available at www.sciencedirect.com
journal homepage: www.europeanurology.com

European Association of Urology

Platinum Priority – Prostate Cancer

*Editorial by Michael Fraser, Paul C. Boutros and Robert G. Bristow on pp. 533–534 of this issue*

# Stromal Gene Expression is Predictive for Metastatic Primary Prostate Cancer☆

*Fan Mo[a,1], Dong Lin[a,b,1], Mandeep Takhar[c,1], Varune Rohan Ramnarine[a], Xin Dong[b], Robert H. Bell[a], Stanislav V. Volik[a], Kendric Wang[a], Hui Xue[b], Yuwei Wang[b], Anne Haegert[a], Shawn Anderson[a], Sonal Brahmbhatt[a], Nicholas Erho[c], Xinya Wang[a], Peter W. Gout[b], James Morris[d], R. Jeffrey Karnes[e], Robert B. Den[f], Eric A. Klein[g], Edward M. Schaeffer[h,i], Ashley Ross[h], Shancheng Ren[j], S. Cenk Sahinalp[k,l], Yingrui Li[m], Xun Xu[m], Jun Wang[m], Jian Wang[m], Martin E. Gleave[a], Elai Davicioni[c], Yinghao Sun[j], Yuzhuo Wang[a,b,1,*], Colin C. Collins[a,k,1,*]*

[a] *Vancouver Prostate Centre & Laboratory for Advanced Genome Analysis, Department of Urologic Sciences, University of British Columbia, Vancouver, BC, Canada;* [b] *Department of Experimental Therapeutics, BC Cancer Agency, Vancouver, BC, Canada;* [c] *Research and Development, GenomeDx Biosciences, Vancouver, BC, Canada;* [d] *Department of Radiation Oncology, BC Cancer Agency, Vancouver, BC, Canada;* [e] *Department of Urology, Mayo Clinic College of Medicine, Rochester, MN, USA;* [f] *Department of Radiation Oncology, Sidney Kimmel Medical College at Thomas Jefferson University, Philadelphia, PA, USA;* [g] *Glickman Urological and Kidney Institute, Cleveland Clinic, Cleveland, OH, USA;* [h] *Department of Urology, James Buchanan Brady Urological Institute, Department of Oncology, Johns Hopkins Hospital, Baltimore, MD, USA;* [i] *Department of Urology, Northwestern University School of Medicine, Chicago, IL, USA;* [j] *Department of Urology, Shanghai Changhai Hospital, Second Military Medical University, Shanghai, China;* [k] *School of Computing Sciences, Simon Fraser University, Burnaby, BC, Canada;* [l] *School of Informatics and Computing, Indiana University, Bloomington, IN, USA;* [m] *BGI-Shenzhen, Shenzhen, China*

**Article info**

***Article history:***
Accepted February 28, 2017

***Associate Editor:***
James Catto

***Keywords:***
Stromal gene
Prostate cancer metastasis
Prognostic biomarkers
Genomic profiling
RNA sequencing

**Abstract**

***Background:*** Clinical grading systems using clinical features alongside nomograms lack precision in guiding treatment decisions in prostate cancer (PCa). There is a critical need for identification of biomarkers that can more accurately stratify patients with primary PCa.
***Objective:*** To identify a robust prognostic signature to better distinguish indolent from aggressive prostate cancer (PCa).
***Design, setting, and participants:*** To develop the signature, whole-genome and whole-transcriptome sequencing was conducted on five PCa patient-derived xenograft (PDX) models collected from independent foci of a single primary tumor and exhibiting variable metastatic phenotypes. Multiple independent clinical cohorts including an intermediate-risk cohort were used to validate the biomarkers.
***Outcome measurements and statistical analysis:*** The outcome measurement defining aggressive PCa was metastasis following radical prostatectomy. A generalized linear model with lasso regularization was used to build a 93-gene stroma-derived metastasis signature (SDMS). The SDMS association with metastasis was assessed using a Wilcoxon rank-sum test. Performance was evaluated using the area under the curve (AUC) for the receiver operating characteristic, and Kaplan-Meier curves. Univariable and multivariable

☆ Raw sequencing data are available at The European Nucleotide Archive (ENA) under accession number PRJEB19256.

[1] These authors contributed equally to this work and should be regarded as joint first and last authors, as appropriate.

* Corresponding authors. Vancouver Prostate Centre, 2660 Oak Street, Vancouver, V6H 3Z6, Canada. Tel. +1 604 875 4818; Fax: +1 604 875 5654.
E-mail addresses: ywang@bccrc.ca (Y. Wang), ccollins@prostatecentre.com (C.C. Collins).

http://dx.doi.org/10.1016/j.eururo.2017.02.038

regression models were used to compare the SDMS alongside clinicopathological variables and reported signatures. AUC was assessed to determine if SDMS is additive or synergistic to previously reported signatures.
***Results and limitations:*** A close association between stromal gene expression and metastatic phenotype was observed. Accordingly, the SDMS was modeled and validated in multiple independent clinical cohorts. Patients with higher SDMS scores were found to have worse prognosis. Furthermore, SDMS was an independent prognostic factor, can stratify risk in intermediate-risk PCa, and can improve the performance of other previously reported signatures.
***Conclusions:*** Profiling of stromal gene expression led to development of an SDMS that was validated as independently prognostic for the metastatic potential of prostate tumors.
***Patient summary:*** Our stroma-derived metastasis signature can predict the metastatic potential of early stage disease and will strengthen decisions regarding selection of active surveillance versus surgery and/or radiation therapy for prostate cancer patients. Furthermore, profiling of stroma cells should be more consistent than profiling of diverse cellular populations of heterogeneous tumors.


## 1. Introduction

Prostate cancer (PCa) is a major source of cancer-related mortality because of metastases from the primary tumor; however, many patients present with indolent disease that is unlikely to develop metastasis and does not require invasive treatment [1,2]. Currently, clinical grading systems using clinical features alongside nomograms lack the precision to guide treatment decisions [3,4]. Therefore, there is a critical need for identification of biomarkers that can more accurately stratify patients with primary PCa. Numerous studies have attempted to address this clinical need using gene expression signatures, genomic alterations, and protein profiling [5–14]. However, it has yet to be convincingly demonstrated that these biomarkers have superior predictive abilities compared to established clinical grading systems, which limits their clinical utility.

