Identification of Prognostic Biomarkers for Prostate Cancer

Farhad Kosari,1,2 Jan Marie A. Munz,1,2 C. Dilara Savci-Heijink,1,2 Craig Spiro,1,2 Eric W. Klee,1,2 Dagmar Marie Kube,1 Lori Tillmans,1,2 Jeff Slezak,1,3 R. Jeffrey Karnes,1,4 Farhad Kosari,1,2 Jan Marie A. Munz,1,2 C. Dilara Savci-Heijink,1,2 Craig Spiro,1,2 Eric W. Klee,1,2 John C. Cheville,1 and George Vasmatzis1

Abstract

Purpose: This paper describes a process for the identification of genes that can report on the aggressiveness of prostate tumors and thereby add to the information provided by current pathologic analysis.

Materials and Methods: Expression profiling data from over 100 laser capture microdissection derived samples from nonneoplastic epithelium; Gleason patterns 3, 4, and 5 and node metastasis prostate cancer were used to identify genes at abnormally high levels in only some tumors. These variably overexpressed genes were stratified by their association with aggressive phenotypes and were subsequently filtered to exclude genes with redundant expression patterns. Selected genes were validated in a case-control study in which cases (systemic progression within 5 years) and controls (no systemic progression at 7 years of follow-up) were matched for all clinical and pathologic criteria from time of prostatectomy (n = 175). Both cases and controls, therefore, could have nodal invasion or seminal vesicle involvement at the time of initial treatment.

Results: A number of candidate variably overexpressed genes selected for their association with aggressive prostate cancer phenotype were evaluated in the case control study. The most prominent candidates were SSTR1 and genes related to proliferation, including TOP2A.

Conclusions: The process described here identified genes that add information not available from current clinical measures and can improve the prognosis of prostate cancer.

The current prognostic indicators predictive of outcome in men with prostate cancer after radical retropubic prostatectomy consist of Gleason score (GS), tumor-node-metastasis stage, surgical margin status, and preoperative serum prostate-specific antigen (1–4). Whereas stratification by these variables is an effective predictor of the course of disease in many instances, there are tumors with similar biochemical, histopathologic, and clinical conditions that behave very differently. For example, a recent study emphasized the heterogeneity among GS7 tumors: there is a statistically significant greater risk for progression to systemic disease among GS(4+3) tumors compared with GS(3+4) tumors (5). Thus, Gleason scoring can help distinguish more aggressive from less aggressive tumors, but, even within each group [GS(3+4) and GS(4+3)], there is uncertainty as to risk of systemic disease (as low as 15% for 3+4 and as high as 40% for 4+3; ref. 5). A supplementary prognostic measure could help identify patients needing postoperative adjuvant therapy or reassure low-risk patients that adjuvant therapy is not necessary.

Molecular profiles can identify characteristics of tumor cells not yet within the scope of pathologic and clinical evaluation. When constructing a molecular profile of prostate cancer, it is important to consider the heterogeneity of the disease. Heterogeneity among prostate tumors was an essential component in the analysis that identified recurrent oncogenic chromosomal rearrangements in prostate cancer (6). Tomlins and coworkers used cancer outlier profile analysis of microarray data to detect associations between recurrent gene fusions and prostate cancer (6). Their analysis selected genes with outlier profiles across tumors by a mathematical transformation of the microarray data based on the median and median absolute deviation of expression profiles (6). The success of this approach suggests that other strategies for finding prognostic markers should take into account the molecular heterogeneity of prostate cancer (PCa).

Optimally, the discovery of molecular markers of progression would occur through the analysis of tissue samples with long-term patient follow-up. However, use of current high-throughput...
expression platforms, such as microarray gene chips, require RNA with quality only available from frozen tissue samples, which often lack associated clinical follow-up. To best exploit frozen samples, the studies described here used two-phase design, wherein (a) candidate markers were identified using expression profiles from frozen tissues, lacking long-term follow-up information, but with Gleason pattern (GP) and node metastasis as surrogate for aggressiveness and (b) selected candidates were subsequently evaluated in formalin-fixed paraffin-embedded (FFPE) samples with long-term follow-up.

Candidate markers were identified by the analysis of gene expression data from 61 primary tumor samples with diverse Gleason architectural patterns, seven lymph node metastatic PCa, and 34 nonneoplastic prostate, benign prostatic hyperplasia (BPH), and prostatic intraepithelial neoplasia cases. From each patient sample, cells of a single GP were captured by laser capture microdissection (LCM) under the guidance of a pathologist to assure that the collected cells were either tumor or nonneoplastic epithelial control cells, avoiding other cell types in the prostate tissue.

