

Integrative Microarray Analysis of Pathways Dysregulated in Metastatic Prostate Cancer

Sunita R. Setlur,^{1,2} Thomas E. Royce,^{4,5} Andrea Sboner,⁴ Juan-Miguel Mosquera,^{1,2} Francesca Demichelis,^{1,2} Matthias D. Hofer,^{1,2} Kirsten D. Mertz,^{1,2} Mark Gerstein,^{4,5,6} and Mark A. Rubin^{1,2,3,7}

¹Department of Pathology, Brigham and Women's Hospital, ²Harvard Medical School, and ³The Dana-Farber Cancer Institute, Boston, Massachusetts; ⁴Department of Molecular Biophysics and Biochemistry, ⁵Program in Computational Biology and Bioinformatics, and ⁶Department of Computer Science, Yale University, New Haven, Connecticut; and ⁷The Broad Institute of MIT and Harvard, Cambridge, Massachusetts

Abstract

Microarrays have been used to identify genes involved in cancer progression. We have now developed an algorithm that identifies dysregulated pathways from multiple expression array data sets without a priori definition of gene expression thresholds. Integrative microarray analysis of pathways (IMAP) was done using existing expression array data from localized and metastatic prostate cancer. Comparison of metastatic cancer and localized disease in multiple expression array profiling studies using the IMAP approach yielded a list of about 100 pathways that were significantly dysregulated ($P < 0.05$) in prostate cancer metastasis. The pathway that showed the most significant dysregulation, HIV-1 NEF, was validated at both the transcript level and the protein level by quantitative PCR and immunohistochemical analysis, respectively. Validation by unsupervised analysis on an independent data set using the gene expression signature from the HIV-1 NEF pathway verified the accuracy of our method. Our results indicate that this pathway is especially dysregulated in hormone-refractory prostate cancer. [Cancer Res 2007;67(21):10296–303]

Introduction

Prostate cancer is the second leading cause of cancer-related deaths in men after lung cancer (1). The mechanism of progression from a clinically localized disease to metastatic cancer is not well understood. Moreover, metastatic cancers inevitably become unresponsive to androgen withdrawal therapies. A clearer understanding of the underlying mechanisms would benefit the design of more effective clinical intervention strategies.

A popular approach in understanding the development of various types of cancers has been through the employment of genome-wide expression array analysis that has yielded a vast amount of information about marker genes involved in disease progression. The conventional method of analyzing microarray data has been to systematically examine the pattern of regulation of individual genes (up-regulation/down-regulation) and then to study the most highly dysregulated genes in greater detail. This approach has been useful both in dissecting the functionality of

various genes in cancer progression and in correlating gene expression with clinical outcome (2–5).

However, focusing on individual genes in a microarray data set is not the most efficient method for making use of the vast amount of genome-wide information because only a few highly up-/down-regulated candidate genes can be validated and studied in detail at any given time. For instance, an expression array comparison between benign and prostate cancer tissue yielded >1,000 genes that were significantly up-regulated (corrected P value, Q value < 0.05; refs. 2, 6). Moreover, because proteins are known to function in regulated networks or pathways, a bioinformatics approach that would help identify the dynamics of these protein interactions would substantially improve the analysis of microarray data. Additionally, genes that are unaltered or minimally altered may become significant in the context of a biological pathway. Hence, such an approach would be useful in understanding the cross-talk between pathways that confer upon cells and the transformed phenotype and would maximize the information that can be gleaned from genome-wide expression array data.

Recently, methods have been developed toward analysis of sets of genes or pathways (7–13) from expression array data. Gene set enrichment analysis (GSEA; ref. 9) ranks gene sets by assigning scores according to their enrichment in a particular microarray data set. Additional methods have been developed on the principles of GSEA (7, 14, 15). Other “module”-based approaches (sets of genes that are biologically related; refs. 8, 10) require the definition of a threshold in terms of fold change, or statistical significance, in expression and, therefore, do not incorporate all the genes in the identification of dysregulated gene sets or pathways. All the methods described are focused on analyzing individual expression array data sets. Even when multiple data sets are used, the results of the pathways are computed separately for each study. We present an alternate method, integrative microarray analysis of pathways (IMAP), that allows for meta-analysis of multiple sets of microarray data and can query multiple pathways simultaneously determining the significance of each pathway. We used IMAP to study pathways dysregulated in metastatic prostate cancer using the data from the profiling studies available on the Oncomine database. We validated our results using quantitative PCR (qPCR) and *in situ* protein expression analysis in a prostate cancer tissue microarray.

Materials and Methods

IMAP. The data from multiple expression array studies was obtained from the Oncomine database.⁸ We included expression array data from

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

S.R. Setlur and T.E. Royce contributed equally to this work.

Requests for reprints: Mark A. Rubin, Room C 410-A, Department of Pathology and Laboratory Medicine, Weill Cornell Medical College, 1300 York Avenue, New York, NY 10021. Phone: 212-746-6313; E-mail: rubinma@med.cornell.edu.

©2007 American Association for Cancer Research.

