Mapping Isoform Abundance and Interactome of the Endogenous TMPRSS2-ERG Fusion Protein by Orthogonal Immunoprecipitation–Mass Spectrometry Assays

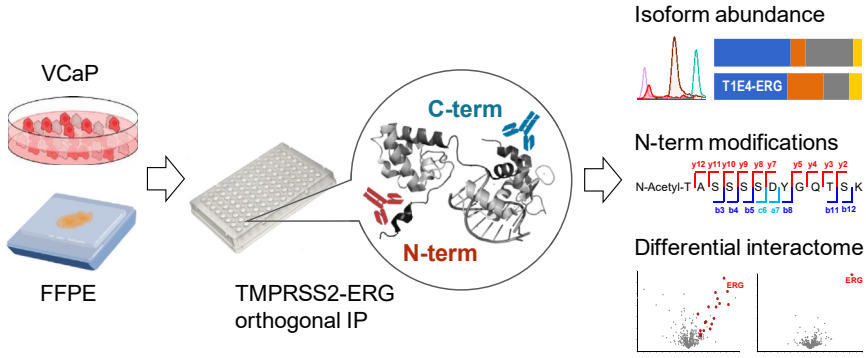
Authors

Zhiqiang Fu, Yasmine Rais, Tarek A. Bismar, M. Eric Hyndman, X. Chris Le, and Andrei P. Drabovich

Correspondence

andrei.drabovich@ualberta.ca

Graphical Abstract



In Brief

Orthogonal immunoprecipitation-mass spectrometry assays quantified TMPRSS2-ERG fusion protein (~27,000 copies/cell) and its four distinct isoforms, and revealed that T1E4-ERG isoform accounted for 52 ± 3% of the total ERG in VCaP cells and 50 ± 11% in FFPE prostate cancer tissues. Methionine-truncated and N-acetylated peptide TASSSSDYGQTSK unique for T1/E4 TMPRSS2-ERG fusion was identified. Unlike the N-terminal antibodies, C-terminal antibodies identified 29 ERG-interacting proteins, including mutually exclusive BRG1- and BRM-associated canonical SWI/SNF chromatin remodeling complexes. Clinical perspectives of assays were discussed.

Highlights

- Orthogonal IP_MS assays revealed four distinct isoforms of TMPRSS2-ERG fusion protein
- T1E4-ERG isoform accounted for 52% of total ERG in VCaP cells, 50% in PCa tissues
- Unique N-terminal modifications of T1/E4 TMPRSS2-ERG fusion protein were identified
- C-term mAb revealed 29 ERG-interacting proteins including canonical SWI/SNF complexes

2021, Mol Cell Proteomics *20*, 100075

https://doi.org/10.1016/j.mcpro.2021.100075

Mapping Isoform Abundance and Interactome of the Endogenous TMPRSS2-ERG Fusion Protein by Orthogonal Immunoprecipitation–Mass Spectrometry Assays

Zhiqiang Fu[1,2], Yasmine Rais[1], Tarek A. Bismar[3], M. Eric Hyndman[4], X. Chris Le[1], and Andrei P. Drabovich[1,*]

TMPRSS2-ERG gene fusion, a molecular alteration found in nearly half of primary prostate cancer cases, has been intensively characterized at the transcript level. However limited studies have explored the molecular identity and function of the endogenous fusion at the protein level. Here, we developed immunoprecipitation–mass spectrometry assays for the measurement of a low-abundance T1E4 TMPRSS2-ERG fusion protein, its isoforms, and its interactome in VCaP prostate cancer cells. Our assays quantified total ERG (~27,000 copies/cell) and its four unique isoforms and revealed that the T1E4-ERG isoform accounted for 52 ± 3% of the total ERG protein in VCaP cells, and 50 ± 11% in formalin-fixed paraffin-embedded prostate cancer tissues. For the first time, the N-terminal peptide (methionine-truncated and N-acetylated TASSSSDYGQTSK) unique for the T1/E4 fusion was identified. ERG interactome profiling with the C-terminal, but not the N-terminal, antibodies identified 29 proteins, including mutually exclusive BRG1- and BRM-associated canonical SWI/SNF chromatin remodeling complexes. Our sensitive and selective IP-SRM assays present alternative tools to quantify ERG and its isoforms in clinical samples, thus paving the way for development of more accurate diagnostics of prostate cancer.

Prostate cancer is the most frequently diagnosed neoplasm and the third leading cause of cancer mortality in men. Introduction of prostate-specific antigen (PSA) testing revolutionized the practice of urologic oncology (1), facilitated earlier detection of localized tumors, and resulted in the active surveillance as a treatment option for many patients with low-grade prostate cancer. PSA test, however, is prone to overdiagnosis and unable to differentiate between indolent and aggressive cancers (2, 3). The race for prognostic biomarkers continues, with numerous genomic, transcriptomic, proteomic, and metabolomic markers being recently discovered and validated (4, 5). The most promising biomarkers are also explored for the molecular mechanisms of their differential expression or regulation (6–8).

Recent genomic studies on the primary prostate adenocarcinoma revealed major subtypes defined by gene fusions of E26 transformation-specific (ETS) transcription factors and mutations in SPOP, FOXA1, and IDH1 genes (9, 10). The most common genomic subtype of primary prostate cancer was represented by the fusion of an androgen-responsive gene TMPRSS2 with a transcription factor ERG (~50% of all cases) (11). While TMPRSS2-ERG rearrangement is heterogeneous, the fusion of TMPRSS2 exon 1 with ERG exon 4 (T1/E4) occurs in ~80% of all TMPRSS2-ERG cases (12). Functionally, TMPRSS2-ERG fusion results in the androgen-dependent overexpression of the N-terminally truncated ERG protein and its isoforms, which may contribute to oncogenic transformation of prostate epithelial cells (Fig. 1*A*) (13).

Numerous studies evaluated TMPRSS2-ERG fusion mRNA as a prognostic biomarker, but the fusion mRNA revealed either positive, or negative, or no association with the clinical significance, progression, or aggressiveness of prostate cancer (14, 15). At the protein level, quantification of the total ERG has been suggested as a surrogate marker for the TMPRSS2-ERG fusion. Measurement of ERG protein expression by

From the [1]Division of Analytical and Environmental Toxicology, Department of Laboratory Medicine and Pathology, Faculty of Medicine and Dentistry, University of Alberta, Edmonton, Alberta, Canada; [2]Key Laboratory of Industrial Ecology and Environmental Engineering (Ministry of Education), School of Environmental Science and Technology, Dalian University of Technology, Dalian, China; [3]Department of Pathology and Laboratory Medicine, University of Calgary Cumming School of Medicine, and Alberta Precision Laboratories, Calgary, Alberta, Canada; [4]Division of Urology, Department of Surgery, Southern Alberta Institute of Urology, University of Calgary, Alberta, Canada

*For correspondence: Andrei P. Drabovich, andrei.drabovich@ualberta.ca.

https://doi.org/10.1016/j.mcpro.2021.100075

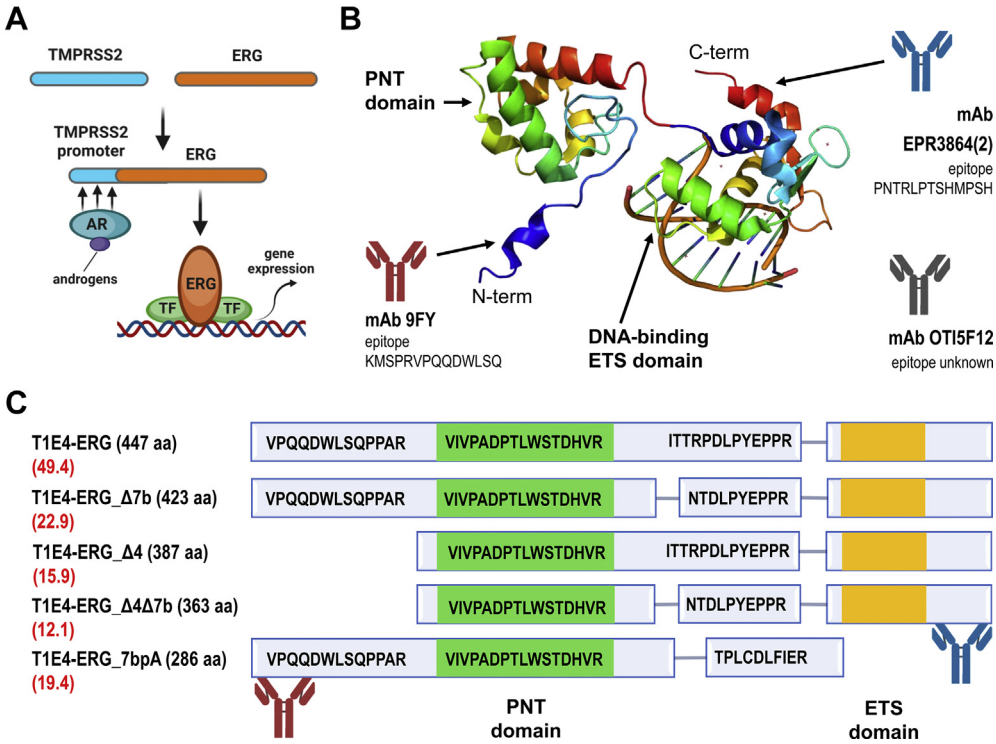


FIG. 1. **Experimental setup.** *A*, androgen receptor (AR)-mediated overexpression of the TMPRSS2-ERG fusion promotes, in coordination with other transcription factors (TF) or transcriptional regulators, oncogenic transformation of prostate epithelial cells. *B*, schematic illustration of ERG protein structure from an arbitrary combination of the N-term (PDB ID: 1SXE) and C-term PDB structures (PDB ID: 4IRI), and corresponding positions of mAb epitopes. 5F12 mAb epitope is not known; *C*, location of unique and shared peptides within different ERG isoforms (transcripts per million abundance of each mRNA isoform is shown in *red*; CCLE data); PNT and DNA-binding ETS domains are highlighted in *green* and *yellow*, respectively.

immunohistochemistry (IHC) in the prostate needle biopsies (16, 17) was proposed to diagnose limited adenocarcinoma, resolve atypical glands suspicious for adenocarcinoma, detect precancerous lesions, or select patients for targeted therapeutic interventions (18). Antibody-based IHC analysis, however, could not provide any details on the heterogeneity of ERG protein isoforms or posttranslational modifications.

Human ERG gene has 12 predicted protein-coding mRNA isoforms, of which six isoforms were experimentally detected by RNA sequencing in VCaP cells (supplemental Table S1). While the wild-type ERG protein is not expressed in VCaP cells, T1/E4 fusion results in overexpression of the N-term truncated isoforms which retain the function the full ERG. Earlier studies suggested that some ERG isoforms could have distinct molecular functions (19, 20). For instance, high levels of a presumably protein-coding mRNA isoform-6 (ENST00000468474) were detected in VCaP cells and patient tissues (12). As a result, isoform-6 protein (P11308-6) lacking DNA-binding ETS domain was suggested as an inhibitor of the transcriptional activation mediated by the full-length ERG isoforms. Overexpression of ERG protein isoforms in HEK293 cells revealed an inhibitory function of isoform-6 protein (20), but expression of the endogenous isoform-6 protein has never been demonstrated in VCaP cells or prostate tissues. In addition, limited experimental data were available for posttranslational modifications (21), protein domains, and interactomes (22–24) of the endogenous ERG protein.

Quantitative proteomic by mass spectrometry (25–28) is a promising tool to generate novel knowledge on TMPRSS2-ERG heterogeneity at the protein level. He *et al.* pioneered measurements of ERG protein by targeted mass spectrometry, utilizing two-dimensional liquid chromatography separations and selected reaction monitoring (SRM) assays (29, 30). Unique peptides of the total ERG protein were quantified in VCaP cell lysate, achieving limits of detection of 1.8 fg/cell or ~3000 cells spiked into urine.

In our study, we aimed at developing immunoprecipitation-shotgun mass spectrometry (IP-MS) and immunoprecipitation-selected reaction monitoring (IP-SRM) assays for the identification and quantification of the endogenous TMPRSS2-ERG fusion protein and its isoforms. As a model cell line, we selected VCaP prostate cancer epithelial cells, which harbored T1/E4 TMPRSS2-ERG gene fusion, overexpressed T1/E4 fusion mRNA, and expressed detectable levels of the endogenous ERG protein. We hypothesized that sensitive IP-MS and

IP-SRM assays, as well as orthogonal assays designed with the N-term and C-term monoclonal antibodies, would provide novel knowledge on the abundance of the endogenousTMPRSS2-ERG and its isoforms, TMPRSS2-ERG posttranslational modifications, and TMPRSS2-ERG protein-protein interactions.

