The Role of *SPINK1* in *ETS* Rearrangement-Negative Prostate Cancers

Scott A. Tomlins,[1,3,22] Daniel R. Rhodes,[1,3,4,22] Jianjun Yu,[1,4,22] Sooryanarayana Varambally,[1,3,6] Rohit Mehra,[1,3,6] Sven Perner,[7,9,12] Francesca Demichelis,[8,9,10] Beth E. Helgeson,[1,3] Bharathi Laxman,[1,3] David S. Morris,[5] Qi Cao,[1,3] Xuhong Cao,[1,2] Ove Andrén,[13] Katja Fall,[14] Laura Johnson,[9,10] John T. Wei,[5,6] Rajal B. Shah,[1,3,5,6] Hikmat Al-Ahmadie,[15] James A. Eastham,[16] Scott E. Eggener,[16] Samson W. Fine,[15] Kristina Hotakainen,[18] Ulf-Håkan Stenman,[18] Alex Tsodikov,[1,4,6,19] William L. Gerald,[15] Hans Lilja,[16,17,20] Victor E. Reuter,[15] Phillip W. Kantoff,[7,9,10,11] Peter T. Scardino,[16] Mark A. Rubin,[8,9,10,11] Anders S. Bjartell,[16,21,23] and Arul M. Chinnaiyan[1,2,3,4,5,6,23,*]

[1]Michigan Center for Translational Pathology
[2]Howard Hughes Medical Institute
[3]Department of Pathology
[4]Center for Computational Medicine and Biology
[5]Department of Urology
[6]Comprehensive Cancer Center
University of Michigan Medical School, Ann Arbor, MI 48109, USA
[7]Department of Medicine
[8]Department of Pathology
Brigham and Women's Hospital, Boston, MA 02115, USA
[9]Brigham and Women's Hospital Boston, MA 02115, USA
[10]Harvard Medical School, Boston, MA 02115, USA
[11]Dana-Farber Cancer Institute, Boston, MA 02115, USA
[12]Institute of Pathology, University Hospital Ulm, D-89081 Ulm, Germany
[13]Department of Urology, Örebro University Hospital, SE-701 85 Örebro, Sweden
[14]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-171 77 Stockholm, Sweden
[15]Department of Pathology
[16]Department of Surgery/Urology Services
[17]Department of Clinical Laboratories and Medicine
Memorial Sloan-Kettering Cancer Center, New York, NY 10065, USA
[18]Department of Clinical Chemistry, Helsinki University Central Hospital, 00290 Helsinki, Finland
[19]Department of Biostatistics, School of Public Health, University of Michigan, Ann Arbor, MI 48109, USA
[20]Department of Laboratory Medicine, University Hospital UMAS, Lund University, SE-205 02 Malmö, Sweden
[21]Department of Urology, University Hospital UMAS, Lund University, SE-205 02 Malmö, Sweden
[22]These authors contributed equally to this work
[23]These authors contributed equally to this work
*Correspondence: arul@umich.edu
DOI 10.1016/j.ccr.2008.04.016

SUMMARY

ETS gene fusions have been characterized in a majority of prostate cancers; however, the key molecular alterations in *ETS*-negative cancers are unclear. Here we used an outlier meta-analysis (meta-COPA) to identify *SPINK1* outlier expression exclusively in a subset of *ETS* rearrangement-negative cancers (~10% of total cases). We validated the mutual exclusivity of *SPINK1* expression and *ETS* fusion status, demonstrated that *SPINK1* outlier expression can be detected noninvasively in urine, and observed that *SPINK1* outlier expression is an independent predictor of biochemical recurrence after resection. We identified the aggressive 22RV1 cell line as a *SPINK1* outlier expression model and demonstrate that *SPINK1* knockdown in 22RV1 attenuates invasion, suggesting a functional role in *ETS* rearrangement-negative prostate cancers.

SIGNIFICANCE

While *ETS* rearrangements play a role in a majority of prostate cancers, little is known about molecular alterations driving *ETS* gene fusion-negative cancers. In this study, we identified *SPINK1* outlier expression exclusively in a subset of *ETS*-negative cancers. *SPINK1* is associated with prostate cancer aggressiveness and can be detected noninvasively in urine. Furthermore, *SPINK1* mediates invasion in a prostate cancer cell line with outlier expression. The mechanism of *SPINK1* outlier expression remains to be characterized and is not explained by chromosomal rearrangement, deletion, or amplification. Thus, *SPINK1* is a biomarker specific to a subset of aggressive *ETS*-negative prostate cancers. Our study also demonstrates the utility of a meta-outlier strategy to identify cancer subtypes.

INTRODUCTION

Recently, we developed a bioinformatics approach termed Cancer Outlier Profile Analysis (COPA) to nominate candidate oncogenes from transcriptomic data based on high expression in a subset of cases ("outlier expression") (Tomlins et al., 2005). When applied to the Oncomine compendium of tumor profiling studies (http://www.oncomine.org) (Rhodes et al., 2004), COPA correctly identified several known oncogenes as outliers, such as *ERBB2* in breast cancer and *PBX1* in leukemia. In addition, COPA identified the *ETS* family members *ERG* and *ETV1* as high-ranking outliers in multiple prostate cancer profiling studies, leading to the discovery of recurrent gene fusions involving the 5′ untranslated region of the androgen-regulated gene *TMPRSS2* with *ERG*, *ETV1*, *ETV4*, or *ETV5* in prostate cancer cases that overexpressed the respective *ETS* family member (Helgeson et al., 2008; Tomlins et al., 2005, 2006). Recently, we identified additional 5′ fusion partners in cases with *ETS* family member outlier expression (Tomlins et al., 2007a).

ETS gene fusions occur in 40%–80% of prostate-specific antigen (PSA)-screened prostate cancers, leaving 20%–60% of prostate cancers in which the key genetic aberration cannot be ascribed to *ETS* gene fusions. Additionally, we have determined that *ETS*-positive and -negative cancers have distinct transcriptional signatures across profiling studies (Tomlins et al., 2007b), suggesting that fusion-negative cancers activate unique oncogenes and downstream targets. Here, we attempted to identify such candidate oncogenes through their outlier expression in *ETS*-negative prostate cancers.

The utility of COPA and other strategies to identify outlier genes from microarray data was recently demonstrated in multiple myeloma and breast cancer (Annunziata et al., 2007; Naderi et al., 2007), suggesting that this strategy can be applied across human cancers to identify relevant subtypes. Here, we refined our COPA strategy based on observations from our initial application of COPA. We observed that correctly identified oncogenes, including *ERG* and *ETV1*, were typically high-ranking outliers in multiple data sets (Tomlins et al., 2005). This suggests that true candidate oncogenes should demonstrate strong outlier profiles across independent studies and supports the use of a meta-analysis-based COPA approach.

RESULTS AND DISCUSSION

Thus, in this study, we performed a focused application of COPA to seven prostate cancer profiling studies (Dhanasekaran et al., 2001; Glinsky et al., 2004; Lapointe et al., 2004; LaTulippe et al., 2002; Vanaja et al., 2003; Welsh et al., 2001; Yu et al., 2004) in the Oncomine database (Rhodes et al., 2004), as described in the Experimental Procedures, to prioritize candidate oncogenes in *ETS*-negative prostate cancers. Twenty-nine genes were nominated as outliers in at least three of the seven data sets (see Table S1 available online), with 11 genes identified as outliers in at least four of the seven data sets (Table 1). Consistent with our previous application of COPA filtered by causal cancer genes (Tomlins et al., 2005), both *ERG* and *ETV1* were high-ranking meta-outliers: *ERG* ranked as the first meta-outlier (seven studies) and *ETV1* as the fifth meta-outlier (four studies).

Table 1. Meta-COPA Analysis of Seven Prostate Cancer Gene Expression Profiling Data Sets in Oncomine

Meta-COPA Rank	Gene	Number of Studies	Average COPA Rank
1	*ERG*[a]	7	19.3
2	*SPINK1*[a,b]	5	29.8
3	*GPR116*[c]	5	46
4	*ORM1*[d]	4	10
5	*ETV1*[a]	4	23
6	*MYL2*[d]	4	26.8
7	*NEB*[d]	4	27
8	*TGM4*[d]	4	30.8
9	*NELL2*[d]	4	33.5
10	*KRT13*[d]	4	49
11	*SLC26A4*[d]	4	63.3

Genes were ranked by the number of studies in which they scored in the top 100 outliers (ranked by COPA) at any of the three predefined percentile cutoffs (75th, 90th, and 95th). Genes were further ranked by their average COPA rank in studies in which they ranked in the top 100.
[a] *ETS* gene.
[b] Gene showing outlier expression exclusively in prostate cancer and mutually exclusive outlier expression with *ETS* genes.
[c] Gene without mutual exclusivity with *ERG* or *ETV1* outlier expression.
[d] Gene showing outlier expression in benign prostate tissue.