Genes expressed at high levels in a subset of tumors [variably overexpressed genes, (VOG)] were identified. Similar approaches have been described for identification of significant genes in cancers (7, 8), including prostate (6). These VOGs were then evaluated for association with aggressiveness (using GP and node metastasis as surrogate). Among the most promising VOGs were genes with nearly identical expression patterns. To minimize apparent redundant information, several of these candidate markers were excluded from the phase II evaluation. Finally, candidates were tested as predictors of tumor behavior in a case-control study. The cases (progression to systemic disease within 5 years of surgery) and controls (no progression within 7 years) were matched for the existing clinical and pathologic prognostic variables, including grade, stage, and preoperative serum levels of prostate-specific antigen. The expression of several genes that correlated with patient outcome in this study provide prognostic information beyond what is currently available, as the case-control was balanced on all the existing clinicopathologic variables. These markers could be included in a clinical assay designed to help clinicians and patients make informed treatment decisions.

Materials and Methods

Discovery phase

The prognostic marker discovery for prostate cancer was based upon comparisons of gene expression profiles obtained from the epithelial cells of nonneoplastic tissue, primary tumor, and metastatic tumor samples. Details on the procedure for sectioning and processing of slides, LCM, and linear amplification have been recently reported (9). This protocol required 2 ng of starting total RNA and used a slightly modified version of the Affymetrix small sample preparation protocol (Affymetrix Corp.).

Microarray gene expression data

The microarray data analyzed included expression profiles of 102 LCM-derived collections of prostate cells from patients with prostate tumors. Nineteen samples were from nonneoplastic prostate epithelium adjacent to tumor, 10 from BPHs, and five from high-grade prostatic intraepithelial neoplasias. The tumor samples included GP3 cells from 28 GS(3+3) cases and 3 GS(3+4) cases; GP4 cells from 10 GS(4+4) cases, 4 GS(4+3) cases, and 6 GS(4+5) cases; and GP5 cells from five GS(5+5) cases and five GS(5+4) cases. Tumor samples also included seven PCa metastases from positive nodes.

Bioinformatics algorithms to identify PCa prognostic genes

Probe set expression values were obtained from the raw microarray data using dChip with invariant set normalization and the PM-only and PM/MM models. Bioinformatics methods developed in this project combined three steps (Fig. 1) to select candidate markers.

In the first step, probe sets with significantly higher expression levels in at least a subset of tumor samples compared with nonneoplastic cells were identified. To do this, the maximum expression level for each probeset in nonneoplastic and BPH cases (\(B_{\text{max}}\)) was determined. Next, probesets with at least 2-fold higher expression than \(B_{\text{max}}\) in at least 10% (seven cases) of tumor samples were identified. Of these probesets, those with mean expression levels in the “overexpressed” cases that were 4-fold higher than mean expression levels in the nonneoplastic and BPH cases were labeled VOGs and selected for further analysis.

In the second step, a surrogate score of aggressiveness was assigned to each VOG by comparing expression data from GP4, GP5, and MET samples (group 1), representing aggressive disease, to expression data from GP3 samples (group 2), representing the nonaggressive disease. For each VOG, the percentage of overexpressed samples was calculated for group 1 (\(P_{\text{ag}}\)) and group 2 (\(P_{\text{nag}}\)). Based on these percentages, a \(P_{\Delta p}\) value was assigned to each VOG, wherein \(P_{\Delta p} = P_{\text{ag}} - P_{\text{nag}}\), and those with high \(P_{\Delta p}\) values were analyzed further.

![Fig. 1](https://www.aacrjournals.org/clinres/2008/14/6/1735_Fig1.png)

**Fig. 1.** An overview of the analytic methods. Data from LCM-derived microarray data are analyzed by a three-step bioinformatics process. The first step selects VOGs that are subsequently ranked by association with aggressive tumors in the second step (see text). Finally, the third step filters the list of selected genes for experimentation by correlation analysis. The most promising genes are validated on FFPE samples with outcome data.
Finally, in the third step, the number of selected VOGs was further reduced by grouping those with the same expression pattern across the 102 microarray samples and eliminating redundancies. To do this, starting from the top of the list produced in step 2, a Pearson’s correlation coefficient was calculated between each probeset in the list and all other selected probesets. Probesets that had a correlation coefficient of >0.7 were grouped into separate clusters. Genes within a group were speculated to have similar prognostic information, and only a subset of these genes was selected for further evaluation.

