Common Variants in *RB1* Gene and Risk of Invasive Ovarian Cancer

Honglin Song,[1] Susan J. Ramus,[3] Danielle Shadforth,[1] Lydia Quaye,[3] Susanne Kruger Kjaer,[4] Richard A. DiCioccio,[6] Alison M. Dunning,[1] Estrid Hogdall,[4] Claus Hogdall,[5] Alice S. Whittemore,[7] Valerie McGuire,[7] Fabienne Lesueur,[1] Douglas F. Easton,[2] Ian J. Jacobs,[3] Bruce A.J. Ponder,[1] Simon A. Gayther,[3] and Paul D.P. Pharoah[1]

[1]Cancer Research UK Department of Oncology and [2]Cancer Research UK Genetic Epidemiology Unit, Strangeways Research Laboratory, University of Cambridge, Cambridge, United Kingdom; [3]Translational Research Laboratory, Department of Gynaecological Oncology, Institute for Women's Health, University College London, London, United Kingdom; [4]Department of Viruses, Hormones, and Cancer, Institute of Cancer Epidemiology, Danish Cancer Society; [5]Gynaecologic Clinic, Juliane Marie Centre, Rigshospitalet, University of Copenhagen, Copenhagen, Denmark; [6]Department of Cancer Genetics, Roswell Park Cancer Institute, Buffalo, New York; and [7]Department of Health Research and Policy, Stanford University School of Medicine, Stanford, California.

Abstract

Somatic alteration of the *RB1* gene is common in several types of cancer, and germ-line variants are implicated in others. We have used a single nucleotide polymorphism (SNP) tagging approach to evaluate the association between common variants (SNP) in *RB1* and risks of invasive ovarian cancer. We genotyped 11 tagging SNPs in three ovarian case-control studies from the United Kingdom, United States, and Denmark, comprising >1500 cases and 4,800 controls. Two SNPs showed significant association with ovarian cancer risk: carriers of the minor allele of rs2854344 were at reduced risk compared with the common homozygotes [odds ratio (OR), 0.73; 95% confidence interval (95% CI), 0.61-0.89; P = 0.0009 dominant model]. Similarly, the minor allele of rs4151620 was found to be associated with reduced risk (rare versus common homozygote; OR, 0.19; 95% CI, 0.07-0.53; P = 0.00005 recessive model). After adjusting for multiple testing, the most significant association (rs4151620) was P = 0.001. A global test comparing common haplotype frequencies in cases and controls was of borderline significance (P_{sdf} = 0.04). There are no common coding SNPs in the *RB1* gene. However, intron 17 of *RB1* contains the open reading frame for the *P2RY5* gene, and rs4151620 is perfectly correlated with rs2227311, which is located in the 5′-untranslated region of *P2RY5* and is predicted to affect *P2RY5* transcription. rs2854344 has been reported previously to be associated with breast cancer risk. The possible associations of rs2854344 and rs4151620 with ovarian cancer risk warrant confirmation in independent case-control studies before studies on their biological mode of action. (Cancer Res 2006; 66(20): 10220-6)

Introduction

Ovarian cancer is one of the most lethal malignancies in women; >204,000 new cases are diagnosed annually worldwide (1). Family and twin studies suggest that inherited factors are an important cause of ovarian cancer (2), but the known ovarian cancer susceptibility genes, such as *BRCA1* and *BRCA2*, explain <30% of the excess familial risk. It is likely that the unexplained excess in familial risk is due to a combination of multiple low/moderate penetrance genetic variants, which are associated with relatively small effects on risk in the individual but contribute substantially to the overall risk in the population.

The task of identifying the genes and genetic variants that might confer susceptibility to ovarian cancer is currently restricted, for the most part, to selecting functionally plausible candidates. For human cancers, pathways involved in the control of replication and cellular proliferation provide many possible candidate genes (3). During mitosis, faithful replication of DNA during S phase and equal distribution of the identical chromosomal copies to the daughter cells during M phase is essential (4). The G_1-S phase of the cell cycle is crucial for the decision of cell whether to commit to growth arrest or proliferation. The function of the retinoblastoma protein, pRb (*RB1* gene product), is to connect the cell cycle control with the transcriptional machinery of cell. pRb is a negative regulator of cellular proliferation, which is achieved by sequestering a variety of nuclear proteins involved in cellular growth. When in a hypophosphorylated state, pRb exerts its antiproliferative function by sequestering and altering the function of the E2F family of transcription factors that control the expression of a bank of genes essential for cells to progress from G_1 into the S phase (5). Conditions that favor phosphorylation of pRb favor cell proliferation (6).

The *RB1* gene spans ~180 kb of genomic DNA on chromosome 13q14 and consists of 27 exons, encoding 928 amino acids. The largest intron (IVS 17) contains an open reading frame encoding the G protein–coupled receptor P2RY5 (purinergic receptor P2Y, G protein coupled, 5 NM_005767) in the reverse orientation relative to the transcription of *RB1* (7). The importance of *RB1* in human cancers and the role of the pRb pathway in cell proliferation and cancer development is well established. Most human cancers show somatic alteration, either in pRb or in one or other components of the pathway (8). Somatic inactivation of *RB1* has been reported in retinoblastomas (9), breast cancer (10), small cell lung cancers (11), and many sarcomas and bladder carcinomas (12). Germ-line mutations in *RB1* predispose to retinoblastoma (13). Very recently, polymorphisms of *RB1*, rs2854344 and rs198580, have been reported to be associated with altered breast cancer risk in the British population; carrying the minor allele of these single nucleotide polymorphisms (SNP) seems to reduce breast cancer risk (14).

