Title

RNA biomarkers associated with metastatic progression in prostate cancer: a multi-institutional high-throughput analysis of SChLAP1

Journal

The Lancet Oncology, 15(13)

ISSN

1470-2045

Authors

Prensner, John R
Zhao, Shuang
Erho, Nicholas
et al.

Publication Date

2014-12-01

DOI

10.1016/s1470-2045(14)71113-1

Peer reviewed

HHS Public Access

Author manuscript

Lancet Oncol. Author manuscript; available in PMC 2015 December 01.

Published in final edited form as:

Lancet Oncol. 2014 December ; 15(13): 1469–1480. doi:10.1016/S1470-2045(14)71113-1.

Nomination and validation of the long noncoding RNA *SChLAP1* as a risk factor for metastatic prostate cancer progression: a multi-institutional high-throughput analysis

John R. Prensner, MD[1,13], **Shuang Zhao, MS**[2,13], **Nicholas Erho, MS**[3], **Matthew Schipper, PhD**[2], **Matthew K. Iyer, PhD**[1], **Saravana M. Dhanasekaran, PhD**[1,4], **Cristina Magi-Galluzzi, MD**[5], **Rohit Mehra, MD**[1,4,11], **Anirban Sahu, BS**[1], **Javed Siddiqui, MS**[1,4], **Elai Davicioni, PhD**[3], **Robert B. Den, MD**[7], **Adam P. Dicker, MD**[7], **R. Jeffrey Karnes, MD**[6], **John T. Wei, MD**[10], **Eric A. Klein, MD**[8], **Robert B. Jenkins, MD**[9], **Arul M. Chinnaiyan, MD**[1,4,10,11,12], and **Felix Y. Feng, MD**[1,2,11]

[1]Michigan Center for Translational Pathology, University of Michigan Medical School, Ann Arbor, Michigan, USA 48109

[2]Department of Radiation Oncology, University of Michigan Medical School, Ann Arbor, Michigan, USA 48109

[3]GenomeDx Biosciences Inc., Vancouver, British Columbia, Canada BC V6B 1B8

[4]Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan, USA 48109

Address correspondence to: Felix Y. Feng, M.D., Department of Radiation Oncology, University of Michigan Medical Center, 1500 E Medical Center Drive, UHB2C490-SPC5010, Ann Arbor, MI 48109-5010, Phone: 734-936-4302; Fax: 734-763-7371, ffeng@med.umich.edu.

[13]These authors contributed equally

Author Contributions

Conception and design: JRP, SZ, AMC, FYF
Development of methodology: SZ, JRP, MKI, NE, FYF
Acquisition of data: JRP, SZ, NE, CMG, RJK, APD, SMD, JTW, RM, AS, FYF
Analysis and interpretation of data: SZ, MKI, JRP, MS, FYF
Administrative, technical, or material support: ED, RBJ, RBD, APD, JTW, JS, EAK, FYF
Study supervision: JRP, SZ, AMC, FYF
Manuscript writing: All authors
Final approval of the manuscript: All authors

Declaration of Interests

JRP, MKI, and AMC hold a patent on Noncoding RNA and Uses Thereof with royalties paid by GenomeDx Biosciences Inc, and a patent Noncoding RNA and Uses Thereof licensed to Warfergen Inc. SZ, RJK, and FYF received funds for travel, accommodations, or expenses from GenomeDx Biosciences Inc. NE and ED are employees of GenomeDx Biosciences Inc. NE has a patent Cancer Diagnostics Using Non-coding Transcripts pending and a patent on Cancer Diagnostics Using Biomarkers pending. RBJ holds a patent Cancer Diagnostics Using Biomarkers with royalties by GenomeDx Biosciences Inc. CMG, RBD, and EAK received research grants from GenomeDx Biosciences Inc. RM, FYF, and EAK have served as consultants for GenomeDx Biosciences Inc. APD received honoraria from Merck EMD, Bayer, and other support from NRG Oncology and the NCI cooperative group. AMC received research grants from the Department of Defense, National Institutes of Health, and Howard Hughes Medical Institute. AMC is a co-founder and advisor to Compendia Bioscience, part of ThermoFisher. AMC served on the Scientific Advisory Board of Wafergen Inc. The other authors declared no conflicts of interest.

[5]Anatomic Pathology, Robert J. Tomsich Pathology and Laboratory Medicine Institute, Cleveland Clinic, Cleveland, Ohio, USA 44195

[6]Department of Radiation Oncology, Thomas Jefferson University, Philadelphia, Pennsylvania, USA 19107

[7]Department of Urology, Mayo Clinic, Rochester, Minnesota, USA 55902

[8]Glickman Urological and Kidney Institute, Cleveland Clinic, Cleveland, Ohio, USA 44195

[9]Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minnesota, USA 55902

[10]Department of Urology, University of Michigan Medical School, Ann Arbor, Michigan, USA 48109

[11]Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, Michigan, USA 48109

[12]Howard Hughes Medical Institute, University of Michigan Medical School, Ann Arbor, Michigan, USA 48109

Abstract

Background—Improved clinical predictors for disease progression are needed for localized prostate cancer, where only a minority of patients experience poor outcomes. We undertake an unbiased large-scale analysis of genes associated with aggressive clinical course.

Methods—Prostate cancer samples, obtained from patients treated with radical prostatectomy at three academic institutions, were analyzed for gene expression using a clinical-grade, high-density Affymetrix GeneChip platform, encompassing >1 million genomic loci. Nomination of prognostic candidate genes was performed on a discovery cohort (n=545) and validated on three independent cohorts (n=463), totaling 1,008 patients. Molecular assays were performed in a CLIA-certified (Clinical Laboratory Improvement Amendments) laboratory facility. Multivariate analyses were performed for the primary endpoint of metastasis. The top prostate-specific gene was evaluated in urine samples from 230 patients using PCR.

Findings—Among all known genes, the long noncoding RNA *SChLAP1* ranked first for elevated expression in patients with metastatic progression by receiver-operator-curve (ROC) area-under-the-curve (AUC) analyses. Of the top five prognostic genes, *SChLAP1* was the only prostate-specific gene. Validation in three independent cohorts confirmed the prognostic value of *SChLAP1* for metastasis. On multivariate modeling, *SChLAP1* expression independently predicted metastasis within ten years (odds ratio (OR) = 2·45, 95% confidence interval (CI) 1·70 – 3·53), death within ten years (OR = 1·93, 95% CI 1·31 – 2·85), and biochemical recurrence within five years (OR = 1·76, 95% CI 1·28 – 2·41) with odds ratios comparable to Gleason score. Evaluation of *SChLAP1* expression in 230 urine sediment samples with either biopsy-confirmed cancer or biopsy-negative tissue demonstrated increased incidence and expression of *SChLAP1* RNA in patients at a higher risk for disease progression.

Interpretation—We perform the largest high-throughput, unbiased study of prostate cancer prognostic biomarkers to date and discover *SChLAP1* as one of the best genes for the prediction of

metastasis. We validate *SChLAP1* extensively using a clinical-grade assay in a CLIA-certified laboratory. We show feasibility of a non-invasive urine test for *SChLAP1*, and suggest that *SChLAP1* represents a very promising biomarker for aggressive clinical course.

Funding—Prostate Cancer Foundation, National Institutes of Health, Department of Defense, Early Detection Research Network, Doris Duke Charitable Foundation, and Howard Hughes Medical Institute.

Keywords

prostate cancer; long noncoding RNA; metastasis; prognosis

Introduction

While a majority of localized prostate cancer patients harbor slow-growing, non-lethal tumors, a smaller fraction of patients experience disease recurrence following definitive first-line therapies, which may lead to metastasis and death.[1–3] To distinguish between aggressive and indolent tumors, current clinical paradigms rely mainly on pre-operative prostate specific antigen (PSA) levels, tumor stage, and biopsy Gleason score, which assesses cancer cell histology, in order to estimate patient risk.[4] Yet, these remain imperfect tools that inaccurately classify some patients.[5,6] Moreover, except for PSA-derived tests, previous prognostic biomarkers, such as Ki67 and TOP2A,[7,8] are expressed non-specifically throughout the body and require invasive sampling of tumor tissue via biopsy or prostatectomy, greatly limiting their use in early-stage disease. Other, more recent tests such as urine measurement of the long noncoding RNA (lncRNA) *PCA3* have demonstrated utility in the diagnosis of prostate cancer but not in risk stratification.[9] Thus, characterization of novel prognostic biomarkers, especially ones suitable for non-invasive detection, represents an important research focus for improving patient management.

