nature genetics

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

Germline *BRCA1* mutations predispose to breast cancer. To identify genetic modifiers of this risk, we performed a genome-wide association study in 1,193 individuals with *BRCA1* mutations who were diagnosed with invasive breast cancer under age 40 and 1,190 *BRCA1* carriers without breast cancer diagnosis over age 35. We took forward 96 SNPs for replication in another 5,986 *BRCA1* carriers (2,974 individuals with breast cancer and 3,012 unaffected individuals). Five SNPs on 19p13 were associated with breast cancer risk ($P_{trend} = 2.3 \times 10^{-9}$ to $P_{trend} = 3.9 \times 10^{-7}$), two of which showed independent associations (rs8170, hazard ratio (HR) = 1.26, 95% CI 1.17–1.35; rs2363956 HR = 0.84, 95% CI 0.80–0.89). Genotyping these SNPs in 6,800 population-based breast cancer cases and 6,613 controls identified a similar association with estrogen receptor–negative breast cancer (rs2363956 per-allele odds ratio (OR) = 0.83, 95% CI 0.75–0.92, $P_{trend} = 0.0003$) and an association with estrogen receptor–positive disease in the opposite direction (OR = 1.07, 95% CI 1.01–1.14, $P_{trend} = 0.016$). The five SNPs were also associated with triple-negative breast cancer in a separate study of 2,301 triple-negative cases and 3,949 controls ($P_{trend} = 1 \times 10^{-7}$ to $P_{trend} = 8 \times 10^{-5}$; rs2363956 per-allele OR = 0.80, 95% CI 0.74–0.87, $P_{trend} = 1.1 \times 10^{-7}$).

Pathogenic *BRCA1* and *BRCA2* mutations confer high risks of breast and ovarian cancer. Variation in risk estimates by degree of family history suggests that these risks are modified by other genetic variants[1–5]. Recent studies from the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA) have demonstrated that common breast cancer susceptibility alleles, identified through genome-wide association studies (GWAS) in the general population[6–9], are also associated with the risk of developing breast cancer in *BRCA1* or *BRCA2* mutation carriers[10,11]. However, although five of six alleles were associated with risk of breast cancer for *BRCA2* mutation carriers, only two polymorphisms (in the *TOX3* and 2q35 regions) were associated with risk for *BRCA1* carriers. These findings are consistent with the distinct pathology of breast cancer in *BRCA1* tumors[12,13] and suggest that the genetic variants that modify breast cancer risk for *BRCA1* mutation carriers may differ from the modifiers of risk for *BRCA2* carriers or for non-carriers.

To search for genetic loci associated with breast cancer in *BRCA1* carriers, we conducted a two-stage GWAS. In stage 1, we genotyped 2,500 *BRCA1* carriers using the Illumina Infinium 610K array, which included 620,901 SNPs. Mutation carriers were selected on the basis of an invasive breast cancer diagnosis at under 40 years of age ($n = 1,250$) or the absence of breast cancer when 35 years of age or older ($n = 1,250$). After quality control exclusions, 2,383 carriers (1,193 unaffected and 1,190 affected) from 20 centers in 11 different countries and 555,616 SNPs were available for analysis (**Supplementary Tables 1** and **2**). Genotype associations were evaluated using a 1 degree-of-freedom (d.f.) score test for trend, based on modeling the retrospective likelihood of the observed genotypes conditional on the disease phenotypes, stratified by country of residence. A kinship-adjusted version of the score test statistic was used to allow for the dependence between related individuals.

There was little evidence for inflation in the test statistic of association (inflation factor (λ) = 1.036; **Supplementary Fig. 1**). Ninety-six SNPs were significant at the $P < 10^{-4}$ level compared with 55.6 SNPs which were expected by chance. In stage 2, we genotyped 86 of these SNPs, seven surrogate SNPs (within 10 kb of the significant SNPs and pair-wise $r^2 > 0.90$) and three additional SNPs in 6,332 *BRCA1* carriers. After quality control exclusions, 89 SNPs and 5,986 *BRCA1* mutation carriers (3,012 unaffected and 2,974 affected) were used in the stage 2 analysis. The most significant associations were for five SNPs on 19p13 ($P < 0.002$), which had hazard ratios in the same direction as in stage 1 (**Table 1** and **Supplementary Table 3**). In the combined analysis of stage 1 and 2, there was strong evidence of association[14] with breast cancer for these SNPs ($P = 2.3 \times 10^{-9}$ to $P = 3.9 \times 10^{-7}$).

The minor alleles of rs8170 and rs4808611 were associated with an increased breast cancer risk for *BRCA1* carriers (per allele HR = 1.26, 95% CI 1.17–1.35 for both SNPs). In contrast, SNPs rs8100241, rs2363956 and rs3745185 were associated with decreased breast cancer risk (HR = 0.84, 95% CI 0.80–0.89 for rs8100241 and rs2363956; HR = 0.86, 95% CI 0.81–0.91 for rs3745185) (**Table 1**). The HR estimates for rs8170 and rs4808611 were similar in stages 1 and 2, but for rs8100241, rs2363956 and rs3745185, the HRs were stronger in stage 1; this may be due to the sample selection criteria for stage 1 or a 'winner's curse' effect[15]. There was no evidence of heterogeneity in the HR estimates among the countries of residence in stages 1 and 2 combined (**Fig. 1**; rs8170, $P = 0.10$; rs4808611, $P = 0.14$; rs8100241, $P = 0.18$; rs2363956, $P = 0.17$; and rs3745185, $P = 0.48$).

The strength of the association with breast cancer could also be affected by the inclusion of prevalent cases if these SNPs were

A full list of authors and affiliations appears at the end of the paper.

Received 30 March; accepted 26 August; published online 19 September 2010; doi:10.1038/ng.669

Table 1 Associations with breast cancer risk in *BRCA1* mutation carriers for the five most significant SNPs on 19p13

SNP, position, allele 1/allele 2	Stage	Number		Allele 2 frequency		HR (95% CI)[b]			
		Unaffected[a]	Affected[a]	Unaffected	Affected	Per allele[c]	Heterozygote	Homozygote[d]	P^{e}_{trend}
rs8170	Stage 1	1,193	1,190	0.16	0.20	1.25 (1.12–1.39)	1.23 (1.08–1.41)	1.61 (1.13–2.30)	1.1×10^{-4}
17,250,704	Stage 2	3,010	2,970	0.17	0.20	1.26 (1.15–1.38)	1.28 (1.14–1.43)	1.54 (1.17–2.03)	4.1×10^{-6}
G/A	Combined	4,203	4,160	0.17	0.20	1.26 (1.17–1.35)	1.26 (1.16–1.37)	1.57 (1.26–1.95)	2.3×10^{-9}
rs4808611	Stage 1	1,191	1,190	0.16	0.19	1.26 (1.13–1.41)	1.23 (1.08–1.41)	1.72 (1.21–2.45)	7.9×10^{-5}
17,215,825	Stage 2	3,000	2,964	0.16	0.19	1.26 (1.15–1.39)	1.30 (1.16–1.46)	1.43 (1.06–1.92)	6.4×10^{-6}
G/A	Combined	4,191	4,154	0.16	0.19	1.26 (1.17–1.35)	1.27 (1.17–1.39)	1.53 (1.22–1.93)	2.7×10^{-9}
rs8100241	Stage 1	1,191	1,189	0.53	0.47	0.81 (0.74–0.88)	0.82 (0.71–0.95)	0.65 (0.55–0.77)	1.8×10^{-6}
17,253,894	Stage 2	3,008	2,972	0.51	0.49	0.86 (0.80–0.92)	0.93 (0.82–1.05)	0.74 (0.65–0.85)	1.1×10^{-4}
G/A	Combined	4,199	4,161	0.52	0.48	0.84 (0.80–0.89)	0.88 (0.81–0.97)	0.71 (0.63–0.79)	3.9×10^{-9}
rs2363956	Stage 1	1,193	1,190	0.53	0.47	0.81 (0.74–0.88)	0.82 (0.71–0.95)	0.65 (0.55–0.77)	1.5×10^{-6}
17,255,124	Stage 2	3,006	2,970	0.51	0.49	0.87 (0.81–0.93)	0.92 (0.82–1.04)	0.75 (0.65–0.86)	1.7×10^{-4}
A/C	Combined	4,199	4,160	0.52	0.48	0.84 (0.80–0.89)	0.88 (0.80–0.97)	0.71 (0.64–0.79)	5.5×10^{-9}
rs3745185	Stage 1	1,193	1,190	0.46	0.40	0.83 (0.76–0.90)	0.81 (0.71–0.93)	0.69 (0.57–0.82)	2.3×10^{-5}
17,245,267	Stage 2	3,009	2,972	0.44	0.41	0.88 (0.82–0.95)	0.89 (0.80–1.00)	0.77 (0.67–0.89)	1.2×10^{-3}
G/A	Combined	4,202	4,162	0.44	0.41	0.86 (0.81–0.91)	0.86 (0.81–0.91)	0.74 (0.66–0.83)	3.9×10^{-7}

[a]Affected, unaffected with breast cancer. [b]Estimated hazard ratio and 95% CI. [c]Per copy of allele 2. [d]Two copies of allele 2. [e]Kinship-adjusted score test.

associated with breast cancer survival. To address this possibility, we excluded breast cancer cases diagnosed with the disease >5 years before study entry. The HR estimates were similar to the overall analysis after this exclusion (**Supplementary Table 4**). This indicates that the inclusion of prevalent breast cancer cases was unlikely to have influenced the overall results.

