in this analysis (Supporting Information Table S1). DNA samples were extracted from buffy coat using the Qiagen method according to the manufacturer's instructions. SNPs were genotyped at the Core Genotyping Facility utilizing a Taqman/Fluidigm platform.

TDLU involution assessment

Digitized images of sections were used for quantitative measurements of TDLU involution as described in detail elsewhere.[25,29] Briefly, H&E slides were scanned as digital images suitable for web-based viewing, electronic marking of regions of interest and image analysis on Digital Image Hub software (SlidePath/Leica, Dublin, Ireland). The lasso tool in Digital Image Hub was used to outline and measure the total tissue area (mm^2) on the slides. The study pathologist (MES) evaluated the images to measure the number of TDLUs ("TDLU count"). TDLU analyzer software[26,33] was used to quantify the acini count per TDLU for up to 10 TDLUs per woman and the median value was selected as a single summary measure for each woman. A high intraobserver agreement (Spearman's $r > 0.90$) for the TDLU measures was previously reported.[25,29]

Statistical analysis

Frequencies and percentages were used to describe selected characteristics of the study populations. Analyses were conducted using pooled data from the KTB study and BREAST Stamp Project. In sensitivity analyses, we also conducted analyses separately by study.

Poisson regression models with robust variance were used to calculate per-allele relative risks (RRs) and 95% confidence intervals (CIs) for the association of TDLU measures (*i.e.*, TDLU count and acini count per TDLU) with each breast cancer susceptibility locus. SNPs were modeled using additive coding for the number of breast cancer risk alleles (0, 1, 2). To account for the tissue area on the slide, an offset variable was included in the model. The Wald test was used to assess the linear trend between breast cancer susceptibility SNPs and morphometric TDLU measures. We additionally fitted linear regression models to confirm that the associations were not driven by model assumptions. Linear regression models were used to estimate the per-allele association (β) and 95% CIs with the log-transformed of TDLU count per 100 mm^2 plus one as the outcome, to better approximate the normal distribution, including zero TDLU count. Finally, the combination of TDLU count and acini count per TDLU, modeled as the product of the two variables, was also considered in the above described models.

All multivariable models were adjusted for study population (KTB or BREAST Stamp Project) and age (in categories: <30; 30–39; 40–49; 50–59; ≥60). Interactions between SNPs and other factors (*e.g.*, menopausal status) were assessed by adding a multiplicative term in the regression model. A woman was considered postmenopausal if menstrual periods had stopped >12 months prior to interview, she had undergone a bilateral oophorectomy, or she had undergone a hysterectomy (or gynecologic surgery associated with cessation of menses) and was 55 years of age or older; otherwise, a woman was considered premenopausal. Associations between SNPs and MD measures were estimated for women in the BREAST Stamp Project for whom MD measures were available[29,31] using linear regression models and the square root of the density measure as outcome.

All statistical tests were two-sided and p-values <0.05 were considered statistically significant. Because SNPs examined in this analysis are well-established breast cancer susceptibility loci, we used a threshold of $p < 0.05$ to define significant associations. Since our power to test for individual associations was limited, we also tested for enrichment of associations between SNPs and measures of TDLU involution.[34] We did this using binomial tests in two ways: (*i*)

Table 1. Selected characteristics of women in the Komen Tissue Bank and BREAST Stamp Project

Characteristics	Komen Tissue Bank ($N = 548$)	BREAST Stamp Project ($N = 324$)
Age (years), *n* (%)		
<30	183 (33.4)	0
30–39	102 (18.6)	0
40–49	124 (22.6)	161 (49.7)
50–59	86 (15.7)	124 (38.3)
≥ 60	53 (9.7)	39 (12.0)
Menopausal Status, *n* (%)		
Premenopausal	393 (71.7)	199 (61.4)
Postmenopausal	155 (28.3)	125 (38.6)
Parity, *n* (%)		
Nulliparous	268 (48.9)	82 (25.3)
Parous	280 (51.1)	242 (74.7)
Family history of breast cancer, *n* (%)		
No	421 (76.8)	240 (75.2)
Yes	127 (23.2)	79 (24.8)
TDLUs observed, *n* (%)		
No	170 (31.0)	87 (26.9)
Yes	378 (69.0)	237 (73.1)
Involution measures, median (IQR)		
TDLU counts per 100 mm^2	22.9 (9.8 – 51.5)	18.0 (6.4 – 38.0)
Median acini count per TDLU	13.0 (7.5 – 2.5)	11 (7.0 – 16.5)
(TDLU counts)* (acini count per TDLU)	287.5 (93.9 – 835.8)	189.7 (48.5 – 477.1)