Advances in high-throughput technologies have now enabled unbiased biomarker discovery approaches. Microarray-based technologies facilitated discovery of ETS gene fusions, *AMACR* overexpression, and other biological subgroups in prostate cancer.[10] Recently, we and others have used next-generation sequencing to define lncRNAs as potential biomarkers in prostate cancer.[11,12] Our work led to the analysis of the lncRNA *SChLAP1* as an oncogenic factor in prostate cancer that associates with poor patient outcomes.[13] Yet, while prior efforts have nominated prognostic genes,[14,15] no unbiased studies have been performed to identify genes correlated with long-term outcomes such as metastasis. Furthermore, there are no validated non-invasive biomarkers for early-stage disease that predict outcome.

Here, we undertake an analysis of 1,008 prostate cancer samples using unbiased approaches to define RNA biomarkers associated with metastatic progression. Using a discovery cohort and three validation cohorts, we show that the lncRNA *SChLAP1* is highly associated for metastasis in multiple cohorts, and *SChLAP1* expression significantly enhances patient risk stratification when combined with other clinicopathological covariates. *SChLAP1* expression was detectable non-invasively in urine samples and associated with higher-risk patients. Finally, addition of *SChLAP1* to established clinical prognostic tools improved the

performance of all tools evaluated. Overall, our study establishes *SChLAP1* as a novel biomarker whose prognostic capacity significantly adds to that of established risk factors.

Methods

Study Design and Tissue Samples

Banked or archived tumor samples were obtained from three prostatectomy patient cohorts enrolled at the Mayo Clinic or the Cleveland Clinic under informed consent protocols approved by local Institutional Review Board (IRB). Analyses were designed in accordance with REMARK criteria.[16] The Mayo Clinic I (MCI) cohort included 212 patients with metastatic progression and a total of 333 patients without metastatic progression (Figure S1) as described.[17] For the Mayo Clinic II (MCII) patients, a case-cohort study design was employed to randomly sample 20% (202/1,010) of patients for analysis, in addition to all who developed metastases, from a cohort of 1,010 high-risk men who underwent radical prostatectomy between 2000–2006 as described (Figure S2).[18] The MCII cohort and its outcomes data represent a modified set of patients overlapping with our previous report of *SChLAP1*with more stringent data quality control filters.[13]

Patients from the Cleveland Clinic (CC) were obtained from a case-control study design sampled from 2,317 conservatively treated radical prostatectomy patients with high-risk features who received no adjuvant or neo-adjuvant therapy from 1987–2008. Patients were sampled to achieve a three:one ratio for non-metastatic (n=134) versus metastatic progression (n=49) patients (Figure S3). The CC cohort has not been previously published. Additional information on cohorts and sample preparation is provided in the Supplementary Methods.

Patient cohorts were designed in accordance with STROBE recommendations for case-control and case-cohort studies.[19]

Microarray Hybridization and Gene Expression

RNA extraction from formalin-fixed paraffin embedded (FFPE) samples and microarray hybridization was performed for Mayo Clinic and Cleveland Clinic samples using clinical-grade techniques in a CLIA-certified (Clinical Laboratory Improvement Amendments) laboratory facility (GenomeDx Biosciences, Inc, San Diego, CA). CLIA certification was obtained through the Centers for Medicare and Medicaid Services (CMS) through standard procedures, and laboratory facilities satisfied all criteria required by the CMS for certification. Details regarding the CLIA requirements can be found online at: http://www.cms.gov/Regulations-and-Guidance/Legislation/CLIA/index.html (last accessed October 17, 2014). RNA purification, hybridization to Affymetrix Human Exon (HuEx) 1.0 ST GeneChips, and gene expression calculations are detailedin the Supplementary Methods. Partition-Around-Medoids unsupervised clustering was used to define expression subgroups in the MCI cohort. This expression threshold was applied to the other cohorts without additional modification. Microarray data are available on the NCBI Gene Expression Omnibus (GEO) as accession numbers GSE46691 (MCI), GSE62116 (MCII), and GSE62667 (CC).

Additional Raw Datasets

We obtained raw Affymetrix HuEx 1.0 ST GeneChip expression data and sample clinical information for Boormans et al.[20] (Erasmus Medical Center; EMC) from the NCBI Gene Expression Omnibus (GSE41408). Gene expression was calculated as above.

Nomination of metastasis-associated genes

We calculated the median expression of each gene in patients experiencing metastasis versus non-metastatic patients. Fold expression change was calculated with the following formula: ($\log_2$(median_expression_metastatic)−$\log_2$(median_expression_no_metastasis)) / ($\log_2$(median_expression_all_samples) + four). The constant four was used to uniformly increase all expression values above zero to avoid a negative denominator.

Urine quantitative PCR

Urine samples were collected from 256 patients with informed consent. Samples were collected following a digital rectal exam at the time of PSA screening. All patients subsequently received a needle biopsy at the University of Michigan with IRB approval as described.[21] RNA processing and quantitative PCR for *KLK3*, *GAPDH*, *TMPRSS2-ERG*, *PCA3*, and *SChLAP1* was performed as described.[21] Data quality control, normalization, and expression calculation were performed according to standard parameters (Supplementary Methods). A total of 230 urine samples passed quality control metrics and were included for data analysis. PCR primers are listed in Table S10.

Outcomes

The primary endpoint of metastatic progression after prostatectomy was defined as a positive CT scan or bone scan. Biochemical recurrence was defined based on the original study protocol for each of the cohorts. In the Mayo Clinic I cohort, biochemical recurrence was defined as two successive increases in PSA measurements above 0.2 ng/mL, with the subsequent measure at least 0.05 ng/mL above the first measurement. In the Mayo Clinic II cohort, biochemical recurrence was defined as a follow-up PSA measurement of 0.4 ng/mL or greater 30 days after prostatectomy. In the Cleveland Clinic cohort, biochemical recurrence was defined as a follow-up PSA concentration greater than 0.2 ng/mL or initiation of salvage therapy.

Statistical Analyses

Fisher's exact test and logistic regression models were used to analyze the relation between each of the three clinical outcomes and clinical factors and biomarkers. In analyses of the primary endpoint, men followed for less than ten years, and who did not have an event during follow-up, were excluded. In MCI, the original definition of cases and controls were used for fold change and AUC calculation.[22] The association between *SChLAP1* and clinical outcomes was assessed separately for each study and overall in a single logistic regression model stratified by study. Multivariate analyses were performed to assess whether *SChLAP1* was able to increase the predictive ability of standard clinical factors. All significant clinical covariates were included in the multivariate models. Details on statistical analyses are found in the Supplementary Methods. A p-value<0·05 was considered statistically significant.

Kaplan-Meier curves and weighted Cox regression comparing time to metastases between groups defined by *SChLAP1* expression are shown only for the MCII case-cohort study and utilize the weighting method as described previously.[23] The case-control study design of MCI and CC cohorts allows for assessment of relative, but not absolute, incidence of events. Time to event data for EMC was not available. Nonparametric AUC values (equivalent to C statistics) were calculated separately for each study. Overall values were calculated as the weighted average of the study-specific values with weights proportional to sample size. Testing for improved AUC value between the full model without *SChLAP1* and the full model with *SChLAP1* was done using the likelihood ratio test for *SChLAP1* in the full model with *SChLAP1*.

Role of the funding source

The funding sources had no role in the study design, data collection, analysis, interpretation, manuscript writing, or manuscript submission. N.E., S.Z., J.R.P., E.D., and F.Y.F. had access to the raw data. The corresponding author had full access to all of the data and the final responsibility to submit for publication.