PCa is recognized as a heterogeneous multifocal disease [5,15]. Both intertumor heterogeneity between patients and intratumor heterogeneity (ITH) within patients present major challenges for both patient stratification and discovery of biomarkers [5,9,16]. To circumvent these, a panel of PCa patient-derived xenograft (PDX) models was established from needle biopsy specimens taken from different foci of a single primary prostate tumor [17]. Although these PDX models have similar genetic and histopathological characteristics, they show marked variations with regard to spontaneous metastatic abilities in vivo [17]. In addition, in accordance with other studies, the stromal components of PDX models are largely replaced by host cells following serial passaging of the xenografts [18–20], which makes it possible to segregate the gene signature of cancer cells (human origin) and that of cancer-associated stromal cells (mouse origin) using species-specific mapping of RNA sequencing (RNA-seq) data. Thus, these xenografts provide a valuable tool for dissection of ITH, studies of metastasis, and identification of potential biomarkers predictive of metastasis.

In this study, we found that gene expression profiling of the cancer-associated stromal cells successfully distinguished nonmetastatic from metastatic PDX models, and further identified a 93-gene stroma-derived metastasis signature (SDMS) with potential for predicting the metastatic risk of primary PCa, including those with intermediate Gleason scores, and may serve as a signature predictive for the metastatic potential of PCa.

## 2. Materials and methods

### 2.1. Xenograft models, sample preparation, and sequencing

The PDX models were maintained via serial transplantation of subrenal capsule xenografts in male NSG mice supplemented with testosterone, as previously described [17]. The lymph nodes, lungs, livers, kidneys, spleens, and bones (femur) of the hosts were fixed for examination of metastases using histological and immunohistochemical staining. Animal care and experiments were carried out in accordance with the guidelines of the Canadian Council on Animal Care. DNA and RNA from frozen tumor tissue were isolated, purified, and sequenced according to standard protocols.

### 2.2. Sequencing data analysis

Raw sequencing data are available at The European Nucleotide Archive (ENA) under accession number PRJEB19256. Whole genome sequencing (WGS) and RNA-seq reads were aligned onto a combined human and mouse genome reference. Genomic alterations and tumor gene expression were profiled using human-specific WGS and RNA-seq reads, while stromal gene expression was analyzed using mouse-specific RNA-seq reads (Supplementary material).

### 2.3. Gene signature development and validation

Microarray data for six radical prostatectomy (RP) cohorts from studies described previously [21–26] were extracted from the Decipher GRID. The SDMS was developed as a generalized linear model with lasso regularization, using the MC I cohort for training. The association with metastasis was assessed separately in each cohort using a Wilcoxon rank-sum, the area under the receiver operating characteristic curve (AUC), and Kaplan-Meier curves, where applicable. Univariable and multivariable regression models were used to compare the SDMS to clinical and pathological variables.

The five previously validated signatures [10–14] were ported to the Affymetrix platform and trained as random forest models using the MC I cohort. To compare them to the SDMS, multivariable analysis was performed using logistic regression models in a pooled validation cohort ($n$ = 621) [22–24] for which the primary outcome was metastasis within

**Table 1 – Characteristics of the LTL-313 patient-derived xenograft (PDX) models and assays performed**

| PDX model | Origin in human prostate | Metastatic ability | Doubling time in mice (d) | Array CGH | WGS depth (×) | WTS reads (million) | |
|---|---|---|---|---|---|---|---|
| | | | | | | PolyA$^+$ | rRNA$^-$ |
| LTL-313A | Left apex | + | 12–15 | Yes | 60.16 | 98.26 | 50.59 |
| LTL-313B | Left base | – | 15–25 | Yes | 60.73 | 92.09 | 45.54 |
| LTL-313C | LM LL | + | 31 | Yes | 66.48 | 87.09 | 49.45 |
| LTL-313D | LM LM | + | 15 | Yes | 60.72 | 98.89 | 46.85 |
| LTL-313H | RM LM | ++ | 13–15 | Yes | 60.72 | 90.51 | 54.76 |

CGH = comparative genomic hybridization; WGS = whole-genome sequencing; WTS = whole-transcriptome sequencing; LM = left median; LL = left-left; RM = right median.

5 yr of RP. To evaluate the additional prognostic value of the SDMS, logistic regression models of the SDMS in combination with the five signatures were trained in the MC I cohort. Each model was then evaluated in the pooled validation cohort using the AUC (Supplementary material).

## 3. Results

### 3.1. Genomic and transcriptomic profiles of tumor cells are not associated with the varying metastatic abilities of LTL-313 PDX models

A panel of transplantable PCa PDX models, the LTL-313 series, was developed from needle biopsies of multiple foci of a single primary tumor [17] (Supplementary Fig. 1). These models retain common histopathological and molecular characteristics of the original cancer tissues, while showing variable metastatic abilities (Table 1). LTL-313H is highly metastatic. The LTL-313A, LTL-313C, and LTL-313D models give rise to lung metastases, but to a lesser extent, and LTL-313B is nonmetastatic.

To assess molecular ITH and metastatic potential, WGS and RNA-seq were performed (Fig. 1). The profiles of chromosomal copy number aberrations (CNAs; Supplementary Table 1, Supplementary Fig. 2), single nucleotide variants (SNVs; Supplementary Table 2), small insertions/deletions (InDels; Supplementary Table 2), and fusion genes (Supplementary Table 3) showed a high level of similarity among these models (Fig. 2A–C). Although certain model-specific CNAs and mutations were detected, corresponding changes in gene expression or mutations were not observed in the mRNA (Fig. 2A). Phylogenetic trees based on chromosomal breakpoint and SNV/InDel similarities failed to separate the nonmetastatic LTL-313B model from the metastatic models (Fig. 2B,C and Supplementary Tables 4 and 5). Reconstruction of a subclonal composition based on SNVs/InDels further emphasized the high degree of similarity among all five models (Fig. 2D and Supplementary Table 6).

Transcriptomic profiles of the LTL313 series and three other independent PCa PDX models (nonmetastatic LTL-418 and metastatic LTL-311 and LTL-331 [17]) were also included in the gene expression profiling. To selectively generate gene expression profiles for the tumors, the RNA-seq reads that were mapped uniquely to the human genome were used (Supplementary material).

