A Large-Scale Analysis of Genetic Variants within Putative miRNA Binding Sites in Prostate Cancer


Published OnlineFirst February 17, 2015; DOI: 10.1158/2159-8290.CD-14-1057
Prostate cancer is the second most common malignancy among men worldwide. Genome-wide association studies have identified 100 risk variants for prostate cancer, which can explain approximately 33% of the familial risk of the disease. We hypothesized that a comprehensive analysis of genetic variations found within the 3′ untranslated region of genes predicted to affect miRNA binding (miRSNP) can identify additional prostate cancer risk variants. We investigated the association between 2,169 miRSNPs and prostate cancer risk in a large-scale analysis of 22,301 cases and 22,320 controls of European ancestry from 23 participating studies. Twenty-two miRSNPs were associated (P < 2.3 × 10^{-5}) with risk of prostate cancer, 10 of which were within 7 genes previously not mapped by GWAS studies. Further, using miRNA mimics and reporter gene assays, we showed that miR-3162-5p has specific affinity for the KLK3 rs1058205 miRSNP T-allele, whereas miR-370 has greater affinity for the VAMP8 rs1010 miRSNP A-allele, validating their functional role.

SIGNIFICANCE: Findings from this large association study suggest that a focus on miRSNPs, including functional evaluation, can identify candidate risk loci below currently accepted statistical levels of genome-wide significance. Studies of miRNAs and their interactions with SNPs could provide further insights into the mechanisms of prostate cancer risk. Cancer Discov; 5(4):368–79. © 2015 AACR.

See related commentary by Yousef, p. 351.
cleaving or degrading mRNA and/or inhibiting its translation (6–8). Most miRNA binding has been observed within the 3′ untranslated region (3′UTR) of their target genes, although there are examples of binding within mRNA coding regions (9). As of March 2014, the miRBase database lists >2,570 mature miRNAs for humans. miRNAs are expressed in a tissue- and cell-specific manner with differential expression profiles in response to disease conditions, with many of these miRNA expression modulations contributing to disease progression (10–15). An impressive effort has been devoted to investigating miRNA dysregulation profiles in prostate cancer. Hence, miRNAs have emerged not only as potential biomarkers for prostate cancer but also as potential therapeutic targets (15–17).

miRNAs negatively regulate their target mRNAs primarily through Watson–Crick base-pairing interactions (18, 19). The most critical region for mRNA binding and repression is miRNA nucleotides 2 to 8, referred to as the miRNA seed site. Experiments have shown that genetic variations within the seed site or in the target mRNA at sites complementary to miRNA seed sites, referred to as miRSNPs, may reduce effectiveness or abolish miRNA-mediated repression, having functional consequences for cancer risk (20, 21). For example, Liu and colleagues (22) recently reported that miRSNPs in ITGA6 are associated with a decreased risk of prostate cancer. In another study assessing 61 putative miRSNPs in a Chinese population, three SNPs were associated with prostate cancer progression, whereas four SNPs were associated with prostate cancer–specific mortality (23). However, all these studies have been conducted using small sample sizes and might not be reflective of true positive association.

To further explore the genetic association of miRSNPs and to derive more reliable risk estimates of previously identified prostate cancer risk miRSNPs, we investigated the association between 2,169 miRSNPs and prostate cancer risk and aggressiveness in 23 studies participating in the Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome (PRACTICAL) Consortium. This effort included 22,301 cases and 22,320 controls of European ancestry. We then validated the functional role of two prostate cancer risk miRSNPs, Kallikrein 3 (KLK3) rs1058205 (T>C) and Vesicle-associated membrane protein 8 (VAMP-8) rs1010 (A>G), as they were most strongly associated with disease aggressiveness. To our knowledge, this is the first large-scale investigation of the association between miRNA-related gene polymorphisms and prostate cancer risk.

RESULTS

Patient Characteristics

The characteristics of the study participants are presented in Table 1. The mean age at diagnosis for cases (64.8 years) was older than the age at interview for controls (60.6 years). Cases (22.1%) were more likely to have a family history of prostate cancer compared with controls (13.9%). As expected, the majority of cases were diagnosed with tumors with a low (≤7) Gleason score (85.5%) that were localized (72.8%) and...
Association of miRSNPs with Prostate Cancer

Figure 1 and Supplementary Table S1 show the results of the association analyses for 2,169 putative miRSNPs with prostate cancer risk. Twenty-five miRSNPs had a minor allele frequency <0.01 in control samples. A total of 22 SNPs (Table 2; Supplementary Fig. S1) representing 16 genes were associated with risk of prostate cancer after correction for multiple testing (P < 2.3 × 10^−5). The most significant association was observed for rs1058205 located within the KLK3 3′UTR with an OR of 8.6 (0.83–0.9), P = 1.7 × 10^−14. This SNP was previously identified in a fine-mapping study (24). Similarly, MDM4 rs4245739 was recently reported in the primary iCOGS analysis (4). Ten SNPs—rs2450975, rs3103353, and rs3127593 (SLC22A2), rs1567669 (NKX3-1), rs1010 (VAMP8), rs1810126 (SLC22A3), rs2647257 (TET2), rs140829 and rs1043853 (PDLIM5), and rs17664 (ITGA6)—were found in the genes/regions previously implicated by prostate cancer GWAS studies. Ten SNPs—rs879161, rs7615039, and rs12492606 (PHC3), rs311497 (GMEB2), rs1530865 and rs2357637 (PDK1), rs12573077 (ARL3), rs7402 (MCAT), rs47340 (TTLL12), and rs4233979 (TMEM17) in seven genes (PHC3, GMEB2, PDK1, ARL3, MCAT, TTLL12, and TMEM17; Table 2), are at least 20 kb away from the previously reported index GWAS SNPs within the locus. Although GMEB2 and ARL3 have been previously reported, these seven genes have not been mapped by previous GWAS as per the National Human Genome Research Institute catalog in November 2014 (25).

