UC Irvine
UC Irvine Previously Published Works

Title

Diagnosis of prostate cancer using differentially expressed genes in stroma.

Permalink

https://escholarship.org/uc/item/0dj047b3

Journal

Cancer research, 71(7)

ISSN

1538-7445

Authors

Jia, Zhenyu
Wang, Yipeng
Sawyers, Anne
et al.

Publication Date

2011-04-01

Peer reviewed

Diagnosis of Prostate Cancer Using Differentially Expressed Genes in Stroma

Zhenyu Jia[1], Yipeng Wang[1,2], Anne Sawyers[1], Huazhen Yao[1,2], Farahnaz Rahmatpanah[1], Xiao-Qin Xia[1,2], Qiang Xu[3], Rebecca Pio[1], Tolga Turan[1], James A. Koziol[4], Steve Goodison[5], Philip Carpenter[1], Jessica Wang-Rodriguez[6,7], Anne Simoneau[8], Frank Meyskens[9], Manuel Sutton[1], Waldemar Lernhardt[10], Thomas Beach[11], Joseph Monforte[3], Michael McClelland[1,2], and Dan Mercola[1]

Abstract

More than one million prostate biopsies are performed in the United States every year. A failure to find cancer is not definitive in a significant percentage of patients due to the presence of equivocal structures or continuing clinical suspicion. We have identified gene expression changes in stroma that can detect tumor nearby. We compared gene expression profiles of 13 biopsies containing stroma near tumor and 15 biopsies from volunteers without prostate cancer. About 3,800 significant expression changes were found and thereafter filtered using independent expression profiles to eliminate possible age-related genes and genes expressed at detectable levels in tumor cells. A stroma-specific classifier for nearby tumor was constructed on the basis of 114 candidate genes and tested on 364 independent samples including 243 tumor-bearing samples and 121 nontumor samples (normal biopsies, normal autopsies, remote stroma, as well as stroma within a few millimeters of tumor). The classifier predicted the tumor status of patients using tumor-free samples with an average accuracy of 97% (sensitivity = 98% and specificity = 88%) whereas classifiers trained with sets of 100 randomly generated genes had no diagnostic value. These results indicate that the prostate cancer microenvironment exhibits reproducible changes useful for categorizing the presence of tumor in patients when a prostate sample is derived from near the tumor but does not contain any recognizable tumor. *Cancer Res; 71(7); 2476–87. ©2011 AACR.*

Introduction

There are more than 1 million prostate biopsy procedures carried out in the United States every year (1). More than 60% of the results are negative (2–4). However even the best current methods, including transrectal ultrasound (TRUS) procedures, may miss up to 30% of clinically significant prostate cancers (5). Indeed, about 20% to 30% of patients who give negative results on initial biopsy are rebiopsied in about 3 to 12 months [∼190,000 patients owing to the presence of prostatic intraepithelial neoplasia (PIN), high-grade PIN (HGPIN), atypical small acinar proliferation (ASAP), or other grounds for clinical suspicion of the presence of tumor (2–4, 6, 7)]. Many repeat biopsies give the diagnosis of adenocarcinoma. For example, 16% to 23% of HGPIN and up to 59% of ASAP cases give the diagnosis of adenocarcinoma upon repeat biopsy (2–4, 8). Patients deferred to repeat biopsy, receive little treatment or guidance during the interim—a period when tumors may continue to progress. Therefore, there is a need for methods that resolve false-negative and equivocal cases.

Equivocal and negative biopsies are, by definition, deficient in diagnostic tumor but contain ample stroma. Moreover, stroma near tumor may contain changes in gene expression that are not found in nontumor samples, which could be the basis for a clinical test. Epithelial cells of prostate cancer infiltrate and propagate in a microenvironment consisting largely of myofibroblast cells as well as inflammatory cells and other supporting cells and structures. It has long been appreciated that this mesenchymal component is not passive but responds to signals from the tumor component and, in turn, alters tumor properties, some of which are essential for tumor growth and progression (9, 10). Indeed, studies of prostate cancer were among the first to demonstrate an important role of the stroma in cancer progression. Mouse model studies showed that sur-

Authors' Affiliations: [1]Department of Pathology & Laboratory Medicine, University of California, Irvine; [2]Vaccine Research Institute of San Diego; [3]Althea Technologies Inc., San Diego; [4]The Scripps Research Institute, La Jolla, California; [5]Cancer Research Institute, MD Anderson Cancer Center Orlando, Orlando, Florida; [6]Veteran Affairs Medical Center, San Diego; [7]Department of Pathology, University of California at San Diego, La Jolla; [8]Department of Urology, [9]Chao Family Comprehensive Cancer Center, University of California at Irvine, Irvine; [10]Proveri Inc., San Diego, California; and [11]Banner Sun Health Research Institute, Sun City, Arizona

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

Z. Jia and Y. Wang are joint first authors.

Corresponding Author: Dan Mercola, Department of Pathology & Laboratory Medicine, University of California, Irvine, CA 92067. Phone: 949-824-1298; Fax: 949-824-2160; E-mail: dmercola@uci.edu or Michael McClelland, Vaccine Research Institute of San Diego, 10835 road to the cure suite 150, San Diego, CA 92121. Phone: 858-581-3960; Fax: 858-581-3970; E-mail: Mcclelland.michael@gmail.com

*American Association for **Cancer Research***

vival and growth of immortalized nontumorigenic human prostate epithelial cells as renal subcapsular xenografts required stroma from tumor-bearing prostate (10). Numerous studies have subsequently demonstrated large numbers of gene expression changes at the RNA level specific to the tumor microenvironment of prostate cancer (e.g., refs. 11–18). Similarly, a variety of protein expression changes have been associated with the microenvironment of prostate cancer. For example, reactive stroma, which is believed to occur in a subset of aggressive tumors, has been shown to correlate with changes in a variety of proteins including FGF2 (fibroblast growth factor); connective tissue growth factor; vimentin; actin, alpha, skeletal muscle; collagen, type I, alpha; and tenascin; some of which have been attributed to epithelial-derived TGF-β (12, 19).

Here, we investigate whether RNA expression changes may be identified that are sufficiently reliable to distinguish normal stroma from stroma near tumor. We have previously developed linear regression method for the identification of cell-type–specific expression of RNA from array data of prostate tumor samples (20). The method was validated using immunohistochemistry and using quantitative PCR applied to laser capture microdissection samples of tumor, stroma, and epithelia of benign prostate hyperplasia for 28 genes involving more than 400 measurements (20). Here, we have extended this approach to identify differentially expressed genes between normal volunteer prostate biopsy samples versus stroma from near tumors. More than a thousand gene expression changes were observed. A subset of stroma-specific genes was used to derive a classifier of 114 genes which accurately identifies tumor or nontumor status of a large number of independent test cases. The classifier may be useful in the diagnosis of stroma-rich biopsies from patients with equivocal pathology.

Materials and Methods

Prostate cancer patient samples and expression analysis

Data sets 1 and 2 (Table 1) are based on postprostatectomy frozen tissue samples obtained by informed consent using Institutional Review Board (IRB)-approved and HIPPA-compliant protocols. All tissues, except where noted, were collected at surgery and escorted to pathology for expedited review, dissection, and snap freezing in liquid nitrogen. In addition, data set 1 contains 27 prostate biopsy specimens obtained as fresh snap-frozen biopsy cores from 18 normal prostates. These samples were obtained from the control untreated subjects of a clinical trial to evaluate the role of difluoromethylornithine (DFMO) to decrease the prostate size of normal men. Eighteen of these were collected before the treatment period and 9 were collected after the treatment period had ended (21). Finally, 13 samples of normal prostate tissue were obtained from the rapid autopsy program of the Sun Health Research Institute (Sun City, AZ) and were frozen within 6 hours of demise.

