# Gene expression correlates of clinical prostate cancer behavior

Dinesh Singh,[1,5,12] Phillip G. Febbo,[1,4,8,11,12] Kenneth Ross,[8] Donald G. Jackson,[10] Judith Manola,[3] Christine Ladd,[8] Pablo Tamayo,[8] Andrew A. Renshaw,[6,14] Anthony V. D'Amico,[7] Jerome P. Richie,[5] Eric S. Lander,[8,9] Massimo Loda,[1,6] Philip W. Kantoff,[1,4] Todd R. Golub,[2,8,13] and William R. Sellers[1,4,11,13]

[1]Department of Adult Oncology
[2]Department of Pediatric Oncology
[3]Department of Biostatistical Sciences
Dana-Farber Cancer Institute
[4]Department of Internal Medicine
[5]Department of Surgery/Urology
[6]Department of Pathology
[7]Department of Radiation Oncology
Brigham and Women's Hospital
Harvard Medical School, Boston, Massachusetts 02115
[8]Whitehead Institute/Massachusetts Institute of Technology, Center for Genome Research, Cambridge, Massachusetts 02139
[9]Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
[10]Pharmaceutical Research Institute, Bristol-Myers Squibb, Inc., Princeton, New Jersey 08543
[11]Correspondence: phil_febbo@dfci.harvard.edu (P.G.F.) and william_sellers@dfci.harvard.edu (W.R.S.)
[12]These authors contributed equally
[13]These authors codirected this work
[14]Present address: Baptist Hospital of Miami, Miami, Florida 33176

## Summary

**Prostate tumors are among the most heterogeneous of cancers, both histologically and clinically. Microarray expression analysis was used to determine whether global biological differences underlie common pathological features of prostate cancer and to identify genes that might anticipate the clinical behavior of this disease. While no expression correlates of age, serum prostate specific antigen (PSA), and measures of local invasion were found, a set of genes was identified that strongly correlated with the state of tumor differentiation as measured by Gleason score. Moreover, a model using gene expression data alone accurately predicted patient outcome following prostatectomy. These results support the notion that the clinical behavior of prostate cancer is linked to underlying gene expression differences that are detectable at the time of diagnosis.**

## Introduction

Prostate cancer is the most common nondermatological cancer in the United States with an estimated 198,100 new cases and 31,500 deaths in 2001 (Greenlee et al., 2000). The adoption of screening based upon the measurement of the serum prostate specific antigen (PSA) has led to the earlier detection of prostate cancer where most tumors now appear confined to the prostate gland at presentation (Han et al., 2001). Early diagnosis provides an opportunity for curative surgery. However, up to 30% of men undergoing radical prostatectomy will relapse, often as a result of micrometastatic disease present at the time of surgery (Roberts et al., 2001a, 2001b). The challenge is to identify those patients at risk for relapse and to better understand the molecular abnormalities that define tumors at risk for relapse.

Several clinical features of prostate cancer including tumor stage (Jewett, 1975), degree of tumor cell differentiation or Gleason score (GS) (Gleason, 1966), and the serum PSA (Stamey et al., 1987) are used in routine clinical practice to separate men into groups at low, intermediate, and high risk for tumor recurrence following local therapy. However, the majority of patients who now undergo prostatectomy have low to intermediate risk clinical features, and determining the prognosis for these patients remains difficult.

The utility of existing prognostic factors might be limited because they largely measure tumor differentiation and bulk but

SIGNIFICANCE

**Improved patient stratification can allow the rational application of current treatments and the selected testing of novel therapeutics in patient populations most likely to benefit. Clinical features including Gleason score, tumor stage, and serum prostate specific antigen (PSA) are used to assess relapse-risk in men with prostate cancer. Such parameters are less useful in guiding therapy for men having intermediate risk disease, 30% of whom recur following local therapy. Our data suggest that expression-based models may help to identify patients at greatest risk for recurrence and thus facilitate the rational application of current therapies. Furthermore, the association of specific genes' expression, such as platelet-derived growth factor β (PDGFRβ), with outcome raises the possibility that expression analysis may prove useful in selecting patients for emerging mechanism-based therapeutics.**

do not otherwise sample the underlying biological properties that likely drive tumor behavior. Attempts to explore genetic correlates of tumor behavior have found alterations in a number of candidate genes associated with prostate cancer progression, including loss of p53, amplification of myc, loss of p27, and loss of PTEN (reviewed in Sellers and Sawyers, 2001). However, no single gene has been shown to have sufficient prognostic utility to warrant clinical implementation.

