NEOPLASIA
www.neoplasia.com

Volume 15 Number 7 July 2013 pp. 720–726 720

Novel 5′ Fusion Partners of *ETV1* and *ETV4* in Prostate Cancer[1,2]

João D. Barros-Silva[*,†,‡,§,3], Paula Paulo[*,†,‡,§,3], Anne Cathrine Bakken[‡,§], Nuno Cerveira[*], Marthe Løvf[‡,§], Rui Henrique[¶,#,], Carmen Jerónimo[*,#,**], Ragnhild A. Lothe[‡,§], Rolf Inge Skotheim[‡,§] and Manuel R. Teixeira[*,†,§,**]**

*Department of Genetics, Portuguese Oncology Institute, Porto, Portugal; †Group of Cancer Genetics, Research Centre of the Portuguese Oncology Institute, Porto, Portugal; ‡Department of Cancer Prevention, Institute for Cancer Research, The Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway; §Centre for Cancer Biomedicine, Faculty of Medicine, University of Oslo, Oslo, Norway; ¶Department of Pathology, Portuguese Oncology Institute, Porto, Portugal; #Group of Cancer Epigenetics, Research Centre of the Portuguese Oncology Institute, Porto, Portugal; **Department of Pathology and Molecular Immunology, Institute of Biomedical Sciences Abel Salazar, University of Porto, Porto, Portugal

Abstract

Gene fusions involving the erythroblast transformation-specific (ETS) transcription factors *ERG*, *ETV1*, *ETV4*, *ETV5*, and *FLI1* are a common feature of prostate carcinomas (PCas). The most common upstream fusion partner described is the androgen-regulated prostate-specific gene *TMPRSS2*, most frequently with *ERG*, but additional 5′ fusion partners have been described. We performed 5′ rapid amplification of cDNA ends in 18 PCas with *ETV1*, *ETV4*, or *ETV5* outlier expression to identify the 5′ fusion partners. We also evaluated the exon-level expression profile of these ETS genes in 14 cases. We identified and confirmed by fluorescent *in situ* hybridization (FISH) and reverse transcription-polymerase chain reaction the two novel chimeric genes *OR51E2-ETV1* and *UBTF-ETV4* in two PCas. *OR51E2* encodes a G-protein–coupled receptor that is overexpressed in PCas, whereas *UBTF* is a ubiquitously expressed gene encoding an HMG-box DNA-binding protein involved in ribosome biogenesis. We additionally describe two novel gene fusion combinations of previously described genes, namely, *SLC45A3-ETV4* and *HERVK17-ETV4*. Finally, we found one PCa with *TMPRSS2-ETV1*, one with *C15orf21-ETV1*, one with *EST14-ETV1*, and two with *14q133-q21.1–ETV1*. In nine PCas (eight *ETV1* and one *ETV5*), exhibiting ETS outlier expression and genomic rearrangement detected by FISH, no 5′ fusion partner was found. Our findings contribute significantly to characterize the heterogeneous group of ETS gene fusions and indicate that all genes described as 5′ fusion partners with one ETS gene can most likely be rearranged with any of the other ETS genes involved in prostate carcinogenesis.

Neoplasia (2013) 15, 720–726

Abbreviations: PCa, prostate carcinoma; FISH, fluorescence *in situ* hybridization; 5′RACE, 5′ rapid amplification of cDNA ends

Address all correspondence to: Manuel R. Teixeira, MD, PhD, Department of Genetics, Portuguese Oncology Institute, Rua Dr. António Bernardino Almeida, 4200-072 Porto, Portugal. E-mail: manuel.teixeira@ipoporto.min-saude.pt

[1]This work was supported by research grants Pest-OE/SAU/UI0776/2011 and PTDC/SAU/OBD/70543/2006 awarded by Fundação para a Ciência e a Tecnologia (FCT). J.D.B.-S. (SFRH/BD/46574/2008) and P.P. (SFRH/BD/27669/2006) are or were research fellows from FCT. The authors have no conflict of interest to declare.
[2]This article refers to supplementary materials, which are designated by Tables W1 to W3 and Figures W1 and W2 and are available online at www.neoplasia.com.
[3]These authors contributed equally to this work and should be regarded as joint first authors.
Received 16 January 2013; Revised 12 April 2013; Accepted 15 April 2013

DOI 10.1593/neo.13232

Introduction

Gene fusions involving the erythroblast transformation-specific (ETS) transcription factor family of genes are a recurrent feature of prostate adenocarcinomas (PCas). These gene aberrations, caused by chromosomal structural abnormalities, originate fusion transcripts that lead to overexpression of N-truncated ETS proteins or, more rarely, to full-length ETS proteins or chimeric fusion proteins [1–3]. Fusion of the androgen-regulated promoter region of the *TMPRSS2* gene with *ERG* is the most common ETS rearrangement, being present in about 50% of PCa and in 20% of high-grade prostatic intraepithelial neoplasia lesions [1,4]. Other rarer fusion events can occur involving the PEA3 subfamily of ETS members, namely, *ETV1*, *ETV4*, and *ETV5* [1,5,6] or the ERG subfamily member *FLI1* [3].

Besides the prostate-specific and androgen-induced *TMPRSS2*, several ETS fusion partners have been described, namely, *HERPUD1*, *NDRG1*, *SLC45A3*, *ACSL3*, *HERV-K_22*, *HERVK17*, *CANT1*, *DDX5*, *KLK2*, *FOXP1*, *EST14*, *HNRPA2B1*, *C15orf21*, and the chromosomal region 14q13.3-14q21.1 [2,7,8], presenting heterogeneous tissue specificities and androgen responsiveness. Fusion partners like the *SLC45A3* gene or the endogenous retroviral element (*HERVK17*) display similar tissue specificity as *TMPRSS2* and are equally androgen-induced. Contrarily, the fusion partner *C15orf21*, despite being overexpressed in PCas, is repressed by androgens. However, there are ubiquitously expressed 5′ fusion partners, such as the *HNRPA2B1* gene, displaying no evidence of androgen regulation. Finally, ETS family genes may be rearranged with prostate-specific enhancers in chromosomal regions such as 14q13.3-14q21.1.

We have previously performed a comprehensive characterization of ETS rearrangements on a series of 200 clinically localized PCas and found rearrangements involving *ERG*, *ETV1*, *ETV4*, *ETV5*, and *FLI1* in 52%, 7%, 1.5%, 0.5%, and 0.5%, respectively [3]. In the present work, we focused on the 18 PCas that showed outlier expression levels of *ETV1*, *ETV4*, or *ETV5* and a genomic rearrangement of the corresponding ETS locus. The combined use of exon-level expression profiles from exon microarrays, 5′ rapid amplification of cDNA ends (5′RACE), and fluorescence *in situ* hybridization (FISH) with bacterial artificial chromosome (BAC)–specific probes allowed us to identify novel 5′ fusion partners for *ETV1* and *ETV4*, as well as to describe novel combinations of genes known to be involved in PCa gene fusions.

