JNCI J Natl Cancer Inst (2017) 109(11): djx058

doi: 10.1093/jnci/djx058
First published online May 26, 2017
Article

ARTICLE

Genome-Wide Association Study to Identify Susceptibility Loci That Modify Radiation-Related Risk for Breast Cancer After Childhood Cancer

Lindsay M. Morton*, Joshua N. Sampson*, Gregory T. Armstrong*, Ting-Huei Chen, Melissa M. Hudson, Eric Karlins, Casey L. Dagnall, Shengchao Alfred Li, Carmen L. Wilson, Deo Kumar Srivastava, Wei Liu, Guolian Kang, Kevin C. Oeffinger, Tara O. Henderson, Chaya S. Moskowitz, Todd M. Gibson, Diana M. Merino†, Smita Bhatia‡, Stephen J. Chanock‡, Margaret A. Tucker‡, Leslie L. Robison‡

Affiliations of authors: Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD (LMM, JNS, THC, EK, CLD, SAL, DMM, SJC, MAT); Department of Epidemiology and Cancer Control (GTA, CLW, TMG, LRL), Division of Cancer Survivorship, Department of Oncology (MMH), and Department of Biostatistics (DKS, WL, GK), St. Jude Children's Research Hospital, Memphis, TN; Cancer Genomics Research Laboratory, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD (EK, CLD, SAL); Departments of Medicine (KCO), Pediatrics (KCO), and Epidemiology and Biostatistics (CSM), Memorial Sloan Kettering Cancer Center, New York, NY; Section of Hematology, Oncology and Stem Cell Transplantation, Department of Pediatrics, University of Chicago, Chicago, IL (TOH); Institute for Cancer Outcomes and Survivorship, University of Alabama at Birmingham, Birmingham, AL (SB)

*Authors jointly led this work.
†For the full list of authors and affiliations, see the Notes section.
‡Authors jointly directed this work.
Correspondence to: Lindsay M. Morton, PhD, Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Institutes of Health, National Cancer Institute, 9609 Medical Center Drive, Room 7E-454, MSC 9778, Bethesda, MD 20892-9778 (e-mail: mortonli@mail.nih.gov).

Abstract

Background: Childhood cancer survivors treated with chest-directed radiotherapy have substantially elevated risk for developing breast cancer. Although genetic susceptibility to breast cancer in the general population is well studied, large-scale evaluation of breast cancer susceptibility after chest-directed radiotherapy for childhood cancer is lacking.
Methods: We conducted a genome-wide association study of breast cancer in female survivors of childhood cancer, pooling two cohorts with detailed treatment data and systematic, long-term follow-up: the Childhood Cancer Survivor Study and St. Jude Lifetime Cohort. The study population comprised 207 survivors who developed breast cancer and 2774 who had not developed any subsequent neoplasm as of last follow-up. Genotyping and subsequent imputation yielded 16 958 466 high-quality variants for analysis. We tested associations in the overall population and in subgroups stratified by receipt of lower than 10 and 10 or higher gray breast radiation exposure. We report *P* values and pooled per-allele risk estimates from Cox proportional hazards regression models. All statistical tests were two-sided.
Results: Among survivors who received 10 or higher gray breast radiation exposure, a locus on 1q41 was associated with subsequent breast cancer risk (rs4342822, nearest gene *PROX1*, risk allele frequency in control subjects [$RAF_{controls}$] = 0.46, hazard ratio = 1.92, 95% confidence interval = 1.49 to 2.44, $P = 7.09 \times 10^{-9}$). Two rare variants also showed potentially promising associations (breast radiation ≥10 gray: rs74949440, 11q23, *TAGLN*, $RAF_{controls} = 0.02$, $P = 5.84 \times 10^{-8}$; <10 gray:

Received: November 7, 2016; **Revised:** January 30, 2017; **Accepted:** March 8, 2017

rs17020562, 1q32.3, *RPS6KC1*, $RAF_{controls} = 0.0005$, $P = 6.68 \times 10^{-8}$). Associations were restricted to these dose subgroups, with consistent findings in the two survivor cohorts.

Conclusions: Our study provides strong evidence that germline genetics outside high-risk syndromes could modify the effect of radiation exposure on breast cancer risk after childhood cancer.

The occurrence of subsequent malignancies is a major cause of morbidity and mortality in childhood cancer survivors (1–4). Risk of breast cancer, one of the most common subsequent malignancies after childhood cancer, is highly dependent on the radiation dose received and volume of breast tissue exposed during radiotherapy (5–7). Approximately 30% of childhood cancer survivors who received 10 or higher gray (Gy) chest radiation develop breast cancer by age 50 years (6). Because of this higher-than-20-fold increased risk, females who received chest radiotherapy prior to age 30 years are recommended to undergo annual breast cancer screening beginning eight years following treatment or at age 25 years, whichever occurs later (8).

Despite an increasing understanding of the heritability of response to ionizing radiation exposure (9,10), as well as discovery of numerous breast cancer susceptibility loci in the general population, genetic predisposition for breast cancer after radiotherapy for childhood cancer remains poorly understood. We therefore initiated a genome-wide association study (GWAS) of subsequent breast cancer following childhood cancer within two cohorts of childhood cancer survivors with detailed treatment data, systematic long-term follow-up, and available DNA: the Childhood Cancer Survivor Study (CCSS) (11) and the St. Jude Lifetime Cohort (SJLIFE) (12).

Methods

Study Population and Phenotype Data

The study population and phenotype data are described in detail in the Supplementary Methods (available online). Briefly, CCSS is a multicenter retrospective cohort with prospective follow-up of individuals who survived five or more years following diagnosis with first primary childhood cancer during 1970 to 1986 (11). Patients eligible for CCSS were diagnosed with the most common forms of childhood cancer before age 21 years at one of 26 participating centers in the United States and Canada. SJLIFE is a clinically assessed cohort of survivors who received treatment for any type of childhood cancer at St. Jude Children's Research Hospital (SJCRH; Memphis, TN) during 1962 to 2005, survived 10 or more years following diagnosis, and were age 18 years or older (12). SJLIFE participants who had already been genotyped as part of the CCSS effort were analyzed in the CCSS cohort. The cohorts were approved by the institutional review boards at each participating center, and participants provided informed consent. The GWAS and pooled analyses were approved by the Institutional Review Board of the National Cancer Institute (Bethesda, MD).

Occurrences of subsequent malignancies were ascertained through self- or proxy-report by questionnaire or death certificate, or by clinical screening (SJLIFE only), and confirmed by review of pathology reports and medical records. Treatment data within the first five years following first primary childhood cancer diagnosis were abstracted from medical records. Radiotherapy data included information on dates of therapy; beam energy; field location and size; use of radiation field blocks to protect normal tissue; and prescribed dose. Radiation exposure was estimated as the maximum treatment dose to the chest, accounting only for direct in-beam contributions (13). Radiation dose reconstruction details and distributions of the estimated radiation doses to the breast are presented in the Supplementary Methods and Supplementary Figure 1 (available online), respectively.

Genotyping, Imputation, and Quality Control

In both cohorts, genotype data for this analysis were derived from a larger effort to genotype all cohort participants with available DNA regardless of sex or ancestry (Supplementary Methods and Supplementary Figures 2–5, available online).

For CCSS, DNA was extracted using standard methods from blood, saliva (Oragene), or buccal cells (collected using mouthwash). For samples with insufficient DNA, whole-genome amplification (WGA) was performed (14). Genotyping of study samples and quality control replicates was conducted at the Cancer Genomics Research Laboratory of the National Cancer Institute on the Illumina (San Diego, CA) HumanOmni5Exome array. We estimated ancestry using the Genotyping Library and Utilities (GLU) struct.admix module with HapMap data as the fixed reference population. We then performed imputation based on the 1000 Genomes Project release version 3 reference haplotypes using IMPUTE version 2.3.0, resulting in a total of 26 135 905 high-quality single nucleotide polymorphisms (SNPs) and small insertions or deletions (InDels).