TDLU: Terminal duct lobular units; IQR: Interquartile range.

comparing the proportion of the SNP associations (regardless of statistical significance) in the direction consistent with our expectation (RR > 1.0) based on the relationship of the SNPs with breast cancer to 50%, which is what would have been expected by chance; and (*ii*) comparing the proportion of SNPs with significant associations to 5%, which is expected by chance alone. We also present individual associations between SNPs and TDLU involution measures. All analyses were conducted in the R software environment (version 3.1.1).

Results

A total of 872 women were included in this analysis, 548 (63%) from the KTB and 324 (37%) from the BREAST Stamp Project. Characteristics of the women are shown in Table 1. Women in both studies tended to be premenopausal (71.7% in KTB; 61.4% in BREAST Stamp) and enriched for family history of breast cancer compared with the general population (23.2% in KTB; 24.8% in BREAST Stamp). Measures of TDLU involution, including TDLU count and acini count per TDLU were similar in both populations. After accounting for age, none of the TDLU involution measures were statistically significantly different between the two populations.

Of the 62 known independent breast cancer susceptibility loci evaluated in this study, 36 (58.1%) were found to be associated with higher TDLU count in the Poisson model (RR > 1.0), albeit the vast majority of associations with the individual loci were not statistically significant (Supporting Information Table S1). Furthermore, statistical evidence of enrichment was not observed (58.1 *vs.* 50%, $p = 0.25$). While the direction of the estimates was generally consistent across the Poisson and linear models, when the linear model was used 40 SNPs (64.5%) were found to be associated with higher TDLU count, which is higher than what would have been expected by chance alone (64.5 *vs.* 50%, $p = 0.031$), suggesting an enrichment of associations with TDLU counts. Under the Poisson model, the top three SNPs associated with TDLU count were rs616488 (*PEX14*), rs11242675 (*FOXQ1*) and rs6001930 (*MKL1*), which were associated with higher TDLU count ($p = 0.047$, 0.045 and 0.031, respectively; Table 2), in directions consistent with their associations with breast cancer risk. None of the above associations showed significant heterogeneity by study (Supporting Information Table S2).

Two breast cancer susceptibility SNPs were significantly associated with higher acini count per TDLU (Table 2; Supporting Information Table S3): rs1353747 (*PDE4D; $p = 0.007$*) and rs6472903 (*8q21.11; $p = 0.027$*). To account for the overall epithelial content in the slide, we studied the relationship between SNPs and the combination of TDLU count and acini count per TDLU, which were modeled as the product of these two TDLU measures (Table 2; Supporting Information Table S4). Two SNPs were associated with the product of TDLU and acini counts: rs1353747 (*PDE4D*) was associated with higher epithelial content ($p = 0.024$) and rs204247 (*RANBP9*) was associated with lower epithelial content ($p = 0.017$).

Five SNPs were statistically significantly associated with TDLU count in premenopausal women [rs616488 (*PEX14*), rs6828523 (*ADAM29*), rs11242675 (*FOXQ1*), rs3817198 (*LSP1*) and rs6001930 (*MKL1*)] (Table 3) and two were significantly associated with TDLU count in postmenopausal women [rs1011970 (*CDKN2A/B*) and rs1292011 (*TBX3*)]. Four SNPs, rs6678914 (*LGR6*), rs6828523 (*ADAM29*), rs1011970 (*CDKN2A/B*) and rs11199914 (*10q26.12*), showed statistically significant heterogeneity in their association with TDLU count by menopausal status (*p*-values: 0.027, 0.036, 0.032 and 0.037, respectively; Table 3).

