SUPPLEMENTARY MATERIALS

Description of Study Populations

The Black Women's Health Study (BWHS): The BWHS is an ongoing prospective follow-up study of health and illness among US black women, with a focus on cancer.[1, 2] The study began in 1995 when 59,000 AA women 21–69 years of age from across the US completed a 14-page postal health questionnaire. Participants complete follow-up questionnaires every two years. Cases (affected) are women with incident invasive breast cancer. Controls (unaffected) were selected randomly from those who had not developed breast cancer and frequency-matched to cases on age at enrollment, most recent questionnaire completed, and geographic region of residence. Blood and saliva samples as a source of genomic DNA were obtained prior to breast cancer diagnosis for 70% of breast cancer cases and after diagnosis for the remainder.

Cancer Prevention Study (CPS)-II Nutrition Cohort: The CPS-II Nutrition Cohort is a prospective cohort study of 184,194 participants (97,786 women) recruited in 1992-1993 from 21 states.[3] Most participants were aged 50-74 years and self-reported white; 1% of participants were self-reported black. At the time of recruitment, participants completed a baseline demographic, lifestyle, and medical history questionnaire. Follow-up questionnaires were sent to cohort members every two years starting in 1997 to update exposure information and to ascertain newly diagnosed cancer outcomes. Self-reported cancers were verified through abstraction of medical records or linkage with state cancer registries. DNA for this study was extracted from blood specimens donated

by 21,963 female participants in 1998-2001 or from buccal cell samples from an additional 38,180 female participants in 2001-2002. For this study, we selected verified breast cancers in black women through 2013 and controls who were cancer-free at the time of the matched case's diagnosis and matched on age, race, and menopausal status. Samples were obtained prior to diagnosis of breast cancer for 72% of the 58 cases.

Cancer Prevention Study (CPS)-3 Cohort: The CPS-3 prospective cohort includes 303,682 participants (198,632 women) who were cancer-free, aged 30-65 years and lived in 35 states or Puerto Rico at the time of recruitment which occurred between 2006 and 2013.[4] Enrollment was conducted primarily at American Cancer Society community events or at community enrollment "drives". Nearly all women (96.9%) provided a blood specimen at the time of recruitment; these specimens were the source of DNA in this study. Most women self-described as non-Hispanic white (85.3%), while 4.1% were black, 6.5% Latina, and 4.1% other. The cohort receives follow-up surveys every three years starting in 2015. For this study, we selected all self-reported breast cancer cases on the 2015 survey and a randomly-selected subcohort of the larger cohort.

California Teachers' Study (CTS): The CTS is a prospective cohort study of 133,000 women who were residents of California and involved in the teaching professions at the time of enrollment in 1995.[5] Approximately 3,000 participants are African American. Follow-up surveys are conducted every five years and cancer cases are identified through the California Cancer Registry. Saliva and/or blood samples were obtained after occurrence of breast cancer for all 55 AA breast cancer cases included in this analysis.

Multiethnic Cohort Study (MEC): The MEC is a prospective cohort study, which enrolled 21,961 African American women from Los Angeles in 1993-1996. [6] Participant were 45-75 years of age at baseline. Follow-up data collection occurs through participant questionnaires and searches of the California Cancer Registry. Blood samples as a source of DNA were obtained during follow-up; samples were obtained prior to diagnosis for approximately 50% of the breast cancer cases.[6] Approximately two cancer-free participants per breast cancer case were included as controls in the present analysis.

Black Women: Etiology and Survival of Triple-negative Breast Cancers (BEST) Study: The BEST Study is a case-only study of early-onset breast cancer in Black women.[7] Black women diagnosed with invasive breast cancer at age ≤50 years between January 1, 2009 and December 31, 2012 and living in Florida at the time of their diagnosis were invited to enroll in the study. Cases were identified through the Florida State Cancer Registry. Saliva samples as a source of genomic DNA were collected at the time of enrollment.

The Northern California Breast Cancer Family Registry (NC-BCFR): The NC-BCFR is a population-based family study conducted in Northern California, and is one of six sites collaborating in the Breast Cancer Family Registry (BCFR), an international consortium funded by NCI. [8] African American breast cancer cases in NC-BCFR were diagnosed between January 1, 1995 and August 31, 2009; population controls were identified through random digit dialing conducted from 1999-2000. Cases with characteristics suggestive of inherited breast cancer were preferentially invited to enroll in the family registry; criteria included breast cancer diagnosis before age 35 years, prior ovarian or childhood cancer, bilateral breast cancer with a first diagnosis before age 50 years, or a first-degree family history of breast, ovarian, or childhood cancer. Cases who did not meet any of these criteria and were

diagnosed at age 35-64 years were randomly sampled and invited to enroll. Among cases included in the present analysis, 296 met high-risk criteria and 391 did not. Epidemiologic data and blood or mouthwash samples were collected at enrollment, usually within 12 months of diagnosis.

UCI Breast Cancer Study (UCIBCS). The UCIBCS is a case-control study in which all cases were diagnosed in Orange County, California in 1995.[9] Controls were age-matched within 5 years of the age at diagnosis of cases. For the present study, 74 cases and 14 controls were included.

Women's Circle of Health Study (WCHS): The WCHS is a case-control study of breast cancer initiated in four New York City boroughs and later expanded to ten counties of New Jersey.[10] Eligible cases were women with incident breast cancer diagnosed between 25 and 74 years of age; controls were identified through random-digit dialing and community health events. Blood or saliva samples as a source of genomic DNA were obtained at the time of enrollment, usually within one year of diagnosis.

Wisconsin Women's Health Study (WWHS): The WWHS is a series of population-based control studies conducted in Wisconsin.[11] Data for this project were collected in 2001-2007; breast cancer cases aged 20-69 years were identified from the Wisconsin mandatory cancer registry and controls were randomly selected from lists of licensed drivers and frequency-matched by age group to cases. Participants were interviewed by telephone about one year after diagnosis for cases, or a similar reference date for controls. At the conclusion of the interview, women were invited to contribute self-collected saliva samples by mail for genotyping.

