Published online 15 September 2008

Nucleic Acids Research, 2008, Vol. 36, No. 20 ***e130***
doi:10.1093/nar/gkn585

A sensitive array-based assay for identifying multiple TMPRSS2:ERG fusion gene variants

Qing Lu[1], Esperanza Nunez[2], Chunrun Lin[2], Kimberly Christensen[3], Tracy Downs[1,4], Dennis A. Carson[1,5], Jessica Wang-Rodriguez[1,3] and Yu-Tsueng Liu[1,5,*]

[1]Moores Cancer Center, [2]Howard Hughes Medical Institute, [3]Department of Pathology, [4]Department of Surgery and [5]Department of Medicine, University of California, San Diego, La Jolla, CA 92093, USA

Received May 7, 2007; Revised August 25, 2008; Accepted August 26, 2008

ABSTRACT

Studies of gene fusions in solid tumors are not as extensive as in hematological malignancies due to several technical and analytical problems associated with tumor heterogeneity. Nevertheless, there is a growing interest in the role of fusion genes in common epithelial tumors after the discovery of recurrent TMPRSS2:ETS fusions in prostate cancer. Among all of the reported fusion partners in the ETS gene family, TMPRSS2:ERG is the most prevalent one. Here, we present a simple and sensitive microarray-based assay that is able to simultaneously determine multiple fusion variants with a single RT–PCR in impure RNA specimens. The assay detected TMPRSS2:ERG fusion transcripts with a detection sensitivity of <10 cells in the presence of more than 3000 times excess normal RNA, and in primary prostate tumors having no >1% of cancer cells. The ability to detect multiple transcript variants in a single assay is critically dependent on both the primer and probe designs. The assay should facilitate clinical and basic studies for fusion gene screening in clinical specimens, as it can be readily adapted to include multiple gene loci.

INTRODUCTION

Chromosome rearrangements are a characteristic feature of cancer. More than 350 gene fusions, as a consequence of chromosome aberrations, have been identified (1). While gene fusions are common in hematological malignancies, their presence in solid tumors is not as well studied due to several technical and analytic problems related to tumor heterogeneity (1). Only very limited gene fusion events were discovered in solid tumors, mostly in sarcomas, until the recent discovery of TMPRSS2:ETS fusion genes in prostate cancer (2). This finding has since changed the general view that gene fusions play only a minor role in the pathogenesis of epithelial tumors. Therefore, there is renewed interest in searching for fusion genes in solid tumors, due to their potential impact on basic research and clinical application as has been demonstrated in chronic myelogenous leukemia (CML) (3,4).

The recurrent gene fusion event in prostate cancer involves an androgen controlled gene, TMPRSS2, and members (ERG, ETV1 and ETV4) of the ETS transcription factor family (2,5,6). Among these fusion genes, TMPRSS2:ERG is the most prevalent and the only member detected in the majority of reports. This fusion transcript results from ~3 Mb interstitial deletion between these two loci at chromosome 21q22. It was found in approximately half (15–78%) of all prostate cancers (2,6–17). As an androgen-related transcription factor controlling cell proliferation, TMPRSS2:ERG has been associated with disease pathogenesis and is a promising biomarker for prostate cancer progression, prognostication and early detection (18–21). While the presence of TMPRSS2:ETS fusion genes is highly prostate cancer-specific, its significance as a prognostic biomarker is still controversial partly because many of the clinical studies have been relatively small scale. Therefore, it is important to develop a simple and robust assay for identifying various TMPRSS2:ETS and potential fusion genes in other solid tumors. However, this could be challenging due to high heterogeneity in prostate cancer and other solid tumors, compared to leukemias and lymphomas (22).

Several approaches that have been used previously for hematological malignancies have been applied to detect TMPRSS2:ERG exon fusion variants. These include fluorescent *in situ* hybridization (FISH) (2,12,14,17,23), RT–PCR and sequencing (2,7,9,13), quantitative PCR (qPCR) (2,8,24) and array-based comparative genome hybridization (array CGH) (10–12). FISH may be the most commonly used method, but it has relatively low resolution, and therefore, cannot accurately determine different fusion variants. Array CGH has a higher

*To whom correspondence should be addressed. Tel: +1 858 534 9972; Fax: +1 858 534 5399; Email: ytliu@ucsd.edu
Correspondence may also be addressed to Jessica Wang-Rodriguez. Tel: +1 858 642 3511; Fax: +1 858 642 3918; Email: jwrodriguez@ucsd.edu

resolution but is costly and often fails when there is normal cell contamination.

RT–PCR and qPCR are relatively easy to perform. However, to assess multiple potential fusion variants requires multiple sets of primers and probes, and a corresponding large quantity of RNA templates. Moreover, sequencing RT–PCR products is laborious and difficult to adapt in routine clinical laboratories. Here, we describe an exon array-based detection system, combined with a RT–PCR reaction, that accurately determines multiple TMPRSS2:ERG fusion transcripts in specimens with only a minor population of tumor cells. The method adopts several features of the Virochip (25) protocol to establish a specific, sensitive and semi-quantitative assay that is very useful for analyzing highly heterogeneous solid tumors.

MATERIALS AND METHODS

RNA isolation

The cell lines described in the article were obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA) and cultured as recommended. Frozen unpurified prostate tissues were obtained during routine surgery, and classified pathologically by one of us. The total RNA was extracted with Trizol (Invitrogen, Carlsbad, CA, USA) following the manufacturer's instructions. The primary tumor samples were purified by Qiagen RNeasy mini kit (Qiagen, Valencia, CA, USA).