To identify candidate oncogenes activated in *ETS*-negative prostate cancers, we analyzed the remaining top meta-outliers for two characteristics: (1) overexpression in prostate cancer compared to benign prostate tissue and (2) mutually exclusive overexpression with *ERG* and *ETV1* (as ~95% of cancers with *ERG* or *ETV1* overexpression have detectable *ETS* fusions [Tomlins et al., 2005, 2007a]). Specific examples of meta-outliers that failed one or both criteria are shown in Figure S1 and Table 1. We identified *SPINK1* (serine peptidase inhibitor, Kazal type 1), the second-ranked meta-outlier, as showing overexpression in prostate cancer compared to benign prostate tissue and mutually exclusive overexpression with *ERG* and *ETV1* across multiple studies. *SPINK1* was not measured in one of the studies in the meta-analysis (Lapointe et al., 2004) and ranked in the top ten in two of the remaining six studies.

The profile of *SPINK1* expression and scatter plots with *ERG* and *ETV1* for two studies (Glinsky et al., 2004; Yu et al., 2004) where *SPINK1* was identified as a top 100 outlier are shown in Figure 1, with plots from the other four studies in the meta-analysis (Dhanasekaran et al., 2001; LaTulippe et al., 2002; Vanaja et al., 2003; Welsh et al., 2001) shown in Figure S2. *SPINK1* expression in an additional unpublished prostate cancer profiling study (NCBI GEO data set GSE8218, where *SPINK1* was the third-ranked outlier at the 90th percentile) and two multicancer studies profiling prostate cancer (Su et al., 2001 and NCBI GEO data set GSE2109) is also shown in Figure S2. In total, from these nine studies, *SPINK1* showed outlier expression (see Experimental Procedures) in only 4 of 136 (2.9%) benign prostate tissue samples and 56 of 376 (14.9%) clinically localized prostate cancers (two-sided Fisher's exact test, p = 9.5E−7). Remarkably, 372 of 376 profiled clinically localized prostate cancers (98.9%) showed mutually exclusive outlier expression
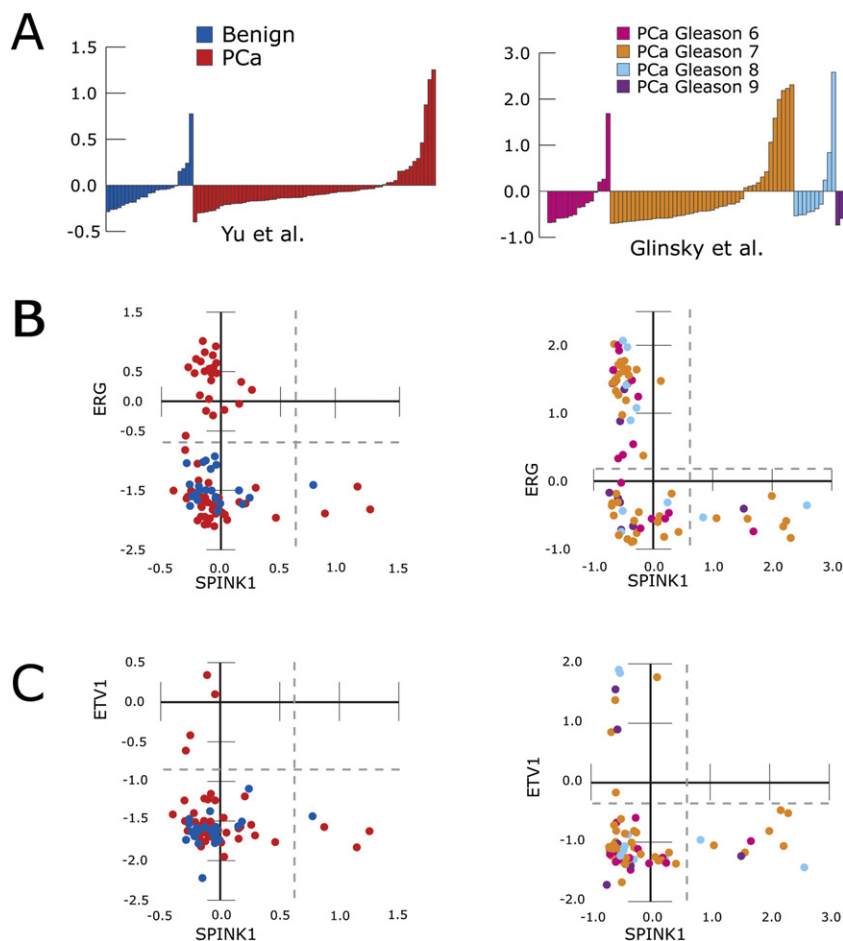
of *SPINK1*, *ERG*, and *ETV1*, as shown in Figure 1, Figure S2, and Table S2.

To confirm the outlier expression of *SPINK1* exclusively in *ETS*-negative prostate cancers, we measured *SPINK1*, *ERG*, and *ETV1* expression by quantitative PCR (qPCR) in an independent cohort of 10 benign prostate tissues and 61 prostate cancers (54 clinically localized and 7 metastatic samples). While *ERG*, *ETV1*, and *SPINK1* showed outlier expression in 25 (41%), 4 (6.5%), and 4 (6.5%) of 61 prostate cancers (clinically localized and metastatic), respectively, no benign prostate tissue samples demonstrated outlier expression of these genes. Consistent with the above microarray studies, *ERG*, *ETV1*, and *SPINK1* showed outlier expression in distinct cancers (Figure S3).

After demonstrating that *SPINK1* outlier expression defines a subset of *ETS* rearrangement-negative prostate cancers at the transcript level, we evaluated the expression of SPINK1 protein in prostate cancers. By immunohistochemical (IHC) analysis of tissue microarrays (TMAs), we evaluated SPINK1 expression in two independent cohorts (University of Michigan [UM] and Swedish Watchful Waiting [SWW]) representing a total of 392 cases of clinically localized prostate cancers. We have previously evaluated both cohorts for *TMPRSS2:ERG* fusion status by fluorescence in situ hybridization (FISH) (Demichelis et al., 2007; Mehra et al., 2007). In both cohorts, prostate cancer epithelia exhibited either strong or no expression of SPINK1, without intermediate staining as observed for many prostate cancer markers. As shown in Figure 2, in the UM cohort, 10 and 36 of 75 cases were positive for SPINK1 expression (13.3%) and *TMPRSS2:ERG* fusions (48%), respectively, with all SPINK1-positive cases being *TMPRSS2:ERG*-negative (one-sided Fisher's exact test, p = 0.0008). In the SWW cohort, 23 and 57 of 312 cases were positive for SPINK1 expression (7.4%) and *TMPRSS2:ERG* fusions (18.3%), respectively, again with all SPINK1-positive cases being *TMPRSS2:ERG*-negative (one-sided Fisher's exact test, p = 0.008).

Approximately 25%–40% of patients treated by radical prostatectomy for clinically localized prostate cancer will experience disease recurrence, initially indicated by an increase in the serum level of PSA (biochemical recurrence) (Han et al., 2001; Hull et al., 2002). Thus, we next sought to determine whether *SPINK1* outlier status was associated with biochemical recurrence after surgical resection. We identified two data sets from the evaluated cohorts for which we had access to follow-up biochemical recurrence information and a sufficient number of *SPINK1*-positive cases (>5). We first examined the Glinsky et al. (2004) gene expression data set, which contained tumors from 79 patients (with 37 recurrences), 10 of which showed outlier mRNA transcript expression of *SPINK1*. These patients had a significantly higher risk of recurrence than patients without *SPINK1* outlier expression (hazard ratio = 2.65; 95% CI = 1.16–6.07; log rank p = 0.016) by Kaplan-Meier analysis (Figure 3A). Multivariate Cox proportional-hazards regression analysis also revealed that *SPINK1* outlier status, independent of Gleason score, lymph node status, surgical margin status, age, and preoperative PSA, was a significant predictor of clinical recurrence of prostate cancer (hazard ratio = 2.5; 95% CI = 1.1–6.0; p = 0.035; Table S3).
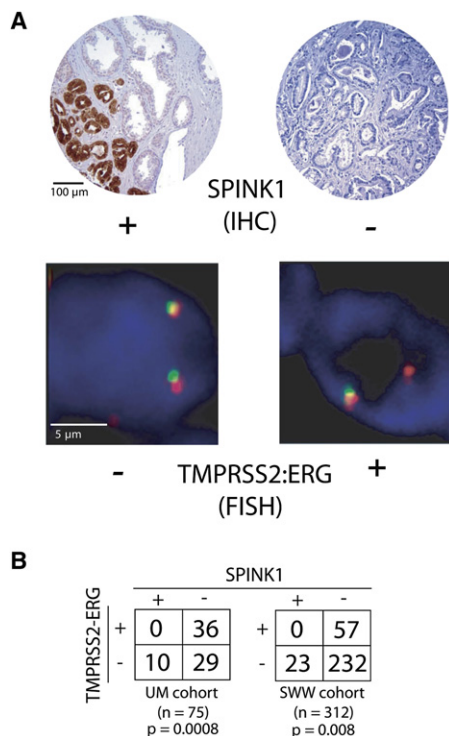
Figure 2. Confirmation of *SPINK1* Outlier Expression Exclusively in *ETS*-Negative Prostate Cancers

