Clinical Cancer Research

Combination of a Novel Gene Expression Signature with a Clinical Nomogram Improves the Prediction of Survival in High-Risk Bladder Cancer

Markus Riester, Jennifer M. Taylor, Andrew Feifer, et al.

Updated Version	Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-11-2271
Supplementary Material	Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2012/01/05/1078-0432.CCR-11-2271.DC1.html

Cited Articles	This article cites 36 articles, 12 of which you can access for free at: http://clincancerres.aacrjournals.org/content/18/5/1323.full.html#ref-list-1

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.

Imaging, Diagnosis, Prognosis

Combination of a Novel Gene Expression Signature with a Clinical Nomogram Improves the Prediction of Survival in High-Risk Bladder Cancer

Markus Riester[1], Jennifer M. Taylor[3], Andrew Feifer[3], Theresa Koppie[5], Jonathan E. Rosenberg[2], Robert J. Downey[4], Bernard H. Bochner[3], and Franziska Michor[1]

Abstract

Purpose: We aimed to validate and improve prognostic signatures for high-risk urothelial carcinoma of the bladder.

Experimental Design: We evaluated microarray data from 93 patients with bladder cancer managed by radical cystectomy to determine gene expression patterns associated with clinical and prognostic variables. We compared our results with published bladder cancer microarray data sets comprising 578 additional patients and with 49 published gene signatures from multiple cancer types. Hierarchical clustering was utilized to identify subtypes associated with differences in survival. We then investigated whether the addition of survival-associated gene expression information to a validated postcystectomy nomogram utilizing clinical and pathologic variables improves prediction of recurrence.

Results: Multiple markers for muscle invasive disease with highly significant expression differences in multiple data sets were identified, such as fibronectin 1 (FN1), NNMT, POSTN, and SMAD6. We identified signatures associated with pathologic stage and the likelihood of developing metastasis and death from bladder cancer, as well as with two distinct clustering subtypes of bladder cancer. Our novel signature correlated with overall survival in multiple independent data sets, significantly improving the prediction concordance of standard staging in all data sets [mean ΔC-statistic: 0.14; 95% confidence interval (CI), 0.01–0.27; $P < 0.001$]. Tested in our patient cohort, it significantly enhanced the performance of a postoperative survival nomogram (ΔC-statistic: 0.08, 95% CI, −0.04–0.20; $P < 0.005$).

Conclusions: Prognostic information obtained from gene expression data can aid in posttreatment prediction of bladder cancer recurrence. Our findings require further validation in external cohorts and prospectively in a clinical trial setting. *Clin Cancer Res; 18(5); 1323–33. ©2012 AACR.*

Introduction

Urothelial carcinoma of the urinary bladder is the fifth most common cancer in the Western world, and the ninth most frequent cancer worldwide, representing 3% of cancers diagnosed globally (1). Treatment of bladder cancer has not changed significantly in more than 20 years, and outcomes for patients remain suboptimal. Approximately 20% to 30% of newly diagnosed patients present with muscle-invasive disease (stages T2–T4) or metastatic disease, whereas up to a third of patients with initially non–muscle-invasive disease (stages Ta/T1/Tis) later progress to muscle-invasive or metastatic disease (2). Clinical features including stage and grade are strongly associated with outcome and play an important role in determining treatment. For example, the 5-year tumor-specific mortality rates range from less than 5% for low-stage and low-grade disease up to approximately 50% for all muscle-invasive lesions (3, 4). However, even though grade and stage are important predictors of outcome, there remains significant variability in the prognosis of patients with similar characteristics. This highlights the need to identify additional tumor characteristics that predict clinical behavior.

In recent years, there has been a growing interest in the use of gene expression signatures for risk stratification of patients with cancer (5). Multiple groups have produced urothelial cancer gene signatures predictive of a range of

Authors' Affiliations: [1]Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute and Department of Biostatistics, Harvard School of Public Health; [2]Bladder Cancer Center, Dana-Farber Cancer Institute, Boston, Massachusetts; [3]Urology Service, [4]Thoracic Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, New York, New York; and [5]Department of Urology, University of California-Davis, Sacramento, California

M. Riester and J.M. Taylor contributed equally to this work.

Corresponding Author: Franziska Michor, Dana-Farber Cancer Institute, 450 Brookline Avenue, Boston, MA 02215. Phone: 671-632-5045; Fax: 617-632-4222; E-mail: michor@jimmy.harvard.edu

Translational Relevance

Urothelial carcinoma of the urinary bladder is the fifth most common cancer in the Western world, representing 3% of the global cancer incidence. Bladder cancers exhibit a heterogeneous clinical behavior highlighted by frequent recurrences in patients with noninvasive tumors and the potential for the development of metastatic disease in those with invasive lesions. Because of this variation, better tools to predict prognosis and refine treatment are needed. Here, we evaluated the utility and reproducibility of expression array differences in bladder tumors as potential prognostic indicators. Our analysis identified novel signatures associated with tumor stage, progression among patients with organ-confined disease, and molecular subtypes of tumors with relevance to disease-specific survival. Incorporation of our gene expression signatures, furthermore, significantly improved the prediction using a postoperative nomogram. Identifying patients at higher risk for progression or death provides additional rationale for selection of multimodality management, using chemotherapy and surgery, in an effort to improve survival.

Table 1. Clinicopathologic characteristics of the tumor samples

Patient age at RC, median (range), y	69.1 (32.1–91.1)
Gender, n (%)	
Female	25 (27)
Male	68 (73)
Histologic type (%)	
Transitional cell carcinoma (TCC)	88 (95)
TCC/squamous	5 (5)
RC stage, n (%)	
pTa	5 (5)
pT1	10 (11)
pT2	17 (18)
pT3	42 (45)
pT4	19 (21)
Lymph node status (%)	
pN0	49 (52)
pN+	28 (30)
pNx[a]	16 (18)
Length of follow-up, median (range), mo	32 (1–175)
Last known status, n (%)	
NED	28 (30)
Death from other cause	27 (29)
Death from bladder cancer	38 (41)
Recurrence, n (%)[b]	
No recurrence	52 (56)
Urothelial	9 (10)
Pelvic	20 (21)
Distant site	34 (37)

Abbreviation: RC, radical cystectomy.
[a]Lymph node dissection not done or no lymph nodes analyzed.
[b]Numbers total more than 100%, as some patients developed recurrence in more than 1 site.

tumor characteristics and outcomes, including stage (6–8), molecular subtype (9, 10), likelihood of recurrence and progression of non–muscle-invasive (11–13) and muscle-invasive disease (14), and survival (7). Here, we aimed the following: (i) to explore whether novel cancer gene expression patterns can be identified that stratify patients undergoing radical cystectomy for urothelial cancer by risk of recurrence and death; (ii) to test our novel signatures and all previously developed signatures in all available bladder cancer microarray data sets; and (iii) to investigate the gain in predictive accuracy of a validated clinical nomogram by combination with gene signature information.

