Tumor Immunobiological Differences in Prostate Cancer between African-American and European-American Men

Tiffany A. Wallace,[1] Robyn L. Prueitt,[1] Ming Yi,[3] Tiffany M. Howe,[1] John W. Gillespie,[2] Harris G. Yfantis,[4] Robert M. Stephens,[3] Neil E. Caporaso,[5] Christopher A. Loffredo,[6] and Stefan Ambs[1]

[1]Laboratory of Human Carcinogenesis and [2]Urologic Oncology Branch, Center for Cancer Research, National Cancer Institute, NIH, Bethesda, Maryland; [3]Advanced Biomedical Computing Center, National Cancer Institute-Frederick/Science Applications International Corporation-Frederick, Inc., Frederick, Maryland; [4]Pathology and Laboratory Medicine, Baltimore Veterans Affairs Medical Center, Baltimore, Maryland; [5]Genetic Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, Rockville, Maryland; and [6]Cancer Genetic and Epidemiology Program, Lombardi Cancer Center, Georgetown University Medical Center, Washington, District of Columbia

Abstract

The incidence and mortality rates of prostate cancer are significantly higher in African-American men when compared with European-American men. We tested the hypothesis that differences in tumor biology contribute to this survival health disparity. Using microarray technology, we obtained gene expression profiles of primary prostate tumors resected from 33 African-American and 36 European-American patients. These tumors were matched on clinical variables. We also evaluated 18 nontumor prostate tissues from seven African-American and 11 European-American patients. The resulting datasets were analyzed for expression differences on the gene and pathway level comparing African-American with European-American patients. Our analysis revealed a significant number of genes, e.g., 162 transcripts at a false-discovery rate of ≤5% to be differently expressed between African-American and European-American patients. Using a disease association analysis, we identified a common relationship of these transcripts with autoimmunity and inflammation. These findings were corroborated on the pathway level with numerous differently expressed genes clustering in immune response, stress response, cytokine signaling, and chemotaxis pathways. Several known metastasis-promoting genes, including *autocrine mobility factor receptor, chemokine (C-X-C motif) receptor 4*, and *matrix metalloproteinase 9*, were more highly expressed in tumors from African-Americans than European-Americans. Furthermore, a two-gene tumor signature that accurately differentiated between African-American and European-American patients was identified. This finding was confirmed in a blinded analysis of a second sample set. In conclusion, the gene expression profiles of prostate tumors indicate prominent differences in tumor immunobiology between African-American and European-American men. The profiles portray the existence of a distinct tumor microenvironment in these two patient groups. [Cancer Res 2008;68(3):927–36]

Introduction

Prostate cancer is the most frequently diagnosed malignancy and the second leading cause of cancer mortality in men older than 40 years in the United States (1). The incidence and mortality rates of prostate cancer vary substantially among different geographic areas and ethnic groups (2). Most notably, African-American men have the highest risk of developing prostate cancer, and due to the manifestation of a more aggressive disease, they have over twice the mortality rate of European-American men (3, 4). Whereas socioeconomic factors contribute to this health disparity (4, 5), they do not fully explain the differences in prostate cancer incidence, aggressiveness, and mortality among different race/ethnic groups in the United States (6, 7).

It has been proposed that differences in tumor biology contribute to the health disparity associated with prostate cancer (4). To date, the contribution of genetic factors to the higher prostate cancer incidence and mortality in African-American men remains controversial (8, 9). Whereas some studies have investigated race/ethnic differences in prostate cancer tumor biology, few differences have been identified (10, 11). Recently, the chromosomal region of 8q24 has been strongly implicated as a site of prostate cancer susceptibility (12, 13). In these studies, men of African ancestry were found to carry at-risk alleles at a higher frequency than men of European ancestry. Those results indicate that genetic factors may account for some of the excess prostate cancer risk that African-American men experience when compared with other race/ethnic groups.

We pursued the hypothesis that the gene expression profiles of prostate tumors from African-American and European-American men may reveal biological differences between the two groups that could explain the more aggressive cancer phenotype in African-American men. Thus, we performed genome-wide gene expression profiling in a large set of tumor samples that were matched on selected clinical variables. The resulting datasets were analyzed on gene and pathway levels to reveal the differences in tumor biology between African-American and European-American patients.