For SJLIFE participants, a blood sample was collected at the initial SJLIFE clinical evaluation, and DNA was isolated using the DNAeasy Blood and Tissue Kit from Qiagen (Hilden, Germany). Genotyping was performed at the SJCRH Hartwell Center for Bioinformatics and Biotechnology using the Affymetrix (Santa Clara, CA) Genome-Wide Human SNP Array 6.0. Ancestry for SJLIFE participants was determined using the program STRUCTURE. We then performed imputation using Minimac with data from the 1000 Genomes Project used as the reference panel (RELEASE STAMP 2012-10-09), resulting in 23 675 718 high-quality variants.

We merged the CCSS and SJLIFE data by genome position and alleles. The final pooled analytic data set included 16 958 466 variants that were present in both cohorts and passed all cohort-specific quality control thresholds.

Statistical Analyses

We performed association analyses in the overall population as well as restricted to populations receiving 10 or higher or lower than 10 Gy radiation exposure to the breast. Details of the analyses are provided in the Supplementary Methods (available online). Briefly, within each population, we first tested for independence between genotype and breast cancer occurrence for all variants using the Mantel-Haenszel (MH) test statistic, counting the number of chromosomes by presence/absence of the minor allele and breast cancer status within three strata defined by cohort and DNA type (SJLIFE, CCSS-gDNA, CCSS-wgaDNA).

For those SNPs with P_{MH} values of less than 1×10^{-6}, we further assessed the association between genotype and breast cancer occurrence using multivariable Cox regression, assuming an

ARTICLE

additive genetic effect (allelic dosage). Age was the underlying time scale, with individuals followed from the age of first primary childhood cancer until the earliest of breast cancer diagnosis, death, or last follow-up. Covariates included cohort, receipt of any alkylating agent- or platinum-containing chemotherapy, and associated principal components. The analysis of the overall population also adjusted for receipt of 10 or higher Gy radiation exposure to the breast. For the top SNP associations (pooled $P_{MH} < 1 \times 10^{-7}$), we obtained a permutation-based *P* value (P_{COX}), based on the Wald test statistic (P_{WALD}), permuting on case status. *P* values of less than $5x10^{-8}$ were considered statistically significant at the genome-wide level. We also tested for departure from a multiplicative effect of genotype (allelic dosage) and radiation (≥10 vs < 10 Gy exposure to the breast) on breast cancer risk by comparing the model fit with and without a genotype×radiation parameter (P_{GR}).

Statistical analyses and data management were conducted using R (version 3.2.3, R Foundation for Statistical Computing, Vienna, Austria) and SAS (version 9.4, SAS Institute Inc., Cary, NC). All statistical tests were two-sided.

Results

The primary analytic population comprised females of European descent who developed breast cancer ($N_{CCSS} = 178$, $N_{SJLIFE} = 29$) or who did not develop a subsequent neoplasm as of the date of death or last follow-up ("control subjects"; $N_{CCSS} = 2200$, $N_{SJLIFE} = 574$) (Table 1; Supplementary Table 1, available online). Among the 207 breast cancer case patients, 73.9% were invasive and 26.1% in situ, median age at first primary cancer (most commonly Hodgkin lymphoma, 64.7%) was 15.6 years (interquartile range [IQR] = 13.5–18.0 years), and median age at breast cancer was 39.2 years (IQR = 35.1–43.2 years). Sixty-three percent of case patients received 10 or higher Gy radiation exposure to the breast. Among the 2774 control subjects, median age at first primary cancer (most commonly leukemia, 34.7%) was 6.1 years (IQR = 2.9–12.7 years), and median age at last follow-up was 32.9 years (IQR = 27.7–39.3 years). Eighteen percent of control subjects received 10 or higher Gy radiation exposure to the breast. Distributions by first primary type and age at first primary cancer differed when we stratified the cohort by 10 or higher and lower than 10 Gy radiation exposure to the breast (Table 1).

Among childhood cancer survivors who received 10 or higher Gy radiation to the breast, the locus at 1q41 marked by rs4342822 was statistically significantly associated with risk of developing subsequent breast cancer (risk allele frequency [RAF] = 0.46 in control subjects, $P_{MH} = 7.09 \times 10^{-9}$) (Table 2; Supplementary Figures 6–7, available online). In Cox regression analyses, rs4342822 was associated with nearly twofold increased risk of breast cancer per G allele (hazard ratio [HR] = 1.92, 95% confidence interval [CI] = 1.49 to 2.44, $P_{COX} = 7.00 \times 10^{-9}$). Results were consistent in the two cohorts. In contrast, no association with this locus was observed among survivors who received lower than 10 Gy breast radiation ($P_{MH} = .80$, HR = 1.04, 95% CI = 0.75 to 1.45, $P_{GR} = .006$) (Supplementary Table 2, available online). In more detailed analyses by radiation exposure level, the presence of the G allele for rs4342822 appeared increasingly detrimental with increasing radiation dose (Figure 2). Similarly, the effect of radiation exposure was increasingly detrimental with an increasing number of G alleles (Supplementary Tables 3 and 4, available online).

In addition, we identified potentially promising associations ($P_{MH} < 1 \times 10^{-7}$) for two rare variants. A locus at 11q23 (rs74949440) was associated with the risk of developing subsequent breast cancer among childhood cancer survivors who received 10 or higher Gy radiation to the breast (RAF = 0.02 in control subjects, $P_{MH} = 5.84 \times 10^{-8}$, HR = 2.59, 95% CI = 1.62 to 4.16, $P_{COX} = 8.40 \times 10^{-8}$, $P_{GR} = .06$), whereas a locus at 1q32.3 (rs17020562) was associated among those who received lower than 10 Gy radiation to the breast (RAF = 0.0005 in control subjects, $P_{MH} = 6.68 \times 10^{-8}$, HR = 44.52, 95% CI = 15.06 to 131.62, $P_{COX} = 5.00 \times 10^{-9}$, $P_{GR} = .002$) (Table 2 and Figure 1; Supplementary Figure 7 and Supplementary Tables 2–4, available online).

Because of the strong associations among treatment, age at exposure, and type of initial childhood cancer (Supplementary Figure 8, available online), we conducted a series of sensitivity analyses for the top three variants in subgroups of survivors defined by these characteristics (Supplementary Table 5, available online). Results appeared consistent regardless of first primary type, age at first primary childhood cancer, and whether ovarian function had been compromised by receipt of pelvic radiation or alkylating agent-containing chemotherapy. Moreover, results were similar when we restricted the case patients to those for whom breast cancer was their first subsequent neoplasm (ie, excluding case patients who developed another intervening neoplasm before breast cancer) and when we included as "control subjects" individuals who developed subsequent neoplasms other than breast cancer. Intriguingly, the associations for all three variants appeared to be specific to the development of subsequent breast cancer because no evidence of risk was observed for the most commonly occurring other types of subsequent neoplasms (basal cell carcinoma of the skin, meningioma, and thyroid cancer), even after accounting for radiation exposure to the tumor locations (Supplementary Table 6, available online).

We also identified seven further regions with promising associations for breast cancer risk after childhood cancer ($P_{MH} < 1 \times 10^{-6}$) (Supplementary Table 7, available online). A locus at 7q36.3 (rs117114682) showed potential evidence for association among survivors who received 10 or higher Gy radiation to the breast. Loci at 2q14.3 (rs1519277) and 17q24.3 (rs11651604) showed potential evidence for association among survivors who received lower than 10 Gy radiation to the breast, whereas loci at 2q37.3 (rs114971217), 4q32.2 (rs139948181), 11q13.2 (rs4930561), and 22q12.1 (rs147512482) showed potential evidence for association in the overall population.

