Two Unique Novel Prostate-Specific and Androgen-Regulated Fusion Partners of *ETV4* in Prostate Cancer

Karin G. Hermans,[1] Anke A. Bressers,[1] Hetty A. van der Korput,[1] Natasja F. Dits,[2] Guido Jenster,[2] and Jan Trapman[1]

Departments of [1]Pathology and [2]Urology, Josephine Nefkens Institute, Erasmus University Medical Center, Rotterdam, the Netherlands

Abstract

Recently, fusion of *ERG* to the androgen-regulated, prostate-specific *TMPRSS2* gene has been identified as the most frequent genetic alteration in prostate cancer. At low frequency, *TMPRSS2-ETV1* and *TMPRSS2-ETV4* fusion genes have been described. In this study, we report two novel *ETV4* fusion genes in prostate cancer: *KLK2-ETV4* and *CANT1-ETV4*. Both gene fusions have important unique aspects. *KLK2* is a well-established androgen-induced and prostate-specific gene. Fusion of *KLK2* to *ETV4* results in the generation of an additional *ETV4* exon, denoted exon 4a. This novel exon delivers an ATG for the longest open reading frame, in this way avoiding translation start in *KLK2* exon 1. Although wild-type *CANT1* has two alternative first exons (exons 1 and 1a), only exon 1a was detected in *CANT1-ETV4* fusion transcripts. We show that *CANT1* transcripts starting at exon 1a have an androgen-induced and prostate-specific expression pattern, whereas *CANT1* transcripts starting at exon 1 are not prostate specific. So, the two novel *ETV4* fusion partners possess as predominant common characteristics androgen-induction and prostate-specific expression. [Cancer Res 2008;68(9):3094–8]

Introduction

Prostate cancer is the most common malignancy in men in Western countries (1). Growth of prostate tumors depends on androgen signaling, mediated by the androgen receptor (AR). Metastatic disease is treated by endocrine therapy; however, all tumors eventually become resistant to this therapy. The majority of resistant tumors still contain a functional active *AR* (2). In part of these tumors, *AR* is overexpressed due to amplification of a small region of the X chromosome, where *AR* maps (3).

Most frequent genomic alterations in primary prostate cancers are losses of large fragments of chromosome arms 6q, 8p, 13q, and 16q and gain of 8q (4, 5). In a subset of tumors, a small region of loss of chromosome 21q22 has been detected. This genomic alteration is associated with recurrent fusion of prostate-specific and androgen-regulated *TMPRSS2* (6) to the ETS transcription factor gene *ERG*, which maps at a distance of 3 Mbp from *TMPRSS2* on 21q (7).

TMPRSS2-ERG fusion is present in 40% to 70% of clinical prostate cancers (7, 8).[3] The gene fusion is an early event that has also been detected in a proportion of precursor lesions (9). Although in many tumors, *TMPRSS2-ERG* overexpression is accompanied by loss of the region between *TMPRSS2* and *ERG*, in others, this region has been retained, indicative of different mechanisms of gene fusion (10). At low frequency, fusion of *TMPRSS2* to a second ETS gene, *ETV1*, which maps to chromosome band 7p21, has been reported (7, 10). *TMPRSS2* seems the only fusion partner of *ERG*, but it has recently been shown that *ETV1* has more fusion partners (11).[4]

For a third ETS gene, *ETV4*, only fusion to *TMPRSS2* has been described (12). In the present study, we identified in clinical prostate cancer samples two unique novel *ETV4* fusion partners: *Kallikrein 2* (*KLK2*) and *Calcium-Activated Nucleotidase 1* (*CANT1*). KLK2, or hGK1, is a well-known prostate marker highly homologous to KLK3 or prostate-specific antigen (13). Like *KLK3*, *KLK2* is prostate-specific and androgen-regulated expressed (13, 14). We show that *CANT1* expression is also androgen regulated. *CANT1* possesses two alternative first exons, but only one of these is present in the *CANT1-ETV4* fusion transcript. In contrast to the majority of *CANT1* transcripts, this mRNA is preferentially expressed in the prostate. The novel fusion genes indicate prostate-specificity and androgen-regulation as important characteristics of *ETV4* fusion partners in prostate cancer.