As for the genomic profiling, unsupervised hierarchical clustering of total gene expression profiles from the tumors (human) did not separate metastatic models from nonmetastatic ones (Fig. 3A). A heatmap generated based on the

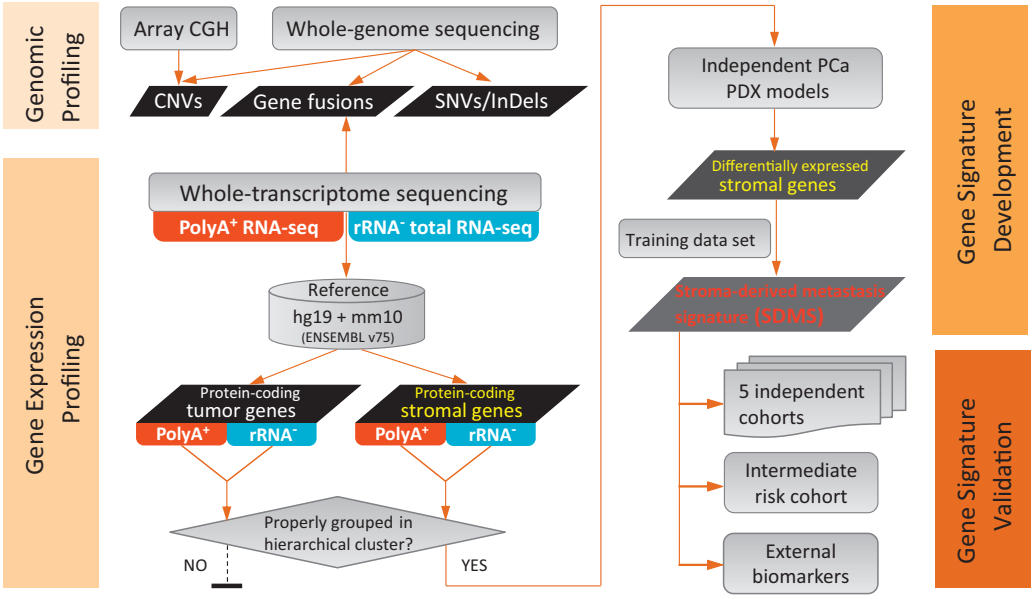


**Fig. 1 – Workflow overview. Comprehensive genomic and transcriptomic profiling led to the discovery of a stromal gene signature. This signature was further validated in multiple large independent clinical cohorts, including a cohort of patients with intermediate-risk tumors.**

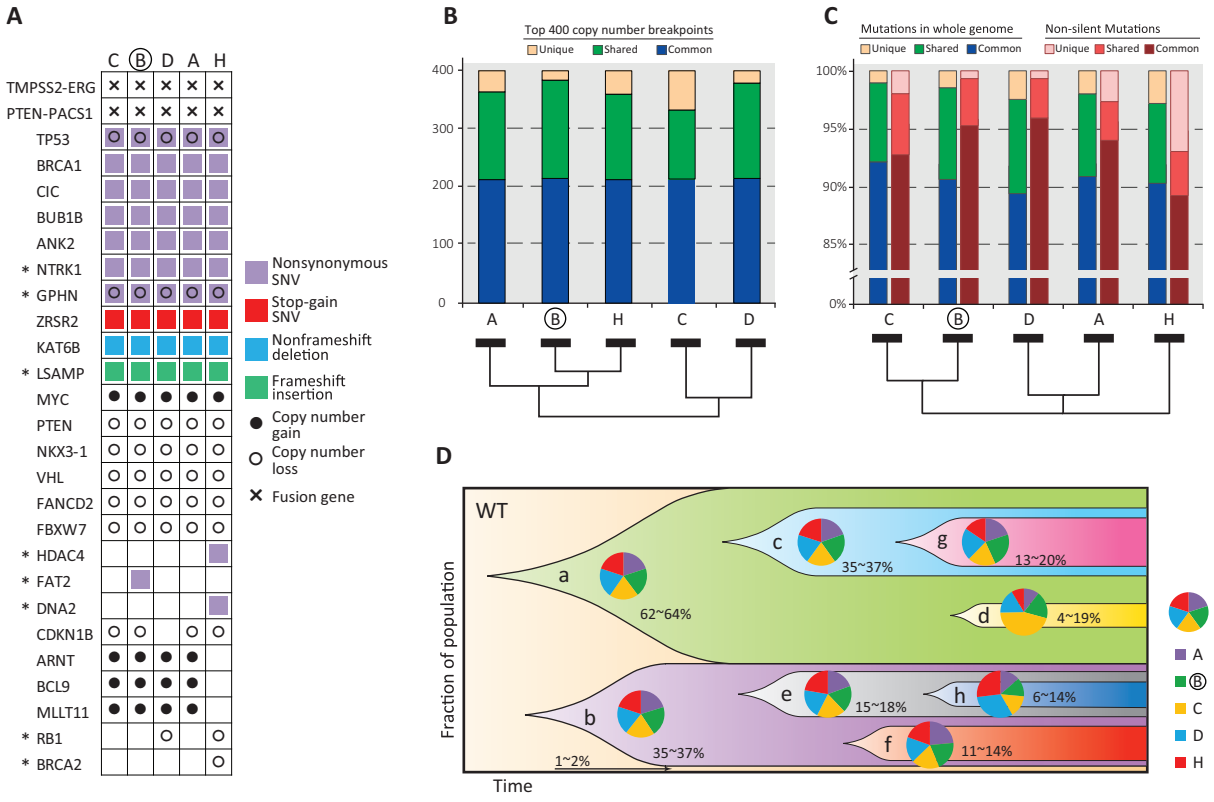


**Fig. 2 – Genomic profiling revealed limited heterogeneity among multiple LTL-313 patient-derived xenograft (PDX) models. (A) Matrix showing the genomic alterations identified in each PDX model. The asterisk before each gene name indicates an alteration event that was not supported by the transcriptome sequencing data. For example, the *RB1* gene had a single copy number loss but no change in expression; the *FAT2* allele contained a nonsynonymous single nucleotide variant (SNV) that was not expressed. (B) The number of unique, shared (by any two models), and common (to all five models) chromosome breakpoints among the five PDX models. The phylogenetic tree was constructed based on the chromosome breakpoints, and demonstrates the inferred evolutionary relationships among PDX models. It shows the copy number profile similarity within each PDX model. (C) The numbers of unique, shared, and common genomic (black bars) and nonsilent (reddish bars) mutations among PDX models. The phylogenetic tree was built based on whole-genome mutations. Again, it did not separate the nonmetastatic model (LTL-313B) from the metastatic models. (D) Reconstruction of the subclonal composition over time was inferred from the mutational profiles of each PDX model. Eight subclones (a–h) were identified; the percentage range indicates the prevalence of each subclone within each PDX model. The pie chart illustrates the relative ratio of each subclone proportion across the five PDX models. In summary, there was no association between the differences in subclonal composition and the differences in metastatic phenotypes of LTL-313 PDX models.**

correlation of the top differentially expressed genes consistently demonstrated the same cluster structures (Fig. 3B and Supplementary Table 7). These results indicate that neither genomic nor RNA-seq features of the tumor can distinguish between nonmetastatic and metastatic models.