A previous GWAS identified variants at 6q21 to be associated with second solid malignancies (>75% breast cancers) after radiotherapy for Hodgkin lymphoma, with the strongest association for rs4946728 (15). Our study included an overlapping set of individuals from CCSS with that case-control study (60 cases, 84 controls). After removal of the overlapping individuals, our results did not support the association between rs4946728 and breast cancer in Hodgkin lymphoma survivors receiving 10 or higher Gy radiation to the breast (CCSS nonoverlapping individuals: HR = 1.07, 95% CI = 0.55 to 2.07, $P_{MH} = .27$; SJLIFE: HR = 0.89, 95% CI = 0.43 to 1.85, $P_{MH} = .52$) (Supplementary Table 8, available online).

Among 122 common breast cancer susceptibility variants reported previously with *P* values of less than 5×10^{-8}, primarily in analyses of sporadic breast cancer in the general population, seven variants from six independent regions ($R^2 < 0.8$) achieved a P_{MH} value of less than .01 in any of the three populations in our study (overall, <10 Gy or ≥10 Gy to the breast) (Supplementary Table 9 and Supplementary Figure 9, available online). These loci included 1q21.1 (*NBPF10, RNF115*), 5p12 (*FGF10*, *MRPS30*), 5q11.2 (*MAP3K1*), 8q24.21 (*POU5F1B*, *CASC8*,

Table 1. Selected characteristics of childhood cancer survivors of European descent included in the GWAS of subsequent breast cancer after childhood cancer, overall and by radiation exposure to the breast

Characteristics	Subsequent breast cancer No. (%)	No subsequent neoplasm No.(%)	By radiation exposure to the breast			
			≥10 Gy		<10 Gy	
			Subsequent breast cancer No.(%)	No subsequent neoplasm No.(%)	Subsequent breast cancer No.(%)	No subsequent neoplasm No.(%)
Total	207 (100.0)	2774 (100.0)	131 (100.0)	493 (100.0)	69 (100.0)	2144 (100.0)
Cohort						
CCSS	178 (86.0)	2200 (79.3)	108 (82.4)	372 (75.5)	63 (91.3)	1691 (78.9)
SJLIFE	29 (14.0)	574 (20.7)	23 (17.6)	121 (24.5)	6 (8.7)	453 (21.1)
Primary cancer						
Retinoblastoma	0 (0.0)	25 (0.9)	0 (0.0)	0 (0.0)	0 (0.0)	25 (1.2)
Leukemia	20 (9.7)	962 (34.7)	3 (2.3)	82 (16.6)	17 (24.6)	827 (38.6)
Hodgkin lymphoma	134 (64.7)	246 (8.9)	106 (80.9)	176 (35.7)	24 (34.8)	60 (2.8)
Non-Hodgkin lymphoma	10 (4.8)	132 (4.8)	6 (4.6)	28 (5.7)	3 (4.3)	97 (4.5)
Neuroblastoma	1 (0.5)	231 (8.3)	0 (0.0)	30 (6.1)	1 (1.4)	188 (8.8)
Soft tissue sarcoma	11 (5.3)	245 (8.8)	4 (3.1)	20 (4.1)	6 (8.7)	212 (9.9)
CNS	3 (1.4)	346 (12.5)	0 (0.0)	71 (14.4)	3 (4.3)	256 (11.9)
Bone	23 (11.1)	224 (8.1)	8 (6.1)	21 (4.3)	14 (20.3)	195 (9.1)
Kidney (Wilms)	4 (1.9)	314 (11.3)	4 (3.1)	60 (12.2)	0 (0.0)	240 (11.2)
Other	1 (0.5)	49 (1.8)	0 (0.0)	5 (1.0)	1 (1.4)	44 (2.1)
Year of primary cancer						
1962–1975	77 (37.2)	480 (17.3)	53 (40.5)	110 (22.3)	24 (34.8)	342 (16.0)
1976–1980	71 (34.3)	693 (25.0)	35 (26.7)	101 (20.5)	30 (43.5)	552 (25.7)
1981–2005	59 (28.5)	1601 (57.7)	43 (32.8)	282 (57.2)	15 (21.7)	1250 (58.3)
Age at primary cancer, y						
0–4	5 (2.4)	1189 (42.9)	3 (2.3)	119 (24.1)	1 (1.4)	1011 (47.2)
5–9	11 (5.3)	628 (22.6)	4 (3.1)	100 (20.3)	7 (10.1)	492 (22.9)
10–14	71 (34.3)	551 (19.9)	44 (33.6)	132 (26.8)	26 (37.7)	397 (18.5)
15–23	120 (58.0)	406 (14.6)	80 (61.1)	142 (28.8)	35 (50.7)	244 (11.4)
Any radiotherapy for primary cancer*						
No	25 (12.1)	1149 (41.4)	0 (0.0)	0 (0.0)	25 (36.2)	1149 (53.6)
Yes	179 (86.5)	1528 (55.1)	131 (100.0)	493 (100.0)	44 (63.8)	995 (46.4)
Unknown	3 (1.4)	97 (3.5)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
AA/PT-containing chemotherapy for primary cancer						
No/unknown	96 (46.4)	1467 (52.9)	64 (48.9)	201 (40.8)	29 (42.0)	1153 (53.8)
Yes	111 (53.6)	1307 (47.1)	67 (51.1)	292 (59.2)	40 (58.0)	991 (46.2)
Age at breast cancer (cases)/last follow-up (controls), y						
11–29	9 (4.3)	982 (35.4)	5 (3.8)	111 (22.5)	4 (5.8)	838 (39.1)
30–39	103 (49.8)	1155 (41.6)	64 (48.9)	195 (39.6)	34 (49.3)	886 (41.3)
40–61	95 (45.9)	637 (23.0)	62 (47.3)	187 (37.9)	31 (44.9)	420 (19.6)
Breast cancer location†						
Upper outer quadrant	73 (35.3)	—	58 (44.3)	—	14 (20.3)	—
Lower outer quadrant	6 (2.9)	—	3 (2.3)	—	3 (4.3)	—
Upper inner quadrant	19 (9.2)	—	4 (3.1)	—	15 (21.7)	—
Lower inner quadrant	10 (4.8)	—	1 (0.8)	—	8 (11.6)	—
Central	7 (3.4)	—	6 (4.6)	—	1 (1.4)	—
Overlapping	39 (18.8)	—	23 (17.6)	—	16 (23.2)	—
Unknown	53 (25.6)	—	36 (27.5)	—	12 (17.4)	—
Breast cancer type						
Invasive	153 (73.9)	—	96 (73.3)	—	53 (76.8)	—
In situ	54 (26.1)	—	35 (26.7)	—	16 (23.2)	—

— = Breast cancer data are not applicable to individuals with no subsequent neoplasm; AA/PT = alkylating agent or platinum; CNS = central nervous system; CCSS = Childhood Cancer Survivor Study; GWAS = genome-wide association study; Gy = gray; NPS = not otherwise specified; SJLIFE = St. Jude Lifetime Cohort Study.

*For cases, represents 10 or higher Gy to the tumor location (categories as listed in the table). For controls, represents 10 or higher Gy to any location in the breast. Radiotherapy was unknown for three cases and 97 controls. Radiotherapy was received, but dose to the breast was unknown for four cases and 40 controls from CCSS. Lower than 10 Gy to the breast includes no radiotherapy ($N_{cases} = 25$, $N_{controls} = 1149$), greater than 0 to less than 1 Gy ($N_{cases} = 16$, $N_{controls} = 681$), and 1 to 9.9 Gy ($N_{cases} = 28$, $N_{controls} = 314$).

†Represents first breast cancer diagnosed. Overlapping sites included upper inner/outer quadrants (n = 15), lower inner/outer quadrants (n = 5), outer upper/lower quadrants (n = 9), inner upper/lower quadrants (n = 5), lower outer quadrant/central (n = 1), and three or more quadrants (n = 4). Second breast cancers were diagnosed in 42 and five women in CCSS and SJLIFE, respectively.