This candidate selection method was contrasted to a more traditional method for selecting candidate markers, which used P value and fold change (PFC approach), to evaluate if these methods identified the same subset of genes. dChip software was used to select probesets present in at least 40% of tumor cases with a mean value overexpression of at least 2.5-fold (P < 0.0005) in tumor cases compared with non-neoplastic and BPH cases. This comparison identified 248 probesets (the 90th percentile false discovery rate was 0%). Probesets detected as present in at least 40% of group 1 (Pag) samples and overexpressed at least 2.3-fold (P < 0.0005) in group 1 compared with group 2 (Pnag) were also selected. This comparison identified 120 probesets (the 90th percentile false discovery rate was 1%).

Evaluation phase

Selected genes were evaluated in a case-control study using independent samples for which follow-up data were available. The study used FFPE samples from men who had prostatectomy surgery at the Mayo Clinic from 1990 to 2000. Cases consisted of men with GS7 or higher prostate cancer who had systemic progression within 5 years of prostatectomy. Controls consisted of men with GS7 or higher prostate cancer who did not present with systemic progression within 7 years of surgery (all but one control sample remained free of systemic progression through the latest available follow-up information). Controls and cases were balanced on GS, age, preoperative serum prostate-specific antigen level, year of surgery, tumor-node-metastasis stage, and margin/nodal invasion. One hundred sample pairs were initially randomly selected for the study. Samples with insufficient tissue for processing or with highly degraded RNA were excluded from analyses. Supplementary Table S1 describes the clinicopathological characteristics of the case-control subjects. All laboratory personnel were blinded to case-control status.

Laboratory methods

**Processing the FFPE samples.** In all experiments during the evaluation phase, the processing of samples was randomized to prevent processing biases. Each case was reviewed by a pathologist, and tumor was identified on H&E-stained sections. Subsequent sections (10 μm) from each case were prepared under RNase free condition and deparaffinized in xylene. Identified tumor areas were scraped into 1.5-mL tubes containing digestion buffer (RecoverAll kit, Ambion). Total RNA was isolated according to the RecoverAll RNA isolation procedure and treated with DNase using Ambion Turbo DNA free kit according to manufacturer’s instructions (Ambion). The amount of RNA in each case was measured by the Quant-iT RiboGreen kit (Invitrogen). Reverse transcription was done using Superscript III First Strand Synthesis System (Invitrogen) and 500 ng of RNA from each case in a 40-μL reaction volume.

Quantitative PCR was done on each sample by adding 12.5 ng total RNA equivalent of cDNA to a 20-μL reaction volume for each gene of interest using SYBR Green PCR Master Mix (Applied Biosystems ABI) on the ABI 7900HT real-time PCR machine using the manufacturer’s default cycling conditions. Primers for quantitative PCR were designed using Primer Express software (ABI) to amplify a 70-85 bp fragment.

**Imaging, Diagnosis, Prognosis**

![Image](http://www.dchip.org)

**Fig. 2.** A, illustrates the selection thresholds used in the algorithm to identify VOGs. \( B_{\text{max}} \) and \( \mu_B \) represent the maximum and the mean expression levels in nonneoplastic and BPH samples, respectively. \( \mu_T \) represents the mean expression level in the tumor samples that overexpress the gene. To qualify as a VOG, the ratios \( \mu_T/B_{\text{max}} > 2 \) and \( \mu_T/\mu_B > 4 \) must be found in at least 10% of the tumor samples (samples in the box). B, two examples of probesets selected as VOGs. Green, cyan, blue, pink, yellow, and red columns represent expression in nonneoplastic, prostatic intraepithelial neoplasia, GP3-5, and metastatic PCa, respectively. The top and bottom bar graphs represent expression of Top2a and ERG, respectively. \( \Delta p \) is associated the predicted prognostic significance of VOGs (see text). Samples from the same patient are connected with lines.

---

\[\text{http://www.dchip.org}\]
from the Affymetrix target sequence for the gene of interest. The primer pair concentrations (final, 0.15 or 0.2 μmol/L) were optimized by generating standard curves using a pool of prostate cDNA from normal and tumor tissue. No RT samples were run in a quantitative PCR reaction, and those with cycle threshold (Ct) of <35 for glyceraldehyde-3-phosphate dehydrogenase were considered contaminated with genomic DNA and were reprocessed. Samples with a Ct of >27 for RSP28 were deemed to have too much degradation in the RNA and genomic DNA and were reprocessed. Samples with a 3-phosphate dehydrogenase were considered contaminated with C
approved protocols.

