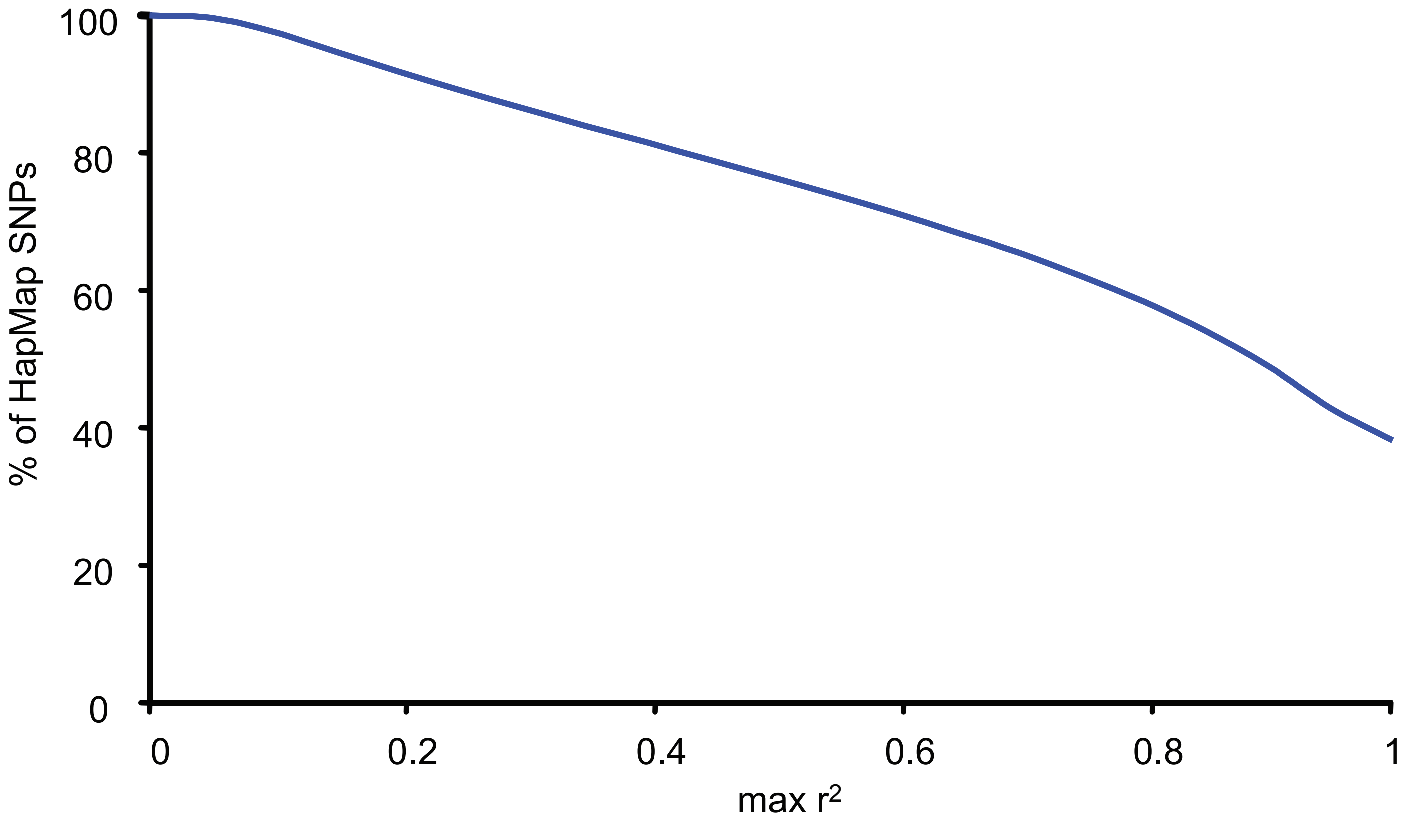
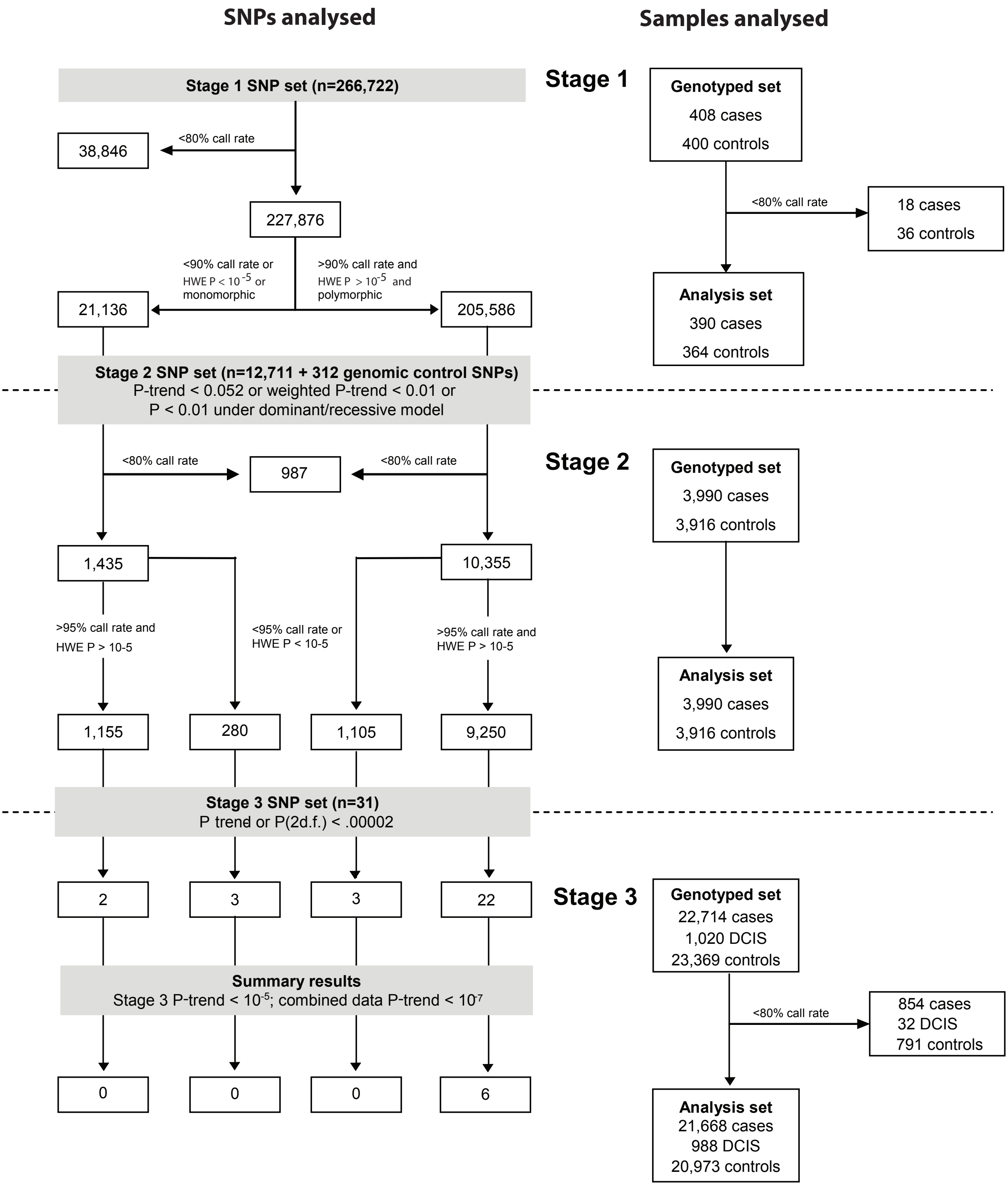
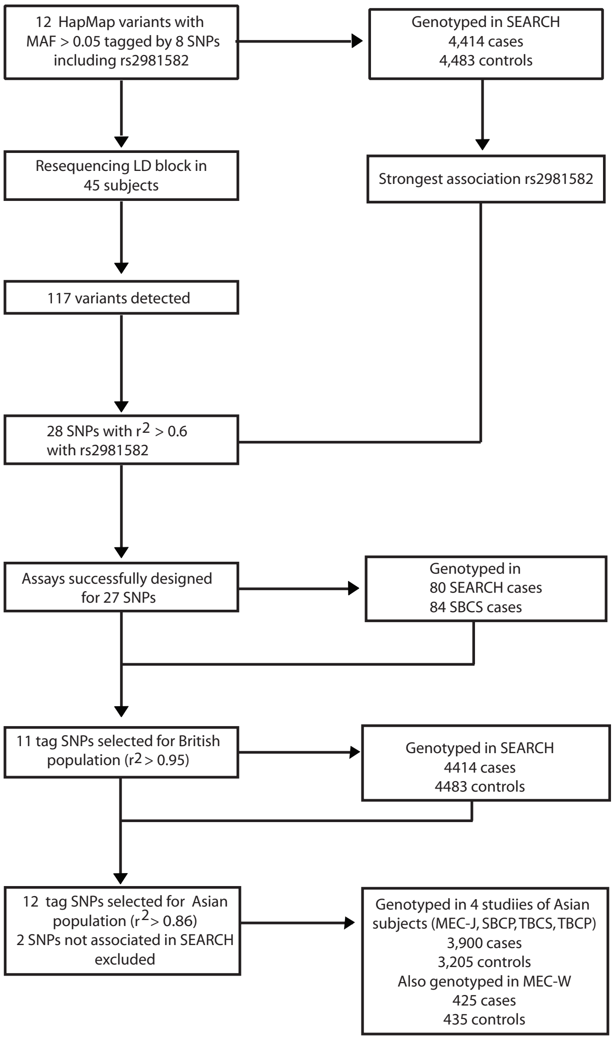