Alterations in the pRb pathway are observed frequently in epithelial ovarian cancer (EOC); ~30% of ovarian cancers exhibit loss of heterozygosity at the *RB1* gene locus, but no mutations have been detected in the remaining *RB1* allele (15). Systematic analysis

Requests for reprints: Honglin Song, Cancer Research UK Department of Oncology, Strangeways Research Laboratory, University of Cambridge, Worts Causeway, Cambridge, United Kingdom CB1 8RN. Phone: 44-1223-740161; Fax: 44-1223-740147; E-mail: honglin@srl.cam.ac.uk.

©2006 American Association for Cancer Research.

doi:10.1158/0008-5472.CAN-06-2222

Downloaded from http://aacrjournals.org/cancerres/article-pdf/66/20/10220/2555798/10220.pdf by guest on 25 February 2024

of the genes involved in the entire pRb pathway (p16-cyclin-dependent kinase 4/cyclin D1-pRb) revealed that 80% of EOCs have abnormalities in this pathway (16). Recent research shows that the overall survival rate for EOC patients with a normal pRb pathway is significantly higher than for patients with an altered pRb pathway (17). Because *RB1* is key in a regulatory pathway that suffers disruption during the pathogenesis of various human tumors and there is some evidence of its association with breast cancer, it is reasonable to consider that common variants in the gene might explain some of the interindividual variability in the risk of EOC.

Population-based genetic association studies using large sets of cases and controls are a powerful method for identifying common low/moderate penetrance disease susceptibility alleles. The aim of this study was to investigate whether common variants in *RB1* gene were associated with ovarian cancer risk using a SNP tagging approach in a large, multicenter case-control study.

Materials and Methods

Study subjects. The cases and controls used for this study were from three different case-controls studies: the SEARCH ovarian cancer study from the United Kingdom, the MALOVA cancer study from Denmark, and the FROC study from the United States. Detailed description of the ovarian study subjects was presented in Song et al. (18). To summarize briefly, the SEARCH ovarian cancer study contains 731 invasive EOC cases collected from East Anglian, West Midlands cancer registries, and 855 controls randomly selected from European Prospective Investigation into Cancer and Nutrition (EPIC)-Norfolk cohort study. Genotype data for up to 4,555 EPIC-Norfolk subjects that had been randomly selected and genotyped as controls for the SEARCH breast cancer case series were also available for analysis. The MALOVA study contains up to 456 invasive ovarian cases and up to 1,231 controls randomly drawn from the same study area. The FROC study contains 327 invasive ovarian cases and 429 age-matched controls; non-White subjects from FROC were excluded from this analysis.

Tag SNP selection. Resequencing data from the National Institute of Environmental Health Sciences Environmental Genome Project (EGP) were available for *RB1* to enable a systematic selection of tagging SNPs (stSNP).[8] The EGP has been resequencing candidate genes for cancer across panels of individuals representative of U.S. ethnicities. The original panel (P1-PDR90) of 90 individuals consists of 24 European Americans, 24 African Americans, 12 Mexican Americans, 6 Native Americans, and 24 Asian Americans, but the ethnic group identifiers are not available. It is known that there is greater genetic and haplotype diversity in individuals of African origin and so we have identified and excluded 28 of the samples with the greatest African ancestry in this population by comparing the genotypes of the PDR90 sample with genotypes for the same SNPs from the National Heart, Lung, and Blood Institute Variation Discovery Resource Project African American panel.[9] Data from the remaining 62 individuals were used to identify stSNPs. Exclusion of the samples from Native American, Hispanic American, and Asian American individuals would be ideal, but because there is less genetic diversity between these groups, this cannot be done with any certainty.

We used the program Tagger to select stSNPs. Tagger uses a strategy that combines the simplicity of pairwise methods with the potential efficiency of multimarker approaches (19).[10] The aim of the SNP tagging was to identify a set of stSNPs that efficiently tags all the known common variants [minor allele frequency (MAF) >0.05] and is likely to tag most of the unknown common variants in the gene. The best measure of the extent to which one SNP tags another SNP is the pairwise correlation coefficient (r^2_p) because the loss in power incurred by using a marker SNP in place of a true causal SNP is directly related to this measure. We aimed to define a set of tagging SNPs such that all known common SNPs had an estimated r^2_p of >0.8, with at least one tagging SNP. However, some SNPs are poorly correlated with other single SNPs but may be efficiently tagged by a haplotype defined by multiple SNPs, so called "aggressive tagging," thus reducing the number of tagging SNPs needed. As an alternative, therefore, we aimed for the correlation between each SNP and a haplotype of tagging SNPs (r^2_S) to be at >0.8.

Genotyping. All samples were genotyped using the Taqman 7900HT Sequence Detection System according to the manufacturer's instructions. Each assay was carried out using 10 ng DNA in a 5 μL reaction using Taqman Universal PCR Master Mix (Applied Biosystems, Warrington United Kingdom), forward and reverse primers, and FAM- and VIC-labeled probes designed by Applied Biosystems (ABI Assay-by-Design). Primer and probe sequences and assay conditions used for each polymorphism analyzed are available from the corresponding author on request. All assays were carried out in 384-well arrays with 12 duplicate samples in each plate for quality control. Where discordant genotypes were observed in duplicates, the genotyping was repeated. Genotypes were determined using Allelic Discrimination Sequence Detection software (Applied Biosystems). DNA samples that did not give a clear genotype result at the first attempt were not repeated. Hence, there are variations in the number of samples successfully genotyped for each polymorphism. Call rates ranged from 94.5% to 99.5% for all the studies and SNPs and overall concordance between duplicate samples was >99%.