Results

Study design

We designed a retrospective biomarker discovery analysis according to REMARK criteria[16] in which prostate cancer patients who developed metastases were compared to those who did not (Figure 1A). We employed 1,008 radical prostatectomy specimens from three academic institutions, comprising four independent patient cohorts. Three cohorts represented case-control study designs; one study was a case-cohort design. Patients were defined as high-risk for recurrence (e.g., pT2 tumor with positive margins or pT3 disease) by current clinical guidelines. The study designs for the MCI, MCII, and CC cohorts further enriched for patients who experienced metastasis (see Methods). For these three main cohorts, a total of 639 (MCI), 256 (MCII), and 197 (CC) patients were selected for analysis, of which 94, 24, and 14, respectively, were excluded due to sample unavailability, poor sample quality, or poor microarray data quality (MCI, n=545 included; MCII, n=232 included; CC, n=183 included; see Figures S1–3). Clinical characteristics of included patients are detailed in Table 1. All cohorts with available information had mean patient follow-up between seven and fourteen years.

A clinical-grade microarray platform, which contains 5 million probes against 1.4 million unique probeset regions (PSRs), was used to measure global gene expression in an unbiased fashion. Tissue samples from three of four cohorts (not for EMC) were processed in a CLIA-certified laboratory, representing 95% (960/1,008) of specimens. We analyzed all known protein-coding genes and lncRNAs previously identified in prostate cancer (PCATs).[11] We used metastasis as the primary endpoint. Whereas localized and locally-recurrent disease is potentially curable, metastatic disease is incurable, requiring intensive treatment such as next generation anti-androgens and chemotherapy, and frequently progresses to mortality.[3,24]

Nomination of SChLAP1 by unbiased expression profiling

Using the MCI cohort (n=545), we performed a global assessment of gene expression differences between tumors from patients who experienced metastasis (n=212) and those who did not (n=333). Mean follow-up was 14 years. We derived median expression values for all genes in each group and compared the relative change in expression between groups. Surprisingly, the top-ranked gene was *SChLAP1* (Figure 1B, q=0·0012, Table S1), which was recently characterized as an oncogenic prostate cancer lncRNA (see Figure S4 for *SChLAP1* PSRs).[13] Overall, there were 230 genes whose expression associated with metastasis at a false discovery rate (FDR) or q≤0·01 (Table S1).

SChLAP1 demonstrated the largest gene expression change between tumors with and without metastatic progression (Figure 1B, Figure S5, Table S1). High *SChLAP1* expression was associated with a higher risk for BCR, metastasis, death from prostate cancer, and death from any cause at ten years post-prostatectomy (p=0·044, p<0·0001, p<0·0001, p<0·0001, respectively, Fisher's exact) (Figure 1C–F).

Validation of SChLAP1

For initial validation of *SChLAP1*, we employed a case-cohort MCII set (n=232) of high-risk localized prostate cancer patients (Table 1, Figure S6) who underwent radical prostatectomy. We observed that *SChLAP1* was again a powerful predictor of time to BCR, metastatic progression, and prostate cancer-specific mortality (p=0·0021, p=0·00016, p=0·0044, respectively, Cox model), with a strong trend for significance in predicting worse overall survival (p=0·066) (Figure 1G–J). Kaplan-Meier curves for all patients in the MCII cohort without stratification by *SChLAP1* status are shown in Figure S7.

We next incorporated data from a third independent cohort of radical prostatectomy tissues from high-risk patients at the CC (n=183, Table 1). We processed the CC data using the same statistical approach as for the Mayo cohorts (Figures S8). Confirming our prior observations, we found a strong association between *SChLAP1* expression and metastatic progression in the CC set (OR=3·1, p=0·021, Figure 1K). To ensure reproducibility of our data, we further experimentally confirmed that expression of *SChLAP1* using PCR in a subset of samples, which demonstrated a high inter-assay correlation for expression levels (Pearson's correlation = 0·75, p=0·00015, Figure S9).

Lastly, we searched for additional publicly-available cohorts with clinical annotation that used the Affymetrix HuEx platform and reported a ≥10% metastasis event rate for statistical robustness. We found one cohort from the Erasmus Medical Center (EMC, n=48)[20] and processed these data for *SChLAP1* expression as above (Table 1). Here, *SChLAP1* expression was again highly associated with metastases (p=0·0022, Figure 1L), with all metastatic events occurring in patients with high *SChLAP1* expression (Table S2). Together, these four datasets represent 1,008 patients, and all cohorts support a strong association between *SChLAP1* expression and metastasis.

A global comparison of SChLAP1 to other genes

To compare *SChLAP1* to other genes, we measured the receiver-operator-curve (ROC) area-under-the-curve (AUC) metric for metastatic disease progression across all annotated protein-coding genes and PCATs using the MCI and MCII cohorts. These cohorts were most enriched for high-risk patients and adverse outcomes. We plotted the AUC values for both cohorts for the top 1,000 genes (Figure 2A, Table S3, Supplementary Methods), of which a small minority displayed substantially higher AUC values in both cohorts (Figure 2A box). A focused analysis of the top genes defined *SChLAP1* as the second best single-gene predictor of metastasis (Figure 2A, right).

Among the top five prognostic genes, *SChLAP1* is the only gene to demonstrate prostate-specific expression, ideal for development as a non-invasive biomarker (Figure 2B).[25,26] Using a large compendium of RNA-Seq data, including >500 samples from >30 tissue types with >13 cancer types,[27] we observed high levels of *SChLAP1* expression only in prostate cancer, with no or minimal expression in other tissues (Figure 2B, Table S4). These data establish *SChLAP1* expression as specific to prostate cancer, which is unique among the top prognostic genes and suggests that *SChLAP1* may represent a more promising non-invasive biomarker compared to other known genes which are more universally expressed.

Clinicopathological associations

Using all four cohorts (MCI, MCII, CC, EMC), we found that *SChLAP1* expression was also significantly associated with other clinical risk factors for aggressive disease. *SChLAP1* expression was significantly associated with extracapsular extension, seminal vesicle invasion, and positive surgical margin status ($p<0{\cdot}05$, Fisher's exact, Figure S10, Table S2). *SChLAP1* performed favorably compared to PSA for the discrimination of metastatic progression of patients (AUC=0.68 for *SChLAP1*, AUC=0.56 for PSA, Figure S11).

Multivariate analyses

We next evaluated the association between *SChLAP1* expression and standard clinicopathological factors with clinical outcomes in univariate and multivariate analyses. Here, we excluded the EMC dataset because this cohort did not indicate the time to metastatic events so that we could not calculate the primary endpoint of metastases within ten years.

SChLAP1 was a significant predictor of the primary endpoint of metastasis in all three cohorts (MCI, MCII, CC) when stratifying by Gleason score and PSA. A pooled analysis for these cohorts showed a highly significant association of *SChLAP1* expression with metastatic progression within ten years of prostatectomy by odds ratio ($p<0{\cdot}0001$) (Figure 3). For secondary endpoints, *SChLAP1* was also significant for BCR within five years ($p<0{\cdot}0001$) and overall survival at ten years ($p<0{\cdot}0001$) (Figure S12). Because our cohorts emphasized high-risk patients, we evaluated BCR at five years since most recurrences were early events.

We then constructed a multivariate model representing all 960 patients from the Mayo and CC cohorts and incorporating all significant clinical factors in addition to *SChLAP1* (Table

2, see Table S5 for the univariate analysis). Strikingly, *SChLAP1* retains its utility as an independent prognostic variable for metastatic progression within ten years post-prostatectomy (OR=2.45 (95% CI 1.70 – 3.53), p<0·0001) (Table 2). *SChLAP1* expression was also an independent covariate for BCR within 5 years (p=0·00044) and death from any cause within 10 years (p=0·00096). Notably, the *SChLAP1* odds ratios were comparable in magnitude to Gleason score. *SChLAP1* remained highly significant for metastasis, BCR and death from any cause when we added significant and non-significant variables on univariate to the model (Table S6). We further confirmed this finding by analyzing each cohort independently (Table S7) and by evaluating all covariates (both significant and non-significant) as continuous variables in the multivariate analysis (Table S8). These investigations demonstrated highly significant p values for *SChLAP1* in all variations of the analyses.