Finally, we examined the association between breast cancer SNPs and MD measures, restricted to women in the BREAST Stamp Project for whom MD measures were available (Supporting Information Table S5). Although some SNPs were suggestively associated with MD, none of the

Table 2. Associations between breast cancer susceptibility loci and measures of TDLU involution with statistically significant associations ($p < 0.05$) among women in the Komen Tissue Bank and BREAST Stamp Project.

			TDLU counts			Acini count per TDLU			(TDLU counts)[1] (acini count per TDLU)		
SNP	Chr	Locus	RR[1]	(95% CI)[1]	*p*-trend[1]	RR[1]	(95% CI)[1]	*p*-trend[1]	RR[1]	(95% CI)[1]	*p*-trend[1]
rs616488	1	*PEX14*	1.16	(1.00,1.35)	0.047	1.08	(0.95,1.22)	0.239	1.07	(0.86,1.33)	0.550
rs1353747	5	*PDE4D*	1.00	(0.80,1.25)	0.981	1.28	(1.07,1.54)	0.007	1.40	(1.05,1.88)	0.024
rs11242675	6	*FOXQ1*	1.13	(1.00,1.28)	0.045	1.10	(0.98,1.23)	0.108	1.15	(0.97,1.37)	0.102
rs204247	6	*RANBP9*	0.93	(0.82,1.05)	0.259	0.92	(0.82,1.03)	0.144	0.81	(0.67,0.96)	0.017
rs6472903	8	*8q21.11*	1.06	(0.87,1.29)	0.563	1.18	(1.02,1.36)	0.027	1.08	(0.85,1.38)	0.515
rs6001930	22	*MKL1*	1.24	(1.02,1.50)	0.031	1.10	(0.92,1.33)	0.292	1.30	(0.99,1.72)	0.058

TDLU: Terminal duct lobular unit; SNP: Single nucleotide polymorphism; Chr: Chromosome; RR: Risk ratio; CI: Confidence interval; TDLU: Terminal duct lobular unit.
[1]Based on a multivariable Poisson regression model with robust variance adjusted for study and age (in categories: <30, 30–39, 40–49, 50–59 and ≥60) and an offset variable accounting for the tissue area on the slide (in log scale). SNPs were modeled using additive coding for the number of breast cancer risk alleles (0, 1, 2).

Table 3. Associations between breast cancer susceptibility loci and TDLU counts by menopausal status (SNPs with $p < 0.05$) among women in the Komen Tissue Bank and BREAST Stamp Project

			Premenopausal[1]			Postmenopausal[1]			
SNP	Chr	Locus	RR	95% CI	*p*-Trend	RR	95% CI	*p*-Trend	*P*-Heterogeneity[2]
rs616488	1	*PEX14*	1.19	(1.01,1.41)	0.040	1.10	(0.81,1.50)	0.548	0.617
rs6678914	1	*LGR6*	0.87	(0.75,1.02)	0.094	1.28	(0.95,1.71)	0.101	0.027
rs6828523	4	*ADAM29*	1.33	(1.07,1.66)	0.011	0.74	(0.45,1.22)	0.236	0.036
rs11242675	6	*FOXQ1*	1.18	(1.03,1.35)	0.020	1.05	(0.79,1.38)	0.749	0.363
rs1011970	9	*CDKN2A/B*	0.94	(0.75,1.18)	0.617	1.43	(1.03,1.98)	0.031	0.032
rs11199914	10	*10q26.12*	1.04	(0.89,1.21)	0.633	0.71	(0.52,0.98)	0.036	0.037
rs3817198	11	*LSP1*	0.83	(0.71,0.97)	0.022	1.08	(0.76,1.52)	0.669	0.165
rs1292011	12	*TBX3*	0.93	(0.80,1.09)	0.381	0.74	(0.55,0.99)	0.044	0.164
rs6001930	22	*MKL1*	1.32	(1.08,1.61)	0.006	0.93	(0.53,1.63)	0.808	0.267

TDLU: Terminal duct lobular unit; SNP: Single nucleotide polymorphism; Chr: Chromosome; RR: Risk ratio; CI: Confidence interval; TDLU: Terminal duct lobular unit.
[1]Based on a multivariable Poisson regression model with robust variance adjusted for study and age (in categories: <30, 30–39, 40–49, 50–59 and ≥60) and an offset variable accounting for the tissue area on the slide (in log scale). SNPs were modeled using additive coding for the number of breast cancer risk alleles (0, 1, 2).
[2]P-heterogeneity was computed as the *p* values associated with the interaction term between menopause and the corresponding SNP.