Materials and Methods

PCa Samples

We studied a set of 18 PCas with outlier mRNA expression levels of *ETV1* (n = 14), *ETV4* (n = 3), and *ETV5* (n = 1) and with a genomic rearrangement previously demonstrated by FISH but with yet unknown fusion partners [3]. These samples were selected from a cohort of 200 patients with clinically localized PCa consecutively diagnosed and treated with radical prostatectomy that were previously typed for ETS rearrangements [3]. This study was approved by the Institutional Review Board, and informed consent was obtained from all subjects.

Gene Expression Microarrays

RNA was extracted from tissue samples using TRIzol (Invitrogen by Life Technologies, Carlsbad, CA), as previously described [3], and 1 μg of RNA was processed into cDNA and hybridized to Affymetrix GeneChip Human Exon 1.0 ST arrays, following the manufacturer's recommendations. The Affymetrix Expression Console v1.1 software was used to obtain exon-level robust multi-array average (RMA)-normalized expression values for the core probe sets only.

5′Rapid Amplification of cDNA Ends

The 5′RACE was performed using the SMARTer RACE cDNA amplification kit and protocol (Clontech Laboratories, Inc, Saint-Germain-en-Laye, France). Briefly, first-strand cDNA was reverse transcribed from 1 μg of total RNA using the SMARTScribe Reverse Transcriptase with the 5′RACE cDNA synthesis primer (5′-CDS) and the SMARTer IIA oligo from the kit. An aliquot of the cDNA was then amplified using a forward gene-specific primer (GSP) and a universal primer mix. Polymerase chain reaction (PCR) conditions used were as described by the manufacturer. Nested PCRs using the nested universal primer as the reverse primer and a nested gene-specific primer (NGSP) were performed to increase the specificity and product yields of 5′RACE. Primers used on 5′RACE and nested PCR are listed in Table W1. Nested PCR products were analyzed on a 2% agarose gel (SeaKem LE Agarose; Lonza, Basel, Switzerland) and the bands were purified using the QIAquick Gel Extraction Kit (Qiagen, Hilden, Germany).

Cloning and Sequencing

Purified 5′ nested race PCR products were cloned into pCR4-TOPO plasmids using the TOPO TA Cloning Kit for sequencing (Invitrogen). Colonies were picked and the plasmids were purified using the Qiagen Plasmid Miniprep Kit (Qiagen) and subsequently sequenced using the M13 forward and T3 primers using BigDye Terminator V3.1 sequencing chemistry on a 3730 DNA Analyzer (Applied Biosystems by Life Technologies, Foster City, CA) according to the manufacturer's recommendations. Genomic alignment of the resulting sequences was performed using BLAST (http://www.ncbi.nlm.nih.gov/blast/) and BLAT (http://genome.ucsc.edu/). All exons identified were numbered according to the longest matching transcripts of the Ensembl database (http://www.ensembl.org/).

Reverse Transcription–PCR

The reverse transcription (RT)–PCR assay for detection of prostate fusion transcripts was performed with the following primer combinations: TMPRSS2_F and ETV1_R (*TMPRSS2-ETV1*), C15orf21_S and ETV1_AS1 (*C15orf21-ETV1*), SLC45A3_S and ETV4_AS (*SLC45A3-ETV4*), UBTF-S and ETV4_AS (*UBTF-ETV4*), OR51E2_S and ETV1_AS2 (*OR51E2-ETV1*), EST14_S and ETV1_AS2 (*EST14-ETV1*), and HERVK17_S and ETV4_AS (*HERVK17-ETV4*; Table W2). PCRs were performed with the Qiagen OneStep RT-PCR Kit (Qiagen) according to the manufacturer's instructions. Amplified products were analyzed on a 2% agarose gel and further validated by sequencing.

Fluorescence In Situ *Hybridization*

To validate the 5′RACE findings, we performed FISH using BAC clones targeting the 5′ fusion partner and the ETS gene on tissue sections from paraffin blocks of the index tumor (Table W3). BAC clones were selected using the University of California, Santa Cruz (UCSC) Human Genome Browser and obtained from the BACPAC Resources Centre (Oakland, CA). BAC DNA was extracted, amplified, labeled, and prepared for hybridization as previously reported [9]. Adequate mapping and probe specificity of all BAC clones was confirmed by hybridization onto human metaphase spreads of normal

lymphocytes. An abnormal signal pattern was considered representative when present in a minimum of 50 morphologically intact, nonoverlapping nuclei.

In Silico *Analysis of Androgen Effect on Gene Expression*

To study the effect of androgen stimulus on *UBTF* and *OR51E2* expression, we used the expression data of the LNCaP cell line that can be accessed from the Gene Expression Omnibus (GSE32875) [10]. The normalized expression values of *OR51E2* and *UBTF* in four replicates of the LNCaP cell line treated with a synthetic androgen (R1881) and four replicates of the LNCaP cell line with no androgen stimulus were compared using the paired sample *t* test.

Results and Discussion

The starting point of this study was a subset of 18 PCas with evidence of rearrangement of *ETV1*, *ETV4*, and *ETV5* and unknown 5′ fusion partner from a larger cohort previously typed for ETS rearrangements [3]. Of these 18 cases, there was evidence for involvement of *ETV1* in 14 cases, of *ETV4* in 3 cases, and of *ETV5* in 1 case, in every instance presenting outlier expression and genomic rearrangement by FISH [3]. After analyzing the exon-level expression data of 51 PCas [11], including 14 of the 18 cases included in this report, we here show that in samples P262T, P274T, P451T, and P542T the exons of *ETV1* are differentially expressed between the 5′ and 3′ ends, pinpointing the breakpoint before exons 4, 7, 8, and 7, respectively (Figure 1). Differences in the exon expression profile of *ETV4* and *ETV5* are present in samples P522T and P261T, respectively, indicating breakpoints before exon 5 of *ETV4* and before exon 6 of *ETV5* (Figure 1). In the remaining cases, all exons of the rearranged ETS gene were similarly overexpressed. To unveil the 5′ fusion partners involved in these ETS gene fusions, we initially performed RT-PCR in all 18 PCas with primers targeting the most commonly involved 5′ fusion partners, namely, *TMPRSS2* and *SLC45A3*. The *TMPRSS2-ETV1* fusion transcript was detected in one of the 14 PCas with *ETV1* rearrangement (P500T), and the *SLC45A3-ETV4* fusion transcript was found in one of the three samples with *ETV4* rearrangement (P236T; Figure W1).