Materials and Methods

MSKCC samples

Characteristics. We utilized previously unpublished cancer gene expression data from 93 patients undergoing radical cystectomy at Memorial Sloan-Kettering Cancer Center (MSKCC), New York, between 1993 and 2004. Specimens were collected with MSKCC Institutional Review Board approval, and a waiver of authorization to review associated clinical data was obtained from the Board. The clinicopathologic characteristics are summarized in Table 1 (see Supplementary Table S1 for details). In those 15 patients with non–muscle-invasive disease on final pathologic analysis after radical cystectomy, 4 had muscle-invasive disease histologically in tissue obtained at the time of prior transurethral resection (TUR) and the remaining patients had high-risk features (e.g., extensive volume of disease, recurrent, or Bacillus Calmette—Guerin (BCG)

refractory disease). Lymph node dissection was done in 77 patients; no patient had metastatic disease at the time of radical cystectomy. Chemotherapy was administered to 3 patients as neoadjuvant, 16 patients as adjuvant, and 19 patients as salvage for recurrent disease. Case selection was restricted to those with frozen specimens with measurable volume of malignancy and adequate percentage of tumor.

Clinical endpoints. Overall survival time was defined as the time from date of radical cystectomy to date of death or last follow-up. Disease-specific survival was defined as the time from radical cystectomy to death attributed to urothelial cancer, when death occurred with known and progressive metastatic disease. If death was recorded in the institutional database without knowledge of a recurrent cancer or with documentation of another malignancy or a non-malignancy cause, death was attributed to other causes. Recurrence was defined as pelvic if within the pelvis below the aortic bifurcation, distant if there was visceral metastasis or recurrence above the aortic bifurcation, and urothelial if within the remaining urinary tract (renal pelvis, ureter, and

urethra). In the analysis of the development of metastases, only pelvic and distance recurrences were included. Primary endpoint in this study was overall survival if not mentioned otherwise.

Sample preparation. All patient specimens collected at the time of surgery were processed expeditiously within the Department of Pathology and stored in an institutional biospecimen bank. Frozen bladder tissues were examined by a genitourinary pathologist to identify tumor content, which was microdissected. Ten 50-μm sections were cut, with confirmation of tumor content by pathologic review. RNA was isolated with TRIzol (Invitrogen) and cleaned with the RNeasy Mini Kits (Qiagen) according to manufacturer protocols. Expression profiling was conducted in the MSKCC Genomics Core Facility. In brief, 5 to 10 μg of RNA was transcribed utilizing a T7-oligo dT primer and converted to cDNA (Invitrogen cDNA synthesis kit). Biotinylated aRNA was produced from the cDNA by an *in vitro* transcription kit (Enzo Diagnostics). Following quality assessment using an Agilent Bioanalyzer, the aRNA was fragmented and hybridized onto Affymetrix U133 Plus 2.0 arrays, containing 54,675 probe sets. Data were uploaded to National Center for Biotechnology Information (NCBI) Geo (15) under accession number GSE31684.

External data sets

We obtained 6 independent bladder cancer microarray data sets (Table 2; refs. 6, 7, 9, 10, 13, 16) totaling 578 additional patients and encompassing a broad spectrum of stages, grades, and histologic types. Overall survival data were available for the Blaveri and colleagues, Sanchez-Carbayo and colleagues, Lindgren and colleagues, and Kim and colleagues cohorts (6, 7, 9, 13; see Supplementary data for preprocessing details). Disease-specific survival time information was only documented for patients in the

cohort from Kim and colleagues. The primary endpoint was, therefore, overall survival in the meta-analysis.

Statistical analysis

Preprocessing. Expression estimates were obtained by the GCRMA algorithm (17).

Differential gene expression. The gene expression profiles were evaluated for significant correlation with multiple clinical variables and outcomes, including stage, lymph node status, recurrence, and overall survival. For the identification of differentially expressed probe sets, we used the LIMMA method (18). If not stated otherwise, the cutoff value of the adjusted P values [the false discovery rate (FDR); ref. 19] was set to 0.01 and the minimum fold change (FC) to 1.5. The probability of observing n or more differentially expressed genes was estimated by permutation tests. For the external data sets, the lists of differentially expressed genes were downloaded from the Supplementary Material of the respective articles (6, 7, 9, 10, 13, 16) when available, otherwise calculated with LIMMA. When comparing lists of differentially expressed genes with corresponding lists from external data sets, the significance of overlaps was calculated in Bioconductor (GSEABase), using 25,000 permutations.

Machine learning. Classifiers were generated with the Fisher linear discriminant and support vector machines (SVM) and were leave-one-out cross-validated. The *R* packages MASS and e1071 were used for building the classifiers. We utilized a linear SVM kernel and default, untuned parameters. Receiver operating characteristic (ROC) curves were generated with the ROCR package (20).

Risk category creation. For each patient, a risk score was calculated by a Cox proportional hazards model that was fitted using the gene expression profiles of all remaining patient samples (leave-one-out cross-validation). The risk

Table 2. Independent bladder cancer microarray data sets

Data set	Accession number	Platform	No. of probes	No. of Samples (no. of invasive)	Median age (range)	% Males	% RC	Median follow-up time (max)	% Death (DOD)
Blaveri and colleagues (6)	GSE1827	cDNA microarray	10,368	80 (53)	66 (28–113)	70.0%	62.5	13 (145)	55.0 (n/a)
Sanchez-Carbayo and colleagues (7)	—	Affymetrix U133A	22,283	90 (65)	69 (37–85)	68.9%	n/a	25 (76)	40.0 (n/a)
Lindgren and colleagues (9)	GSE19915	cDNA microarray	2,506	144 (47)	n/a	n/a	32.6	46 (180)	18.1 (n/a)
Dyrskjøt and colleagues (10)	GSE89	Affymetrix Hu6800	7,129	40 (9)	n/a	n/a	n/a	n/a	n/a
Kim and colleagues (13)	GSE13507	Illumina human-6 v2.0	43,148	165 (62)	66 (24–88)	81.8%	n/a	37 (137)	41.8 (19.4)
Stransky and colleagues (16)	E-TABM-147	Affymetrix U95A/U95Av2	12,626	59 (33)	n/a	83%	n/a	n/a	n/a