SPINK1 protein expression was evaluated in two cohorts (University of Michigan [UM] and Swedish Watchful Waiting [SWW]) using immunohistochemistry (IHC) on tissue microarrays previously evaluated for *TMPRSS2:ERG* status by fluorescence in situ hybridization (FISH).

(A) Representative SPINK1-positive and -negative cores, along with cells from the same cores negative and positive for *TMPRSS2:ERG* rearrangement by FISH. A *TMPRSS2:ERG* rearrangement through intrachromosomal deletion is indicated by loss of one 5′ (green) ERG signal.

(B) Contingency tables for SPINK1 expression and *TMPRSS2:ERG* status and p values for Fisher's exact tests for both cohorts.

We next performed the same analysis on the UM cohort (75 cases, 28 recurrences) evaluated for SPINK1 status by IHC. By Kaplan-Meier analysis, SPINK1-positive staining was significantly associated with biochemical recurrence (hazard ratio = 2.49; 95% CI = 1.01–6.18; p = 0.04; Figure 3B). Multivariate Cox proportional-hazards regression analysis again confirmed that SPINK1 status predicted recurrence independently of other clinical parameters (Table S3). With an adjusted hazard ratio of 4.1 (95% CI = 1.4–11.7; p = 0.009), it was the strongest predictor in this model.

As a final validation, we performed IHC for SPINK1 status on an independent cohort of 817 evaluable prostate cancers (200 recurrences) from the Memorial Sloan-Kettering Cancer Center (MSKCC). In this MSKCC cohort, with IHC performed independently from the UM and SWW cohorts using a different SPINK1 antibody, 297 of the 817 cases (36%) showed positive SPINK1 immunoreactivity in at least one of three triplicate cores. In addition, staining intensity was more variable than that observed in the UM and SWW cohorts. As the percentage of cases in this cohort with SPINK1 staining (36%) was far greater than the other IHC cohorts (13% and 7%) or the percentage of *SPINK1* outlier samples from DNA microarray and qPCR studies (15% and 7%;

see Table S4), we defined SPINK1-positive cases in the MSKCC cohort as those with at least one core showing >80% of cells showing positive SPINK1 immunoreactivity, resulting in 75 SPINK1-positive cases (9%), consistent with the other studies.

By Kaplan-Meier analysis, SPINK1-positive cases in the MSKCC cohort showed significantly shorter time to biochemical recurrence (hazard ratio = 2.32; 95% CI = 1.59–3.39; p = 6.96E–06; Figure 3C). Multivariate Cox proportional-hazards regression analysis again confirmed that *SPINK1* outlier status, independent of Gleason score, lymph node status, surgical margin status, seminal vesicle invasion, extracapsular extension, and preoperative PSA, was a significant predictor of clinical recurrence (hazard ratio = 2.02; 95% CI = 1.37–2.99; p = 0.0004; Table S3). Clinically, nomograms are commonly used to predict the likelihood of biochemical recurrence after surgical resection by optimally incorporating clinical and pathological parameters. To determine whether the addition of SPINK1 improves a validated nomogram for predicting the 7-year postprostatectomy probability of biochemical recurrence (Kattan et al., 1999), we assessed the concordance index (Kattan et al., 2003) (the probability that given two randomly selected patients, the patient with the worse outcome is indeed predicted to have a worse outcome) of the nomogram and the nomogram plus SPINK1 status. The bootstrap-corrected concordance index was minimally improved in all three data sets by the addition of SPINK1 status to the nomogram (Glinsky et al., [2004], 0.772 versus 0.762; UM IHC, 0.698 versus 0.676; MSKCC, 0.775 versus 0.765). Thus, while SPINK1 does not dramatically add to the predictive ability of an optimized multivariate model, we demonstrated by analyzing 971 cancers from three independent cohorts that *SPINK1* outlier status identifies an aggressive subset of prostate cancers.

We next sought to determine whether outlier expression of *SPINK1* could be detected noninvasively. As increased serum levels of SPINK1 occur in multiple malignancies (Paju and Stenman, 2006; Stenman, 2002), and as 44% of patients with prostate cancer are reported to have elevated serum levels of SPINK1 (Paju et al., 2007), we sought to establish a more specific assay to identify patients with tumors showing *SPINK1* outlier expression. We have recently described the detection of *TMPRSS2:ERG* fusion transcripts in the urine of men with prostate cancer (Laxman et al., 2006), and this assay allows us to more directly assess transcripts contributed by prostatic cells. Thus, we assessed *SPINK1* expression from a cohort of 148 urine samples collected from men with prostate cancer that we have characterized as *TMPRSS2:ERG* positive (43) or negative (105). As expected, *SPINK1* expression was higher in *TMPRSS2:ERG*-negative versus -positive samples (Mann-Whitney U test, p = 5E–5), and 21 of the 22 samples with the highest *SPINK1* expression were *TMPRSS2:ERG* negative. Using the same method to identify *SPINK1* outlier samples as described for our tissue qPCR cohort, 1 of the 43 *TMPRSS2:ERG*-positive samples (2.3%) showed *SPINK1* outlier expression, and 10 of the 105 *TMPRSS2:ERG*-negative samples (10%) showed *SPINK1* outlier expression (Figure 4) (Fisher's exact test, p = 0.12). In addition, compared to urine collected from 96 men presenting for evaluation of prostate cancer with negative needle biopsies, *SPINK1* expression is a significant predictor of prostate cancer in both univariate and multivariate analyses, and no negative samples show *SPINK1* outlier expression (Laxman et al., 2008).
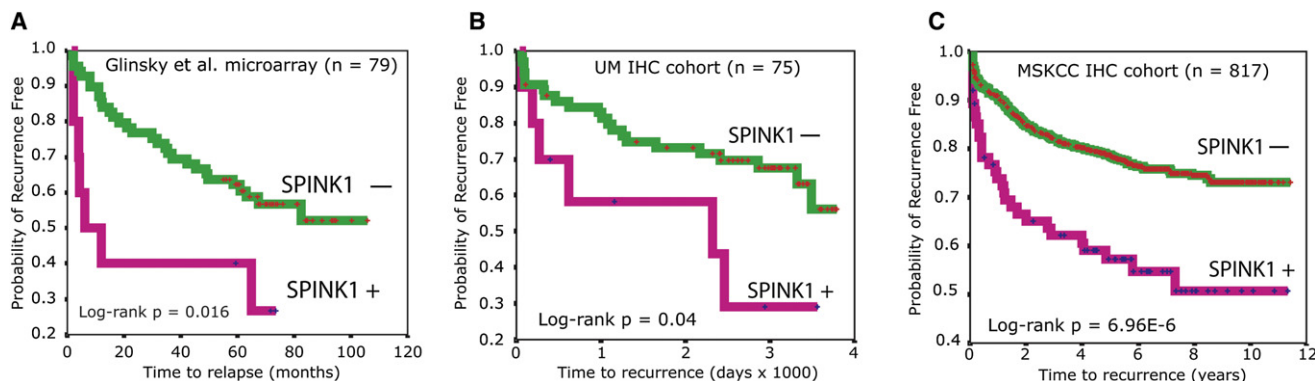
Figure 3. *SPINK1* Outlier Expression Identifies an Aggressive Subtype of *ETS*-Negative Prostate Cancers

Relationship between *SPINK1* outlier expression and biochemical recurrence after surgical resection. Kaplan-Meier analyses of *SPINK1* outlier expression from the Glinsky et al. (2004) DNA microarray data set (A) and SPINK1 IHC from the UM (B) and Memorial Sloan-Kettering Cancer Center (MSKCC) (C) cohorts and biochemical recurrence after surgical resection are shown.