Materials and Methods

Samples. Two series of clinical prostate cancer samples were obtained from the tissue bank of the Erasmus University Medical Center. Samples were snap frozen and stored in liquid nitrogen. All samples contained at least 70% tumor cells. Collection of patient samples has been performed according to national legislation concerning ethical requirements. Use of these samples has been approved by the Erasmus MC Medical Ethics Committee according to the Medical Research Involving Human Subjects Act (MEC-2004-261). Prostate cancer xenografts were propagated by serial transplantation on male nude mice as described (10, 15).

DNA and RNA isolation. RNA from clinical prostate cancer specimens was isolated using the RNA-Bee kit (Campro Scientific). DNA was isolated using the DNeasy DNA extraction kit (Qiagen). Xenograft RNA was isolated according to the LiCl protocol. RNA from the prostate cancer cell line LNCaP cultured in the presence of 10-9M R1881, or in the absence of hormone, was isolated using the RNeasy RNA extraction kit (Qiagen).

mRNA expression analysis. Analysis of mRNA expression was performed by reverse transcription-PCR (RT-PCR) or by QPCR. Two micrograms of RNA were reverse transcribed using 400 U M-MLV RT (Invitrogen) and an oligo-dT12 primer. cDNAs of 16 different tissues were purchased from Clontech. RT-PCR products were analyzed on a 1.5% agarose gel. QPCR was performed in an ABI Prism 7700 Sequence Detection System (Applied Biosystems), using Power SYBR Green PCR Master Mix (Applied Biosystems), containing 330 nmol/L forward and reverse primer,

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

Requests for reprints: Jan Trapman, Department of Pathology, Josephine Nefkens Institute, Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, the Netherlands. Phone: 31-107043933; Fax: 31-107044762; E-mail: j.trapman@erasmusmc.nl.

doi:10.1158/0008-5472.CAN-08-0198

[3] K.G. Hermans, J.L. Boormans, D. Gasi, G. Jenster, P. Verhagen, and J. Trapman, unpublished.

[4] K.G. Hermans, H.A. van der Korput, R. van Marion, A. Ziel-van der Made, N.F. Dits, T.H. van der Kwast, H. van Dekken, C.H. Bangma, H. Korsten, G. Jenster, and J. Trapman, unpublished.

in a total volume of 25 μL. Thermocycling conditions were according to the recommendations of the manufacturer. Messenger RNA expression was determined relative to *Porphobilinogen deaminase* (*PBDG*) by the Standard curve method (Applied Biosystems). Primers are listed in Supplementary Table S1.

RNA ligase–mediated rapid amplification of cDNA ends. 5′ RNA ligase–mediated rapid amplification of cDNA ends (RLM-RACE) was performed using the GeneRacer kit from Invitrogen according to instructions of the manufacturer. To obtain the 5′-sequence, cDNA was amplified with Qiagen Taq (Qiagen) using the Generacer 5′-primer and an *ETV4* exon 6 primer. PCR products were separated over a 1.5% agarose gel, bands were excised, purified, and sequenced on an ABI 3100 genetic analyzer (Applied Biosystems).

Array comparative genomic hybridization. Arrays were produced from the human 3600 BAC/PAC genomic clone set of the Welcome Trust Sanger Institute, covering the full genome at ~1-Mb spacing. DNA labeling and hybridization were performed as previously described (10). After hybridization, arrays were scanned in a ScanArray Express HT (Perkin-Elmer). The resulting images were analyzed with GenePix Pro 5.0 software (Axon Instruments) and subsequently visualized with an excel macro.

Interphase fluorescent *in situ* hybridization. Interphase fluorescence *in situ* hybridization (FISH) was done on 5-μm frozen tissue sections as described previously (10). BAC clones RP11-100E5 and RP11-209M4 (both flanking *ETV4*) were purchased from BacPac Resources. Specificity of BACs was confirmed on metaphase chromosome spreads. BAC DNAs were either Spectrum Orange– or Spectrum Green–labeled using a Nick Translation Reagent kit (Vysis). Tissue sections were counterstained with 4′,6-diamidino-2-phenylindole in antifade solution (Vector Laboratories). Images of the three fluorochromes were collected on an epifluorescence microscope (Leica DM) equipped with appropriate filter sets (Leica) and a CCD cooled camera (Photometrics).