Statistics. The three studies were treated as separate strata in the analyses. Deviations of the genotype frequencies in the controls from those expected under Hardy-Weinberg equilibrium (HWE) were assessed by χ^2 tests [1 degree of freedom (*df*)]. The primary tests of association were the univariate analyses between each tagging SNP and ovarian cancer. Genotype frequencies in cases and controls were compared for each study separately using χ^2 tests for heterogeneity (2 *df*). The data from the three studies were then pooled and genotype frequencies were compared in cases and controls using unconditional logistic regression with terms for genotype and study and an appropriate likelihood-ratio test. We tested for heterogeneity between study strata by comparing logistic regression models with and without a genotype-stratum interaction term using likelihood ratio tests. Genotypic specific risks with the common homozygote as the baseline comparator were estimated as odds ratios (OR) with associated 95% confidence intervals (95% CI) by unconditional logistic regression. We also tested for rare allele dose effect (assuming a multiplicative model) using χ^2 tests (1 *df*) for each study separately and unconditional logistic regression for the pooled data.

In addition to the univariate analyses, we carried out a global haplotype test for association of the common haplotypes with EOC. Haplotype frequencies and subject-specific expected haplotype indicators were calculated separately for each study using the program TagSNPs, which implements an expectation-substitution approach to account for haplotype uncertainty given unphased genotype data (20). Subjects missing >50% genotype data were excluded. We considered haplotypes with >2% frequency in at least one study to be "common." Rare haplotypes were pooled. We used unconditional logistic regression to test the global null hypothesis of no association between haplotype frequency and ovarian cancer stratified by study, by comparing a model with multiplicative effects for each common haplotype (treating the most common haplotype as the referent) to the intercept-only model. Haplotype-specific ORs were also estimated with their associated confidence intervals.

Results

stSNP selection for ovarian cancer association study. The EGP resequencing data cover 38% of the genomic sequence of *RB1*. There were 25 common (MAF >0.05) variants in this data set with another 254 rare variants. Fourteen SNPs were initially chosen as stSNPs, but three assays failed design. There were no alternatives to these three (singletons). A fourth stSNP, rs4151580, was not

[8] http://egp.gs.washington.edu/.
[9] http://pga.gs.washington.edu/finished_genes.html.
[10] http://www.broad.mit.edu/mpg/tagger/.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/66/20/10220/2555798/10220.pdf by guest on 25 February 2024

polymorphic in our population. Thus, 10 stSNPs were genotyped. These tagged the remaining 11 common SNPs with minimum r_p^2 of 0.8; mean r_p^2 was 0.88. The SNPs that failed assays design were still tagged by the stSNPs with r_p^2 of 0.19, 0.28, and 0.46, respectively. Although not identified in the EGP data set, we also genotyped rs2854344 in our sample sets because it has been associated previously with breast cancer risk.

Association analysis of SNPs with ovarian cancer risk. We have genotyped these 11 SNPs in three ovarian cancer case-control studies. Genotype distributions in controls were consistent with

Table 1. *RB1* polymorphisms and genotype distributions by study

dbsnp*	Study	Controls				Cases				MAF		$P_{heterogeneity}$ †
		AA ‡	Aa§	aa‖	Total	AA ‡	Aa§	aa‖	Total	Control	Case	
rs1981434	SEARCH	1,608	1,250	252	3,110	377	287	61	725	0.28	0.28	0.94
	STAN	203	181	30	414	164	134	21	319	0.29	0.28	0.80
	MALOVA	586	494	115	1,195	210	183	40	433	0.30	0.30	0.94
	Combined	2,397	1,925	397	4,719	751	604	122	1,477	0.29	0.29	0.97
rs2854345	SEARCH	2,063	857	109	3,029	480	219	28	727	0.18	0.19	0.56
	STAN	280	135	8	423	208	104	11	323	0.18	0.20	0.42
	MALOVA	757	396	61	1,214	288	133	19	440	0.21	0.19	0.50
	Combined	3,100	1,388	178	4,666	976	456	58	1,490	0.19	0.19	0.92
rs399413	SEARCH	1,608	1,191	223	3,022	387	285	59	731	0.27	0.28	0.82
	STAN	210	188	29	427	170	131	23	324	0.29	0.27	0.61
	MALOVA	367	282	62	711	124	124	26	274	0.29	0.32	0.20
	Combined	2,185	1,661	314	4,160	681	540	108	1,329	0.28	0.29	0.73
rs4151540	SEARCH	1,672	1,212	232	3,116	386	292	51	729	0.27	0.32	0.81
	STAN	216	179	24	419	175	124	21	320	0.27	0.30	0.54
	MALOVA	572	502	114	1,188	228	164	42	434	0.31	0.29	0.25
	Combined	2,460	1,893	370	4,723	789	580	114	1,483	0.28	0.29	0.63
rs4151551	SEARCH	2,481	533	30	3,044	598	124	7	729	0.10	0.10	0.95
	STAN	365	56	2	423	267	52	4	323	0.07	0.09	0.27
	MALOVA	998	202	12	1,212	357	74	3	434	0.09	0.09	0.84
	Combined	3,844	791	44	4,679	1,222	250	14	1,486	0.09	0.09	0.95
rs2854344	SEARCH	4,450	780	31	5,261	642	75	5	722	0.08	0.06	0.006
	STAN	360	52	5	417	284	25	3	312	0.07	0.05	0.14
	MALOVA	608	107	3	718	241	39	4	284	0.08	0.08	0.22
	Combined	5,418	939	39	6,396	1,167	139	12	1,318	0.08	0.06	**0.0015**
rs425834	SEARCH	2,933	183	4	3,120	680	44	3	727	0.03	0.03	0.48
	STAN	380	42	0	422	300	21	2	323	0.05	0.04	0.08
	MALOVA	1,136	59	3	1,198	404	31	1	436	0.03	0.04	0.23
	Combined	4,449	284	7	4,740	1,384	96	6	1,486	0.03	0.04	0.33
rs4151611	SEARCH	4,926	460	11	5,397	667	57	0	724	0.04	0.04	0.44
	STAN	375	46	2	423	298	23	2	323	0.06	0.04	0.21
	MALOVA	1,089	114	2	1,205	385	42	3	430	0.05	0.06	0.23
	Combined	6,390	620	15	7,025	1,350	122	5	1,477	0.05	0.05	0.61
rs4151620	SEARCH	2,261	703	52	3,016	538	183	2	724	0.13	0.13	0.027
	STAN	311	83	5	399	231	81	0	314	0.12	0.13	0.05
	MALOVA	547	157	11	715	216	62	2	280	0.13	0.12	0.59
	Combined	3,119	943	68	4,130	985	326	4	1,318	0.13	0.13	**0.0001**
rs3092904	SEARCH	1,698	1,146	205	3,049	406	278	45	729	0.26	0.25	0.85
	STAN	229	171	20	420	188	114	20	322	0.25	0.24	0.28
	MALOVA	373	285	64	722	141	113	27	281	0.27	0.26	0.89
	Combined	2,300	1,602	289	4,191	735	505	92	1,332	0.27	0.26	0.95
rs4151636	SEARCH	2,859	258	6	3,123	661	65	2	728	0.05	0.05	0.76
	STAN	372	47	2	421	296	20	1	317	0.06	0.04	0.07
	MALOVA	1,109	107	1	1,217	399	37	0	436	0.04	0.04	0.90
	Combined	4,340	412	9	4,761	1,356	122	3	1,481	0.05	0.04	0.82

