Evaluation of the 8q24 Prostate Cancer Risk Locus and MYC Expression

Mark M. Pomerantz,1 Christine A. Beckwith,1 Meredith M. Regan,2 Stacia K. Wyman,4 Gyorgy Petrovics,5 Yongmei Chen,6 Dorota J. Hawksworth,9 Fredrick R. Schumacher,5,6 Lorelei Mucci,7 Kathryn L. Penney,4 Meir J. Stampfer,9 Jennifer A. Chan,10 Kristin G. Ardlie,11 Brian R. Fritz,1 Rachael K. Parkin,4 Daniel W. Lin,5,7,8 Michelle Dyke,9 Paula Herman,1 Steve Lee,1 William K. Oh,1 Philip W. Kantoff,1 Muneesh Tewari,11 David G. McLeod,3 Shiv Srivastava,3 and Matthew L. Freedman1

Departments of Medical Oncology and Biostatistics and Computational Biology, Dana-Farber Cancer Institute; and Department of Epidemiology, Harvard School of Public Health, Boston, Massachusetts; Division of Human Biology, Public Health Sciences Division, and Division of Clinical Research, Fred Hutchinson Cancer Research Center; Department of Urology, University of Washington; and Department of Veterans Affairs Puget Sound Health Care System, Seattle, Washington; Center for Prostate Disease Research, Department of Surgery, Uniformed Services University of the Health Sciences, Rockville, Maryland; Departments of Pathology and Laboratory Medicine, Clinical Neurosciences, and Oncology, University of Calgary, Calgary, Alberta, Canada; and Broad Institute of Harvard and MIT, Cambridge, Massachusetts

Abstract
Polymorphisms at 8q24 are robustly associated with prostate cancer risk. The risk variants are located in nonprotein coding regions and their mechanism has not been fully elucidated. To further dissect the function of this locus, we tested two hypotheses: (a) unannotated microRNAs (miRNA) are transcribed in the region, and (b) this region is a cis-acting enhancer. Using next generation sequencing, 8q24 risk regions were interrogated for known and novel miRNAs in histologically normal radical prostatectomy tissue. We also evaluated the association between the risk variants and transcript levels of multiple genes, focusing on the proto-oncogene, MYC. RNA expression was measured in histologically normal and tumor tissue from 280 prostatectomy specimens (from 234 European American and 46 African American patients), and paired germline DNA from each individual was genotyped for six 8q24 risk single nucleotide polymorphisms. No evidence was found for significant miRNA transcription within 8q24 prostate cancer risk loci. Likewise, no convincing association between RNA expression and risk allele status was detected in either histologically normal or tumor tissue. To our knowledge, this is one of the first and largest studies to directly assess miRNA in this region and to systematically measure MYC expression levels in prostate tissue in relation to inherited risk variants. These data will help to direct the future study of this risk locus. [Cancer Res 2009;69(13):5568–74]

Introduction
Until recently, the genetic etiology of prostate cancer remained largely unknown. Difficulty in validating loci discovered by linkage mapping led to the hypothesis that the genetic basis of prostate cancer was largely due to the actions of many loci, each of modest effect. The effort to identify and to catalogue common human genetic variation resulted in dense genetic maps and facilitated the mapping of lower penetrant risk loci through genome-wide association studies (1). In 2006, a prostate cancer locus on chromosome 8q24 was among the first genetic risk loci to be reproducibly validated. Initially described by two groups using different methodologies, the region confers an elevated risk across multiple ethnic groups (2, 3), and multiple studies involving thousands of cases and controls have confirmed the finding (4–7). Since the original discovery in prostate cancer, risk for other cancers, such as those of the colon, breast, and bladder, have also been mapped to 8q24 (8–11).

Fine mapping of the locus has revealed a set of six single nucleotide polymorphisms (SNP) and a microsatellite variant that independently contribute to prostate cancer risk (12). These alleles cluster in three regions (defined by linkage disequilibrium) spanning a distance of ~490 kilobases (Supplementary Fig. S1). Intriguingly, the risk variants are located in a nonprotein coding region of the genome and the mechanism by which they contribute to disease is unknown. Possible mechanisms of action include the presence of unannotated transcripts whose expression is influenced by inherited variation, and/or regulation of genes beyond the risk regions.

