The Landscape of Prognostic Outlier Genes in High-Risk Prostate Cancer

Shuang G. Zhao¹,², Joseph R. Evans¹, Vishal Kothari¹, Grace Sun¹, Ashley Larm¹, Victor Mondine¹, Edward M. Schaeffer⁵, Ashley E. Ross³, Eric A. Klein⁴, Robert B. Den⁵, Adam P. Dicker⁵, R. Jeffrey Karnaš⁶, Nicholas Erho⁷, Paul L. Nguyen⁸, Elai Davicioni⁹, and Felix Y. Feng¹,⁹,¹⁰

Abstract

**Purpose:** There is a clear need to improve risk stratification and to identify novel therapeutic targets in aggressive prostate cancer. The goal of this study was to investigate genes with outlier expression with prognostic association in high-risk prostate cancer patients as potential biomarkers and drug targets.

**Experimental Design:** We interrogated microarray gene expression data from prostatectomy samples from 545 high-risk prostate cancer patients with long-term follow-up (mean 13.4 years). Three independent clinical datasets totaling an additional 545 patients were used for validation. Novel prognostic outlier genes were interrogated for impact on oncogenic phenotypes in vitro using siRNA-based knockdown. Association with clinical outcomes and comparison with existing prognostic instruments was assessed with multivariable models using a prognostic outlier score.

**Results:** Analysis of the discovery cohort identified 20 prognostic outlier genes. Three top prognostic outlier genes were novel prostate cancer genes; NVL, SMC4, or SQLE knockdown reduced migration and/or invasion and outlier expression was significantly associated with poor prognosis. Increased prognostic outlier score was significantly associated with poor prognosis independent of standard clinicopathologic variables. Finally, the prognostic outlier score prognostic association is independent of, and adds to existing genomic and clinical tools for prognostication in prostate cancer (Decipher, the cell-cycle progression signature, and CAPRA-S).

**Conclusions:** To our knowledge, this study represents the first unbiased high-throughput investigation of prognostic outlier genes in prostate cancer and demonstrates the potential biomarker and therapeutic importance of this previously unstudied class of cancer genes. Clin Cancer Res; 1–10. ©2015 AACR.

Introduction

While a majority of prostate cancer patients will be cured of disease with primary treatment, high-risk prostate cancer represents a subset of patients which still carry a significant risk of metastatic progression (1–3). High-risk prostate cancer is local-ized by definition and thus potentially curable, but cure rates are only 50% to 60%. Of the patients who experience recurrence, 30% to 40% progress to metastatic disease that ultimately leads to death (4, 5). Therefore, there is a critical need to identify which patients are at greatest risk for metastatic progression, and to identify biologic drivers that can be targeted to improve outcomes for these patients (1). Current clinical prognostic paradigms, based on PSA level, Gleason score, tumor stage, and prostatectomy pathology provide guidance but are imperfect (6, 7). Molecular biomarkers are a promising strategy to improve stratification of prostate cancer patients at high risk of metastatic progression from those at reduced or standard risk.

Numerous prognostic molecular biomarkers to risk-stratify prostate cancer patients have been selected manually based on a known role in molecular pathways important in prostate cancer (8–10). Genomics technology has revolutionized the search for prognostic biomarkers with unbiased techniques based on differential expression, leading to the nomination of numerous single-gene biomarkers (11–14) and molecular subtypes (15). Prognostic studies in prostate cancer are hampered by the long disease course, resulting in follow-up of insufficient length to study associations with the most meaningful outcomes of metastatic progression, prostate cancer death, and overall survival. Thus, many studies have been limited to the intermediate outcome of biochemical recurrence (16).

