Priority Report

Cancer Research

Transcriptome Sequencing Reveals *PCAT5* as a Novel ERG-Regulated Long Noncoding RNA in Prostate Cancer

Antti Ylipää[1,2], Kati Kivinummi[1,2], Annika Kohvakka[2,3], Matti Annala[1,2], Leena Latonen[2,3], Mauro Scaravilli[2,3], Kimmo Kartasalo[1,2], Simo-Pekka Leppänen[1,2], Serdar Karakurt[2,3], Janne Seppälä[1,2], Olli Yli-Harja[1], Teuvo L.J. Tammela[4], Wei Zhang[5], Tapio Visakorpi[2,3], and Matti Nykter[1,2]

Abstract

Castration-resistant prostate cancers (CRPC) that arise after the failure of androgen-blocking therapies cause most of the deaths from prostate cancer, intensifying the need to fully understand CRPC pathophysiology. In this study, we characterized the transcriptomic differences between untreated prostate cancer and locally recurrent CRPC. Here, we report the identification of 145 previously unannotated intergenic long noncoding RNA transcripts (lncRNA) or isoforms that are associated with prostate cancer or CRPC. Of the one third of these transcripts that were specific for CRPC, we defined a novel lncRNA termed *PCAT5* as a regulatory target for the transcription factor ERG, which is activated in approximately 50% of human prostate cancer. Genome-wide expression analysis of a *PCAT5*-positive prostate cancer after *PCAT5* silencing highlighted alterations in cell proliferation pathways. Strikingly, an *in vitro* validation of these alterations revealed a complex integrated phenotype affecting cell growth, migration, invasion, colony-forming potential, and apoptosis. Our findings reveal a key molecular determinant of differences between prostate cancer and CRPC at the level of the transcriptome. Furthermore, they establish *PCAT5* as a novel oncogenic lncRNA in ERG-positive prostate cancers, with implications for defining CRPC biomarkers and new therapeutic interventions. *Cancer Res; 75(19); 4026–31. ©2015 AACR.*

Introduction

The most frequent genomic lesion in prostate cancers is deletion of 21q22 in 50% of cases resulting in overexpression of ERG, an ETS family transcription factor. A translocation following the deletion fuses the regulatory sequence of an androgen-regulated gene, most often TMPRSS2, with the protein coding sequence of ERG bringing it under androgen regulation (1). *ERG* is a critical proto-oncogene that disrupts the ability of the cells to differentiate when activated. *ERG* fusions also contribute to development of androgen independence in prostate cancer through inducing repressive epigenetic programs via activation of a Polycomb methylatransferase EZH2, inhibiting androgen receptor (AR) expression, and disruption of AR signaling (2). Overexpression of ERG, or other ETS transcription factors, such as ETV1, and ETV4 activates cell invasion programs (3). ETS-negative prostate cancers have rare alternate driving events, such as SKIL-activating rearrangements (4). Generally, the molecular mechanisms of action for ERG are yet to be fully understood.

Recently, long noncoding RNA (lncRNA) molecules that are mainly transcribed from the intergenic regions of the genome (lncRNAs) have become a focus in transcriptome studies of cancers (5). These molecules form an integral part of many biologic processes, often through interactions with the Polycomb complex, which lead to silencing tumor-suppressive functions (6), but many other mechanisms have also been described previously (7). Few prostate cancer–specific lncRNAs have been well characterized to date, particularly *PCGEM1* (8), *PRNCR1* (9), *PCAT1* (10) and *SChLAP1* (11). *PCAT1* is a regulator of cell proliferation and a target of the Polycomb-Repressive Complex 2 (PRC2) that represses *BRCA2* tumor-suppressor and controls homologous recombination (12). *SChLAP1* antagonizes the regulatory functions of the SWI/SNF chromatin-modifying complex leading to increased invasiveness and metastasis *in vitro*, and its expression predicts poor outcome in clinical setting (11).

We hypothesized that there is still a significant amount of previously unexplored transcriptomic differences between hormone-naive and castration-resistant prostate cancer (CRPC), especially in the expression patterns of lncRNAs. To conduct the first comprehensive characterization of protein-coding genes, small RNAs and lncRNAs in these prostate cancers, we

[1]Department of Signal Processing, Tampere University of Technology, Tampere, Finland. [2]Institute of Biosciences and Medical Technology—BioMediTech, University of Tampere, Tampere, Finland. [3]Fimlab Laboratories, Tampere University Hospital, Tampere, Finland. [4]Department of Urology, Tampere University Hospital and Medical School, University of Tampere, Tampere, Finland. [5]Department of Pathology, University of Texas MD Anderson Cancer Center, Houston, Texas.

