Low penetrance breast cancer predisposition SNPs are site specific

Niall Mcinerney, Gabrielle Colleran, Andrew Rowan, Axel Walther, Ella Barclay, Sarah Spain, Angela M. Jones, Stephen Tuohy, Catherine Curran, Nicola Miller, et al.

Breast Cancer Res Treat (2009) 117:151–159
DOI 10.1007/s10549-008-0235-7

EPIDEMIOLOGY

Low penetrance breast cancer predisposition SNPs are site specific

Niall Mcinerney · Gabrielle Colleran · Andrew Rowan · Axel Walther · Ella Barclay · Sarah Spain · Angela M. Jones · Stephen Tuohy · Catherine Curran · Nicola Miller · Michael Kerin · Ian Tomlinson · Elinor Sawyer

Received: 17 October 2008 / Accepted: 20 October 2008 / Published online: 13 November 2008

Abstract Large scale association studies have identified low penetrance susceptibility alleles that predispose to breast cancer. A locus on chromosome 8q24.21 has been shown to harbour variants that predispose to breast, ovarian, colorectal and prostate cancer. The finding of risk variants clustering at 8q24 suggests that there may be common susceptibility alleles that predispose to more than one epithelial cancer. The aim of this study was firstly to determine whether previously identified breast cancer susceptibility alleles are associated with sporadic breast cancer in the West of Ireland and secondly to ascertain whether there are susceptibility alleles that predispose to all three common epithelial cancers (breast, prostate, colon). We genotyped a panel of 24 SNPs that have recently been shown to predispose to prostate, colorectal or breast cancer in 988 sporadic breast cancer cases and 1,016 controls from the West of Ireland. We then combined our data with publicly available datasets using standard techniques of meta-analysis. The known breast cancer SNPs rs13281615, rs2981582 and rs3803662 were confirmed as associated with breast cancer risk ($P_{\text{allelic test}} = 1.8 \times 10^{-2}$, OR = 1.17; $P_{\text{allelic test}} = 2.2 \times 10^{-3}$, OR = 1.22; $P_{\text{allelic test}} = 5.1 \times 10^{-2}$, OR = 1.15, respectively) in the West of Ireland cohort. For the remaining five breast cancer SNPs that were studied there was no evidence of an association with breast cancer in the West Ireland population ($P_{\text{allelic test}} > 6.5 \times 10^{-2}$). There was also no association between any of the prostate or colorectal susceptibility SNPs, whether at 8q24 or elsewhere, with breast cancer risk. Meta-analysis confirmed that all susceptibility SNPs were site specific, with the exception of rs6983269 which is known to predispose to both colorectal and prostate cancer. This study confirms that susceptibility loci at FGFR2, 8q24 and TNCR9 predispose to sporadic breast cancer in the West of Ireland. It also suggests that low penetrance susceptibility SNPs for breast, prostate and colorectal cancer are distinct. Although 8q24 harbours variants that predispose to all three cancers, the susceptibility loci within the region appear to be specific for the different cancer types with the exception of rs6983269 in colon and prostate cancer.

Keywords Breast cancer · Colorectal cancer · Prostate cancer · Genetic susceptibility · 8q24 · West Ireland

Electronic supplementary material The online version of this article (doi:10.1007/s10549-008-0235-7) contains supplementary material, which is available to authorized users.

N. Mcinerney · G. Colleran · S. Tuohy · C. Curran · N. Miller · M. Kerin
Department of Surgery, Clinical Science Institute, University College Hospital, Galway, Ireland

N. Mcinerney · G. Colleran · A. Rowan · A. Walther · E. Barclay · S. Spain · A. M. Jones · S. Tuohy · I. Tomlinson · E. Sawyer
Molecular and Population Genetics Laboratory, Cancer Research UK, 44, Lincoln's Inn Fields, London WC2A 3PX, UK

E. Sawyer (✉)
Department of Academic Oncology, Guy's, King's, St Thomas' Cancer Centre, Guy's Hospital, St Thomas's Street, London SE1 9RT, UK
e-mail: elinor.sawyer@cancer.org.uk

Introduction

The risk of breast cancer is raised about twofold in the sisters and mothers of affected women. Only a small proportion of this risk is accounted for by known breast cancer predisposition genes such as *BRCA1* and *BRCA2*. The

search for further high penetrance genes has failed to identify loci that predispose to breast cancer, suggesting that less penetrant genes may be important. Despite the identification of rare alleles that confer a modestly increased risk of breast cancer-such as *CHEK2, ATM, BRIP* and *PALB2*-about 75% of the familial risk remains unexplained [1]. A polygenic model for susceptibility to breast cancer and other common cancers, including colorectal and prostate cancer, has been suggested, where risk is conferred by a large number of alleles, each conferring a small risk, but which, when combined, confer a range of susceptibilities within the population [2].

