Supplementary Table 1. Per-allele OR and 95% CIs for the association of established cancer susceptibility loci and breast cancer risk overall in PBCS

Locus (Neighborhood genes)	rs number	MAF[a]	Controls, N	Cases, N	OR (95% CI)[b]	*P*	Reference[c]
10q26 (FGFR2)	rs2981582 (C/T)[d]	0.353	2249	1698	1.29 (1.18-1.41)	1.28E-08	Easton DF, 2007[1]
11q13 (ORAOV1)	rs614367 (G/A)	0.195	2174	1639	1.24 (1.12-1.37)	3.95E-05	Turnbull C, 2010[2]
10q21.2 (ZNF365)	rs10995190 (G/A)	0.207	2173	1633	0.79 (0.71-0.89)	5.92E-05	Turnbull C, 2010[2]
2q35 (TNP1/IGFBP5/IGFBP2/TNS1)	rs13387042 (T/C)	0.446	2324	1850	0.86 (0.79-0.93)	3.16E-04	Stacey SN, 2007[3]
8q24 (FAM84B/c-MYC/POU5F1P1)	rs13281615 (A/G)	0.459	2339	1860	1.13 (1.04-1.23)	0.005	Easton DF, 2007[1]
1p11.2 (NOTCH2/FCGR1B)	rs11249433 (A/G)	0.369	2340	1856	1.12 (1.03-1.22)	0.007	Thomas G, 2009[4]
6q25.1 (ESR1)	rs2046210 (G/A)	0.294	2291	1770	1.13 (1.03-1.24)	0.01	Zheng W, 2009[5]
5p12 (MRPS30/FGFR10)	rs10941679 (A/G)	0.259	2366	1885	1.13 (1.03-1.24)	0.01	Stacey SN, 2008[6]
9p21 (CDKN2B)	rs1011970 (G/T)	0.128	2163	1637	1.17 (1.03-1.33)	0.01	Turnbull C, 2010[2]
14q24.1 (RAD51L1)	rs10483813 (T/A)	0.239	2362	1878	0.90 (0.82-0.99)	0.04	Thomas G, 2009[4]
5p12 (MRPS30/FGFR10)	rs4415084 (C/T)	0.421	2232	1751	1.09 (1.00-1.18)	0.05	Stacey SN, 2008[6]
11p15 (LSP1)	rs3817198 (T/C)	0.324	2349	1863	1.07 (0.98-1.17)	0.13	Easton DF, 2007[1]
5q11 (MAP3K1)	rs889312 (A/C)	0.261	2285	1775	1.07 (0.98-1.18)	0.14	Easton DF, 2007[1]
10q22.3 (ZMIZ1)	rs704010 (C/T)	0.330	2151	1631	1.07 (0.98-1.17)	0.16	Turnbull C, 2010[2]
17q23 (COX11)	rs6504950 (G/A)	0.259	2361	1878	0.94 (0.85-1.03)	0.16	Ahmed S, 2009[7]
19p13.1 (C19orf62)	rs8170 (G/A)	0.208	2156	1639	1.02 (0.92-1.14)	0.71	Antoniou AC, 2010[8]
10p15 (ANKRD16)	rs2380205 (G/A)	0.478	2185	1643	0.99 (0.9-1.07)	0.75	Turnbull C, 2010[2]
3p24 (NEK10/SLC4A7)	rs4973768 (G/A)	0.490	2269	1767	1.00 (0.92-1.09)	0.97	Ahmed S, 2009[7]

[a]Minor allele frequency over all controls in PBCS. [b]ORs estimated from logistic regression models adjusted for age in 5-year categories and study site. ORs for each SNP are for the minor allele. [c]Reference for manuscript in which SNP was first identified. [d]Major/minor allele.

Supplementary Table 2. Per-allele OR and 95% CI for the association of established cancer susceptibility loci and breast cancer risk by E-cadherin tumor expression in PBCS