Recently, several high-risk prostate cancer patient cohorts with prostatectomy sample high-density microarray gene expression
data have accrued follow-up long enough for correlation with metastatic progression and survival endpoints. This invaluable resource has allowed for several very exciting developments for prostate cancer biomarkers. A 22-gene genomic classifier developed and validated in this dataset stratifies patients by metastatic progression risk 5 years after prostatectomy [22.5% for high risk (19% of patients) vs. 2.4% for low risk (60%)], and provides prognostic information independent of the clinical data-based Cancer of the Prostate Risk Assessment Post-surgical (CAPRA-S) prognostic instrument (17, 18). On the basis of the high-density of the microarrays in this dataset, our group discovered and validated a novel long noncoding RNA biomarker called SChLAP1 that correlates with aggressive prostate cancer, and may be useful as a urine diagnostic biomarker (13). These examples demonstrate the value of this large cohort with long follow-up and gene expression data to discover and validate prostate cancer prognostic biomarkers.

Outlier analysis was developed to identify genes with very high expression in a subset of samples, which is often not detected by standard differential expression analysis. This novel approach to analyzing gene expression patterns led to the discovery of ERG family gene fusions and SPINK-positive prostate cancer (19, 20). Outlier analyses have traditionally been limited to comparisons of gene expression in cancer versus normal samples, and have not been analyzed with respect to clinical outcomes. Recently, two studies analyzed outlier expression within cancer cohorts, and correlated gene outlier status in breast cancer and melanoma (21, 22). These studies indicate prognostic outlier genes to be a promising class of cancer genes for biomarker and therapeutic discovery.

Here, we applied prognostic outlier gene analysis in high-risk prostate cancer to better identify patients at risk of metastatic progression after prostatectomy, and the genes implicated in progression. We analyzed a large cohort of high-risk prostate cancer patients with long follow-up after prostatectomy and high-density microarray gene expression data from the prostatectomy specimen. Twenty prognostic outlier genes were identified. Several prognostic outlier genes unstudied in prostate cancer were analyzed for contribution to oncogenic phenotypes in vitro using RNAi in two prostate cancer cell lines and for differential overexpression in metastatic prostate cancer samples in Oncomine. Prognostic significance of outlier expression of these 20 prognostic outlier genes was validated in three independent high-risk prostate cancer prostatectomy cohorts and compared with existing clinical and molecular prognostic signatures. This is the first study of prognostic outlier genes in prostate cancer, which appears to provide a powerful platform for cancer biomarker and therapeutic target discovery.

Translational Relevance

Although the majority of localized prostate cancer patients are effectively managed with active surveillance or primary therapy, high-risk prostate cancer carries a substantial risk of metastatic progression and death. Prognostic biomarkers would allow for personalized therapy intensification in high-risk prostate cancer, but their discovery is hampered by the long disease course. We applied a biomarker approach novel to prostate cancer to identify genes with outlier expression that correlates with outcome. Using large cohorts of high-risk prostate cancer patients and microarray gene expression data with postprostatectomy follow-up sufficient for correlation with metastatic progression, we identified 20 prognostic outlier genes, and validated outlier status of these genes as prognostic alone and in combination with existing clinical and genomic tools. In addition, we show that three prognostic outlier genes are novel prostate cancer genes that promote invasion and/or migration. Prognostic outlier genes are a novel class of cancer genes for biomarker and therapeutic discovery.

Materials and Methods

Clinical cohorts and microarray processing

Four cohorts of high-risk prostate cancer patients treated with prostatectomy were assembled from three different institutions: two from Mayo Clinic (MCI and MCII), a third from Cleveland Clinic Foundation (CCF), and a fourth from Thomas Jefferson University (TJU, Philadelphia, PA). Gene expression of these samples was analyzed by microarray as previously reported (5). Informed consent protocols were approved by local Institutional Review Boards (IRB). Briefly, MCI was a nested case–control study with 545 patients, MCII was a case–cohort study of 232 patients, CCF was a case–control study of 183 patients who did not receive neoadjuvant or adjuvant therapy, and TJU was a cohort of 130 patients who received postoperative radiotherapy. RNA extraction from formalin-fixed paraffin-embedded prostatectomy samples and microarray hybridization and normalization was performed at a CLIA laboratory (GenomeDx Biosciences; refs. 5, 13, 17, 23, 24). Microarray data are available with NCBI GEO accession numbers GSE46691, GSE62116, GSE72291, and GSE62667.