SPINK1 encodes a 56 amino acid secreted peptide, also known as PSTI or TATI. Originally isolated from bovine pancreas and human pancreatic juice, its normal function is thought to be the inhibition of serine proteases such as trypsin (Greene et al., 1976; Haverback et al., 1960; Kazal et al., 1948; Paju and Stenman, 2006). SPINK1 levels are strongly elevated during inflammation and pancreatitis (Paju and Stenman, 2006). Like the pancreas, the prostate gland also secretes a variety of serine proteases, most notably the kallikrein enzyme PSA, but also trypsin, the expression of which is increased in prostate cancer (Bjartell et al., 2005). Thus, *SPINK1* outlier expression may have a role in modulating the activity of cancer-related proteases. Additionally, SPINK1 has been reported to stimulate DNA synthesis in rat pancreatic cancer cells and human fibroblasts, suggesting additional roles in oncogenesis (Freeman et al., 1990; Ogawa et al., 1985). *SPINK1* mRNA and protein have been detected in a variety of benign and cancerous tissues, and its expression in prostate and prostate cancer has recently been described (Paju et al., 2007; Paju and Stenman, 2006;

Figure 4. *SPINK1* Outlier Expression Can Be Detected Noninvasively in Urine

Noninvasive detection of *SPINK1* outlier expression in men with *TMPRSS2:ERG*-negative prostate cancers. Total RNA was isolated from the urine of 148 men with prostate cancer and assessed for *TMPRSS2:ERG* and *SPINK1* expression by quantitative PCR. Samples above the dashed red line show *SPINK1* outlier expression (see Experimental Procedures).

Stenman, 2002). It is notable that SPINK1 is also overexpressed in other cancers, and elevated serum level is an independent prognostic sign in many of these (reviewed in Paju et al., 2007; Paju and Stenman, 2006).

To investigate a functional role for *SPINK1* in prostate cancer, we generated adenoviruses expressing *SPINK1* and infected the benign immortalized prostate epithelial cell line RWPE to generate RWPE-*SPINK1* cells. Overexpression of *SPINK1* had no significant effect on the proliferation or invasion of RWPE cells (Figures 5A and 5B). As *SPINK1* overexpression had no effect on benign prostate cells, we hypothesized that *SPINK1* overexpression may occur later in prostate cancer progression in the presence of coexisting genetic lesions, consistent with its association with aggressive prostate cancer.

Thus, we analyzed a panel of prostate cancer cell lines to identify an appropriate in vitro model for *SPINK1* outlier expression. We identified marked overexpression of *SPINK1* exclusively in the 22RV1 cell line (Figure 5C), consistent with previous work reporting high expression in this cell line (Paju et al., 2007). The aggressive 22RV1 prostate cancer cell line was derived from a human prostate carcinoma xenograft that was serially propagated in nude mice after castration-induced regression and relapse of the parental, androgen-dependent CWR22 xenograft (Sramkoski et al., 1999). Importantly, 22RV1 does not overexpress *ERG* or *ETV1* (Figure 5C), similar to clinical *SPINK1* outlier cases, supporting its use as a cell line model of *SPINK1* outlier expression. To assess the function of *SPINK1* in 22RV1, we utilized siRNA knockdown. While *SPINK1* knockdown had no effect on 22RV1 proliferation (Figure 5D), *SPINK1* knockdown markedly attenuated the invasiveness of 22RV1 cells through a modified basement membrane (Figures 5E and 5F). Similar results were obtained with two additional siRNA duplexes targeting *SPINK1* (Figure S4).

Consistent with the mutually exclusive overexpression of *ERG*, *ETV1*, and *SPINK1*, siRNA knockdown of *ERG* or *ETV1* in 22RV1 had no effect on invasion, while *SPINK1* knockdown had no effect on the invasiveness of VCaP (*TMPRSS2:ERG+*, *SPINK1−*) or LNCaP (*ETV1* rearrangement+, *SPINK1−*) (Figures 5G and 5H). Importantly, siRNA knockdown of *ERG* in VCaP and *ETV1* in LNCaP similarly attenuated invasion (Figures 5G and 5H)
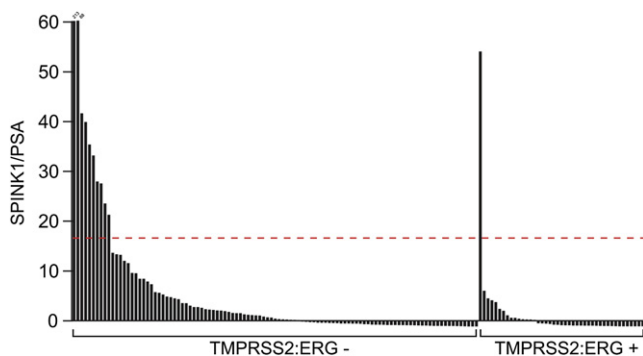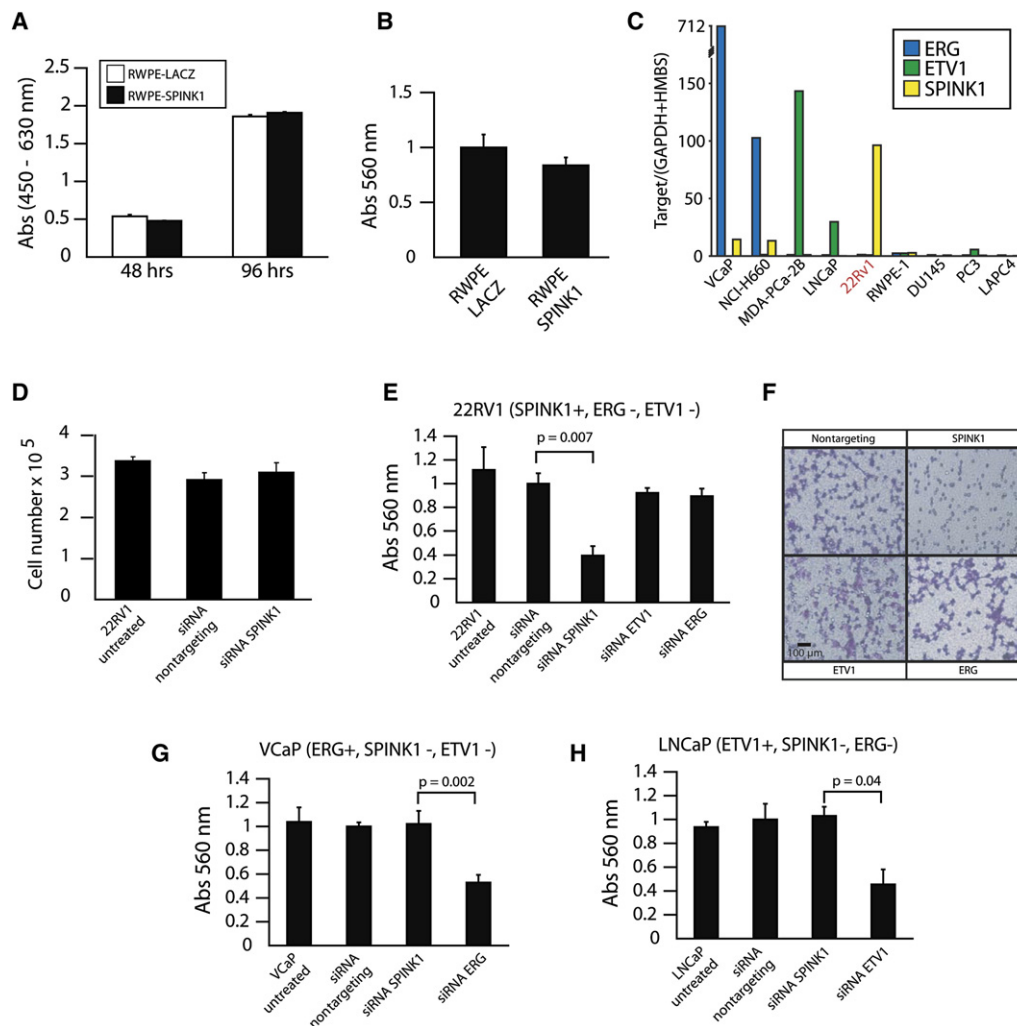
Figure 5. Knockdown of *SPINK1* in 22RV1 Prostate Cancer Cells Attenuates Invasiveness
(A and B) To recapitulate the outlier expression of *SPINK1*, we generated adenoviruses expressing *SPINK1* or *LACZ* (control). The benign immortalized prostate cell line RWPE was infected with *SPINK1* or *LACZ* adenovirus as indicated and assayed for proliferation (A) or invasion (B) through a modified basement membrane.
(C) As these results suggest that *SPINK1* may require coexisting genetic lesions to function in prostate cancer, we assayed prostate cancer cell lines by quantitative PCR for *SPINK1* (yellow), *ERG* (blue), and *ETV1* (green) outlier expression.
(D–F) *SPINK1* mediates invasiveness in 22RV1 cells. To investigate the role of *SPINK1* in the outlier-expressing cell line 22RV1, cells were treated with transfection reagent alone (untreated) or transfected with nontargeting or siRNA against *SPINK1*, *ETV1*, or *ERG* as indicated. Cells were assayed for proliferation (D) and invasion (E). Photomicrographs of invaded cells treated with the indicated siRNAs are shown in (F).
(G and H) VCaP (*TMPRSS2:ERG*-positive) (G) and LNCaP (*ETV1* rearrangement-positive) (H) prostate cancer cell lines were treated with transfection reagent alone (untreated) or transfected with nontargeting or siRNA against *SPINK1*, *ETV1*, or *ERG* as indicated and assayed for invasion.
For all proliferation and invasion experiments, means (n = 3) + SEM are shown, and p values < 0.05 are given.

