Modern Pathology (2007) 20, 921–928
© 2007 USCAP, Inc All rights reserved 0893-3952/07 $30.00
www.modernpathology.org

Gene fusions between *TMPRSS2* and *ETS* family genes in prostate cancer: frequency and transcript variant analysis by RT-PCR and FISH on paraffin-embedded tissues

Jiangling J Tu*, Stephen Rohan*, Jean Kao, Naoki Kitabayashi, Susan Mathew and Yao-Tseng Chen

Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, New York, NY, USA

Recurrent gene fusions between *TMPRSS2* and *ETS* family genes have recently been shown to occur at a high frequency in prostate cancer. In this study, we used formalin-fixed paraffin-embedded tissue and evaluated both *TMPRSS2-ERG* and *TMPRSS2-ETV1* fusions by reverse transcription polymerase chain reaction (RT-PCR) and fluorescence *in situ* hybridization (FISH). The results were correlated to overexpression of the downstream *ERG* and *ETV1* sequences. Of 82 cases examined, *TMPRSS2-ETV1* fusion was seen in only one case, by FISH. In comparison, *TMPRSS2-ERG* fusion was documented in 35 cases (43%) by either RT-PCR or FISH. Deletion, rather than translocation, was found to be the main mechanism for *TMPRSS2-ERG* gene fusion (81 *vs* 19%). RT-PCR and FISH results correlated well, with most positive cases resulting in overexpression of downstream *ERG* sequences. Several *TMPRSS2-ERG* fusion transcript variants were identified, most of which are predicted to encode truncated *ERG* proteins. Prostate cancer of Gleason's scores 6 or 7 had more frequent *TMPRSS2-ERG* fusions than higher-grade tumors, but this difference was not statistically significant ($P = 0.42$). On the other hand, mucin-positive carcinomas more often harbor such gene fusions when compared to mucin-negative tumors ($P = 0.004$). These morphological correlates, and more importantly the potential correlation of such fusions to clinical outcome and treatment responses, should be further explored.

Modern Pathology (2007) **20,** 921–928; doi:10.1038/modpathol.3800903; published online 13 July 2007

Keywords: *TMPRSS2*; *ERG*; translocation; microdeletion

The role of recurrent chromosomal translocations in the pathogenesis of many hematopoetic malignancies is well established.[1] In lymphomas and leukemias these translocation events lead either to the activation of the 3′ partner gene—as in the activation of *MYC* following fusion to *IgH* in Burkitt's lymphoma —or to the formation of fusion proteins, such as the BCR-ABL fusion protein resulting from the t(9;22) translocation in chronic myeloid leukemia. In recent years, a variety of recurrent chromosomal translocations, most of which result in the formation of fusion proteins, have been described in sarcomas.[2] It is now known that approximately one-third of all soft tissue sarcomas are characterized by specific recurrent translocations, and in many cases these are the only detectable cytogenetic abnormality.[2] In both sarcomas and hematopoetic malignancies, many of these recurrent translocations have proven to be pathogenetically important.[1,2]

In contrast to lymphomas/leukemias and sarcomas, carcinomas are generally believed to arise via multi-step carcinogenesis, and recurrent chromosomal trasnslocations, with the exception of *MECT1-MAML2* t(11;19) translocation in low-grade mucoepidermoid carcinoma, are very uncommon in carcinomas.[3–5] However, a recently described gene fusion between *TMPRSS2* and *ETS* family genes in prostate carcinoma stands to challenge the old paradigm that epithelial malignancies arise by nonspecific chromosomal aberrations. Through bioinformatic analysis of DNA microarray data and subsequent fluorescence *in situ* hybridization (FISH) and reverse transcription polymerase chain reaction (RT-PCR) analysis, Tomlins *et al*[6] identified *TMPRSS2-ERG* fusion in 16/29 (55%) and *TMPRSS2-ETV1* fusion in 7/29 (27%) cases of prostate cancer. *ETV4*, another member of the *ETS* family, was later identified as another minor partner

Correspondence: Dr Y-T Chen, MD, PhD, Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, 1300 York Avenue, F504B, New York, NY 10021, USA.
E-mail: ytchen@med.cornell.edu
*These authors contributed equally to this work.
Received 12 March 2007; accepted 3 April 2007; published online 13 July 2007

in this gene fusion.[7] The high prevalence of prostate cancer in western societies, in combination with the reported frequency of fusion between *TMPRSS2* and *ETS* family genes, would make this one of the most common genetic alterations identified in human malignancy.