Non-invasive detection of SChLAP1 in urine sediments

Next, we sought to evaluate *SChLAP1* in prostate cancer patients early in their disease course. We employed a University of Michigan cohort of 230 patient urine sediments[21] obtained post-digital rectal examination at the time of PSA screening for asymptomatic men. All men subsequently received a diagnostic prostate biopsy to determine whether cancer was present. Although urine sediments also contain bladder cells, *SChLAP1* expression is specific to prostate cells (Figure 2B).

We then measured *SChLAP1* expression in our cohort of urine sediment samples (Table S9). We observed high *SChLAP1* expression only in a subset of patients (Figure 4A), which is consistent with the *SChLAP1* expression profile seen in all previous tissue cohorts (Figures S5,6,8).[13] When integrated with measurement of the lncRNA biomarker *PCA3* and the *TMPRSS2-ERG* gene fusion, *SChLAP1* was able to identify 8% (11/141) of cancers missed by these other two tests in this cohort (Figure S13), although *SChLAP1* was less sensitive overall. Among patients with biopsy-confirmed cancer, expression of *SChLAP1* was both more frequent and more highly elevated in Gleason 7 patients compared to Gleason 6 patients (p=0·029, Fisher's exact, Figure 4B, Figure S14A). We were unable to evaluate Gleason ≥8 due to low numbers of patients. Finally, we stratified patients into low, intermediate, and high-risk categories according to standard PSA and Gleason thresholds. We found that *SChLAP1* expression was significantly elevated in intermediate and high risk patients compared to low risk patients (p=0·0022, Mann Whitney U test, Figure 4C, Figure S14B, C) and was able to effectively discriminate between these two groups of patients via ROC analyses (AUC=0·68, Figure S14D). These data provide proof-of-principle analyses that *SChLAP1* expression is detectable non-invasively in prostate cancer patient urine samples. However, additional validation is needed to confirm the clinical utility of a urine-based *SChLAP1* test.

SChLAP1 expression improves established clinical tools

Lastly, we interrogated the role of *SChLAP1* in clinical decision making through its utility in conjunction with the Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S) score, a validated prognostic tool incorporating clinicopathological parameters,[28,29] the Decipher test, a clinically validated genomic classifier comprised of a prognostic gene

signature,[30,31] or the cell cycle progression (CCP) expression signature.[32] To do this, we integrated *SChLAP1* into these models using standard analyses.[33] First, we added *SChLAP1* to the CAPRA-S tool on the Mayo Clinic I, Mayo Clinic II, and Cleveland Clinic cohorts and evaluated whether the addition of *SChLAP1* improved its prognostic power. We observed a statistically significant increase in prognostic utility with an AUC of 0·69 for CAPRA-S alone improved to 0·74 with the addition of *SChLAP1* (p=0·015, Figure S15). Second, we added *SChLAP1* to the Decipher genomic classifier and we determined that addition of *SChLAP1* led to a statistically significant increase in the prognostic potential of Decipher (p=0·048, Figure S16). Finally, we observed that addition of *SChLAP1* to the genes of the CCP signature also significantly enhanced the prognostic power of that signature using the HuEx array data (p=0·00027, Figure S17). Please note that the CCP signature was originally measured using PCR (not the HuEx array),[32] and therefore we have not formally recapitulated the Polaris assay, which is based on the CCP score. Taken together, these three analyses lend further evidence that measurement of *SChLAP1* is able to provide additional prognostic information in conjunction with the existing, established tools and assays used for clinical and molecular risk stratification.

Discussion

Here, we perform the largest biomarker discovery project to date in prostate cancer, employing over 1,000 patients with one discovery cohort and three validation cohorts. Notably, our use of high-density microarrays enables broad surveillance of non-coding transcripts not measured with conventional microarrays. Important aspects of this study include our ability to: (1) nominate and validate using a high-throughput clinical-grade assay in a CLIA-certified laboratory; (2) use distant metastasis as the primary endpoint; and (3) systematically evaluate non-coding elements in the transcriptome. Our use of metastasis as a primary endpoint is highly clinically significant, since biochemical recurrence is not an accurate reflection of systemic disease progression. Indeed, all prior studies have either used biochemical recurrence as the primary endpoint or evaluated patients who had already experienced recurrence.[14,15,17,34] Biochemical recurrence is weaker and dependent on clinical factors, such as surgical margins, as well as biological factors. Finally, we employ a clinical-grade assay performed in a CLIA-certified laboratory, making our results highly translatable for clinical practice.

We find that *SChLAP1* is one of the best genes for predicting metastatic progression, and is highly tissue specific, unlike most other prognostic biomarkers. This association between *SChLAP1* and metastasis is robust across multiple independent cohorts with both pooled and individual univariate and multivariate analyses incorporating standard clinical risk factors. We note that *SChLAP1* is robust in these analyses when only significant covariates as well as when all covariates are incorporated into our multivariate model. Remarkably, the odds ratio of *SChLAP1* positivity on multivariate analysis is comparable to Gleason score, which has remained the single best predictor of metastatic disease since its description over 40 years ago.[3,35] We note that our multivariate analysis incorporated patients who both received (MCI, MCII) and did not receive (CC) adjuvant therapy; we believe that this reflects a strength of our model as *SChLAP1* remains strongly significant regardless, suggesting that the advantageous effect of adjuvant therapy on patient outcomes does not

eliminate the prognostic significance of *SChLAP1*. Moreover, we are unaware of any other RNA that independently predicts for prostate cancer metastasis, although Ki67 immunohistochemistry does.[8] While the *PCA3* urine RNA assay predicts biopsy status and histopathological characteristics,[36,37] it does not predict outcomes such as recurrence and metastasis.

More broadly, the goal of the present work identifies *SChLAP1* as a potential adjunct test in order to stratify patient risk for adverse outcomes at an early clinical stage. The fact that *SChLAP1* expression functions as an independent risk factor for metastatic progression and overall survival lends clinical utility to *SChLAP1* measurement as a potential clinical predictor that could be evaluated along standard clinical parameters such as PSA and Gleason score during the early evaluation and staging of prostate cancer patients. As such, our work has uniquely looked at time to metastasis as a primary endpoint, but has also found that *SChLAP1* expression associates with a higher risk of death from any cause, which may be more clinically insightful.

Our previous work on *SChLAP1* also warrants attention. We initially nominated *SChLAP1* as an outlier (based on gene expression only) in prostate cancer[11] and observed that cancers with *SChLAP1* expression associated with worse clinical outcomes.[13] Yet independent validation, multivariate analyses, and global comparisons to all other genes were not conducted at that time. In this regard, our current study is both an extension and an expansion of our prior work in order to more fully define novel biomarkers for prostate cancer aggressiveness.

One essential aspect of our findings is the fact that *SChLAP1* expression is specific to prostate cancer, with minimal expression in all other tumor and tissue types. This is a striking contrast to other biomarkers (e.g. Ki67, TOP2A), which have non-specific expression patterns and thus require immunohistochemistry staining on biopsy or prostatectomy tissues to evaluate protein abundance in specific cell types. *SChLAP1*, therefore, is uniquely suitable as a non-invasive biomarker, whereas all other prognostic genes are not appropriate for non-invasive detection.

To this end, we show proof-of-principle data of *SChLAP1* in patient urine samples. Here, we demonstrate that *SChLAP1* expression in the urine associates with a Gleason 7 over Gleason 6 histology. Although it would be optimal to compare urine *SChLAP1* expression with long-term outcomes (i.e. metastasis), such clinical information is not available for our urine cohort at this time. We have therefore used Gleason score as a proxy measure, recognizing that the majority of Gleason 7 patients do not experience disease progression after prostatectomy and therefore a low sensitivity of *SChLAP1* for Gleason 7 histology may suggest that *SChLAP1* expression identifies only the subset of Gleason 7 patients who will have progressive disease. *SChLAP1* was further able to discriminate effectively between low risk patients and patients with intermediate or high-risk features defined by Gleason score and screening serum PSA values.