Gene nomination

To nominate outliers in metastatic prostate cancer, we compared gene expression profiles in patients that metastasized versus those that did not in the MCI cohort. The Cancer Outlier Profile Analysis (COPA) algorithm (19) was used with two minor modifications. Centering was performed using the median of the nonmetastatic samples, as the median expression of genes associated with metastatic progression would shift depending on the prevalence of metastatic progression in the cohort. Second, we used a COPA cutoff of 2.5 (half of original threshold) to increase sensitivity, as there were no prognostic genes meeting the original COPA cutoff. We focused our search on genes which had outlier expression in 4% or more of the cohort, and positive predictive value of 0.7 or higher.

Clinical validation and comparison with published prognostic models

We examined prognostic outlier gene expression in the three additional validation cohorts, MCII, TJU, and CCF, using the outlier cut-off value from the MCI discovery cohort. Prognostic outlier score was defined as the number of prognostic outlier genes with outlier status. Scores were also generated using existing prognostic instruments Decipher, CAPRA-S, and cell-cycle progression (CCP) signature (18, 25–28). As the CCP signature was ported to our microarray platform, we will refer to the microarray...
version as mCCP. In the three validation cohorts, individual prognostic outlier gene status and prognostic outlier score were analyzed with a pooled multivariable logistic regression model. The independence of the prognostic outlier score and the existing prognostic signatures was modeled using logistic regression model including both tools, and prognostic scores were scaled by the SD to make them comparable. To evaluate the addition of the prognostic outlier score to existing clinical and genomic tools, a logistic regression model of the prognostic outlier score in combination with Decipher, mCCP, or CAPRA-S was trained in the discovery (MCI) cohort. Each combined model was then evaluated in the pooled validation cohort for improvement in receiver-operating curves (ROC) area under the curve (AUC).

The independence of the prognostic outlier score and the existing prognostic signatures was modeled using logistic regression models and ROC calculations used 10-year metastatic progression as the primary endpoint, and models were stratified by cohort (13). Oncomine prostate cancer datasets with n > 20 were interrogated for differential expression of the four candidate genes in metatases versus primary tumors using the standard web analysis (29).

Cell culture and invasion/migration assays
DU145 and LNCaP-AR cells were propagated in RPMI1640 base media supplemented with 10% FBS (HI, Invitrogen) and 1% penicillin/streptomycin (Invitrogen). DU145 cells were purchased from ATCC 6/2011 and LNCaP-AR cells were a kind gift from Charles Sawyers in June 2014. Upon receipt, both cell lines were confirmed authentic with the ThermoFisher Identifier Plus kit by the University of Michigan DNA Sequencing Core Facility and immediately frozen. For experiments, both cell lines were used within 6 months of thawing from those original vials. Lipofectamine RNAiMAX was used for transient transfection of siRNA pools targeting NVL, SMC4, SQLE, or TPX2 (GE/Dharmacon ON-TARGET SMARTpool). Two days later, cells were plated in serum-free media for invasion in Matrigel-coated micropore transwells (Corning Bio-Coat) and in micropore transwells without Matrigel for migration. RNA was isolated from the remaining cells to confirm knockdown by qRT-PCR. Two days after invasion/migration plating, transwells were fixed in 4% formaldehyde, stained with crystal violet, and cells on the upper side of the micropore membrane wiped off. Cells on the lower side of the membrane were photographed with a microscope, the dye was dissolved in 10% acetic acid and optical density measured at 560 nm (OD560 nm).

qRT-PCR assays
High-Capacity cDNA Reverse Transcription Kit (Life Technologies) was used to generate cDNA, which was used as template for RT-PCR using standard cycling parameters with Fast SYBR Master Mix (ABI) and gene-specific primers (sequences from Origene, purchased from IDT). ΔΔCt was used for relative quantititation analysis.