without affecting proliferation (Tomlins et al., 2007a, 2008). Additionally, microarray analysis of 22RV1-si*SPINK1* cells revealed only limited transcriptional effects (76 features overexpressed, 14 features underexpressed; Table S5 and Figure S5), suggesting that *SPINK1* knockdown directly affects cellular invasiveness. Together, these results support a role for *SPINK1* in prostate cancer invasion, consistent with its overexpression in aggressive prostate cancers.

The outlier expression of *SPINK1* in a subset of prostate cancers suggested that *SPINK1* expression may be activated by a unique molecular event, similar to *TMPRSS2:ETS*-positive

prostate cancers. However, FISH studies using locus/control and 5′/3′ split probes demonstrated no evidence of amplification or gross rearrangements, respectively, in samples with *SPINK1* overexpression (data not shown). Additionally, sequencing of the *SPINK1* coding region identified no mutations in samples with *SPINK1* outlier expression (data not shown). Thus, *SPINK1* may be activated by increased transcription, possibly through promoter mutations affecting regulatory elements. Alternatively, *SPINK1* may be activated by a unique upstream genetic event. However, few genes show consistent correlation with *SPINK1* across data sets (Figure S6), suggesting that *SPINK1* would be

an exclusive downstream target. It is also possible that *SPINK1* may be downregulated in *TMPRSS2:ETS*-positive cancers; however, we would expect high *SPINK1* expression in benign prostatic epithelium, and the in vitro data described above support a role for *SPINK1* overexpression in prostate cancer progression.

Future studies will be directed at determining the mechanism by which *SPINK1* is overexpressed in *TMPRSS2:ETS*-negative prostate cancers and whether determination of SPINK1 in serum is of diagnostic and prognostic use.

Although conflicting reports of *TMPRSS2:ETS* fusion status and aggressiveness have been reported, recent large-cohort studies have shown that *TMPRSS2:ERG* fusion-positive prostate cancers harboring an intrachromosomal deletion between the *TMPRSS2* and *ERG* loci on chromosome 21 are associated with aggressiveness (Attard et al., 2008; Demichelis et al., 2007; Lapointe et al., 2007; Mehra et al., 2007; Nam et al., 2007a, 2007b; Perner et al., 2006; Rajput et al., 2007; Wang et al., 2006; Winnes et al., 2007; Yoshimoto et al., 2006). Supporting this hypothesis, in a cohort of patients undergoing rapid autopsy after death from hormone-refractory metastatic prostate cancer, we found that all *TMPRSS2:ERG* fusion-positive patients were deletion positive (R.M. et al., unpublished data). These "deletion-positive" *TMPRSS2:ERG*-positive cases, representing ~25% of all prostate cancers, likely account for the association of *TMPRSS2:ERG* positivity with aggressiveness. In this report, we identify *SPINK1*-positive samples as defining an aggressive subset of *TMPRSS2:ETS*-negative prostate cancers (~10% of all prostate cancers). Future studies will be needed to identify the molecular mechanisms, including response or resistance to current therapies, that drive the aggressiveness of *TMPRSS2:ERG* deletion-positive and *SPINK1*-positive prostate cancers.

In conclusion, using a combination of in silico bioinformatics analysis coupled with independent experimental validation, we analyzed data on ~1800 prostate cancers, demonstrating the consistent outlier expression of *SPINK1* in *TMPRSS2:ETS*-negative prostate cancers (Table S4). We provide evidence that *SPINK1* outlier expression defines an aggressive molecular subtype of prostate cancer (~10% of cases) not attributable to known gene fusion events. We hypothesize that the molecular lesion or lesions that initially drive *ETS*-negative tumors, which are presently unclear, may predispose to activation of *SPINK1* expression later in prostate cancer progression. Additionally, *SPINK1*-positive tumors may arise from a different prostate progenitor cell type than *ETS*-positive tumors, and *SPINK1* expression may be a marker of this cell type. We demonstrate that *SPINK1* may be monitored noninvasively in urine and thus could serve to complement gene-fusion-based urine testing for prostate cancer. Additionally, we demonstrate the utility of 22RV1 as a cell line model for *SPINK1* outlier expression. Finally, we extend the utility of our original COPA approach by using a meta-COPA strategy to nominate candidate oncogenes in specific cancer types.

EXPERIMENTAL PROCEDURES

Cancer Outlier Profile Analysis and Outlier Analysis
Cancer Outlier Profile Analysis (COPA) analysis was performed on seven prostate cancer gene expression data sets (Dhanasekaran et al., 2001; Glinsky et al., 2004; Lapointe et al., 2004; LaTulippe et al., 2002; Vanaja et al., 2003; Welsh et al., 2001; Yu et al., 2004) in Oncomine 3.0 (http://www.oncomine.org) as described previously (Tomlins et al., 2005). (1) For each data set considering all samples, gene expression values are median centered per gene, setting each gene's median expression value to 0. (2) The median absolute deviation (MAD) is calculated per gene and scaled to 1 by dividing each gene expression value by its MAD. Of note, median and MAD are used for transformation as opposed to mean and standard deviation so that outlier expression values do not unduly influence the distribution estimates and are thus preserved postnormalization. (3) For each gene in each data set, COPA scores are computed as the 75th, 90th, and 95th percentile of ascending transformed gene expression values. Thus, each gene in each data set has three COPA scores, one at each percentile cutoff, representing the degree of overexpression in decreasing subsets of cases. (4) In each data set, all genes are rank ordered by the three COPA scores, generating three rank-ordered lists of genes per data set. (5) For each data set, we defined outlier genes as those that ranked in the top 100 COPA scores in any one of the three rank-ordered lists. (6) To identify "meta-outlier" genes, we ranked genes by the number of data sets in which the gene was identified as an outlier gene. Genes identified as outliers in the same number of studies were further ranked by their average outlier rank across those studies. This process is summarized in Figure S7. Data sets can be accessed in Oncomine by searching for "[author last name]_prostate" (e.g., "Yu_prostate").

SPINK1 expression was also interrogated in prostate cancer specimens from two multicancer profiling studies (Su et al., 2001 and the International Genomics Consortium's expO data set GSE2109) and the Yang et al. "Gene expression data from prostate cancer samples" data set (NCBI GEO data set GSE8218). The two multicancer studies were not included in the meta-analysis, as prostate cancer samples comprised a minority of the profiled samples, and GSE8218 was not available at the time the meta-analysis was performed.