Because *SChLAP1* is intended as a non-invasive prognostic biomarker, as opposed to a diagnostic biomarker, our studies have not focused on sensitivity and specificity of urine

SChLAP1 expression for the diagnosis of prostate cancer, as prior *PCA3* urine studies have.[9] Thus, we have provided initial evidence that *SChLAP1* expression may complement existing urine diagnostic assays, including *PCA3* and *TMPRSS2-ERG*, and that clinical application of a *SChLAP1* urine test would be most effective in conjunction with these, and potentially other, urine assays. Ongoing urine studies evaluating the concordance between urine *SChLAP1* expression and tissue-based measurement of *SChLAP1* expression will also be important. Although formalized optimization of urine biomarker assays requires substantial investment and resources, we are encouraged by these data and argue that prioritization of *SChLAP1* during future biomarker development studies may be appropriate.

We further find that *SChLAP1* is able to improve upon established clinical algorithms for the risk stratification of prostate cancer. Specifically, *SChLAP1* improves the CAPRA-S score,[29,30] which is one of the best clinicopathological models to date. *SChLAP1* further improves upon both the Decipher test[15,22] and the CCP gene signature.[32] Thus, we believe that *SChLAP1* has the potential in prostate cancer to advance the cause of "precision medicine", which has been pioneered in breast cancer with clinical prognostic tests such as the OncotypeDx and Mammaprint gene expression platforms. We emphasize that, unlike the OncotypeDx and Mammaprint assays in breast cancer, molecular biomarkers have not been routinely integrated into clinicopathological models in prostate cancer, and so our analysis of *SChLAP1* with the CAPRA-S score is particularly revealing. Thus, we anticipate that *SChLAP1* will be most effective in conjunction with other prognostic tools, such the CAPRA-S, or other gene-based assays incorporating multiple markers, such as Decipher or the CCP, in order to provide the most accurate prognosis for patients with prostate cancer.

Regarding cellular biology, *SChLAP1* is a lncRNA, which is an emerging class of RNA molecules that do not encode for a protein.[38] We have previously shown that *SChLAP1* is an essential mediator of aggressive disease processes, including tumor invasion and hematogenous spread.[13] *SChLAP1* operates through transcriptional regulation via antagonism of the SWI/SNF epigenetic complex,[13] which is responsible for the positioning of histone proteins at gene promoters.[39] SWI/SNF is a well-defined tumor suppressor in numerous cancer types, including prostate cancer,[40] and is inactivated by genetic mutation or deletion of core subunits.[39] By disrupting SWI/SNF function, *SChLAP1* contributes to the altered expression of hundreds-to-thousands of genes,[13] which may facilitate the metastatic cascade globally rather than through a single signaling pathway, potentially enhancing early castrate resistance and risk of mortality. Although selected SWI/SNF-associated proteins have been suggested to promote prostate cancer proliferation,[41,42] it is unclear whether these proteins are functioning in conjunction with, or independently of, endogenous core SWI/SNF enzymatic subunits.

Lastly, this study has limitations. We perform extensive retrospective studies, but not prospective studies. We also have not evaluated *SChLAP1* in the context of androgen-deprivation therapy or radiotherapy. In addition, our urine PCR assay for *SChLAP1* is a preliminary analysis which does not qualify as a clinical-grade test according to established criteria. Finally, because of the extremely good outcomes for localized prostate cancer patients, we used case-control and case-cohort studies to enrich patient cohorts for

individuals with progressive disease. Thus, we emphasized high-risk patients, and additional studies aimed at other patient populations are warranted.

Overall, our work provides compelling evidence that *SChLAP1* expression is highly prognostic and an independent risk factor for metastasis. We show that non-invasive detection of *SChLAP1* in urine is feasible. As such, this study provides insight into the pathogenesis of aggressive prostate cancer and identifies *SChLAP1* as a potential new clinical biomarker for metastatic progression. We hope that measurement of *SChLAP1* expression becomes a component of future prospective clinical trials, and we believe that prostate cancer patients with high *SChLAP1* expression may be appropriate for clinical trials intensifying adjuvant therapies.

Research in context

Systematic review

Many expression profiling studies for prognostic prostate cancer biomarkers have been performed.[14,15,17,34] A PubMed search for “expression profiling”, “prostate cancer”, and “biomarker” yields 77 primary research publications. All studies used biochemical recurrence as the primary endpoint. No study used a clinical-grade assay in a CLIA-certified laboratory. No study evaluated long noncoding RNAs systematically.

Interpretation

Our study was able to: (1) nominate and validate a biomarker using a high-throughput clinical-grade assay in a CLIA-certified laboratory; (2) use distant metastasis as a primary endpoint; (3) systematically evaluate non-coding elements in the transcriptome. With over 1,000 patients, our study is also the largest biomarker discovery project to date in prostate cancer. Our use of metastasis as a primary endpoint is highly clinically significant, since biochemical recurrence is not an accurate reflection of systemic disease progression, with less than half of patients with biochemical recurrence experiencing metastasis.[3] There are currently no established biomarkers for prostate cancer metastasis for early stage tumors nor non-invasive assays for long-term aggressive outcomes. The fact that *SChLAP1* is tissue-specific is highly significant and enables the development of non-invasive urine tests. As such, our study provides highly novel findings of direct clinical importance with a pathway for rapid clinical translation of the findings.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding Sources

The Prostate Cancer Foundation (PCF) Young Investigator Award (J.R.P., R.M.), the PCF-Movember Challenge Award (F.Y.F.), NIH Prostate Specialized Program of Research Excellence grant P50CA69568, Department of Defense grant PC100171 (A.M.C.), the Early Detection Research Network grant UO1 CA111275 (A.M.C.),the US National Institutes of Health R01CA132874-01A1 (A.M.C.), the Doris Duke Charitable Foundation (A.M.C.) and the Howard Hughes Medical Institute (A.M.C.).