Given our recent findings of occasional heterogeneity of PCa gene fusions in different areas of the same tumor [3], all samples were further investigated by 5′RACE PCR. Sequencing of the 5′RACE PCR products of cases P451T and P522T unveiled *OR51E2* and *UBTF* as novel 5′ fusion partners of *ETV1* and *ETV4*, respectively. The *OR51E2-ETV1* transcript was composed by the untranslated exon 1 of *OR51E2* (ENST00000396950), with an additional untranslated 292-bp sequence downstream of exon 2 of *OR51E2*, fused to exon 8 of *ETV1* (Figure 2), which includes an in-frame ATG start codon (amino acid 132). As previously described [12,13], gene fusions involving noncoding regions of the 5′ partner with exon 8 of *ETV1* results in an N-truncated ETV1 protein with 345 amino acids. An alternative transcript fusing exon 9 of *ETV1* with the same *OR51E2* sequences was also found, with the predicted resulting protein starting at an internal ATG in exon 9 of *ETV1* (amino acid 203) and consisting of 274 amino acids. The presence of both *OR51E2-ETV1* fusion transcripts was confirmed by RT-PCR. The *OR51E2* and *ETV1* genes map to chromosome bands 11p15 and 7p21, respectively, and have the same transcriptional orientation, which indicates that this novel fusion gene was originated by a chromosomal translocation. The *OR51E2* (alias *PSGR*) is named as olfactory receptor, family 51, subfamily E, member 2, and encodes a member of a large family of G-protein–coupled receptors arising from single coding exon

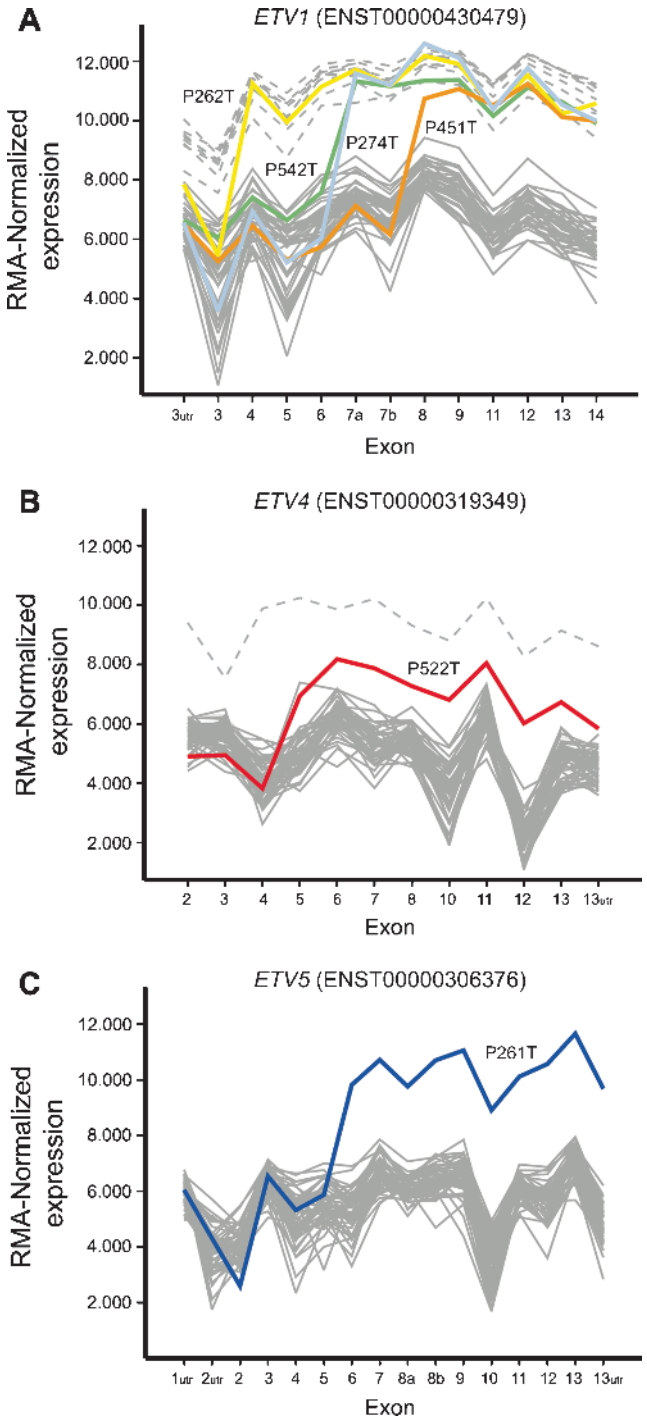

Figure 1. Exon-level expression profiles. RMA-normalized exon-level expression profile of *ETV1*, *ETV4*, and *ETV5* in 51 PCas obtained using Affymetrix GeneChip Human Exon 1.0 ST arrays. Differential expression of exons at the 5′ and 3′ ends of *ETV1* (A) was found in four samples (P262T, P275T, P451T, and P542T), indicating breakpoint before exons 4, 7, 8, and 7, respectively, whereas samples P522T and P261T displayed differential expression of exons at the 5′ and 3′ ends of *ETV4* (B) and *ETV5* (C), respectively, with a breakpoint before exon 5 of *ETV4* and exon 6 of *ETV5*. Gray lines represent samples with normal ETS expression, whereas gray dashed lines represent samples with overexpression of the whole transcript. For all genes, exons are numbered according to the indicated Ensembl transcript identifiers. The array included two probe sets targeting proximal and distal regions of exon 7 of *ETV1* and exon 8 of *ETV5* depicted in the plot as 7a and 7b and 8a and 8b, respectively.