In secondary analysis, seven SNPs showed significant differences in per-allele ORs between aggressive and nonaggressive disease (Supplementary Table S2). The most significant difference was observed for the KLK3 rs1058205 SNP; however, this SNP was more strongly associated with nonaggressive disease, which is in line with previous reports on other Kallikrein SNPs (24, 26). Interestingly, only two SNPs, rs1010 in VAMP8 and rs311497 in GMEB2, showed stronger association with aggressive disease. The rs1567669 SNP (NKX3-1) was associated with PSA levels in the patient cohort (Supplementary Table S3). Six SNPs (rs1043853, rs1058205, rs14082, rs2450975, rs3103353, and rs3127593), including KLK3 rs1058205, were marginally associated with PSA levels in controls (Supplementary Table S3). Six SNPs showed a trend with respect to age at diagnosis, including rs1058205, rs1043853, rs12492606, rs14082, rs7615039, and rs879161 (Supplementary Table S4).

Gene Expression and eQTL Analysis

Using the Oncomine (Compendia Bioscience) analysis tool, we compared the expression levels of the 16 genes harboring 22 significant miRSNPs. Using the Grasso dataset (27) of 59 tumor and 28 nontumor samples, expression of 7 genes was found to be deregulated in prostate cancer versus matched benign tissue analysis (Fig. 2). We found KLK3 and VAMP8 to be the second- and seventh-highest deregulated genes within the 16 genes analyzed.

We undertook a cis-eQTL analysis of the 22 prostate cancer–associated miRSNPs using The Cancer Genome Atlas (TCGA) data. Three SNPs [rs1058205 (KLK3), rs1530865, and rs2357637 (PDK1)] were not covered by the TCGA genotyping platform (Affymetrix 6.0) and did not have an appropriate proxy SNP identified by SNA Annotation and Proxy search (SNAP; ref. 28), and thus could not be analyzed. The remaining 19 SNPs (six index and 13 proxy SNPs identified by SNAP) were assessed for correlation with the respective gene expression harboring these SNPs. SNP rs2450975 (indexed by rs316000, r^2 = 0.95) in the SLC22A2 gene was found to be associated with mRNA transcript expression (P = 1.76 × 10^−5), whereas the SLC22A2 SNP, rs10945656 (indexing rs3103353 and rs3127593, r^2 = 1.0), and VAMP8 rs1010 SNP showed a trend (P = 0.09) toward genotype transcript expression (Supplementary Table S5). None of the other SNPs were associated with transcript levels of the gene harboring the SNP.

Functional Validation of the KLK3 rs1058205 and VAMP8 rs1010 miRSNPs

Using a range of computational prediction algorithms, we identified three miRNAs predicted to have differences in binding affinity between the KLK3 rs1058205 SNP-alleles. SNPinfo (29) and mirsnpscore (30) predicted miR-219-1-3p to target the T-allele, MicroSNiPer (31) and mirsnpscore predicted miR-3162-5p to target the T-allele, and MicroSNiPer and mirsnpscore predicted miR-4278 to target the C-allele. Two miRNAs were predicted to have differences in binding affinity between the VAMP8 rs1010 SNP-alleles. SNPinfo predicted both miR-103 and miR-370-5p to target the A-allele (sense strand = T-allele).

Reporter vector assays were then used to test the validity of the in silico–predicted mRNA binding potential to their target genes with specific genotype. For KLK3 rs1058205 SNP, miRNA miR-3162-5p induced an approximate 29% (P = 0.048) decrease in luciferase levels for the T-allele compared with the C-allele, suggesting that miR-3162-5p has specific affinity for the T-allele (Fig. 3A). No significant changes were observed for miR-219-1-3p or miR-4278 with either of the alleles for SNP rs1058205 (Supplementary Fig. S2A and S2B). For VAMP8 rs1010 (A>G), though miR-370-5p induced a change in luciferase activity for both alleles, there was a decrease in luciferase levels for the A-allele was approximately 2-fold (P = 0.0067) stronger than for the G-allele (Fig. 3B). Although miR-103 was found to regulate VAMP expression, it showed comparable results for both alleles (Supplementary Fig. S2C).

Expression of miR-3162-5p and miR-370-5p in Prostate Cancer

Though previously reported in melanoma, breast cancer, and cervical cancer (with expression upregulated in cervical cancer; refs. 32–34), prostatic expression for miR-3162-5p has not been determined. However, the miR-3162-5p gene is located within intron seven of the oxysterol binding protein gene, which is known to be expressed in the normal and cancerous prostate (32–35). Using qPCR, we specifically confirmed miR-3162-5p expression in a range of cancers and noncancerous prostatic cell lines (Supplementary Fig. S3A) as well as in patient tissue samples (Supplementary Fig. S3B). miR-3162-5p was detected in all the model cell lines and patient samples with varying expression levels. Prostatic expression for miR-370-5p has been reported previously to be upregulated in cancer (17, 36).