The annotated protein coding region closest to any risk allele is the well-known oncogene, MYC, >250 kilobases from the nearest prostate cancer risk SNP. Attempts have been made to detect an association between MYC expression and genotype at cancer risk SNPs. To date, no study has definitively found an association, but each has been limited by sample size or tissue type. For example, some have measured MYC RNA expression in lymphoblastoid cell lines rather than the tissue from which cancer formed. Expression of many genes differs markedly across tissues (13). Therefore, when evaluating the influence of a risk SNP for a particular disease such as prostate cancer, it is important to study the gene in a tissue-specific context.

In the present study, two hypotheses concerning the mechanism of inherited prostate cancer risk are evaluated: first, that 8q24 prostate cancer risk regions harbor previously unannotated microRNA (miRNA) species and, second, that elements in the risk regions participate in the regulation of distal genes, affecting RNA expression. We searched for miRNA transcripts within risk loci in prostate tissue using a next generation sequencing approach.

Note: Supplementary data for this article are available at Cancer Research Online (http://cancerres.aacrjournals.org/).

Requests for reprints: Matthew L. Freedman, Dana-Farber Cancer Institute, Dana Bldg Rm 710A, 44 Binney Street, Boston, MA 02115. Phone: 617-582-8598; Fax: 617-582-8157; E-mail: matthew_freedman@dfci.harvard.edu.

Published OnlineFirst June 23, 2009; DOI: 10.1158/0008-5472.CAN-09-0387

Cancer Res 2009; 69: (13). July 1, 2009 5568 www.aacrjournals.org

Downloaded from cancerres.aacrjournals.org on April 15, 2017. © 2009 American Association for Cancer Research.
Three small RNA cDNA libraries were generated using 50, 10, and 5 manufacturers recommended protocol and RNA samples were pooled. Total RNA was extracted with the Ambion miRVana miRNA isolation kit, using the mirVana Lysis/Binding buffer using a Qiagen Tissuelyzer. Approximately 100 mg of fresh-frozen, histologically normal prostate tissue from each of two individuals were homogenized in Ambion protocols. Approximately 100 mg of fresh-frozen, histologically normal prostate tissue was obtained from individuals undergoing prostatectomy at one of three institutions: Dana-Farber/Partners Cancer Institute (DFCI), Brigham and Women’s Hospital, and the Center for Prostate Disease Research (CPDR), a collection of databases derived from nine American and 26 African American) via LCM at the Center for Prostate Disease Research (CPDR), a collection of databases derived from nine military hospitals (15). Procedures for tissue processing, LCM, and RNA preparation were performed as described previously (16–18). RNA was isolated from 62 European American and 26 African American) via LCM at the Center for Prostate Disease Research (CPDR), a collection of databases derived from nine military hospitals (15). Procedures for tissue processing, LCM, and RNA preparation were performed as described previously (16–18). RNA was isolated from 62 European normal tissue and tumor from normal tissue was isolated in 88 RP samples (62 European American and 26 African American) via LCM at the Center for Prostate Disease Research (CPDR), a collection of databases derived from nine military hospitals (15). Procedures for tissue processing, LCM, and RNA preparation were performed as described previously (16–18). RNA was isolated from 67 RP tumor samples from subjects in the Physicians’ Health Study (PHS), initiated in 1982 and comprising 22,071 U.S. male physicians, ages 40 to 84 y (19). RNA was reverse transcribed and each sample was subjected to one of three methods of gene expression analysis—competitive reverse transcription-PCR (RT-PCR), quantitative real-time-PCR, and the cDNA-mediated Ampealing, Selection, Extension, and Ligation (DASL) expression assay. In competitive RT-PCR, used for DFCI samples, seven transcripts of interest were chosen based on proximity to risk polymorphisms and previously reported oncogenic activity. Seven normalization genes were chosen based on known expression in prostate tissue. All 14 assays and their competitive oligos were plexed into a single reaction mix, using the Sequenom iPLEX mass spectrometry platform. Sequences are available upon request. Reactions were performed in quadruplicate using eight serial dilutions of competitor, ranging from 10⁻¹⁸ to 10⁻¹² M. Thus, a total of 32 reactions were performed for each individual cDNA species. Resulting spectra were analyzed and the EC₅₀—the point at which cDNA and competitor concentrations are equal—was calculated using QGE Analyzer software (Sequenom). A gene expression normalization factor was calculated for each sample using the geNorm algorithm (20).