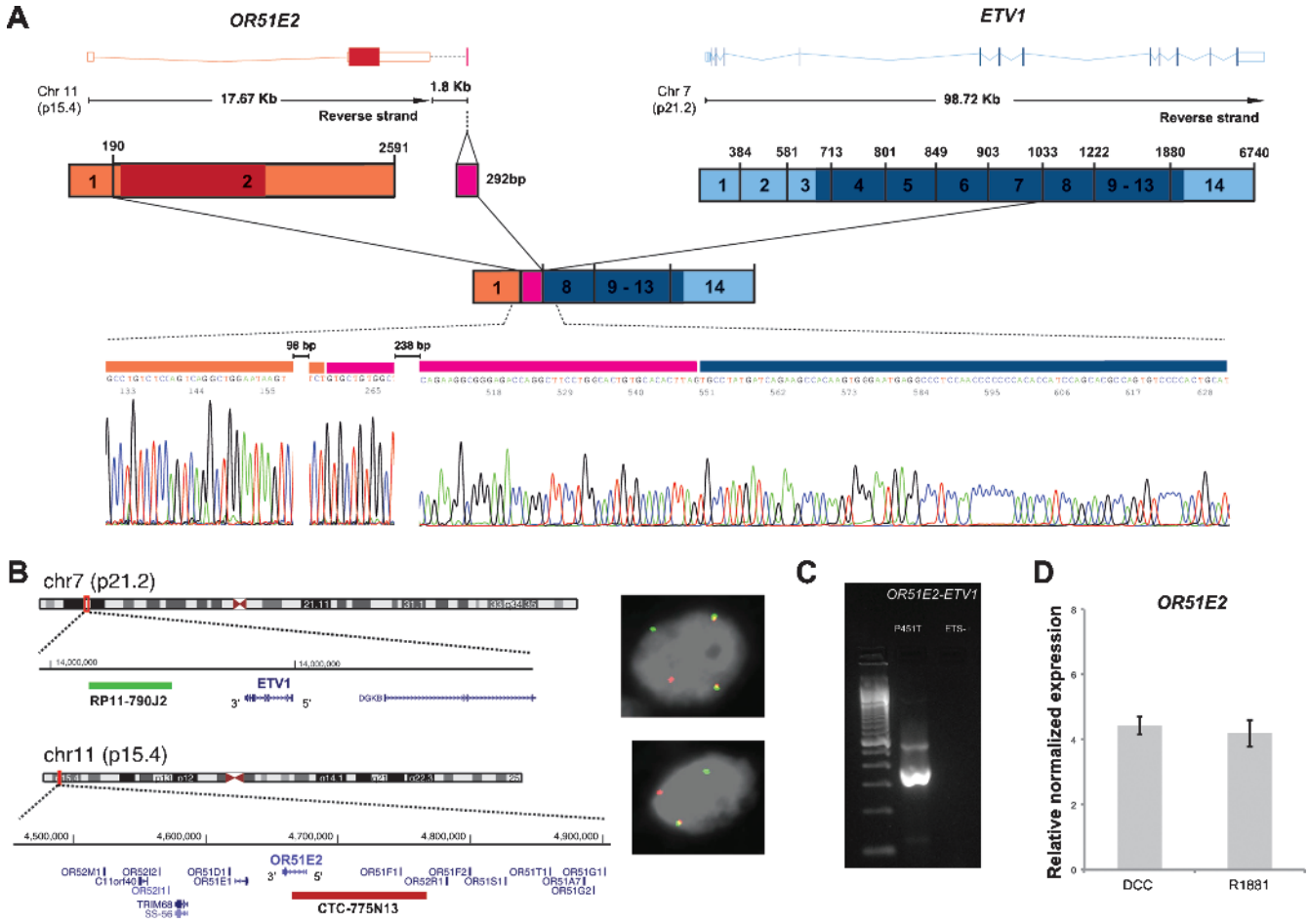

Figure 2. Characterization of the gene fusion involving *OR51E2* and *ETV1* in the clinical prostate cancer P451T with *ETV1* outlier expression and with an *ETV1* genomic rearrangement previously demonstrated by FISH. (A) Fusion transcript of the untranslated exon 1 of *OR51E2* and a downstream sequence of 292 bp of *OR51E2* with exon 8 of *ETV1* detected by sequencing of the 5′RACE PCR products using a reverse primer on exon 10 of *ETV1*. The upper representation depicts wild-type transcripts of *OR51E2* and *ETV1*, indicating its length and chromosomal localization. Exons are represented by boxes (colored if translated) and introns by horizontal lines. The scheme below represents the wild-type cDNA, where untranslated regions are in lighter shades. (B) Interphase FISH on formalin-fixed, paraffin-embedded tissue of the PCa P451T. Co-localization of probes for *OR51E2* (red) and *ETV1* (green) confirms genomic *OR51E2-ETV1* fusion in the cancer cells. (C) RT-PCR assay for detection of *OR51E2-ETV1* fusion transcript in the PCa P451T and in an ETS-negative PCa using a forward primer in the 292-bp sequence downstream of *OR51E2* and a reverse primer in exon 9 of *ETV1*. (D) Expression of *OR51E2* is relatively unchanged in LNCaP cell lines treated with an androgen analog (R1881) *versus* cells grown in dextran-coated charcoal culture medium as a control. The expression data were obtained from the Gene Expression Omnibus (GSE32875).

genes. It has been described by several authors to have a prostate-specific expression pattern and has also been found overexpressed in PCa [14] (Figure W2). The *OR51E2-ETV1* fusion retains the tissue specific regulatory activity present in exon 1 of *OR51E2*, one of the two described *OR51E2* promoters [15]. A recent study of the gene expression profile of benign and malignant prostate tissues from three radical prostatectomies and from three prostate biopsies of surgically castrated patients showed higher *OR51E2* expression in benign/cancer tissues compared to castrated prostate tissues [16], suggesting androgen regulation. However, it has been shown that the two described *OR51E2* promoters are regulated by interleukin-6 but not by the synthetic androgen R1881 [15]. Furthermore, we performed *in silico* analysis of public microarray data of R1881-treated LNCaP cells and found no alteration of *OR51E2* expression after androgenic stimulation (Figure 2; microarray data originally published in [10]).

The second novel fusion transcript *UBTF-ETV4* is composed by the untranslated exon 1 and exon 2 of *UBTF* (ENST00000393606) fused to exon 5 of *ETV4* (Figure 3). Sequencing also revealed a partial deletion of exon 6 and exon 7 of *ETV4* consisting in the loss of 29 and 148 bp, respectively. The presence of the *UBTF-ETV4* fusion transcript was confirmed by RT-PCR. Both genes are located in the same chromosome band, 17q21, and are about 700 kb apart (Figure 3). Break apart probes flanking *ETV4* showed deletion of the probe targeting the 5′ *ETV4* region, suggesting a cryptic deletion in chromosome 17 between the two genes as the mechanism of *UBTF-ETV4* fusion. Exon 2 of *UBTF* contains the ATG start codon and is in-frame with exon 5 of *ETV4*, so this gene fusion presumably originates a chimeric protein containing 19 amino acids encoded by exon 2 of *UBTF* and an ETV4 counterpart lacking the first 67 N-terminal amino acids as well as additional 59 amino acids resulting from the partial deletion on exons 6 and 7 of *ETV4*. Interestingly, a previously reported gene fusion in PCa involving exon 5 of *ETV4* also translates into a chimeric protein [17]. Although relatively rare, there is evidence of other ETS fusion transcripts encoding chimeric proteins, namely,

TMPRSS2 fused with *ERG*, *ETV1*, *ETV4*, and *ETV5*, and the fusion transcripts *FOXP1-ETV1*, *EST14-ETV1*, *HNRPA2B1-ETV1*, and *DDX5-ETV4* [2]. The *UBTF* (upstream binding transcription factor, RNA polymerase I) gene encodes a member of the HMG-box DNA-binding protein family [18,19], which plays a critical role in ribosomal RNA (rRNA) transcription as a key component of the preinitiation complex, mediating the recruitment of RNA polymerase I to ribosomal DNA promoter regions. Synthesis of rRNA by Pol I, which drives ribosome biogenesis, is an important determinant of the cellular growth response [20]. This gene is ubiquitously expressed in different tissues and does not display an increased expression in PCa (Figure W2). Moreover, we here show by *in silico* analysis of public microarray data that the expression of *UBTF* is not altered by androgen stimulus in the LNCaP cell line (Figure 3; microarray data originally published in [10]). Interestingly, it has been shown that the oncoprotein MYC directly influences Pol I transcription of rRNA genes [20] and that it positively regulates UBTF expression [21].