Table 2. Top SNP associations (pooled $P_{MH} < 1 \times 10^{-7}$) in the GWAS of subsequent breast cancer after childhood cancer in the pooled population and by cohort

Location, nearest gene, and SNP information*	Population†	Cohort	Genotyping status*	Controls/cases	RAF		P_{MH}‡	HR (95% CI)‡	P_{COX}‡
					Controls	Cases			
1q41, *PROX1*	≥10 Gy	Pooled	N/A	493/131	0.46	0.66	7.09×10^{-9}	1.92 (1.49 to 2.44)	7.00×10^{-9}
rs4342822 (214024225)		CCSS	Imputed (0.996)	372/108	0.47	0.66	2.23×10^{-7}	1.89 (1.31 to 2.50)	1.38×10^{-7}
Risk allele = G, referent allele = T		SJLIFE	Genotyped	121/23	0.45	0.65	.02	2.33 (1.27 to 4.35)	.003
11q23, *TAGLN*	≥10 Gy	Pooled	N/A	493/131	0.02	0.09	5.84×10^{-8}	2.59 (1.62 to 4.16)	8.40×10^{-8}
rs74949440 (117070361)		CCSS	Imputed (0.905)	372/108	0.02	0.08	1.75×10^{-5}	2.47 (1.45 to 4.23)	1.24×10^{-4}
Risk allele = T, referent allele = C		SJLIFE	Imputed (0.674)	121/23	0.01	0.13	7.09×10^{-4}	3.14 (0.97 to 10.19)	.001
1q32.3, *RPS6KC1*	<10 Gy	Pooled	N/A	2144/69	0.0005	0.04	6.68×10^{-8}	44.52 (15.06 to 131.62)	5.00×10^{-9}
rs17020562 (213542706)		CCSS	Imputed (0.988)	1691/63	0.00	0.02	4.36×10^{-5}	—	
Risk allele = C, referent allele = T		SJLIFE	Genotyped	453/6	0.002	0.17	9.27×10^{-4}	41.01 (6.59 to 255.10)	7.44×10^{-5}

— = Cox model results are not shown if <2 cases or controls have the variant allele; CI = confidence interval; HR = hazard ratio; MH = Mantel-Haenszel; NA = not applicable; RAF = risk allele frequency.

*Position (in parentheses after rs number) according to Genome Reference Consortium Human Build 37 (hg19). For imputed variants, number in parentheses after genotyping status represents the quality score (CCSS = IMPUTE2 info_score; SJLIFE = Minimac R^2).

†We performed association analyses in the overall population as well as restricted to survivors receiving 10 or higher or lower than 10 Gy radiation exposure to the breast. Results for all three populations for these SNPs are presented in Supplementary Table 2 (available online).

‡P_{MH} represents the two-sided *P* value for the Mantel-Haenszel test statistic, calculated from the exact conditional distribution. Hazard ratio (95% confidence interval) and two-sided permutation-based P_{COX} calculated using multivariable Cox regression. Adjustment variables in the Cox model included cohort, receipt of any alkylating agent- or platinum-containing chemotherapy, and associated principal components to adjust for potential population stratification.

CASC21), and 10q26.13 (*FGFR2*). Although a previous study reported an association between rs1219648 (*FGFR2*) and breast cancer after Hodgkin lymphoma (16), we found no statistically significant difference in the risk estimates for this SNP in subgroups stratified by radiation exposure to the breast ($P_{GR} = .44$).

Discussion

We conducted the first large-scale GWAS of subsequent breast cancer risk among childhood cancer survivors by combining two cohorts with detailed treatment data, systematic second cancer ascertainment, and long-term follow-up. We found a locus at 1q41 to be associated with subsequent breast cancer risk only among survivors who received 10 or higher Gy breast radiation exposure. Two additional rare variants also showed potentially promising associations with subsequent breast cancer risk, with the locus at 11q23 most evident among survivors who received 10 or higher Gy breast radiation exposure and the locus at 1q32.3 most evident among those with lower than 10 Gy breast radiation exposure.

Each of these three loci map to the vicinity of biologically plausible candidate genes for breast cancer (Figure 2; Supplementary Tables 9 and 10, available online). The variant rs4342822 maps near *PROX1* (prospero homeobox 1), a transcription factor involved in early embryonic development and implicated in cellular proliferation and migration (17–19). Previous studies have reported altered *PROX1* expression in breast tumors because of DNA hypermethylation (20), and the Cancer Genome Atlas (TCGA) data demonstrate a high frequency of *PROX1* alterations (13%), particularly amplifications, in breast cancers (21). Additionally, several correlated variants ($R^2 > 0.4$, $D' > 0.9$) have regulatory potential, binding transcription factors such as *CTCF* and *ETS1*, both of which have been implicated in breast carcinogenesis (22,23). The variant rs74949440 lies intronic to *TAGLN* (transgelin), an actin binding protein involved in cellular migration (24) near a region of open chromatin (25). *TAGLN* overexpression has been observed in triple-negative breast cancer (26) and promotes cellular transformation and proliferation through its role in the inactivation of p53 via metalloprotein isozymes, which has been shown to increase tumor cell tolerance to chemotherapy and gamma radiation (27–29). The association of these variants with breast cancer risk only after radiation exposure raises the hypothesis that these germline genetic variants could create a pro-proliferative, pro-invasive phenotype that supports the growth of malignant cells following transformation by ionizing radiation. In contrast, the variant rs17020562, which maps near *RPS6KC1* (ribosomal protein S6 kinase, 52kDa, polypeptide 1), showed the strongest association among survivors who received lower than 10 Gy radiation to the breast. TCGA data demonstrate that *RPS6KC1* is amplified or mutated in approximately 13% of breast cancers (21), and *RPS6KC1* has been proposed as an oncogene in endometrial cancer, possibly because of its role in endosomal trafficking (30). Clarification of the mechanisms of action, accounting for varying levels of radiation exposure, will require laboratory follow-up.

We also identified seven further regions with promising associations for breast cancer risk after childhood cancer. Of these regions, only the association for the locus at 7q36.3 (rs117114682) was associated with breast cancer risk among survivors who received 10 or higher Gy radiation to the breast. This association is particularly intriguing because rs117114682 is intronic to *PTPRN2* (protein tyrosine phosphatase, receptor type N2), which has been associated with breast cancer metastasis via actin-remodeling-dependent migration (31,32), further supporting our hypothesis that germline genetic variants contribute to a favorable growth environment for transformed cells in the breast. Additional data will be needed to clarify the possible role of these promising loci in the development of breast cancer after childhood cancer.

A previous GWAS identified rs4946728 at 6q21 in association with second solid malignancies after radiotherapy for Hodgkin lymphoma (15). The discovery phase of that study utilized CCSS, selecting as case patients Hodgkin lymphoma survivors

Dose to the breast (Gy)	<1 *	1-9	10-29	30-39	≥40
Controls, N	1830	314	292	105	96
Cases, N	41	28	42	40	49