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

A. Ylipää and K. Kivinummi contributed equally to this work.

Corresponding Author: Matti Nykter, University of Tampere, Biokatu 6, Tampere 33520, Finland. Phone: 358-50-318-6869; Fax: 358-3-364-1291; E-mail: matti.nykter@uta.fi and Tapio Visakorpi, University of Tampere, Biokatu 6, Tampere 33520 Finland. Phone: 358-40-717-4402; Fax: 358-3-364-1247; E-mail: tapio.visakorpi@uta.fi

doi: 10.1158/0008-5472.CAN-15-0217

deep-sequenced transcriptomes of 12 benign prostatic hyperplasias (BPH), 28 untreated prostate cancers, and 13 CRPCs. In addition to identifying several CRPC-specific lncRNAs, we discovered PCAT5, an ERG-regulated tumor growth–associated lncRNA, that is exclusively expressed in ERG-positive prostate cancers and CRPCs. Its functional association in prostate cancer progression may partly explain how ERG exerts its widespread effect in gene regulation.

Materials and Methods

Patient samples and sequencing

Fresh-frozen tissue specimens from 12 BPHs, 28 prostate cancers, and 13 CRPCs were acquired from Tampere University Hospital (Tampere, Finland). The BPHs included both transition zone ($n = 4$) and peripheral zone ($n = 8$) samples received either by transurethral resection of the prostate (TURP) or cystoprostatectomy, respectively, from patients without prostate cancer diagnosis. All cancer samples contained a minimum of 70% cancerous or hyperplastic epithelial cells. Prostate cancer samples were obtained by radical prostatectomy and locally recurrent CRPCs by TURP sequenced. Libraries were prepared for paired-end analysis on the Illumina HiSeq 2000. On average, we obtained 110 million 90 bp-long paired-end reads from the whole transcriptome sequencing (RNA-seq), and 8.2 million 50 bp-long single-end reads from the small RNA sequencing (sRNA-seq). The sequencing reads were subsequently aligned to the human genome, and expression estimates for all expressed transcripts were computed. On average were able to align 92% of the reads (and minimum of 84%), indicating sufficient quality reads from all samples (Supplementary Table S1). More detailed description of the experimental setup can be found in Supplementary Methods.

Transcriptome assembly

To fully characterize the wealth of expressed transcripts in different stages of prostate cancer, we assembled a consensus transcriptome from all the samples using Cufflinks (13) RABT (reference annotation–based transcript) assembly approach with NCBI 37.2/hg19 genome build. Comparing the assembled prostate cancer transcriptome with all the exonic and intronic sequences in human reference transcriptomes (UCSC hg19, NCBI build 37.2, Ensembl GRCh37, Gencode version 12e) resulted in identification of 99,120 novel loci of expression. Transcripts overlapping exonic sequences were labeled as known sequences (and not included in the 99,120 novel loci), transcripts fully contained in an intron were labeled as intragenic (32,744; 33%), and transcripts not overlapping with exonic or intronic sequences were labeled as intergenic (66,376; 67%). To reduce the effect of noise, we filtered the lowly expressed transcripts (maximum normalized read count under 500), and included only transcripts that were differentially expressed across the tumor types using a negative binomial test and Mann–Whitney *U* test (adjusted $P < 0.001$ for both tests). Filtering reduced the number of loci to 152 intergenic and 25 intragenic prostate cancer–associated novel loci of transcription that were expressed at a significant level and were differentially expressed between BPH and prostate cancer or prostate cancer and CRPC samples (Supplementary Table S2). More detailed description of the data analysis can be found in Supplementary Methods.

While taking into account the computationally predicted sequences of the transcripts, we manually inferred putative exon structures, different isoforms, and strandedness for 145 transcripts or isoforms merging some of the adjacent loci of transcription. The curation from 152 loci into 145 isoforms was made based on the recurrent splice junctions in the paired-end read data coinciding with canonical intron splice site motifs. We were able to infer these structural details only for about half of the loci. The rest of the loci may either encode functional single-exon transcripts or be parts of ambiguously expressed large genomic regions such as *SChLAP1* (11). Following the previously adopted nomenclature, we named the transcripts tentatively as TPCATs (Tampere prostate cancer–associated transcripts) followed by chromosome and locus identifications (Supplementary Table S2). The annotation process is described in Supplementary Methods.