Cell proliferation and cell death assays
Cell proliferation after siRNA transfection was measured in the IncuCyte ZOOM (Essen BioScience) automated cell culture microscope system with image analysis software. Cell death was measured by flow cytometry after Annexin V-fluorescein/propidium iodide (PI) staining, as the percentage of cells positive for Annexin V or Annexin V and PI.

Table 1. High-risk prostate cancer prognostic outlier genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Patients with outlier expression</th>
<th>PPV</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVL</td>
<td>4.0%</td>
<td>0.86</td>
<td>0.99</td>
</tr>
<tr>
<td>CENPF</td>
<td>5.3%</td>
<td>0.86</td>
<td>0.99</td>
</tr>
<tr>
<td>SMC4</td>
<td>4.4%</td>
<td>0.85</td>
<td>0.99</td>
</tr>
<tr>
<td>CAMK2N1</td>
<td>8.8%</td>
<td>0.81</td>
<td>0.97</td>
</tr>
<tr>
<td>STMN1</td>
<td>4.6%</td>
<td>0.80</td>
<td>0.98</td>
</tr>
<tr>
<td>TFPX2</td>
<td>4.0%</td>
<td>0.77</td>
<td>0.98</td>
</tr>
<tr>
<td>MKB7</td>
<td>4.0%</td>
<td>0.77</td>
<td>0.98</td>
</tr>
<tr>
<td>SMC2</td>
<td>4.6%</td>
<td>0.76</td>
<td>0.98</td>
</tr>
<tr>
<td>NSMCE2</td>
<td>5.3%</td>
<td>0.76</td>
<td>0.98</td>
</tr>
<tr>
<td>HDAC9</td>
<td>7.2%</td>
<td>0.74</td>
<td>0.97</td>
</tr>
<tr>
<td>SQLE</td>
<td>5.0%</td>
<td>0.74</td>
<td>0.98</td>
</tr>
<tr>
<td>LCLAT1</td>
<td>4.2%</td>
<td>0.74</td>
<td>0.98</td>
</tr>
<tr>
<td>LPGA1</td>
<td>4.2%</td>
<td>0.74</td>
<td>0.98</td>
</tr>
<tr>
<td>TOP2A</td>
<td>8.3%</td>
<td>0.73</td>
<td>0.96</td>
</tr>
<tr>
<td>MTF4A</td>
<td>4.0%</td>
<td>0.73</td>
<td>0.98</td>
</tr>
<tr>
<td>MDH1B</td>
<td>6.6%</td>
<td>0.72</td>
<td>0.97</td>
</tr>
<tr>
<td>NPR1</td>
<td>4.6%</td>
<td>0.72</td>
<td>0.98</td>
</tr>
<tr>
<td>ERO1LB</td>
<td>5.7%</td>
<td>0.71</td>
<td>0.97</td>
</tr>
<tr>
<td>PRKCA</td>
<td>4.4%</td>
<td>0.71</td>
<td>0.98</td>
</tr>
<tr>
<td>POSTN</td>
<td>4.4%</td>
<td>0.71</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Statistical analysis
Differences among gene expression by qRT-PCR, and invasion and migration by OD560 nm were analyzed by one-way ANOVA, with pairwise comparison of each knockdown sample to the nontargeting control siRNA. Statistical analysis of clinical data is described above. Missing data was handled with case-wise deletion.