### *3.2. Stromal cell transcriptomic profiles are linked to the variable metastatic phenotypes of LTL-313 models*

The stromal components of PCa PDX models were largely replaced by host cells, which is consistent with previous studies on other types of cancer [18–20] (Supplementary Fig. 3). Stromal gene expression profiles were generated using RNA-seq reads that were mapped uniquely to the mouse genome. The proportion of murine reads within PDX models ranged from 3% to 8.6%, in agreement with histopathological analysis of the PDX tumors, and remained consistent between PolyA$^+$ and rRNA$^-$ RNA-seq data (Supplementary Fig. 4). Unsupervised hierarchical clustering of total stromal gene expression revealed two major groups that were also observed in the subcluster heatmap of top differentially expressed stromal genes (Fig. 3C,D and Supplementary Table 8). The first and smaller group comprised data from the two nonmetastatic models (LTL-418 and LTL-313B). However, the second and larger group was formed from all metastatic models. The distinct separation of stromal transcriptomes between the nonmetastatic and metastatic models suggests a link between the expression patterns of the cancer-associated stromal genes and metastatic potential.

### *3.3. A stromal gene signature for stratifying indolent primary PCa from PCa with metastatic potential*

A total of 124 differentially expressed murine stromal genes (60 upregulated and 64 downregulated in metastatic models) were identified by applying stringent filters that ensure the reliability of results (Supplementary material and Supplementary Table 9). IPA analysis (www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis) revealed that upregulated genes were significantly enriched for cellular

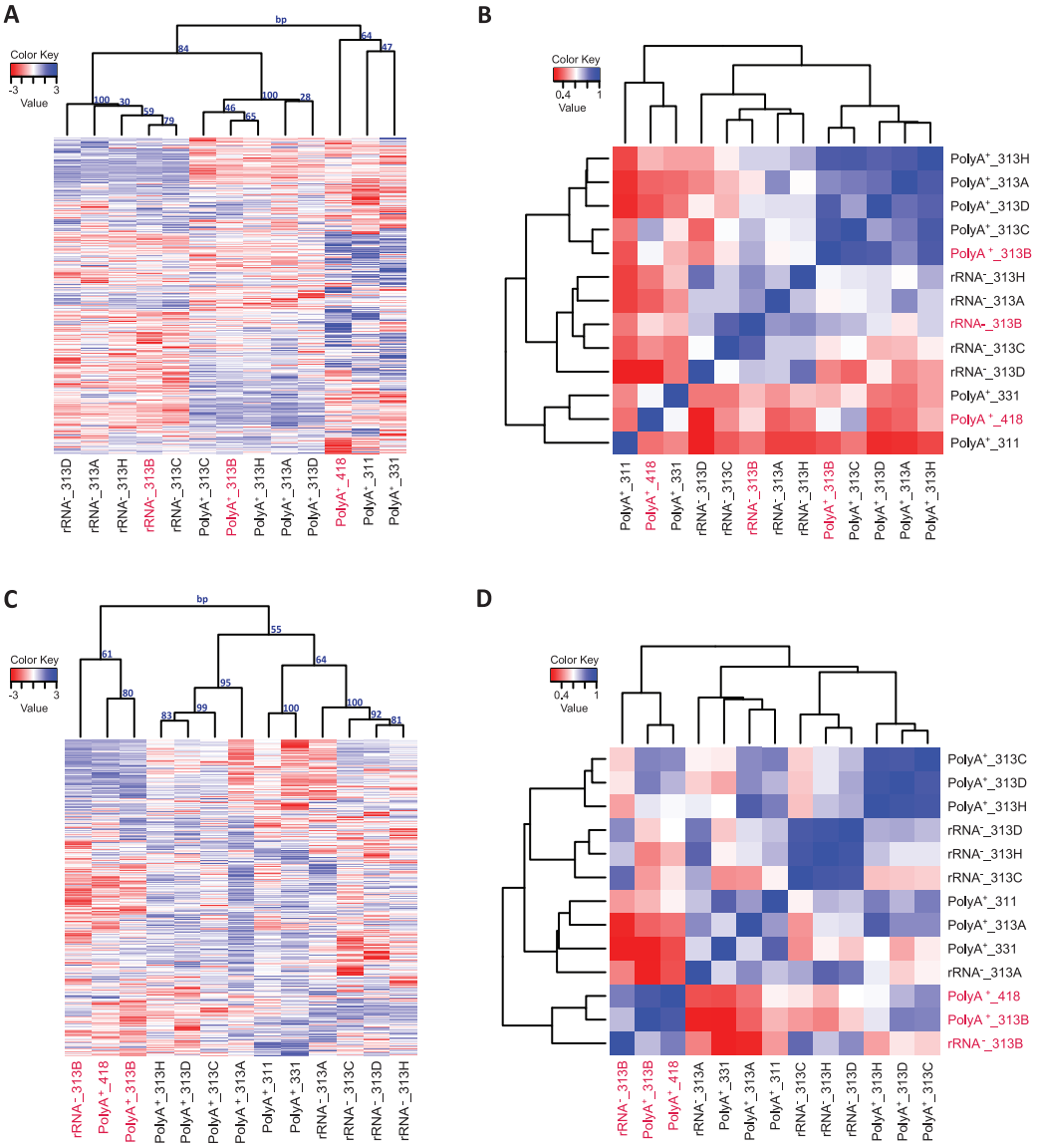


**Fig. 3 – Unsupervised hierarchical clustering based on protein-coding gene expression showed the distinct separation between the stromal transcriptomes of nonmetastatic and metastatic xenografts. (A) Unsupervised hierarchical clustering of total human tumor cell gene expression and 4361 top differentially expressed human tumor cell genes. (B) Subcluster heatmap showing the correlation among patient-derived xenograft (PDX) models. (C) Unsupervised hierarchical clustering of total murine stromal gene expression and 3471 top differentially expressed murine stromal genes. (D) Subcluster heatmap showing the correlation among PDX models. The nonmetastatic xenografts (LTL-313B and LTL-418) were separated from the metastatic xenografts by the top differentially expressed murine stromal genes. Bootstrap probability values are indicated in blue. Nonmetastatic PDX models are highlighted in red.**

movement and cell-to-cell signaling and interaction, implying that a gain in these particular stromal cell functions may promote metastasis (Supplementary Tables 10 and 11, Supplementary Figs. 5 and 6).