Published OnlineFirst February 17, 2015; DOI: 10.1158/2159-8290.CD-14-1057
<table>
<thead>
<tr>
<th>miRSNP</th>
<th>Gene</th>
<th>Transcript variant</th>
<th>Predicted miRNA(s)</th>
<th>chr</th>
<th>Position</th>
<th>Risk OR (95% CI)</th>
<th>Risk P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4245739</td>
<td>MDM4(^d)</td>
<td>NM_002393, NM_001204171, NM_001204172</td>
<td>miR-191-5p, miR-887, miR-3669</td>
<td>1</td>
<td>204518842</td>
<td>0.92 (0.89–0.95)</td>
<td>7.81E−08</td>
</tr>
<tr>
<td>rs4233979</td>
<td>TME17</td>
<td>NM_198276</td>
<td>miR-299-5p</td>
<td>2</td>
<td>62727902</td>
<td>1.08 (1.04–1.12)</td>
<td>1.61E−05</td>
</tr>
<tr>
<td>rs1010</td>
<td>VAMP8</td>
<td>NM_003761</td>
<td>miR-370, miR-103</td>
<td>2</td>
<td>85808982</td>
<td>0.91 (0.89–0.94)</td>
<td>3.79E−11</td>
</tr>
<tr>
<td>rs17664</td>
<td>ITGA6(^e)</td>
<td>NM_001079818, NM_000210</td>
<td>miR-548c-3p, miR-548a, miR-4691-3p</td>
<td>2</td>
<td>173369231</td>
<td>0.93 (0.91–0.96)</td>
<td>5.51E−07</td>
</tr>
<tr>
<td>rs1530865</td>
<td>PDK1</td>
<td>NM_002610</td>
<td>miR-877-5p, miR-3125, miR-494</td>
<td>4</td>
<td>95586224</td>
<td>1.09 (1.06–1.12)</td>
<td>1.27E−09</td>
</tr>
<tr>
<td>rs2357637</td>
<td>PDK1</td>
<td>NM_002610</td>
<td>miR-3916, miR-3125, miR-887-5p</td>
<td>2</td>
<td>17346138</td>
<td>0.81 (0.76–0.88)</td>
<td>3.66E−08</td>
</tr>
<tr>
<td>rs7615039</td>
<td>PHC3</td>
<td>NM_024947</td>
<td>miR-208a, miR-208b</td>
<td>3</td>
<td>169806170</td>
<td>0.86 (0.83–0.89)</td>
<td>8.72E−12</td>
</tr>
<tr>
<td>rs12492606</td>
<td>PHC3</td>
<td>NM_024947</td>
<td>miR-939, miR-362-5p</td>
<td>3</td>
<td>169808354</td>
<td>0.86 (0.83–0.89)</td>
<td>1.23E−11</td>
</tr>
<tr>
<td>rs879161</td>
<td>PHC3</td>
<td>NM_024947</td>
<td>miR-27a-5p, miR-320, miR-3158-3p</td>
<td>3</td>
<td>169812115</td>
<td>0.86 (0.83–0.89)</td>
<td>6.57E−12</td>
</tr>
<tr>
<td>rs14082</td>
<td>PDLIM5(^e)</td>
<td>NM_006457, NM_001011513, NM_001256425, NM_001256426, NM_001256428</td>
<td>miR-128, miR-494</td>
<td>4</td>
<td>95588274</td>
<td>1.08 (1.05–1.11)</td>
<td>9.65E−09</td>
</tr>
<tr>
<td>rs1043853</td>
<td>PDLIM5(^e)</td>
<td>NM_006457, NM_001011513, NM_001256425, NM_001256426, NM_001256427, NM_001256428</td>
<td>miR-567, miR-3120, miR-4310</td>
<td>4</td>
<td>95588274</td>
<td>1.08 (1.05–1.11)</td>
<td>9.65E−09</td>
</tr>
<tr>
<td>rs2647257</td>
<td>TET2(^e)</td>
<td>NM_001127208</td>
<td>miR-301a, miR-301b, miR-4330</td>
<td>4</td>
<td>106199905</td>
<td>1.1 (1.06–1.13)</td>
<td>2.05E−10</td>
</tr>
<tr>
<td>rs2450975</td>
<td>SLCA2A(^e)</td>
<td>NM_003058, NM_003058</td>
<td>miR-412, miR-4282</td>
<td>6</td>
<td>160637975</td>
<td>1.12 (1.09–1.16)</td>
<td>9.36E−13</td>
</tr>
<tr>
<td>rs3127593</td>
<td>SLCA2A(^e)</td>
<td>NM_003058</td>
<td>miR-200a, miR-302a, miR-488-3p</td>
<td>6</td>
<td>160638003</td>
<td>1.12 (1.08–1.17)</td>
<td>2.22E−08</td>
</tr>
<tr>
<td>rs3103353</td>
<td>SLCA2A(^e)</td>
<td>NM_003058</td>
<td>miR-942, miR-4268</td>
<td>6</td>
<td>160638076</td>
<td>1.12 (1.08–1.17)</td>
<td>1.75E−08</td>
</tr>
<tr>
<td>rs1810126</td>
<td>SLCA2A(^e)</td>
<td>NM_021977</td>
<td>miR-1205, miR-124-3p, miR-216b</td>
<td>6</td>
<td>160872151</td>
<td>1.1 (1.07–1.13)</td>
<td>2.81E−11</td>
</tr>
<tr>
<td>rs1567669</td>
<td>NKX3-1(^e)</td>
<td>NM_006167, NM_001256339</td>
<td>miR-637, miR-1275, miR-625</td>
<td>8</td>
<td>23538533</td>
<td>1.11 (1.08–1.15)</td>
<td>3.15E−13</td>
</tr>
<tr>
<td>rs12573077</td>
<td>ARL3</td>
<td>NM_004311</td>
<td>miR-432, miR-1258, miR-1224-5p</td>
<td>10</td>
<td>104434630</td>
<td>0.93 (0.91–0.96)</td>
<td>6.43E−06</td>
</tr>
<tr>
<td>rs1058205</td>
<td>KLK3(^e)</td>
<td>NM_001648, NM_001030047, NM_001030048</td>
<td>miR-3162-5p, miR-219-1-3p, miR-4278</td>
<td>19</td>
<td>51363398</td>
<td>0.87 (0.83–0.9)</td>
<td>1.73E−14</td>
</tr>
<tr>
<td>rs311497</td>
<td>GMEB2</td>
<td>NM_012384</td>
<td>miR-26c, miR-492, miR-619, miR-4648</td>
<td>20</td>
<td>62221249</td>
<td>0.92 (0.89–0.94)</td>
<td>1.28E−09</td>
</tr>
<tr>
<td>rs7402</td>
<td>MCA7</td>
<td>NM_173467, NM_014507</td>
<td>miR-616-3p</td>
<td>22</td>
<td>43529029</td>
<td>0.93 (0.91–0.96)</td>
<td>4.41E−06</td>
</tr>
<tr>
<td>rs47340</td>
<td>TILL12</td>
<td>NM_015140</td>
<td>let-7f, let-7g, let-7l, miR-103, miR-107, miR-764</td>
<td>22</td>
<td>43562829</td>
<td>1.07 (1.04–1.1)</td>
<td>5.79E−06</td>
</tr>
</tbody>
</table>

\(^a\) miRSNPs are present within the 3′ UTRs of the most common splice variants of the miRNA target genes.

\(^b\) miRNAs were predicted using four algorithms (Methods).