While the concept of a pathogenetically important recurrent genetic aberration in prostate carcinoma may be novel, the 3′ genes involved in this fusion event are well known. *ERG*, *ETV1* and *ETV4* are all members of the *ETS* gene family, and *ERG* and *ETV1* have been previously characterized as minor fusion partners (<10% cases) of *EWS-ETS* gene fusions in Ewing's sarcoma/primitive neuroectodermal tumors.[8] In addition, the *ERG* gene has also been shown to be the 3′ fusion partner in translocation events seen in a subset of acute myeloid leukemia, with the 5′ partners being *FUS*, *EWS* or *ELF4*.[9] The 5′ partner of this fusion, *TMPRSS2*, encodes a transmembrane serine protease that is constitutively expressed in prostate under the transcriptional control of androgens.[10,11]

In prostate cancer, multiple *TMPRSS2-ERG* fusion variants have been identified at the transcript level, most of them involving only the 5′ untranslated sequence of *TMPRSS2*.[6,12–15] The *TMPRSS2-ETS* fusions thus lead to overexpression of the *ERG* or *ETV* protein induced by the *TMPRSS2* promoter. In contrast to other translocations involving the *ERG* gene, for example, *EWS-ERG* fusion, no fusion protein is expected in most *TMPRSS2-ERG* fusions. Unlike the *TMPRSS2-ETV1* fusion, which involves a t(7;21) translocation, the *TMPRSS2* and *ERG* genes are located close to each other on chromosome 21, approximately 3 Mb apart.[15] Thus, the *TMPRSS2-ERG* fusion could result either from balanced or unbalanced translocation, or from deletion of the intervening DNA segment, and the literature to date suggests deletion as the more frequent event.[16,17]

In the present study, we evaluated a series of 82 formalin-fixed paraffin-embedded prostate cancer specimens, by RT-PCR and by FISH analyses. *TMPRSS2-ERG* fusion was observed in 43% of the cases, more commonly resulting from deletion than translocation. Novel fusion transcript variants were identified, and overexpression of *ERG* downstream sequence was seen in most fusion-positive cases. Possible correlation between the fusion status, histological features and grade of the tumor was also observed. In comparison to the frequent *TMPRSS2-ERG* fusion, only one *TMPRSS2-ETV1* fusion was identified, indicating that this fusion is less common than previously reported.

Materials and methods

Tissue Specimens

Tissue specimens, derived from radical prostatectomy, were obtained from the Department of Pathology at the Weill Medical College of Cornell University, following a protocol approved by the Institutional Review Board.

RNA Extraction

One representative block was identified from each case, and four 8 μm sections were used for RNA extraction, using Optimum FFPE RNA isolation kit (Ambion, Austin, TX, USA). The non-tumor areas on the slides were manually removed with surgical blades, and the remaining tissue was scraped into an Eppendorf tube for RNA extraction.

RT-PCR and DNA Sequencing

All primer sequences are listed in Table 1. *TMPRSS2-ERG* and *TMPRSS2-ETV1* gene fusions were evaluated by nested RT-PCR. For *TMPRSS2-ERG* fusion, two sets of RT-PCR were performed, coupling a 5′*TMPRSS* exon 1 primer to a 3′*ERG* primer located at either exon 4 or exon 5. For *TMPRSS2-ETV1* fusion, the 5′*TMPRSS2* exon 1 primer was coupled to a 3′*ETV1* exon 6 primer. Although fusion products were often detectable by gel electrophoresis after the first PCR, nested PCRs were performed on all cases with corresponding internal primers, and PCR products were identified by 1% agarose gel electrophoresis and ethidium bromide visualization. In cases where discrete PCR products were detected, direct PCR sequencing was performed on the purified PCR products. DNA elution from the gel fragments was performed if more than one DNA species was present. Overexpression of downstream *ERG* or *ETV1* sequences was evaluated by quantitative RT-PCR assays, using *ERG* exon 5–exon 6 primer pair or *ETV1* exon 6–exon 7 primer pair, coupled to corresponding TaqMan FAM-labeled *ERG* or *ETV1* probe (Applied Biosystems gene expression assay ID Hs00171666 and Hs00231877, respectively). 18S ribosomal RNAs were used as endogenous controls for RNA quality. *ETV1* exon 2–exon 3 primer pair amplification (Applied Biosystems assay ID Hs00951945) was also found to be a reliable endogenous control and was used to normalize the *ERG* amplification results.

Table 1 RT-PCR primers for detection of *TMPRSS2-ERG* and *TMPRSS2-ETV1* fusions[a]

Primer	Sequence
TMPRSS2-1A	TAGGCGCGAGCTAAGCAGGAG
TMPRSS2-1A1	CGCGAGCTAAGCAGGAGGC
ERG-4B	GTAGGCACACTCAAACAACGACTGG
ERG-4B1	TGGTCCTCACTCACAACTGATAAG
ERG-5B	TCATCCCAACGGTGTCTGGGCTG
ERG-5B1	CAACGGTGTCTGGGCTGCCCACC
ETV1-5B	CAGGCCATGAAAAGCCAAACTT

[a]All 'A' primers are forward primers and 'B' primers are reverse primers; 'A1' and 'B1' primers are nested primers.