Sister Study. The Sister Study is a prospective cohort study of women who have a sister diagnosed with breast cancer.[12] The Sister Study began in July 2003, enrolling volunteers without a history of breast cancer themselves, aged 35-74, residing in the United States and Puerto Rico who had a sister with breast cancer. Enrollment of the cohort of 50,884 closed in March 2009. Baseline data collection included a comprehensive risk factor oriented Computer Assister Telephone Interview; self-administered questionnaires on family medical history, early life exposures, diet, and personal care products; and a home visit for collecting biospecimens (first morning void urine, blood, toenail clippings, and household dust). Cases included African American women diagnosed during follow-up; controls were African American women in a random subcohort of all participants who did not develop breast cancer.

Absolute risk estimation for breast cancer

Absolute risk is the probability that an individual with a measured set of risk factors (e.g. mutation status) and disease-free at age a will be diagnosed with the disease in the subsequent τ years[13]. Let Z be the set of measured risk factors; in our calculations, Z is an indicator for carrying a pathogenic variant in a specified risk gene. We can express the absolute risk as

$$R(a,\tau,Z) = \int_a^{a+\tau} h_1\,(u|Z) \times \exp\left(-\int_a^u \{h_1(v|Z) + h_2(v|Z)\}\,\mathrm{d}v\right)\mathrm{d}u$$

where $h_1(a|Z)$ is the conditional breast-cancer hazard at age a and $h_2(a|Z)$ is the hazard for competing risks (mortality) at age a. The breast-cancer hazard can be parameterized with a proportional hazards model as:

$$h_1(a|Z) = h_{10}(a)\exp(\beta Z)$$

where $h_{10}(a)$ is the baseline hazard and β represents the breast-cancer relative risk for Z. Estimates for β can be estimated from the case-control data and the baseline hazards estimated by the following relationship between the baseline hazard and the marginal hazard

$$h_1^*(a) = h_{10}(a)\mathrm{E}(\exp(\beta Z)) \approx \int h_{10}(a)\exp(\beta z)dF(z)$$

where $F(Z)$ denotes the distribution of the risk factors in the population. For each age, we solve for $h_{10}(a)$ by using the expected distribution of the risk factor in the population[14, 15]. We assume the risk factors, Z, are independent of the hazard for the competing events and directly use the baseline hazard.

For each of four genes, we estimated absolute risk of breast cancer for an African American who carried a pathogenic mutation in that gene. Odds ratio estimates from affected versus unaffected women were used as estimates for β and the 95% confidence intervals for the odds ratios were used to generate confidence bands for the absolute risk curves. SEER18 age-specific incidence rates[16] were used for estimation of the baseline hazard combined with the observed carrier frequency in the controls for the population frequency. The age-specific competing hazards model for all-cause mortality, $h_2(a|Z)$, was assumed to be independent of the pathogenic variant status—i.e. $h_2(a|Z) = h_2(a)$—and was estimated from CDC WONDER US mortality rates among African American women, subset to the same states as the SEER18 registries. Lifetime absolute risk curves were then estimated using the absolute risk equation above and computing the cumulative risk.

Supplementary Table 1. Contributing studies

Study Acronym	Study Name	Study design	Region of USA	Mean age (SD)	% with first degree family hx of breast cancer	
					Cases	Controls
BWHS	Black Women's Health Study[1, 2]	Prospective cohort study: nested case-control	Nationwide	53.9 (10.7)	14.9	9.5
CPSII	Cancer Prevention Study-II Nutrition Cohort[3]	Prospective cohort study: nested case-control	Nationwide	69.7 (7.6)	15.5	16.7
CPS3	Cancer Prevention Study-3[4]	Prospective cohort study: nested case-cohort	Nationwide	48.0 (10.1)	31.2	20.3
CTS	California Teachers Study[5]	Prospective cohort study: nested case-control	California	62.4 (9.9)	16.7	14.3
MEC	Multiethnic Cohort[6]	Prospective cohort study: nested case-control	California	67.1 (9.4)	19.1	12.0
BEST	Black Women: Etiology and Survival of TNBC[7]	Case-only study with preferential selection at young age at diagnosis	Florida	42.1 (6.1)	19.6	--
NC-BCFR	Northern California Breast Cancer Family Registry[8]	Population-based case-cohort study and population controls, preferential selection of high-risk women	Northern California	51.2 (8.8)	26.7	9.3
UCIBCS	UCI Breast Cancer Study[9]	Population-based case-control study with preferential selection at young age at diagnosis	California	45.4 (9.7)	0	0
WCHS	Women's Circle of Health Study[10]	Population-based case control study	NJ, NY	52.5 (10.6)	16.7	12.2
WWHS	Wisconsin Women's Health Study[11]	Population-based case-control study	Wisconsin	54.0 (9.6)	18.9	16.1
SISTER	SISTER study[12]	Prospective cohort of sisters of breast cancer patients: case-cohort	Nationwide	55.4 (8.9)	100	100

Supplementary Table 2. Genes in the custom QIAseq panel evaluated for presence of mutations in AA women with breast cancer

HGNC symbol	RefSeq	Ensemble Transcript Id	Chromosome name	Start position	End position	strand	Ensemble gene id
ATM	NM000051.3	ENST00000278616	11	108093211	108239829	1	ENSG00000149311
BARD1	NM000465.3	ENST00000260947	2	215590370	215674428	-1	ENSG00000138376
BRCA1	NM007294.3	ENST00000357654	17	41196312	41277500	-1	ENSG00000012048
BRCA2	NM000059.3	ENST00000544455	13	32889611	32973805	1	ENSG00000139618
BRIP1	NM032043.2	ENST00000259008	17	59758627	59940882	-1	ENSG00000136492
CDH1	NM004360.4	ENST00000261769	16	68771128	68869451	1	ENSG00000039068
CDKN2A	NM000077.4	ENST00000304494	9	21967751	21995300	-1	ENSG00000147889
CHEK2	NM007194.3	ENST00000328354	22	29083731	29138410	-1	ENSG00000183765
ERCC3	NM000122.1	ENST00000285398	2	128014866	128051752	-1	ENSG00000163161
FANCC	NM000136.2	ENST00000289081	9	97861336	98079991	-1	ENSG00000158169
FANCM	NM020937.3	ENST00000267430	14	45605143	45670093	1	ENSG00000187790
MLH1	NM000249.3	ENST00000231790	3	37034823	37107380	1	ENSG00000076242
MRE11A	NM005591.3	ENST00000323929	11	94152895	94227074	-1	ENSG00000020922
MSH2	NM000251.2	ENST00000233146	2	47630108	47789450	1	ENSG00000095002
MSH6	NM000179.2	ENST00000234420	2	47922669	48037240	1	ENSG00000116062
NBN	NM002485.4	ENST00000265433	8	90945564	91015456	-1	ENSG00000104320
NF1	NM001042492.2	ENST00000358273	17	29421945	29709134	1	ENSG00000196712
PALB2	NM024675.3	ENST00000261584	16	23614488	23652631	-1	ENSG00000083093
PTEN	NM000314.6	ENST00000371953	10	89622870	89731687	1	ENSG00000171862
RAD51C	NM058216.2	ENST00000337432	17	56769934	56811703	1	ENSG00000108384
RAD51D	NM001142571	ENST00000345365	17	33426811	33448541	-1	ENSG00000185379
RECQL	NM002907	ENST00000444129	12	21621845	21654603	-1	ENSG00000004700
TP53	NM000546.5	ENST00000269305	17	7565097	7590856	-1	ENSG00000141510