doi:10.1158/0008-5472.CAN-07-2173

⁸ <http://www.oncomine.org>

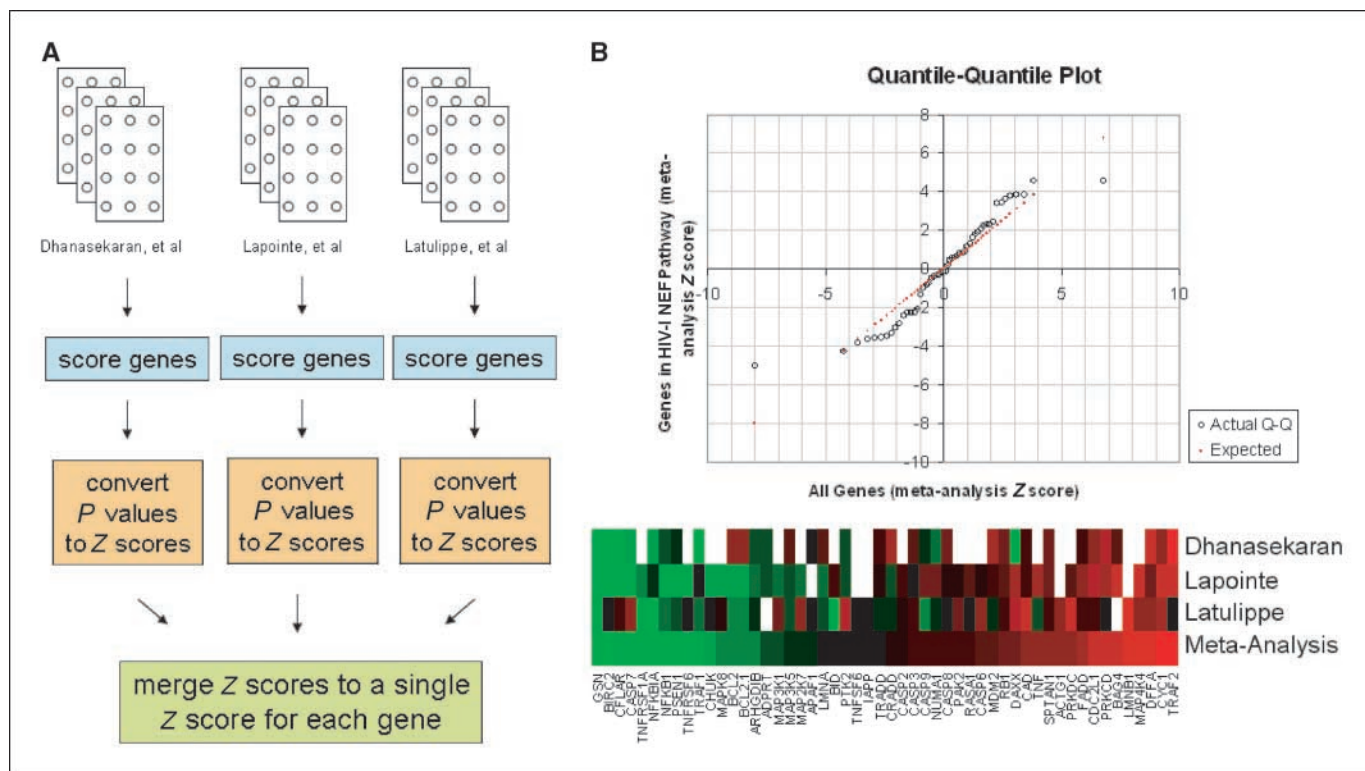


Figure 1. Identification of dysregulated pathways by IMAP. *A*, meta-analysis of gene expression. *P* values are obtained via the Mann-Whitney *U* test. The *P* values are then converted into corresponding *Z* scores, and these *Z* scores are combined using the Liptak-Stouffer method for each gene. A combined *Z* score is used to assign a score for each pathway, and the significance of this score is computed by iterative random permutations. *B*, quantile-quantile plot of the genes constituting the HIV-1 NEF pathway and heat map depiction of gene expression. *Red*, overexpression; *green*, underexpression; *white*, missing data.

three different studies that were found to provide consistent information when analyzed for similarity. To assess for consistency between the three studies, Pearson correlation was computed pair-wise between the mean values of common genes. The three studies showed significant positive pair-wise correlation. The characteristics of the samples included in the three data sets are presented in Supplementary Table S1. The pathways on Biocarta⁹ (*n* = 315) and KEGG¹⁰ (*n* = 137) databases were used as pathway references for the analysis.

Meta-analysis. *P* values were calculated for all genes within each data set using the Mann-Whitney *U* test by comparing localized prostate cancer and metastatic samples. The *P* values were then converted to *Z* scores by mapping onto a standard normal curve. The sign of a gene's *Z* score is then defined to be positive if the median expression level within metastatic samples is greater than that in the localized samples and negative otherwise.

For each gene, *Z* scores from across studies were combined using the Liptak-Stouffer method (16) to obtain a single *Z* score (Fig. 1A). This method for combining different studies requires computing a weighted *Z* score for each gene:

$$Z_{\text{comb},g} = \frac{\sum_{i=1}^m w_i Z_{g,i}}{\sqrt{\sum_{i=1}^m w_i^2}}$$

where *m* is the number of studies, *Z_{gi}* is the *Z* score of the gene *g* in the *i*-th study, and *w_i* is the weight of the *i*-th study. The weights can be assigned according to different criteria like the sample size, recent studies, etc. In the analysis conducted, the criterion of sample size was selected to assign weights.

Pathway analysis. The basic idea of our approach is to assign a score to each pathway and to compute its significance. To begin with, a score *s_g* is assigned to every gene in the pathway *P*.

$$s_g = -\log(p_g)$$

where *p_g* is the *P* value computed from the corresponding *Z_{comb,g}*. In this case, a gene with a low *P* value will have a high score. The gene scores are then summed up for all the genes constituting the pathway *P*, obtaining therefore a pathway score *S*:

$$S = \sum_{g \in P} s_g$$

The significance of *S* was then computed by iteratively computing a score *S** on randomly sampled genes, following the above equation, and counting the number of times *S** is >*S*. One thousand random permutations were done, and a *P* value for the pathway is therefore defined as this count divided by 1,000.

Pathway validation on an independent data set. The Glinsky et al. data set (5) was used for the validation of the HIV-1 NEF pathway. This data set includes 79 localized prostate cancer samples and 8 metastatic samples. The cases are well annotated by Gleason grade and biochemical recurrence (PSA recurrence following prostatectomy). Hierarchical clustering was done using the dChip software.¹¹ We used centroid linkage and correlation as distance metric. The *P* value of the sample cluster is computed via the hypergeometric distribution.

Tissue samples and tissue microarray. Four unmatched localized and hormone-naïve lymph node metastatic samples obtained from the radical prostatectomy program (University of Ulm) were used for the validation

⁹ <http://www.biocarta.com>

¹⁰ <http://www.genome.jp/kegg/pathway.html>

¹¹ <http://biosun1.harvard.edu/complab/dchip/>

Table 1. Top 10 maximally dysregulated pathways in prostate cancer metastasis as identified by IMAP ($P < 0.001$)

Pathway	Database	Number of genes	P
HIV-I NEF pathway	Biocarta	51	<0.001
Cell cycle pathway	Biocarta	23	<0.001
NFKB pathway	Biocarta	20	<0.001
Rho pathway	Biocarta	19	<0.001
MAPK signaling pathway	Kegg	237	<0.001
Cell cycle	Kegg	94	<0.001
TGF- β signaling pathway	Kegg	78	<0.001
Integrin-mediated cell adhesion	Kegg	79	<0.001
Regulation of actin cytoskeleton	Kegg	178	<0.001
Cardiac EGF Pathway	Biocarta	14	0.001

studies by qPCR. A progression tissue microarray (TMA) was constructed from benign prostatic tissues ($n = 10$), localized ($n = 36$), and metastatic prostate cancer samples (18 hormone-naïve metastases, 18 hormone-refractory metastases). Each sample was represented by four TMA cores.