PCR were performed for 45 cycles as previously described.[18]

FISH

Interphase FISH was performed on the same set of cases analyzed by RT-PCR. Tissue microarray slides were prepared from the blocks used for RNA extractions, with each case represented by three 1 mm tissue cores. *TMPRSS2-ERG* fusion was evaluated using break-apart probes, consisting of rhodamine-labeled 5′*ERG* probe (BAC RP11-95I21) and FITC-labeled 3′*ERG* probe (BAC RP11-476D17). Using this break-apart approach, a cell with two normal copies of chromosome 21 would have two yellow signals, due to the close proximity of the two probes. A cell with a translocation would have one yellow, one green and one red signal, and a cell in which *TMPRSS2-ERG* fusion was due to a deletion of intervening DNA would have one green and one yellow signal—the 5′ (red) signal would be lost. *TMPRSS2-ETV1* fusion was evaluated with a two-color two-signal approach, using rhodamine-labeled *TMPRSS2* probe (BAC RP11-35C4) and FITC-labeled *ETV1* probe (BAC RP11-124L22). *TMPRSS2-ETV1* translocation results in fusion of these two signals, generating a yellow signal. BAC clones were obtained from Children's Hospital of Oakland Research Institute (CHORI) and from Invitrogen (Carlsbad, CA, USA). Specificity and quality of the probes were confirmed by hybridization to the metaphase spread of normal peripheral lymphocytes. In a previous study,[6] the same 5′*ERG* probe that we used showed a weak signal on chromosome 2 due to cross-hybridization. However, we detect this nonspecific signal neither in our metaphase spread nor in the interphase of non-neoplastic cells on tissue sections. Tissue pretreatment was performed using Paraffin Pretreatment kit I (Vysis, Des Plaines, IL, USA), and hybridization and washing were performed using Vysis hybridization reagents, following the manufacturer's protocols. An average of 100 cells were evaluated, and the FISH results were independently scored by a cytogeneticist (S Mathew) and a pathologist (S Rohan).

Results

Frequency of *TMPRSS2-ETS* Fusions

Of 82 cases, eight cases showed poor RNA quality, evidenced by suboptimal 18S rRNA amplification and lack of amplification with the *ETV1* exon 2/exon 3 primer pair, and these eight cases were evaluated by FISH only. Excluding these cases, *TMPRSS2-ERG* fusion transcripts were detected in 33 of 74 cases (45%). Examples of RT-PCR results are shown in Figure 1.

FISH analysis for *ERG*-related translocations and/or deletions was successfully performed on tissue microarray in 59 of the 82 cases. Gene fusion signals were detected in 27 cases (46%), resulting from either translocations or deletions (see below). Combining both assays, 35 of 82 cases (43%) were shown to have *TMPRSS2-ERG* fusion by one of the two assays. Both RT-PCR and FISH data were available in 52 cases, including 25 concordant positive and 25 concordant negative cases. Two discordant cases were seen, one PCR-negative, FISH-positive, and the other PCR-positive, FISH-negative. RT-PCR was repeated and confirmed in both cases. Using FISH as the gold standard, RT-PCR using FFPE thus had 93% sensitivity and 93% specificity.

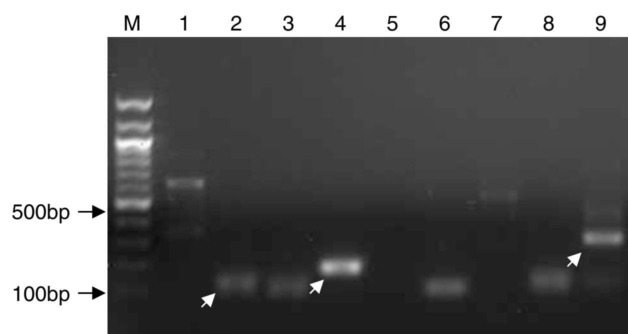

Figure 1 Representative *TMPRSS2-ERG* fusion transcripts by RT-PCR. The arrows indicate the following three transcript variants in this panel: T1/E4 (lanes 2, 3, 6, 8), T1-4/E4 (lane 4) and T1/E2 (lane 9). Lanes 1, 5 and 7 were negative for gene fusion. The higher-molecular-weight DNA bands in lanes 1 and 7 were sequenced and shown to be unrelated to *TMPRSS2* and *ERG*, representing nonspecific amplifications.

In contrast to the frequent *TMPRSS2-ERG* fusion, the *TMPRSS2-ETV1* fusion was not found in any case by RT-PCR. However, one case did show evidence of *TMPRSS-ETV1* fusion by FISH analysis (Figure 2d).

Deletion, Translocation and Other FISH Findings

Of the 27 FISH-positive cases, evidenced by split *ERG* signals, 22 (81%) cases showed loss of the split 5′ *ERG* signal (red), indicating a deletion of the intervening sequence between *TMPRSS2* and the *ERG* genes (Figure 2a). The remaining five cases, in contrast, retained both 5′ (red) and 3′ (green) split signals, indicating a translocation instead of deletion (Figure 2b).

In addition to deletion and translocations, one case showed additional *ERG* copies, with up to five 3′ (green) *ERG* signals seen in the tumor cells. This case was *TMPRSS2-ERG* fusion positive by RT-PCR. Three *ERG* signals (one yellow and two green) were seen as the most common karotype in this case, indicating one intact copy (yellow) and two copies with the 5′*ERG* sequence deleted, presumably resulting from duplication of the copy containing *TMPRSS2-ERG* fusion (Figure 2c).