SNPs associated with TDLU measures was significantly associated with any of the MD measures analyzed.

Discussion

In this first study evaluating the association between established genetic variants associated with breast cancer risk and TDLU involution, we found six SNPs (9.7%) to be nominally associated with at least one standardized TDLU measure using data from two studies with benign tissues, the KTB and the BREAST Stamp Project. Of the six SNPs, TDLU associations for five were in the same direction as their relationship with breast cancer risk. However, there was no evidence for statistical enrichment by either of the two definitions used in our analysis, although this may be due to the small samples size. This suggests that the breast cancer loci included in this analysis do not strongly influence TDLU involution. Future genome-wide association analysis may help to elucidate the biologic underpinnings of TDLU involution and its relationship with breast cancer risk.

Three variants were nominally associated with TDLU count. The risk allele of rs6001930 (*MKL1*), which we found to be associated with higher TDLU count, had been previously associated with increased risk of both estrogen receptor (ER)-positive and ER-negative breast cancers.[21] In addition, rs6001930 (*MKL1*) has also been significantly associated with MD,[35,36] a strong breast cancer risk factor.[37] However, the direction of the association of rs6001930 (*MKL1*) with MD was not consistent with its associations with breast cancer risk: in a large pooled analysis, rs6001930 (*MKL1*) was associated with lower measures of MD, including absolute dense area and nondense area.[36] Among women in our analysis for whom MD measures were available, we also found rs6001930

(*MKL1*) to be inversely associated with absolute dense and non-dense areas. These divergent directions of association are consistent with analyses showing lower levels of TDLU involution and higher MD to be independent risk factors for breast cancer development.[37] Our findings suggest that the association between this SNP and breast cancer risk may be in part due to its influence on higher epithelial cell content and lower levels of TDLU involution, rather than mediated through higher levels of MD, which is thought to predominantly reflect stromal tissue.[38] This result highlights the importance of conducting genetic analysis in relation to both histologic and radiologic measures of breast tissue composition in order to better understand the role of these intermediate endpoints in breast carcinogenesis.

The other two variants associated with TDLU count have been differentially associated with ER-positive and ER-negative breast cancers.[20,21] SNP rs616488 (*PEX14*) has been found to be associated with overall[21] and ER-negative breast cancer specifically.[39] This SNP has been found to be more strongly associated with *PEX14* expression in tumor tissues compared to normal tissues.[21] Our finding that rs616488 (*PEX14*) is associated with higher TDLU count may indicate that the association of this SNP with breast cancer risk may be in part mediated through TDLU involution. We explored the interaction between SNPs and menopausal status and found that several SNPs had different associations with TDLU count for premenopausal and postmenopausal women. In particular, rs616488 (*PEX14*) variant was associated with increased TDLU count only in premenopausal women and has previously been associated with triple-negative breast cancer, which is more common in younger women. If replicated, future studies should further examine whether the association of *PEX14* with breast cancer risk may be due in part to an influence of involution levels in younger women.

The SNP rs11242675 (*FOXQ1*) has been previously found to be associated with both ER-positive and ER-negative breast cancers,[21] and *FOXQ1* has been related to tumor aggressiveness.[40] Consistent with this relationship, we noted that the association with higher TDLU count was only observed among premenopausal women, who tend to be diagnosed with more aggressive breast cancers, suggesting a biologically plausible relationship with TDLU involution. In contrast, the rs6828523 (*ADAM29*) variant, which was previously found to increase risk of ER-positive breast cancer, was also associated with higher TDLU count. The ADAM family of proteins has been associated with cell proliferation and invasion.[41]

We found SNP rs1353747 (*PDE4D*) was significantly associated with both higher acini count per TDLU and overall higher epithelial content, but not with TDLU count alone. While the reasons for this are unclear, TDLU involution is thought to be a sequential process, with initial disappearance of acini and reduction of TDLU size, followed by the disappearance of TDLUs. Therefore, different genetic variants may be involved in different aspects of this complex process.

