Multiple Loci With Different Cancer Specificities Within the 8q24 Gene Desert


Recent studies based on genome-wide association, linkage, and admixture scan analysis have reported associations of various genetic variants in 8q24 with susceptibility to breast, prostate, and colorectal cancer. This locus lies within a 1.18-Mb region that contains no known genes but is bounded at its centromeric end by FAM84B and at its telomeric end by c-MYC, two candidate cancer susceptibility genes. To investigate the associations of specific loci within 8q24 with specific cancers, we genotyped the nine previously reported cancer-associated single-nucleotide polymorphisms across the region in four case-control sets of prostate (1854 case subjects and 1894 control subjects), breast (2270 case subjects and 2280 control subjects), colorectal (2299 case subjects and 2284 control subjects), and ovarian (1975 case subjects and 3411 control subjects) cancer. Five different haplotype blocks within this gene desert were specifically associated with risks of different cancers. One block was solely associated with risk of breast cancer, three others were associated solely with the risk of prostate cancer, and a fifth was associated with the risk of prostate, colorectal, and ovarian cancer, but not breast cancer. We conclude that there are at least five separate functional variants in this region.

J Natl Cancer Inst 2008;100:962–966

Recently, genome-wide association studies have been effective at identifying common genetic variants or single-nucleotide polymorphisms (SNPs) associated with common disease risk without any presumption about their localization or function. Recent studies have identified and confirmed associations of breast, prostate, and colorectal cancer with several variants within a 600-Kb region of a longer, 1.18-Mb, sequence that does not code for any known genes on chromosome 8q24 (1–10). Large chromosomal regions devoid of genes (often referred to as gene deserts) have been discovered to be associated with several diseases, indicating that they may have a function. Here, we have genotyped the nine previously reported cancer-associated SNPs across the region: rs13254738, rs6983561, rs16901979, rs13281615, rs10505477, rs10808556, rs6983267, rs7000448, and rs1447295 (or a good surrogate SNP for each fourth footnote to Table 1) in four large case-control sets of prostate (1854 case subjects and 1894 control subjects), breast (2270 case subjects and 2280 control subjects), colorectal (2299 case subjects and 2284 control subjects), and ovarian (1975 case subjects and 3411 control subjects) cancer (Table 1 and Supplementary Table 1). Case subjects with colorectal and breast cancer were drawn from SEARCH, an ongoing population-based study in East Anglia, UK. Control subjects were randomly selected from the Norfolk, UK, component of EPIC (European Prospective Investigation of Cancer). Case subjects and control subjects for prostate cancer were also drawn from the UK population, whereas case subjects and control subjects for ovarian cancer were selected from four different studies from the United Kingdom, United States, and Denmark. To test the association between these nine variants and the four types of cancer, we performed univariate analysis and compared genotype frequencies in case subjects and control subjects using unconditional logistic regression.

The data we generated from the above case-control studies show that there are at
least five different cancer susceptibility loci within the 8q24 “desert,” each separated from the others by recombination hot spots and each specific for cancer of particular tissue type (Table 1 and Figure 1, A).

Region 1, the most centromeric block, spans base positions 128.14–128.28 Mb (NCBI Build 35). SNP rs16901979 (1.3, Table 1) was reported to be associated with prostate cancer by two independent studies (4,5). More recently, rs13254738 (1.1) and rs6983561 (1.2) have also been found to be associated with prostate cancer (5). However, SNPs 1.2 and 1.3 are highly correlated; thus, they reflect the same association (Figure 1, B). We confirmed the association of these SNPs with prostate cancer (odds ratio [OR] = 1.12, 95% confidence interval [CI] = 1.01 to 1.24, P-value from Cochran Armitage test for trend = 0.029 for rs13254738; OR = 2.11, 95% CI = 1.65 to 2.71, P-value = 1.4 × 10^-4 for rs6983561; and OR = 2.06, 95% CI = 1.61 to 2.65, P-value = 4.9 × 10^-4 for rs16901979) but found no evidence for their association with risks of breast, colorectal, or ovarian cancers. The only published study that addressed the association of these SNPs with risk of colon cancer also found no evidence for an association (6). To our knowledge, no other studies have specifically addressed the association of these SNPs with breast, ovarian, or other cancer types. Thus, variants in region 1 appear to be specifically associated with the risk of prostate cancer.