Recently, genomic methodologies have been used to discover consistent gene expression patterns associated with a given histological or clinical phenotype (Golub et al., 1999; Perou et al., 2000; van't Veer et al., 2002). Here, gene expression patterns from 52 tumor and 50 normal prostate specimens were studied in order to ask whether such patterns could be used to predict common clinical and pathological phenotypes relevant to the treatment of men diagnosed with this disease. In addition to expression patterns that correlated with GS and with the distinction of tumor from normal, an expression-based model was built that accurately predicted patient outcome. These data suggest that it may be possible to predict the clinical behavior of prostate cancer based upon gene expression analysis of primary tumors. Such prediction strategies, if generalizable, would allow for the rational application of additional postsurgical therapeutics to high-risk individuals.

## Results

### Tumor versus normal classification

To investigate whether robust gene expression differences could be found that distinguished common clinical and pathological features of prostate cancer, 235 radical prostatectomy specimens were analyzed from patients undergoing surgery between 1995 and 1997. Of these samples, 65 had tumor on opposing sides of the tissue specimen. High-quality expression profiles were successfully derived from 52 of these prostate tumors and 50 nontumor prostate samples (referred to as normal hereafter) using oligonucleotide microarrays containing probes for approximately 12,600 genes and ESTs (raw data available at http://www-genome.wi.mit.edu/MPR/prostate). The clinical and pathological features of the 52 patients and tumors included in this study were indistinguishable from those of all patients treated with radical prostatectomy during the collection period (Table 1).

Genes were ranked according to their differential expression across the two classes (tumor versus normal) using a variation of a signal-to-noise metric (S2N) (Golub et al., 1999). The statistical significance of these gene expression correlations was determined by comparing the observed correlations to the results derived from 1000 permutations of the class labels (tumor or normal). This analysis indicated that 317 genes had higher expression in the tumor samples ($p \leq 0.001$) whereas 139 genes were more highly expressed in normal prostate samples ($p \leq 0.001$) (Supplemental Figure S1 [see Supplemental data, below]).

Gene expression differences between tumor and normal prostate samples have been previously reported (Chetcuti et al., 2001; Dhanasekaran et al., 2001; Luo et al., 2001; Welsh et al., 2001); however, the feasibility of using such differences to predict the identity of prostate samples has not been tested. To this end, we built predictors using a *k*-nearest neighbor (*k*-NN) supervised machine learning algorithm. Models that utilized 4 or more genes classified samples with greater than 90% accuracy in leave-one-out cross-validation testing ($p < 0.001$ as measured by permutation testing) (Suppl. Figure S2A and S2B). The 4-gene and 16-gene models were tested on an independent data set of 8 normal and 27 tumor prostate samples provided by G. Hampton (Welsh et al., 2001). Despite a nearly 10-fold difference in overall microarray intensity between these datasets (see Supplemental Experimental procedures [below]), the classifier performed with relatively high accuracy (4-gene model 77%; 16-gene model 86%; $p < 0.05$, Fisher's exact test) (Suppl. Figure S3). Thus, expression differences can be used to predict the identity of unknown prostate samples and these gene expression differences are conserved across independent data sets.

### Prediction of pathological features of prostate cancer

In order to ask whether gene expression patterns exist that describe and or predict the differences in clinical behavior apparent among prostate tumors, the expression patterns within the 52 tumors were analyzed. Correlations between gene expression and known clinical and pathological parameters were determined for dichotomous variables (e.g., the presence or absence of capsular penetration, perineural invasion, or positive surgical margins), as well as for factors treated as continuous variables (e.g., patient age, serum PSA, and GS). Statistical significance was determined by comparing the observed correlations to those correlations measured in randomly permuted datasets. With the exception of GS (see below), no statistically significant gene expression correlates of these clinical and pathological features were observed (see Suppl. Figure S4). Specifically, no expression signature discriminated between locally invasive and noninvasive phenotypes (e.g., capsular penetration, positive surgical margins, and perineural invasion). Thus, while these features are often associated with different clinical outcomes, they are not reflected by global gene expression differences.

A gene expression signature of GS, however, was detectable. Fifteen genes had expression positively correlated with GS (Type I) and 14 genes had expression negatively correlated with GS (Type II) beyond what would be expected by chance alone ($p \leq 0.001$) (Figure 1 and Suppl. Figure S4). As these genes were the most positively and negatively correlated with GS, when used in hierarchical clustering, the 29 Type I and Type II were, as expected, separated into two groups (Figure 1). The correlation of these genes with GS and their coordinate expression in tumors, nonetheless, may have occurred by random chance alone in the initial dataset. However, when the same 29 genes were used to drive hierarchical clustering of the independent data set, Type I and Type II genes remained highly cosegregated suggesting that this coexpression is reproducible ($p < 0.0001$) (Suppl. Figure S5).

Strikingly, in both data sets, while most high-grade tumors expressed the Type I genes, a subset of intermediate grade tumors also expressed many of the Type I genes (Figure 1 and Suppl. Figure S5). This indicates that some tumors of intermediate histological grade share the gene expression signature of higher grade tumors. Thus, the coexpression of these genes may identify tumors that are of intermediate histological grade, yet share the molecular phenotype of high-grade tumors.