In order to identify common, low penetrance cancer susceptibility alleles, large genome-wide association studies (GWAS) have been performed. Easton et al. identified five loci associated with increased risk of developing breast cancer. These were initially identified in breast cancer patients with a family history and verified in a much larger cohort of patients with both sporadic and familial breast cancer [3]. The SNPs with the strongest association with breast cancer in this study were rs2981582 (10q: FGFR2), rs3803662 (16q:TNRC9) and rs13281615 (chr8q24) with a per allele OR typically of 1.1–1.26. Several other studies have also identified an association with SNPs within FGFR2 and TNRC9 and breast cancer [4, 5]. More recent studies have also identified SNPs in other regions that predispose to breast cancer. SNPs on 2q35 (rs13387042) and 5p12 (rs4415084 and rs10941679) have been shown to be associated with estrogen-receptor (ER) positive breast cancer [5, 6] and SNPs lying within cell cycle genes have been shown to be associated with breast cancer in the British population (rs997669 in CCNE1, rs3176336 in CDKN1A, rs34330 in CDKN1B and rs3731239 in CDKN2A/2B) [7]. These associations have yet to be confirmed in other studies or populations, with the exception of rs34330, which has been shown to be associated with breast cancer in the Chinese population [8].

One of the breast cancer susceptibility SNPs identified by Easton et al. rs13281615 (chr8:128,424,800), lies at chromosome 8q24.21, a region that is also associated with colorectal and prostate cancer risk. Within 8q24.21, several GWASs have shown an association between prostate cancer and SNP rs1447295 (chr8:128,553,970) [9–13]. This locus has been studied in breast cancer cohorts, but no association found [14]. A second independent locus within 8q24 conferring prostate cancer risk, rs16901979 (chr8:128,194,098), has been identified [11, 12]. Another nearby SNP, rs6983267 (chr8:128,482,237), confers an increased risk of colorectal cancer [15, 16] and also confers some independent prostate cancer risk and ovarian cancer risk [17]. The association with prostate cancer is weaker than that with rs1447295 [18], but again independent of both rs1447295 and rs16901979. All four SNPs are found within 400 kb of each other: rs13281615 is 130 kb proximal of rs1447295, 60 kb proximal of rs6983267 and 230 kb distal of rs16901979.

Many of the common cancers are associated with a twofold increase in familial relative risk across the different cancer types [19]. This together with the finding of risk variants for breast, prostate, colon and ovarian cancer clustering at 8q24 suggests that there may be low penetrance susceptibility alleles elsewhere in the genome predisposing to all of these cancers. Large genome wide association studies of prostate and colorectal cancer have now identified new prostate [20–22] and colorectal [23–27] susceptibility alleles in addition to the alleles at 8q24. It is not known whether these alleles also predispose to breast cancer.

The aims of this study were:

1. To confirm whether recently identified breast cancer predisposition SNPs also predispose to breast cancer in the Irish population.
2. To determine whether any of the other known cancer SNPs, also predispose, perhaps more weakly, to other common cancer types.

Methods

A total of 988 breast cancer cases and 1,016 controls from the West of Ireland were collected with appropriate ethical approval as part of the Breast cancer in Galway Genetics Study (BIGGS). This population shares similar ethnic ancestry to the population in the UK, but has been subject to fewer demographic movements [28]. Consequently, this population is relatively homogenous, which reduces allelic and genotypic heterogeneity in case–control studies. The results are also likely to be relevant to several other populations of European ancestry. The mean age of the cases at diagnosis was 53 years (range 24–90 years). They were not selected with regard to family history of breast or ovarian cancer, personal history of ovarian cancer, the presence of a contralateral breast cancer or other second primary cancer (122 patients had a first degree relative with breast cancer, mean age 53.5 years; 153 had a second degree relative with breast cancer, mean age 50.5 years; and 40 had bilateral breast cancer, mean age 53.5 years). All controls were from a West of Ireland lineage (as were cases) and comprised women over the age of 60 years, with no self-reported personal history of any cancer and no family history of breast or ovarian cancer.

The power of our study, based on reasonable assumptions from previous studies, for the breast cancer SNPs rs2981582, rs3803662 and rs13281615, was as follows. Using an allelic test, 988 cases and 1,016 controls provided >90% power at

$P < 0.05$ (two-tailed test) to detect a SNP allele with a frequency of 0.4 associated with an OR of 1.3; >60% power to detect an allele of frequency of 0.25 associated with an OR of 1.2; and ~30% power to detect an allele of frequency 0.4 associated with an OR of 1.1.

DNA was extracted from blood using the Chemagic Magnetic Separation Module (Chemagen, Baesweiler, Germany) using the manufacturers reagents. A panel of 24 breast cancer, colorectal cancer and prostate cancer susceptibility loci were genotyped using the KASPar SNP genotyping system (KBiosciences), Table 1. Primers are available on request. A subset of samples were sequenced for each SNP to confirm the results. Sequencing was performed using the ABI 3100 automated sequencer and sequence analysis software (Applied Biosystems, CA).

Deviation of genotype frequencies from those expected under Hardy–Weinberg equilibrium (HWE) was assessed by χ^2 test, or Fisher's exact test where an expected cell count was <5. The risk associated with each SNP was estimated by allelic, dominant and recessive odds ratio (OR) and associated 95% confidence intervals (CI). Genotypic ORs were determined by logistic regression.