RNA extraction and cDNA synthesis. Frozen tissue blocks with at least 90% tumor were selected. The tissues were sectioned and grossly dissected to enrich for tumor glands before RNA isolation. Six 10- μ m sections were used for RNA extraction. Total RNA was isolated using the Ribopure RNA isolation kit (Ambion Inc.) according to the manufacturer's instructions. The RNA was reverse transcribed to generate first-strand cDNA using the TaqMan reverse transcription kit (Applied Biosystems) with random hexamer priming.

Real-time qPCR. Prevalidated TaqMan gene expression assay probes and primers (Applied Biosystems) were used for the validation of a set of genes [*nuclear factor of κ light polypeptide gene enhancer in B cells 1* (NFKB1), *nuclear factor of κ light polypeptide gene enhancer in B cells inhibitor α* (NFKBIA/I κ B α), *BCL2-associated athanogene 4* (BAG4)/*Silencer of death domains* (SODD), *baculoviral IAP repeat-containing 2* (BIRC2/cIAP1), *caspase-7* (CASP7), *TNF receptor-associated factor 1* (TRAF1), *TNF receptor-associated factor 2* (TRAF2)] from the pathway showing the highest dysregulation (HIV-I NEF). Real-time PCR was done using Opticon monitor 2 (Bio-Rad Laboratories). Quantitation was carried out using the $\Delta\Delta$ CT method. Average expression of the localized prostate cancer cases was considered for the calculation of fold changes. GraphPad Prism software was used for plotting the data.

Immunohistochemistry. The antibodies used for immunohistochemistry were NFKB1 (Biolegend), NFKBIA/I κ B α (Cell Signaling), SODD/BAG4 (Oncogene Research products), TRAF1 (Santa Cruz Biotechnology), and TRAF2 (Cell Signaling). Immunohistochemistry was done as described earlier (17). Briefly, 6- μ m sections were deparaffinized and treated with 10 mmol/L citrate buffer (pH, 6.0). Antigen retrieval was done by pressure cooking. The following dilutions of the different antibodies were used: SODD (1:100); I κ B α (1:50, overnight PC); TRAF1 (1:50, overnight PC); NFKB1 (1:50); and TRAF2 (1:50). The TMAs were evaluated using a semiautomated quantitative image analysis system ACIS II (Chromavision Medical Systems Inc.) that has been previously validated (18, 19). The areas for analysis were selected by the study pathologist (J.-M.M). Cores/cases that were deemed non-assessable by the pathologist were excluded from the analysis. The results were exported and analyzed using the SPSS software.

Results

Pathway analysis. We used three data sets for this study (2, 20, 21). These studies used either the Affymetrix or the spotted cDNA

