Supplementary Materials

Supplementary Materials

Results

The twelve markers identified in this study were taken forward into another independent study of prostate cancer FFPE biopsy samples to develop a locked down model for clinical use (manuscript submitted). In this new study, we identified the best marker subset of the 12 markers and locked the resulting 8-marker model down, containing the following biomarkers: SMAD4, FUS, CUL2, YBX1, DERL1, PDSS2, HSPA9 and pS6. In the interest of completeness, we analyzed this set of markers on the TMA samples in this study, with the understanding that the TMA cohort contributed to the marker selection process. We again used the same patient partition, and trained on the L TMA followed by testing on both L TMA and H TMA samples. We analyzed 268 patients containing 40 dead-from-disease events. The resulting test AUC based

on L TMA for prediction of aggressive disease was 0.64 (95% CI: 0.56–0.71) with a test odds ratio for aggressive disease of 13 per unit change in risk score (95% CI: 2.3–341). The test hazard ratio for lethal outcome prediction was 14 per unit change in risk score (95%CI: 1.3–393). To confirm the ability to generalize across sampling error, the model derived from L TMA train was also tested on the test H TMA with consistent results for both indications. The H TMA test AUC was 0.70 (95% CI: 0.62–0.78) with an odds ratio for aggressive disease of 46 per unit change in risk score (95% CI: 5.6–1290). The H TMA test hazard ratio for prediction of lethal outcome was 19 per unit change in risk score (95% CI: 1.4–620).

Materials and Methods

Generation of TMA blocks

TMA blocks were prepared using a modified agarose block procedure(Yan *et al*, 2007). To generate the test TMA (MPTMA10), we selected 72 FFPE tissue blocks of prostatectomy samples with available annotations for GS and pathological stage. Of these, 37 had a GS of 3 + 3 = 6 with T2 stage, while 35 had a GS of 4 + 3 = 7 or a GS of either 3 + 3 = 6 or 3 + 4 = 7 with T3b stage. One 1 mm core per patient sample was taken from areas of lowest Gleason pattern and placed into an acceptor block.

For construction of H and L TMAs, we used the cohort of FFPE human prostate cancer tissue blocks with clinical annotations and long-term patient outcome information. For each patient sample, a core was taken from an area with the highest Gleason pattern and deposited into an H acceptor block. A second core was then taken from an area with the lowest Gleason pattern and put into an L acceptor block. The order of sample core placement into H block was randomized,

and core positions in the L block were identical to those in the H block. In addition, cores from FFPE blocks of cell-line controls (Table S5) were placed in the upper and lower parts of all H and L TMA blocks. Upon completion, 5 μm serial sections were cut from each block and representative sections were stained with H&E and scanned with the ScanScope XT system. Images of H&E-stained cores were then independently annotated for observed Gleason pattern by a board-certified anatomical pathologist in a blinded manner.

The resulting H and L TMA blocks were identical for a set of patient samples, but differed in observable Gleason pattern (Figure 1, bottom). For this study, two pairs of TMA blocks (MPTMAF5H and 5L, 6H and 6L) were generated with cores from 380 patient samples.

Biomarker selection

To identify biomarkers for prostate cancer aggressiveness, we developed a selection and evaluation process that could be broadly applicable across diseases and conditions. The process, shown in Figure 2, had biological, technical, performance and validation stages.

In the biological stage, an initial list of potential biomarkers for prostate cancer aggressiveness was compiled from publically available data. The list was then prioritized based on biological relevance, *in silico* analysis, review of the Human Protein Atlas (www.proteinatlas.org), and commercial availability of requisite MAbs. Biological relevance review was based on mechanism of action in cells and, in particular, in the disease. *In silico* analysis was based on previously known gene amplifications, deletions and mutations, and univariate performance or

progression correlation between these genetic alterations and the disease. The Human Protein Atlas contains data on protein expression levels in various tissues across disease states.

In the technical stage, commercial MAbs were obtained and tested for their ability to detect biomarkers from clinical samples. Initially, we stained samples of malignant and benign prostatic tissue using a DAB-based IHC staining procedure and selected candidate antibodies that exhibited a good signal:noise ratio and were specific for epithelial cell staining. We further tested successful candidates on malignant and benign prostatic tissue samples using IF along with region-of-interest markers, epithelial cytokeratins CK8 and CK18 and basal markers CK5 and Trim29, as described (Supplementary Materials and Methods). Antibodies and biomarkers that met the IF criteria were taken forward to the performance stage.

In the performance stage, MAbs were tested on TMAs. Performance was evaluated for a univariate correlation between tumor epithelium expression and disease state. The MAbs and biomarkers that demonstrated univariate correlation between expression and disease state were then evaluated on a larger H and L TMA set for both univariate correlation and performance in combination with other markers.

The quantitative multiplex immunofluorescence (QMIF) staining procedure

The QMIF was composed of two initial blocking steps followed by four MAb incubation steps with appropriate washes in between. Blocking consisted of biotin blocking steps followed by treatment with Sniper reagent (Biocare Medical), according to the manufacturer's instructions. The first MAb incubation step consisted of a mixture of anti-biomarker 1 mouse MAb and anti-

biomarker 2 rabbit MAb, followed by a second step containing a mixture of anti-mouse IgG Fab–fluorescein isothiocyanate (FITC) and anti-rabbit IgG Fab–biotin. A third "visualization" step included a mixture of anti-FITC MAb–Alexa 568, streptavidin–Alexa 633, as well as MAbs against epithelium (anti-CK8–Alexa 488 and anti-CK18–Alexa 488) and basal epithelium (anti-CK5–Alexa 555 and anti-Trim29–Alexa 555), respectively. A final, fourth step comprised a brief incubation with 4′,6-diamidino-2-phenylindole (DAPI) for nuclear staining. After final washes, slides were mounted with Prolong Gold™ (Life Technologies) before coverslips were added. Slides were kept permanently at –20°C before and after imaging.

For diaminobenzidine (DAB)-based IHC staining, slides with tissue were processed as described above, blocked with Sniper reagent™ (Biocare Medical) and incubated with primary antibody solution. UltraVision (Thermo Scientific) was used as a secondary reagent. Finally, tissue was counterstained with hematoxylin and coverslips were added.