References

1. Etzioni R, Cha R, Feuer EJ, Davidov O. Asymptomatic incidence and duration of prostate cancer. American Journal of Epidemiology. 1998; 148:775–85. [PubMed: 9786232]
2. Makarov DV, Humphreys EB, Mangold LA, Carducci MA, Partin AW, Eisenberger MA, et al. The natural history of men treated with deferred androgen deprivation therapy in whom metastatic prostate cancer developed following radical prostatectomy. J Urol. 2008; 179:156–61. discussion 61–2. [PubMed: 18001801]
3. Pound CR, Partin AW, Eisenberger MA, Chan DW, Pearson JD, Walsh PC. Natural history of progression after PSA elevation following radical prostatectomy. JAMA. 1999; 281:1591–7. [PubMed: 10235151]
4. Epstein JI. An update of the Gleason grading system. J Urol. 2010; 183:433–40. [PubMed: 20006878]
5. D'Amico AV, Chen MH, Roehl KA, Catalona WJ. Identifying patients at risk for significant versus clinically insignificant postoperative prostate-specific antigen failure. J Clin Oncol. 2005; 23:4975–9. [PubMed: 16051949]
6. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC, et al. Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. JAMA. 2005; 294:433–9. [PubMed: 16046649]
7. Malhotra S, Lapointe J, Salari K, Higgins JP, Ferrari M, Montgomery K, et al. A tri-marker proliferation index predicts biochemical recurrence after surgery for prostate cancer. PLoS One. 2011; 6:e20293. [PubMed: 21629784]
8. Berney DM, Gopalan A, Kudahetti S, Fisher G, Ambroisine L, Foster CS, et al. Ki-67 and outcome in clinically localised prostate cancer: analysis of conservatively treated prostate cancer patients from the Trans-Atlantic Prostate Group study. Br J Cancer. 2009; 100:888–93. [PubMed: 19293807]
9. Roobol MJ, Schroder FH, van Leeuwen P, Wolters T, van den Bergh RC, van Leenders GJ, et al. Performance of the prostate cancer antigen 3 (PCA3) gene and prostate-specific antigen in prescreened men: exploring the value of PCA3 for a first-line diagnostic test. Eur Urol. 2010; 58:475–81. [PubMed: 20637539]
10. Tomlins SA, Rubin MA, Chinnaiyan AM. Integrative biology of prostate cancer progression. Annu Rev Pathol. 2006; 1:243–71. [PubMed: 18039115]
11. Prensner JR, Iyer MK, Balbin OA, Dhanasekaran SM, Cao Q, Brenner JC, et al. Transcriptome sequencing across a prostate cancer cohort identifies PCAT-1, an unannotated lincRNA implicated in disease progression. Nat Biotechnol. 2011; 29:742–9. [PubMed: 21804560]
12. Du Z, Fei T, Verhaak RG, Su Z, Zhang Y, Brown M, et al. Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. Nat Struct Mol Biol. 2013; 20:908–13. [PubMed: 23728290]
13. Prensner JR, Iyer MK, Sahu A, Asangani IA, Cao Q, Patel L, et al. The long noncoding RNA SChLAP1 promotes aggressive prostate cancer and antagonizes the SWI/SNF complex. Nat Genet. 2013; 45:1392–8. [PubMed: 24076601]
14. Henshall SM, Afar DE, Hiller J, Horvath LG, Quinn DI, Rasiah KK, et al. Survival analysis of genome-wide gene expression profiles of prostate cancers identifies new prognostic targets of disease relapse. Cancer Res. 2003; 63:4196–203. [PubMed: 12874026]
15. Glinsky GV, Glinskii AB, Stephenson AJ, Hoffman RM, Gerald WL. Gene expression profiling predicts clinical outcome of prostate cancer. J Clin Invest. 2004; 113:913–23. [PubMed: 15067324]
16. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. REporting recommendations for tumour MARKer prognostic studies (REMARK). Br J Cancer. 2005; 93:387–91. [PubMed: 16106245]
17. Nakagawa T, Kollmeyer TM, Morlan BW, Anderson SK, Bergstralh EJ, Davis BJ, et al. A tissue biomarker panel predicting systemic progression after PSA recurrence post-definitive prostate cancer therapy. PLoS One. 2008; 3:e2318. [PubMed: 18846227]

18. Karnes RJ, Bergstralh EJ, Davicioni E, Ghadessi M, Buerki C, Mitra AP, et al. Validation of a genomic classifier that predicts metastasis following radical prostatectomy in an at risk patient population. J Urol. 2013; 190:2047–53. [PubMed: 23770138]
19. Vandenbroucke JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. PLoS Med. 2007; 4:e297. [PubMed: 17941715]
20. Boormans JL, Korsten H, Zielvan der Made AJ, van Leenders GJ, de Vos CV, Jenster G, et al. Identification of TDRD1 as a direct target gene of ERG in primary prostate cancer. Int J Cancer. 2013; 133:335–45. [PubMed: 23319146]
21. Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R, et al. A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. Cancer Res. 2008; 68:645–9. [PubMed: 18245462]
22. Erho N, Crisan A, Vergara IA, Mitra AP, Ghadessi M, Buerki C, et al. Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. PLoS One. 2013; 8:e66855. [PubMed: 23826159]
23. Lin DY, Ying Z. Cox regression with incomplete covariate measurements. Journal of the American Statistical Association. 1993; 88:1341–9.
24. Ferraldeschi R, Sharifi N, Auchus RJ, Attard G. Molecular pathways: Inhibiting steroid biosynthesis in prostate cancer. Clin Cancer Res. 2013; 19:3353–9. [PubMed: 23470964]
25. Zhou B, Yuan T, Liu M, Liu H, Xie J, Shen Y, et al. Overexpression of the structural maintenance of chromosome 4 protein is associated with tumor de-differentiation, advanced stage and vascular invasion of primary liver cancer. Oncol Rep. 2012; 28:1263–8. [PubMed: 22842912]
26. Akhtar J, Wang Z, Yu C, Li CS, Shi YL, Liu HJ. STMN-1 is a potential marker of lymph node metastasis in distal esophageal adenocarcinomas and silencing its expression can reverse malignant phenotype of tumor cells. BMC Cancer. 2014; 14:28. [PubMed: 24433541]
27. Kalyana-Sundaram S, Kumar-Sinha C, Shankar S, Robinson DR, Wu YM, Cao X, et al. Expressed pseudogenes in the transcriptional landscape of human cancers. Cell. 2012; 149:1622–34. [PubMed: 22726445]
28. Punnen S, Freedland SJ, Presti JC Jr, Aronson WJ, Terris MK, Kane CJ, et al. Multi-institutional validation of the CAPRA-S score to predict disease recurrence and mortality after radical prostatectomy. Eur Urol. 2014; 65:1171–7. [PubMed: 23587869]
29. Cooperberg MR, Hilton JF, Carroll PR. The CAPRA-S score: A straightforward tool for improved prediction of outcomes after radical prostatectomy. Cancer. 2011; 117:5039–46. [PubMed: 21647869]
30. Cooperberg MR, Davicioni E, Crisan A, Jenkins RB, Ghadessi M, Karnes RJ. Combined Value of Validated Clinical and Genomic Risk Stratification Tools for Predicting Prostate Cancer Mortality in a High-risk Prostatectomy Cohort. Eur Urol. 2014
31. Ross AE, Feng FY, Ghadessi M, Erho N, Crisan A, Buerki C, et al. A genomic classifier predicting metastatic disease progression in men with biochemical recurrence after prostatectomy. Prostate Cancer Prostatic Dis. 2014; 17:64–9. [PubMed: 24145624]
32. Cuzick J, Swanson GP, Fisher G, Brothman AR, Berney DM, Reid JE, et al. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. Lancet Oncol. 2011; 12:245–55. [PubMed: 21310658]
33. Etzioni R, Kooperberg C, Pepe M, Smith R, Gann PH. Combining biomarkers to detect disease with application to prostate cancer. Biostatistics. 2003; 4:523–38. [PubMed: 14557109]
34. Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci U S A. 2004; 101:811–6. [PubMed: 14711987]
35. Cheng L, Koch MO, Juliar BE, Daggy JK, Foster RS, Bihrle R, et al. The combined percentage of Gleason patterns 4 and 5 is the best predictor of cancer progression after radical prostatectomy. J Clin Oncol. 2005; 23:2911–7. [PubMed: 15860849]
36. Nakanishi H, Groskopf J, Fritsche HA, Bhadkamkar V, Blase A, Kumar SV, et al. PCA3 molecular urine assay correlates with prostate cancer tumor volume: implication in selecting candidates for active surveillance. J Urol. 2008; 179:1804–9. discussion 9–10. [PubMed: 18353398]

37. Whitman EJ, Groskopf J, Ali A, Chen Y, Blase A, Furusato B, et al. PCA3 score before radical prostatectomy predicts extracapsular extension and tumor volume. J Urol. 2008; 180:1975–8. discussion 8–9. [PubMed: 18801539]

38. Prensner JR, Chinnaiyan AM. The emergence of lncRNAs in cancer biology. Cancer Discov. 2011; 1:391–407. [PubMed: 22096659]

39. Wilson BG, Roberts CW. SWI/SNF nucleosome remodellers and cancer. Nat Rev Cancer. 2011; 11:481–92. [PubMed: 21654818]

40. Grasso CS, Wu YM, Robinson DR, Cao X, Dhanasekaran SM, Khan AP, et al. The mutational landscape of lethal castration-resistant prostate cancer. Nature. 2012; 487:239–43. [PubMed: 22722839]

41. Link KA, Balasubramaniam S, Sharma A, Comstock CE, Godoy-Tundidor S, Powers N, et al. Targeting the BAF57 SWI/SNF subunit in prostate cancer: a novel platform to control androgen receptor activity. Cancer Res. 2008; 68:4551–8. [PubMed: 18559499]

42. van de Wijngaart DJ, Dubbink HJ, Molier M, de Vos C, Trapman J, Jenster G. Functional screening of FxxLF-like peptide motifs identifies SMARCD1/BAF60a as an androgen receptor cofactor that modulates TMPRSS2 expression. Mol Endocrinol. 2009; 23:1776–86. [PubMed: 19762545]