\(^c\) adjusted for 6 principal components and study group.

\(^d\) SNP previously identified in prostate cancer GWAS and fine-mapping studies.

\(^e\) Other SNPs within or around the genes for the listed SNPs have been identified in previous GWAS studies.
Association of miRSNPs with Prostate Cancer

**Figure 1.** miRSNP association with prostate cancer risk. Manhattan plot with −log $P$ values adjusted for study site and principal components. A total of 2,169 miRSNPs were assessed for association with prostate cancer risk. Twenty-two miRSNPs representing 16 genes were found to be associated with risk of prostate cancer after correction for multiple testing ($P < 2.3 \times 10^{-5}$). (The 10 most significant SNPs are labeled.) KLK3 rs1058205 and VAMP8 rs1010 (in bold) were selected for further functional validation studies.

**Figure 2.** Expression levels of 16 genes harboring 22 significant miRSNPs in cancerous and normal tissue from patients with prostate cancer. Oncomine analysis of the Grasso dataset (26) of 59 tumor and 28 nontumor (normal) samples shows the expression of 7 genes (PDLIM5–VAMP8; in bold) to be deregulated ($P < 0.05$) in prostate cancer. Colors are z-score normalized to depict relative values within rows. They cannot be used to compare values between rows.
Levels by miR-3162-5p ancestry. We identified 22 SNPs within the 3′UTR of the 16 genes to be associated with risk of prostate cancer, 7 of which, although within the previously identified GWAS locus, were not previously mapped by the GWAS studies. The most significant association was for the KLK3 rs1058205 SNP, previously identified to be associated with prostate cancer risk in a recent study (24).

Seven of these SNPs, including KLK3 (rs1058205) and VAMP8 (rs1010), a gene for which little is known about its prostatic function, showed significant differences between aggressive and nonaggressive disease. As expected, these results were not as robust due to the small sample size in the aggressive disease subcohort. Thus, it was not surprising that, as has been the case for many previously GWAS-identified SNPs (37), we could identify only two SNPs to be more significantly associated with aggressive disease.

Using Oncomine, we compared the expression levels of these 16 genes in a dataset consisting of 59 tumor and 28 nontumor samples, revealing the expression of seven of these genes to be deregulated in prostate cancer. KLK3 was the second most deregulated gene. We then chose to validate the functional role of the KLK3 rs1058205 SNP, previously identified as having greater affinity for the T-allele. Our results support the previously identified GWAS locus, were not previously mapped by the GWAS studies. The most significant association was for the KLK3 rs1058205 SNP, previously identified to be associated with prostate cancer risk in a recent study (24).

Seven of these SNPs, including KLK3 (rs1058205) and VAMP8 (rs1010), a gene for which little is known about its prostatic function, showed significant differences between aggressive and nonaggressive disease. As expected, these results were not as robust due to the small sample size in the aggressive disease subcohort. Thus, it was not surprising that, as has been the case for many previously GWAS-identified SNPs (37), we could identify only two SNPs to be more significantly associated with aggressive disease.

Using Oncomine, we compared the expression levels of these 16 genes in a dataset consisting of 59 tumor and 28 nontumor samples, revealing the expression of seven of these genes to be deregulated in prostate cancer. KLK3 was the second most deregulated gene. We then chose to validate the functional role of the KLK3 rs1058205 SNP, previously identified as having greater affinity for the T-allele. Our results support the

**DISCUSSION**

In this study, we investigated the association between 2,169 putative miRSNPs and prostate cancer risk in a large sample set including 22,301 cases and 22,320 controls of European ancestry. We identified 22 SNPs within the 3′UTR of the 16 genes to be associated with risk of prostate cancer, 7 of which, although within the previously identified GWAS locus, were not previously mapped by the GWAS studies. The most significant association was for the KLK3 rs1058205 SNP, previously identified to be associated with prostate cancer risk in a recent study (24).

Seven of these SNPs, including KLK3 (rs1058205) and VAMP8 (rs1010), a gene for which little is known about its prostatic function, showed significant differences between aggressive and nonaggressive disease. As expected, these results were not as robust due to the small sample size in the aggressive disease subcohort. Thus, it was not surprising that, as has been the case for many previously GWAS-identified SNPs (37), we could identify only two SNPs to be more significantly associated with aggressive disease.

Using Oncomine, we compared the expression levels of these 16 genes in a dataset consisting of 59 tumor and 28 nontumor samples, revealing the expression of seven of these genes to be deregulated in prostate cancer. KLK3 was the second most deregulated gene. We then chose to validate the functional role of the KLK3 rs1058205 SNP, previously identified as having greater affinity for the T-allele. Our results support the

**Figure 3.** miR-3162-5p directly targets the KLK3 rs1058205 SNP T-allele, and miR-370-5p targets the VAMP8 rs1010 SNP A-allele with greater affinity. Following overexpression with miR-3162-5p, reporter vector assays demonstrated an approximate 29% decrease in luciferase levels (P = 0.048) for the KLK3 rs1058205 SNP T-allele compared with the C-allele (A). Overexpression of miR-370-5p resulted in a change in luciferase activity for both VAMP8 rs1010 SNP variants with the decrease for the A-allele approximately 2-fold (P = 0.0067) stronger than for the C-allele (B). Mean ± SD, n = 3. *P < 0.05; **P < 0.01.

**Figure 4.** miR-3162-5p induces a reduction in KLK3 mRNA and KLK3 protein expression in LNCaP cells homozygous for the rs1058205 T SNP-allele. A, qPCR analyses in LNCaP cells revealed a 25% decrease in KLK3 mRNA (P = 0.016) following overexpression of miR-3162-5p compared with the negative control miRNA mimic treatment. B, Western blot analyses in LNCaP cells revealed a 32% decrease in cellular KLK3 protein (P = 0.007) following overexpression of miR-3162-5p compared with the negative control. C, representative Western blot. Mean ± SD, n = 3. *P < 0.05; **P < 0.01.
Association of miRSNPs with Prostate Cancer

emerging “miRNA network” that contributes to prostate cancer by regulating kallikrein and non-kallikrein genes (38).