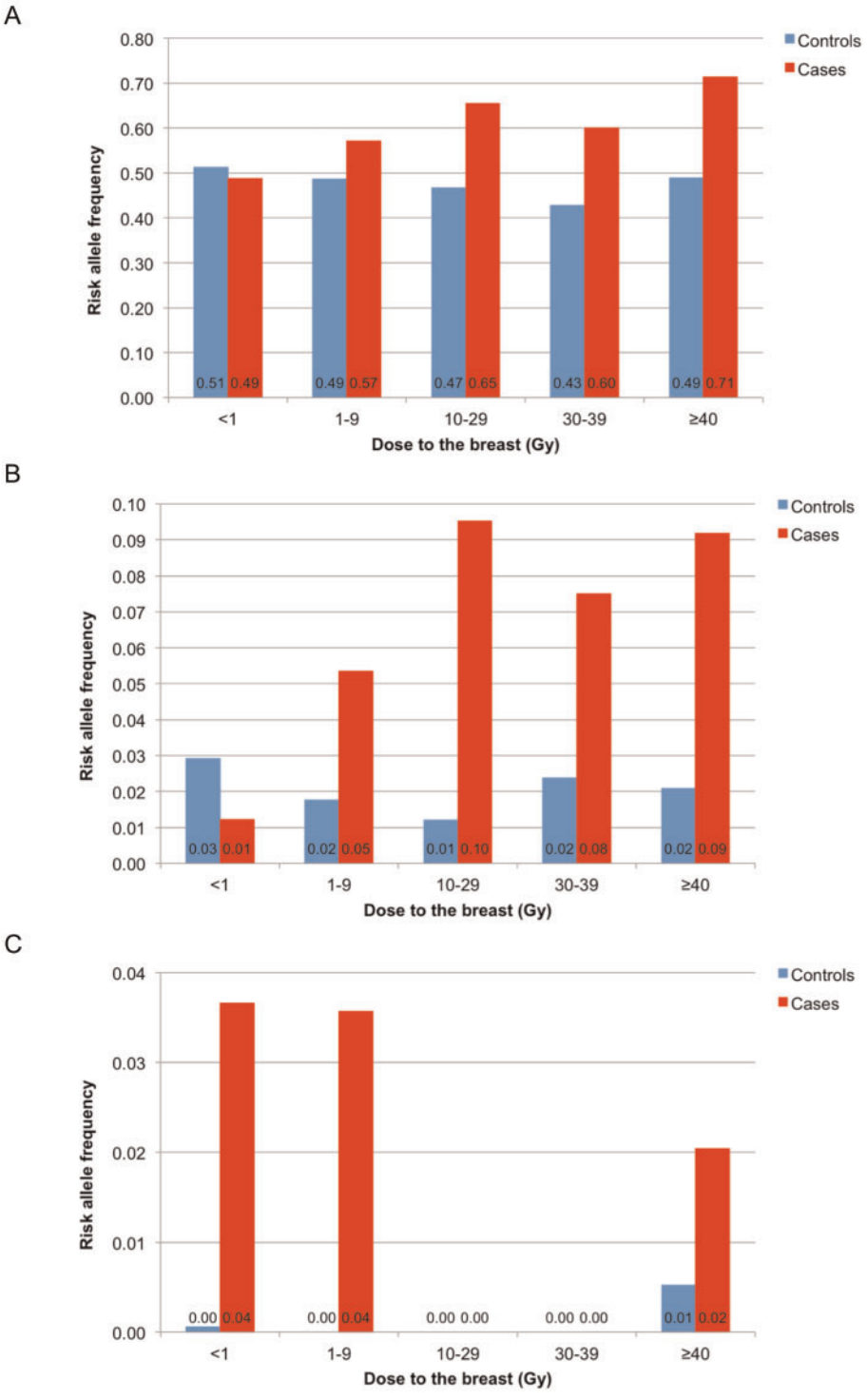

Figure 1. Risk allele frequency (RAF) by radiation dose to the breast for rs4342822 **(A)**, rs74949440 **(B)**, and rs17020562 **(C)** in the genome-wide association study of subsequent breast cancer after childhood cancer. Radiation-related risks by genotype and the joint effects of genotype and radiation exposure to the breast are presented in Supplementary Tables 3 and 4 (available online). *Lower than 1 Gy to the breast includes no radiotherapy ($N_{cases} = 25$, $N_{controls} = 1149$) and radiotherapy with estimated dose to the breast of more than 0 to less than 1 Gy ($N_{cases} = 16$, $N_{controls} = 681$). GWAS = genome-wide association study; Gy = gray; RAF = risk allele frequency.

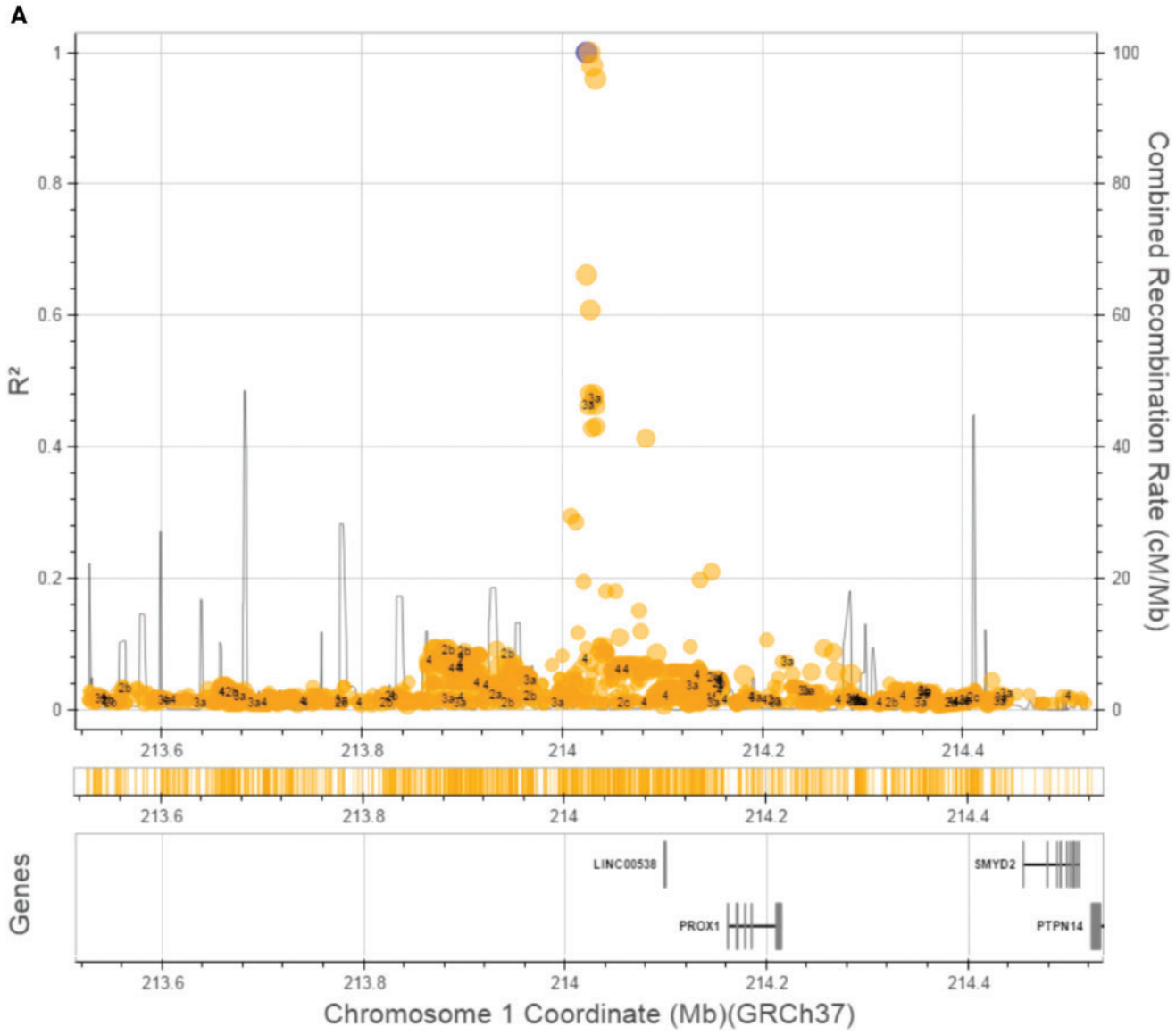

Figure 2. Genomic mapping of variants (-/+ 500 kilobases) correlated with rs4342822 **(A)**, rs74949440 **(B)**, and rs17020562 **(C)** using LDLink (35). The **darker circle** represents the variant of interest, **remaining circles** represent correlated variants according to data from the 1000 Genomes Project CEU population (Northern Europeans from Utah), and the **number inside the circle** represents the regulatory potential (1 = highest, 7 = lowest; numbers are only shown for variants with scores of 1–4) (35,36). See Supplementary Table 10 (available online) for further detail, including the score for regulatory potential and associations for correlated variants within the pooled study population.

who developed a second solid malignancy after receiving radiotherapy to the body region where that tumor developed (n = 96, of whom 59 [61.5%] developed breast cancer) and as control subjects Hodgkin lymphoma survivors who were followed for 27 or more years (median = 32 years, range = 27–38 years) without developing a second cancer (n = 82). Results from our pooled study population restricted to the cohort of Hodgkin lymphoma survivors who received 10 or higher Gy radiation exposure to the breast provided only suggestive evidence for the association between rs4946728 and breast cancer, but no evidence once overlapping case patients were excluded. The difference in results can be primarily attributed to the use of selected CCSS case patients and control subjects in Best et al. (15) vs the cohort of individuals in CCSS with available DNA in the current study, as well as a lack of association in SJLIFE (Supplementary Table 8, available online). However, small sample sizes in both our study and the previous study could also contribute to the differences, underscoring the importance of additional data sets for replication.