FFPE tissue block quality evaluation

A 5 μm section from each FFPE block was manually stained with anti-phospho STAT3(T705) rabbit MAb, anti-STAT3 mouse MAb and region-of-interest markers, as described above. Slides were visually examined under a fluorescence microscope. Based on the staining intensities and autofluorescence, the sections and their corresponding FFPE blocks were graded into four quality categories.

Image acquisition

Two Vectra Intelligent Slide Analysis Systems (PerkinElmer) were used for automated image acquisition as described elsewhere. DAPI, FITC, tetramethylrhodamine isothiocyanate (TRITC)

and Cy5 long pass filter cubes were optimized for maximal multiplexing capability. Vectra 2.0 and Nuance 2.0 software packages (PerkinElmer) were used for automated image acquisition and development of the spectral library, respectively.

TMA acquisition protocols were run in an automated mode according to the manufacturer's instructions (PerkinElmer). Two 20× fields per core were imaged using a multispectral acquisition protocol that included consecutive exposures with DAPI, FITC, TRITC and Cy5 filters. For maximal reproducibility, light source intensity was adjusted with the help of an X-Cite Optical Power Measurement System (Lumen Dynamics) before image acquisition for each TMA slide. Identical exposure times were used for all slides containing the same antibody combination. A set of TMA slides stained with the same antibody combinations was imaged on the same Vectra microscope.

A spectral profile was generated for each fluorescent dye as well as for FFPE prostate tissue autofluorescence. Interestingly, two types of autofluorescence were observed in FFPE prostate tissue. A typical autofluorescence signal was common in both benign and tumor tissue, whereas an atypical "bright" type of autofluorescence was specific for bright granules present mostly in epithelial cells of benign tissue. A spectral library containing a combination of these two spectral profiles was used to separate or "unmix" individual dye signals from the autofluorescent background.

Definiens automated image analysis

We developed an automated image analysis algorithm using Definiens Developer XD for tumor identification and biomarker quantification. For each 1.0 mm TMA core, two 20× image fields were acquired. The Vectra multispectral image files were first converted into multilayer TIFF files using inForm (PerkinElmer) and a customized spectral library, and then converted to single-layer TIFF files using BioFormats (OME). The single-layer TIFF files were imported into the Definiens workspace using a customized import algorithm so that, for each TMA core, both of the image field TIFF files were loaded and analyzed as "maps" within a single "scene".

Autoadaptive thresholding was used to define fluorescent intensity cut-offs for tissue segmentation in each individual tissue sample in our image analysis algorithm. Cell-line control cores within the TMA were automatically identified in the Definiens algorithm based on predefined core coordinates. The tissue samples were segmented using the fluorescent epithelial and basal cell markers, along with 4′,6-diamidino-2-phenylindole (DAPI) for classification into epithelial cells, basal cells, and stroma, and further compartmentalized into cytoplasm and nuclei. Individual gland regions were classified as malignant or benign based on the relational features between basal cells and adjacent epithelial structures combined with object-related features, such as gland thickness. Epithelial markers are not present in all cell lines, therefore the cell-line controls were segmented into tissue versus background using the autofluorescence channel. Fields with artifact staining, insufficient epithelial tissue, or out-of-focus images were removed by a rigorous multi-parameter quality-control algorithm.

Epithelial marker and DAPI intensities were quantified in malignant and nonmalignant epithelial regions as quality-control measurements. Biomarker intensity levels were measured in the cytoplasm, nucleus, or whole cell in the malignant tissue based on predetermined subcellular localization criteria. The mean biomarker pixel intensity in the malignant compartments was averaged across the maps with acceptable quality parameters, to yield a single value for each tissue sample and cell line control core.

Univariate analysis of aggressiveness and lethality

Our objectives for univariate analysis were twofold: to characterize univariate behavior as a performance assessment for potential inclusion in the final marker set, and to provide a reduced set of markers for exhaustive multivariable model exploration. All modeling was done in R 3.0 using standard functions and packages, including glm, survival, KMsurv, binom, and pROC. Biomarkers were assessed based on two outcomes: prediction of Surgical GS and prediction of death (lethality). Prediction of Surgical GS, categorized as indolent or more severe, was modeled with both ORs (logistic regression) and biomarker means (linear regression). Lethality was modeled using HRs (traditional Cox proportional hazards), ORs (logistic regression), and marker means (linear regression). In addition, to provide nonparametric and robust assessments, Wilcoxon and permutation tests were applied. Figure 3 shows the key results. Univariate results were also directly considered in selection of the final marker set, as seen in Figure 5.

Biomarker ranking for aggressiveness via exhaustive search of multimarker models

We sought to rank the biomarkers by importance in multimarker models; 31 biomarkers, refined from the original set of 39 to improve technical performance further, were used in an exhaustive

biomarker search. We considered all combinations of up to five biomarkers from the 31 biomarkers tested in the L TMA in the H and L TMA analysis. For each biomarker combination, 500 training sets were generated by bootstrapping, and associated complementary test sets were obtained. A logistic regression model was applied to each training set and then tested on each of the associated test sets. Training and test AUC (i.e. C statistic) and training AIC were obtained in each round. Medians and 95% CIs were obtained for all three statistics.

We then considered biomarker selection frequency in the models and sorted them by their AIC and, separately, by their test AUC. For each of the resulting rankings of the models, the frequency of biomarker utilization in the top 1% and the top 5% of the lists was determined. The biomarkers that were included in at least 50% of models were then identified.
Table S2 shows biomarker frequency in the prediction of aggression assessment. The performance of the top-ranking models was similar. Moreover, the number of biomarkers in the top-ranking models varied. To resolve this issue, which appeared to relate to model size, we considered the top 1% of the models sorted by test AUC. We studied the resulting distributions for a number of different population assumptions, including cases where intermediate core GSs were excluded from analysis, or were included with indolent scores, or were included with high scores. In the final analysis, we concluded that an eight-biomarker model provided the best trade-off between performance and complexity in this experimental data set.