EXPERIMENTAL PROCEDURES

Hypothesis, Study Design, and Objectives

We hypothesized that the orthogonal IP-MS assays with the N-term and C-term ERG antibodies will provide novel knowledge on the identity and abundance of ERG isoforms, interactome, and posttranslational modifications. Our study was designed to measure the endogenous T1E4 TMPRSS2-ERG fusion protein in VCaP cells. Specific objectives included: (1) develop IP-MS and IP-SRM assays; (2) quantify the relative abundance of TMPRSS2-ERG protein isoforms in VCaP cells and formalin-fixed paraffin-embedded (FFPE) prostate cancer tissues (exploratory stage), (3) identify the N-term modifications of TMPRSS2-ERG protein, (4) identify TMPRSS2-ERG protein interactome in VCaP cells, and (5) investigate if synthetic N-term epitope peptides could disrupt TMPRSS2-ERG interactome.

Chemicals and Cell Culture Reagents

Iodoacetamide, dithiothreitol, L-methionine, and trifluoracetic acid (TFA) were purchased from Thermo Fisher Scientific. Optima water and acetonitrile were purchased from Fisher Scientific. Formic acid (FA) was obtained from Sigma-Aldrich. Stable isotope-labeled peptides (SpikeTides_L and SpikeTide_TQL) were obtained from JPT Peptide Technologies GmbH. VCaP and LNCaP prostate cancer lines were obtained from the American Type Culture Collection. Cell lines were cultured in a humidified incubator at 37 °C and 5% CO_2. Dulbecco's Modified Eagles Medium (HyClone) and RPMI 1640 medium (Gibco) were used to culture VCaP and LNCaP cells, respectively. Media were supplemented with 10% fetal bovine serum (Invitrogen) and 1% penicillin-streptomycin (Invitrogen).

Cell Lysis and Sample Preparation

Cell pellets were lysed in 50 μl of 0.1% RapiGest SF (Waters) with repeated pipetting, vortexing, and probe sonication at 20 kHz. EDTA-free protease inhibitor cocktail (Roche) combined with Benzonase Nuclease (Fisher Scientific) was added prior to cell lysis to reduce proteolysis and digest nucleic acids. To remove cell debris, lysates were centrifuged at 16,000*g* and 4 °C for 10 min. Total protein of lysates was measured with Pierce BCA protein assay kit (Fisher Scientific). Proteins were denaturated, and disulfide bonds were reduced by 10 mM dithiothreitol at 70 °C for 15 min and alkylated with 20 mM iodoacetamide at room temperature (RT) in the dark for 45 min. Digestion was completed overnight at 37 °C using recombinant dimethylated SOLu-trypsin (Sigma-Aldrich) with a trypsin:protein ratio 1:20. Trifluoroacetic acid (1%) was added to cleave and precipitate Rapigest SF, and 1 μl of 0.4 M L-methionine was added to limit methionine oxidation during storage. OMIX C18 10 μl tips (Agilent Technologies) were used for desalting and microextraction of tryptic peptides. Finally, samples were diluted in 5% acetonitrile with 0.1% formic acid.

RT-PCR

Total RNA was extracted from LNCaP and VCaP cells using TRIzoL (Thermo Fisher Scientific). RNA was reversely transcribed to cDNA *via* iScriptTM Reverse Transcription Supermix (Bio-Rad Laboratories). After quantification by NanoVue Plus spectrophotometer (GE Healthcare), 500 ng cDNA was utilized as a template for amplification of TMPRSS2-ERG fusion using Hot Start Taq 2X Master Mix (New England Biolabs) and GeneAmp PCR System 2700 thermal cycler (Applied Biosystems). The forward and reverse PCR primers for T1/E4 fusion included 5'-TAGGCGCGAGCTAAGCAGGAG-3' and 5'-CCATAT TCTTTCACCGCCCACTCC-3' (Integrated DNA Technologies). ERG isoform-6 primers were 5'-GGTAC-GAAAACACCCCTGTG-3' (forward) and 5'-CCAAATCAACA-GAGGCAGAA-3' (reverse); the total ERG primers were 5'-AA CGAGCGCAGAGTTATCGT-3' (forward) and 5'-GTGAGCCTC TGGAAGTCGTC-3' (reverse). The final volume was 25 μl, and an initial denaturation step of 95 °C for 5 min was followed by 40 cycles of 95 °C for 30 s, 59 °C for 30 s, 72 °C for 30 s, and one cycle at 72 °C for 5 min. T1/E4 fusion, isoform-6, and total ERG cDNA were detected by 2% agarose gel electrophoresis (supplemental Fig. S1).

Immunoprecipitation

Rabbit monoclonal anti-ERG antibody EPR3864(2) (C-term epitope PNTRLPTSHMPSH; Abcam), mouse monoclonal antibody 9FY (N-term epitope KMSPRVPQQDWLSQ; BioCare Medical), and mouse monoclonal antibody OTI5F12 (epitope unknown; Origene Technologies) were used as capture antibodies for the enrichment of ERG protein from cell lysates (Fig. 1*B*). Rabbit polyclonal antibody GTX129433 (Genetex) was used to enrich ARI1A_HUMAN. Two microliters of antibody stock was diluted in PBS (pH 7.4), coated (100 μl per well) onto high-binding 96-well polystyrene microplates (#07000128; Greiner Bio-One), and incubated overnight at RT. Since 9FY antibody was provided in Renoir Red solution with carrier proteins, a goat anti-mouse Fcγ fragment-specific antibody (Jackson Immunoresearch Labs) was first coated on plates for 9FY antibody pull-down. After washing with 0.1% Tween 20 in PBS (washing buffer; three times with 200 μl), the plate was blocked for 1 h at RT with 200 μl of blocking buffer (2% BSA in wash buffer). The washing step was repeated, followed by addition of 80 μg total protein VCaP lysates per well and dilution to 100 μl with the dilution buffer (0.1% BSA in wash buffer, 0.2 μm-filtered). After 2 h incubation

with continuous shaking, the plate was finally washed three times with 200 μl washing buffer and three times with 50 mM ammonium bicarbonate. Proteins were digested with trypsin (0.25 ng per well). Heavy isotope-labeled peptide internal standards were mixed and diluted to 100 *f*mol/μl. Three microliters of the internal standard mixture was spiked into each sample before (SpikeTides_TQL peptides) or after (SpikeTides_L peptides) trypsin digestion, and each digest was analyzed in triplicates (10 μl per injection). For the interactome studies, antibody isotype controls included: (i) anti-FOLH1 mouse monoclonal antibody (clone 3B5, Abnova) as an IgG_1 isotype control for OTI5F12 and 9FY antibodies; (ii) anti-KLK3 rabbit monoclonal antibody (Sino Biological) as an IgG isotype control for EPR3864(2) antibody.

Selection of Tryptic Peptides and Development of SRM Assays

Peptide Atlas, neXtProt database, and our in-house IP-shotgun MS data were used to select best peptides for the total ERG, or peptides unique and shared by specific isoforms. Excision of the N-term methionine, N-term acetylation, and serine and threonine phosphorylation were selected as potential posttranslational modifications of a unique N-term peptide MTASSSSDYGQTSK of T1/E4 fusion. In total, 14 synthetic heavy isotope-labeled SpikeTides_L peptides were used as internal standards for the Tier 2 SRM assay development (supplemental Table S2). Initially, 28 heavy and light peptide pairs (280 transitions) were included into a multiplex unscheduled SRM method with 5 ms scan time per transition. Based on the SRM peak area, main charge states and collision energies for each peptide were determined, and poorly performing peptides were removed. Low-intensity transitions and transitions with interferences with the VCaP cell lysate were removed. Finally, six best peptides with three or four most intense and reproducible transitions per peptide (42 transitions) were scheduled within 2-min acquisition windows (supplemental Table S3). Scan time of 10 ms ensured acquisition of at least 20 points per peak. Superposition of light and heavy peptide peaks, peak shapes, and the order of y-ion transition intensities ensured the correct identities of peptides in the cell lysate (31, 32). Amounts of the light endogenous peptides were calculated using the peak area ratio of the spiked-in heavy peptide internal standards. To enable accurate and absolute quantification, heavy isotope-labeled SpikeTides_TQL peptides with the trypsin-cleavable JPT-tags [serine-alanine-(3-nitro)tyrosine-glycine] were finally used as internal standards.

Chromatography and Targeted Mass Spectrometry

A quadrupole ion-trap mass spectrometer (AB SCIEX QTRAP 5500) coupled to EASY-nLC II liquid chromatography (Thermo Scientific) *via* a NanoSpray III ion source (AB SCIEX) was used for SRM assays. The tryptic peptides were loaded at 5 μl/min onto a C18 trap column (Thermo Scientific, 100 μm ID × 2 cm, 5 μm, 120 Å). Peptides were separated with PicoFrit columns (New Objective, 15 cm × 75 μm ID, 8 μm tip, PepMap C18, 3 μm, 100 Å) and 28 min gradients (300 nl/min). The gradient started with 5% buffer B and ramped to 65% buffer B over 20 min, then to 100% buffer B within 1 min, and continued for 7 min. QTRAP 5500 parameters were: 2300 V ion spray; 75 °C source temperature; 2.0 arbitrary units for gas 1 (N_2), 0 arbitrary units for gas 2; 25 arbitrary units for curtain gas (N_2); and 100 V declustering potential. Q Exactive Hybrid Quadrupole-Orbitrap (Thermo Scientific) coupled to EASY-nLC 1000 (Thermo Scientific) was used for PRM assays. The mobile phase consisted of 0.1% formic acid in water (buffer A) and 0.1% formic acid in acetonitrile (buffer B). Acclaim PepMap 100 nanoViper C18 precolumn (Thermo Scientific, 100 μm ID×2 cm, 5 μm, 100 Å) was used for sample loading, while EASY-Spray C18 (Thermo Scientific, 15 cm × 75 μm ID, 3 μm, 5 μm) was used as an analytical column. An 18-min gradient (400 nl/min) started with 0% buffer B and ramped to 50% buffer B over 15 min, then to 100% buffer B within 1 min, and continued for 2 min. PRM scans were performed at 17.5 K resolution with 27% normalized collision energy. Automatic Gain Control target value was set to 3×10^6 (100 ms injection time; 2.0 m/z isolation width). The performance of the nanoLC-MS systems was assessed every six runs with BSA digest solution (10 μl of 20 *f*mol/μl).

Chromatography and Shotgun Mass Spectrometry

ERG interactome was identified using Orbitrap Elite Hybrid Ion Trap-Orbitrap mass spectrometer (Thermo Scientific) coupled to EASY-nLC II liquid chromatography. The peptides were eluted at 300 nl/min with a 2-h gradient: 5% B for 5 min, 5 to 35% B for 95 min, 35 to 65% B for 10 min, 65 to 100% B for 1 min, and 100% B for 9 min. LC-MS/MS data were acquired with XCalibur (v. 2.2). MS1 scans (400–1250 m/z) were performed at 60 K resolution in the profile mode, followed by top 20 ion trap centroid MS/MS, acquired at 33% normalized collision energy. Ion counts were set to 1×10^6 (FTMS; 200 ms injection time) and 9000 (MS/MS; 100 ms injection time). MS/MS acquisition settings included 500 minimum signal threshold, 2.0 m/z isolation width, 10 ms activation time, and 60 s dynamic exclusion. Monoisotopic precursor selection and charge state rejection were enabled; +1 and unknown charge states were rejected. Instrument parameters included 230 °C capillary temperature, 2.0 kV source voltage, and 67% S-lens RF level.

IP-SRM Assessment of Interactome Disruption by Synthetic 9FY Epitope Peptides

Synthetic 9FY epitope peptides (HCl salts; Biomatik) included the minimal (RVPQQDWL) and extended (KMSPRVPQQDWLSQ) epitopes. The C-term antibody EPR3864(2) (1 μg per well) was used to enrich ERG interactome from VCaP lysate (80 μg total protein). Following ERG interactome enrichment on 96-well plates (in analytical

duplicates), 9FY epitope peptides were added at 0, 200, 400, and 800 μM (in 100 μl PBS) and incubated overnight at RT, followed by three washings and trypsin digestion (0.25 ng per well). SpikeTides_L peptides (200 *f*mol per well) were used for the accurate relative quantification of seven ERG-interacting proteins (supplemental Table S4). Each analytical replicate was analyzed in technical duplicates using the Tier 2 PRM assays and Q Exactive mass spectrometer.

Lysis and Sample Preparation of Prostate Cancer FFPE Tissues

Thirteen FFPE prostate cancer prostatectomy tissue blocks (ten TMPRSS2-ERG positive and three negative) were obtained from the University of Calgary (Ethics approval #HREBA.CC-14-0085). TMPRSS2-ERG positive areas were marked based on ERG IHC staining at Alberta Precision Laboratories, as previously reported (33). Fifteen 10 μm slices were cut from each block, deparaffinized with toluene (three times for 5 min), centrifuged at 14,000 rpm, and rehydrated with 100% ethanol for 2 min, then 95% ethanol for 2 min, 70% ethanol for 2 min, and distilled water for 1 min. Protocol for tissue pellet homogenization, antigen retrieval, and protein extraction included: (i) 90 min incubation at 90 °C in 0.2% Rapigest SF (100 mM tris-HCl, pH 8) and 1 mM DTT, (ii) repeated probe sonication at 45 kHz for 2 min, (iii) repeated incubation at 90 °C for 90 min; and (iv) centrifugation at 14,000 rpm. One sample (C13) could not be well homogenized, while the limited amount of sample C33 was sufficient only for the IP-SRM measurements.