Supplementary Table 3. Allele count (AC) of pathogenic mutations in African American women with breast cancer (affected) and unaffected African American women

Gene with pathogenic variants	Affected	Unaffected	Total
ATM	41	16	57
c.1017delT	1	1	2
c.1208C>A_p.Ser403X	3	—	3
c.1339C>T_p.Arg447X	—	1	1
c.1464G>A_p.Trp488X	1	—	1
c.1564_1565delGA	1	—	1
c.1915_1916insT	1	1	2
c.2023C>T_p.Gln675X	—	1	1
c.2062G>T_p.Glu688X	1	—	1
c.2124+1G>A	1	—	1
c.2250G>A_p.=	1	—	1
c.2806_2809dupCTAG	—	1	1
c.289delA	1	—	1
c.2921+1G>A	1	—	1
c.3049C>T_p.Gln1017X	1	1	2
c.3078-1G>A	1	—	1
c.3372C>G_p.Tyr1124X	1	—	1
c.3539_3540delTG	1	—	1
c.3865A>T_p.Lys1289X	1	—	1
c.3G>A_p.Met1?	—	1	1
c.4358delT	1	—	1
c.4804_4805delGT	1	—	1
c.5118_5121delAGAA	1	—	1
c.5798G>A_p.Trp1933X	1	1	2
c.5890A>T_p.Lys1964X	1	—	1
c.680C>G_p.Ser227X	—	1	1
c.6813_6814delGinsCA	1	—	1

c.6829delC	1	—	1
c.7000_7003delTACA	1	1	2
c.72+1G>A	1	—	1
c.7220C>A_p.Ser2407X	1	—	1
c.7271T>G_p.Val2424Gly	1	—	1
c.742C>T_p.Arg248X	1	—	1
c.748C>T_p.Arg250X	1	2	3
c.7671_7674delGTTT	—	1	1
c.7913G>A_p.Trp2638X	5	1	6
c.7998dupT	1	—	1
c.8545C>T_p.Arg2849X	—	1	1
c.8725A>T_p.Arg2909X	1	—	1
c.8835_8836delGT	—	1	1
c.902-1G>A	1	—	1
c.9139C>T_p.Arg3047X	3	—	3
BARD1	7	8	15
c.1023delG	1	—	1
c.1061C>G_p.Ser354X	1	—	1
c.1270delA	2	1	3
c.1385G>A_p.Trp462X	1	1	2
c.1652C>G_p.Ser551X	1	—	1
c.188T>G_p.Leu63X	1	—	1
c.1921C>T_p.Arg641X	—	4	4
c.334C>T_p.Arg112X	—	2	2
BRCA1	90	3	93
c.1386delG	—	1	1
c.1636_1654del19	1	—	1
c.182G>A_p.Cys61Tyr	1	—	1
c.190T>G_p.Cys64Gly	1	—	1
c.2071delA	1	—	1

c.213-11T>G	1	—	1
c.2740G>T_p.Glu914X	3	—	3
c.2880dupC	1	—	1
c.3016delC	1	—	1
c.3358_3359delGT	1	—	1
c.3481_3491del11	1	—	1
c.3598C>T_p.Gln1200X	—	1	1
c.3748G>T_p.Glu1250X	1	—	1
c.3756_3759delGTCT	3	—	3
c.391A>T_p.Arg131X	1	—	1
c.3G>T_p.Met1?	1	—	1
c.4041_4042delAG	1	—	1
c.4065_4068delTCAA	1	—	1
c.4163_4166delAGAG	1	—	1
c.4327C>T_p.Arg1443X	2	—	2
c.4357+1G>A	9	—	9
c.4389C>A_p.Tyr1463X	2	—	2
c.4408G>T_p.Glu1470X	1	—	1
c.4484G>T_p.Arg1495Met	3	—	3
c.4603G>T_p.Glu1535X	1	—	1
c.4668dupA	1	—	1
c.4986+6T>C	5	—	5
c.5152+1G>A	1	—	1
c.5165C>T_p.Ser1722Phe	1	—	1
c.5177_5180delGAAA	8	—	8
c.5251C>T_p.Arg1751X	4	—	4
c.5278-1G>C	1	—	1
c.5324T>G_p.Met1775Arg	7	—	7
c.5387C>A_p.Ser1796X	4	—	4
c.5467+1G>A	3	—	3

c.5562_5563insGG	1	—	1
c.68_69delAG	1	—	1
c.815_824dup10	11	1	12
c.981_982delAT	1	—	1
del exon 13-15	2	—	2
BRCA2	107	12	119
c.1103C>G_p.Ser368X	1	—	1
c.115delG	1	—	1
c.1310_1313delAAGA	6	—	6
c.1570_1571delAT	1	—	1
c.1705_1706delCA	1	—	1
c.1796_1800del5	2	—	2
c.1800T>A_p.Tyr600X	2	—	2
c.1887_1893del7	2	—	2
c.2092delC	1	—	1
c.2099T>A_p.Leu700X	1	—	1
c.2564_2565delCA	1	—	1
c.2808_2811delACAA	2	—	2
c.2808delA	—	2	2
c.2830A>T_p.Lys944X	1	—	1
c.2957_2958insG	1	—	1
c.3009_3010delCA	1	—	1
c.3450dupT	1	—	1
c.3599_3600delGT	1	—	1
c.3680_3681delTG	3	2	5
c.3847_3848delGT	1	—	1
c.3860_3863delATAA	1	—	1
c.3922G>T_p.Glu1308X	1	—	1
c.4211delC	1	—	1
c.4456_4459delGTTA	1	—	1