Microarrays

The exon and junction probes are 30-mer oligonucleotides synthesized by Integrated DNA Technologies (Coralville, IA, USA) or Illumina/Invitrogen (San Diego, CA, USA) and printed on poly-L-lysine slides at 50 μM along with Human Cot-1 DNA (Invitrogen), which is enriched for repetitive sequences, and herring sperm DNA (Promega, Madison, WI, USA), which was used as nonspecific control. The printing procedure has been described and essentially follows the manual of the DeRisi arrayer with silicon microcontact printing pins (Parallel Synthesis Technologies, Inc. Santa Clara, CA, USA) (25–27). Arrays were postprocessed with succinic anhydride-based method for blocking before hybridization as previously described (27). The protocols related to array printing and hybridization in this article generally can be found in the following link: http://cat.ucsf.edu/equipment/arrayer/index.html.

Probe labeling

The RT–PCR reaction was performed with an OneStep RT-PCR kit (Qiagen) essentially following the manufacturer's protocol, except that the final reaction volume was scaled down to 20 μl. The forward (GTT TCC CAG TCA CGA TCC AGG AGG CGG AGG CGG A) and reverse primers (GTT TCC CAG TCA CGA TCG GCG TTG TAG CTG GGG GTG AG) are located at exon 6 of ERG and exon 1 of TMPRSS2 respectively, as described (2,9). The 5′-ends of both primers have the sequence of primer B (GTT TCC CAG TCA CGA TC) for the subsequent step of PCR labeling with a single primer B as described previously (25).

The procedure is a modification of the previously reported 'Round ABC' protocol (25). Briefly, the RT–PCR reaction was assembled at 4°C in a PCR workstation and transferred to a thermocycler with the block preheated to 50°C. The initial reverse transcription was performed at 50°C for 30 min and followed by 95°C for 15 min to activate HotStarTaq DNA polymerase as well as to inactivate the reverse transcriptases (Round A). The PCR conditions were 35 cycles at 92°C for 30 s, 55°C for 30 s and 68°C for 1.5 min with a final extension step at 68°C for 5 min (Round B). One microliter of unpurified product was subsequently used as a template for another 20 cycles of amplification to label the amplicons via a previously described 'Round C' PCR protocol (94°C for 30 s, 40°C for 30 s, 50°C for 30 s and 72°C for 1 min) with primer B and a 4:1 mixture of aminoallyl dUTP (Ambion, Austin, TX, USA) and dTTP for probe labeling (25). The labeled amplicons were purified with DNA Clean-up and Concentrator-5 columns (Zymo Research, Orange, CA, USA), eluted in 9 μl of sodium bicarbonate (pH 9.0) and coupled with 1 μl of DMSO dissolved Cy3 NHS esters (GE Healthcare, Piscataway, NJ, USA) for 30–60 min. The Cy3-labeled amplicons were purified with DNA Clean-up and Concentrator-5 columns and eluted with 10 μl of 10 mM Tris–HCl (pH 8.0). Then, the Cy3-labeled amplicons were diluted in water and combined with 3.6 μl of 20× SSC, 0.5 μl of HEPES (pH 7.0) and finally 0.5 μl of 10% SDS to reach final volume of 25 μl. The mixed solution was heated for 2 min at 95°C, cooled to room temperature and hybridized to the exon mapping arrays at 63°C overnight essentially as previously described (25–27). The hybridized arrays were washed and scanned with a GenePix 4000B scanner (Molecular Device, Sunnyvale, CA, USA) and analyzed by GenePix Pro 6.0 software.

2D DNA-FISH

Frozen prostate cancer samples were sectioned onto slides. Cell nuclei were isolated *in situ* with ice-cold cytoskeleton buffer (CSK: 300 mM sucrose, 100 mM NaCl, 10 mM PIPES, 3 mM $MgCL_2$, 1 mM EGTA and 0.5% Triton X-100) (28). The slides were fixed by dipping in ice-cold methanol for 3 min, followed by ice-cold acetone. After air drying the slides, they were allowed to age for at least 1 week in ethanol.

DNA-FISH was carried out according to a method for single copy loci detection (28). The protocol was adapted with few modifications, using 50-mer oligonucleotides specific to the loci of interest and labeled with a desired hapten. Two probes were used for FISH for break-apart assay. A probe located at the promoter region of TMPRSS2 was labeled with biotin (Bio/GACTCCA GGAGCGCTCCCCAGAATCCCCTTCCTTAACCCA AACTCGAGCC). The other probe at exon 2 of ERG was labeled with 5′-6-carboxyfluorescein (56FAM) (56FAM/ GATCTTTGGAGACCCGAGGAAAGCCGTGTTGA CCAAAAGCAAGACAAATG). Detection of probes was achieved by using antibodies conjugated to quantum dots

(Qdot) against the hapten label. Conditions were optimized to use a combination of two antibodies (1 : 200) obtained from Invitrogen–Molecular Probes™ (Qdot 655 * sheep anti-Bio primary antibody conjugate; Qdot 525 * Goat anti-FITC whole IgG primary antibody conjugate). Image acquisition was done with a Zeiss Axioplan 2e microscope (Carl Zeiss, Inc.). All pictures in the corresponding three channels were deconvolved and optical sections merged to produce 2D pictures using Axiovision 4.0 software (Carl Zeiss, Inc.) and Image J (Rasband, W.S., ImageJ, U. S. National Institutes of Health, Bethesda, MD, USA, http://rsb.info.nih.gov/ij/, 1997–2006.)

RESULTS

Microarray-based TMPRSS2:ERG exon fusion mapping

To develop a multiplexing assay that is highly sensitive in clinical samples of high complexity, we adopted our Virochip system (25). The key protocol, Round ABC, designed for unbiased amplification (29), is crucial for identifying various fusion variants in this application.