			Case-control analyses				Case-only
Locus (Gene)	**rs number**	**Controls, N**	**N**	**OR (95% CI)**	**N**	**OR (95% CI)**	***P*-het[a]**
6q25.1 (ESR1)	rs2046210 (G/A)[b]	2291	239	1.43 (1.18-1.75)	949	1.06 (0.94-1.19)	0.007
1p11.2 (NOTCH2/FCGR1B)	rs11249433 (A/G)	2340	250	1.30 (1.08-1.56)	989	1.06 (0.95-1.18)	0.05
10p15 (ANKRD16)	rs2380205 (G/A)	2185	215	0.85 (0.70-1.04)	881	1.03 (0.93-1.16)	0.06
10q22.3 (ZMIZ1)	rs704010 (C/T)	2151	216	0.88 (0.71-1.09)	880	1.08 (0.96-1.21)	0.06
10q26 (FGFR2)	rs2981582 (C/T)	2249	228	1.09 (0.89-1.33)	918	1.26 (1.13-1.42)	0.14
19p13.1 (C19orf62)	rs8170 (G/A)	2156	216	1.17 (0.92-1.50)	884	1.03 (0.89-1.18)	0.28
11q13 (ORAOV1)	rs614367 (G/A)	2174	215	1.40 (1.12-1.74)	881	1.30 (1.14-1.47)	0.53
9p21 (CDKN2B)	rs1011970 (G/T)	2163	216	1.37 (1.05-1.79)	881	1.12 (0.95-1.31)	0.14
2q35 (TNP1/IGFBP5/IGFBP2/TNS1)	rs13387042 (T/C)	2324	246	0.78 (0.64-0.94)	997	0.88 (0.79-0.97)	0.24
5q11 (MAP3K1)	rs889312 (A/C)	2285	237	1.09 (0.88-1.36)	955	1.03 (0.91-1.17)	0.70
11p15 (LSP1)	rs3817198 (T/C)	2349	248	1.02 (0.84-1.25)	999	1.06 (0.95-1.19)	0.86
10q21.2 (ZNF365)	rs10995190 (G/A)	2173	212	0.75 (0.57-0.98)	876	0.83 (0.72-0.96)	0.51
17q23 (COX11)	rs6504950 (G/A)	2361	248	0.89 (0.72-1.10)	1010	0.93 (0.83-1.05)	0.71
5p12 (MRPS30/FGFR10)	rs4415084 (C/T)	2232	233	1.07 (0.88-1.29)	945	1.07 (0.96-1.20)	0.91
5p12 (MRPS30/FGFR10)	rs10941679 (A/G)	2366	251	1.06 (0.86-1.30)	1012	1.11 (0.99-1.25)	0.60
8q24 (FAM84B/c-MYC/POU5F1P1)	rs13281615 (A/G)	2339	248	1.06 (0.88-1.28)	1000	1.09 (0.98-1.21)	0.70
14q24.1 (RAD51L1)	rs10483813 (T/A)	2362	251	0.84 (0.67-1.06)	1007	0.90 (0.79-1.02)	0.62
3p24 (NEK10/SLC4A7)	rs4973768 (G/A)	2269	237	1.01 (0.83-1.21)	951	1.02 (0.92-1.14)	0.98

[a]Case-only P-value was used to test for heterogeneity (Phet) and was estimated using a polytomous logistic regression model with E-cadherin (E-cad) status as the outcome adjusted for age in 5-year categories. [b]Major/minor allele.

Supplementary Table 3. Distribution of select clinicopathologic features among breast cancer cases by E-cadherin tumor tissue expression levels in PBCS and SEARCH

	PBCS			SEARCH		
Characteristics	E-cad low (N=267)	E-cad high (N=1080)	P-value	E-cad low (N=471)	E-cad high (N=1535)	P-value
Histology, n (%)						
Ductal	76 (28.5)	677 (62.7)	**<0.001**	241 (51.2)	1291 (84.2)	**<0.001**
Lobular	132 (49.4)	110 (10.2)		179 (38.0)	124 (8.0)	
Other	59 (22.1)	292 (27.0)		51 (10.8)	119 (7.8)	
Grade, n (%)						
Well/Moderately differentiated	209 (80.1)	780 (73.7)	**0.03**	271 (70.0)	978 (69.8)	0.93
Poorly differentiated	52 (19.9)	278 (26.3)		116 (30.0)	423 (30.2)	
Tumor size (cm), n (%)						
≤2	119 (45.1)	571 (53.7)	**0.01**	140 (59.6)	539 (66.8)	**0.04**
>2	145 (54.9)	492 (46.3)		95 (40.4)	268 (33.2)	
Axillary node involvement, n (%)						
Negative	150 (58.4)	622 (59.0)	0.85	141 (59.8)	471 (60.0)	0.94
Positive	107 (41.6)	432 (41.0)		95 (40.2)	314 (40.0)	
ER status, n (%)						
Negative	93 (36.2)	307 (29.8)	0.05	85 (20.0)	344 (25.2)	**0.03**
Positive	164 (63.8)	724 (70.2)		341 (80.0)	1022 (74.8)	

P-values were calculated using the chi-squared test. Abbreviations: E-cad, E-cadherin; ER, estrogen receptor.