c.4471_4474delCTGA	2	—	2
c.4477_4478delGA	2	—	2
c.4552delG	1	—	1
c.4712_4713delAG	1	—	1
c.476-4_476-1delCCAGinsT	1	—	1
c.489_490insG	1	—	1
c.4912A>T_p.Lys1638X	1	—	1
c.4936_4939delGAAA	3	—	3
c.4965C>G_p.Tyr1655X	1	—	1
c.5073dupA	2	—	2
c.517-2A>G	1	—	1
c.518delG	1	—	1
c.5350_5351delAA	2	1	3
c.5351dupA	2	—	2
c.5576_5579delTTAA	1	—	1
c.5616_5620del5	4	1	5
c.5621_5624delTTAA	2	—	2
c.5946delT	1	—	1
c.5979dupA	1	—	1
c.6124C>T_p.Gln2042X	1	—	1
c.6137C>A_p.Ser2046X	1	—	1
c.6468_6469delTC	1	—	1
c.6486_6489delACAA	1	—	1
c.658_659delGT	—	1	1
c.6600_6601delTT	1	—	1
c.6611delC	1	—	1
c.67+2T>G	—	1	1
c.6938-1G>A	—	1	1
c.7024C>T_p.Gln2342X	2	—	2
c.7115C>A_p.Ser2372X	1	—	1

c.7115C>G_p.Ser2372X	1	—	1
c.71delT	1	—	1
c.7208_7211delCCAA	2	—	2
c.7480C>T_p.Arg2494X	1	—	1
c.7485dupT	—	1	1
c.7543dupA	1	—	1
c.7558C>T_p.Arg2520X	3	1	4
c.7757G>A_p.Trp2586X	1	—	1
c.7805+1G>A	1	—	1
c.793+1G>A	1	—	1
c.8009C>T_p.Ser2670Leu	—	1	1
c.8777T>A_p.Leu2926X	2	—	2
c.8817_8820delGAAA	1	—	1
c.8969G>A_p.Trp2990X	2	—	2
c.9041C>G_p.Ser3014X	2	—	2
c.9253dupA	6	—	6
c.9294C>G_p.Tyr3098X	1	—	1
c.9329dupA	1	—	1
c.9382C>T_p.Arg3128X	5	—	5
c.9435_9436delGT	1	—	1
c.9924C>G_p.Tyr3308X	1	—	1
BRIP1	10	6	16
c.1045G>C_p.Ala349Pro	3	1	4
c.112dupA	1	—	1
c.1156A>T_p.Lys386X	—	1	1
c.1162C>T_p.Gln388X	1	—	1
c.141delC	—	1	1
c.1430T>G_p.Leu477X	—	1	1
c.1871C>A_p.Ser624X	1	—	1
c.1936-2A>G	1	—	1

c.2038_2039dupTT	—	1	1
c.2114_2118del5	1	—	1
c.2392C>T_p.Arg798X	1	—	1
c.376C>T_p.Gln126X	1	—	1
c.622C>T_p.Gln208X	—	1	1
CDH1	4	2	6
c.1137+1G>A	1	—	1
c.187C>T_p.Arg63X	—	1	1
c.1979dupT	1	1	2
c.454C>T_p.Gln152X	1	—	1
c.532-1G>C	1	—	1
CDKN2A	1	—	1
c.159G>C_p.Met53Ile	1	—	1
CHEK2	19	6	25
c.1011C>A_p.Tyr337X	—	1	1
c.1100delC	9	2	11
c.1116dupC	5	—	5
c.1139_1140delTC	1	—	1
c.1263delT	1	—	1
c.1462-1G>A	1	—	1
c.1502_1503dupAG	1	—	1
c.339C>G_p.Tyr113X	—	2	2
c.684-1G>C	—	1	1
c.847-1G>A	1	—	1
ERCC3	14	11	25
c.1354C>T_p.Arg452X	6	6	12
c.1678_1679delGT	1	—	1
c.1720C>T_p.Arg574X	2	1	3
c.1725delG	1	—	1
c.1735delT	1	—	1

c.1757_1758delAG	1	—	1
c.1757delA	—	1	1
c.1828-2A>G	—	1	1
c.1841C>A_p.Ser614X	—	1	1
c.235-1G>A	—	1	1
c.657+1G>A	1	—	1
c.700C>T_p.Arg234X	1	—	1
FANCC	22	10	32
c.1043_1044delTT	1	—	1
c.1417C>T_p.Gln473X	1	—	1
c.1642C>T_p.Arg548X	2	—	2
c.319C>T_p.Gln107X	—	1	1
c.355_360delTCTCATinsA	13	7	20
c.45G>A_p.Trp15X	2	—	2
c.520C>T_p.Arg174X	1	—	1
c.553C>T_p.Arg185X	1	2	3
del exon 5-7	1	—	1
FANCM	13	11	24
c.1213C>T_p.Arg405X	—	1	1
c.1362_1363insCAAAGTTAAA	1	—	1
c.1506_1507insTA	—	2	2
c.1777C>T_p.Arg593X	—	1	1
c.1879C>T_p.Arg627X	—	1	1
c.1972C>T_p.Arg658X	2	—	2
c.2003-1G>A	1	—	1
c.2465_2469del5	—	—	—
c.2586_2589delAAAA	—	1	1
c.3088C>T_p.Arg1030X	1	—	1
c.3146T>A_p.Leu1049X	—	1	1
c.3628C>T_p.Gln1210X	2	—	2