Results

Comprehensive transcriptome analysis reveals alterations in key regulatory pathways

We integrated the sequencing data into a comprehensive view of the prostate cancer and CRPC transcriptomes. Hierarchical clustering (Fig. 1A) and principal component analysis (PCA; Fig. 1B) of gene-expression profiles separated BPH, prostate cancer, and CRPC samples into distinct clusters. From PCA analysis, we observed two additional clusters that represented cancers with special features: One cluster contained two AR-negative tumors, whereas another contained tumors with strong AR amplification. We looked for differentially expressed genes using the Mann–Whitney *U* test with threshold for significant difference $P < 0.0001$, and absolute difference between medians of length-normalized read counts above 200 and $\log_2$-ratio above 1. In total, we identified 798 genes and 20 small RNAs differentially expressed between BPH and prostate cancer, and 330 genes and 43 small RNAs between prostate cancer and CRPC (Supplementary Table S3). When pathway analysis was run using the genes that were differentially expressed between prostate cancer and BPH, cytochrome p450 metabolism, cell adhesion, and TGFβ signaling pathways were identified as aberrated. Altered processes between CRPC and prostate cancer were dominated by regulatory pathways in which NR4A1, EGR family, FOS, DUSP1, and ATF3 play a key role (Supplementary Table S3). These genes were overexpressed in prostate cancers but not in CRPCs, and their mutual correlation (Pearson correlation > 0.9) indicated potentially shared regulation.

To highlight the pathway level changes in cell-cycle and androgen regulation, we constructed pathway models of these processes and projected the observed expression changes onto these models. In cell cycle, we noted a strong combined overexpression of the proliferation markers *MKI67*, *TOP2A*, *AURKA*, and *EZH2* in half of CRPCs, suggesting a high proliferation rate in these tumors. This high proliferation rate was also reflected in the expression of cyclins *CCNB1*, *CCNB2* and *CCNE2*, and the cyclin-dependent kinase *CDK1* (Supplementary Fig. S1). In the androgen regulation pathway, we observed overexpression of AR in 7 of 13 CRPCs. The AR coactivator *FOXA1* was overexpressed in untreated prostate cancer relative to BPH. Isozymes *SRD5A1* and *SRD5A2*, responsible for testosterone-to-DHT conversion, showed respective up- and downregulation in CRPC. Enzymes *AKR1C3* and *AKR1C2*, responsible for canonical androstenedione-to-testosterone