NOTE: Data in bold highlight the statistic significant results.

*SNPs are presented in sequential order on chromosome 13 configuration NT_024524.13 between chromosome position 47779155 to 47954735.

†Comparison of genotype frequencies in cases and controls (χ^2, 2 *df*).

‡Common homozygous.

§Heterozygous.

‖Rare homozygous.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/66/20/10220/2555798/10220.pdf by guest on 25 February 2024

Table 2. Genotypic specific risks (OR and 95% CI) for each of polymorphism by study

dbSNP	Study	HeOR* (95% CI)	HomOR* (95% CI)	P_{trend}
rs1981434	SEARCH	0.98 (0.83-1.16)	1.03 (0.76-1.40)	1.00
	STAN	0.92 (0.68-1.24)	0.87 (0.48-1.57)	0.51
	MALOVA	1.03 (0.82-1.30)	0.97 (0.66-1.44)	0.97
	Combined	0.98 (0.87-1.11)	0.99 (0.79-1.23)	0.82
rs2854345	SEARCH	1.10 (0.92-1.31)	1.10 (0.72-1.69)	0.30
	STAN	1.04 (0.76-1.42)	1.85 (0.73-4.68)	0.40
	MALOVA	0.88 (0.70-1.12)	0.82 (0.48-1.39)	0.24
	Combined	1.02 (0.89-1.16)	1.06 (0.78-1.44)	0.69
rs399413	SEARCH	0.99 (0.84-1.18)	1.10 (0.81-1.50)	0.71
	STAN	0.86 (0.64-1.16)	0.98 (0.55-1.76)	0.51
	MALOVA	1.30 (0.97-1.75)	1.24 (0.75-2.05)	0.12
	Combined	1.02 (0.89-1.17)	1.10 (0.87-1.40)	0.47
rs4151540	SEARCH	1.04 (0.88-1.24)	0.95 (0.69-1.31)	0.92
	STAN	0.86 (0.63-1.16)	1.08 (0.58-2.00)	0.61
	MALOVA	0.82 (0.65-1.04)	0.92 (0.63-1.36)	0.24
	Combined	0.94 (0.83-1.07)	0.96 (0.77-1.21)	0.44
rs4151551	SEARCH	0.97 (0.78-1.20)	0.97 (0.42-2.21)	0.75
	STAN	1.27 (0.84-1.91)	2.73 (0.50-15.04)	0.13
	MALOVA	1.02 (0.76-1.37)	0.70 (0.20-2.49)	0.93
	Combined	1.02 (0.87-1.20)	1.03 (0.56-1.91)	0.76
rs2854344	SEARCH	0.67 (0.52-0.86)	1.12 (0.43-2.89)	0.005
	STAN	0.61 (0.37-1.01)	0.76 (0.18-3.21)	0.07
	MALOVA	0.92 (0.62-1.37)	3.36 (0.75-15.1)	0.76
	Combined	**0.71 (0.58-0.86)**	1.32 (0.65-2.68)	**0.006**
rs425834	SEARCH	1.04 (0.74-1.46)	3.23 (0.72-14.49)	0.46
	STAN	0.63 (0.37-1.09)	N/A	0.29
	MALOVA	1.48 (0.94-2.32)	0.94 (0.10-9.04)	0.12
	Combined	1.03 (0.81-1.32)	2.68 (0.88-8.19)	0.39
rs4151611	SEARCH	0.92 (0.69-1.22)	N/A	0.37
	STAN	0.63 (0.37-1.06)	1.26 (0.18-8.99)	0.15
	MALOVA	1.04 (0.72-1.51)	4.24 (0.71-25.49)	0.43
	Combined	0.92 (0.74-1.14)	1.41 (0.46-4.34)	0.59
rs4151620	SEARCH	1.09 (0.91-1.32)	0.16 (0.04-0.67)	0.65
	STAN	1.31 (0.93-1.87)	N/A	0.42
	MALOVA	1.00 (0.72-1.40)	0.46 (0.10-2.09)	0.65
	Combined	1.11 (0.96-1.29)	**0.20 (0.07-0.55)**	0.81
rs3092904	SEARCH	1.01 (0.86-1.20)	0.92 (0.65-1.29)	0.83
	STAN	0.81 (0.60-1.10)	1.22 (0.64-2.33)	0.59
	MALOVA	1.05 (0.78-1.40)	1.12 (0.68-1.82)	0.63
	Combined	0.98 (0.86-1.12)	1.01 (0.78-1.30)	0.87
rs4151636	SEARCH	1.09 (0.82-1.45)	1.44 (0.29-7.16)	0.49
	STAN	0.56 (0.33-0.96)	0.63 (0.06-6.94)	0.03
	MALOVA	1.01 (0.69-1.48)	N/A	0.79
	Combined	0.93 (0.75-1.16)	0.96 (0.25-3.71)	0.54