### Table 1. Summary of study populations and genotyping results

<table>
<thead>
<tr>
<th>Population</th>
<th>Source</th>
<th>Method</th>
<th>No.</th>
<th>Risk SNPs: no. homozygote no risk (%)</th>
<th>no. heterozygote (%)</th>
<th>no. homozygote risk (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EA normal tissue</td>
<td>DFCI Comp PCR</td>
<td>105</td>
<td>61 (58)</td>
<td>43 (41)</td>
<td>1 (1)</td>
<td>46 (44)</td>
</tr>
<tr>
<td>CPDR LCM qRT-PCR</td>
<td>62</td>
<td>48 (77)</td>
<td>11 (18)</td>
<td>0 (0)</td>
<td>21 (34)</td>
<td>32 (52)</td>
</tr>
<tr>
<td>AA normal tissue</td>
<td>DFCI/CPDR Comp PCR</td>
<td>20</td>
<td>5 (25)</td>
<td>14 (70)</td>
<td>1 (5)</td>
<td>3 (15)</td>
</tr>
<tr>
<td>CPDR LCM qRT-PCR</td>
<td>62</td>
<td>14 (54)</td>
<td>9 (35)</td>
<td>3 (12)</td>
<td>4 (15)</td>
<td>10 (38)</td>
</tr>
<tr>
<td>EA tumor tissue</td>
<td>DFCI Comp PCR</td>
<td>33</td>
<td>22 (67)</td>
<td>11 (33)</td>
<td>0 (0)</td>
<td>12 (36)</td>
</tr>
<tr>
<td>CPDR LCM qRT-PCR</td>
<td>62</td>
<td>48 (77)</td>
<td>11 (18)</td>
<td>0 (0)</td>
<td>21 (34)</td>
<td>32 (52)</td>
</tr>
<tr>
<td>AA tumor tissue</td>
<td>PHS DASL qRT-PCR</td>
<td>67</td>
<td>55 (82)</td>
<td>12 (18)</td>
<td>0 (0)</td>
<td>25 (37)</td>
</tr>
<tr>
<td>CPDR LCM qRT-PCR</td>
<td>26</td>
<td>14 (54)</td>
<td>9 (35)</td>
<td>3 (12)</td>
<td>4 (15)</td>
<td>10 (38)</td>
</tr>
</tbody>
</table>

NOTE: Study subjects were first categorized by ancestry and normal prostate vs prostate tumor tissue in RP specimens. Subjects were further characterized by institution, method of tissue isolation (LCM vs macrodissection), and method of quantifying gene expression. Genotyping results for six risk SNPs are presented in the columns to the right. The Broad1 1994905 SNP was not tested in all subjects. For each SNP, listed are the number of men carrying no risk alleles, the number carrying one copy, and the number carrying two copies of the risk allele.

Abbreviations: qRT-PCR, quantitative real-time PCR; HMZ, homozygote; HET heterozygote.

Association between 8q24 risk allele status and MYC mRNA expression levels was then measured in 401 normal and tumor prostate tissue samples derived from 280 European and African American prostate cancer patients. A subset of 176 samples had histologically normal and tumor isolated via laser capture microdissection (LCM) for RNA extraction. Lastly, in a subset of 158 samples, associations between risk allele carrier status and mRNA expression of six other annotated genes in the 8q24 region were evaluated.

### Materials and Methods

#### Sequencing for detection of miRNA species.
Histologically normal segments of prostate tissue were obtained from individuals undergoing retropubic radical prostatectomy (RP) for prostate cancer. Specimens were collected under informed consent with institutional review board approved protocols. Approximately 100 mg of fresh-frozen, histologically normal prostate tissue from each of two individuals were homogenized in Ambion mirVana Lysis/Binding buffer using a Qiagen TissueLyzer. Total RNA was extracted with the Ambion mirVana miRNA isolation kit, using the manufacturers recommended protocol and RNA samples were pooled. Three small RNA cDNA libraries were generated using 50, 10, and 5 μg of pooled RNA, respectively, using a protocol described previously (14) with modified linker and primer sequences to adapt for sequencing on an Illumina Genome Analyzer. The 3’ cloning linker 5’(SfAppCTGTAGG-CACCATCAAT/3’dC/3’) was purchased from IDT and the 5’ Illumina linker was synthesized by IDT as a custom oligo. The fully ligated libraries were reverse transcribed and PCR amplified with primer sequence necessary for cloning linker 1(5’TACCGACCGACCATCAAT/3’dC/3’).