RNA for expression analysis was prepared directly from frozen tissue following dissection of OCT (optimum cutting temperature compound) blocks with the aid of a cryostat. For expression analysis, 50 μg (10 μg for biopsy tissue) of total RNA samples were processed for hybridization to Affymetrix GeneChips. Expression analysis for all samples for data set 1 were assessed using the U133 Plus 2.0 platform, whereas for data set 2, the U133A platform was used. The data have been deposited in the Gene Expression Omnibus (GEO) database with accession numbers GSE17951 (data set 1) and GSE8218 (data set 2). For data sets 1 and 2, the distributions for the 4 principal cell types (tumor epithelial cells, stroma cells, epithelial cells of

Table 1. Datasets used in the study[a]

Data	Use	Platform	Number of Subjects	Number of Arrays	Array: Tumor/Nontumor/Normal	Reference
1A	Training and test	U133Plus2	P = 13	13	13/0/0	
			B = 15	15	0/0/15	
						GSE17951
			P = 87	95	55/40/0	
1B	Test only	U133Plus2	B = 9[b]	12	0/0/12	
			A = 13	13	0/0/13	
2	Test only	U133A	P = 82	136	65/71/0	GSE08218
3	Test only	U133A	P = 79	79	79/0/0	Unpublished, see ref. 22[c] GSE25136
4	Test only	U133A	P = 44	57	44/13/0	E-TABM-26 (24)

[a]P, samples from prostate cancer patients; B, biopsies from normal donors; A, prostate donated by rapid autopsy. Datasets 1 and 2 were collected from five participating institutions in San Diego County, CA. Demographic, pathology, and clinical values are individually recorded in shadow charts and maintained in the UCI SPECS consortium database.
[b]Seven patients from 1A were rebiopsied one year apart and applied to arrays.
[c]Data Set 3 was provided by William L. Gerald (Stephenson et al., ref. 22).

benign prostatic hyperplasia, and epithelial cells of dilated cystic glands) were estimated by 3 (data set 1) or 4 pathologists (data set 2), whose estimates were averaged as described (20).

Data sets 3 and 4 were independently developed and used as test sets (Table 1). Data set 3 consists of a series of 79 samples (22, 23), whereas data set 4 (24) is composed of 57 samples from 44 patients including 13 samples of stroma near tumor and 44 tumor-bearing samples. Expression analysis of the data sets was determined using the U133A platform.

Manual microdissection

Seventy-one of the tumor-bearing samples of data set 2 were manually microdissected to obtain tumor-adjacent stroma which was used for validation of the diagnostic classifier. For manual microdissection, the tumor-bearing tissue was embedded in an OCT block then mounted in a cryostat. Frozen sections were stained using hematoxylin and eosin (H&E) to visualize the location of the tumor. A border between tumor and adjacent stroma was marked on the glass slide using a Pilot Ultrafine Point Pen which was used as a guide to locate the border on the OCT block surface. Then the OCT-embedded block was etched with a single straight cut with a scalpel ($\sim$1-mm deep) to divide the embedded tissue into a tumor zone and tumor-adjacent stroma. Subsequent cryosections produced 2 halves at the site of the etched cut and were separately used for H&E staining and examined to confirm their composition. Multiple subsequent frozen sections of the tumor-adjacent stroma half were then pooled and used for RNA preparation and microarray hybridization. A final frozen section was used for H&E staining and examined to confirm that the tumor-adjacent stroma remained free of tumor cells.

Statistical tools implemented in R

The U133 Plus 2.0 platform used for data set 1 has about 55,000 probe sets whereas the U133A used for data sets 2, 3, and 4, contains 22,000 probe sets. Normalization was carried out across multiple data sets using the $\sim$22,000 probe sets in common to all data sets. First, data set 1 was quantile-normalized using the function "normalizeQuantiles" of LIMMA (Linear Models for Microarray Data) routine (25). Data sets 2 to 4 were then quantile-normalized by referencing normalized data set 1 using a modified function "REFnormalizeQuantiles" which was coded by Z.J. and is available at the SPECS Web site (26).

The LIMMA package from Bioconductor was used to detect differentially expressed genes.

Prediction Analysis of Microarray (PAM; ref. 27), implemented in R, was used to develop an expression-based classifier from the training sets and then applied to the test sets without further change.

A multiple linear regression (MLR) model was used to fit gene expression data and known percent cell-type composition for 4 cell types to estimate expression coefficients for each cell component (see Supplement for details). Percent cell-type distributions were estimated by 3 (data set 1) or 4 (data set 2) pathologists and exhibited an overall agreement of 4.3% SD for the 4 estimated cell types. The resulting significantly differentially expressed genes for the comparison of normal prostate biopsies to tumor-bearing prostate tissue were used for development of the diagnostic classifier.

Results

Development of a stroma-derived diagnostic classifier

We hypothesized that stroma within and directly adjacent to prostate cancer epithelial cells exhibits significant RNA expression changes compared with normal prostate stroma. To generate candidate biomarkers, we developed a 3-step strategy. First, we identified genes that are differentially expressed between tumor-adjacent stroma and normal stroma. Second, these differences were filtered by removing the age-related genes and removing the genes that are also expressed in tumor cells to create a stroma-specific set of differentially expressed genes. Finally, owing to the limited number of normal biopsies, we repeated steps 1 and 2 using a permutation procedure which greatly enhanced the extraction of information in the normal biopsies.

In step 1, Affymetrix gene expression data were acquired from normal frozen biopsies from each of the 15 subjects that were judged to be free of cancer by histologic examination of the 6 cores of the volunteer biopsies (21). Data from 13 of these samples (with 2 held in reserve as explained later) were compared with the gene expression data for 13 tumor-bearing patient cases from data set 1 selected with tumor cell content (T) greater than 0% but less than 10% (the average stroma content is $\sim$80%). These criteria ensured that the majority of stroma tissues included from the cancer-positive patients was close to tumor, whereas T less than 10% ensured that the impact from tumor cells is minimal to allow capture of altered expression signals from stroma cells rather than tumor cells. Using a moderated t test implemented in the LIMMA package of R (25), this comparison yielded 3,888 significant expression changes between these 2 groups with a value of $P < 0.05$. We used a relatively relaxed P value cutoff for the first step of feature selection to allow more genes to enter subsequent screening steps. The 3,888 probe sets were composed of a nearly equal number of up- and downregulated genes.

There was a substantial difference in age between the normal stroma group (average age = 51.9 years) and the near-tumor stroma group (average age = 60.6 years). In step 2, we compared the overall gene expression of the 13 normal stroma samples used for training versus 13 normal prostate specimens obtained by rapid autopsy (see "Materials and Methods") with an average age of 82 years. Prostate glands from the rapid autopsy series with an average age of 84 years exhibited a markedly increased heterogeneity of gland shapes with stroma containing increased fibroblast and myofibroblast-like cells. The comparison revealed 8,898 significant expression changes ($P < 0.05$). Of these probe sets, 1,678 were also detected in the comparison of normal stroma samples to stroma near tumor. After eliminating all of these potential aging-related genes, the remaining 2,210 probe sets consisted of nearly equal numbers of up- and downregulated genes.