Through literature review, we found that most of the TMPRSS2:ERG fusion junctions are between exons 1 or 2 of the TMPRSS2 and exons 2–5 of the ERG (2,7,9,13). Such constraints perhaps are related to whether a functional ERG protein can be made from the gene fusions (9). Therefore, we initially used a pair of primers at exon 1 of the TMPRSS2 and exon 6 of the ERG for RT–PCR. As shown in Figure 1A, PCR products were only generated when there was a gene fusion, since the two primers are located at different genes. Subsequently, the PCR products were labeled and hybridized to an exon array

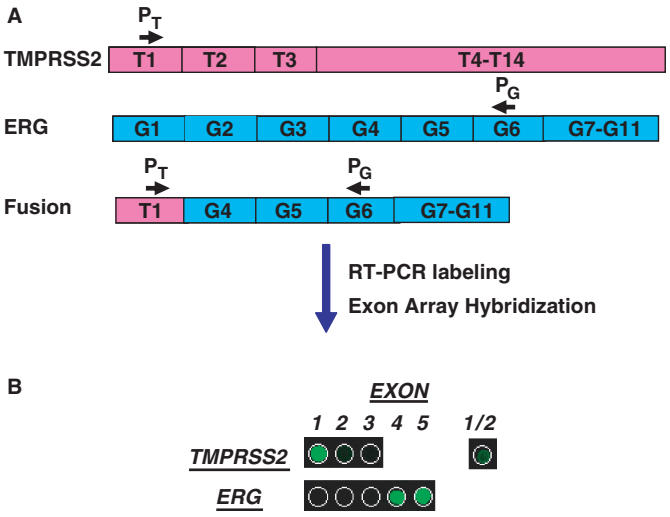

Figure 1. TMPRSS2:ERG exon mapping strategy. (**A**) The RT–PCR is performed using a 3′ primer from exon 6 of the ERG and a 5′ primer from exon 1 of the TMPRSS2. Only fusion transcripts can be exponentially amplified since the two primers are at different genes. The probes on the array are derived from exons 1–3 of the TMPRSS2 and exons 1–5 of the ERG (Table 1). (**B**) The hybridization pattern of RT–PCR labeled amplicons with total RNA derived from the VCaP cell line. The result clearly shows that the fusion junction is at exon 1 of the TMPRSS2 and exon 4 of the ERG, as illustrated in the fusion scenario in (A). A probe that spans on the junction of exon 1 and 2 of the TMPRSS2 is labeled as '1/2'.

for mapping the exons near the fusion junction. Printed on the array are 30-mer oligonucleotide probes derived from exons 1–3 of the TMPRSS2 and exons 1–5 of the ERG (Table 1). Each selected sequence is represented by two complementary probes (F: forward and R: reverse complement) since sometimes PCR-labeled amplicons may bind to only one strand of the probe, based on empirical observations (25). We observed that probes with reverse complementary (R) orientation worked consistently with our RT–PCR labeling protocol.

A prostate cancer cell line, VCaP, (30) with a TMPRSS2 and ERG fusion (2) was used for initial feasibility testing. The total RNA was subjected to RT–PCR with a pair of primers located at exon 6 of ERG and exon 1 of TMPRSS2 (2,9). The unpurified product was labeled and hybridized on the microarray (Figure 1B). Only spots corresponding to exon 1 of TMPRSS2 and exons 4–5 of ERG developed strong signals. This result indicates the fusion junction is at the exon 1 of TMPRSS2 and exon 4 of ERG, which is consistent with a previous report (2).

To mimic a typical clinical situation, in which small population of cancer cells are present among normal host cells in a primary tumor, we spiked decreasing amounts of total RNA extracted from VCaP cells into an excess of HeLa RNA, which does not have the fusion transcripts. The detection limit is 32 pg of VCaP RNA in the presence of 100 ng of HeLa RNA (Figure 2). This translates into only 1–3 cancer cells in the presence of 3000 times more normal cells. The level of sensitivity is superior to previous methods for detecting fusion transcripts (24).

TMPRSS2:ERG fusion junction mapping in primary prostate cancer

To test the ability of the exon mapping array to detect and characterize TMPRSS2:ERG fusion transcripts in clinical samples, we isolated total RNA from frozen unpurified primary prostate tissues obtained during surgery.

Table 1. Exon probes

Name	Sequence
T1F	GGGCGGGGAGCGCCGCCTGGAGCGCGGCAG
T2F	ACATTCCAGATACCTATCATTACTCGATGC
T3F	GGTCACCACCAGCTATTGGACCTTACTATG
T1/2F	TGGAGCGCGGCAGGTCATATTGAACATTCC
G1F	AGGGACATGAGAGAAGAGGAGCGGCGCTCA
G2F	AGACCCGAGGAAAGCCGTGTTGACCAAAAG
G3F	GCTGGTAGATGGGCTGGCTTACTGAAGGAC
G4F	TTATCAGTTGTGAGTGAGGACCAGTCGTTG
G5F	CTCTCCTGATGAATGCAGTGTGGCCAAAGG
T1R	CTGCCGCGCTCCAGGCGGCGCTCCCCGCCC
T2R	GCATCGAGTAATGATAGGTATCTGGAATGT
T3R	CATAGTAAGGTCCAATAGCTGGTGGTGACC
T1/2R	GGAATGTTCAATATGACCTGCCGCGCTCCA
G1R	TGAGCGCCGCTCCTCTTCTCTCATGTCCCT
G2R	CTTTTGGTCAACACGGCTTTCCTCGGGTCT
G3R	GTCCTTCAGTAAGCCAGCCCATCTACCAGC
G4R	CAACGACTGGTCCTCACTCACAACTGATAA
G5R	CCTTTGGCCACACTGCATTCATCAGGAGAG

T, TMPRSS2; G, ERG. F, forward probe; R, reverse complement probe.

Many of these tumors had a substantial fraction of normal stromal cells. Total RNA (5–50 ng) from prostate cancers ($n = 20$) and nonmalignant hyperplastic prostate tissues ($n = 10$) were subjected to RT–PCR labeling and array hybridization. The results showed that 7/20 cancers but 0/10 nonmalignant samples had TMPRSS2:ERG fusion genes. To confirm the presence of the gene fusions, direct sequencing was performed for the seven samples. The sequencing data validated the exon fusion findings revealed by the array assays. Similar to other reports (7,12,13), some samples clearly showed two or more bands on the agarose gel when the PCR products were subjected to electrophoresis, corresponding to two or more fusion transcripts in the same specimens.

