A Four-Gene Signature Predicts Disease Progression in Muscle Invasive Bladder Cancer

Wun-Jae Kim,1,2 Seon-Kyu Kim,1,2 Pildu Jeong,1,2 Seok-Joong Yun,1,2 In-Chang Cho,3 Isaac Yi Kim,4 Sung-Kwon Moon,5 Hong-Duck Um,6 and Yung Hyun Choi7

1Department of Urology, Chungbuk National University College of Medicine, Cheongju, Chungbuk, South Korea; 2BK21 Chungbuk Biomedical Science Center, School of Medicine, Chungbuk National University, Cheongju, Chungbuk, South Korea; 3Center for Prostate Cancer, National Cancer Center, South Korea, 4Section of Urologic Oncology, Cancer Institute of New Jersey, Robert Wood Johnson Medical School, New Brunswick, New Jersey, United States of America; 5Department of Food and Biotechnology, Chungju National University, Chungju, Chungbuk, South Korea; 6Laboratories of Radiation Tumor Physiology, Korea Institute of Radiological and Medical Sciences, Seoul, South Korea; and 7Department of Biochemistry, Dongeui University College of Oriental Medicine and Department of Biomaterial Control (BK21 Program), Dongeui University Graduate School, Busan, South Korea

There are no reliable criteria to handle disease progression of muscle invasive bladder cancer (MIBC), which strongly influences patient survival. Therefore, an accurate predicting method to identify progressive MIBC patients is greatly needed. The aim of this study was to identify a genetic signature associated with disease progression in MIBC. To address this issue, we analyzed three independent cohorts (a training set, test set 1 and test set 2) comprising a total of 128 MIBC patients. Microarray gene expression profiling, including gene network analysis, was performed in the training set to identify a gene expression signature associated with disease progression. The prognostic value of the signature was validated in test set 1 and test set 2 by microarray and real-time reverse transcriptase polymerase chain reaction (RT-PCR), respectively. The determination of gene expression patterns by microarray data analysis identified 1,320 genes associated with disease progression. Gene network analysis of the 1,320 genes suggested that IL1B, S100A8, S100A9 and EGFR were important mediators of MIBC progression. We validated this putative four-gene signature in two independent cohorts (log-rank test, P < 0.05 each, respectively) and estimated the predictive value of the signature by multivariate Cox regression analysis (hazard ratio [HR], 6.24; 95% confidence interval [CI], 1.58–24.61; P = 0.009). Finally, signature-based stratification demonstrated that the four-gene signature was an independent predictor of MIBC progression. In conclusion, a molecular signature defined by four genes represents a promising diagnostic tool for the identification of MIBC patients at high risk of progression.

© 2011 The Feinstein Institute for Medical Research, www.feinsteininstitute.org
Online address: http://www.molmed.org
doi: 10.2119/molmed.2010.00274

INTRODUCTION

Bladder cancer is responsible for 150,000 deaths annually and is the seventh most prevalent type of cancer worldwide (1,2). Transitional cell carcinoma of the urinary bladder represents over 90% of all bladder cancers, approximately 80% of which are nonmuscle invasive bladder cancer (NMIBC). Although only 20% of bladder cancer patients are diagnosed with muscle invasive bladder cancer (MIBC), the vast majority of cancer-specific deaths are due to MIBC. Moreover, nearly 50% of patients with MIBC already have occult distant metastases at the time of diagnosis (3).

Conventional histopathological parameters such as tumor stage or grade are generally viewed as prognostic factors, and numerous biomarkers have been investigated as prognostic indicators of MIBC (4–8). However, there are no reliable urinary, histological or stage/grading criteria that can predict disease progression or metastasis of MIBC adequately, or inform the choice of a more aggressive approach to therapy, for example, radical cystectomy. Thus, there is a great need for robust methods capable of identifying patients with high-risk MIBC that is likely to metastasize, a facet of the disease that strongly influences patient survival.