NOTE: Data in bold highlight the statistic significant results.
*Compared with common homozygote.

HWE, except for rs425834, in the MALOVA study ($P = 0.02$). This is likely to be a chance observation because the discrimination of genotypes for this assay was good and deviation from HWE was not seen in cases.

The genotype frequencies in controls were similar for all three study populations with minor differences for rs2854345, rs425834, and rs4151540 (Table 1). The observed genotype frequencies by study and for the combined data set are presented in Table 1, which also shows the result of the test for the comparison of genotype frequencies ($P_{heterogeneity}$) between cases and controls. The genotypic specific risks for the individual studies and combined data and the result of the trend test for association are presented in Table 2. There was no association in controls between age and genotype frequency for any of the SNPs, and age-adjusted genotype-specific ORs were similar to unadjusted ORs (data not shown). There was no significant heterogeneity between strata for any of the SNPs studied (data not shown). There was a significant difference in genotype frequency in ovarian cancer cases and controls for rs2854344 ($P_{heterogeneity} = 0.0015$) and rs4151620 ($P_{heterogeneity} = 0.0001$). There was no evidence for heterogeneity

Downloaded from http://aacrjournals.org/cancerres/article-pdf/66/20/10220/2555798/10220.pdf by guest on 25 February 2024

Table 3. Estimated haplotype frequencies in cases and controls and haplotypic specific risks

Haplotype*	Control frequency	Case frequency	OR (95% CI)	χ^2 test, *P*	Global test, *P*
h00000000000	0.38	0.39	1		$\chi^2 = 16.07$
h11110000010	0.17	0.18	1.03 (0.91-1.16)	0.66	$P_{8df} = 0.04$
h00000000100	0.12	0.13	1.02 (0.88-1.18)	0.83	
h00001000000	0.09	0.09	1.01 (0.87-1.18)	0.88	
h10110100010	0.07	0.06	0.85 (0.71-1.02)	0.09	
h00000001000	0.04	0.04	0.93 (0.75-1.15)	0.48	
h00000000001	0.03	0.03	0.92 (0.71-1.19)	0.52	
h10100010000	0.03	0.04	1.15 (0.91-1.46)	0.24	
rare	0.06	0.04	0.73 (0.59-0.90)	0.003	

*In haplotype, 0 represents common allele and 1 represents rare allele. SNPs used in haplotypes are in the order rs1981434, rs2854345, rs399413, rs4151540, rs4151551, rs2854344, rs425834, rs4151611, rs4151620, rs3092904, and rs4151636.

between studies for these two SNPs ($P = 0.19$ and 0.62, respectively). There was no difference in genotype frequencies for the other nine SNPs. The best fitting model for rs4151620 was recessive; rare allele homozygote being at ~80% reduced risk of ovarian cancer (OR, 0.19; 95% CI, 0.07-0.53; $P = 0.00005$). For rs2854344, the best fitting model was dominant, this model fitting the data slightly better than a codominant one ($P = 0.0009$ versus $P = 0.005$), with 27% reduced risk for minor allele carriers (OR, 0.73; 95% CI, 0.61-0.89). The risks were similar after exclusion of the additional controls from the SEARCH breast cancer study but less significant as would be expected with a smaller sample size. The OR (95% CI) for rs4151620 rare allele homozygote compared with common homozygote was 0.22 (0.08-0.62), and the OR (95% CI) for rs2854344 rare allele carriers was 0.77 (0.62-0.95). The two associated SNPs are only weakly correlated with each other ($r^2 = -0.07$) and their effects seem to be independent. In a stepwise logistic regression, including all 11 SNPs, only the terms for the minor allele homozygotes of rs4151620 ($P = 0.003$) and the rs2854344 heterozygote ($P = 0.001$) remained in the final model.

The haplotype frequencies for *RB1* gene in cases and controls were estimated after stratification and the haplotype-specific risks and association test are presented in Table 3. No significant differences in individual haplotype frequencies were seen between cases and controls and the global test was only marginally significant ($P_{8df} = 0.04$).