Subjects and Methods

Collection of tissue specimens. A total of 69 fresh-frozen prostate tumors were obtained from the National Cancer Institute Cooperative Prostate Cancer Tissue Resource (CPCTR) and the Department of Pathology at the University of Maryland. All tumors were resected adenocarcinomas that had not received any therapy before prostatectomy. The macrodissected CPCTR tumor specimens (n = 59) were reviewed by a

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

Requests for reprints: Stefan Ambs, Laboratory of Human Carcinogenesis, National Cancer Institute, Building 37/Room 3050B, Bethesda, MD 20892-4258. Phone: 301-496-4668; E-mail: ambss@mail.nih.gov.

doi:10.1158/0008-5472.CAN-07-2608

CPCTR-associated pathologist, who confirmed the presence of tumor in the specimens. These tissues were collected between 2002 and 2004 at four different sites, with each site providing tissues from both African-American and European-American patients. Information on race/ethnicity (33 African-Americans and 36 European-Americans) was either extracted from medical records (CPCTR) or obtained through an epidemiologic questionnaire in which race/ethnicity was self-reported (University of Maryland). Only one patient, a European-American, was also Hispanic. Surrounding nontumor prostate tissue was collected from 18 of the recruited patients in this study. Of those, seven were African-American men and 11 were European-American men. We also isolated total RNA from 10 needle biopsy specimens collected from patients at the National Naval Medical Center (one African-American and nine European-Americans) who did not have prostate cancer. From those, we prepared two RNA pools, each representing five patients. Clinicopathologic characteristics of the patients, including age at prostatectomy, histology, Gleason score, pathologic stage, PSA at diagnosis, tumor size, extraprostatic extension, margin involvement, and seminal vesicle invasion were obtained from CPCTR. For University of Maryland cases, this information was extracted from the medical and pathology records, if available. Written informed consent was obtained from all donors. Tissue collection and study design were approved by the institutional review boards of the participating institutions.

Second tumor set for validation. For validation of a two-gene tumor signature that distinguished between African-American and European-American prostate cancer patients, we obtained 39 prostate tumor samples that were collected between 1991 and 1994 (14). The characteristics of this sample set were not revealed to the authors to assure a fully blinded analysis. These samples were processed like all other samples in the study.

RNA extraction. Total RNA was isolated using the TRIZOL reagent according to the manufacturer's instructions (Invitrogen). RNA integrity for each sample was confirmed with the Agilent 2100 Bioanalyzer (Agilent Technologies).

Affymetrix microarrays. RNA labeling and hybridization were performed according to Affymetrix standard protocols. Briefly, 5 μg of total RNA were reverse-transcribed with an oligo (dT) primer that has a T7 RNA polymerase promoter at the 5′ end. Second-strand synthesis was followed by cRNA production with incorporation of biotinylated ribonucleotides using the BioArray high-yield RNA transcript labeling kit T3 from Enzo Life Sciences. The labeled cRNA was fragmented and hybridized to Affymetrix GeneChip HG-U133A 2.0 arrays. This array contains 22,283 probe sets that represent ~13,000 human genes. Hybridization signals were visualized with phycoerythrin-conjugated streptavidin (Invitrogen) and scanned using a GeneChip Scanner 3000 7G (Affymetrix). In accordance with Minimum Information About a Microarray Experiment guidelines, we deposited the CEL files for the microarray data and additional patient information into the GEO repository.[7] The GEO submission accession number is GSE6956.

Data normalization and statistical analysis. All chips were normalized using the robust multichip analysis procedure (15). To generate lists of significantly differently expressed genes, the resulting data set was subjected to the significance analysis of microarray procedure (16). We generated gene lists based on both P values from two-sided t tests and intended false discovery rates (FDR). The FDR calculation followed the method described by Storey and Tibshirani (17). Prediction analysis for microarrays (PAM) was used to classify patients as either African-American or European-American (18). In this analysis, the threshold δ was chosen based on the best compensation for both training error rates and coefficient of variation error rates. Cross-validation was performed by leaving out 10% of samples to determine the appropriate threshold variable in PAM.