Individual samples showing outlier expression in each data set were identified by a two-step process that recreates the visual process of identifying the natural "gap" between nonoutlier and outlier sample populations. First, Oncomine-generated gene expression values (*ERG*, *ETV1*, and *SPINK1*) for all prostate samples in each data set (non-COPA transformed, excluding metastatic prostate cancer) were median centered. Next, for each gene, all samples were rank ordered in ascending order, and the difference between each rank-ordered sample and the preceding sample was calculated. In each data set, *ERG* showed two distributions of expression separated by a natural gap in expression levels. This visual gap for each data set was quantified after ordering the samples as just described and ranged from 0.22 to 1.0 (median 0.63) normalized expression units. This same method was then applied to define *ETV1* outlier expression, with the natural gap for *ETV1* populations ranging from 0.25 to 2.1 (median 0.48), except for the GSE2109 study, which showed no *ETV1* outlier population. *SPINK1* populations showed a similar distribution in all data sets, with the natural gap ranging from 0.27 to 1.3 (median 0.41). Hence, formally described, the first sample with a positive median-centered value and a difference of >0.22 normalized expression units compared to the preceding sample marked the transition to the outlier population for all genes in each data set (Figure S8). Specific reporters used and the number of *SPINK1*, *ERG*, and *ETV1* outliers for each data set are shown in Table S2. Outlier expression in quantitative PCR (qPCR) samples (tissue and urine) was determined similarly, except that normalized expression values for each target gene were log transformed before median centering and rank ordering. Metastatic prostate cancer samples were also included in the qPCR tissue cohort.

Samples
Tissues used for qPCR were from the radical prostatectomy series at the University of Michigan and from the Rapid Autopsy Program, both of which are part of the University of Michigan Prostate Cancer Specialized Program of Research Excellence (SPORE) Tissue Core. For combined fluorescence in situ hybridization (FISH) and immunohistochemistry (IHC) evaluation, the University of Michigan (UM) cohort consisted of samples from the radical prostatectomy series. The Swedish Watchful Waiting (SWW) cohort consisted of samples from a Swedish population-based cohort of men with localized prostate cancer diagnosed incidentally by transurethral resection of the prostate for symptomatic benign prostatic hyperplasia as described previously (Andren et al., 2006; Johansson et al., 2004). The Memorial Sloan-Kettering Cancer

Center (MSKCC) cohort consisted of patients with localized or locally advanced prostate cancer treated by radical prostatectomy at MSKCC between 1985 and 2003. All samples were obtained with institutional review board approval from the respective institutions (UM, MSKCC, or Örebro Medical Center for the SWW cohort). The prostate cancer cell line 22RV1 was provided by Jill Macoska (University of Michigan).

Quantitative PCR from Tissue Samples
qPCR was performed using SYBR green dye on an Applied Biosystems 7300 Real-Time PCR system (Applied Biosystems, Foster City, CA, USA) essentially as described previously (Tomlins et al., 2005, 2006). Briefly, total RNA was isolated from tissues using TRIzol (Invitrogen, Carlsbad, CA, USA). RNA was quantified using a ND-1000 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) and 3–5 μg of total RNA was reverse transcribed into cDNA using SuperScript III (Invitrogen) in the presence of random primers. All qPCR reactions were performed with Power SYBR Green Master Mix (Applied Biosystems) and 25 ng of both the forward primer and the reverse primer using the manufacturer's recommended thermocycling conditions. For each experiment, threshold levels were set during the exponential phase of the qPCR reaction using Sequence Detection Software version 1.2.2 (Applied Biosystems). The amount of *ERG*, *ETV1*, and *SPINK1* relative to the average of the housekeeping genes *GAPDH* and *HMBS* for each sample was determined using the comparative threshold cycle (C_t) method (according to Applied Biosystems User Bulletin #2, http://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_040980.pdf). All oligonucleotide primers were synthesized by Integrated DNA Technologies (Coralville, IA, USA). *GAPDH*, *HMBS*, *ERG* (exon5_6), and *ETV1* (exon6_7) primers were as described previously (Tomlins et al., 2005). Sequences for *SPINK1* are as follows: *SPINK1*_f, 5′-CAAAAATCTGGGCCTTGCTGAGAAC-3′; *SPINK1*_r, 5′-AGGCCTCGCGGTGACCTGAT-3′. Approximately equal efficiencies of the primers were confirmed using serial dilutions of pooled prostate cancer cDNA in order to use the comparative C_t method. All reactions were subjected to melt-curve analysis.

Immunohistochemistry and Fluorescence In Situ Hybridization
IHC for the UM and SWW cohorts was performed using a mouse monoclonal antibody against SPINK1 (H00006690-M01; Abnova, Taipei City, Taiwan) on tissue microarrays (TMAs) containing cores from 75 (UM) and 312 (SWW) evaluable cases of localized prostate cancer. Cases with staining in any cancerous epithelial cells were deemed positive (median 40%, range 1%–90%). Previously, we have evaluated cases on these tissue microarrays for *TMPRSS2:ERG* fusion status by FISH using break-apart *ERG* assays as described (Demichelis et al., 2007; Mehra et al., 2007; Tomlins et al., 2005). A one-sided Fisher's exact test was used to evaluate the relationship between SPINK1 and fusion status, as these studies were performed with the prior hypothesis that there was an inverse correlation between SPINK1 expression and fusion status.

MSKCC Immunohistochemistry
IHC for the MSKCC cohort was performed using an in-house mouse monoclonal antibody against SPINK1 (code 6E8; Osman et al., 1993) on tissue microarrays containing triplicate cores from 817 evaluable cases of localized prostate cancer. The percentage of tumor cells in each core was estimated and assigned values of 0%, 5%, or multiples of 10%. The intensity of the expression was assigned a value of 0, 1, 2, or 3. Triplicate cores from each specimen were scored separately, and the presence of tumorous tissue in at least two interpretable cores was required to include a case for analysis. We considered cases as SPINK1 positive if any of the three cores exhibited >80% of cancerous cells showing positive SPINK1 immunoreactivity (intensity 1–3).

Outcome Analyses
For Kaplan-Meier analysis of the Glinsky et al. (2004) and UM data sets, biochemical recurrence was defined as a 0.2 ng/ml increase in PSA or recurrence of disease after prostatectomy, such as development of metastatic cancer, if biochemical recurrence information was not available. For the MSKCC cohort, only biochemical recurrence, defined as PSA > 0.2 ng/ml after surgical resection with a second confirmatory PSA measurement > 0.2 ng/ml, was considered, as all patients with a clinical failure had previously had a biochemical recurrence. For outcome analysis from the Glinsky et al. (2004) data set,

samples positive for outlier expression of *SPINK1* were defined as described above. For the IHC analysis of the UM and MSKCC cohorts, positive cases were defined as described above. Kaplan-Meier analysis and multivariate Cox proportional-hazards regression were then used to examine the association of SPINK1 with biochemical PSA recurrence. To predict the probability of disease recurrence, we used the Kattan 7-year postoperative nomogram (Kattan et al., 1999), and the concordance index of the nomogram and the nomogram plus SPINK1 status was evaluated using 1000-times bootstrapping as described (Kattan et al., 2003).

Urine-Based Detection of *SPINK1* Expression
Collection of urine, isolation of RNA, RNA amplification, and qPCR for *TMPRSS2:ERG* from men with prostate cancer was performed as described previously (Laxman et al., 2006, 2008). Briefly, 25 ng of isolated RNA was amplified using a TransPlex Whole Transcriptome Amplification (WTA) kit (Rubicon Genomics, Ann Arbor, MI, USA) according to the manufacturer's instructions. For each qPCR reaction, 10 ng of WTA-amplified cDNA was used as template. 2× Power SYBR Green Master Mix (Applied Biosystems) and 25 ng of both the forward and reverse primers were used for *SPINK1*, *ERG* (primers as described above), and *PSA* (Laxman et al., 2006). For all experiments, the same threshold and baseline were set using Sequence Detection Software version 1.2.2 (Applied Biosystems). All samples with a C_t value greater than 26 for *PSA* were excluded to remove samples with insufficient prostate cell recovery. Samples were considered *TMPRSS2:ERG* positive if both *ERG* and *TMPRSS2:ERG* assays showed C_t values less than 37. The amount of *SPINK1* relative to *PSA* was determined for each sample using the comparative C_t method. Outlier samples were identified as described above. One-sided Fisher's exact test and Mann-Whitney U tests were used to evaluate the relationship between *SPINK1* and *TMPRSS2:ERG* status, as this study was performed with the prior hypothesis that there was an inverse correlation between *SPINK1* expression and fusion status.

In Vitro Overexpression of *SPINK1*
cDNA of *SPINK1* (NM_003122.2), as present in a clinical prostate cancer specimen overexpressing *SPINK1*, was amplified by RT-PCR using the following primers, with the forward primer including a consensus Kozak sequence (start and stop codons underlined): *SPINK1*_full-f, 5′-ACCACC<u>ATG</u>AAGGTAACAGGCATCTTTCTT-3′; *SPINK1*_full-r, 5′-<u>TCA</u>GCAAGGCCCAGATTTTTGA-3′. The cDNA product was TOPO cloned into the Gateway entry vector pCR8/GW/TOPO (Invitrogen), yielding pCR8-*SPINK1*. To generate adenoviral constructs, pCR8-*SPINK1* was recombind with pAD/CMV/V5 (Invitrogen) using LR Clonase II (Invitrogen). Control pAD/CMV/*LACZ* clones were obtained from Invitrogen. Adenoviruses were generated by the University of Michigan Vector Core. The benign immortalized prostate cell line RWPE was infected with *SPINK1* or *LACZ* adenoviruses, generating RWPE-*SPINK1* and RWPE-*LACZ* for transient overexpression.