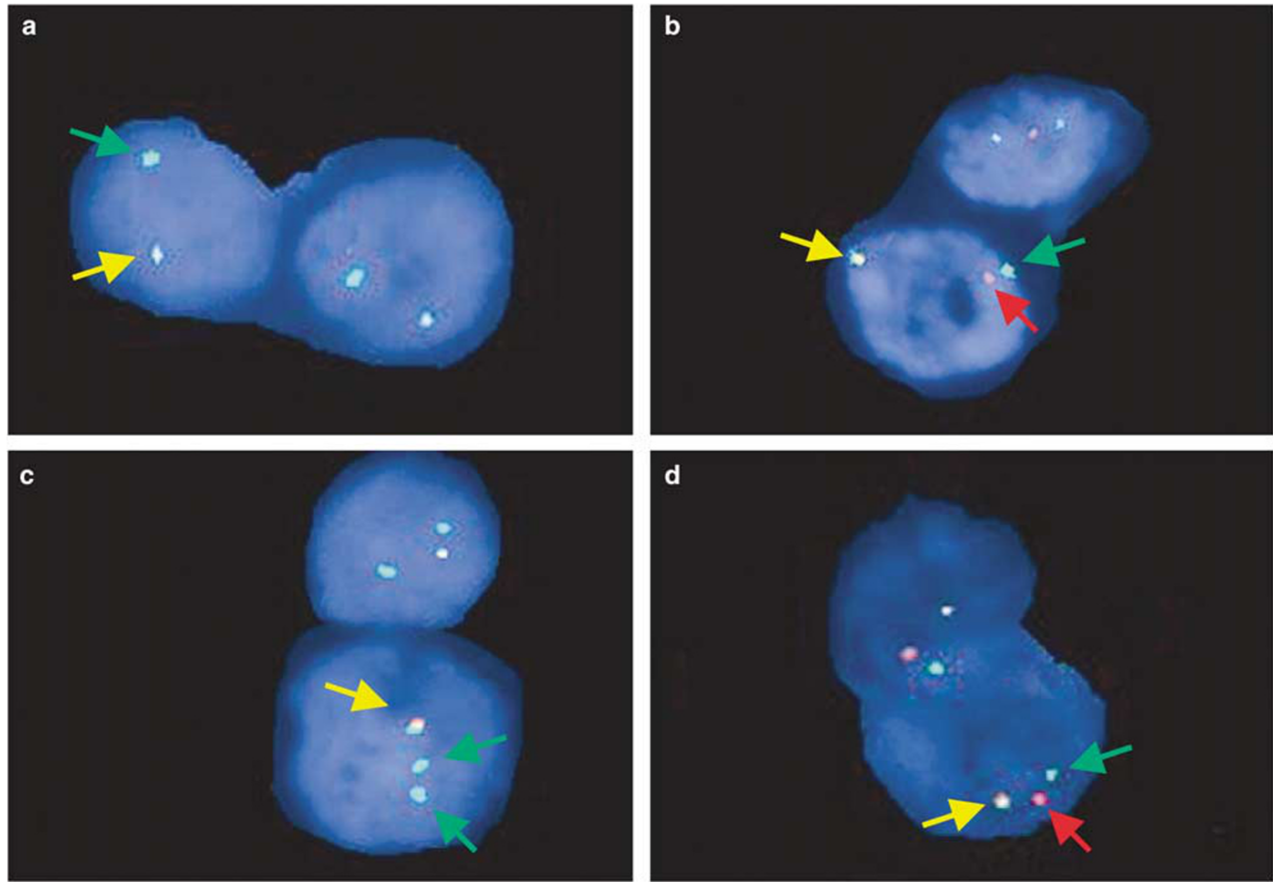

Figure 2 *TMPRSS2-ETS* fusion analysis by FISH. (**a–c**) *TMPRSS2-ERG* fusion was analyzed using break-apart *ERG* probes. *TMPRSS2-ERG* fusion resulting from deletion of the intervening sequence led to an isolated 3′*ERG* signal (green) and no 5′*ERG* (red) signal (**a**). In contrast, a red signal is detected in cases of *TMPRSS2-ERG* fusion due to translocation (**b**). Low-copy *ERG* amplification coexisted with *TMPRSS2-ERG* fusion in one case (**c**), which in most cells showed an intact copy of *ERG* (yellow) and two copies with only 3′*ERG* signal (green), likely due to duplication of the copy carrying *TMPRSS2-ERG* fusion. (**d**) The only case of *TMPRSS-ETV1* fusion, represented by the fused (yellow) signal.

Correlation of Gene Fusion to Downstream *ETS* Overexpression

The downstream *ERG* mRNA expression level was evaluated using qRT-PCR assay and *ERG* exon 5 and exon 6 primers, 3′ to all fusion junctions (see below). PCR amplification using *ETV1* exon 2/exon 3 primers, 5′ to any potential *ETV1*-related gene fusion, was found to produce consistent amplification results across most specimens, and this was used as endogenous control to evaluate potential *ERG* and *ETV1* mRNA overexpression. Figure 3a shows the correlation between the *TMPRSS2-ERG* gene fusion status and the expression levels of the downstream *ERG* sequences. Setting a normalized $-\Delta Ct_{(ERG\ \mathrm{exon}\ 5/6\text{-}ETV1\ \mathrm{exon}\ 2/3)}$ value of 4 as the cut-off threshold, *ERG* overexpression was found in 32 of 75 cases with valid RT-PCR data, among which 29 were positive for *TMPRSS2-ERG* fusion by RT-PCR and/or by FISH. FISH was negative in two of the remaining three cases, and failed in the third case.