Determination of a general cancer risk allele was performed by meta-analysis using publicly available SNP data for breast and prostate cancer (CGEMS, http://cgems.cancer.gov/: breast study-1,183 postmenopausal cases and 1,185 controls; prostate study-1,177 cases and 1,105 controls) and SNP data from the first phase of our colorectal cancer GWAS (930 familial colorectal tumour cases and 960 controls) [15]. Data from these breast, prostate and colorectal cancer samples were combined with the West of Ireland cohort to derive an overall estimate of cancer risk using a fixed effects model with inverse variance weighting of individual studies. Reported predisposition SNPs not found on the Illumina platform data were imputed using IMPUTE [29]. Bias was assessed using the I^2 and Q estimates. For any significant findings in the fixed effects model with values $I^2 \geq 50\%$ (considered moderate heterogeneity [30]), the analysis was repeated using a random effects model. All statistical analyses were undertaken using STATA 9.2 (Stata Corp, College Station, TX).

Results

Verification of reported breast cancer SNPs in the BIGGS study

Samples that consistently failed in >20% of the SNPs studied were excluded from the analysis (25 cases, 20

Table 1 Panel of SNPs genotyped in the breast cancer population from Galway, West Ireland

SNP	Association	Gene	Locus
rs13281615	Breast [3]	POU5F1P1/MYC	8q (128424800)
rs2981582	Breast [3]	FGFR2	10q (123342307)
rs889312	Breast [3]	MAP3K1	5q (56067641)
rs3817198	Breast [3]	LSP1	11p (1865582)
rs3803662	Breast [3]	TNRC9/LOC643714	16q (51143842)
rs997669	Breast [7]	CCNE1	19q (34996323)
rs3176336	Breast [7]	CDKN1A	6p (36756544)
rs34330	Breast [7]	CDKN1B	12p (12761962)
rs4415084	Breast [6]		5p (44698272)
rs6983267	Colorectal/Prostate [15, 16]	POU5F1P1	8q (128482487)
rs16892766	Colorectal [26]		8q (117699864)
rs4779584	Colorectal [23]	GREM1, SGNE1	15q (30782048)
rs4939827	Colorectal [25]	SMAD7	18q (44707461)
rs3802842	Colorectal [24]	LOC120376	11q (110676919)
rs9929218	Colorectal [27]	CDH1	16q (67378697)
rs16901979	Prostate [11]	POU5F1P1	8q (128194098)
rs6465657	Prostate [20]	LMTK2	7q (97654013)
rs7920517	Prostate [20]	MSMB	10q (51202627)
rs7931342	Prostate [20]		11q (68751073)
rs2660753	Prostate [20]	LOC285232	3 (87193364)
rs5945619	Prostate [20]	NUDT11	Xp (51074708)
rs266849	Prostate [20]	KLK3 KLK15	19q (56040902)
rs2659056	Prostate [20]	KLK1 KLK15	19q (56027755)
rs10993994	Prostate [20]	MSMB	10q (51219502)

controls). Genotypes were scored in >94% of samples for each SNP; frequencies did not deviate significantly from Hardy–Weinberg equilibrium ($P > 0.01$ for controls) at any of the 24 SNPs. There was no linkage disequilibrium (LD) between any of the four SNPs at 8q24, with all pairwise r^2 values being zero and D′ values being <0.20.

We used a threshold of $P \leq 0.05$ (two-tailed) for replication of previous studies. Of the five breast cancer loci previously identified by Easton et al. only rs2981582 (FGFR2) and rs13281615 (8q24) showed an association with breast cancer in the West of Ireland population at $P < 0.05$, Table 2a. There was a borderline association with rs3803662 (TNRC9) at $P_{\text{allelic test}} = 5.1 \times 10^{-2}$. rs2981582 showed the strongest association with breast cancer, allelic test of association ($P = 1.8 \times 10^{-2}$, Table 2a) and the genotypic test ($P = 2 \times 10^{-3}$, Table 2a). The odds ratio (OR) for the TT homozygote was 1.489 (95% CI 1.153–1.922, $P = 2 \times 10^{-3}$) and that for the CT heterozygote was 1.176 (95% CI 0.958–1.443, $P = 1 \times 10^{-1}$), relative to the CC homozygote in each case. The association between rs13281615 and breast cancer was also confirmed using the allelic test of association ($P = 1.8 \times 10^{-2}$, Table 2a) and the genotypic test ($P = 0.026$, Table 2a). The odds ratio (OR) for the GG homozygote was 1.28 (95% CI 0.98–1.66, $P = 7 \times 10^{-2}$) and that for the AG heterozygote was 1.34 (95% CI 1.09–1.64, $P = 6 \times 10^{-3}$) relative to the AA homozygote. There was a borderline association between rs3803662 and breast cancer (allelic test of association, $P = 1.8 \times 10^{-2}$ and the genotypic test, $P = 4.7 \times 10^{-2}$, Table 2a). The odds ratio (OR) for the TT homozygote was 1.554 (95% CI 1.09–2.23, $P = 1.6 \times 10^{-2}$) and that for the CT heterozygote was 1.060 (95% CI 0.88–1.278, $P = 5.7 \times 10^{-1}$), relative to the CC homozygote in each case.