100
80
60
40
20
0
% of HapMap SNPs
0
0.2
0.4
0.6
0.8
1
max r²


SNPs analysed
Samples analysed
Stage 1 SNP set (n=266,722)
<80% call rate
38,846
227,876
<90% call rate or HWE P < 10-5 or monomorphic
>90% call rate and HWE P > 10-5 and polymorphic
21,136
205,586
Stage 2 SNP set (n=12,711 + 312 genomic control SNPs)
P-trend < 0.052 or weighted P-trend < 0.01 or
P < 0.01 under dominant/recessive model
<80% call rate
987
<80% call rate
1,435
10,355
>95% call rate and HWE P > 10-5
<95% call rate or HWE P < 10-5
>95% call rate and HWE P > 10-5
1,155
280
1,105
9,250
Stage 3 SNP set (n=31)
P trend or P(2d.f.) < .00002
2
3
3
22
Summary results
Stage 3 P-trend < 10-5; combined data P-trend < 10-7
0
0
0
6
Stage 1
Genotyped set
408 cases
400 controls
<80% call rate
18 cases
36 controls
Analysis set
390 cases
364 controls
Stage 2
Genotyped set
3,990 cases
3,916 controls
Analysis set
3,990 cases
3,916 controls
Stage 3
Genotyped set
22,714 cases
1,020 DCIS
23,369 controls
<80% call rate
854 cases
32 DCIS
791 controls
Analysis set
21,668 cases
988 DCIS
20,973 controls

Fine mapping of FGFR2 LD block



Supplementary table 1. Characteristics of markers showing a case-control difference significant at p<.01 in stages 1 and 2 (Cochran-Armitage test).

	Stage 1		Stage 2	
	p<.01	p>.01	p<.01	p>.01
Total	2388 (1.2%)	203198	171 (1.6%)	10,234
Call rate				
90-95%	290 (1.6%)	18050	-	-
95-99%	809 (1.1%)	69764	43 (1.5%)	2,747
>99%	1289 (1.1%)	115384	128 (1.7%)	7,487
Call rate difference >2%	374 (1.4%)	26692	4 (2.9%)	132
Hardy-Weinberg p<.001	17 (0.9%)	1847	4 (3.1%)	126

Supplementary table 2. Summary of the breast cancer case-control studies used in stages 2 and 3.

Study	Abbreviation	Ref	General Setting	Number of subjects				Age range		Genotyping platform(s)
				Invasive cases	CIS cases	Controls	positive FH (cases)	Cases	Controls	
Australian Breast Cancer Family Study	ABCFS	1,2	Australia; population-based case-control-family study	1485		839	191	20-68	23-69	Sequenom iPLEX
Breast Cancer Study in Taiwan[2]	BCST	3,4	Taiwan; hospital-based case-control study	1000	0	1000	21	30-85	30-85	Taqman
Copenhagen Breast Cancer Study and General Population Study	CGPS	5,6	Denmark (Copenhagen); population-based	1738	59	3045	383	29-93	20-90	Taqman
Gene Environment Interaction and Breast Cancer in Germany	GENICA	7	Germany; population-based case-control study	608		626	71	23-80	24-80	Sequenom iPLEX
Hannover Breast Cancer Study	HBCS	8	Germany; hospital-based	1098	4	1017	162	27-83	18-68	Taqman
Helsinki Breast Cancer Project	HBCP	9-11	Finland; hospital-based case-control study + additional familial cases	2099	164	1065	805	21-96	18-65	Taqman
IARC – Thai Breast Cancer Study	TBCS		Thailand, hospital based case-control study	474		390	16	17 – 81	18 - 82	Taqman
Kathleen Cuningham Foundation Consortium for Familial Breast Cancer (kConFab) and Australian Ovarian Cancer Study (AOCS)	KConFab/ AOCS	12	Australia and New Zealand; clinic-based recruitment of familial breast cancer patients (cases) Australia; population-based case-control study of ovarian cancer (controls only)	387		770	82	25-78	20-81	Sequenom iPLEX
Kuopio Breast Cancer Project	KBCP	13	Finland; hospital-based prospective clinical cohort	490	22	527	52	44-91	37-77	Taqman
Leiden University Medical Centre Breast Cancer Study	LUMCBCS	14	Netherlands (Leiden and Rotterdam), hospital-based	1206	152	600	228	21-87	18-94	Taqman
Mayo Clinic Breast Cancer Study	MCBCS	15	US; clinic-based case-control study	796		842	234	25-87	27-85	Taqman
Melbourne Collaborative Cohort Study	MCCS	16	Australia: population based prospective cohort study	519	1	578	84	41-79	39-69	Taqman
Multi-ethnic cohort (White)	MEC-W	17	Case-control study nested in prospective cohort	425		435	69			Taqman
Multi-ethnic cohort (Japanese)	MEC-J	17	Case-control study nested in prospective cohort	448		394	77			Taqman
Nurses Health Study	NHS	18	Case-control nested in prospective cohort	1029	210	1761	209	43-69	43-69	Taqman