PCR experiments.

Represented by two probesets with shown in purple diamonds (KHRDSB3 is red dots, and candidates that were not in the case-control study are shown in to be associated with systemic progression.

All studies were carried out under Mayo Institutional Review Board–approved protocols.

Results

Processing expression data to find predictors of prostate cancer aggressiveness. Previously acquired gene expression profiles were analyzed from 102 tumor and nonneoplastic LCM-derived samples (9) to identify a group of genes that can be used for prognosis of prostate cancer. The analysis was based on the premise that deranged gene expression underlies the cancer phenotype, but, because prostate cancer is a heterogeneous disease, the pattern of deranged gene expression would not be the same in all tumors. As such, the pool of candidates for predicting the course of disease should include genes expressed at distinctly different levels in a subset of tumors compared with normal cells to account for the heterogeneity of prostate cancer. After this line of reasoning, VOGs in a subset of tumors compared with normal tissues were selected for evaluation (see Materials and Methods). In the absence of long-term follow-up and outcome data in the discovery phase samples, GPs and METs were used as a surrogate measure of aggressiveness. Finally, a filtering step was used to eliminate candidates with suspected information overlap. The resulting list of genes was experimentally evaluated on independent patient samples. The methodology described here identifies a set of genes that is tested for its ability to predict clinical outcome (Fig. 1).

Analysis identified 270 probe sets (representing about 220 genes) that had elevated expression levels in a subset of tumor samples. To be selected as a candidate, a gene had to be overexpressed in at least 10% of tumors. Overexpression was defined by two criteria: (a) the level of gene expression had to be at least twice the highest gene expression level of any nonneoplastic sample (including BPH) and (b) the mean expression level of overexpressed tumor samples had to be at least four times the mean expression level in the nonneoplastic samples (Fig. 2A). TOP2a and ERG are two examples of genes identified as VOGs (Fig. 2B). Even for tumor samples with the same GP, expression levels varied from very low (comparable with expression levels in nonneoplastic samples) to at least four fold higher than normal (Fig. 2B).

To identify genes that are associated with aggressive tumors, the relationship between overexpression and a surrogate measure of aggressiveness, which included GP and (nodal) metastasis, was quantified (Fig. 3). The difference (Δp) between the percentage of overexpressed cases in a group representing aggressive PCa (GP4, GP5, and metastasis) and a group representing nonaggressive tumors (GP3) was calculated. Probesets overexpressed in a higher percentage of aggressive tumors possessed a higher Δp and were suspected to have higher prognostic value. To visualize the potential of VOGs to predict aggressiveness, the expression level of VOGs was plotted against the Δp for probe sets of interest (Fig. 3). Note that Δp varies widely among the overexpressed probesets (−39 to +50). That is, overexpression in cancer cells (fold change) does not predict association with aggressiveness (Δp). ERG and AMACR, previously identified markers of prostate cancer (6, 10), have negative Δp values, yet some of the corresponding probe sets show high levels of overexpression in tumors compared with nonneoplastic cases. This is consistent with the model that not all markers of cancer are effective predictors of aggressiveness (see Discussion). Genes with Δp values of >30 were selected for evaluation.

Information content reduction to identify shortlist of candidates. Among the VOG-Δp candidate list are probesets representing
the same gene, as well as probesets representing different
genes with similar or identical expression patterns (Fig. 4).
Because the criterion for identifying a gene of interest is based
on the level of mRNA expression, genes with identical patterns
of expression can be considered redundant with respect to
prognostic value (such “redundant” genes may contain
valuable mechanistic information). Using genes with the
highest $\Delta p$ as seeds, probesets with a Pearson correlation of
$>0.7$ with the seed were grouped into clusters. For example, one
cluster included TOP2A and ~50 other probesets. A number of
genes in this cluster have been identified as part of a
“proliferation cluster” (11). Within the list of VOGs with high
$\Delta p$, probe sets with the same pattern of expression were
excluded from further experimental analysis.

Candidate markers provide prognostic information beyond the existing clinicopathologic variables. A case-control study was
used to evaluate the association between candidate genes and
PCa outcome. Cases consisted of tissue samples (from original
treatment) from patients who progressed to systemic disease in
<5 years after the initial treatment, and controls consisted of
tissue samples from patients without systemic disease for at
least 7 years after treatment. Cases were matched with controls
for other current clinical and pathologic measures (including margin status, seminal vesicle or nodal involvement, and
preoperative prostate-specific antigen; see Materials and
Methods). By matching cases and controls, only candidate
genes that discriminate patients likely to progress to systemic
disease within 5 years follow-up from those at low risk of
progression would be identified. Furthermore, such candidates
would provide information not currently available to clinicians
(as all currently available prognostic criteria were used to match
cases and controls).