c.3732_3735dupAATA	1	—	1
c.3985delT	1	—	1
c.428_430delCGAinsGT	1	—	1
c.4317+1G>A	—	1	1
c.4504_4505delAG	1	—	1
c.473dupT	1	—	1
c.4812delT	—	2	2
c.919-1G>C	—	1	1
del exon 7-11	1	—	1
MLH1	—	1	1
c.1517T>C_p.Val506Ala	—	1	1
MRE11A	2	3	5
c.1015A>T_p.Lys339X	2	1	3
c.1714C>T_p.Arg572X	—	1	1
c.1771C>T_p.Gln591X	—	1	1
MSH6	4	3	7
c.3226C>T_p.Arg1076Cys	1	—	1
c.3557-1G>A	1	—	1
c.3739delA	—	1	1
c.3934_3937dupGTTA	1	—	1
c.3991C>T_p.Arg1331X	—	1	1
c.741delA	—	1	1
c.973C>T_p.Gln325X	1	—	1
NBN	8	10	18
c.1154_1155delAA	1	—	1
c.11delT	—	2	2
c.1255_1258delAATA	—	1	1
c.127C>T_p.Arg43X	2	—	2
c.1484delC	—	1	1
c.1741C>T_p.Gln581X	1	—	1

c.535_537delTTCinsAA	—	1	1
c.580G>T_p.Glu194X	—	1	1
c.657_661del5	—	2	2
c.698_701delAACA	1	—	1
c.808_809delGT	1	—	1
c.83_89del7	—	1	1
c.897-2A>T	2	1	3
NF1	6	1	7
c.2033delC	1	—	1
c.4333-2A>G	1	—	1
c.5488C>T_p.Arg1830Cys	1	—	1
c.6004C>T_p.Gln2002X	1	—	1
c.611dupT	—	1	1
c.663G>A_p.Trp221X	1	—	1
c.7115delA	1	—	1
PALB2	57	5	62
c.1039G>T_p.Glu347X	—	1	1
c.12dupT	1	—	1
c.1479delC	1	—	1
c.1490delA	1	—	1
c.156delA	1	—	1
c.1633G>T_p.Glu545X	2	—	2
c.172_175delTTGT	3	1	4
c.196C>T_p.Gln66X	1	—	1
c.1972delG	1	—	1
c.2120delC	1	—	1
c.212-2A>G	1	—	1
c.2167_2168delAT	1	—	1
c.226delA	3	—	3
c.2386G>T_p.Gly796X	2	—	2

c.2674G>T_p.Glu892X	1	—	1
c.2730T>A_p.Tyr910X	1	—	1
c.2835-1G>C	1	—	1
c.2878delC	2	—	2
c.2888delC	1	—	1
c.2938delA	—	1	1
c.3048delT	1	—	1
c.3058C>T_p.Gln1020X	1	—	1
c.3113G>A_p.Trp1038X	2	—	2
c.3116delA	2	—	2
c.3323delA	9	1	10
c.347T>A_p.Leu116X	1	—	1
c.3549C>G_p.Tyr1183X	2	—	2
c.451C>T_p.Gln151X	1	—	1
c.519delG	1	—	1
c.527_531del5	1	—	1
c.758dupT	1	—	1
c.79G>T_p.Glu27X	4	—	4
c.844_847delAGAT	1	—	1
del exon 1-10	1	—	1
del exon 1-6	1	—	1
del exon 8-13	1	—	1
(blank)	2	1	3
RAD51C	9	3	12
c.186_187delAA	—	1	1
c.405-1G>C	2	—	2
c.472dupA	1	—	1
c.535delC	1	—	1
c.625_628delTATT	—	1	1
c.630T>G_p.Tyr210X	1	—	1

c.904+5G>T	1	—	1
c.914G>A_p.Trp305X	2	—	2
c.97C>T_p.Gln33X	1	1	2
RAD51D	8	2	10
c.326dupC	3	1	4
c.461_462insTT	—	1	1
c.473_480+1del9	1	—	1
c.739-1G>A	1	—	1
del exon 7-10	2	—	2
del exon 9-10	1	—	1
RECQL	17	5	22
c.1194T>A_p.Tyr398X	1	—	1
c.1267dupA	1	—	1
c.1353_1355+1delCAAG	—	3	3
c.1461_1462delGA	2	—	2
c.1489dupA	5	1	6
c.1667_1667+3delAGTA	3	—	3
c.1756_1759dupATTA	—	1	1
c.1870C>T_p.Gln624X	1	—	1
c.427_428delCT	1	—	1
c.950-2A>C	1	—	1
c.962_965delGTTT	2	—	2
TP53	5	1	6
c.375+1G>A	—	1	1
c.473G>A_p.Arg158His	1	—	1
c.714T>A_p.Cys238X	1	—	1
c.743G>A_p.Arg248Gln	2	—	2
c.818G>A_p.Arg273His	1	—	1
Grand Total	444	119	563

Supplementary Table 4. Frequency of pathogenic mutations in known or suspected breast cancer susceptibility genes and associations with breast cancer risk in African American women from population-based* studies, with adjustment for individual study

Gene	Affected (N=3916)		Unaffected (N=4925)		Odds ratio† (95% CI)	p value‡
	Mutated alleles No.	Mutation frequency%	Mutated alleles No.	Mutation frequency %		
ATM	28	0.72	16	0.33	1.72 (0.92-3.32	0.096
BARD1	7	0.18	8	0.16	0.94 (0.32-2.71)	0.91
BRCA1	41	1.05	1	0.02	39.92 (8.60-710)	<0.001
BRCA2	72	1.84	12	0.24	8.24 (4.54-16.37)	<0.001
BRIP1	6	0.15	6	0.12	1.08 (0.32-3.65)	0.89
CHEK2	15	0.38	6	0.12	3.42 (1.33-10.00)	0.015
ERCC3	13	0.33	9	0.18	2.51 (1.06-6.20)	0.038
FANCC	16	0.41	10	0.20	2.40 (1.08-5.61)	0.035
FANCM	11	0.28	11	0.22	1.08 (0.45-2.63)	0.86
NBN	4	0.10	9	0.18	0.50 (0.13-1.61)	0.27
PALB2	39	1.00	5	0.10	8.12 (3.43-23.91)	<0.001
RAD51C	7	0.18	3	0.06	2.75 (0.75-12.94)	0.15
RAD51D	6	0.15	2	0.04	4.09 (0.87-29.78)	0.10
RECQL	12	0.31	5	0.10	3.02 (1.07-9.82)	0.046
Total	277	7.07	103	2.09		

*Studies that did not preferentially enroll cases based on family history or age; included were BWHS, CPSII, CPS3, CTS, MEC, WCHS, WWHS

†Odds ratios adjusted for study, age, and first-degree family history of breast cancer. Reference group is women who have no mutations in the given gene.