The biologic mechanisms that are associated with TDLU involution remain poorly understood. Animal models and *in vitro* studies suggest the involvement of insulin-like growth factors (IGFs) and their binding proteins in mammary gland development and involution.[42,43] Recent data also suggest that higher levels of circulating IGFs are related to decreased lobular involution.[44] Fine mapping data of the 2q35 susceptibility locus suggest that breast cancer risk may be mediated through regulation of the IGF binding protein 5 (*IGFBP5*),[45] but whether *IGFBP5* plays a specific role in modulating TDLU involution in humans is unknown. In our analysis, we did not find the variant rs13387042 in 2q35 to be significantly associated with measures of TDLU involution, although a positive relationship was observed. Interestingly, we also saw positive, significant associations of rs13387042 with both percent density and dense area measures (Supporting Information Table S5), which was not observed in a recent pooled analysis of MD.[36] The reasons for these contrary results with density compared to involution measures are unclear. It is possible that this SNP is not associated with involution or we did not have enough power to detect an association with involution measures, as rs13387042 is only associated with a 12% decreased breast cancer risk.[45] Alternatively, this SNP may influence breast cancer risk through other mechanisms such as stromal-epithelial signaling. Future research is needed to clarify whether the IGF family members beyond *IGF1* and *IGFBP3* influence breast tissue composition in humans.

The strengths of this study are the use of well-characterized data of non-malignant tissues and standardized, reproducible measures of lobular involution. While additional studies were not available for replication, we found that the reported associations were similar in the two study populations included in this analysis and generally in the same direction as the SNP-breast cancer risk association. One of the major limitations of our study was the sample size, which may have precluded us from observing additional significant associations and none of the associations remained significant after Bonferroni correction for multiple testing. Sample size calculations indicated that double the sample size would be needed to observed significant associations at the 5% alpha level with 80% power. Another limitation was the differences between the KTB and the BREAST Stamp project. While the former consisted of volunteers for whom collection of tissue samples did not specifically target areas of the breast enriched by epithelial cells, the latter population consisted of women undergoing diagnostic image-guided breast biopsies that were likely enriched for TDLUs. However, associations were in the same direction in both studies and did not differ significantly for the reported SNPs, which gives confidence in the robustness of the results. For our sensitivity analysis, we evaluated these associations using two different models, each one with different underlying statistical assumptions, to further confirm that the reported associations were not driven by statistical properties of the models. We found that the majority of associations, including those reported, were in the same

direction regardless of the model used, although the statistical significance varied by the underlying assumptions of the model. Future research should explore under which circumstances each model may be better suited.

In conclusion, we observed six of 62 breast cancer susceptibility loci nominally associated with TDLU involution, with limited evidence of statistical enrichment. Future approaches with larger sample sizes may provide further insights into the genetics of TDLU involution and its role in breast cancer etiology.

Acknowledgement

This study was supported by the Intramural Research Program of the Division of Cancer Epidemiology and Genetics of the US National Cancer Institute. Breast Cancer Research Stamp Funds (awarded to M.E. Sherman and L.A. Brinton) and cooperative agreement U01CA70013 (B.M. Geller, P.M. Vacek, D.L. Weaver and R.E. Chicoine) from the National Cancer Institute funded some of the data collection for this study. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