Statistical power to identify subgroup effects is limited, but we did carry out subgroup analysis on the serous histopathologic subtype of cases. The genotypic specific risks estimated from the combined data are presented in Table 4. These were similar to the overall risks, although the protective effect for the rare allele of rs4151620 was somewhat stronger [OR, 0.11; 95% CI, (0.02-0.80)].

Discussion

The *RB1* gene product is a crucial component of the cell cycle control pathways. Loss of pRb function deprives the pathway and thus the cell of an important mechanism for disrupting cell proliferation through modulation of gene expression (12). We chose a SNP tagging approach to investigate the association of invasive ovarian cancer with the *RB1* gene region. To our knowledge, this is the first case-control study to investigate common polymorphisms in the *RB1* gene region in relation to ovarian cancer susceptibility.

Table 4. Genotypic specific risks (OR and 95% CI) for each of polymorphism among serous type EOC in combined data

dbSNP	Controls	Cases	Heterozygote OR* (95% CI)	Homozygote OR* (95% CI)	$P_{heterogeneity}$	P_{trend}
rs1981434	4,719	684	1.01 (0.85-1.20)	1.03 (0.76-1.40)	0.97	0.82
rs2854345	4,666	693	0.93 (0.78-1.12)	1.22 (0.83-1.81)	0.41	0.97
rs399413	4,160	584	1.07 (0.88-1.28)	1.16 (0.84-1.62)	0.60	0.32
rs4151540	4,723	688	0.87 (0.73-1.03)	0.97 (0.72-1.32)	0.28	0.30
rs4151551	4,679	688	0.98 (0.79-1.23)	1.15 (0.51-2.62)	0.93	0.99
rs2854344	6,396	581	0.76 (0.58-1.01)	0.97 (0.33-2.85)	0.14	0.08
rs425834	4,740	689	1.13 (0.81-1.56)	4.74 (1.42-15.8)	**0.048**	0.09
rs4151611	7,025	687	0.85 (0.63-1.14)	1.14 (0.24-5.32)	0.54	0.34
rs4151620	4,130	582	1.09 (0.89-1.34)	0.11 (0.02-0.81)	**0.004**	0.68
rs3092904	4,191	587	0.93 (0.77-1.13)	1.08 (0.76-1.52)	0.65	0.88
rs4151636	4,761	686	0.85 (0.63-1.16)	0.62 (0.07-5.18)	0.53	0.27

NOTE: Data in bold highlight the statistically significant results.
*Compared with common homozygote.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/66/20/10220/2555798/10220.pdf by guest on 25 February 2024

Table 5. FPRP values for two results on association of the variants with invasive ovarian cancer

SNP	OR (95% CI)	Statistical power	*P*	Prior probability					
				0.1	0.01	0.001	0.0001	0.00001	0.000001
Rs2854344	0.73 (0.61-0.89)	0.43*	0.0009	0.02	0.18	0.68	0.96	1.0	1.0
Rs4151620	0.19 (0.07-0.53)	0.27 †	0.00005	0.00	0.03	0.21	0.73	0.96	0.96

*Using dominant model for rs2854344.
†Using recessive model for rs4151620.

The SNPs tested here were chosen to efficiently tag all the known common variants in the gene and not for their likely functional effect. Our stSNP selection was based on EGP resequencing data, but only 38% of the gene had been resequenced; therefore common SNPs in the nonsequenced regions will have been missed. However, the HapMap project has genotyped over 300 SNPs in *RB1* (phase II release no. 20). Of these, only 22 SNPs in a single linkage disequilibrium (LD) block were common (MAF >0.05) in the samples of north-western European ancestry by the centre d'Etude du polymorphisme Humain (CEPH). Nine HapMap SNPs were also in the EGP data set (5 stSNPs). The 5 stSNPs (i.e., rs1981434, rs520342, rs399413, rs4151551, and rs3092904) tagged the 22 HapMap SNPs with mean r^2_p of 0.83 and 17 SNPs were captured with $r^2 > 0.8$. The full set of 11 stSNPs would be expected to do better. stSNPs selected from HapMap phase II data provide good power to detect all common variation (21). We are therefore confident that our set of stSNPs has adequately tagged the known and unknown common variants within the gene.

We found a significant association for the minor allele of two SNPs, rs2854344 and rs4151620; both were associated with reduced EOC risks. Similarly, rs2854344 has also been reported to be associated with reduced breast cancer risk (14). Carriers of the minor allele of rs2854344 seems to have reduced breast cancer risk compared with common allele homozygote (OR, 0.86; 95% CI, 0.76-0.96). Nevertheless, these results ought to be interpreted with caution. The *P*-value presented above have not been adjusted for multiple hypothesis testing. Because SNPs within the same gene may be in LD, the test statistics for the 11 SNPs are not independent, and standard methods for adjusting for multiple testing, such as the Bonferroni correction, would be too conservative. We therefore used a simulation to determine an empirical *P* for the most significant result (i.e., $P_{recessive} = 0.00005$ for rs4151620). In this analysis, we randomly shuffled the case-control status among individuals multiple times and estimated how frequently a $P < 0.00005$ was obtained from the randomly permuted data. This method also accounts for the testing of multiple genetic models with each SNP. In 10,000 permutations, a $P < 0.00005$ was observed on 11 occasions, giving the most significant *P* corrected for multiple testing of 0.0011. Nevertheless, the possibility that this result is a false positive remains substantial. One way of assessing this probability is to use the false-positive report probability (FPRP). The FPRP is the estimated probability that a specific result is a false positive. It depends on the prior probability of a true association, the observed level of significance (α), and the statistical power to detect the OR of the alternative hypothesis at the given α (22). As there is a very large number of common SNPs in the genome, the overall prior probability of association is very low (say 1 in a million). However, the prior is likely to be more favorable because we selected *RB1* as a candidate based on known biology and one of the two associated SNPs (rs2854344) has also been reported to be associated with breast cancer with similar protective effect. The FRPRs for the two associated SNPs under different prior probabilities are presented in Table 5. Hidden population stratification is an alternative explanation for a spurious association. This occurs when allele frequencies differ between population subgroups and case and controls are drawn differentially from those subgroups. It seems unlikely that population stratification is relevant in this association study because the cases and controls in the three studies reported here were drawn from the same ethnic groups. Furthermore, if stratification were present, it is unlikely that the same degree of stratification would be seen in all three studies.