Breakpoint mapping. Fusion points were mapped by standard long-range PCR on 200 ng genomic DNA in the presence of 0.5 μmol/L of each forward (fusion partner) and reverse (*ETV4*) primer with Taq polymerase and Proofstart DNA polymerase (Qiagen). Primers are given in Supplementary Table S1. PCR products were separated on a 1% agarose gel. Specific amplified fragments were isolated and sequenced.

Results and Discussion

ETV4 expression was studied by QPCR on two sets of clinical prostate cancer samples. Set 1 was composed of 84 clinical prostate cancer samples (49 primary prostate tumors, 11 lymph node metastases, and 24 recurrent tumors) and set 2 of 36 primary prostate tumors and 29 recurrent tumors. In primary tumor (98), and in recurrent tumor (206), both from set 1, *ETV4* overexpression was detected (Fig. 1*A*). In the second set, we did not detect *ETV4* overexpression (data not shown), indicating that overexpression is a rare event, occurring in <2% clinical prostate cancers. Follow-up experiments showed that *ETV4* overexpression was caused by fusions to two different genes with unique properties.

First, RT-PCR experiments with *TMPRSS2*- and *ETV4*-specific primers excluded *TMPRSS2* as fusion partner (data not shown). Next, we performed RLM-RACE using a reverse primer in *ETV4* exon 6. Sequencing of the amplified fragments showed *KLK2* (sample 98) and *CANT1* (sample 206) as novel *ETV4* fusion partners. The presence of *KLK2-ETV4* and *CANT1-ETV4* fusion transcripts in the individual samples was confirmed by RT-PCR (Fig. 1*B*).

The *KLK2-ETV4* mRNA fragment, detected by RLM-RACE, was composed of *KLK2* exon 1 linked to a new *ETV4* exon (here denoted exon 4a), followed by *ETV4* exon 5 and 6 sequences (Fig. 1*C*). The novel *ETV4* exon 4a has a length of 133 bp and delivers the ATG start codon of the longest predicted open reading frame in the *KLK2-ETV4* fusion transcript (Fig. 1*D*). *KLK2* maps to chromosome band 19q13 and *ETV4* on chromosome band 17q21. Because of the orientations of *KLK2* and *ETV4, KLK2-ETV4* gene fusion cannot be explained by a single chromosomal translocation.

The *CANT1-ETV4* fragment detected by RLM-RACE contained one of the two described exons 1 of *CANT1* (here denoted exon 1a). This exon maps ~4 Kbp downstream of the other first exon (here denoted exon 1; Fig. 1*C*). *CANT1* exon 1a delivers the ATG start