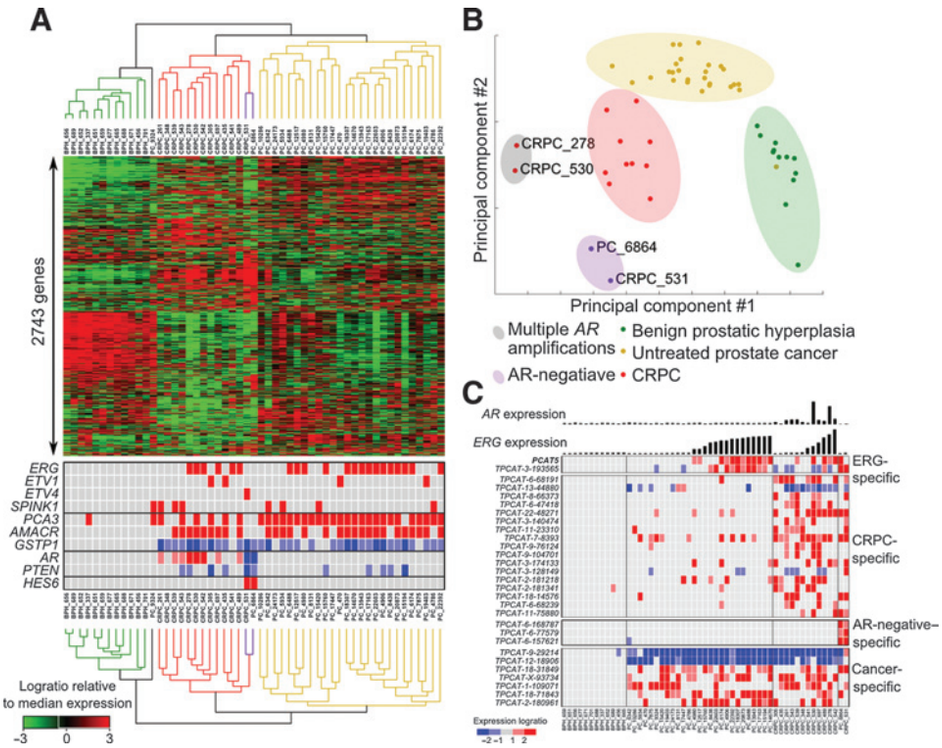

Figure 1.
Expression level characterization of prostate cancers. A, hierarchical clustering of annotated genes reveals distinct gene-expression signatures for BPH (green), prostate cancer (PC; yellow), and locally recurrent CRPC (red). High expressions of key marker genes, such as *ERG* and *AR*, have been highlighted for all the tumors in red and low expression in blue, different levels of shade indicating the level of expression difference from the median. Two tumors (PC_6864 and CRPC_531, purple) were negative for *AR* expression and positive for neuronal differentiation marker *HES6*. B, principal component analysis. BPH, prostate cancer, and CRPC samples are well separated into distinct clusters based on their gene-expression profiles. AR negative tumors (PC_6864 and CRPC_531) as well as tumors with strong *AR* amplification and overexpression (CRPC_530 and CRPC_278) formed separate clusters. C, association of found transcripts to disease phenotypes. A number of transcripts were associated with *ERG*-positive prostate cancers and CRPCs, including *PCAT5*. In addition, many transcripts were CRPC-specific and few showed specificity to untreated prostate cancers. Most transcripts were found in all cancer tissue types.

reduction, were overexpressed in 30% to 50% of CRPCs, with associated overexpression of *UGT2B15* and *UGT2B17*, enzymes responsible for glucuronidation of testosterone and DHT. Transcription factors *ERG* and *ETV1* were overexpressed in 25 of 41 prostate cancers corresponding to previously established frequency of fusions with the androgen regulated *TMPRSS2* (Supplementary Fig. S2; ref. 1).

The expression patterns of novel lncRNAs differentiate between prostate cancer and CRPC

Majority of the novel expressed loci were detected in CRPC only or in both prostate cancer and CRPC, but a few loci were expressed in all three sample groups, albeit at different rates, or were specific to the AR-negative samples (Fig. 1C). More than 30% of the transcripts were expressed on average at more than 10 times higher level in CRPCs than in prostate cancers or BPHs, which we considered highly CRPC-specific expression pattern. Some of the transcripts were expressed in only few samples corresponding to outlier expression pattern. The specificities of the TPCAT expression patterns were further validated using available RNA-sequencing data from 21 prostate cancer cell lines (10), 24 normal tissues (14), 2 human embryonic stem cells (PolyA-selected and non-selected; ref. 15), and two independent cohorts of prostate cancer tumors ($n = 30$ and $n = 34$, respectively; Fig. 2A and Supplementary Table S2; refs. 10, 16). Generally, TPCATs were minimally expressed in normal tissues, with testes most commonly being the normal tissue with the highest expression level. The expression patterns in cancer tissue were generally concordant in all three prostate cancer cohorts.

To investigate whether changes in DNA methylation or copy number bring about the expression of the novel transcripts in the samples that express them, we integrated DNA-sequencing and MeDIP-sequencing data from the same samples with the RNA-seq data (See Supplementary Methods). We computed Spearman correlations between transcript expression values and the copy number of the locus, and expression values and methylation values of nearby differentially methylated regions. In addition, we tested for differential expression between samples that had copy-number aberrations at the locus versus samples that had normal copy number for each TPCAT using the *t* test. We required a significant correlation between the expression and copy number, and significantly differential expression between samples with normal copy number and samples with copy-number aberration. Similar requirements were applied to methylation. None of the TPCATs were found to be significant taking account both criteria, indicating that the expression differences of TPCATs were

