nature genetics

Identification of ten variants associated with risk of estrogen-receptor-negative breast cancer

Most common breast cancer susceptibility variants have been identified through genome-wide association studies (GWAS) of predominantly estrogen receptor (ER)-positive disease[1]. We conducted a GWAS using 21,468 ER-negative cases and 100,594 controls combined with 18,908 *BRCA1* mutation carriers (9,414 with breast cancer), all of European origin. We identified independent associations at $P < 5 \times 10^{-8}$ with ten variants at nine new loci. At $P < 0.05$, we replicated associations with 10 of 11 variants previously reported in ER-negative disease or *BRCA1* mutation carrier GWAS and observed consistent associations with ER-negative disease for 105 susceptibility variants identified by other studies. These 125 variants explain approximately 16% of the familial risk of this breast cancer subtype. There was high genetic correlation (0.72) between risk of ER-negative breast cancer and breast cancer risk for *BRCA1* mutation carriers. These findings may lead to improved risk prediction and inform further fine-mapping and functional work to better understand the biological basis of ER-negative breast cancer.

GWAS have identified 107 SNPs that are independently associated with breast cancer risk[2–32]. Association studies focused on ER-negative disease, or *BRCA1* mutation carriers, who are more likely to develop ER-negative disease (70–80% of cases)[33], have identified 11 of these SNPs[3,9,12,19,29,30]. We aimed to discover additional susceptibility variants for ER-negative breast cancer by performing a GWAS in women of European origin.

New genotyping data were generated for 9,655 ER-negative cases and 45,494 controls from 68 Breast Cancer Association Consortium (BCAC) studies and 15,566 *BRCA1* mutation carriers (7,784 with breast cancer) from 58 Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA) studies (**Supplementary Tables 1** and **2**) using the Illumina OncoArray BeadChip, a 570,000-SNP custom array with genome-wide coverage[34]. Imputation was used to derive estimated genotypes for ~21 million SNPs, using the 1000 Genomes Project (Phase 3) as the reference; ~11.5 million of these SNPs with imputation r^2 >0.3 and minor allele frequency (MAF) >0.005 were included in further analyses. For BCAC data, we estimated per-allele odds ratios (ORs) using logistic regression, adjusting for country and principal components. For CIMBA data, we estimated per-allele hazard ratios (HRs) using a retrospective cohort analysis framework, modeling time to breast cancer and stratifying by country, Ashkenazi Jewish origin and birth cohort[35,36] (Online Methods). These analyses were also applied to an independent set of previously generated data from other genome-wide genotyping of additional European participants in 44 BCAC studies (11,813 ER-negative cases and 55,100 controls)[9,12,16,20,37,38] and 54 CIMBA studies (3,342 *BRCA1* mutation carriers, 1,630 with breast cancer) (**Supplementary Tables 1** and **2**). Fixed-effects meta-analysis was used to combine results across genotyping initiatives within consortia and, assuming that the odds ratio and hazard ratio estimates approximate the same underlying relative risk, across consortia[39].

Results from the combined meta-analysis are summarized in **Supplementary Figures 1** and **2**. There was minimal inflation of test statistics ($\lambda_{1,000} = 1.004$; **Supplementary Fig. 3**). We identified ten variants at nine new loci that were independently associated with risk of ER-negative breast cancer at $P < 5 \times 10^{-8}$ (**Table 1**, **Supplementary Figs. 4–11** and **Supplementary Table 3**). Two independent signals were observed within 12 kb of each other at 11q22.3, for rs74911261 (MAF = 0.02) and rs11374964 (MAF = 0.42); odds ratio estimates and statistical significance were largely unchanged when each variant was adjusted for the other (**Supplementary Table 4**). The association with rs66823261 at 8p23.3 was not observed for *BRCA1* mutation carriers ($P = 0.32$, P value for heterogeneity (P_{het}) = 0.030).

For each of these ten new signals, we identified candidate causal SNPs analytically[40,41] (Online Methods) and combined multiple sources of *in silico* functional annotation from public databases[42–52] to identify likely functional variants and target genes. Results are summarized in **Figure 1**, **Supplementary Table 5** (including UCSC Genome Browser links; also see the **Supplementary Note**) and **Supplementary Figures 4–11** (data sources in **Supplementary Table 6**). Many candidate causal SNPs lie in predicted regulatory regions and are associated with expression of nearby genes in blood or other tissues. At 2p23, the predicted target genes include *ADCY3* and *NCOA1* (**Supplementary Fig. 4**). At 6q23.1 (**Supplementary Fig. 5**), the most plausible target gene is *L3MBTL3* (ref. 53). A predicted target at 8q24.13 is *FBXO32*, which is expressed in ER-negative human mammary epithelial cells (HMECs) but not ER-positive MCF7 breast cancer cells (**Supplementary Fig. 7**) and has a known role in cancer cachexia[54]. At 11q22.3 (**Fig. 1**), a predicted target gene of common risk-associated variants is *NPAT*[55]. The rarer SNPs underlying the other 11q22.3 signal are predicted to target *ATM*, a known breast cancer susceptibility gene[56]. Three rare coding variants (MAF ≤ 0.03) in *ATM*, *NPAT* and *KDELC2* are also among the candidate causal

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

Received 30 May 2016; accepted 11 January 2017; published online 23 October 2017; doi:10.1038/ng.3785

Table 1 Ten new loci associated with risk of estrogen-receptor-negative breast cancer using meta-analysis of BCAC and CIMBA data

						BCAC ER-negative disease[b]			CIMBA *BRCA1* mutation carriers[c]				Meta-analysis	Heterogeneity
Location	SNP	Chr.	Position (bp)	Nearest gene	Alleles[a]	MAF	OR	95% CI	*P* value	MAF	HR	95% CI	*P* value	*P* value	*P* value[d]
2p23.3	rs200648189	2	24,739,694	*NCOA1*	CT/C	0.19	0.94	0.91–0.97	4.7×10^{-4}	0.20	0.88	0.84–0.92	3.3×10^{-7}	9.7×10^{-9}	2.0×10^{-2}
6q23.1	rs6569648	6	130,349,119	*L3MBTL3*	T/C	0.23	0.93	0.90–0.95	4.3×10^{-8}	0.22	0.94	0.90–0.98	5.4×10^{-3}	8.3×10^{-10}	0.64
8p23.3	rs66823261	8	170,692	*RPL23AP53*	T/C	0.23	1.09	1.06–1.12	5.6×10^{-9}	0.22	1.02	0.98–1.07	0.32	3.3×10^{-8}	3.0×10^{-2}
8q24.13	rs17350191	8	124,757,661	*ANXA13*	C/T	0.34	1.07	1.04–1.09	2.0×10^{-8}	0.34	1.08	1.04–1.12	1.9×10^{-4}	1.7×10^{-11}	0.81
11q22.3	rs11374964	11	108,345,515	*KDELC2*	G/GA	0.42	0.94	0.92–0.96	3.6×10^{-8}	0.43	0.91	0.88–0.95	1.3×10^{-6}	4.1×10^{-13}	0.26
11q22.3	rs74911261	11	108,357,137	*KDELC2*	G/A	0.02	0.82	0.75–0.89	2.3×10^{-6}	0.02	0.74	0.65–0.84	2.0×10^{-6}	5.4×10^{-11}	0.17
16p13.3	rs11076805	16	4,106,788	*ADCY9*	C/A	0.25	0.92	0.90–0.95	2.2×10^{-8}	0.25	0.96	0.92–1.00	0.073	1.4×10^{-8}	0.14
18q12.1	rs36194942	18	25,401,204	*CDH2*	A/AT	0.30	0.94	0.91–0.96	2.5×10^{-7}	0.31	0.95	0.91–0.99	1.4×10^{-2}	1.4×10^{-8}	0.50
19p13.2	rs322144	19	11,423,703	*TSPAN16*	C/G	0.47	0.95	0.93–0.97	2.4×10^{-5}	0.46	0.92	0.89–0.96	3.7×10^{-5}	7.4×10^{-9}	0.23
19q12	rs113701136	19	30,277,729	*CCNE1*	C/T	0.32	1.07	1.04–1.09	1.7×10^{-7}	0.32	1.05	1.01–1.09	1.2×10^{-2}	6.8×10^{-9}	0.57

Chr., chromosome; MAF, minor allele frequency; OR, odds ratio per copy of the minor allele; CI, confidence interval; HR, hazard ratio per copy of the minor allele.

[a]More common allele listed first, minor allele listed second. [b]Combined data from 21,468 ER-negative cases and 100,594 controls of European ancestry from the Breast Cancer Association Consortium (BCAC). [c]Combined data from 18,908 *BRCA1* mutation carriers from the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA), 9,414 of whom had developed breast cancer. [d]Test for heterogeneity in effect size for ER-negative disease and overall disease for *BRCA1* mutation carriers.

SNPs at this locus. At 16p13, predicted target genes include *ADCY9* and *CREBBP* (**Supplementary Fig. 8**). At 19q12 (**Supplementary Fig. 11**), a potential target gene (*CCNE1*) encodes cyclin E1, which is involved in cell cycle control and phosphorylation of NPAT[57].

Expression quantitative trait locus (eQTL) associations were assessed for each candidate causal variant and genes within 1 Mb using 79 ER-negative breast tumors from The Cancer Genome Atlas (TCGA) and 135 normal breast tissue samples from the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC)[58–60]. The strongest associations identified were for rs6569648 at 6q23.1 with *L3MBTL3* ($P = 4.3 \times 10^{-6}$) and for rs12965632 at 18q12.1 with *CDH2* ($P = 1.0 \times 10^{-4}$), both in METABRIC (**Supplementary Table 5**). SNP rs6569648 was the top *cis*-eQTL (of all imputed variants within 1 Mb) for *L3MBTL3*, while the *P* value for the eQTL effect of rs12965632 on *CDH2* was within two orders of magnitude of those for the top *cis*-eQTLs for this gene (**Supplementary Figs. 12** and **13**).

For 10 of the 11 variants previously identified through GWAS of ER-negative disease or overall disease for *BRCA1* mutation carriers[3,9,12,18,19,30,31] or reported as more strongly associated with ER-negative breast cancer[29], associations with ER-negative disease were replicated ($P < 0.05$) using OncoArray data from BCAC, which do not overlap with any of the discovery studies (**Table 2**). Effect sizes were generally similar to those originally reported. Using all available CIMBA data, 6 of these 11 variants were associated with breast cancer risk ($P < 0.05$) for *BRCA1* mutation carriers (**Table 2**). No evidence of association was observed for rs2284378 at 20q11 (ref. 12) in either BCAC or CIMBA ($P \geq 0.46$).

On the basis of odds ratios estimated using BCAC data for all cases with known ER status (16,988 ER negative and 65,275 ER positive), all ten new and ten previously reported and replicated susceptibility SNPs for ER-negative disease were more strongly associated with risk of the ER-negative subtype than the ER-positive subtype ($P_{het} < 0.05$, except for the new signal for rs322144 at 19p13.2; **Supplementary Table 7**). Two variants (rs4245739 at 1q32.1 and rs67397200 at 19p13.11) were not associated with ER-positive disease. For four variants (rs11374964 (11q22.3), rs74911261 (11q22.3), rs6678914 (1q32.1) and rs4577244 (2p23.2)), the risk-associated allele for ER-negative disease was associated with reduced risk of ER-positive disease ($P < 0.05$).

For these 20 susceptibility SNPs for ER-negative breast cancer, we also assessed association by triple-negative status (negative for ER, progesterone receptor and HER2; **Table 3**), tumor grade (**Table 4**) and age at diagnosis (**Supplementary Table 8**) using BCAC data only. Five SNPs, including the new susceptibility variants at 11q22.3 (rs11374964 and rs74911261), were more strongly associated with both risk of triple-negative disease and risk of higher-grade disease ($P < 0.05$); however, after adjustment for triple-negative status, heterogeneity in effect by tumor grade was observed only for the variants at 11q22.3 (rs74911261) and 1q32.1 (rs4245739) ($P < 0.05$). For rs4577244 at 2p23.3, heterogeneity was observed for tumor grade only, while rs2747652 at 6q25.2 was more strongly associated with risk of other (non-triple-negative) ER-negative breast cancer subtypes ($P < 0.05$). At younger ages at diagnosis, associations appeared to be stronger for two variants (rs10069690 (5p15.33) and rs67397200 (19p13.11)) and weaker for one (rs2747652 (6q25.2)) ($P < 0.05$).

Elsewhere, we report 65 new susceptibility loci for overall breast cancer[1]. Three of these are located within 500 kb of the new susceptibility loci for ER-negative disease reported here (variants rs200648189 (2p23.3), rs6569648 (6q23.1) and rs17350191 (8q24.13)). We assessed associations with risk of ER-negative disease, and with risk of overall breast cancer for *BRCA1* mutation carriers, for SNPs at the remaining 62 loci, as well as for the 96 previously reported breast cancer susceptibility