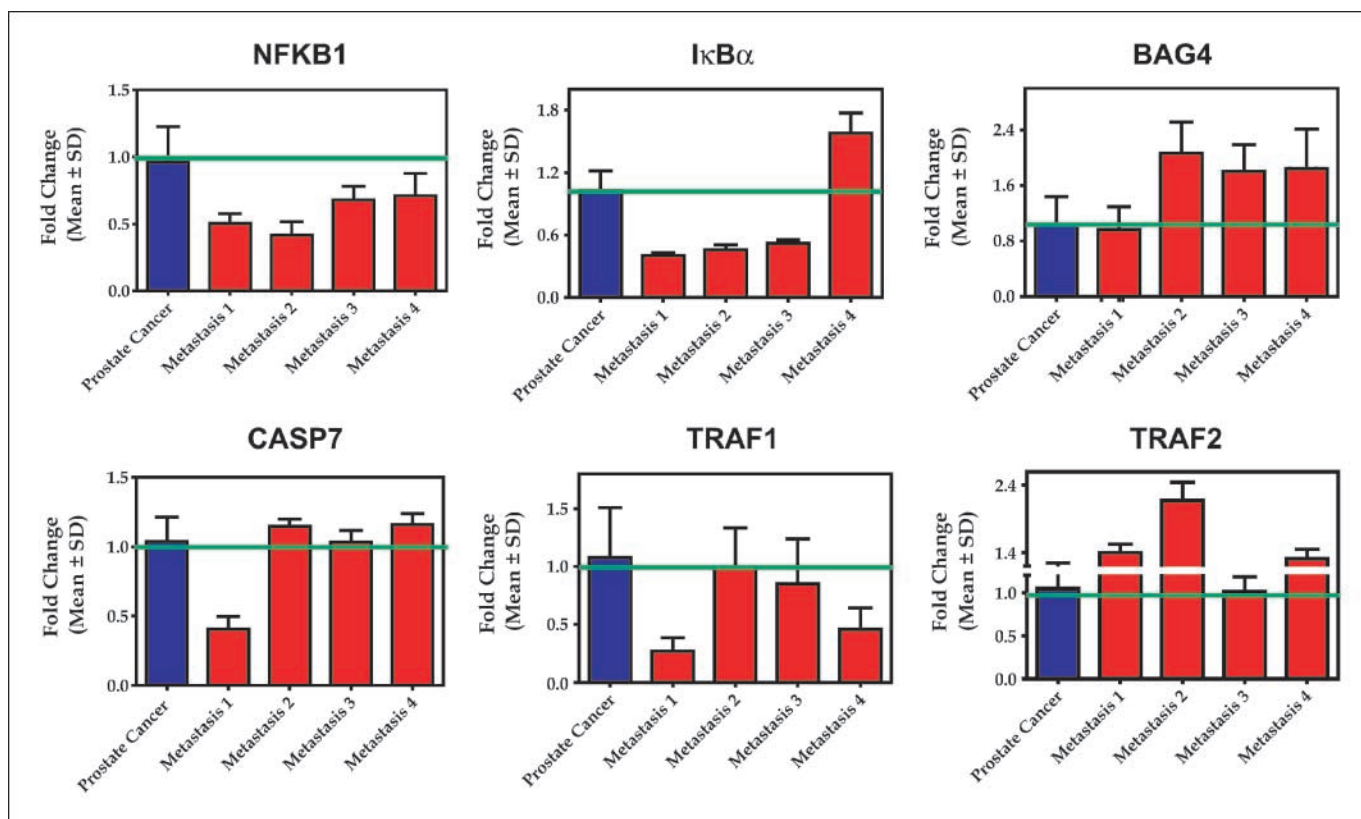


Figure 2. Real-time qPCR analysis of the HIV-I NEF pathway. Expression of all the selected genes (NFKB1/p50, I κ B α /NFKBIA, SODD/BAG4, CASP7, TRAF1, and TRAF2) was consistent with the microarray expression data. Green horizontal line, normalized expression equalized to 1. The results are representative of repeat experiments.

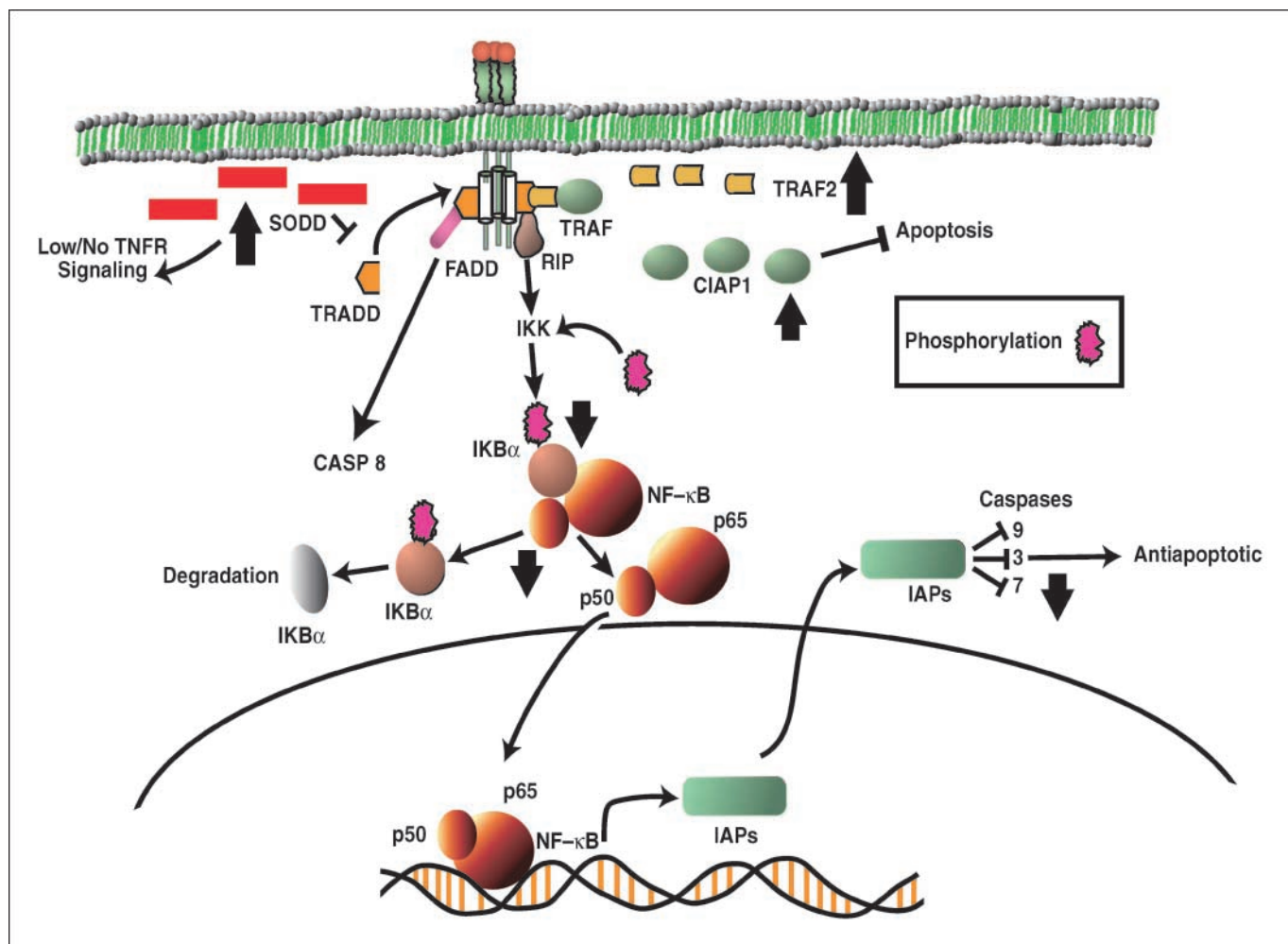


Figure 3. HIV-1 NEF pathway summary. Binding of $\text{TNF}\alpha$ to its receptor results in the activation of $\text{NF-}\kappa\text{B}$ following phosphorylation and dissociation of its inhibitor $\text{NF}\kappa\text{BIA}$. This process is mediated by adaptors (TRADD , TRAF2 , RIP) that recruit the kinase signaling cascade, which is responsible for the phosphorylation of the $\text{I}\kappa\text{B}$ proteins. The microarray and the qPCR data revealed that although the expression of $\text{NF}\kappa\text{B1}$ is low, the $\text{NF-}\kappa\text{B}$ dimer may be constitutively active because the expression of $\text{NF}\kappa\text{BIA}$ is also down-regulated. This would thus lead to the overexpression of IAP like cIAP1 as seen from the expression array and qPCR data and an increase in survival.

microarray platforms for the gene expression analysis. We used the gene symbols to map the different genes across the platforms. The data were downloaded from the Oncomine database. This database provides preprocessed data sets, i.e., "all data are log-transformed, median centered per array, and the SD is normalized to 1." The three data sets used contained benign samples, localized prostate cancer, and a mixture of hormone-naïve and hormone-refractory metastatic samples. The Mann-Whitney U test was used to score genes for being differentially expressed in metastases compared with localized prostate cancers. The gene lists obtained from each study were then merged as described in Materials and Methods (Fig. 1A) to arrive at a meta-gene list. This list was then used for pathway analysis. We found 104 pathways significantly dysregulated in metastasis ($P < 0.05$). Table 1 shows the top 10 dysregulated pathways, which include the cell cycle pathway, the nuclear factor- κB ($\text{NF-}\kappa\text{B}$) pathway, the mitogen-activated protein kinase (MAPK) signaling, and the cell adhesion pathways that have been implicated to play a role in the development of prostate cancer (22). The pathway that was found to be maximally altered was the HIV-1 NEF pathway, which comprises the tumor necrosis factor receptor (TNFR) and the fas

receptor signaling pathways (Fig. 1B). It was seen that the TNFR arm was present in multiple pathways that were significant in the analysis, indicative of its importance. The TNFR pathway involves the activation of $\text{NF-}\kappa\text{B}$ heterodimer by $\text{TNF}\alpha$ and mediates cell survival.