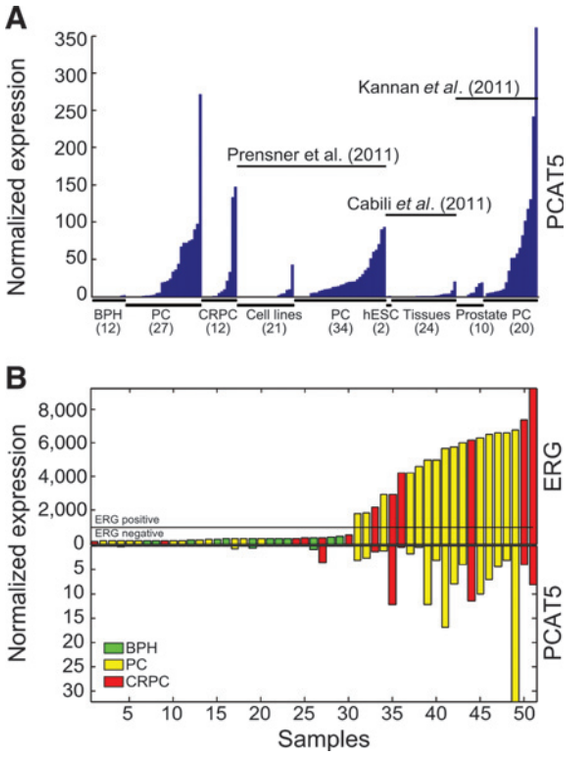

Figure 2.
Expression patterns of *PCAT5*. A, association of *PCAT5* expression to prostate cancer was confirmed by analyzing additional prostate cancer and healthy tissue RNA-seq datasets. B, *ERG* and *PCAT5* expressions derived from our RNA-seq data were plotted for each sample. A cutoff of 1,000 RPKs was chosen to differentiate ERG-negative and ERG-positive samples. The *ERG*-negative pool contained all the BPH samples, 12 prostate cancers, and 6 CRPC whereas the ERG-positive pool contained 15 prostate cancers and 6 CRPC samples.

not explained by these factors. This suggesting that, at least in general, TPCATs are transcriptionally regulated and that their expression does not arise due to genetic alterations or changes in DNA methylation (Supplementary Table S4).

We wanted to find prostate cancer–associated transcription factors that could act partly by regulating some of the lncRNAs we discovered. Spearman correlations were computed between the expression of the lncRNAs and eight transcription factors (*ERG*, *AR*, *FOXA1*, *EZH2*, *HDAC1*, *HDAC2*, *HDAC3*, and *RUNX2*), for which public ChIP-sequencing data in prostate cancer cell lines were available for validating the regulatory association. The correlation analysis associated the expression of several TPCATs with the expression of these transcriptional regulators (Supplementary Table S4). The strongest positive correlation (r = 0.69) was observed between *ERG* and transcript *TPCAT-10-36067* (officially termed *PCAT5*; Fig. 1C). Concordantly with *ERG* expression, *PCAT5* was expressed in a subset of prostate cancers and CRPCs (Fig. 2A and B). The expression was detected at a comparable frequency in both independent cohorts of prostate cancer, but not significantly in healthy tissues, including BPHs. The expression of *PCAT5* was quantified and validated in independent cohort of 76 primary prostate cancer samples as well as *ETV4*-positive PC-3 and *ERG*-positive VCaP cells using qRT-PCR (Supplementary Fig. S3). In addition, and the expression correlation between *PCAT5* and *ERG* was validated in this 76 sample set with ERG immunohistochemistry, and in LuCaP xenografts with qRT-PCR ($r = 0.78$; Supplementary Fig. S3). Expressions of four additional CRPC-expressed multiexon TPCATs were also validated with RT-PCR (Supplementary Figs. S4 and S5). Because ERG is a dominant feature in prostate cancers, we decided to concentrate our validation efforts to deciphering the exact structure of *PCAT5*, elucidating the regulatory connection between *PCAT5* and ERG, and investigating the functional relevance of *PCAT5*.

Inhibition of *PCAT5* expression reduces growth, migration, and invasion of ERG-positive prostate cancer cells

On the basis of the spliced read alignments, we inferred a three-exon structure for *PCAT5* with no components of viral open reading frames or other repetitive elements located on the exons (Fig. 3A). Both exon–exon junctions were validated with RT-PCR and Sanger sequencing in three clinical samples (Supplementary Fig. S3). To accurately identify both termini of the transcript, we performed 5′ and 3′ rapid amplification of cDNA ends (RACE). Open reading frame analysis indicated that the