FFPE cell-line controls

Selected cell lines were grown in standard conditions with and without treatment before harvesting as indicated (Table S5). Cells were washed with phosphate-buffered saline (PBS),

fixed directly on plates with 10% formalin for 5 min, then scraped and collected in PBS with continued fixation at room temperature for 1 hour. Cells were washed twice with PBS, resuspended in Histogel (Thermo Scientific) at 70°C and quickly spun down in a1.5 ml microfuge tube to form a condensed cell–Histogel pellet. The pellets were then embedded in paraffin and placed into standard paraffin blocks that served as donor blocks for TMA construction. DU145 cells with inducible knock down of CCND1 and SMAD4 were established according to manufacturer's instructions using the 'Tet-one' system (Clontech).

Antibody specificity assays

Several MAbs, including anti-ACTN1, anti-CUL2, anti-Derlin1, anti-FUS, anti-PDSS2, anti-SMAD2, anti-VDAC1, anti-YBX1, and anti-HSPA9, were validated by Western blotting (WB) and immunohistochemistry (IHC) assay of target-specific knockdown and control cells (Figure S1). Details of the small interfering RNA (siRNA) sequences and host cell lines are listed in Table S6. Cells were seeded into 12-well plates and transfected with 25 nM of siRNAs and DharmaFect transfection reagent (Thermo Scientific Dharmacon); mock transfection included only the transfection reagent. Cells transfected with two nontargeting sequences were also included as controls.

For WB assay, transfected cells were harvested at 72 hours and lyzed with Pierce RIPA buffer (Thermo Scientific) supplemented with Halt protease inhibitor cocktail (Thermo Scientific). Protein concentration was measured using Pierce BCA reagent (Thermo Scientific). Samples were adjusted to equal protein concentrations and then mixed with sample buffer (Boston BioProducts) and run on precast Criterion TGX 4–15% SDS-PAGE gels (Bio-Rad). The samples were transferred onto PVDF or nitrocellulose membranes using the IBlot apparatus (Life

Technologies), and immunoblotted with antibodies at 4°C overnight, followed by incubation with secondary mouse or rabbit MAbs (Sigma Aldrich). The blots were developed with SuperSignal West Femto reagents (Thermo Scientific), and visualized by exposure to the FluorChem Q system (Protein Simple).

For the IHC assay, cells grown on coverslips in a 12-well plate were fixed with methanol on ice for 20 min at 72 hours post-transfection. This was followed by permeabilization with 0.2% Triton X-100 on ice for 10 min. UltraVision LP Detection System HRP Polymer/DAB Plus Chromogen Kit (Thermo Scientific) was used for the subsequent IHC assay according to the manufacturer's instructions.

The SMAD4 antibody was validated by WB and IHC assays of the SMAD4-positive cell line PC3 and the SMAD4-negative cell line BxPC3. The phospho-S6 antibody was validated by WB and IHC of naïve and LY294002-treated DU145 cells.

Cell proliferation assay

HeLa cells were transiently transfected with two nontargeting siRNAs as well as si9-11, specific for HSPA9 (see Table S6 for details of siRNA sequences). Cells were replated 48 hours after transfection and seeded in triplicate at 1000 cells per well in a 96-well plate. Cell proliferation was monitored using a CellTiter-Glo® Luminescent Cell Viability Kit (Promega) according to the manufacturer's instructions at 0, 24, 72 and 120 hours after replating.

Clonogenic assay

At 48 hours post-transfection, HeLa cells were replated at 500 cells per well in a 6-well plate with 2 ml of cell medium. The cells were fixed with Crystal Violet Solution (Sigma) 7 days after plating. The images of each well were captured using AlphaView software in the FluorChem Q system (Protein Simple) and processed using ImageJ software.

Cell vitality assay

HeLa cells were harvested at 120 hours post-transfection. Cells were collected using trypsin. The cell pellets from each well of a 12-well plate were suspended in 500 µl of cell medium. Cell suspension (95 µl) was mixed with 5 µl of Solution 5 (VB-48/PI/AO), and 30 µl of the mixture was loaded onto an NC-Slide A2 (both from ChemoMetec). Cell vitality was measured by a NucleoCounter NC-3000™ (ChemoMetec) according to the manufacturer's instructions.

Caspase assay

HeLa cells were harvested at 120 hours after siRNA transfection using trypsin. Cells were suspended at 2×10^6 cells/ml. An aliquot of 93 µl of the cell suspension was mixed with 5 µl diluted FLICA reagent (ImmunoChemistry Technologies) and 2 µl of Hoechst 33342 (Life Technologies). The mixture was incubated at 37°C for 1 hour. HeLa cells were washed twice with 1× Apoptosis Buffer (ImmunoChemistry Technologies). The cell pellets were suspended in 100 µl 1× Apoptosis Buffer and 2 µl of propidium iodide. A 30 µl aliquot of the mixture was loaded onto an NC-Slide A2 and read using NucleoCounter NC-3000 software for caspase assay. Cells positive for FLICA staining were counted as apoptotic cells.

Identification of HSPA9 (Mortalin)

For identification of the Leica “anti-DCC” antibody target (Figure S1), a preparative immunoprecipitation was performed. Ten p100 plates of confluent A549 cells were harvested with 5 ml of RIPA buffer (Thermo Scientific) with added protease inhibitors. The cell lysate was spun down at 14,000 rpm for 5 min; the supernatant was heated for 5 min at 80°C, then chilled on ice, and spun down again at 14,000 rpm for 5 min. Supernatant was collected and, after addition of 50 µl of Protein A/G beads (Thermo Scientific) with 2 µg of pre-bound “anti-DCC” antibody, was incubated with rocking at 4°C for 2 hours. Beads were washed three times with TBS + 1% Triton X100, and boiled with 30 µl of 1× SDS-PAGE loading buffer. Supernatant was loaded onto a 10% SDS-PAGE gel, and separated under standard SDS-PAGE conditions. The gel was stained with a silver stain kit for mass spectrometry (Thermo Scientific); the specific band was cut out, digested with trypsin, and subjected to MS/MS sequencing mass spectrometry at the Taplin Mass Spectrometry Facility (Harvard Medical School). Identified peptides were aligned with Human Protein reference databases. The identified protein HSPA9 was further validated as described.