### Prediction of clinical outcome

In this data set, 21 patients were evaluable with respect to recurrence following surgery with 8 patients having relapsed

**Table 1.** Clinical and pathological features

| Variable | | Study group | All | p | Recurrent | Nonrecurrent | p |
|---|---|---|---|---|---|---|---|
| #Patients | | 52 | 393 | | 8 | 13 | |
| Age | Median | 58.5 | 61 | 0.32 | 58.5 | 60.0 | 0.74 |
| | Range | 47–72 | 40–79 | | 51–72 | 47–72 | |
| PSA | Median | 6.3 | 6.7 | 0.62 | 6.8 | 6.3 | 0.64 |
| | Range | 1.0–27.8 | 0.7–46.0 | | 5.0–24.3 | 3.6–18.0 | |
| Gleason | 2–6 | 19 (37%) | 190 (51%) | 0.10 | 2 (25%) | 6 (46%) | 0.45 |
| Score | 7 | 29 (56%) | 146 (39%) | | 5 (63%) | 6 (46%) | |
| (Clinical) | 8–10 | 4 (8%) | 34 (9%) | | 1 (12%) | 1 (8%) | |
| | Unknown | | 23 | | | | |
| Gleason | 2–6 | 24 (46%) | | 0.46* | 1 (12%) | 7 (54%) | 0.01 |
| Score | 7 | 22 (42%) | | | 3 (38%) | 6 (46%) | |
| (Sample) | 8–10 | 6 (12%) | | | 4 (50%) | 0 | |
| Clinical | T1–T2a | 38 (88%) | 285 (79%) | 0.10 | 5 (100%) | 10 (91%) | 1.00 |
| Stage | T2b | 5 (12%) | 24 (7%) | | 0 | 9 (9%) | |
| | ≥T2c | 0 | 50 (14%) | | 0 | 0 | |
| | Unknown | 9 | 34 | | 3 | 2 | |
| Pathologic | T2a | 7 (13%) | 49 (15%) | 0.15 | 1 (13%) | 2 (15%) | 0.90 |
| Stage | T2b | 25 (48%) | 189 (58%) | | 4 (50%) | 3 (38%) | |
| | T3a | 16 (31%) | 74 (23%) | | 2 (25%) | 4 (31%) | |
| | T3b | 4 (8%) | 12 (4%) | | 1 (13%) | 2 (15%) | |
| | T4a | 0 | 2 (1%) | | 0 | 0 | |
| | Unknown | | 53 | | | | |
| Gland | Median | 51.75 | 53.0 | 0.89 | 67.5 | 50.0 | 0.15 |
| Vol. | Range | 35–191 | 18–191 | | 35.5–191 | 35–169 | |
| Ext. Cap. | No | 32 (62%) | 239 (73%) | 0.10 | 5 (63%) | 7 (54%) | 1.00 |
| | Yes | 20 (38%) | 88 (27%) | | 3 (37%) | 6 (46%) | |
| | Unknown | | 66 | | | | |
| SVInv. | No | 49 (94%) | 315 (96%) | 0.44 | 7 (88%) | 12 (92%) | 1.00 |
| | Yes | 3 (6%) | 12 (4%) | | 1 (12%) | 1 (8%) | |
| | Unknown | | 66 | | | | |
| Pos. Mar. | No | 39 (75%) | 283 (76%) | 0.86 | 5 (62%) | 7 (54%) | 1.00 |
| | Yes | 13 (25%) | 89 (24%) | | 3 (38%) | 6 (46%) | |
| | Unknown | | 21 | | | | |

PSA, serum prostate specific antigen; Vol., Volume; SV Inv., seminal vesicle invasion; Ext Cap., Extension through capsule; Pos. Margin, positive surgical resection margin. Gleason Score (Clinical) indicates the Gleason Score recorded from the radical prostatectomy specimen. Gleason Score (Sample) indicates the Gleason Score of the frozen sections from the tumor specimens used in RNA preparation. *p value resulting from the comparison of the Gleason Score (Sample) of the 52 tumors to the Gleason Score (Clinical) from the entire population.

(defined as two successive PSA values $> 0.2$ ng/ml) and 13 patients having remained relapse free for at least 4 years. While these two groups did not differ with respect to the Clinical GS, serum PSA, or tumor stage, the GS of the sections adjacent to tissue used for RNA extraction was ≥8 in a greater proportion of recurrent patients (4/8 versus 0/13) (Table 1).