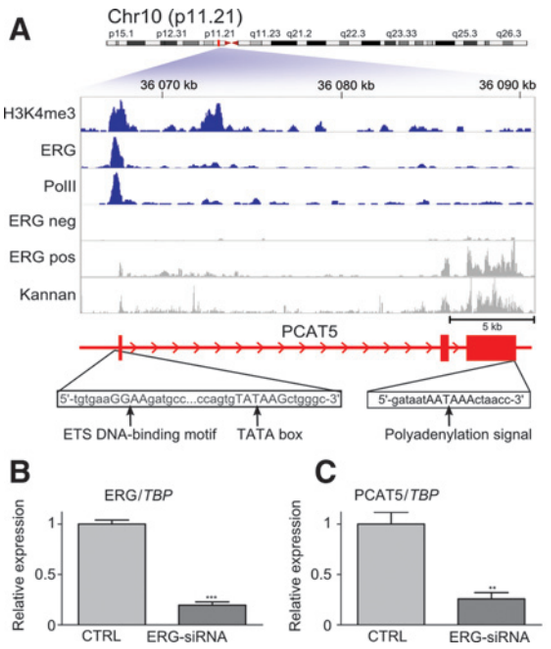

Figure 3.
Sequence analysis of *PCAT5*, located in 10p11.21. A, regulatory properties of *PCAT5* were investigated using ChIP-sequencing data of H3K4me3, ERG, and POL2 from *PCAT5*-positive VCaP cells (blue), and detailed expression profiles in *ERG*-negative and *ERG*-positive samples, and in an independent prostate cancer cohort (gray). The data were overlaid onto the inferred exon structure of *PCAT5* (in red), indicating open chromatin coinciding with ERG- and POL2-binding events in the promoter of VCaP cells, and active expression in *ERG*-positive samples compared with *ERG*-negative samples. An ETS DNA-binding domain and TATA box were identified at the suspected promoter region and a Poly-A signal sequence at the end of third exon. *ERG* (B) and *PCAT5* (C) expressions after 50 nmol/L ERG-siRNA or scrambled CTRL-siRNA treatment in VCaP cells; error bars, SEM; **, $P < 0.0083$; ***, $P < 0.001$, unpaired two-tailed *t* test ($n = 3$).

transcript lacks protein-coding potential. From available ChIP-sequencing data measured from ERG-positive VCaP cells (Supplementary Table S4), we identified open chromatin histone markers, such as H3K4 trimethylation, and binding events of ERG and RNA polymerase II at proximal promoter of *PCAT5* (Fig. 3A). Conversely, no ERG binding or H3K4 trimethylation was found at the *PCAT5* promoter in LNCaP cells, which do not express *PCAT5* (Supplementary Fig. S6). Sequence analysis revealed a canonical ETS family DNA-binding motif and a TATA-box coinciding with the locus that ERG was bound to, and a polyadenylation signal at the 3′-end of the transcript (Fig. 3A). We further validated the regulatory association by knocking down *ERG* in VCaP cells using an siRNA (Fig. 3B), which led to 75% inhibition of *PCAT5* expression (Fig. 3C). Similarly, we validated the association between *PCAT5* and another ETS-family transcription factor, ETV4, by knocking it down in ERG-negative PC-3 cells, leading to comparable inhibition of PCAT5 expression (Supplementary Fig. S7). These data indicate that *PCAT5* is under direct regulation by ERG, and likely other ETS family transcription factors as well.

To characterize and validate the function of *PCAT5*, we suppressed its expression with two different siRNAs in two cell lines: PC-3 cells, which expressed the transcript (Fig. 4A), and 22Rv1 cells, which did not express it. The genome-wide expression changes that the suppression induced to PC-3 cells were studied using expression arrays. Gene ontology enrichment analysis indicated that cell-cycle, mitosis, and Aurora signaling were the most extensively affected processes (Supplementary Table S5). Several functional assays validated the computationally identified biologic processes: The knockdown dramatically decreased cell growth (Fig. 4B) and invasiveness (Fig. 4C), and increased the rate of apoptosis (Fig. 4D). In addition, colony formation (Fig. 4E) and migration potential (Fig. 4F) of the transfected PC-3 cells decreased substantially compared with non-transfected PC-3 cells. Conversely, the growth rate of 22Rv1 cells that do not express *PCAT5* was unaffected by the siRNA suppression as expected (Supplementary Fig. S8) whereas the growth of ERG-positive DuCaP cells decreased after siRNA suppression of *PCAT5* (Supplementary Fig. S9). These results suggest that *PCAT5* has a key role in regulating tumor growth and malignancy in ETS-positive prostate cancers.