Supplementary Figures and Tables

Figure S1. Antibody specificity. The specificity of ACTN1 (A), CUL2 (B), Derlin1 (C), FUS (D), PDSS2 (E), SMAD2 (F), VDAC1 (G), and YBX1 (H) antibodies were validated by Western blotting (WB) and immunohistochemistry (IHC) of siRNA-treated cells and control cells. Marker-specific siRNA treatment significantly reduced the intensity of the band on WB, and the specific IHC staining in cells confirmed the specificity of the antibodies. The specificity of the SMAD4 antibody (I) was validated by WB and IHC of the SMAD4-positive control cell line PC3 and SMAD4-negative control cell line BxPC3. The specificity of the pS6 antibody (J) was validated by WB and IHC of naïve DU145 cells and DU145 cells treated with PI3K inhibitor LY294002. LY294002 treatment significantly reduced phosphorylation of S6, as shown by WB and IHC. The Leica anti-DCC antibody (K) detected a band on WB that did not match the expected size for the DCC protein (marked "X" in K); IHC staining was also not reduced in DCC siRNA-treated cells (left panel in K). The Leica anti-DCC antibody appeared to recognize the HSPA9 protein, as shown by WB and IHC of HSPA9 siRNA-treated cells and control cells (right panel in K). β-Actin was used as a WB loading control.

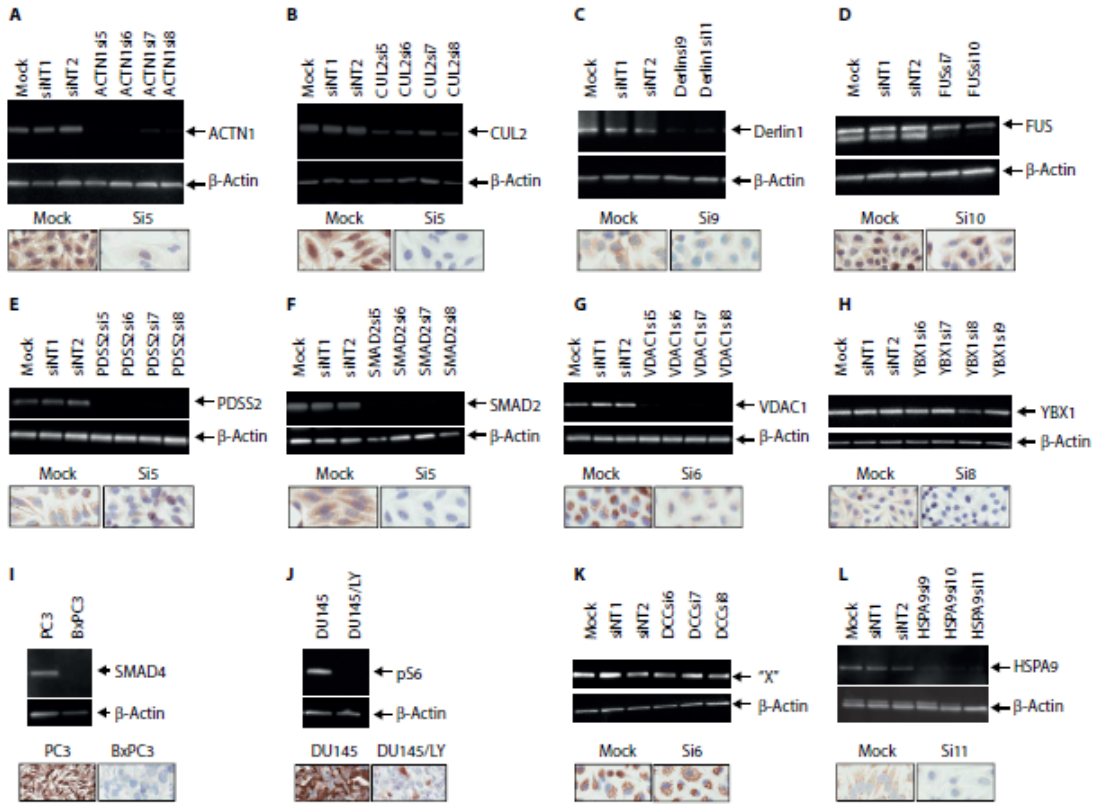
A
Mock
siNT1
siNT2
ACTN1si5
ACTN1si6
ACTN1si7
ACTN1si8
ACTN1
β-Actin
Mock
Si5
B
Mock
siNT1
siNT2
CUL2si5
CUL2si6
CUL2si7
CUL2si8
CUL2
β-Actin
Mock
Si5
C
Mock
siNT1
siNT2
Derlinsi9
Derlin1si11
Derlin1
β-Actin
Mock
Si9
D
Mock
siNT1
siNT2
FUSsi7
FUSsi10
FUS
β-Actin
Mock
Si10
E
Mock
siNT1
siNT2
PDSS2si5
PDSS2si6
PDSS2si7
PDSS2si8
PDSS2
β-Actin
Mock
Si5
F
Mock
siNT1
siNT2
SMAD2si5
SMAD2si6
SMAD2si7
SMAD2si8
SMAD2
β-Actin
Mock
Si5
G
Mock
siNT1
siNT2
VDAC1si5
VDAC1si6
VDAC1si7
VDAC1si8
VDAC1
β-Actin
Mock
Si6
H
Mock
siNT1
siNT2
YBX1si6
YBX1si7
YBX1si8
YBX1si9
YBX1
β-Actin
Mock
Si8
I
PC3
BxPC3
SMAD4
β-Actin
PC3
BxPC3
J
DU145
DU145/LY
pS6
β-Actin
DU145
DU145/LY
K
Mock
siNT1
siNT2
DCCsi6
DCCsi7
DCCsi8
"X"
β-Actin
Mock
Si6
L
Mock
siNT1
siNT2
HSPA9si9
HSPA9si10
HSPA9si11
HSPA9
β-Actin
Mock
Si11