Study	Abbreviation	Ref	General Setting	Number of subjects				Age range		Genotyping platform(s)
				Invasive cases	CIS cases	Controls	positive FH (cases)	Cases	Controls	
Polish Breast Cancer Study	PBCS	19	Poland, Warsaw and Łódź; population-based case-control study	1987	136	2374	212	20-74	20-74	Taqman
Rotterdam Breast Cancer Study	RBCS	20	Hospital based case-control study	799		800	799	18-84	20+	Taqman
Singapore and Sweden Breast Cancer Study	SASBAC	21	Sweden, population-based case-control study	1314		1515	205	50-74	50-74	Sequenom iPLEX, Taqman
SEARCH – stage 2[1]	SEARCH2	22	England; population-based case-control study	3814	169	3890	347	25-69	42-80	Taqman
SEARCH – stage 3[1]	SEARCH3	22	England; population-based case-control study	612	20	554	50			
Seoul Breast Cancer Project	SBCP	23	Korea (Seoul); hospital-based case-control study	2442		1780	-	≥ 40	≥ 40	Taqman
Sheffield Breast Cancer Study	SBCS	24,25	England; hospital-based case-control study	1115	92	1265	165	28-92	45-80	Taqman
Spanish National Cancer Centre Breast Cancer Study	CNIOBCS	26	Spain; hospital-based case-control study	732	33	829	224	22-86	23-86	Taqman
U.S. Radiologic Technologist Study	USRT	27,28	US (national); nested case–control study within a cohort of approximately 140,000 radiologic technologists	733	126	1059	173	19-84 at dx 41-99 at interview	41-91 at interview	Taqman

[1] SEARCH study contributed all the samples to stage 2 plus additional samples to stage 3 (see methods)

[2] BCST was only used in the fine-scale mapping of the FGFR2 locus.

Supplementary table 3. Summary odds ratios for the 31 SNPs tested in stage 3

LocalID	rs number	Alleles[1]	Maf [2] SEARCH (Asian)	Per allele OR	HetOR (95%CI)	HomOR (95%CI)	p-trend			p-het			Hom[3]
							Stages1 &2	Stage3	Combined	Stages1 &2	Stage3	Combined	
BCAC3_01	2981582	G/A	0.38 *(0.30)*	1.26 (1.23-1.30)	1.23 (1.18-1.28)	1.63 (1.53-1.72)	4×10^{-16}	4×10^{-62}	2×10^{-76}	3×10^{-15}	10^{-61}	5×10^{-76}	.45
BCAC3_02	12443621	A/G	0.46 *(0.60)*	1.11 (1.08-1.14)	1.14 (1.09-1.20)	1.23 (1.17-1.30)	10^{-7}	9×10^{-14}	2×10^{-19}	3×10^{-7}	3×10^{-13}	3×10^{-19}	.18
BCAC3_03	3857481	G/A	0.30 *(0.29)*	1.00 (0.98-1.04)	1.01 (0.96-1.05)	1.01 (0.94-1.08)	5×10^{-7}	.75	.009	2×10^{-6}	.95	.03	.01
BCAC3_04	1318703	T/C	0.42 *(0.50)*	1.03 (1.00-1.06)	1.03 (0.99-1.08)	1.06 (0.99-1.13)	.00003	.08	.64	.0002	.20	.90	.36
BCAC3_05	16998733	C/T	0.12 *(0.14)*	1.03 (0.98-1.08)	1.03 (0.97-1.08)	1.04 (0.86-1.25)	.00002	.30	.26	0.0001	.57	0.53	.45
BCAC3_09	2314099	G/T	0.06 *(0.03)*	0.98 (0.92-1.05)	0.97 (0.91-1.04)	1.16 (0.83-1.64)	.53	.57	.43	0.39	.44	0.21	.53
BCAC3_10	4841365	C/G	0.30 *(0.77)*	1.00 (0.97-1.03)	0.96 (0.89-1.03)	0.98 (0.91-1.05)	5×10^{-6}	1	.03	2×10^{-5}	.52	.006	.15
BCAC3_11	13281615	T/C	0.40 *(0.56)*	1.08 (1.05-1.11)	1.06 (1.01-1.11)	1.18 (1.10-1.25)	2×10^{-7}	6×10^{-7}	5×10^{-12}	4×10^{-8}	2×10^{-06}	4×10^{-12}	.73
BCAC3_12	3817198	T/C	0.30 *(0.14)*	1.07 (1.04-1.11)	1.06 (1.02-1.11)	1.17 (1.08-1.25)	8×10^{-6}	10^{-5}	3×10^{-9}	2×10^{-5}	6×10^{-5}	10^{-8}	.48
BCAC3_16	889312	A/C	0.28 *(0.54)*	1.13 (1.10-1.16)	1.13 (1.09-1.18)	1.27 (1.19-1.36)	4×10^{-6}	3×10^{-15}	7×10^{-20}	2×10^{-5}	3×10^{-14}	6×10^{-19}	.70
BCAC3_18	4666451	G/A	0.41 *(0.04)*	0.97 (0.94-1.00)	0.98 (0.93-1.02)	0.93 (0.87-0.99)	5×10^{-6}	.04	.00006	2×10^{-7}	.10	4×10^{-5}	.30
BCAC3_19	2107425	C/T	0.31 *(0.44)*	0.96 (0.93-0.99)	0.94 (0.90-0.98)	0.95 (0.89-1.01)	7×10^{-6}	.01	.00002	4×10^{-5}	.01	5×10^{-5}	.02
BCAC3_20	7313833	G/A	0.32 *(0.45)*	1.03 (1.00-1.06)	1.03 (0.99-1.08)	1.05 (0.98-1.12)	6×10^{-6}	.08	.0003	6×10^{-6}	.20	.001	.67
BCAC3_23	981782	A/C	0.47 *(0.37)*	0.96 (0.93-0.99)	0.96 (0.92-1.01)	0.92 (0.87-0.97)	.00008	.003	.000009	.0004	.01	5×10^{-5}	.07
BCAC3_25	11235127	G/A	0.16 *(0.41)*	0.98 (0.94-1.02)	0.97 (0.92-1.01)	1.00 (0.90-1.11)	.00006	.26	.41	.00005	.29	0.72	.67
BCAC3_26	17157070	A/G	0.05 *(0.12)*	1.03 (0.97-1.10)	1.02 (0.95-1.09)	1.24 (0.90-1.70)	.08	.33	.10	.07	.37	.08	.22