Time-Resolved Fluorescence (TRF) ELISA

TRF-ELISA was developed as previously described (4). Briefly, C-term EPR3864(2) mAb (500 ng in 100 μl PBS per well) was coated onto a 96-well plate and incubated overnight at RT. The plate was washed six times with 250 μl wash buffer (125 mM Tris, 22.5% NaCl, 1.25% Tween 20, pH 7.8) using microplate washer (Biotek ELx50), followed by 1 h blocking (6% BSA in 250 μl wash buffer), and subsequent washing. To generate a calibration curve (7.8–1000 pg/ml), VCaP lysate with the known amount of total ERG (as measured by SRM) was diluted in assay diluent (6% BSA, 0.6% Tris, 20 mM KCl, 0.25% Tween 20). Twelve FFPE sample lysates were diluted in assay diluent (1:5 and 1:20). Calibrants and samples (100 μl per well) were incubated for 2 h with continuous shaking. Detection antibody OTI5F12 was biotinylated in-house with EZ-Link Sulfo-NHS-LC-LC-Biotin (Thermo Scientific, A35358). OTI5F12 (40 ng per well; in 100 μl blocking buffer) was incubated for 1 h. After washing, alkaline phosphatase-conjugated streptavidin (1 mg/ml; Jackson Immuno-Research) diluted in blocking buffer (1:1000 final) was added and incubated for 30 min at RT with continuous shaking. Following the final wash, 100 μl of 10 mM diflunisal phosphate (Toronto Research Chemicals) diluted in substrate buffer (10 mM Tris, 150 mM NaCl, 2 mM $MgCl_2$) was added to each well. After 10 min incubation, 100 μl developing solution (1 M Tris, 0.12 M NaOH, 2 mM Tb^{3+}, 3 mM EDTA) was added. Following 2 min shaking, luminescence of Tb-diflunisal-EDTA complex was measured with FilterMax F5 multi-mode microplate reader (Molecular Devices) with excitation at 370 nm (0.1 ms; TRF excitation filter 370NM BW80) and emission at 625 nm (0.1 ms delay and 0.6 ms integration time; fluorescence emission filter 625NM BW35).

Experimental Design and Statistical Rationale

For the preliminary quantification of the TMPRSS2-ERG protein isoforms, VCaP cell lysates were subjected to IP by the C-term, N-term, and 5F12 mAbs, or matched no-Ab controls, and were analyzed by the single-point titration with 100 fmol/injection SpikeTides_TQL heavy peptide internal standards, three analytical replicates, and two technical replicates (36 injections in total). According to effect size calculations, IP-SRM analysis with three analytical replicates could detect at least 17% change in ERG peptide abundance, respectively, assuming 80% power, $\alpha = 0.05$, average CV 5% for analytical replicates in our IP-SRM experiments, and a two-tailed paired *t*-test (G*Power software, v3.1.7, Heinrich Heine University Dusseldorf). For more accurate titration, VCaP cell lysates were subjected to IP by the C-term and N-term mAbs and were analyzed with 1, 5, and 20 fmol/injection SpikeTides_TQL standards, two analytical replicates, and two technical replicates (24 injections in total). To measure ERG and its isoforms in clinical samples (exploratory stage), 13 FFPE samples (ten positive and three negative for TMPRSS2-ERG fusion) were selected. Due to the low amounts of tissue, only one analytical and two technical replicates were analyzed per sample. For the interactome study, VCaP cell lysates were subjected to IP by the C-term, N-term, and 5F12 mAbs, or matched isotype mAb controls, and were analyzed with three analytical and two technical replicates (36 injections). According to effect size calculations, IP-MS analysis with three analytical replicates could identify proteins up- or down-regulated at least 1.09-fold, assuming 80% power, $\alpha = 0.05$, average 2.7% coefficient of variation (CV) for log2-transformed LFQ (Label-Free Quantification with MaxQuant software) values and a two-tailed *t*-test (G*Power software, v3.1.7, Heinrich Heine University Dusseldorf).

MS Data Analysis

SRM and PRM raw files were analyzed with Skyline Targeted Proteomics Environment v20.1.0.76 (MacCoss Lab) (34). Peak boundaries were adjusted manually, and the integrated areas of all transitions were extracted. Light-to-heavy peak area ratios were used for accurate relative or absolute quantification of the endogenous peptides. Shogun MS data were searched against the reviewed human UniProtKB/Swiss-prot database (20,365 entries, Uniprot release 2020_03) using MaxQuant software (v1.6.3.4) (35). Search parameters

TABLE 1
TMPRSS2-ERG protein isoforms identified and quantified in VCaP cells

TMPRSS2-ERG protein isoform	Length(aa[a])	Relative abundance (%)	Corresponding wildtype ERG protein isoforms (UniProt)	Corresponding mRNA isoforms (Ensembl)
T1E4-ERG	447	52.3 ± 2.7	Isoform 1 (P11308-4) Isoform 2 (P11308-3)	ENST00000288319.12 ENST00000398919.6 ENST00000417133.6
T1E4-ERG_Δ7b	423	9.5 ± 1.0	Isoform 3 (P11308-1) B5MDW0_HUMAN	ENST00000398911.5 ENST00000442448.5 ENST00000398905.5
T1E4-ERG_Δ4	387	32.9 ± 2.1	A0A0C4DG41_HUMAN	ENST00000453032.6
T1E4-ERG_Δ4Δ7b	363	5.3 ± 2.3	Isoform 4 (P11308-2)	ENST00000398897.5
T1E4-ERG_7bpA	286	n/d	Isoform 6 (P11308-6)	ENST00000468474.5

Δ4 and Δ7b, isoforms arising due to missing exons 4 or 7b; 7bpA, an isoform arising due to an alternative polyadenylation site; n/d, not detected; T1E4, fusion of TMPRSS2 exon 1 with ERG exon 4.
Protein isoform naming is according to Zammarchi *et al.*(64).
[a]Protein length includes the N-terminal methionine.

included: trypsin enzyme specificity, two missed cleavages, 7 aa minimum peptide length, top eight MS/MS peaks per 100 Da, 20 ppm MS1, and 0.5 Da MS/MS tolerance. Fixed modifications included cysteine carbamidomethylation; variable modifications included methionine oxidation, N-terminal acetylation, and deamidation (N). False-discovery rates for peptide and protein identifications were set to 1.0%. Label-free quantification (LFQ) algorithm was used for protein quantification. MaxQuant search results file proteinGroups.txt was uploaded to Perseus software (v1.6.12.0) to complete statistical analysis and data visualization (36). Protein identifications marked as "Only identified by site," "Reverse," and "Contaminants" were excluded. LFQ intensities were log2-transformed, and missing LFQ values were imputed with values representing normal distribution (0.2 width and 1.8 down shift). Proteins significantly enriched by anti-ERG antibodies *versus* isotype controls were selected using the log2 fold cutoffs (one-tail *t*-test p-value < 0.05). In addition, IP-shotgun MS data was searched against the custom database with all potential noncanonical ERG isoforms (total 46 unreviewed TrEMBL ERG isoforms and 20,365 entries of the reviewed human Swiss-Prot release 2020_03).

Assessment of Synthetic 9FY Epitope Peptides in Live VCaP Cells

To facilitate cell penetration and investigate synthetic 9FY epitope peptides in live VCaP cells, RVPQQDWL and KMSPRVPQQDWLSQ peptides were synthesized with the TAT sequence (GRKKRRQRRRG) at N- and C- terms (HCl salts; Biomatik). Peptides (final 400 μM in PBS) were incubated with VCaP cells (15,000 cells per well) on a 96-well plate. After 4 days of incubation, cell morphology was monitored *via* live cell imaging using the MetaXpress XLS system (Molecular Devices). Cell confluency was estimated with cell counting (average of four images) by the FIJI cell segmentation algorithm (ImageJ v1.51w).

RESULTS

Selection of ERG Protein Isoforms

RNA sequencing data available at the Cancer Cell Line Encyclopedia (CCLE) and DeepMap database (v2; https://depmap.org (37)) revealed that only six protein-coding mRNA isoforms were expressed in VCaP cells (supplemental Table S1). Our rationale for consideration of the endogenous ERG isoforms was based on the following facts: (i) VCaP cells had four copies of chromosome 21 (38), of which only two harbored T1/E4 TMPRSS2-ERG fusion and expressed TMPRSS2-ERG mRNA and a truncated ERG protein; (ii) similar to other prostate cancer cells, wild-type chromosomes 21 did not express any full length ERG (supplemental Table S5 and supplemental Fig. S2); (iii) mRNA T1/E4 isoforms 1 and 2, as well as isoforms 3 and B5MDW0, were identical at the protein level (Table 1); (iv) protein-coding mRNA isoforms 5, A0A088AWP2, A8MX39, and A8MZ24 were not expressed in VCaP cells. As a result, we considered five protein isoforms for further investigation (Table 1).

To measure the expression and relative abundance of ERG protein isoforms, tryptic peptides for the total ERG, peptides shared by some isoforms, or peptides unique for specific isoforms were selected. Among five protein isoforms, only T1E4-ERG_7bpA had unique peptides (Fig. 1*C*). The remaining four isoforms could be distinguished using combinations of the shared tryptic peptides VPQQDWLSQPPAR, NTDLPYEPPR, and ITTRPDLPYEPPR (Fig. 1*C*). Since isoforms T1E4-ERG_Δ4Δ7b and T1E4-ERG_Δ4 were lacking the N-term exon 4 (including KMSPRVPQQDWLSQ epitope of the 9FY mAb), immunoprecipitation with either the N-term 9FY or C-term EPR3864(2) antibodies followed by quantification of the shared and total ERG peptides would unambiguously resolve relative abundances of these four ERG isoforms.

Quantification of TMPRSS2-ERG Protein Isoforms by Orthogonal IP-SRM Assays

ERG protein was immunoprecipitated from VCaP cell lysate with three different anti-ERG mAbs. Following trypsin digestion,

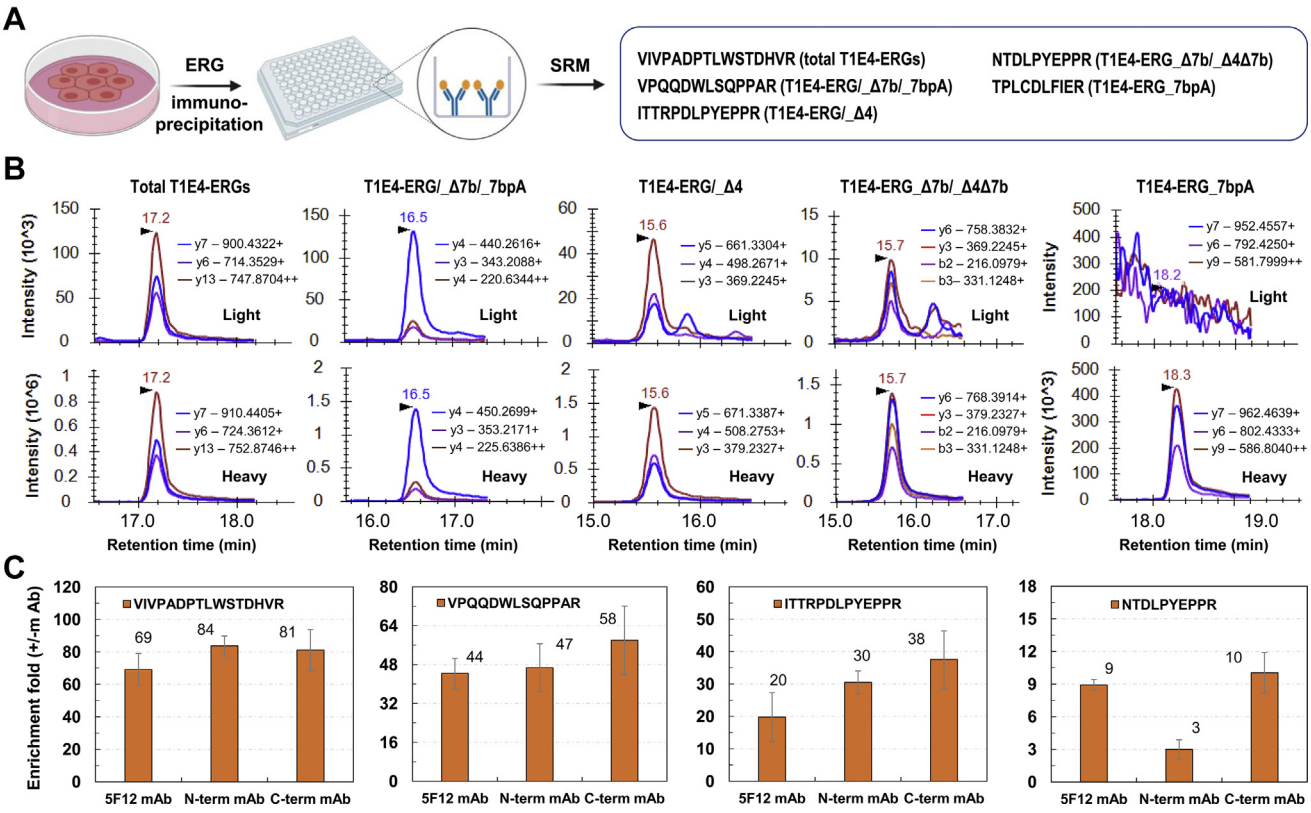


FIG. 2. **IP-SRM measurement of the shared and unique ERG isoform peptides in VCaP cells.** *A*, schematic workflow of the IP-SRM experiments and sequences of targeted peptides; *B*, SRM measurement of the endogenous ERG peptides; *C*, enrichment fold changes ((+) mAb/(–)mAb peak area ratios) revealed that three mAbs efficiently captured the total T1E4-ERGs from VCaP lysate, while the N-term mAb failed to capture the T1E4-ERG_Δ4 and T1E4-ERG _ Δ4Δ7b isoforms lacking the N-term epitope.