The multiple fusion transcripts in a single prostate cancer sample may reflect tumor heterogeneity or alternative splicing events. In order to map multiple fusion junctions in a single assay, we redesigned the exon array to include junction probes between exons 1 and 2 of the TMPRSS2 gene and exons 1–6 of the ERG gene (Table 2). The modified probe set showed very clearly that the patient sample #15 had two fusion transcripts and also revealed the relative ratios of the two fusion transcripts through their respective signal intensities (Figure 3). In this case, the two fusion transcripts are between exon 4 of the ERG fused to either exon 1 (T1G4) or exon 2 (T2G4) of the TMPRSS2. The signal intensity of T2G4 junction probe is weaker than that of the T1G4 junction probe (Figure 3A), consistent with the intensities of the probes within the exons. These two fusion transcripts are very likely due to alternative splicing. Figure 4 summarizes the cluster analysis (31) of the seven arrays; the results are shown in Table 3. Multiple fusion variants were found in 4/7 positive samples.

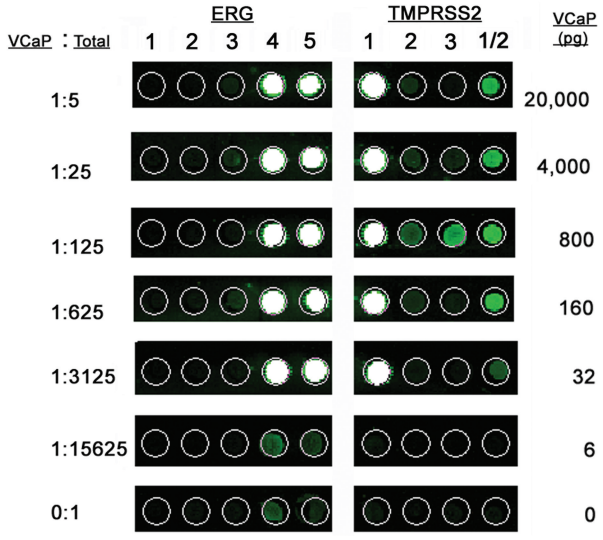

Figure 2. Assay sensitivity. The VCaP total RNA was serially diluted in a solution containing HeLa RNA to mimic the heterogeneous cell population in primary tumors or human body fluids. The total amount of RNA for each reaction is 100 ng. The laser power (PMT 600, 100% output) was adjusted to maximize the sensitivity of detection. Therefore, the intensity of the each expected feature (T1, G4, G5) is at the saturated level. The signal disappeared when the VCaP RNA was diluted from 1:3125 (32 pg) to 1:15625 (6.4 pg).

Table 3 also lists the percentages of cancer cells in the tumors, the Gleason tumor grades and the detected variants of TMPRSS2:ERG fusion transcripts. In this small sample set, there is no clear association between tumor

Table 2. Junction probes

Name	Sequence
T1G1F	CCTGGAGCGCGGCAGCCCCCGAGGGACATG
T1G2F	CCTGGAGCGCGGCAGGTTATTCCAGGATCT
T1G3F	CCTGGAGCGCGGCAGCCGTCAGGTTCTGAA
T1G4F	CCTGGAGCGCGGCAGGAAGCCTTATCAGTT
T1G5F	CCTGGAGCGCGGCAGATGCCACCCCCAAAC
T1G6F	CCTGGAGCGCGGCAGATCCTACGCTATGGA
T2G1F	ATGGCTTTGAACTCACCCCCGAGGGACATG
T2G2F	ATGGCTTTGAACTCAGTTATTCCAGGATCT
T2G3F	ATGGCTTTGAACTCACCGTCAGGTTCTGAA
T2G4F	ATGGCTTTGAACTCAGAAGCCTTATCAGTT
T2G5F	ATGGCTTTGAACTCAATGCCACCCCCAAAC
T2G6F	ATGGCTTTGAACTCAATCCTACGCTATGGA
T1G1R	CATGTCCCTCGGGGGCTGCCGCGCTCCAGG
T1G2R	AGATCCTGGAATAACCTGCCGCGCTCCAGG
T1G3R	TTCAGAACCTGACGGCTGCCGCGCTCCAGG
T1G4R	AACTGATAAGGCTTCCTGCCGCGCTCCAGG
T1G5R	GTTTGGGGGTGGCATCTGCCGCGCTCCAGG
T1G6R	TCCATAGCGTAGGATCTGCCGCGCTCCAGG
T2G1R	CATGTCCCTCGGGGGTGAGTTCAAAGCCAT
T2G2R	AGATCCTGGAATAACTGAGTTCAAAGCCAT
T2G3R	TTCAGAACCTGACGGTGAGTTCAAAGCCAT
T2G4R	AACTGATAAGGCTTCTGAGTTCAAAGCCAT
T2G5R	GTTTGGGGGTGGCATTGAGTTCAAAGCCAT
T2G6R	TCCATAGCGTAGGATTGAGTTCAAAGCCAT

T, TMPRSS2; G, ERG. F, forward probe; R, reverse complement probe.

