SUPPLEMENTAL EXPERIMENTAL PROCEDURES

Generation of Tmprss2$^{-/-}$;TRAMP Mice. The generation and characterization of *Tmprss2* deficient mice has been previously described (40). *Tmprss2$^{-/-}$* mice on a hybrid 129X1/SvJ, 129S1/SV, C57Bl/6J background were crossed to the previously characterized TRAMP transgenic mice on the inbred C57Bl/6 background (41, 78). The resulting progeny were appropriately crossed to generate *Tmprss2$^{+/+}$;*TRAMP and *Tmprss2$^{-/-}$;*TRAMP mice. Analyses were performed on virgin males at the ages indicated and TRAMP negative littermates were used as controls.

Histology and Immunohistochemistry. Tissue microarrays (TMAs) were constructed from prostate cancer metastasis acquired in the context of a tissue acquisition necropsy under an IRB-approved protocol. At least three metastasis from 44 different patients were included for a total of 166 metastasis. TMA blocks were cut to 5 micron thickness and placed on charged glass slides, deparaffinized and rehydrated following standard protocols. The slides were initially treated with 0.3% hydrogen peroxide at room temperature for 10 minutes. Following antigen retrieval, slides were incubated in a 1:10000 dilution of the purified TMPRSS2 monoclonal P5H9-A3 antibody, or 1:100 ERG antibody (clone EPR3864–Epitomics, Burlingame, CA) at room temperature for 1 hour. Biotinylated secondary antibodies were utilized at a 1:200 dilution for 30 minutes. Bound antibodies were visualized using the Vector Elite ABC reagent (Vector Laboratories, Burlingame, CA) according to the manufacturer's directions. Negative controls for the immunohistochemical studies consisted of substituting a solution of normal mouse IgG (at 0.2 μg/ml) for the primary antibody. Staining was scored by assigning a single value to each core, based on a 0 to 3 scale; 0 (no staining), 1 (faint/equivocal staining), 2 (definite staining of a minority of cells) and 3 (definite staining of a majority of cells). Each metastasis was assigned a single expression score.

Mouse Pathology and Histology*.* All mice were anesthetized and weighed prior to sacrifice. The genitourinary tract (GU) consisting of the seminal vesicles, bladder, urethra, ampullary gland, and the prostate was removed and weighed (GU+SV). The seminal vesicles were then removed and the GU tract and re-weighed (GU-SV). The GU-SV block was bisected along the urethra and one half was frozen in OCT using liquid nitrogen while the other one half was fixed in 10% neutral buffered formalin. The pelvic lymph nodes, kidney, liver, and lung were routinely collected for histology. Fixed samples were routinely incubated in 10% neutral buffered formalin

at room temperature, transferred to 70% ethanol, dehydrated, and embedded in paraffin. Hematoxylin and eosin (HE) staining was performed on 5 μm sections. HE stained slides were reviewed by two pathologists (F. V. L. and L.T.) without knowledge of the genotype of the animals. Each slide was scored for the percentage area corresponding to normal prostate glands, well differentiated adenocarcinoma (WD), moderately differentiated adenocarcinoma (MD), poorly differentiated adenocarcinoma (PD), and phyllodes-like cancer (PHY) in accordance with previously described TRAMP tumor pathology (43, 78). Tissues were stained for the following antigens, as described above,: N-Cadherin (ECM Biosciences), Vimentin (AbCam), Androgen Receptor (GeneTex), caspase III (Cell Signaling) and Ki67 (DAKO).

qRT-PCR. Quantitative reverse-transcription PCR analysis was performed on laser captured microdissected prostate epithelia from the ventral and anterior lobes. For tumor bearing mice, only adenocarcinoma was collected. No samples from phyllodes-like regions were collected. Ventral lobe epithelia from 28-32 week old *Tmprss2*$^{+/+}$;TRAMP and *Tmprss2*$^{-/-}$;TRAMP mice was isolated and RNA extracted as previously described (79). Primers for the various genes analyzed are provided in Supplemental Table 1. Expression levels were normalized to the expression level of S16. All QRT-PCR reactions utilized Power SybrGreen and were performed on an Applied Biosystems (Foster City, CA) 7900HT Fast Real-Time PCR System. Reaction cycling conditions included a 10 minute incubation at 95°C followed by 40 cycles of 95°C for 15 seconds with a one minute extension at 60°C. Product disassociation curves were generated using the machines default conditions.