BCAC3_27	6469633	T/C	0.22 *(0.29)*	1.02 (0.99-1.06)	1.01 (0.97-1.06)	1.06 (0.96-1.16)	3x10^{-7}	.26	.002	7x10^{-7}	.47	.009	.31
BCAC3_28	4331913	C/T	0.42 *(0.32)*	1.01 (0.98-1.05)	0.98 (0.94-1.03)	1.04 (0.98-1.12)	.00003	.41	.005	.0001	.16	.008	.61
BCAC3_29	8051542	C/T	0.44 *(0.20)*	1.09 (1.06-1.13)	1.10 (1.05-1.16)	1.19 (1.12-1.27)	4x10^{-6}	4x10^{-8}	10^{-12}	2x10^{-5}	2x10^{-7}	10^{-11}	.08
BCAC3_30	6843340	C/T	0.48 *(0.47)*	1.00 (0.97-1.03)	0.99 (0.94-1.05)	1.00 (0.94-1.06)	.0009	.92	.13	.004	.97	.30	.29
BCAC3_31	2049621	C/G	0.09 *(0.02)*	1.04 (0.99-1.09)	1.06 (1.00-1.12)	0.92 (0.74-1.14)	.0009	.15	.009	.002	.10	.02	.45
BCAC3_33	10508468	T/C	0.39 *(0.38)*	0.98 (0.96-1.01)	0.99 (0.95-1.03)	0.97 (0.91-1.03)	.002	.30	.02	7x10^{-5}	.56	.02	.80
BCAC3_35	13110927	C/T	0.47 (0.38)	1.01 (0.98-1.04)	1.00 (0.96-1.05)	1.02 (0.96-1.07)	.0004	.60	.04	3x10^{-5}	.86	.04	.61
BCAC3_36	4954956	C/T	0.26 (0.29)	0.97 (0.94-1.00)	0.98 (0.94-1.02)	0.94 (0.86-1.02)	.001	.08	.002	8x10^{-6}	.21	.001	,89
BCAC3_37	6463266	T/G	0.19 (0.21)	1.00 (0.97-1.04)	1.01 (0.97-1.06)	0.97 (0.88-1.08)	.0002	.92	.07	.0008	.70	.17	.44
BCAC3_38	3852789	A/C	0.21 (0.27)	1.00 (0.96-1.03)	1.01 (0.97-1.06)	0.96 (0.87-1.06)	.00003	.89	.04	5x10^{-5}	.53	.03	.72
BCAC3_39	12658840	G/A	0.29 (0.36)	0.99 (0.96-1.02)	0.99 (0.94-1.03)	0.98 (0.92-1.05)	.00005	.48	.20	.0002	.76	.43	.61
BCAC3_40	2298075	C/A	0.23 (0.15)	0.97 (0.94-1.01)	0.99 (0.94-1.03)	0.92 (0.83-1.02)	.27	.17	.08	.0002	.28	.005	.68
BCAC3_41	7307700	G/A	0.47 (0.49)	1.02 (0.99-1.05)	1.03 (0.98-1.07)	1.04 (0.98-1.09)	.01	.21	.02	.002	.43	.06	.36
BCAC3_42	30099	G/A	0.08 *(0.39)*	1.05 (1.01-1.10)	1.06 (1.00-1.11)	1.09 (0.96-1.24)	.003	.02	.001	.002	.07	.005	.74
BCAC3_50	3803662	C/T	0.25 *(0.60)*	1.20 (1.16-1.24)	1.23 (1.18-1.29)	1.39 (1.29-1.49)	3x10^{-12}	10^{-26}	10^{-36}	2x10^{-11}	3x10^{-26}	7x10^{-36}	.95

[1] Coded to the minor allele in the UK-Search study. Reference allele/alternate allele (+ strand).

[2] Minor allele frequency.

[3] Test of homogeneity of the per allele odds ratio among studies.

Supplementary table 4. Haplotype analysis of seven tag SNPs in FGFR2.

Haplotype[1]	Control freq[2]	Case freq[2]	OR[3] (95%CI)	OR[3] (95%CI)
1111111	0.184	0.169	1.04 (0.94-1.13)	0.80 (0.72-0.88)
2111111	0.022	0.0169	0.87 (0.69-1.09)	0.66 (0.53-0.84)
1211111	0.281	0.248	1	0.77 (0.70-0.84)
1221111	0.114	0.114	1.13 (1.02-1.26)	0.87 (0.78-0.97)
1111221	0.161	0.186	1.31 (1.19-1.42)	1
2112121	0.120	0.130	1.23 (1.11-1.36)	0.94 (0.85-1.05)
2112122	0.081	0.094	1.31 (1.19-1.50)	1.03 (0.91-1.16)

[1] Loci in order are rs11200012, rs3750817, rs17542768, rs1219643, rs17102287, rs2981582, rs1219639. 1 represents the commoner allele in the UK Search study controls. Positions shown in figure 3c.

[2] Estimated frequencies in 4,470 cases and 4,560 controls (Search study).

[3] Odds ratios relative to the 1211111 haplotype (the commonest haplotype carrying the allele 1 at rs2981582) as baseline.

[4] Odds ratios relative to the 1111221 haplotype (the commonest haplotype carrying the allele 2 at rs2981582) as baseline.