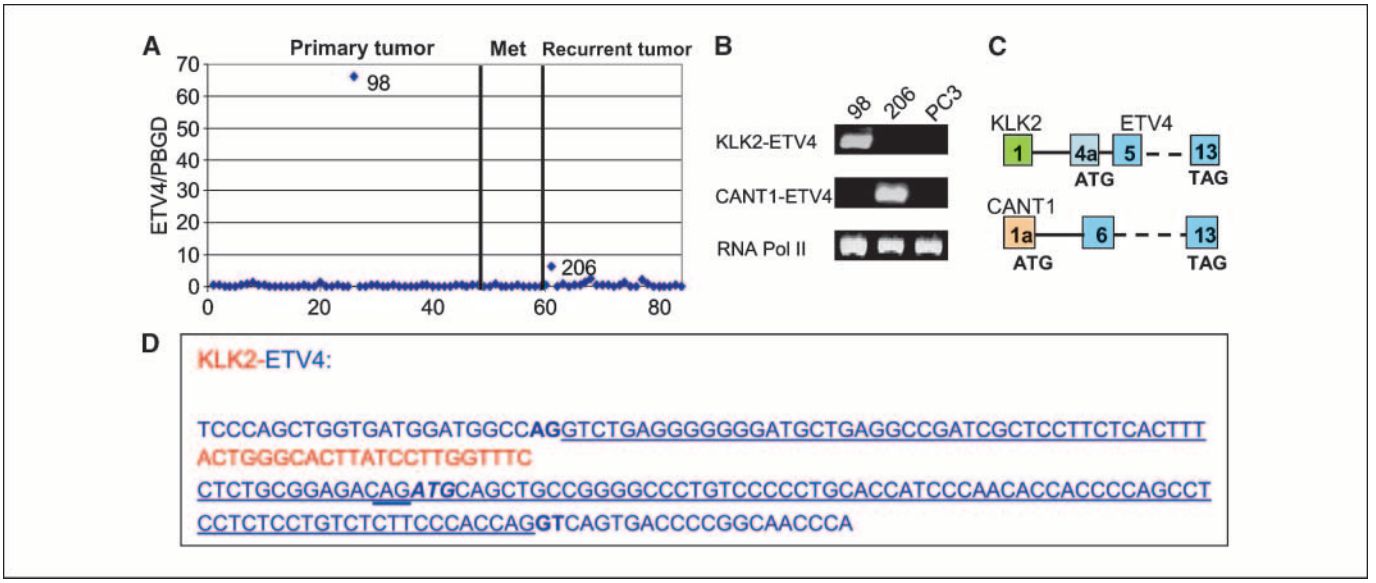

Figure 1. Expression of *ETV4* and characterization of *ETV4* fusion transcripts in clinical prostate cancer samples. *A,* QPCR analysis of *ETV4* expression in clinical prostate cancer samples compared with *PBGD* expression. Overexpression of *ETV4* was detected in samples 98 and 206. *Met,* regional lymph node metastasis. *B,* confirmation of *KLK2-ETV4* (sample 98) and *CANT1-ETV4* (sample 206) fusion transcripts by RT-PCR, using *ETV4* and fusion partner specific primers. RNA from prostate cancer cell line PC3 was used as a negative control. An RNA Pol II–amplified fragment is shown as a loading control. *C,* schematic representation of *KLK2-ETV4* and *CANT1-ETV4* fusion transcripts. Exons are indicated by colored boxes. In both transcripts, ATG start codons and TAG stop codons of longest predicted open reading frames are indicated. *D,* sequence of *ETV4* exon 4a and flanking sequences. The *ETV4* exon 4a sequence is underlined. *KLK2* intron 1 sequence present in the fusion gene is in red. Splice sites are bold. The ATG start codon in *ETV4* exon 4a is depicted in bold and italic.