In the present study, we investigated putative genetic signatures associated with disease progression in patients with MIBC. Based on the success of previous genome-wide gene expression profiling studies, including work from our laboratory (9–12), we analyzed gene expression patterns in a training and two test co-
horts of MIBC patients to address the heterogeneity of these tumors, and identified a potential four-gene signature that could be applied to identify distinct patient subclasses with different prognoses.

MATERIALS AND METHODS

Patients and Tissue Samples

We used previously published microarray data of bladder cancer specimens from 165 primary bladder cancer patients with histologically verified transitional cell carcinoma (12,13). Among these, 38 MIBC patients who underwent radical cystectomy were selected as an original cohort (training set), and 25 MIBC patients who did not undergo radical cystectomy were chosen as an independent cohort (test set 1) for validation of the predictive model. To verify the significance of the candidate genes even in the heterogeneous MIBC patients regardless of operation or chemotherapy, another independent cohort (test set 2), of which patients were neither included in the training set nor the test set 1, was populated with the mixture of cystectomy and noncystectomy patients (65 primary cases; 36 cystectomized and 29 noncystectomized). To reduce confounding factors for affecting the analyses, any patients diagnosed with a concomitant carcinoma in situ (CIS) lesion were excluded. In total, 128 primary MIBC patients were included in the analysis. Microarray gene expression data were used for analysis of the training set and test set 1 separately, and the final predictive value of the putative gene signature was evaluated using real time reverse transcriptase polymerase chain reaction (RT-PCR) in test set 2. The study design and validation strategy are shown in Figure 1.

All tumors were macrodissected, typically within 15 min of surgical resection. Each bladder cancer specimen was confirmed as representative by analysis of adjacent tissue in fresh frozen sections from cystectomy and transurethral resection specimens, and then frozen in liquid nitrogen and stored at −80°C until use. Collection and analysis of all samples was approved by the Institutional Review Board of Chungbuk National University, and informed consent was obtained from each subject.

Tumors were staged according to the 2002 TNM classification and the 1973 WHO grading system, respectively (14). Patients with clinically localized or locally advanced tumors and good performance status (ECOG 0 or 1) underwent radical cystectomy and complete pelvic lymph node dissection. Patients who were not eligible for radical cystectomy, that is, those with metastatic disease, poor life expectancy or poor performance status (ECOG greater than 2), underwent transurethral resection or biopsy for histopathological diagnosis. Patients with pT3, pT4 or node-positive disease based on the analysis of radical cystectomy specimens, or with metastatic disease with good performance status, received at least four cycles of cisplatin-based chemotherapy. Alternatively, patients of poor general condition, old age and who were reluctant to undergo chemotherapy did not receive chemotherapy. MIBC patients who received radiation therapy for any reason or who had serious complications associated with surgery were excluded from the study. Regardless of the radical cystectomy, the patients who refused or did not perform the imaging workup such as CT scan or MRI at least once per 3 months also were excluded from the study. In this study, progression was defined as local regional recurrence or new distant metastasis in the cystectomized group and ≥20% increment of mass or new distant metastasis in the noncystectomized group.

RNA Extraction, Microarray Data Processing and RT-PCR Analysis

Total RNA was isolated with TRIzol reagent (Life Technologies, NY, USA), according to the manufacturer’s protocol. The quality and integrity of the RNA were confirmed by agarose gel electrophoresis and ethidium bromide staining, followed by visual examination under ultraviolet light. The microarray data were normalized using quantile normalization. Measured gene expression values were log2 transformed and median centered across genes and samples. Primary microarray data are available in the National Center for Biotechnology (NCBI) Gene Expression Omnibus public database (microarray platform, GPL6102; microarray data, GSE13507).

For critical validation of the prognostic value of the molecular signature identified by microarray, RT-PCR was performed using a Rotor Gene 6000 PCR system (Corbett Research, Mortlake, Australia) on test set 2 specimens. RT-PCR reactions in micro-reaction tubes (Corbett Research) contained primers and SYBR Premix EX Taq (Takara Bio Inc., Otsu, Japan). Spectral data were captured and analyzed using Rotor-Gene Real-Time Analysis Software 6.0 Build 14 (Corbett Research). Gene expression was normalized to B-globin expression.