Assuming the results to be real, they may either be due to a direct causative effect of the two SNPs tested or maybe because these SNPs are markers for the real determinant of a disease in LD with them. Both rs2854344 and rs4151620 are noncoding, in intron 17 and intron 24, respectively, and neither is in a highly conserved region.[11] Intron 17 contains an open reading frame encoding the G protein–coupled receptor P2RY5 in the reverse orientation relative to the transcription of *RB1* (7). P2RY5 belongs to a family of purine and pyrimidine nucleotide receptors that are coupled to G proteins. Activation results in mobilization of inositol 1.4.5-trisphosphate-sensitive Ca^{2+} stores, activation of inward plasma membrane currents, and stimulation of diacylglyceride-dependent protein kinases (23). There is increasing evidence that purinergic signaling can have prolonged effects on cell growth and proliferation (24, 25). The presence of P2Y receptors has been described in various cancer types, including melanoma, prostate, lung, esophageal, and colorectal cancers (26–30). The *P2RY5* gene consists of only one coding exon and rs2854344 lies 8-kb 5′-untranslated region (UTR) of it. The bioinformatics tool PupaSNP suggests that the variant rs2854344 does not have a functional effect or alters dramatically the structure of *RB1* or *P2RY5* (data not shown).[12] Thus, it seems likely that any true causal variant(s) will be in LD with rs2854344 and rs4151620. No coding SNPs in *RB1* were identified by the EGP and we found no common variants in the unique coding exon of *P2RY5* or 500 bp of its flanking sequences in a panel of 48 controls. rs2227311, located in 5′-UTR of *P2RY5*, is in perfect

[11] http://genome.ucsc.edu/cgi-bin/hgGateway.
[12] http://pupasnp.bioinfo.cnio.es.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/66/20/10220/2555798/10220.pdf by guest on 25 February 2024

LD with rs4151620 (r^2_p = 1). PupaSNP predicts that rs2227311 disrupts the exonic splicing enhancer (ESE) consensus sequence that is responsive to human serine/arginine-rich (SR) proteins of sc35 and srp40 (31). ESE sequences are the binding sites for specific SR proteins involved in the splicing machinery and growing evidence suggests that alternative splicing produced by ESEs is related to different diseases (32, 33). It has been estimated that 15% of point mutations that result in human genetic diseases cause RNA splicing defects (34). Therefore, rs2227311, which is tagged by rs4151620, is a good candidate causal variant for the observed association. However, we cannot exclude the possible involvement of other variants in *RB1*. Apparently, unidentified variants tagged by rs2854344 and rs4151620 present in the promoter or regulatory region or intron-exon boundaries affect the transcription of *RB1*.

In conclusion, we have genotyped 11 stSNPs tagging the known common variants in *RB1* in three ovarian cancer case-control studies. We found some evidence of association for rs2854344 and rs4151620 with invasive ovarian cancer risk. The observed associations with ovarian cancer risk warrant confirmation in independent studies before further functional studies.

Acknowledgments

Received 6/16/2006; revised 7/18/2006; accepted 8/7/2006.