Cell Culture and Phenotypic Assays. *Constructs, Cell Lines and Expression*. TMPRSS2 was cloned via the PCR using KOD high fidelity polymerase (Novagen, San Diego, CA) from pooled human prostate cDNA (Clontech, Mountain View, CA) into the pCR2.1-TOPO vector following the manufacturers protocol (Invitrogen, Carlsbad, CA). The cloning primers were 5'-AGCAAGATGGCTTTGAACTCAGGGT-3' and 5'-GGATTAGCCGTCTGCCCTCATTTGTCGA-3'. A S441A point mutations for TMPRSS2 was generated using the QuikChange II sitedirected mutagenesis kit and protocol (Stratagene, La Jolla, CA). The pCR2.1-TMPRSS2 vector was used as template and the mutagenesis primers for, S441A mutation were 5'-TCTTGCCAGGGTGACGCTGGAGGGCCTC-3'and5' CTGGTCTCCGGAGGTCGCAGTG-GACCG-3'. Wildtype TMPRSS2 and TMPRSS2-S441A were subcloned into the pLenti6 mammalian expression lentiviral vector (Invitrogen, Carlsbad, CA). Infectious viral particles were packaged following the manufacturers protocol.

The human prostate cancer cell lines LNCaP, PC-3, DU145 and HepG2 cells were obtained from the American Type Culture Collection. BPH1 cells, C4-2B cells and TRAMP C2 cells were kindly provided by Dr. Simon Hayward, Dr. Leland Chung, and Dr. Norman Greenberg, respectively, and were used within 6 passages after receipt or were authenticated by matching transcript profiles of the cells used in these experiments with transcript profiles generated from the cell stocks originally provided or from public gene expression datasets corresponding to the specific cell line in order to confirm the identity of the cell line. Cells were maintained in RPMI 1640 with 10% FBS. Cells were transduced with viral particles at 50% confluence in the presence of 6 µg/ml of Polybrene. Stably transduced cells were selected for 14 days using 5 µg/ml of Blasticidin. Individual clonal populations were isolated using cloning disks and expanded before TMPRSS2 expression analysis. Cells were analyzed for TMPRSS2 expression levels by both Western blot, using an antibody raised against a portion of the protease domain of TMPRSS2, and by QRT-PCR. A shRNA lentiviral vector, targeting sequence CCGGCCTCTTAACAATCCATGGCATCTCGAGATGCCATGGATTGTTAAGAGGTTTT, to TMPRSS2 as well as a scrambled control sequence was obtained from Open Biosystems (Hunstville, AL). The shRNA lentiviral vectors were transfected and virus was generated in HEK293T cells following the company's protocol. Individual viral pools were tested in LNCaP C4-2B cells for their ability to knock down TMPRSS2 expression. Two separate and distinct lines of stably knocked down TMPRSS2 LNCaP C4-2B cells were generated through the use of the puromycin drug selectable marker in the vector.

In vitro Assays. In order to determine the effect of TMPRSS2 expression on prostate cancer cell proliferation, individually selected TMPRSS2 expressing cell line clones were assayed by two different methods. First, cells were plated at a density of 10,000 viable cells/well in triplicate in 6-well tissue culture plates. Cells were trypsinized, stained with trypan blue and counted every 2 days for 8 days. The growth media was changed on days 3 and 6 of the assay. Second, cells were plated at a density of 2000 viable cells/well in a 96-well tissue culture plate in replicates of 8. The number of metabolically active cells was measured every day for 8 days using the Celltiter 96 Cell Proliferation Assay from Promega (Madison, WI). Media was changed on days 3 and 6 of the assay. Each assay was performed a total of 4 times.