ERG protein was quantified by SRM using trypsin-cleavable SpikeTides_TQL internal standard peptides (supplemental Table S6), to enable absolute quantification (Fig. 2*A*). Figure 2*B* shows extracted-ion chromatograms of transitions monitored for the unique ERG peptides in VCaP lysate. ERG peptides detected in VCaP lysate by three different antibodies are listed in Table 2 and supplemental Table S7. The recovery of ERG after IP was estimated at ~90% (relative to the total ERG measured in the direct digest of VCaP cells). IP-SRM assays revealed excellent reproducibility, with coefficient of variation (CV) below 10% (Table 2). Limit of detection of IP-SRM, as estimated with serial dilutions of VCaP lysate, was 93.6 amol/μg total protein or ~10,000 cells (supplemental Fig. S3). Three antibodies displayed similar efficiency of capturing total ERG protein (69–84%; Fig. 2*C* and supplemental Table S8). Based on the cell count, the amount of total ERG was estimated at 2.2 fg per VCaP cell (~27,000 copies/cell), which was in good agreement with the previously reported levels (1.8 fg per cell) (29).

While mRNA isoform-6 was expressed at high levels (supplemental Fig. S4), none of the unique peptides of the corresponding protein isoform T1E4-ERG_7bpA were detected in VCaP cells (supplemental Fig. S5). Detection of the mutually exclusive peptides ITTRPDLPYEPPR and NTDLPYEPPR indicated expression of at least two distinct ERG isoforms (Fig. 2*B*). Based on our data (Table 2, supplemental Table S9, and supplemental Text), the relative abundance of four ERG isoforms was estimated at 52.3 ± 2.7% (T1E4-ERG), 9.5 ± 1.0% (T1E4-ERG_Δ7b), 32.9 ± 2.1% (T1E4-ERG_Δ4), and 5.3 ± 2.3% (T1E4-ERG_Δ4Δ7b). To the best of our knowledge, the relative abundance of endogenous ERG isoforms, as well as lack of expression of the shortest isoform T1E4-ERG_7bpA (P11308-6) missing ETS domain was quantified for the first time.

Identification of the N-terminal Peptide of T1/E4 TMPRSS2-ERG Protein

Tryptic digestion of T1E4-ERG and T1E4-ERG_Δ7b isoforms (~62% of the total ERG in VCaP cells) should generate a unique N-terminal peptide MTASSSSDYGQTSK (as opposed to the wild-type ERG peptide TEMTASSSSDYGQTSK). However, this unique peptide has never been identified in previous studies (29, 30). We hypothesized that the N-terminus of ERG could be further modified by acetylation, methionine cleavage, or threonine and serine phosphorylation. The challenge to detect a low-abundance fusion-specific peptide was our initial

TABLE 2
Quantification of TMPRSS2-ERG isoform razor peptides by IP-SRM in VCaP cells

Shared and unique peptides	ERG isoforms	C-term mAb; L/H ratio (CV%)			N-term mAb; L/H ratio (CV%)		
		20 fmol	*5 fmol*	*1 fmol*	*20 fmol*	*5 fmol*	*1 fmol*
VIVPADPTLWSTDHVR	Total ERG (all five isoforms)	0.363 (1.4)	1.680 (4.2)	7.320 (3.4)	0.424 (3.8)	1.817 (3.9)	7.854 (4.2)
VPQQDWLSQPPAR	T1E4-ERG; T1E4-ERG_Δ7b; T1E4-ERG_7bpA	0.236 (1.6)	1.027 (4.1)	4.338 (9.0)	0.451 (2.4)	1.960 (2.4)	9.001 (6.4)
ITTRPDLPYEPPR	T1E4-ERG; T1E4-ERG_Δ4	0.104 (2.7)	0.395 (4.4)	1.721 (4.6)	0.116 (13)	0.441 (6.3)	2.041 (1.8)
NTDLPYEPPR	T1E4-ERG_Δ7b; T1E4-ERG_Δ4Δ7b	0.017 (12)	0.069 (5.6)	0.322 (12)	0.023 (9.5)	0.074 (14)	0.369 (4)
TPLCDLFIER	T1E4-ERG_7bpA	0	0	0	0	0	0

The amounts (L/H ratios) and coefficients of variation (CV) were calculated based on analytical duplicates with spiked heavy standards of 1, 5, and 20 fmol per injection. Detailed data are presented in supplemental Table S9.

motivation. We then realized that the knowledge of the N-terminal modifications of ERG could be important to explore a presumably increased stability of TMPRSS2-ERG protein (39), or facilitate development of therapeutic strategies for degradation (40) of the fusion-derived, but not the wild-type ERG.

To reveal the N-terminal identity of T1/E4 TMPRSS2-ERG protein, we searched IP-MS/MS data for peptides modified with the N-terminal acetylation, N-terminal methionine truncation, and phosphorylation of threonine, serine, and tyrosine. Our search identified a peptide N-acetyl-TASSSSDYGQTSK, with the near-complete series of y- and b-fragment ions (Fig. 3*A*). Additional search of unreviewed TrEMBL sequences did not provide any nonstandard isoforms. We then confirmed the N-terminal identity of the T1/E4 TMPRSS2-ERG protein with more sensitive IP-SRM assays using internal standards with several potential modifications (Fig. 3*B*). IP-SRM confirmed the presence of N-acetyl-TASSSSDYGQTSK peptide (Fig. 3*C*). No phosphorylation of threonine or serine, as opposed to previous studies on ERG overexpressed in 293T cells (41), were found.

Identification of ERG Interactome

We hypothesized that orthogonal IP-MS/MS assays with the N-term and C-term antibodies could elucidate the most complete interactome of the endogenous ERG. First, we optimized conditions for the mild cell lysis. Several nondenaturing detergents (24), including (i) 0.1% Rapigest; (ii) modified IP lysis buffer (0.1% Rapigest, 1% NP-40, 0.1% sodium deoxycholate), and (iii) 0.5% SDS with 0.5% NP-40 were tested, and 0.1% Rapigest was selected based on the completeness of the cell lysis, total protein amount, and ERG recovery.

Following concurrent immunoprecipitation with the N-term, C-term, and 5F12 mAbs, ERG interactomes were identified by shotgun MS with the label-free quantification. MaxQuant search resulted in identification and quantification of 449 proteins (supplemental Table S10). As opposed to shotgun MS of the VCaP direct digest (no ERG identified), IP-MS identified ERG with eight unique peptides, including some isoform-specific peptides (supplemental Fig. S6). Proteins identified with LFQ fold change >1.3 and p-value <0.05 were selected as ERG-interacting proteins (Fig. 4 and supplemental Table S11). As a result, IP-MS with the C-term antibody identified 29 ERG-interacting proteins, including canonical BAF (SWI/SNF) chromatin remodeling complex subunits, androgen receptor (42), Ewing's Sarcoma breakpoint protein EWS_HUMAN (43), and nuclear receptor co-activators NCOA2_HUMAN and NCOA6_HUMAN (Fig. 4*A* and supplemental Fig. S7). Identification of eight proteins of the canonical BAF complexes (~2000 kDa polymorphic assemblies (44)) was in agreement with a recent study by Sandoval *et al.* (24). Interestingly, unique protein subunits of the polybromo-associated BAF complexes (including a unique subunit ARID2_HUMAN), noncanonical BAF (BRD9_HUMAN), embryonic stem cell-specific BAF (BC11A_HUMAN), or neuron-specific BAF (ACTL6B_HUMAN) were not identified in the ERG interactome, thus suggesting that expression of ERG-regulated genes in VCaP cells was mediated through the canonical BAF complexes, including mutually exclusive BRG1 (SMARCA4)- and BRM (SMARCA2)-associated BAF complexes (44). We believe that our C-term IP-MS identified one of the most comprehensive ERG interactomes (supplemental Table S11).

Interestingly, while ERG protein alone has been efficiently enriched 81-fold (N-term mAb, supplemental Table S12), 67-fold (C-term mAb), and 48-fold (5F12 mAb), immunoprecipitation with the N-term mAb and 5F12 mAb provided mostly high-abundance ribosomal, mitochondrial, and housekeeping proteins (most common contaminating proteins (45)), but failed to provide any meaningful candidates, such as proteins involved in transcriptional regulation (Fig. 4, *B* and *C*). We believe that such difference was due to the antibodies specificity, rather than stochastic nature of co-IP-MS. For example, detailed examination of MS1 data for ARI1A_HUMAN protein, the subunit of the canonical BAF complexes, revealed the high reproducibility of the C-term IP-MS assays and confirmed the absence of ARI1A_HUMAN in the IP experiments with N-term mAb and 5F12 mAb (supplemental Fig. S8). Such unexpected

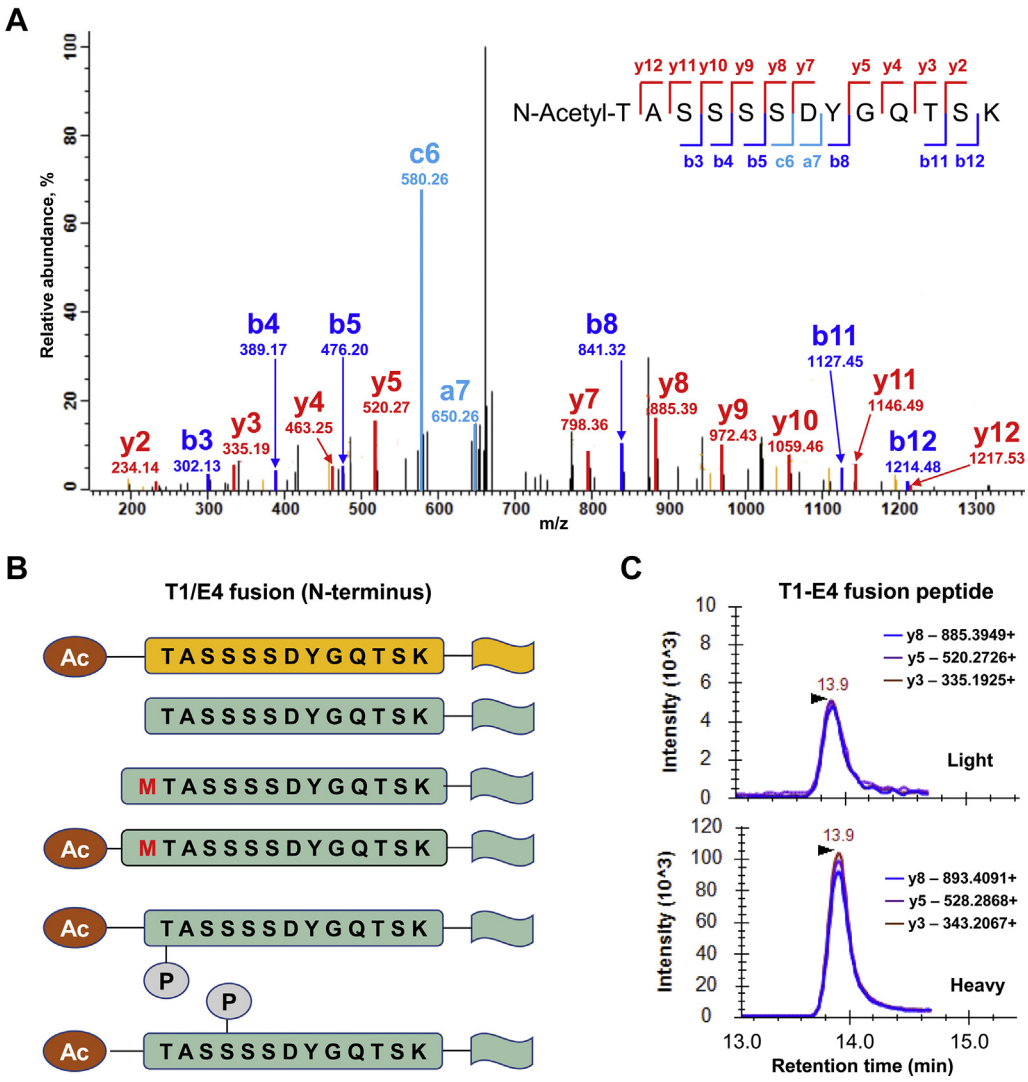


FIG. 3. **Detection of a unique fusion peptide of TMPRSS2-ERG protein.** *A*, MS/MS spectrum of N-acetyl-TASSSSDYGQTSK peptide identified in the VCaP cell lysate by shotgun IP-MS. *B*, six different N-terminal peptides, including N-term acetylated (Ac) and phosphorylated (P) peptides, were measured by SRM. *C*, IP-SRM measurement of a unique fusion peptide N-acetyl-TASSSSDYGQTSK in the VCaP cell lysate. The peptide represented two most abundant isoforms T1E4-ERG and T1E4-ERG_Δ7b (~62% total ERG).

result on interactome differences suggested that some antibodies could interfere or disrupt protein–protein interactions of ERG. We previously observed similar antibody-mediated disruptions of protein–protein interactions for the TEX101-DPEP3 complex (46). Here, this finding motivated us to investigate further if the ERG interactome could be disrupted by short synthetic peptides representing the binding epitope of the N-term mAb, thus paving the way for development of potential disruptors of the oncogenic ERG interactomes.