As prostatic expression for miR-3162-5p had not been determined previously, we confirmed its expression in a range of cancerous and noncancerous prostatic cell lines and in patient tissue samples. It is interesting to note that prostate cancer PC3 cell lines with the highest miR-3162-5p expression do not produce any endogenous KLK3. In previous studies, this differential expression has been attributed to absence of an androgen receptor in the PC3 cell lines because KLK3 expression is androgen dependent in other prostate cancer cell lines. However, its regulation by miR-3162-5p could be an alternative regulatory mechanism and rationale for no KLK3 expression in these cells.

KLK3 is also referred to as a prostate-specific antigen (PSA), as it is expressed at orders of magnitude higher in the prostate compared with other tissues (39). As KLK3 serum levels are often elevated in prostate cancer, largely due to leakage associated with a loss of tissue architecture, KLK3 is thus utilized as the major serum biomarker for this disease (39, 40). Given the potential importance of KLK3 for prostate cancer diagnosis, we further characterized the miRNA-KLK3 rs1058205 interaction, demonstrating that miR-3162-5p is able to cause a decrease in KLK3 mRNA and KLK3 protein expression in LNCaP cells homozygous for the T-allele. Interestingly, the KLK3 rs1058205 SNP was associated with PSA levels in the control population (Supplementary Table S3), which may reflect some effects of strong regulatory factors (such as miRNA), exerting genotype-specific effects for this locus. The rs1058205 SNP, in addition to other PSA-associated SNPs, may therefore have implications for PSA-based diagnosis, hence requiring adjustments to PSA ranges for specific genotype. Furthermore, as KLK3 belongs to a family of 15 homologous genes, it would be important to consider potential additional effects of miR-3162-5p miRNA on other kallikrein and non-kallikrein targets in future studies.

The KLK3 rs1058205 T-allele was previously shown to be associated with increased prostate cancer risk (24). Here, we demonstrate that decreased KLK3 expression induced by miR-3162-5p targeting of the T-allele represents a mechanism by which the rs1058205 T-allele may be associated with increased prostate cancer risk. Interestingly, it has been shown that more aggressive prostate tumors have lower tissue levels of KLK3 (41). One mode via which KLK3 may act in a protective capacity in prostate cancer is through inhibition of angiogenesis (42, 43). Although the full mechanism is unclear, the antiangiogenic effect of KLK3 has been attributed to its proteolytic function on various angiogenic and antiangiogenic proteins (44, 45). However, due to its additional proteolytic function and its subsequent potential to target components of cell–cell adhesion and the extracellular matrix, high levels of KLK3 have also been proposed as a risk for prostate cancer. It is therefore possible that the effects of KLK3 on tumor development are stage-specific, with low KLK3 contributing to increased localized tumor growth (as observed in genetic risk analysis), whereas high KLK3 poses a risk at later metastatic stages of tumor development.

The VAMP8 rs1010 SNP was also selected for functional validation due to its significant association with aggressive prostate cancer, where miR-370-5p was found to have greater affinity for the VAMP8 rs1010 A-allele versus the minor G-allele. Interestingly, prostatic expression for miR-370-5p has been reported previously to be upregulated in cancer (17, 35).

To our knowledge, this is the first study to report an association and mechanism of action between a VAMP8 miRSNP and prostate cancer risk. VAMP8 is an integral membrane protein that is involved in the fusion of synaptic vesicles with the presynaptic membrane. It also plays a complex role in the control of granule secretion, transport vesicle trafficking, phagocytosis, and endocytosis (46, 47). Loss of VAMP8 has been shown to affect glucose metabolism, energy expenditure, and insulin sensitivity in mice (48). Though a direct role of VAMP8 in cancer is unknown, its ability to influence glucose metabolism and energy expenditure makes it a potential candidate in carcinogenesis, in relation to the shift in cellular metabolism from oxidative phosphorylation to glycolysis (the Warburg effect) that occurs in cells undergoing malignant transformation (48, 49). Hence, the role of VAMP8 may be important for prostate cancer. The VAMP8 rs1010 SNP was previously associated with risk for early onset myocardial infarction (50) and is in high linkage disequilibrium (LD); \( r^2 = 0.98 \) with an intergenic SNP (rs10187424) identified in a previous GWAS for prostate cancer risk \( \text{OR}, 0.92; 95\% \text{CI}, 1.05–0.90; 95\% \text{confidence interval (CI)}, 0.89–0.94; P = 2.1 \times 10^{-6} \) (ref. 37). No functional relevance has been assigned to the rs10187424 SNP. In the current study, we demonstrated that miR-370-5p has greater affinity for the VAMP8 rs1010 A-allele, thus identifying the likely causal variant behind the GWAS marker SNP. Nevertheless, the possibility of another functional variant in LD with rs1010, or any other putative functional miRSNPs (including KLK3) identified in our study cannot be ruled out.

Larger sample sizes are now required to provide additional power to assess true independence and/or the effect of these SNPs as modifiers of the unknown functional variants and/or top risk GWAS SNPs using conditional regression and/or haplotype analysis. Furthermore, it should be noted that the size effects of these variants are very small though comparable with previous GWAS studies. Thus, once independent causal variants or haplotypes at each of the known GWAS loci are identified, it would be interesting to undertake risk score calculations to assess the additive effects of all GWAS-identified SNPs, including the miRSNPs identified in our study.

Although our analysis has identified several miRSNPs previously not reported by the GWAS analysis, functional validation of these variants is required. Our eQTL analysis did not yield any significant results for genotype-mRNA expression correlation except for an SNP in SLC22A2, which is not surprising given that the miRNA machinery might not affect the mRNA levels in situ but will only inhibit the translation of these genes. Future studies are warranted to correlate genotypes with protein expression using immunohistochemistry and/or Western blot analysis. Additional functional studies may further clarify the role of these novel miRSNPs in prostate cancer etiology.