Validation of the HIV-1 NEF pathway on the transcript level.

To validate the IMAP results, the expression of a few key genes in the $\text{NF-}\kappa\text{B}$ pathway that were significantly dysregulated in all three data sets was determined by real-time qPCR analysis using unmatched localized ($n = 4$) and metastatic samples ($n = 4$). The details of the regulation of the individual genes are shown in Fig. 2. Agreement was seen in six of the seven validated genes. The regulation of gene expression within the pathway validated by qPCR is depicted in Fig. 3. The solid up and down black arrows indicate the up- or down-regulation of gene expression. The survival pathway mediated by $\text{NF-}\kappa\text{B}$ seems to be up-regulated as seen by the increase in TRAF2 expression and BIRC2 expression (data not shown).

Validation of the HIV-1 NEF pathway at the protein level. We further examined the *in situ* protein expression of the genes that were consistently dysregulated by qPCR in a larger number of

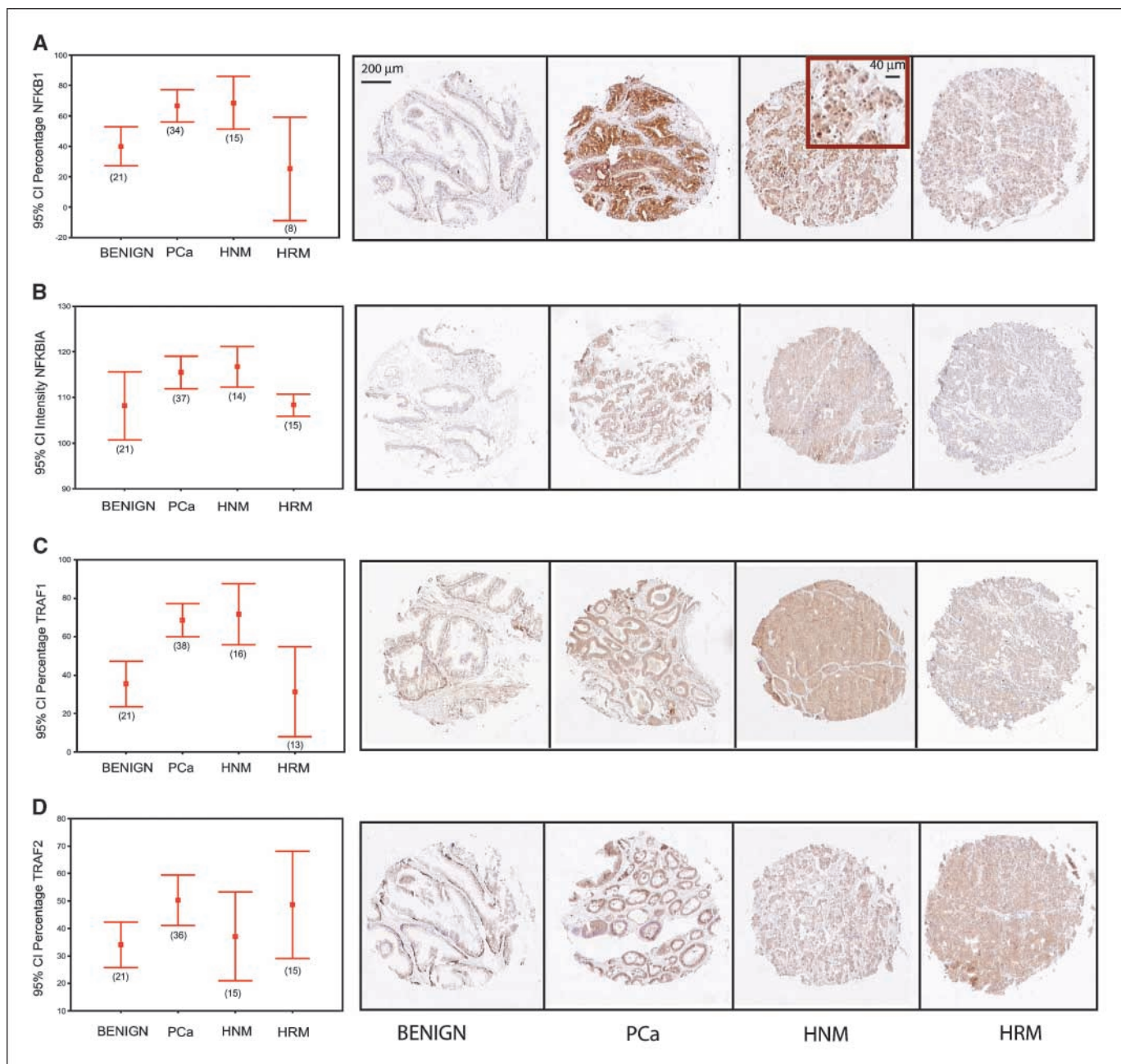


Figure 4. *In situ* analysis of protein expression. The protein expression of (A) NFKB1, (B) NFKBIA, (C) TRAF1, and (D) TRAF2 was determined by immunohistochemistry (magnification, 20 \times). The proteins were seen to be dysregulated particularly in hormone-refractory metastasis. An increase in the nuclear localization was observed for NFKB1 in metastatic samples. *Brown boxed inset*, magnification, 40 \times . *PCa*, prostate cancer; *HNM*, hormone naive metastasis; *HRM*, hormone refractory metastasis.

samples using a progression TMA (Fig. 4). The TMA data showed that NFKB1, NFKBIA/ $\text{I}\kappa\text{B}\alpha$, and TRAF1 were up-regulated in localized prostate cancer and down-regulated in hormone-refractory metastasis. The results with TRAF2 were not conclusive. BAG4/SODD showed an increasing trend in expression from benign tissues to hormone-naïve metastasis, but was also down-regulated in hormone-refractory metastasis (Supplementary Fig. S1). Hence, it seems that this pathway is dysregulated particularly in the hormone-refractory metastasis compared with localized prostate cancer. In addition, increase in nuclear localization of NFKB1 was seen in the metastatic samples, indicating that although the levels of NFKB1 are decreased in

hormone-refractory metastasis, the protein present is transcriptionally active, which could be mediated by low levels of $\text{I}\kappa\text{B}\alpha$.