Independent Verification of ERG-Interacting Proteins by PRM

To verify some ERG-interacting proteins, we developed parallel reaction monitoring (PRM) assays for the most interesting ERG-interacting proteins, including some cBAF complex subunits (ARI1A_HUMAN, SMCA2_HUMAN, and SMCE1_HUMAN), androgen receptor (ANDR_HUMAN), nuclear receptor coactivators (NCOA2_HUMAN and NOCA6_HUMAN), and a transcriptional coregulator EWS. Targeted assays with the heavy peptide internal standards facilitated precise measurements of protein relative abundances and more accurate estimation of differential enrichments, as previously demonstrated (25, 47). Interestingly, independent verification revealed potentially three groups of ERG-interacting proteins: 1) strong and moderate ERG binders disrupted by the N-term antibody (NCOA2_HUMAN, NCOA6_HUMAN, and ARI1A_HUMAN); 2) a strong ERG binder not disrupted by the N-term antibody (EWS_HUMAN); and 3) weak ERG binders (supplemental Table S12).

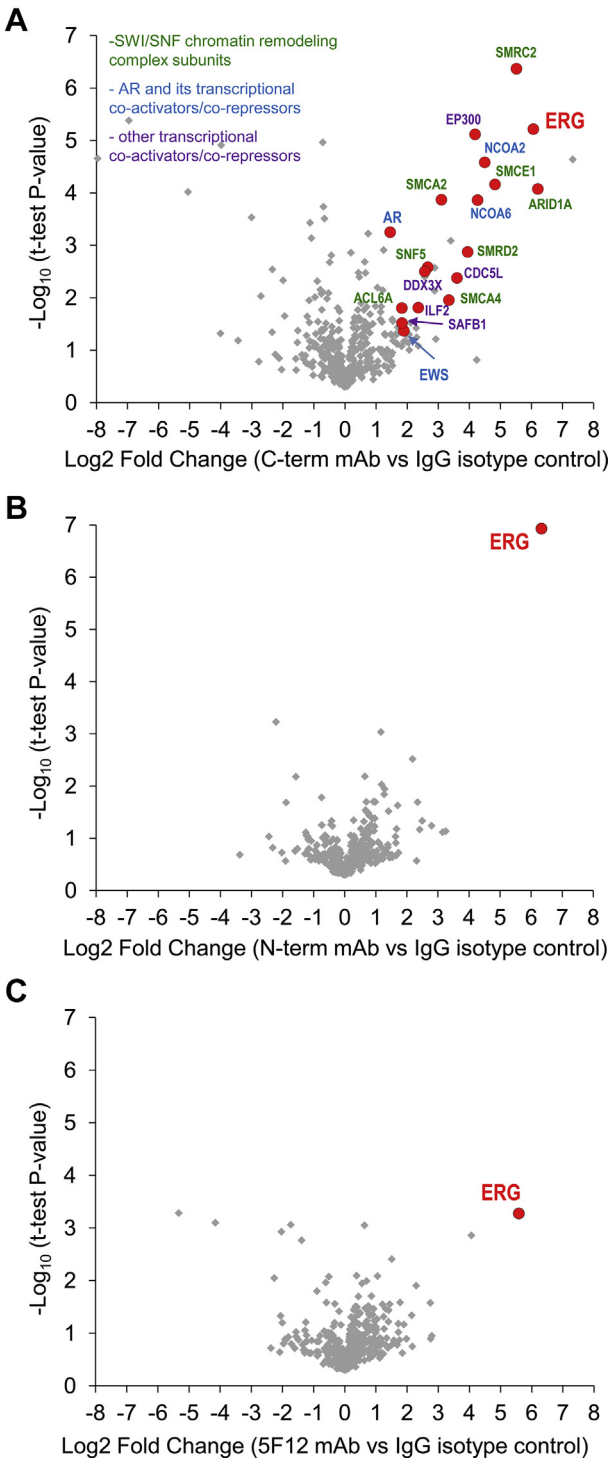


FIG. 4. **TMPRSS2-ERG interactome identified in VCaP cells.** IP experiments were completed simultaneously with the C-term (*A*), N-term (*B*), and 5F12 (*C*) mAbs. VCaP cell lysate, IP conditions, and proteomic sample preparation were identical; analytical replicates for all mAbs were simultaneously enriched and digested on a single 96-well plate. As a result, IP-MS with the C-term antibody identified 29 ERG-interacting proteins, while the N-term and 5F12 mAb enriched ERG, but did not provide any meaningful candidates.

It should be noted that while additional experiments with reciprocal IP-SRM and polyclonal anti-ARI1A antibodies successfully enriched ARI1A_HUMAN (110-fold *versus* isotype control), neither total ERG or its isoforms were detected. That result could be explained by the lower sensitivity of the ERG peptide assays, saturation of polyclonal anti-ARI1A antibodies with ARI1A_HUMAN paralogs (VCaP cells expressed six proteins of the AT-rich interactive domain-containing family (48), with 30–60% sequence identity), or lower stability of massive cBAF complexes (~2000 kDa) during IP-SRM. Overall, protein–protein interactions of cBAF are heterogeneous and dynamic; cBAF complexes simultaneously interact with numerous transcriptional activators. Our ERG C-term IP-SRM data estimated that only 1 in 50 ERG molecules interacted with ARI1A_HUMAN. We could suggest that the reciprocal IP-SRM assays for verification of cBAF interactions might utilize mAbs generated against known and unique epitopes with no cross-reactivity to the numerous paralogs of cBAF subunit proteins.

Evaluation of ERG Interactome Disruption by the N-term Epitope Peptides

Since the N-term mAb could potentially disrupt ERG interactome, we hypothesized that synthetic peptides representing minimal or extended epitopes (RVPQQDWL and KMSPRVPQQDWLSQ (49), respectively) could also disrupt interactions between ERG protein and its interactome.

First, we confirmed by IP-PRM that both peptides had no effect on ERG enrichment by the C-term mAb, but disrupted ERG enrichment by the N-term mAb. Following that, we evaluated the impact of a minimal epitope peptide RVPQQDWL on co-IP of some ERG-interacting proteins. As a result, only small changes were observed at very high RVPQQDWL concentrations (800 μM; Fig. 5*A*, supplemental Table S13). We concluded that the minimal epitope peptide was not a strong candidate for the disruption of ERG-interacting proteins (EC50 in mM range, if any at all).

Impact of the N-term Epitope Peptides on Proliferation and Morphology of VCaP Cells

We also evaluated the impact of a minimal epitope peptide RVPQQDWL on proliferation and morphology of VCaP cells. The cell-permeable N-term epitope peptides were conjugated to the cationic HIV-TAT motif GRKKRRQRRRG, to facilitate uptake into VCaP cells (50). As a result, we found that the cell-permeable N-term epitope peptides (400 μM) had no impact on VCaP proliferation (Fig. 5*B*, supplemental Table S14) or morphology (Fig. 5*C*).

Measurement of TMPRSS2-ERG Isoforms in Prostate Cancer FFPE Tissues by IP-SRM Assays

Quantification of four distinct TMPRSS2-ERG isoforms in VCaP cells motivated us to explore isoform identity in prostate

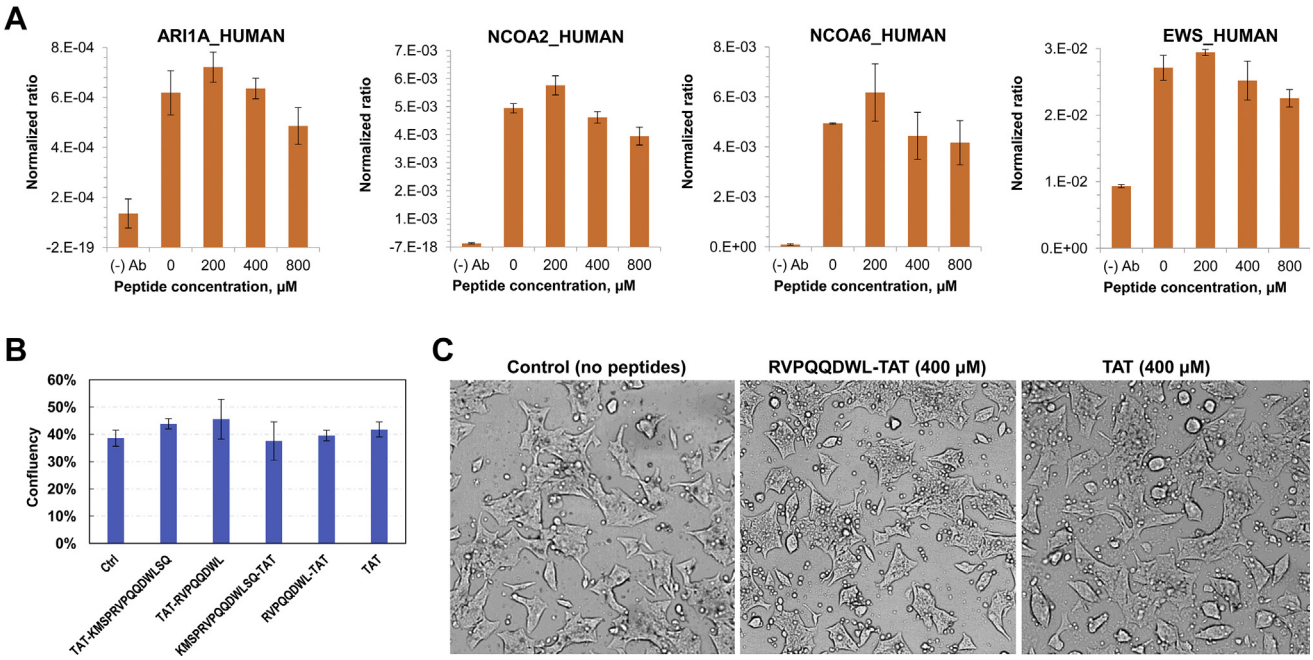


FIG. 5. **Assessment of the ERG interactome disruption by synthetic 9FY epitope peptides.** *A*, ERG-interacting proteins were enriched from the VCaP lysate using the C-term mAb, incubated with the increasing amounts of 9FY epitope peptide RVPQQDWL, and quantified by PRM assay. No significant differences were observed. *B*, confluency of VCaP cells treated with 400 μM epitope peptides (including peptides with the cell-permeable HIV-TAT sequences) and grown for 4 days. *C*, comparison of VCaP cell morphology untreated, treated with 400 μM RVPQQDWLSQ-TAT and TAT peptides for 4 days. No differences in cell morphology were observed.

cancer clinical samples. While FFPE tissues appear as the most challenging samples for proteomics, FFPE samples are widely available, have detailed clinical information, and include large amounts of tissue (radical prostatectomy FFPEs). Here, we obtained radical prostatectomy FFPEs blocks with positive (N = 10) and negative (N = 3) tissues for TMPRSS2-ERG fusion (as assessed by ERG IHC staining), and developed an IP-SRM-compatible protocol for protein extraction and quantification. Both C-term and N-term mAbs could enrich ERG protein even after intense tissue homogenization, probe sonication, and rigorous protein denaturation at high temperature in the presence of detergents.

As a result, total ERG protein was quantified by IP-SRM in six and ten fusion-positive FFPEs, with C-term and N-term mAbs, respectively (Fig. 6, *A* and *B*). LOD of the N-term IP-SRM (0.39 fmoles on column; S/N = 3) was low enough to differentiate between fusion-positive and negative FFPEs. Median ERG levels in ERG-positive tissues were 1.5 [IQR 1.2–1.6] fmoles on column. In addition, our C-term and N-term IP-SRM were sensitive enough to measure low-abundance isoform-specific peptides in five FFPE samples (Fig. 6*B* and supplemental Table S15). Interestingly, IP-SRM with the N-term mAb could also detect a low-abundance fusion-specific peptide Ac-TASSSSDYGQTSK in three fusion-positive FFPEs. Collectively, our data (Fig. 6) revealed that T1E4-ERG was the dominant isoform both in VCaP cells (52 ± 3%) and prostate-cancer FFPE tissues (50 ± 11%). These new data on expression of distinct ERG isoforms warrant further investigation on their association with progression and aggressiveness of prostate cancer.