Abbrevaiations: DOD, death of disease; n/a, not available; RC, radical cystectomy.

score was defined as the sum of the gene signature expression values, weighted by the Cox regression coefficients (21). Each individual was then classified into low-risk or high-risk categories, based on the median leave-one-out cross-validated risk score: patients were classified as low-risk when their risk score was smaller than the cohort median, otherwise as high-risk. This was done independently for all data sets for which survival information was available. The significance of survival curve differences of cross-validated models was estimated with permutation tests. The process from cross-validation to risk stratification was repeated 500 times with shuffled survival labels. This empirical χ^2 distribution was then utilized to estimate P values. For noncross-validated models, the log-rank test was used. Multivariate survival prediction models were compared by the C-statistic, an estimator of the model concordance (22) and by likelihood ratio tests. The concordance represents the probability that given 2 random noncensored individuals, the one with the higher risk score has a shorter survival time.

Clustering analysis. For unsupervised clustering, we used the Ward clustering method and the Pearson correlation distance as implemented in the pvclust R package (23). The significance of a cluster was reported as its bootstrap value, which is the proportion of 10,000 bootstrap samples showing this particular cluster topology. The clustering was applied to the 6 external data sets of bladder cancer samples.

Meta-analysis using published data sets

Published signatures. We compiled 49 published gene signatures (Supplementary Table S2) associated with malignancy: 39 from Lauss and colleagues (24) as well as other bladder (11–14) and melanoma signatures (25, 26); the melanoma signatures were included to test whether signatures from other solid tumors would conduct well in bladder cancer. The 49 gene signatures were tested in our data set and the 4 external gene expression data sets with survival information. Associations of gene expression signatures with outcome and other covariates were calculated using global test (27); this test can be used to estimate whether the expression of a group of genes is significantly associated with a particular response variable, for example, stage or survival. Gene signatures were further tested by leave-one-out cross-validation. To avoid problems with highly correlated covariates in multivariate Cox proportional hazards models, expression values were scaled by principal component analysis (PCA). The number of components was chosen so that at least 99% of the expression variance was included in the model, with a maximum of 20 components. The choice of this cutoff value was examined by varying the maximum number of components from 3 to 30 (Supplementary Table S2). The performance of these cross-validated models was reported as the C-statistic of a univariate Cox model with the risk score as covariate.

Feature selection. The signatures were optimized by stepwise selection, an iterative procedure which serially removes and adds probe sets from a pool of candidates. The procedure was terminated when adding, removing, or replacing a probe set did not further significantly improve

the mean leave-one-out cross-validated area under receiver operating curve (AUC; for the stage and subtype signatures) or C-statistic (for the survival signature) over all data sets (see Supplemental data for details). Thus all data sets were used for training and validation. In total, over a million promising signatures were cross-validated in all data sets.

Addition of new survival signature to existing predictive model for recurrence-free survival

We calculated the recurrence-free survival probabilities in our patients using the postoperative nomogram developed by the International Bladder Cancer Nomogram Consortium (IBCNC; ref. 28). This nomogram, developed from data of more than 9,000 patients, includes age, sex, time from diagnosis to surgery, pathologic tumor stage and grade, tumor histologic subtype, and regional lymph note status. The contribution of our newly curated gene signature for overall survival to the predictive ability of the nomogram was then evaluated. Multivariate modeling with nomogram information applied the leave-one-out cross-validated gene signature risk score as a second covariate. Thus 2 Cox models were used: the first model estimated the risk as described for the Kaplan–Meier analysis using the gene signature alone; the second model combined the nomogram and the signature score. The right endpoint of this second model was set to be recurrence-free survival. Concordance was estimated with the C-statistics approach (22) for a univariate model with the nomogram alone and then for a multivariate model with the addition of the survival gene signature. Prediction models were compared by the likelihood ratio test.

In Fig. 1, we provide an overview of the analyses conducted in this study.

Results

Differentially expressed genes are associated with clinical features and outcome

Comparison by pathologic stage. When comparing non–muscle-invasive with muscle-invasive samples in our data set, we found 636 significantly differentially expressed probe sets, which significantly overlapped with the corresponding lists of genes in the 6 independent studies ($P <$ 0.0001, Supplementary Tables S3A, S4, and S5, Supplementary Fig. S1). Fibronectin 1 (FN1), a member of the integrin signaling pathway and several other members or close interaction partners of this pathway were upregulated in muscle-invasive as compared with non–muscle-invasive tumors in most data sets we investigated; these genes include ACTN1, COL1A2, COL3A1, COL5A2, COL6A3, COL11A1, COL16A1, FBN1, FLNA, LUM, TGFBI, and TNC. Furthermore, we found that members of the TGF-β signaling pathway were present in most lists of differentially expressed genes; for example, SMAD3, SMAD6, and BMP7 were overexpressed in non–muscle-invasive as compared with muscle-invasive tumors, whereas INHBA, NNMT, and POSTN were overexpressed in muscle-invasive samples. To further characterize the association of gene expression with