Region 2, spanning base positions 128.35–128.51 Mb, was first identified as a potential breast cancer susceptibility locus by a genome-wide scan; this identification was confirmed by a study of 21,860 case subjects and 22,578 control subjects (2). In follow-up fine mapping, we have studied 23 SNPs that tag the common variation in this haplotype block in the SEARCH study. None of these SNPs showed a stronger association with breast cancer than that shown by the original tag SNP rs13281615 (data not shown). This SNP (2.1, Table 1) was not associated with prostate, colorectal, or ovarian cancer. To date, the only published study that tested the association of these SNPs with risks of other cancers (prostate and colorectal) found no evidence of an association (6). Taken together, these data suggest that region 2 is specific for breast cancer susceptibility.

Region 3, spanning base positions 128.47–128.54 Mb, was originally detected in African Americans by an admixture scan (a method for localizing disease-causing genetic variants that differ in frequency across populations) for prostate cancer (Table 1, rs6983267, 3.3; rs7000448, 4.1) (5). Subsequently, two genome-wide scans found that SNP 3.3 and rs10505477 (3.1) (8,10) were associated with colorectal cancer, and these associations have been consistently replicated in independent case–control studies (6,8,10,11). Another SNP in the same block, rs10808556 (3.2), has also been associated with colorectal cancer (6). We found that SNPs 3.1, 3.2, and 3.3 were all associated with risks of prostate (OR = 1.43, 95% CI = 1.30 to 1.56, P-value = 7.7 × 10^-9), colorectal (OR = 1.27, 95% CI = 1.16 to 1.37, P-value = 3.6 × 10^-5), and ovarian cancers (OR = 1.11, 95% CI = 1.03 to 1.23, P-value = 9.9 × 10^-4) (ORs, 95% CIs, and P values are given for SNP 3.2). This is the strongest evidence, to date, reporting an association between ovarian cancer risk and a common allele. The three SNPs in this block are highly correlated with each other in control subjects (r^2 values >0.65, Figure 1, B). Using stepwise logistic regression, the associations for each disease could be explained by a single SNP (data not shown). We found no evidence that one of these SNPs was more strongly associated with risk of prostate and colon cancer than the other two. It is therefore likely that there is common underlying factor that increases the risk of the three cancers. None of the SNPs in this region were associated with breast cancer risk. Our data suggest that the prostate, colorectal, and ovarian cancer locus is smaller than the one originally defined (5) and only spans base positions 128.47–128.50 Mb. Therefore, we have designated the remaining portion of the original locus, spanning positions 128.50–128.54 Mb, as region 4.

Region 4 (prostate cancer) contains SNP rs7000448 (4.1), which has been shown to be associated with prostate cancer (5). This SNP is only weakly correlated with the region 3 and region 5 SNPs (r^2 < 0.13, Figure 1, B). Furthermore, we confirmed an association of this variant with prostate cancer risk (OR = 1.23, 95% CI = 1.11 to 1.35, P-value = 2.8 × 10^-4) but found no association with risks of colorectal, ovarian, or breast cancers, suggesting that this is a separate prostate cancer–specific locus.

Region 5 is the closest of the five regions to the c-MYC oncogene and spans base positions 128.54–128.62 Mb. SNP rs1447295 (5.1, Table 1) was originally found to be associated with prostate cancer through linkage and association analyses in the Icelandic population (1). This association has subsequently been replicated in other populations (3,7,9,12,13). A second SNP, rs10090154, which is perfectly correlated with rs1447295 (5.1) in Europeans (r^2 = 1 in CEU HapMap) but not in Africans (r^2 = 0.64), was subsequently identified (5). A weak association of rs10090154 with colorectal cancer was reported as provisional, pending independent confirmation (6). We found SNP 5.1 (rs1447295) to be statistically significantly associated with prostate cancer (OR = 1.86, 95% CI = 1.60 to 2.15, P value = 6.9 × 10^-7) but not with breast, colorectal, or ovarian cancer. A large study, nested in seven US and European cohorts, has also noted the

**CONTEXT AND CAVEATS**

Prior knowledge
Genetic variants in a region of chromosome 8 had been associated with the risk of breast, colorectal, and prostate cancer.

Study design
Case subjects with each of four cancers (breast, colorectal, prostate, and ovarian) and control subjects were examined for the presence of previously identified risk variants that span the chromosomal region previously associated with cancer risk. Genotype frequencies were compared using unconditional logistic regression.