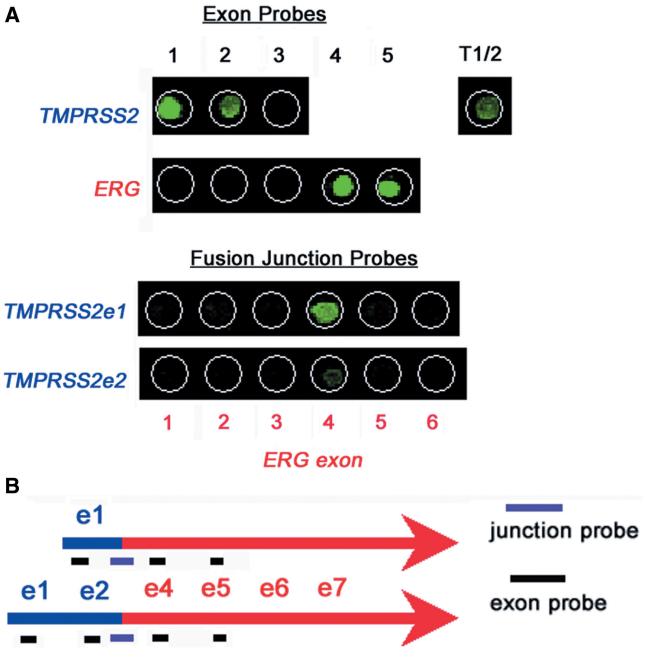

Figure 3. Detecting multiple fusion transcripts with junction probes. (**A**) The exon probes alone show the sum of the signal derived from individual transcripts. The junction probes reveal the species of the transcripts. These two sets of data together are very useful to distinguish weak but true signals from otherwise random background signals. (**B**) A scheme is presented to assist with data interpretation for (A), which shows coexistence of T1G4 and T2G4 transcripts.

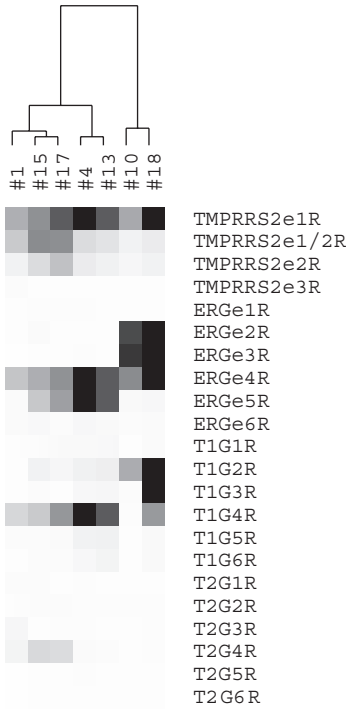

Figure 4. Cluster analysis of seven prostate cancer samples having fusion transcripts. The signal intensity of each feature is divided by the intensity of a nonspecific control (herring sperm DNA) to normalize the data for cluster analysis. The result is shown in Table 3. The samples having similar fusion transcript variants were clustered together by the program.

Table 3. The percentage of cancer cells in the tumors, the Gleason tumor grades and the detected variants of TMPRSS2:ERG fusion transcripts

Sample no.	Cancer (%)	Gleason grade	Fusion transcripts
1	30	7	T1-G4; T2-G4
2	20	5	
3	50	5	
4	20	6	T1-G4
5	80	9	
6	1	6	
7	90	8	
8	20	4	
9	80	8	
10	1	6	T1-G2
11	2	6	
12	70	7	
13	20	9	T1-G4
14	1	6	
15	70	8	T1-G4; T2-G4
16	20	8	
17	80	8	T1-G4; T2-G4
18	50	7	T1-G2; T1-G3; T1-G4
19	80	7	
20	80	7	

grade and the presence of fusion transcripts. A relatively larger study also showed that the presence of fusion transcript was associated with tumor stage but not tumor grade (12). Consistent with the VCaP titration study

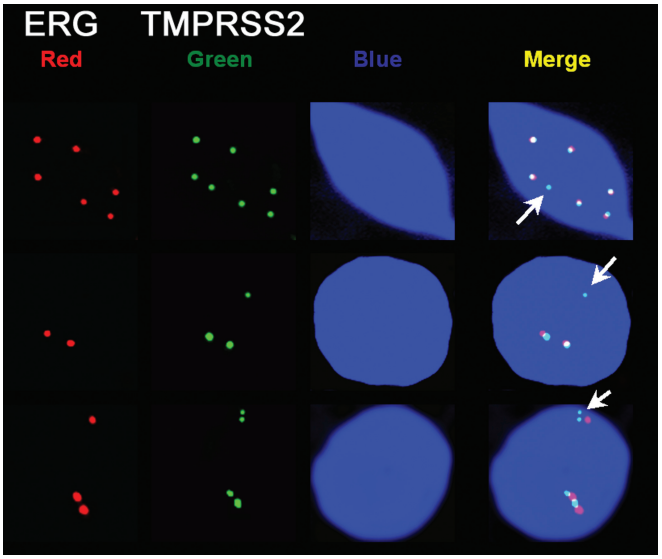

Figure 5. Heterogeneity of FISH patterns of interstitial deletion between TMPRSS2 and ERG in a primary prostate tumor. An unpaired green dot (TMPRSS2 probe, indicated by arrows) suggests an interstitial deletion. Nonrandom variation of FISH patterns is shown by the fact that most of the green and red signals (two different but nearby probes) are paired in each panel. This variation is expected on a heterogeneous aneuploid cancer cell population, which often makes it difficult to distinguish meaningful events from random background aberrations.

(Figure 2), the clinical assay can detect the fusion transcript when only 1% tumor cells is present in the prostate tissue (sample 10).

FISH analysis

We used FISH analysis to independently confirm our array approach. Our FISH procedure (28) employed 50-mer probes that were labeled with small haptens for target hybridization in conjunction with individual hapten-specific, quantum dot conjugated antibodies for signal detection. The resolution of this method is ~50 kb. We designed two probes for the FISH assays, one at the promoter region of TMPRSS2 (green in Figure 5) and the other at exon 2 of ERG (red in Figure 5). We observed heterogeneity of the FISH patterns in some primary prostate cancer samples (Figure 5). It is more difficult to find interstitial deletions between TMPRSS2 and ERG in tumor samples containing low percentages of cancer cells by FISH. Therefore, we used samples that contained >80% cancer cells without detectable fusion genes to confirm the results of the arrays. The FISH experiments revealed no genomic deletions at this location for all the selected samples (#5, #7, #9, #19 and #20) that were similarly nondeleted by array hybridization.