While no single gene was statistically associated with recurrence (at $p = 0.05$) (data not shown), when a *k*-NN classification approach was applied, a 5-gene model with 2 nearest neighbors ($k = 2$) reached 90% accuracy in predicting recurrence during leave-one-out cross validation. When Kaplan-Meier survival analysis was performed based upon the predicted outcome, the results compared favorably with known prognostic indicators in this data set (Figure 2B). However, the standard Kaplan-Meier log-rank statistic, while demonstrating a difference in the survival curves, does not account for the multiple hypothesis testing that occurred during model optimization. To further assess the statistical significance of this prediction model, we performed 1000 permutations of the class labels (recurrence versus nonrecurrence), and for each permutation attempted to find multigene expression classifiers using the same range of gene numbers. Only 37 of the 1000 permutations yielded models whose accuracy matched or exceeded 90%. Thus, the likelihood of matching the success of the observed 5-gene model simply by chance alone was estimated at $p = 0.037$ (Figure 2A).

While there were too few tumor samples to allow for multivariate analysis, as mentioned above, only the Sample GS was significantly different between patients who recurred and those who did not recur (Figure 2B and Table 1). Nonetheless, 4 recurrent tumors were of intermediate grade ($GS \leq 7$) raising the possibility that gene expression-based models might provide additional prognostic information not currently described by existing clinical and pathological parameters.

The genes that were used by the 5-gene outcome predictors during leave-one-out cross validation are shown in Figure 3. The top 5 genes were each used in over half of the models, and included chromogranin A, platelet-derived growth factor receptor β (PDGFRβ), HOXC6, inositol triphosphate receptor 3 (IPTR3) and sialyltransferase-1.

## Discussion

There is an immediate need for robust prognostic markers capable of identifying patients at risk of relapse following local therapy; conventional and experimental therapeutics could then be focused on this subpopulation, rather than the general population of prostate cancer patients, 70% of whom are cured by surgery alone.

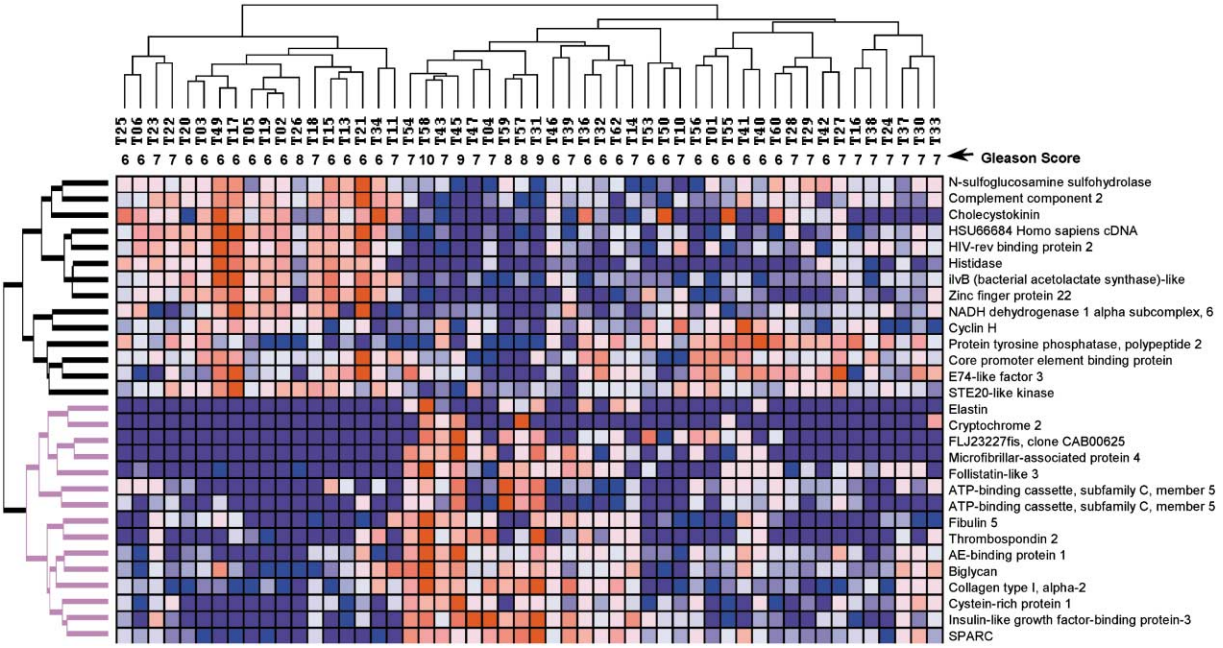



**Figure 1.** Gene expression correlates of Gleason score

Hierarchical clustering of tumors and the 29 genes statistically associated with GS. Genes and samples are shown as ordered by Gene Cluster and Treeview. The expression of each gene in each sample is represented by the number of standard deviations above (red) or below (blue) the mean for that gene across all 52 samples.