Supplementary table 5. Estimated haplotype frequencies in 80 UK and 84 Korean samples, for the 27 SNPs in *FGFR2* that were correlated with rs2981582 and could be genotyped by Taqman or Pyrosequencing.

rs4752569	rs35012336	rs7895676	rs10736303	rs11200014	rs2912781	rs2912780	rs2912779	rs2981579	rs2912778	rs1078806	rs2981578	rs35054928	rs11599804	rs34032268	rs4752571	rs2981575	rs1219648	rs1219642	rs2912774	rs2936870	rs2981584	rs2860197	rs2420946	rs2162540	rs2981582	rs3135718	UK	KOREAN
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.425	0.234
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.331	0.367
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.066	0.019
2	2	1	1	2	1	2	2	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.025	0.082
1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	1	2	2	2	2	2	2	2	2	2	2	0.025	n/a
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	0.016	n/a
2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0.013	n/a
2	1	1	1	2	1	2	2	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.012	n/a
1	1	1	1	2	1	2	2	2	1	2	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	0.012	n/a
1	2	1	1	2	1	2	2	2	1	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	0.010	n/a
2	1	1	1	2	1	1	1	1	1	2	1	1	2	1	2	1	1	1	1	1	1	1	1	1	1	1	n/a	0.038
2	1	1	1	2	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	1	n/a	0.032
2	2	1	1	2	1	2	2	2	1	2	1	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	n/a	0.025
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	1	1	1	1	1	1	1	1	1	1	n/a	0.019
2	2	1	1	2	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	1	1	n/a	0.019
2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	n/a	0.013
2	2	1	1	2	1	1	1	1	1	2	1	1	2	1	1	1	1	1	1	1	1	1	1	1	2	1	n/a	0.013
2	2	1	1	2	1	1	1	1	1	2	1	1	2	1	2	1	2	2	2	2	2	2	2	2	2	2	n/a	0.010

Supplementary table 6. Estimated r^2 for 27 SNPs in intron 2 of FGFR2, in 80 samples from the UK Search study (lower left triangle) and 84 samples from Korea (upper right triangle). SNPs in bold italic were chosen as tag SNPs. rs4752569 and rs35012336 were not considered in the Korean case-control study since they were strongly excluded in the UK study.

	rs4752569	rs35012336	rs7895676	***rs10736303***	***rs11200014***	rs2912781	***rs2912780***	rs2912779	rs2981579	rs2912778	rs1078806	***rs2981578***	rs35054928	rs11599804	rs34032268	***rs4752571***	***rs2981575***	rs1219648	***rs1219642***	***rs2912774***	***rs2936870***	rs2981584	rs2860197	***rs2420946***	rs2162540	***rs2981582***	***rs3135718***
rs4752569		0.49	0.22	0.22	0.81	0.17	0.38	0.38	0.38	0.22	0.83	0.24	0.40	0.82	0.34	0.54	0.34	0.44	0.42	0.39	0.39	0.42	0.39	0.40	0.41	0.54	0.35
rs35012336	0.66		0.45	0.46	0.61	0.42	0.70	0.71	0.70	0.43	0.62	0.45	0.67	0.59	0.68	0.56	0.60	0.75	0.74	0.69	0.68	0.73	0.68	0.69	0.68	0.56	0.61
rs7895676	0.46	0.68		0.90	0.33	0.83	0.61	0.60	0.60	0.93	0.33	0.89	0.63	0.31	0.60	0.41	0.53	0.54	0.50	0.54	0.51	0.54	0.46	0.46	0.48	0.36	0.45
rs10736303	0.49	0.71	0.97		0.33	0.86	0.67	0.67	0.67	0.81	0.34	0.95	0.61	0.29	0.56	0.42	0.57	0.53	0.54	0.58	0.53	0.53	0.47	0.48	0.48	0.33	0.38
rs11200014	0.65	0.84	0.73	0.77		0.28	0.49	0.49	0.49	0.29	0.99	0.32	0.49	0.93	0.48	0.61	0.40	0.52	0.51	0.48	0.48	0.50	0.45	0.46	0.46	0.60	0.41
rs2912781	0.50	0.73	0.95	0.97	0.73		0.58	0.57	0.57	0.77	0.30	0.79	0.53	0.26	0.52	0.34	0.49	0.50	0.48	0.56	0.49	0.47	0.44	0.42	0.43	0.31	0.43
rs2912780	0.60	0.83	0.76	0.79	0.97	0.76		1.00	1.00	0.58	0.51	0.65	0.95	0.46	0.95	0.66	0.88	0.83	0.83	0.78	0.77	0.82	0.76	0.77	0.76	0.54	0.67
rs2912779	0.60	0.83	0.76	0.79	0.97	0.76	1.00		1.00	0.57	0.51	0.65	0.94	0.46	0.95	0.66	0.88	0.82	0.83	0.78	0.77	0.82	0.76	0.77	0.76	0.54	0.68
rs2981579	0.57	0.81	0.74	0.78	0.97	0.74	1.00	1.00		0.57	0.51	0.64	0.95	0.46	0.95	0.66	0.88	0.83	0.83	0.78	0.77	0.82	0.76	0.77	0.76	0.54	0.67
rs2912778	0.50	0.73	0.95	0.97	0.73	1.00	0.76	0.76	0.74		0.29	0.90	0.58	0.27	0.59	0.35	0.49	0.49	0.48	0.50	0.48	0.47	0.43	0.44	0.41	0.33	0.46
rs1078806	0.64	0.84	0.75	0.78	1.00	0.75	0.97	0.97	0.97	0.75		0.32	0.50	0.95	0.48	0.62	0.41	0.52	0.53	0.50	0.49	0.51	0.47	0.47	0.48	0.61	0.42
rs2981578	0.50	0.72	0.97	0.95	0.72	0.97	0.75	0.75	0.74	0.97	0.74		0.64	0.32	0.60	0.44	0.58	0.55	0.55	0.59	0.55	0.55	0.48	0.50	0.50	0.35	0.42
rs35054928	0.60	0.83	0.76	0.79	0.97	0.76	1.00	1.00	1.00	0.76	0.97	0.74		0.52	0.97	0.72	0.90	0.85	0.84	0.79	0.78	0.84	0.77	0.80	0.79	0.60	0.68
rs11599804	0.61	0.85	0.75	0.77	1.00	0.75	0.97	0.97	0.97	0.75	1.00	0.74	0.97		0.50	0.66	0.44	0.55	0.54	0.51	0.52	0.54	0.49	0.50	0.50	0.66	0.41
rs34032268	0.51	0.71	0.67	0.69	0.84	0.67	0.88	0.88	0.87	0.67	0.85	0.65	0.88	0.85		0.73	0.90	0.92	0.92	0.84	0.94	0.92	0.87	0.87	0.87	0.66	0.79
rs4752571	0.56	0.76	0.67	0.69	0.90	0.67	0.88	0.88	0.87	0.67	0.90	0.66	0.88	0.90	0.95		0.65	0.82	0.82	0.78	0.76	0.82	0.75	0.76	0.77	0.55	0.66
rs2981575	0.62	0.85	0.74	0.77	1.00	0.74	0.98	0.98	0.97	0.74	1.00	0.73	0.98	1.00	0.85	0.90		0.81	0.81	0.76	0.74	0.81	0.74	0.74	0.75	0.54	0.65
rs1219648	0.57	0.76	0.67	0.69	0.90	0.67	0.88	0.88	0.87	0.67	0.90	0.65	0.88	0.90	0.95	1.00	0.90		1.00	0.94	0.92	1.00	0.93	0.94	0.94	0.69	0.83
rs1219642	0.56	0.76	0.67	0.69	0.89	0.67	0.88	0.88	0.87	0.67	0.90	0.65	0.88	0.90	0.95	1.00	0.90	1.00		0.94	0.92	0.99	0.93	0.94	0.93	0.68	0.81
rs2912774	0.52	0.73	0.68	0.70	0.86	0.68	0.90	0.90	0.89	0.68	0.87	0.66	0.90	0.87	0.97	0.97	0.87	0.97	0.97		0.86	0.93	0.89	0.90	0.89	0.64	0.77
rs2936870	0.40	0.60	0.61	0.64	0.77	0.61	0.78	0.78	0.79	0.61	0.74	0.59	0.78	0.75	0.86	0.86	0.76	0.86	0.86	0.91		0.90	0.85	0.86	0.85	0.63	0.73
rs2981584	0.53	0.74	0.69	0.71	0.87	0.69	0.91	0.91	0.90	0.69	0.87	0.67	0.90	0.87	0.98	0.98	0.88	0.98	0.98	1.00	0.89		0.92	0.92	0.94	0.67	0.81
rs2860197	0.50	0.70	0.65	0.68	0.82	0.65	0.86	0.86	0.85	0.65	0.82	0.64	0.86	0.83	0.92	0.93	0.83	0.93	0.92	0.97	0.86	0.95		0.99	0.98	0.72	0.88
rs2420946	0.48	0.63	0.65	0.66	0.78	0.65	0.80	0.80	0.78	0.65	0.79	0.63	0.80	0.77	0.87	0.87	0.77	0.90	0.87	0.92	0.80	0.90	0.95		0.99	0.73	0.88
rs2162540	0.50	0.70	0.65	0.68	0.82	0.65	0.86	0.86	0.85	0.65	0.82	0.64	0.86	0.83	0.92	0.93	0.83	0.93	0.92	0.97	0.86	0.95	1.00	0.95		0.72	0.86
rs2981582	0.49	0.68	0.64	0.66	0.79	0.64	0.84	0.84	0.85	0.64	0.80	0.64	0.84	0.81	0.90	0.90	0.81	0.90	0.90	0.95	0.89	0.93	0.98	0.93	0.98		0.62
rs3135718	0.47	0.67	0.62	0.64	0.79	0.62	0.83	0.83	0.82	0.62	0.79	0.60	0.83	0.80	0.90	0.90	0.80	0.90	0.90	0.95	0.83	0.93	0.98	0.92	0.98	0.95	