Cell invasion assays were performed using the Cultrex 96-Well BME Invasion Assay from Trevigen (Gaithersburg, MD). Briefly, 50,000 cells were plated with serum free media into the upper matrigel coated invasion chamber and allowed to invade for 24 hours toward the lower chamber which contained media with 10% FBS or various test compounds such as: 50 ng of

sc-HGF, 50 ng of *in vitro* TMPRSS2-digested HGF, 50 ng of *in vitro* matriptase-digested HGF, 50 ng *in vitro* TMPRSS2-digested HGF plus 500 ng of anti-HGF neutralizing antibody, conditioned media from prostate fibroblast cells or 250 µM bromhexine. After a number of rinses the invaded cells were detached, lysed, stained and counted in a fluorescence plate reader.

Anchorage-independent growth differences were assayed following the procedure of Arteaga, et al. (1988). Briefly, 10,000 cells were dispersed in 1 ml of 0.4% Sea-Plaque agarose (Cambrex, Rockland, ME) made with RPMI 1640 and 20% FBS. This single cell suspension was layered over an already hardened layer of 1ml of 0.8% Sea-plaque agarose made with complete media in 6-well tissue culture plates. After 21 days colonies were stained with crystal violet, counted and measured using an inverted microscope.

Microarray Methods. Total RNA from LCM experimental samples were isolated using the Arcturus PicoPure RNA Isolation kit (Molecular Devices, Sunnyvale, CA) incorporating DNAse-treatment using the RNase-Free DNase Set (Qiagen Inc, Valencia, CA.). To provide a reference standard RNA for use on two-color cDNA microarrays, we pooled total RNA isolated from normal adult male Swiss-Webster mice (10% prostate and 30% each testis, liver, and kidney). Reference RNA was purified using Trizol (Life Technologies, Rockville, MD) following the manufacturer's protocol followed by further purification by RNeasy maxi kit (Qiagen Inc, Valencia, CA) including DNAse treatment using the RNase-Free DNase Set (Qiagen Inc, Valencia, CA). Total RNA from experimental samples were amplified two rounds and reference total RNA was amplified one round using the Ambion MessageAmp aRNA Kit (Ambion Inc, Austin, TX), according to the manufacturer's specifications. The amplified RNA was used as template for cDNA probe synthesis followed by hybridization to a custom mouse prostate cDNA array (MPEDB array) composed of approximately 8,300 genes expressed in the developing and adult mouse prostate. cDNA probe pairs were prepared by amino-allyl reverse transcription using 2 µg of amplified RNA and labeling with Cy3-dCTP or Cy5-dCTP fluorescent dyes (Amersham Bioscience, Piscataway, NJ). Experimental and reference probes were combined and competitively hybridized to MPEDB microarrays under a coverslip for 16 h at 63°C. Fluorescent array images were collected for Cy3 and Cy5 emissions using a GenePix 4000B fluorescent scanner (Axon Instruments, Foster City, CA). Image intensity data were extracted and analyzed using GenePix Pro v4.1.1.44 software. Log2 ratios were normalized using the printtiploess function from the Limma package in R.

Metastasis and Xenograft Assays. Primary prostate tumors were removed from Tmprss2$^{-/-}$;

TRAMP or Tmprss2$^{+/+}$; TRAMP mice at 28-32 weeks of age and transferred into phenol red free RPMI with 10% fetal bovine serum and antibiotics. Part of each tumor was fixed or frozen for histological analysis. The remaining tumor tissue was minced with scalpels and passed through a 100 μm cell strainer (BD Falcon). Live cell number was estimated by trypan blue exclusion using a hemocytometer. Approximately 5x 10^5 live cells were injected into the tail vein of 6-8 week old male ICR SCID mice (Taconic). SCID recipients were euthanized 8 weeks after injection. Intratibial injection of primary tumor cells was performed essentially as previously described (80). Approximately 2 x 10^5 live tumor cells in total volume of 20 μl was injected into the proximal end of the tibia of 6-8 week old male ICR SCID mice. Mice were euthanized 8 weeks after injection. Radiographic images were obtained using a Faxitron with Kodak minR2000 film with a single Kodak minR screen at 25 kV for 5 sec. For histological analysis, hind limbs were removed and cleaned of residual tissue, fixed in neutral buffered formalin, transferred to 70% ethanol, and decalcified in EDTA prior to embedding in paraffin. HE staining was performed on 5 μm sections.