Proliferation Assay
Proliferation for RWPE-*LACZ* and RWPE-*SPINK1* cells was measured by a colorimetric assay based on the cleavage of the tetrazolium salt WST-1 by mitochondrial dehydrogenases (cell proliferation reagent WST1; Roche Diagnostics, Mannheim, Germany) at the indicated time points in triplicate. Cell counts for 22RV1 cells were estimated by trypsinizing cells and analysis by Coulter counter (Beckman Coulter, Fullerton, CA, USA) at 72 hr in triplicate.

Invasion Assays
For invasion assays, RWPE-*SPINK1* and RWPE-*LACZ* cells (48 hr after infection with adenoviruses) or 22RV1 cells were used. Equal numbers of the indicated cells were seeded onto the basement membrane matrix (EC matrix; Chemicon, Temecula, CA, USA) present in the insert of a 24-well culture plate, with fetal bovine serum added to the lower chamber as a chemoattractant. After 48 hr, noninvading cells and EC matrix were removed using a cotton swab. Invaded cells were stained with crystal violet and photographed. The inserts were treated with 10% acetic acid, and absorbance was measured at 560 nm.

SPINK1 Knockdown
For siRNA knockdown of *SPINK1* in 22RV1 cells, the individual siRNAs composing the Dharmacon SMARTpool against *SPINK1* (LQ-019724-00; Chicago)

were tested for *SPINK1* knockdown by qPCR, and the most effective single siRNA (J-019724-07) was used for further experiments. siCONTROL Non-Targeting siRNA #1 (D-001210-01) or siRNA against *SPINK1* was transfected into 22RV1 cells using Oligofectamine (Invitrogen). After 24 hr, we carried out a second identical transfection, and cells were harvested 24 hr later for RNA isolation, invasion assays, or proliferation assays as described above. Invasion experiments using two other siRNAs directed against *SPINK1* (J-019724-05 and J-019724-06; SPINK1-b and -c, respectively) were also performed (Figure S4).

Expression Profiling

Expression profiling was performed using the Agilent Whole Human Genome Oligo Microarray (Santa Clara, CA, USA). Total RNA isolated using TRIzol was purified using the QIAGEN RNeasy Micro Kit (Valencia, CA, USA). One microgram of total RNA was converted to cRNA and labeled according to the manufacturer's protocol (Agilent). Hybridizations were performed for 16 hr at 65°C, and arrays were scanned on an Agilent DNA microarray scanner. Images were analyzed and data extracted using Agilent Feature Extraction Software 9.1.3.1, with linear and lowess normalization performed for each array. For 22RV1-si*SPINK1* hybridizations, the reference was 22RV1 cells infected with nontargeting siRNA. Duplicate hybridizations were performed with duplicate dye flips, for a total of four arrays. Over- and underexpressed signatures were generated by filtering to include only features with significant differential expression (pValueLogRatio < 0.01) in all hybridizations and Cy5/Cy3 ratios (LogRatio) greater than or less than 1 (unlogged) in all hybridizations, after correction for the dye flip.

ACCESSION NUMBERS

The 22RV1 expression profiling data are available at the NCBI GEO (http://www.ncbi.nlm.nih.gov/geo/) under the accession number GSE11132.

SUPPLEMENTAL DATA

The Supplemental Data include five tables and eight figures and can be found with this article online at http://www.cancercell.org/cgi/content/full/13/6/519/DC1/.

REFERENCES

Andren, O., Fall, K., Franzen, L., Andersson, S.O., Johansson, J.E., and Rubin, M.A. (2006). How well does the Gleason score predict prostate cancer death? A 20-year followup of a population based cohort in Sweden. J. Urol. *175*, 1337–1340.

Annunziata, C.M., Davis, R.E., Demchenko, Y., Bellamy, W., Gabrea, A., Zhan, F., Lenz, G., Hanamura, I., Wright, G., Xiao, W., et al. (2007). Frequent engagement of the classical and alternative NF-kappaB pathways by diverse genetic abnormalities in multiple myeloma. Cancer Cell *12*, 115–130.

Attard, G., Clark, J., Ambroisine, L., Fisher, G., Kovacs, G., Flohr, P., Berney, D., Foster, C.S., Fletcher, A., Gerald, W.L., et al. (2008). Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. Oncogene *27*, 253–263.

Bjartell, A., Paju, A., Zhang, W.M., Gadaleanu, V., Hansson, J., Landberg, G., and Stenman, U.H. (2005). Expression of tumor-associated trypsinogens (TAT-1 and TAT-2) in prostate cancer. Prostate *64*, 29–39.

Demichelis, F., Fall, K., Perner, S., Andren, O., Schmidt, F., Setlur, S.R., Hoshida, Y., Mosquera, J.M., Pawitan, Y., Lee, C., et al. (2007). TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. Oncogene *26*, 4596–4599.

Dhanasekaran, S.M., Barrette, T.R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K.J., Rubin, M.A., and Chinnaiyan, A.M. (2001). Delineation of prognostic biomarkers in prostate cancer. Nature *412*, 822–826.

Freeman, T.C., Curry, B.J., Calam, J., and Woodburn, J.R. (1990). Pancreatic secretory trypsin inhibitor stimulates the growth of rat pancreatic carcinoma cells. Gastroenterology *99*, 1414–1420.

Glinsky, G.V., Glinskii, A.B., Stephenson, A.J., Hoffman, R.M., and Gerald, W.L. (2004). Gene expression profiling predicts clinical outcome of prostate cancer. J. Clin. Invest. *113*, 913–923.

Greene, L.J., Pubols, M.H., and Bartelt, D.C. (1976). Human pancreatic secretory trypsin inhibitor. Methods Enzymol. *45*, 813–825.

Han, M., Partin, A.W., Pound, C.R., Epstein, J.I., and Walsh, P.C. (2001). Long-term biochemical disease-free and cancer-specific survival following anatomic radical retropubic prostatectomy. The 15-year Johns Hopkins experience. Urol. Clin. North Am. *28*, 555–565.

Haverback, B.J., Dyce, B., Bundy, H., and Edmondson, H.A. (1960). Trypsin, trypsinogen and trypsin inhibitor in human pancreatic juice. Am. J. Med. *29*, 421–433.

Helgeson, B.E., Tomlins, S.A., Shah, N., Laxman, B., Cao, Q., Prensner, J.R., Cao, X., Singla, N., Montie, J.E., Varambally, S., et al. (2008). Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. Cancer Res. *68*, 73–80.

Hull, G.W., Rabbani, F., Abbas, F., Wheeler, T.M., Kattan, M.W., and Scardino, P.T. (2002). Cancer control with radical prostatectomy alone in 1,000 consecutive patients. J. Urol. *167*, 528–534.

Johansson, J.E., Andren, O., Andersson, S.O., Dickman, P.W., Holmberg, L., Magnuson, A., and Adami, H.O. (2004). Natural history of early, localized prostate cancer. JAMA *291*, 2713–2719.

Kattan, M.W., Wheeler, T.M., and Scardino, P.T. (1999). Postoperative nomogram for disease recurrence after radical prostatectomy for prostate cancer. J. Clin. Oncol. *17*, 1499–1507.

Kattan, M.W., Shariat, S.F., Andrews, B., Zhu, K., Canto, E., Matsumoto, K., Muramoto, M., Scardino, P.T., Ohori, M., Wheeler, T.M., and Slawin, K.M. (2003). The addition of interleukin-6 soluble receptor and transforming growth factor beta1 improves a preoperative nomogram for predicting biochemical progression in patients with clinically localized prostate cancer. J. Clin. Oncol. *21*, 3573–3579.

Kazal, L.A., Spicer, D.S., and Brahinsky, R.A. (1948). Isolation of a crystalline trypsin inhibitor-anticoagulant protein from pancreas. J. Am. Chem. Soc. *70*, 3034–3040.