Results
Demographics
The demographics of the discovery and validation cohorts are shown in Supplementary Table S1. The discovery dataset consisted of 545 patients, 212 (39%) of which developed metastatic progression after a mean follow-up of 13.4 years. This cohort had a significant rate of adverse disease features, consistent with the high rate of metastatic disease, including 24% of patients with PSA >20 ng/mL, 39% with Gleason 8 to 10, 50% with extracapsular extension, 32% with seminal vesicle invasion, and 49% with positive margin. The validation cohorts consist of 232, 130, and 183 patients, with metastatic disease in 75 (32%), 10 (8%), and 49 (27%) patients, after mean follow-up of 6.7, 8.7, and 9.7 years. The rate of adverse disease features was slightly lower in the validation datasets, concomitant with the variably lower rates of metastatic progression.

Outlier analysis
In the discovery cohort (MCI), we used modified COPA analysis to find genes with outlier expression preferentially in patients who developed metastatic progression, as compared with patients with no evidence of metastatic disease at last follow-up. We identified 20 prognostic outlier genes that have outlier expression in 4% or more of the discovery cohort and have PPV of 0.7 or more (Table 1). We chose PPV as the primary measure of outlier gene prognostic association because we wanted to select outlier genes which identify subsets of patients at high risk for metastatic progression, that is, ruling in patients at risk for metastatic progression. However, any individual outlier gene is, by definition, only expressed in a subset of patients who will develop metastatic progression. Several of the
outlier genes have known prognostic significance across multiple cancer types (TOP2A and KI67), or specifically in prostate cancer (CAMK2N1 (30), NRP1 (31), and STMN1 (32)), thus serving as positive controls to confirm that our prognostic outlier gene nomination strategy identifies relevant genes. Interestingly, a number of the prognostic outlier genes are little studied in prostate cancer, or cancer biology in general, and we chose to focus further studies on three of these little-studied candidates: NVL (nuclear VCP-like), SMC4 (structural maintenance of chromosomes 4), and SQLE (squalene epoxidase); we included TPX2 as a positive control prostate cancer gene (33). SMC2 and NSMCE2 are structurally and functionally related to SMC4, so we chose to investigate SMC4 only (34).

Figure 1.
RNAi knockdown, cell death, and proliferation. A and B, expression knockdown of NVL, SMC4, SQLE, and TPX2 after siRNA transfection by qRT-PCR in DU 145 (A) and LNCaP-AR (B) cells. C and D, cell death as measured by Annexin V/PI flow cytometry in DU145 (C) and LNCaP-AR (D) cells. E and F, proliferation as measured by the IncuCyte system in DU145 (E) and LNCaP-AR (F) cells. ***, P < 0.001 versus siNT; ***, P < 0.001; **, P < 0.01; *, P < 0.05; #, P = 0.09.

In vitro validation assays
Prognostic genes may be bystanders in the metastatic cascade, as is the case for KI67, or may contribute biologically and therefore be putative therapeutic targets. We investigated the mechanistic involvement of the three less-studied prognostic outlier genes...
NVL, SMC4, and SQLE in the metastatic cascade using transwell invasion and migration assays after transient siRNA knockdown in the prostate cancer cell lines DU145 and LNCaP-AR. In addition, we included TPX2 in our in vitro studies as a positive control for a prognostic outlier gene with a demonstrated impact on oncogenic phenotypes (33). Knockdown efficiency was 64% for NVL in DU145 cells, and otherwise >75% for all genes in both cell lines (Fig. 1A and B). Knockdown of TPX2 strongly induced cell death (Fig. 1C and D) and inhibited proliferation (Fig. 1E and F) in both cell lines, as expected (33). NVL and SMC4 knockdown had a mild inhibitory effect on proliferation in LNCaP-AR cells (Fig. 1F). We next tested these three genes in an in vitro model of the metastatic process; the cell death induction by TPX2 was such that there were insufficient cells for invasion or migration assays for both cell lines. We observed that knockdown of NVL resulted in decreased migration in DU145 cells (Fig. 2A), and decreased migration and invasion in LNCaP-AR cells (Fig. 2B). SMC4 knockdown significantly decreased both migration and invasion in both cell lines (Fig. 2C and D). Knockdown of SQLE nonsignificantly increased migration and invasion in DU145 cells (Fig. 2E), and significantly decreased migration and invasion in LNCaP-AR cells (Fig. 2F). These results confirm that NVL, SMC4, and SQLE have biologic roles in the metastatic process in prostate cancer.