Recombinant TMPRSS2*.* To facilitate expression in *Pichia pastoris* a single N-linked glycosylation site was mutated (N249G) using the QuickChange site directed mutagenesis protocol (Stratagene). pCR2.1-TMPRSS2 served as the template for the reaction and the mutagenesis primers were TMPRSS2-N249G-Forward (5'TGCGGGTCAACTTGGGCTCAAGCCGCCAGAG) and TMPRSS2-N249G-Reverse (5'CTCTGGCGGCTTGAGCCCAAGTTGACCCCGCA). The scavenger receptor cysteine rich and serine protease domain regions of TMPRSS2-N249G (amino acids 148-492) were subcloned into the *Pichia pastoris* secretory expression vector pPIC9K (Invitrogen). The KM71 strain of *P. pastoris* was transfected and individual clones were assayed for expression of TMPRSS2.

Optimal TMPRSS2 expression was achieved by growing the *P. pastoris* clone at 25°C with agitation to a density of OD_{600}=1 and then inducing with 0.5% methanol each day for 4 days. Secreted TMPRSS2 was purified by pelleting the *P. pastoris* at 15,000G for 30 minutes. The media was then filtered through a 0.4 micron filter and the TMPRSS2 precipitated with 40% ammonium sulfate overnight at 4°C with constant stirring. Precipitated protein was collected by centrifugation at 22,000G for 1 hour. The protein was solubilized in 0.1M Tris (8.8) / 0.5M NaCl / 0.1% CHAPS and dialyzed for 2 days at 4°C against the same buffer without the CHAPS. Insoluble material was filtered out and the clarified supernatant was incubated with a Benzamidine sepharose matrix (GE Healthcare) for 4 hours. The matrix was passed in column and washed exhaustively with 0.1M Tris (8.8) / 0.5M NaCl. TMPRSS2 was eluted with 5 column volumes of

20mM p-aminobenzamidine in the wash buffer. TMPRSS2 was then purified by size exclusion chromatography using Superdex 75 (GE Healthcare) pre-equilibrated with 0.1M Tris (8.8) / 0.5M NaCl. Isolated fractions were dialyzed against the equilibration buffer and assayed for activity against the fluorogenic peptide Boc-Gln-Ala-Arg-MCA.

TMPRSS2 Substrate Identification.

Recombinant TMPRSS2. To facilitate expression in *Pichia pastoris* a single N-linked glycosylation site was mutated (N249G) using the QuickChange site directed mutagenesis protocol (Stratagene). pCR2.1-TMPRSS2 served as the template for the reaction and the mutagenesis primers were TMPRSS2-N249G-Forward (5'TGCGGGTCAACTTGGGCTCAAGCCGCCAGAG) and TMPRSS2-N249G-Reverse (5'CTCTGGCGGCTTGAGCCCAAGTTGACCCCGCA). The scavenger receptor cysteine rich and serine protease domain regions of TMPRSS2-N249G (amino acids 148-492) were subcloned into the *Pichia pastoris* secretory expression vector pPIC9K (Invitrogen). The KM71 strain of *P. pastoris* was transfected and individual clones were assayed for expression of TMPRSS2. The expression of TMPRSS2 was optimized by growing the *P. pastoris* clone at 25°C with agitation to a density of OD_{600}=1 and then inducing with 0.5% methanol each day for 4 days. Secreted TMPRSS2 was purified by pelleting the *P. pastoris* at 15,000G for 30 minutes. The media was then filtered through a 0.4 micron filter and the TMPRSS2 precipitated with 40% ammonium sulfate overnight at 4°C with constant stirring. Precipitated protein was collected by centrifugation at 22,000G for 1 hour. The protein was solubilized in 0.1M Tris (8.0) / 0.5M NaCl / 0.1% CHAPS and dialyzed for 2 days at 4°C against the same buffer without the CHAPS. Insoluble material was filtered out and the clarified supernatant was incubated with a Benzamidine sepharose matrix (GE Healthcare) for 4 hours. The matrix was passed in column and washed exhaustively with 0.1M Tris (8.0) / 0.5M NaCl. TMPRSS2 was eluted with 5 column volumes of 20mM p-aminobenzamidine in the wash buffer. TMPRSS2 was then purified by size exclusion chromatography using Superdex 75 (GE Healthcare) pre-equilibrated with 0.1M Tris (8.0) / 0.5M NaCl. Isolated fractions were dialyzed against the equilibration buffer and assayed for activity.