Figure S2. Identification of HSPA9 instead of DCC as a prostate cancer prognosis biomarker. The Leica anti-DCC antibody was not validated by DCC siRNA knockdown cells by WB and IHC (A), because the size of the band detected by the antibody on WB was much smaller than what was expected for DCC protein (75 kDa vs 158 kDa) and the IHC staining intensity was not reduced in DCC siRNA-treated cells. Mass spectrometry identified the protein recognized by the Leica anti-DCC antibody on WB to be HSPA9. To confirm that the Leica anti-DCC antibody was indeed an anti-HSPA9 antibody, it was tested by WB and IHC on HSPA9 siRNA-treated cells and control cells; both the WB band and the IHC staining detected by the Leica anti-DCC antibody were significantly reduced in HSPA9 siRNA-treated cells (B). The WB and IHC patterns of the Leica anti-DCC antibody on HSPA9 siRNA-treated cells were similar to those detected by a Santa Cruz anti-HSPA9 antibody (C). Silencing HSPA9 by siRNA appeared to decrease the proliferation of HeLa cells (D), reduced HeLa cell colony formation in a clonogenic assay (E), and caused increased cell death (F) and caspase activity (G).

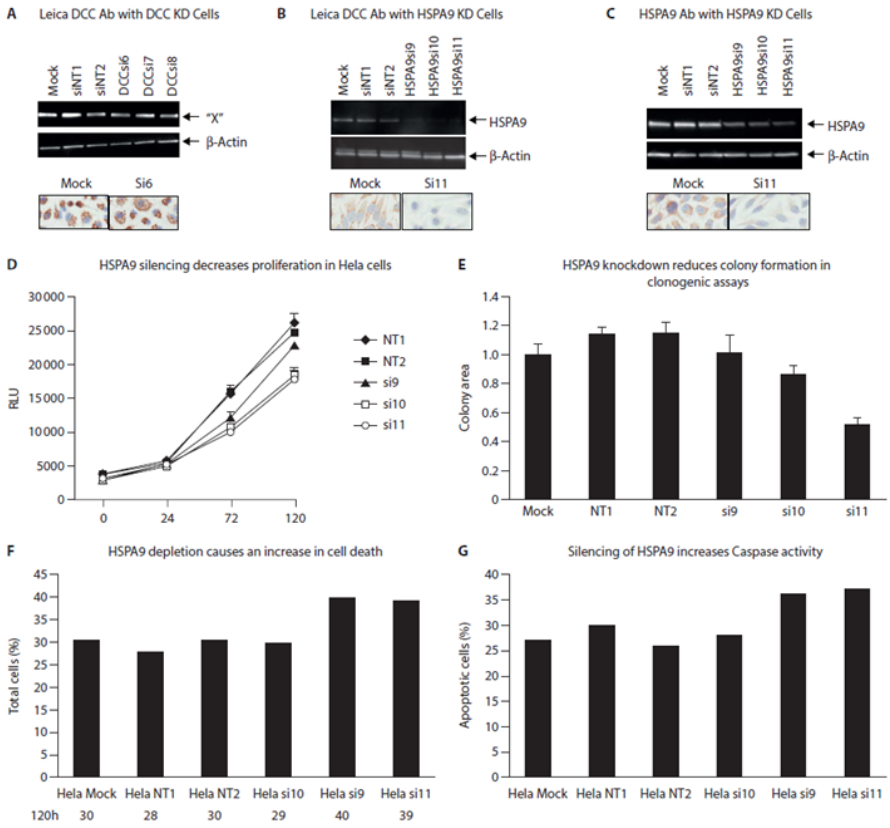
A
Leica DCC Ab with DCC KD Cells
Mock
siNT1
siNT2
DCCsi6
DCCsi7
DCCsi8
"X"
β-Actin
Mock
Si6
B
Leica DCC Ab with HSPA9 KD Cells
Mock
siNT1
siNT2
HSPA9si9
HSPA9si10
HSPA9si11
HSPA9
β-Actin
Mock
Si11
C
HSPA9 Ab with HSPA9 KD Cells
Mock
siNT1
siNT2
HSPA9si9
HSPA9si10
HSPA9si11
HSPA9
β-Actin
Mock
Si11
D
HSPA9 silencing decreases proliferation in Hela cells
RLU
30000
25000
20000
15000
10000
5000
0
0
24
72
120
NT1
NT2
si9
si10
si11
E
HSPA9 knockdown reduces colony formation in clonogenic assays
Colony area
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0
Mock
NT1
NT2
si9
si10
si11
F
HSPA9 depletion causes an increase in cell death
Total cells (%)
45
40
35
30
25
20
15
10
5
0
Hela Mock
Hela NT1
Hela NT2
Hela si10
Hela si9
Hela si11
120h
30
28
30
29
40
39
G
Silencing of HSPA9 increases Caspase activity
Apoptotic cells (%)
40
35
30
25
20
15
10
5
0
Hela Mock
Hela NT1
Hela NT2
Hela si10
Hela si9
Hela si11

Table S1. Candidate biomarkers identified from published literature and gene expression databases. Notes: **DAB staining and specificity.** Passed: antibodies that with DAB-based immunohistochemical staining demonstrated signal intensity and staining pattern of benign and tumor prostate tissue commensurate with published literature. **Immunofluorescence signal and specificity.** Passed: antibodies that with immunofluorescent staining exhibited high level of signal and with staining pattern of benign and tumor prostate tissue commensurate with published literature. **Marker stability in tissue.** Passed: antibodies that showed signal intensities correlating with epithelial marker staining intensities across tissue areas of variable quality. **MPTMA 10.** Passed: antibodies that demonstrated correlation between expression and Surgical (prostatectomy) Gleason score.