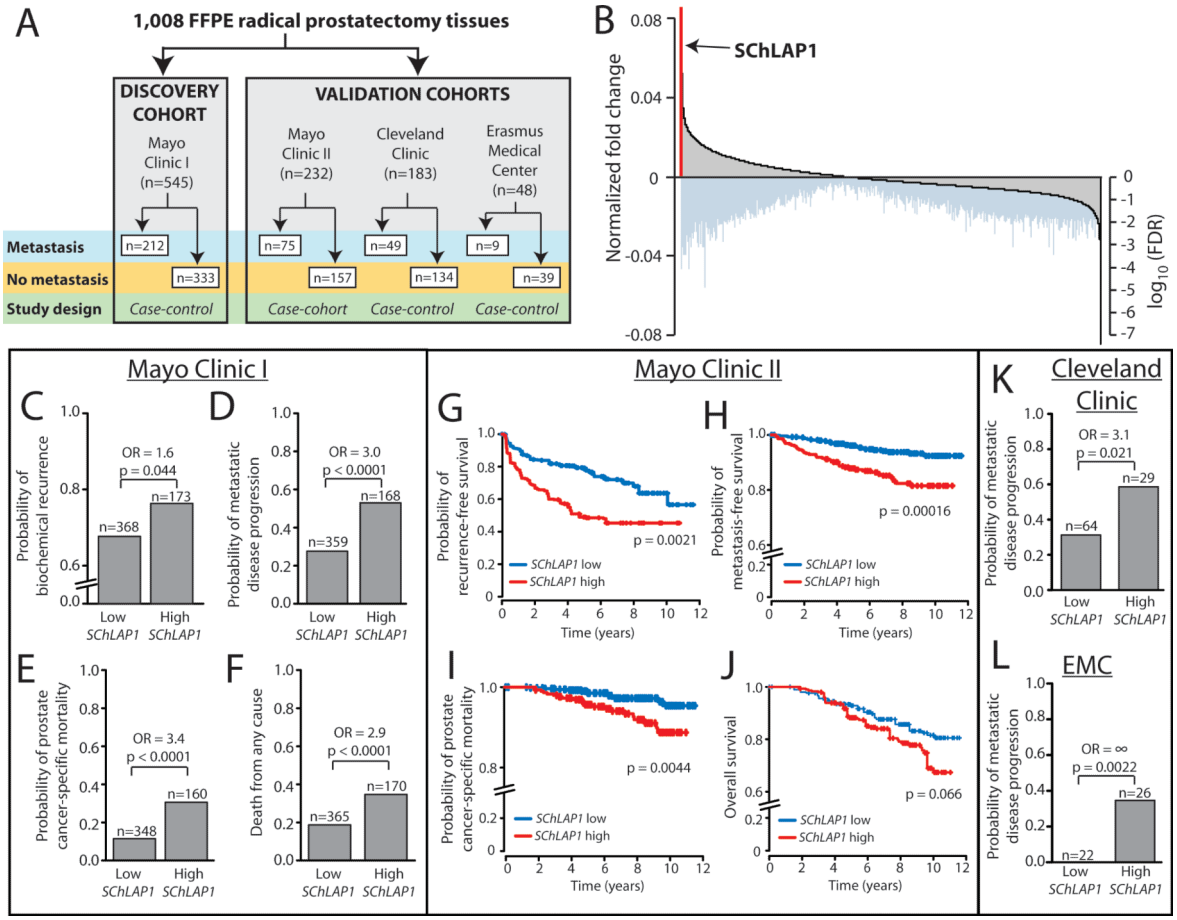

Figure 1. Nomination and validation of *SChLAP1* as a top-ranked prognostic gene

(A) A schematic overview of the patient specimens and cohort study designs employed in this analysis.**(B)** A global view of gene expression changes associated with metastatic progression. In the Mayo Clinic I cohort (n=545), gene expression was determined with Affymetrix Exon microarrays and differential expression analysis was performed for patients who experienced metastatic progression compared to those who did not. Ranking all genes according to the fold change of expression between metastatic and non-metastatic samples nominates *SChLAP1* as the top-ranked outlier gene associated with metastatic progression. Log_{10} false discovery rate (FDR) values for each corresponding gene are displayed below. **(C–F)** Patient outcomes in the Mayo Clinic I cohort (n=545) were stratified by *SChLAP1* expression for biochemical recurrence **(C)**, progression to metastatic disease **(D)**, prostate cancer-specific mortality **(E)**, and death from any cause **(F)**. P values in **C–F** were determined by a two-tailed Fisher's exact test.**(G–J)** Patient outcomes in the Mayo Clinic II cohort (n=232) were stratified by *SChLAP1* expression for biochemical recurrence **(G)**, progression to metastatic disease **(H)**, prostate cancer-specific mortality **(I)**, and overall survival **(J). (K, L)** Patient outcomes for metastasis were stratified by *SChLAP1* expression in the Cleveland Clinic **(K)** and Erasmus Medical Center **(L)** datasets. P values in **K–L** were determined by a two-tailed Fisher's exact test.

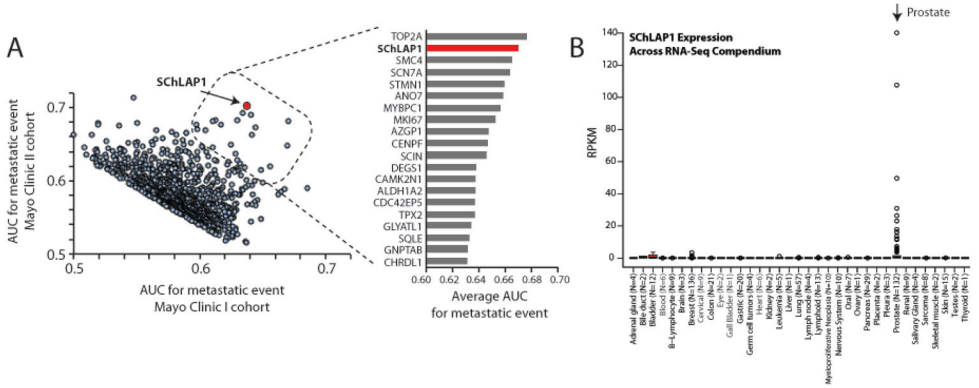

Figure 2. A global analysis of *SChLAP1* in the Mayo Clinic I and II cohorts

(A) The Mayo Clinic I and II cohorts were independently analyzed to determine the receiver-operator-curve (ROC) area under the curve (AUC) metric for all genes for the development of metastatic disease. This global analysis generated independent AUC values for each cohort. AUC values for the top 1000 genes in both cohorts were plotted (*left*). A small subset of genes demonstrated superior AUC values in both cohorts (*see outlined box*). A detailed analysis of these top-20 genes is plotted (*right*) using the averaged AUC value between both cohorts. *SChLAP1* ranks #2 overall for prediction of metastatic spread. **(B)** Expression of *SChLAP1* across 552 samples, representing 35 cancer and tissue types. Expression was determined by RNA-Seq and *SChLAP1* level is represented as reads per kilobase per million mapped reads (RPKM).