Measurement of TMPRSS2-ERG in FFPE Tissues by TRF-ELISA

To facilitate independent validation of IP-SRM data in FFPE tissues, we developed in-house a highly sensitive (60 pg/ml LOD) and reproducible time-resolved fluorescence ELISA using the C-term and 5F12 mAbs for ERG capture and detection, respectively. TRF-ELISA revealed ERG levels below LOD in three ERG-negative FFPE lysates, and ERG levels ranging from 0.4 to 2.7 ng/ml in eight ERG-positive FFPE lysates (Fig. 6*A*, supplemental Fig. S9 and supplemental Table S16).

It should be noted that colorimetric ELISA for ERG protein (300 pg/ml LOD) has previously been reported (29), with the N-term mAb (9FY) used as a capture antibody (lacking detection of Δ4 ERG isoforms: T1E4-ERG_Δ4 and T1E4-ERG_Δ4Δ7b). Even though the epitope of 5F12 mAb was not known, our IP-SRM data suggested that 5F12 enriched Δ4 ERG isoforms (based on VPQQDWLSQPPAR/VIVPADPTLWSTDHVR ratios of 0.85 ± 0.04, 0.78 ± 0.01, and 1.28 ± 0.02 for the C-term, 5F12, and N-term mAbs, respectively; supplemental Table S7). In future, our ELISA and IP-SRM assays may facilitate development of ERG isoform-specific ELISA and evaluate clinical significance of total ERG (ELISA with the C-term capture and 5F12 detection mAbs) *versus* clinical significance of isoforms

A

FFPE tissue	IHC status	TRF-ELISA (ng/mL)	IP-SRM (N-term, fmoles on column)
C39	**+**	**2.7**	**2.1**
D39	**+**	**1.7**	**1.5**
D45	**+**	**1.4**	**1.9**
D27	**+**	**1.3**	**1.2**
D32	**+**	**1.0**	**1.5**
A40	+	0.7	0.9
C35	+	0.8	0.4
B22	+	0.4	1.3
C33	+	n/a	1.6
C13	+	< LOD	< LOD
C38	-	< LOD	< LOD
B33	-	< LOD	< LOD
C30	-	< LOD	< LOD

B

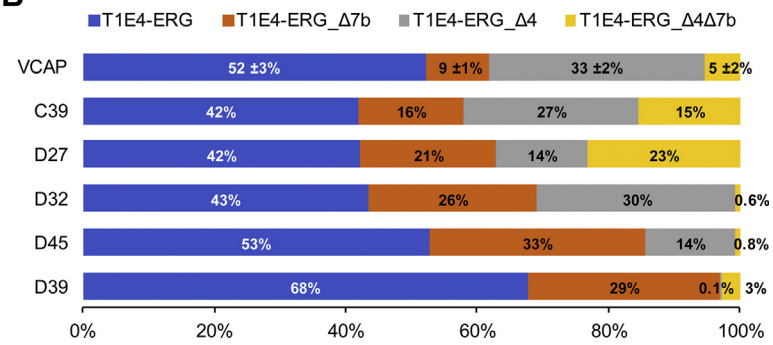


FIG. 6. **Measurement of the TMPRSS2-ERG protein isoforms in prostate cancer FFPE tissues.** *A*, thirteen prostatectomy FFPE tissue blocks with the known ERG IHC status were subjected to lysis and proteins extraction, followed by ERG quantification with TRF ELISA and IP-SRM. Three ERG IHC-negative samples (–) revealed total ERG below LOD of ELISA (60 pg/ml) and IP-SRM (0.39 fmoles on column; S/N = 3). One sample (C13) could not be homogenized. n/a, not available for measurements. *B*, relative abundances of ERG isoforms in VCaP cells and five FFPE tissues were calculated based on IP with the C-term and N-term mAbs and SRM measurements.

T1E4-ERG and T1E4-ERG_Δ7b (ELISA with the N-term capture and 5F12 detection mAbs).

DISCUSSION

Proteomics by mass spectrometry advanced to the level of identification and quantification of nearly whole proteomes of human cells (~12,000 proteins per cell) (48). However, quantification of low-abundance cellular proteins, such as transcription factors, may still require extensive sample fractionation or protein enrichment approaches. Distinct isoforms of transcription factors may present particular interest due to their different, or even opposite, molecular functions (for example, dominant-negative effects of the N-terminally truncated isoform ERα46 of the estrogen receptor alpha (51)). While mRNA isoforms are routinely measured (37), elucidation of the identity, abundance, and function of the distinct protein isoforms is still challenging and may be considered as one of the milestones of proteomics and proteogenomics (52–54).

IP-MS and IP-SRM assays have recently gained considerable interest due to their high sensitivity and selectivity for quantification of low-abundance proteins in cells and biological fluids (55). Immunoprecipitation substantially reduces sample complexity and facilitates quantification of multiple peptides per protein, thus enabling resolution of splicing isoforms and in-depth analysis of posttranslational modifications. In our previous studies, IP-SRM assays were successfully utilized to quantify low-abundance kallikrein-related peptidases (56), resolve protein isoforms (57), screen for antibody clones (58), discover the TEX101-DPEP3 complex (46), and detect a low-abundance missense variant of TEX101 protein (59). In this study, we focused on development of IP-MS and IP-SRM assays for a low-abundance fusion protein TMPRSS2-ERG. While the TMPRSS2-ERG gene fusion was found in nearly 50% of prostate cancer cases (60), and while its mRNA expression has been well characterized, clinical significance of TMPRSS2-ERG (for example, association with the more aggressive cancer (14, 15)) is still conflicting. Here, we hypothesized that IP-MS and IP-SRM would emerge as novel assays to characterize TMPRSS2-ERG fusion protein, quantify its levels in cells and tissues, elucidate its N-terminal posttranslational modifications, resolve the identity and abundance of TMPRSS2-ERG protein isoforms, and discover TMPRSS2-ERG interactome.

Wild-type ERG is a transcription factor expressed in endothelial cells and implicated in vascular development and angiogenesis (61). ERG expression is negligible in normal prostate tissues and prostate epithelial and stromal cells. Only two prostate cancer cell lines harbor TMPRSS2-ERG gene fusion and express measurable amounts of ERG transcripts: VCaP (~50,000-fold higher expression *versus* baseline expression in prostate stromal cells) and NCI-H660 (~6,000-fold) (62). Oncogenic nature of ERG protein is mediated through its function as a transcription factor, which promotes cell migration and cancer progression (63).

While ERG protein can now be identified in the direct digest of VCaP cells using the most powerful mass spectrometry instruments (48), ERG isoforms still cannot be fully resolved without extensive fractionation or protein enrichment approaches. He *et al.* previously pioneered quantification of ERG protein with two-dimensional liquid chromatography and targeted SRM assays (29, 30). ERG levels in VCaP cells were estimated at 1.8 fg per cell, which was in agreement with our IP-SRM data (2.2 fg, or ~27,000 copies per cell).

ERG mRNA isoforms were previously characterized in cells and prostate cancer tissues (64). As opposed to the total ERG mRNA, isoform-6 (previously known as ERG8 or isoform-8) was associated with more favorable outcomes of prostate cancer (12). Corresponding T1E4-ERG_7bpA protein (lacking DNA-binding ETS domain essential for transcriptional activation) was suggested as a hypothetical inhibitor of the ERG-mediated gene expression (19, 20). Here, we resolved for the first time the relative abundance of four ERG isoforms in

VCaP cells and demonstrated that isoform-6 was not expressed at the protein level (supplemental Figs. S4 and S5).

Recently, it has been demonstrated that degradation of wild-type ERG protein was mediated through the SPOP ubiquitin ligase-binding site 42-ASSSS-46 ("degron") located at the N-term of ERG. The N-term truncated TMPRSS2-ERG fusion protein displayed significantly reduced interaction with SPOP *in vivo* and *in vitro*. (39, 41) This fact encouraged us to characterize the N-terminal composition of the endogenous TMPRSS2-ERG fusion protein in VCaP cells. Our sensitive IP-SRM assays revealed the N-term methionine truncation and threonine acetylation (peptide N-acetyl-TASSSSDYGQTSK), but no phosphorylation of threonine or serine. In future, development of affinity ligands recognizing the N-terminal motif N-acetyl-TASSSS (none of the human canonical UniProt protein isoforms have the N-term motifs MTASSSS or MTASSS) could be utilized as strategy to target TMPRSS2-ERG fusion protein for degradation (65).

Finally, we employed our orthogonal IP-MS assays to identify interactome of the endogenous TMPRSS2-ERG protein in VCaP cells. Previous co-IP-MS studies completed with polyclonal antibodies identified partial ERG interactomes (lacking BAF complex subunits (66), EWS protein (24), and others). We believe that our IP-MS assay with the C-term mAb identified one of the most complete ERG interactomes (including BRG1-and BRM-associated canonical BAF complexes, EWS (43), androgen receptor (22, 67), and numerous transcriptional regulators).

Our data revealed that the N-term and 5F12 mAbs could enrich ERG, but not its interactome. We suggested that these high-affinity mAbs could interfere with ERG protein–protein interactions, and that the region crucial for interactions with cBAF complexes could be located between the N-term epitope (46-KMSPRVPQQDWLSQ-59 of P11308-4) and the DNA-binding ETS domain (aa 320–391). PNT domain (aa 120–199) located in that region was previously suggested as the protein–protein interaction domain of the ETS family of transcription factors (23). Our experimental methods, however, were lacking any three-dimensional structural perspectives to explore the hypothesis that interactions of large cBAF complexes (~2000 kDa) and transcriptional regulators (NCOA2_HUMAN and NCOA6_HUMAN, ~200 kDa) could be disrupted due to steric hindrance with the N-term or 5F12 mAbs. In future, the exact regions of ERG protein–protein interactions could be identified with the series of recombinant truncated ERG proteins, epitope mapping arrays, or cross-linking mass spectrometry (68). It would also be worth investigating whether the shorter isoforms T1E4-ERG_Δ7b and T1E4-ERG_Δ4Δ7b (lacking exon 7 of 26 aa in the region between the N-term epitope and ETS domain; representing 15% of total ERG) could reveal alternative interactomes or exhibit any dominant-negative inhibition of the ERG transcriptional activity (similar to the hypothetical isoform T1E4-ERG_7bpA (20)). In line with these investigations, short peptides disrupting ERG interactome, similar to peptides disrupting ERG-DNA interactions (50), could be identified and explored as potential targeted therapies of prostate cancer.

Finally, we demonstrated that our IP-SMR assays were useful for quantification of ERG and its four isoforms in prostatectomy FFPE tissues. Sensitivity of the N-term IP-SRM was sufficient to differentiate between fusion-positive FFPEs (median ERG levels 1.5 [IQR 1.2–1.6] fmoles on column) and fusion-negative FFPEs (values below LOD of 0.39 fmoles on column). These data were in agreement with our in-house TRF-ELISA results for fusion-positive (median 1.2 [IQR 0.8–1.5] ng/ml) and fusion-negative FFPEs (values below LOD of 0.06 ng/ml). In future, our IP-SRM assays could be utilized to quantify ERG and its isoforms in fresh-frozen tissues, semen, and exfoliated prostate epithelial cells in urine. Some post-digital rectal examination urine samples of prostate cancer patients were found to contain up to 27,000 exfoliated cells (69), which was above LOD of our IP-SRM assay (~10,000 VCaP cells). Our ELISA and IP-SRM assays could also facilitate development of ERG isoform-specific immunoassays and evaluate clinical significance (for example, the risk of prostate cancer progression during active surveillance (70)) of total ERG *versus* isoforms T1E4-ERG and T1E4-ERG_Δ7b in the large cohorts of clinical samples. Collectively, future applications of our IP-SRM assays may well contribute to the precision diagnostics needs prioritized by the Movember Prostate Cancer Landscape Analysis (71).

CONCLUSIONS

We developed IP-SRM and IP-MS assays for the quantification of a low-abundant transcriptional factor TMPRSS2-ERG fusion protein (~27,000 copies/cell), its isoforms, and its interactome in VCaP cells. Our orthogonal IP-SRM assays quantified for the first time the relative abundance of four isoforms and revealed that the T1E4-ERG isoform accounted for 52 ± 3% of the total ERG protein in VCaP cells and 50 ± 11% in prostate cancer FFPE tissues. For the first time, the N-terminal peptide (methionine-truncated and N-acetylated TASSSSDYGQTSK) unique for the T1/E4 fusion was identified. ERG interactome mapping with the C-terminal antibodies identified 29 proteins, including the mutually exclusive BRG1- and BRM-associated canonical SWI/SNF chromatin remodeling complexes, and numerous transcriptional regulators. Our sensitive and selective IP-SRM assays present alternative tools to quantify ERG and its isoforms in clinical samples, thus paving the way for development of more accurate diagnostics of prostate cancer.