rs889312 (MAP3K1) and rs3817198 (LSP1) showed no association with breast cancer in this population ($P_{\text{allelic test}} = 2.4 \times 10^{-1}$, $P_{\text{allelic test}} = 5.4 \times 10^{-1}$, respectively). There was also no evidence of an association with the cell cycle gene SNPs (rs997669, rs3176336, rs34330) or rs4415084 (5p12) that have recently been suggested as breast cancer predisposition SNPs. It is not clear whether these differences are due to the different population studied or due to a lack of power in this study to detect these weaker associations. In order to answer this question we performed a meta-analysis on those SNPs where there was publicly available data in breast cancer, Table 3. This confirmed the association with rs4415084 (5p12) and rs3176336 (CDKN1A) also showed a significant association with breast cancer. However, the latter finding was based only on data from two studies (SEARCH and BIGGS) as there was no data available from CGEMS for this SNP. There was no evidence of an association between breast cancer and rs34330 (CDKN1B), rs997669 (CCNE1) or rs3817198 (LSP1), Table 3.

Association between breast cancer SNPs and clinico-patholgical features of breast cancer

Histological information was available on 795 (80%) cases and is summarised in Table 4. Univariate analysis showed evidence of an association between rs2981582 and PR positivity ($P = 3.5 \times 10^{-2}$), but not ER positivity ($P = 1.0 \times 10^{-1}$). Univariate analysis also showed a negative association between rs3803662 and her two positivity ($P = 3.7 \times 10^{-2}$), but no evidence of an association with ER/PR status. There was no evidence of an association between the rs13281615 risk allele and ER status, Her2 positivity, grade, nodal status or bilaterality. There was however an association with age and rs13281615 with younger cases more likely to have the at risk allele ($P = 4.1 \times 10^{-2}$).

Analysis of colorectal and prostate SNPs in the BIGGS study

The prostate and colorectal SNPs on 8q24 showed no association with breast cancer in our cohort and there was no association with any of the other low penetrance colorectal and prostate SNPs that were studied for any of the genetic models tested (Table 2b, c for allelic and genotypic tests). Logistic regression analysis provided no evidence for an independent effect of the colorectal or prostate SNPs studied on breast cancer risk (details not shown).

Meta-analysis to search for general cancer predisposition alleles for the three common cancers

We performed a meta-analysis in order to try and ascertain whether there was any evidence of a general cancer predisposition allele. A comprehensive list of SNPs known to confer an increased cancer risk of either breast, colorectal and prostate cancer was compiled (Table 5). As all publicly available data were based on the Illumina platform, reported predisposition SNPs not found on the Illumina platform data were imputed using IMPUTE [29] (rs4666451, rs981782, rs2180341, rs13281615, rs1045485). In the case of rs2981582 (FGFR2), we used a previously identified Illumina SNP (rs1219648) at this locus which has also been shown to be associated with the breast cancer [4]. Two SNPs (rs3176336 and rs889312) previously shown to predispose to breast cancer could not be imputed due to lack of HapMap data.

Data from publicly available sources (see "Methods") on genotype frequencies in cases and controls for these risk SNPs were combined using a fixed effects model for allelic, genotypic, recessive and dominant modes of risk. Only one SNP (rs6983267 on 8q24, $P < 5 \times 10^{-4}$) conferred risk to cancer in two tissue types (CRC and prostate cancer), as

Table 2 Genotype frequencies and tests of association for (a) Previously reported breast cancer association SNPs in BIGGS; (b) Colorectal cancer association SNPs in BIGGS; (c) Prostate cancer association SNPs in BIGGS