It remained likely that some differential expression in this comparison included expression changes specific to the residual tumor cells or epithelium cells in some samples, rather than changes between more than one type of stromal cell. To reduce the possibility that epithelial cell–derived expression changes might influence subsequent results, we removed genes that appeared to be expressed in tumor at 10% or more of the expression in stroma. However, even "pure" tumor samples are contaminated with stroma thereby risking the elimination of genes expressed only in stroma. So, identification of genes expressed in tumor was achieved using MLR analysis (described in "Materials and Methods" and Supplement). The percent cell composition of 108 samples from 87 patients in data set 1 intentionally encompassing a wide range of tissue percentages was determined by a panel of 3 pathologists (20). The distribution is shown in Figure 1A. Model diagnostics showed that the fitted model for genes significantly expressed in tumor or stroma accounted for more than 70% of the total variation (i.e., the variation of error was less than 30% of the total variation), indicating a plausible modeling scheme.
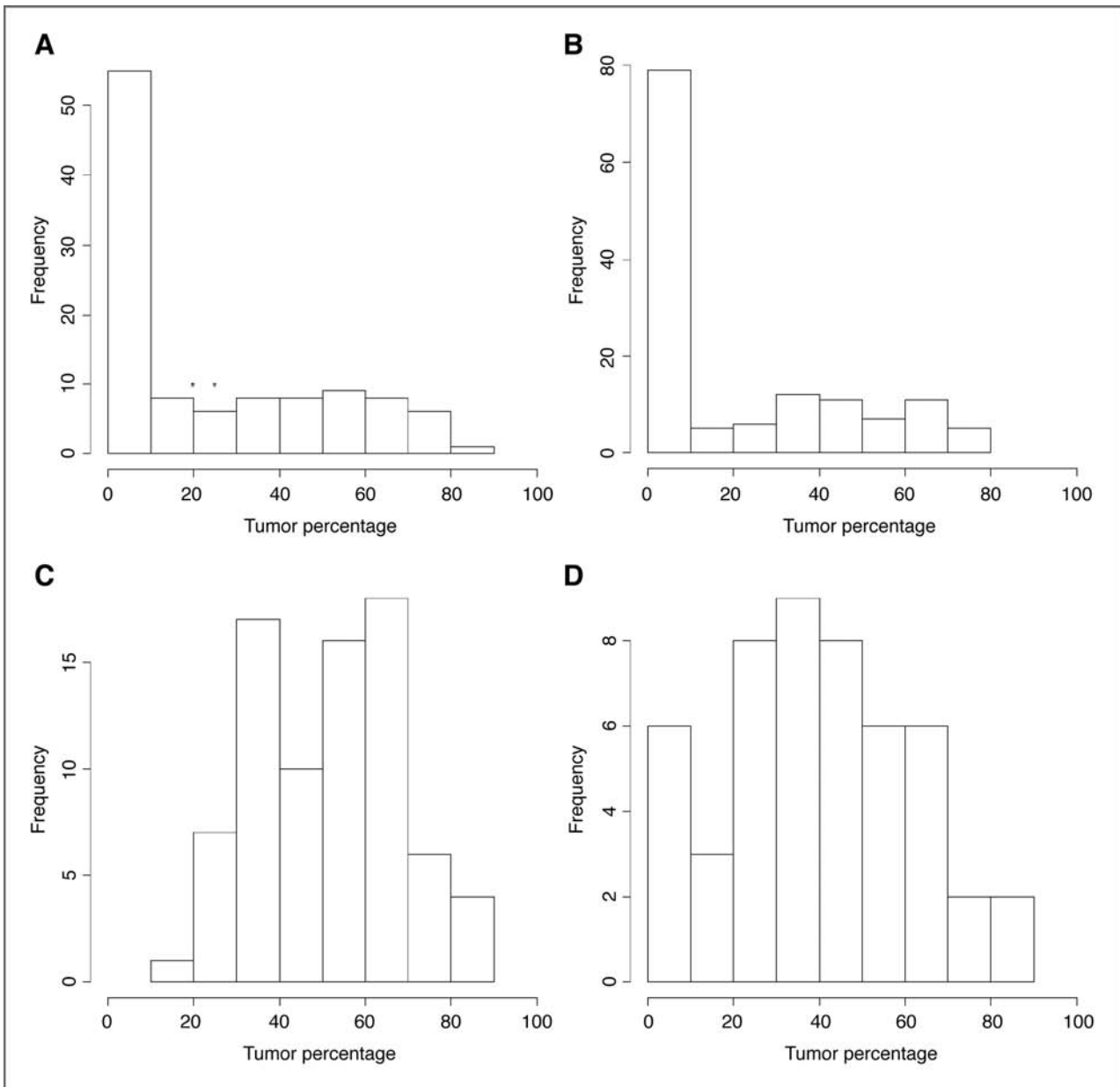

Figure 1. Histogram of tumor percentage for data sets 1 to 4. The tumor percentage data of (A) and (B) were provided by SPECS pathologists, whereas the tumor percentage data of (C) and (D) were estimated by CellPred program (29). The stars in (A) mark the tumor percentages of the misclassified tumor-bearing cases in data set 1, which CellPred indicates may actually be nontumor samples.

Of the 2,210 probe sets, derived above, we obtained 160 probe sets that were predominantly expressed in stroma cells and also show differential expression between near-tumor stroma and normal stroma. The average expression of these 160 probe sets was estimated to be more than 2-fold greater than the average of all genes expressed in stroma, which is a consequence for the filtering steps for robustness, and also favors good sensitivity.

Finally in step 3, a permutation analysis was performed. The above procedure for the generation of differentially expressed genes between 13 of the 15 normal stroma biopsies and the 13 biopsies of stroma near tumor was repeated using a different selection of 13 biopsy samples from 15, until all 105 possible combinations of 13 normal biopsy samples drawn from 15 samples ($C_{15}^{13} = 105$, where C_n^m is the number of combinations of m elements chosen from a total of n elements) was complete. After filtering for genes associated with aging (discussed earlier), a total of 339 probe sets that were differentially expressed between stroma near tumor and normal stroma were generated by the 105-fold gene selection procedure (the frequency of selection is summarized in Supplementary Fig. S1). Thus, the permutation increased the basis set by 339/160 or more than 2-fold. One hundred forty-six probe sets with at least 50 occurrences in the 105-fold permutation were selected for classifier construction (listed in Table 3)

PAM (28) was used to build a diagnostic classifier. The training set (Table 2, line 1) included all the 15 normal biopsies and the initial 13 samples of stroma near tumor. Of the 146 PAM input probe sets, 131 probe sets–-corresponding to 114 genes—were retained following the 10-fold cross-validation procedure of PAM (28) leading to a prediction accuracy of 96% (Table 2). Supplementary Figure S2 presents a "heatmap" of the relative expression of the 131 probe sets among all training samples. The separation of normal and near-tumor stroma samples of the training set by the classifier is illustrated by the 2 distinct populations shown in Figure 2.

Testing with independent data sets

The 131 probe set classifier was then tested on 243 samples that had not been used for training, and that all contained tumor, though usually very little tumor (Table 2, lines 2–5). Almost all the 243 samples were recognized as being from cancer patients with high average accuracy of about 99% (see Supplementary Table S1 for derived operating characteristics). Only 2 cases were misclassified. In Figure 1A, the 2 misclassified test are marked with asterisks. Although these samples are ostensibly given tumor percentages of 20% and 25% by pathologists, they are predicted to possibly contain little or no tumor using the CellPred program which estimates the tissue components using an *in silico* multiple-variate linear regression model (29). It is possible that these 2 exceptions were archived incorrectly and are not from patients with cancer or are from a very distant location relative to the tumor.