Conversely, of the 33 *TMPRSS2-ERG* RT-PCR fusion-positive cases, 28 showed *ERG* overexpression, with five cases showing *ERG* levels similar to those of fusion-negative cases. Using fusion-negative, *ERG* overexpression-negative cases as control and assuming 100% PCR amplification efficiency, the calculated downstream *ERG* mRNA levels in the fusion-positive group was on average 11-fold greater than that of the control group.

Downstream *ETV1* overexpression was similarly investigated, and only three cases showed a clear increase of *ETV1* downstream transcript level (Figure 3b). Both RT-PCR and FISH were negative in these cases. The only *TMPRSS2-ETV1* translocation-positive case detected by FISH showed no detectable *ETV1* overexpression.

TMPRSS2-ERG Transcript Variants

Six *TMPRSS2-ERG* fusion transcript variants were identified in this study. Table 2 shows a compilation of these six and additional 13 variants reported in the literature. The most predominant variant, seen in 29/33 (88%) of fusion-positive cases, fused

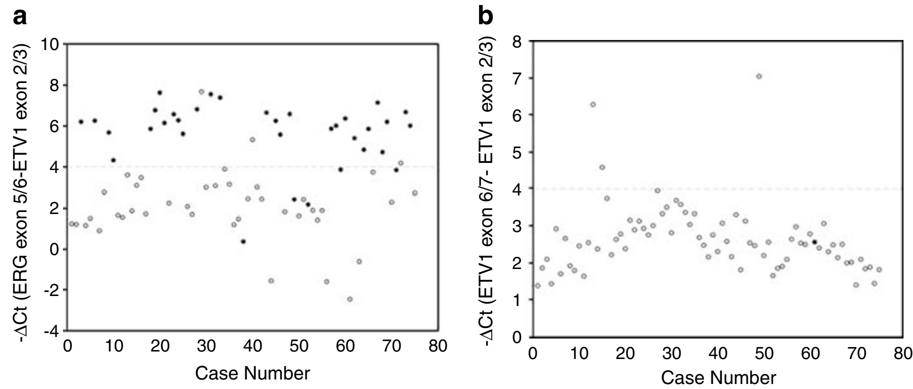

Figure 3 Correlation between mRNA expression of ETS downstream sequences (*ERG* (**a**) and *ETV1* (**b**)) to *TMPRSS2-ETS* fusion status. Cases positive for *TMPRSS2-ERG* or *TMPRSS2-ETV1* fusion (most by RT-PCR, two by FISH only, see text) are shown in solid circles, whereas negative cases are in open circles. The *y*-axis depicts relative expression levels to *ETV1* 5′mRNA, the endogenous control, and higher numbers indicate higher mRNA levels. By setting the threshold for overexpression at 4 (dashed lines), most *TMPRSS2-ERG* fusion-positive and -negative cases can be separated ((**a**), see text). All but one case were negative for *TMPRSS2-ETV1* fusion, and similar *ETV1* downstream mRNA levels were seen in all but three fusion-negative cases.

Table 2 *TMPRSS2-ERG* fusion transcript variants identified in current study and in the literature

	Variant	*Clark et al*[12]	*Wang et al*[14]	*Soller et al*[13]	*Yoshimoto et al*[15]	*Tomlins et al*[6]	*Current study (#)*
1	T1/E2	6	I	—	√	√	√ (2)
2	T1/E234-6	5	—	—	—	—	—
3	T1/E3	7	II	—	—	—	—
4	T1/E3-5	3	—	—	—	—	—
5	T1/E-int3a	12	—	—	—	—	—
6	T1-int1/E-int3b	13	—	—	—	—	—
7	T1/E-int3c	14	—	—	—	—	—
8	T1/E4	9	III	1	√	√	√ (29)
9	T1/E5	10	IV	—	√	—	√ (6)
10	T1/E6	11	—	—	—	—	—
11	T2/E2	2	V	—	—	—	—
12	T2/E4	8	VI	2	—	—	√ (1)
13	T2/E5	4	VII	—	√	—	—
14	T3/E4	1	VIII	—	—	—	—
15	T4/E4	—	—	5	—	—	—
16	T1-4/E4	—	—	—	—	—	√ (1)
17	T4/E5	—	—	4	—	—	—
18	T1-4/E5	—	—	—	—	—	√ (1)
19	T5/E4	—	—	3	—	—	—

Number of cases.