To investigate whether any of these SNPs were associated with ovarian cancer risk for *BRCA1* carriers, we analyzed the data within a competing risks framework and estimated HR simultaneously for breast and ovarian cancer. There was no evidence of association with ovarian cancer risk for any of the SNPs, and the breast cancer associations were virtually identical to the primary analysis both in terms of significance and in the HR estimates (**Table 2**). We repeated the breast cancer association analysis after excluding all individuals who developed ovarian cancer either before or after a breast cancer diagnosis. Despite the sample size reduction, the top four SNPs remained significant at $P < 10^{-7}$ and the HR estimates were identical to the analysis which included individuals with ovarian cancer as unaffected individuals (**Supplementary Table 4**). We also evaluated ovarian cancer associations after excluding individuals with ovarian cancer who were recruited >3 years after their cancer diagnosis in order to account for a potential survival bias. No significant associations were observed after this exclusion ($P_{trend} = 0.44$ to $P_{trend} = 0.96$ using competing risk analysis). We conclude that the associations with breast cancer were not confounded by the competing risk of ovarian cancer.

We evaluated the SNP associations by the predicted functional consequences of *BRCA1* mutation type[16–18]. Class 1 mutations correspond to loss-of-function mutations and are expected to result in a reduced transcript or protein level due to nonsense-mediated RNA decay, whereas class 2 mutations are likely to generate stable proteins with potential residual or dominant negative function[18–20]. Among class 1 mutation carriers (combined stage 1 and 2, $n = 5{,}732$), the five most significant associated SNPs included rs6994019, an intronic SNP in *MMP16* on chromosome 8 ($P_{trend} = 2.9 \times 10^{-6}$) and four SNPs in the 19p13 region ($P_{trend} = 7.6 \times 10^{-6}$ to $P_{trend} = 1.6 \times 10^{-4}$). The *MMP16* SNP rs6994019 was the ninth most significant SNP in the primary analysis of all mutations combined ($P_{trend} = 2.7 \times 10^{-4}$ in stage 1 and 2 combined; **Supplementary Table 3**). The strongest association with breast cancer risk for carriers of class 2 mutations was at the five SNPs in the 19p13 region ($P_{trend} = 1.8 \times 10^{-6}$ to $P_{trend} = 1.2 \times 10^{-4}$; **Supplementary Table 3**). The HR estimates for the five SNPs in 19p13 were larger for class 2 mutations, but the differences between class 1 and class 2 mutations were significant for only rs8170 and rs3745185 ($P = 0.03$ and $P = 0.004$, respectively). These differences might reflect a stronger modifying effect on breast cancer risk for tumors retaining residual or dominant negative BRCA1 function.

Tumor estrogen or progesterone receptor status was available for 1,197 breast cancer cases in stage 1 and 2 combined. A case-only analysis revealed significant differences in the associations for the 19p13 SNPs between estrogen receptor–positive and estrogen receptor–negative disease and between estrogen receptor– or progesterone receptor–positive and estrogen receptor– and progesterone receptor–negative disease, particularly for SNPs rs8100241, rs2363956 and rs3745185 ($P = 0.002$ to $P = 0.04$; **Supplementary Table 5**).

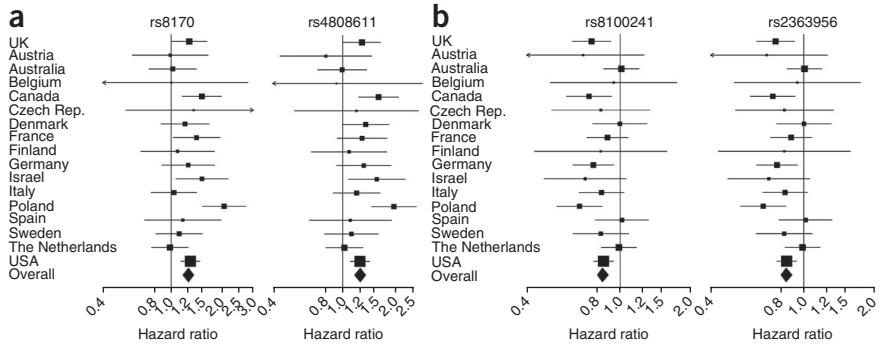

Figure 1 Forest plots of the associations by country of residence of *BRCA1* mutation carriers in the combined stage 1 and stage 2 samples. (**a**,**b**) Squares indicate the country specific per-allele HR estimates for SNPs rs8170, rs4808611 (**a**) and rs8100241, rs2363956 (**b**). The area of the square is proportional to the inverse of the variance of the estimate. Horizontal lines indicate 95% CIs.

Table 2 Competing risk analysis; associations with breast and ovarian cancer risk for *BRCA1* mutation carriers in the combined stage 1 and 2 samples

					Ovarian cancer			Breast cancer		
SNP	Genotype	Unaffected (%)	Breast cancer (%)	Ovarian cancer (%)	HR	95% CI	P^a	HR	95% CI	P^a
rs8170	GG	2,306 (68.4)	2,631 (63.4)	584 (69.3)	1.00			1.00		
	GA	973 (28.9)	1,360 (32.8)	238 (28.2)	1.10	0.92–1.31		1.27	1.17–1.39	
	AA	91 (2.7)	159 (3.8)	21 (2.5)	1.06	0.68–1.66		1.58	1.27–1.97	
	Per allele				1.07	0.93–1.24	0.33	1.27	1.18–1.36	1.5×10^{-10}
rs4808611	GG	2,353 (70.0)	2,696 (65.1)	593 (70.7)	1.00			1.00		
	GA	923 (27.5)	1,307 (31.5)	229 (27.3)	1.14	0.96–1.36		1.29	1.18–1.41	
	AA	86 (2.6)	141 (3.4)	17 (2.0)	0.99	0.58–1.69		1.54	1.22–1.94	
	Per allele				1.10	0.94–1.27	0.34	1.27	1.18–1.37	1.6×10^{-10}
rs8100241	GG	793 (23.6)	1,100 (26.5)	188 (22.4)	1.00			1.00		
	GA	1,676 (49.8)	2,118 (51.0)	428 (50.9)	1.01	0.83–1.23		0.89	0.81–0.98	
	AA	899 (26.7)	933 (22.5)	225 (26.8)	0.89	0.71–1.11		0.70	0.62–0.78	
	Per allele				0.94	0.84–1.05	0.28	0.84	0.79–0.88	1.6×10^{-10}
rs2363956	AA	793 (23.6)	1,100 (26.5)	188 (22.3)	1.00			1.00		
	AC	1,678 (49.8)	2,116 (51.0)	429 (51.0)	1.01	0.83–1.23		0.89	0.80–0.97	
	CC	896 (26.6)	934 (22.5)	225 (26.7)	0.89	0.71–1.12		0.70	0.63–0.78	
	Per allele				0.94	0.85–1.05	0.30	0.84	0.79–0.88	2.4×10^{-10}
rs3745185	GG	1,051 (31.2)	1,423 (34.3)	245 (29.1)	1.00			1.00		
	GA	1,675 (49.7)	2,048 (49.3)	437 (51.8)	1.03	0.85–1.23		0.86	0.79–0.94	
	AA	643 (19.1)	681 (16.4)	161 (19.1)	0.92	0.73–1.15		0.73	0.65–0.82	
	Per allele				0.97	0.86–1.08	0.54	0.86	0.81–0.91	7.1×10^{-8}

[a]Robust Wald statistic.

The OR estimates suggest that these SNPs are more strongly associated with estrogen receptor–negative disease.

The two most significant SNPs (rs8170 and rs4808611) were strongly correlated ($r^2 = 0.87$) in the *BRCA1* samples but displayed a low correlation with the other associated SNPs ($r^2 < 0.23$). rs8100241 and rs2363956 were perfectly correlated ($r^2 = 1$), whereas the least significant SNP, rs3745185, had weaker correlations with both sets of SNPs ($r^2 = 0.17$ and $r^2 = 0.74$ with rs8170 and rs8100241, respectively).

To evaluate the contribution of the 19p13 locus to breast cancer risk in the general population, we genotyped rs8170 and rs2363956 in 6,800 breast cancer cases and 6,613 controls from the SEARCH (Studies of Epidemiology and Risk Factors in Cancer Heredity) study in the UK. Neither SNP was associated with overall breast cancer risk ($P = 0.65$ and $P = 0.79$; **Table 3**). However, stratification of tumors by estrogen receptor status indicated that both SNPs were associated with estrogen receptor–negative breast cancer (rs8170, per-allele OR = 1.21, 95% CI 1.07–1.37, $P = 0.0029$ and rs2363956, OR = 0.83, 95% CI 0.75–0.92, $P = 0.0003$; **Table 3**). These effect sizes were similar to the estimated HRs for *BRCA1* carriers, consistent with the observation that *BRCA1* mutations predispose predominately to estrogen receptor–negative disease. Weaker associations were observed in the opposite direction for estrogen receptor–positive disease (rs8170, per-allele OR = 0.91, 95% CI 0.84–0.98, $P = 0.011$ and rs2363956, OR = 1.07, 95% CI 1.01–1.14, $P = 0.016$). Similar patterns were observed when tumors were stratified by progesterone receptor status or estrogen receptor and progesterone receptor status combined (**Table 3**).

The majority of breast tumors in *BRCA1* carriers exhibit a triple-negative (estrogen receptor, progesterone receptor and HER2 negative) phenotype. To evaluate the association of the 19p13 locus with triple-negative disease in the general population, we obtained genotype data for the five SNPs from up to 2,301 cases from 15 centers in six countries involved in the triple-negative breast cancer consortium (TNBCC). Genotype data from up to 3,949 geographically matched controls were also available (**Supplementary Table 5**). All SNPs were associated with triple-negative breast cancer, and the ORs were comparable to the HRs seen in the *BRCA1* carriers and the ORs for estrogen receptor–negative breast cancer seen in the SEARCH population-based study (rs2363956, per-allele OR = 0.80, 95% CI 0.74–0.87, $P = 1.1 \times 10^{-7}$ and rs8170, OR = 1.28, 95% CI 1.16–1.41, $P = 1.2 \times 10^{-6}$; **Table 3** and **Supplementary Table 5**).