DATA AVAILABILITY

Raw shotgun MS data have been deposited to ProteomeXchange Consortium *via* PRIDE (www.ebi.ac.uk/pride/

archive/login) with the data set identifier PXD021236. All annotated spectra, including single-peptide identification, can be inspected with MS-Viewer (https://msviewer.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msviewer) and the search key: 2elnmyocus. Raw SRM and PRM data, as well as processed Skyline files, were deposited to Peptide Atlas with the dataset identifier PASS01624 (www.peptideatlas.org/PASS/PASS01624 or ftp://PASS01624:SL727hk@ftp.peptideatlas.org).

Supplemental Data—This article contains supplemental data (21, 23).

Acknowledgments—We thank Prof. Xingfang Li for the access to QTRAP 5500 mass spectrometer and Dr. Konstantin Stoletov for providing VCaP and LNCaP cells.

Author contributions—A. P. D. designed the research project. Z. F. and Y. R. performed all major experiments and analyzed the data. T. A. B. and M. E. H. provided clinical samples and clinical information. Z. F. and A. P. D. wrote the article, and all the authors contributed to the revisions.

Z. F.: Investigation, Methodology, Validation, Formal Analysis, Writing—Original Draft, Writing—Review and Editing, Data Curation, Visualization, Y. R.: Investigation, Writing—Review and Editing, T. A. B.: Resources, M. E. H.: Resources, X. C. L.: Resources, Writing—Review and Editing, Funding Acquisition, A. P. D.: Conceptualization, Methodology, Validation, Formal Analysis, Data Curation, Writing—Original Draft, Writing—Review and Editing, Visualization, Supervision, Project Administration, Funding Acquisition

Funding and additional information—This work was supported by the Prostate Cancer Canada grant (RS2015-01) to A. P. D.; Z. F. acknowledges the support from the National Natural Science Foundation of China (21806018) and Globalink Early Career Fellowship (Mitacs, Canada and China Scholarship Council, China; 201806065018).

Conflict of interest—The authors declare no potential conflicts of interest.

Abbreviations—The abbreviations used are: aa, amino acids; CV, coefficient of variation; ELISA, enzyme-linked immunosorbent assay; ERG, v-Ets avian erythroblastosis virus E26 oncogene homolog; FFPE, formalin-fixed paraffin-embedded tissues; IHC, immunohistochemistry; LC-MS/MS, liquid chromatography-tandem mass spectrometry; LFQ, label-free quantification; mAb, monoclonal antibody; PCa, prostate cancer; PRM, parallel reaction monitoring; RT, room temperature; SRM, selected reaction monitoring.

Received February 23, 2021 Published, MCPRO Papers in Press, March 23, 2021, https://doi.org/10.1016/j.mcpro.2021.100075

REFERENCES

1. Pinsky, P. F., Prorok, P. C., and Kramer, B. S. (2017) Prostate cancer screening - a perspective on the current state of the evidence. *N. Engl. J. Med.* **376**, 1285–1289
2. Grossman, D. C., Curry, S. J., Owens, D. K., Bibbins-Domingo, K., Caughey, A. B., Davidson, K. W., Doubeni, C. A., Ebell, M., Epling, J. W., Jr., Kemper, A. R., Krist, A. H., Kubik, M., Landefeld, C. S., Mangione, C. M., Silverstein, M., *et al*. (2018) Screening for prostate cancer: US preventive services task force recommendation statement. *JAMA* **319**, 1901–1913
3. Saraon, P., Drabovich, A. P., Jarvi, K. A., and Diamandis, E. P. (2014) Mechanisms of androgen-independent prostate cancer. *EJIFCC* **25**, 42–54
4. Drabovich, A. P., Saraon, P., Drabovich, M., Karakosta, T. D., Dimitromanolakis, A., Hyndman, M. E., Jarvi, K., and Diamandis, E. P. (2019) Multi-omics biomarker pipeline reveals elevated levels of protein-glutamine gamma-glutamyltransferase 4 in seminal plasma of prostate cancer patients. *Mol. Cell. Proteomics* **18**, 1807–1823
5. Koo, K. M., Mainwaring, P. N., Tomlins, S. A., and Trau, M. (2019) Merging new-age biomarkers and nanodiagnostics for precision prostate cancer management. *Nat. Rev. Urol.* **16**, 302–317
6. Drabovich, A. P., Pavlou, M. P., Batruch, I., and Diamandis, E. P. (2013) Proteomic and mass spectrometry technologies for biomarker discovery. In: Issaq, H. J., Veenstra, T. D., eds. *Proteomic and Metabolomic Approaches to Biomarker Discovery*, Academic Press (Elsevier), Waltham, MA: 17–37
7. Cho, C. K., Drabovich, A. P., Karagiannis, G. S., Martinez-Morillo, E., Dason, S., Dimitromanolakis, A., and Diamandis, E. P. (2013) Quantitative proteomic analysis of amniocytes reveals potentially dysregulated molecular networks in Down syndrome. *Clin. Proteomics* **10**, 2
8. Konvalinka, A., Zhou, J., Dimitromanolakis, A., Drabovich, A. P., Fang, F., Gurley, S., Coffman, T., John, R., Zhang, S. L., Diamandis, E. P., and Scholey, J. W. (2013) Determination of an angiotensin II-regulated proteome in primary human kidney cells by stable isotope labeling of amino acids in cell culture (SILAC). *J. Biol. Chem.* **288**, 24834–24847
9. Abeshouse, A., Ahn, J., Akbani, R., Ally, A., Amin, S., Andry, C. D., Annala, M., Aprikian, A., Armenia, J., Arora, A., Auman, J. T., Balasundaram, M., Balu, S., Barbieri, C. E., Bauer, T., *et al*. (2015) The molecular taxonomy of primary prostate cancer. *Cell* **163**, 1011–1025
10. Armenia, J., Wankowicz, S. A. M., Liu, D., Gao, J., Kundra, R., Reznik, E., Chatila, W. K., Chakravarty, D., Han, G. C., Coleman, I., Montgomery, B., Pritchard, C., Morrissey, C., Barbieri, C. E., Beltran, H., *et al*. (2018) The long tail of oncogenic drivers in prostate cancer. *Nat. Genet.* **50**, 645–651
11. Tomlins, S. A., Laxman, B., Dhanasekaran, S. M., Helgeson, B. E., Cao, X., Morris, D. S., Menon, A., Jing, X., Cao, Q., Han, B., Yu, J., Wang, L., Montie, J. E., Rubin, M. A., Pienta, K. J., *et al*. (2007) Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448**, U595–U599
12. Hu, Y., Dobi, A., Sreenath, T., Cook, C., Tadase, A. Y., Ravindranath, L., Cullen, J., Furusato, B., Chen, Y., Thangapazham, R. L., Mohamed, A., Sun, C., Sesterhenn, I. A., McLeod, D. G., Petrovics, G., *et al*. (2008) Delineation of TMPRSS2-ERG splice variants in prostate cancer. *Clin. Cancer Res.* **14**, 4719–4725
13. Tomlins, S. A., Rhodes, D. R., Perner, S., Dhanasekaran, S. M., Mehra, R., Sun, X. W., Varambally, S., Cao, X. H., Tchinda, J., Kuefer, R., Lee, C., Montie, J. E., Shah, R. B., Pienta, K. J., Rubin, M. A., *et al*. (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**, 644–648
14. Fine, S. W., Gopalan, A., Leversha, M. A., Al-Ahmadie, H. A., Tickoo, S. K., Zhou, Q., Satagopan, J. M., Scardino, P. T., Gerald, W. L., and Reuter, V. E. (2010) TMPRSS2-ERG gene fusion is associated with low Gleason scores and not with high-grade morphological features. *Mod. Pathol.* **23**, 1325–1333
15. Toubaji, A., Albadine, R., Meeker, A. K., Isaacs, W. B., Lotan, T., Haffner, M. C., Chaux, A., Epstein, J. I., Han, M., Walsh, P. C., Partin, A. W., De Marzo, A. M., Platz, E. A., and Netto, G. J. (2011) Increased gene copy number of ERG on chromosome 21 but not TMPRSS2-ERG fusion predicts outcome in prostatic adenocarcinomas. *Mod. Pathol.* **24**, 1511–1520