Illumina primary analysis software for image analysis and base calling. Illuminalinker

Deep sequencing data sets, we trimmed the final three nucleotides from each sequence [National Center for Biotechnology Information (NCBI) build 36.1]. As 3’ sequence was trimmed from the 3’ end, following which the reads were collapsed to generate nonredundant sequences. These were then mapped to the reference human genome sequence; no additional reads mapping to the 8q24 region were identified by this operation. Unique sequences mapping to the segment chr8:128,100,000-128,700,000 were selected for examination in this study. As described in Results, all sequences in this segment were present at three reads or fewer in the data set, with the exception of one sequence present at five reads.

#### Gene expression analysis in three RP patient populations.
Fresh frozen RP specimens from 108 subjects at the Dana-Farber Cancer Institute (DFCI) and Brigham and Women’s Hospital were reviewed by a pathologist (J.C.) to isolate areas prostatic adenocarcinoma and benign tissue. Areas of tumor were selected where >60% of cells consisted of tumor cells. Areas of benign tissue were selected where >50% of cells consisted of nonneoplastic epithelium and were at least 5 mm away from any area of tumor focus. Two millimeter punch biopsy cores of frozen tissue were processed RNA extraction using a modified Qiagen Allprep DNA/RNA protocol. RNA from tumor and from normal tissue was isolatetion in 88 RP samples (62 European American and 26 African American) via LCM at the Center for Prostate Disease Research (CPDR), a collection of databases derived from nine military hospitals (15). Procedures for tissue processing, LCM, and RNA preparation were performed as described previously (16–18). RNA was isolated from 67 RP tumor specimens from subjects in the Physicians’ Health Study (PHS), initiated in 1982 and comprising 22,071 U.S. male physicians, ages 40 to 84 y (19). RNA was reverse transcribed and each sample was subjected to one of three methods of gene expression analysis—competitive reverse transcription-PCR (RT-PCR), quantitative real-time-PCR, and the cDNA-mediated Ampealing, Selection, Extension, and Ligation (DASL) expression assay. In competitive RT-PCR, used for DFCI samples, seven transcripts of interest were chosen based on proximity to risk polymorphisms and previously reported oncogenic activity. Seven normalization genes were chosen based on known expression in prostate tissue. All 14 assays and their competitive oligos were plexed into a single reaction mix, using the Sequenom iPLEX mass spectrometry platform. Sequences are available upon request. Reactions were performed in quadruplicate using eight serial dilutions of competitor, ranging from 10⁻¹⁸ to 10⁻¹² M. Thus, a total of 32 reactions were performed for each individual cDNA species. Resulting spectra were analyzed and the EC₅₀—the point at which cDNA and competitor concentrations are equal—was calculated using QGE Analyzer software (Sequenom). A gene expression normalization factor was calculated for each sample using the geNorm algorithm (20).

CPDR samples were analyzed via TaqMan-based qRT-PCR on ABI 7700 (Applied Biosystems), as previously described (16, 18). MYC gene expression in each sample was normalized to glyceraldehyde-3-phosphate dehydrogenase (GAPDH) expression.
dehydrogenase expression levels and results were plotted as average threshold cycle values of duplicate samples.

Analyses across DFCI and CPDR cohorts were performed separately by ancestry. To compare samples based on total number of risk alleles, each cohort was split into five risk groups. The European Americans were grouped as 0 to 1, 2, 3, 4, and 5 to 6 risk allele carriers. The African Americans were grouped as carrying 2 to 6, 7 to 8, and 9 to 10 risk alleles. Expression levels were also conditioned on number of risk SNPs carried within a given risk region in a separate analysis. Transcript levels for each message were regressed on risk allele carrier status. The analysis used Kruskal-Wallis tests for a global test of difference among compared samples; P values were adjusted for multiple comparisons using permutation testing. To compare samples based on genotype status at an individual polymorphism, each individual was classified as carrying 0, 1, or 2 risk alleles (reflecting homozygote wild-type, heterozygote, and homozygote risk genotypes). For SNPs at low frequency in the population, only two groups were feasible—0 versus 1 or 2 alleles.