We further investigated the reciprocal behavior of the proteins by evaluating protein ratios of NFKB1 and its inhibitor or target proteins (NFKB1: $\text{I}\kappa\text{B}\alpha$, NFKB1:TRAF1, NFKB1:TRAF2, and TRAF1:TRAF2; Supplementary Fig. S2). We considered intraclass and interclass expression heterogeneity. We found heterogeneity in the protein expression among the samples belonging to each class. In general, a large percentage of samples expressed low ratios of proteins in benign tissues. The majority of samples expressed high ratios of the proteins in hormone-naïve metastases and localized cancers. The hormone-refractory metastasis, on the other hand,

resembled the benign tissues. In the case of TRAF1:TRAF2 ratios, it was seen that about 40% of hormone-refractory metastasis had high TRAF2 levels, which indicates that the NF-κB pathway is activated in hormone-refractory metastasis.

Validation on an independent data set. The gene signature constituting the HIV-I NEF pathway was further validated on an independent prostate cancer microarray data set (ref. 5; Fig. 5). The Glinsky et al. (23) data set is well characterized and has been used for other validation studies. By reducing the expression array signature to the gene space including the gene signature in the HIV-I NEF pathway and performing hierarchical clustering, the localized prostate cancer segregated from metastatic samples. By further reducing the gene space, it was seen that 26 genes were sufficient to distinguish the two groups (data not shown). Principal component analysis showed that the metastatic samples were clustered closely with the samples that had biochemical recurrence (Supplementary

Fig. S3). Furthermore, these genes were able to significantly distinguish the samples that did not recur (Supplementary Fig. S4). Validation of additional top ranking pathways (Supplementary Fig. S5) showed that the pathways identified by IMAP could distinguish the metastatic samples from localized prostate cancer.

Discussion

We have developed a method to carry out meta-analysis of altered pathways from multiple micorarray data sets. In the current study, we show how this approach can be used to identify critical pathways in prostate cancer progression. Several other approaches have been developed toward determining the regulation of molecular interactions from microarray data sets. Some of these like the GSEA depend on defining sets of genes and querying them on a single data set (7, 9, 12–14), whereas others have focused on finding sets of genes