Grant support: Cancer Research UK, Roswell Park Alliance, Danish Cancer Society, National Cancer Institute (CA71766, core grants CA16056 and RO1 CA61107), WellBeing grant (H. Song), and Mermaid component of the Eve Appeal (S.J. Ramus). B.A.J. Ponder is a Gibb Fellow, P.D.P. Pharoah is a Senior Clinical Research Fellow, S.A. Gayther is a Higher Education Funding Council for England–Funded Senior lecturer, and D.F. Easton is a Principal Research Fellow of Cancer Research UK.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Joan MacIntosh, Hannah Munday, Barbara Perkins, Clare Jordan, Kristy Driver, Mitul Shah, the local general practices and nurses, and the East Anglian Cancer Registry for recruitment of the United Kingdom cases; the EPIC-Norfolk investigators for recruitment of the United Kingdom controls; Craig Luccarini and Don Conroy for expert technical assistance; and all the study participants who contributed to this research.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P. Global cancer statistics, 2002. CA Cancer J Clin 2005;55:74–108.
2. Pharoah PD, Ponder BA. The genetics of ovarian cancer. Best Pract Res Clin Obstet Gynaecol 2002;16:449–68.
3. Tessema M, Lehmann U, Kreipe H. Cell cycle and no end. Virchows Arch 2004;444:313–23.
4. Sherr CJ. Cancer cell cycles. Science 1996;274:1672–7.
5. Hanahan D, Weinberg RA. The hallmarks of cancer. Cell 2000;100:57–70.
6. Cobrinik D, Dowdy SF, Hinds PW, Mittnacht S, Weinberg RA. The retinoblastoma protein and the regulation of cell cycling. Trends Biochem Sci 1992;17:312–5.
7. Herzog H, Darby K, Hort YJ, Shine J. Intron 17 of the human retinoblastoma susceptibility gene encodes an actively transcribed G protein-coupled receptor gene. Genome Res 1996;6:858–61.
8. Yaginuma Y, Hayashi H, Kawai K, et al. Analysis of the Rb gene and cyclin-dependent kinase 4 inhibitor genes (p16INK4 and p15INK4B) in human ovarian carcinoma cell lines. Exp Cell Res 1997;233:233–9.
9. Benedict WF, Fung YK, Murphree AL. The gene responsible for the development of retinoblastoma and osteosarcoma. Cancer 1998;62:1691–4.
10. T'Ang A, Varley JM, Chakraborty S, Murphree AL, Fung YK. Structural rearrangement of the retinoblastoma gene in human breast carcinoma. Science 1988;242:263–6.
11. Yokota J, Akiyama T, Fung YK, et al. Altered expression of the retinoblastoma (RB) gene in small-cell carcinoma of the lung. Oncogene 1988;3:471–5.
12. Weinberg RA. The retinoblastoma protein and cell cycle control. Cell 1995;81:323–30.
13. Friend SH, Bernards R, Rogelj S, et al. A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma. Nature 1986;323:643–6.
14. Lesueur F, Song H, Ahmed S, et al. Single-nucleotide polymorphisms in the RB1 gene and association with breast cancer in the British population. Br J Cancer, 2006.
15. Li SB, Schwartz PE, Lee WH, Yang-Feng TL. Allele loss at the retinoblastoma locus in human ovarian cancer. J Natl Cancer Inst 1991;83:637–40.
16. Hashiguchi Y, Tsuda H, Yamamoto K, Inoue T, Ishiko O, Ogita S. Combined analysis of p53 and RB pathways in epithelial ovarian cancer. Hum Pathol 2001;32:988–96.
17. Hashiguchi Y, Tsuda H, Inoue T, Nishimura S, Suzuki T, Kawamura N. Alteration of cell cycle regulators correlates with survival in epithelial ovarian cancer patients. Hum Pathol 2004;35:165–75.
18. Song H, Ramus S, Quaye L, et al. Common variants in mismatch repair genes and risk of invasive ovarian cancer. Carcinogenesis, 2006; doi:10.1093/carcin/bgl089.
19. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet 2005;37:1217–23.
20. Zaykin DV, Westfall PH, Young SS, Karnoub MA, Wagner MJ, Ehm MG. Testing association of statistically inferred haplotypes with discrete and continuous traits in samples of unrelated individuals. Hum Hered 2002;53:79–91.
21. Altshuler D, Brooks LD, Chakravarti A, Collins FS, Daly MJ, Donnelly P. A haplotype map of the human genome. Nature 2005;437:1299–320.
22. Wacholder S, Chanock S, Garcia-Closas M, El Ghormli L, Rothman N. Assessing the probability that a positive report is false: an approach for molecular epidemiology studies. J Natl Cancer Inst 2004;96:434–42.
23. Adrian K, Bernhard MK, Breitinger HG, Ogilvie A. Expression of purinergic receptors (ionotropic P2X1-7 and metabotropic P2Y1-11) during myeloid differentiation of HL60 cells. Biochim Biophys Acta 2000;1492:127–38.
24. Burnstock G. Purinergic signaling and vascular cell proliferation and death. Arterioscler Thromb Vasc Biol 2002;22:364–73.
25. Greig AV, Linge C, Cambrey A, Burnstock G. Purinergic receptors are part of a signaling system for keratinocyte proliferation, differentiation, and apoptosis in human fetal epidermis. J Invest Dermatol 2003;121:1145–9.
26. Hopfner M, Maaser K, Barthel B, et al. Growth inhibition and apoptosis induced by P2Y2 receptors in human colorectal carcinoma cells: involvement of intracellular calcium and cyclic adenosine monophosphate. Int J Colorectal Dis 2001;16:154–66.
27. Janssens R, Boeynaems JM. Effects of extracellular nucleotides and nucleosides on prostate carcinoma cells. Br J Pharmacol 2001;132:536–46.
28. Maaser K, Hopfner M, Kap H, et al. Extracellular nucleotides inhibit growth of human oesophageal cancer cells via P2Y(2)-receptors. Br J Cancer 2002;86:636–44.
29. Schafer R, Sedehizade F, Welte T, Reiser G. ATP- and UTP-activated P2Y receptors differently regulate proliferation of human lung epithelial tumor cells. Am J Physiol Lung Cell Mol Physiol 2003;285:L376–85.
30. White N, Ryten M, Clayton E, Butler P, Burnstock G. P2Y purinergic receptors regulate the growth of human melanomas. Cancer Lett 2005;224:81–91.
31. Graveley BR. Sorting out the complexity of SR protein functions. RNA 2000;6:1197–211.
32. Caceres JF, Kornblihtt AR. Alternative splicing: multiple control mechanisms and involvement in human disease. Trends Genet 2002;18:186–93.
33. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. Nat Rev Genet 2002;3:285–98.
34. Krawczak M, Reiss J, Cooper DN. The mutational spectrum of single base-pair substitutions in mRNA splice junctions of human genes: causes and consequences. Hum Genet 1992;90:41–54.

Downloaded from http://aacrjournals.org/cancerres/article-pdf/66/20/10220/2555798/10220.pdf by guest on 25 February 2024