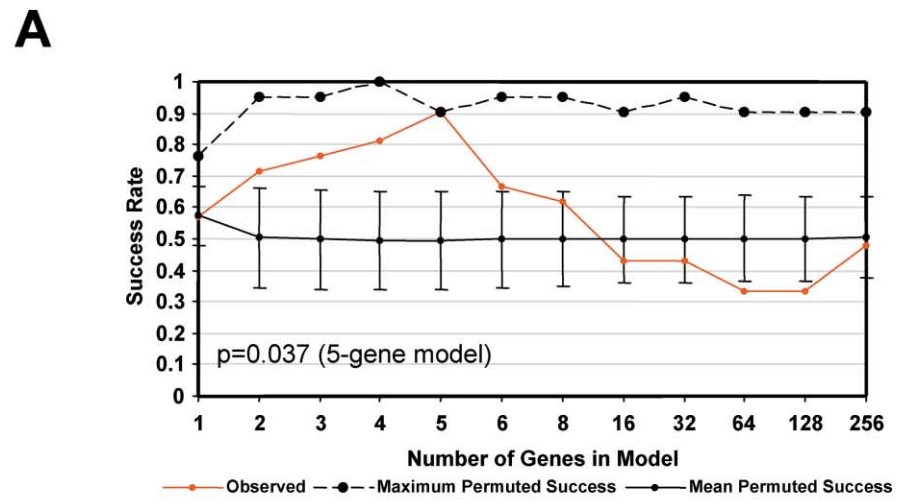
Our analysis revealed global gene expression differences that were sufficiently robust to distinguish tumor from normal in both training and validation sets. While the level of accuracy (86%–92%) is not sufficient to *replace* histological examination, these molecular markers may be useful adjuncts to morphology-based diagnostics. In addition, while certain genes differentially expressed between normal and tumor prostate specimens in microarray experiments have been correlated with outcome in large data sets (Dhanasekaran et al., 2001), in our data such differentially expressed genes were not highly correlated with outcome.

Among prognostic factors for prostate cancer, serum PSA and measures of local invasion were not associated with robust gene expression signatures. The lack of an associated gene expression signature does not exclude the possibility that such signatures exist. It is possible that analyzing genes beyond those present on the current microarray, or extending the experiment to larger data sets might reveal such patterns. However, there was a readily detectable and statistically significant signature of GS. The expression pattern of these genes separated tumors into distinct groups during hierarchical clustering in both our initial and in a validation data set and grouped some of the intermediate-grade tumors with high-grade tumors. In so doing, this set of genes may help identify a subset of histologically intermediate-grade tumors that may have more aggressive clinical behavior.

In this data set, GS was associated with patient outcome (Table 1 and Figure 2B); however, only two of the genes correlated with GS (IGFBP-3 and COL1A2) contributed to the outcome prediction model (Figure 3). Instead, genes whose expression was not correlated with GS were the most frequently used in the outcome prediction model, suggesting that while GS is associated with patient outcome, GS-independent markers and determinants of prostate cancer behavior exist.

Attempts to build a gene expression-based predictor of recurrence following prostatectomy led to a model that correctly predicted the outcome of 19 of the 21 evaluable patients in this study. While the result reached statistical significance based on permutation testing ($p = 0.037$), the performance of the model may be due, at least in part, to overoptimization. As such, this is a preliminary model and larger datasets will be required to reach model stability, to minimize the possibility of model overfitting and ultimately allow the independent validation of such a predictor.

Despite these limitations, the identity of the genes comprising the outcome prediction model support the existence of measurable outcome determinants for prostate cancer recurrence. For example, chromogranin A, one of the 5 genes most frequently used in the prediction model, has previously been associated with poor outcome in prostate cancer (Theodorescu et al., 1997). The utility of PDGFRβ expression in the recurrence predictor is also intriguing in light of the recent observations that PDGFR (α and β) are expressed in advanced prostate cancer (Chott et al., 1999). The successful prediction of patient outcome will ultimately lead to improved decision making regarding current therapeutic options and the rational selection of patients at high risk for relapse for clinical trials testing adjuvant therapeutics. Furthermore, the identification of genes whose expression drives outcome and whose protein products are tractable targets for small molecules may contribute to the development and selective application of novel mechanism-based treatments. Ongoing trials of Gleevec, an inhibitor of the abl,