In conclusion, our study has identified putative functional SNPs associated with prostate cancer risk in several genes that further show differential expression in tumor versus normal tissue from patients with prostate cancer. The functional validation for the rs1058205 and rs1010 miRSNPs herein provides increasing evidence that miRSNPs may be associated with prostate cancer risk.
METHODS

Study Populations

The Collaborative Oncological Gene-environment Study (COGS) is a large collaborative effort among different consortia, including PRACTICAL, to evaluate genetic variants for associations with the risk of prostate, ovarian, and breast cancers. Details of the study have been reported previously (4). Briefly, 32 studies participating in the PRACTICAL consortium contributed samples from 25,074 prostate cancer cases and 24,727 controls to COGS. The majority of studies were nested, population-based, or hospital-based case-control studies. Individuals were excluded from the study based on strict quality control criteria, including overall genotype call rate <95%, genotypically non-European origin, samples that were XX or XXY and therefore not genotypically males, or samples not concordant with previous genotyping within PRACTICAL. The present analysis included 44,621 samples (22,301 cases and 22,320 controls) of European ancestry. Demographic and clinical information on study participants, including age at diagnosis, Gleason score, stage of disease, PSA, and cause of death, was obtained through in-person interviews or medical or death records. Aggressive disease was defined as Gleason score ≥8, PSA >20ng/ml, disease stage of “distant” (outside the pelvis), or prostate cancer-associated death. Study was approved by each Institutional Review Board, and informed consent was obtained from each participant. Patient studies were conducted in accordance with the Declaration of Helsinki.

miRSNP Selection and Genotyping

A total of 2,169 miRSNPs within the 3′UTRs of the cancer-associated genes were selected for genotyping. An SNP was selected if differential miRNA binding potential for the alternative allele was predicted by at least two of four algorithms: (i) Mirnspscore (30); (ii) Miranda and (iii) Sanger (both available through SNPinfo; ref. 29); and (iv) MicroSNiPer (30). Genotyping was performed using a custom Illumina Infinium array that included 211,115 SNPs (the iCOGS chip; ref. 4). Genotypes were called using Illumina’s proprietary GenCall algorithm. SNPs were excluded from further analysis if the call rate was <95%, deviated from Hardy-Weinberg equilibrium in controls at P < 10⁻⁶, or if genotypes were discrepant in more than 2% of duplicate samples.

cis-eQTL Analysis

For each index miRSNP, we retrieved all the correlated (r² ≥ 0.8) variants in European populations from 1000 Genomes using SNAP (28). The preprocessed (Level 2) germline genotypes of the index or correlated SNPs were downloaded from the TCGA data portal, and the expression levels of genes harboring these SNPs were obtained via the cBio Portal for Cancer Genomics. Using standard QC analysis, 6 samples were removed either due to discordant sex information (X-chromosome homozygosity rate between 0.2 and 0.8) or due to a heterozygosity rate >3 SDs from the mean. An additional 45 individuals were removed due to ethnic heterogeneity as calculated using principal component analysis. Data from 178 Caucasian individuals were used for the final genotype expression correlation analysis by the Kruskal-Wallis test using IBM SPSS Statistics (version 22).

miRNA Target Reporter Vector Assays

To assess the validity of in silico predictions for miRNA-miRNA affinity, miRNA target luciferase reporter vector assays were performed. Reporter vectors were constructed for each major and minor SNP-allele variants for both KLK3 and VAMP8 using the pmirGLO Dual-Luciferase miRNA Target Expression Vector (Promega) for KLK3 constructs, and the pMIR REPORT vector (Ambion) for VAMP8 constructs. (Portions of KLK3 and VAMP8 pertaining to predicted miRNA binding regions were synthesized by Integrated DNA Technologies.) LNCaP cells were cotransfected with vector and mirVana miRNA Mimics (Life Technologies) using FuGENE transfection reagent (Promega) and then analyzed 24 hours later using the Dual-Luciferase Reporter Assay System (Promega) according to the manufacturer’s instructions. Luciferase levels were normalized against Renilla coexpressed from the same vector or against the β-galactosidase derived from coexpression with a second vector measured using Galacto-Light (Tropix) for VAMP8. For KLK3, on the day of transfection, cells were cultured in 1% charcoal-stripped serum for the remainder of the experiment to induce lower levels of endogenous KLK3 to minimize the impact of miRNA mimic-endogenous KLK3 binding on reporter vector assay sensitivity. A negative control mirVana miRNA Mimic, Negative Control #1 (Life Technologies), was used for analysis alongside candidate miRNAs. A single experiment consisted of each miRNA/vector treatment cultured in triplicate. Three independent experiments were conducted in total.

RT-qPCR Analysis to Assess miRNA Expression

miRNA was extracted as total RNA from cell lines (LNCaP, LAPC4, DUCAP, DU145, PC3, 22Rv1, RWPE1, and BPH1) obtained from ATCC unless otherwise stated (see Acknowledgments) using TRIzol (Fairfield, Ohio). Formalin-fixed and paraffin-embedded (FFPE) blocks from prostate tumors and their adjacent noncancer prostate were obtained from the Australian Prostate Cancer BioResource tumor bank. Tissue blocks containing the tumor cells were serially sectioned (20-μm sections) and transferred to glass slides. Slides were stained with methyl green, and the tumor areas were marked by a pathologist. Marked areas were then manually microdissected under a microscope using a sterile injection needle (size, 0.65 × 25 mm). RNA was extracted using the miRNeasy FFPE Kit (Qiagen).

To assess expression, reverse transcription and qPCR were performed using the TaqMan MicroRNA Reverse Transcription Kit and TaqMan MicroRNA Assays (Life Technologies). The small nuclear RNA RNU24 was used as an endogenous quantitative normalization control (16). Relative expression levels were calculated using the comparative C method. As TaqMan MicroRNA Assays were not commercially available for miR-3162-5p, we designed in-house assays for miR-3162-5p (Supplementary Methods S1) following the TaqMan methodology (S1).