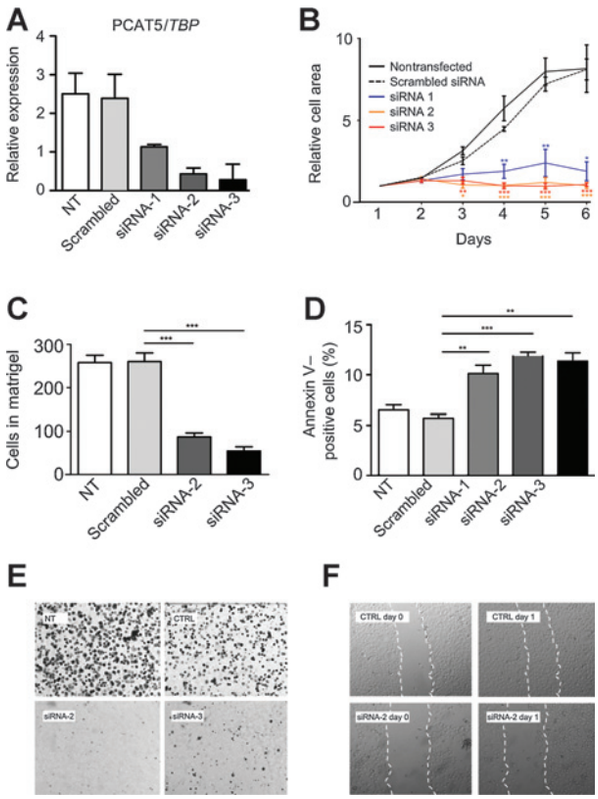

Figure 4.
Functional validation of *PCAT5*. A, successful silencing of *PCAT5* using multiple siRNA knockdowns was validated using qRT-PCR. B, growth of the *PCAT5*-positive PC-3 cell line was completely inhibited by siRNAs targeting the transcript. C, invasiveness of PC-3 cells was reduced by the siRNA knockdown. D, Annexin V assay indicated an increased rate of apoptosis after the knockdown; error bars, SEM; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, unpaired two-tailed t test. E, colony formation was significantly reduced in the *PCAT5*-deficient PC-3 cells. F, wound-healing assay (triplicates, representative experiment shown) indicated a significantly impaired migration ability.

Discussion

Hundreds of lncRNAs, for which little more than an expression pattern is known, have been discovered by RNA-sequencing and stored in databases such as NONCODE (17). A growing interest toward lncRNAs in cancer research is sparked by the dozens of molecules that have been implicated as key players in cancer cells (5). In prostate tumorigenesis, differential expression of hundreds of lncRNAs is already a recognized phenomenon (8, 9, 11). However, the functional role of many cancer-associated lncRNAs remains undetermined. The expression of lncRNA may confer clinical information about disease outcomes, and thus have utility as diagnostic tests. One prostate cancer–specific biomarker lncRNA, PCA3, is currently in use (18). Evidence for effectively targeting tumor-specific lncRNAs as a therapeutic regimen (19), such as the telomerase lncRNA *TERC*, are accumulating. Therefore, the characterization of the noncoding RNA species and their functions are clinically important.

To identify novel transcripts, a comparable experimental and computational approach was taken by Prensner and colleagues (10), which resulted in discovery of 121 lncRNAs in untreated prostate cancer. Here, we extended the list of prostate tumor–specific transcripts with 145 distinct molecular entities by including CRPC samples to the cohort and performing much deeper sequencing. Outlier-type expression patterns of many lncRNAs discovered in this article may explain why many of them had not been discovered to this date despite the use of RNA-sequencing technologies. Furthermore, many lncRNAs were expressed in moderate-to-low levels, which may have caused previous studies to overlook them. Also, PCAT5 that was identified as a key molecule in ERG-positive prostate cancers is quite lowly but consistently transcribed in other cohorts.

Integration with DNA-seq and MeDIP-seq data indicated that the expression of none of the novel transcripts correlated with copy-number or DNA methylation status, and thus it seems that the expression of these transcripts is not driven by copy-number changes or differential DNA methylation. For example, in cell lines that have open chromatin at *PCAT5* promoter likely express

it due to a binding of an ETS family transcription factor. The mechanism is intriguing because *ERG* overexpression is one of the hallmark events of prostate cancer, whereas the mechanisms downstream of ERG remain poorly understood. Our siRNA experiments revealed that *PCAT5* affects both cell growth and invasiveness, which suggests that the transcript may be an integral mediator in the regulatory cascade downstream of ERG. Although low-expression level is probably not ideal for a biomarker, the strong phenotype combined with prostate cancer specificity, makes *PCAT5* a prospective target for therapy.

In conclusion, we performed the first transcriptomic analysis on CRPC and identified more than hundred novel lncRNAs that seem to be specific for either prostate cancer or CRPC. One of the lncRNAs, *PCAT5*, was shown to be regulated by ERG and has a dramatic effect on prostate cancer cells. Inclusion of more specimens, especially CRPCs, would likely result in identification of even more novel lncRNAs with outlier type of expression pattern. Biopsies of metastases might also reveal novel lncRNAs that are specific to these tumors and ones that are not expressed in primary tumors. However, the identified transcripts form an interesting pool of putative biomarker and mechanisms for prostate cancer progression.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Authors' Contributions

Conception and design: A. Ylipää, K. Kivinummi, W. Zhang, T. Visakorpi, M. Nykter