Pathway and disease association analyses. The pathway analysis was performed with the in-house WPS software (19). Biological processes and pathways were annotated according to Gene Ontology Biological Processes (GOBP; Gene Ontology Consortium[8]) and the BioCarta pathway collection.[9] These two annotation methods are complementary because GOBP uses substantially more genes for functional annotation than does BioCarta, but the GOBP terms are not as detailed and well-defined as the pathway mappings in BioCarta. Our database had 16,762 human genes annotated for GOBP and 1,429 genes for BioCarta pathways. A one-sided Fisher's exact test was used to determine which biological processes and pathways had a statistically significant enrichment of differently expressed genes. Genes were included into the pathway analysis based on the FDR (≤30%) of their corresponding probesets on the microarray. If several probesets encoded the same gene, the software recognized this and assured that the gene was counted only once for significance testing at the pathway level. In addition, a permutated P value or FDR for enrichment in a biological process/pathway was calculated with a permutation method. To calculate the permutated P value, randomly selected genes were assigned to a biological process/pathway without changing the total number of genes that were originally assigned to this pathway, and a P value for enrichment was calculated. This procedure was iterated 300 times, and the permutated P value for enrichment was calculated as the fraction of random trials resulting in a Fisher's exact test P value less than the original Fisher's exact P value without the permutations. We then compiled the Fisher's exact test results for cluster analyses and displayed the results in color-coded heat maps to reveal the patterns of significantly altered biological processes and pathways. The color coding of the heat maps is related to the enrichment of genes in a pathway [−log(P value)–based] with red indicating a higher enrichment.

The disease association analysis was conducted using both our WPS software and information provided by the genetic association database (20). The WPS software assessed whether differently expressed genes (FDR, ≤30%) between prostate tumors from African-American and European-American patients had previously been associated with other diseases, as indicated by the genetic association database, and whether these disease-associated genes, e.g., systemic lupus erythematosus genes, were statistically significantly enriched in the prostate tumor gene signature using the one-sided Fisher's exact test. In addition, we calculated a permutated P value or FDR for disease association of the differently expressed genes. The permutated P value was calculated following the procedures described in the previous paragraph. The color coding of the heat map that displays the results of the disease association analysis is related to the enrichment of disease genes [−log(P value)–based], e.g., systemic lupus erythematosus genes, among the differently expressed genes between prostate tumors from African-American and European-American patients.

Quantitative real-time PCR. One hundred nanograms of total RNA were reverse-transcribed using the high-capacity cDNA archive kit (Applied Biosystems). Quantitative real-time PCR (qRT-PCR) was subsequently performed in triplicate using TaqMan Gene Expression Assays (Applied Biosystems), which include preoptimized probe and primer sets specific for the genes being validated. The assay ID numbers of the validated genes are as follows: phosphoserine phosphatase homologue (*PSPHL*), Hs00863464_m1; matrix metallopeptidase 9 (*MMP9*), Hs00234579_m1; autocrine mobility factor receptor (*AMFR*), Hs00181609_m1; and chemokine (C-X-C motif) receptor 4 (*CXCR4*), Hs00607978_s1. Data were collected using the ABI PRISM 7500 sequence detection system. The 18s RNA was used as the internal standard reference. Normalized expression was calculated using the comparative C_t method as described, and fold changes were derived from the $2^{-\Delta\Delta Ct}$ values for each gene (21). Graphs were prepared using relative C_t values that were calculated by subtracting the 18s C_t values from the corresponding C_t values for the gene being measured. For illustration purposes only, an arbitrary value of 20 was uniformly added to all relative C_t values. The statistical analysis of these data was performed with a two-sided t test if the expression data showed normal distribution or

[7] http://www.ncbi.nlm.nih.gov/geo/

[8] http://www.geneontology.org

[9] http://www.biocarta.com/genes/allPathways.asp

with a two-sided Wilcoxon rank-sum test if the expression data did not follow normal distribution.

Results

Tumor selection and gene expression analysis. Prostate tumors were collected from 33 African-American and 36 European-American patients with localized disease. Tumor samples were similar for age at prostatectomy, pathologic stage of the disease, and Gleason score between the African-American and European-American patients (Table 1). African-American patients tended to have higher occurrence of seminal vesicle invasion than European-American patients, which is consistent with the literature (22).