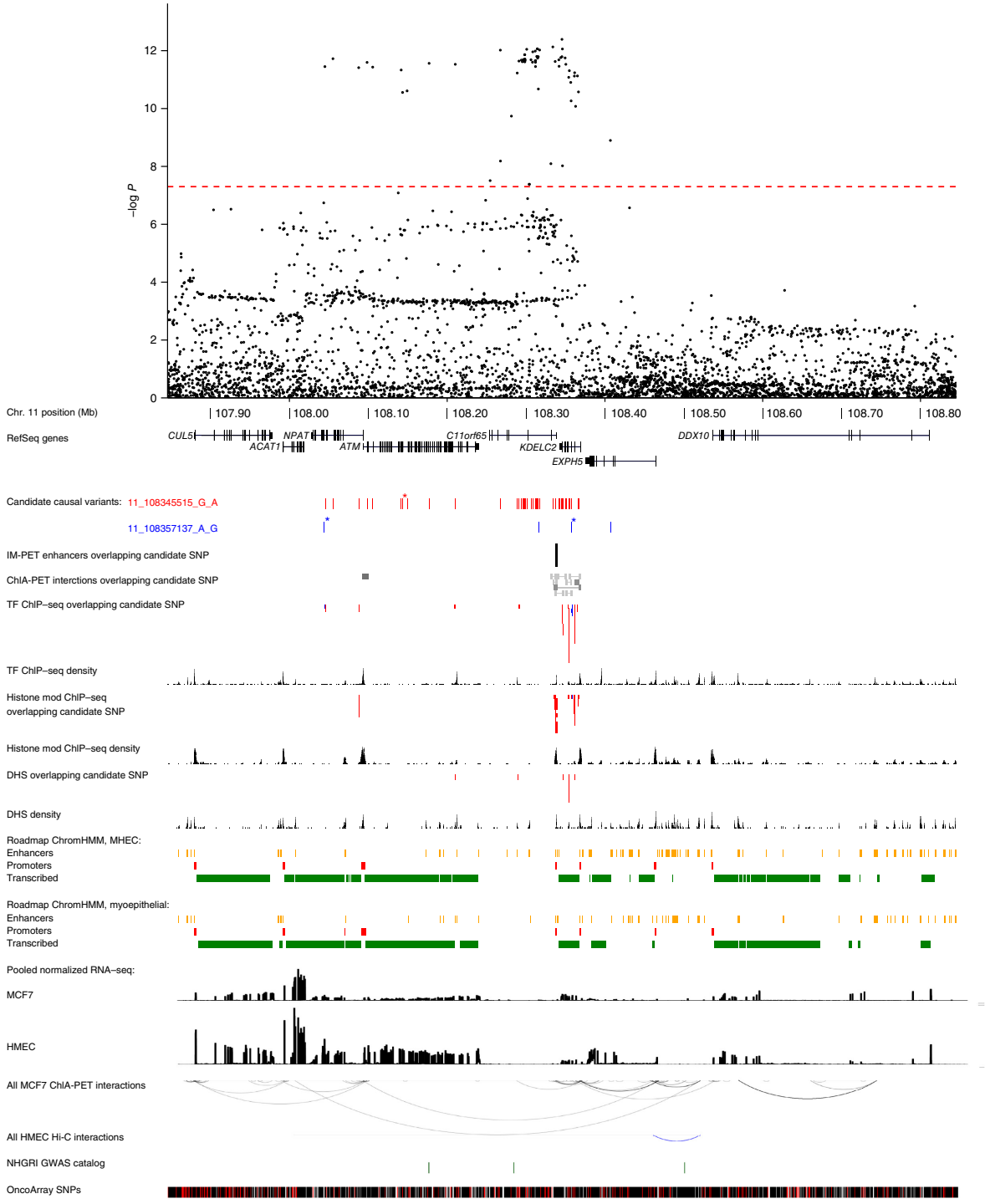

Figure 1 Genomic region around the independent ER-negative risk-associated variants 11_108345515_G_A (rs11374964) and 11_108357137_G_A (rs74911261). A 1-Mb regional association plot shows the statistical significance of all genotyped and imputed SNPs with the genome-wide significance level ($P = 5 \times 10^{-8}$) represented by the dashed red line. The positions of candidate causal variants for two independent signals (depicted as red and blue ticks) are shown in relation to RefSeq genes. Missense variants are labeled with asterisks. Mammary cell enhancers overlapping candidate SNPs predicted to target nearby genes by IM-PET[46] are depicted as black bars. Chromatin interactions from ENCODE ChIA-PET data in MCF7 cells overlapping candidate variants are shaded to reflect interaction confidence scores with darker shading indicating greater confidence. Epigenomic features (derived from publicly available ChIP–seq and DNase–seq data sets) that overlap candidate variants are shown as red or blue segments, depending on the intersected signal. Density tracks show the summed occurrence of ChIP–seq and DNase–seq peak signals at each position. Roadmap Epigenomics Project chromatin state models for HMECs and myoepithelial cells grouped into enhancer, promoter or transcribed annotations are shown as yellow, red and green segments, respectively. Transcript levels in MCF7 cells and HMECs are represented by histograms depicting mean normalized RNA-seq expression. All MCF7 ChIA-PET (ENCODE) and HMEC Hi-C[47] chromatin interactions are represented by black and blue arcs, respectively. NHGRI GWAS catalog SNPs are shown as green ticks. All OncoArray SNPs (genotyped or imputed) are shown as black ticks, and uninterrogated, common SNPs (dbSNP138, European (EUR) MAF > 1%) are shown as red ticks. Features may be examined in detail via exploration of a custom UCSC Genome Browser session accessible via hyperlinks within **Supplementary Table 5**. TF, transcription factor; DHS, DNase I–hypersensitive site.

Table 2 Previously reported associations with estrogen-receptor-negative disease: replication using independent data from BCAC and combined results using all BCAC and CIMBA data

						Independent replication						All available data combined					
						BCAC ER-negative disease (OncoArray)[b]				BCAC ER-negative disease[c]			CIMBA *BRCA1* mutation carriers[d]		
Location	SNP	Chr.	Position (bp)	Ref.	Nearest gene	Alleles[a]	MAF	OR	95% CI	*P* value	OR	95% CI	*P* value	HR	95% CI	*P* value
1q32.1	rs6678914	1	202,187,176	19	*LGR6*	G/A	0.41	0.94	0.91–0.97	1.1×10^{-4}	0.92	0.90–0.94	2.6×10^{-12}	0.98	0.95–1.02	0.31
1q32.1	rs4245739	1	204,518,842	19	*MDM4*	A/C	0.26	1.12	1.09–1.17	9.2×10^{-11}	1.14	1.11–1.16	3.1×10^{-23}	1.09	1.04–1.13	7.3×10^{-5}
2p24.1	rs12710696	2	19,320,803	19	*MIR4757*	C/T	0.37	1.04	1.00–1.07	2.5×10^{-2}	1.06	1.04–1.09	6.5×10^{-8}	1.01	0.98–1.05	0.49
2p23.2	rs4577244	2	29,120,733	30	*WDR43*	C/T	0.34	0.93	0.89–0.96	9.6×10^{-5}	0.92	0.90–0.95	1.5×10^{-9}	0.92	0.88–0.96	1.3×10^{-4}
5p15.33	rs10069690	5	1,279,790	9,18	*TERT*	C/T	0.26	1.19	1.14–1.23	3.8×10^{-21}	1.18	1.15–1.21	1.5×10^{-35}	1.18	1.14–1.23	3.7×10^{-16}
6q25.1	rs3757322	6	151,942,194	29	*ESR1*	T/G	0.32	1.14	1.10–1.18	5.5×10^{-14}	1.15	1.12–1.18	2.8×10^{-31}	1.14	1.10–1.19	2.9×10^{-12}
6q25.2	rs2747652	6	152,437,016	29	*ESR1*	C/T	0.48	0.92	0.89–0.95	1.1×10^{-7}	0.91	0.89–0.93	1.9×10^{-18}	1.00	0.97–1.04	0.96
13q22.1	rs6562760	13	73,957,681	30	*KLF5*	G/A	0.24	0.92	0.88–0.95	5.0×10^{-6}	0.92	0.90–0.95	8.7×10^{-10}	0.89	0.86–0.93	3.5×10^{-7}
16q12.2	rs11075995	16	53,855,291	19	*FTO*	T/A	0.30	1.07	1.03–1.11	3.3×10^{-4}	1.09	1.06–1.12	1.0×10^{-10}	1.01	0.97–1.06	0.49
19p13.11	rs67397200	19	17,401,404	3,31	*ANKLE1*	C/G	0.32	1.17	1.13–1.21	7.0×10^{-20}	1.17	1.14–1.19	2.7×10^{-37}	1.18	1.14–1.23	2.7×10^{-17}
20q11.21	rs2284378	20	32,588,095	12	*RALY*	C/T	0.32	0.99	0.95–1.02	0.46	1.03	1.01–1.06	1.7×10^{-2}	1.00	0.97–1.04	0.81

Chr., chromosome; ref., publication(s) in reference list in which the association was identified; MAF, minor allele frequency; OR, odds ratio per copy of the minor allele; CI, confidence interval; HR, hazard ratio per copy of the minor allele. [a]More common allele listed first, minor allele listed second. [b]Includes Breast Cancer Association Consortium (BCAC) OncoArray data from 9,655 ER-negative cases and 45,494 controls; cases and controls were not included in previously published studies. [c]Combined data from 21,468 ER-negative cases and 100,594 controls of European ancestry from BCAC, including samples overlapping with previous publications for all SNPs. [d]Combined data from 18,908 *BRCA1* mutation carriers from the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA), 9,414 of whom had developed breast cancer; including samples overlapping with previous publications for SNPs rs4577244, rs3757322, rs2747652 and rs6562760.

variants that were not specific to ER-negative disease. Of these 158 SNPs, 105 were associated ($P < 0.05$) with risk of ER-negative breast cancer and 24 were associated with overall risk for *BRCA1* mutation carriers (**Supplementary Tables 9** and **10**). Results for *BRCA2* mutation carriers are presented in **Supplementary Table 11**.

Pathway analysis based on mapping each SNP to the nearest gene was performed using summary association statistics from the meta-analysis of BCAC and CIMBA data combined[61–64] (Online Methods). This identified several pathways implicated in ER-negative disease (enrichment score (ES) ≥ 0.41; **Supplementary Fig. 14** and **Supplementary Tables 12** and **13**), including a subset of pathways that were not enriched in susceptibility to ER-positive disease (ES < 0; **Supplementary Table 14**). One of the latter subsets was the adenylate cyclase (AC)-activating pathway (ES = 0.62; **Supplementary Fig. 15**). Two of the predicted target genes for the ten new susceptibility variants for ER-negative breast cancer, based on the eQTL analysis (**Supplementary Table 5**), *ADCY3* ($P_{TCGA} = 6.7 \times 10^{-3}$) and *ADCY9* ($P_{METABRIC} = 1.3 \times 10^{-4}$), are part of this pathway, and their association signals were critical to the elevated enrichment scores observed (**Supplementary Fig. 14**). ADCY9 is stimulated by β2 adrenergic receptor (β2AR) signaling[65] in ER-negative breast cancer[66] and in turn drives AC–cAMP signaling, including, for example, mitogenic signaling through the β-arrestin–Src–ERK pathway[67].

To further explore the functional properties of the genome that contribute to heritability of ER-negative breast cancer, we conducted a partitioned heritability analysis using linkage disequilibrium (LD) score regression[68]. Considering 52 'baseline' genomic features, we observed the greatest enrichment for super-enhancers (2.5-fold, $P = 2 \times 10^{-7}$) and the H3K4me3 histone mark (2.4-fold, $P = 0.0005$), with 33% depletion ($P = 0.0002$) observed for repressed regions (**Supplementary Table 15**). No differences in enrichment for these features were observed between susceptibility to ER-negative and ER-positive breast cancer, but baseline genomic features are not specific to cell type[68]. The estimated correlation between ER-negative and ER-positive breast cancer based on ~1 million common genetic variants[69,70] was 0.60 (standard error (SE) = 0.03), indicating that, although these two breast cancer subtypes have a shared genetic component, a substantial proportion of their genetic bases is distinct. The estimated correlation between ER-negative disease in the general population and overall breast cancer for *BRCA1* mutation carriers was 0.72 (SE = 0.11).

In summary, in this study of women of European origin, we have identified ten new susceptibility variants for ER-negative breast cancer and replicated associations with ER-negative disease for ten SNPs identified by previous GWAS. Most of these variants were not associated or were more weakly associated with ER-positive disease, consistent with the findings from pathway and partitioned heritability analyses showing that ER-negative breast cancer has a partly distinct genetic etiology. We also observed consistent associations with ER-negative disease for a further 105 susceptibility SNPs for breast cancer overall. Together, these 125 variants explain ~14% of an assumed twofold increased risk of developing ER-negative disease for the first-degree female relatives of women affected with this subtype (the newly identified SNPs explain ~1.5%; **Supplementary Table 16**) and ~40% of the estimated familial risk that is attributable to all variants imputable from the OncoArray (Online Methods). We have also identified 9 new breast cancer susceptibility variants for *BRCA1* mutation carriers and confirmed associations for a further 30 previously reported SNPs; these 39 variants explain ~8% of the variance in polygenic risk for carriers of these mutations (**Supplementary Table 17**). However, the lower number of risk-associated variants in disease with *BRCA1* mutation may merely be a consequence of the smaller sample size, as the genetic

Table 3 Associations for ten new and ten previously reported (and replicated) susceptibility loci for estrogen-receptor-negative breast cancer, by triple-negative status

Location	SNP	Triple-negative disease			Other ER-negative disease			Heterogeneity *P* value[a]
		OR	95% CI	*P* value	OR	95% CI	*P* value	
Loci identified by the present study								
2p23.3	rs200648189	0.95	0.90–1.00	4.8×10^{-2}	0.96	0.91–1.03	0.24	0.36
6q23.1	rs6569648	0.93	0.89–0.97	1.4×10^{-3}	0.93	0.88–0.98	5.6×10^{-3}	0.91
8p23.3	rs66823261	1.11	1.05–1.16	3.3×10^{-5}	1.12	1.07–1.19	2.4×10^{-5}	0.91
8q24.13	rs17350191	1.07	1.03–1.11	7.9×10^{-4}	1.07	1.02–1.12	4.0×10^{-3}	0.67
11q22.3	rs11374964	0.88	0.85–0.91	1.9×10^{-11}	0.99	0.95–1.04	0.75	1.5×10^{-5}
11q22.3	rs74911261	0.76	0.66–0.87	1.1×10^{-4}	0.98	0.84–1.13	0.76	3.0×10^{-2}
16p13.3	rs11076805	0.91	0.87–0.96	1.5×10^{-4}	0.95	0.90–1.00	4.5×10^{-2}	0.20
18q12.1	rs36194942	0.93	0.89–0.96	2.4×10^{-4}	0.92	0.88–0.97	9.9×10^{-4}	0.94
19p13.2	rs322144	0.94	0.91–0.98	5.9×10^{-3}	0.94	0.90–0.98	9.7×10^{-3}	0.68
19q12	rs113701136	1.10	1.06–1.15	9.1×10^{-7}	1.07	1.02–1.12	4.4×10^{-3}	0.12
Previously reported loci (associations replicated by the present study)								
1q32.1	rs6678914	0.94	0.91–0.98	2.1×10^{-3}	0.91	0.87–0.95	2.0×10^{-5}	0.45
1q32.1	rs4245739	1.18	1.13–1.23	4.3×10^{-15}	1.04	1.00–1.10	7.5×10^{-2}	6.5×10^{-4}
2p24.1	rs12710696	1.07	1.03–1.11	1.1×10^{-3}	1.04	1.00–1.09	6.1×10^{-2}	0.52
2p23.2	rs4577244	0.90	0.86–0.94	5.3×10^{-6}	0.94	0.89–0.99	1.9×10^{-2}	0.15
5p15.33	rs10069690	1.28	1.23–1.33	2.4×10^{-33}	1.07	1.02–1.12	5.4×10^{-3}	5.6×10^{-8}
6q25.1	rs3757322	1.15	1.10–1.19	4.3×10^{-12}	1.14	1.10–1.20	4.8×10^{-9}	0.35
6q25.2	rs2747652	0.93	0.89–0.96	5.7×10^{-5}	0.87	0.83–0.91	2.9×10^{-10}	9.6×10^{-3}
13q22.1	rs6562760	0.94	0.90–0.98	2.8×10^{-3}	0.92	0.87–0.96	8.8×10^{-4}	0.46
16q12.2	rs11075995	1.06	1.02–1.11	6.5×10^{-3}	1.08	1.03–1.13	3.1×10^{-3}	0.81
19p13.11	rs67397200	1.27	1.22–1.32	2.0×10^{-32}	1.05	1.01–1.10	2.7×10^{-2}	4.7×10^{-10}

Only Breast Cancer Association Consortium (BCAC) data are shown. BCAC data were combined from 6,877 triple-negative and 4,467 other ER-negative cases and 83,700 controls. OR, odds ratio per copy of the minor allele; CI, confidence interval.
[a]Analysis of ER-negative cases only, by triple-negative status.

correlation with ER-negative breast cancer is high. These findings may inform improved risk prediction, both for the general population and *BRCA1* mutation carriers[30,71,72]. Further investigation is required for other populations of non-European origin. Fine-mapping and functional studies should lead to a better understanding of the biological basis of ER-negative breast cancer and may perhaps inform the design of more effective preventive interventions, early detection and treatments for this disease.