We examined whether the PAM classification results correlated with cell composition (Fig. 1). For the test cases of data sets 1 and 2, these values are known from the pathologists

Table 2. Operating characteristics for training and testing

Line		Data set	Number of Samples	Accuracy, %
1	**Training set**	1	28 (15 + 13)[a]	96.4
	Test set:			
	Tumor			
2	Tumor-bearing	1	55	96.4
3	Tumor-bearing	2	65	100
4	Tumor-bearing	3	79	100
5	Tumor-bearing	4	44	100
	Normal			
6	Biopsies (1)	1	7[b]	100
7	Biopsies (2)	1	5	60
8	Rapid autopsies	1	13	92.3
	Microdissected			
9	Stroma adjacent to tumor	2	71	97.1
10	Stroma adjacent to tumor	4	13	100
11	Stroma close to tumor	1	12	75
12	Stroma > 15 mm from tumor	1	28	35.7

[a]Fifteen normal biopsy samples and thirteen tumor-adjacent stroma samples.
[b]Patients from 1A of Table 1 were rebiopsied one year apart and applied to arrays.

estimates, whereas for data sets 3 and 4 (Fig. 1C and D, respectively), these tumor cell contents were estimated using the CellPRed program (29). Examining the tumor cell percentages in all the samples in Figure 1, it is clear that the PAM classification is successful on independent test samples with a broad range of tumor epithelial cells including samples with just a few percent of epithelial cells. These observations argue that the classifier is accurate in the categorization of prostate cancer cases independent of the presence or amount of the tumor epithelial component.

The classifier was then tested using specimens composed of *normal* prostate stroma and epithelium. Twelve biopsies from the DFMO study, all of them different from the 15 samples used earlier for training, were separated into 2 groups. In group 1, there were 7 second biopsies from the same participants whose first biopsy samples were included in the training set, taken 12 months later. These were accurately (100%) identified as nontumor (Table 2, line 6). In group 2, there were 5 biopsy samples not from subjects previously used for training. Two of these 5 biopsy samples were categorized as being from cancer patients (Table 2, line 7). When the histories for these volunteers were investigated, it was found that both donors had consistently exhibited elevated prostate-specific antigen (PSA) levels of 6.1 and

Table 3. One hundred forty-six diagnostic probe sets with incidence number greater than 50 for 105-fold gene selection procedure

Probe set	Gene symbol	Gene title	LogFC[a]
213764_s_at	MFAP5	Microfibrillar associated protein 5	−1.73
209758_s_at	MFAP5	Microfibrillar associated protein 5	−1.48
213765_at	MFAP5	Microfibrillar associated protein 5	−1.36
210280_at	MPZ	Myelin protein zero (Charcot-Marie-Tooth neuropathy 1B)	−1.20
210198_s_at	PLP1	Proteolipid protein 1 (Pelizaeus-Merzbacher disease, spastic paraplegia 2, uncomplicated)	−1.18
215104_at	NRIP2	Nuclear receptor interacting protein 2	−0.94
213847_at	PRPH	Peripherin	−0.93
214767_s_at	HSPB6	Heat shock protein, alpha-crystallin-related, B6	−0.88
209843_s_at	SOX10	SRY (sex determining region Y)-box 10	−0.61
209686_at	S100B	S100 calcium binding protein B	−0.94
209915_s_at	NRXN1	Neurexin 1	−0.80
214023_x_at	TUBB2B	Tubulin, beta 2B	−0.75
214954_at	SUSD5	Sushi domain containing 5	−0.98
204584_at	L1CAM	L1 cell adhesion molecule	−1.20
204777_s_at	MAL	Mal, T-cell differentiation protein	−0.99
205132_at	ACTC1	Actin, alpha, cardiac muscle 1	−0.99
203151_at	MAP1A	Microtubule-associated protein 1A	−0.69
210869_s_at	MCAM	Melanoma cell adhesion molecule	−0.71
204627_s_at	ITGB3	Integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	−0.82
209086_x_at	MCAM	Melanoma cell adhesion molecule	−0.61
219314_s_at	ZNF219	Zinc finger protein 219	−0.51
221204_s_at	CRTAC1	Cartilage acidic protein 1	−0.56
212886_at	CCDC69	Coiled-coil domain containing 69	−0.59
210814_at	TRPC3	Transient receptor potential cation channel, subfamily C, member 3	−0.75
212793_at	DAAM2	Dishevelled associated activator of morphogenesis 2	−0.56
212565_at	STK38L	Serine/threonine kinase 38 like	−0.58
214606_at	TSPAN2	Tetraspanin 2	−0.54
336_at	TBXA2R	Thromboxane A2 receptor	−0.65
218660_at	DYSF	Dysferlin, limb girdle muscular dystrophy 2B (autosomal recessive)	−0.55
214434_at	HSPA12A	Heat shock 70-kDa protein 12A	−0.57
212274_at	LPIN1	Lipin 1	−0.48
206874_s_at	–	–	−0.44
203939_at	NT5E	5′-nucleotidase, ecto (CD73)	−0.49
205954_at	RXRG	Retinoid X receptor, gamma	−0.53
219909_at	MMP28	Matrix metallopeptidase 28	−0.54
206425_s_at	TRPC3	Transient receptor potential cation channel, subfamily C, member 3	−0.57
205433_at	BCHE	Butyrylcholinesterase	−0.93
35846_at	THRA	Thyroid hormone receptor, alpha (erythroblastic leukemia viral (v-erb-a) oncogene homolog, avian)	−0.46
204736_s_at	CSPG4	Chondroitin sulfate proteoglycan 4	−0.55
202806_at	DBN1	Drebrin 1	−0.43
212097_at	CAV1	Caveolin 1, caveolae protein, 22kDa	−0.38
201841_s_at	HSPB1	Heat shock 27-kDa protein 1	−0.44

(Continued on the following page)

Table 3. One hundred forty-six diagnostic probe sets with incidence number greater than 50 for 105-fold gene selection procedure (Cont'd)