Gene expression profiles from the African-American and European-American prostate tumors were generated using the Affymetrix GeneChip HG-U133A 2.0 array that represents ~13,000 human genes. In addition, 18 nontumor tissues and two RNA pools from cancer-free prostates were used for genetic comparisons. In an initial analysis of our dataset, we generated the gene expression profiles from tumor and nontumor tissues using the combined dataset (69 tumors and 20 nontumor samples) and obtained the gene signature that differentiated tumor and nontumor tissues. For cross-validation of our dataset, we compared these results with those from a published meta-analysis of four prostate cancer gene expression datasets (23). Despite the use of different technology platforms, we found a very good agreement between our results and the results of the meta-analysis. Of the top 40 overexpressed genes in prostate tumors across the meta-analysis, 21 were also found to be significantly up-regulated (FDR, ≤5%) in our study. Of the top 40 underexpressed genes in prostate tumors across the meta-analysis, six were significantly down-regulated (FDR, ≤5%) in our study (Supplementary Table S1). Specifically, previously validated marker genes for prostate tumors, such as the genes encoding α-methylacyl-CoA racemase, hepsin, and fatty acid synthase, were highly up-regulated in the tumors of our dataset. Signature tumor genes of other studies, such as *MUC1* and *HOXC6* (24, 25), were also found be overexpressed in these tumors.

The comparison of gene expression profiles from African-American and European-American tumors revealed 162 transcripts, 280 transcripts, and 489 transcripts, to be differently expressed between the two groups at FDRs of ≤5%, ≤ 10%, and ≤ 30%, respectively (Supplementary Table S2). We compared this gene list with the list of genes that were described to be differently expressed between tumor and nontumor tissue in the aforementioned meta-analysis (23). Of the top 80 tumor genes in the meta-analysis, none was found to be differently expressed between African-American patients and European-American patients at the FDR of ≤5% threshold. Only one of those 80 genes, the gene encoding fibrillin 1, was differently expressed at a less stringent FDR of ≤30%. These data suggest that previously identified marker genes for prostate cancer do not significantly differ between African-American and European-American prostate tumors.

Disease and pathway association of transcripts that are differently expressed by race/ethnicity. Bioinformatics enables us to determine common associations between a gene list and disease classes and pathways. We used our in-house software (19) to assess disease and pathway associations of genes that were differently expressed in tumors by race/ethnicity. The disease association analysis revealed associations of these genes with

Table 1. Clinical characteristics of the study population

	All cases (n = 69)	African-American (n = 33)	European-American (n = 36)	P^* (t test)
Age at prostatectomy [median (range)] n = 69	60 (44–73)	61 (46–72)	60 (44–73)	0.77
PSA at diagnosis [median (range)] n = 50 †	6.1 (1.3–47.7)	6.1 (1.3–47.7)	6.1 (4.0–20.0)	0.23
Largest nodule (g) [median (range)] n = 59 †	1.6 (0.2–2.9)	1.5 (0.8–2.9)	1.6 (0.2–2.8)	0.97
	N (%)	N (%)	N (%)	P (Fisher's exact test)
Source of tissue				
National Cancer Institute CPCTR, n (%)	59 (86)	30 (90)	29 (81)	0.31
University of Maryland, n (%)	10 (14)	3 (10)	7 (19)	
Pathologic stage				
pT2, n (%)	38 (55)	18 (55)	20 (55)	1.0
≥pT3 ‡, n (%)	31 (44)	15 (45)	16 (42)	
Gleason sum score				
<7 (5–6), n (%)	18 (26)	9 (27)	9 (25)	1.0
≥7 (7–9), n (%)	51 (74)	24 (73)	27 (75)	
Seminal vesicle invasion †				
No, n (%)	49 (83)	22 (73)	27 (93)	0.08
Yes, n (%)	10 (17)	8 (27)	2 (7)	
Surgical margin status †				
Negative, n (%)	35 (59)	18 (60)	17 (59)	1.0
Positive, n (%)	24 (41)	12 (40)	12 (41)	

*P value for the difference between African-Americans and European-Americans. All tests were two-sided.
†Cases with unknown status are not included.
‡One European-American patient was staged pT4.