Supplementary table 7. Summary odds ratios for the 14 SNPs typed in the fine-scale mapping genotyping of the FGFR2 locus.

	European[1]				Asian[2]			
SNP	maf[3]	Per allele OR (95%CI)	p-trend	Homogeneity test[4]	maf[3]	Per allele OR (95%CI)	p-trend	Homogeneity test[4]
rs4752569	0.53	1.13 (1.07-1.20)	3x10^{-5}	-	-	-	-	-
rs35012336	0.44	1.19 (1.12-1.26)	10^{-8}	-	-	-	-	-
rs10736303	0.46	1.25 (1.18-1.32)	5x10^{-14}	.30	0.52	1.26 (1.18-1.35)	2x10^{-11}	.65
rs11200014	-	-	-	-	0.26	1.16 (1.05-1.28)	.004	.23
rs29812780	0.40	1.25 (1.18-1.33)	3x10^{-14}	.23	0.42	1.21 (1.13-1.31)	10^{-7}	.34
rs2981578	0.46	1.26 (1.19-1.34)	10^{-15}	.54	0.50	1.25 (1.17-1.34)	2x10^{-11}	.92
rs4752571	-	-	-	-	0.33	1.17 (1.09-1.26)	10^{-5}	.44
rs2981575	0.39	1.27 (1.20-1.35)	9x10^{-16}	.11	0.40	1.21 (1.13-1.30)	4x10^{-8}	.18
rs1219642	0.39	1.27 (1.20-1.35)	3x10^{-16}	.10	0.37	1.21 (1.13-1.30)	9x10^{-8}	.50
rs2912774	0.39	1.26 (1.19-1.34)	3x10^{-15}	.13	0.38	1.19 (1.11-1.27)	8x10^{-7}	.23
rs2936870	0.39	1.26 (1.19-1.34)	5x10^{-15}	.10	0.37	1.20 (1.12-1.29)	2x10^{-7}	.32
rs2420946	0.38	1.27 (1.20-1.35)	5x10^{-16}	.06	0.37	1.17 (1.09-1.26)	7x10^{-6}	.48
rs2981582	0.38	1.27 (1.20-1.34)	8x10^{-16}	.07	0.30	1.19 (1.11-1.28)	2x10^{-6}	.05
rs3135718	-	-	-	-	0.38	1.15 (1.07-1.23)	7x10^{-5}	.59

[1] SEARCH and MEC-W studies.
[2] MEC-J, TBCS, SBCP, BCST studies.
[3] Minor allele frequency.
[4] p-value for homogeneity of the per allele OR among studies.

Supplementary table 8. Details of 29 SNPs correlated with rs2981582 at $r^2>0.6$.

rs Number	Alleles[1]	Position[2]	Log-likelihood Difference[3]			% conservation over 75 nucleotides[4]							Allele[5]						
			European	Asian	Total	Chi	Rhe	Dog	Cow	Rat	Mou	Opo	Chi	Rhe	Dog	Cat	Rat	Mou	Opo
rs4752569	T/A	123321680	8.43		8.43	96	84	0	0	0	0	0	2	2	-	-	-	-	-
rs35012336	-/C	123323185	4.70		4.70	95	93	0	23	0	45	0	2	2	-	1N	-	1	
rs33971856	- /G	123323189				99	93	0	23	0	51	0	1	1	-	g	-	g	-
rs7895676	T/C	123323987	0	0	0	99	91	0	15	0	0	0	2	2N	-	-	-	-	-
rs10736303	A/G	123324447	0.86	0.64	1.50	99	91	72	68	52	44	0	2	2	2	2	1	1	-
rs11200014	C/T	123324920	0.41	4.22	4.62	100	92	*62*	50	37	52	0	1	1	1	0	g	g	-
rs35393331	-/ TAAA	123325378				100	80	0	0	0	0	0	1	g	-	-	-	-	-
rs2912781	A/G	123326484	0.07	1.55	1.62	97	84	0	21	0	0	0	2	1	-	2N	-	-	-
rs2912780	A/G	123327107	0.49	3.68	4.17	99	89	65	59	0	0	0	1	2	1	1N	-	-	-
rs2912779	G/A	123327172	-0.83	5.19	4.36	99	93	69	60	0	0	0	1	1	1	1	-	-	-
rs2981579	G/A	123327325	4.71	1.98	6.69	100	95	73	76	0	0	0	1	1	2	2N	-	-	-
rs2912778	T/C	123328644	0.07	0.37	0.44	99	93	65	57	0	37	0	2	2	2N	2N	-	g	-
rs1078806	A/G	123328965	-0.34	4.25	3.91	97	88	75	71	*58*	75	0	1	1	1	1N	*1*	1	-
rs2981578	A/G	123330301	-0.78	2.41	1.63	99	91	64	81	75	69	0	2	2	1	2	2N	2N	-
rs35054928	-/C	123330423	-0.83	2.05	1.22	100	93	48	50[$]	53	52	0	1	1	1N	1[$]	1N	0	-