PHS samples were analyzed via the DASL expression assay (Illumina, Inc.), as described previously (19). Molecular data generated using the DASL approach were compared with prostate cancer expression array data generated using frozen tissue samples on conventional microarray platforms to verify results (19). Samples were compared based on number of risk alleles at an individual polymorphism (0, 1, or 2 risk alleles or 0 versus 1 to 2 for low frequency risk alleles). ANOVA analysis (F test) was used to detect associations between MYC expression and risk allele status.

Genomic DNA was prepared from peripheral blood using QIamp DNA Blood mini kit (QIAGEN, Inc.). Subjects were genotyped for six SNPs identified in a fine mapping analysis of 8q24-associated prostate cancer risk and eight SNPs in high or complete linkage disequilibrium with the risk SNPs. Genotyping was carried out using Sequenom iPLEX mass spectrometry platform. The error rate on this platform is estimated to be <0.03%.

Results

Assessment of miRNA transcripts in the 8q24 prostate risk regions. miRNAs are an important layer of regulatory control and therefore we investigated the possibility that the 8q24 locus risk allele may be related to a known or novel miRNA encoded in the region.

To identify potential novel miRNAs expressed in the prostate that may have not yet been discovered, we characterized known and novel miRNAs in histologically normal prostate tissue using a next generation sequencing approach. We pooled RNA from histologically normal prostate tissue samples from two individuals collected at the time of RP. We isolated small RNA (18–24 nt) in triplicate by PAGE and generated three cDNA libraries that we subjected to ultrahigh-throughput sequencing using an Illumina Genome Analyzer. A total of 17,552,100 reads were generated.

We first analyzed seven computationally defined miRNAs that had been reported previously as 8q24 miRNAs and confirmed by RT-PCR to determine whether they were present in our data set (21). None of the seven miRNAs perfectly matched any of our reads. Given that it is difficult to computationally identify the 5′ and 3′ ends of miRNAs, we trimmed the sequence of these seven miRNAs by 1, 2, and then 3 nucleotides from the 5′ and 3′ ends and searched for a match. Even by that analysis, none of the computationally identified miRNAs were found in our expression data.

We proceeded to search for potential novel miRNAs in the 8q24 region that might be relevant to the risk allele. We processed our sequence reads by trimming linker sequences, combining redundant reads and mapping to the reference human genome sequence 8q24 region at chr8:128,100,000-128,700,000 (NCBI build 36.1; Supplementary Fig. S1). Removing sequences that mapped to 10 or more loci (i.e., those that correspond to repetitive sequence and are difficult to ascribe to a particular locus), gave us 459 unique sequences with 487 reads mapping to 1,283 loci in the specific region of interest. All the matching sequences were present at extremely low abundance. For example, 443 of the 471 were singleton reads, 24 had 2 reads, 3 had 3 reads, and 1 had five reads. Minimal criteria for annotation of novel miRNAs generally require multiple reads and origin from a hairpin precursor. We examined each of the nonsingleton sequences for origin from a putative hairpin precursor by folding each read with 100 nt of up- and down-stream genomic sequence at a given locus [as described previously in Bar and colleagues (14)]. None of the nonsingleton reads met criteria for origin from a hairpin precursor.

Taken together, our analyses did not find evidence for a significant miRNA transcript in the 8q24 risk region studied.

Genotype and MYC expression in three prostate cancer cohorts. We sought to identify a gene or set of genes whose expression is influenced by risk variants. We focused on the proto-oncogene MYC, testing the hypothesis that MYC transcript abundance is associated with risk allele status. A total of 280 individuals were evaluated from three independent prostate cancer populations (Table 1).

Given the range of plausible biological models and the genetic complexity of the 8q24 locus, MYC expression was analyzed across a range of scenarios (Table 1). First, histologically normal prostate tissue and prostate tumor tissue were each examined because risk alleles may exert their effect more profoundly in a particular tissue type. To minimize confounding from tissue cell admixture, a subset of RP samples whose normal and tumor epithelial cells were isolated by LCM were analyzed separately (n = 88). Furthermore,
European American and African American prostate tissues were distinguished from one another because risk allele frequencies differ significantly across the two populations, and one of the risk alleles is expressed only in men of African ancestry.