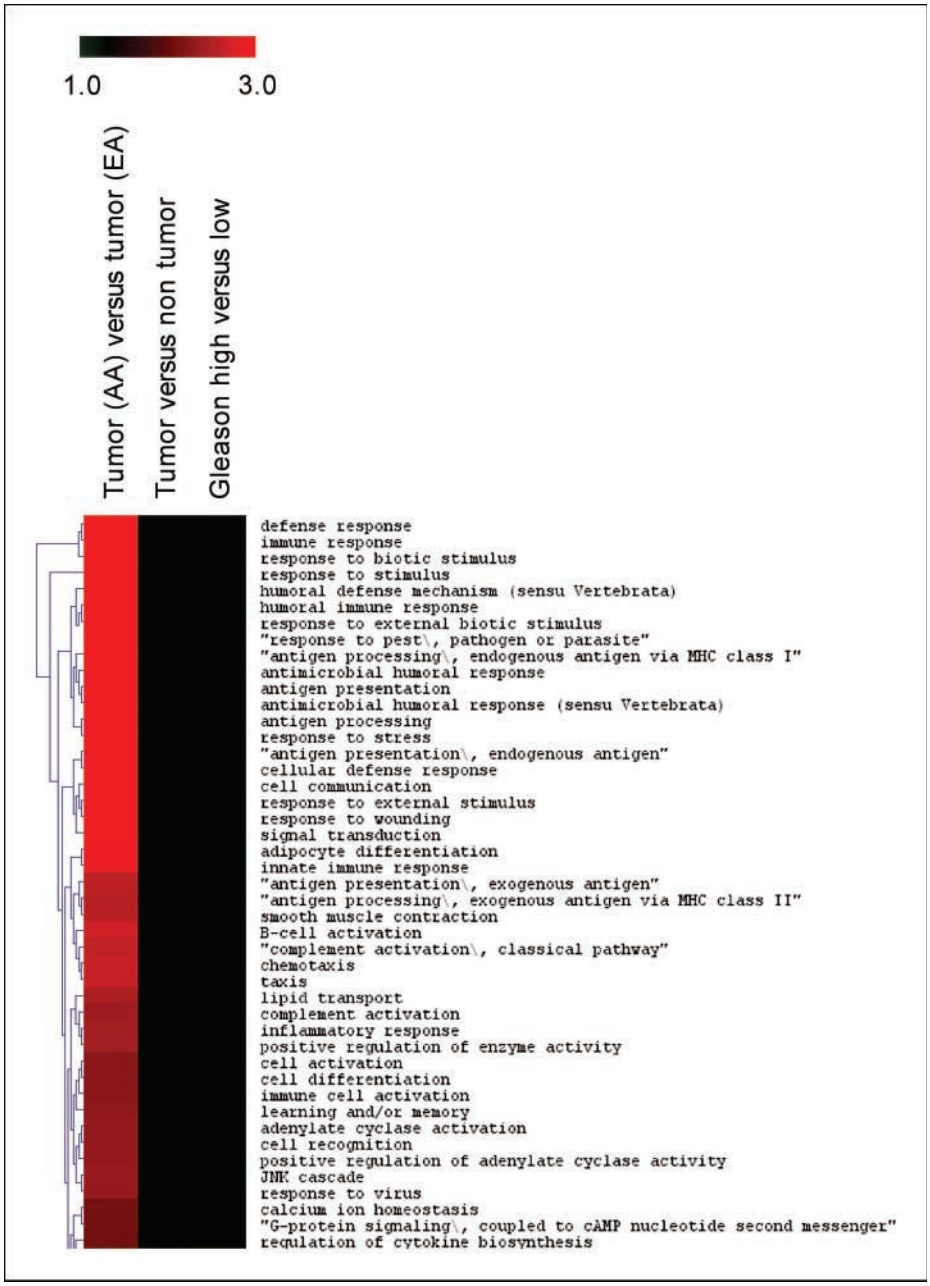

Figure 1. GOBP terms that are uniquely enriched for differently expressed genes comparing prostate tumors from African-Americans (*AA*; $n = 33$) with tumors from European-Americans (*EA*; $n = 36$). For comparison, we also show the contrast of tumor ($n = 69$) versus nontumor ($n = 20$) and the contrast of high ($n = 51$) versus low ($n = 18$) Gleason sum score. Gleason sum score was dichotomized into high (having a sum score of 7–9) and low (having a sum score of 5–6). The results of our analysis are displayed by a heat map with the red color indicating an enrichment of differently expressed genes in a GOBP term.