‡Two-sided p-value from logistic regression analysis

Supplementary Table 5. Associations between pathogenic mutations in breast cancer susceptibility genes and triple negative breast cancer risk

	Women with triple negative breast cancer (N=654)			
Gene	Mutated alleles No.	Mutation frequency %	OR (95% CI)*	p value†
ATM	0	-		
BARD1	3	0.45		
BRCA1	43	6.57	180 (37.9 - 3238)	<0.001
BRCA2	13	1.99	6.23 (2.65-14.84)	<0.001
BRIP1	2	0.22		
CDH1	1	0.15		
CHEK2	1	0.15		
ERCC3	1	0.15		
FANCC	2	0.30		
NF1	2	0.30		
PALB2	14	2.14	23.5 (8.35-76.71)	<0.001
RAD51C	2	0.30		
RAD51D	1	0.15		
RECQL	3	0.45		
TP53	0	-		

*Odd ratios adjusted for study design, age, and first-degree family history of breast cancer. Reference group is women who have no mutations in the given gene.

† Two-sided p-value from logistic regression analysis

Supplementary Table 6. Enrichment of pathogenic susceptibility gene mutations in ER-positive relative to ER-negative breast cancer

	OR (95% CI)*	p value†
ATM	1.94 (0.85-5.22)	0.14
BARD1	0.45 (0.083-2.46)	0.33
BRCA1	0.14 (0.080-0.24)	< 0.001
BRCA2	0.84 (0.54-1.32)	0.44
BRIP1	0.57 (0.13-2.93)	0.47
CHEK2	2.98 (0.83-19.01)	0.15
ERCC3	1.50 (0.45-6.75)	0.54
FANCC	1.16 (0.43-3.66)	0.78
PALB2	0.40 (0.22-0.74)	0.0030
RAD51C	0.45 (0.10-1.93)	0.26
RAD51D	0.13 (0.019-0.59)	0.015
RECQL	0.82 (0.25-3.09)	0.74

*Adjusted for study design, age and first-degree family history of breast cancer. Reference group is women who have no mutations in the given gene.
†Two-sided p-value from logistic regression analysis

Supplementary Table 7. Comparisons of pathogenic mutations in breast cancer susceptibility genes in African American affected and unaffected women with pathogenic mutations in gnomAD AFR reference unaffected women

	Affected		Unaffected		gnomAD AFR		Affected vs. gnomAD		Unaffected vs. gnomAD	
Gene	AC*	AN*	AC	AN	AC	AN	Odds ratio* (95% CI)	p-value†	Odds ratio* (95% CI)	p-value†
ATM	39	10108	16	9986	18	15277	3.28 (1.87-5.79)	<0.001	1.36 (0.68-2.69)	0.38
BARD1	7	10108	8	9986	7	15030	1.49 (0.46-4.76)	0.59	1.72 (0.63-5.27)	0.30
BRCA1	79	10108	1	9986	9	15287	13.4 (6.66-27.5)	<0.001	0.17 (0.01-1.23)	0.10
BRCA2	98	10108	12	9986	20	15188	7.42 (4.56-12.44)	<0.001	0.91 (0.42-1.89)	0.86
BRIP1	9	10108	6	9986	17	15279	0.80 (0.42-3.95)	0.69	0.54 (0.21-1.40)	0.21
CDH1	4	10108	2	9986	1	14947	—	—	—	—
CHEK2	19	10108	6	9986	10	14912	2.81 (1.28-6.56)	0.0074	0.90 (0.32-2.53)	0.90
ERCC3	14	10108	9	9986	9	15290	2.36 (0.96-5.61)	0.053	1.53 (0.57-4.12)	0.47
FANCC	21	10108	10	9986	12	15256	2.65 (1.25-5.73)	0.0069	1.27 (0.54-2.99)	0.66
NF1	6	10108	1	9986	4	15240	2.26 (0.62-8.55)	0.21	—	—
PALB2	49	10108	4	9986	15	15282	4.96 (2.80-8.95)	<0.001	0.41 (0.12-1.23)	0.16
PTEN	0	10108	0	9986	1	15159	—	—	—	—
RAD51C	9	10108	3	9986	2	15268	6.80 (1.51-43.86)	0.0095	2.29 (0.36-18.48)	0.39
RAD51D	5	10108	2	9986	2	15227	3.77 (0.78-26.73)	0.12	—	—
RECQL	16	10108	5	9986	12	15037	1.99 (0.91-4.29)	0.082	0.63 (0.21-1.80)	0.46
TP53	5	10108	1	9986	1	15265	—	—	—	—

*AC (Allele Count); AN (Allele Number)

*Reference group is women who have no mutations in the given gene.

† P-value from two-sided Fisher's exact text.

Supplementary Table 8. Associations between pathogenic mutations in predisposition genes and breast cancer risk among African American women under age 50 years and African American women aged 50 years and older

	Age <50 years		Age ≥50 years	
Gene	OR (95% CI)*	p-value†	OR (95% CI)*	p-value†
ATM	1.40 (0.55-3.71)	0.49	2.23 (1.01-5.44)	0.06
BARD1	0.76 (0.09-6.63)	0.79	0.87 (0.24-3.05)	0.83
BRCA1	—	—	9.23 (1.76-170)	0.03
BRCA2	14.61 (5.25-60.8)	< 0.001	4.02 (1.97-9.06)	<0.001
BRIP1	0.76 (0.09-6.63)	0.79	1.21 (0.30-5.17)	0.78
CHEK2	4.23 (1.02-28.6)	0.07	2.92 (0.95-10.8)	0.08
FANCC	3.87 (1.19-17.3)	0.04	1.21 (0.40-3.65)	0.73
PALB2	12.38 (2.42-226)	0.02	7.45 (2.90-25.3)	<0.001
RAD51C	5.85 (0.84-115)	0.12	1.38 (0.21-11.1)	0.73
RAD51D	1.05 (0.17-8.30)	0.96	—	—
RECQL	2.97 (0.62-21.2)	0.20	3.21 (0.90-15.0)	0.09