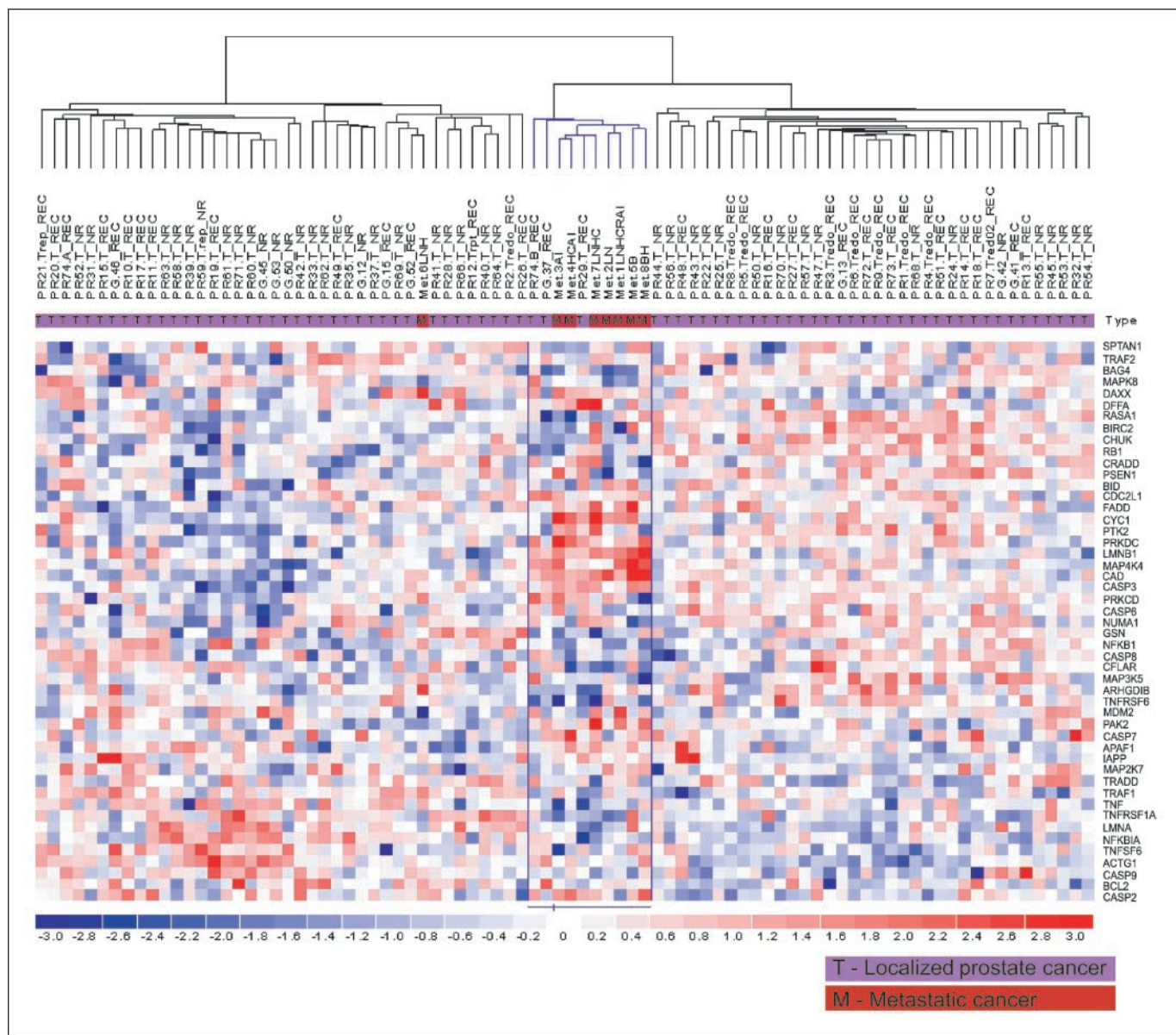


Figure 5. Validation of the HIV-I NEF pathway gene signature: Unsupervised clustering analysis of the Glinsky data set using the gene signature from the HIV-I NEF pathway showed a clear separation of metastatic samples from the localized tumor samples. *Red*, overexpression; *blue*, down-regulation. The cluster identified by the blue-colored dendrogram branch is enriched for metastases (P value < 0.00005).

that are biologically related, e.g., "modules" that are "over-represented" in the up- or down-regulated gene lists after defining a threshold for fold change in expression (8, 10). IMAP differs from these methods by allowing the combination of data from multiple data sets, leading to the identification of significantly dysregulated biological pathways without a priori definition of thresholds. Hence, it takes into account genes that show a marginal variation but become significant in the context of the pathway as a whole. It is therefore possible to identify genes that were not previously described in literature because of their apparent lack of association with prostate cancer when considered individually. However, those genes can function together in a pathway and become relevant for disease progression. Pathways are identified as being dysregulated instead of being overexpressed or underexpressed.

Using this strategy, we found several pathways as being dysregulated in metastatic prostate cancer, like NF κ B, transforming growth factor- β (TGF β) signaling, integrin signaling, and cell cycle pathways, that have been previously implicated in prostate cancer metastasis (22, 24, 25). The most significantly dysregulated pathway, the NF- κ B arm of the HIV-I NEF pathway (Fig. 3), consists of activation of NF- κ B by TNF α . The binding of TNF α to TNFR results in the dissociation of the inhibitory protein SODD and recruitment of the adapter protein TRADD. TRADD in turn binds to additional adapter proteins, RIP1 and TRAF2, which leads to the recruitment and activation of the IKK complex (26). This activation results in the phosphorylation and dissociation of I κ B α from the NF- κ B heterodimer and subsequent nuclear translocation of the active NF- κ B heterodimer. Nuclear translocation results in the transcription of the target survival pathway genes that include the TRAF proteins and the inhibitors of apoptosis (IAP).

Validation by qPCR and immunohistochemistry showed this pathway to be dysregulated especially in hormone-refractory metastasis. The differences that we detected between localized prostate cancer and hormone-naïve metastasis on the transcript (mRNA) level were not apparent at the protein level as detected by immunohistochemistry. We have previously shown that transcript and protein expression are discordant in 48% to 60% of cases (27). These discrepancies may arise due to tissue heterogeneity. The three microarray data sets used for IMAP included both hormone-naïve and hormone-refractory cases, and it has been reported that the metastatic lesions are not homogeneous and differ between individuals, sometimes within the same individual depending on the site of occurrence (28). Our results showed that the percentage of samples having a particular ratio of proteins varies within each class (Supplementary Fig. S2). This further shows that the tissues are heterogeneous with respect to protein expression. *In situ* hybridization to probe for transcript expression in tissues could be used on the TMAs to help confirm these results, but is not suitable for high-throughput analysis as is immunohistochemistry. Because the microarray pathway analysis identifies critical nodes in prostate cancer progression, immunohistochemistry is the best approach to translate the data into a clinically applicable test.

We found an increase in nuclear localization of NF κ B1 in both hormone-naïve metastasis and hormone-refractory metastasis consistent with other studies that have shown that NF- κ B p65 is localized in the nucleus and constitutively activated in cell lines (29–33) and metastatic tissue samples (31, 33, 34). The down-regulation of I κ B α may result in the constitutive activation of NF- κ B as reported by Tergaonkar et al. in the case of p65, where depletion of the I κ B proteins led to constitutive activation, with a fraction of NF- κ B localized to the nucleus at the basal steady-state level (35). The nuclear localization has also been reported to be associated with increased invasion (36) and biochemical relapse (37). Hormone-refractory metastasis are androgen-independent tumors, and there have been some studies that have shown that androgen receptor (*AR*) and NF- κ B are inversely related (22), and the loss of *AR* is accompanied by an increase in NF- κ B activity. Thus, the increase in the nuclear localization and activation of NF κ B1 in metastasis might allow for better survival. The analysis also identified TRAF1, TRAF2, and SODD as being involved in prostate cancer metastasis.

These results suggest that the NF- κ B pathway is important for the development and spread of prostate cancer, and that the low levels of NF κ B1 in hormone-refractory metastasis are active and sufficient to sustain the survival of metastatic lesions. *NFKB1* is known to be autoregulated and is also regulated by members of the ETS family (38), which have been recently shown to play a critical role in prostate cancer (39). Further studies will be needed to understand the mechanism underlying the down-regulation of *NFKB1* expression because this pathway seems to be important for early development of prostate cancer and in advanced disease (hormone-refractory metastasis).

In conclusion, we have developed a robust method for analyzing pathways altered in prostate cancer metastasis using three expression array data sets. The gene signature of the top dysregulated pathway, HIV-I NEF, when validated on an independent data set, was found to correctly classify all the metastatic samples. The genes in the pathway were also able to distinguish samples that had a good prognosis (i.e., no biochemical failure). Although this method has been used on three microarray studies, it can be extended to multiple expression array data sets. This study sets the stage for further discovery of the basic mechanisms that underlie a diseased state and would help in identifying critical nodes in the pathway that can be targeted for diagnosis and therapeutic intervention.

Acknowledgments

Received 6/12/2007; revised 8/28/2007; accepted 9/11/2007.

Grant support: Department of Defense Prostate Cancer Research Program fellowship award PC050965 (S.R. Setlur).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Robert Kim and Cara Pina for their technical assistance and Monideepa Roy for critical reading of the manuscript.