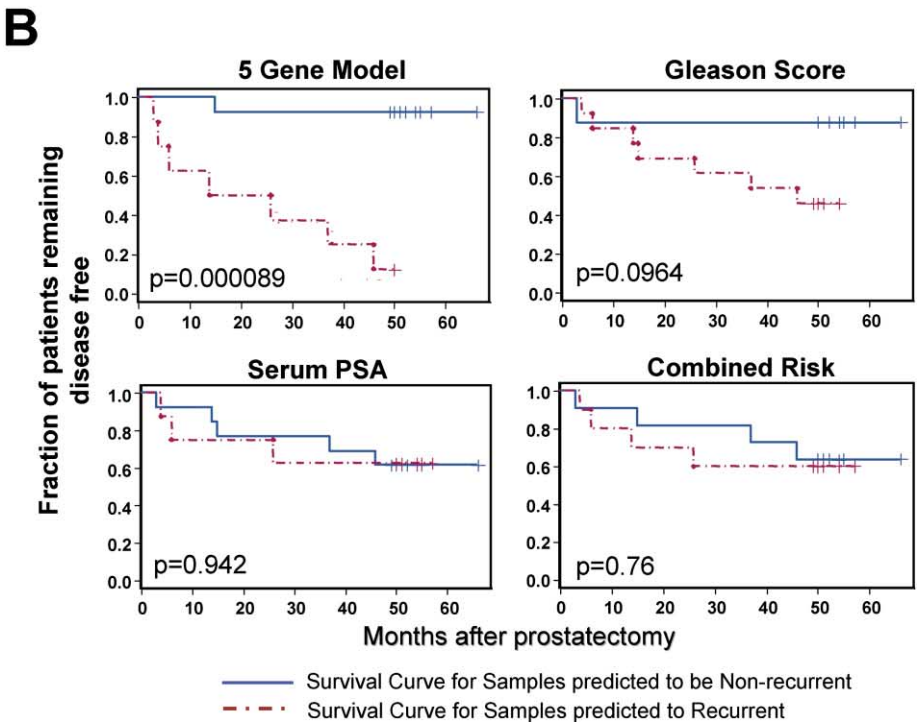




**Figure 2.** Outcome prediction models

**A:** The success rates of models predicting outcome. Leave-one-out cross validation was used to build outcome prediction models (recurrent versus nonrecurrent) using from 1 to 256 genes. The *x* axis indicates number of genes used in model building, and the *y* axis indicates the frequency of success. Shown is the number of correct predictions divided by the total number of predictions (red line) in the observed data using leave-one-out cross validation. The mean success rate ± the standard deviation (bottom dashed line) and maximum success rate (top dashed line) obtained using the permuted data is shown.

**B:** Disease-free survival of patients stratified based on the 5-gene model, GS, serum PSA, or combined risk. Kaplan-meir analysis was used to plot the fraction of at-risk patients remaining free of disease (*y* axis) at the indicated time after prostatectomy (*x* axis). Shown is patient stratification based on the 5-gene model, GS ($\leq$ 6 versus $\geq$ 7), serum PSA ($<$ 10 versus PSA $\geq$ 10), and a combination of GS, serum PSA, and surgical stage (low and intermediate versus high risk). High risk was defined as a GS $>$ 7, PSA $\geq$ 20, and surgical stage T3 or higher, the remaining samples were considered low or intermediate risk. P values were calculated using a log-rank test (Mantel-Hwenszel test).

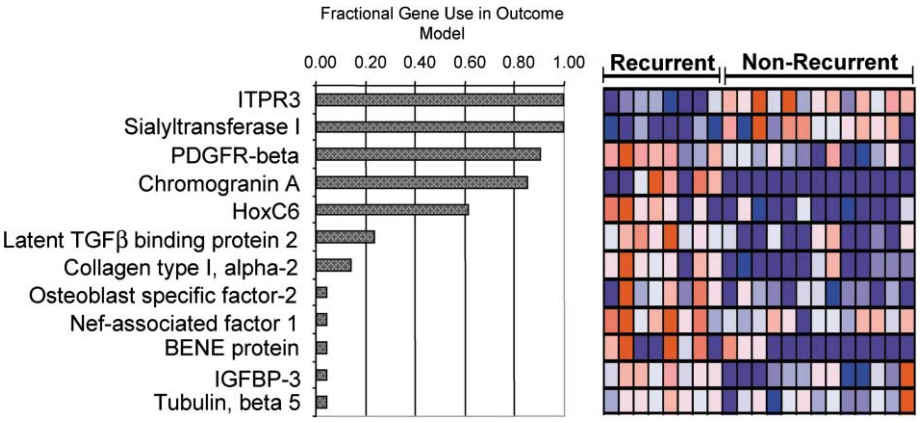



**Figure 3.** Genes used to build an outcome prediction model

The genes most commonly used in the 5-gene model are shown as described for Figure 2B. The expression of each gene (rows) in each recurrent or nonrecurrent sample (columns) is represented by the number of standard deviations above (red) or below (blue) the mean for that gene across all 21 samples.

kit, and PDGFR tyrosine kinases, in prostate cancer will test the hypothesis that PDGFRβ falls into this latter category.

The samples in this study were derived from patients diagnosed after the widespread adoption of PSA screening. As such the findings in this study are expected to be relevant to patients diagnosed today. Clearly, larger, confirmatory studies will be required prior to the implementation of any changes in the clinical care of patients with prostate cancer. Nevertheless, these studies provide evidence that the clinical phenotypes and behavior of prostate cancer can be anticipated by the analysis of the gene expression profiles.