Our study was not designed to evaluate variants previously reported to be associated with breast cancer in the general population because, given our sample size, we did not have the statistical power to detect the small effect sizes typically observed for such variants (per-allele risk ratio < 1.5). Moreover, the etiology of breast cancers arising after childhood cancer might differ from the etiology of breast cancers arising in case patients of previous GWAS, which primarily included women sampled from the general population who tended to be diagnosed at older ages and would rarely have received chest radiotherapy. Nevertheless, for completeness, we evaluated these variants in our study population and did not find a strong relationship.

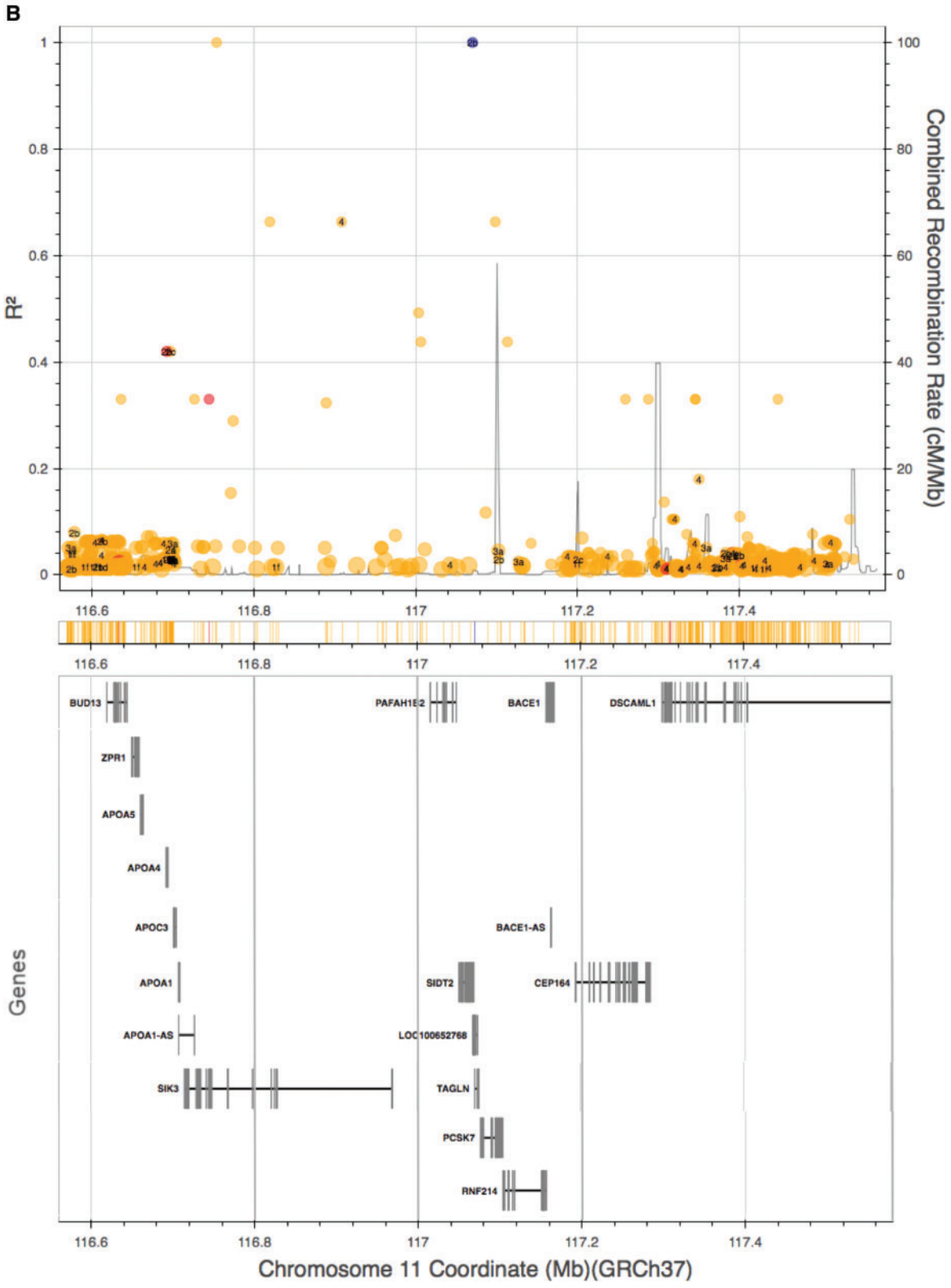

Figure 2. Continued

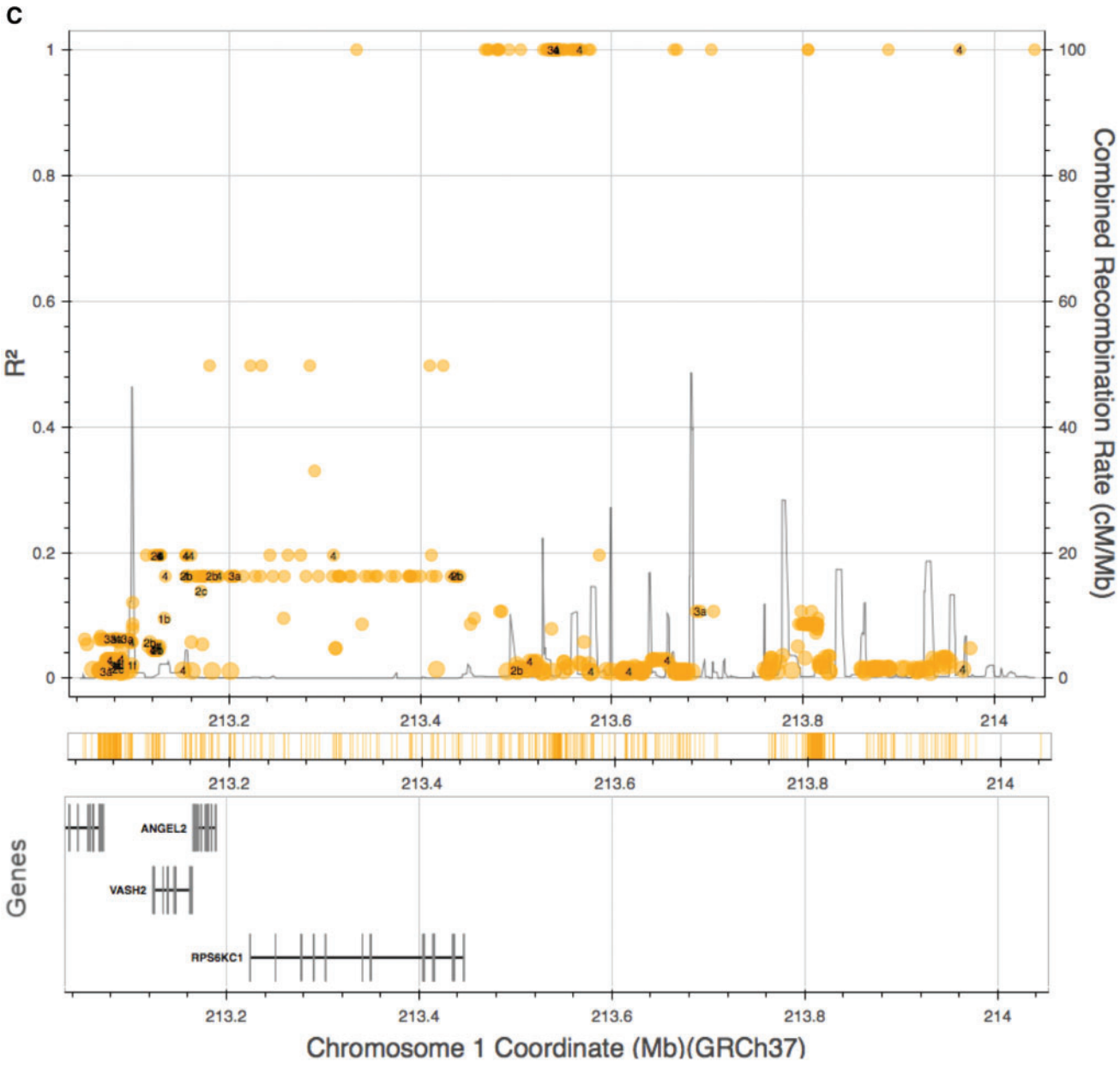

Figure 2. Continued

Before the findings from this study can be translated to the clinical setting, the results require further replication in independent populations, the magnitude of the risks need to be more precisely quantified, and the functional significance of the variants in the context of specific radiation exposures needs to be better understood. Our study was limited because of small sample sizes in some subgroups and because of the potential for misclassification of radiation exposure due to lack of information on 1) the specific breast tumor location, 2) precise location of radiotherapy treatment field blocks (used to protect normal tissue), and/or 3) treatments received five or more years following first primary childhood cancer diagnosis. Caution is particularly warranted in interpreting the results for rare variants, particularly for the radiation subgroups because of small numbers, potential genotype misclassification from imputation, and, for rs17020562, because of potential residual confounding by population substructure because the RAF varies substantially by ancestry. We could not fully disentangle the complex relations of genetic susceptibility, treatment, age at exposure, attained age, and type of initial childhood cancer with subsequent breast cancer risk. Further research is needed to evaluate whether our findings generalize to adults treated with radiotherapy to the chest (ie, after breast development associated with puberty, pregnancy, and lactation), and also whether our findings persist among childhood cancer survivors treated with current approaches (33). Finally, we could not determine breast tumor receptor status and molecular subtype, which have been suggested to have different genetic etiologies (34).

In summary, this discovery study presents strong evidence that germline variants beyond identified high-risk cancer

susceptibility genes may interact with radiation exposure to modify risk for breast cancer after childhood cancer. With further replication of our results in additional patient cohorts, knowledge of such germline variants could ultimately influence clinical practice for choosing frontline therapy and/or post-treatment surveillance for breast cancer.

Funding

The Childhood Cancer Survivor Study (CCSS) genome-wide association study and pooled analyses were supported by the Intramural Research Program of the National Cancer Institute, National Institutes of Health. This work utilized the computational resources of the National Institutes of Health High-Performance Computing Biowulf cluster (http://hpc.nih.gov).