To evaluate risk allele status, each subject was genotyped for six prostate cancer risk SNPs at chromosome 8q24 (Table 1). The SNPs chosen for analysis independently influence risk and reside across three distinct linkage disequilibrium regions (Supplementary Fig. S1). Given 6 risk SNPs, an individual can carry a maximum of 12 risk alleles (2 alleles per locus x 6 loci). Because the Broad1193405 risk allele is present only in individuals of African ancestry, European Americans can carry a maximum of 10 risk alleles. Across all populations genotyped, European American subjects (n = 234) carried a total of 0 to 6 alleles, and the total number of risk alleles carried by African American subjects (n = 46) ranged from 2 to 10.

Three models for the relationship between inherited variation and gene expression were tested because the mechanism by which 8q24 variants confer risk is unknown. One hypothesis is that the risk alleles act collectively to influence MYC expression. In this model, European American and African American subjects were categorized into subgroups based on the number (of a possible total of 12) of risk alleles carried (see Materials and Methods). No significant association exists between MYC expression and total risk allele status in normal prostate tissue or tumor tissue (P > 0.05). Another hypothesis is that each of the three 8q24 linkage...
disequilibrium risk regions [as described by Haiman and colleagues (ref. 12)] behaves as a unit that influences a target gene. In this model, risk alleles within a given risk region are summed and analyzed for association with MYC expression. Under this model, no statistically significant associations (P > 0.05) were found between steady-state MYC mRNA expression and genotype within any risk region. This included normal and tumor tissue in both ancestral populations.

A third model is that each risk SNP acts independently. Based on the SNP allele frequencies, subjects were categorized as carrying zero, one, or two risk alleles, or as carriers or noncarriers of the risk allele. (Table 1; Fig. 1) Only one assay showed a nominally statistically significant association between risk status at an individual SNP and MYC expression: rs13254738 in laser capture microdissected European American tumor tissue (P = 0.02). The statistical significance was driven by increased expression among the seven individuals homozygous for the rs13254738 risk allele. No difference in expression was identified between homozygotes for the wild-type allele and heterozygotes. The association was not observed in two other European American tumor tissue sample sets or in normal prostate tissue from the same population.

**Association between risk alleles and six candidate 8q24 transcripts.** To detect associations between risk allele status and expression at other candidate genes, six transcripts of interest in the 8q24 region were selected for analysis (Supplementary Fig. S1). Annotated genes located within 1 Mb for the nearest risk region, such as FAM84B and PVT1, were included. PVT1 seems to encompass a wide area of transcription and exists in multiple splice forms. Three annotated transcripts in the region were selected to survey PVT1 expression: TMEM75, M34330, and BC033263. Also, the genes MTSS1 and KIAA0196 were analyzed. Although >1 Mb from any single risk SNP, these genes do decrease within the originally described 3.8-Mb admixture peak identifying 8q24 as a risk locus (3) and were added to the analysis because they have been implicated in prostate carcinogenesis (22, 23).

Associations between risk allele status and transcript abundance at these six additional 8q24 transcripts (Supplementary Fig. S1) were assessed in 105 histologically normal European American and 20 histologically normal African American RP tissues. The same analysis was performed in paired tumor tissue specimens in a 33-person subset of the European American samples. As with MYC, associations were sought between total risk alleles, the number of risk alleles within a risk region and the number of risk alleles at an individual SNP. Among the European Americans, no significant association was seen between expression of any of the six transcripts and risk allele status in normal and tumor tissue. Among African Americans, a statistically significant association (P = 0.0068) was detected with expression levels of PVT1 transcript BC033263 and risk allele status across region 2 (Supplementary Fig. S1). In this instance, expression increased when comparing those with two risk alleles in this region to those with four. However, expression decreased when comparing carriers of four risk alleles to carriers of five. Thus, no consistent trend in gene expression correlated with risk allele status.