Besides *UBTF*, we here describe two additional novel *ETV4* fusion partners by 5′RACE PCR (Table 1). In sample P352T, we found a fusion transcript of the prostate-specific androgen-repressed endogenous retroviral element, *HERVK17*, fused to exon 5 of *ETV4* (Figures 4 and W1). *HERVK17* has previously been described as a fusion partner for *ETV1* [13]. Confirming the earlier findings by RT-PCR, sequencing of the 5′RACE PCR product of sample P236T showed an *SLC45A3-ETV4* fusion transcript, in which exon 1 of *SLC45A3* (ENST00000367145) is fused in-frame with exon 3 of *ETV4* (Figures 4 and W1). *ETV4* rearrangements involving exon 3 have been previously described [5]. Interestingly, this tumor also harbors an *SLC45A3-ERG* fusion in a different area of the same cancerous focus [3], evidencing intratumor genetic heterogeneity in the form of two different ETS genes fused with the same 5′ fusion partner, which most likely resulted from collision of initially separated neoplasms/foci. With the findings we here report, *SLC45A3* has now been found fused with all five ETS members known to be involved in

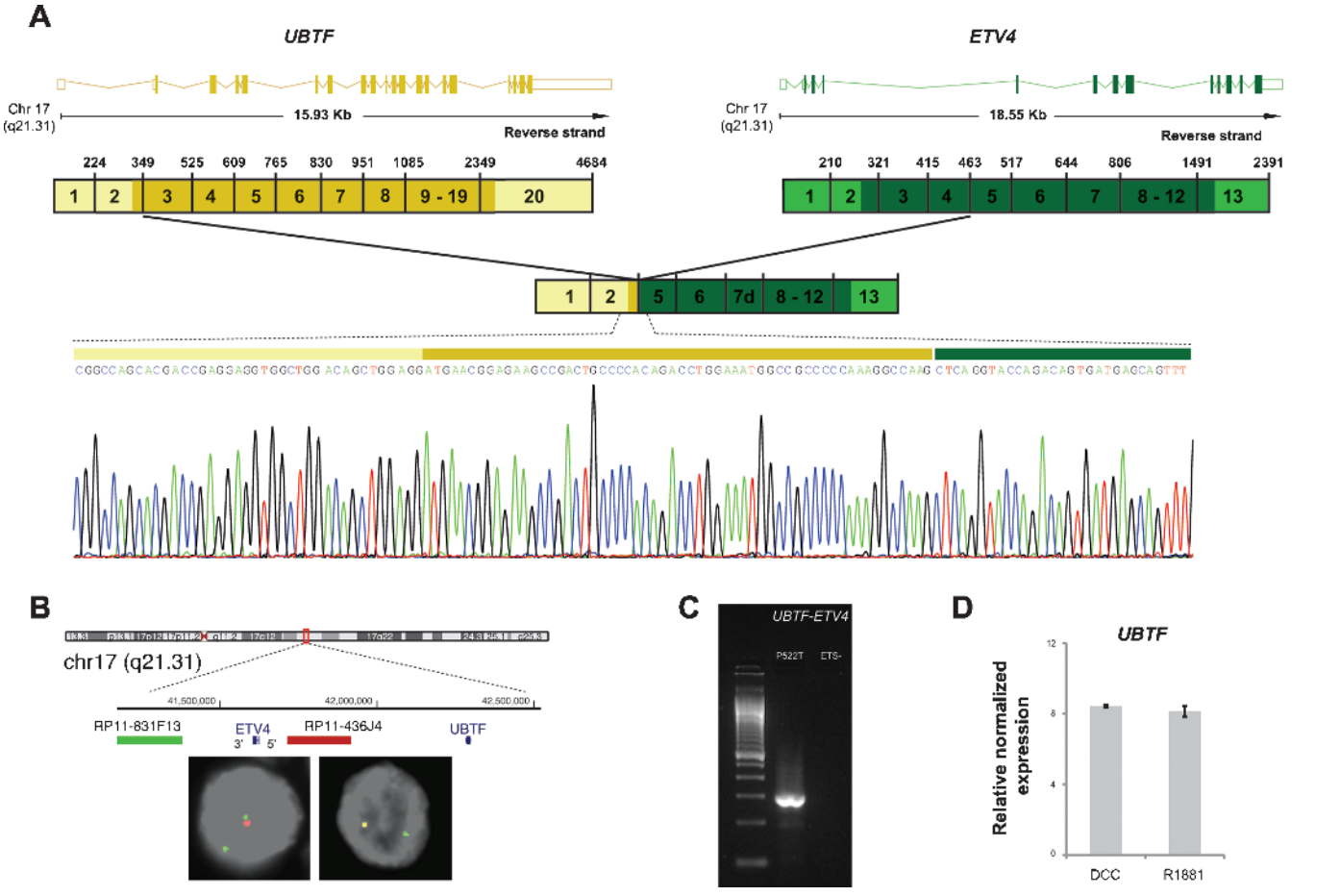

Figure 3. Characterization of the gene fusion involving *UBTF* and *ETV4* in the clinical prostate cancer P522T with *ETV4* outlier expression and with an *ETV4* genomic rearrangement previously demonstrated by FISH. (A) Fusion transcript of the untranslated exon 1 and exon 2 of *UBTF* with exon 5 of *ETV4* detected by sequencing of the 5′RACE PCR products using a reverse primer on exon 10 of *ETV4*. The upper representation depicts wild-type transcripts of *UBTF* and *ETV4* indicating their length and chromosomal localization. Exons are represented by boxes (colored if translated) and introns by horizontal lines. The scheme below represents the wild-type cDNA, where untranslated regions are in lighter shades. (B) Interphase FISH on formalin-fixed, paraffin-embedded tissue of the PCa P522T. Deletion of the probe spanning the chromosome region between *UBTF* and *ETV4* (red) indicates a cryptic deletion at chromosome 17 as the genomic mechanism for *UBTF-ETV4* gene fusion. (C) RT-PCR assay for detection of *UBTF-ETV4* fusion transcript in the PCa P522T and in an ETS-negative PCa using a forward primer in exon 2 of *UBTF* and a reverse primer in exon 6 of *ETV4*. (D) Expression of *UBTF* is relatively unchanged in LNCaP cell lines treated with an androgen analog (R1881) *versus* cells grown in dextran-coated charcoal culture medium as a control. The expression data were obtained from the Gene Expression Omnibus (GSE32875).

Table 1. Summary of the Experimental Findings in 18 Prostate Tumors with *ETV1*, *ETV4*, or *ETV5* Rearrangements.

Sample	TLDA Outlier Expression	FISH	Exon Array	RT-PCR (TMPRSS2/ SLC45A3-ETV1)	RT-PCR (TMPRSS2/ SLC45A3-ETV4)	RT-PCR (TMPRSS2/ SLC45A3-ETV5)	FISH (14q13.3-q21.1–ETV1)	5′RACE
P262T	ETV1	ETV1+	D	neg.	–	–	*14q13.3-q21.1–ETV1*	neg.
P272T	ETV1	ETV1+	W	neg.	–	–	neg.	neg.
P305T	ETV1	ETV1+	W	neg.	–	–	neg.	neg.
P344T	ETV1	ETV1+	W	neg.	–	–	n. a.	neg.
P456T	ETV1	ETV1+	W	neg.	–	–	*14q13.3-q21.1–ETV1*	neg.
P488T	ETV1	ETV1+	W	neg.	–	–	neg.	neg.
P499T	ETV1	ETV1+	W	neg.	–	–	neg.	neg.
P525T	ETV1	ETV1+	W	neg.	–	–	–	*C15orf21-ETV1*
P542T	ETV1	ETV1+	D	neg.	–	–	neg.	neg.
P274T	ETV1	ETV1+	D	neg.	–	–	–	*EST14-ETV1*
P451T	ETV1	ETV1+	D	neg.	–	–	–	*OR51E2-ETV1*
P271T	ETV1	ETV1+	n. a.	neg.	–	–	neg.	neg.
P498T	ETV1	ETV1+	n. a.	neg.	–	–	neg.	neg.
P500T	ETV1	ETV1+	n. a.	*TMPRSS2-ETV1*	–	–	–	neg.
P236T	ETV4	ETV4+	D	–	*SLC45A3-ETV4*	–	–	*SLC45A3-ETV4*
P352T	ETV4	ETV4+	n. a.	–	neg.	–	–	*HERVK17-ETV4*
P522T	ETV4	ETV4+	D	–	neg.	–	–	*UBTF-ETV4*
P262T	ETV5	ETV5+	D	–	–	neg.	–	neg.