References

1. Russo J, Hu YF, Yang X, et al. Developmental, cellular, and molecular basis of human breast cancer. *J Natl Cancer Inst Monogr* 2000;17–37.
2. Milanese TR, Hartmann LC, Sellers TA, et al. Age-related lobular involution and risk of breast cancer. *J Natl Cancer Inst* 2006;98:1600–7.
3. Baer HJ, Collins LC, Connolly JL, et al. Lobule type and subsequent breast cancer risk: results from the Nurses' Health Studies. *Cancer* 2009; 115:1404–11.
4. Cox A, Dunning AM, Garcia-Closas M, et al. A common coding variant in CASP8 is associated with breast cancer risk. *Nat Genet* 2007;39:352–8.
5. Easton DF, Pooley KA, Dunning AM, et al. Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* 2007;447: 1087–93.
6. Hunter DJ, Kraft P, Jacobs KB, et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* 2007;39:870–4.
7. Stacey SN, Manolescu A, Sulem P, et al. Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2007;39:865–9.
8. Stacey SN, Manolescu A, Sulem P, et al. Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. *Nat Genet* 2008;40:703–6.
9. Ahmed S, Thomas G, Ghoussaini M, et al. Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat Genet* 2009;41:585–90.
10. Thomas G, Jacobs KB, Kraft P, et al. A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (RAD51L1). *Nat Genet* 2009;41:579–84.
11. Zheng W, Long J, Gao Y-T, et al. Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat Genet* 2009;41: 324–8.
12. Antoniou AC, Wang X, Fredericksen ZS, et al. A locus on 19p13 modifies risk of breast cancer in BRCA1 mutation carriers and is associated with hormone receptor-negative breast cancer in the general population. *Nat Genet* 2010;42:885–92.
13. Turnbull C, Ahmed S, Morrison J, et al. Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat Genet* 2010;42:504–7.
14. Fletcher O, Johnson N, Orr N, et al. Novel Breast Cancer Susceptibility Locus at 9q31.2: Results of a Genome-Wide Association Study. *J Natl Cancer Inst* 2011;103:425–35.
15. Haiman CA, Chen GK, Vachon CM, et al. A common variant at the TERT-CLPTM1L locus is associated with estrogen receptor-negative breast cancer. *Nat Genet* 2011;43:1210–4.
16. Ghoussaini M, Fletcher O, Michailidou K, et al. Genome-wide association analysis identifies three new breast cancer susceptibility loci. *Nat Genet* 2012;44:312–8.
17. Siddiq A, Couch FJ, Chen GK, et al. A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum Mol Genet* 2012;21: 5373–84.
18. Bojesen SE, Pooley KA, Johnatty SE, et al. Multiple independent variants at the TERT locus are associated with telomere length and risks of breast and ovarian cancer. *Nat Genet* 2013;45: 371–84.
19. French Juliet D, Ghoussaini M, Edwards Stacey L, et al. Functional Variants at the 11q13 Risk Locus for Breast Cancer Regulate Cyclin D1 Expression through Long-Range Enhancers. *Am J Hum Genet* 2013;92:489–503.
20. Garcia-Closas M, Couch FJ, Lindstrom S, et al. Genome-wide association studies identify four ER negative-specific breast cancer risk loci. *Nat Genet* 2013;45:392–8.
21. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat Genet* 2013;45:353–61.
22. Lambrechts D, Truong T, Justenhoven C, et al. 11q13 is a susceptibility locus for hormone receptor positive breast cancer. *Human Mutation* 2012; 33:1123–32.
23. Stevens KN, Fredericksen Z, Vachon CM, et al. 19p13.1 Is a Triple-Negative–Specific Breast Cancer Susceptibility Locus. *Cancer Res* 2012;72: 1795–803.
24. Figueroa JD, Garcia-Closas M, Humphreys M, et al. Associations of common variants at 1p11.2 and 14q24.1 (RAD51L1) with breast cancer risk and heterogeneity by tumor subtype: findings from the Breast Cancer Association Consortium. *Hum Mol Genet* 2011;20:4693–706.
25. Figueroa JD, Pfeiffer RM, Patel DA, et al. duct lobular unit involution of the normal breast: implications for breast cancer etiology. *J Natl Cancer Inst* 2014;106.