Because the cancer-associated stroma appears to play an important role in metastasis, we examined whether human homologs of these differentially expressed murine stromal genes can predict the metastatic propensity of primary PCa. First, ten murine genes without human homologs were removed (Supplementary Table 8). Next, microarray data sets for six RP cohorts used in previous studies [21–26] were assembled (Supplementary Table 12). The Mayo Clinic (MC) I cohort [21], containing 545 patients, was used as a training set to construct a gene signature based on these differentially expressed stromal genes. A generalized linear model with lasso regularization resulted in identification of a 93-gene SDMS as described in the Supplementary material (Supplementary Table 13). Higher SDMS scores indicate a higher metastatic potential.

### 3.4. *Validation of the 93-gene SDMS for risk stratification*

The SDMS was validated for risk stratification using five independent cohorts of patients with primary prostate tumors [22–26] and containing 856 patients in total. For all five cohorts, patients who developed metastases had significantly higher signature scores compared to those who did not develop metastasis following RP ($p < 0.001$;

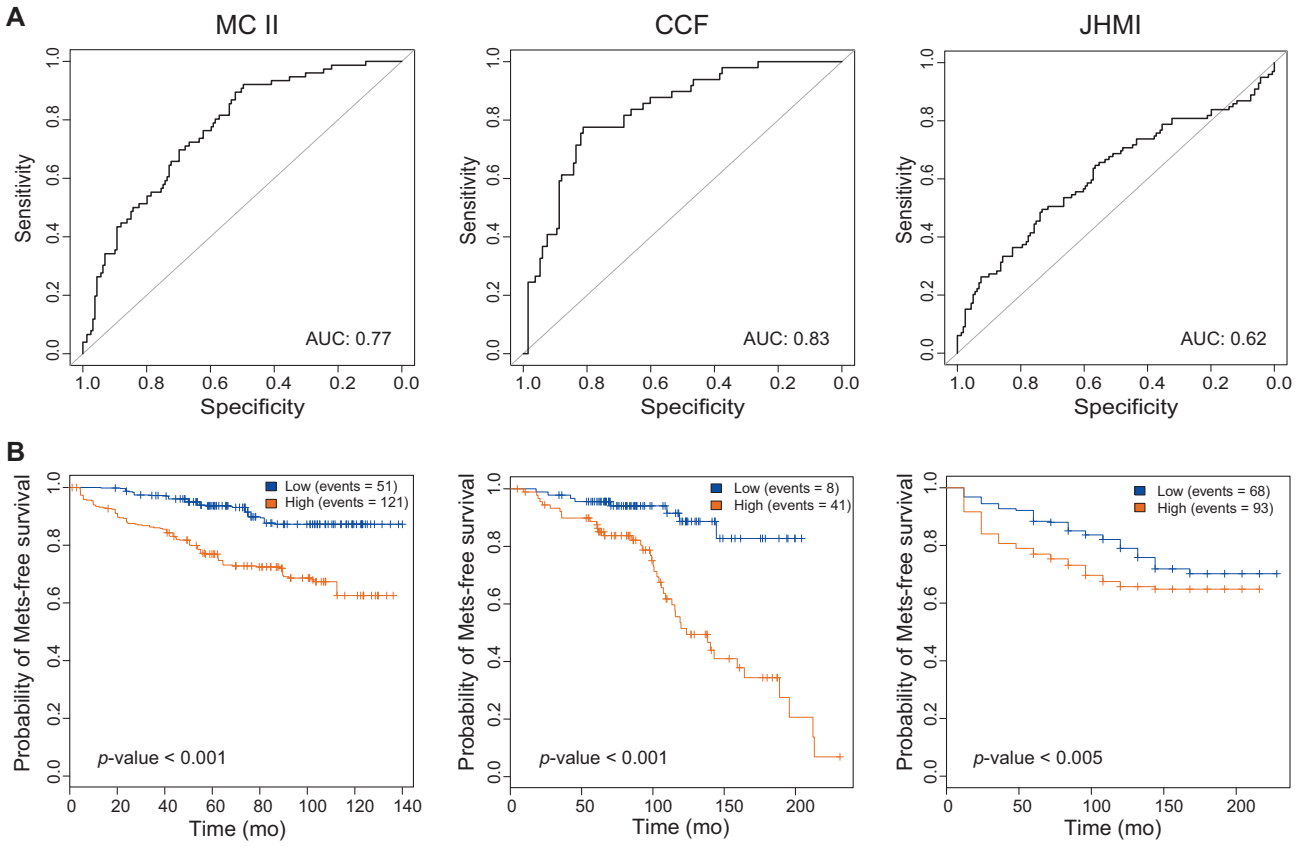


**Fig. 4 – Area under the curve (AUC) and Kaplan-Meier survival analysis for the stroma-derived metastasis signature (SDMS) in multiple independent clinical cohorts demonstrated that the SDMS can distinguish indolent primary prostate tumors from those with metastatic potential. (A) Receiver operating characteristic curves show that the SDMS can separate patients with metastatic potential from patients with indolent tumors: Mayo Clinic (MC) II, AUC = 0.77; Cleveland Clinic Foundation (CCF), AUC = 0.83; Johns Hopkins Medical Institutions (JHMI) AUC = 0.62. (B) Kaplan-Meier curves show that patients from the high-score group, based on a median split of SDMS scores within each cohort (low/high), have worse outcome in terms of metastasis (Mets)-free survival: MC II, $p < 0.001$; CCF, $p < 0.001$; JHMI, $p < 0.005$.**

Fig. 4A, Table 2, and Supplementary Fig. 7). Kaplan-Meier (KM) curves were produced for the three applicable cohorts; curves for the MC II [22] ($p < 0.001$), Cleveland Clinic Foundation (CCF) [25] ($p < 0.001$), and Johns Hopkins Medical Institutions (JHMI) [24] ($p < 0.01$) cohorts demonstrated significantly worse prognosis for patients with higher SDMS scores based on a median split (Fig. 4B). Univariable regression analysis comparing the unadjusted performance of the 93-gene SDMS for predicting metastasis to that of clinicopathological risk factors revealed that the SDMS was more prognostic than the clinicopathological factors in the five validation cohorts (Table 2 and Supplementary Table 14). Using multivariable analyses in which the prognostic significance of the SDMS was adjusted for the clinical and pathological variables, the SDMS was a significant predictor of metastasis, independent of clinicopathological variables, in the MC II and CCF cohorts (Table 2, Supplementary Table 14, and Supplementary material).