Figure 1. Overview of the study. This study represents a massive meta-analysis of our cohort of high-risk patients with bladder cancer (Table 1) and 6 independent studies (Table 2). All data were subject to comprehensive differential expression and survival analyses and to hierarchical clustering. Machine learning algorithms were used to find gene signatures predictive of stage, survival, and molecular subtype. The survival gene signature was then included in a validated postoperative nomogram, developed by the IBCNC (28).
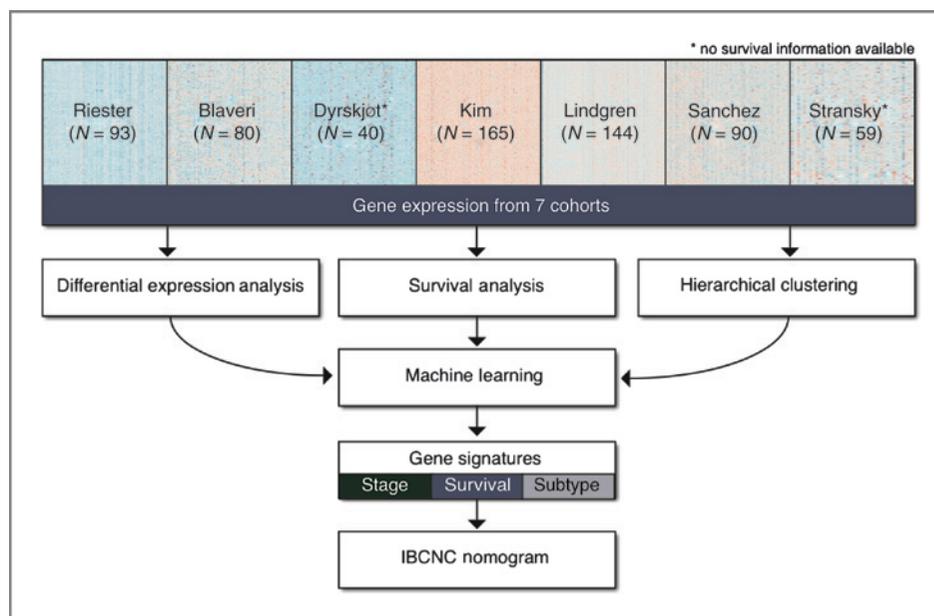
pathologic stage, we created a classifier based on consistently differentially expressed genes among all data sets. The optimized classifier (Supplementary Table S6) identified muscle-invasive tumors with a mean accuracy of 89.0% (88.4% SVM, Supplementary Fig. S3) over all data sets.

Comparison by recurrence status. We then compared the gene expression patterns of patients who developed metastasis (pelvic or distant recurrence) and died of bladder cancer with those of patients who did not. When considering samples of all stages, we identified no significantly differentially expressed genes. However, when restricting this comparison to patients with pathologically organ-confined muscle-invasive (pT2N0) tumors (17 patients, 5 of whom developed metastasis), we identified 53 differentially expressed genes with FDR less than 10% in our data set ($P <$ 0.06, Supplementary Table S3B). In the external data sets, the number of pT2N0 tumors with later metastasis was too small for the identification of significantly differentially expressed genes in this comparison. Among those genes that were differentially expressed in patients who developed metastasis and died were PERP (a TP53 apoptosis factor), ATXN10 (an activator of the Ras-MAPK pathway; ref. 24); and PCM1 (associated with papillary thyroid cancer, chronic myeloid leukemia, and myeloproliferative disorders; ref. 25).

Identification of a robust survival signature for muscle-invasive tumors

Validation of published signatures. We then evaluated 49 published gene signatures (Supplementary Table S2) alongside our classifiers described earlier for correlation with clinical features in our data set and with the 4 independent gene expression data sets with available survival information (6, 7, 9, 13). Because of the smaller number of genes on the Lindgren and colleagues and Blaveri and colleagues cDNA microarrays, only fractions of these sig-

natures' genes could be mapped. We observed overfitting of published survival signatures so that most signatures achieved significance, in terms of survival information, only in the data sets used to identify them (Table 3). Applying previously published signatures to our cohort, we found no significant association between survival and any signatures. Nevertheless, when we analyzed all genes from all survival and progression signatures in a univariate fashion, several markers that achieved low P values in multiple data sets emerged; for example BST2, HLA-G, ICAM1, and LIMCH1 from the Smith and colleagues signature (14); DGCR2, DSC3, and SCEL from Kim and colleagues (13); UBE2C, VCAM1, and WNT5B from Blaveri and colleagues (6); APOBEC3B from Lindgren and colleagues (9); and ATXN10, C17orf39, and INPP4B from Sanchez-Carbayo and colleagues (7; Supplementary Table S2). Especially striking are mitogen—activated protein (MAP) kinases, with different members present in various signatures (7, 9, 11, 14, 29).

Validation of published signatures, stratified by molecular subtype. As reported in multiple other studies (6, 7, 9), bladder tumors cluster in 2 very distinct molecular subtypes. We reproduced this finding in all analyzed data sets (Supplementary Fig. S4), and further, found significant differences in survival between the 2 subtypes (Supplementary Fig. S5). The expression differences between the subtypes were remarkable and overlapped significantly across data sets ($P < 0.00004$, Supplementary Table S5). This allowed a robust classification of tumors by subtype with a novel 19-gene signature (Supplementary Table S7, Supplementary Fig. S6). The very heterogeneous landscape of muscle-invasive bladder tumors motivated us to evaluate the performance of all signatures in the 2 subtypes separately. Two progression signatures, one developed for bladder (6) and one for breast cancer (30), were significantly associated with survival in our cohort when

Table 3. Performance of published gene signatures

Signature	Endpoint	Stage	Blaveri and colleagues (6)		Kim and colleagues (13)		Lindgren and colleagues (9)		Sanchez-Carbayo and colleagues (7)		Riester	
			P	C	P	C	P	C	P	C	P	C
Blaveri and colleagues (6)	OS	MI	**0.01**	0.66	0.66	0.57	0.15	0.60	0.66	0.61	0.73	0.56
Kim and colleagues (13)	OS	MI	0.41	0.60	**0.00**	0.75	0.76	0.53	0.40	0.54	0.84	0.50
Kim and colleagues (13)	DSS	MI	0.11	0.57	**0.00**	0.72	0.84	0.63	0.26	0.60	0.87	0.56
Kim and colleagues (13)	Progression	non-MI	0.70	0.50	0.09	0.56	0.53	0.53	**0.04**	0.59	0.27	0.56
Lindgren and colleagues (9)	OS	all	0.29	0.55	0.17	0.55	0.10	0.72	0.48	0.50	0.93	0.54
Sanchez-Carbayo and colleagues (7)	OS	MI	0.15	0.56	0.56	0.54	0.33	0.50	**0.02**	0.60	0.39	0.50
Sanchez-Carbayo and colleagues (7)	Progression	MI	0.21	0.50	**0.02**	0.58	0.42	0.73	0.15	0.51	0.95	0.54
Smith and colleagues (14)	Progression	MI	0.42	0.54	0.17	0.54	0.15	0.58	0.15	0.60	0.79	0.50
Riester	OS	MI	**0.04**	0.76	**0.04**	0.75	**0.002**	0.87	**0.005**	0.74	0.06	0.71

NOTE: The table shows performance measures of all published bladder cancer survival gene signatures and 3 progression signatures. All signatures were tested in all public data sets with survival information (muscle-invasive tumors only). The P values represent the probabilities that there is no association between expression and survival (27). The C-statistic of a leave-one-out cross-validated Cox model with the gene intensities as covariates is reported as a second performance measure. A C-statistic of 0.5 corresponds to a random model, one of 1.0 to a perfect model. All signatures were originally identified with the corresponding data set, except for the Smith and colleagues signature (14), which was obtained in that study by training with the Sanchez-Carbayo and colleagues data set (7) and with publicly unavailable data. Our signature was obtained by using all data sets for training. Tested right endpoint was overall survival in all data sets, not necessarily the endpoint for which the signature was designed for. See Supplementary Fig. S6 for a detailed analysis of 41 other signatures.