**Discussion**

Inherited variants at chromosome 8q24 are associated with prostate cancer risk, a finding validated in multiple cohorts and several ethnic groups (2–7). This discovery, made possible by the sequencing of the human genome, the International HapMap Project, and technological advances, provides a unique opportunity to gain insight into the pathogenesis of prostate cancer. However, the observation that many risk variants reside in nonprotein coding regions presents a formidable challenge to identifying the mechanism by which these variants cause disease. Because there are no known protein-coding genes at the risk loci, we sought to identify whether one or more miRNAs are transcribed in the region using a high throughput sequencing approach. No robust evidence of miRNA activity was found. This observation further supports the notion that the risk variants may be regulatory elements.

Previous studies have used gene expression as an intermediate trait to understand the mechanism through which risk alleles are acting (24, 25). Several factors point to MYC as a prime candidate for being the transcript under regulation by risk variants (26–28). Yet, no study has comprehensively quantified MYC expression levels in prostate tissue of men across all of the known risk alleles. A previous study genotyped 32 individuals for the SNP rs1447295 and reported a statistically significant difference in histologically normal prostate MYC expression between individuals homozygous for the wild-type allele and those heterozygous for the risk SNP. However, the analysis was based only on six heterozygotes (29).

Upon discovery of an 8q24 risk variant for colon cancer (rs983267, in prostate cancer risk region 3), cytoplasmic and nuclear immunohistochemical staining of MYC in 86 colon cancer samples were analyzed based on risk allele status (8). No significant differences were identified. In a recent study identifying 8q24 variants conferring risk for bladder cancer—located only 30 kilobase upstream of MYC—MYC expression was measured as a function of risk allele status; RNA expression levels in blood and adipose tissue were examined, and no significant association with risk allele status was found (11).

The data in the present analysis strongly substantiate that steady-state levels of MYC in normal and tumor prostate tissues are not associated with risk allele status. One nominally statistically significant association was observed for MYC expression in a microdissected tumor tissue sample set. The association was detected between MYC expression in European American patients carrying the risk allele at SNP rs13254738. This finding, however, must be carefully interpreted. The signal seen in this subgroup is primarily driven by an increase in MYC expression among homozygotes for the risk allele; however, only seven individuals were homozygous for this variant. Expression levels of heterozygotes for the risk SNP in this subgroup actually decreased slightly relative to those homozygous for the nonrisk allele. Moreover, two other tumor data sets presented here did not show this association, nor did laser capture microdissected African American tumor samples. If the rs13254738 finding is a true positive, it suggests that the other risk polymorphisms are associated with risk via mechanisms other than through MYC. Although this is a possibility, and one that was modeled in our analysis, it seems an unlikely scenario. Given the number of tests performed, the possibility that this observation is due to chance must be considered.

Our study had 80% power to detect a minimal difference of 2- to 2.5-fold in mean expression levels, depending on allele frequencies, among European American subjects at an α level of 0.05. Among African American subjects, the study had 80% power to detect a minimal difference of 2.4- to 3.3-fold in mean expression levels, depending on allele frequencies.

Expression of 6 other candidate transcripts at 8q24 in a total of 125 RP specimens was also evaluated for association with risk allele status. These included three transcripts within PVT1. Recent
evidence suggests that PVT1, a noncoding RNA complex, plays a significant role in cancer pathogenesis as an activator of MYC and/or as an oncogene itself (30, 31). No consistently significant trend in gene expression for these transcripts or for genes FAM84B, KIAA0196, and MTS1 were found to be associated with risk SNP genotype. Power considerations for this aspect of the study were similar to those for MYC.

Despite these findings, the 8q24 locus, and MYC in particular, should continue to be the target of further investigation. Recent work shows that risk loci may act as enhancer elements and these elements come into contact with MYC (32). Influence on MYC expression therefore is a likely mechanism by which the 8q24 risk polymorphisms exert their effects. Mechanisms other than changes in steady state levels of MYC should be considered. Risk alleles may influence the rate of MYC mRNA expression, rather than total abundance, with steady-state RNA remaining relatively constant throughout the cell. Risk alleles may increase MYC expression in a nonepithelial component, such as stromal cells; previous work has suggested that certain signals can dramatically shift MYC expression from one cell type to another despite little change in overall mRNA expression (33). Additionally, steady-state levels may remain consistent across genetic risk groups but response to cellular insult or stimulation may differ. Risk alleles also may influence transcription at an earlier stage in development and the effect may no longer be present at a later stage of life when prostate cancer is diagnosed (34). In addition, it is possible that the risk alleles affect the transcription of MYC isoforms that remain to be annotated. The ENCODE project has revealed a complex landscape of transcription and that many loci have a surprising number of previously unannotated exons (13). Finally, risk alleles may influence MYC expression so subtly that it is very difficult given current technology to detect the causative changes. There may be selective pressure keeping expression of MYC at relatively low levels to avoid activation of intrinsic tumor suppression (35).