SNP	Genotype			Frequency		Allelic test			Genotypic test		
						P (2 tail)	OR	95% CI	*P* (2 tail)	OR	95% CI
(*a*)											
rs2981582	CC	CT	TT	C	T						
Cases	269	458	214	0.53	0.47	2.2×10^{-3}	1.22	1.07–1.39	2.00×10^{-3}	1.21	1.07–1.38
Controls	335	483	179	0.58	0.42						
rs13281615	AA	AG	GG	A	G						
Cases	272	467	178	0.55	0.45	1.8×10^{-2}	1.17	1.03–1.33	2.63×10^{-2}	1.16	1.02–1.31
Controls	355	456	182	0.59	0.41						
rs3803662	CC	CT	TT	C	T						
Cases	486	382	82	0.71	0.29	5.1×10^{-2}	1.15	1.00 –1.32	4.70×10^{-2}	1.16	1.00–1.33
Controls	532	396	58	0.74	0.26						
rs889312	AA	AC	CC	A	C						
Cases	449	400	88	0.69	0.31	2.38×10^{-1}	1.08	0.95 –1.25	2.34×10^{-1}	1.09	0.95–1.25
Controls	496	411	81	0.71	0.29						
rs3817198	TT	TC	CC	T	C						
Cases	363	454	121	0.63	0.37	5.36×10^{-1}	1.04	0.91–1.19	5.18×10^{-1}	0.96	0.84–1.09
Controls	379	453	116	0.64	0.36						
rs997669	AA	AG	GG	A	G						
Cases	282	497	152	0.57	0.43	3.20×10^{-1}	0.94	0.82–1.06	2.91×10^{-1}	0.93	0.82–1.06
Controls	305	471	200	0.55	0.45						
rs3176336	AA	AT	TT	A	T						
Cases	365	455	115	0.63	0.37	6.54×10^{-1}	0.94	0.85–1.16	6.38×10^{-1}	0.97	0.85–1.11
Controls	380	452	135	0.63	0.37						
rs34330	CC	CT	TT	C	T						
Cases	567	335	45	0.78	0.22	2.0×10^{-1}	0.91	0.78–1.05	2.15×10^{-1}	0.90	0.78–1.06
Controls	558	355	57	0.76	0.24						
rs4415084	CC	TC	TT	C	T						
Cases	290	417	175	0.57	0.43	8.28×10^{-1}	1.02	0.89–1.15	9.02×10^{-1}	1.01	0.86–1.15
Controls	323	488	186	0.57	0.43						
(*b*)											
rs6983267	GG	GT	TT	G	T						
Cases	251	464	230	0.51	0.49	5.56×10^{-1}	0.96	0.85–1.09	5.61×10^{-1}	0.96	0.85–1.09
Controls	248	464	245	0.50	0.50						
rs16892766	AA	AC	CC	A	C						
Cases	839	115	3	0.94	0.06	4.74×10^{-1}	1.10	0.85–1.43	4.72×10^{-1}	1.10	0.85–1.43
Controls	891	110	3	0.94	0.06						
rs4779584	CC	CT	TT	C	T						
Cases	671	262	30	0.83	0.17	1.16×10^{-1}	0.88	0.74–1.03	1.12×10^{-1}	0.89	0.74–1.03
Controls	655	301	34	0.81	0.19						
rs4939827	TT	TC	CC	T	C						
Cases	264	473	212	0.53	0.47	7.33×10^{-1}	0.99	0.86–1.11	7.36×10^{-1}	1.02	0.90–1.16
Controls	278	468	235	0.52	0.48						
rs3802842	AA	AC	CC	A	C						
Cases	412	447	70	0.75	0.25	3.65×10^{-1}	0.94	0.82–1.08	3.47×10^{-1}	0.93	0.81–1.08
Controls	416	428	96	0.72	0.28						

Table 2 continued

SNP	Genotype			Frequency		Allelic test			Genotypic test		
						P (2 tail)	OR	95% CI	*P* (2 tail)	OR	95% CI
rs9929218	GG	AG	AA	G	A						
Cases	488	399	77	0.72	0.28	1.31×10^{-1}	0.90	0.78–1.03	7.49×10^{-2}	0.85	0.71–1.02
Controls	467	451	84	0.69	0.31						
(*c*)											
rs16901979	CC	AC	AA	C	A						
Cases	909	54	3	0.97	0.03	7.74×10^{-1}	0.95	0.66–1.36	7.76×10^{-1}	0.95	0.67–1.35
Controls	930	65	0	0.97	0.03						
rs6465657	TT	CT	CC	TT	CC						
	285	441	198	0.55	0.45	7.16×10^{-1}	0.89	0.76–1.05	7.95×10^{-1}	1.02	0.90–1.15
	319	457	217	0.55	0.45						
rs7920517	AA	AG	GG	A	G						
Cases	259	473	223	0.52	0.48	1.83×10^{-1}	0.92	0.81–1.04	1.88×10^{-1}	0.92	0.81–1.04
Controls	251	472	256	0.50	0.50						
rs7931342	GG	GT	TT	G	T						
Cases	296	478	187	0.56	0.44	1.17×10^{-1}	0.90	0.80–1.03	1.16×10^{-1}	1.11	0.98–1.26
Controls	281	494	218	0.53	0.47						
rs2660753	CC	CT	TT	C	T						
Cases	786	176	7	0.90	0.10	4.17×10^{-1}	0.92	0.75–1.12	4.16×10^{-1}	0.92	0.75–1.12
Controls	787	182	13	0.89	0.11						
rs5945619	TT	CT	CC	T	C						
Cases	397	433	136	0.64	0.36	6.45×10^{-2}	1.13	0.99–1.30	1.06×10^{-1}	1.12	0.98–1.27
Controls	394	396	103	0.66	0.34						
rs266849	AA	AG	GG	A	G						
Cases	598	316	39	0.79	0.21	1.17×10^{-1}	0.90	0.80–1.03	1.59×10^{-1}	0.89	0.76−1.05
Controls	642	328	24	0.81	0.19						
rs2659056	AA	AG	GG	A	G						
Cases	522	368	56	0.75	0.25	7.40×10^{-2}	1.14	0.80–1.03	6.40×10^{-2}	0.87	0.75–1.01
Controls	498	410	68	0.72	0.28						
rs10993994	CC	CT	TT	C	T						
Cases	335	470	150	0.60	0.40	2.33×10^{-1}	0.93	0.81–1.05	2.36×10^{-1}	0.93	0.82–1.05
Controls	342	464	187	0.58	0.42						
rs16892766	AA	AC	CC	A	C						
Cases	839	115	3	0.94	0.06	4.74×10^{-1}	1.10	0.85–1.43	4.72×10^{-1}	1.10	0.85–1.43
Controls	891	110	3	0.94	0.06						