TLDA indicates TaqMan low-density array; W, whole gene overexpression; D, differential exon expression; neg., negative; n. a., not analyzable.

prostate carcinogenesis, whereas *TMPRSS2* has been described as 5′ fusion partner with four of these (all except *FLI1*).

Finally, sequencing of the 5′RACE PCR fragments showed gene fusions that have previously been described in two additional cases. Sample P274T showed the fusion of *EST14* with exon 7 of *ETV1*, as previously described by Hermans and co-workers [13], whereas in sample P525T we found that the untranslated exon 1 of *C15orf21* (ENST00000409454) is fused with exon 3 of *ETV1*, thus giving rise to a full-length ETV1 protein (Figures 4 and W1). The latter is different from the previously described *C15orf21-ETV1* fusion [22], where exon 2 of *C15orf21* was fused with exon 8 of *ETV1*. There is evidence in the literature of fusion transcripts involving ETS genes

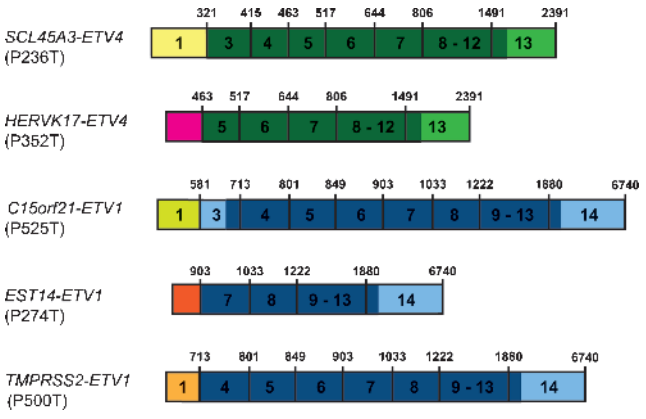

Figure 4. Schematic representation of gene fusions involving known ETS 5′fusion partners detected in PCas using 5′RACE, including the two novel *ETV4* fusion partners. *ETV1* exons are represented in blue and *ETV4* exons in green, and noncoding regions are in lighter shades. From top to bottom: Exon 1 of *SLC45A3* fused with exon 3 of *ETV4* (prostate cancer P236T); transcript of the endogenous retroviral element *HERVK17* fused with exon 4 of *ETV4* (prostate cancer P352T); exon 1 of *C15orf21* fused with exon 3 of *ETV1* (prostate cancer P525T); exon 1 of *EST14* fused with exon 7 of *ETV1* (prostate cancer P274T); exon 1 of *TMPRSS2* fused with exon 4 of *ETV1* (prostate cancer P500T).

that include all their coding exons. Helgeson and co-workers described a prostate tumor with four fusion transcripts of *TMPRSS2-ETV5*, each including all coding exons of *ETV5* [6]. More recently, Gasi and colleagues described a prostate tumor harboring *EST14-ETV1* fusion transcripts including all coding exons of *ETV1* and overexpressing full-length ETV1 [23].

Contrarily to case P236T, in which the initial RT-PCR finding of the *SLC45A3-ETV4* gene fusion was later confirmed by 5′RACE PCR, in one PCa the *TMPRSS2-ETV1* fusion transcript was detected by RT-PCR but not by 5′RACE PCR, which may be due to technical difficulties of the 5′RACE methodology. The 11 cases where no 5′ fusion partner was found using both RT-PCR and 5′RACE included 10 samples with evidence of *ETV1* rearrangement and 1 with evidence of *ETV5* rearrangement. The exon-level expression profile of the prostate tumor P261T displayed differential expression of *ETV5* exons 6 to13 compared to exons 1 to 5 (Figure 1); however, no 5′ fusion partner was found either by RT-PCR or 5′RACE. Moreover, this sample was tested using a FISH approach with probes flanking *ETV5* combined with two additional probes targeting the 5′ region of the two known *ETV5* fusion partners (*TMPRSS2* and *SLC45A3*), but no evidence of involvement of these genes was found. Further FISH analysis lead to the identification of two cases with 14q13.3-q21.1–*ETV1* rearrangement. Regarding the eight *ETV1* cases where no 5′ fusion partner was found, exon-level expression data were available in six cases, of which one displayed differential exon-level expression and the remaining five showed overexpression of full-length *ETV1* transcripts. This is in agreement with a recent study of Gasi et al. [23], where, of five prostate tumors with overexpression of full-length *ETV1* transcripts and genomic rearrangement of *ETV1* determined by FISH, only in one case the 5′RACE was efficient in identifying a fusion transcript, which turned out to be *EST14-ETV1*.

Recent studies have brought new insights on the mechanisms underlying chromosomal rearrangements in human cancer. Reports describing androgen receptor–driven chromatin looping [24] and ligand-bound androgen receptor recruitment of enzymes capable of inducing double-strand breaks [25] suggest that androgen receptor may play an important role in gene fusion formation in PCa.

Another critical aspect of chromosome rearrangements is the spatial proximity of the partners involved. Growing evidence indicates that the formation of tumor translocations is partially affected by higher order genomic organization in nuclear territory [26]. In light of this, future work may show that the new fusion partners here described occupy the same nuclear domain.

In summary, this study contributes significantly to characterize the pattern of fusion genes in PCa, as we report two novel 5′ fusion partners (*OR51E2-ETV1* and *UBTF-ETV4*) of ETS rearrangements, as well as two novel gene fusion combinations involving previously described genes (*SLC45A3-ETV4* and *HERVK17-ETV4*). Our findings suggests that not only may there be more 5′ partner genes yet to be identified in PCa but also that all the genes described as 5′ fusion partners with one ETS gene can most likely be rearranged with any of the other ETS genes known to be involved in PCa.

References

[1] Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, et al. (2005). Recurrent fusion of *TMPRSS2* and ETS transcription factor genes in prostate cancer. *Science* **310**, 644–648.

[2] Clark JP and Cooper CS (2009). ETS gene fusions in prostate cancer. *Nat Rev Urol* **6**, 429–439.

[3] Paulo P, Barros-Silva JD, Ribeiro FR, Ramalho-Carvalho J, Jeronimo C, Henrique R, Lind GE, Skotheim RI, Lothe RA, and Teixeira MR (2012). FLI1 is a novel ETS transcription factor involved in gene fusions in prostate cancer. *Genes Chromosomes Cancer* **51**, 240–249.

[4] Cerveira N, Ribeiro FR, Peixoto A, Costa V, Henrique R, Jeronimo C, and Teixeira MR (2006). TMPRSS2-ERG gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. *Neoplasia* **8**, 826–832.

[5] Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE, Cao X, Wei JT, Rubin MA, Shah RB, et al. (2006). *TMPRSS2:ETV4* gene fusions define a third molecular subtype of prostate cancer. *Cancer Res* **66**, 3396–3400.