Development of methodology: A. Ylipää, K. Kivinummi, S. Karakurt, M. Nykter

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): K. Kivinummi, A. Kohvakka, L. Latonen, M. Scaravilli, K. Kartasalo, S.-P. Leppänen, T.L.J. Tammela, T. Visakorpi

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A. Ylipää, K. Kivinummi, M. Annala, L. Latonen, M. Scaravilli, K. Kartasalo, J. Seppälä, W. Zhang, T. Visakorpi, M. Nykter

Writing, review, and/or revision of the manuscript: A. Ylipää, K. Kivinummi, A. Kohvakka, T.L.J. Tammela, W. Zhang, T. Visakorpi

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): K. Kartasalo, S.-P. Leppänen, O. Yli-Harja, T. Visakorpi

Study supervision: T.L.J. Tammela, T. Visakorpi, M. Nykter

Acknowledgments

The authors express gratitude to Dr. Robert L. Vessella for the LuCaP xenograft data. They also thank Paula Kosonen, Riina Kylätie, Päivi Martikainen, and Marika Vähä-Jaakkola for their technical assistance and CSC—IT Center for Science for providing computational resources.

Grant Support

The work was supported by grants from the Finnish Funding Agency for Technology and Innovation Finland Distinguished Professor program (O. Yli-Harja and M. Nykter), Academy of Finland (project nos. 269474 and 132877 to M. Nykter and project no. 251790 to O. Yli-Harja, and project no. 127187 to T. Visakorpi), Sigrid Juselius Foundation (M. Nykter and T. Visakorpi), Emil Aaltonen Foundation (M. Annala), Cancer Society of Finland (M Nykter and T. Visakorpi). Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital (grant 9P053 to T.L.J. Tammela and grant 9N087 to T. Visakorpi), and EU-FP7 Marie Curie Integrated Training Network, PRO-NEST (T. Visakorpi), the NIH (U24CA143835; to W. Zhang).

Received January 29, 2015; revised May 19, 2015; accepted July 2, 2015; published OnlineFirst August 17, 2015.

References

1. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–8.
2. Yu J, Yu J, Mani RS, Cao Q, Brenner CJ, Cao X, et al. An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. Cancer Cell 2010;17:443–54.
3. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 2007;448:595–9.
4. Annala M, Kivinummi K, Tuominen J, Karakurt S, Granberg K, Latonen L, et al. Recurrent SKIL-activating rearrangements in ETS negative prostate cancer. Oncotarget 2015;6:6235–50.
5. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. Cancer Discov 2011;1:391–407.
6. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007;129:1311–23.
7. Wang KC, Chang HY. Molecular mechanisms of long noncoding RNAs. Mol Cell 2011;43:904–14.
8. Srikantan V, Zou Z, Petrovics G, Xu L, Augustus M, Davis L, et al. PCGEM1, a prostate-specific gene, is overexpressed in prostate cancer. Proc Natl Acad Sci U S A 2000;97:12216–21.
9. Chung S, Nakagawa H, Uemura M, Piao L, Ashikawa K, Hosono N, et al. Association of a novel long non-coding RNA in 8q24 with prostate cancer susceptibility. Cancer Sci 2011;102:245–52.
10. Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, et al. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotechnol 2011;29:742–9.
11. Prensner JR, Iyer MK, Sahu A, Asangani IA, Cao Q, Patel L, et al. The long noncoding RNA SchLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. Nat Genet 2013;45:1392–8.
12. Prensner JR, Chen W, Iyer MK, Cao Q, Ma T, Han S, et al. PCAT-1, a long noncoding RNA, regulates BRCA2 and controls homologous recombination in cancer. Cancer Res 2014;74:1651–60.
13. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 2010;28:511–5.
14. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 2011;24:1915–27.
15. ENCODE Project Consortium Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, et al. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74.
16. Kannan K, Wang L, Wang J, Ittmann MM, Li W, Yen L. Recurrent chimeric RNAs enriched in human prostate cancer identified by deep sequencing. Proc Natl Acad Sci U S A 2011;108:9172–7.
17. Bu D, Yu K, Sun S, Xie C, Skogerbø G, Miao R, et al. NONCODE v3.0: integrative annotation of long noncoding RNAs. Nucleic Acids Res 2012;40:D210–5.
18. Hessels D, Schalken JA. The use of PCA3 in the diagnosis of prostate cancer. Nat Rev Urol 2009;6:255–61.
19. Li CH, Chen Y. Targeting long non-coding RNAs in cancers: progress and prospects. Int J Biochem Cell Biol 2013;45:1895–910.