CCSS is supported by the National Cancer Institute (CA55727: GTA, Principal Investigator). A portion of the CCSS genotyping also was supported by the Leukemia and Lymphoma Society (K. Kamdar, Principal Investigator). Review and confirmation of CCSS breast cancer diagnoses were supported in part by the National Cancer Institute (R01CA136783: CSM, Principal Investigator; R01CA134722: KCO, Principal Investigator). The St. Jude Lifetime Cohort (SJLIFE) is supported by the National Cancer Institute (U01 CA195547: MMH, Principal Investigator; Cancer Center Support CORE grant CA21765: C. Roberts, Principal Investigator) and the American Lebanese Syrian Associated Charities, Memphis, Tennessee.

Notes

Authors: Lindsay M. Morton*, Joshua N. Sampson*, Gregory T. Armstrong*, Ting-Huei Chen, Melissa M. Hudson, Eric Karlins, Casey L. Dagnall, Shengchao Alfred Li, Carmen L. Wilson, Deo Kumar Srivastava, Wei Liu, Guolian Kang, Kevin C. Oeffinger, Tara O. Henderson, Chaya S. Moskowitz, Todd M. Gibson, Diana M. Merino, Jeannette R. Wong, Sue Hammond, Joseph P. Neglia, Lucie M. Turcotte, Jeremy Miller, Laura Bowen, William A. Wheeler, Wendy M. Leisenring, John A. Whitton, Laurie Burdette, Charles Chung, Belynda D. Hicks, Kristine Jones, Mitchell J. Machiela, Aurelie Vogt, Zhaoming Wang, Meredith Yeager, Geoffrey Neale, Matthew Lear, Louise C. Strong, Yutaka Yasui, Marilyn Stovall, Rita E. Weathers, Susan A. Smith, Rebecca Howell, Stella M. Davies, Gretchen A. Radloff, Kenan Onel, Amy Berrington de González, Peter D. Inskip, Preetha Rajaraman, Joseph F. Fraumeni Jr., Smita Bhatia†, Stephen J. Chanock†, Margaret A. Tucker†, Leslie L. Robison†

*Authors jointly led this work.

†Authors jointly directed this work.

Affiliations of authors: Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services, Bethesda, MD (LMM, JNS, THC, EK, CLD, SAL, DMM, JRW, LBu, CC, BDH, KJ, MJM, AV, ZW, MY, ABdG, PDI, PR, JFFJr, SJC, MAT); Department of Epidemiology and Cancer Control (GTA, CLW, TMG, YY, LRL), Division of Cancer Survivorship, Department of Oncology (MMH), Department of Biostatistics (DKS, WL, GK), Hartwell Center for Bioinformatics and Biotechnology (GN), and Department of Pathology (ML), St. Jude Children's Research Hospital, Memphis, TN; Cancer Genomics Research Laboratory, Leidos Biomedical Research, Inc., Frederick National Laboratory for Cancer Research, Frederick, MD (EK, CLD, SAL, LBu, CC, BDH, KJ, AV, ZW, MY); Departments of Medicine (KCO), Pediatrics (KCO), and Epidemiology and Biostatistics (CSM), Memorial Sloan Kettering Cancer Center, New York, NY; Section of Hematology, Oncology and Stem Cell Transplantation, Department of Pediatrics, University of Chicago, Chicago, IL (TOH, KO); Nationwide Children's Hospital and the Ohio State University School of Medicine, Columbus, OH (SH); Department of Pediatrics, University of Minnesota, Minneapolis, MN (JPN, LMT); Information Management Services, Inc., Calverton, MD (JM, LBo, WAW); Cancer Prevention and Clinical Statistics Programs (WML) and Cancer Prevention Program (JAW), Fred Hutchinson Cancer Research Center, Seattle, WA; Department of Genetics (LCS) and Department of Radiation Physics (MS, REW, SAS, RH), The University of Texas at MD Anderson Cancer Center, Houston, TX; Department of Pediatrics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH (SMD, GAR); Institute for Cancer Outcomes and Survivorship, University of Alabama at Birmingham, Birmingham, AL (SB).