**Table 2 – Validation of the 93-gene stroma-derived metastasis signature and its performance and potential utility for patient risk stratification in multiple independent clinical cohorts**

| Cohort | Patients (*n*) | Wilcox (*p* value) | AUC | UVA (*p* value) | MVA (*p* value) | KM (*p* value) |
|---|---|---|---|---|---|---|
| Mayo Clinic I [a] | 545 | 6.36E–46 | 0.86 | 1.72E–30 | 4.37E–20 | N/A |
| Mayo Clinic II | 235 | 4.78E–11 | 0.77 | 5.77E–07 | 1.49E–04 | 5.77E–07 |
| CCF | 182 | 1.06E–11 | 0.83 | 5.00E–07 | 2.28E–02 | 2.89E–06 |
| JHMI | 260 | 1.47E–03 | 0.62 | 4.22E–03 | 1.12E–01 | 4.22E–03 |
| Rotterdam | 48 | 6.56E–04 | 0.85 | 2.77E–02 | 9.77E–02 | N/A |
| MSKCC | 131 | 1.88E–04 | 0.87 | 3.99E–02 | 5.97E–01 | N/A |

AUC = area under the receiver operating characteristic curve; UVA = univariable analysis; MVA = multivariable analysis; KM = Kaplan-Meier analysis; N/A = not applicable; CCF = Cleveland Clinic Foundation; JHMI = Johns Hopkins Medical Institutions; MSKCC = Memorial Sloan Kettering Cancer Center.
[a] Training data set.

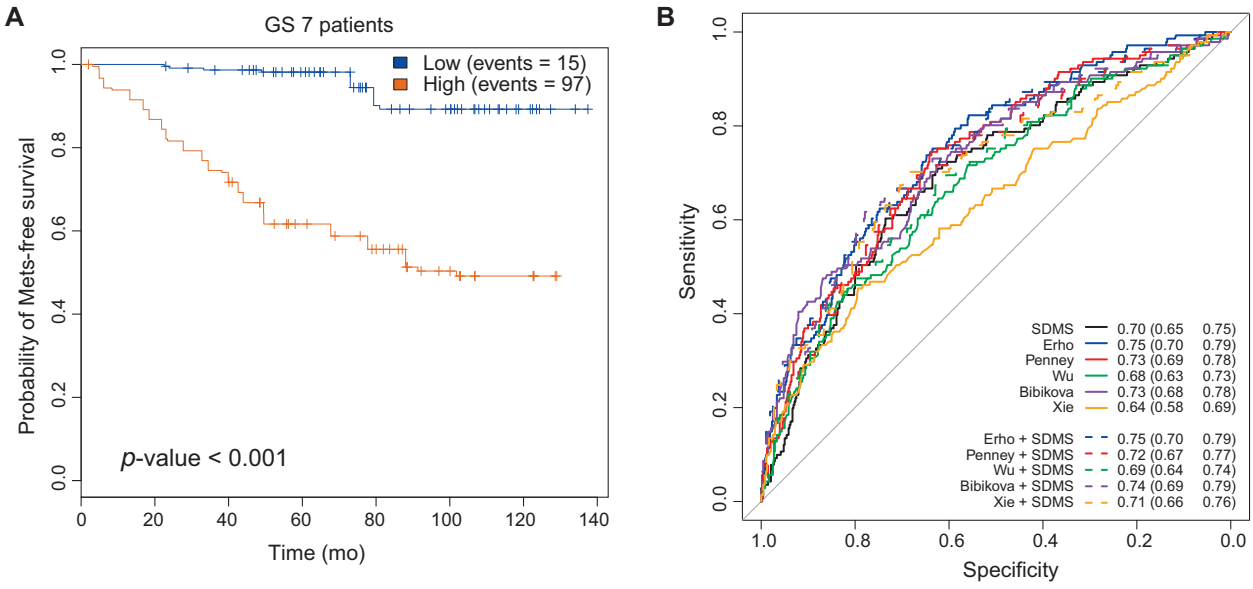


**Fig. 5 – Kaplan-Meier survival analysis for the stroma-derived metastasis signature (SDMS) for patients with intermediate-risk Gleason 7 tumors and improvement in prognostic performance of previously validated prognostic signatures. (A) Kaplan-Meier curves show that intermediate-risk patients with high SDMS scores, based on a cohort median split (low/high), have worse outcome. (B) Addition of the SDMS to previously validated signatures improves their predictive power. Combined logistic regression models were trained in the Mayo Clinic I cohort and evaluated in the pooled validation cohort. An improvement in the area under the receiver operating characteristic curve is observed on addition of the SDMS to the Wu, Bibikova, and Xie signatures.**

### *3.5. The SDMS provides a biomarker for risk stratification of patients with intermediate-risk PCa*

Treatment decisions are particularly problematic for patients with intermediate-risk primary PCa because of the wide range of biochemical and/or clinical recurrences in this group [27,28]. To investigate the prognostic performance of the SDMS in this group, KM curves were generated for a group of patients with Gleason 7 PCa extracted from the MC II cohort. On the basis of a median split, patients with high signature scores had worse prognosis ($p < 0.001$) than those with low signature scores (Fig. 5A). The results suggest that the SDMS is predictive for metastatic potential within the intermediate-risk group.

### *3.6. The 93-gene SDMS provides additional independent information for prediction of PCa risk*

The SDMS was further compared to five other previously reported gene signatures associated with aggressive PCa in predicting metastasis [10–14] in an additional multivariable analysis for a pool of patients from three validation cohorts (total 621 patients) [22–24]. The five signatures selected for comparison were the top performing signatures by *C*-index as reported by Ross et al [24]. The results show that the SDMS is independently prognostic of metastasis within 5 yr following RP. Furthermore, to evaluate whether the SDMS could improve on the predictive power of the previously reported signatures, combined logistic regression models (trained in the MC I cohort) were evaluated in the pooled validation cohort to evaluate any improvement in AUC. The SDMS improved the AUC for three previously reported gene signatures: a significant improvement in AUC was observed for the Xie [13] and Wu [12] signatures; an improvement was also seen with the Bibikova [10] signature, but this did not reach statistical significance (Fig. 5B and Table 3). These results suggest that the SDMS captures additional prognostic information (Table 3 and Supplementary Table 15) and could improve the predictive power of the previously reported signatures.