DISCUSSION

We have established a simple assay that can concurrently profile variants of TMPRSS2:ERG fusion transcripts by combining a single RT–PCR with an exon array. The modified 'Round ABC' protocol, which was

originally designed for genomic amplification (29) and has been widely adopted for chromatin immunoprecipitation (ChIP) and whole-genome DNA microarrays (ChIP-chip) (32,33) and Virochip (25,34) experiments, is a simple and relatively unbiased amplification procedure to semi-quantitatively measure the fusion variants in a complex sample. Previously, the same simple procedure was used to obtain 83% (25 kb/30 kb) of the SARS coronavirus genome with total nucleic acids isolated from a viral culture (25).

The inclusion within the array of probes derived from individual exons and potential fusion junctions simplifies the breakpoint mapping and increases the confidence of data interpretation (Figures 3 and 4). In contrast to some reports that used multiple primers targeted to every potential fusion junction in hematological malignancies (35–38), we used a single set of primers for target amplification (Figure 1). The fusion junctions were subsequently decoded by array. This design significantly reduces the problems associated with primer dimers in the multiplex PCR reaction, and creates more room for future assays to include additional fusion genes. Furthermore, most searches for fusion genes have focused on blood cancers, because the cells can be purified before analysis. The application of the previous methodologies is less useful for highly complex solid tumors that are inevitably admixed with normal cells. For example, a previously reported MLLFusionChip could not be applied to samples with cancer cells of <5–10% in 1 μg of total RNA (39).

The current assay should facilitate a thorough compilation of the gene fusion variants in primary prostate specimens, which may be useful for stratifying the aggressiveness of prostate cancer (13). In this regard, fusion variants of EWS with another member of the ETS family, FLI1, have been shown to be an independent predictor of disease progression in Ewing's sarcoma (40,41). It will be of interest to compare in transfected cells the biological activities of the different TMPRSS2:ERG variants from patients with contrasting clinical outcomes (41).

While some studies have suggested that the presence of TMPRSS2:ERG fusions is associated with more aggressive disease or higher Gleason tumor grade, other investigators did not reach the same conclusion (12,14,17,20,23,42). We also did not find such an association in a small series of samples. However, all of these results are defective due to small sample size. The technology described here should make possible a larger scale investigation to find whether there is a correlation between the aggressiveness of the disease and the presence of specific fusion genes.

It is crucial to have true cancer-specific biomarkers for early cancer detection as well as for minimal residual disease monitoring, which has been extensively demonstrated in hematologic maligancies (43). Such biomarkers could help to avoid under- or over-treatment. Thus, there is past interest (24,44) in applying TMPRSS2:ERG fusion assays for such application since PSA and many other markers in development are not truly prostate cancer-specific (45). A recent study reported a TMPRSS2:ERG assay with a sensitivity of detecting 1600 VCaP cells (24). However, this level of sensitivity might not be sufficient for broad clinical application, especially with small biopsy specimens or urine samples. We were able to achieve an assay sensitivity of <32 pg of total RNA derived from VCaP cells, an equivalent to 1–3 cells (Figure 2). Because our assay is simple and amenable to automation, it is readily adaptable to clinical studies. While it has been challenging to adapt microarray-based technology to the clinic, some tests (e.g. AmpliChip CYP450 and MammaPrint) have been approved by FDA (46).

The same strategy can be applied to detect other less prevalent fusion transcripts (TMPRSS2:ETV1 and TMPRSS2:ETV4) in prostate cancer (2,5,6). In addition, the exon array approach can also be applied to other fusion genes, such as BCR-ABL in CML and clonal Ig/TCR rearrangements in lymphocytic malignancies. While this methodology development was motivated by the clinical need, it is generally applicable to other research requirements that are analogous to the situation for detecting fusion genes in the single cell level when a large excess of normal cells are present. For example, a developmental biologist may use a similar approach to screen mutants that have a desirable gene fusion when direct gene targeting is not feasible.

There are some shortcomings of using RNA transcripts as prostate cancer biomarkers, despite our ability to achieve very sensitive detection of TMPRSS2:ERG fusion variants. First, RNA is unstable and difficult to process in routine clinical assays. Second, commonly used drugs that inhibit androgen growth pathways, including GnRH agonists and testosterone antagonists, may diminish the production of the TMPRSS2:ERG mRNA fusion transcript, thereby producing false-negative results in patients on hormonal therapy with evolving androgen-independent tumors. Indeed, it has been reported that TMPRSS2:ERG mRNA fusion transcripts are not expressed in androgen-independent tumors in spite of the presence of interstitial deletions in between TMPRSS2 and ERG at chromosome 21q22 (10). While FISH is useful for identifying genomic rearrangements, it has relatively lower resolution and is difficult to use in highly heterogeneous samples with small percentages of tumor cells. We have recently developed a technology, designated Primer Approximation Multiplex PCR (PAMP) for identifying breakpoints in genomic DNA without the need to purify cancer cells from normal tissues (26). We are optimizing this assay for detecting the breakpoints between TMPRSS2 and ERG loci for primary prostate tumors to overcome any potential problems associated with RNA based biomarkers. In addition, the DNA-based assay will provide information about whether multiple fusion transcripts in a sample are derived from alternative splicing or tumor heterogeneity. The best approach may ultimately be to combine DNA and RNA based assays in a common format.

FUNDING

National Institutes of Health (CA119335 to UCSD NanoTumor Center of Excellence for Cancer Nanotechnology, CA133634 to Y. T. Liu, NS034934 to

M.G. Rosenfeld); NCI Cancer Center (2 P30 CA023100-23 to DNA Sequencing Shared Resource, UCSD Cancer Center) partially. Funding for open access charge: CA133634.

Conflict of interest statement. None declared.