Clinical validation

To validate the prognostic significance of the outlier genes, we examined them in three independent clinical datasets from different institutions: MCII, TJU, and CCF. Focusing on the four prognostic outlier genes we studied in vitro, outlier expression of each gene individually had strong prognostic association with metastatic progression with HRs 3.7–11 (Fig. 3A–D). To understand the role of these four prognostic outlier genes after metastasis, we examined the expression of each candidate gene in metastases versus primary tumors in the Oncomine prostate cancer datasets. Indeed, each gene has significant overexpression in metastases versus primary tumor in nearly all datasets, indicating an ongoing contribution to metastatic prostate cancer (Fig. 3E). These results indicate that outlier expression of NVL, SMC4, SQLE, or TPX2 has strong prognostic association with metastatic progression after prostatectomy, and that these prognostic outlier genes continue to have a role in metastatic prostate cancer.
Prognostic outlier score

While we focused in vitro studies on a small subset of the prognostic outlier genes, our initial hypothesis that prognostic outlier status could be an important prognostic tool was not limited to NVL, SMC4, SQLE, and TPX2. Therefore, we extended our clinical validation analyses to the entire list of 20 prognostic outlier genes. As individual patients may have outlier status of more than one prognostic outlier gene, we first defined a prognostic outlier score as the number of prognostic outlier genes with outlier status. Clinicopathologic variables Gleason score, positive surgical margins, seminal vesicle invasion, and lymph node involvement were significantly different among the prognostic
outlier score groups, whereas age, PSA, stage, and extracapsular extension were not (Supplementary Table S2). The prognostic outlier score results in well-stratified patient groups in both the discovery (MCI; Fig. 4A) and the pooled validation cohorts (Fig. 4B). Multivariable analysis in both the discovery (MCI) cohort and the pooled validation cohort (Supplementary Table S3) showed outlier score to be one of the most significant predictors of 10-year metastatic progression with an OR 1.5 per each additional outlier in both the discovery and the pooled validation cohorts ($P<0.0001$). In addition, we found that this was also true for overall survival (Supplementary Fig. S4).

Given the novelty of our prognostic outlier gene approach compared with prior biomarker discovery efforts in prostate cancer, we hypothesized that the prognostic outlier score could supplement existing prostate cancer prognostic instruments. In fact, the prognostic outlier score was a significant predictor of metastatic progression independent of Decipher, CAPRA-S, or mCCP score in a multivariable model (Table 2) and significantly increased ROC AUC when integrated into each score (Supplementary Fig. S5; Table 2; refs. 18, 25).

### Discussion

High-risk prostate cancer carries a significant risk of metastatic progression after prostatectomy, and therefore better prognostic biomarkers and therapeutic strategies are needed (2, 3). Molecular biomarkers are a promising class of prognostic biomarker candidates, but these studies in prostate cancer have been limited by follow-up insufficient for study of metastatic progression or survival because of the long disease course (16). Separately, outlier analysis has identified preeminent prostate cancer genes TMPRSS2-ERG fusions and SPINK1 (19, 20), but has been limited to comparisons of cancer versus normal samples. These two approaches have been very fruitful in the prostate cancer translational research.