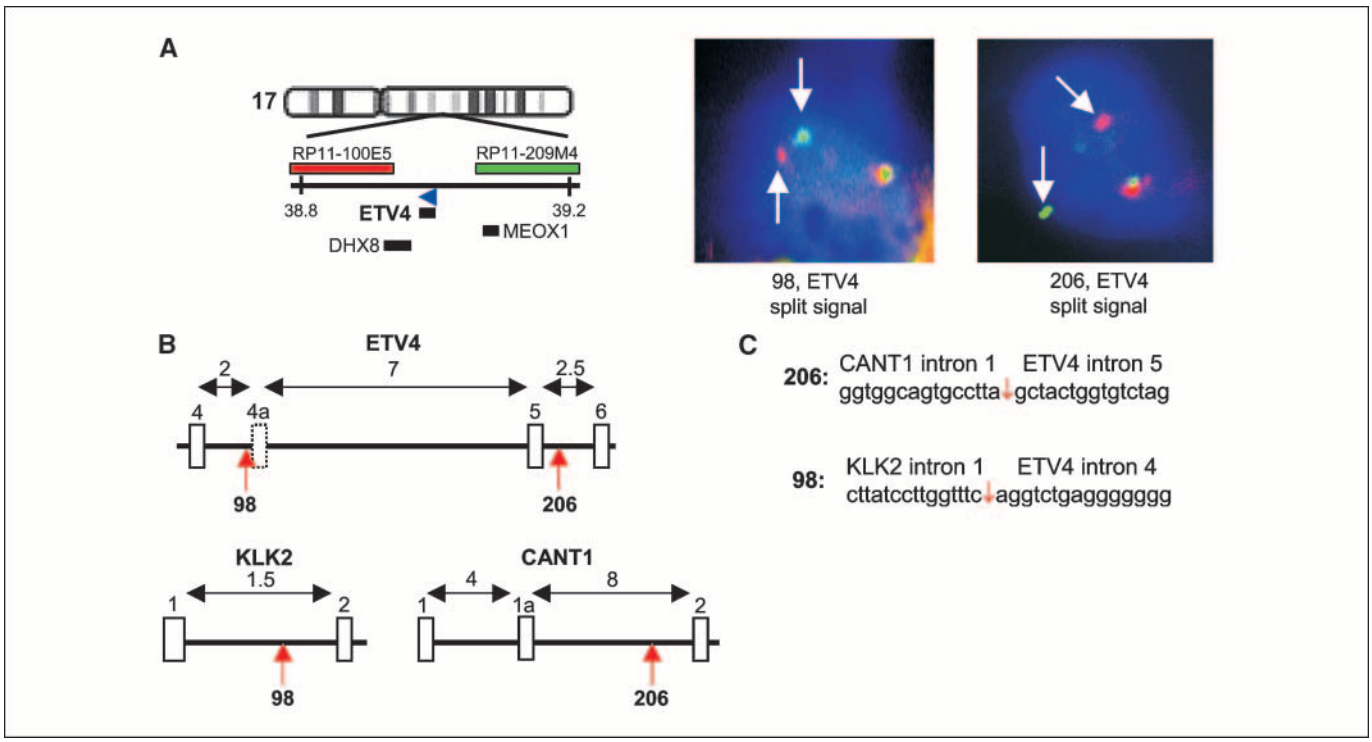

Figure 2. Characterization of *ETV4* fusion genes in clinical prostate cancer samples. *A,* schematic representation of the *ETV4* genomic region on chromosome 17. Distance from the top of chromosome is indicated in Mbp. *Blue arrowhead,* orientation of *ETV4*. BAC probes used in split signal FISH are indicated in colors corresponding to FISH staining. *Arrows,* split signal FISH on frozen tissue sections confirms *ETV4* rearrangements in samples 98 and 206. Schematic representation of genomic breakpoints in *ETV4, KLK2*, and *CANT1* as found by long-range PCR and sequencing. *Open boxes,* numbered exons. *Red arrows,* positions of genomic breakpoints. Distances in kbp between exons are indicated. *C,* sequences of *CANT1-ETV4* and *KLK2-ETV4* fusion points. *Red arrow,* fusion point.

codon of the predicted *ETV4* open reading frame. Remarkably, *CANT1-ETV4* fusion transcripts starting at *CANT1* exon 1 were not detected, although wild-type *CANT1* transcripts starting either at exon 1 or 1a of the nonrearranged second allele were present in the tumor sample and in normal prostate (data not shown). *CANT1* and *ETV4* map in the same orientation on 17q, at a distance of ~35 Mbp. However, array-based comparative genomic hybridization showed that the genomic region between *ETV4* and *CANT1* was not lost (data not shown), indicating that either an internal rearrangement of 17q had occurred or that this region was reintegrated in another part of the genome.

ETV4 rearrangements in samples 98 and 206 were confirmed by split signal FISH with probes flanking *ETV4* at both sites (Fig. 2*A*). Next, the genomic fusion points were mapped by long-range PCR and sequencing (Fig. 2*B* and *C*). As expected, the breakpoints of *CANT1-ETV4* were located in *CANT1* intron 1, downstream from exon 1a, and in *ETV4* intron 5, respectively. This latter breakpoint was located in a MER20 repeat, a low copy repetitive element known to be involved in unstable genomic regions prone to chromosomal rearrangement (16). Breakpoints in *KLK2* and *ETV4* in sample 98 were not in repetitive sequences. The breakpoint in *ETV4* intron 4 of the *KLK2-ETV4* rearrangement was 2 bp upstream of *ETV4* exon 4a (Figs. 1*D* and 2*C*). Due to the gene fusion the purine-rich *ETV4*, intron sequence was replaced by a more pyrimidine-rich sequence of *KLK2* intron 1, generating a novel splice acceptor sequence. To confirm that the rearrangement induced *ETV4* exon 4a, a splicing assay was performed (Supplementary Fig. S1). *ETV4* exon 4a was indeed retained if the *KLK2-ETV4* fusion sequence was used, but not if intron 4 sequence of wild-type *ETV4* was used in the assay (Supplementary Fig. S1).