REFERENCES

1. Mitelman,F., Johansson,B. and Mertens,F. (2007) The impact of translocations and gene fusions on cancer causation. *Nat. Rev. Cancer*, **7**, 233–245.
2. Tomlins,S.A., Rhodes,D.R., Perner,S., Dhanasekaran,S.M., Mehra,R., Sun,X.W., 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.
3. Melo,J.V. and Barnes,D.J. (2007) Chronic myeloid leukaemia as a model of disease evolution in human cancer. *Nat. Rev. Cancer*, **7**, 441–453.
4. Hughes,T., Deininger,M., Hochhaus,A., Branford,S., Radich,J., Kaeda,J., Baccarani,M., Cortes,J., Cross,N.C., Druker,B.J. *et al.* (2006) Monitoring CML patients responding to treatment with tyrosine kinase inhibitors: review and recommendations for harmonizing current methodology for detecting BCR-ABL transcripts and kinase domain mutations and for expressing results. *Blood*, **108**, 28–37.
5. Tomlins,S.A., Mehra,R., Rhodes,D.R., Smith,L.R., Roulston,D., Helgeson,B.E., Cao,X., Wei,J.T., Rubin,M.A., Shah,R.B. *et al.* (2006) TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. *Cancer Res.*, **66**, 3396–3400.
6. Lapointe,J., Kim,Y.H., Miller,M.A., Li,C., Kaygusuz,G., van de Rijn,M., Huntsman,D.G., Brooks,J.D. and Pollack,J.R. (2007) A variant TMPRSS2 isoform and ERG fusion product in prostate cancer with implications for molecular diagnosis. *Mod. Pathol.*, **20**, 467–473.
7. Soller,M.J., Isaksson,M., Elfving,P., Soller,W., Lundgren,R. and Panagopoulos,I. (2006) Confirmation of the high frequency of the TMPRSS2/ERG fusion gene in prostate cancer. *Genes Chromosomes Cancer*, **45**, 717–719.
8. Cerveira,N., Ribeiro,F.R., Peixoto,A., Costa,V., Henrique,R., Jeronimo,C. and Teixeira,M.R. (2006) TMPRSS2-ERG gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. *Neoplasia*, **8**, 826–832.
9. Clark,J., Merson,S., Jhavar,S., Flohr,P., Edwards,S., Foster,C.S., Eeles,R., Martin,F.L., Phillips,D.H., Crundwell,M. *et al.* (2007) Diversity of TMPRSS2-ERG fusion transcripts in the human prostate. *Oncogene*, **26**, 2667–2673.
10. Hermans,K.G., van Marion,R., van Dekken,H., Jenster,G., van Weerden,W.M. and Trapman,J. (2006) 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.*, **66**, 10658–10663.
11. Iljin,K., Wolf,M., Edgren,H., Gupta,S., Kilpinen,S., Skotheim,R.I., Peltola,M., Smit,F., Verhaegh,G., Schalken,J. *et al.* (2006) TMPRSS2 fusions with oncogenic ETS factors in prostate cancer involve unbalanced genomic rearrangements and are associated with HDAC1 and epigenetic reprogramming. *Cancer Res.*, **66**, 10242–10246.
12. Perner,S., Demichelis,F., Beroukhim,R., Schmidt,F.H., Mosquera,J.M., Setlur,S., Tchinda,J., Tomlins,S.A., Hofer,M.D., Pienta,K.G. *et al.* (2006) TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res.*, **66**, 8337–8341.
13. Wang,J., Cai,Y., Ren,C. and Ittmann,M. (2006) Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. *Cancer Res.*, **66**, 8347–8351.
14. Demichelis,F., Fall,K., Perner,S., Andren,O., Schmidt,F., Setlur,S.R., Hoshida,Y., Mosquera,J.M., Pawitan,Y., Lee,C. *et al.* (2007) TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene*, **26**, 4596–4599.
15. Mehra,R., Tomlins,S.A., Shen,R., Nadeem,O., Wang,L., Wei,J.T., Pienta,K.J., Ghosh,D., Rubin,M.A., Chinnaiyan,A.M. *et al.* (2007) Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. *Mod. Pathol.*, **20**, 538–544.
16. Nam,R.K., Sugar,L., Wang,Z., Yang,W., Kitching,R., Klotz,L.H., Venkateswaran,V., Narod,S.A. and Seth,A. (2007) Expression of TMPRSS2 ERG gene fusion in prostate cancer cells is an important prognostic factor for cancer progression. *Cancer Biol. Ther.*, **6**, 40–45.
17. Rajput,A.B., Miller,M.A., De Luca,A., Boyd,N., Leung,S., Hurtado-Coll,A., Fazli,L., Jones,E.C., Palmer,J.B., Gleave,M.E. *et al.* (2007) Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. *J. Clin. Pathol.*, **60**, 1238–1243.
18. Tomlins,S.A., Laxman,B., Dhanasekaran,S.M., Helgeson,B.E., Cao,X., Morris,D.S., 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.
19. Tomlins,S.A., Laxman,B., Varambally,S., Cao,X., Yu,J., Helgeson,B.E., Cao,Q., Prensner,J.R., Rubin,M.A., Shah,R.B. *et al.* (2008) Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia*, **10**, 177–188.
20. Kumar-Sinha,C., Tomlins,S.A. and Chinnaiyan,A.M. (2008) Recurrent gene fusions in prostate cancer. *Nat. Rev. Cancer*, **8**, 497–511.
21. Rubin,M.A. (2008) Targeted therapy of cancer: new roles for pathologists–prostate cancer. *Mod. Pathol.*, **21 (**Suppl. 2), S44–S55.
22. Stuart,R.O., Wachsman,W., Berry,C.C., Wang-Rodriguez,J., Wasserman,L., Klacansky,I., Masys,D., Arden,K., Goodison,S., McClelland,M. *et al.* (2004) In silico dissection of cell-type-associated patterns of gene expression in prostate cancer. *Proc. Natl Acad. Sci. USA*, **101**, 615–620.
23. Yoshimoto,M., Joshua,A.M., Chilton-Macneill,S., Bayani,J., Selvarajah,S., Evans,A.J., Zielenska,M. and Squire,J.A. (2006) Three-color FISH analysis of TMPRSS2/ERG fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. *Neoplasia*, **8**, 465–469.
24. Laxman,B., Tomlins,S.A., Mehra,R., Morris,D.S., Wang,L., Helgeson,B.E., Shah,R.B., Rubin,M.A., Wei,J.T. and Chinnaiyan,A.M. (2006) Noninvasive detection of TMPRSS2:ERG fusion transcripts in the urine of men with prostate cancer. *Neoplasia*, **8**, 885–888.
25. Wang,D., Urisman,A., Liu,Y.T., Springer,M., Ksiazek,T.G., Erdman,D.D., Mardis,E.R., Hickenbotham,M., Magrini,V., Eldred,J. *et al.* (2003) Viral discovery and sequence recovery using DNA microarrays. *PLoS Biol.*, **1**, E2.
26. Liu,Y.T. and Carson,D.A. (2007) A novel approach for determining cancer genomic breakpoints in the presence of normal DNA. *PLoS ONE*, **2**, e380.
27. Eisen,M.B. and Brown,P.O. (1999) DNA arrays for analysis of gene expression. *Methods Enzymol.*, **303**, 179–205.
28. Lunyak,V.V., Prefontaine,G.G., Nunez,E., Cramer,T., Ju,B.G., Ohgi,K.A., Hutt,K., Roy,R., Garcia-Diaz,A., Zhu,X. *et al.* (2007) Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science.*, **317**, 248–251.
29. Bohlander,S.K., Espinosa,R. III, Le Beau,M.M., Rowley,J.D. and Diaz,M.O. (1992) A method for the rapid sequence-independent amplification of microdissected chromosomal material. *Genomics*, **13**, 1322–1324.
30. Korenchuk,S., Lehr,J.E., Mclean,L., Lee,Y.G., Whitney,S., Vessella,R., Lin,D.L. and Pienta,K.J. (2001) VCaP, a cell-based model system of human prostate cancer. *In Vivo*, **15**, 163–168.
31. Eisen,M.B., Spellman,P.T., Brown,P.O. and Botstein,D. (1998) Cluster analysis and display of genome-wide expression patterns. *Proc. Natl Acad. Sci. USA*, **95**, 14863–14868.
32. Lieb,J.D., Liu,X., Botstein,D. and Brown,P.O. (2001) Promoter-specific binding of Rap1 revealed by genome-wide maps of protein-DNA association. *Nat. Genet.*, **28**, 327–334.
33. Buck,M.J. and Lieb,J.D. (2004) ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics*, **83**, 349–360.