In summary, one of the first reproducible and robust genetic associations for prostate cancer has been identified at chromosome 8q24 (2, 3, 8, 9, 11, 36). Data presented here suggest that transcription of miRNA is not present within risk regions. Steady-state expression levels of multiple transcripts at 8q24, including MYC, across many different conditions are not associated with the risk alleles. Identification of the gene(s) involved in this process will lend critical insight into the pathways that, when deregulated, result in prostate cancer. In fact, the majority of risk alleles discovered to date by genetic association studies are located in nonprotein coding regions. Establishing a framework for understanding the functional consequences of inheriting these alleles will become increasingly important.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 2/3/09; revised 4/2/09 accepted 4/3/09; published OnlineFirst 6/23/09. Grant support: NIH RO1 CA129435 (M.L. Freedman), the Mayer Foundation (M.L. Freedman), the Harry L. Snyder Medical Foundation (M.L. Freedman), the Dana-Farber/Harvard Cancer Center Prostate Cancer Specialized Programs of Research Excellence (National Cancer Institute Grant no. 5P50CA93081), the American Society of Clinical Oncology (M.L. Freedman), the Prostate Cancer Foundation (M.M. Pomerantz), and the Fred Hutchinson Cancer Research Center New Development funds, including support from the Canada Foundation (M. Tewari), Pilot Grant from the Pacific Northwest Prostate Cancer Specialized Program of Research Excellence Grant P50 CA97186 (M. Tewari), and Chromosome Metabolism Training Grant 5 T32 CA09657-16 (S.K. Wyman).

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 David Reich for his guidance in designing and analyzing ancestry-informative markers and Oliver Sartor for expert advice in project design.

References

25. Emilsson V, Thorleifsson G, Zhang B, et al. Genetic associations for prostate cancer has been identified at chromosome 8q24 (2, 3, 8, 9, 11, 36). Data presented here suggest that transcription of miRNA is not present within risk regions. Steady-state expression levels of multiple transcripts at 8q24, including MYC, across many different conditions are not associated with the risk alleles. Identification of the gene(s) involved in this process will lend critical insight into the pathways that, when deregulated, result in prostate cancer. In fact, the majority of risk alleles discovered to date by genetic association studies are located in nonprotein coding regions. Establishing a framework for understanding the functional consequences of inheriting these alleles will become increasingly important.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 2/3/09; revised 4/2/09 accepted 4/3/09; published OnlineFirst 6/23/09. Grant support: NIH RO1 CA129435 (M.L. Freedman), the Mayer Foundation (M.L. Freedman), the Harry L. Snyder Medical Foundation (M.L. Freedman), the Dana-Farber/Harvard Cancer Center Prostate Cancer Specialized Programs of Research Excellence (National Cancer Institute Grant no. 5P50CA93081), the American Society of Clinical Oncology (M.L. Freedman), the Prostate Cancer Foundation (M.M. Pomerantz), and the Fred Hutchinson Cancer Research Center New Development funds, including support from the Canada Foundation (M. Tewari), Pilot Grant from the Pacific Northwest Prostate Cancer Specialized Program of Research Excellence Grant P50 CA97186 (M. Tewari), and Chromosome Metabolism Training Grant 5 T32 CA09657-16 (S.K. Wyman).

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 David Reich for his guidance in designing and analyzing ancestry-informative markers and Oliver Sartor for expert advice in project design.
Evaluation of the 8q24 Prostate Cancer Risk Locus and MYC Expression

Mark M. Pomerantz, Christine A. Beckwith, Meredith M. Regan, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-09-0387

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2009/12/30/0008-5472.CAN-09-0387.DC1

Cited articles
This article cites 35 articles, 13 of which you can access for free at:
http://cancerres.aacrjournals.org/content/69/13/5568.full.html#ref-list-1

Citing articles
This article has been cited by 27 HighWire-hosted articles. Access the articles at:
/content/69/13/5568.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.