Contribution
At least five distinct cancer susceptibility loci were found within the chromosomal region, each separated by recombination hot spots and specific for one or more of the four cancers.

Implications
Fine mapping of the identified loci may help elucidate molecular mechanisms that contribute to carcinogenesis.

Limitations
It is unknown whether any of the cancer-associated polymorphisms examined are causal variants or simply markers of unknown causal variants.
Table 1. Association of 8q24 single nucleotide polymorphisms with colorectal, ovarian, breast, and prostate cancer*

<table>
<thead>
<tr>
<th>Marker SNP (region, relative position)</th>
<th>Reference allele (frequency in controls subjects)</th>
<th>Colorectal cancer OR (95% CI)</th>
<th>P value‡</th>
<th>Ovarian cancer OR (95% CI)</th>
<th>P value‡</th>
<th>Breast cancer OR (95% CI)</th>
<th>P value§</th>
<th>Prostate cancer OR (95% CI)</th>
<th>P value§</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs13254738 (A/C) (1.1) (region 1, 128173525)</td>
<td>A (0.70)</td>
<td>1.06 (0.99 to 1.13)</td>
<td>0.22</td>
<td>1.02 (0.94 to 1.11)</td>
<td>0.64</td>
<td>0.96 (0.88 to 1.05)</td>
<td>0.35</td>
<td>1.12 (1.01 to 1.24)</td>
<td>0.029</td>
</tr>
<tr>
<td>rs6983561 (A/C) (1.2) (region 1, 128176062)</td>
<td>A (0.97)</td>
<td>0.95 (0.81 to 1.11)</td>
<td>0.65</td>
<td>0.90 (0.72 to 1.13)</td>
<td>0.36</td>
<td>0.96 (0.77 to 1.21)</td>
<td>0.76</td>
<td>2.11 (1.65 to 2.71)</td>
<td>1.4 × 10⁻³</td>
</tr>
<tr>
<td>rs16901979 (G/T) (1.3) (region 1, 128194098)</td>
<td>G (0.97)</td>
<td>0.89 (0.77 to 1.06)</td>
<td>0.36</td>
<td>0.89 (0.71 to 1.11)</td>
<td>0.30</td>
<td>0.98 (0.80 to 1.25)</td>
<td>0.98</td>
<td>2.06 (1.61 to 2.65)</td>
<td>4.9 × 10⁻³</td>
</tr>
<tr>
<td>rs13281615 (A/G) (2.1) (region 2, 128424800)§</td>
<td>A (0.60)</td>
<td>0.94 (0.89 to 1.00)</td>
<td>0.17</td>
<td>0.99 (0.91 to 1.07)</td>
<td>0.75</td>
<td>1.21 (1.11 to 1.32)</td>
<td>1 × 10⁻⁵</td>
<td>0.95 (0.87 to 1.05)</td>
<td>0.33</td>
</tr>
<tr>
<td>rs10505477 (G/A) (3.1) (region 3, 128476622)</td>
<td>G (0.50)</td>
<td>1.27 (1.19 to 1.33)</td>
<td>2.9 × 10⁻⁸</td>
<td>1.14 (1.04 to 1.23)</td>
<td>2.0 × 10⁻³</td>
<td>0.96 (0.88 to 1.04)</td>
<td>0.35</td>
<td>1.43 (1.30 to 1.56)</td>
<td>7.7 × 10⁻¹⁰</td>
</tr>
<tr>
<td>rs10809556 (A/G) (3.2) (region 3, 128482329)</td>
<td>A (0.59)</td>
<td>1.26 (1.16 to 1.37)</td>
<td>5.1 × 10⁻⁸</td>
<td>1.13 (1.04 to 1.22)</td>
<td>1.7 × 10⁻³</td>
<td>0.99 (0.91 to 1.08)</td>
<td>0.80</td>
<td>1.31 (1.19 to 1.44)</td>
<td>4.2 × 10⁻³</td>
</tr>
<tr>
<td>rs69832676 (A/G) (3.3) (region 3, 128482487)§</td>
<td>A (0.49)</td>
<td>1.27 (1.16 to 1.37)</td>
<td>3.6 × 10⁻⁸</td>
<td>1.11 (1.03 to 1.20)</td>
<td>9.9 × 10⁻³</td>
<td>0.97 (0.89 to 1.05)</td>
<td>0.50</td>
<td>1.43 (1.30 to 1.56)</td>
<td>7.7 × 10⁻¹⁰</td>
</tr>
<tr>
<td>rs7000448 (G/A) (4.1) (region 4, 128510352)</td>
<td>(0.98)</td>
<td>1.04 (0.96 to 1.13)</td>
<td>0.32</td>
<td>1.04 (0.96 to 1.13)</td>
<td>0.33</td>
<td>0.96 (0.88 to 1.05)</td>
<td>0.38</td>
<td>1.23 (1.11 to 1.35)</td>
<td>2.8 × 10⁻⁵</td>
</tr>
<tr>
<td>rs1447295 (G/T) (5.1) (region 5, 128554220)</td>
<td>A (0.90)</td>
<td>0.98 (0.89 to 1.08)</td>
<td>0.82</td>
<td>1.07 (0.93 to 1.22)</td>
<td>0.35</td>
<td>0.92 (0.80 to 1.07)</td>
<td>0.28</td>
<td>1.86 (1.60 to 2.15)</td>
<td>6.9 × 10⁻¹⁷</td>
</tr>
</tbody>
</table>