The funders had no role in the design of the study; the collection, analysis, or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication. Author contributions: LMM, JNS, GTA, SB, SJC, MAT, and LLR organized and designed the study and led the data interpretation and manuscript drafting. GTA, MMH, CLW, DKS, WL, GK, KCO, TOH, CSM, TMG, SH, JPN, LMT, WML, JAW, GN, ML, LCS, YY, MS, REW, SAS, RH, SMD, GAR, SB, and LLR conducted the CCSS and SJLIFE cohort studies, contributing the phenotype data, follow-up data, radiation dose reconstruction, and/or DNA. LMM, GTA, EK, CLD, SAL, CLW, LB, CC, BDH, KJ, AV, ZW, MY, GN, ML, SB, SJC, MAT, and LLR conducted and/or supervised the genotyping of samples. LMM, JNS, GTA, THC, DKS, WL, GK, JM, LB, WAW, WML, JAW, YY, SB, SJC, MAT, and LLR contributed to the design and execution of the statistical analysis. All authors contributed to the interpretation of the data and the writing of the manuscript.

References

1. Reulen RC, Frobisher C, Winter DL, et al. Long-term risks of subsequent primary neoplasms among survivors of childhood cancer. *JAMA*. 2011;305(22):2311–2319.
2. Olsen JH, Moller T, Anderson H, et al. Lifelong cancer incidence in 47,697 patients treated for childhood cancer in the Nordic countries. *J Natl Cancer Inst*. 2009;101(11):806–813.
3. Cardous-Ubbink MC, Heinen RC, Bakker PJ, et al. Risk of second malignancies in long-term survivors of childhood cancer. *Eur J Cancer*. 2007;43(2):351–362.
4. Friedman DL, Whitton J, Leisenring W, et al. Subsequent neoplasms in 5-year survivors of childhood cancer: The Childhood Cancer Survivor Study. *J Natl Cancer Inst*. 2010;102(14):1083–1095.
5. Henderson TO, Amsterdam A, Bhatia S, et al. Systematic review: Surveillance for breast cancer in women treated with chest radiation for childhood, adolescent, or young adult cancer. *Ann Intern Med*. 2010;152(7):444–455, W144–W154.
6. Moskowitz CS, Chou JF, Wolden SL, et al. Breast cancer after chest radiation therapy for childhood cancer. *J Clin Oncol*. 2014;32(21):2217–2223.
7. Inskip PD, Robison LL, Stovall M, et al. Radiation dose and breast cancer risk in the childhood cancer survivor study. *J Clin Oncol*. 2009;27(24):3901–3907.
8. Mulder RL, Kremer LC, Hudson MM, et al. Recommendations for breast cancer surveillance for female survivors of childhood, adolescent, and young adult cancer given chest radiation: A report from the International Late Effects of Childhood Cancer Guideline Harmonization Group. *Lancet Oncol*. 2013;14(13):e621–e629.
9. Barnett GC, West CM, Dunning AM, et al. Normal tissue reactions to radiotherapy: Towards tailoring treatment dose by genotype. *Nat Rev Cancer*. 2009;9(2):134–142.
10. Burnet NG, Barnett GC, Elliott RM, et al. RAPPER: The radiogenomics of radiation toxicity. *Clin Oncol*. 2013;25(7):431–434.
11. Robison LL, Armstrong GT, Boice JD, et al. The Childhood Cancer Survivor Study: A National Cancer Institute-supported resource for outcome and intervention research. *J Clin Oncol*. 2009;27(14):2308–2318.
12. Hudson MM, Ness KK, Nolan VG, et al. Prospective medical assessment of adults surviving childhood cancer: Study design, cohort characteristics, and feasibility of the St. Jude Lifetime Cohort study. *Pediatr Blood Cancer*. 2011;56(5):825–836.

ARTICLE

13. Stovall M, Weathers R, Kasper C, et al. Dose reconstruction for therapeutic and diagnostic radiation exposures: Use in epidemiological studies. *Radiat Res*. 2006;166(1 pt 2):141–157.
14. Silander K, Saarela J. Whole genome amplification with Phi29 DNA polymerase to enable genetic or genomic analysis of samples of low DNA yield. *Methods Mol Biol*. 2008;439:1–18.
15. Best T, Li D, Skol AD, et al. Variants at 6q21 implicate PRDM1 in the etiology of therapy-induced second malignancies after Hodgkin's lymphoma. *Nat Med*. 2011;17(8):941–943.
16. Ma YP, van Leeuwen FE, Cooke R, et al. FGFR2 genotype and risk of radiation-associated breast cancer in Hodgkin lymphoma. *Blood*. 2012;119(4):1029–1031.
17. Chang TM, Hung WC. Transcriptional repression of TWIST1 gene by Prospero-related homeobox 1 inhibits invasiveness of hepatocellular carcinoma cells. *FEBS Lett*. 2012;586(20):3746–3752.
18. Rodrigues MF, de Oliveira Rodini C, de Aquino Xavier FC, et al. PROX1 gene is differentially expressed in oral cancer and reduces cellular proliferation. *Medicine*. 2014;93(28):e192.
19. Elsir T, Smits A, Lindstrom MS, et al. Transcription factor PROX1: Its role in development and cancer. *Cancer Metastasis Rev*. 2012;31(3-4):793–805.
20. Versmold B, Felsberg J, Mikeska T, et al. Epigenetic silencing of the candidate tumor suppressor gene PROX1 in sporadic breast cancer. *Int J Cancer*. 2007; 121(3):547–554.
21. Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: An open platform for exploring multidimensional cancer genomics data. *Cancer Discov*. 2012;2(5):401–404.
22. Dittmer J. The role of the transcription factor Ets1 in carcinoma. *Semin Cancer Biol*. 2015;35:20–38.
23. Mustafa M, Lee JY, Kim MH. CTCF negatively regulates HOXA10 expression in breast cancer cells. *Biochem Biophys Res Commun*. 2015;467(4):828–834.
24. Assinder SJ, Stanton JA, Prasad PD. Transgelin: An actin-binding protein and tumour suppressor. *Int J Biochem Cell Biol*. 2009;41(3):482–486.
25. Shlyueva D, Stampfel G, Stark A. Transcriptional enhancers: From properties to genome-wide predictions. *Nat Rev Genet*. 2014;15(4):272–286.
26. Rao D, Kimler BF, Nothnick WB, et al. Transgelin: A potentially useful diagnostic marker differentially expressed in triple-negative and non-triple-negative breast cancers. *Hum Pathol*. 2015;46(6):876–883.
27. Meplan C, Richard MJ, Hainaut P. Metalloregulation of the tumor suppressor protein p53: Zinc mediates the renaturation of p53 after exposure to metal chelators in vitro and in intact cells. *Oncogene*. 2000;19(46):5227–5236.
28. Zhang ZW, Yang ZM, Zheng YC, et al. Transgelin induces apoptosis of human prostate LNCaP cells through its interaction with p53. *Asian J Androl*. 2010; 12(2):186–195.
29. Kim TR, Moon JH, Lee HM, et al. SM22alpha inhibits cell proliferation and protects against anticancer drugs and gamma-radiation in HepG2 cells: Involvement of metallothioneins. *FEBS Lett*. 2009;583(20):3356–3362.
30. Liang H, Cheung LW, Li J, et al. Whole-exome sequencing combined with functional genomics reveals novel candidate driver cancer genes in endometrial cancer. *Genome Res*. 2012;22(11):2120–2129.
31. Sorokin AV, Nair BC, Wei Y, et al. Aberrant expression of proPTPRN2 in cancer cells confers resistance to apoptosis. *Cancer Res*. 2015;75(9):1846–1858.
32. Sengelaub CA, Navrazhina K, Ross JB, et al. PTPRN2 and PLCbeta1 promote metastatic breast cancer cell migration through PI(4,5)P2-dependent actin remodeling. *EMBO J*. 2015;35(1):62–76.
33. Schaapveld M, Aleman BM, van Eggermond AM, et al. Second cancer risk up to 40 years after treatment for Hodgkin's lymphoma. *N Engl J Med*. 2015; 373(26):2499–2511.
34. 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(4):392–398, 398e1–398e2.
35. Machiela MJ, Chanock SJ. LDlink: A web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*. 2015;31(21):3555–3557.
36. Boyle AP, Hong EL, Hariharan M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22(9): 1790–1797.