KLK2 is a well-known androgen-regulated and prostate-specific gene (13, 14). To establish the specific characteristics of *CANT1* expression, QPCR analysis on cDNAs from 11 human prostate cancer xenografts derived from various stages of the disease and with different *AR* status was performed (10, 15). We compared expression of transcripts starting at *CANT1* exon 1 and transcripts starting at exon 1a to *KLK2* expression. *KLK2* mRNA expression was completely restricted to xenografts with highest *AR* expression (Fig. 3*A*). *CANT1* transcripts starting at exon 1a showed highest expression in the more differentiated, androgen-dependent xenografts (Fig. 3*B*). In contrast, transcripts starting at exon 1 were expressed at variable levels in all xenografts. Further analysis of RNA from LNCaP prostate cancer cells, which were *in vitro* cultured in the presence of the synthetic androgen R1881 or in the absence of hormone, showed that both transcripts starting at *CANT1* exon 1 and exon 1a are induced by androgens (Fig. 3*C*). As expected, *KLK2* expression was strongly induced by androgens. Next, we tested the tissue specificity of the *CANT1* and *KLK2* transcripts in a cDNA panel from 16 different normal tissue samples. *KLK2* showed an expected strictly prostate-specific expression pattern (Fig. 3*D*). Remarkably, transcripts starting at *CANT1* exon 1 were ubiquitously expressed, but transcripts starting at exon 1a had a much more restricted expression pattern, with highest expression in the prostate (Fig. 3*E*).

Both *KLK2-ETV4* and *CANT1-ETV4* have specific, unique characteristics: *KLK2-ETV4* because a novel *ETV4* exon is generated (exon 4a) and *CANT1-ETV4* because *CANT1* exon 1a is

exclusively used in *CANT1-ETV4* fusion transcripts. A KLK2-ETV4 fusion protein containing the NH_2-terminal KLK2 signal peptide would be secreted and could not function as a transcription factor. However, the start codon in the novel *ETV4* exon 4a, which is preceded by an in frame stop codon, prevents the generation of such a fusion protein. Instead, synthesis of a truncated ETV4 protein, starting in *ETV4* exon 4a can now be predicted (Fig. 1*D*). Exclusive usage of *CANT1* exon 1a as first exon in *CANT1-ETV4* fusion transcripts might have various explanations, including the positions of breakpoints of the specific genomic rearrangement and the prostate-specific expression of transcripts starting at exon 1a.

In prostate cancer, *TMPRSS2* is the common fusion partner of the ETS family members *ERG, ETV1*, and *ETV4*. Recently, several novel *ETV1* fusion partners have been identified: *SLC45A3, HERV-K_22q11.23, C15orf21*, and *HNRPA2B1* (11), with different expression characteristics. Based on expression profiles, the fusion partners were divided into distinct classes of *ETV1* rearrangements, separating prostate-specific, androgen up-regulated genes and down-regulated genes, and ubiquitously regulated genes. Recently, we identified three novel *ETV1* fusion partners: *FOXP1, HERVK17*, and *EST14*.[4] Both *HERVK17* and *EST14* have a prostate-specific and androgen-induced expression pattern; *FOXP1* seems expressed in many tissues. The three *ETV4* partners identified, *TMPRSS2, KLK2*, and *CANT1*, show identical expression profiles: up-regulation by androgens and prostate-specificity.

The common expression pattern of *ETV4* fusion partners indicates a similar mechanism of gene fusion. It is tempting to speculate that the genomic region of prostate-specific genes is unstable in a specific cell type. Based on the stem cell/progenitor cell concept of tumor development (17), progenitor cells of the luminal epithelial prostate cells, in which expression of the prostate-specific fusion partners is expected to be activated, are interesting candidates. In this regard, we could confirm *TMPRSS2* induction during development of the mouse prostate.[4] Unfortunately, there are no appropriate mouse counterparts of *KLK2* and *CANT1* (exon 1a).