16. Park, K., Tomlins, S. A., Mudaliar, K. M., Chiu, Y. L., Esgueva, R., Mehra, R., Suleman, K., Varambally, S., Brenner, J. C., MacDonald, T., Srivastava, A., Tewari, A. K., Sathyanarayana, U., Nagy, D., Pestano, G., *et al*. (2010) Antibody-based detection of ERG rearrangement-positive prostate cancer. *Neoplasia* **12**, U590–U595
17. Pettersson, A., Graff, R. E., Bauer, S. R., Pitt, M. J., Lis, R. T., Stack, E. C., Martin, N. E., Kunz, L., Penney, K. L., Ligon, A. H., Suppan, C., Flavin, R., Sesso, H. D., Rider, J. R., Sweeney, C., *et al*. (2012) The TMPRSS2:ERG rearrangement, ERG expression, and prostate cancer outcomes: A cohort study and meta-analysis. *Cancer Epidemiol. Biomarkers Prev*. **21**, 1497–1509
18. Shah, R. B. (2013) Clinical applications of novel ERG immunohistochemistry in prostate cancer diagnosis and management. *Adv. Anat. Pathol.* **20**, 117–124
19. Carrere, S., Verger, A., Flourens, A., Stehelin, D., and Duterque-Coquillaud, M. (1998) Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes via two distinct domains. *Oncogene* **16**, 3261–3268
20. Hoesel, B., Malkani, N., Hochreiter, B., Basilio, J., Sughra, K., Ilyas, M., and Schmid, J. A. (2016) Sequence-function correlations and dynamics of ERG isoforms. ERG8 is the black sheep of the family. *Biochim. Biophys. Acta* **1863**, 205–218
21. Kedage, V., Strittmatter, B. G., Dausinas, P. B., and Hollenhorst, P. C. (2017) Phosphorylation of the oncogenic transcription factor ERG in prostate cells dissociates polycomb repressive complex 2, allowing target gene activation. *J. Biol. Chem.* **292**, 17225–17235
22. Yu, J. D., Yu, J. J., Mani, R. S., Cao, Q., Brenner, C. J., Cao, X. H., Wang, X. J., Wu, L. T., Li, J., Hu, M., Gong, Y. S., Cheng, H., Laxman, B., Vellaichamy, A., Shankar, S., *et al*. (2010) An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell* **17**, 443–454
23. Hollenhorst, P. C., McIntosh, L. P., and Graves, B. J. (2011) Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu. Rev. Biochem.* **80**, 437–471
24. Sandoval, G. J., Pulice, J. L., Pakula, H., Schenone, M., Takeda, D. Y., Pop, M., Boulay, G., Williamson, K. E., McBride, M. J., Pan, J., St Pierre, R., Hartman, E., Garraway, L. A., Carr, S. A., Rivera, M. N., *et al*. (2018) Binding of TMPRSS2-ERG to BAF chromatin remodeling complexes mediates prostate oncogenesis. *Mol. Cell* **71**, 554–566.e557
25. Drabovich, A. P., Dimitromanolakis, A., Saraon, P., Soosaipillai, A., Batruch, I., Mullen, B., Jarvi, K., and Diamandis, E. P. (2013) Differential diagnosis of azoospermia with proteomic biomarkers ECM1 and TEX101 quantified in seminal plasma. *Sci. Transl. Med.* **5**, 212ra160
26. Drabovich, A. P., Pavlou, M. P., Schiza, C., and Diamandis, E. P. (2016) Dynamics of protein expression reveals primary targets and secondary messengers of estrogen receptor alpha signaling in MCF-7 breast cancer cells. *Mol. Cell. Proteomics* **15**, 2093–2107
27. Drabovich, A. P., Pavlou, M. P., Dimitromanolakis, A., and Diamandis, E. P. (2012) Quantitative analysis of energy metabolic pathways in MCF-7 breast cancer cells by selected reaction monitoring assay. *Mol. Cell. Proteomics* **11**, 422–434
28. Huttenhain, R., Soste, M., Selevsek, N., Rost, H., Sethi, A., Carapito, C., Farrah, T., Deutsch, E. W., Kusebauch, U., Moritz, R. L., Nimeus-Malmstrom, E., Rinner, O., and Aebersold, R. (2012) Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. *Sci. Transl. Med.* **4**, 142ra194
29. He, J., Schepmoes, A. A., Shi, T., Wu, C., Fillmore, T. L., Gao, Y., Smith, R. D., Qian, W. J., Rodland, K. D., Liu, T., Camp, D. G., 2nd, Rastogi, A., Tan, S. H., Yan, W., Mohamed, A. A., *et al*. (2015) Analytical platform evaluation for quantification of ERG in prostate cancer using protein and mRNA detection methods. *J. Transl. Med.* **13**, 54
30. He, J., Sun, X., Shi, T., Schepmoes, A. A., Fillmore, T. L., Petyuk, V. A., Xie, F., Zhao, R., Gritsenko, M. A., Yang, F., Kitabayashi, N., Chae, S. S., Rubin, M. A., Siddiqui, J., Wei, J. T., *et al*. (2014) Antibody-independent targeted quantification of TMPRSS2-ERG fusion protein products in prostate cancer. *Mol. Oncol.* **8**, 1169–1180
31. Martinez-Morillo, E., Cho, C. K., Drabovich, A. P., Shaw, J. L., Soosaipillai, A., and Diamandis, E. P. (2012) Development of a multiplex selected reaction monitoring assay for quantification of biochemical markers of Down syndrome in amniotic fluid samples. *J. Proteome Res.* **11**, 3880–3887
32. Drabovich, A. P., and Diamandis, E. P. (2010) Combinatorial peptide libraries facilitate development of multiple reaction monitoring assays for low-abundance proteins. *J. Proteome Res.* **9**, 1236–1245
33. Lee, S. L., Yu, D., Wang, C., Saba, R., Liu, S., Trpkov, K., Donnelly, B., and Bismar, T. A. (2015) ERG expression in prostate needle biopsy: Potential diagnostic and prognostic implications. *Appl. Immunohistochem. Mol. Morphol.* **23**, 499–505
34. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968
35. Tyanova, S., Temu, T., and Cox, J. (2016) The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319
36. Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., and Cox, J. (2016) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* **13**, 731–740
37. Ghandi, M., Huang, F. W., Jane-Valbuena, J., Kryukov, G. V., Lo, C. C., McDonald, E. R., Barretina, J., Gelfand, E. T., Bielski, C. M., Li, H., Hu, K., Andreev-Drakhlin, A. Y., Kim, J., Hess, J. M., Haas, B. J., *et al*. (2019) Next-generation characterization of the cancer cell line encyclopedia. *Nature* **569**, 503
38. van Bokhoven, A., Caires, A., Maria, M. D., Schulte, A. P., Lucia, M. S., Nordeen, S. K., Miller, G. J., and Varella-Garcia, M. (2003) Spectral karyotype (SKY) analysis of human prostate carcinoma cell lines. *Prostate* **57**, 226–244
39. Gan, W., Dai, X., Lunardi, A., Li, Z., Inuzuka, H., Liu, P., Varmeh, S., Zhang, J., Cheng, L., Sun, Y., Asara, J. M., Beck, A. H., Huang, J., Pandolfi, P. P., and Wei, W. (2015) SPOP promotes ubiquitination and degradation of the ERG oncoprotein to suppress prostate cancer progression. *Mol. Cell* **59**, 917–930
40. Bogaert, A., Fernandez, E., and Gevaert, K. (2020) N-terminal proteoforms in human disease. *Trends Biochem. Sci.* **45**, 308–320
41. An, J., Ren, S., Murphy, S. J., Dalangood, S., Chang, C., Pang, X., Cui, Y., Wang, L., Pan, Y., Zhang, X., Zhu, Y., Wang, C., Halling, G. C., Cheng, L., Sukov, W. R., *et al*. (2015) Truncated ERG oncoproteins from TMPRSS2-ERG fusions are resistant to SPOP-mediated proteasome degradation. *Mol. Cell* **59**, 904–916
42. Wasmuth, E. V., Hoover, E. A., Antar, A., Klinge, S., Chen, Y., and Sawyers, C. L. (2020) Modulation of androgen receptor DNA binding activity through direct interaction with the ETS transcription factor ERG. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 8584–8592
43. Kedage, V., Selvaraj, N., Nicholas, T. R., Budka, J. A., Plotnik, J. P., Jerde, T. J., and Hollenhorst, P. C. (2016) An interaction with Ewing's sarcoma breakpoint protein EWS defines a specific oncogenic mechanism of ETS factors rearranged in prostate cancer. *Cell Rep.* **17**, 1289–1301
44. Kadoch, C., Hargreaves, D. C., Hodges, C., Elias, L., Ho, L., Ranish, J., and Crabtree, G. R. (2013) Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601
45. Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J. P., St-Denis, N. A., Li, T., Miteva, Y. V., Hauri, S., Sardiu, M. E., Low, T. Y., Halim, V. A., Bagshaw, R. D., Hubner, N. C., Al-Hakim, A., Bouchard, A., *et al*. (2013) The CRAPome: A contaminant repository for affinity purification-mass spectrometry data. *Nat. Methods* **10**, 730–736
46. Schiza, C., Korbakis, D., Panteleli, E., Jarvi, K., Drabovich, A. P., and Diamandis, E. P. (2018) Discovery of a human testis-specific protein complex TEX101-DPEP3 and selection of its disrupting antibodies. *Mol. Cell. Proteomics* **17**, 2480–2495
47. Drabovich, A. P., Jarvi, K., and Diamandis, E. P. (2011) Verification of male infertility biomarkers in seminal plasma by multiplex selected reaction monitoring assay. *Mol. Cell. Proteomics* **10**, M110.004127
48. Nusinow, D. P., Szpyt, J., Ghandi, M., Rose, C. M., McDonald, E. R., Kalocsay, M., Jane-Valbuena, J., Gelfand, E., Schweppe, D. K., Jedrychowski, M., Golji, J., Porter, D. A., Rejtar, T., Wang, Y. K., Kryukov, G. V., *et al*. (2020) Quantitative proteomics of the cancer cell line encyclopedia. *Cell* **180**, 387–402
49. Rastogi, A., Tan, S. H., Banerjee, S., Sharad, S., Kagan, J., Srivastava, S., McLeod, D. G., Srivastava, S., and Srinivasan, A. (2014) ERG monoclonal antibody in the diagnosis and biological stratification of prostate cancer: Delineation of minimal epitope, critical residues for binding, and

molecular basis of specificity. *Monoclon. Antib. Immunodiagn. Immunother.* **33**, 201–208
50. Wang, X. J., Qiao, Y. Y., Asangani, I. A., Ateeq, B., Poliakov, A., Cieslik, M., Pitchiaya, S., Chakravarthi, B., Cao, X. H., Jing, X. J., Wang, C. X., Apel, I. J., Wang, R., Tien, J. C. Y., Juckette, K. M., *et al*. (2017) Development of peptidomimetic inhibitors of the ERG gene fusion product in prostate cancer. *Cancer Cell* **31**, 532–548
51. Metivier, R., Penot, G., Carmouche, R. P., Hubner, M. R., Reid, G., Denger, S., Manu, D., Brand, H., Kos, M., Benes, V., and Gannon, F. (2004) Transcriptional complexes engaged by apo-estrogen receptor-alpha isoforms have divergent outcomes. *EMBO J.* **23**, 3653–3666
52. Smith, L. M., and Kelleher, N. L. (2018) Proteoforms as the next proteomics currency. *Science* **359**, 1106–1107
53. Aebersold, R., Agar, J. N., Amster, I. J., Baker, M. S., Bertozzi, C. R., Boja, E. S., Costello, C. E., Cravatt, B. F., Fenselau, C., Garcia, B. A., Ge, Y., Gunawardena, J., Hendrickson, R. C., Hergenrother, P. J., Huber, C. G., *et al*. (2018) How many human proteoforms are there? *Nat. Chem. Biol.* **14**, 206–214
54. Dimitrakopoulos, L., Prassas, I., Diamandis, E. P., Nesvizhskii, A., Kislinger, T., Jaffe, J., and Drabovich, A. (2016) Proteogenomics: Opportunities and caveats. *Clin. Chem.* **62**, 551–557
55. Mohammed, H., Taylor, C., Brown, G. D., Papachristou, E. K., Carroll, J. S., and D'Santos, C. S. (2016) Rapid immunoprecipitation mass spectrometry of endogenous proteins (RIME) for analysis of chromatin complexes. *Nat. Protoc.* **11**, 316–326
56. Karakosta, T. D., Soosaipillai, A., Diamandis, E. P., Batruch, I., and Drabovich, A. P. (2016) Quantification of human kallikrein-related peptidases in biological fluids by multiplatform targeted mass spectrometry assays. *Mol. Cell. Proteomics* **15**, 2863–2876
57. Korbakis, D., Schiza, C., Brinc, D., Soosaipillai, A., Karakosta, T. D., Legare, C., Sullivan, R., Mullen, B., Jarvi, K., Diamandis, E. P., and Drabovich, A. P. (2017) Preclinical evaluation of a TEX101 protein ELISA test for the differential diagnosis of male infertility. *BMC Med.* **15**, 60
58. Korbakis, D., Brinc, D., Schiza, C., Soosaipillai, A., Jarvi, K., Drabovich, A. P., and Diamandis, E. P. (2015) Immunocapture-selected reaction monitoring screening facilitates the development of ELISA for the measurement of native TEX101 in biological fluids. *Mol. Cell. Proteomics* **14**, 1517–1526
59. Schiza, C., Korbakis, D., Jarvi, K., Diamandis, E. P., and Drabovich, A. P. (2019) Identification of TEX101-associated proteins through proteomic measurement of human spermatozoa homozygous for the missense variant rs35033974. *Mol. Cell. Proteomics* **18**, 338–351
60. Tomlins, S. A., Bjartell, A., Chinnaiyan, A. M., Jenster, G., Nam, R. K., Rubin, M. A., and Schalken, J. A. (2009) ETS gene fusions in prostate cancer: From discovery to daily clinical practice. *Eur. Urol.* **56**, 275–286
61. Birdsey, G. M., Shah, A. V., Dufton, N., Reynolds, L. E., Osuna Almagro, L., Yang, Y., Aspalter, I. M., Khan, S. T., Mason, J. C., Dejana, E., Gottgens, B., Hodivala-Dilke, K., Gerhardt, H., Adams, R. H., and Randi, A. M. (2015) The endothelial transcription factor ERG promotes vascular stability and growth through Wnt/beta-catenin signaling. *Dev. Cell* **32**, 82–96
62. Mertz, K. D., Setlur, S. R., Dhanasekaran, S. M., Demichelis, F., Perner, S., Tomlins, S., Tchinda, J., Laxman, B., Vessella, R. L., Beroukhim, R., Lee, C., Chinnaiyan, A. M., and Rubin, M. A. (2007) Molecular characterization of TMPRSS2-ERG gene fusion in the NCI-H660 prostate cancer cell line: A new perspective for an old model. *Neoplasia* **9**, 200–U203
63. Carver, B. S., Tran, J., Gopalan, A., Chen, Z., Shaikh, S., Carracedo, A., Alimonti, A., Nardella, C., Varmeh, S., Scardino, P. T., Cordon-Cardo, C., Gerald, W., and Pandolfi, P. P. (2009) Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nat. Genet.* **41**, 619–624
64. Zammarchi, F., Boutsalis, G., and Cartegni, L. (2013) 5′ UTR control of native ERG and of Tmprss2:ERG variants activity in prostate cancer. *PLoS One* **8**, e49721
65. Lai, A. C., Toure, M., Hellerschmied, D., Salami, J., Jaime-Figueroa, S., Ko, E., Hines, J., and Crews, C. M. (2016) Modular PROTAC design for the degradation of oncogenic BCR-ABL. *Angew. Chem. Int. Ed. Engl.* **55**, 807–810
66. Mounir, Z., Korn, J. M., Westerling, T., Lin, F., Kirby, C. A., Schirle, M., McAllister, G., Hoffman, G., Ramadan, N., Hartung, A., Feng, Y., Kipp, D. R., Quinn, C., Fodor, M., Baird, J., *et al*. (2016) ERG signaling in prostate cancer is driven through PRMT5-dependent methylation of the androgen receptor. *Elife* **5**, e13964
67. Zhang, Z., Chng, K. R., Lingadahalli, S., Chen, Z., Liu, M. H., Do, H. H., Cai, S., Rinaldi, N., Poh, H. M., Li, G., Sung, Y. Y., Heng, C. L., Core, L. J., Tan, S. K., Ruan, X., *et al*. (2019) An AR-ERG transcriptional signature defined by long-range chromatin interactomes in prostate cancer cells. *Genome Res.* **29**, 223–235
68. Mashtalir, N., D'Avino, A. R., Michel, B. C., Luo, J., Pan, J., Otto, J. E., Zullow, H. J., McKenzie, Z. M., Kubiak, R. L., Pierre, R. St., Valencia, A. M., Poynter, S. J., Cassel, S. H., Ranish, J. A., and Kadoch, C. (2018) Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell* **175**, 1272–1288.e1220
69. Pal, R. P., Kockelbergh, R. C., Pringle, J. H., Cresswell, L., Hew, R., Dormer, J. P., Cooper, C., Mellon, J. K., Barwell, J. G., and Hollox, E. J. (2016) Immunocytochemical detection of ERG expression in exfoliated urinary cells identifies with high specificity patients with prostate cancer. *BJU Int.* **117**, 686–696
70. Berg, K. D., Vainer, B., Thomsen, F. B., Roder, M. A., Gerds, T. A., Toft, B. G., Brasso, K., and Iversen, P. (2014) ERG protein expression in diagnostic specimens is associated with increased risk of progression during active surveillance for prostate cancer. *Eur. Urol.* **66**, 851–860
71. Kouspou, M. M., Fong, J. E., Brew, N., Hsiao, S. T. F., Davidson, S. L., Choyke, P. L., Crispino, T., Jain, S., Jenster, G. W., Knudsen, B. S., Millar, J. L., Mittmann, N., Ryan, C. J., Tombal, B., and Buzza, M. (2020) The movember prostate cancer landscape analysis: An assessment of unmet research needs. *Nat. Rev. Urol.* **17**, 499–512