Statistical Analysis

Differences in continuous variables between groups were assessed by the two-sample t test. Categorical variables were
compared using the chi-square test. Univariate Cox regression analysis was performed to evaluate the association between disease progression and gene signature. To classify patients into two groups, a hierarchical clustering algorithm, using the uncentered correlation coefficient as the measure of similarity and average linkage clustering, was applied as described in Eisen et al. (15). The Kaplan-Meier method was used to calculate time to progression, and differences between the times were assessed using log rank statistics.

To explore the relationships between progression-related genes, we examined functional associations among the genes and generated gene networks based on whether there were more interconnected genes than would be expected to occur by chance. The significance of each network was estimated using the scoring system provided by the Ingenuity Pathway Analysis Tool (version 8.5, Ingenuity Systems Inc., Redwood City, CA, USA). The scores were determined by the number of differentially expressed genes within each of the networks and the strength of the associations among the network members.

After gene-to-gene network analysis of progression-related genes in the original cohort, we chose four genes for validation in tumor samples obtained from the independent validation cohorts. Using these four selected genes, a risk score of disease progression for each patient was calculated as the sum of the levels of expression of each gene multiplied by the corresponding regression coefficients (13,16–22). Receiver-operating-characteristic (ROC) curves were used to determine the optimal cutoff point of each risk score that yielded the highest combined sensitivity and specificity for progression. Using these values, patients were classified into good- or poor-prognostic gene signature groups. The prognostic value of the gene expression signature was determined by multivariate Cox proportional hazard regression analysis. Statistical analysis was performed by using R (version 2.10.0, available at http://www.r-project.org/) and SPSS 12.0 software (SPSS Inc., Chicago, IL, USA), and a P value of < 0.05 was considered statistically significant.

All supplementary materials are available online at www.molmed.org.

RESULTS

Gene Expression Signature Associated with Disease Progression in MIBC

The baseline characteristics of the training and validation cohorts are presented in Table 1, and poor general conditions of patients not receiving cystectomy are shown in Supplementary Table S1. The median follow-up periods after surgery among training set, test set 1 and test set 2 were 37.5 months (range: 1.4–154.4 months), 11.1 months (range: 1.0–80.9 months), and 13.4 months (range: 0.2–181.0 months), respectively. In the training set, we identified 1,320 genes for which changes in expression correlated significantly with progression-free survival in MIBC patients (Cox regression analysis, P < 0.05). Based on hierarchical clustering analysis of the expression patterns of these genes, we divided the MIBC samples into two groups: good- and poor-prognostic signature groups. The progression rate of the poor-prognostic signature group was significantly higher than that of the good-prognostic signature group (log-rank test, P = 1.15 × 10^{-4}; Supplementary Figure S1).

Biological Interpretation of a Gene Expression Signature for Disease Progression

To identify the predominant signaling networks that are active in the progression of MIBC, gene network analysis of the 1,320 genes of the progression signature (Supplementary Figure S1) was carried out using Ingenuity Pathways Anal-
the predominant activation of the signaling networks (Supplementary Table S2), the most striking of which was the involvement of several important signaling pathways. Genes enriched in the gene networks, such as cellular growth and proliferation, cell death and cell cycle genes, were enriched in the gene networks. Genes involved in inflammatory responses, infection mechanisms, immunological disease, immune cell trafficking and infectious disease also were present in significant numbers (Supplementary Figure S2).