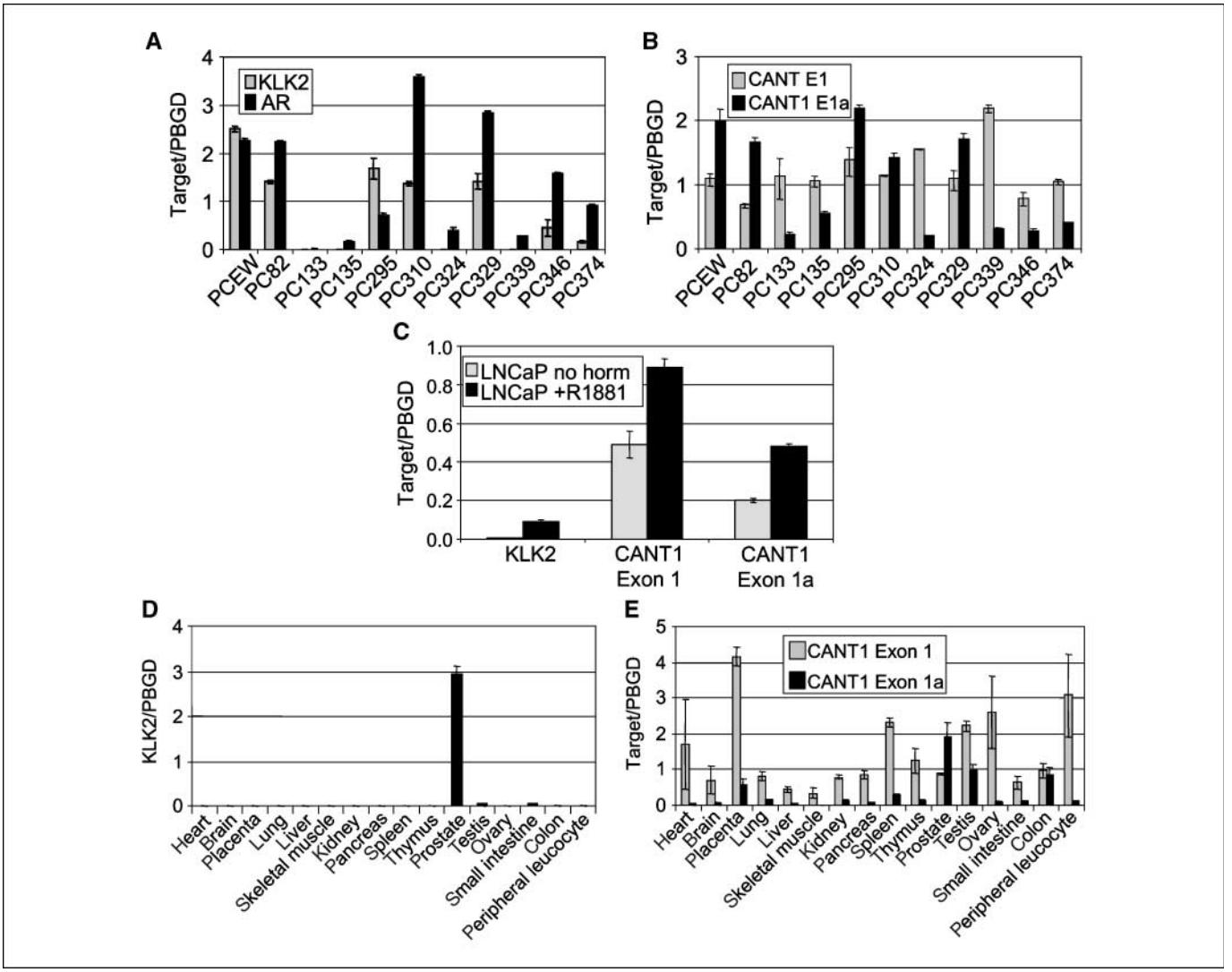

Figure 3. Characteristics of *KLK2* and *CANT1* mRNA expression. Expression of (*A*) *KLK2* and *AR* mRNA, and (*B*) transcripts starting in *CANT1* exon 1 and *CANT1* exon 1a, respectively, in 11 human prostate cancer xenografts. *PBGD* expression is used as control. *C,* androgen-regulated expression of *KLK2* and *CANT1* transcripts assessed by QPCR analysis in LNCaP cells grown in the absence or in the presence of the synthetic androgen R1881 (1 nmol/L). Expression relative to *PBGD* is presented. Tissue-specific expression of (*D*) *KLK2* and (*E*) *CANT1* transcripts. Tissue-specific expression was tested on a cDNA panel from 16 different normal tissues by QPCR analysis and is shown relative to *PBGD* expression. *Columns,* mean calculated from two independent experiments; *bars,* SD.