The evaluation of biomarkers across samples from many
individuals requires careful selection of reference genes. We
searched for genes with minimum variation in the expression
microarray data. RSP28 was identified by this process, as its
expression across samples had a SD of ~0.35. Hypoxanthine
phosphoribosyltransferase and glyceraldehyde-3-phosphate
derhydrogenase were also examined based on their use in many
previous studies (12). RSP28 showed minimum variation
across cases and controls (Fig. 5A) and was used in all
subsequent analyses to normalize data.

The case-control samples were analyzed for expression of
the selected genes with $\Delta p$ values of $>30$. Four genes with

---

Fig. 4. A, TOP2a and RRM2, two genes with similar pattern of
expression in PIN GP3 to GP5 and PCa metastasis samples represented by
cyrian blue, pink, yellow, and red dots, respectively. Correlation was not as
high in nonneoplastic and BPH samples (green dots). Probesets within
a cluster with similar pattern of expression (B) and probesets from
different clusters with different pattern of expression (C).
Δp of >30 (RRM2, TOP2A, HSPC150, and CDC2) were grouped in the same cluster. RRM2 and TOP2A were examined by reverse transcription–PCR (RT-PCR), and as expected, a significant correlation between the expression levels of the two genes across the case control samples was observed ($R^2 \approx 0.6$; Fig. 5B). This suggests that these two genes possess significantly overlapping information for PCa outcome prediction.

Approximately 30% of the genes selected for quantitative RT-PCR evaluation had significantly higher expression in cases (systemic progression in <5 years) than controls ($P < 0.02$; Fig. 5C and Table 1; performance of tested genes). To ascertain whether genes showing association with outcome in the case-control analysis were likely to have been identified by standard statistical methods involving $P$ value and fold changes (pFC method; see Materials and Methods), 120 probesets identified by pFC comparison of nonaggressive versus aggressive tumors (Supplementary Fig. S1B) were ranked by fold change (Table 1A; rank 1 corresponds to highest fold change). Four of the five genes listed in Table 1 as having significant association with outcome were not on the list generated by pFC (Table 1). Thus, the analysis method presented in this paper identified a set of candidates not likely to be identified by other means. More importantly, the resulting gene list includes candidates possessing prognostic information not currently available to clinicians.

We also analyzed the 10 highest ranked genes (13 probesets) by the pFC approach in our case-control set. As shown in Table 1B, two of these genes were found to have association with systemic progression. This shows that alternative...
approaches are capable of identifying different markers for aggressive tumors, and there is no one prescribed method for prognostic marker identification.

Among candidates with high $\Delta p$, some failed to show association with systemic progression (Fig. 5D). The computed $\Delta p$ value for HOXC6 was 40, and yet this gene was not significantly overexpressed in cases (see Discussion). ERG ($\Delta p = -14$) and AMACR ($\Delta p = -7$) were predicted to be weakly, negatively correlated with progression. These predictions were confirmed by the RT-PCR results, which did not show a significant difference in the expression between cases and controls.

### Discussion

The aim of this study was to expand the number of molecular predictors of the course of prostate cancer progression. A body of published reports shows the interest in extracting from expression data molecular signatures that associate with clinical or pathologic variables, including recent publications describing molecular correlates for Gleason grading (13, 14). In this study, however, by controlling for clinical and pathologic variables, candidate markers were selected to provide information beyond that currently available to clinicians.

The presented approach included high throughput gene expression profiling in the discovery phase and quantitative measurements of expression levels of candidate biomarkers in the evaluation phase. The state of technology requires intact RNA for expression profiling, which currently can be obtained only from fresh-frozen archives. Because our frozen prostate specimens lacked outcome information, GPs and PCa node metastasis were used as surrogates for prostate cancer aggressiveness. LCM was used to collect phenotypically pure non-neoplastic and prostate adenocarcinoma cells, so that markers could be correlated with tumor architectural pattern. The use of LCM in the discovery phase was essential. The heterogeneity of bulk samples or tissue scrapes compounds difficulties in