## Experimental procedures

### Prostate tissue samples

From 1995 to 1997 samples of prostate tumors and adjacent prostate tissue not containing tumor (referred to as "normal") were collected from patients undergoing radical prostatectomy at the Brigham and Women's Hospital. From 235 "tumor" samples, 65 had cancer present on opposing sides of the OCT embedded specimens. Each of these samples was reviewed by a single pathologist to determine the "Sample" GS. Other pathological features from the radical prostatectomy specimens included in this analysis as well as from all contemporary prostatectomies were abstracted from the pathology reports including the "Clinical" GS.

The Wilcoxon rank sum test (Wilcoxin, 1945) and Fisher's exact test (Cox, 1970) was used to test for differences in continuous variables and dichotomous variables, respectively, between the study sample and all patients treated from 1993 to 1997 and between patients who recurred and those who did not. Tests for differences in these groups on ordered, categorical variables were done using the methods described by Mehta (Mehta and Patel, 1984). Kaplan-Meir survival plots and log-rank statistics (Mantel-Haenszel test) were generated using the S-Plus statistical software package (Insightful Corp).

### Gene expression measurements

Total RNA extraction, generation of labeled cRNA, fragmentation, hybridization to U95Av2 arrays (Affymetrix) and wash steps were performed as previously described (Bhattacharjee et al., 2001; Golub et al., 1999). Raw expression values were normalized to the median array intensity and thresholds were set at 10 and 16,000 units. Genes whose expression varied less than 5-fold between any two samples in any given experiment were removed.

### Gene ranking, class prediction by *k*-nearest neighbors, and permutation testing for dichotomous variables

Gene expression differences associated with a particular dichotomous class distinction were measured and ranked using a variation of the S2N statistic as previously described (Golub et al., 1999). Measured S2N values were compared to calculated S2N values obtained in 1000 data sets where a given class label was randomly permuted (Good, 2000). For these comparisons, p values represent the frequency at which the S2N statistic from randomly permuted data exceeded the measured S2N statistic.

*K*-nearest neighbor (*k*-NN) class prediction models were built as previously described (Pomeroy et al., 2002). Briefly, for each class distinction tested, after exclusion of one sample, genes were ranked using the S2N metric derived from the remaining samples. The Euclidean distances (ED) between the withheld sample and the remaining samples were calculated using a given number of genes (as in the figures). The identity of the left-out sample was predicted based upon the class membership of the *k*-closest samples weighted by the reciprocal of the EDs. P-values were then assigned based upon the frequency with which models generated and tested on 1000 randomly permuted data sets performed better than models generated using the observed data.

### Correlation of gene expression with continuous variables

The Pearson coefficient was used to measure correlations between gene expression patterns and patient age, serum PSA, and GS (treated as a continuous variable). Pearson coefficients were also used to measure the same correlations in data sets where the sample label of each variable tested (age, PSA, or GS) was randomly permuted 10,000 times. P-values were then assigned based upon the frequency with which Pearson coefficients generated from randomly permuted data exceeded those generated from the observed data.

### Independent prostate expression data used for validation

Oligonucleotide array-based expression data (Affymetrix Hum95Av2) and clinical data for 8 normal and 27 prostate tumors were provided G. Hampton as a validation set (Welsh et al., 2001). Global differences between the initial and validation data sets were quantified by determining the means of the mean array intensities. Validation of the tumor normal prediction models and of the coexpression observed for the genes highly correlated to GS was performed as described in the Supplemental experimental procedures (see Supplemental data, below).

### Supplemental data

Supplemental experimental procedures and Figures S1–S5 can be found at http://www.cancercell.org/cgi/content/full/1/2/203/DC1 and at http://www-genome.wi.mit.edu/MPR/prostate.

### Acknowledgments

The authors thank David Livingston and Matthew Meyerson for their assistance in this research and for critical reading of this manuscript, and Garret Hampton for providing the independent prostate cancer expression data set. This work was supported by the National Cancer Institute (CA84995) and in part by the Gelb Center at DFCI, Bristol-Meyers Squibb, Affymetrix and Millennium Pharmaceuticals. P.G.F. also receives support from the National Cancer Institute (CA89031) and from CaP CURE. W.R.S. also receives support from the Damon-Runyon Lilly Clinical Investigator Award and CaP CURE.

Received: January 10, 2002
Revised: March 1, 2002

### References

Bhattacharjee, A., Richards, W.G., Staunton, J., Li, C., Monti, S., Vasa, P., Ladd, C., Beheshti, J., Bueno, R., Gillette, M., et al. (2001). Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc. Natl. Acad. Sci. USA *98*, 13790–13795.

Chetcuti, A., Margan, S., Mann, S., Russell, P., Handelsman, D., Rogers, J., and Dong, Q. (2001). Identification of differentially expressed genes in organ-confined prostate cancer by gene expression array. Prostate *47*, 132–140.