TMPRSS2 exon 1 (5′ untranslated exon) to exon 4 of *ERG*, and this transcript was designated T1/E4. Using this terminology, the other transcript variants found in this study were T1/E5 (six cases), T1/E2 (two cases), T2/E4 (one case), T1-4/E4 (one case) and T1-4/E5 (one case). All six cases positive for T1/E5 also expressed T1/E4 transcript. In contrast, the T1/E2 and T2/E4 cases did not express T1/E4 or T1/E5. One single case showed two novel transcripts consisting of *TMPRSS2* exon 1 alternatively spliced to exon 4, which was fused to the *ERG* exon 4 or exon 5 (T1-4/E4 and T1-4/E5).

Of the different variant transcripts described, five variants were reported by at least three groups and constituted the main variants isolated to date (Figure 4). As the *TMPRSS2* native translational initiation site is located in exon 2, only the T2/E4 transcript would result in a *TMPRSS2-ERG* fusion protein. The T2/E4 variant potentially encodes a fusion protein of 454 amino acids, of which the N-terminal five residues being derived from *TMPRSS2*, a region outside of any known functional domain. Among the other four common transcripts, only the T1/E2 variant can encode the full-length *ERG* protein (462 amino acids), the other three variants can only encode truncated *ERG* proteins using an internal methionine as translational initiation site, with predicted putative proteins of 423 (T1/E4) and 363 (T1/E5 and T2/E5) residues, both still containing the *ETS* functional domain.

Correlating Genetic Changes to Pathological Stages and Histological Parameters

The 36 fusion-positive cases, by either RT-PCR or FISH, included 23 of 54 (43%) stage 2 (T2) tumors, 13 of 28 (46%) stage 3 (T3) tumors and 2 of 5 (40%)

lymph node-positive (N1) tumors. Therefore, the fusion status did not correlate to the pathological stages. In contrast, these 36 fusion-positive cases contained eight of 12 (67%) Gleason's 6, 22 of 52 (42%) Gleason's 7, one of five Gleason's 8, and five of 13 of Gleason's 9 carcinomas. The frequency of *TMPRSS2-ERG* fusion was lower in poorly differentiated carcinoma (Gleason's 8 and above, 6/18, 33%) than in the better differentiated ones (Gleason's 6 and 7, 30/64, 47%), but the difference was not statistically significant ($P=0.42$). The differences between tumors of individual grades were also not statistically significant.

Possible morphological-genetic correlations were explored, and the results are shown in Table 3. Tumors with *TMPRSS2-ERG* fusions more frequently had focal intraluminal mucin, amphophilic cytoplasm and cribriform architecture. In contrast, tumors without *TMPRSS2-ERG* fusions more often had focal foamy gland change, signet ring-like feature or ductal differentiation. However, the only statistically significant difference was focal intraluminal mucin, seen at a much higher frequency in tumors with the fusion (78% *vs* 46%, $P=0.004$).

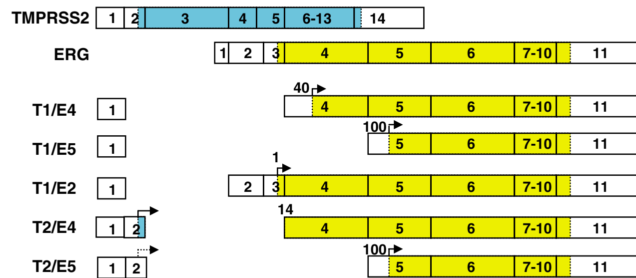

Figure 4 Major *TMPRSS2-ERG* fusion transcript variants and their putative protein products. T1/E4 is by far the most commonly detected transcript (see text). Exons are depicted in open boxes and protein coding sequences are in color boxes. Except for T1/E2 variant, all other variants code for proteins with truncated *ERG* sequences. The predicted translational initiation sites are represented by solid arrows, with the ERG amino-acid numbers of the full-length ERG protein indicated. Although the T2/E5 variant could initiate from the *TMPRSS2* sequence (dashed arrow), this would not be in frame with the *ERG* and likely not translated.

Discussion

The clinical and pathological significance of recurrent chromosomal translocation in hematopoetic malignancies and certain sarcoma subtypes has been recognized for many years. The fusion of *TMPRSS2* to *ETS* family genes in prostate cancer, first described by Tomlins *et al*,[6] raised the possibility that similar recurrent genetic alterations may also be important in prostate carcinogenesis. However, a few substantial differences have now emerged based on our present study and other recent studies.[6,12–14,16,19]

First, the frequency of gene fusion in prostate cancer is lower than in translocation-positive sarcomas. In synovial sarcoma, for instance, the fusion of *SYT-SSX* is present in essentially all cases, to the extent that the absence of *SYT-SSX* fusion would raise serious doubt about the diagnosis. In the initial study of prostate cancer, a very high *TMPRSS2-ETS* fusion frequency (79.3%, 23/29) was observed, including *TMPRSS2-ERG* fusion in 55% (16/29) and *TMPRSS2-ETV1* in 24% (7/29) of cases. Such a high fusion frequency was similarly observed by Soller *et al*,[13] who found *TMPRSS2-ERG* fusion in 14/18 cases (78%), but no *TMPRSS2-ETV* fusions. In comparison to these two studies, we observed a significantly lower frequency (43%) of *TMPRSS2-ERG* fusion in our larger series, and only a single *TMPRSS2-ETV1* fusion was detected by FISH. To exclude the possibility of false-negative results due to RT-PCR assays on paraffin-embedded tissues, we analyzed the same set of cases for *ERG* overexpression and by FISH, and both showed similar results. This lower frequency of gene fusion was indeed comparable to that observed in the more recent studies,[14,16] indicating *TMPRSS2-ERG* fusion as the crucial event in about 40–55% of prostate cancer cases. Additionally, similar to the study by Soller *et al*,[13] we found *TMPRSS2-ETV1* fusion to be a rare event. Although it is possible that the remaining half of the cases might have involved currently unknown gene fusion partners, the finding that *ERG* and *ETS* genes are the only genes found to be overexpressed so far[6,20] argues against such a possibility and implies the existence of more than one pathway in prostate carcinogenesis.