### Table 1.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Affy probeset</th>
<th>$\Delta p$</th>
<th>pFC rank</th>
<th>$P$ (RT-PCR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Genes selected for further processing by VOG-$\Delta p$ approach and their performance characteristics by RT-PCR analysis in the case-control set</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NRP1</td>
<td>210510_s_at</td>
<td>48</td>
<td>16</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>TOP2A</td>
<td>201291_s_at</td>
<td>30</td>
<td>—</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>RRM2</td>
<td>209773_s_at</td>
<td>33</td>
<td>—</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>KHRDSB3</td>
<td>209781_s_at</td>
<td>50</td>
<td>—</td>
<td>0.015</td>
</tr>
<tr>
<td>SST1R1</td>
<td>235591_s_at</td>
<td>37</td>
<td>6</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>F5</td>
<td>204714_s_at</td>
<td>32</td>
<td>—</td>
<td>Marginal</td>
</tr>
<tr>
<td>PHCA</td>
<td>222689_at</td>
<td>37</td>
<td>—</td>
<td>Marginal</td>
</tr>
<tr>
<td>B3GnT6</td>
<td>1552834_at</td>
<td>43</td>
<td>37</td>
<td>Marginal</td>
</tr>
<tr>
<td>DSC54</td>
<td>220014_at</td>
<td>30</td>
<td>28</td>
<td>Marginal</td>
</tr>
<tr>
<td>PPFIA2</td>
<td>206973_at</td>
<td>32</td>
<td>—</td>
<td>Marginal</td>
</tr>
<tr>
<td>HOXC6</td>
<td>206858_s_at</td>
<td>40</td>
<td>120</td>
<td>Marginal</td>
</tr>
<tr>
<td>C010A1</td>
<td>217428_s_at</td>
<td>40</td>
<td>35</td>
<td>Marginal</td>
</tr>
<tr>
<td>C20orf102</td>
<td>226973_at</td>
<td>40</td>
<td>—</td>
<td>Marginal</td>
</tr>
<tr>
<td>TEMEM45</td>
<td>230323_s_at</td>
<td>40</td>
<td>—</td>
<td>Marginal</td>
</tr>
<tr>
<td>AMACR</td>
<td>209426_s_at</td>
<td>-14</td>
<td>—</td>
<td>Marginal</td>
</tr>
<tr>
<td>ERG</td>
<td>213541_s_at</td>
<td>-7</td>
<td>—</td>
<td>Marginal</td>
</tr>
</tbody>
</table>

| B. The 10 highest ranked genes/ESTs by the pFC ($p < 0.02$) increase expression in tumors that developed systemic progression within 5 y are shown in bold (Fig. 5C). Genes in the lower part of the table were not associated with systemic progression in our case control study. F5 had a marginal ($p = 0.045$) association with systemic progression. Significant at $p$ values ($>0.02$) are reported in the last column. ST6GALNAC5 had a marginal ($p = 0.03$) association with systemic progression. |
|-----------|---------------|------------|-----------|--------------|
| PPFIA2    | 243273_at     | —          | 1         |           |
| FLJ20016  | 1566882_at    | 82         | 2         |           |
| PPFIA2    | 232073_at     | 38         | 3         |           |
| op94a07.x5| 237973_at     | —          | 4         |           |
| PPFIA2    | 1558295_a_at  | 26         | 5         |           |
| SST1R1    | 235591_at     | 9          | 6         | <0.001     |
| LRRN1     | 226884_at     | —          | 7         |           |
| DIRAS2    | 219619_at     | —          | 8         | <0.001     |
| DIRAS2    | 240122_at     | 60         | 9         | <0.001     |
| ST6GALNAC5| 220979_s_at   | —          | 10        | Marginal   |
| DSCAM     | 1562821_a_at  | —          | 11        |           |
| SLC12A1   | 220281_at     | —          | 12        |           |
| SERPIN1   | 205352_at     | —          | 13        |           |

NOTE: A. most genes that were shown to be associated with systemic progression by quantitative RT-PCR were not on the list of genes selected by the pFC approach. Genes with statistically significant ($p < 0.02$) increase expression in tumors that developed systemic progression within 5 y are shown in bold (Fig. 5C). Genes in the lower part of the table were not associated with systemic progression in our case control study. F5 had a marginal ($p = 0.045$) association with systemic progression.

B. significant at $p$ values ($>0.02$) are reported in the last column. ST6GALNAC5 had a marginal ($p = 0.03$) association with systemic progression.
interpreting expression data, as it is difficult to deconvolute contributions from different sources of heterogeneity in the expression profiles. The use of LCM reduced the contribution of noise from other sources compared with the heterogeneity of prostate cancer itself.