Table 4 Associations for ten new and ten previously reported (and replicated) susceptibility loci for ER-negative breast cancer, by tumor grade

Location	SNP	Grade 1			Grade 2			Grade 3			Heterogeneity *P* value[a]
		OR	95% CI	*P* value	OR	95% CI	*P* value	OR	95% CI	*P* value	
Loci identified by the present study											
2p23.3	rs200648189	1.11	0.92–1.33	0.28	0.95	0.88–1.03	0.23	0.96	0.91–1.00	6.8×10^{-2}	0.70
6q23.1	rs6569648	0.93	0.79–1.09	0.37	0.93	0.87–0.99	1.6×10^{-2}	0.94	0.91–0.98	3.8×10^{-3}	0.34
8p23.3	rs66823261	1.13	0.96–1.34	0.14	1.12	1.04–1.19	1.2×10^{-3}	1.10	1.05–1.15	1.3×10^{-5}	0.11
8q24.13	rs17350191	1.16	1.01–1.34	3.0×10^{-2}	1.05	0.99–1.11	0.10	1.09	1.05–1.12	4.1×10^{-6}	0.94
11q22.3	rs11374964	0.91	0.79–1.04	0.16	0.99	0.94–1.05	0.85	0.93	0.90–0.96	1.3×10^{-5}	3.0×10^{-2}
11q22.3	rs74911261	1.22	0.81–1.84	0.35	0.89	0.73–1.07	0.21	0.74	0.65–0.85	7.4×10^{-6}	6.7×10^{-4}
16p13.3	rs11076805	0.90	0.76–1.06	0.21	0.93	0.87–0.99	3.2×10^{-2}	0.92	0.88–0.95	4.5×10^{-5}	0.71
18q12.1	rs36194942	0.97	0.84–1.13	0.73	0.93	0.88–0.99	2.2×10^{-2}	0.96	0.92–0.99	2.3×10^{-2}	0.98
19p13.2	rs322144	0.94	0.81–1.08	0.38	0.95	0.90–1.01	0.11	0.96	0.93–1.00	6.4×10^{-2}	0.48
19q12	rs113701136	1.02	0.89–1.18	0.77	1.06	1.01–1.13	3.0×10^{-2}	1.10	1.06–1.14	2.5×10^{-7}	0.12
Previously reported loci (associations replicated by the present study)											
1q32.1	rs6678914	0.95	0.83–1.09	0.46	0.90	0.85–0.95	9.3×10^{-5}	0.92	0.89–0.95	1.2×10^{-6}	0.75
1q32.1	rs4245739	1.02	0.88–1.19	0.75	1.05	0.99–1.12	8.7×10^{-2}	1.1)	1.14–1.22	2.5×10^{-18}	4.3×10^{-5}
2p24.1	rs12710696	1.08	0.94–1.23	0.28	1.10	1.04–1.16	9.6×10^{-4}	1.04	1.01–1.08	1.6×10^{-2}	0.28
2p23.2	rs4577244	1.02	0.88–1.20	0.77	0.95	0.89–1.01	9.4×10^{-2}	0.90	0.86–0.93	1.2×10^{-7}	4.0×10^{-2}
5p15.33	rs10069690	0.96	0.83–1.12	0.64	1.07	1.01–1.14	2.2×10^{-2}	1.21	1.17–1.26	1.5×10^{-24}	7.3×10^{-4}
6q25.1	rs3757322	1.16	1.01–1.34	0.04	1.13	1.07–1.20	7.5×10^{-6}	1.18	1.14–1.22	4.5×10^{-20}	0.16
6q25.2	rs2747652	0.86	0.75–0.98	0.02	0.92	0.87–0.97	1.9×10^{-3}	0.90	0.87–0.93	1.6×10^{-9}	0.61
13q22.1	rs6562760	0.98	0.84–1.15	0.82	0.92	0.87–0.98	1.4×10^{-2}	0.91	0.88–0.95	1.2×10^{-5}	0.52
16q12.2	rs11075995	1.16	1.00–1.35	4.7×10^{-2}	1.09	1.02–1.15	7.5×10^{-3}	1.08	1.04–1.13	5.2×10^{-28}	0.42
19p13.11	rs67397200	1.01	0.87–1.16	0.91	1.08	1.02–1.14	9.8×10^{-3}	1.22	1.18–1.26	5.3×10^{-37}	1.3×10^{-3}

Only Breast Cancer Association Consortium (BCAC) data are shown. BCAC data were combined from 492 grade 1, 3,243 grade 2 and 8,568 grade 3 cases and 82,347 controls. OR, odds ratio per copy of the minor allele; CI, confidence interval.
[a]Analysis of ER-negative cases only, by tumor grade (trend test, 1 degree of freedom).

URLs. Database of Genotypes and Phenotypes (dbGaP), https://www.ncbi.nlm.nih.gov/gap; Breast Cancer Association Consortium (BCAC), http://bcac.ccge.medschl.cam.ac.uk/; Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA), http://cimba.ccge.medschl.cam.ac.uk/; PCcalc software, http://ccge.medschl.cam.ac.uk/software/pccalc/; SNPTEST, https://mathgen.stats.ox.ac.uk/genetics_software/snptest/snptest.html; GeneSets, http://baderlab.org/GeneSets; GenGen package, http://gengen.openbioinformatics.org/en/latest/.

METHODS

Methods, including statements of data availability and any associated accession codes and references, are available in the online version of the paper.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

ACKNOWLEDGMENTS

We thank all the individuals who took part in these studies and all the researchers, clinicians, technicians and administrative staff who have enabled this work to be carried out.

Genotyping for the OncoArray was funded by the government of Canada through Genome Canada and the Canadian Institutes of Health Research (GPH-129344), the Ministère de l'Économie, de la Science et de l'Innovation du Québec through Génome Québec, the Quebec Breast Cancer Foundation for the PERSPECTIVE project, the US National Institutes of Health (NIH) (1 U19 CA 148065 for the Discovery, Biology and Risk of Inherited Variants in Breast Cancer (DRIVE) project and X01HG007492 to the Center for Inherited Disease Research (CIDR) under contract HHSN268201200008I), Cancer Research UK (C1287/A16563), the Odense University Hospital Research Foundation (Denmark), the National R&D Program for Cancer Control–Ministry of Health and Welfare (Republic of Korea) (1420190), the Italian Association for Cancer Research (AIRC; IG16933), the Breast Cancer Research Foundation, the National Health and Medical Research Council (Australia) and German Cancer Aid (110837).

Genotyping for the iCOGS array was funded by the European Union (HEALTH-F2-2009-223175), Cancer Research UK (C1287/A10710, C1287/A10118 and C12292/A11174]), NIH grants (CA128978, CA116167 and CA176785) and the Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 (GAME-ON initiative)), an NCI Specialized Program of Research Excellence (SPORE) in Breast Cancer (CA116201), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, the Ministère de l'Économie, Innovation et Exportation du Québec (PSR-SIIRI-701), the Komen Foundation for the Cure, the Breast Cancer Research Foundation and the Ovarian Cancer Research Fund.

Combination of the GWAS data was supported in part by the NIH Cancer Post-Cancer GWAS initiative (1 U19 CA 148065) (DRIVE, part of the GAME-ON initiative). LD score regression analysis was supported by grant CA194393. BCAC is funded by Cancer Research UK (C1287/A16563) and by the European Union via its Seventh Framework Programme (HEALTH-F2-2009-223175, COGS) and the Horizon 2020 Research and Innovation Programme (633784, B-CAST; 634935, BRIDGES). CIMBA is funded by Cancer Research UK (C12292/A20861 and C12292/A11174).

For a full description of funding and acknowledgments, see the **Supplementary Note**.

AUTHOR CONTRIBUTIONS

Writing group: R.L.M., K.B.K., K. Michailidou, J. Beesley, S. Kar, S. Lindström, S. Hui, G.D.B., P.D.P.P., F.J.C., D.F.E., P.K., G.C.-T., M.G.-C., M.K.S., A.C.A., J. Simard. Conception and coordination of OncoArray synthesis: D.F.E., A.C.A., J. Simard, C.I.A., J. Byun, S.J.C., E.D., D.J.H., A. Lee, P.D.P.P., J.T., Z.W. OncoArray genotyping: M.A., A.C.A., S.E.B., M.K.B., F.B., G.C.-T., J.M.C., K.F.D., D.F.E., N. Hammell, B. Hicks, K.J., C. Luccarini, L.M., J.M., E.P., J. Romm, M.K.S., X.S., J. Simard, P. Soucy, D.C.T., D.V., J. Vollenweider, L.X., B.Z. OncoArray genotype calling and quality control: X.C., J.D., E.D., D.F.E., K.B.K., J. Lecarpentier, A. Lee, M. Lush. Database management: D. Barrowdale, M.K.B., M.L., L.M., Q.W., R. Keeman, M.K.S. Statistical analysis: K.B.K., K. Michailidou, S. Hui, S. Kar, X.J., A. Rostamianfar, H. Finucane, S. Lindström, D. Barnes, P.K., P.D.P.P., G.D.B., R.L.M., A.C.A., D.F.E. Bioinformatic analysis: J. Beesley, P. Soucy, A. Lemaçon, D. Barnes, F.A.-E., A.D., J. Simard, G.C.-T. Provision of DNA samples and/or phenotypic data: ABCTB Investigators, C.M.A., J. Adlard, S. Agata, S. Ahmed, H.A., J. Allen, K.A., C.B.A., I.L.A., H.A.-C., N.N.A., A.C.A., V.A., N.A., K.J.A., B.A., P.L.A., M.G.E.M.A., J. Azzollini, J. Balmaña, M. Barile, L. Barjhoux, R.B.B., M. Barrdahl, D. Barnes, D. Barrowdale, C. Baynes, M.W.B., J. Beesley, J. Benitez, M. Bermisheva, L. Bernstein, Y.-J.B., K.R.B., M.J.B., C. Blomqvist, W.B., K.B., B. Boeckx, N.V.B., A. Bojesen, S.E.B., M.K.B., B. Bonanni, A. Bozsik, A.R.B., J.S.B., H. Brauch, H. Brenner, B.B.-d.P., C. Brewer, L. Brinton, P.B., A.B.-W., J. Brunet, T.B., B. Burwinkel, S.S.B., A.-L.B.-D., Q.C., T. Caldés, M.A.C., I. Campbell, F.C., O.C., A. Carracedo, B.D.C., J.E.C., L.C., V.C.-M., S.B.C., J.C.-C., S.J.C., X.C., G.C.-T., T.-Y.D.C., J. Chiquette, H.C., K.B.M.C., C.L.C., NBSS Collaborators, T. Conner, D.M.C., J. Cook, E.C.-D., S.C., F.J.C., I. Coupier, D.G.C., A. Cox, S.S.C., K. Cuk, K. Czene, M.B.D., F.D., H.D., R.D., J.D., P.D., O.D., Y.C.D., N.D., S.M.D., C.M.D., S.D., P.-A.D., M. Dumont, A.M.D., L.D., M. Dwek, B.D., T.D., EMBRACE, D.F.E., D.E., R.E., H. Ehrencrona, U.E., B.E., A.B.E., A.H.E., C.E., M.E., L. Fachal, L. Faivre, P.A.F., U.F., J.F., D.F.-J., O.F., H. Flyger, W.D.F., E.F., L. Fritschi, D.F., GEMO Study Collaborators, M. Gabrielson, P. Gaddam, M.D.G., M.G.-D., P.A.G., S.M.G., J. Garber, V.G.-B., M.G.-C., J.A.G.-S., M.M.G., M.G.-V., A. Gehrig, V.G., A.-M.G., G.G.G., G.G., A.K.G., M.S.G., D.E.G., A.G.-N., P. Goodfellow, M.H.G., G.I.G.A., M. Grip, J. Gronwald, A. Grundy, D.G.-K., Q.G., P. Guénel, HEBON, L.H., E. Hahnen, C.A.H., P. Hall, E. Hallberg, U.H., S. Hankinson, T.V.O.H., P. Harrington, S.N.H., J.M.H., C.S.H., A. Hein, S. Helbig, A. Henderson, J.H., P. Hillemanns, S. Hodgson, F.B.H., A. Hollestelle, M.J.H., B. Hoover, J.L.H., C.H., G.H., P.J.H., K.H., D.J.H., N. Håkansson, E.N.I., C.I., M.I., L.I., A.J., P.J., R.J., W.J., U.B.J., E.M.J., N.J., M.J., A.J.-V., R. Kaaks, M. Kabisch, K. Kaczmarek, D.K., K. Kast, R. Keeman, M.J.K., C.M.K., M. Keupers, S. Khan, E.K., J.I.K., J.A.K., I.K., V.-M.K., S.-W.K., P.K., V.N.K., T.A.K., K.B.K., A.K., Y.L., F. Lalloo, K.L., D.L., C. Lasset, C. Lazaro, L.l.M., J. Lecarpentier, M. Lee, A. Lee, E.L., J. Lee, F. Lejbkowicz, F. Lesueur, J. Li, J. Lilyquist, A. Lincoln, A. Lindblom, S. Lindström, J. Lissowska, W.-Y.L., S. Loibl, J. Long, J.T.L., J. Lubinski, C. Luccarini, M. Lush, A.-V.L., R.J.M., T.M., E.M., K.E.M., I.M.K., A. Mannermaa, S. Manoukian, J.E.M., S. Margolin, J.W.M.M., M.E.M., K. Matsuo, D.M., S. Mazoyer, L.M., C. McLean, H.M.-H., A. Meindl, P.M., H.M., K. Michailidou, A. Miller, N.M., R.L.M., G.M., M.M., K. Muir, A.M.M., C. Mulot, S.N., K.L.N., S.L.N., H.N., I.N., D.N., S.F.N., B.G.N., A.N., R.L.N., K. Offit, E.O., O.I.O., J.E.O., H.O., C.O., K. Ong, J.C.O., N.O., A.O., L.O., V.S.P., L.P., S.K.P., T.-W.P.-S., Y.P.-K., R.L., I.S.P., B. Peissel, A.P., J.I.A.P., P.P., J.P., G.P., P.D.P.P., C.M.P., M.P., D.P.-K., B. Poppe, M.E.P., R.P., N.P., D.P., M.A.P., K.P., B.R., P.R., N.R., J. Rantala, C.R.-F., H.S.R., G.R., V.R., K.R., A. Richardson, G.C.R., A. Romero, M.A.R., A. Rudolph, T.R., E.S., J. Sanders, D.P.S., S. Sangrajrang, E.J.S., D.F.S., M.K.S., R.K.S., M.J. Schoemaker, F.S., L. Schwentner, P. Schürmann, C. Scott, R.J.S., S. Seal, L. Senter, C. Seynaeve, M.S., P. Sharma, C.-Y.S., H. Shimelis, M.J. Shrubsole, X.-O.S., L.E.S., J. Simard, C.F.S., C. Sohn, P. Soucy, M.C.S., J.J.S., A.B.S., C. Stegmaier, J. Stone, D.S.-L., G.S., H. Surowy, C. Sutter, A.S., C.I.S., R.M.T., Y.Y.T., J.A.T., M.R.T., M.-I.T., L. Tong, M. Tengström, S.H.T., M.B.T., A.T., M. Thomassen, D.L.T., K. Thöne, M.G.T., L. Tihomirova, M. Tischkowitz, A.E.T., R.A.E.M.T., I.T., D.T., M. Tranchant, T.T., K. Tucker, N.T., H.-U.U., C.V., D.v.d.B., L.V., R.V.-M., A. Vega, A. Viel, J. Vijai, L.W., Q.W., S.W.-G., B.W., C.R.W., J.N.W., C.W., J.W., A.S.W., J.T.W., W.W., R.W., A.W., A.H.W., X.R.Y., D.Y., D.Z., W.Z., A.Z., E.Z., K.K.Z., I.d.-S.-S., kConFab AOCS Investigators, C.J.v.A., E.v.R., A.M.W.v.d.O. All authors read and approved the final version of the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

1. Michailidou, K. *et al.* Association analysis identifies 65 new breast cancer risk loci. *Nature* (in press).
2. Ahmed, S. *et al.* Newly discovered breast cancer susceptibility loci on 3p24 and 17q23.2. *Nat. Genet.* **41**, 585–590 (2009).
3. Antoniou, A.C. *et al.* A locus on 19p13 modifies risk of breast cancer in *BRCA1* mutation carriers and is associated with hormone receptor–negative breast cancer in the general population. *Nat. Genet.* **42**, 885–892 (2010).
4. Cai, Q. *et al.* Genome-wide association study identifies breast cancer risk variant at 10q21.2: results from the Asia Breast Cancer Consortium. *Hum. Mol. Genet.* **20**, 4991–4999 (2011).
5. Cox, A. *et al.* A common coding variant in *CASP8* is associated with breast cancer risk. *Nat. Genet.* **39**, 352–358 (2007).
6. Easton, D.F. *et al.* Genome-wide association study identifies novel breast cancer susceptibility loci. *Nature* **447**, 1087–1093 (2007).
7. Fletcher, O. *et al.* Novel breast cancer susceptibility locus at 9q31.2: results of a genome-wide association study. *J. Natl. Cancer Inst.* **103**, 425–435 (2011).
8. Ghoussaini, M. *et al.* Genome-wide association analysis identifies three new breast cancer susceptibility loci. *Nat. Genet.* **44**, 312–318 (2012).
9. Haiman, C.A. *et al.* A common variant at the *TERT–CLPTM1L* locus is associated with estrogen receptor–negative breast cancer. *Nat. Genet.* **43**, 1210–1214 (2011).