Two of the SNPs (rs8170 and rs2366956) were genotyped in 2,486 *BRCA2* mutation carriers as part of an ongoing GWAS. Neither SNP was associated with breast cancer risk for *BRCA2* carriers ($P_{trend} = 0.17$ and $P_{trend} = 0.07$), but the HR estimates were in line with the ORs estimated for estrogen receptor–positive disease in the SEARCH study (**Table 3**).

All five SNPs were located in a region that spans 39 kb on 19p13 (**Fig. 2**). In an analysis for the joint effect of these SNPs on breast cancer risk for *BRCA1* carriers, it was not possible to distinguish between rs8170 and rs4808611, as neither SNP improved the model fit significantly when the other was included ($P = 0.11$ and $P = 0.22$ for rs8170 and rs4808611, respectively). rs8100241 was retained in preference to rs3745185 (P for inclusion of rs3745185 in model = 0.79). Thus, the most parsimonious model included SNPs rs8170 and either rs8100241 or rs2363956 (P for inclusion = 7.7×10^{-5} and $P = 6.7 \times 10^{-5}$ for rs8170 and rs8100241, respectively) and had a 2 d.f. $P = 6.3 \times 10^{-13}$ for inclusion of both SNPs. This suggests that these associations may be driven by a single causative variant that is partially correlated with all five SNPs. To investigate this further, we evaluated the associations for SNPs identified through the 1000 Genomes Project using imputation. 1,055 SNPs in a 300-kb interval with a minor allele frequency >0.01 in samples of European ancestry, were evaluated. Thirty-one SNPs, none of which were genotyped in stage 1, displayed $P < 1.76 \times 10^{-9}$ (**Fig. 2** and **Supplementary Table 3**). The most significant associations with the imputed genotypes in stage 1 and 2 combined were for eight perfectly correlated SNPs within a 13-kb region (the most significant SNP was rs4808075, $P = 9.4 \times 10^{-12}$; **Supplementary Table 3**). These SNPs were correlated with the four

Table 3 Associations with breast cancer risk in the SEARCH study overall and by tumor subtype, associations with triple negative breast cancer in the TNBCC study and associations with overall breast cancer risk for *BRCA2* mutation carriers

	rs8170				rs2363956			
Study/subtype		Controls (%)	Cases (%)	OR/HR[a] (95% CI)	*P*		Controls (%)	Cases (%)	OR/HR[a] (95%CI)	*P*
SEARCH										
All cases										
	GG	4,288 (65.8)	4,227 (66.5)	1.00		AA	1,628 (24.7)	1,556 (24.3)	1.00	
	GA	1,999 (30.7)	1,885 (29.7)	0.96 (0.89–1.03)		AC	3,261 (49.4)	3,174 (49.7)	1.02 (0.93–1.11)	
	AA	229 (3.5)	241 (3.8)	1.07 (0.89–1.29)		CC	1,714 (26.0)	1,660 (26.0)	1.01 (0.92–1.12)	
	Per allele			0.99 (0.93–1.05)	0.65	Per allele			1.01 (0.96–1.06)	0.79
Estrogen receptor status										
Estrogen receptor positive										
	GG	4,288 (65.8)	2,437 (68.7)	1.00		AA	1,628 (24.7)	817 (22.7)	1.00	
	GA	1,999 (30.7)	988 (27.9)	0.87 (0.79–0.95)		AC	3,261 (49.4)	1,791 (49.8)	1.09 (0.99–1.21)	
	AA	229 (3.5)	123 (3.5)	0.95 (0.75–1.18)		CC	1,714 (26.0)	992 (27.6)	1.15 (1.03–1.29)	
	Per allele			0.91 (0.84–0.98)	0.011	Per allele			1.07 (1.01–1.14)	0.016
Estrogen receptor negative										
	GG	4,288 (65.8)	503 (61.4)	1.00		AA	1,628 (24.7)	240 (28.8)	1.00	
	GA	1,999 (30.7)	272 (33.2)	1.16 (0.99–1.36)		AC	3,261 (49.4)	421 (50.5)	0.88 (0.74–1.04)	
	AA	229 (3.5)	44 (5.4)	1.64 (1.17–2.29)		CC	1,714 (26.0)	172 (20.7)	0.68 (0.55–0.84)	
	Per allele			1.21 (1.07–1.37)	0.0029	Per allele			0.83 (0.75–0.92)	0.0003
	Heterogeneity[b]				2.9×10^{-5}					1.6×10^{-6}
Progesterone receptor status										
Progesterone receptor positive										
	GG	4,288 (65.8)	1,087 (68.1)	1.00		AA	1,628 (24.7)	368 (23.3)	1.00	
	GA	1,999 (30.7)	447 (28.0)	0.88 (0.78–1.00)		AC	3,261 (49.4)	759 (48.0)	1.03 (0.90–1.18)	
	AA	229 (3.5)	62 (3.9)	1.07 (0.80–1.43)		CC	1,714 (26.0)	454 (28.7)	1.17 (1.01–1.37)	
	Per allele			0.94 (0.85–1.04)	0.21	Per allele			1.08 (1.00–1.17)	0.038
Progesterone receptor negative										
	GG	4,288 (65.8)	451 (62.4)	1.00		AA	1,628 (24.7)	199 (27.5)	1.00	
	GA	1,999 (30.7)	237 (32.8)	1.13 (0.95–1.33)		AC	3,261 (49.4)	375 (51.7)	0.94 (0.78–1.13)	
	AA	229 (3.5)	35 (4.8)	1.45 (1.01–2.10)		CC	1,714 (26.0)	151 (20.8)	0.72 (0.58–0.90)	
	Per allele			1.16 (1.01–1.33)	0.031	Per allele			0.85 (0.77–0.95)	0.004
	Heterogeneity[b]				0.0088					0.0002
Estrogen receptor and progesterone receptor status										
Estrogen receptor or progesterone receptor positive										
	GG	4,288 (65.8)	2,515 (68.6)	1.00		AA	1,628 (24.7)	848 (22.8)	1.00	
	GA	1,999 (30.7)	1,019 (27.8)	0.87 (0.79–0.95)		AC	3,261 (49.4)	1,838 (49.5)	1.08 (0.98–1.20)	
	AA	229 (3.5)	130 (3.6)	0.97 (0.78–1.21)		CC	1,714 (26.0)	1,026 (27.6)	1.15 (1.03–1.29)	
	Per allele			0.91 (0.85–0.98)	0.014	Per allele			1.07 (1.01–1.13)	0.017
Estrogen receptor and progesterone receptor negative										
	GG	4,288 (65.8)	280 (59.5)	1.00		AA	1,628 (24.7)	134 (28.3)	1.00	
	GA	1,999 (30.7)	169 (35.9)	1.29 (1.07–1.58)		AC	3,261 (49.4)	256 (54.0)	0.95 (0.77–1.19)	
	AA	229 (3.5)	22 (4.7)	1.47 (0.93–2.32)		CC	1,714 (26.0)	84 (17.7)	0.60 (0.45–0.79)	
	Per allele			1.26 (1.07-1.48)	0.0054	Per allele			0.79 (0.69–0.90)	0.0004
	Heterogeneity[b]				0.0002					8.3×10^{-6}
TNBCC										
Estrogen receptor, progesterone receptor and HER2 negative										
	GG	2,610 (66.2)	1,388 (60.7)	1.00		AA	890 (22.6)	614 (26.9)	1.00	
	GA	1,200 (30.5)	791 (34.6)	1.30 (1.15–1.47)		AC	1,938 (49.3)	1,115 (48.9)	0.83 (0.72–0.95)	
	AA	131 (3.3)	106 (4.6)	1.55 (1.16–2.07)		CC	1,103 (28.1)	550 (24.1)	0.65 (0.55–0.76)	
	Per allele			1.28 (1.16–1.41)	1.2×10^{-6}	Per allele			0.80 (0.74–0.87)	1.1×10^{-7}
BRCA2										
	GG	784 (65.1)	864 (69.5)	1.00		AA	302 (24.9)	297 (23.8)	1.00	
	GA	373 (31.0)	337 (27.1)	0.86 (0.71–1.04)		AC	608 (50.2)	599 (47.9)	1.03 (0.82–1.28)	
	AA	47 (3.9)	43 (3.5)	0.92 (0.58–1.46)		CC	301 (24.9)	354 (28.3)	1.25 (0.98–1.61)	
	Per allele			0.90 (0.77–1.05)	0.17[c]	Per allele			1.12 (0.99–1.27)	0.07[c]

[a]OR estimates for the SEARCH and TNBCC studies and HR estimates for the *BRCA2* associations. [b]Difference in OR between hormone receptor–positive and hormone receptor–negative breast cancer tumors. [c]Based on the kinship-adjusted score test statistic.