The candidates were tested for association with disease outcome in a case-control study using quantitative RT-PCR on FFPE samples with long-term follow-up information. The case-control study was designed so that currently available clinical and pathologic variables could not distinguish the cases from the controls. They were matched for GS, age, preoperative prostate-specific antigen level, nodal/semenal vesicle invasion, and margin status. Thus, using the battery of available measures, the cases and controls had the same prognosis. The test of the candidate genes was to see whether they contained information not available to clinicians, information that could help predict the course of the patient's disease.

Genes with potential association with aggressive disease ($\Delta p > 30$) were evaluated, and $\sim 30\%$ of the candidates showed promise. The cutoff of $\Delta p > 30$ was used to limit the number of genes for experimental validation based on practical considerations. Some other genes, with lower $\Delta p$, were also found to be associated with systemic progression. For example, GRIN3A with $\Delta p$ of 21, was found to be associated with systemic progression in the case-control study ($P = 0.01$; data not shown).

We have shown that the described approach to identifying candidate genes has little overlap with an alternative analysis based on fold change in expression and $P$ value (see Results and online Supplementary information). Candidates identified by both approaches were not necessarily among the highest ranked, and some of the candidates selected by the pFC approach were found to be associated with aggressive behavior in our case-control analysis (Table 1B). Therefore, it cannot be ruled out that alternative approaches also select promising prognostic markers.

We examined the expression of the genes that validated in our study in publicly available prostate cancer expression profiling data bases on Gene Expression Omnibus (see Supplementary information). We identified three studies that included outcome information (15–17). For the most part, we observed variable overexpression of the selected genes in prostate tumor cases, especially in aggressive or metastatic prostate tumors, as shown by our analyses. These findings strengthen the argument that the methodology presented in this paper is robust as the validated genes were verified in other studies.

Tomlins et al. developed cancer outlier profile analysis to help search for genes overexpressed due to translocation and fusion. The methodology accentuated outliers’ profiles and penalized biomarkers profiles (6). Using a compendium of microarray expression data, their analysis assigned a low priority score to known prostate cancer biomarkers and a high score to genes with outlier profiles in a subset of cases, such as ERG. There are important differences in selection philosophy between cancer outlier profile analysis and the VOG-$\Delta p$ approach presented here. The selection of VOGs does not suppress biomarker profiles but rather selects genes that are significantly overexpressed in a subset or the entire set of tumor samples compared with the nonneoplastic controls. Furthermore, cancer outlier profile analysis was not designed to identify aberrant gene expression within a subset of tumors (i.e., aggressive PCa cases). It is therefore not surprising that the list of top 100 cancer outlier profile analysis candidates had minimal overlap (~5\%) with the candidate lists from both pFC and the VOG-$\Delta p$ approaches (data not shown).

In the discovery data set, ERG is overexpressed in GP3, GP4, and GP5 and PCa metastatic cases compared with nonneoplastic prostate, suggesting that the fusion of TMPRSS2 with ETS gene family contributes to the transformed phenotype in aggressive and nonaggressive tumors. In the case-control study, a univariate association of ERG with systemic progression was not found. Studies that used a different patient sample or clinical end point suggest that the TMPRSS2/ERG fusion may influence PCa outcome (18, 19). A recent manuscript provided evidence that both the specific isoform and the expression levels of TMPRSS2/ERG are associated with aggressive prostate cancer (20).

Several factors may explain why some of the candidates failed to be associated with systemic progression in the case-control study. First, the assignment of aggressive and nonaggressive phenotypes was based on a surrogate rather than on actual patient outcome. Another important factor is that sample selection differed in the discovery and the evaluation/validation phases. Both cases and controls in the evaluation phase included tumors with positive margin status or nodal involvement. Positive margin and nodal involvement are important indicators of aggressiveness. As the aim of the case-control study was to test for information that could supplement current clinical measures, the criteria for markers of aggressiveness were stringent. A gene strongly associated with an invasive phenotype could be lost in the case-control design used to test the candidate genes (as invasion was a property of both case and control samples). One such candidate marker is HOXC6, with a high $\Delta p$ but without significantly altered expression in the case-control study. HOXC6 may have prognostic significance in other settings. Expression of HOXC6 increases in prostate tumor compared with nonneoplastic prostate, and high-level expression is associated with a decrease in apoptosis of prostate tumor cells (21). It is also important to note that our analysis was focused on the measurements of the candidate genes transcripts. It is plausible that some of the candidates that were not found to be significantly associated with systemic progression have protein products that are correlated with prostate cancer outcome. In our study, we did not find a significant association between AMACR transcript levels in the primary tumors and PCa systemic progression. However, a report by Rubin et al. suggests that decreased protein expression level of AMACR measured by scoring of the stained areas in immunohistochemistry experiments is associated with increased risk of biochemical failure and cancer-specific death (10). These apparent discrepancies can be explained by the fact that the two studies used different sample population and expression measurement methodologies.