In this study, we sought to address both prognostic and therapeutic needs in high-risk prostate cancer by applying the novel prognostic outlier gene approach. We identified 20 prognostic outlier genes in the discovery cohort, many of which are well-known cancer genes, but several of which have not been studied in prostate cancer. We defined a prognostic outlier score as the number of prognostic outlier genes with outlier status, and demonstrated that each additional prognostic outlier gene with outlier status associates with more aggressive clinical behavior. Indeed, increasing prognostic outlier score was one of the most significant predictors of metastatic progression and overall survival after prostatectomy. The prognostic outlier score predicted outcomes independently from the published prognostic instruments Decipher, CAPRA-S, and mCCP score, and added to the prognostic ability of these tools.

### Table 2. Outlier score and other prostate cancer signatures

<table>
<thead>
<tr>
<th>Published signatures</th>
<th>Decipher</th>
<th>mCCP</th>
<th>CAPRA-S</th>
</tr>
</thead>
<tbody>
<tr>
<td>ROC AUC</td>
<td>0.768</td>
<td>0.710</td>
<td>0.715</td>
</tr>
<tr>
<td>+ Outlier score</td>
<td>0.791</td>
<td>0.783</td>
<td>0.785</td>
</tr>
<tr>
<td>$P$ value</td>
<td>0.042</td>
<td>0.0032</td>
<td>0.00098</td>
</tr>
<tr>
<td>MVA with published signature and outlier score</td>
<td>2.56E–05</td>
<td>0.000997</td>
<td>0.000084</td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td>2.6 (1.7–4.3)</td>
<td>2.4 (1.4–3.9)</td>
<td>2.2 (1.5–3.4)</td>
</tr>
<tr>
<td>Outlier score, $P$ value</td>
<td>0.00557</td>
<td>4.03E–05</td>
<td>8.68E–05</td>
</tr>
<tr>
<td>OR (95% CI)</td>
<td>2.8 (1.4–5.8)</td>
<td>4.8 (2.3–10.1)</td>
<td>4.2 (2.1–8.2)</td>
</tr>
</tbody>
</table>

Abbreviations: AUC, area under the curve; MVA, multivariable analysis; ROC, receiver-operating curve.
We found that three prognostic outlier genes that are unstudied in prostate cancer contribute to oncogenic phenotypes in two prostate cancer cell lines. TPX2 knockdown induced cell death as our positive control for a prognostic outlier gene with an impact on oncogenic phenotypes (33); NVL, SMC4, or SQLE knockdown significantly decreased migration and/or invasion in one or both cell lines used. The molecular mechanisms of these prognostic outlier genes are not yet clear. NVL is a member of the AAA ATPase superfamily with two predominant isoforms in the nucleolus and nucleoplasm, is reportedly involved in ribosomogenesis, and is an essential component of the TERT holoenzyme (42, 43). Interruption of either of these functions could potentially have significant impact on tumor cell biology. However, this is the first report to implicate NVL in cancer biology, and further study is required. SMC4 is an ATPase core protein of the condensin I and II complexes that capture and condense chromatids in the early stages of mitosis (44). SMC4 has been implicated in colorectal, liver, and breast cancer (45–47), but has not been studied in prostate cancer or implicated in the metastatic cascade. In addition, SMC4/condensin interacts with the genomic transcriptional insulator CTCF, and thus may be required for oncogenic gene silencing (48). SQLE is a cholesterol biosynthesis enzyme that has been implicated in several cancers other than prostate, and, interestingly, is located in the chromosome 8q24 Myc oncogene amplicon (49–51). SQLE knockdown had an impact in LNCaP-AR but not DU145 cells, which is likely reflective of the much more significant role of steroid hormone signaling in LNCaP-AR cells. Although some molecular details about each gene are known, the prognostic performance demonstrated in our study provides strong motivation for further investigation in prostate cancer.

Although our study is able to leverage the power of several large cohorts, there are potential limitations with our data. We cannot account for all inter- and intracohort variation from factors such as warm/cold ischemia time, different tissue fixation and processing procedures, relative ratios of different cells (e.g., tumor, stroma, benign glands, inflammatory cells, etc.) and different patient characteristics and treatment protocols at each institution. However, we attempted to mitigate these effects by performing all RNA processing and microarray hybridization in the same CLIA-certified laboratory, and by using multivariable analysis and stratification to account for measured and unmeasured confounders, respectively. In addition, we note that a study specifically examining the effects of ischemia time on gene expression found <1% of genes significantly changed, and there is no overlap with the prognostic outlier genes identified in this study (52).