*Odd ratios adjusted for study design, age and first-degree family history of breast cancer. Reference group is women who have no mutations in the given gene

† Two-sided p-value from logistic regression analysis

Supplementary Table 9. Frequency of pathogenic mutations in known or suspected breast cancer susceptibility genes and associations with breast cancer risk in African American women without a first-degree family history of breast cancer

Gene	Affected (N=4038)		Unaffected (N=4403)		Odds Ratio* (95% CI)	p value†
	Mutated Alleles No.	Mutation Frequency%	Mutated Alleles No.	Mutation Frequency %		
ATM	27	0.67	12	0.27	1.96 (0.98-4.14)	0.064
BARD1	6	0.15	8	0.18	0.71 (0.22-2.14)	0.54
BRCA1	43	1.07	0	0	-	-
BRCA2	69	1.71	9	0.20	8.03 (4.15-17.52)	< 0.001
BRIP1	7	0.17	5	0.11	1.32 (0.39-4.75)	0.66
CDH1	2	0.05	2	0.05	1.01 (0.11-9.18)	0.99
CDKN2A	1	0.03	0	0	-	-
CHEK2	15	0.37	6	0.14	2.76 (1.07-7.99)	0.044
ERCC3	13	0.32	9	0.20	2.26 (0.96-5.57)	0.065
FANCC	17	0.42	8	0.18	2.29 (0.98-5.82)	0.065
FANCM	11	0.27	8	0.18	1.61 (0.62-4.31)	0.33
MRE11A	2	0.05	3	0.07	-	-
MSH6	2	0.05	2	0.05	1.00 (0.11-9.17)	0.999
NBN	6	0.15	8	0.18	0.72 (0.22-2.20)	0.56
NF1	5	0.12	0	0	-	-
PALB2	34	0.84	4	0.09	8.39 (3.25-28.62)	< 0.001
RAD51C	9	0.22	3	0.07	3.02 (0.87-13.91)	0.11
RAD51D	4	0.10	1	0.02	4.19 (0.56-85.60)	0.22
RECQL	14	0.35	5	0.11	2.91 (1.08-9.20)	0.046
TP53	3	0.07	1	0.02	-	-
Total	290	7.2	95	2.2		

* Adjusted for study design and age. Reference group is women who have no mutations in the given gene.

† Two-sided p-value from logistic regression analysis

Supplementary Table 10. Frequency of pathogenic mutations in known or suspected breast cancer susceptibility genes and associations with breast cancer risk in African American women with a first-degree family history of breast cancer

Gene	Affected (N=1049)		Unaffected (N=700)		Odds ratio* (95% CI)	p value
	Mutated alleles No.	Mutation frequency %	Mutated alleles No.	Mutation frequency %		
ATM	13	1.24	4	0.57	2.11 (0.72-7.65)	0.204
BARD1	1	0.10	0	0.16	-	-
BRCA1	42	4.00	3	0.43	9.02 (3.21-37.76)	< 0.001
BRCA2	36	3.43	3	0.43	7.97 (2.83-33.30)	< 0.001
BRIP1	3	0.29	1	0.14	1.96 (0.24-40.34)	0.564
CDH1	2	0.19	0	0	-	-
CHEK2	4	0.38	0	0	-	-
ERCC3	1	0.10	2	0.29	0.26 (0.01-2.89)	0.286
FANCC	3	0.29	2	0.29	1.40 (0.22-11.1)	0.718
FANCM	2	0.19	3	0.43	0.40 (0.05-2.52)	0.326
MSH6	2	0.19	1	0.14	1.31 (0.12-28.6)	0.828
NBN	2	0.19	2	0.29	0.54 (0.06-4.64)	0.545
NF1	1	0.10	1	0.14	1.10 (0.04-28.1)	0.949
PALB2	22	2.10	1	0.14	13.55 (2.81-243)	0.011
RAD51D	4	0.38	1	0.14	2.26 (0.33-44.7)	0.471
RECQL	14	0.30	5	0.10	3.19 (1.22-9.90)	0.027
TP53	2	0.19	0	0	-	-
Total	154	14.7	29	4.1		

* ORs adjusted for study design and age. Reference group is women who have no mutations in the given gene.

† Two-sided p-value from logistic regression analysis

Supplementary Table 11. Prevalence of pathogenic mutations in African American women with breast cancer according to age at diagnosis and first-degree family history of breast cancer

	Age at breast cancer diagnosis (years)											
	≤40			41-50			51-60			>60		
	Cases N	Mut. N	Mut. %	Cases N	Mut. N	Mut. %	Cases N	Mut. N	Mut. %	Cases N	Mut. N	Mut. %
BRCA1 mutations												
All breast cancer												
Famhx-No	505	18	3.6%	1131	18	1.6%	1162	6	0.5%	1229	1	0.08%
Famhx-Yes	75	17	22.7%	263	17	6.5%	322	6	1.8%	384	2	0.5%
Overall	607	38	6.3%	1440	25	2.5%	1510	13	0.8%	1633	3	0.02%
Estrogen receptor negative breast cancer												
Famhx-No	157	14	8.9%	327	13	4.0%	323	3	0.9%	266	1	0.4%
Famhx-Yes	26	14	53.8%	79	10	12.7%	78	6	7.7%	71	2	2.8%
Overall	194	30	15.5%	423	24	5.6%	409	10	2.4%	342	3	0.9%
Estrogen receptor positive breast cancer												
Famhx-No	277	4	1.4%	667	4	0.3%	691	3	0.4%	773	0	0.0%
Famhx-Yes	44	3	6.8%	151	5	3.3%	201	0	0.0%	257	0	0.0%
Overall	337	8	2.4%	846	9	1.1%	908	3	0.3%	1039	0	0.0%
BRCA2 mutations												
All breast cancer												
Famhx-No	505	21	4.2%	1131	29	2.6%	1162	15	1.3%	1229	4	0.3%
Famhx-Yes	75	2	2.7%	263	10	4.6%	322	7	3.7%	384	7	2.3%
Overall	607	25	4.1%	1440	41	2.9%	1510	27	1.8%	1633	13	0.8%
Estrogen receptor negative breast cancer												
Famhx-No	157	6	3.8%	327	8	2.4%	323	7	2.2%	266	1	0.4%
Famhx-Yes	26	1	3.9%	79	1	1.3%	78	1	1.3%	71	5	7.0%
Overall	194	7	3.6%	423	9	2.1%	409	8	2.0%	342	7	2.1%
Estrogen receptor positive breast cancer												
Famhx-No	277	14	5.1%	667	18	2.7%	691	6	0.9%	773	2	0.2%
Famhx-Yes	44	1	2.3%	151	8	5.3%	201	11	5.5%	257	3	1.2%
Overall	337	17	5.0%	846	26	3.1%	908	17	1.9%	1039	5	0.5%
Mutations in any of 12 genes (*ATM, BARD1, BRCA1, BRCA2, CDH1, CHEK2, NF1, PALB2, PTEN, RAD51C, RAD51D, TP53)*												
All breast cancer												
Famhx-No	505	52	10.3%	1131	75	6.6%	1162	56	4.8%	1229	31	2.5%
Famhx-Yes	75	22	30.7%	263	29	14.4%	322	25	10.2%	384	23	7.3%
Overall	607	80	13.2%	1440	115	8.0%	1510	90	6.0%	1633	60	3.7%
Estrogen receptor negative breast cancer												
Famhx-No	157	24	15.3%	327	30	9.2%	323	20	6.2%	266	10	3.8%
Famhx-Yes	26	16	61.5%	79	14	17.7%	78	13	16.7%	71	15	21.1%
Overall	194	42	21.6%	423	45	10.6%	409	34	8.3%	342	25	7.3%
Estrogen receptor positive breast cancer												
Famhx-No	277	24	8.7%	667	37	5.5%	691	30	4.3%	773	17	2.2%
Famhx-Yes	44	7	15.9%	151	17	11.3%	201	19	9.5%	257	10	3.9%
Overall	337	34	10.1%	846	55	6.5%	908	49	5.4%	1039	28	2.7%