Positional Scanning of Synthetic Combinatorial peptide Libraries: PS-SCLs were assayed with recombinant human TMPRSS2. Combinatorial libraries were prepared in 96-well Microfluor Black 'U'-bottom plates (Dynex Technologies, Chantilly, VA, U.S.A.) as described previously (44). Libraries were assayed under optimal buffer conditions in 50 mM Tris/HCl, pH 8.2, 2% (v/v) PEG [poly(ethylene glycol)]-8000 and 0.2% (v/v) NP40 (Nonidet P40). PS-SCLs were diluted to 0.25 mM in a final 100 µl assay volume. Thus, the individual tetrapeptides in each pool

were at a concentration far below the expected K_m of TMPRSS2, and the initial rates were proportional to the specificity constant k_{cat}/K_m (30 nM each peptide per well for P1-library). Proteolysis reactions were initiated by the addition of TMPRSS2 (1 nM). Fluorescence was monitored with a SpectraMAX Gemini fluorimetric 96-well plate reader (Molecular Devices Corporation, Sunnyvale, CA, U.S.A.) at 21 °C and the rate of substrate hydrolysis was analyzed with the SOFTmax PRO software (version 3.1.1, Molecular Devices Corporation). Excitation and emission was measured at 380 nm and 460 nm respectively (21 °C for 10 min).

Candidate Substrate Identification. In order to identify preferred protein substrates for TMPRSS2, we used the information obtained from the PS-SCL peptide profiling and the testing of individual peptide substrates to create a database of human sequences that contain a trypsin-like fold as extracted from Pfam (http://pfam.wustl.edu) and MEROPS. We then selected serine protease-fold like proteins that had (i) a known or predicted P4–P1 sequence in their activation domain, (ii) an arginine at the P1 position, and (iii) a hydrophobic amino acid (isoleucine or valine) at the P1' position to bind to the large, hydrophobic S1 pocket of TMPRSS2. These proteins were clustered and ranked based on matches with the preferred amino acids at the P1, P2, P3 and P4 positions as obtained by the PS-SCL screen. Cluster A contained candidate substrates where all four cleavage site amino acids in the activation domain matched favored ones, such as glycine, isoleuce, methionine or proline at P4, glutamic acid, methionine or glutamine at P3, threonine, phenylalanine, tryptophan, alanine or valine at P2 and arginine at P1. Clusters B and C contained precursors where only three or two positions matched a favored amino acid respectively. We then tested the ability of TMPRSS2 to cleave selected candidate substrates from each cluster such as the HGF precursor, tPA and KLK2. These candidates were subjected to standard *in vitro* digests. The digests were analyzed by SDS/PAGE under reducing conditions.

HGF Assays. Single chain pro-Hepatocyte growth factor was purchased from R&D Systems (Minneapolis, MN). HGF was incubated with TMPRSS2 for 60 minutes at 37°C in 0.1M Tris (8.8) / 0.5M NaCl . The products were electrophoresed under reducing conditions on a 4-12% polyacrylamide gradient gel (Invitrogen) and transferred to nitrocellulose. The blot was probed with a polyclonal antibody against HGF (R&D systems). The prostate cancer cell line DU-145 was seeded at 600000 cells per 35 mm culture plate in RPMI 1640 with 10%FBS. After 6 hours at 37°C the cells were washed twice with PBS and incubated for 15 min in 1 ml of RPMI 1640 medium containing no FBS, 10% FBS, 50 ng of sc-HGF, 50 ng of *in vitro* TMPRSS2-digested HGF, 50 ng of *in vitro* matriptase-digested HGF or 50 ng *in vitro* TMPRSS2-digested HGF plus

500 ng of anti-HGF neutralizing antibody. Whole cell lysates were made using RIPA buffer containing protease and phosphatase inhibitors (Roche Applied Sciences). Proteins were electrophoresed and blotted onto nitrocellulose and probed with monoclonal anti-human Met (25H2) and polyclonal rabbit anti-human phospho-Met antibodies from Cell Signaling Technology (Beverly, MA, U.S.A.).