Biomarker	DAB staining and specificity	Immunofluorescence signal and specificity	Marker Stability in tissue	MPTMA 10	Tested on H and L TMAs
PIK3R1	Failed				
PHLPP1 (Poly)	Passed	Passed	Passed	Failed	
CDKN1B (p27kip1)	Passed	Passed	Passed	Passed	Yes
SPRY2	Failed				
NCOR2	Passed	Failed			
E2F1	Failed				
Top2A	Failed				
IGF1	Failed				
EGR1	Failed				
SRF	Passed	Failed			
CTGF	Failed				
CCL2	Failed				
FUS	Passed	Passed	Passed	Passed	Yes
LKB1 (STK11)	Passed	Failed			
CD142	Passed	Failed			
MTHFD1L	Failed				

SHMT2	MAb not available				
KRT6A	Failed				
LOX	Failed				
CD53	Passed	Failed			
CUL2	Passed	Passed	Passed	Passed	Yes
MBD2	Failed				
MTERF	MAb not available				
PARD3	Failed				
RBL2	Passed	Passed	Passed	Failed	
SMAD2	Passed	Passed	Passed	Passed	Yes
SMAD7	Failed				
DSC2	Passed	Passed	Failed		
EMD	Passed	Passed	Passed	Failed	
PRMT1	Passed	Passed	Passed	Failed	
REV1	Passed	Failed			
StAR	Passed	Passed	Passed	Passed	Yes
CPNE3	MAb not available				
CML66	Passed	Passed	Passed	Failed	
GRINA	Passed	Failed			
SPAG1	MAb not available				
ANPTL4	MAb not available				
TGS1	MAb not available				
WWP1	Passed	Passed	Passed	Failed	
ATF2	Passed	Failed			
COPB2	Passed	Passed	Passed	Failed	
DERL1	Passed	Passed	Passed	Passed	Yes
FAM91A1	MAb not available				
FOLH1	Passed	Failed			
KIF5C	Passed	Passed	Passed	Failed	
NPC2	Failed				
OXCT1	MAb not available				
RAB18	Failed				
RHOA	Passed	Passed	Passed	Failed	

UNC13B	Failed				
YIPF6	MAb not available				
ST6GAL1 (CD75)	Passed	Passed	Passed	Passed	Yes
BHLHE40 (Dec1)	Passed	Passed	Passed	Passed	Yes
BHLHE41 (Dec2)	Passed	Failed			
EIF2C2	Passed	Failed			
PUF60	Failed				
WDR67	MAb not available				
SQLE	Passed	Failed			
RNF19A	Failed				
UBR5	Passed	Failed			
PABPC1	Passed	Passed	Passed	Failed	
EIF3H	Passed	Passed	Passed	Passed	Yes
ARMC1	Failed				
WDYHV1	MAb not available				
ANKRD46	MAb not available				
AKAP9	Failed				
AKAP8	Passed	Passed	Passed	Passed	Yes
EEF1D	Failed				
TMEM68	MAb not available				
SRI	Passed	Failed			
HOXB13	Passed	Passed	Passed	Passed	Yes
NCOA2 (clone 29)	Passed	Passed	Passed	Passed	Yes
SLC2A4/GLUT4	Failed				
GRIP-1	Passed	Passed	Passed	Passed	Yes
SCRIB	Passed	Passed	Failed	Failed	
PXN	Passed	Passed	Passed	Passed	Yes
ARHGEF7	Passed	Failed			
RAVER1	Failed				
PTBP1	Passed	Passed	Failed	Failed	
KHDRBS2	MAb not available				
KHDRBS3	Passed	Passed	Passed	Failed	

UBE2L3	Failed				
UBE2L6	Failed				
SNCG	Passed	Failed			
MT-CO2	Passed	Passed	Passed	Failed	
RTN4	Failed				
COMT	Passed	Passed	Passed	Failed	
PNMT	Failed				
ABL2	Failed				
ACTN1	Passed	Passed	Passed	Passed	Yes
CDC7	Failed				
CPNE3	MAb not available				
DAB2	Failed				
FKBP5	Passed	Passed	Passed	Passed	Yes
HMMR	Failed				
ITGB3BP	Failed				
KIAA0196	MAb not available				
KIF11	Passed	Passed	Passed	Failed	
MAP2K6	Failed				
MRPL37	MAb not available				
MTHFD2	Failed				
NRP1	Failed				
OXCT1	MAb not available				
ST14	Passed	Passed	Passed	Failed	
PDSS2	Passed	Passed	Passed	Passed	Yes
DIABLO	Passed	Passed	Passed	Passed	Yes
ATP6V1F	Failed				
AZGP1	Passed	Failed			
CAPZA2	Passed	Failed			
COX6C	Passed	Passed	Passed	Passed	Yes
DAD1	Failed				
HSD17B4	Passed	Passed	Passed	Passed	Yes
PRDX5	Passed	Failed			
SLC22A3	Passed	Passed	Failed		
YBX1	Passed	Passed	Passed	Passed	Yes
MAOA	Passed	Passed	Passed	Passed	Yes
SHMT2	Failed				

ECHS1	Failed				
TMEM16G	Failed				
VCAN	Failed				
PDIA3	Passed	Passed	Failed		
MAP3K5	Passed	Passed	Passed	Passed	Yes
ANXA5	Failed				
TRAF4	Passed	Failed			
VCP	Failed				
VDAC1	Passed	Passed	Passed	Passed	Yes
COL1A2	Failed				
SSTR1	Failed				
LACTB2	Passed	Failed			
XKR9	Passed	Failed			
PEBP4	Failed				
PPP3CC	Failed				
SLC39A14	Passed	Failed			
LATS2	Passed	Passed	Passed	Passed	Yes
PLAG1	Passed	Passed	Passed	Passed	Yes
Stat 5	Failed				
cMyc	Passed	Failed			
ANO7	Passed	Passed	Failed		
AGPAT6	Passed	Passed	Passed	Passed	Yes
ROCK1	Passed	Failed			
RAD21	Passed	Failed			
FASN	Passed	Passed	Failed		
PECI	Passed	Failed			
Stathmin	Failed				
SLC16A1	Passed	Passed	Failed		
TGM2	Failed				
Ubc2H10	Passed	Failed			
EZH2	Passed	Passed	Passed	Passed	Yes
AR	Passed	Passed	Failed		
FOXA1	Passed	Passed	Failed		
HSPA9	Passed	Passed	Passed	Passed	Yes
FAK1	Passed	Passed	Passed	Passed	Yes
LMO7	Passed	Passed	Passed	Passed	Yes
MTDH2	Passed	Passed	Passed	Passed	Yes
AGK	Passed	Passed	Passed	Passed	Yes
CDH10	Passed	Passed	Passed	Failed	

COBP2	Passed	Passed	Passed	Failed	
CRLF1	Passed	Passed	Passed	Failed	
RASSF1	Passed	Passed	Passed	Failed	
RRM2	Passed	Passed	Passed	Failed	
PRMT16	Passed	Passed	Passed	Failed	
pS6	N/A	N/A			Yes
SMAD4	N/A	N/A			Yes
CCND1	N/A	N/A			Yes
pPRAS40	N/A	N/A			Yes
PTEN	N/A	N/A			Yes
SPP1	N/A	N/A			Yes

N/A, not applicable.