References

- Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. *CA Cancer J Clin* 2007;57:43–66.
- Dhanasekaran SM, Barrette TR, Ghosh D, et al. Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001;412:822–6.
- Beer DG, Kardia SL, Huang CC, et al. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 2002;8:816–24.
- van 't Veer LJ, Dai H, van de Vijver MJ, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002;415:530–6.
- Glinsky GV, Glinskii AB, Stephenson AJ, Hoffman RM, Gerald WL. Gene expression profiling predicts clinical outcome of prostate cancer. *J Clin Invest* 2004; 113:913–23.
- Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Barrette TR, Ghosh D, Chinnaiyan AM. Mining for regulatory programs in the cancer transcriptome. *Nat Genet* 2005;37:579–83.

7. Saxena V, Orgill D, Kohane I. Absolute enrichment: gene set enrichment analysis for homeostatic systems. *Nucleic Acids Res* 2006;34:e151.
8. Segal E, Friedman N, Koller D, Regev A. A module map showing conditional activity of expression modules in cancer. *Nat Genet* 2004;36:1090–8.
9. Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
10. Tomlins SA, Mehra R, Rhodes DR, et al. Integrative molecular concept modeling of prostate cancer progression. *Nat Genet* 2007;39:41–51.
11. Zeeberg BR, Feng W, Wang G, et al. GoMiner: a resource for biological interpretation of genomic and proteomic data. *Genome Biol* 2003;4:R28.
12. Goeman JJ, van de Geer SA, de Kort F, van Houwelingen HC. A global test for groups of genes: testing association with a clinical outcome. *Bioinformatics* 2004;20:93–9.
13. Efron B, Tibshirani R. On testing the significance of sets of genes. *Ann Appl Stat* 2007;1:107–29.
14. Jiang Z, Gentleman R. Extensions to gene set enrichment. *Bioinformatics* 2007;23:306–13.
15. Kim SY, Volsky DJ. PAGE: parametric analysis of gene set enrichment. *BMC Bioinformatics* 2005;6:144.
16. Folks JL. Handbook of statistics v.4: Nonparametric methods. In: Krishnaiah PR and Sen PK, editors. *Combination of independent tests*. North-Holland, Amsterdam: Elsevier Science Pub. Co.; 1984. p.113–21.
17. Rubin MA, Varambally S, Beroukheim R, et al. Overexpression, amplification, and androgen regulation of TPD52 in prostate cancer. *Cancer Res* 2004;64:3814–22.
18. Rubin MA, Bismar TA, Andren O, et al. Decreased α -methylacyl CoA racemase expression in localized prostate cancer is associated with an increased rate of biochemical recurrence and cancer-specific death. *Cancer Epidemiol Biomarkers Prev* 2005;14:1424–32.
19. Perner S, Hofer MD, Kim R, et al. Prostate-specific membrane antigen expression as a predictor of prostate cancer progression. *Hum Pathol* 2007;38:696–701.
20. LaTulippe E, Satagopan J, Smith A, et al. Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastatic disease. *Cancer Res* 2002;62:4499–506.
21. Lapointe J, Li C, Higgins JP, et al. Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci U S A* 2004;101:811–6.
22. Suh J, Rabson AB. NF- κ B activation in human prostate cancer: important mediator or epiphenomenon? *J Cell Biochem* 2004;91:100–17.
23. Bismar TA, Demichelis F, Riva A, et al. Defining aggressive prostate cancer using a 12-gene model. *Neoplasia* 2006;8:59–68.
24. Danielpour D. Functions and regulation of transforming growth factor- β (TGF- β) in the prostate. *Eur J Cancer* 2005;41:846–57.
25. Slack-Davis JK, Parsons JT. Emerging views of integrin signaling: implications for prostate cancer. *J Cell Biochem* 2004;91:41–6.
26. Hacker H, Karin M. Regulation and function of IKK and IKK-related kinases. *Sci STKE* 2006;2006:re13.
27. Varambally S, Yu J, Laxman B, et al. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell* 2005;8:393–406.
28. Shah RB, Mehra R, Chinnaiyan AM, et al. Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program. *Cancer Res* 2004;64:9209–16.
29. Gasparian AV, Yao YJ, Kowalczyk D, et al. The role of IKK in constitutive activation of NF- κ B transcription factor in prostate carcinoma cells. *J Cell Sci* 2002;115:141–51.
30. Palayoor ST, Youmell MY, Calderwood SK, Coleman CN, Price BD. Constitutive activation of I κ B kinase α and NF- κ B in prostate cancer cells is inhibited by ibuprofen. *Oncogene* 1999;18:7389–94.
31. Shukla S, MacLennan GT, Fu P, et al. Nuclear factor- κ B/p65 (Rel A) is constitutively activated in human prostate adenocarcinoma and correlates with disease progression. *Neoplasia* 2004;6:390–400.
32. Suh J, Payvandi F, Edelstein LC, et al. Mechanisms of constitutive NF- κ B activation in human prostate cancer cells. *Prostate* 2002;52:183–200.
33. Ismail HA, Lessard L, Mes-Masson AM, Saad F. Expression of NF- κ B in prostate cancer lymph node metastases. *Prostate* 2004;58:308–13.
34. Lessard L, Begin LR, Gleave ME, Mes-Masson AM, Saad F. Nuclear localization of nuclear factor- κ B transcription factors in prostate cancer: an immunohistochemical study. *Br J Cancer* 2005;93:1019–23.
35. Tergaonkar V, Correa RG, Ikawa M, Verma IM. Distinct roles of I κ B proteins in regulating constitutive NF- κ B activity. *Nat Cell Biol* 2005;7:921–3.
36. Lessard L, Karakiewicz PI, Bellon-Gagnon P, et al. Nuclear localization of nuclear factor- κ B p65 in primary prostate tumors is highly predictive of pelvic lymph node metastases. *Clin Cancer Res* 2006;12:5741–5.
37. Domingo-Domenech J, Mellado B, Ferrer B, et al. Activation of nuclear factor- κ B in human prostate carcinogenesis and association to biochemical relapse. *Br J Cancer* 2005;93:1285–94.
38. Lambert PF, Ludford-Menting MJ, Deacon NJ, Kola I, Doherty RR. The NF κ B1 promoter is controlled by proteins of the Ets family. *Mol Biol Cell* 1997;8:313–23.
39. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* 2005;310:644–8.