10. Hein, R. *et al.* Comparison of 6q25 breast cancer hits from Asian and European genome wide association studies in the Breast Cancer Association Consortium (BCAC). *PLoS One* **7**, e42380 (2012).
11. 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).
12. Siddiq, A. *et al.* A meta-analysis of genome-wide association studies of breast cancer identifies two novel susceptibility loci at 6q14 and 20q11. *Hum. Mol. Genet.* **21**, 5373–5384 (2012).
13. 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).
14. Stacey, S.N. *et al.* Common variants on chromosome 5p12 confer susceptibility to estrogen receptor–positive breast cancer. *Nat. Genet.* **40**, 703–706 (2008).
15. Thomas, G. *et al.* A multistage genome-wide association study in breast cancer identifies two new risk alleles at 1p11.2 and 14q24.1 (*RAD51L1*). *Nat. Genet.* **41**, 579–584 (2009).
16. Turnbull, C. *et al.* Genome-wide association study identifies five new breast cancer susceptibility loci. *Nat. Genet.* **42**, 504–507 (2010).
17. Zheng, W. *et al.* Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1. *Nat. Genet.* **41**, 324–328 (2009).
18. Bojesen, S.E. *et al.* Multiple independent variants at the *TERT* locus are associated with telomere length and risks of breast and ovarian cancer. *Nat. Genet.* **45**, 371–384, e1–e2 (2013).
19. Garcia-Closas, M. *et al.* Genome-wide association studies identify four ER negative–specific breast cancer risk loci. *Nat. Genet.* **45**, 392–398, e1–e2 (2013).
20. Michailidou, K. *et al.* Large-scale genotyping identifies 41 new loci associated with breast cancer risk. *Nat. Genet.* **45**, 353–361, e1–e2 (2013).
21. Cai, Q. *et al.* Genome-wide association analysis in East Asians identifies breast cancer susceptibility loci at 1q32.1, 5q14.3 and 15q26.1. *Nat. Genet.* **46**, 886–890 (2014).
22. Long, J. *et al.* Genome-wide association study in east Asians identifies novel susceptibility loci for breast cancer. *PLoS Genet.* **8**, e1002532 (2012).
23. Michailidou, K. *et al.* Genome-wide association analysis of more than 120,000 individuals identifies 15 new susceptibility loci for breast cancer. *Nat. Genet.* **47**, 373–380 (2015).
24. Milne, R.L. *et al.* Common non-synonymous SNPs associated with breast cancer susceptibility: findings from the Breast Cancer Association Consortium. *Hum. Mol. Genet.* **23**, 6096–6111 (2014).
25. Gaudet, M.M. *et al.* Identification of a BRCA2-specific modifier locus at 6p24 related to breast cancer risk. *PLoS Genet.* **9**, e1003173 (2013).
26. Meyer, K.B. *et al.* Fine-scale mapping of the *FGFR2* breast cancer risk locus: putative functional variants differentially bind FOXA1 and E2F1. *Am. J. Hum. Genet.* **93**, 1046–1060 (2013).
27. Orr, N. *et al.* Fine-mapping identifies two additional breast cancer susceptibility loci at 9q31.2. *Hum. Mol. Genet.* **24**, 2966–2984 (2015).
28. French, J.D. *et al.* Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. *Am. J. Hum. Genet.* **92**, 489–503 (2013).
29. Dunning, A.M. *et al.* Breast cancer risk variants at 6q25 display different phenotype associations and regulate *ESR1*, *RMND1* and *CCDC170*. *Nat. Genet.* **48**, 374–386 (2016).
30. Couch, F.J. *et al.* Identification of four novel susceptibility loci for oestrogen receptor negative breast cancer. *Nat. Commun.* **7**, 11375 (2016).
31. Lawrenson, K. *et al.* Functional mechanisms underlying pleiotropic risk alleles at the 19p13.1 breast–ovarian cancer susceptibility locus. *Nat. Commun.* **7**, 12675 (2016).
32. Wyszynski, A. *et al.* An intergenic risk locus containing an enhancer deletion in 2q35 modulates breast cancer risk by deregulating *IGFBP5* expression. *Hum. Mol. Genet.* **25**, 3863–3876 (2016).
33. Mavaddat, N. *et al.* Pathology of breast and ovarian cancers among *BRCA1* and *BRCA2* mutation carriers: results from the Consortium of Investigators of Modifiers of *BRCA1/2* (CIMBA). *Cancer Epidemiol. Biomarkers Prev.* **21**, 134–147 (2012).
34. Amos, C.I. *et al.* The OncoArray Consortium: a network for understanding the genetic architecture of common cancers. *Cancer Epidemiol. Biomarkers Prev.* **26**, 126–135 (2017).
35. 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).
36. Barnes, D.R., Lee, A., Easton, D.F. & Antoniou, A.C. Evaluation of association methods for analysing modifiers of disease risk in carriers of high-risk mutations. *Genet. Epidemiol.* **36**, 274–291 (2012).
37. Ahsan, H. *et al.* A genome-wide association study of early-onset breast cancer identifies *PFKM* as a novel breast cancer gene and supports a common genetic spectrum for breast cancer at any age. *Cancer Epidemiol. Biomarkers Prev.* **23**, 658–669 (2014).
38. Stevens, K.N. *et al.* 19p13.1 is a triple-negative-specific breast cancer susceptibility locus. *Cancer Res.* **72**, 1795–1803 (2012).
39. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
40. Maller, J.B. *et al.* Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).
41. Udler, M.S., Tyrer, J. & Easton, D.F. Evaluating the power to discriminate between highly correlated SNPs in genetic association studies. *Genet. Epidemiol.* **34**, 463–468 (2010).
42. ENCODE Project Consortium. A user's guide to the Encyclopedia of DNA Elements (ENCODE). *PLoS Biol.* **9**, e1001046 (2011).
43. Kheradpour, P. & Kellis, M. Systematic discovery and characterization of regulatory motifs in ENCODE TF binding experiments. *Nucleic Acids Res.* **42**, 2976–2987 (2014).
44. Kundaje, A. *et al.* Integrative analysis of 111 reference human epigenomes. *Nature* **518**, 317–330 (2015).
45. Boyle, A.P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res.* **22**, 1790–1797 (2012).
46. He, B., Chen, C., Teng, L. & Tan, K. Global view of enhancer–promoter interactome in human cells. *Proc. Natl. Acad. Sci. USA* **111**, E2191–E2199 (2014).
47. Rao, S.S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).
48. Corradin, O. *et al.* Combinatorial effects of multiple enhancer variants in linkage disequilibrium dictate levels of gene expression to confer susceptibility to common traits. *Genome Res.* **24**, 1–13 (2014).
49. Forrest, A.R. *et al.* A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).
50. GTEx Consortium. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science* **348**, 648–660 (2015).
51. Hnisz, D. *et al.* Super-enhancers in the control of cell identity and disease. *Cell* **155**, 934–947 (2013).
52. Westra, H.J. *et al.* Systematic identification of *trans* eQTLs as putative drivers of known disease associations. *Nat. Genet.* **45**, 1238–1243 (2013).
53. James, L.I. *et al.* Small-molecule ligands of methyl-lysine binding proteins: optimization of selectivity for L3MBTL3. *J. Med. Chem.* **56**, 7358–7371 (2013).
54. Sukari, A., Muqbil, I., Mohammad, R.M., Philip, P.A. & Azmi, A.S. F-BOX proteins in cancer cachexia and muscle wasting: emerging regulators and therapeutic opportunities. *Semin. Cancer Biol.* **36**, 95–104 (2016).
55. Ling Zheng, L. *et al.* Interaction of heat shock protein Cpn10 with the cyclin E/Cdk2 substrate nuclear protein ataxia-telangiectasia (NPAT) is involved in regulating histone transcription. *J. Biol. Chem.* **290**, 29290–29300 (2015).
56. Easton, D.F. *et al.* Gene-panel sequencing and the prediction of breast-cancer risk. *N. Engl. J. Med.* **372**, 2243–2257 (2015).
57. Rogers, S. *et al.* Cyclin E2 is the predominant E-cyclin associated with NPAT in breast cancer cells. *Cell Div.* **10**, 1 (2015).
58. Li, Q. *et al.* Integrative eQTL-based analyses reveal the biology of breast cancer risk loci. *Cell* **152**, 633–641 (2013).
59. Cancer Genome Atlas Network. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
60. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).
61. Merico, D., Isserlin, R., Stueker, O., Emili, A. & Bader, G.D. Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* **5**, e13984 (2010).
62. Wang, K., Li, M. & Bucan, M. Pathway-based approaches for analysis of genomewide association studies. *Am. J. Hum. Genet.* **81**, 1278–1283 (2007).
63. Wang, K., Li, M. & Hakonarson, H. Analysing biological pathways in genome-wide association studies. *Nat. Rev. Genet.* **11**, 843–854 (2010).
64. Wang, L., Jia, P., Wolfinger, R.D., Chen, X. & Zhao, Z. Gene set analysis of genome-wide association studies: methodological issues and perspectives. *Genomics* **98**, 1–8 (2011).
65. Hacker, B.M. *et al.* Cloning, chromosomal mapping, and regulatory properties of the human type 9 adenylyl cyclase (ADCY9). *Genomics* **50**, 97–104 (1998).
66. Melhem-Bertrandt, A. *et al.* β-blocker use is associated with improved relapse-free survival in patients with triple-negative breast cancer. *J. Clin. Oncol.* **29**, 2645–2652 (2011).
67. Pon, C.K., Lane, J.R., Sloan, E.K. & Halls, M.L. The β2-adrenoceptor activates a positive cAMP–calcium feedforward loop to drive breast cancer cell invasion. *FASEB J.* **30**, 1144–1154 (2016).
68. Finucane, H.K. *et al.* Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
69. Bulik-Sullivan, B. *et al.* An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
70. Bulik-Sullivan, B.K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
71. Milne, R.L. & Antoniou, A.C. Genetic modifiers of cancer risk for *BRCA1* and *BRCA2* mutation carriers. *Ann. Oncol.* **22** (Suppl. 1), i11–i17 (2011).
72. Mavaddat, N. *et al.* Prediction of breast cancer risk based on profiling with common genetic variants. *J. Natl. Cancer Inst.* **107**, djv036 (2015).

Roger L Milne[1,2,334], Karoline B Kuchenbaecker[3,4,334], Kyriaki Michailidou[3,5,334], Jonathan Beesley[6], Siddhartha Kar[7], Sara Lindström[8,9], Shirley Hui[10], Audrey Lemaçon[11], Penny Soucy[11], Joe Dennis[3], Xia Jiang[9], Asha Rostamianfar[10], Hilary Finucane[9,12], Manjeet K Bolla[3], Lesley McGuffog[3], Qin Wang[3], Cora M Aalfs[13], ABCTB Investigators[14], Marcia Adams[15], Julian Adlard[16], Simona Agata[17], Shahana Ahmed[7], Habibul Ahsan[18], Kristiina Aittomäki[19], Fares Al-Ejeh[20], Jamie Allen[3], Christine B Ambrosone[21], Christopher I Amos[22], Irene L Andrulis[23,24], Hoda Anton-Culver[25], Natalia N Antonenkova[26], Volker Arndt[27], Norbert Arnold[28], Kristan J Aronson[29], Bernd Auber[30], Paul L Auer[31,32], Margreet G E M Ausems[33], Jacopo Azzollini[34], François Bacot[35], Judith Balmaña[36], Monica Barile[37], Laure Barjhoux[38], Rosa B Barkardottir[39,40], Myrto Barrdahl[41], Daniel Barnes[3], Daniel Barrowdale[3], Caroline Baynes[7], Matthias W Beckmann[42], Javier Benitez[43–45], Marina Bermisheva[46], Leslie Bernstein[47], Yves-Jean Bignon[48], Kathleen R Blazer[49], Marinus J Blok[50], Carl Blomqvist[51], William Blot[52,53], Kristie Bobolis[54], Bram Boeckx[55,56], Natalia V Bogdanova[26,57,58], Anders Bojesen[59], Stig E Bojesen[60–62], Bernardo Bonanni[37], Anne-Lise Børresen-Dale[63], Aniko Bozsik[64], Angela R Bradbury[65], Judith S Brand[66], Hiltrud Brauch[67–69], Hermann Brenner[27,69,70], Brigitte Bressac-de Paillerets[71], Carole Brewer[72], Louise Brinton[73], Per Broberg[74], Angela Brooks-Wilson[75,76], Joan Brunet[77], Thomas Brüning[78], Barbara Burwinkel[79,80], Saundra S Buys[81], Jinyoung Byun[22], Qiuyin Cai[52], Trinidad Caldés[82], Maria A Caligo[83], Ian Campbell[84,85], Federico Canzian[86], Olivier Caron[71], Angel Carracedo[87,88], Brian D Carter[89], J Esteban Castelao[90], Laurent Castera[91], Virginie Caux-Moncoutier[92], Salina B Chan[93], Jenny Chang-Claude[41,94], Stephen J Chanock[73], Xiaoqing Chen[6], Ting-Yuan David Cheng[95], Jocelyne Chiquette[96], Hans Christiansen[57], Kathleen B M Claes[97], Christine L Clarke[98], Thomas Conner[99], Don M Conroy[7], Jackie Cook[100], Emilie Cordina-Duverger[101], Sten Cornelissen[102], Isabelle Coupier[103], Angela Cox[104], David G Cox[105,106], Simon S Cross[107], Katarina Cuk[27], Julie M Cunningham[108], Kamila Czene[66], Mary B Daly[109], Francesca Damiola[38], Hatef Darabi[66], Rosemarie Davidson[110], Kim De Leeneer[97], Peter Devilee[111,112], Ed Dicks[7], Orland Diez[113], Yuan Chun Ding[47], Nina Ditsch[114], Kimberly F Doheny[15], Susan M Domchek[65], Cecilia M Dorfling[115], Thilo Dörk[58], Isabel dos-Santos-Silva[116], Stéphane Dubois[11], Pierre-Antoine Dugué[1,2], Martine Dumont[11], Alison M Dunning[7], Lorraine Durcan[117,118], Miriam Dwek[119], Bernd Dworniczak[120], Diana Eccles[118], Ros Eeles[121], Hans Ehrencrona[122], Ursula Eilber[41], Bent Ejlertsen[123], Arif B Ekici[124], A Heather Eliassen[125,126], EMBRACE[14], Christoph Engel[127,128], Mikael Eriksson[66], Laura Fachal[7], Laurence Faivre[129,130], Peter A Fasching[42,131], Ulrike Faust[132], Jonine Figueroa[73,133], Dieter Flesch-Janys[134,135], Olivia Fletcher[136], Henrik Flyger[137], William D Foulkes[138], Eitan Friedman[139,140], Lin Fritschi[141], Debra Frost[3], Marike Gabrielson[66], Pragna Gaddam[142], Marilie D Gammon[143], Patricia A Ganz[144], Susan M Gapstur[89], Judy Garber[145], Vanesa Garcia-Barberan[82], José A García-Sáenz[82], Mia M Gaudet[89], Marion Gauthier-Villars[92], Andrea Gehrig[146], GEMO Study Collaborators[14], Vassilios Georgoulias[147], Anne-Marie Gerdes[148], Graham G Giles[1,2], Gord Glendon[23], Andrew K Godwin[149], Mark S Goldberg[150,151], David E Goldgar[152], Anna González-Neira[43], Paul Goodfellow[153], Mark H Greene[154], Grethe I Grenaker Alnæs[63], Mervi Grip[155], Jacek Gronwald[156], Anne Grundy[157], Daphne Gschwantler-Kaulich[158], Pascal Guénel[101], Qi Guo[159], Lothar Haeberle[42], Eric Hahnen[160–162], Christopher A Haiman[163], Niclas Håkansson[164], Emily Hallberg[165], Ute Hamann[166], Nathalie Hamel[35], Susan Hankinson[167], Thomas V O Hansen[168], Patricia Harrington[7], Steven N Hart[165], Jaana M Hartikainen[169–171], Catherine S Healey[7], HEBON[14], Alexander Hein[42], Sonja Helbig[58], Alex Henderson[172], Jane Heyworth[173], Belynda Hicks[174], Peter Hillemanns[58], Shirley Hodgson[175], Frans B Hogervorst[176], Antoinette Hollestelle[177], Maartje J Hooning[177], Bob Hoover[73], John L Hopper[2], Chunling Hu[108], Guanmengqian Huang[166], Peter J Hulick[178,179], Keith Humphreys[66], David J Hunter[9,126], Evgeny N Imyanitov[180], Claudine Isaacs[181], Motoki Iwasaki[182], Louise Izatt[183], Anna Jakubowska[156], Paul James[85,184], Ramunas Janavicius[185], Wolfgang Janni[186], Uffe Birk Jensen[187], Esther M John[188,189], Nichola Johnson[136], Kristine Jones[174], Michael Jones[190], Arja Jukkola-Vuorinen[191], Rudolf Kaaks[41], Maria Kabisch[166], Katarzyna Kaczmarek[156], Daehee Kang[192–194], Karin Kast[195], kConFab/AOCS Investigators[14], Renske Keeman[102], Michael J Kerin[196], Carolien M Kets[197], Machteld Keupers[198], Sofia Khan[199], Elza Khusnutdinova[46,200], Johanna I Kiiski[199], Sung-Won Kim[158], Julia A Knight[201,202], Irene Konstantopoulou[203], Veli-Matti Kosma[169–171], Vessela N Kristensen[63,204,205], Torben A Kruse[206], Ava Kwong[207–209], Anne-Vibeke Lænkholm[210], Yael Laitman[139], Fiona Lalloo[211], Diether Lambrechts[55,56], Keren Landsman[212], Christine Lasset[213], Conxi Lazaro[214], Loic Le Marchand[215],