**Table 3 – Multivariable analysis (MVA) of the stroma-derived metastasis signature (SDMS) and external signatures, and improvements in the area under the receiver operating characteristic curve (AUC) by combining SDMS with external signatures**

| | SDMS | Erho [14] | Penney [11] | Wu [12] | Bibikova [10] | Xie [13] |
|---|---|---|---|---|---|---|
| **MVA** | | | | | | |
| Odds ratio | 1.85 | 2.06 | 1.88 | 1.98 | 1.72 | 1.17 |
| 95% confidence interval | 1.24–2.75 | 1.35–3.15 | 1.20–2.96 | 1.31–2.99 | 1.14–2.62 | 0.77–1.76 |
| *p* value | 0.002 | <0.001 | 0.01 | 0.002 | 0.01 | 0.46 |
| **AUC** | | | | | | |
| External signature only | – | 0.75 | 0.73 | 0.68 | 0.73 | 0.64 |
| External signature + SDMS | – | 0.75 | 0.72 | 0.69 | 0.74 | 0.71 |
| *p* value | – | 0.97 | 0.48 | <0.001 | 0.61 | <0.001 |

## 4. Discussion

There is mounting evidence to support the concept that metastasis is mediated by a dynamic and bidirectional interaction between cancer cells and surrounding stromal cells. In light of this, a deeper investigation of the metastasis-related gene expression of tumor-stromal cells may lead to identification of prognostic gene signatures that can be used to improve patient risk stratification.

We used a panel of PDX models originally derived from multiple foci of a single prostate tumor. Although genetically similar, these models exhibit markedly different metastatic abilities. Since the stromal genes within established PDX models are largely of mouse origin [18–20] (Supplementary Fig. 3), these models allow for distinct separation between stromal and tumor gene expression. To the best of our knowledge, this is the first study of global stromal gene expression in PCa PDX models using RNA-seq. Of major interest is our finding that a stromal gene signature can be used to distinguish indolent prostate tumors from those with metastatic propensity in clinical cohorts. This observation is consistent with results from a recent study that reported an association between a lower presence of lymphatic vessels or reduced immune cytotoxicity and metastasis, while patients both with and without distant metastases showed no major differences in cancer cell–related gene expression levels [29].

A large proportion of genes included in the SDMS are related to cellular movement and migration, which was also the most enriched biofunction category. Furthermore, it has been reported that some of the genes in the SDMS are involved in the modulation of cell-cell and cell-matrix interactions (such as *LGALS1*, *LAGLS3*, *SPP1*, *ADAM12*, and *COL14A*), which is consistent with the concept of cancer-stromal interactions during the process of metastasis. Our results support the potential utility of stromal gene signatures as a source of predictive biomarkers for metastasis in primary tumors. This proposed use of stroma-based biomarkers might also reduce the problems associated with tumor heterogeneity and biopsy undersampling. Therefore, the discovery of a stromal gene signature as a prognostic biomarker for primary PCa has the potential to be transformative for various aspects of PCa, ranging from a deeper understanding of tumor biology to patient stratification and drug development. Unfortunately, although whole transcriptomic sequencing was able to distinguish between human and murine transcripts, it was still unable to determine the specific cellular origins of the gene signature. Further investigations using immunohistochemistry and fluorescence-activated cell sorting analysis will be critical for clarification of the cellular origin(s) of the SDMS signature to gain a more nuanced understanding of the cancer-stroma interaction in PCa.

The prognostic value of this gene panel was successfully validated in five independent clinical cohorts using GenomeDx data sets involving a total of 856 patients. More importantly, this SDMS is independent of other clinicopathological factors, which suggests that it can provide additional prognostic value. Furthermore, multivariable analysis demonstrated that the SDMS is a strong independent factor compared to five previously reported signatures. SDMS improved the AUC for three of these signatures, reiterating its independent prognostic value. An association has also been reported between stromal gene signatures and tumor molecular subtypes [30]. However, in the present study we did not find any association between tumor subtypes and SDMS strength, which further indicates the independent nature of the information provided by the SDMS (Supplementary Fig. 8).

In addition, the SDMS may help to fill a critical gap in PCa patient management if its ability to stratify intermediate-grade Gleason 7 PCa is further validated. This study did not include sufficient Gleason 6 cases for validation of the SDMS predictive value in low-risk tumors. Since sequencing of biopsies is becoming common, evaluation of the SDMS in PCa biopsy samples, particularly in PCa with low to intermediate Gleason score, will be possible. Further development of an easy-to-use assay to test the SDMS in formalin-fixed, paraffin-embedded tissue and in fresh-frozen samples using targeted RNA-seq techniques will expand the potential utility and applicability of this signature. Finally, if it is confirmed that the SDMS includes genes shared among different tumor types, this signature might serve as a pan-cancer biomarker for metastatic risk beyond PCa.

## 5. Conclusions

We used species-specific transcriptomic profile analysis of a unique panel of PDX models to develop a 93-gene signature that accurately predicts the metastatic potential of primary PCa, including the PCa subpopulation with intermediate Gleason score. Its predictive value was successfully validated in five large independent clinical cohorts. If implemented in combination with other established biomarkers and clinicopathological factors, this SDMS could provide invaluable information to facilitate better PCa patient management.

***Author contributions:*** Colin C. Collins had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Mo, Lin, Yuzhuo Wang, Collins.
*Acquisition of data:* Davicioni, Karnes, Den, Klein, Schaeffer, Ross.
*Analysis and interpretation of data:* Mo, Lin, Takhar, Ramnarine, Dong, Bell, Volik, K. Wang, Anderson.
*Drafting of the manuscript:* Mo, Lin, Takhar, X. Wang, Gout.
*Critical revision of the manuscript for important intellectual content:* Gleave, Davicioni, Sahinalp.
*Statistical analysis:* Takhar, Erho.
*Obtaining funding:* Collins, Yuzhuo Wang, Yinghao Sun, Ren.
*Administrative, technical, or material support:* Xue, Yuwei Wang, Haegert, Brahmbhatt, Morris, Ren, Li, Xu, Jun Wang, Jian Wang, Sun.
*Supervision:* Collins, Yuzhuo Wang.
*Other:* None.

***Financial disclosures:*** Colin C. Collins certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria,

stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: Elai Davicioni is the founder, president, chief strategy officer, and director of GenomeDx. Mandeep Takhar and Nicholas Erho are employees of GenomeDx. The remaining authors have nothing to disclose.