rs11599804	C/T	123330654	-0.79	5.37	4.58	84	83	39	63	36	36	0	1	2	1N	g	g	g	-
rs34032268	A/C	123331515	-1.09	3.43	2.34	99	87	0	0	0	0	0	2	2	-	-	-	-	-
rs4752571	T/C	123332557	-1.05	6.38	5.33	100	95	60	84	80	32	0	2	2	2N	2	2N	2N	-
rs2981575	A/G	123336106	-1.44	3.80	2.35	99	0	69	63	_49_	_46*_	51	2	-	1	1	_1N_	-	1N
rs1219648	A/G	123336180	-0.56	4.31	3.75	100	0	77	71	_52*_	-	55	1	-	1	1	_2N_	-	1N
rs1219642	A/G	123338379	-0.95	5.00	4.05	96	0	0	0	29[$]	31[$]	_48_	1	-	-	-	1N[$]	1N[$]	_g_
rs2912774	G/T	123338652	-1.09	5.49	4.40	100	0	63	67	-	69*	32	2	-	g	0	-	-	-
rs2936870	G/A	123338892	-0.59	5.30	4.71	100	0	71	89	87	88	65	2	-	2N	2	2	2N	2N
rs2981584	C/A	123340206	-1.08	4.36	3.29	99	0	55	64*	61	61	4	2	-	2N	0	1N	1N	-
rs2860197	A/G	123341292	-0.82	5.93	5.11	97	0	64	69	65	63	0	1	-	0	1	1N	1N	-
rs2420946	G/A	123341314	-1.15	6.51	5.36	99	0	64	69	65	63	_42_	2	-	0	2	2N	2N	_2N_
rs2162540	A/G	123342126	-0.82	6.18	5.36	97	0	63	49	39*	51	49	1	-	1N	2N	0	g	2N
rs2981582	G/A	123342307	-0.68	6.78	6.10	100	0	87	85	88	87	75	1	-	1N	1N	1N	1N	2N
rs3135718	A/G	123343859	-1.59	7.63	6.04	100	0	79	85	_76_	_77_	63	2	-	2	2	_0_	0	2N

[1] Allele that is in LD with the commoner G allele in rs2981582 (and hence associated with the lower risk) listed first

[2] build 36.2position on chromosome 10

[3] log-likelihood difference relative to rs7895676, in European and Asian populations (see methods for computation of these likelihoods).

[4] Chi=Chimp, Rhe=Rhesus, Mou=mouse, Opo=opossum (http://pipeline.lbl.gov/methods.shtml) [29-31]

[5] Chi=Chimp, Rhe=Rhesus, Mou=mouse, Opo=opossum; 2=rare/high-risk allele, 1=common allele, 0=neither allele, g=gap in sequence, - =no data (no alignment), N=no conservation of adjacent nucleotide, ie SNP site. Data from Ensembl (http://www.emsembl.org) except: italics=data from UCSC (http:// genome.ucsc.edu/), [$]=data from Vista (http://pipeline.lbl.gov/methods.shtml), *homology ignoring gap in alignment.

Supplementary table 9. Odds ratios for breast carcinoma-in-situ (CIS), versus controls, in stage 3, for the 5 SNPs reaching $p<10^{-7}$.

SNP rs number	Per allele OR (95%CI)	HetOR (95%CI)	HomOR (95%CI)	p-trend (CIS vs controls)	p-trend (CIS vs invasive)
rs2981582	1.25 (1.14-1.37)	1.11 (0.96-1.30)	1.62 (1.34-1.96)	6×10^{-7}	.67
rs3803662	1.18 (1.05-1.33)	1.14 (0.97-1.35)	1.45 (1.10-1.91)	.0006	.95
rs8051542	1.30 (1.17-1.44)	1.31 (1.14-1.51)	1.66 (1.31-2.09)	10^{-7}	.007
rs13281615	1.07 (0.96-1.19)	1.18 (1.00-1.39)	1.10 (0.88-1.37)	.06	.21
rs3817198	1.07 (0.97-1.19)	1.04 (0.90-1.20)	1.18 (0.94-1.49)	.25	.71

Supplementary table 10. Per allele odds ratios by age at breast cancer diagnosis in stage 3, for the five independent SNPs reaching $p<10^{-7}$.

SNP rs number	<40	40-49	50-59	60+	p-trend
rs2981582	1.39 (1.23-1.56)	1.24 (1.16-1.34)	1.21 (1.15-1.28)	1.26 (1.20-1.32)	.40
rs3803662	1.36 (1.16-1.60)	1.26 (1.16-1.36)	1.22 (1.15-1.29)	1.20 (1.14-1.26)	.13
rs8051542	1.11 (0.97-1.27)	1.17 (1.08-1.27)	1.17 (1.11-1.24)	1.08 (1.03-1.14)	.13
rs13281615	1.06 (0.91-1.23)	1.07 (0.99-1.16)	1.14 (1.08-1.20)	1.11 (1.06-1.17)	.47
rs3817198	1.10 (0.96-1.27)	1.14 (1.05-1.24)	1.05 (1.00-1.11)	1.06 (1.01-1.11)	.21

[1] Test for trend in odds ratios by age

Supplementary table 11. Per allele odds ratios for the presence of a positive family history of breast cancer and for bilaterality in relation to genotype, based on a case only analysis in stage 3, for the five independent SNPs reaching $p<10^{-7}$ in the case-control analysis.