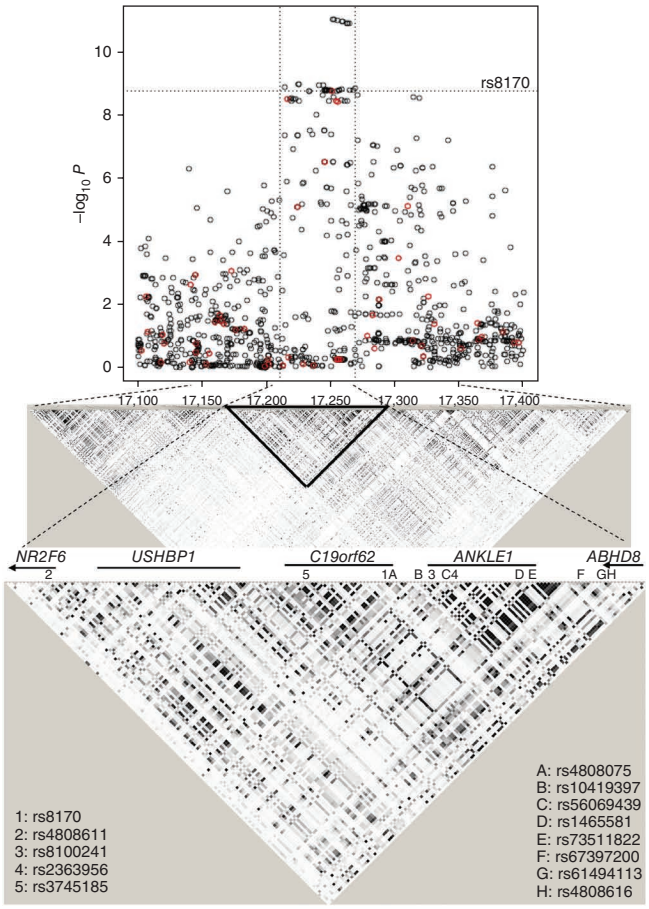

Figure 2 Above, results of the kinship-adjusted score test statistic (1 d.f.) by position (kb) in stage 1 and 2 samples combined for genotyped and imputed SNPs in the associated region (chromosome 19, positions 17,100–17,400 kb). Genotyped SNPs in stages 1 or 2 are shown in red and imputed SNPs are shown in black. The horizontal dotted line indicates the *P* values for the strongest association among genotyped SNPs (rs8170). At middle, the linkage disequilibrium (LD) blocks around the top five associated SNPs (chromosome 19, positions 17,150–17,350 kb) in the combined analysis of stage 1 and stage 2 samples based on the 1000 Genomes Project data for the samples of European ancestry. Squares in the LD blocks indicate pairwise correlations between the SNPs (r^2) by grayscale (darker symbols indicate correlations closer to 1). Below, details of the region containing the most significantly associated genotyped and imputed SNPs (chromosome 19, positions 17,210–17,268 kb). Location of genotyped SNPs shown by numbers 1–5 and the eight most significantly associated imputed SNPs are shown in letters A–H ($P = 9.0 \times 10^{-12}$ to $P = 1.0 \times 10^{-11}$).

genotyped SNPs ($r^2 = 0.37$ to $r^2 = 0.58$ based on the 1000 Genomes Project data; **Supplementary Fig. 2**). This suggests that one or more of these imputed SNPs may be causally associated with breast cancer risk. However, some rare SNPs may have been missed because the 1000 Genome Project data used were based on the resequencing of only 56 individuals. Therefore, the possibility that the association is driven by a rarer variant, or a more cryptic common variant not detected in the resequencing, cannot yet be ruled out.

Of the five genotyped SNPs in the region and the eight most significant imputed SNPs, only rs8170 and rs2363956 were located in coding regions. The smaller 13-kb region, defined by the most strongly associated SNPs, contains three genes: *ABHD8* (encoding abhydrolase domain containing 8), *ANKLE1* (encoding ankyrin repeat and LEM domain containing 1) and *C19orf62*. The eight most significant imputed SNPs were clustered in and around *ANKLE1*, which encodes a protein of undefined function. However, *C19orf62*, which encodes MERIT40 (Mediator of Rap80 Interactions and Targeting 40 kD), is a more plausible genetic modifier of breast cancer in *BRCA1* carriers because MERIT40 interacts with BRCA1 in a protein complex. MERIT40 is a component of the BRCA1 A complex containing BRCA1-BARD1, Abraxas1, RAP80, BRCC36 and BRCC45 that is required for recruitment and retention of the BRCA1-BARD1 ubiquitin ligase and the BRCC36 deubiquitination enzyme at sites of DNA damage[21–23]. Thus, a variant that modifies MERIT40 function or expression might influence BRCA1-dependent DNA repair and checkpoint activity in mammary epithelial cells of *BRCA1* carriers sufficiently, before loss of the wildtype BRCA1 allele, to increase the risk of breast cancer. However, until the SNPs that increase risk of cancer have been definitively linked to MERIT40, it remains possible that the other genes in the region or genes influenced by long range chromatin remodeling or by transcriptional events account for the breast cancer association.

Genetic variation at this locus, in combination with other risk modifiers, may prove useful in individual cancer risk assessment for breast cancer in *BRCA1* carriers. In addition, understanding the functional basis of this association may provide important insights into the etiology of *BRCA1*-associated breast cancer and hormone receptor–negative breast cancer in the general population. Our results suggest that GWAS in *BRCA1* mutation carriers or GWAS restricted to specific breast cancer subtypes may identify further breast cancer susceptibility variants.

URLs. 1000 Genomes Project, http://www.1000genomes.org; MACH software, http://www.sph.umich.edu/csg/yli/mach/index.html/.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

Financial support for this study was provided by the Breast Cancer Research Foundation (BCRF), Susan G. Komen for the Cure and US National Institutes of Health grant CA128978 to F.J.C. and by Cancer Research UK to D.F.E. and A.C.A. A.C.A. is a Cancer Research UK Senior Cancer Research Fellow and D.F.E. is a Cancer Research UK Principal Research Fellow. The authors thank Cancer Genetic Markers of Susceptability (CGEMS) and Wellcome Trust Case Control Consortium (WTCCC) for provision of genotype data from controls. Study specific acknowledgments listed in **Supplementary Note**.

AUTHOR CONTRIBUTIONS

F.J.C., A.C.A. and D.F.E. designed the study and obtained financial support. G.C.-T. founded CIMBA in order to provide the infrastructure for the *BRCA1* GWAS. F.J.C. and X.W. coordinated collection of samples. A.C.A. directed the statistical analysis. D.F.E. advised on the statistical analysis. C.K., Z.S.F. and T.L. carried out analyses. Z.S.F., R.T., J.M., L.M. and D.B. provided bioinformatics and database support. F.J.C., H. Hakonarson and X.W. directed the genotyping of the *BRCA1* carrier and triple-negative samples. M.G. directed the genotyping of the UK case-control samples. A.C.A., F.J.C. and D.F.E. drafted the manuscript. F.J.C. was the overall project leader.

O.M.S. and S.H. coordinated the *BRCA1* mutation classification. T.K., J.V., M.M.G., D.A. and C.G. were involved in the *BRCA2* GWAS genotyping and coordination. K.O. led the *BRCA2* GWAS.

S.P., M.C., C.O., D.F., D.E., D.G.E., R.E., L.I., C.C., F.D., J.P., O.M.S., D.S.-L., C.H., S.M., S.G., C.L., A.R., O.C., A.H., P.B., F.B.L.H., M.A.R., A.J., A.v.d.O., N.H., R.B.v.d.L., H.M.-H., E.B.G.G., P.D., M.P.G.V., J.L., A.J., J.G., T.H., T.B., B.G., C.C., A.B.S., H.H., D.E., E.M.J., J.L.H., M.S., S.S.B., M.B.D., M.-B.T., R.K.S., B.W., C.E., A.M., S.P.-A., N.A., D.N., C.S., S.M.D., K.L.N., T.R., J.L.B., M.P., G.C.R., K.W., J.F.B., J.B., S.V.B., E.F., B.K., Y.L., R.M., I.L.A., G.G., H.O., N.L., K.H., J.R., H.E., A.-M.G., M.T., L.S., P.P., S.M., B.B., A.V., P.R., T.C., M.d.l.H., C.F.S., A.F.-R., M.H.G., P.L.M.,

J.T.L., L.G., N.M.L., T.V.O.H., F.C.N., I.B., C.L., J.G., S.J.R., S.A.G., C.P., S.N., C.I.S., J.B., A.O., H.N., T.H., M.A.C., M.S.B., U.H., A.K.G., M.M., C.C., S.L.N., B.Y.K., N.T., A.E.T., J.W., O.O., J.S., P.S., W.S.R., A.A. and G.R. collected data and samples on *BRCA1* and or *BRCA2* mutation carriers.

N.G.M., G.W.M., J.C.-C., D.F.-J., H.B., G.S., L.B., A.C., S.S.C., P.M., S.M.G., W.T., D.Y., G.F., P.A.F., M.W.B., I.d.S.S., J.P., D.L., R.P., T.R., A.F., R.W., K.P., R.B.D., A.M.L., J.E.-P., C.V., F.B., K.D., A.D. and P.P.D.P. collected data and samples for the TNBCC case-control and/or the SEARCH studies.