Table S2. Performance-based biomarker ranking: aggressiveness. Combinations of up to five biomarkers were generated and tested for their ability to predict severe disease (aggressiveness). The frequency of each biomarker in the best models was used for ranking.

Sort by AIC		**Sort by Test**	
YBX1	70.80	ACTN1	99.94
CUL2	65.72	FUS	34.18
ACTN1	44.09	SMAD2	26.13
AKAP8	20.74	CUL2	25.00
SMAD2	17.43	DIABLO	21.59
DEC1	16.37	HSPA9	20.79
DIABLO	15.29	PLAG1	20.4
CD75	15.12	DERL1	17.42
FUS	14.31	PDSS2	16.21
HOXB13	14.17	AKAP8	14.94
PLAG1	14.02	VDAC1	14.08
HSPA9	13.76	HOXB13	12.84
PDSS2	13.29	CD75	11.93
EIF3H	11.93	LATS2	10.44
PXN	11.65	HSD17B4	10.00
DERL1	11.20	DEC1	9.40
LATS2	10.80	LMO7	9.20
pS6	10.51	YBX1	9.18
pPRAS40	10.38	MTDH2	8.76
HSD17B4	10.33	CDKN1B	8.67
MAOA	9.10	PXN	8.65
FAK1	8.71	SMAD4	8.49
VDAC1	8.35	EIF3H	8.48
FKBP5	8.14	CCND1	8.39
MTDH2	7.45	COX6C	8.37
MAP3K5	7.42	pS6	8.28
CCND1	7.25	FKBP5	8.22
LMO7	7.14	pPRAS40	6.97
COX6C	6.50	MAOA	6.81
CDKN1B	6.13	MAP3K5	6.75

SMAD4	5.57	FAK1	5.83

Table S3. Performance-based biomarker ranking: lethal outcome. Combinations of up to five markers were generated and tested for their ability to predict lethal outcome (lethality). The frequency of each biomarker in the best models was used for ranking.

Sort by AIC (%)		**Sort by Test (%)**	
ACTN1	95.20	ACTN1	97.55
PLAG1	41.99	PLAG1	40.62
MTDH2	37.97	MTDH2	32.79
DERL1	21.86	HOXB13	29.65
HOXB13	20.76	DERL1	16.26
CD75	17.49	PDSS2	16.18
PDSS2	16.69	CD75	15.56
FAK1	16.19	COX6C	13.85
FUS	12.99	FAK1	13.30
AKAP8	12.39	FUS	12.72
COX6C	11.51	AKAP8	11.97
SMAD4	11.06	CUL2	11.29
MAP3K5	10.90	pS6	10.96
pS6	10.25	EIF3H	10.04
LMO7	10.20	CCND1	9.62
FKBP5	9.97	DIABLO	9.41
CUL2	9.67	YBX1	9.36
EIF3H	9.57	HSPA9	9.32
VDAC1	9.54	pPRAS40	9.27
CDKN1B	9.28	HSD17B4	9.26
MAOA	9.23	LATS2	9.21
pPRAS40	8.95	SMAD4	9.16
YBX1	8.90	PXN	9.08
HSPA9	8.78	CDKN1B	9.06
DEC1	8.76	MAP3K5	8.84
DIABLO	8.63	DEC1	8.78
SMAD2	8.29	LMO7	8.77
LATS2	8.24	SMAD2	8.46
CCND1	8.24	MAOA	8.33
HSD17B4	7.95	FKBP5	8.12
PXN	7.81	VDAC1	7.54

Table S4. Antibody sources.

Protein	H and L TMAs	Source	Cat #	Clonality	Host	Clone ID
CDKN1B (p27kip1)	Yes	Epitomics	1591-1	Mono	Rabbit	Y236
FUS	Yes	Epitomics/Abcam	5321-1/ab133571	Mono	Rabbit	EPR5813
CUL2	Yes	Invitrogen	700179	Mono	Rabbit	50H17L12
SMAD2	Yes	Invitrogen	700048	Mono	Rabbit	31H15L54
StAR	Yes	Santa Cruz	sc-166821	Mono	Mouse	D-2
DERL1	Yes	Sigma	SAB4200148	Mono	Mouse	Derlin1-1
ST6GAL1 (CD75)	Yes	Novus	NB100-78091	Mono	Mouse	LN1
BHLHE40 (Dec1)	Yes	Santa Cruz	sc-101023	Mono	Mouse	S-8
EIF3H	Yes	Cell Signaling	3413	Mono	Rabbit	
AKAP8	Yes	Epitomics	6620-1	Mono	Rabbit	EPR8978(B)
HOXB13	Yes	Santa Cruz	sc-28333	Mono	Mouse	F-9
NCOA2 (clone 29)	Yes	Santa Cruz	81280	Mono	Mouse	
GRIP-1	Yes	Santa Cruz	136244	Mono	Mouse	clone29
PXN	Yes	Epitomics	1500-1	Mono	Rabbit	Y113
ACTN1	Yes	Santa Cruz	sc-17829	Mono	Mouse	H-2
FKBP5	Yes	Epitomics	5532-1	Mono	Rabbit	EPR6617
PDSS2	Yes	Abcam	ab119768	Mono	Mouse	1D12
DIABLO	Yes	Epitomics	1012-1	Mono	Rabbit	Y12
COX6C	Yes	Santa Cruz	sc-65240	Mono	Mouse	3G5
HSD17B4	Yes	Santa Cruz	sc-365167	Mono	Mouse	A-6
YBX1	Yes	Epitomics/Abcam	2397-1/76149	Mono	Rabbit	EP2708Y
MAOA	Yes	Epitomics	5530-1	Mono	Rabbit	EPR7101
MAP3K5	Yes	Epitomics	1772-1	Mono	Rabbit	EP553Y
VDAC1	Yes	Santa Cruz	sc-58649	Mono	Mouse	20B12
LATS2	Yes	Abcam	ab54073	Mono	Mouse	
PLAG1	Yes	Sigma	SAB1404215	Mono	Mouse	
AGPAT6	Yes	Sigma/Protein Tech	SAB1403460/16762-1-AP	Mono	Mouse	
EZH2	Yes	Cell Signaling	5246	Mono	Rabbit	DC29
DCC (HSPA9)	Yes	Leica(Novocastra)	NCL-DCC	Mono	Mouse	DM51
FAK1	Yes	Epitomics	2146-1	Mono	Rabbit	EP1831Y
LMO7	Yes	Santa Cruz	sc-365515	Mono	Mouse	C-5
MTDH2	Yes	Epitomics	3674-1	Mono	Rabbit	EP4445
AGK	Yes	Santa Cruz	sc-374390	Mono	Mouse	F-3
pS6 (POC)	Yes	Epitomics/Abcam	2268-1/ab157359	Mono	Rabbit	EP1338(2)Y
SMAD4 (POC)	Yes	Santa Cruz	sc-7966	Mono	Mouse	B-8