26. Khodr ZG, Sherman ME, Pfeiffer RM, et al. Circulating sex hormones and terminal duct lobular unit involution of the normal breast. *Cancer Epidemiol Biomarkers Prev* 2014;23:2765–73.
27. Horne HN, Sherman ME, Pfeiffer RM, et al. Circulating insulin-like growth factor-I, insulin-like growth factor binding protein-3 and terminal duct lobular unit involution of the breast: a cross-sectional study of women with benign breast disease. *Breast Cancer Res* 2016;18:24
28. Oh H, Khodr ZG, Sherman ME, et al. Relation of serum estrogen metabolites with terminal duct lobular unit involution among women undergoing diagnostic image-guided breast biopsy. *Horm Cancer* 2016;7:305–315.
29. Gierach GL, Patel DA, Pfeiffer RM, et al. Relationship of Terminal Duct Lobular Unit Involution of the Breast with Area and Volume Mammographic Densities. *Cancer Prev Res (Phila)* 2016;9:149–58.
30. Sherman ME, Figueroa JD, Henry JE, et al. The Susan G. Komen for the Cure Tissue Bank at the IU Simon Cancer Center: A Unique Resource for Defining the "Molecular Histology" of the Breast. *Cancer Prev Res* 2012;5:528–35.
31. Gierach GL, Geller BM, Shepherd JA, et al. Comparison of Mammographic Density Assessed as Volumes and Areas among Women Undergoing Diagnostic Image-Guided Breast Biopsy. *Cancer Epidemiol Biomarkers Prev* 2014;23:2338–48.
32. Garcia-Closas M, Egan KM, Abruzzo J, et al. Collection of genomic DNA from adults in epidemiological studies by buccal cytobrush and mouthwash. *Cancer Epidemiol Biomarkers Prev* 2001;10:687–96.
33. Rosebrock A, Caban JJ, Figueroa J, et al. Quantitative Analysis of TDLUs using Adaptive Morphological Shape Techniques. *Proc SPIE* 2013; 8676:86760N.
34. Oh H, Bodelon C, Palakal M, et al. Ages at menarche- and menopause-related genetic variants in relation to terminal duct lobular unit involution in normal breast tissue. *Breast Cancer Res Treat* 2016;158:341–50.
35. Lindstrom S, Thompson DJ, Paterson AD, et al. Genome-wide association study identifies multiple loci associated with both mammographic density and breast cancer risk. *Nat Commun* 2014;5:5303.
36. Stone J, Thompson DJ, Dos Santos Silva I, et al. Novel Associations between Common Breast Cancer Susceptibility Variants and Risk-Predicting Mammographic Density Measures. *Cancer Res* 2015;75:2457–67.
37. Ghosh K, Vachon CM, Pankratz VS, et al. Independent association of lobular involution and mammographic breast density with breast cancer risk. *J Natl Cancer Inst* 2010;102:1716–23.
38. Li T, Sun L, Miller N, et al. The association of measured breast tissue characteristics with mammographic density and other risk factors for breast cancer. *Cancer Epidemiol Biomarkers Prev* 2005;14:343–9.
39. Purrington KS, Slager S, Eccles D, et al. Genome-wide association study identifies 25 known breast cancer susceptibility loci as risk factors for triple-negative breast cancer. *Carcinogenesis* 2014;35: 1012–9.
40. Zhang H, Meng F, Liu G, et al. Forkhead transcription factor foxq1 promotes epithelial-

mesenchymal transition and breast cancer metastasis. *Cancer Res* 2011;71:1292–301.
41. Mochizuki S, Okada Y. ADAMs in cancer cell proliferation and progression. *Cancer Sci* 2007;98: 621–8.
42. Ning Y, Hoang B, Schuller AG, et al. Delayed mammary gland involution in mice with mutation of the insulin-like growth factor binding protein 5 gene. *Endocrinology* 2007;148: 2138–47.
43. Sureshbabu A, Tonner E, Flint DJ. Insulin-like growth factor binding proteins and mammary gland development. *Int J Dev Biol* 2011;55:781–9.
44. Horne HN, Sherman ME, Pfeiffer RM, et al. Circulating insulin-like growth factor-I, insulin-like growth factor binding protein-3 and terminal duct lobular unit involution of the breast: a cross-sectional study of women with benign breast disease. *Breast Cancer Res* 2016; 18:1–12.
45. Ghoussaini M, Edwards SL, Michailidou K, et al. Evidence that breast cancer risk at the 2q35 locus is mediated through IGFBP5 regulation. *Nat Commun* 2014;4.