[6] Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q, Prensner JR, Cao X, Singla N, Montie JE, Varambally S, et al. (2008). Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. *Cancer Res* **68**, 73–80.

[7] Maher CA, Palanisamy N, Brenner JC, Cao X, Kalyana-Sundaram S, Luo S, Khrebtukova I, Barrette TR, Grasso C, Yu J, et al. (2009). Chimeric transcript discovery by paired-end transcriptome sequencing. *Proc Natl Acad Sci USA* **106**, 12353–12358.

[8] Pflueger D, Rickman DS, Sboner A, Perner S, LaFargue CJ, Svensson MA, Moss BJ, Kitabayashi N, Pan Y, de la Taille A, et al. (2009). N-myc downstream regulated gene 1 (NDRG1) is fused to ERG in prostate cancer. *Neoplasia* **11**, 804–811.

[9] Barros-Silva JD, Ribeiro FR, Rodrigues A, Cruz R, Martins AT, Jeronimo C, Henrique R, and Teixeira MR (2011). Relative 8q gain predicts disease-specific survival irrespective of the *TMPRSS2-ERG* fusion status in diagnostic biopsies of prostate cancer. *Genes Chromosomes Cancer* **50**, 662–671.

[10] Rajan P, Dalgliesh C, Carling PJ, Buist T, Zhang C, Grellscheid SN, Armstrong K, Stockley J, Simillion C, Gaughan L, et al. (2011). Identification of novel androgen-regulated pathways and mRNA isoforms through genome-wide exon-specific profiling of the LNCaP transcriptome. *PLoS One* **6**, e29088.

[11] Paulo P, Ribeiro FR, Santos J, Mesquita D, Almeida M, Barros-Silva JD, Itkonen H, Henrique R, Jeronimo C, Sveen A, et al. (2012). Molecular subtyping of primary prostate cancer reveals specific and shared target genes of different ETS rearrangements. *Neoplasia* **14**, 600–611.

[12] Attard G, Clark J, Ambroisine L, Mills IG, Fisher G, Flohr P, Reid A, Edwards S, Kovacs G, Berney D, et al. (2008). Heterogeneity and clinical significance of ETV1 translocations in human prostate cancer. *Br J Cancer* **99**, 314–320.

[13] Hermans KG, van der Korput HA, van Marion R, van de Wijngaart DJ, Ziel-van der Made A, Dits NF, Boormans JL, van der Kwast TH, van Dekken H, Bangma CH, et al. (2008). Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. *Cancer Res* **68**, 7541–7549.

[14] Xu LL, Stackhouse BG, Florence K, Zhang W, Shanmugam N, Sesterhenn IA, Zou Z, Srikantan V, Augustus M, Roschke V, et al. (2000). *PSGR*, a novel prostate-specific gene with homology to a G protein-coupled receptor, is over-expressed in prostate cancer. *Cancer Res* **60**, 6568–6572.

[15] Weng J, Ma W, Mitchell D, Zhang J, and Liu M (2005). Regulation of human prostate-specific G-protein coupled receptor, PSGR, by two distinct promoters and growth factors. *J Cell Biochem* **96**, 1034–1048.

[16] Vaarala MH, Hirvikoski P, Kauppila S, and Paavonen TK (2012). Identification of androgen-regulated genes in human prostate. *Mol Med Rep* **6**, 466–472.

[17] Han B, Mehra R, Dhanasekaran SM, Yu J, Menon A, Lonigro RJ, Wang X, Gong Y, Wang L, Shankar S, et al. (2008). A fluorescence *in situ* hybridization screen for E26 transformation-specific aberrations: identification of DDX5-ETV4 fusion protein in prostate cancer. *Cancer Res* **68**, 7629–7637.

[18] Jones KA, Black DM, Griffiths BL, and Solomon E (1995). Localization of the human RNA polymerase I transcription factor gene (*UBTF*) to the *D17S183* locus on chromosome 17q21 and construction of a long-range restriction map of the region. *Genomics* **30**, 602–604.

[19] Matera AG, Wu W, Imai H, O'Keefe CL, and Chan EK (1997). Molecular cloning of the RNA polymerase I transcription factor hUBF/NOR-90 (*UBTF*) gene and localization to 17q21.3 by fluorescence *in situ* hybridization and radiation hybrid mapping. *Genomics* **41**, 135–138.

[20] Grandori C, Gomez-Roman N, Felton-Edkins ZA, Ngouenet C, Galloway DA, Eisenman RN, and White RJ (2005). c-Myc binds to human ribosomal DNA and stimulates transcription of rRNA genes by RNA polymerase I. *Nat Cell Biol* **7**, 311–318.

[21] Poortinga G, Hannan KM, Snelling H, Walkley CR, Jenkins A, Sharkey K, Wall M, Brandenburger Y, Palatsides M, Pearson RB, et al. (2004). MAD1 and c-MYC regulate UBF and rDNA transcription during granulocyte differentiation. *EMBO J* **23**, 3325–3335.

[22] Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, et al. (2007). Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448**, 595–599.

[23] Gasi D, van der Korput HA, Douben HC, de Klein A, de Ridder CM, van Weerden WM, and Trapman J (2011). Overexpression of full-length ETV1 transcripts in clinical prostate cancer due to gene translocation. *PLoS One* **6**, e16332.

[24] Lin C, Yang L, Tanasa B, Hutt K, Ju BG, Ohgi K, Zhang J, Rose DW, Fu XD, Glass CK, et al. (2009). Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translocations in cancer. *Cell* **139**, 1069–1083.

[25] Haffner MC, Aryee MJ, Toubaji A, Esopi DM, Albadine R, Gurel B, Isaacs WB, Bova GS, Liu W, Xu J, et al. (2010). Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nat Genet* **42**, 668–675.

[26] Engreitz JM, Agarwala V, and Mirny LA (2012). Three-dimensional genome architecture influences partner selection for chromosomal translocations in human disease. *PLoS One* **7**, e44196.

Supplementary Reference

[1] Kilpinen S, Autio R, Ojala K, Iljin K, Bucher E, Sara H, Pisto T, Saarela M, Skotheim RI, Björkman M, et al. (2008). Systematic bioinformatic analysis of expression levels of 17,330 human genes across 9,783 samples from 175 types of healthy and pathological tissues. *Genome Biol* **9**, R139.

Table W1. Oligonucleotide Primer Sequences for 5′RACE PCR.

Primer	Oligonucleotide Sequence (5′→3′)	Ensembl Transcript	Exon
ETV1_R_GSP	GTTCTGCTGGGATGAGCCAGGAAGC	ENST00000430479	10
ETV1_F_IC	TCCCTCCATCGCAGTCCATACCAGA	ENST00000430479	8
ETV1_R_NGSP	AAGCCTTGTGGTGGGAAGGGGATGT	ENST00000430479	9
ETV4_R_GSP	CAGGGACAACGCAGACATCATCTGG	ENST00000319349	10
ETV4_F_IC	CATGGCGAGCAGTGCCTTTACTCCA	ENST00000319349	6
ETV4_R_NGSP	GGCGAAGTCCGTCTGTTCCTGTTTGA	ENST00000319349	8
ETV5_R_GSP	CTCTCAGGCACAACACAAGTGTCGTCAA	ENST00000306376	10
ETV5_F_IC	CCCTTCCCTCCTCAGCCAGGAGTTC	ENST00000306376	8
ETV5_R_NGSP	TGGCTGCTGGAGAAATAACCCCCTCT	ENST00000306376	9

Table W2. Oligonucleotide Primers Used for RT-PCR.