***Funding/Support and role of the sponsor:*** This work was supported by Terry Fox Foundation (201012TFF to CC), Prostate Cancer Canada Team Grant (T2013-01 to CC), Canadian Institutes of Health Research (MOP-123449 to YuzhuoW), Prostate Cancer Canada and the Movember Foundation (D2014-24 to YuzhuoW), Urology Foundation for LAST project (889494399 to YuzhuoW), and BC Cancer Foundation (1NSRG030 to YuzhuoW), National Natural Science Foundation of China (81472397 to Ren), National Key Research and Development Program (2016YFC0902200 to Sun), Shanghai Municipal Science and Technology Commission Grant (14411950100 to Sun), Shanghai Municipal Health and Family Planning Commission (2013ZYJB0101 to Sun). The sponsors played no direct role in the study.

***Acknowledgments:*** We thank Stephane Le Bihan for helpful coordination and Michael Cox for helpful discussions.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.eururo.2017.02.038.

## References

[1] VanderWeele DJ, Brown CD, Taxy JB, et al. Low-grade prostate cancer diverges early from high grade and metastatic disease. Cancer Sci 2014;105:1079–85.
[2] Daskivich TJ, Chamie K, Kwan L, et al. Overtreatment of men with low-risk prostate cancer and significant comorbidity. Cancer 2011;117:2058–66.
[3] Welch HG, Albertsen PC. Prostate cancer diagnosis and treatment after the introduction of prostate-specific antigen screening: 1986-2005. J Natl Cancer Inst 2009;101:1325–9.
[4] Prensner JR, Rubin MA, Wei JT, Chinnaiyan AM. Beyond PSA: the next generation of prostate cancer biomarkers. Sci Transl Med 2012;4:127rv3.
[5] Lalonde E, Ishkanian AS, Sykes J, et al. Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: a retrospective cohort study. Lancet Oncol 2014;15:1521–32.
[6] Paris PL, Weinberg V, Albo G, et al. A group of genome-based biomarkers that add to a Kattan nomogram for predicting progression in men with high-risk prostate cancer. Clin Cancer Res 2010;16:195–202.
[7] Ramaswamy S, Ross KN, Lander ES, Golub TR. A molecular signature of metastasis in primary solid tumors. Nat Genet 2003;33:49–54.
[8] Ross RW, Galsky MD, Scher HI, et al. A whole-blood RNA transcript-based prognostic model in men with castration-resistant prostate cancer: a prospective study. Lancet Oncol 2012;13:1105–13.
[9] Klein EA, Cooperberg MR, Magi-Galluzzi C, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. Eur Urol 2014;66:550–60.
[10] Bibikova M, Chudin E, Arsanjani A, et al. Expression signatures that correlated with Gleason score and relapse in prostate cancer. Genomics 2007;89:666–72.
[11] Penney KL, Sinnott JA, Fall K, et al. mRNA expression signature of Gleason grade predicts lethal prostate cancer. J Clin Oncol 2011;29:2391–6.
[12] Wu C-L, Schroeder BE, Ma X-J, et al. Development and validation of a 32-gene prognostic index for prostate cancer progression. Proc Natl Acad Sci U S A 2013;110:6121–6.
[13] Xie BX, Zhang H, Wang J, et al. Analysis of differentially expressed genes in LNCaP prostate cancer progression model. J Androl 2011;32:170–82.
[14] Erho N, Crisan A, Vergara IA, et al. Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. PLoS One 2013;8:e66855.
[15] Wyatt AW, Mo F, Wang K, et al. Heterogeneity in the inter-tumor transcriptome of high risk prostate cancer. Genome Biol 2014;15:1–14.
[16] Ruijter ET, Van De Kaa CA, Schalken JA, Debruyne FM, Ruiter DJ. Histological grade heterogeneity in multifocal prostate cancer. Biological and clinical implications. J Pathol 1996;180:295–9.
[17] Lin D, Wyatt AW, Xue H, et al. High fidelity patient-derived xenografts for accelerating prostate cancer discovery and drug development. Cancer Res 2014;74:1272–83.
[18] DeRose YS, Wang G, Lin Y-C, et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. Nat Med 2011;17:1514–20.
[19] Cutz J-C, Guan J, Bayani J, et al. Establishment in severe combined immunodeficiency mice of subrenal capsule xenografts and transplantable tumor lines from a variety of primary human lung cancers: potential models for studying tumor progression–related changes. Clin Cancer Res 2006;12:4043–54.
[20] Delitto D, Pham K, Vlada AC, et al. Patient-derived xenograft models for pancreatic adenocarcinoma demonstrate retention of tumor morphology through incorporation of murine stromal elements. Am J Pathol 2015;185:1297–303.
[21] Nakagawa T, Kollmeyer TM, Morlan BW, et al. A tissue biomarker panel predicting systemic progression after PSA recurrence post-definitive prostate cancer therapy. PLoS One 2008;3:e2318.
[22] Karnes RJ, Bergstralh EJ, Davicioni E, 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.
[23] Klein EA, Yousefi K, Haddad Z, et al. A genomic classifier improves prediction of metastatic disease within 5 years after surgery in node-negative high-risk prostate cancer patients managed by radical prostatectomy without adjuvant therapy. Eur Urol 2015;67:778–86.
[24] Ross AE, Johnson MH, Yousefi K, et al. Tissue-based genomics augments post-prostatectomy risk stratification in a natural history cohort of intermediate-and high-risk men. Eur Urol 2016;69:157–65.
[25] Taylor BS, Schultz N, Hieronymus H, et al. Integrative genomic profiling of human prostate cancer. Cancer Cell 2010;18:11–22.
[26] Boormans JL, Korsten H, Ziel-van der Made AJ, et al. Identification of TDRD1 as a direct target gene of ERG in primary prostate cancer. Int J Cancer 2013;133:335–45.
[27] Mitchell JA, Cooperberg MR, Elkin EP, et al. Ability of 2 pretreatment risk assessment methods to predict prostate cancer recurrence after radical prostatectomy: data from CaPSURE. J Urol 2005;173:1126–31.
[28] D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA 1998;280:969–74.
[29] Mlecnik B, Bindea G, Kirilovsky A, et al. The tumor microenvironment and Immunoscore are critical determinants of dissemination to distant metastasis. Science Transl Med 2016;8:327ra26.
[30] Hägglöf C, Hammarsten P, Strömvall K, et al. TMPRSS2-ERG expression predicts prostate cancer survival and associates with stromal biomarkers. PLoS One 2014;9:e86824.