genetic pathways that are linked to autoimmune diseases, allergies, and inflammatory diseases (Supplementary Fig. S1). The enrichment of disease genes among the differently expressed genes was statistically significant, e.g., inflammatory urogenital disease genes (permutated $P < 0.001$), skin-prick test genes (permutated $P = 0.003$), systemic lupus erythematosus genes (permutated $P = 0.006$), and asthma (permutated $P = 0.01$). Autoimmune disease modulators, such as *PTPN22* and components of the HLA complex, and key genes in antigen presentation, such as *TAP1* and *TAP2*, were higher expressed in African-American tumors. In contrast, genes differently expressed between tumor and nontumor tissue, and between tumors with high and low Gleason sum score, showed common associations with other epithelial cancers as one may expect (Supplementary Fig. S1).

These observations were further corroborated by pathway analyses using GO biological processes and BioCarta annotations. Differently expressed genes between tumors of African-Americans and European-Americans clustered uniquely in GO biological processes (Fig. 1) and BioCarta pathways (Supplementary Fig. S2) that are associated with immune response, defense response, antigen presentation, B-cell/T-cell function, cytokine signaling, chemotaxis, and inflammatory response. The clustering of genes in these biological processes was highly significant (Table 2), and immune response genes were enriched at the $P < 1.0 \times 10^{-30}$ significance level. The enrichment of differently expressed genes in the 20 highest-ranked GOBP terms (Table 2) had a low FDR (<1% for each term). Most commonly, the involved genes were more highly expressed in tumors of African-American patients when compared with those of European-American patients, as detailed by an example for the immune response biological process (Supplementary Table S3). An analogue analysis of nontumor tissue did not generate these biological process/pathway associations, suggesting

that this gene signature is derived from the tumor microenvironment (Supplementary Fig. S3).

Differences in expression of metastasis-related genes. Next, we directed our attention to specific genes that were most significantly differently expressed in prostate tumors (fold change, ≥1.5; FDR, ≤5%) comparing African-American with European-American patients. This list included 47 genes (Table 3). The two genes that had the greatest expression differences were *PSPHL* (4.64-fold; $P = 1.0 \times 10^{-11}$), a homologue of a putative phosphoserine phosphatase (26), and *CRYBB2* (3.02-fold; $P = 1.3 \times 10^{-11}$). The list also contained a number of genes that encode chemokines and chemokine receptors. Several of those, including the genes encoding *CXCR4* and chemokine receptor 7 (*CCR7*), have been previously linked to cancer metastasis (27, 28). These and other metastasis-associated genes, e.g., *AMFR* and *MMP9*, were more highly expressed in tumors of African-American patients than those of European-Americans (Table 3). Three metastasis-associated genes were chosen for validation via qRT-PCR. Consistent with the microarray data, we found a higher expression of *AMFR* (1.5-fold), *CXCR4* (1.8-fold), and *MMP9* (2.0-fold) in primary tumors of African-Americans (Fig. 2) when compared with the tumors from European-Americans.

Further analyses were performed to determine the expression of these genes in tumor and surrounding nontumor tissue. The data showed that *CXCR4* was differently expressed between tumor and surrounding nontumor tissue in African-Americans, but not in European-Americans (Supplementary Table S4). None of the other genes followed this pattern. Instead, these genes seemed to be higher expressed in both tumor and surrounding nontumor tissue of African-American patients when compared with the European-American patients (Supplementary Table S4). Additional pathway analysis reiterated the potential significance of *CXCR4* in prostate cancer biology among the African-American patients. That analysis identified the *CXCR4* pathway as the highest-ranked pathway with the most significant enrichment of differently expressed genes comparing tumor with surrounding nontumor tissue among the African-American patients (Supplementary Table S5).