Abbreviations: C, C-statistic; DSS, disease-specific survival; MI, muscle invasive; OS, overall survival.

Table 4. Gene signature predictive of overall survival for patients with muscle-invasive tumors (stage T2–T4)

Gene symbol	Probe set	Gene name
APOBEC3B	206632_s_at	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 3B
ATF3	202672_s_at	Activating transcription factor 3
CCL5	204655_at	Chemokine (C–C motif) ligand 5
DGCR2	214198_s_at	DiGeorge syndrome critical region gene 2
ENDOD1	212573_at	Endonuclease domain containing 1
FADD	202535_at	Fas (TNFRSF6) associated via death domain
JUNB	203022_at	Ribonuclease H2, subunit A
LMO7	242722_at	LIM domain 7
MAP2K1	202670_at	Mitogen-activated protein kinase kinase 1
MAP3K1	225927_at	mitogen-activated protein kinase kinase kinase 1
PDGFC	218718_at	Platelet-derived growth factor C
PEA15	200787_s_at	Phosphoprotein enriched in astrocytes 15
PFN1	214617_at	Perforin 1 (pore-forming protein)
PPP1R12A	201604_s_at	Protein phosphatase 1, regulatory (inhibitor) subunit 12A
PRDX1	208680_at	Peroxiredoxin 1
PRMT1	206445_s_at	Protein arginine methyltransferase 1
SLC1A5	208916_at	Solute carrier family 1 (neutral amino acid transporter), member 5
TNFAIP6	206025_s_at	TNF, α-induced protein 6
TSG101	201758_at	Tumor susceptibility gene 101
TSPAN5	209890_at	Tetraspanin 5

applied to one of the 2 subtypes (Supplementary Fig. S7). By analyzing primary and recurrent tumors from the Kim and colleagues cohort (13), we found evidence that tumors can progress from one molecular subtype to the other (Supplementary Table S8).

Curation of a novel survival signature. We then curated a new 20-gene overall survival signature for muscle-invasive tumors (Table 4) by stepwise optimization and tested its predictive accuracy in our data set and the 4 external data sets with available survival information. A leave-one-out cross-validated multivariate Cox proportional hazards model based on this signature classified tumors into 2 equally sized risk groups with significantly different survival in all data sets (Fig. 2A–E). As an example, with the model applied to the data set by Kim and colleagues (13), there was a 5-fold increase in the risk of death in those classified as high-risk [HR: 5.05; 95% confidence interval (CI), 2.26–11.3; $P < 0.008$] compared with those classified as low risk. See Supplementary Figs. S8 and S9 for details of the signature. Neoadjuvant chemotherapy, administered to 3 patients in our cohort, had no impact on the prediction accuracies (Supplementary Table S9). We next compared the prediction accuracies of the signature with pathologic stage and grade in all cohorts and found highly significant improvements of model concordance and likelihood in all data sets (mean ΔC-statistic: 0.14; 95% CI, 0.01–0.27; likelihood ratio test, $P < 0.001$; Fig. 2F, Supplementary Table S10B).

Addition of survival signature to the nomogram increases its predictive accuracy

We then incorporated the new signature developed for muscle-invasive tumors (Table 4) into the IBCNC nomogram (28), which was designed to predict 5-year recurrence-free survival after radical cystectomy. Necessary data points for the nomogram were not available in the external data sets. When tested in our muscle-invasive patient cohort, a multivariate Cox model predicting recurrence-free survival with both nomogram and gene signature score (C-statistic: 0.66; 95% CI, 0.56–0.76) was a significant improvement (ΔC-statistic: 0.08; 95% CI, −0.04–0.20; $P < 0.005$, likelihood ratio test) over a model based on the nomogram score alone (C-statistic: 0.58; 95% CI, 0.43–0.72). As the signature was designed for overall survival due to the lack of recurrence dates in the independent data, the prediction accuracies increased profoundly in a model predicting overall survival (ΔC-statistic: 0.15; 95% CI, 0.07–0.23; Supplementary Table S10C–D).