Chott, A., Sun, Z., Morganstern, D., Pan, J., Li, T., Susani, M., Mosberger, I., Upton, M.P., Bubley, G.J., and Balk, S.P. (1999). Tyrosine kinases expressed in vivo by human prostate cancer bone marrow metastases and loss of the type 1 insulin-like growth factor receptor. Am. J. Pathol. *155*, 1271–1279.

Cox, D. (1970). Analysis of Bianary Data (London: Mechuen and Co.).

Dhanasekaran, S.M., Barrette, T.R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K.J., Rubin, M.A., and Chinnaiyan, A.M. (2001). Delineation of prognostic biomarkers in prostate cancer. Nature *412*, 822–826.

Gleason, D. (1966). Classification of prostatic carcinomas. Cancer Chemother. Rep. *50*, 125–128.

Golub, T.R., Slonim, D.K., Tamayo, P., Huard, C., Gaasenbeek, M., Mesirov, J.P., Coller, H., Loh, M.L., Downing, J.R., Caligiuri, M.A., et al. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. Science *286*, 531–537.

Good, P. (2000). Permutation tests: A practical guide to resampling methods for testing hypotheses, Second Edition, S.S.i. Statistics (New York: Springer-Verlag).

Greenlee, R.T., Murray, T., Bolden, S., and Wingo, P.A. (2000). Cancer statistics, 2000. CA Cancer J. Clin. *50*, 7–33.

Han, M., Partin, A.W., Piantadosi, S., Epstein, J.I., and Walsh, P.C. (2001). Era

specific biochemical recurrence-free survival following radical prostatectomy for clinically localized prostate cancer. J. Urol. *166*, 416–419.

Jewett, H.J. (1975). The present status of radical prostatectomy for stages A and B prostatic cancer. Urol. Clin. North Am. *2*, 105.

Luo, J.H., Yu, Y.P., Cieply, K., Lin, F., Deflavia, P., Dhir, R., Finkelstein, S., Michalopoulos, G., and Becich, M. (2001). Gene expression analysis of prostate cancers. Mol. Carcinog. *33*, 25–35.

Mehta, C., and Patel, A.T. (1984). Exact significance testing for ordered categorical data. Biometrics *30*, 819–825.

Perou, C.M., Sorlie, T., Eisen, M.B., van de Rijn, M., Jeffrey, S.S., Rees, C.A., Pollack, J.R., Ross, D.T., Johnsen, H., Akslen, L.A., et al. (2000). Molecular portraits of human breast tumours. Nature *406*, 747–752.

Pomeroy, S.L., Tamayo, P., Gaasenbeek, M., Sturla, L.M., Angelo, M., McLaughlin, M.E., Kim, J.Y., Goumnerova, L.C., Black, P.M., Lau, C., et al. (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. Nature *415*, 436–442.

Roberts, S.G., Blute, M.L., Bergstralh, E.J., Slezak, J.M., and Zincke, H. (2001a). PSA doubling time as a predictor of clinical progression after biochemical failure following radical prostatectomy for prostate cancer. Mayo Clin. Proc. *76*, 576–581.

Roberts, W.W., Bergstralh, E.J., Blute, M.L., Slezak, J.M., Carducci, M., Han, M., Epstein, J.I., Eisenberger, M.A., Walsh, P.C., and Partin, A.W. (2001b). Contemporary identification of patients at high risk of early prostate cancer recurrence after radical retropubic prostatectomy. Urology *57*, 1033–1037.

Sellers, W., and Sawyers, C. (2001). Molecular abnormalities in prostate cancer. In Prostate Cancer: Principles & Practice, P. Kantoff, P. Carroll, and A. D'Amico, eds. (Philadelphia: Lippincott Williams & Wilkins), pp. 53–76.

Stamey, T.A., Yang, N., Hay, A.R., McNeal, J.E., Freiha, F.S., and Redwine, E.A. (1987). Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. N. Engl. J. Med. *317*, 909.

Theodorescu, D., Broder, S.R., Boyd, J.C., Mills, S.E., and Frierson, H.F., Jr. (1997). Cathepsin D and chromogranin A as predictors of long term disease specific survival after radical prostatectomy for localized carcinoma of the prostate. Cancer *80*, 2109–2119.

van't Veer, L., Dai, H., Vijver, M.v.D., He, Y., Hart, A., Moa, M., Peterse, H., Kooy, K.v.D., Marton, M., Witteveen, A., Schreiber, G., Kerkhoven, R., Roberts, C., Linsley, P., Bernards, R., and Friend, S. (2002). Gene expression profiling predicts clinical outcome of breast cancer. Nature *415*, 530–536.

Welsh, J.B., Sapinoso, L.M., Su, A.I., Kern, S.G., Wang-Rodriguez, J., Moskaluk, C.A., Frierson, H.F., Jr., and Hampton, G.M. (2001). Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. Cancer Res. *61*, 5974–5978.

Wilcoxin, F. (1945). Individual comparisons by ranking methods. Biometrics *1*, 80–83.