All authors provided critical review of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Antoniou, A. *et al.* Average risks of breast and ovarian cancer associated with *BRCA1* or *BRCA2* mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am. J. Hum. Genet.* **72**, 1117–1130 (2003).
2. Antoniou, A.C. *et al.* The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br. J. Cancer* **98**, 1457–1466 (2008).
3. Begg, C.B. *et al.* Variation of breast cancer risk among *BRCA1/2* carriers. *J. Am. Med. Assoc.* **299**, 194–201 (2008).
4. Milne, R.L. *et al.* The average cumulative risks of breast and ovarian cancer for carriers of mutations in *BRCA1* and *BRCA2* attending genetic counseling units in Spain. *Clin. Cancer Res.* **14**, 2861–2869 (2008).
5. Simchoni, S. *et al.* Familial clustering of site-specific cancer risks associated with *BRCA1* and *BRCA2* mutations in the Ashkenazi Jewish population. *Proc. Natl. Acad. Sci. USA* **103**, 3770–3774 (2006).
6. Easton, D.F. *et al.* Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* **447**, 1087–1093 (2007).
7. Hunter, D.J. *et al.* A genome-wide association study identifies alleles in *FGFR2* associated with risk of sporadic postmenopausal breast cancer. *Nat. Genet.* **39**, 870–874 (2007).
8. Stacey, S.N. *et al.* Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor–positive breast cancer. *Nat. Genet.* **39**, 865–869 (2007).
9. Stacey, S.N. *et al.* Common variants on chromosome 5p12 confer susceptibility to estrogen receptor–positive breast cancer. *Nat. Genet.* **40**, 703–706 (2008).
10. Antoniou, A.C. *et al.* Common breast cancer-predisposition alleles are associated with breast cancer risk in *BRCA1* and *BRCA2* mutation carriers. *Am. J. Hum. Genet.* **82**, 937–948 (2008).
11. Antoniou, A.C. *et al.* Common variants in LSP1, 2q35 and 8q24 and breast cancer risk for *BRCA1* and *BRCA2* mutation carriers. *Hum. Mol. Genet.* **18**, 4442–4456 (2009).
12. Lakhani, S.R. *et al.* The pathology of familial breast cancer: predictive value of immunohistochemical markers estrogen receptor, progesterone receptor, HER-2, and p53 in patients with mutations in *BRCA1* and *BRCA2*. *J. Clin. Oncol.* **20**, 2310–2318 (2002).
13. Lakhani, S.R. *et al.* Prediction of *BRCA1* status in patients with breast cancer using estrogen receptor and basal phenotype. *Clin. Cancer Res.* **11**, 5175–5180 (2005).
14. Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661–678 (2007).
15. Zollner, S. & Pritchard, J.K. Overcoming the winner's curse: estimating penetrance parameters from case-control data. *Am. J. Hum. Genet.* **80**, 605–615 (2007).
16. Buisson, M., Anczukow, O., Zetoune, A.B., Ware, M.D. & Mazoyer, S. The 185delAG mutation (c.68_69delAG) in the *BRCA1* gene triggers translation reinitiation at a downstream AUG codon. *Hum. Mutat.* **27**, 1024–1029 (2006).
17. Mazoyer, S. *et al.* A *BRCA1* nonsense mutation causes exon skipping. *Am. J. Hum. Genet.* **62**, 713–715 (1998).
18. Perrin-Vidoz, L., Sinilnikova, O.M., Stoppa-Lyonnet, D., Lenoir, G.M. & Mazoyer, S. The nonsense-mediated mRNA decay pathway triggers degradation of most BRCA1 mRNAs bearing premature termination codons. *Hum. Mol. Genet.* **11**, 2805–2814 (2002).
19. Antoniou, A.C. *et al.* RAD51 135G→C modifies breast cancer risk among *BRCA2* mutation carriers: results from a combined analysis of 19 studies. *Am. J. Hum. Genet.* **81**, 1186–1200 (2007).
20. Liu, H.X., Cartegni, L., Zhang, M.Q. & Krainer, A.R. A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA1* and other genes. *Nat. Genet.* **27**, 55–58 (2001).
21. Feng, L., Huang, J. & Chen, J. MERIT40 facilitates BRCA1 localization and DNA damage repair. *Genes Dev.* **23**, 719–728 (2009).
22. Shao, G. *et al.* MERIT40 controls BRCA1-Rap80 complex integrity and recruitment to DNA double-strand breaks. *Genes Dev.* **23**, 740–754 (2009).
23. Wang, B., Hurov, K., Hofmann, K. & Elledge, S.J. NBA1, a new player in the Brca1 A complex, is required for DNA damage resistance and checkpoint control. *Genes Dev.* **23**, 729–739 (2009).

Antonis C Antoniou[1], Xianshu Wang[2], Zachary S Fredericksen[3], Lesley McGuffog[1], Robert Tarrell[3], Olga M Sinilnikova[4,5], Sue Healey[6], Jonathan Morrison[1], Christiana Kartsonaki[1], Timothy Lesnick[3], Maya Ghoussaini[6], Daniel Barrowdale[1], EMBRACE[1,133], Susan Peock[1], Margaret Cook[1], Clare Oliver[1], Debra Frost[1], Diana Eccles[7], D Gareth Evans[8], Ros Eeles[9], Louise Izatt[10], Carol Chu[11], Fiona Douglas[12], Joan Paterson[13], Dominique Stoppa-Lyonnet[14], Claude Houdayer[14], Sylvie Mazoyer[5], Sophie Giraud[4], Christine Lasset[15,16], Audrey Remenieras[17], Olivier Caron[18], Agnès Hardouin[19], Pascaline Berthet[19], GEMO Study Collaborators[20,133], Frans B L Hogervorst[21], Matti A Rookus[22], Agnes Jager[23], Ans van den Ouweland[24], Nicoline Hoogerbrugge[25], Rob B van der Luijt[26], Hanne Meijers-Heijboer[27], Encarna B Gómez García[28], HEBON[29,133], Peter Devilee[30,31], Maaike P G Vreeswijk[32], Jan Lubinski[33], Anna Jakubowska[33], Jacek Gronwald[33], Tomasz Huzarski[33], Tomasz Byrski[33], Bohdan Górski[33], Cezary Cybulski[33], Amanda B Spurdle[6], Helene Holland[6], kConFab[34,133], David E Goldgar[35], Esther M John[36], John L Hopper[37], Melissa Southey[37], Saundra S Buys[38], Mary B Daly[39], Mary-Beth Terry[40], Rita K Schmutzler[41,42], Barbara Wappenschmidt[41,42], Christoph Engel[43], Alfons Meindl[44], Sabine Preisler-Adams[45], Norbert Arnold[46], Dieter Niederacher[47], Christian Sutter[48], Susan M Domchek[49], Katherine L Nathanson[49], Timothy Rebbeck[49], Joanne L Blum[50], Marion Piedmonte[51], Gustavo C Rodriguez[52], Katie Wakeley[53], John F Boggess[54], Jack Basil[55], Stephanie V Blank[56], Eitan Friedman[57], Bella Kaufman[58], Yael Laitman[57], Roni Milgrom[57], Irene L Andrulis[59–61], Gord Glendon[59], Hilmi Ozcelik[60], Tomas Kirchhoff[62,63], Joseph Vijai[62,63], Mia M Gaudet[64,65], David Altshuler[66], Candace Guiducci[66], SWE-BRCA[67,133], Niklas Loman[68], Katja Harbst[68], Johanna Rantala[69], Hans Ehrencrona[70], Anne-Marie Gerdes[71], Mads Thomassen[72], Lone Sunde[73], Paolo Peterlongo[74,75], Siranoush Manoukian[76], Bernardo Bonanni[77], Alessandra Viel[78], Paolo Radice[74,75], Trinidad Caldes[79], Miguel de la Hoya[77], Christian F Singer[80], Anneliese Fink-Retter[80], Mark H Greene[81], Phuong L Mai[81], Jennifer T Loud[81], Lucia Guidugli[2], Noralane M Lindor[82], Thomas V O Hansen[83], Finn C Nielsen[83], Ignacio Blanco[84], Conxi Lazaro[84], Judy Garber[85], Susan J Ramus[86], Simon A Gayther[86], Catherine Phelan[87], Stephen Narod[88], Csilla I Szabo[2],

MOD SQUAD[89,133], Javier Benitez[90], Ana Osorio[90], Heli Nevanlinna[91], Tuomas Heikkinen[91], Maria A Caligo[92], Mary S Beattie[93–95], Ute Hamann[96], Andrew K Godwin[39], Marco Montagna[97], Cinzia Casella[97], Susan L Neuhausen[98], Beth Y Karlan[99,100], Nadine Tung[101], Amanda E Toland[102], Jeffrey Weitzel[103], Olofunmilayo Olopade[104], Jacques Simard[105,106], Penny Soucy[105], Wendy S Rubinstein[107], Adalgeir Arason[108], Gad Rennert[109], Nicholas G Martin[110], Grant W Montgomery[110], Jenny Chang-Claude[111], Dieter Flesch-Janys[112], Hiltrud Brauch[113,114], GENICA[115,133], Gianluca Severi[116], Laura Baglietto[116], Angela Cox[117], Simon S Cross[118], Penelope Miron[119], Sue M Gerty[7], William Tapper[7], Drakoulis Yannoukakos[120], George Fountzilas[121,122], Peter A Fasching[123], Matthias W Beckmann[124], Isabel dos Santos Silva[125], Julian Peto[125], Diether Lambrechts[126], Robert Paridaens[127], Thomas Rüdiger[128], Asta Försti[96,129], Robert Winqvist[130], Katri Pylkäs[130], Robert B Diasio[131], Adam M Lee[131], Jeanette Eckel-Passow[3], Celine Vachon[3], Fiona Blows[1], Kristy Driver[1], Alison Dunning[1], Paul P D Pharoah[1], Kenneth Offit[61], V Shane Pankratz[3], Hakon Hakonarson[132], Georgia Chenevix-Trench[6], Douglas F Easton[1] & Fergus J Couch[2]