RT-qPCR and Western Blot Analysis to Assess KLK3 mRNA and Protein Expression

LNCaP cells were plated at 150,000 cells per well on a 6-well plate overnight, then cultured in 1% charcoal-stripped serum (for the remaining duration of the experiment), and transiently transfected with 30 nmol/L of miRNA mimics using Lipofectamine RNAiMAX transfection reagent (Life Technologies) followed by a treatment with 10 nmol/L dihydrotestosterone (DHT) to stimulate KLK3 expression via the androgen receptor pathway. Total RNA was isolated after 24 hours incubation after DHT stimulation using the RNeasy Mini Kit (Qiagen) and assessed for quality and yield using a Nanodrop ND-1000 spectrophotometer. RNA was reverse transcribed using oligo dt primers and 300 ng of total RNA. qPCR for KLK3 was then performed using the SYBR Green PCR Master Mix (Life Technologies) follows the same sample in triplicate with β-actin used as an endogenous quantitative normalization control. Relative expression levels were calculated using the comparative C method. Primers were synthesized by Integrated DNA Technologies. Primer sequences for KLK3 were as follows: forward primer 5′-gaggaggaagactcacaac-3′, reverse primer 5′-ccagcagatgctcctttgct-3′. Primer sequences for...
β-actin were as follows: forward primer 5′-ggtgacacctccctgctgcaaac 3′, reverse primer 5′-gtggtagcttgcaacctgttctta-3′. Total protein was isolated using SDS lysis buffer (1% SDS, 5% glycerol, 10 mmol/L Tris, Roche Complete protease inhibitor), concentration was assessed via the bicinchoninic acid (BCA) method, and 10 μg of total protein was run using standard techniques on a 12% resolving polyacrylamide gel. Western blotting was performed using standard techniques with primary antibodies, including Rb anti-KLK3 (0.6 μg/mL; Dako; A0562) and Rb anti-β-actin (0.6 μg/mL; Abcam; ab25894). Western blots were imaged on an Odyssey Imaging System (LI-COR Biosciences) using fluorescently labeled secondary antibodies (Alexa Fluor 680 and 790; Invitrogen) with protein band intensities analyzed via densitometry using Odyssey Imaging System software.

Statistical Analysis

Demographic, clinical, and mortality information was summarized by mean (SD) and number (%). Ethnic groups were defined based on a subset of 37,000 uncorrelated markers that passed quality control (including ~1,000 selected as ancestry informative markers). The COGS data were combined with the three Hapmap2 populations, and multidimensional scaling was used to identify and exclude ethnic outliers (4). After exclusion of ethnic outliers, principal component analyses were carried out for Europeans. The first six principal components were used to control for population substructure as additional principal components did not reduce inflation further (4). Associations between individual SNPs and prostate cancer risk or aggressive disease were evaluated using logistic regression models to estimate per minor allele ORs and 95% CIs. Associations between individual SNPs and prostate cancer risk were also evaluated in a similar manner for different age categories. The associations between SNP genotypes and PSA level were assessed using linear regression, after log-transformation of PSA level to correct for skewness. Analyses were performed using SPSS and R. All models included study site and principal components as covariates. Unless otherwise stated, for all other analyses, three independent experiments were conducted with results presented as mean ± SD, and analyzed using a Student t test, with a P value of <0.05 considered statistically significant for the functional studies.

Disclosure of Potential Conflicts of Interest

R. Eles has received an honorarium from the speakers bureau of the UK Cancer Convention. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions


Development of methodology: S. Stegeman, D. Kim, S. Srinivasan


Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): R. Eles, S. Benlloch, K. Muir, B.G. Nordestgaard, P. Pharoah, K.-T. Khaw, C. Maier, A.S. Kibel, PRACTICAL Consortium, A.B. Spurdle

Study supervision: F. Wiklund, J.A. Clements, J. Batra

Other (provided human prostate cancer biospecimens and clinical data): Australian Prostate Cancer BioResource

Acknowledgments

The authors thank John Lai, Farhana Matin, and LeBre Moya for technical support; Charles Sawyer for providing the prostate cancer cell line LAPC4; and Matthias Nees for providing the prostate cancer cell line DuCAP. In addition, the results published here are based partly on data generated by TCGA, established by the NCI and the National Human Genome Research Institute. The authors also thank the specimen donors and relevant research groups associated with this project.

COGS Acknowledgments

This study would not have been possible without the contributions of the following: Per Hall (COGS); Douglas F. Easton, Paul Pharoah, Kyrkais Michailidou, Manjel K. Boilla, and Qin Wang (BCAC); Andrew Berchuck (OACR); Rosalind A. Eles, Douglas F. Easton, Ali Amin Al Olama, Zsofia Kote-Jarai, and Sara Benlloch (PRACTICAL); Georgia Chenevix-Trench, Antonis Antoniou, Lesley McGuigoff, Ferguson Couch and Ken Offit (CIMBA); Joe Dennis, Alison M. Dunning, Andrew Lee, and Ed Dicks, Craig Luccarini, and the staff of the Centre for Genetic Epidemiology Laboratory; Javier Benitez, Anna Gonzalez-Neira, and the staff of the CNIO genotyping unit; Jacques Simard and Daniel C. Tessier; Francois Bacot, Daniel Vincent, Sylvie LaBoissiere, and Frederic Robidoux and the staff of the McGill University and Genome Quebec Innovation Centre; Stig E. Bojesen, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory; and Julie M. Cunningham, Sharon A. Winderbank, Christopher A. Hilker, Jeffrey Meyer, and the staff of Mayo Clinic Genotyping Core Faculty.