34. Palacios,G., Quan,P.L., Jabado,O.J., Conlan,S., Hirschberg,D.L., Liu,Y., Zhai,J., Renwick,N., Hui,J., Hegyi,H. *et al.* (2007) Panmicrobial oligonucleotide array for diagnosis of infectious diseases. *Emerg. Infect. Dis.*, **13**, 73–81.
35. Shi,R.Z., Morrissey,J.M. and Rowley,J.D. (2003) Screening and quantification of multiple chromosome translocations in human leukemia. *Clin. Chem.*, **49**, 1066–1073.
36. Nasedkina,T.V., Zharinov,V.S., Isaeva,E.A., Mityaeva,O.N., Yurasov,R.N., Surzhikov,S.A., Turigin,A.Y., Rubina,A.Y., Karachunskii,A.I., Gartenhaus,R.B. *et al.* (2003) Clinical screening of gene rearrangements in childhood leukemia by using a multiplex polymerase chain reaction-microarray approach. *Clin. Cancer Res.*, **9**, 5620–5629.
37. Zhang,Q.Y., Garner,K. and Viswanatha,D.S. (2002) Rapid detection of leukemia-associated translocation fusion genes using a novel combined RT-PCR and flow cytometric method. *Leukemia*, **16**, 144–149.
38. Nasedkina,T., Domer,P., Zharinov,V., Hoberg,J., Lysov,Y. and Mirzabekov,A. (2002) Identification of chromosomal translocations in leukemias by hybridization with oligonucleotide microarrays. *Haematologica*, **87**, 363–372.
39. Maroc,N., Morel,A., Beillard,E., De La Chapelle,A.L., Fund,X., Mozziconacci,M.J., Dupont,M., Cayuela,J.M., Gabert,J., Koki,A. *et al.* (2004) A diagnostic biochip for the comprehensive analysis of MLL translocations in acute leukemia. *Leukemia*, **18**, 1522–1530.
40. de Alava,E., Kawai,A., Healey,J.H., Fligman,I., Meyers,P.A., Huvos,A.G., Gerald,W.L., Jhanwar,S.C., Argani,P., Antonescu,C.R. *et al.* (1998) EWS-FLI1 fusion transcript structure is an independent determinant of prognosis in Ewing's sarcoma. *J. Clin. Oncol.*, **16**, 1248–1255.
41. Lin,P.P., Brody,R.I., Hamelin,A.C., Bradner,J.E., Healey,J.H. and Ladanyi,M. (1999) Differential transactivation by alternative EWS-FLI1 fusion proteins correlates with clinical heterogeneity in Ewing's sarcoma. *Cancer Res.*, **59**, 1428–1432.
42. Saramaki,O.R., Harjula,A.E., Martikainen,P.M., Vessella,R.L., Tammela,T.L. and Visakorpi,T. (2008) TMPRSS2:ERG fusion identifies a subgroup of prostate cancers with a favorable prognosis. *Clin. Cancer Res.*, **14**, 3395–3400.
43. Szczepanski,T. (2007) Why and how to quantify minimal residual disease in acute lymphoblastic leukemia? *Leukemia*, **21**, 622–626.
44. Smit,F., Verhaegh,G.W. and Schalken,J.A. (2006) The TMPRSS2-ERG fusion gene can be found in urinary sediments after DRE. *Eur. Urol. Suppl.*, **5**, 787–787.
45. Bradford,T.J., Tomlins,S.A., Wang,X. and Chinnaiyan,A.M. (2006) Molecular markers of prostate cancer. *Urol. Oncol.*, **24**, 538–551.
46. Jordan,B.R. (2007) DNA microarrays in the clinic: how soon, how extensively? *Bioessays*, **29**, 699–705.