Substrate Assays. 2.5 µg of pro-UPA (Cortex Biochem, San Leandro, CA)(not shown), pro-TPA (Meridian Life Science,) Glu-plasminogen (Heamatologic Technologies Inc., Essex Junction, VT), MMP-2, MMP-9 (R&D Systems) and pro-KLK2 (Fitzgerald Industries, Acton, MA) were incubated with TMPRSS2 for 5 or 60 minutes at 37°C in 0.1M Tris (8.8) / 0.5M NaCl. Resulting products were visualized on coomassie blue stained polyacrylamide gels. Additionally, pro-UPA and pro-TPA that had been pre-incubated with TMPRSS2 were incubated with Glu-plasminogen for 60 minutes.

Supplemental Table 1. PCR Primers.

TMPRSS2	
Forward	5'-CAT GAT CTG TGC CGG CTT CCT GCA GG-3'
Reverse	5'-CTT GTA TCC CCT ATC AGC CAC CAG ATA-3'
RPL13A	
Forward	5'-CCT GGA GGA GAA GAG GAA AGA GA-3'
Reverse	5'-TTG AGG ACC TCT GTG TAT TTG TCA A-3'
S16	
Forward	5'-AGGAGCGATTTGCTGGTGTGGA-3'
Reverse	5'-GCTACCAGGCCTTTGAGATGGA-3'.
Twist	
Forward	5' - AGT CTG AAC ACT CGT TTG TGT CCC – 3'
Reverse	5' - TGC CTT TCC TGT CAG TGG CTG ATT – 3'
Twist2	
Forward	5'- ACT GGA CCA AGG CTC TCA GAA CAA - 3'
Reverse	5'- TTC CAG GCT TCC TCG AAA CAG TCA - 3'
Vimentin	
Forward	5' - AGATGGCTCGTCACCTTCGTGAAT - 3'
Reverse	5' - TTGAGTGGGTGTCAACCAGAGGAA - 3'
Cxcl 12α	
Forward	5' – TTCAACAGTGGCTCTATGGGCTCT – 3'

Reverse 5’ – GGCTTCATGGCAAGATTCTGGCTT – 3’

PCNA
Forward 5’ – GGCCGTGAACCTCACCAGTAT – 3’
Reverse 5’ – TCTCGGCATATACGTGCAAA – 3’

p21
Forward 5’ – GACACCACTGGAGGGTGACT – 3’
Reverse 5’ – CAGGTCCACATGGTCTTCCT – 3’

p16
Forward 5’ – CACCGAATAGTTACGGTCGG – 3’
Reverse 5’ – GCACGGGTCGGGTGAGAGTG – 3’

p27
Forward 5’ – CAGTATTTCATTGCCTGTGTATGG - 3’
Reverse 5’ – ATGCCACTTTGGCTTGTATATTG – 3’

SV40T
Forward 5’ – CAGAGCAGAATTGTGGAGTGG – 3’
Reverse 5’ – GGACAAACCACAACTAGAATGCAGTG – 3’

TCR
Forward 5’ – CACTGTTGCTTGTCTTGTG – 3’
Reverse 5’ – GTCAGTCGAGTGCACAGTTT – 3’

Hepsin
Forward 5’ – CCAGGGTCCTCTCTTCCA – 3’
Reverse 5’ – GTCCCAGACAGCAGAACAATA – 3’

Matriptase
Forward 5’ – AGGACACTGGTGGTTCTACT – 3’
Reverse 5’ – GGGAGGCAGATATGCACAAA – 3’