Julie Lecarpentier[3], Andrew Lee[3], Eunjung Lee[163], Jong Won Lee[216], Min Hyuk Lee[217], Flavio Lejbkowicz[212], Fabienne Lesueur[218–221], Jingmei Li[66], Jenna Lilyquist[222], Anne Lincoln[223], Annika Lindblom[224], Jolanta Lissowska[225], Wing-Yee Lo[67,68], Sibylle Loibl[226], Jirong Long[52], Jennifer T Loud[154], Jan Lubinski[156], Craig Luccarini[7], Michael Lush[3], Robert J MacInnis[1,2], Tom Maishman[117,118], Enes Makalic[2], Ivana Maleva Kostovska[227], Kathleen E Malone[228], Siranoush Manoukian[34], JoAnn E Manson[229], Sara Margolin[230], John W M Martens[177], Maria Elena Martinez[231,232], Keitaro Matsuo[233,234], Dimitrios Mavroudis[147], Sylvie Mazoyer[235], Catriona McLean[236], Hanne Meijers-Heijboer[13,237], Primitiva Menéndez[238], Jeffery Meyer[108], Hui Miao[239], Austin Miller[240], Nicola Miller[196], Gillian Mitchell[85,184], Marco Montagna[17], Kenneth Muir[241,242], Anna Marie Mulligan[243,244], Claire Mulot[245], Sue Nadesan[54], Katherine L Nathanson[65], NBSC Collaborators[14], Susan L Neuhausen[47], Heli Nevanlinna[199], Ines Nevelsteen[198], Dieter Niederacher[246], Sune F Nielsen[60,61], Børge G Nordestgaard[60–62], Aaron Norman[165], Robert L Nussbaum[247], Edith Olah[64], Olufunmilayo I Olopade[248], Janet E Olson[165], Curtis Olswold[165], Kai-ren Ong[249], Jan C Oosterwijk[250], Nick Orr[136], Ana Osorio[44,45], V Shane Pankratz[251], Laura Papi[252], Tjoung-Won Park-Simon[58], Ylva Paulsson-Karlsson[253], Rachel Lloyd[254], Inge Søkilde Pedersen[255], Bernard Peissel[34], Ana Peixoto[256], Jose I A Perez[257], Paolo Peterlongo[258], Julian Peto[116], Georg Pfeiler[159], Catherine M Phelan[259], Mila Pinchev[212], Dijana Plaseska-Karanfilska[227], Bruce Poppe[97], Mary E Porteous[260], Ross Prentice[31], Nadege Presneau[119], Darya Prokofieva[200], Elizabeth Pugh[15], Miquel Angel Pujana[261], Katri Pylkäs[262,263], Brigitte Rack[114,186], Paolo Radice[264], Nazneen Rahman[265], Johanna Rantala[266], Christine Rappaport-Fuerhauser[158], Gad Rennert[212,267], Hedy S Rennert[212,267], Valerie Rhenius[7], Kerstin Rhiem[160–162], Andrea Richardson[268], Gustavo C Rodriguez[269], Atocha Romero[82,270], Jane Romm[15], Matti A Rookus[271], Anja Rudolph[41], Thomas Ruediger[272], Emmanouil Saloustros[273], Joyce Sanders[274], Dale P Sandler[275], Suleeporn Sangrajrang[276], Elinor J Sawyer[277], Daniel F Schmidt[2], Minouk J Schoemaker[190], Fredrick Schumacher[278], Peter Schürmann[58], Lukas Schwentner[186], Christopher Scott[165], Rodney J Scott[279,280], Sheila Seal[265], Leigha Senter[281], Caroline Seynaeve[177], Mitul Shah[7], Priyanka Sharma[282], Chen-Yang Shen[283,284], Xin Sheng[163], Hermela Shimelis[108], Martha J Shrubsole[52], Xiao-Ou Shu[52], Lucy E Side[285], Christian F Singer[158], Christof Sohn[286], Melissa C Southey[287], John J Spinelli[288,289], Amanda B Spurdle[6], Christa Stegmaier[290], Dominique Stoppa-Lyonnet[92], Grzegorz Sukiennicki[156], Harald Surowy[79,80], Christian Sutter[291], Anthony Swerdlow[190,292], Csilla I Szabo[293], Rulla M Tamimi[9,125,126], Yen Y Tan[158], Jack A Taylor[275,294], Maria-Isabel Tejada[295], Maria Tengström[169,296,297], Soo H Teo[298,299], Mary B Terry[300], Daniel C Tessier[35], Alex Teulé[301], Kathrin Thöne[135], Darcy L Thull[302], Maria Grazia Tibiletti[303], Laima Tihomirova[304], Marc Tischkowitz[138,305], Amanda E Toland[306], Rob A E M Tollenaar[307], Ian Tomlinson[308], Ling Tong[18], Diana Torres[166,309], Martine Tranchant[11], Thérèse Truong[101], Kathy Tucker[310], Nadine Tung[311], Jonathan Tyrer[7], Hans-Ulrich Ulmer[312], Celine Vachon[165], Christi J van Asperen[313], David Van Den Berg[163], Ans M W van den Ouweland[314], Elizabeth J van Rensburg[115], Liliana Varesco[315], Raymonda Varon-Mateeva[316], Ana Vega[317,318], Alessandra Viel[319], Joseph Vijai[223], Daniel Vincent[35], Jason Vollenweider[108], Lisa Walker[320], Zhaoming Wang[73,321], Shan Wang-Gohrke[186], Barbara Wappenschmidt[160–162], Clarice R Weinberg[322], Jeffrey N Weitzel[49], Camilla Wendt[230], Jelle Wesseling[102,274], Alice S Whittemore[189,323], Juul T Wijnen[112,314], Walter Willett[126,324], Robert Winqvist[262,263], Alicja Wolk[164], Anna H Wu[163], Lucy Xia[163], Xiaohong R Yang[73], Drakoulis Yannoukakos[203], Daniela Zaffaroni[34], Wei Zheng[52], Bin Zhu[174], Argyrios Ziogas[25], Elad Ziv[325], Kristin K Zorn[303], Manuela Gago-Dominguez[87,231], Arto Mannermaa[169–171], Håkan Olsson[74], Manuel R Teixeira[256,326], Jennifer Stone[254,327], Kenneth Offit[328,329], Laura Ottini[330], Sue K Park[192–194], Mads Thomassen[206], Per Hall[66,331], Alfons Meindl[332], Rita K Schmutzler[160–162], Arnaud Droit[11], Gary D Bader[10,335], Paul D P Pharoah[3,7,335], Fergus J Couch[108,335], Douglas F Easton[3,7,335], Peter Kraft[9,126,335], Georgia Chenevix-Trench[6,335], Montserrat García-Closas[73,335], Marjanka K Schmidt[102,333,335], Antonis C Antoniou[3,335] & Jacques Simard[11,335]

[1]Cancer Epidemiology and Intelligence Division, Cancer Council Victoria, Melbourne, Victoria, Australia. [2]Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Victoria, Australia. [3]Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [4]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, UK. [5]Department of Electron Microscopy/Molecular Pathology, The Cyprus Institute of Neurology and Genetics, Nicosia, Cyprus. [6]Cancer Division, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [7]Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK. [8]Department of Epidemiology, University of Washington School of Public Health, Seattle, Washington, USA. [9]Program in Genetic Epidemiology and Statistical Genetics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA. [10]Donnelly Centre, University of Toronto, Toronto, Ontario, Canada. [11]Genomics Center, Centre Hospitalier Universitaire de Québec Research Center, Laval University, Québec City, Québec, Canada. [12]Department of Mathematics, Massachusetts Institute of Technology,