Probe set	Gene symbol	Gene title	LogFC[a]
206382_s_at	*BDNF*	Brain-derived neurotrophic factor	−0.62
219091_s_at	*MMRN2*	Multimerin 2	−0.44
205076_s_at	*MTMR11*	Myotubularin-related protein 11	−0.57
204159_at	*CDKN2C*	Cyclin-dependent kinase inhibitor 2C (p18, inhibits CDK4)	−0.46
212992_at	*AHNAK2*	AHNAK nucleoprotein 2	−0.60
206024_at	*HPD*	4-hydroxyphenylpyruvate dioxygenase	−0.57
218094_s_at	*DBNDD2 /// SYS1-DBNDD2*	Dysbindin (dystrobrevin binding protein 1) domain containing 2 /// SYS1-DBNDD2	−0.41
211276_at	*TCEAL2*	Transcription elongation factor A (SII)-like 2	−0.52
209191_at	*TUBB6*	Tubulin, beta 6	−0.51
213675_at	–	CDNA FLJ25106 fis, clone CBR01467	−0.44
211340_s_at	*MCAM*	Melanoma cell adhesion molecule	−0.46
210632_s_at	*SGCA*	Sarcoglycan, alpha (50-kDa dystrophin-associated glycoprotein)	−0.58
218651_s_at	*LARP6*	La ribonucleoprotein domain family, member 6	−0.34
207876_s_at	*FLNC*	Filamin C, gamma (actin binding protein 280)	−0.45
218877_s_at	*TRMT11*	tRNA methyltransferase 11 homolog (*S. cerevisiae*)	+0.44
219416_at	*SCARA3*	Scavenger receptor class A, member 3	−0.57
209981_at	*CSDC2*	Cold shock domain containing C2, RNA binding	−0.56
214212_x_at	*FERMT2*	Fermitin family homolog 2 (Drosophila)	−0.42
207554_x_at	*TBXA2R*	Thromboxane A2 receptor	−0.44
205231_s_at	*EPM2A*	Epilepsy, progressive myoclonus type 2A, Lafora disease (laforin)	−0.42
215306_at	–	mRNA; cDNA DKFZp586N2020 (from clone DKFZp586N2020)	−0.48
218435_at	*DNAJC15*	DNAJ (Hsp40) homolog, subfamily C, member 15	−0.49
203597_s_at	*WBP4*	WW domain binding protein 4 (formin binding protein 21)	−0.34
205303_at	*KCNJ8*	Potassium inwardly-rectifying channel, subfamily J, member 8	−0.42
201389_at	*ITGA5*	Integrin, alpha 5 (fibronectin receptor, alpha polypeptide)	−0.50
204940_at	*PLN*	Phospholamban	−0.49
220765_s_at	*LIMS2*	LIM and senescent cell antigen-like domains 2	−0.41
203299_s_at	*AP1S2*	Adaptor-related protein complex 1, sigma 2 subunit	−0.41
201344_at	*UBE2D2*	Ubiquitin-conjugating enzyme E2D 2 (UBC4/5 homolog, yeast)	−0.38
218648_at	*CRTC3*	CREB-regulated transcription coactivator 3	−0.33
204939_s_at	*PLN*	Phospholamban	−0.45
201431_s_at	*DPYSL3*	Dihydropyrimidinase-like 3	−0.40
215534_at	–	mRNA; cDNA DKFZp586C1923 (from clone DKFZp586C1923)	−0.46
209169_at	*GPM6B*	Glycoprotein M6B	−0.34
209651_at	*TGF-B1I1*	Transforming growth factor beta 1 induced transcript 1	−0.42
218711_s_at	*SDPR*	Serum deprivation response (phosphatidylserine binding protein)	+0.41

(Continued on the following page)

Table 3. One hundred forty-six diagnostic probe sets with incidence number greater than 50 for 105-fold gene selection procedure (Cont'd)

Probe set	Gene symbol	Gene title	LogFC[a]
212358_at	CLIP3	CAP-GLY domain containing linker protein 3	−0.47
218691_s_at	PDLIM4	PDZ and LIM domain 4	−0.42
218266_s_at	FREQ	Frequenin homolog (Drosophila)	−0.46
210319_x_at	MSX2	Msh homeobox 2	+0.45
218545_at	CCDC91	Coiled-coil domain containing 91	−0.31
44702_at	SYDE1	Synapse defective 1, Rho GTPase, homolog 1 (C. elegans)	−0.38
221014_s_at	RAB33B	RAB33B, member RAS oncogene family	−0.38
221246_x_at	TNS1	Tensin 1	−0.27
208789_at	PTRF	Polymerase I and transcript release factor	−0.42
220722_s_at	SLC5A7	Solute carrier family 5 (choline transporter), member 7	−0.41
209087_x_at	MCAM	Melanoma cell adhesion molecule	−0.40
221667_s_at	HSPB8	Heat shock 22kDa protein 8	−0.40
205561_at	KCTD17	Potassium channel tetramerisation domain containing 17	−0.32
213808_at	–	Clone 23688 mRNA sequence	−0.43
202565_s_at	SVIL	Supervillin	−0.36
211964_at	COL4A2	Collagen, type IV, alpha 2	−0.39
219563_at	C14orf139	Chromosome 14 open reading frame 139	−0.38
214122_at	PDLIM7	PDZ and LIM domain 7 (enigma)	−0.30
212589_at	RRAS2	Related RAS viral (r-ras) oncogene homolog 2	−0.29
205973_at	FEZ1	Fasciculation and elongation protein zeta 1 (zygin I)	−0.35
218818_at	FHL3	Four and a half LIM domains 3	−0.36
212120_at	RHOQ	Ras homolog gene family, member Q	−0.31
219073_s_at	OSBPL10	Oxysterol binding protein-like 10	−0.37
221480_at	HNRNPD	Heterogeneous nuclear ribonucleoprotein D (AU-rich element RNA binding protein 1, 37kDa)	−0.36
207071_s_at	ACO1	Aconitase 1, soluble	−0.27
211717_at	ANKRD40	Ankyrin repeat domain 40	−0.28
201313_at	ENO2	Enolase 2 (gamma, neuronal)	−0.36
204628_s_at	ITGB3	Integrin, beta 3 (platelet glycoprotein IIIa, antigen CD61)	−0.31
204303_s_at	KIAA0427	KIAA0427	−0.35
214439_x_at	BIN1	Bridging integrator 1	−0.29
209015_s_at	DNAJB6	DNAJ (Hsp40) homolog, subfamily B, member 6	−0.29
213547_at	CAND2	Cullin-associated and neddylation-dissociated 2 (putative)	−0.31
204058_at	ME1	Malic enzyme 1, NADP(+)-dependent, cytosolic	−0.34
219902_at	BHMT2	Betaine-homocysteine methyltransferase 2	−0.33
214306_at	OPA1	Optic atrophy 1 (autosomal dominant)	−0.27
210201_x_at	BIN1	Bridging integrator 1	−0.29
212509_s_at	MXRA7	Matrix-remodelling associated 7	−0.27
213231_at	DMWD	Dystrophia myotonica, WD repeat containing	−0.30
201843_s_at	EFEMP1	EGF-containing fibulin-like extracellular matrix protein 1	−0.32
206289_at	HOXA4	Homeobox A4	−0.29
203501_at	PGCP	Plasma glutamate carboxypeptidase	−0.30
216894_x_at	CDKN1C	Cyclin-dependent kinase inhibitor 1C (p57, Kip2)	−0.27
216500_at	–	HL14 gene encoding beta-galactoside-binding lectin, 3′ end, clone 2	−0.29
220050_at	C9orf9	Chromosome 9 open reading frame 9	−0.32

(Continued on the following page)

Table 3. One hundred forty-six diagnostic probe sets with incidence number greater than 50 for 105-fold gene selection procedure (Cont'd)