An examination of enriched genes revealed the involvement of several important signaling networks (Supplementary Table S2), the most striking of which was the predominant activation of the signaling pathway between IL1B and its associated genes (Figure 2). IL1B formed the primary hub of the gene network, with the highest connectivity to the rest of the genes in the network, suggesting that IL1B could be a critical factor of disease progression. We divided MIBC patients into two groups based on the expression level of IL1B. The frequency of progression was significantly higher in the group with IL1B expression levels in the upper 50th percentile as compared with the group with IL1B expression levels in the lower 50th percentile (log-rank test, \( P = 0.046 \); Supplementary Figure S3A). Many of the satellite genes associated with IL1B (that is, CEBPB, CSF2, GBP6, RNASE7, S100A8 and S100A9) (see Figure 2) participate in inflammatory and immune responses, which are the best-characterized functions of IL1B. The expression levels of S100A8 and S100A9 were associated significantly with progression of MIBC (log-rank test, each \( P < 0.05 \); Supplementary Figure S3B, C, respectively), indicating their potential role in MIBC progression. Another satellite hub gene, EGFR, is a downstream effector of IL1B (see Figure 2). Patients with EGFR expression levels in the upper 50th percentile experienced earlier progression as compared with those with EGFR expression levels in the lower 50th percentile (log-rank test, \( P = 0.007 \), Supplementary Figure S3D). These results indicated that EGFR also may be involved in the progression of MIBC.

**Validation of the Predictive Gene Signature**

We used four genes (IL1B, S100A8, S100A9 and EGFR) identified by gene network analysis to validate the putative progression-associated gene expression signature in tumor samples from test sets 1 and 2. The primer sequence information of four genes for RT-PCR analysis is provided in Supplementary Table S3. First, we tested the expression signature on test set 1, which consisted of patients who did not undergo radical cystectomy (Table 1). According to the calculated risk score of progression, patients in test set 1 were classified into good- or poor-prognostic signature groups. Patients in the poor-prognostic signature group had a significantly shorter time to progression than those in the good-prognostic signature group (log-rank test, \( P = 0.03 \), Figure 3A).

We also validated our gene expression signature in tumor samples obtained from test set 2 (Table 1). Patients in test set 2 were divided into good- or poor-prognostic signature groups according to their calculated risk score of progression. The frequency of progression was significantly higher in the poor-prognostic signature group than those in the good-prognostic signature group (log-rank test, \( P = 0.004 \); Figure 3B).

In the validation cohort (test set 2), multivariate Cox regression analysis revealed that the molecular signature of four genes (hazard ratio [HR], 6.24; 95% confidence interval [CI], 1.58–24.61; \( P = 0.009 \)) was an independent risk factor for progression-free survival (Table 2).

**Estimation of the Predictive Value of a Four-Gene Signature for Progression-Free Survival according to Cystectomy and Chemotherapy**

Since cystectomy and chemotherapy are the most influential treatment tools for MIBC, patients were stratified according to these two modalities to create more homogeneous patient groups, and then the prognostic value of the newly-identified gene signature was tested in each group. When the signature-based stratification was applied to patients who received radical cystectomy (training set in Table 1), we successfully identified a population of high risk patients in both the chemotherapy and nonchemotherapy groups (log-rank test, each \( P < 0.05 \) respectively, Figure 4A, B). In the patient group without radical cystectomy (test set 1 in Table 1), we did not observe a significant association between gene signature and chemotherapy (Supplementary Figure S4A, B).

Similar results were obtained when the validation cohort (test set 2 in Table 1) was analyzed by RT-PCR. Among patients that received radical cystectomy, the frequency of progression in the chemotherapy and nonchemotherapy groups was significantly higher for those with a poor prognostic signature than those with a good prognostic signature (log-rank test, \( P < 0.05 \) in each, Figure 4C, D). However, in patients without radical cystectomy, the gene signature did not predict disease progression in either the chemotherapy or nonchemotherapy group (Supplementary Figure S4C, D).

**DISCUSSION**

In the current study, we identified a gene expression signature associated with MIBC progression and demonstrated that this signature predicts the likelihood of bladder cancer progression. Based on the results of gene network analysis, we identified a putative molecular mechanism involving four genes (IL1B, S100A8, S100A9 and EGFR) that might be responsible for disease progression. The validity of this signature as a prognostic indicator was confirmed in independent analyses of two MIBC cohorts.