Cambridge, Massachusetts, USA. [13]Department of Clinical Genetics, Academic Medical Center, Amsterdam, the Netherlands. [14]A list of members and affiliations appears in the **Supplementary Note**. [15]Center for Inherited Disease Research (CIDR), Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA. [16]Yorkshire Regional Genetics Service, Chapel Allerton Hospital, Leeds, UK. [17]Immunology and Molecular Oncology Unit, Istituto Oncologico Veneto (IOV), IRCCS, Padua, Italy. [18]Center for Cancer Epidemiology and Prevention, University of Chicago, Chicago, Illinois, USA. [19]Department of Clinical Genetics, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [20]Personalised Medicine Team, QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia. [21]Roswell Park Cancer Institute, Buffalo, New York, USA. [22]Center for Genomic Medicine, Department of Biomedical Data Science, Geisel School of Medicine, Dartmouth College, Lebanon, New Hampshire, USA. [23]Fred A. Litwin Center for Cancer Genetics, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada. [24]Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada. [25]Department of Epidemiology, University of California, Irvine, Irvine, California, USA. [26]N.N. Alexandrov Research Institute of Oncology and Medical Radiology, Minsk, Belarus. [27]Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Heidelberg, Germany. [28]Institute of Clinical Molecular Biology / Department of Gynecology and Obstetrics, University Hospital of Schleswig-Holstein, Campus Kiel, Christian-Albrechts University Kiel, Kiel, Germany. [29]Department of Public Health Sciences and Cancer Research Institute, Queen's University, Kingston, Ontario, Canada. [30]Institute of Human Genetics, Hannover Medical School, Hannover, Germany. [31]Cancer Prevention Program, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [32]Zilber School of Public Health, University of Wisconsin–Milwaukee, Milwaukee, Wisconsin, USA. [33]Department of Medical Genetics, University Medical Center Utrecht, Utrecht, the Netherlands. [34]Unit of Medical Genetics, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto di Ricovero e Cura a Carattere Scientifico), Istituto Nazionale dei Tumori (INT), Milan, Italy. [35]McGill University and Génome Québec Innovation Centre, Montréal, Québec, Canada. [36]Department of Medical Oncology, University Hospital, Vall d'Hebron, Barcelona, Spain. [37]Division of Cancer Prevention and Genetics, Istituto Europeo di Oncologia, Milan, Italy. [38]Bâtiment Cheney D, Centre Léon Bérard, Lyon, France. [39]Laboratory of Cell Biology, Department of Pathology, Landspitali, Reykjavik, Iceland. [40]BMC (Biomedical Centre), Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [41]Division of Cancer Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany. [42]Department of Gynaecology and Obstetrics, University Hospital Erlangen, Friedrich Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany. [43]Human Genotyping Unit -Centro Nacional de Genotipado (CEGEN), Human Cancer Genetics Programme, Spanish National Cancer Research Centre (CNIO), Madrid, Spain. [44]Human Genetics Group, Human Cancer Genetics Programme, Spanish National Cancer Centre (CNIO), Madrid, Spain. [45]Spanish Network on Rare Diseases (CIBERER), Madrid, Spain. [46]Institute of Biochemistry and Genetics, Ufa Scientific Center of the Russian Academy of Sciences, Ufa, Russian Federation. [47]Department of Population Sciences, Beckman Research Institute of City of Hope, Duarte, California, USA. [48]Université Clermont Auvergne, INSERM, U1240, Imagerie Moléculaire et Stratégies Théranostiques, Centre Jean Perrin, Clermont-Ferrand, France. [49]Clinical Cancer Genetics, City of Hope, Duarte, California, USA. [50]Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, the Netherlands. [51]Department of Oncology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [52]Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University School of Medicine, Nashville, Tennessee, USA. [53]International Epidemiology Institute, Rockville, Maryland, USA. [54]City of Hope Clinical Cancer Genomics Community Research Network, Duarte, California, USA. [55]Vesalius Research Center, VIB, Leuven, Belgium. [56]Laboratory for Translational Genetics, Department of Oncology, University of Leuven, Leuven, Belgium. [57]Department of Radiation Oncology, Hannover Medical School, Hannover, Germany. [58]Gynaecology Research Unit, Hannover Medical School, Hannover, Germany. [59]Department of Clinical Genetics, Vejle Hospital, Vejle, Denmark. [60]Copenhagen General Population Study, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark. [61]Department of Clinical Biochemistry, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark. [62]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark. [63]Department of Cancer Genetics, Institute for Cancer Research, Oslo University Hospital Radiumhospitalet, Oslo, Norway. [64]Department of Molecular Genetics, National Institute of Oncology, Budapest, Hungary. [65]Department of Medicine, Abramson Cancer Center, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA. [66]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. [67]Dr. Margarete Fischer-Bosch-Institute of Clinical Pharmacology, Stuttgart, Germany. [68]University of Tübingen, Tübingen, Germany. [69]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg, Germany. [70]Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Heidelberg, Germany. [71]Gustave Roussy, Biopathology Department, Villejuif, France. [72]Department of Clinical Genetics, Royal Devon and Exeter Hospital, Exeter, UK. [73]Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, Maryland, USA. [74]Department of Cancer Epidemiology, Clinical Sciences, Lund University, Lund, Sweden. [75]Genome Sciences Centre, BC Cancer Agency, Vancouver, British Columbia, Canada. [76]Department of Biomedical Physiology and Kinesiology, Simon Fraser University, Burnaby, British Columbia, Canada. [77]Genetic Counseling Unit, Hereditary Cancer Program, IDIBGI (Institut d'Investigació Biomèdica de Girona), Catalan Institute of Oncology, Girona, Spain. [78]Institute for Prevention and Occupational Medicine of the German Social Accident Insurance, Institute of the Ruhr University Bochum (IPA), Bochum, Germany. [79]Department of Obstetrics and Gynecology, University of Heidelberg, Heidelberg, Germany. [80]Molecular Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. [81]Department of Medicine, Huntsman Cancer Institute, Salt Lake City, Utah, USA. [82]Medical Oncology Department, CIBERONC, Hospital Clínico San Carlos, Madrid, Spain. [83]Section of Molecular Genetics, Department of Laboratory Medicine, University of Pisa and University Hospital of Pisa, Pisa, Italy. [84]Research Department, Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia. [85]Sir Peter MacCallum Department of Oncology, University of Melbourne, Melbourne, Victoria, Australia. [86]Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. [87]Genomic Medicine Group, Galician Foundation of Genomic Medicine, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Complejo Hospitalario Universitario de Santiago, Servizo Galego de Saúde SERGAS, Santiago de Compostela, Spain. [88]Centro de Investigación en Red de Enfermedades Raras (CIBERER) and Centro Nacional de Genotipado (CEGEN-PRB2), Universidade de Santiago de Compostela, Santiago de Compostela, Spain. [89]Epidemiology Research Program, American Cancer Society, Atlanta, Georgia, USA. [90]Oncology and Genetics Unit, Instituto de Investigación Biomédica (IBI) de Orense-Pontevedra-Vigo, Xerencia de Xestión Integrada de Vigo, Servizo Galego de Saúde SERGAS, Vigo, Spain. [91]Centre François Baclesse, Caen, France. [92]Service de Génétique Oncologique and INSERM U830, Institut Curie, Paris, France - Université Paris Descartes, Sorbonne Paris Cité. [93]Cancer Genetics and Prevention Program, University of California, San Francisco, San Francisco, California, USA. [94]University Cancer Center Hamburg (UCCH), University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [95]Division of Cancer Prevention and Population Sciences, Roswell Park Cancer Institute, Buffalo, New York, USA. [96]Unité de Recherche en Santé des Populations, Centre des Maladies du Sein Deschênes-Fabia, Hôpital du Saint-Sacrement, Québec City, Québec, Canada. [97]Center for Medical Genetics, Ghent University, Ghent, Belgium. [98]Westmead Institute for Medical Research, University of Sydney, Sydney, New South Wales, Australia. [99]Huntsman Cancer Institute, Salt Lake City, Utah, USA. [100]Sheffield Clinical Genetics Service, Sheffield Children's Hospital, Sheffield, UK. [101]Cancer and Environment Group, Center for Research in Epidemiology and Population Health (CESP), INSERM, University Paris-Sud, University Paris-Saclay, Villejuif, France. [102]Division of Molecular Pathology, Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands. [103]Unité d'Oncogénétique, CHU Arnaud de Villeneuve, Montpellier, France. [104]Academic Unit of Molecular Oncology, Department of Oncology and Metabolism, University of Sheffield, Sheffield, UK. [105]Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK. [106]INSERM U1052, Cancer Research Center of Lyon, Lyon, France. [107]Academic Unit of Pathology, Department of Neuroscience, University of Sheffield, Sheffield, UK. [108]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA. [109]Department of Clinical Genetics, Fox Chase Cancer Center, Philadelphia, Pennsylvania, USA. [110]Department of Clinical Genetics, South Glasgow University Hospitals, Glasgow, UK. [111]Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands. [112]Department of Human Genetics, Leiden University Medical Center, Leiden, the Netherlands. [113]Oncogenetics Group, Vall d'Hebron Institute of Oncology (VHIO), Clinical and Molecular Genetics Area, Vall d'Hebron University Hospital, Barcelona, Spain. [114]Department of Gynecology and Obstetrics, Ludwig Maximilians University of Munich, Munich, Germany. [115]Cancer Genetics Laboratory, Department of Genetics, University of Pretoria, Arcadia, South Africa. [116]Department of Non-Communicable Disease Epidemiology, London School of Hygiene and Tropical Medicine, London, UK. [117]Southampton Clinical Trials Unit, Faculty of Medicine, University of Southampton, Southampton, UK. [118]Cancer Sciences Academic Unit, Faculty of Medicine, University of Southampton, Southampton, UK. [119]Department of Biomedical Sciences, Faculty of Science and Technology, University of Westminster, London, UK. [120]Institute of Human Genetics, University of Münster, Münster, Germany. [121]Oncogenetics Team, The Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK. [122]Department of Clinical Genetics, Lund University Hospital, Lund, Sweden. [123]Department of Oncology, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [124]Institute of Human Genetics, University Hospital Erlangen, Friedrich Alexander University Erlangen-Nuremberg, Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany. [125]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. [126]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA. [127]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany. [128]LIFE–Leipzig Research Centre for Civilization Diseases, University of Leipzig, Leipzig, Germany.

[129]Genetics Department, Dijon University Hospital, Dijon, France. [130]Oncogenetics, Centre Georges-François Leclerc, Dijon, France. [131]Division of Hematology and Oncology, Department of Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, California, USA. [132]Institute of Medical Genetics and Applied Genomics, University of Tübingen, Tübingen, Germany. [133]Usher Institute of Population Health Sciences and Informatics, CRUK Edinburgh Centre, University of Edinburgh Medical School, Edinburgh, UK. [134]Institute for Medical Biometrics and Epidemiology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [135]Department of Cancer Epidemiology, Clinical Cancer Registry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany. [136]Breast Cancer Now Toby Robins Research Centre, The Institute of Cancer Research, London, UK. [137]Department of Breast Surgery, Herlev and Gentofte Hospital, Copenhagen University Hospital, Herlev, Denmark. [138]Program in Cancer Genetics, Departments of Human Genetics and Oncology, McGill University, Montréal, Québec, Canada. [139]Susanne Levy Gertner Oncogenetics Unit, Institute of Human Genetics, Chaim Sheba Medical Center, Ramat Gan, Israel. [140]Sackler Faculty of Medicine, Tel Aviv University, Ramat Aviv, Israel. [141]School of Public Health, Curtin University, Perth, Western Australia, Australia. [142]Clinical Cancer Genetics Laboratory, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [143]Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. [144]Division of Cancer Prevention and Control Research, Jonsson Comprehensive Cancer Center, Schools of Medicine and Public Health, University of California, Los Angeles, Los Angeles, California, USA. [145]Cancer Risk and Prevention Clinic, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [146]Centre of Familial Breast and Ovarian Cancer, Department of Medical Genetics, Institute of Human Genetics, University Würzburg, Würzburg, Germany. [147]Department of Medical Oncology, University Hospital of Heraklion, Heraklion, Greece. [148]Department of Clinical Genetics, Rigshospitalet, Copenhagen, Denmark. [149]Department of Pathology and Laboratory Medicine, University of Kansas Medical Center, Kansas City, Kansas, USA. [150]Department of Medicine, McGill University, Montréal, Québec, Canada. [151]Division of Clinical Epidemiology, Royal Victoria Hospital, McGill University, Montréal, Québec, Canada. [152]Department of Dermatology, Huntsman Cancer Institute, University of Utah School of Medicine, Salt Lake City, Utah, USA. [153]Department of Obstetrics and Gynecology, Ohio State University James Comprehensive Cancer Center, Columbus, Ohio, USA. [154]Clinical Genetics Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, US National Institutes of Health, Bethesda, Maryland, USA. [155]Department of Surgery, Oulu University Hospital, University of Oulu, Oulu, Finland. [156]Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland. [157]Centre de Recherche du Centre Hospitalier de Université de Montréal (CHUM), Montréal, Québec, Canada. [158]Department of Obstetrics and Gynaecology and Comprehensive Cancer Centre, Medical University of Vienna, Vienna, Austria. [159]Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK. [160]Center for Familial Breast and Ovarian Cancer, University Hospital of Cologne, Cologne, Germany. [161]Center for Integrated Oncology (CIO), University Hospital of Cologne, Cologne, Germany. [162]Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany. [163]Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California, USA. [164]Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden. [165]Department of Health Sciences Research, Mayo Clinic, Rochester, Minnesota, USA. [166]Molecular Genetics of Breast Cancer, German Cancer Research Center (DKFZ), Heidelberg, Germany. [167]Department of Biostatistics and Epidemiology, University of Massachusetts, Amherst, Amherst, Massachusetts, USA. [168]Center for Genomic Medicine, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark. [169]Translational Cancer Research Area, University of Eastern Finland, Kuopio, Finland. [170]Institute of Clinical Medicine, Pathology and Forensic Medicine, University of Eastern Finland, Kuopio, Finland. [171]Imaging Center, Department of Clinical Pathology, Kuopio University Hospital, Kuopio, Finland. [172]Institute of Genetic Medicine, Centre for Life, Newcastle-upon-Tyne Hospitals NHS Trust, Newcastle-upon-Tyne, UK. [173]School of Population Health, University of Western Australia, Perth, Western Australia, Australia. [174]Cancer Genomics Research Laboratory (CGR), Division of Cancer Epidemiology and Genetics, National Cancer Institute, Bethesda, Maryland, USA. [175]Medical Genetics Unit, St George's, University of London, London, UK. [176]Family Cancer Clinic, Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands. [177]Department of Medical Oncology, Family Cancer Clinic, Erasmus MC Cancer Institute, Rotterdam, the Netherlands. [178]Center for Medical Genetics, NorthShore University HealthSystem, Evanston, Illinois, USA. [179]Pritzker School of Medicine, University of Chicago, Evanston, Illinois, USA. [180]N.N. Petrov Institute of Oncology, St. Petersburg, Russian Federation. [181]Lombardi Comprehensive Cancer Center, Georgetown University, Washington, DC, USA. [182]Division of Epidemiology, Center for Public Health Sciences, National Cancer Center, Tokyo, Japan. [183]Clinical Genetics, Guy's and St Thomas' NHS Foundation Trust, London, UK. [184]Familial Cancer Centre, Peter MacCallum Cancer Centre, Melbourne, Victoria, Australia. [185]State Research Institute Centre for Innovative Medicine, Vilnius, Lithuania. [186]Department of Gynaecology and Obstetrics, University of Ulm, Ulm, Germany. [187]Department of Clinical Genetics, Aarhus University Hospital, Aarhus, Denmark. [188]Department of Epidemiology, Cancer Prevention Institute of California, Fremont, California, USA. [189]Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California, USA. [190]Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK. [191]Department of Oncology, Oulu University Hospital, University of Oulu, Oulu, Finland. [192]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea. [193]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Republic of Korea. [194]Cancer Research Institute, Seoul National University, Seoul, Republic of Korea. [195]Department of Gynecology and Obstetrics, Medical Faculty and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany. [196]School of Medicine, National University of Ireland, Galway, Ireland. [197]Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands. [198]Leuven Multidisciplinary Breast Center, Department of Oncology, Leuven Cancer Institute, University Hospitals Leuven, Leuven, Belgium. [199]Department of Obstetrics and Gynecology, Helsinki University Hospital, University of Helsinki, Helsinki, Finland. [200]Department of Genetics and Fundamental Medicine, Bashkir State University, Ufa, Russian Federation. [201]Prosserman Centre for Health Research, Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada. [202]Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada. [203]Molecular Diagnostics Laboratory, INRASTES, National Centre for Scientific Research 'Demokritos', Athens, Greece. [204]Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway. [205]Department of Clinical Molecular Biology, Oslo University Hospital, University of Oslo, Oslo, Norway. [206]Department of Clinical Genetics, Odense University Hospital, Odense, Denmark. [207]Hong Kong Hereditary Breast Cancer Family Registry, Hong Kong. [208]Department of Surgery, University of Hong Kong, Hong Kong. [209]Department of Surgery, Hong Kong Sanatorium and Hospital, Hong Kong. [210]Department of Pathology, University Hospital of Region Zealand, Division Slagelse, Slagelse, Denmark. [211]Genetic Medicine, Manchester Academic Health Sciences Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK. [212]Clalit National Cancer Control Center, Haifa, Israel. [213]Unité de Prévention et d'Epidémiologie Génétique, Centre Léon Bérard, Lyon, France. [214]Molecular Diagnostic Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. [215]University of Hawaii Cancer Center, Honolulu, Hawaii, USA. [216]Department of Surgery, University of Ulsan College of Medicine and Asan Medical Center, Seoul, Republic of Korea. [217]Department of Surgery, Soonchunhyang University and Hospital, Seoul, Republic of Korea. [218]Institut Curie, Paris, France. [219]PSL Research University, Paris, France. [220]INSERM U900, Paris, France. [221]Mines Paris Tech, Fontainebleau, France. [222]Department of Health Sciences Research, Mayo Clinic, Scottsdale, Arizona, USA. [223]Clinical Genetics Research Laboratory, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [224]Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, Sweden. [225]Department of Cancer Epidemiology and Prevention, M. Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. [226]German Breast Group, Neu Isenburg, Germany. [227]Research Centre for Genetic Engineering and Biotechnology 'Georgi D. Efremov', Macedonian Academy of Sciences and Arts, Skopje, Macedonia. [228]Division of Public Health Sciences, Epidemiology Program, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. [229]Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts, USA. [230]Department of Oncology, Pathology, Karolinska Institutet, Stockholm, Sweden. [231]Moores Cancer Center, University of California, San Diego, La Jolla, California, USA. [232]Department of Family Medicine and Public Health, University of California, San Diego, La Jolla, California, USA. [233]Division of Molecular and Clinical Epidemiology, Aichi Cancer Center Research Institute, Nagoya, Japan. [234]Department of Epidemiology, Nagoya University Graduate School of Medicine, Nagoya, Japan. [235]Lyon Neuroscience Research Center–CRNL, INSERM U1028, CNRS UMR 5292, University of Lyon, Lyon, France. [236]Anatomical Pathology, Alfred Hospital, Melbourne, Victoria, Australia. [237]Department of Clinical Genetics, VU University Medical Centre, Amsterdam, the Netherlands. [238]Servicio de Anatomía Patológica, Hospital Monte Naranco, Oviedo, Spain. [239]Saw Swee Hock School of Public Health, National University of Singapore, Singapore. [240]NRG Oncology, Statistics and Data Management Center, Roswell Park Cancer Institute, Buffalo, New York, USA. [241]Institute of Population Health, University of Manchester, Manchester, UK. [242]Division of Health Sciences, Warwick Medical School, Warwick University, Coventry, UK. [243]Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada. [244]Laboratory Medicine Program, University Health Network, Toronto, Ontario, Canada. [245]Université Paris Sorbonne Cité, INSERM UMRS 1147, Paris, France. [246]Department of Gynecology and Obstetrics, University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf, Germany. [247]Department of Medicine, University of California, San Francisco,San Francisco, California, USA. [248]Center for Clinical Cancer Genetics and Global Health, University of Chicago, Chicago, Illinois, USA. [249]West Midlands Regional Genetics Service, Birmingham Women's Hospital Healthcare NHS Trust, Edgbaston, Birmingham, UK. [250]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands. [251]University of New Mexico Health Sciences Center, University of New Mexico, Albuquerque, New Mexico, USA. [252]Unit of Medical