Probe set	Gene symbol	Gene title	LogFC[a]
209362_at	MED21	Mediator complex subunit 21	−0.26
202931_x_at	BIN1	Bridging integrator 1	−0.27
213480_at	VAMP4	Vesicle-associated membrane protein 4	−0.24
205611_at	TNFSF12	Tumor necrosis factor (ligand) superfamily, member 12	−0.29
204365_s_at	REEP1	Receptor accessory protein 1	−0.29
203389_at	KIF3C	Kinesin family member 3C	−0.26
205368_at	FAM131B	Family with sequence similarity 131, member B	−0.27
217066_s_at	DMPK	Dystrophia myotonica-protein kinase	−0.29
212457_at	TFE3	Transcription factor binding to IGHM enhancer 3	−0.25
200685_at	SFRS11	Splicing factor, arginine/serine-rich 11	−0.16
200788_s_at	PEA15	Phosphoprotein enriched in astrocytes 15	−0.22
202522_at	PITPNB	Phosphatidylinositol transfer protein, beta	−0.16
208869_s_at	GABARAPL1	GABA(A) receptor-associated protein like 1	−0.19
209524_at	HDGFRP3	Hepatoma-derived growth factor, related protein 3	−0.14
211347_at	CDC14B	CDC14 cell division cycle 14 homolog B (S. cerevisiae)	−0.21
211677_x_at	CADM3	Cell adhesion molecule 3	−0.21
212610_at	PTPN11	Protein tyrosine phosphatase, non-receptor type 11 (Noonan syndrome 1)	−0.23
212848_s_at	C9orf3	Chromosome 9 open reading frame 3	−0.27
214643_x_at	BIN1	Bridging integrator 1	−0.23
217820_s_at	ENAH	Enabled homolog (Drosophila)	−0.19
218597_s_at	CISD1	CDGSH iron sulfur domain 1	−0.18
221502_at	KPNA3	Karyopherin alpha 3 (importin alpha 4)	−0.20
222221_x_at	EHD1	EH-domain containing 1	−0.20
32625_at	NPR1	Natriuretic peptide receptor A/guanylate cyclase A (atrionatriuretic peptide receptor A)	−0.22

[a]logFC is the logarithm fold change as tumoros stroma being compared with normal stroma. +/− represents up- or downregulated expression level in tumoros stroma.
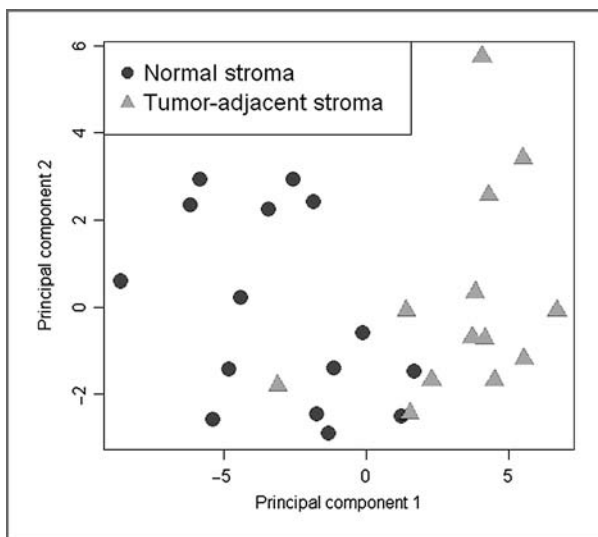

Figure 2. Plot of 2 principal components of training cases using the 131 probe set diagnostic classifier.

8 ng/mL (normal values < 3 ng/mL), respectively, although no tumor was observed in either of 2 sets of sextant biopsies obtained from these volunteers. The volunteers also had a history of prostate cancer in the family. All other donors of the normal biopsy volunteers exhibited normal PSA values. The IRB-approved protocol precluded following up further to establish that these patients had cancer that had been missed in the biopsies.

The classifier was then tested on 13 specimens obtained by rapid autopsy of individuals dying of unrelated causes (Table 2, line 8). Twelve of 13 of these samples, 92% accuracy, were classified as nontumor. Histologic examination of all embedded tissue of the one "misclassified" case revealed multiple foci of small "latent" tumors.

In summary, 25 nominally normal samples were classified as being from donors without prostate cancer or were classified in accordance with abnormal features that were subsequently uncovered. These results provide further support for the ability of the classifier to discriminate among normal and abnormal prostate tissue in the absence of histologically recognizable tumor cells in the samples studied.

Validation by manual microdissection, random classifiers, and the published literature

We sought to validate the classifier by developing histologic confirmed samples of stroma adjacent to tumor. An etching procedure was used to prepare 71 samples of tumor-adjacent stroma from patient tissues of data set 2, and 13 samples from data set 4. An additional 12 samples from data set 1 were obtained from OCT blocks entirely by manual microdissection, that is, without etching but leaving a margin of tissue between tumor and stroma, followed by histologically examining frozen section analysis of the OCT surface and bottom side of the pieces, to ensure the absence of tumor. These 12 manually excised pieces are termed "close stroma" (~3 mm). The expression values for all 96 samples were used to test the 131 probe set classifier using the PAM procedure. The accuracy in classifying that the samples were from patients with tumor was 97% for the 71 adjacent stroma samples from data set 2, 100% for 13 adjacent stroma samples from data set 4, and 75% for the 12 "close" stroma samples from data set 1 (Table 2, lines 9–11). This is an overall accuracy of 95% for the 96 independent samples.

Five of the 96 samples appeared "misclassified" as normal. Three of these misclassifications were among the 12 "close" stroma samples in data set 1. These 12 samples were obtained by manual excision and therefore some of the samples may not have been as near to tumor as the samples obtained by the etching method. Therefore, we examined how far the expression changes characteristic of tumor stroma may extend away from the tumor; we obtained 28 samples greater than 15 mm from any known tumor and generally from the contralateral lobe (Table 2, line 12). Only 10 of the 28 samples (36%) were categorized as tumor-associated stroma. Using the Fisher's exact test, the distribution for the 28 "remote" samples was significantly different from the 12 stroma samples from "close" to tumor of the *same* patient tissues ($P = 0.038$). This result, as well as the observation of a gradient of classification frequency values from 98%, 75%, and 36% for samples adjacent, close, and greater than 15 mm from tumor, suggests that the expression changes recognized by the classifier decline with increasing distance of stroma from tumor. Such observation bears on the likely mechanism for the production of differential gene expression in tumor-adjacent stroma which is generally believed to involve the influence of "paracrine" factors emanating from tumor foci (10, 30, 31).

We found that the normal samples and rapid autopsy samples can be easily distinguished from samples containing tumor using many of the individual genes (e.g., heatmap, Supplementary Fig. S3). However, the differences that allow near stroma to be distinguished from control stroma are more subtle and vary between patients, requiring a classifier based on a number of genes.

Further validation included a comparison with 100 random classifiers generated by arbitrarily sampling 131 probe sets for each classifier. The results (Supplementary Table S1 and Supplement) showed that these random classifiers had no diagnostic value, further indicating that the results obtained with the 131 probe set classifier cannot be attributed to chance.

Finally, we sought to validate that representative genes were in fact preferentially expressed in stroma by PCR. In addition, to test the translational relevance, we utilized independent cases from a formalin-fixed and paraffin-embedded (FFPE) clinical collection. Gene expression was assessed by a modified quantitative PCR procedure (see "Materials and Methods"). In a limited survey, 4 genes were found to have reliably preserved short amplicons. Blocks of 63 tumor cases were examined and tumor and stroma regions in H&E sections were demarcated by a pathologist (D.A.M.). Punches were removed from adjacent unstained sections and used for PCR for 63 tumor portions and 38 stroma portions. For all 4 genes, highly significant preferential expression in stroma was observed (Supplementary Table S4). These results for independent cases and by an independent method further support the preferential expression of these genes in tumor stroma and further argue that the classifier may be adapted to clinical biopsies preserved in FFPE, the standard method of archiving patient biopsies.

Finally, we also reviewed 2 recent studies describing expression analysis results for subclasses of the stroma of prostate cancer (16, 17), which showed consistent findings (see Supplement). In particular, the 339 probe sets (Affymetrix arrays) we identified map to 557 genes on Agilent arrays which have been used for deriving profiles for "reactive" stroma, a special case of adjacent stroma associated with poor outcome disease (17). A total of 31 genes or probe sets appeared to be concordant (in terms of gene identity and the direction of expression alteration) between the 339 probe sets (Affymetrix arrays) we identified in this study and the 557 mapped genes (Agilent arrays) in the "reactive" stroma study (17) with P value of 0.0001 (Supplementary Table S2). The formation of reactive stroma has been associated with poor prognosis (32). Thus, it is possible that some diagnostic markers in stroma could also be of prognostic interest.