Primer	Oligonucleotide Sequence (5′→3′)	Ensembl Transcript	Exon
C15orf21_S	TCTTTTGGTTGGTGATTTTGATGA	ENST00000409454	1
SLC45A3_S	GGAACCAGCCTGCACGC	ENST00000367145	1
UBTF_S	TTGACCCCGGGTTGCC	ENST00000393606	2
TMPRSS2	TAGGCGCGAGCTAAGCAGGAG	ENST00000332149	1
OR51E2_S	ACCACCACCAGGAATTGGC		
EST14_S	GAAGAGAAGAGAATTCACAGAGTGAAAA		
HERVK17_S	TTTCCACACTCTCATTCCGGA		
ETV1_AS1	TCGTTACAATTTCTCCCACGCT	ENST00000430479	4
ETV1_AS2	GCTTAAAGCCTTGTGGTGGG	ENST00000430479	9
ETV4_AS	GGAGTAAAGGCACTGCTCGC	ENST00000319349	6

Table W3. BAC Probes Used to Assess Genomic Rearrangements of ETS Transcription Factors and Novel 5′ Fusion Partners.

Gene Symbol	5′ Probe (Red)	3′ Probe (Green)
ETV1	RP11-643C13	RP11-790J2
ETV4	RP11-436J4	RP11-831F13
OR51E2	CTC-775N13	NA
Chr14 (q13.3-21.1)	RP11-945C4	NA
	RP11-666J24	

NA indicates not applicable.

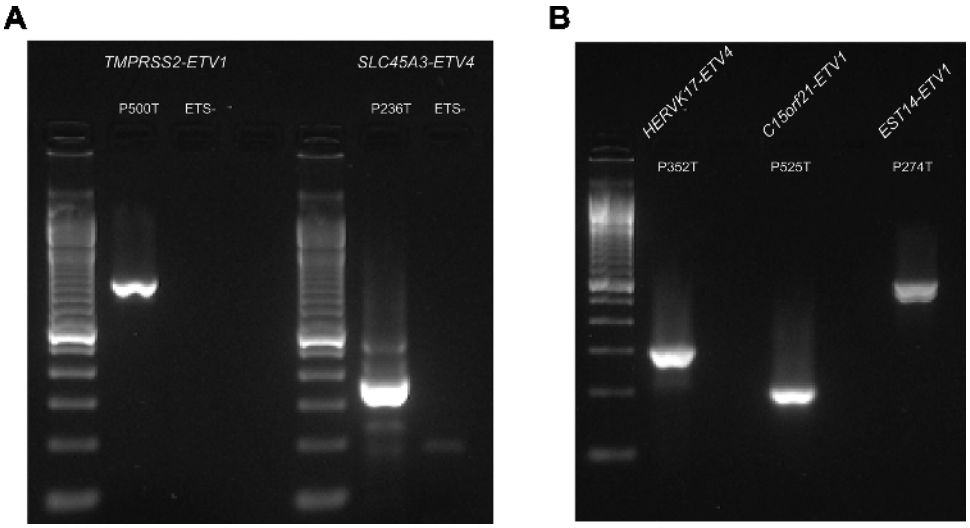

Figure W1. RT-PCR screening of gene fusions involving known 5′ fusion partners. Detection of *TMPRSS2-ETV1* and *SLC45A3-ETV4* fusion transcripts in PCas P500T and P236T, respectively (A). RT-PCR validation of 5′RACE findings of gene fusions involving known 5′ partners (B).

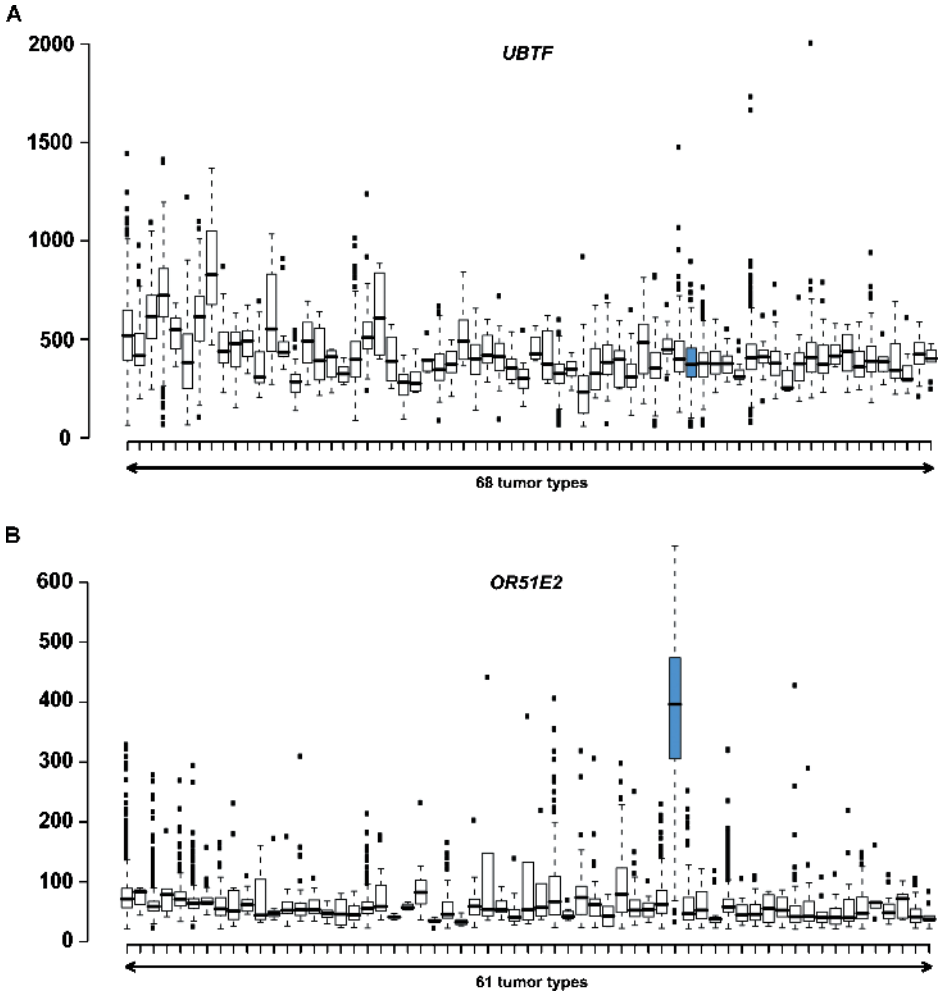

Figure W2. Expression of *UBTF* and *OR51E2* in different tumor types. The plots representing *UBTF* (a) and *OR51E2* expression (b), respectively, were obtained from the publicly available GeneSapiens database [1]. The blue box represents PCas.