Genetics, Department of Biomedical, Experimental and Clinical Sciences, University of Florence, Florence, Italy. [253]Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. [254]Curtin UWA Centre for Genetic Origins of Health and Disease, Curtin University and University of Western Australia, Perth, Western Australia, Australia. [255]Section of Molecular Diagnostics, Clinical Biochemistry, Aalborg University Hospital, Aalborg, Denmark. [256]Department of Genetics, Portuguese Oncology Institute, Porto, Portugal. [257]Servicio de Cirugía General y Especialidades, Hospital Monte Naranco, Oviedo, Spain. [258]IFOM, FIRC (Italian Foundation for Cancer Research) Institute of Molecular Oncology, Milan, Italy. [259]Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, Florida, USA. [260]South East of Scotland Regional Genetics Service, Western General Hospital, Edinburgh, UK. [261]ProCURE, Catalan Institute of Oncology, IDIBELL (Bellvitge Biomedical Research Institute), Barcelona, Spain. [262]Laboratory of Cancer Genetics and Tumor Biology, Cancer and Translational Medicine Research Unit, Biocenter Oulu, University of Oulu, Oulu, Finland. [263]Laboratory of Cancer Genetics and Tumor Biology, Northern Finland Laboratory Centre Oulu, Oulu, Finland. [264]Unit of Molecular Bases of Genetic Risk and Genetic Testing, Department of Preventive and Predictive Medicine, Fondazione IRCCS (Istituto di Ricovero e Cura a Carattere Scientifico), Istituto Nazionale dei Tumori (INT), Milan, Italy. [265]Section of Cancer Genetics, The Institute of Cancer Research, London, UK. [266]Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. [267]Carmel Medical Center and B. Rappaport Faculty of Medicine-Technion, Haifa, Israel. [268]Brigham and Women's Hospital, Dana-Farber Cancer Institute, Boston, Massachusetts, USA. [269]Division of Gynecologic Oncology, NorthShore University HealthSystem, University of Chicago, Evanston, Illinois, USA. [270]Medical Oncology Department, Hospital Universitario Puerta de Hierro, Madrid, Spain. [271]Department of Epidemiology, Netherlands Cancer Institute, Amsterdam, the Netherlands. [272]Institute of Pathology, Staedtisches Klinikum Karlsruhe, Karlsruhe, Germany. [273]Hereditary Cancer Clinic, University Hospital of Heraklion, Heraklion, Greece. [274]Department of Pathology, Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands. [275]Epidemiology Branch, National Institute of Environmental Health Sciences, US National Institutes of Health, Research Triangle Park, North Carolina, USA. [276]National Cancer Institute, Bangkok, Thailand. [277]Department of Population and Quantitative Health Sciences, Case Western Reserve University, Cleveland, Ohio, USA. [278]Research Oncology, Guy's Hospital, King's College London, London, UK. [279]Division of Molecular Medicine, Pathology North, John Hunter Hospital, Newcastle, New South Wales, Australia. [280]Discipline of Medical Genetics, School of Biomedical Sciences and Pharmacy, Faculty of Health, University of Newcastle, Callaghan, New South Wales, Australia. [281]Clinical Cancer Genetics Program, Division of Human Genetics, Department of Internal Medicine, Comprehensive Cancer Center, Ohio State University, Columbus, Ohio, USA. [282]Department of Internal Medicine, University of Kansas Medical Center, Kansas City, Kansas, USA. [283]School of Public Health, China Medical University, Taichung, Taiwan. [284]Taiwan Biobank, Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan. [285]North East Thames Regional Genetics Service, Great Ormond Street Hospital for Children NHS Trust, London, UK. [286]National Center for Tumor Diseases, University of Heidelberg, Heidelberg, Germany. [287]Department of Pathology, University of Melbourne, Melbourne, Victoria, Australia. [288]Cancer Control Research, BC Cancer Agency, Vancouver, British Columbia, Canada. [289]School of Population and Public Health, University of British Columbia, Vancouver, British Columbia, Canada. [290]Saarland Cancer Registry, Saarbrücken, Germany. [291]Institute of Human Genetics, University Hospital Heidelberg, Heidelberg, Germany. [292]Division of Breast Cancer Research, The Institute of Cancer Research, London, UK. [293]National Human Genome Research Institute, US National Institutes of Health, Bethesda, Maryland, USA. [294]Epigenetic and Stem Cell Biology Laboratory, National Institute of Environmental Health Sciences, US National Institutes of Health, Research Triangle Park, North Carolina, USA. [295]Molecular Genetics Laboratory, Clinical Genetics Service, Cruces University Hospital and BioCruces Health Research Institute, Barakaldo, Spain. [296]Cancer Center, Kuopio University Hospital, Kuopio, Finland. [297]Institute of Clinical Medicine, Oncology, University of Eastern Finland, Kuopio, Finland. [298]Cancer Research Malaysia, Subang Jaya, Malaysia. [299]Breast Cancer Research Unit, Cancer Research Institute, University Malaya Medical Centre, Kuala Lumpur, Malaysia. [300]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, USA. [301]Genetic Counseling Unit, Hereditary Cancer Program, IDIBELL (Bellvitge Biomedical Research Institute), Catalan Institute of Oncology, Barcelona, Spain. [302]Magee-Womens Hospital, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania, USA. [303]Ospedale di Circolo ASST Settelaghi, Varese, Italy. [304]Latvian Biomedical Research and Study Centre, Riga, Latvia. [305]Department of Medical Genetics, Addenbrooke's Treatment Centre, Addenbrooke's Hospital, Cambridge, UK. [306]Department of Molecular Virology, Immunology and Medical Genetics, Comprehensive Cancer Center, Ohio State University, Columbus, Ohio, USA. [307]Department of Surgery, Leiden University Medical Center, Leiden, the Netherlands. [308]Wellcome Trust Centre for Human Genetics and Oxford NIHR Biomedical Research Centre, University of Oxford, Oxford, UK. [309]Institute of Human Genetics, Pontificia Universidad Javeriana, Bogotá, Colombia. [310]Hereditary Cancer Clinic, Department of Medical Oncology, Prince of Wales Hospital, Randwick, New South Wales, Australia. [311]Department of Medical Oncology, Beth Israel Deaconess Medical Center, Boston, Massachusetts, USA. [312]Frauenklinik der Stadtklinik Baden Baden, Baden Baden, Germany. [313]Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands. [314]Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands. [315]Unit of Hereditary Cancer, Department of Epidemiology, Prevention and Special Functions, IRCCS (Istituto di Ricovero e Cura a Carattere Scientifico) AOU San Martino, IST Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy. [316]Institute of Human Genetics, Campus Virchov Klinikum, Charite Berlin, Berlin, Germany. [317]Fundación Pública Galega de Medicina Xenómica, Servizo Galego de Saúde SERGAS, Instituto de Investigaciones Sanitarias (IDIS), Santiago de Compostela, Spain. [318]Grupo de Medicina Xenómica, Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), Universidade de Santiago de Compostela, Santiago de Compostela, Spain. [319]Unit of Functional Onco-genomics and Genetics, CRO Aviano, National Cancer Institute, Aviano, Italy. [320]Oxford Regional Genetics Service, Churchill Hospital, Oxford, UK. [321]Department of Computational Biology, St. Jude Children's Research Hospital, Memphis, Tennessee, USA. [322]Biostatistics and Computational Biology Branch, National Institute of Environmental Health Sciences, US National Institutes of Health, Research Triangle Park, North Carolina, USA. [323]Department of Biomedical Data Sciences, Stanford University School of Medicine, Stanford, California, USA. [324]Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA. [325]Department of Medicine, Institute for Human Genetics, UCSF Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, San Francisco, California, USA. [326]Biomedical Sciences Institute (ICBAS), University of Porto, Porto, Portugal. [327]Department of Obstetrics and Gynaecology, University of Melbourne and the Royal Women's Hospital, Melbourne, Victoria, Australia. [328]Clinical Genetics Research Laboratory, Cancer Biology and Genetics Program, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [329]Clinical Genetics Service, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, New York, USA. [330]Department of Molecular Medicine, University La Sapienza, Rome, Italy. [331]Department of Oncology, South General Hospital, Stockholm, Sweden. [332]Division of Gynaecology and Obstetrics, Technische Universität München, Munich, Germany. [333]Division of Psychosocial Research and Epidemiology, Netherlands Cancer Institute, Antoni van Leeuwenhoek Hospital, Amsterdam, the Netherlands. [334]These authors contributed equally to this work. [335]These authors jointly directed this work. Correspondence should be addressed to R.L.M. (roger.milne@cancervic.org.au).

ONLINE METHODS

Study subjects. Supplementary Table 1 summarizes the studies from BCAC that contributed data. The majority were case–control studies. Sixty-eight BCAC studies participated in the ER-negative breast cancer component of the OncoArray, contributing 9,655 cases and 45,494 controls. All studies provided core data on disease status and age at diagnosis/observation, and the majority provided information on clinicopathological and lifestyle factors, which has been curated and incorporated into the BCAC database (version 6). ER status for most (~70%) cases was obtained from clinical records. After removal of overlapping participants, genotype data were also available from eight GWAS[9,12,16,37,38] (4,480 ER-negative cases and 12,632 controls) and 40 studies previously genotyped using the Illumina iCOGS custom array[20] (7,333 ER-negative cases and 42,468 controls).

A total of 21,468 ER-negative cases were included in the combined analyses. Of these, 5,793 had tumors that were also negative for progesterone receptor (PR) and human epidermal growth factor receptor 2 (HER2) and were defined as triple negative. PR and HER2 status were also obtained predominantly from clinical records. A further 4,217 cases were positive for PR or HER2 and were considered non–triple negative. The remainder had unknown PR or HER2 status. All participating studies were approved by their appropriate ethics review boards, and all subjects provided informed consent.

Subjects included from CIMBA were women of European ancestry aged 18 years or older with a pathogenic variant for *BRCA1*. The majority of the participants were sampled through cancer genetics clinics. Multiple members of the same family were included in some instances. Fifty-eight studies from 24 countries contributed OncoArray genotype data. After quality control and removal of participants overlapping with the BCAC OncoArray study, data were available on 15,566 *BRCA1* mutation carriers, of whom 7,784 were affected with breast cancer (**Supplementary Table 2**). We also obtained iCOGS genotype data on 3,342 *BRCA1* mutation carriers (1,630 with breast cancer) from 54 studies through CIMBA. All mutation carriers provided written informed consent and participated under ethically approved protocols.

OncoArray SNP selection. Approximately 50% of the SNPs for the OncoArray were selected as a 'GWAS backbone' (Illumina HumanCore), which aimed to provide high coverage for the majority of common variants through imputation. The remaining allocation was selected from lists supplied by each of six disease-based consortia, together with a seventh list of SNPs of interest to multiple disease study groups. Approximately 72,000 SNPs were selected specifically for their relevance to breast cancer, on the basis of prior evidence of association with overall or subtype-specific disease, with breast density or with breast-tissue-specific gene expression. Lists were merged as described previously[34].

Genotype calling and quality control. Details of the genotype calling and quality control for the iCOGS and GWAS are described elsewhere[19,20,23,30], and those for OncoArray are described in the **Supplementary Note**.

Imputation. Genotypes for ~21 million SNPs were imputed for all samples using the October 2014 (Phase 3) release of the 1000 Genomes Project data as the reference panel and based on 800 haplotypes. The iCOGS, OncoArray and six of the GWAS data sets were imputed using a two-stage imputation approach, with SHAPEIT[73] for phasing and IMPUTEv2 (ref. 74) for imputation. Imputation was performed in 5-Mb non-overlapping intervals. All subjects were split into subsets of ~10,000 samples, with subjects from the same study grouped in the subset. The Breast and Prostate Cancer Cohort Consortium (BPC3) and Breast Cancer Family Registry (BCFR) GWAS performed imputation separately using MACH and Minimac[75,76]. We imputed genotypes for all SNPs that were polymorphic (MAF > 0.1%) in either European or Asian samples. For the BCAC GWAS, data were included in the analysis for all SNPs with MAF >0.01 and imputation r^2 >0.3. For iCOGS and OncoArray, we included data for all SNPs with imputation r^2 >0.3 and MAF >0.005.

Statistical analyses of BCAC data. Per-allele odds ratios and standard errors were generated for the OncoArray and iCOGS data sets and for each GWAS, adjusting for principal components, using logistic regression. The OncoArray and iCOGS analyses were additionally adjusted for country and study, respectively. For the OncoArray data set, principal-components analysis was performed using data for 33,661 SNPs (which included the 2,318 markers of continental ancestry) with MAF ≥0.05 and maximum correlation of 0.1, using purpose-written software (PCcalc; see URLs) to allow standard calculations to be performed sufficiently rapidly on a very large data set. We used the first ten principal components, as the inclusion of additional components did not further reduce inflation of the test statistics. We used nine principal components for the iCOGS data set and up to ten principal components for the other GWAS, where this was found to reduce inflation.

Odds ratio estimates were derived using MACH for the BCFR GWAS, ProbABEL[77] was used for the BPC3 GWAS, SNPTEST (see URLs) was used for the remaining GWAS and purpose-written software was applied for the iCOGS and OncoArray data sets. Odds ratio estimates and standard errors were combined by a fixed-effects inverse-variance-weighted meta-analysis using METAL[39]. This was first done across the eight GWAS, applying genomic control, as described previously[20]. It was then applied (without genomic control) to combine findings from the three BCAC genotyping initiatives (GWAS, iCOGS and OncoArray).

The independence of the signals from two variants at 11q22.3 was assessed by fitting the logistic regression models described above with both variants as covariates. This was done separately for iCOGS and OncoArray data, and results for each variant were combined by meta-analysis.

For selected SNPs, we estimated per-allele odds ratios by ER status using all available BCAC data for 82,263 cases with known ER status and 87,962 controls from the iCOGS and OncoArray studies. We also estimated per-allele odds ratios by triple-negative status (triple-negative versus other ER-negative subtypes) and tumor grade, using available BCAC data for ER-negative cases and corresponding controls. Tests for heterogeneity by subtype were derived by applying logistic regression to cases only. This was done separately for the iCOGS and OncoArray data sets, adjusting as before, and results were then combined in a fixed-effects meta-analysis. Multinomial regression was applied to cases only to test for a linear trend for tumor grade, with the model constrained so that the difference between grade 1 and grade 3 was double that for the difference between grade 2 and grade 3; this method was also used to test for a linear trend with age at diagnosis, with ordinal values 1, 2, 3 and 4 representing ages <40, 40–49, 50–59 and ≥60 years, respectively.

Statistical analyses of CIMBA data. Associations between genotype and breast cancer risk for *BRCA1* mutation carriers were evaluated using a 1-degree-of-freedom per-allele trend test (P_{trend}), based on modeling the retrospective likelihood of the observed genotypes conditional on breast cancer phenotypes[36]. This was done separately for iCOGS and OncoArray data. To allow for non-independence among related individuals, an adjusted test statistic was used that took into account correlation in genotypes[3]. All analyses were stratified by country of residence and, for countries where the strata were sufficiently large (United States and Canada), by Ashkenazi Jewish ancestry. The results from the iCOGS and OncoArray data sets were then pooled using fixed-effects meta-analysis. We repeated these analyses modeling ovarian cancer as a competing risk and observed no substantial difference in the results obtained.