* Genotype results were obtained for more than 95% of all subjects. rs1000154 was not evaluated because it was perfectly correlated with rs1447295 in our European population sample. All genotyping was performed by Taqman assay unless otherwise indicated. No deviation from Hardy–Weinberg equilibrium was observed in the genotype distributions of the control subjects for any of the SNPs. OR = odds ratio; CI = confidence interval; SNP = single nucleotide polymorphism. The bold font refers to significant P values (P < .05) and their corresponding OR.
† Sample sets consisted of 2299 colorectal cancer case subjects and 2284 control subjects, 1975 ovarian cancer case subjects and 3411 control subjects, 2270 breast cancer case subjects and 2280 control subjects, 1854 prostate cancer case subjects and 1894 control subjects.
‡ P value from the Cochran–Armitage trend test.
§ Genotyped in the prostate study using the illumina 550K chip covering approximately 550,000 SNPs across the genome. Hence, the SNPs were replaced by alternative tags on the illumina chip: rs13281615 by rs672888 (r² = 0.97) and rs10505477 by rs6983267 (r² = 0.93).

absence of association of this SNP with breast cancer susceptibility (7).
To date, three risk-associated regions at 8q24 (regions 1, 3, and 5) have been reported to confer independent risks of prostate cancer. In this study, we found a total of eight SNPs, distributed across four regions, to be associated with the risk of prostate cancer. To test how many of these associations were independent, we performed a stepwise logistic regression that included all eight SNPs in the model. Five SNPs (two in region 1 and one in each of regions 3, 4, and 5) were independently associated with prostate cancer (rs13254738, P = .008; rs6983561, P = 1.6 × 10⁻1⁰; rs6983267, P = 1.6 × 10⁻2; rs7000448, P = .022; rs1447295, P = 2.0 × 10⁻¹³). Theoretically, each of these independent SNPs may be markers for a separate causative factor in prostate cancer development.
Thus, we have shown there are at least five independent loci within this gene desert with different associations with particular cancers. Further studies of the region may identify additional loci associated with specific cancers and possibly refine our understanding of the mechanisms underlying the associations reported here. A recent publication has reported that none of the above SNPs were associated with risk of endometrial cancer (14).
The biologic mechanisms underlying these associations with different cancers are unknown. This region is a frequent site of somatic amplification in several cancers (15,16). It is possible that these variants affect tissue-specific enhancers in the region, thus altering expression of one or more genes an unknown distance away. The known genes that are closest to 8q24 are FAM84B and c-MYC. Overexpression of c-MYC occurs in both breast and prostate cancers (17–19), and reduction of c-MYC expression by RNA interference inhibits tumor growth both in vivo and in vitro (20). FAM84B is described as a breast cancer membrane-associated protein, but little more is known about its function (18). However, SNPs located in the c-MYC and FAM84B genes were not found to be associated with prostate cancer (1,4). Furthermore, SNPs in regions 1, 3, and 5
found to be associated with prostate cancer do not appear to be associated with changes in expression of these genes in prostate or colorectal tumors (1,4,10). Several other genes were predicted to exist in 8q24 (1,10), although there is no evidence for any protein-coding transcripts (1,10). One is a putative pseudogene of the transcription factor POU5F1P1 in region 3. One study has confirmed the expression of this transcript in cancer tissues, including colon cancer, although its physiological role is unknown (8).