A second process that might play a role in a common mechanism of gene fusion includes the nuclear compartmentalization of gene expression. It might be proposed that prostate-specific genes are all expressed in a limited number of nuclear compartments, or so-called transcription factories (18, 19). Possibly, *ETV4* is expressed in the same compartments. In accordance with the latter hypothesis, it has recently been shown that *IGH* and *cMYC*, which are fusion partners in Burkitt lymphoma and plasmacytoma, are preferentially expressed in the same nuclear regions at *in vitro* B cell stimulation (20).

Taken together, the results from this study show that *ETV4* has multiple fusion partners and highlight the importance of meticulous examination of gene fusions. Importantly, key determinants of *ETV4* fusion partners are not chromosomal location or expression level, but their androgen-regulated and prostate-specific expression pattern.

Acknowledgments

Received 1/16/2008; revised 3/5/2008; accepted 3/6/2008.

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 Theo van der Kwast and Arno van Leenders for pathology, Chris Bangma and Wilma Teubel for clinical samples, Wytske van Weerden for xenograft tissues, and Anieta Siewerts for RNA isolation of clinical samples.

References

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. CA Cancer J Clin 2007;57:43–66.
2. van der Kwast TH, Schalken J, Ruizeveld de Winter JA, et al. Androgen receptors in endocrine-therapy-resistant human prostate cancer. Int J Cancer 1991;48:189–93.
3. Visakorpi T, Hyytinen E, Koivisto P, et al. *In vivo* amplification of the androgen receptor gene and progression of human prostate cancer. Nat Genet 1995;9:401–6.
4. Cher ML, Bova GS, Moore DH, et al. Genetic alterations in untreated metastases and androgen-independent prostate cancer detected by comparative genomic hybridization and allelotyping. Cancer Res 1996;56:3091–102.
5. Visakorpi T, Kallioniemi AH, Syvanen AC, et al. Genetic changes in primary and recurrent prostate cancer by comparative genomic hybridization. Cancer Res 1995;55:342–7.
6. Lin B, Ferguson C, White JT, et al. Prostate-localized and androgen-regulated expression of the membrane-bound serine protease TMPRSS2. Cancer Res 1999;59:4180–4.
7. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–8.
8. Mehra R, Tomlins SA, Shen R, et al. Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. Mod Pathol 2007;20:538–44.
9. Cerveira N, Ribeiro FR, Peixoto A, et al. TMPRSS2-ERG gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. Neoplasia 2006;8:826–32.
10. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM, Trapman J. TMPRSS2: ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. Cancer Res 2006;66:10658–63.
11. Tomlins SA, Laxman B, Dhanasekaran SM, et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 2007;448:595–9.
12. Tomlins SA, Mehra R, Rhodes DR, et al. TMPRSS2: ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res 2006;66:3396–400.
13. Chapdelaine P, Paradis G, Tremblay RR, Dube JY. High level of expression in the prostate of a human glandular kallikrein mRNA related to prostate-specific antigen. FEBS Lett 1988;236:205–8.
14. Riegman PH, Vlietstra RJ, van der Korput HA, Romijn JC, Trapman J. Identification and androgen-regulated expression of two major human glandular kallikrein-1 (hGK-1) mRNA species. Mol Cell Endocrinol 1991;76:181–90.
15. van Weerden WM, de Ridder CM, Verdaasdonk CL, et al. Development of seven new human prostate tumor xenograft models and their histopathological characterization. Am J Pathol 1996;149:1055–62.
16. Shaw CJ, Lupski JR. Implications of human genome architecture for rearrangement-based disorders: the genomic basis of disease. Hum Mol Genet 2004;13:57–64.
17. Lawson DA, Witte ON. Stem cells in prostate cancer initiation and progression. J Clin Invest 2007;117:2044–50.
18. Fraser P, Bickmore W. Nuclear organization of the genome and the potential for gene regulation. Nature 2007;447:413–7.
19. Misteli T. Beyond the sequence: cellular organization of genome function. Cell 2007;128:787–800.
20. Osborne CS, Chakalova L, Mitchell JA, et al. Myc dynamically and preferentially relocates to a transcription factory occupied by Igh. PLoS Biol 2007;5:1763–72.

Two Unique Novel Prostate-Specific and Androgen-Regulated Fusion Partners of *ETV4* in Prostate Cancer

Karin G. Hermans, Anke A. Bressers, Hetty A. van der Korput, et al.

Cancer Res 2008;68:3094-3098.