In summary, we have performed a novel analysis in a large high-risk prostate cancer cohort with long follow-up to identify 20 prognostic outlier genes as potential biomarkers and therapeutic targets. In vitro validation of three prognostic outlier genes unstudied in prostate cancer, NVL, SMC4, and SQLE, confirmed that these genes play a causal role in the metastatic cascade, and therefore may be further explored for therapeutic intervention. Importantly, the prognostic performance of outlier status of the 20 prognostic outlier genes was validated as one of the strongest prognostic predictors of metastatic progression across three independent clinical datasets comprising more than 500 patients, and provided prognostic information independent of three published prognostic instruments. On the basis of these findings, prognostic outlier genes may serve as an important novel class of cancer genes for biomarker and therapeutic discovery in prostate cancer.

Disclosure of Potential Conflicts of Interest

Shuang G. Zhao reports receiving travel reimbursement from Genome Dx Biosciences. E.M. Schaeffer, A.E. Ross, A.P. Dicker, P.L. Nguyen, and F.Y. Feng are consultant/advisory board members for Genome Dx Biosciences. E.A. Klein reports receiving research funding from Genome Dx Biosciences. R.J. Karnes reports receiving research travel reimbursement and funding from Genome Dx Biosciences. No potential conflicts of interest were disclosed by the other authors.

Authors’ Contributions

Conception and design: S.G. Zhao, J.R. Evans, E.M. Schaeffer, E.A. Klein, A.P. Dicker, E. Davicioni, F.Y. Feng
Development of methodology: S.G. Zhao, J.R. Evans, V. Kothari, A.P. Dicker, E. Davicioni, F.Y. Feng
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J.R. Evans, V. Kothari, G. Sun, A. Larm, V. Mondine, E.M. Schaeffer, A.E. Ross, R.B. Den, R.J. Karnes, E. Davicioni
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): S.G. Zhao, J.R. Evans, V. Kothari, G. Sun, A. Larm, V. Mondine, E.M. Schaeffer, E.A. Klein, R.B. Den, A.P. Dicker, R.J. Karnes, N. Erho, P.L. Nguyen, E. Davicioni, F.Y. Feng
Writing, review, and/or revision of the manuscript: S.G. Zhao, J.R. Evans, A.E. Ross, E.A. Klein, R.B. Den, A.P. Dicker, R.J. Karnes, N. Erho, P.L. Nguyen, E. Davicioni, F.Y. Feng
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S.G. Zhao, J.R. Evans, V. Kothari, V. Mondine, A.P. Dicker, N. Erho, E. Davicioni, F.Y. Feng
Study supervision: S.G. Zhao, J.R. Evans, P.L. Nguyen, F.Y. Feng

Acknowledgments

The authors thank Steven Kronenberg for expertly preparing the figures.

Grant Support

This research was supported by Challenge Grants awarded (to F.Y. Feng and J.R. Evans) by the Prostate Cancer Foundation and the Evans Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 2, 2015; revised October 19, 2015; accepted November 3, 2015; published OnlineFirst December 2, 2015.

References


Clinical Cancer Research

The Landscape of Prognostic Outlier Genes in High-Risk Prostate Cancer

Shuang G. Zhao, Joseph R. Evans, Vishal Kothari, et al.

Clin Cancer Res  Published OnlineFirst December 2, 2015.

Updated version  Access the most recent version of this article at:

Supplementary Material  Access the most recent supplemental material at:
http://clincancerres.aacrjournals.org/content/suppl/2015/12/02/1078-0432.CCR-15-1250.DC1

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.