Table 3 Results of meta-analysis of breast cancer studies for rs3176336, rs34330, rs3817198, rs4415084, rs997669 using random effects model

SNP	Studies used in meta-analysis	Allelic test (*P* value-2 tailed)
rs3176336	BIGGS SEARCH[a]	$<5.00 \times 10^{-4}$
rs34330	BIGGS SEARCH[a] CGEMS	3.32×10^{-1}
rs3817198	BIGGS CGEMS	2.88×10^{-1}
rs4415084	BIGGS CGEMS	4.00×10^{-3}
rs997669	BIGGS SEARCH[a] CGEMS	9.88×10^{-1}

[a] Study which found a positive association

previously published [15, 16]. Of the remaining 31 SNPs analysed (5 associated with CRC risk, 13 with breast cancer, 13 with prostate cancer), none were significantly associated with cancer risk in another tumour type, either on its own or in combination of those tissue types in which the SNP was not discovered (Supplementary Table 1).

Table 4 Histological characteristics in 794 cases

Histological characteristics	Percentage of cases
Ductal	86.5
Lobular	13.5
Grade 1	13
Grade 2	39
Grade 3	35
Unknown grade	13
ER positive	60
ER negative	21
ER unknown	19
Node positive	46
Node negative	54
Bilateral cases	5
Family history of first degree relative with breast cancer	15
Family history of second degree relative with breast cancer	19

Discussion

We have provided replication of the association between rs2981582, rs13281615, and rs3803662 and breast cancer susceptibility in a sporadic breast cancer cohort from the West of Ireland. Unlike many previous studies, cases were not selected for family history of breast cancer [3] or enriched for bilateral breast cancers [31]. The allelic ORs detected in this study for both rs2981582 and rs3803662 were similar to that detected by Easton et al. For rs13281615 the OR of 1.17 (95% CI = 1.03–1.33) is higher than that estimated by Easton et al. (OR = 1.08, 95% CI = 1.05–1.11) in populations of European origin [3], but not as high as that estimated by Fletcher et al. (OR = 1.24, 95% CI = 1.12–1.38) in a series enriched for bilateral breast cancer cases [31], although there is overlap of the 95% CIs.

A previous study showed that the rs2981582 risk allele had a stronger association with ER and PR positive breast cancers than ER negative breast cancers [32]. A similar but weaker association was found for rs13281615 and rs3803662 [32]. We only found an association between PR-positive tumours and rs2981582 on univariate analysis. ER and PR status showed no association with rs13281615 and rs3803662 in this study. We did however find a negative association between rs3803662 and her two positivity and an association between age of breast cancer diagnosis and rs13281615, with younger breast cancer patients more likely to have the risk allele.

In the study by Easton et al. rs889312 (MAP3K1) and rs3817198 (LSP1) showed a weaker association with breast cancer than rs2981582, rs13281615, and rs3803662. Neither of these SNPs showed evidence of an association in BIGGS. We also found no evidence of an association between SNPs found within cell cycle genes and the West of Ireland breast cancer population, and did not verify the findings of Stacey et al. of a breast cancer predisposition SNP on 5p12. The latter association was only found in ER positive breast cancers. When we confined our analysis to this subgroup of breast cancer, we found no association with the 5p12 SNP, rs4415084, but did find an association between rs3176336, rs4779584 and rs7931342 and ER positive breast cancer ($P_{\text{allelic test}} = 0.02$, $P_{\text{allelic test}} = 0.03$, $P_{\text{allelic test}} = 0.03$, respectively), but this may represent modulation of phenotype, not susceptibility. The explanation for some of these differences is due to the power of our study as when we performed a meta-analysis combining our data with publicly available data we did confirm the association with rs4415084 (5p12), but not for rs34330, rs997669 or rs3817198.

We have performed a comprehensive meta-analysis of alleles that predispose to breast, prostate and colorectal cancer in order to test whether they are general carcinoma susceptibility loci. This showed that only one SNP (rs6983267 on 8q24) conferred risk to cancer in two tissue types (colorectal cancer and prostate cancer). Importantly, rs6983267 was not associated with an increase in breast cancer risk despite the known association of breast cancer with that region, overall, our data suggest that low penetrance susceptibility SNPs for breast, prostate and colorectal cancer are distinct leaving open the exploration for the increase in familial relative risk across the cancer types.

The most recent SNP found to predispose to colorectal cancer is rs9929218 which is located in intron1 of *CDH1* (E-cadherin) [27]. A SNP in the E-cadherin promoter (−160C → A) has previously been shown to be associated with prostate cancer [33]. However, the CGEMS prostate cancer study showed no evidence of an association with rs9929218 and prostate cancer. Loss of E-cadherin expression is characteristic of invasive lobular breast cancers [34] and germline mutations have been found in some familial lobular carcinomas [35]. There was no evidence of an association between rs9929218 and lobular cancer in the BIGGS cohort, however the study does not have the power to detect such an association if it existed with only 106 invasive lobular cancers in this series.