Discussion

This study shows the utility of adding molecular information to an existing prognostic tool in patients with bladder cancer who undergo radical cystectomy. Gene expression data from a new cohort of patients with bladder cancer was analyzed and compared with gene expression
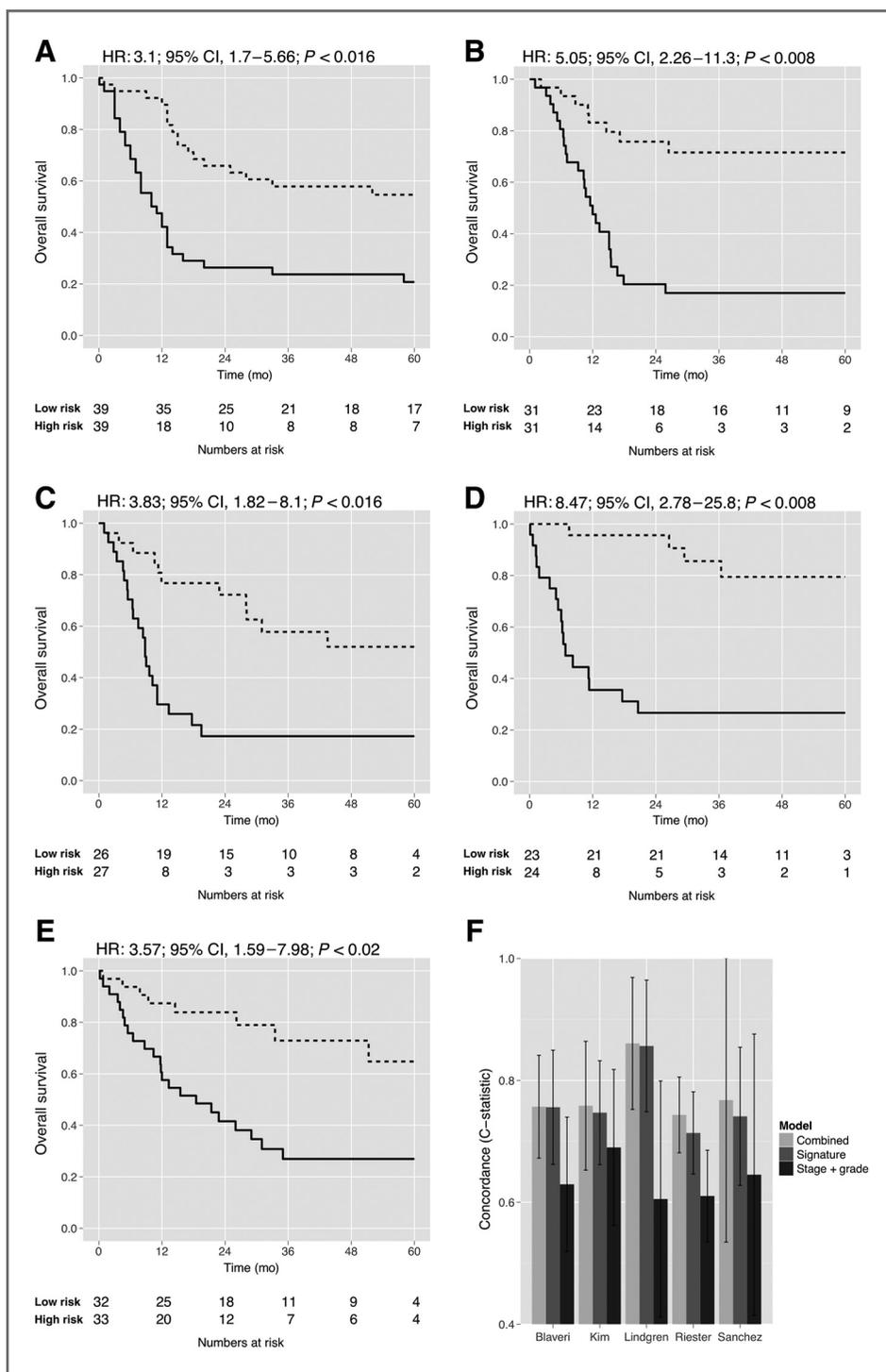
Figure 2. Prediction accuracy of our novel gene signature. A–E, Kaplan–Meier plots for low- and high-risk groups in 5 data sets. HRs indicate how well the signature separates tumors. An HR of 3.1, for example, corresponds to 3.1-fold increase in risk for the high-risk group. Some genes were not present on all microarray platforms. The novel signature is shown in Table 4. A, Riester (all 20 genes); B, Blaveri (18 genes); C, Kim (20 genes); D, Lindgren (10 genes); and E, Sanchez-Carbayo (20 genes). In all plots, patients were stratified in 2 equally sized groups by the cohort medians of the calculated risk scores. We further tested whether the gene signature improves the predictive utility of stage and grade, 2 established predictors available for all 5 data sets. This comparison was done by analyzing the concordance (C-statistic, ref. 22) of the models. A concordance of 0.5 corresponds to a random model, one of 1.0 of a perfect model. F, C-statistics for all 5 data sets for a multivariate model with stage, grade, and the gene signature risk score; for model with gene signature risk score alone; and for a model with only stage and grade. Error bars indicate 95% CIs.

data from multiple published studies. A robust expression signature was identified that improved the predictive outcome of a well-accepted nomogram and was independently predictive of survival in all cohorts where survival data were available. These findings are particularly striking given the relative homogeneity of the population analyzed, as our

cohort, having all undergone the most aggressive surgical therapy, is highly selected for being at high risk of death from disease.

One important contribution of this study is the clinical validation of published gene signatures. To be clinically relevant, new survival markers must (i) stratify patients in

groups with significantly different survival; (ii) deliver survival information that is not already included in established clinically used predictors such as grade and stage; and (iii) increase the accuracies of existing prediction models to an extent that warrants the cost and effort of obtaining biomarker status (31). Previously published gene signatures (Supplementary Table S2), however, did poorly in our patient cohort; no gene signature provided more survival information than would be expected by chance. This finding can partly be explained by cross-platform issues like incomplete mapping of genes to available probe sets or probe sets targeting different exons. Another reason is a likely overfitting of published signatures due to rather small training data sets for the very heterogeneous group of muscle-invasive bladder cancers. To circumvent these limitations, we optimized the survival signatures for the subgroup of muscle-invasive tumors, and showed that gene expression clearly stratifies patients into 2 groups, denoted as low-risk and high-risk. The new signature further segregated by risk 305 patients with muscle-invasive tumors in the external data sets containing gene expression and survival information. Interestingly, 4 genes in our muscle-invasive survival signature, APOBEC3B, DGCR2, PRMT1, SLC1A5, are located on chromosomes 22 and 19, in neighboring chromosome bands of single-nucleotide polymorphisms (SNP) recently described as associated with an increased risk of developing bladder cancer (32). In addition to our signature, several other genes (Supplementary Table S2) significantly correlated with survival in multiple independent cohorts. These genes identified are potential therapeutic targets and worthy of further investigation to understand their role in bladder cancer. For example, inhibitors of MAP2K1 (i.e., MEK1), which is part of our overall survival signature (Table 4), are currently being evaluated in several clinical trials and our results provide a rationale for this in bladder cancer. Our screen for potential biomarkers in multiple data sets thus also has important applications in gene selection for technologies such as NanoString, which can target only a limited number of genes.

To place our signature in the context of currently known clinical data, we applied the survival signature to a robust nomogram that predicts outcome after radical cystectomy (28). This postoperative nomogram was shown to have a significantly better predictive accuracy (concordance index, 0.75) than standard tumor-node-metastasis (TNM; concordance index, 0.68; $P < 0.001$) or pathologic (concordance index, 0.62; $P < 0.001$) subgroupings. When this model was applied to our patient cohort with the addition of our newly developed survival signature, we observed statistically significant improvements in the prediction of risk of death as compared with the use of the nomogram only. As the signature reflects unique molecular features associated with risk, its addition to the nomogram thus provides more insights into the biologic behavior of a malignancy of patient, and leads to improved prognostication as compared with the use of the nomogram alone. A more accurate prediction of progression may help select the higher-risk patients for

additional treatment, such as adjuvant systemic therapy after surgery. External validation of these findings is ongoing.