Discussion

We compared the expression profiles of 15 normal biopsy samples and 13 tumor-adjacent stroma samples from prostatectomies using a permutation strategy to enhance detection of significant differences. About 3,800 significant gene expression changes were observed, which were then filtered to exclude genes known to be expressed at similar levels in epithelial tumor cells and to remove genes that change with age. The top-ranked 146 probe sets remaining after applying these filters were used for the 10-fold cross-validation procedure of PAM using the same 28 samples used for the initial training. The PAM procedure led to a 131 probe set classifier, which had a training accuracy of 96%. We then tested the classifier on a number of independent expression microarray data sets of tumor-bearing tissue including data from 110 samples generated by us (Table 2, data sets 1 and 2) and data from 123 samples generated elsewhere (Table 2, data sets 3 and 4). These samples were classified as being from cancer patients with an overall accuracy of 98%, a value compares favorably with the diagnostic accuracy of PSA-based methods of about 70% (33). Only 2 samples recorded as containing tumor cells were misclassified. Upon further investigation of these 2 samples using CellPred, a method to determine the tumor percentage of

samples based solely on their expression profile (29), these samples were predicted to have little or no tumor, although they had been recorded as having more than 20% tumor, indicating their assignment as tumor may have been a bookkeeping error. Similarly, we generated data from 25 samples of normal prostate, which were recognized as nontumor with an accuracy of 92%. Only 3 samples were "misclassified". Two of these samples were biopsies donated by men with abnormally high PSA levels and a family history of prostate cancer, although no tumor was recognized in any of the sextant biopsies taken at the beginning and end of the study period for which these volunteers were controls. In addition, one sample derived from the rapid autopsy donors was potentially "misclassified" as cancerous. Examination of multiple blocks of the glands taken from both lobes and all zones revealed tumor foci in the misclassified case. Thus, the "misclassifications" correlate well with the unusual clinical and pathologic features of the cases. In summary, the handful of misclassifications of tumor and normal each had evidence that they had been mislabeled before the test, potentially raising the actual sensitivity and specificity for classifying these samples to 100%.

Finally, for validation, we used 153 samples from data sets 1 and 4 to prepare "pure" stroma adjacent, close, and far (>15 mm) from known tumor foci. These data sets were able to detect the presence of tumor in the prostate with a decreasing accuracy of 98%, 75%, and 36%, respectively. The observation of a gradual reduction in the sensitivity of the classifier as the distance increases bears on the likely mechanism for the production of differential gene expression in tumor-adjacent stroma which is generally believed to involve the influence of "paracrine" factors emanating from tumor foci (10, 30, 31). Indeed, the tumor microenvironment is likely the source of factors that are required for tumor formation by the epithelial component (10). The amount of diffusible paracrine factors of this complex interaction mechanism likely declines with separation of target cells from the secreting cells. Indeed, a simple radial dilution model would predict a decline of effects of tumor-derived factors by at least the square of the distance of target stroma cells from a tumor focus. On the basis of this simple model, the decrease in the frequency of categorization stroma taken from more than 15 mm from a known tumor focus to 36% suggests a 50% recognition distance of about 13 mm in fresh frozen tissue. In view of the modest average fold-change of the 131 probe sets of the classifier (Table 3), the distance at which "presence of tumor" is recognized suggests a surprisingly large range of "influence" of tumor more than steady-state gene expression changes in nearby stroma. Systematic studies of differential expression as a function of known distances will be required to confirm and refine this inference.

The classifier developed here used highly selective methods to enrich for mesodermal and ectodermal derivatives compared with endoderm/epithelial derivatives. Computer-assisted gene enrichment analysis classification using DAVID (34) identified a number of statistically significant gene enrichment categories. The 10 most significant categories are summarized in Supplementary Table S3. Numerous genes associated with expression in nerve and muscle are apparent, such as the nine genes of the *actin cytoskeleton*

enrichment category, and in the *disease mutation* category including MPZ (Charcot-Maire-Tooth neuropathy 1b), optic atrophy 1, EPM2a (Lafora disease), BDGF, PLN (phospholamban), SGCA (dystophin-associated glycoprotein), and EFEMP. Biochemical associations include genes related to the TGF-β pathway (SMAD3, TGFIT, ID4, and CKDN1C/p57), the Wnt pathway (FZD7, SMAD3, DAAM1, and WISP2) and interacting genes (*PCH12*, *PCDH7*, and *CDH19*). These pathways are associated with tumor–stroma paracrine interactions (16, 17, 32, 35, 36). Given that reactive stroma has been associated with poor prognosis (32), it is possible that some of the 131 diagnostic markers identified in stroma could also be of prognostic interest. Nevertheless, we have not ruled that classifier developed here can distinguish other prostate conditions such as acute and chronic inflammation of the prostate and, therefore, stroma near these lesions may conceivably be misdiagnosed. Additional work with samples containing such lesions could identify genes that distinguish inflammation from cancer.

Our preclinical results suggest practical applications. Assessment of suspicious initial biopsies for expression of the classifier genes was identified here by microarray but could also potentially by any number of other biomarker methods, including those available for assessment of RNA, protein, or epigenetic markers in FFPE samples. Such quantitation may have use in defining "presence of tumor" based solely on the detection of changes in the microenvironment near a focus of tumor by quantitative criteria. Such a method would be applicable to cases with an initial negative biopsy results that would otherwise be referred for rebiopsy owing to the presence of ASAP or PIN. The determination of "presence of tumor" may strengthen guidance for neoadjuvant therapy or prevention therapy or an accelerated scheduling of rebiopsy. Finally, because stroma facilitates tumor growth (10), the expression changes that occur in stroma indicating the presence of tumor might be targets for therapeutic intervention that could leave normal stroma relatively unaffected.

Disclosure of Potential Conflicts of Interest

Acknowledgments

Grant Support

References

1. Marks LS, Bostwick DG. Prostate cancer specificity of PCA3 gene testing: examples from clinical practice. Rev Urol 2008;10:175–81.

2. O'Dowd G J, Miller MC, Orozco R, Veltri RW. Analysis of repeated biopsy results within 1 year after a noncancer diagnosis. Urology 2000;55:553–9.

3. Che M, Sakr W, Grignon D. Pathologic features the urologist should expect on a prostate biopsy. Urol Oncol 2003;21:153–61.

4. Pepe P, Aragona F. Saturation prostate needle biopsy and prostate cancer detection at initial and repeat evaluation. Urology 2007; 70:1131–5.

5. Andriole GL, Bullock TL, Belani JS, Traxel E, Yan Y, Bostwick DG, et al. Is there a better way to biopsy the prostate? Prospects for a novel transrectal systematic biopsy approach. Urology 2007;70 Suppl:22–6.

6. Mian BM, Naya Y, Okihara K, Vakar-Lopez F, Troncoso P, Babaian RJ. Predictors of cancer in repeat extended multisite prostate biopsy in men with previous negative extended multisite biopsy. Urology 2002;60:836–40.

7. Leite KR, Camara-Lopes LH, Cury J, Dall'oglio MF, Sanudo A, Srougi M. Prostate cancer detection at rebiopsy after an initial benign diagnosis: results using sextant extended prostate biopsy. Clinics 2008;63:339–42.