[1]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [2]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA. [3]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA. [4]Unité Mixte de Génétique Constitutionnelle des Cancers Fréquents, Centre Hospitalier Universitaire de Lyon/Centre Léon Bérard, Lyon, France. [5]Equipe labellisée LIGUE 2008, UMR5201 CNRS, Centre Léon Bérard, Université de Lyon, Lyon, France. [6]Queensland Institute of Medical Research, Brisbane, Australia. [7]Wessex Clinical Genetics Service, Princess Anne Hospital, Southampton, UK. [8]Genetic Medicine, Manchester Academic Health Sciences Centre, Central Manchester University Hospitals National Health Service (NHSFT) Foundation Trust, Manchester, UK. [9]Oncogenetics Team, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, Sutton, UK. [10]Clinical Genetics, Guy's and St. Thomas' NHS Foundation Trust, London, UK. [11]Yorkshire Regional Genetics Service, Leeds, UK. [12]Institute of Human Genetics, Centre for Life, Newcastle Upon Tyne Hospitals NHS Trust, Newcastle upon Tyne, UK. [13]Department of Clinical Genetics, East Anglian Regional Genetics Service, Addenbrookes Hospital, Cambridge, UK. [14]INSERM U509, Service de Génétique Oncologique, Institut Curie, Université Paris-Descartes, Paris, France. [15]CNRS UMR5558, Université Lyon 1, Lyon, France. [16]Unité de Prévention et d'Epidémiologie Génétique, Centre Léon Bérard, Lyon, France. [17]Genetics Department, Institut de Cancérologie Gustave Roussy, Villejuif, France. [18]Consultation de Génétique, Département de Médecine, Institut de Cancérologie Gustave Roussy, Villejuif, France. [19]Centre François Baclesse, Caen, France. [20]GEMO study, Cancer Genetics Network 'Groupe Génétique et Cancer', Fédération Nationale des Centres de Lutte Contre le Cancer, Paris, France. [21]Family Cancer Clinic, Netherlands Cancer Institute, Amsterdam, The Netherlands. [22]Department of Epidemiology, Netherlands Cancer Institute, Amsterdam, The Netherlands. [23]Department of Medical Oncology, Rotterdam Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, The Netherlands. [24]Department of Clinical Genetics, Rotterdam Family Cancer Clinic, Erasmus University Medical Center, Rotterdam, The Netherlands. [25]Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [26]Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands. [27]Department of Clinical Genetics, Vrije Universiteit (VU) Medical Center, Amsterdam, The Netherlands. [28]Department of Clinical Genetics, School for Oncology and Developmental Biology, Maastricht University Medical Center, Maastricht, The Netherlands. [29]HEBON, Hereditary Breast and Ovarian Cancer Research Group, The Netherlands. [30]Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands. [31]Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands. [32]Department of Toxicogenetics, Leiden University Medical Center, Leiden, The Netherlands. [33]International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland. [34]kConFab, Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer, Peter MacCallum Cancer Centre, Melbourne, Australia. [35]Department of Dermatology, University of Utah School of Medicine, Salt Lake City, Utah, USA. [36]Cancer Prevention Institute of California, Stanford University School of Medicine, Stanford, California, USA. [37]The University of Melbourne, Melbourne, Australia. [38]Huntsman Cancer Institute, University of Utah Health Sciences Centre, Salt Lake City, Utah, USA. [39]Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA. [40]Columbia University, New York, New York, USA. [41]Centre of Familial Breast and Ovarian Cancer, Department of Obstetrics and Gynaecology, University of Cologne, Cologne, Germany. [42]Centre for Integrated Oncology (CIO), University of Cologne, Cologne, Germany. [43]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany. [44]Department of Obstetrics and Gynaecology, Division of Tumor Genetics, Klinikum rechts der Isar, Technical University Munich, Munich, Germany. [45]Institute of Human Genetics, University of Muenster, Muenster, Germany. [46]Department of Obstetrics and Gynaecology, University Hospital of Schleswig-Holstein, Campus Kiel, Christian-Albrechts University, Kiel, Germany. [47]Department of Obstetrics and Gynaecology, Division of Molecular Genetics, University Hospital Düsseldorf, Heinrich-Heine University Düsseldorf, Dusseldorf, Germany. [48]Institute of Human Genetics, Division of Molecular Diagnostics, University Heidelberg, Heidelberg, Germany. [49]University of Pennsylvania, Philadelphia, Pennsylvania, USA. [50]Baylor-Charles A. Sammons Cancer Center, Dallas, Texas, USA. [51]Gynecology Oncology Group Statistical and Data Center, Roswell Park Cancer Institute, Buffalo, New York, USA. [52]Northshore University Health System, Evanston Northwestern Healthcare, Evanston, Illinois, USA. [53]Tufts University, New England Medical Center, Boston, Massachusetts, USA. [54]University of North Carolina, Chapel Hill, North Carolina, USA. [55]St. Elizabeth Medical Center, Edgewood, Kentucky, USA. [56]New York University School of Medicine, New York, New York, USA. [57]The Susanne Levy Gertner Oncogenetics Unit, Sheba Medical Center, Tel-Hashomer, Israel. [58]Oncology Institute, Sheba Medical Center, Tel-Hashomer, Israel. [59]Ontario Cancer Genetics Network, Cancer Care Ontario, Toronto, Ontario, Canada. [60]Fred A. Litwin Center for Cancer Genetics, Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, Ontario, Canada. [61]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [62]Clinical Genetics Service, Memorial Hospital, New York, New York, USA. [63]Cancer Biology and Genetics Program, Sloan-Kettering Institute, Memorial Sloan-Kettering Cancer Center, New York, New York, USA. [64]Department of Epidemiology and Population Health, Albert Einstein College of Medicine, New York, New York, USA. [65]Department of Obstetrics and Gynecology and Women's Health, Albert Einstein College of Medicine, New York, New York, USA. [66]Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA. [67]SWE-BRCA, Swedish Breast Cancer Study, Sweden. [68]Department of Oncology, Lund University, Lund, Sweden. [69]Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. [70]Department of Genetics and Pathology Rudbeck Laboratory, Uppsala University, Uppsala, Sweden. [71]Department of Clinical Genetics, Rigshospitalet, Copenhagen, Denmark. [72]Department of Clinical Genetics, Odense University Hospital, Odense, Denmark. [73]Department of Clinical Genetics, Aalborg and Aarhus University Hospital, Aarhus, Denmark. [74]Unit of Genetic Susceptibility to Cancer, Department of Experimental Oncology and Molecular Medicine, Fondazione Instituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Istituto Nazionale Tumori (INT), Milan, Italy. [75]IFOM, Fondazione Istituto Fondazione Italiana per la Ricera sul Cancro (FIRC) di Oncologia Molecolare, Milan, Italy. [76]Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori (INT), Milan, Italy. [77]Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia (IEO), Milan, Italy. [78]Division of Experimental Oncology, Centro di Riferimento Oncologico (CRO), IRCCS, Aviano (PN), Italy. [79]Hospital Clinico San Carlos, Madrid, Spain. [80]Department of Obstetrics/Gynocology, Medical University of Vienna, Vienna, Austria. [81]Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA. [82]Medical Genetics, Mayo Clinic Rochester, Rochester, Minnesota, USA. [83]Department of Clinical Biochemistry, Copenhagen University Hospital, Copenhagen, Denmark. [84]Hereditary Cancer Program, Catalan Institute of Oncology, Gran Via de l'Hospitalet, Barcelona, Spain. [85]Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [86]Department of Gynaecological Oncology, University College London, Elizabeth Garrett Anderson (EGA) Institute for Women's Health, London, UK. [87]Women's College Research Institute, Toronto, Ontario, Canada. [88]Department of Epidemiology and Genetics, H. Lee Moffitt Cancer Center, Tampa, Florida, USA. [89]MOD SQUAD, Modifier Study of Quantitative Effects on Disease. [90]Human Cancer Genetics Programme, Spanish National Cancer Research Centre, Madrid, Spain. [91]Helsinki University Central Hospital, Department of Obstetrics and Gynecology, Helsinki, Finland. [92]Division of Surgical, Molecular and Ultrastructural Pathology, Department of Oncology, University of Pisa, Pisa University Hospital, Pisa, Italy. [93]Department of Medicine, University of California, San Francisco, San Francisco, California, USA.