In addition, our analysis led to the identification of 53 genes significantly associated with the development of metastasis and death of patients with pathologically organ-confined muscle-invasive disease (pT2N0). This list of genes establishes a molecular signature of less favorable outcome, which may help stratify patients in this identically staged subgroup of bladder tumors to more or less intensive multimodal therapy. Patients with this stage of disease are typically followed expectantly, but evaluation of gene expression data or testing for downstream products of expression may identify those patients at higher risk for progression who may benefit from systemic therapy after radical surgery. The investigation of this list in a prospective collection of similar cystectomy samples is essential. An interesting additional opportunity for evaluation would be among patients with muscle-invasive disease at the time of TUR—by characterizing a greater risk of aggressive behavior, we may more carefully select patients for chemotherapy prior to radical surgery as well, to improve our survival outcomes using multimodal therapy.

We found that FN1 was significantly overexpressed in muscle-invasive tumors as compared with non–muscle-invasive samples in all data sets we analyzed. The role of FN1 in cancer invasion and metastasis is still poorly understood, but its increased expression might be the result of an enhanced recruitment of cancer-associated fibroblasts (33, 34). The consistency across data sets suggests that the abundant differential expression of FN1 and other extracellular matrix (ECM) genes is unlikely due to stromal contamination. Our stage classifier (Supplementary Table S6) could identify patients as having muscle-invasive disease with a mean accuracy of 89%, which is within the range of previously published classifiers (6, 24). However, our stage classifier was validated in more data sets and consists of fewer genes than the published classifiers, which offer advantages for its potential clinical use. Furthermore, the accuracy rate of 89% does not imply a false classification rate of 11% as some tumors might have been incorrectly classified histologically. Urothelial cancer of the bladder is exceptionally difficult to stage accurately prior to cystectomy, given technical limitations of endoscopic resection and imaging. Significant rates of upstaging from time of TUR to radical cystectomy have been reported, ranging from 31% to 61% (35–37). In our patient cohort, more than 80% of the tumors were determined to have higher pathologic stage at the time of cystectomy, when compared with the preoperative TUR specimens, with many tumors found to be locally advanced (>pT2) when clinically staged as organ confined (≤pT2).

The unsupervised clustering produced 2 strikingly different groups of patients, who were not clearly distinguishable by many clinical variables such as the presence of lymph node metastases or recurrence status. Gene expression showed large differences between the 2

groups (Supplementary Table S3C) with highly significant overlaps in all data sets (Supplementary Table S5), allowing a robust identification of cluster membership by a 19-gene signature (Supplementary Table S7). Lindgren and colleagues (9) similarly identified 2 distinct molecular subtypes (denoted as MS1 and MS2) by analyzing 144 tumor samples obtained by transurethral bladder biopsy, representing a wide range of tumor stage and grade. Their investigation, which included whole-genome and specific mutational analyses, identified a greater number of genomic alterations in the MS2 group, containing nearly all of the muscle-invasive samples. The genomic instability in the MS2 group was also associated with expression of genes involved in cell-cycle pathways and cellular transformation (9), all indicative of more aggressive behavior and worse prognosis. Consistent with their results, we found a progression from MS1 to MS2 to be more likely than a progression from MS2 to MS1 when analyzing 23 recurrent tissue samples from Kim and colleagues (13). While only stage had a significant association with subtype, other covariates not available in this study, for example response to treatments, could be tested for association in prospective data. Not surprisingly, survival signatures often showed very different prediction accuracies in the 2 subtypes. This heterogeneity of bladder cancer tumors thus points to the necessity of large training data to establish robust signatures. Interestingly, from all 49 considered previously published signatures, a breast cancer progression signature (30) showed the highest association with overall survival in the MS2 group of our patient cohort (Supplementary Fig. S7A). This finding suggests that optimizing thoroughly validated signatures designed for other solid tumors might yield more robust signatures than developing new signatures from scratch, in particular when only small training data sets are available. The identification of different molecular signatures for subtype suggests that these subtypes, which are strongly associated with invasion status, may be a building block for future work to noninvasively determine the aggressiveness of tumors; for example, these changes in expression may be detectable in urine.

Our data set represents a high-risk population of patients with bladder cancer who underwent radical cystectomy at a high-volume academic center. The sample size is large for a single-center report of microarray analysis and comprises, to our knowledge, the largest number of high-risk patients with bladder cancer of all published microarray data sets. One limitation of our findings is, however, a possible element of overfitting, resulting from the use of the same data in testing and validation steps, although this concern is mitigated through the use of statistical methods such as cross-validation. This points to the imperative for validation in independent data of any gene signature associated with survival. Such a validation is necessary because signatures established from single, moderately sized data sets often did poorly in external data due to the relatively high probability that genes correlate with survival by chance. However, extending the sample size with independent data, as carried out here, can minimize this problem. For example, a gene found to be predictive in one data set could arise due to chance, but if one gene is predictive with consistent expression in multiple data sets, it is more likely to be a robust survival marker.

A considerable strength of our investigation is its comprehensive effort to combine the majority of published data on gene expression in bladder cancer, all arrayed on different platforms, in a meaningful way. With this study, we have shown that meta-analyses across platforms are feasible. Especially in rare cancers or subtypes, the increase in sample size might often outweigh the disadvantages of meta-analyses, such as the heterogeneity of patient cohorts, the different treatments before and after surgery, and the decrease in potential markers when focusing on genes present on all platforms. Nonetheless, large independent patients cohorts are needed for further validation and optimization (i.e., for other endpoints such as recurrence-free survival or for particular microarray platforms) of the prognostic signatures, before warranting a clinical trial (38). The results of this study strongly motivate such sample collection for urothelial tumors.

Conclusions

Our gene expression analysis identifies novel signatures predictive of tumor stage, progression among patients with organ-confined disease, molecular subtypes of tumors, and overall survival. We further incorporated 6 published gene expression data sets for a large-scale comparison and validation of predictive signatures and found a significant increase in predictive accuracy of a postoperative nomogram by addition of our gene signature. Multiple genes identified across several sets have emerged as highly promising candidates for further investigation. The identification of patients at higher risk for progression or death will provide additional rationale for multimodality management, using chemotherapy and surgery, in an effort to improve survival.

References

1. Jemal A SR, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. CA Cancer J Clin 2009;59:225–49.

2. Kaufman DS, Shipley WU, Feldman AS. Bladder cancer. Lancet 2009;374:239–49.

3. Donat SM. Evaluation and follow-up strategies for superficial bladder cancer. Urol Clin North Am 2003;30:765–76.

4. Herr HW, Dotan Z, Donat SM, Bajorin DF. Defining optimal therapy for muscle invasive bladder cancer. J Urol 2007;177:437–43.