8. Amin MM, Jeyaganth S, Fahmy N, Bégin L, Aronson S, Jacobson S, et al. Subsequent prostate cancer detection in patients with prostatic intraepithelial neoplasia or atypical small acinar proliferation. Can Urol Assoc J 2007;1:245–9.

9. Cunha GR, Hayward SW, Wang YZ. Role of stroma in carcinogenesis of the prostate. Differentiation 2002;70:473–85.

10. Cunha GR, Hayward SW, Wang YZ, Ricke WA. Role of the stromal microenvironment in carcinogenesis of the prostate. Int J Cancer 2003;107:1–10.

11. Ernst T, Hergenhahn M, Kenzelmann M, Cohen CD, Bonrouhi M, Weninger A, et al. Decrease and gain of gene expression are equally discriminatory markers for prostate carcinoma: a gene expression analysis on total and microdissected prostate tissue. Am J Pathol 2002;160:2169–80.

12. Tuxhorn JA, Ayala GE, Smith MJ, Smith VC, Dang TD, Rowley DR. Reactive stroma in human prostate cancer: induction of myofibroblast phenotype and extracellular matrix remodeling. Clin Cancer Res 2002;8:2912–23.

13. Chandran UR, Dhir R, Ma C, Michalopoulos G, Becich M, Gilbertson J. Differences in gene expression in prostate cancer, normal appearing prostate tissue adjacent to cancer and prostate tissue from cancer free organ donors. BMC Cancer 2005;5:45.

14. Yang SZ, Dong JH, Li K, Zhang Y, Zhu J. Detection of AFPmRNA and melanoma antigen gene-1mRNA as markers of disseminated hepatocellular carcinoma cells in blood. Hepatobiliary Pancreat Dis Int 2005;4:227–33.

15. Verona EV, Elkahloun AG, Yang J, Bandyopadhyay A, Yeh IT, Sun LZ. Transforming growth factor-beta signaling in prostate stromal cells supports prostate carcinoma growth by up-regulating stromal genes related to tissue remodeling. Cancer Res 2007;67:5737–46.

16. Richardson AM, Woodson K, Wang Y, Rodriguez-Canales J, Erickson HS, Tangrea MA, et al. Global expression analysis of prostate cancer-associated stroma and epithelia. Diagn Mol Pathol 2007;16: 189–97.

17. Dakhova O, Ozen M, Creighton CJ, Li R, Ayala G, Rowley D, et al. Global gene expression analysis of reactive stroma in prostate cancer. Clin Cancer Res 2009;15:3979–89.

18. Van Der Heul-Nieuwenhuijsen L, Dits N, Van Ijcken W, de Lange D, Jenster G. The FOXF2 pathway in the human prostate stroma. Prostate 2009;69:1538–47.

19. Yang F, Tuxhorn JA, Ressler SJ, McAlhany SJ, Dang TD, Rowley DR. Stromal expression of connective tissue growth factor promotes angiogenesis and prostate cancer tumorigenesis. Cancer Res 2005;65:8887–95.

20. Stuart RO, Wachsman William, Berry Charles C, Arden Karen, Goodison Steven, Klacansky Igor, , et al. In silico dissection of cell-type associated patterns of gene expression in prostate cancer. Proc Natl Acad Sci U S A 2004;101:615–20.

21. Simoneau AR, Gerner EW, Nagle R, Ziogas A, Fujikawa-Brooks S, Yerushalmi H, et al. The effect of difluoromethylornithine on decreasing prostate size and polyamines in men: results of a year-long phase IIb randomized placebo-controlled chemoprevention trial. Cancer Epidemiol Biomarkers Prev 2008;17:292–9.

22. Stephenson AJ, Smith A, Kattan MW, Satagopan J, Reuter VE, Scardino PT, et al. Integration of gene expression profiling and clinical variables to predict prostate carcinoma recurrence after radical prostatectomy. Cancer 2005;104:290–8.

23. Sun Y, Goodison S. Optimizing molecular signatures for predicting prostate cancer recurrence. Prostate 2009;69:1119–27.

24. Liu P, Ramachandran S, Ali Seyed M, Scharer CD, Laycock N, Dalton WB, et al. Sex-determining region Y box 4 is a transforming oncogene in human prostate cancer cells. Cancer Res 2006;66:4011–9. Available from: http://www.ebi.ac.uk/arrayexpress/browse.html?keywords=E-TABM-26.

25. Dalgaard P. Statistics and Computing: Introductory Statistics with R. New York: Springer-Verlag Inc.; 2002. p. 260.

26. SPECS Web site. Available from: http://www.pathology.uci.edu/faculty/mercola/UCISPECSHome.html [cited 2007 May 21].

27. Guo Y, Hastie T, Tibshirani R. Regularized linear discriminant analysis and its application in microarrays. Biostatistics 2007;8:86–100.

28. Tibshirani R, Hastie T, Narasimhan B, Chu G. Diagnosis of multiple cancer types by shrunken centroids of gene expression. Proc Natl Acad Sci U S A 2002;99:6567–72.

29. Wang Y, Xiao-Qin Xia, Zhenyu Jia, Anne Sawyers, Huazhen Yao, Jessica Wang-Rodriquez, et al. In silico estimates of tissue components in surgical samples based on expression profiling data using. Cancer Res In press. Available from: http://webarraydborg/webarray/indexhtml.

30. Tuxhorn JA, Ayala GE, Rowley DR. Reactive stroma in prostate cancer progression. J Urol 2001;166:2472–83.

31. Rowley DR. What might a stromal response mean to prostate cancer progression? Cancer Metastasis Rev 1998;17:411–9.

32. Yanagisawa N, Li R, Rowley D, Liu H, Kadmon D, Miles BJ, et al. Stromogenic prostatic carcinoma pattern (carcinomas with reactive stromal grade 3) in needle biopsies predicts biochemical recurrence-free survival in patients after radical prostatectomy. Hum Pathol 2007;38:1611–20.

33. Shariat SF, Scardino PT, Lilja H. Screening for prostate cancer: an update. Can J Urol 2008;15:4363–74.

34. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, et al. DAVID: Database for Annotation, Visualization, and Integrated Discovery. Genome Biol 2003;4:P3.

35. Tuxhorn JA, McAlhany SJ, Yang F, Dang TD, Rowley DR. Inhibition of transforming growth factor-beta activity decreases angiogenesis in a human prostate cancer-reactive stroma xenograft model. Cancer Res 2002;62:6021–5.

36. Zhang Q, Helfand BT, Jang TL, Zhu LJ, Chen L, Yang XJ, et al. Nuclear factor-kappaB-mediated transforming growth factor-beta-induced expression of vimentin is an independent predictor of biochemical recurrence after radical prostatectomy. Clin Cancer Res 2009;15: 3557–67.

Cancer Research

Diagnosis of Prostate Cancer Using Differentially Expressed Genes in Stroma

Zhenyu Jia, Yipeng Wang, Anne Sawyers, et al.

Updated version	Access the most recent version of this article at: http://cancerres.aacrjournals.org/content/71/7/2476
Supplementary Material	Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2011/03/25/71.7.2476.DC1.html

Cited Articles	This article cites by 33 articles, 11 of which you can access for free at: http://cancerres.aacrjournals.org/content/71/7/2476.full.html#ref-list-1
Citing articles	This article has been cited by 2 HighWire-hosted articles. Access the articles at: http://cancerres.aacrjournals.org/content/71/7/2476.full.html#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions	To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.