[94]Department of Epidemiology and Biostatistics, University of California, San Francisco, San Francisco, California, USA. [95]Cancer Risk Program, Helen Diller Family Cancer Center, University of California, San Francisco, San Francisco, California, USA. [96]German Cancer Research Center (DKFZ), Heidelberg, Germany. [97]Istituto Oncologico Veneto-IRCCS, Immunology and Molecular Oncology Unit, Padua, Italy. [98]Department of Population Sciences, The Beckman Research Institute of the City of Hope, Duarte, California, USA. [99]Women's Cancer Research Institute at the Samuel Oschin Cancer Institute, Division of Gynecologic Oncology, Cedars-Sinai Medical Center, Los Angeles, California, USA. [100]Department of Obstetrics and Gynecology, David Geffen School of Medicine at the University of California Los Angeles, Los Angeles, California, USA. [101]Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA. [102]Comprehensive Cancer Center, Department of Molecular Virology, Immunology and Medical Genetics, Division of Human Genetics, Department of Internal Medicine, Ohio State University, Columbus, Ohio, USA. [103]Division of Clinical Cancer Genetics, City of Hope Cancer Center, Duarte, California, USA. [104]University of Chicago Medical Center, Chicago, Illinois, USA. [105]Cancer Genomics Laboratory, Centre Hospitalier Universitaire de Quebec, Québec City, Quebec, Canada. [106]Université Laval, Centre de recherche du Centre Hospitalier Universitaire de Quebéc (CHUQ), Québec City, Quebec, Canada. [107]Northshore University Health System, Evanston, Illinois, USA. [108]Department of Pathology, Landspitali University Hospital, Reykjavik, Iceland. [109]QIMR GWAS Collective, Queensland Institute of Medical Research, Brisbane, Australia. [110]Clalit Health Services National Cancer Control Center, Department of Community Medicine and Epidemiology, Carmel Medical Center and B. Rappaport Faculty of Medicine, Technion, Haifa, Israel. [111]Division of Cancer Epidemiology, German Cancer Research Center, Heidelberg, Germany. [112]Institute for Medical Biometrics and Epidemiology, University Clinic Hamburg-Eppendorf, Hamburg, Germany. [113]Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany. [114]University Tübingen, Tübingen, Germany. [115]GENICA, Gene Environment Interaction and Breast Cancer in Germany. [116]Melbourne Collaborative Cohort Study (MCCS), Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Australia. [117]Institute for Cancer Studies, Department of Oncology, Faculty of Medicine, Dentistry and Health, University of Sheffield, Sheffield, UK. [118]Academic Unit of Pathology, Department of Neuroscience, Faculty of Medicine, Dentistry and Health, University of Sheffield, Sheffield, UK. [119]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [120]Molecular Diagnostics Laboratory Institute of Radioisotopes and Radiodiagnostic Products, National Centre for Scientific Research 'Demokritos', Athens, Greece. [121]Department of Medical Oncology, Papageorgiou Hospital, Aristotle University of Thessaloniki School of Medicine, Thessaloniki, Greece. [122]Hellenic Cooperative Oncology Group, Greece. [123]University of California at Los Angeles, David Geffen School of Medicine, Department of Medicine, Division of Hematology and Oncology, Los Angeles, California, USA. [124]University Hospital Erlangen, Department of Gynecology and Obstetrics, Erlangen, Germany. [125]Cancer Research UK Epidemiology and Genetics Group, Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, UK. [126]Vesalius Research Center (VRC), VIB, Leuven, Belgium. [127]Multidisciplinary Breast Center, University Hospitals Leuven, Leuven, Belgium. [128]Institute of Pathology, Städtisches Klinikum Karlsruhe, Karlsruhe, Germany. [129]Center for Primary Health Care Research, University of Lund, Lund, Sweden. [130]Laboratory of Cancer Genetics, Department of Clinical Genetics and Biocenter Oulu, University of Oulu, Oulu University Hospital, Oulu, Finland. [131]Department of Pharmacology, Mayo Clinic, Rochester, Minnesota, USA. [132]Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA. [133]A full list of members is provided in the **Supplementary Note**. Correspondence should be addressed to F.J.C. (couch.fergus@mayo.edu).

ONLINE METHODS

Subjects. *BRCA1 mutation carriers.* Carriers of pathogenic mutations in *BRCA1* used in stages 1 and 2 were drawn from 39 centers from North America, Europe and Australia. Thirty-six centers are participating in the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA). The majority of mutation carriers were recruited through cancer genetics clinics offering genetic testing and were enrolled into national or regional studies. The remainder of the carriers were identified by population-based sampling of cases or by community recruitment. Eligibility to participate was restricted to female carriers of pathogenic *BRCA1* mutations who were 18-years-old or older at recruitment and who were of 'white' self-reported ancestry. Information collected included the year of birth; mutation description, including nucleotide position and base change; age at last follow up; ages at breast and ovarian cancer diagnosis; and age or date at bilateral prophylactic mastectomy. Information was also available on the country of residence, defined as the country of the clinic at which the carrier family was recruited to the study. Related individuals were identified through a unique family identifier. Carriers of unclassified variants of uncertain clinical significance were excluded. All carriers participated in clinical or research studies at the host institutions under ethically approved protocols.

Stage 1 sample selection. To improve power, stage 1 of the GWAS included *BRCA1* mutation carriers diagnosed with breast cancer at a young age and older mutation carriers who had not developed the disease. We selected 1,250 carriers diagnosed with invasive breast cancer when younger than 40-years-old and 1,250 carriers who had not developed breast cancer or who had developed a first ovarian cancer when 35 years of age or older. Related pairs of individuals were eligible for inclusion in stage 1 if they were discordant with respect to their breast cancer disease status (211 pairs). Stage 1 mutation carriers resided in eleven different countries (**Supplementary Table 2**).

Stage 2 samples. We selected stage 2 samples from the remaining *BRCA1* mutation carriers from each participating center. Individuals who developed ductal carcinoma *in situ* (DCIS) were not eligible for inclusion in either stage 1 or stage 2. Carriers from an additional six countries were included in stage 2 (**Supplementary Table 2**).

Population-based breast cancer cases and controls. Breast cancer cases were drawn from the Studies in Epidemiology and Risks of Cancer Heredity (SEARCH), an ongoing population based study of cancer patients ascertained through the East Anglian Cancer Registry in the UK. Eligible cases were those diagnosed with breast cancer under age 55 years between the years 1991 and 1996 and were still alive in 1996, and all incident breast cancer cases diagnosed under age 70 years between 1996 and 2006. Controls from the same geographical region were drawn at random from the European Prospective Investigation of Cancer (EPIC-Norfolk) cohort and from general practices recruiting patients to SEARCH. The majority of cases and controls were self reported as 'white' (>98% of the population) and were broadly of the same age. A more detailed description of these studies has been previously described[24]. Both of these studies were approved by local ethical committees and participants gave written informed consent.

Triple-negative breast cancer cases and controls. Five centers contributed triple-negative breast cancer cases and controls, and ten centers contributed triple-negative breast cancer cases as part of the Triple-Negative Breast Cancer Consortium (TNBCC). Eligible subjects were female individuals with breast cancer of 'white' self-reported ancestry. A triple-negative breast cancer case was defined as an individual with an estrogen receptor–negative, progesterone receptor–negative and HER2–negative (0 or 1 by immunohistochemical staining or 2+ by immunohistochemical and FISH negative) breast cancer diagnosed after age 18. The studies were approved by local ethical committees. A total of 2,301 cases from TNBCC passed genotyping quality control checks (**Supplementary Table 1**). Controls were unaffected women contributed by each center or genotyped in other GWAS including the Wellcome Trust Case Control Consortium UK 1958 Birth Cohort[14] (n = 1,421), the Cancer Genetic Markers of Susceptibility[7] (CGEMS, n = 1,142), Kooperative Gesundheitsforschung in der Region Augsburg (KORA) (n = 215) and the Australian Twin Cohort[25] (n = 659) (**Supplementary Table 2**). All available controls were used; a categorical variable defined by center (for centers with both cases and controls) or country of residence (for centers contributing cases only combined with controls from other GWAS) was included as an adjustment variable in the primary analysis.

BRCA2 mutation carriers. Carriers of pathogenic mutations in *BRCA2* were drawn from an ongoing two-stage GWAS of genetic modifiers for *BRCA2* mutation carriers. Individuals were recruited through 33 studies, which were largely the same as the studies that contributed to the *BRCA1* GWAS and had similar eligibility criteria. Genotype data for the SNPs under investigation in this report were available only for the stage 2 samples of the *BRCA2* GWAS (2,486 individuals in total).

Genotyping and quality control. *BRCA1 mutation carrier samples.* All *BRCA1* DNA samples were assessed for quality and concentration by PicoGreen analysis and by visualization after electrophoresis in E-Gel (Invitrogen) agarose gels. Samples selected for stage 1 were genotyped at The Center for Applied Genomics, Children's Hospital of Philadelphia, using the Illumina Infinium 610K array. Of the 2,500 samples genotyped, 55 failed to produce genotyping data and were excluded from all further assessment. Of the 620,901 markers genotyped, 22,545 were excluded from subsequent quality control analysis because of low quality genotyping data (SNP call rate <90%), were Y-chromosome SNPs or were copy number variants. An iterative quality control process was applied to the remaining samples to exclude samples with low call rates (<99%), excess heterozygosity, sex errors and sample duplications, and to exclude SNPs with call rates <95%, minor allele frequencies <0.01, minor allele frequencies between 0.01 and 0.05 and call rate <0.98, or Hardy-Weinberg equilibrium $P < 10^{-7}$) in the entire sample (**Supplementary Table 1**). We further excluded individuals of non-European ancestry using multi-dimensional scaling. For this purpose we selected 37,804 autosomal SNPs that were not strongly correlated (pair-wise $r^2 < 0.10$). These SNPs were used to compute the genomic kinship between all pairs of *BRCA1* mutation carriers in our sample, along with 210 HapMap samples (from the CHB, YRI and CEU populations). These were converted to distances (by subtracting from 0.5) and subjected to multidimensional scaling. **Supplementary Figure 1** shows a graphical representation of the first two components for all *BRCA1* mutation carriers and for the subgroups defined by the common Ala185delGly *BRCA1* Jewish founder mutation and the 5382insCys Eastern European founder mutation. Using the first two components, we calculated the proportion of European ancestry for each individual to be:

$$\frac{(x - x_2)(y_1 - y_2) - (y - y_2)(x_1 - x_2)}{(x_3 - x_2)(y_1 - y_2) - (y_3 - y_2)(x_1 - x_2)}$$

where x is the first principal component coordinate, y is the second component coordinate, and x_i and y_i represent the mean coordinates among the CHB (i = 1), YRI (i = 2) and CEU (i = 3) HapMap samples. Individuals with >15% non-European ancestry were excluded from the analysis.

Stage 2. In stage 2, 6,332 BRCA1 mutation carriers were genotyped for 96 SNPs. These included 86 SNPs with $P < 10^{-4}$ in stage 1 with Illumina design scores suggestive of assay conversion. Seven other SNPs with low design scores were replaced with SNPs within 10 kb and which had $r^2 > 0.90$. One SNP within 60 base pairs of another candidate SNP and two SNPs with low design scores and no SNPs with $r^2 > 0.75$ were replaced by the next most significant candidate SNPs from stage 1. Genotyping was performed using the Illumina VeraCode platform in the Genotyping Shared Resource from the Mayo Clinic. Of the samples genotyped, 118 failed genotyping, 50 had call rates of <0.95 and 152 were sample duplications. Another 26 samples were excluded because they could not be allocated in one of the strata used in the analysis based on their reported country of residence, thus leaving 5,986 samples for inclusion in the stage 2 analysis.