Supplementary Table 12. Age at diagnosis and tumor characteristics in African American women with breast cancer in CARRIERS population-based studies and SEER18 data, 2010-2016

	CARRIERS (N=3916)	SEER (N=75899)
Age at diagnosis (years)		
Mean (SD)	56.4 (12.1)	59.6 (13.2)
Range	22 - 91	16 - 108
Age category (years) (%)		
Missing	16	0
<40	373 (9.6)	5279 (7.0)
40-49	910 (23.3)	14451 (19.0)
50-59	1167 (29.9)	20679 (27.2)
60-69	918 (23.5)	19483 (25.7)
≥70	532 (13.6)	16007 (21.1)
Behavior (%)		
Missing	299	0
Invasive	3009 (83.2)	61084 (80.5)
in situ	608 (16.8)	14815 (19.5)
Hormone receptor status (%)		
Missing	620	17940
negative	861 (26.1)	14476 (25.0)
positive	2435 (73.9)	43483 (75.0)
HER2 status (%)		
Missing	1651	17940
negative	1828 (80.7)	47644 (82.2)
positive	437 (19.3)	10315 (17.8)

REFERENCES

1. Palmer JR, Ruiz-Narvaez EA, Rotimi CN, *et al.* Genetic susceptibility loci for subtypes of breast cancer in an African American population. Cancer Epidemiol Biomarkers Prev 2013;22(1):127-34.
2. Rosenberg L, Adams-Campbell L, Palmer JR. The Black Women's Health Study: a follow-up study for causes and preventions of illness. J Am Med Womens Assoc (1972) 1995;50(2):56-8.
3. Calle EE, Rodriguez C, Jacobs EJ, *et al.* The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. Cancer 2002;94(2):500-11.
4. Patel AV, Jacobs EJ, Dudas DM, *et al.* The American Cancer Society's Cancer Prevention Study 3 (CPS-3): Recruitment, study design, and baseline characteristics. Cancer 2017;123(11):2014-2024.
5. Bernstein L, Allen M, Anton-Culver H, *et al.* High breast cancer incidence rates among California teachers: results from the California Teachers Study (United States). Cancer Causes Control 2002;13(7):625-35.
6. Kolonel LN, Henderson BE, Hankin JH, *et al.* A multiethnic cohort in Hawaii and Los Angeles: baseline characteristics. Am J Epidemiol 2000;151(4):346-57.
7. Pal T, Bonner D, Cragun D, *et al.* A high frequency of BRCA mutations in young black women with breast cancer residing in Florida. Cancer 2015;121(23):4173-80.
8. John EM, Sangaramoorthy M, Koo J, *et al.* Enrollment and biospecimen collection in a multiethnic family cohort: the Northern California site of the Breast Cancer Family Registry. Cancer Causes Control 2019;30(4):395-408.
9. Anton-Culver H, Cohen PF, Gildea ME, *et al.* Characteristics of BRCA1 mutations in a population-based case series of breast and ovarian cancer. Eur J Cancer 2000;36(10):1200-8.
10. Ambrosone CB, Ciupak GL, Bandera EV, *et al.* Conducting Molecular Epidemiological Research in the Age of HIPAA: A Multi-Institutional Case-Control Study of Breast Cancer in African-American and European-American Women. J Oncol 2009;2009:871250.
11. Trentham-Dietz A, Sprague BL, Hampton JM, *et al.* Modification of breast cancer risk according to age and menopausal status: a combined analysis of five population-based case-control studies. Breast Cancer Res Treat 2014;145(1):165-75.
12. Sandler DP, Hodgson ME, Deming-Halverson SL, *et al.* The Sister Study Cohort: Baseline Methods and Participant Characteristics. Environ Health Perspect 2017;125(12):127003.
13. Gail MH, Brinton LA, Byar DP, *et al.* Projecting individualized probabilities of developing breast cancer for white females who are being examined annually. J Natl Cancer Inst 1989;81(24):1879-86.
14. Costantino JP, Gail MH, Pee D, *et al.* Validation studies for models projecting the risk of invasive and total breast cancer incidence. J Natl Cancer Inst 1999;91(18):1541-8.
15. Chatterjee N, Shi J, Garcia-Closas M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. Nat Rev Genet 2016;17(7):392-406.
16. National Cancer Institute, DCCPS, Surveillance Research Program, *et al.* Surveillance, Epidemiology, and End Results (SEER) Program Research Data (1973-2014). In; April 2017 (Based on November 2016 Submission).