Table 3 The correlation of the histological features and *TMPRSS2-ERG* fusion

TMPRSS2-ETS fusion	*Histological features*[a]					
	*Mucin***	*Amphophilic*	*Cribriform*	*Foamy cells*	*Signet ring*	*Ductal diff.*
Positive ($n=46$)	21 (46%)	17 (37%)	11 (24%)	22 (48%)	7 (15%)	10 (22%)
Negative ($n=36$)	28 (78%)	20 (56%)	16 (44%)	11 (31%)	4 (11%)	5 (14%)

[a]Features evaluated: intraluminal mucin, amphophilic cytoplasm, cribriform architecture, foamy cell changes, presence of signet ring cells and ductal differentiation.
**$P=0.004$

Another difference between *ERG* fusions in prostate cancer and gene fusions in other types of malignancies is the lack of a fusion protein in the former. The *ERG* gene is also involved in chromosomal translocations of Ewing's sarcoma and acute myeloid leukemia, with *EWS* being the most common 5′ partner. The *EWS-ERG* fusion, however, leads to fusion proteins that include functional DNA/RNA binding domains of both partners, presumably resulting in a synergistic biological effect. In contrast, the *TMPRSS2-ERG* fusion will create fusion proteins in rare cases, for example, in the T2/E4 transcript variant. Even in these cases, the *TMPRSS2* sequence is minimal and not anticipated to be biologically active. The result of the *TMPRSS2-ETS* fusion is thus analogous to the *IgH-MYC* fusion in B-cell lymphomas, in that the 5′ partner serves the function of providing a tissue-specific promoter, resulting in constitutive expression of the 3′ gene.

One unique feature of the *TMPRSS2-ERG* fusion is the vast diversity of fusion transcripts. We identified six variants in the present study, including two novel variants T1-4/E4 and T1-4/E5. Compiling our and other recent RT-PCR studies, a total of 19 transcript variants have been identified to date, variably and confusingly named types I–VIII,[14] 1–14,[12] etc, by investigators. In light of this complexity, we would propose that a generic nomenclature system be used, and these variants would be named as T1/E4, T1/E5, T1-4/E4, etc, as was first used by Clark *et al*.[12] From Table 2, it is clear that T1/E4 is the predominant transcript observed in all studies, and other common variants are T1/E5, T1/E2, T2/E4 and T2/E5. Intriguingly, except for the T2/E4 variant, all other fusion transcripts, including the most common T1/E4 variant, can only encode truncated *ERG* proteins. The translational efficiency of these aberrant RNAs and the biological activity of the truncated proteins in prostate cancer remain to be investigated. In addition, it is unclear how these transcript variants would correspond to different fusion junctions at the genomic DNA level. The observation that T1/E5 coexisted with T1/E4 in most cases (Clark *et al*[12] and this study) indicates that these two variants are presumably alternatively spliced products from a single fusion. On the other hand, T1/E4 does not appear to coexist with T1/E2, T2/E4 or T1-4/E4, suggesting that these might represent different gene fusion junctions. It is of interest that Wang *et al*[14] proposed that the T2/E4 variant might be associated with a more aggressive phenotype of prostate cancer. This variant was only observed in one case in our series that was of Gleason's grade 7 and lymph node negative. Elucidation of the fusion junctions by long-range PCR, chromosomal walking or other cloning methodology should shed light on this issue.

Another unique feature of the *TMPRSS2-ERG* gene fusion is that unlike *TMPRSS2-ETV1* or other gene fusions, *TMPRSS2* and *ERG* are located 3 Mb apart on chromosome 21, and deletion instead of translocation appears to be the main mechanism of gene fusion. First demonstrated by Yoshimoto *et al*[15] with three-color FISH, Perner and co-workers showed deletion in ~70% of cases, similar to our observation (21/27, 81%). This finding raises the possibility that the *TMPRSS2-ERG* fusion event, in addition to *ERG* activation, is often accompanied by variable losses of intervening genes. It is currently unclear whether this microdeletion of chromosome 21q would carry additional biological significance beyond the *ERG* activation, a view that was recently entertained.[16,17]