Radical cystectomy is the gold standard in the treatment of MIBC and provides excellent local control of the disease with a better prognosis than not undergoing radical cystectomy (22,23). Thus, the prognosis is significantly different between cystec-
Figure 2. Gene networks enriched with genes associated with disease progression. Upregulated and downregulated genes in the poor-prognostic signature group are indicated in red and green, respectively. The intensity of color is indicative of the degree of over- or underexpression. Genes without color are not part of the progression signature but are associated with the regulated genes. Each line and arrow represents functional and physical interactions between the genes and the direction of regulation reported in the literature. The networks were generated through the use of Ingenuity Pathways Analysis (Ingenuity® Systems, www.ingenuity.com).
tomized and noncystectomized MIBC patients. To overcome this heterogeneity between patient groups, we analyzed microarray gene expression data separately in an original cohort (training set), which consisted of cystectomized patients, and a validation cohort (test set 1), which consisted of patients without radical cystectomy. We also constructed another independent cohort (test set 2) that contained both cystectomized and noncystectomized patients to estimate the overall predictive value of the gene signature (see Figure 1).

Although considerable effort has been devoted to identifying a prognostic model of MIBC that can provide useful information about survival and treatment options at diagnosis (4–8), the ability to predict the course of disease progression for patients with bladder cancer remains a major clinical challenge. Thus, there is a critical need for methods that are capable of identifying patients with MIBC that are likely to experience disease progression or metastasis. In the present study, we developed a method to predict the progression of primary MIBC based on a specific gene expression signature. We showed that this gene signature has strong predictive value by validation in two independent cohorts (Figure 3) and multivariate regression analysis (Table 2). The results of the current study underscore the effectiveness of this molecular signature as a prognostic indicator in MIBC, and suggest that this signature could be useful in clinical settings.

Expression of a single gene is rarely an indicator of disease progression because progression is regulated by many different mechanisms. The identification of stable and reliable gene-to-gene relationships is an essential step in unraveling the interactions and functional correlations between genes (24). We identified 1,320 genes for which a change in expression was associated with progression-free survival (Supplementary Figure S1), and then performed gene network analysis to identify putative associations among these genes. Interestingly, we found that functional connectivity between IL1B and its downstream effectors was the most prominent pathway associated with MIBC progression (see Figure 2). Among the genes of the IL1B network, the expression signatures of S100A8, S100A9 and EGFR were strongly associated with disease progression (Supplementary Figure S3). The results of recent studies have suggested that IL1B is associated with tumor invasiveness and metastasis (25,26). S100A8 is reportedly upregulated in many cancers, including bladder cancer (27–34), and has been implicated in the regulation of tumor cell proliferation and metastasis (27,35–37). The expression of S100A9 is upregulated in conjunction with S100A8 in many cancers, including gastric cancer (31), prostate cancer (27,28) and colorectal cancer (29,30). Both S100A8 and S100A9 have been implicated in the regulation of cell proliferation (27,35) and metastatic processes (36). Thus, S100A8 and S100A9 may be valuable targets for the prevention of tumor cell migration to premetastatic sites (37). EGFR has been identified as a strong prognostic indicator in many different cancer types, including bladder cancer, in which increased EGFR expression was associated significantly with reduced relapse-free survival (38). Thus, the use of the four-gene signature identified in the current

Table 2. Multivariate Cox regression analysis for prediction of disease progression.