The finding of risk variants clustering at 8q24 is intriguing, our data and that of Fletcher et al. suggest that these variants are organ specific. Telomeric to the rs16901979 and rs13281615 haplotype blocks on 8q24 lie the expressed transcript *POU5F1P1* (128.50 Mb), that has potential to encode a homologue of the transcription factor *OCT4*, and the oncogene *MYC* (128.82 Mb). It remains possible that combinations of different genomic control

Table 5 Predisposition SNPs used in meta-analysis

	Chromosome	Position	Gene
Prostate predisposition SNPs			
rs6983267	8	128482487	POU5F1P1
rs7920517	10	51202627	MSMB
rs7931342	11	68751073	
rs2660753	3	87193364	LOC285232
rs5945619	×	51074708	NUDT11
rs266849	19	56040902	KLK3|KLK15
rs10993994	10	51219502	MSMB
rs9364554	6	160804075	SLC22A3
rs6465657	7	97460978	LMTK2
rs902774	12	51560171	
rs2659056	19	56027755	KLK1|KLK15
rs2735839	19	56056435	KLK3|KLK2
rs4430796	17	33172153	TCF2
Breast predisposition SNPs			
rs13281615	8	128424800	POU5F1P1
rs981782	5	45321225	
rs30099	5	52454339	MOCS2
rs4666451	2	19150174	
rs3803662	16	51143842	TNRC9
rs2180341	6	127642073	RNF146
rs1219648	10	123336180	FGFR2
rs1045485	2	201857584	casp8
rs997669	19	34996323	CCNE1
rs34330	12	12761962	CDKN1B
rs3731239	9	21964218	CDKN2A
rs13387042	2	217731338	
rs4415084	5	44698272	
Colorectal predisposition SNPs			
rs6983267	8	128482487	POU5F1P1
rs16892766	8	117699864	
rs4779584	15	30782048	GREM1|SGNE1
rs4939827	18	44707461	SMAD7
rs3802842	11	110676919	FLJ45803
rs9929218	16	67378697	CDH1

regions involved in regulating *MYC* or *POU5F1P1* expression are associated with susceptibility to different types of the common cancers. It has been proposed that variation at putative 8q24 cis-regulator(s) of transcription can significantly alter germline c-MYC expression levels and, thus, contribute to cancer susceptibility [36].

In conclusion we have confirmed that SNPs close to FGFR2 and TNCR9 and at 8q24(128424800) predispose to sporadic breast cancer in the West of Ireland. However, we have failed to replicate the association of SNPs close to MAP3K1, LSP1, CCNE1, CDKN1A, CDKN1B and at 5p12 in this population. We have shown that there is no association between the recently identified colorectal and prostate cancer SNPs and breast cancer. In general, prostate and colorectal cancer susceptibility alleles are site specific, despite some clustering at 8q24.

Acknowledgments We are grateful to the patients in the study and colleagues involved in patient ascertainment and sample collection. Help was also kindly provided by the Equipment Park, London Research Institute, Cancer Research UK. NM, GC and ST were funded by the National Breast Cancer Research Institute, Ireland. ES and AJ are funded by a Department of Health Clinician Scientist Grant, UK.

References

1. Antoniou AC, Easton DF (2006) Models of genetic susceptibility to breast cancer. Oncogene 25(43):5898–5905. doi:10.1038/sj.onc.1209879
2. Pharoah PD, Antoniou A, Bobrow M et al (2002) Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31(1):33–36. doi:10.1038/ng853
3. Easton DF, Pooley KA, Dunning AM et al (2007) Genome-wide association study identifies novel breast cancer susceptibility loci. Nature 447(7148):1087–1093. doi:10.1038/nature05887
4. Hunter DJ, Kraft P, Jacobs KB et al (2007) A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. Nat Genet 39(7):870–874. doi:10.1038/ng2075
5. Stacey SN, Manolescu A, Sulem P et al (2007) Common variants on chromosomes 2q35 and 16q12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 39(7):865–869. doi:10.1038/ng2064
6. Stacey SN, Manolescu A, Sulem P et al (2008) Common variants on chromosome 5p12 confer susceptibility to estrogen receptor-positive breast cancer. Nat Genet 40(6):703–706. doi:10.1038/ng.131
7. Driver KE, Song H, Lesueur F et al (2008) Association of single-nucleotide polymorphisms in the cell cycle genes with breast cancer in the British population. Carcinogenesis 29(2):333–341. doi:10.1093/carcin/bgm284
8. Ma H, Jin G, Hu Z et al (2006) Variant genotypes of CDKN1A and CDKN1B are associated with an increased risk of breast cancer in Chinese women. Int J Cancer 119(9):2173–2178. doi:10.1002/ijc.22094
9. Amundadottir LT, Sulem P, Gudmundsson J et al (2006) A common variant associated with prostate cancer in European and African populations. Nat Genet 38(6):652–658. doi:10.1038/ng1808
10. Freedman ML, Haiman CA, Patterson N et al (2006) Admixture mapping identifies 8q24 as a prostate cancer risk locus in African–American men. Proc Natl Acad Sci USA 10338:14068–14073. doi:10.1073/pnas.0605832103
11. Gudmundsson J, Sulem P, Manolescu A et al (2007) Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. Nat Genet 39(5):631–637. doi:10.1038/ng1999
12. Haiman CA, Patterson N, Freedman ML et al (2007) Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet 39(5):638–644. doi:10.1038/ng2015
13. Yeager M, Orr N, Hayes RB et al (2007) Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 39(5):645–649. doi:10.1038/ng2022
14. Schumacher FR, Feigelson HS, Cox DG et al (2007) A common 8q24 variant in prostate and breast cancer from a large nested