The decision to select significantly overexpressed genes was based on practical considerations for clinical prognostic assays, which often involve FFPE bulk tissue biopsies. Such specimens inevitably include a mixed collection of adenocarcinoma and nonneoplastic cells. Detection of overexpressed rather than underexpressed genes produces a more sensitive assay against such background. If a sample includes a fixed percentage of aggressive and nonaggressive cell types, a gene that is
overexpressed in aggressive cell types produces a stronger discriminating signal than another gene that is underexpressed by the same fold change. Another consideration is that, when the extent of overexpression of a gene, be it in a subset of aggressive tumors, is substantial (>4-fold), it is more likely to produce a detectable signal in clinical settings. Whereas genes overexpressed by a smaller magnitude may be critical for understanding the natural history of prostate cancer, they are likely to prove of little use in a clinical setting.

$\Delta p$ and correlative analysis were used to reduce the number of candidate genes evaluated in the case-control without losing critical information from the expression analysis. The selection of thresholds and the ordering of the analyses were empirical and based on the data derived from tumor samples in the discovery set. Biological and statistical criteria were considered in treating the data. The methodology should be applicable to other diseases as well, where the ordering of $\Delta p$, correlative analysis, and the threshold selection will be appropriate for the particular disease. For example, for more homogeneous disease, correlative analysis would more aptly be done before the $\Delta p$ analysis, so that the genes with highest $\Delta p$ values will represent clusters with different expression patterns and information content.

Of the genes that were found to be associated with systemic progression in PCa, TOP2A and RRM2 are identified as part of a cell proliferation cluster described by Tabach et al. (11). Members of this cluster, including TOP2A, are likely to be associated with poor outcome in breast (22, 23) and ovarian carcinomas (24, 25) carcinomas. Our microarray data from clear cell renal cell carcinoma also indicates that RRM2 and TOP2A are associated with aggressive clear cell renal cell carcinoma (publication in progress). In prostate cancer, TOP2A has been found to correlate with GS (14). SSTR1 is one of five somatostatin receptor subtypes. These receptors are expressed in a number of tumors, including breast, renal cell, and prostate tumors, and most tumors of neuroendocrine origin (26). Because of their antiproliferative action, somatostatin analogues have been used in clinical trials as a therapeutic agent in somatostatin receptor-positive tumors. In prostate cancer, however, the utility of somatostatin analogues for therapy is still unclear (27). Neuruplin receptor 1 (NRP1) is implicated in a number of developmental processes, including angiogenesis and formation of certain neuronal circuits. Increased expression of NRP1 has been associated with an aggressive angiogenic phenotype (glomeruloid microvascular proliferation) in malignant melanoma (28). In vivo experiments on rats has shown that induced expression of NRP1 in prostate carcinoma cells using a tetracycline-inducible promoter resulted in enlarged tumors with significantly increased tumor angiogenesis. These findings corroborate results from this study and reports by Latil et al. (29), indicating an association between NRP1 overexpression and aggressive PCa. However, a report on prostate cancer cell line has suggested that increased expression of NRP1 is negatively correlated with invasion (30). KH domain containing RNA binding signal transduction associated 3 is an RNA-binding protein that participates in the regulation of alternative splicing and in the control of splice site selection and exon inclusion. KH domain containing RNA binding signal transduction associated 3 is located on 8q chromosome, whose amplification is associated with poor survival in PCa (31, 32). To the best of our knowledge, this is the first report linking SSTR1 and KH domain containing RNA binding signal transduction associated 3 expressions to systemic progression in prostate cancer.

Univariate receiver operating characteristic curves, which provide a measure of the sensitivity and specificity of candidate markers, show a significant association between some candidates and systemic PCa (data not shown). The use of several markers in a multivariate model currently being developed may provide a useful supplement to current clinical measures. The inclusion of other molecular measures, such as genetic and epigenetic alterations, may significantly influence PCa progression and therefore support a more complete model for prognosis.

**Acknowledgments**

We thank the Minnesota Department of Employment and Economic Development from the state’s legislative appropriation for the Minnesota Partnership for Biotechnology and Medical Genomics that provided the microarray data for analysis and Dr. Franklyn G. Prendergast for the critical evaluation of this manuscript.

---

**References**