The PRACTICAL Consortium, in Addition to Those Named in the Author List

Margaret Cook, PRACTICAL coordination; Angela Morgan, Arityia Lophathananon, Cyril Fisher, Daniel Leongamornlert, Edward J. Saunders, Emma J. Sawyer, Kyoevela Gniodasami, Malgorzata Tymrekowska; Michelle Guy, Naomi Livni, Rosemary Wilkinson, Sara Jurgunthart, Steve Hazel, and Tokhir Dadaev, UKGPCS; John Pedersen, John L Hopper, and Melissa C. Southey, MICS; Ami Karlsson, Carin Cavalli-Bjorkman, Jan-Erik Johansson, Jan Adolfsen, Markus Ayl, Michael Broms, and Peter Statin, CAPS and STHM1; Brian E. Henderson and Fredrick Schumacher, MEC; Anssi Auvinen, Kimmo.Taan, Lisa Maeattaennan, Paula Kujala, Teemu Murtola, and Tiina Wahlfor, TAMPERE; Andreas Roder, Peter Iversen, Krasnikh Kranck, Sune F. Nielsen, Maren Weisheier, CPS1 and CPS2, Tim J. Key, Hans Wallinder, and Sven Gustafsson, EPIC; Jenny L Donovan, Freddie Hamdy, Anne George, Athene Lane, Gemma Marsden, Michael Davis, and Paul Brown, ProtecT and ProMPT; Nora Pashayan, SEARCH; Sarah Holt, FHCR; Lisa B. Signoretti and Wei Zheng, SCS; Liang Wang, Lori Tillman, and Shun Rika, MAYO; Antje Rinckleb, Kathleen Herkomer, Manuel Lueddeke, Walther Vogel Poland Dominika Wokoroczky, Jan Lubsik, and Wojciech Klimkunak, ULM; Aida K. Dieffenbach, Christa Stegmaier, and Volker Arndt, ESTHER, Bahar Zachariah, Hyun Park, Juno Pow-Sang, Maria Rincon, and Selina Radlgen, MOFFITT, Aleksandra Vlahova, Atanaska Mitkova, Chavdar Slavov, Darina Kachakova, Elenko Popov, Svetlana Christova, Tihomir Dikov, and Vanio Mitev, PCMUS; Felicity Lose, APCB, QLD; Joana Santos, Joao Barros-Silva, Paula Paio, Pedro Pinto, Rui Henriques, and Sofia Maa, IPO-Porto.
RESEARCH ARTICLE

Stegeman et al.

The Australian Prostate Cancer BioResource, in Addition to Those Named in the Author List

Prof. Gail Risbridger and Dr. Renea Taylor (Prostate Cancer Research Group, Monash University, Clayton, Victoria, Australia; and Australian Prostate Cancer BioResource, Melbourne, Victoria, Australia); Prof. Wayne Tilley and A/Prof. Lisa Butler (Adelaide Prostate Cancer Research Centre, Adelaide, South Australia, Australia; and Australian Prostate Cancer BioResource, Adelaide, South Australia, Australia); A/Prof. Lisa Horvath (Garvan Institute of Medical Research/The Kinghorn Cancer Centre, Darlinghurst, New South Wales, Australia; Chits O’Brien Lifehouse, Campendonk, New South Wales, Australia; and Australian Prostate Cancer BioResource, Sydney, New South Wales, Australia); Dr. Trina Yeadon and Ms. Allison Eckert (Cancer Program, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia; Australian Prostate Cancer Research Centre-Queensland at the Princess Alexandra Hospital, Queensland University of Technology, Brisbane, Queensland, Australia; Australian Prostate Cancer BioResource, Brisbane, Queensland, Australia); Dr. Glen Wood, Dr. Peter Heathcote, and Dr. Greg Malone (The Brisbane Urology Clinic Central Queensland Urology Clinic, Queensland, Australia); Dr. Kris Kerr, Dr. Megan Turner, and Dr. Angus Collins (Sullivan and Nicholson Pathology, Brisbane, Queensland, Australia).

Grant Support

J. Batra was supported by a National Health and Medical Research Council (NHMRC) Career Development Fellowship. A.B. Spurdle is an NHMRC Senior Research Fellow. J.A. Clements is an NHMRC Principal Research Fellow. This work was supported by Cure Cancer Australia Foundation and Cancer Australia PiCCRS grant 1068321 and NHMRC grant 1050742 (Principal Investigator, J. Batra); Department of Defense grant W81XWH-12-1-0113 (Principal Investigator, J.Y. Park); and NCIN01CA128813 (Principal Investigator, J.Y. Park). E.K. Amankwah was supported by a Cancer Prevention Fellowship from the NCI (R25T CA147531) while this study was conducted. The Australian Prostate Cancer BioResource (APCB) funding was obtained from the NHMRC (enabling grant 614296) and an infrastructure grant from the Prostate Cancer Foundation of Australia. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby indicated in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Funding for the CRUK Study and PRACTICAL Consortium

This work was supported by the Canadian Institutes of Health Research; European Commission’s Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175); Cancer Research UK (CRUK) grants C5047/A7357, C1287/A10118, C1235/A10692, and C16915/A16135; and NIH Cancer Post-Cancer GWAS initiative grant no. 1 U19 CA148537-01 (the GAME-ON initiative). The authors acknowledge support from the NHMRC to the Biomedical Research Centre at The Institute of Cancer Research and Royal Marsden NHS Foundation Trust.

Funding for COGS

Funding for the iCOGS infrastructure came from the European Commission’s Seventh Framework Programme grant agreement n° 223175 (HEALTH-F2-2009-223175); Cancer Research UK grants C1287/A10118, C1235/A10710, C12292/A11174, C1261/A12014, C5047/A10384, C5047/A15007, and C5047/A10692; NIH CA128978; and the Post-Cancer GWAS initiative (U1U9 CA148537, U1U9 CA148065, and U1U9 CA148112—the GAME-ON initiative). This study was also supported by the Department of Defense (W81XWH-12-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund.

Received September 15, 2014; revised January 23, 2015; accepted February 2, 2015; published OnlineFirst February 17, 2015.

REFERENCES

Association of miRSNPs with Prostate Cancer


A Large-Scale Analysis of Genetic Variants within Putative miRNA Binding Sites in Prostate Cancer

Shane Stegeman, Ernest Amankwah, Kerenafitli Klein, et al.

Cancer Discovery 2015;5:368-379. Published OnlineFirst February 17, 2015.

Updated version
Access the most recent version of this article at:
doi:10.1158/2159-8290.CD-14-1057

Supplementary Material
Access the most recent supplemental material at:
http://cancerdiscovery.aacrjournals.org/content/suppl/2015/02/17/2159-8290.CD-14-1057.DC1

Cited articles
This article cites 51 articles, 12 of which you can access for free at:
http://cancerdiscovery.aacrjournals.org/content/5/4/368.full#ref-list-1

Citing articles
This article has been cited by 2 HighWire-hosted articles. Access the articles at:
http://cancerdiscovery.aacrjournals.org/content/5/4/368.full#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.