The independence of the signals from two variants at 11q22.3 was assessed using OncoArray data only, fitting a Cox regression model with per-allele effects for both variants, adjusting for birth cohort, stratifying by country of residence, and using robust standard errors and clustered observations for relatives. This approach provides valid significance tests of association, although the resulting hazard ratio estimates can be biased[35].

Meta-analysis of BCAC and CIMBA. A fixed-effects meta-analysis of the results from BCAC and CIMBA was conducted using an inverse-variance-weighted approach assuming fixed effects, as implemented in METAL[39]. The effect estimates used were the logarithm of the per-allele hazard ratio estimate for association with breast cancer risk for *BRCA1* mutation carriers from CIMBA and the logarithm of the per-allele odds ratio estimate for association with risk of ER-negative breast cancer based on BCAC data, both of which were assumed to approximate the same relative risk. We assessed genomic inflation using common (MAF > 1%) GWAS backbone variants. As λ is influenced by sample size, we calculated $\lambda_{1,000}$ so that the values were comparable with those from other studies.

All statistical tests conducted were two-sided.

Definition of known association signals. We identified all associations previously reported from genome-wide or candidate analyses at a significance level of $P < 5 \times 10^{-8}$ in overall breast cancer, in ER-negative or ER-positive breast cancer, for *BRCA1* or *BRCA2* mutation carriers, or in meta-analyses of these categories. We included only one SNP in any 500-kb interval unless joint analysis provided genome-wide significant evidence (conditional $P < 5 \times 10^{-8}$) of more than one independent signal. Where multiple studies reported associations in the same region, we considered the first reported association unless a later study identified a different variant in the same region that was more strongly associated with breast cancer risk. One hundred and seven previously reported association signals were identified, 11 of which were found through GWAS of ER-negative disease or of breast cancer for *BRCA1* mutation carriers or were reported as being more strongly associated with ER-negative breast cancer. These are listed in **Table 2**. The other 96 previously reported association signals are listed in **Supplementary Table 10**.

Definition of new association signals. To search for new loci, we assessed all SNPs excluding those within 500 kb of a known association signal. This approach identified 206 SNPs in nine regions that were associated with disease risk at $P < 5 \times 10^{-8}$ in the meta-analysis of BCAC ER-negative breast cancer and breast cancer for CIMBA *BRCA1* mutation carriers. The SNP with the lowest *P* value from this analysis was considered to be the lead SNP. No additional loci were detected from analysis of BCAC data only. Imputation quality, as assessed by the IMPUTE2 imputation r^2 value in the OncoArray data set, was ≥0.89 for the ten lead SNPs reported (**Supplementary Table 3**).

Candidate causal SNPs. To define the set of potentially causal variants at each of the new susceptibility loci, we selected all variants with *P* values within two orders of magnitude of the *P* value for the most significant SNP at each of the ten new loci. This is approximately equivalent to selecting variants whose posterior probability of causality is within two orders of magnitude of that of the most significant SNP[40,41]. This approach was applied to identify potentially causal variants for the signal given by the more common lead SNP at 11q22.3 (rs11374964). A similar approach was applied for the rarer lead SNP at this locus (rs74911261), but this was based on *P* values from analyses adjusted for rs11374964.

Proportion of familial risk explained. The relative risk of ER-negative breast cancer for the first-degree female relative of a woman with ER-negative disease has not been estimated. We therefore assumed that the twofold increase in risk observed for overall disease also applied to ER-negative disease. To estimate the proportion of this risk explained by the 125 variants associated with ER-negative disease, we used MAF and odds ratio estimates from the OncoArray-based genotype data and applied the formula

$$\sum_i p_i(1-p_i)(\beta_i^2 - \tau_i^2)/\ln(\lambda)$$

where p_i is the MAF for variant *i*, β_i is the log(OR) estimate for variant *i*, τ_i is the standard error of β_i and $\lambda = 2$ is the assumed overall familial relative risk.

The corresponding estimate for the familial relative risk due to all variants is the frailty-scale heritability, defined as

$$h_f^2 = \sum_i 2p_i(1-p_i)\gamma_i^2$$

where γ_i is the true relative risk conferred by variant *i*, assuming a log-additive model. We first obtained the estimated heritability based on the full set of summary estimates using LD score regression[68], which derives a heritability estimate on the observed scale. We then converted this to an estimate on the frailty scale using the formula

$$h_f^2 = h_{\text{obs}}^2 / P(1-P)$$

where *P* is the proportion of samples in the population that are cases.

Proportion of polygenic risk-modifying variance explained for *BRCA1* mutation carriers. The proportion of the variance in polygenic frailty modifying risk for *BRCA1* mutation carriers explained by the set of associated SNPs was estimated by

$$\sum_i \ln c_i / \sigma^2$$

where c_i is the squared estimated coefficient of variation in incidence associated with SNP *i* (ref. 78) and σ^2 is the total polygenic variance, estimated from segregation data[79].

***In silico* annotation of candidate causal variants.** We combined multiple sources of *in silico* functional annotation from public databases to help identify potential functional SNPs and target genes, on the basis of previous observations that breast cancer susceptibility alleles are enriched in *cis*-regulatory elements and alter transcriptional activity[28,80–82]. The influence of candidate causal variants on transcription factor binding sites was determined using the ENCODE-Motifs resource[43]. To investigate functional elements enriched across the region encompassing the strongest candidate causal SNPs, we analyzed chromatin biofeatures data from the Encyclopedia of DNA Elements (ENCODE) Project[42] and the Roadmap Epigenomics Project[44] together with other data obtained through the NCBI Gene Expression Omnibus (GEO), namely chromatin state segmentation by hidden Markov models (chromHMM), DHSs, and histone modifications for the epigenetic marks H3K4, H3K9 and H3K27 in HMECs, myoepithelial cells, and T47D and MCF7 breast cancer cells and transcription factor ChIP–seq in a range of breast cell lines (**Supplementary Table 6**). To identify the SNPs most likely to be functional, we used RegulomeDB[45]; to identify putative target genes, we examined potential functional chromatin interactions between distal and proximal regulatory transcription factor binding sites and promoters in the risk regions, using Hi-C data generated in HMECs[47] and ChIA-PET data generated in MCF7 cells. This approach detects genome-wide interactions brought about by or associated with CCCTC-binding factor (CTCF), RNA polymerase II (Pol II) and ER, all of which are involved in transcriptional regulation[47]. Annotation of putative *cis*-regulatory regions and predicted target genes used the integrated method for predicting enhancer targets (IM-PET)[46], the predicting specific tissue interactions of genes and enhancers (PreSTIGE) algorithm[48], Hnisz[51] and FANTOM[49]. Intersections between candidate causal variants and regulatory elements were identified using Galaxy, BedTools v2.24 and HaploReg v4.1 and were visualized in the UCSC Genome Browser. Publically available eQTL databases, including Gene-Tissue Expression (GTEx[50]; version 6, multiple tissues) and Westra[52] (blood), were queried for candidate causal variants.

eQTL analyses. eQTL analyses were performed using data from the TCGA and METABRIC projects[59,60].

The TCGA eQTL analysis was based on 79 ER-negative breast tumors that had matched gene expression, copy number and methylation profiles together with corresponding germline genotypes available. All 79 individuals were of European ancestry, as ascertained using the genotype data and the local ancestry in admixed populations (LAMP) software package (LAMP estimate cutoff > 95% European)[83]. Germline genotypes were imputed into the 1000 Genomes Project reference panel (October 2014 release) using IMPUTE2 (refs. 75,84). Gene expression had been measured on the Illumina HiSeq 2000 RNA-seq platform (gene-level RSEM normalized counts[85]), copy number estimates were derived from the Affymetrix SNP 6.0 array (somatic copy number alteration minus germline copy number variation called using the GISTIC2 algorithm[86]) and methylation β values were measured on the Illumina Infinium HumanMethylation450 array, as previously described[59]. Primary TCGA eQTL analysis focused on all potentially causal variants in the ten new regions associated with breast cancer risk in the meta-analysis of ER-negative cases and controls from BCAC and *BRCA1* mutation carriers from CIMBA. We considered all genes located up to 1 Mb away on either side of each of these variants. The effects of tumor copy number and methylation on gene expression were first removed using a method described previously[58], and eQTL analysis was performed by linear regression as implemented in the R package Matrix eQTL[87].

The METABRIC eQTL analysis was based on 135 normal breast tissue samples resected from patients with breast cancer of European ancestry. Germline genotyping for the METABRIC study was also performed on the Affymetrix

doi:10.1038/ng.3785

SNP 6.0 array, and ancestry estimation and imputation for this data set were conducted as described for TCGA. Gene expression in the METABRIC study had been measured using the Illumina HT12 microarray platform, and we used probe-level estimates. As for TCGA, we considered all genes in ten regions using Matrix eQTL.

We also performed additional eQTL analyses using the METABRIC data set for all variants within 1 Mb of *L3MBTL3* and *CDH2* and expression of these specific genes.

Global genomic enrichment analyses. We performed stratified LD score regression analyses[68] for ER-negative breast cancer using the summary statistics based on meta-analyses of the OncoArray, GWAS, iCOGS and CIMBA data sets. We used all SNPs in the 1000 Genomes Project Phase 1 v3 release that had MAF >1% and imputation quality score R^2 >0.3 in the OncoArray data. LD scores were calculated using the 1000 Genomes Project Phase 1 v3 EUR panel. Further details are provided in the **Supplementary Note**.

We tested the differences in functional enrichment between ER-positive and ER-negative subsets for individual features through a Wald test, using the regression coefficients and standard errors for the two subsets based on the models described above. Finally, we assessed the heritability due to genotyped and imputed SNPs[70] and estimated the genetic correlation between ER-positive and ER-negative breast cancer[69]. The genetic correlation analysis was restricted to the ~1 million SNPs included in HapMap 3.

Pathway enrichment analyses. The pathway gene set database Human_GOBP_AllPathways_no_GO_iea_January_19_2016_symbol.gmt (GeneSets; see URLs)[61] was used for all analyses. Pathway size was determined by the total number of genes in the pathway to which SNPs in the imputed GWAS data set could be mapped. To provide more biologically meaningful results and reduce false positives, only pathways that contained between 10 and 200 genes were considered.

SNPs were mapped to the nearest gene within 500 kb; SNPs that were further than 500 kb away from any gene were excluded. Gene significance was calculated by assigning the lowest *P* value observed across all SNPs mapped to a gene[63,64], on the basis of the meta-analysis of BCAC and CIMBA data described above.

The gene set enrichment analysis (GSEA)[61] algorithm, as implemented in the GenGen package (see URLs)[62,63], was used to perform pathway analysis. Briefly, the algorithm calculates an enrichment score (ES) for each pathway based on a weighted Kolmogorov–Smirnov statistic[62]. Pathways that have most of their genes at the top of the ranked list of genes obtain higher ES values.

We defined an ES threshold (ES ≥ 0.41) to yield a true positive rate (TPR) of 0.20 and a false positive rate (FPR) of 0.14, with true positive pathways defined as those observed with false discovery rate (FDR) < 0.05 in a prior analysis carried out using the analytic approach defined above applied to iCOGS data for ER-negative disease.

To visualize the pathway enrichment analysis results, an enrichment map was created using the Enrichment Map (EM) v 2.1.0 app[61] in Cytoscape v3.30 (ref. 88), applying an edge-weighted force-directed layout. To measure the contribution of each gene to enriched pathways and annotate the map, we reran the pathway enrichment analysis multiple times, each time excluding one gene. A gene was considered to drive the enrichment if the ES dropped to zero or less (pathway enrichment driver) after it was excluded. Pathways were grouped in the map if they shared >70% of their genes or their enrichment was driven by a shared gene.

See the **Supplementary Note** for further details.

Data availability. A subset of the data that support the findings of this study is publically available via dbGaP (see URLs; accessions phs001265.v1.p1 for BCAC data and phs001321.v1.p1 for CIMBA data). Requests for data can be made to the corresponding author or the Data Access Coordination Committees (DACCs) of BCAC (see URLs) and CIMBA (see URLs). BCAC DACC approval is required to access data from the ABCFS, ABCS, ABCTB, BBCC, BBCS, BCEES, BCFR-NY, BCFR-PA, BCFR-UT, BCINIS, BSUCH, CBCS, CECILE, CGPS, CTS, DIETCOMPLYF, ESTHER, GC-HBOC, GENICA, GEPARSIXTO, GESBC, HABCS, HCSC, HEBCS, HMBCS, HUBCS, KARBAC, KBCP, LMBC, MABCS, MARIE, MBCSG, MCBCS, MISS, MMHS, MTLGEBCS, NC-BCFR, OFBCR, ORIGO, pKARMA, POSH, PREFACE, RBCS, SKKDKFZS, SUCCESSB, SUCCESSC, SZBCS, TNBCC, UCIBCS, UKBGS and UKOPS studies (**Supplementary Table 1**). CIMBA DACC approval is required to access data from the BCFR-ON, CONSIT TEAM, DKFZ, EMBRACE, FPGMX, GC-HBOC, GEMO, G-FAST, HEBCS, HEBON, IHCC, INHERIT, IOVHBOCS, IPOBCS, MCGILL, MODSQUAD, NAROD, OCGN, OUH and UKGRFOCR studies (**Supplementary Table 2**).

73. Delaneau, O., Marchini, J. & Zagury, J.F. A linear complexity phasing method for thousands of genomes. *Nat. Methods* **9**, 179–181 (2011).
74. Howie, B.N., Donnelly, P. & Marchini, J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet.* **5**, e1000529 (2009).
75. Howie, B., Fuchsberger, C., Stephens, M., Marchini, J. & Abecasis, G.R. Fast and accurate genotype imputation in genome-wide association studies through pre-phasing. *Nat. Genet.* **44**, 955–959 (2012).
76. Li, Y., Willer, C.J., Ding, J., Scheet, P. & Abecasis, G.R. MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genet. Epidemiol.* **34**, 816–834 (2010).
77. Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).
78. Antoniou, A.C. & Easton, D.F. Polygenic inheritance of breast cancer: implications for design of association studies. *Genet. Epidemiol.* **25**, 190–202 (2003).
79. 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).
80. Darabi, H. *et al.* Polymorphisms in a putative enhancer at the 10q21.2 breast cancer risk locus regulate *NRBF2* expression. *Am. J. Hum. Genet.* **97**, 22–34 (2015).
81. Glubb, D.M. *et al.* Fine-scale mapping of the 5q11.2 breast cancer locus reveals at least three independent risk variants regulating *MAP3K1*. *Am. J. Hum. Genet.* **96**, 5–20 (2015).
82. Ghoussaini, M. *et al.* Evidence that breast cancer risk at the 2q35 locus is mediated through *IGFBP5* regulation. *Nat. Commun.* **4**, 4999 (2014).
83. Baran, Y. *et al.* Fast and accurate inference of local ancestry in Latino populations. *Bioinformatics* **28**, 1359–1367 (2012).
84. Abecasis, G.R. *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
85. Li, B. & Dewey, C.N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
86. Mermel, C.H. *et al.* GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **12**, R41 (2011).
87. Shabalin, A.A. Matrix eQTL: ultra fast eQTL analysis via large matrix operations. *Bioinformatics* **28**, 1353–1358 (2012).
88. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).