The finding that *TMPRSS2-ETS* fusion is seen in approximately half of the prostate cancer also raises the issue of whether the fusion-positive and fusion-negative cases would differ biologically, for instance, in their histomorphology, response to treatment or prognosis. Previous studies have shown no correlation to Gleason's grades, lymph node status or clinical stages,[16] and Demichelis *et al*[21] suggested that *TMPRSS2-ERG* fusion-positive carcinomas might represent a more aggressive phenotype. In contrast to this notion, we confirmed that the gene fusion did not correlate with the tumor's stages. Although a trend of more frequent *TMPRSS2-ERG* fusion was observed in moderately differentiated tumor (*vs* poorly differentiated tumor) in this study, this finding was not statistical significant and needs to be further evaluated. We also found carcinomas with focal visible intraluminal mucin to have a significantly higher frequency of *TMPRSS2-ETS* fusion. Other histologically features that we examined, for example, amphophilic cytoplasm, cribriform architecture, focal foamy gland change, etc, were not statistically different between the fusion-positive and -negative groups. Such correlations between the fusion status and morphologic features could impact on pathology practice if the presence (or absence) of gene fusion carries therapeutic or prognostic implications. Since the *TMPRSS2-ETS* fusions would place the *ERG/ETV* genes under the regulation of androgen control, one might indeed speculate that the fusion status might predict the response to hormonal treatment, and studies to address this possible correlation, such as the recent *in vitro* study by Hermans *et al*,[22] would clearly be clinically important.

Acknowledgement

We thank Ms Andrea Panczykowski for assistance with the FISH analysis.

References

1 Jaffe ES, Harris NL, Stein H, *et al*. Pathology and Genetics of Tumors of Hematopietic and Lymphoid Tissue. IARC Press: Lyon, 2001.

2 Antonescu CR. The role of genetic testing in soft tissue sarcoma. Histopathology 2006;48:13–21.
3 Okabe M, Miyabe S, Nagatsuka H, *et al.* MECT1-MAML2 fusion transcript defines a favorable subset of mucoepidermoid carcinoma. Clin Cancer Res 2006;12:3902–3907.
4 Argani P, Ladanyi M. Translocation carcinomas of the kidney. Clin Lab Med 2005;25:363–378.
5 Zettl A, Strobel P, Wagner K, *et al.* Recurrent genetic aberrations in thymoma and thymic carcinoma. Am J Pathol 2000;157:257–266.
6 Tomlins SA, Rhodes DR, Perner S, *et al.* Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–648.
7 Tomlins SA, Mehra R, Rhodes DR, *et al.* TMPRSS2: ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res 2006;66:3396–3400.
8 Ladanyi M. The emerging molecular genetics of sarcoma translocations. Diagn Mol Pathol 1995;4: 162–173.
9 Moore SD, Offor O, Ferry JA, *et al.* ELF4 is fused to ERG in a case of acute myeloid leukemia with a t(X;21)(q25-26;q22). Leuk Res 2006;30:1037–1042.
10 Vaarala MH, Porvari KS, Kellokumpu S, *et al.* Expression of transmembrane serine protease TMPRSS2 in mouse and human tissues. J Pathol 2001;193:134–140.
11 Paoloni-Giacobino A, Chen H, Peitsch MC, *et al.* Cloning of the TMPRSS2 gene, which encodes a novel serine protease with transmembrane, LDLRA, and SRCR domains and maps to 21q22.3. Genomics 1997; 44:309–320.
12 Clark J, Merson S, Jhavar S, *et al.* Diversity of TMPRSS2-ERG fusion transcripts in the human prostate. Oncogene 2007;26:2667–2673.
13 Soller MJ, Isaksson M, Elfving P, *et al.* Confirmation of the high frequency of the TMPRSS2/ERG fusion gene in prostate cancer. Genes Chromosomes Cancer 2006;45:717–719.
14 Wang J, Cai Y, Ren C, *et al.* Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. Cancer Res 2006;66: 8347–8351.
15 Yoshimoto M, Joshua AM, Chilton-Macneill S, *et al.* Three-color FISH analysis of TMPRSS2/ERG fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. Neoplasia 2006;8:465–469.
16 Perner S, Demichelis F, Beroukhim R, *et al.* TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. Cancer Res 2006;66:8337–8341.
17 Iljin K, Wolf M, Edgren H, *et al.* TMPRSS2 fusions with oncogenic ETS factors in prostate cancer involve unbalanced genomic rearrangements and are associated with HDAC1 and epigenetic reprogramming. Cancer Res 2006;66:10242–10246.
18 Chen YT, Tu JJ, Kao J, *et al.* Messenger RNA expression ratios among four genes predict subtypes of renal cell carcinoma and distinguish oncocytoma from carcinoma. Clin Cancer Res 2005;11:6558–6566.
19 Ahlers CM, Figg WD. ETS-TMPRSS2 fusion gene products in prostate cancer. Cancer Biol Ther 2006;5:254–255.
20 Petrovics G, Liu A, Shaheduzzaman S, *et al.* Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. Oncogene 2005;24: 3847–3852.
21 Demichelis F, Fall K, Perner S, *et al.* TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. Oncogene 2007 (in press).
22 Hermans KG, van Marion R, van Dekken H, *et al.* 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: 10658–10663.