Genotyping. *Non-BRCA1 samples.* DNA samples from the SEARCH breast cancer subjects and regional controls were genotyped for SNPs rs8170 and rs2363956 using the 5′exonuclease assay (TaqMan) using the ABI Prism 7900HT sequence detection system. *BRCA2* mutation carriers were genotyped using the Sequenom

iPlex platform. A total of 1,385 TNBCC cases were genotyped for the five SNPs as part of an ongoing GWAS using the Illumina Infinium 660K array. Another 916 TNBCC cases were genotyped for rs8170, rs2363956 and rs4808611 by iPlex (Sequenom). The QIMR and CGEMS controls were genotyped using the Illumina Infinium 550K array, and the WTCCC controls were genotyped using a custom Illumina Infinium 1.2M array. More details of the genotyping in the control samples have been previously described[1–3]. Only observed genotypes for the five SNPs on 19p13 were used in the analyses.

Statistical methods. *Analysis of genotype data from BRCA1 carriers*. The main analyses were focused on the evaluation of associations between each genotype and breast cancer risk. The phenotype of each individual was defined by age at diagnosis of breast cancer or age at last follow up. For this purpose, individuals were censored at the age of the first breast cancer diagnosis, ovarian cancer diagnosis or bilateral prophylactic mastectomy, whichever occurred first, or at the age at last observation. Mutation carriers censored at ovarian cancer diagnosis were considered to be unaffected controls. Analyses were carried out within a survival analysis framework. Because mutation carriers were not sampled randomly with respect to their disease status, standard methods of survival analysis (such as Cox regression) may have led to biased estimates[26]. We therefore conducted the analysis by modeling the retrospective likelihood of the observed genotypes conditional on the disease phenotypes. The associations between genotype and breast cancer risk at both stages were assessed using the 1 d.f. score test statistic based on this retrospective likelihood as described previously[19]. This statistic has the form

$$U = \sum_{k=1}^{m} (g_k - \bar{g})(y_k - \bar{y})$$

where g_k is the genotype of individual k and $y_k = O_k - \Lambda_o(t_k)$, where $O_k = 1$ if the individual is affected and $O_k = 0$ if the individual is unaffected and m is the number of individuals in the sample. $\Lambda_o(t_k)$ is the *BRCA1* cumulative incidence rate to age t_k, which was obtained from previously published data[2]. To allow for the non-independence among related individuals, an adjusted version of the score test was used, in which the variance of the score was derived by taking into account the correlation between the genotypes. The variance of the score is given by:

$$Var(U) = \sum_{k=1}^{m}\sum_{l=1}^{m}(y_k - \bar{y})(y_l - \bar{y})2\phi_{kl}\,\mathrm{var}(\underline{g})$$

where ϕ_{kl} is the kinship coefficient for the pair of individuals k and l, which can be estimated from the available genomic data[27,28]. A robust variance approach using a Huber-White sandwich estimator, based on reported family membership, was also used for comparison purposes[29]. We chose to present the *P* values based on the kinship-adjusted score test as it utilizes the degree of relationship between individuals in our sample. *P* values based on the robust score test statistic are also quoted in **Supplementary Table 3a**. In practice, the differences in significance levels between the two methods were slight (with the kinship-adjusted *P* values being slightly less significant for the key SNPs). These analyses were performed in R using the GenABEL[30] and SNPMatrix[31] libraries, as well as custom-written functions.

To estimate the magnitude of the associations, the effect of each SNP was modeled either as a per-allele HR (multiplicative model) or as separate HRs for heterozygotes and homozygotes, and these effects were estimated on the log scale. The HRs were assumed to be independent of age (that is, we used a Cox proportional-hazards model). The retrospective likelihood was modeled in the pedigree-analysis software MENDEL[32] as previously described[19]. As sample sizes varied substantially between studies or contributing centers (including some with small numbers) heterogeneity was examined at the country level. This was assessed by comparing models that allowed for country-specific log-hazard ratios against models in which the same log-hazard ratio was assumed to apply to all countries. All analyses were stratified by country of residence and used calendar year and cohort-specific breast cancer incidence rates for *BRCA1* (ref. 2). The combined stage 1 and stage 2 analysis was also stratified by stage. Similar methods were used to assess the associations with breast cancer risk for *BRCA2* mutation carriers.

To evaluate the combined effects of the significant SNPs on breast cancer risk, we fitted for each pair of SNPs retrospective likelihood models of the form

$$\lambda(t) = \lambda_0(t)\exp(\beta_1 x_1 + \beta_2 x_2)$$

where β_1 is the per-allele log-hazard ratio for SNP 1, β_2 is the per-allele log-hazard ratio for SNP 2, and x_1 and x_2 represent the number of minor alleles at locus 1 and 2 respectively (for example, 0, 1 or 2), while at the same time allowing for linkage disequilibrium between the loci. To test whether the fit of the model was significantly improved by the inclusion of a locus into the model, we tested for the significance of parameters β_1 and β_2.

Competing risk analysis. In the competing risk analysis, we estimated HR simultaneously for breast and ovarian cancer. For this purpose, we extended the retrospective likelihood model[4] so that each individual was at risk of developing either breast or ovarian cancer by assuming that the probabilities of developing each disease were independent and were conditional on the underlying genotype. A different censoring process was used in this case, whereby individuals were followed up to the age of the first breast or ovarian cancer diagnosis and were considered to have then developed the corresponding disease. No follow up was considered after the first cancer diagnosis. Individuals were censored for breast cancer at the age of bilateral prophylactic mastectomy and for ovarian cancer at the age of bilateral oophorectomy and were assumed to be unaffected for the corresponding disease. The remaining individuals were censored at the age at last observation and were assumed to be unaffected for both diseases.

Associations by tumor characters in SEARCH and TNBCC studies. Associations in SEARCH were evaluated using logistic regression by estimating both genotype-specific ORs and per-allele ORs. Differences in the associations between different groups of individuals with breast cancer defined by the breast cancer tumor characteristics (both in *BRCA1* mutation carriers and case-control study) were investigated using a case-only analysis by logistic regression. For this purpose, the SNPs were assumed to be the independent variable and the tumor characteristic was the outcome variable. Analyses were performed for estrogen receptor, progesterone receptor and for estrogen receptor and progesterone receptor combined. The number of *BRCA1* mutation carriers with HER2 data was too small to assess the associations reliably. All analyses in the *BRCA1* cohort were adjusted for country of residence and the age at diagnosis of the subject. Associations for triple-negative breast cancer in TNBCC were also estimated using logistic regression, adjusted for strata defined by center or country of residence. Homogeneity of the per-allele OR among strata was tested using the *Q* statistic.

Imputation. We imputed genotype data for all the SNPs identified in CEU individuals in the 1000 Genomes Project (see URLs) with minor allele frequency >0.01 in the region located at 17,100–17,400 kb on chromosome 19 (build 36). The MACH software (see URLs) was used to impute non-genotyped SNPs for stage 1 and stage 2 samples based on the phased haplotypes from the 1000 Genomes Project (pilot 1, data release 07/05/09) for the samples of European ancestry (CEU). The imputation process included 2,383 samples from stage 1 and 5,986 samples from stage 2. For this purpose, genotype data on a total of 59 SNPs in the region were available for the stage 1 samples and 5 SNPs were available for the stage 2 samples (rs4808611, rs3745185, rs8170, rs8100241 and rs2363956). The imputation was conducted using the stage 1 and stage 2 samples combined using the 1000 Genomes Project data as the reference panel. Associations between each marker and breast cancer risk were then assessed using a similar score test to that used for the observed SNPs, but these assessments were based on the posterior genotype probabilities at each imputed marker for each individual.

24. Thompson, D.J. *et al.* Identification of common variants in the *SHBG* gene affecting sex hormone-binding globulin levels and breast cancer risk in postmenopausal women. *Cancer Epidemiol. Biomarkers Prev.* **17**, 3490–3498 (2008).
25. Medland, S.E. *et al.* Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am. J. Hum. Genet.* **85**, 750–755 (2009).
26. Antoniou, A.C. *et al.* A weighted cohort approach for analysing factors modifying disease risks in carriers of high-risk susceptibility genes. *Genet. Epidemiol.* **29**, 1–11 (2005).

doi:10.1038/ng.669

27. Amin, N., van Duijn, C.M. & Aulchenko, Y.S. A genomic background based method for association analysis in related individuals. *PLoS ONE* **2**, e1274 (2007).
28. Leutenegger, A.L. *et al.* Estimation of the inbreeding coefficient through use of genomic data. *Am. J. Hum. Genet.* **73**, 516–523 (2003).
29. Boos, D.D. On generalised score tests. *Am. Stat.* **46**, 327–333 (1992).
30. Aulchenko, Y.S., Ripke, S., Isaacs, A. & van Duijn, C.M. GenABEL: an R library for genome-wide association analysis. *Bioinformatics* **23**, 1294–1296 (2007).
31. Clayton, D. & Leung, H.T. An R package for analysis of whole-genome association studies. *Hum. Hered.* **64**, 45–51 (2007).
32. Lange, K., Weeks, D. & Boehnke, M. Programs for pedigree analysis: MENDEL, FISHER, and dGENE. *Genet. Epidemiol.* **5**, 471–472 (1988).