Despite their strong associations with cancer, it is not known whether the SNPs tested here are causal variants or are simply markers that are correlated with the causal variants in each region. Resequencing and fine mapping of each of the haplotype blocks, followed by functional characterization studies, may ultimately identify the causal variants and reveal their mechanisms in cancer susceptibility and pathogenesis. If this 8q24 locus is truly a gene desert, it points to a very long-range mode of action for these variants that had previously been considered unlikely.

References


**Funding**

This work was supported by Cancer Research UK. The ProtecT study which provided control subjects for the prostate analyses is funded by the Health Technology Assessment Programme (projects 96/20/06, 96/20/99). We also like to thank the following for funding support: The Institute of Cancer Research and The Everyman Campaign, The Prostate Cancer Research Foundation, Prostate Research Campaign UK, The National Cancer Research Network UK, and The National Cancer Research Institute UK and grants from the National Health and Medical Research Council, Australia (209057, 251533, 450104); VicHealth; The Cancer Council Victoria; The Whitten Foundation; Tattersall’s; The Roswell Park Alliance; The Danish Cancer Society; National Cancer Institute (CA1766 and Core Grant CA16056 and RO1 CA61107); and Fondation Dr Dubois-Ferrière Dinu Lipatti.

**Notes**

H. Song, T. Koessler, and A. A. A. Olama contributed equally to the work.

List of members of The UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons’ Section of Oncology is available on request.

UK ProtecT Study Collaborators: Prasad Bollina, Sue Bonnington, Debbie Cooper, Andrew Doble, Alan Doherty, Garett Durkan, Emma Elliott, David Gillatt, Pippa Herbert, Peter Holding, Joanne Howson, Mandy Jones, Roger Kockelbergh, Howard Kynaston, Teresa Lennon, Norma Lyons, Hing Leung, Hilary Moody, Philip Powell, Stephen Prescott, Pauline Thompson—Care of Surgical Oncology (Uro-Oncology:54), University of Cambridge, Box 279, Addenbrooke’s Hospital, Hills Road, Cambridge, UK.

We would like to thank all the patients and control subjects who took part in this study. We would also like to thank Hannah Munday, Barbara Perkins, Helen Imogen Field, Mitul Shah, Clare Jordan, Judy West, Anabel Simpson, Sue Irvine, the search team: the local general practices and nurses and the East Anglian Cancer Registry for recruitment of the UK case subjects and the EPIC-Norfolk investigators for recruitment of the UK control subjects; Claus K. Hogdal and Jan Blaakcer for their additional contribution to the MALignant OVArian cancer collection; Aleksandra Gentry-Maharaj, Eva Wozniak, Usha Menon, and the UK Ovarian-cancer Population Study (UKOPS) team of research nurses for their contribution to the UKOPS ovarian cancer collection (funded by the OAK foundation). D. F. Easton is a principal research fellow of Cancer Research UK, P. D. P. Pharoah is Cancer Research UK senior clinical research fellow, and B. A. J. Ponder is a Gibb fellow of CRUK. J. L. Hopper is a principal investigator of the UK Prostate Cancer Research Foundation, The Prostate Cancer Research Foundation, The National Cancer Research Network UK, and the EPIC-Norfolk investigators for recruitment of the UK control subjects; Claus K. Hogdal and Jan Blaakcer for their additional contribution to the MALignant OVArian cancer collection; Aleksandra Gentry-Maharaj, Eva Wozniak, Usha Menon, and the UK Ovarian-cancer Population Study (UKOPS) team of research nurses for their contribution to the UKOPS ovarian cancer collection (funded by the OAK foundation). D. F. Easton is a principal research fellow of Cancer Research UK, P. D. P. Pharoah is Cancer Research UK senior clinical research fellow, and B. A. J. Ponder is a Gibb fellow of CRUK. J. L. Hopper is an Australia fellow of the National Health and Medical Research Council. The authors had full responsibility for the analysis and interpretation of the data and for the writing and submission of the manuscript.

Manuscript received November 27, 2007; revised April 20, 2008; accepted May 14, 2008.