Lapointe, J., Li, C., Higgins, J.P., van de Rijn, M., Bair, E., Montgomery, K., Ferrari, M., Egevad, L., Rayford, W., Bergerheim, U., et al. (2004). Gene

expression profiling identifies clinically relevant subtypes of prostate cancer. Proc. Natl. Acad. Sci. USA *101*, 811–816.

Lapointe, J., Kim, Y.H., Miller, M.A., Li, C., Kaygusuz, G., van de Rijn, M., Huntsman, D.G., Brooks, J.D., and Pollack, J.R. (2007). A variant TMPRSS2 isoform and ERG fusion product in prostate cancer with implications for molecular diagnosis. Mod. Pathol. *20*, 467–473.

LaTulippe, E., Satagopan, J., Smith, A., Scher, H., Scardino, P., Reuter, V., and Gerald, W.L. (2002). Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastatic disease. Cancer Res. *62*, 4499–4506.

Laxman, B., Tomlins, S.A., Mehra, R., Morris, D.S., Wang, L., Helgeson, B.E., Shah, R.B., Rubin, M.A., Wei, J.T., and Chinnaiyan, A.M. (2006). Noninvasive detection of TMPRSS2:ERG fusion transcripts in the urine of men with prostate cancer. Neoplasia *8*, 885–888.

Laxman, B., Morris, D.S., Yu, J., Siddiqui, J., Cao, J., Mehra, R., Lonigro, R.J., Tsodikov, A., Wei, J.T., Tomlins, S.A., and Chinnaiyan, A.M. (2008). A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res. *68*, 645–649.

Mehra, R., Tomlins, S.A., Shen, R., Nadeem, O., Wang, L., Wei, J.T., Pienta, K.J., Ghosh, D., Rubin, M.A., Chinnaiyan, A.M., and Shah, R.B. (2007). Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. Mod. Pathol. *20*, 538–544.

Naderi, A., Teschendorff, A.E., Beigel, J., Cariati, M., Ellis, I.O., Brenton, J.D., and Caldas, C. (2007). BEX2 is overexpressed in a subset of primary breast cancers and mediates nerve growth factor/nuclear factor-kappaB inhibition of apoptosis in breast cancer cell lines. Cancer Res. *67*, 6725–6736.

Nam, R.K., Sugar, L., Wang, Z., Yang, W., Kitching, R., Klotz, L.H., Venkateswaran, V., Narod, S.A., and Seth, A. (2007a). Expression of TMPRSS2:ERG gene fusion in prostate cancer cells is an important prognostic factor for cancer progression. Cancer Biol. Ther. *6*, 40–45.

Nam, R.K., Sugar, L., Yang, W., Srivastava, S., Klotz, L.H., Yang, L.Y., Stanimirovic, A., Encioiu, E., Neill, M., Loblaw, D.A., et al. (2007b). Expression of the TMPRSS2:ERG fusion gene predicts cancer recurrence after surgery for localised prostate cancer. Br. J. Cancer *97*, 1690–1695.

Ogawa, M., Tsushima, T., Ohba, Y., Ogawa, N., Tanaka, S., Ishida, M., and Mori, T. (1985). Stimulation of DNA synthesis in human fibroblasts by human pancreatic secretory trypsin inhibitor. Res. Commun. Chem. Pathol. Pharmacol. *50*, 155–158.

Osman, S., Turpeinen, U., Itkonen, O., and Stenman, U.H. (1993). Optimization of a time-resolved immunofluorometric assay for tumor-associated trypsin inhibitor (TATI) using the streptavidin-biotin system. J. Immunol. Methods *161*, 97–106.

Paju, A., and Stenman, U.H. (2006). Biochemistry and clinical role of trypsinogens and pancreatic secretory trypsin inhibitor. Crit. Rev. Clin. Lab. Sci. *43*, 103–142.

Paju, A., Hotakainen, K., Cao, Y., Laurila, T., Gadaleanu, V., Hemminki, A., Stenman, U.H., and Bjartell, A. (2007). Increased expression of tumor-associated trypsin inhibitor, TATI, in prostate cancer and in androgen-independent 22Rv1 cells. Eur. Urol. *52*, 1670–1679.

Perner, S., Demichelis, F., Beroukhim, R., Schmidt, F.H., Mosquera, J.M., Setlur, S., Tchinda, J., Tomlins, S.A., Hofer, M.D., Pienta, K.G., et al. (2006). TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. Cancer Res. *66*, 8337–8341.

Rajput, A.B., Miller, M.A., De Luca, A., Boyd, N., Leung, S., Hurtado-Coll, A., Fazli, L., Jones, E.C., Palmer, J.B., Gleave, M.E., et al. (2007). Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. J. Clin. Pathol. *60*, 1238–1243.

Rhodes, D.R., Yu, J., Shanker, K., Deshpande, N., Varambally, R., Ghosh, D., Barrette, T., Pandey, A., and Chinnaiyan, A.M. (2004). ONCOMINE: a cancer microarray database and integrated data-mining platform. Neoplasia *6*, 1–6.

Sramkoski, R.M., Pretlow, T.G., 2nd, Giaconia, J.M., Pretlow, T.P., Schwartz, S., Sy, M.S., Marengo, S.R., Rhim, J.S., Zhang, D., and Jacobberger, J.W. (1999). A new human prostate carcinoma cell line, 22Rv1. In Vitro Cell. Dev. Biol. Anim. *35*, 403–409.

Stenman, U.H. (2002). Tumor-associated trypsin inhibitor. Clin. Chem. *48*, 1206–1209.

Su, A.I., Welsh, J.B., Sapinoso, L.M., Kern, S.G., Dimitrov, P., Lapp, H., Schultz, P.G., Powell, S.M., Moskaluk, C.A., Frierson, H.F., Jr., and Hampton, G.M. (2001). Molecular classification of human carcinomas by use of gene expression signatures. Cancer Res. *61*, 7388–7393.

Tomlins, S.A., Rhodes, D.R., Perner, S., Dhanasekaran, S.M., Mehra, R., Sun, X.W., Varambally, S., Cao, X., Tchinda, J., Kuefer, R., et al. (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science *310*, 644–648.

Tomlins, S.A., Mehra, R., Rhodes, D.R., Smith, L.R., Roulston, D., Helgeson, B.E., Cao, X., Wei, J.T., Rubin, M.A., Shah, R.B., and Chinnaiyan, A.M. (2006). TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res. *66*, 3396–3400.

Tomlins, S.A., Laxman, B., Dhanasekaran, S.M., Helgeson, B.E., Cao, X., Morris, D.S., Menon, A., Jing, X., Cao, Q., Han, B., et al. (2007a). Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature *448*, 595–599.

Tomlins, S.A., Mehra, R., Rhodes, D.R., Cao, X., Wang, L., Dhanasekaran, S.M., Kalyana-Sundaram, S., Wei, J.T., Rubin, M.A., Pienta, K.J., et al. (2007b). Integrative molecular concept modeling of prostate cancer progression. Nat. Genet. *39*, 41–51.

Tomlins, S.A., Laxman, B., Varambally, S., Cao, X., Yu, J., Helgeson, B.E., Cao, Q., Prensner, J.R., Rubin, M.A., Shah, R.B., et al. (2008). The role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia *10*, 177–188.

Vanaja, D.K., Cheville, J.C., Iturria, S.J., and Young, C.Y. (2003). Transcriptional silencing of zinc finger protein 185 identified by expression profiling is associated with prostate cancer progression. Cancer Res. *63*, 3877–3882.

Wang, J., Cai, Y., Ren, C., and Ittmann, M. (2006). Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. Cancer Res. *66*, 8347–8351.

Welsh, J.B., Sapinoso, L.M., Su, A.I., Kern, S.G., Wang-Rodriguez, J., Moskaluk, C.A., Frierson, H.F., Jr., and Hampton, G.M. (2001). Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. Cancer Res. *61*, 5974–5978.

Winnes, M., Lissbrant, E., Damber, J.E., and Stenman, G. (2007). Molecular genetic analyses of the TMPRSS2-ERG and TMPRSS2-ETV1 gene fusions in 50 cases of prostate cancer. Oncol. Rep. *17*, 1033–1036.

Yoshimoto, M., Joshua, A.M., Chilton-Macneill, S., Bayani, J., Selvarajah, S., Evans, A.J., Zielenska, M., and Squire, J.A. (2006). Three-color FISH analysis of TMPRSS2/ERG fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. Neoplasia *8*, 465–469.

Yu, Y.P., Landsittel, D., Jing, L., Nelson, J., Ren, B., Liu, L., McDonald, C., Thomas, R., Dhir, R., Finkelstein, S., et al. (2004). Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. J. Clin. Oncol. *22*, 2790–2799.