CCND1 (POC)	Yes	Spring Bio	M3044	Mono	Rabbit	SP4
pPRAS40 (POC)	Yes	Cell Signaling	2997	Mono	Rabbit	C77D7
PTEN (POC)	Yes	Cell Signaling	9188	Mono	Rabbit	D4.3
SPP1 (POC)	Yes	Abcam	ab91655	Mono	Rabbit	EPR3688

Table S5. Cell-line controls. The cell lines listed were included as samples on the TMA to provide positive controls for the antibodies used. (Dox = doxycycline)

Cell line	shRNA knockdown or treatment
DU145	None
PC-3	None
WM266-4	None
RPMI7951	None
BxPC-3	None
RWPE-1	None
SK-MEL-5	None
DU145	SMAD4 knockdown; 0 µg/ml Dox
DU145	LY-treated for 1 hour
DU145	SMAD4 knockdown; 1 µg/ml Dox
PC-3	LY treated for 1 hour
DU145	CCND1 knockdown; 0 µg/ml Dox
DU145	CCND1 knockdown; 1 µg/ml Dox

Table S6. siRNA sequences used for antibody validation. siRNAs were used to reduce expression of the expected targets of the antibodies used to detect biomarkers. Sequences for the siRNAs used in validation are given.

Gene name	Gene ID	Cell line	Catalog no.	siRNA sequences	Antibody source
ACTN1	87	HeLa	LQ-011195	si5: GAGACAGCCGACACAGAUA	Santa Cruz
				si6: UGACUUACGUGUCUAGCUU	sc-17829
				si7: GAACUGCCCGACCGGAUGA	
				si8: GAAUACGGCUUUUGACGUG	
CUL2	8453	HeLa	LQ-007277	si5: GGAAGUGCAUGGUAAAUUU	Invitrogen
				si6: CAUCCAAGUUCAUAUACUA	700179
				si7: GCAGAAAGACACACCACAA	
				si8: UGGUUUACCUCAUAUGAUU	
Derlin1	79139	DU145	LQ-010733	si9: GGGCCAGGGCUUUCGACUU	Sigma
				si11: CAACAAUCAUAUUCACGUU	SAB4200148
FUS	2521	A375	LQ-009497	si7: GAUCAAUCCUCCAUGAGUA	Epitomics
				si10: GAGCAGCUAUUCUUCUUAU	5321-1
PDSS2	57107	HeLa	LQ-018550	si5: GGAAGAGAUUUGUGGAUUA	Abcam
				si6: GGCCAGAUCUGCUUUAGAA	ab119768
				si7: GAAUAUGGCAUUUCAGUAU	
				si8: GAAGAUUGGACUAUGCUAA	
SMAD2	4087	HeLa	LQ-003561	si5: GAAUUGAGCCACAGAGUAA	Invitrogen
				si6: GGUUUACUCUCCAAUGUUA	700048
				si7: UCAUAAAGCUUCACCAAUC	
				si8: ACUAGAAUGUGCACCAUAA	
VDAC1	7416	A549	LQ-019764	si5: UAACACGCGCUUCGGAAUA	Abcam
				si6: GAAACCAAGUACAGAUGGA	ab139752
				si7: GAGUACGGCCUGACGUUUA	
				si8: CCUGAUAGGUUUAGGAUAC	
YBX1	4904	A375	LQ-010213	si6: CUGAGUAAAUGCCGGCUUA	Epitomics
				si7: CGACGCAGACGCCCAGAAAA	2397-1
				si8: GUAAGGAACGGAUAUGGUU	
				si9: GCGGAGGCAGCAAAUGUUA	
DCC	1630	A549	LQ-003880	si6: GGAAGCAACUUACGGAUAC	Leica
				si7: GAUUCUGGCUCAAUUAUUA	NCL-DCC

				si8: GAAGUCAGAUGAAGGCUUU	
				si9: GUGAACAAAUGGGAAGUUU	
HSPA9	3313	HeLa	LQ-004750	si9: GGAAUGGCCUUAGUCAUGA	Santa Cruz
				si10: CCAAUGGGAUAGUACAUGU	sc-13967
				si11: CCUAUGGUCUAGACAAAUC	
SMAD4	4089				Santa Cruz
					sc-7966
pS6					Epitomics
					2268-1
			D-001810-01	NT1: ON-TARGETplus Non-targeting siRNA1	
			D-001810-02	NT2: ON-TARGETplus Non-targeting siRNA2	