case–control study. Cancer Res 67(7):2951–2956. doi:10.1158/0008-5472.CAN-06-3591
15. Tomlinson I, Webb E, Carvajal-Carmona L et al (2007) A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet 39(8):984–988. doi:10.1038/ng2085
16. Zanke BW, Greenwood CM, Rangrej J et al (2007) Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet 39(8):989–994. doi:10.1038/ng2089
17. Ghoussaini M, Song H, Koessler T et al (2008) Multiple loci with different cancer specificities within the 8q24 gene desert. J Natl Cancer Inst 100(13):962–966. doi:10.1093/jnci/djn190
18. Witte JS (2007) Multiple prostate cancer risk variants on 8q24. Nat Genet 39(5):579–580. doi:10.1038/ng0507-579
19. Houlston RS, Peto J (2004) Genetics and the common cancers. In Eeles RA, Easton DF, Ponder BAJ, Eng C (eds) Genetic predisposition to cancer, 2nd edn, Arnold, London
20. Eeles RA, Kote-Jarai Z, Giles GG et al (2008) Multiple newly identified loci associated with prostate cancer susceptibility. Nat Genet 40(3):316–321. doi:10.1038/ng.90
21. Thomas G, Jacobs KB, Yeager M et al (2008) Multiple loci identified in a genome-wide association study of prostate cancer. Nat Genet 40(3):310–315. doi:10.1038/ng.91
22. Gudmundsson J, Sulem P, Rafnar T et al (2008) Common sequence variants on 2p15 and Xp11.22 confer susceptibility to prostate cancer. Nat Genet 40(3):281–283. doi:10.1038/ng.89
23. Jaeger E, Webb E, Howarth K et al (2008) Common genetic variants at the CRAC1 (HMPS) locus on chromosome 15q13.3 influence colorectal cancer risk. Nat Genet 40(1):26–28. doi:10.1038/ng.2007.41
24. Tenesa A, Farrington SM, Prendergast JG et al (2008) Genome-wide association scan identifies a colorectal cancer susceptibility locus on 11q23 and replicates risk loci at 8q24 and 18q21. Nat Genet 40(5):631–637. doi:10.1038/ng.133
25. Broderick P, Carvajal-Carmona L, Pittman AM et al (2007) A genome-wide association study shows that common alleles of SMAD7 influence colorectal cancer risk. Nat Genet 39(11):1315–1317. doi:10.1038/ng.2007.18
26. Tomlinson IP, Webb E, Carvajal-Carmona L et al (2008) A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14 and 8q23.3. Nat Genet 40(5):623–630. doi:10.1038/ng.111
27. Houlston RS, Webb E, Broderick P et al (2008) Meta-analysis of genome-wide association data identifies 4 novel susceptibility loci for colorectal cancer. Nat Genet (in press)
28. Hill EW, Jobling MA, Bradley DG (2000) Y-chromosome variation and Irish origins. Nat Genet 404(6776):351–352. doi:10.1038/35006158
29. Marchini J, Howie B, Myers S et al (2007) A new multipoint method for genome-wide association studies by imputation of genotypes. Nat Genet 39(7):906–913. doi:10.1038/ng2088
30. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327(7414):557–560. doi:10.1136/bmj.327.7414.557
31. Fletcher O, Johnson N, Gibson L et al (2008) Association of genetic variants at 8q24 with breast cancer risk. Cancer Epidemiol Biomarkers Prev 17(3):702–705. doi:10.1158/1055-9965.EPI-07-2564
32. Garcia-Closas M, Hall P, Nevanlinna H et al (2008) Heterogeneity of breast cancer associations with five susceptibility Loci by clinical and pathological characteristics. PLoS Genet 4(4):1–10. doi:10.1371/journal.pgen.1000054
33. Wang GY, Lu CQ, Zhang RM et al (2008) The E-cadherin gene polymorphism 160C-> A and cancer risk: a HuGE review and meta-analysis of 26 case-control studies. Am J Epidemiol 167(1):7–14
34. Qureshi HS, Linden MD, Divine G et al (2006) E-cadherin status in breast cancer correlates with histologic type but does not correlate with established prognostic parameters. Am J Clin Pathol 125(3):377–385
35. Masciari S, Larsson N, Senz J et al (2007) Germline E-cadherin mutations in familial lobular breast cancer. J Med Genet 44(11):726–731. doi:10.1136/jmg.2007.051268
36. Solé X, Hernández P, de Heredia ML et al (2008) Genetic and genomic analysis modeling of germline c-MYC overexpression and cancer susceptibility. BMC Genomics 9:12. doi:10.1186/1471-2164-9-12