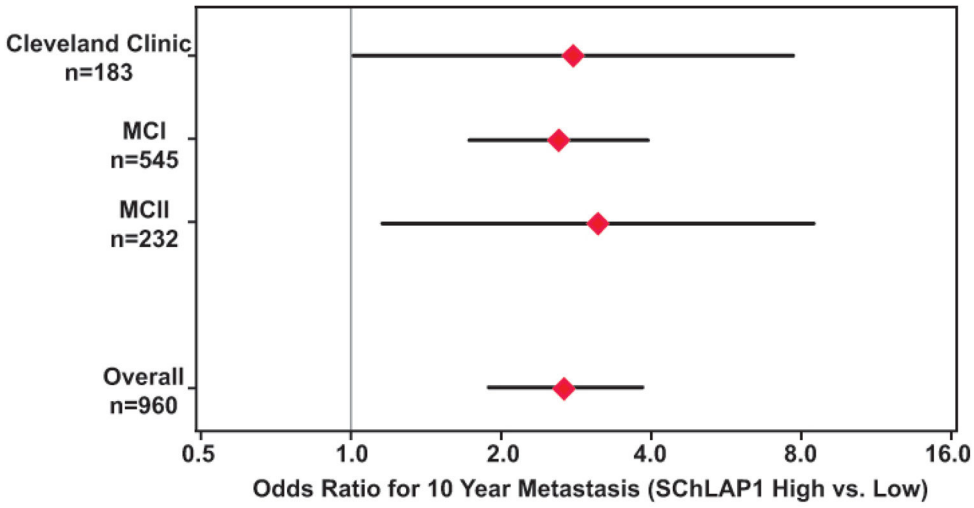

Figure 3. Univariate analysis of *SChLAP1* in three cohorts

A forest plot for individual univariate analyses as well as a pooled analysis of all cohorts for metastatic progression at 10 years post-prostatectomy. The odds ratio for patient outcomes based on *SChLAP1* expression is shown (red diamond, odds ratio; black bar, 95% confidence interval). For the overall odds ratio in this figure, data were calculated from a logistic regression model, p<0·0001.

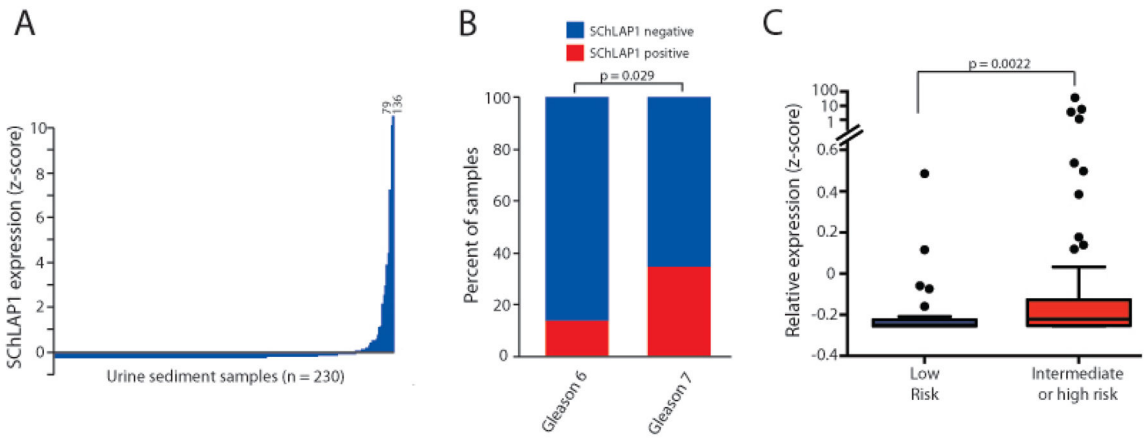

Figure 4. Detection of *SChLAP1* in patient urine samples

(A) Detection of *SChLAP1* RNA in patient urine sediments. Samples are ordered according to *SChLAP1* expression. Expression is represented as the z-score. **(B)** The fraction of Gleason 6 (n=44) and Gleason 7 (n=49) urine sediments that demonstrate positive *SChLAP1* expression. P value was determined by a two-tailed Fisher's exact test. **(C)** *SChLAP1* expression in urine sediments from low risk (n=37) and intermediate/high risk patients (n=68). P value was determined by a Mann Whitney U test.

TABLE 1

Cohort clinical characteristics

	Mayo Clinic I (n=545)	Mayo Clinic II (n=232)	Cleveland Clinic (n =183)	Erasmus Medical Center (n = 48)
Age (Years, mean ± SD)	65·3 ± 6·4	63·1 ± 7·4	61·6 ± 6·3	NA
Follow-up (Months, mean ± SD)	160·7 ± 56·2	80·6 ± 30·1	116·6 ± 50·1	NA
Metastatic progression				
No	333 (61%)	157 (68%)	134 (73%)	39 (81%)
Yes	212 (39%)	75 (32%)	49 (27%)	9 (19%)
Pre-operative PSA				
<10	282 (52%)	126 (54%)	127 (69%)	21 (44%)
10 to 20	117 (22%)	62 (27%)	41 (23%)	17 (35%)
>20	131 (24%)	44 (19%)	12 (7%)	8 (17%)
Not available	15 (3%)	0 (0%)	3 (1%)	2 (4%)
Gleason score				
5	0 (0%)	0 (0%)	0 (0%)	0 (0%)
6	60 (11%)	17 (7%)	25 (17%)	23 (48%)
7	271 (49%)	117 (50%)	113 (62%)	16 (33%)
8	68 (13%)	39 (17%)	23 (13%)	8 (17%)
9	134 (24%)	57 (25%)	22 (12%)	1 (2%)
10	9 (2%)	1 (1%)	0 (0%)	0 (0%)
Not available	3 (1%)	1 (1%)	0 (0%)	0 (0%)
Tumour stage				
I	0 (0%)	0 (0%)	0 (0%)	0 (0%)
II	219 (40%)	97 (42%)	0 (0%)	16 (33%)
III	253 (46%)	102 (44%)	0 (0%)	26 (54%)
IV	0 (0%)	0 (0%)	0 (0%)	6 (13%)
Not available	73 (13%)	33 (14%)	183 (100%)	0 (0%)
Extracapsular extension				
Negative	272 (50%)	136 (59%)	51 (28%)	0 (0%)
Positive	273 (50%)	96 (41%)	132 (72%)	0 (0%)
Not available	0 (0%)	0 (0%)	0 (0%)	48 (100%)
Seminal vesicle invasion				
Negative	369 (68%)	149 (64%)	152 (83%)	0 (0%)
Positive	176 (32%)	83 (36%)	31 (17%)	0 (0%)
Not available	0 (0%)	0 (0%)	0 (0%)	48 (100%)
Lymph node invasion				
Negative	472 (87%)	199 (86%)	183 (100%)	0 (0%)
Positive	73 (13%)	33 (14%)	0 (0%)	0 (0%)
Not available	0 (0%)	0 (0%)	0 (0%)	48 (100%)
Surgical margin status				
Negative	279 (51%)	99 (43%)	92 (50%)	0 (0%)

	Mayo Clinic I (n=545)	Mayo Clinic II (n=232)	Cleveland Clinic (n =183)	Erasmus Medical Center (n = 48)
Positive	266 (49%)	133 (57%)	91 (50%)	0 (0%)
Not available	0 (0%)	0 (0%)	0 (0%)	48 (100%)

TABLE 2

Multivariate analyses

	Biochemical recurrence at 5 years		Metastatic progression at 10 years		Death from any cause at 10 years	
	Odds Ratio (± 95% CI)	**P value**	**Odds Ratio (± 95% CI)**	**P value**	**Odds Ratio (± 95% CI)**	**P value**
SChLAP1 (high vs. low)	1·76 (1·28 – 2·41)	0·00044	2·45 (1·70 – 3·53)	<0·0001	1·93 (1·31 – 2·85)	0·00096
Gleason score	1·56 (1·33 – 1·82)	<0·0001	2·14 (1·77 – 2·58)	<0·0001	2·06 (1·69 – 2·51)	<0·0001
PSA	1·01 (0·99 – 1·02)	0·061	1·01 (0·99 – 1·02)	0·26	1·00 (0·99 – 1·01)	0·89
Seminal vesicle invasion	1·76 (1·25 – 2·49)	0·0013	1·44 (0·96 – 2·16)	0·074	1·79 (1·16 – 2·75)	0·0083
Surgical margin status	1·38 (1·03 – 1·86)	0·030	Not included		1·28 (0·87 – 1·89)	0·21
Extracapsular extension	1·20 (0·87 – 1·63)	0·27	1·08 (0·74 – 1·58)	0·69	1·05 (0·69 – 1·61)	0·81
Lymph node invasion	Not included		1·14 (0·65 – 2·01)	0·64	1·31 (0·74 – 2·32)	0·35