5. McDermott U, Downing JR, Stratton MR. Genomics and the continuum of cancer care. N Engl J Med 2011;364:340–50.

6. Blaveri E, Simko J, Korkola JE, Brewer JL, Baehner F, Mehta K, et al. Bladder cancer outcome and subtype classification by gene expression. Clin Cancer Res 2005;11:4044–55.

7. Sanchez-Carbayo M SN, Lozano J, Saint F, Cordon-Cardo C. Defining molecular profiles of poor outcome in patients with invasive bladder cancer using oligonucleotide microarrays. J Clin Oncol 2006;24: 778–89.

8. Elsamman E, Fukumori T, Ewis AA, Ali N, Kajimoto K, Shinohara Y, et al. Differences in gene expression between noninvasive and invasive transitional cell carcinoma of the human bladder using complementary deoxyribonucleic acid microarray: preliminary results. Urol Oncol 2006;24:109–15.

9. Lindgren D, Frigyesi A, Gudjonsson S, Sjödahl G, Hallden C, Chebil G, Veerla S, et al. Combined gene expression and genomic profiling define two intrinsic molecular subtypes of urothelial carcinoma and gene signatures for molecular grading and outcome. Cancer Res 2010;70:3463–72.

10. Dyrskjøt L, Thykjaer T, Kruhøffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S, et al. Identifying distinct classes of bladder carcinoma using microarrays. Nat Genet 2003;33:90–6.

11. Birkhahn M, Mitra A, Williams AJ, Lam G, Ye W, Datar RH, et al. Predicting recurrence and progression of noninvasive papillary bladder cancer at initial presentation based on quantitative gene expression profiles. Eur Urol 2010;57:12–20.

12. Catto JW, Abbod MF, Wild PJ, Linkens DA, Pilarsky C, Rehman I, et al. The application of artificial intelligence to microarray data: identification of a novel gene signature to identify bladder cancer progression. Eur Urol 2010;57:398–406.

13. Kim WJ, Kim EJ, Kim SK, Kim YJ, Ha YS, Jeong P, et al. Predictive value of progression-related gene classifier in primary non-muscle invasive bladder cancer. Mol Cancer 2010;9:3.

14. Smith SC, Baras AS, Dancik G, Ru Y, Ding KF, Moskaluk CA, et al. A 20-gene model for molecular nodal staging of bladder cancer: development and prospective assessment. Lancet Oncol 2011;12:137–43.

15. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207–10.

16. Stransky N, Vallot C, Reyal F, Bernard-Pierrot I, de Medina SG, Segraves R, et al. Regional copy number-independent deregulation of transcription in cancer. Nat Genet 2006;38:1386–96.

17. Wu Z, Irizarry R, Gentleman R, Martinez-Murillo F, Spencer F. A model-based background adjustment for oligonucleotide expression arrays. J Am Stat Assoc 2004;99:909–17.

18. Smyth G. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Stat Appl Genet Mol Biol 2004;3:Article 3.

19. Benjamini Y HY. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Statist Soc 1995;B:289–300.

20. Sing T, Sander O, Beerenwinkel N, Lengauer T. ROCR: visualizing classifier performance in R. Bioinformatics 2005;21:3940–1.

21. van Houwelingen HC, Bruinsma T, Hart AA, Van't Veer LJ, Wessels LF. Cross-validated Cox regression on microarray gene expression data. Stat Med 2006;25:3201–16.

22. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med 2011;30:1105–17.

23. Suzuki R, Shimodaira H. Pvclust: an R package for assessing the uncertainty in hierarchical clustering. Bioinformatics 2006;22:1540–2.

24. Lauss M, Ringnér ÄM, Höglund M. Prediction of stage, grade, and survival in bladder cancer using genome wide expression data: a validation study. Clin Cancer Res 2010;16:4421–33.

25. Mandruzzato S, Callegaro A, Turcatel G, Francescato S, Montesco MC, Chiarion-Sileni V, et al. A gene expression signature associated with survival in metastatic melanoma. J Transl Med 2006;4:50.

26. Soikkeli J, Podlasz P, Yin M, Nummela P, Jahkola T, Virolainen S, Krogerus L, et al. Metastatic outgrowth encompasses COL-I, FN1, and POSTN up-regulation and assembly to fibrillar networks regulating cell adhesion, migration, and growth. Am J Pathol 2010;177:387–403.

27. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. Bioinformatics 2004;20:93–9.

28. Bochner BH, Kattan MW, Vora KC. Postoperative nomogram predicting risk of recurrence after radical cystectomy for bladder cancer. J Clin Oncol 2006;24:3967–72.

29. Mitra AP, Pagliarulo V, Yang D, Waldman FM, Datar RH, Skinner DG, et al. Generation of a concise gene panel for outcome prediction in urinary bladder cancer. J Clin Oncol 2009;27:3929–37.

30. van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. Nature 2002;415:530–6.

31. Steyerberg EW, Vickers AJ, Cook NR, Gerds T, Gonen M, Obuchowski N, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–38.

32. Rothman N, Garcia-Closas M, Chatterjee N, Malats N, Wu X, Figueroa JD, et al. A multi-stage genome-wide association study of bladder cancer identifies multiple susceptibility loci. Nat Genet 2010;42: 978–84.

33. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell 2011;144:646–74.

34. Kalluri R, Zeisberg M. Fibroblasts in cancer. Nat Rev Cancer 2006;6:392–401.

35. Ficarra V, Dalpiaz O, Alrabi N, Novara G, Galfano A, Artibani W. Correlation between clinical and pathological staging in a series of radical cystectomies for bladder carcinoma. BJU Int 2005;95: 786–90.

36. Shariat SF, Palapattu GS, Karakiewicz PI, Rogers CG, Vazina A, Bastian PJ, et al. Discrepancy between clinical and pathologic stage: impact on prognosis after radical cystectomy. Eur Urol 2007;51:137–49; discussion 149–51.

37. Millikan R, Dinney C, Swanson D, Sweeney P, Ro JY, Smith TL, et al. Integrated therapy for locally advanced bladder cancer: final report of a randomized trial of cystectomy plus adjuvant M-VAC versus cystectomy with both preoperative and postoperative M-VAC. J Clin Oncol 2001;19:4005–13.

38. Koscielny S. Why most gene expression signatures of tumors have not been useful in the clinic. Sci Transl Med 2010;2:14ps2.