<table>
<thead>
<tr>
<th>Variable</th>
<th>HR (95% CI)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (male versus female)</td>
<td>0.54 (0.14–2.05)</td>
<td>0.37</td>
</tr>
<tr>
<td>Age</td>
<td>1.01 (0.95–1.08)</td>
<td>0.66</td>
</tr>
<tr>
<td>Stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>Reference</td>
<td>–</td>
</tr>
<tr>
<td>T3</td>
<td>1.65 (0.50–5.43)</td>
<td>0.41</td>
</tr>
<tr>
<td>T4</td>
<td>1.25 (0.39–4.05)</td>
<td>0.71</td>
</tr>
<tr>
<td>Node status (no versus yes)</td>
<td>2.01 (0.68–5.93)</td>
<td>0.21</td>
</tr>
<tr>
<td>Metastatic status (no versus yes)</td>
<td>1.96 (0.63–6.16)</td>
<td>0.25</td>
</tr>
<tr>
<td>Grade (G2 versus G3)</td>
<td>2.2 (0.74–6.57)</td>
<td>0.16</td>
</tr>
<tr>
<td>Chemotherapy (no versus yes)</td>
<td>0.27 (0.09–0.8)</td>
<td>0.02</td>
</tr>
<tr>
<td>Radical cystectomy (no versus yes)</td>
<td>0.44 (0.15–1.34)</td>
<td>0.15</td>
</tr>
<tr>
<td>Four gene signature (good versus poor)</td>
<td>6.24 (1.58–24.61)</td>
<td>0.009</td>
</tr>
</tbody>
</table>

aPredicted outcome in Figure 3B was used for analysis (good- or poor-prognostic signature).
study as a predictive indicator could potentially enable more accurate prognosis of heterogeneous MIBC patients at diagnosis, which would allow for individualized treatment and evaluation.

While cisplatin-based chemotherapy is a standard treatment in metastatic bladder cancer, there is some debate about the use of adjuvant chemotherapy for patients with locally advanced disease after radical cystectomy (39,40). In recent trial updates, cisplatin-based combination chemotherapy was able to produce long-term disease-free survival in some groups of patients (41–43). Therefore, markers that can predict responsiveness before chemotherapy would be valuable in terms of being able to discriminate between responders and nonresponders. To date, however, markers that can accurately predict response to chemotherapy have yet to be introduced. We analyzed whether our newly identified signature had predictive prognostic value in relation to chemotherapy. Among cystectomized patients, the four-gene signature successfully predicted disease progression by differentiating between good and poor risk groups in either the chemotherapy or nonchemotherapy group (log-rank test, \( P < 0.05 \) in each, Figure 4). These results demonstrate that this gene signature is a powerful predictor of disease progression in MIBC patients regardless of chemotherapy status. The gene signature might also be useful in determining patients for whom chemotherapy is necessary. In patients without radical cystectomy, however, the gene signature was not predictive of disease progression, regardless of chemotherapy status. Because patients without cystectomy were mainly those of poor general condition or old age and with advanced disease (Supplementary Table S1), life expectancy might be very short in this group, which may explain why the gene signature was not able to predict disease progression or responsiveness to chemotherapy.

For evaluating whether classification methods have statistical significance, a large sample size is needed before expression profiling can be utilized in a clinical setting. A possible limitation of the present study is that the patient cohorts have relatively small cases compared with other genome-wide gene expression studies for non-MIBC tissues. To our knowledge, however, we have not found MIBC-related studies using more patient samples than the present study on a total of 128 MIBC patients. Moreover, the patients and samples in this study have more advantages than other investigations: all materials including clinical information of patients, RNA extraction, and data processing were handled by a single institute, implying that heterogeneity between patient samples was optimally minimized; and to reduce confounding factors for affecting the analyses, any patients diagnosed with a concomitant CIS lesion or only CIS lesion were excluded, as opposed to other studies. Currently, we are continuing advanced validation study of our newly found signature using an enormous patient cohort collected from multiple institutes, hoping to report strongly significant prognostic values of the four-gene signature in near future.

In conclusion, we have demonstrated that an expression signature consisting of four genes (\( \text{IL1B}, \text{S100A8}, \text{S100A9} \) and \( \text{EGFR} \)) is a reliable prognostic indicator...
of progression in MIBC, independent of traditional pathologic prognostic parameters. Identification of patients with high-risk MIBC may improve the effectiveness of treatments currently available and provide opportunities for the development of new treatment modalities.

ACKNOWLEDGMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2010-0001729).

DISCLOSURE

The authors declare that they have no competing interests as defined by Molecular Medicine, or other interests that might be perceived to influence the results and discussion reported in this paper.

REFERENCES