SNP rs number	Family History[1]		Bilaterality[2]	
	Per allele OR	p-trend	Per allele OR	p-trend
rs2981582	1.07 (1.02-1.12)	.02	1.18 (1.03-1.37)	.02
rs3803662	1.08 (1.01-1.15)	.03	1.13 (0.88-1.44)	.36
rs8051542	1.01 (0.96-1.07)	.64	1.09 (0.93-1.28)	.26
rs13281615	1.06 (1.00-1.12)	.05	0.99 (.85-1.16)	.91
rs3817198	1.01 (0.95-1.07)	.77	1.06 (0.91-1.23)	.46

[1] Based on 3275 cases with a positive family history from 20 studies, and 14,028 cases without a family history from these studies.
[2] Based on 422 known bilateral cases from 9 studies, and 8012 unilateral cases from these studies.

References

1. Spurdle,A.B. *et al.* CYP17 promoter polymorphism and breast cancer in Australian women under age forty years. *Journal of the National Cancer Institute* **92**, 1674-1681 (2000).

2. Spurdle,A.B. *et al.* Androgen receptor exon 1 CAG repeat length and breast cancer in women before age forty years. *Journal of the National Cancer Institute* **91**, 961-966 (1999).

3. Lo,Y.L. *et al.* Breast cancer risk associated with genotypic polymorphism of the mitosis-regulating gene Aurora-A/STK15/BTAK. *International Journal of Cancer* **115**, 276-283 (2005).

4. Cheng,T.C. *et al.* Breast cancer risk associated with genotype polymorphism of the catechol estrogen-metabolizing genes: A multigenic study on cancer susceptibility. *International Journal of Cancer* **113**, 345-353 (2005).

5. Bojesen,S.E., Tybjaerg-Hansen,A., Axelsson,C.K. & Nordestgaard,B.G. No association of breast cancer risk with integrin beta(3) (ITGB3) Leu33Pro genotype. *British Journal of Cancer* **93**, 167-171 (2005).

6. Weischer,M., Bojesen,S.E., Tybjaerg-Hansen,A., Axelsson,C.K. & Nordestgaard,B.G. Increased risk of breast cancer associated with CHEK2*1100delC. Journal of Clinical Oncology . 2007.
   Ref Type: In Press

7. Pesch,B. *et al.* Factors modifying the association between hormone-replacement therapy and breast cancer risk. *European Journal of Epidemiology* **20**, 699-711 (2005).

8. Dork,T. *et al.* Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients. *Cancer Research* **61**, 7608-7615 (2001).

9. Kilpivaara,O. *et al.* Correlation of CHEK2 protein expression and c.1100delC mutation status with tumor characteristics among unselected breast cancer patients. *International Journal of Cancer* **113**, 575-580 (2005).

10. Tommiska,J. *et al.* Breast cancer patients with p53 pro72 homozygous genotype have a poorer survival. *Clinical Cancer Research* **11**, 5098-5103 (2005).

11. Syrjakoski,K. *et al.* Population-based study of BRCA1 and BRCA2 mutations in 1035 unselected Finnish breast cancer patients. *Journal of the National Cancer Institute* **92**, 1529-1531 (2000).

12. Mann,G.J. *et al.* Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Research* **8**, (2006).

13. Hartikainen,J.M. *et al.* An autosome-wide scan for linkage disequilibrium-based association in sporadic breast cancer cases in eastern Finland: Three candidate regions found. *Cancer Epidemiology Biomarkers & Prevention* **14**, 75-80 (2005).

14. de Bock,G.H. *et al.* Tumour characteristics and prognosis of breast cancer patients carrying the germline CHEK2*1100delC variant. *Journal of Medical Genetics* **41**, 731-735 (2004).

15. Olson,J.E. *et al.* A comprehensive examination of CYP19 variation and risk of breast cancer using two haplotype-tagging approaches. Breast Cancer Research and Treatment . 2007.
Ref Type: In Press

16. Giles GG & English DR. Nutrition and lifestyle: opportunities for cancer prevention. IARC Scientific Publications No 156. Riboli E,L.R. (ed.) (IARC, Lyon,2002).

17. Kolonel,L.N. *et al.* A multiethnic cohort in Hawaii and Los Angeles: Baseline characteristics. *American Journal of Epidemiology* **151**, 346-357 (2000).

18. Hankinson,S.E. *et al.* Plasma sex steroid hormone levels and risk of breast cancer in postmenopausal women. *Journal of the National Cancer Institute* **90**, 1292-1299 (1998).

19. Garcia-Closas,M. *et al.* Polymorphisms in DNA double-strand break repair genes and risk of breast cancer: two population-based studies in USA and Poland, and meta-analyses. *Human Genetics* **119**, 376-388 (2006).

20. Verhoog,L.C. *et al.* Large regional differences in the frequency of distinct BRCA1/BRCA2 mutations in 517 Dutch breast and/or ovarian cancer families. *European Journal of Cancer* **37**, 2082-2090 (2001).

21. Wedren,S. *et al.* Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: a case control study. *Breast Cancer Research* **6**, R437-R449 (2004).

22. Lesueur,F. *et al.* Allelic association of the human homologue of the mouse modifier Ptprj with breast cancer. *Human Molecular Genetics* **14**, 2349-2356 (2005).

23. Lee,K.M. *et al.* Genetic polymorphisms of ataxia telangiectasia mutated and breast cancer risk. *Cancer Epidemiology Biomarkers & Prevention* **14**, 821-825 (2005).

24. MacPherson,G. *et al.* Association of a common variant of the CASP8 gene with reduced risk of breast cancer. *Journal of the National Cancer Institute* **96**, 1866-1869 (2004).

25. Rafii,S. *et al.* A potential role for the XRCC2 R188H polymorphic site in DNA-damage repair and breast cancer. *Human Molecular Genetics* **11**, 1433-1438 (2002).

26. Milne,R.L. *et al.* ERCC4 associated with breast cancer risk: A two-stage case-control study using high-throughput genotyping. *Cancer Research* **66**, 9420-9427 (2006).

27. Sigurdson,A.J. *et al.* Kin-cohort estimates for familial breast cancer risk in relation to variants in DNA base excision repair, BRCA1 interacting and growth factor genes. *Bmc Cancer* **4**, (2004).

28. Sigurdson,A.J. *et al.* Cancer incidence in the US radiologic technologists health study, 1983-1998. *Cancer* **97**, 3080-3089 (2003).

29. Frazer,K.A., Pachter,L., Poliakov,A., Rubin,E.M. & Dubchak,I. VISTA: computational tools for comparative genomics. *Nucleic Acids Research* **32**, W273-W279 (2004).

30. Mayor,C. *et al.* VISTA: visualizing global DNA sequence alignments of arbitrary length. *Bioinformatics* **16**, 1046-1047 (2000).

31. Kent,W.J. BLAT - The BLAST-like alignment tool. *Genome Research* **12**, 656-664 (2002).