A two-gene predictor is associated with race/ethnicity. We interrogated our datasets for genes that can differentiate between African-Americans and European-Americans using the gene expression profile of prostate tumors and surrounding nontumor tissue. We identified a two-gene signature that accurately differentiated between these two groups. These two genes were *PSPHL* (probeset 205048_s_at) and *CRYBB2* (probeset 206777_s_at). Based on the expression of these two genes in prostate tumors, PAM classified correctly 30 of 33 African-Americans (91%) and 34 of 36 European-Americans (94%). We could not obtain this accuracy with the two-gene signature using the gene expression profile of the nontumor tissue. To achieve similar prediction accuracy (86% for African-Americans and 100% for European-Americans) with the gene expression profile of nontumor tissue, a classifier consisting of 19 probesets was required (data not shown). We confirmed the ability of *PSPHL* and *CRYBB2* expression levels to predict race using a blinded analysis of 39 additional tumors (14). The two-gene classifier correctly predicted 30 of 34 African-Americans

Table 2. Twenty highest-ranked GOBP terms enriched for differently expressed genes comparing tumors from African-American and European-American patients

GOBP term	Term hits*	All genes†	Annotated genes for term‡	All annotated genes§	*P* (Fisher's exact test)
Immune response	66	217	944	16,762	5.66E-31
Defense response	68	217	1,045	16,762	4.38E-30
Response to biotic stimulus	70	217	1,190	16,762	2.44e−28
Organismal physiologic process	74	217	2,111	16,762	1.46e−16
Response to stimulus	77	217	2,281	16,762	2.23e−16
Response to pest/pathogen or parasite	27	217	472	16,762	9.70e−11
Humoral immune response	16	217	162	16,762	3.38e−10
Response to external biotic stimulus	27	217	502	16,762	3.83e−10
Humoral defense mechanism	13	217	114	16,762	2.75e−09
Response to stress	34	217	929	16,762	3.76e−08
Antigen processing	8	217	45	16,762	9.90e−08
Antigen processing via MHC class I	5	217	14	16,762	6.32e−07
Antigen presentation	7	217	44	16,762	1.41e−06
Antimicrobial humoral response	9	217	85	16,762	1.52e−06
Antimicrobial humoral response	9	217	88	16,762	2.04e−06
Antigen presentation	4	217	13	16,762	1.78e−05
Cellular defense response	8	217	93	16,762	2.75e−05
Signal transduction	67	217	3,465	16,762	2.40e−04
Cell communication	77	217	4,177	16,762	3.11e−04
Apoptosis	16	217	458	16,762	3.20e−04

*Annotated genes in a GOBP term that are differently expressed (FDR, ≤30%) comparing tumors from African-American with those from European-American.

†All GOBP-annotated genes that are differently expressed in this comparison.

‡All annotated genes in a GOBP term.

§All GOBP-annotated genes.

Table 3. Most differently expressed tumor genes by race/ethnicity

Genes higher expressed in prostate tumors of African-American patients (at least 1.5-fold)

No.	Gene symbol	Genbank ID	Gene title	Fold change*
1	ADAMDEC1	NM_014479	ADAM-like, decysin 1	1.81
2	AMFR	NM_001144	Autocrine motility factor receptor	1.9
3	APOLD1	NM_030817	Apolipoprotein L domain containing 1	1.56
4	BRDG1	NM_012108	BCR downstream signaling 1	2.06
5	C1orf38	AB035482	Chromosome 1 open reading frame 38	1.58
	C1orf38	NM_004848	Chromosome 1 open reading frame 38	1.51
6	C8orf4	NM_020130	Chromosome 8 open reading frame 4	1.71
7	CCL5	M21121	Chemokine (C-C motif) ligand 5	1.61
	CCL5	NM_002985	Chemokine (C-C motif) ligand 5	1.51
8	CCR7	NM_001838	Chemokine (C-C motif) receptor 7	1.66
9	CD52	N90866	CD52 molecule	1.7
10	CRYBB2	NM_000496	crystallin, β B2	3.02
11	CXCR4	AJ224869	Chemokine (C-X-C motif) receptor 4	1.72
	CXCR4	AF348491	Chemokine (C-X-C motif) receptor 4	1.66
	CXCR4	L01639	Chemokine (C-X-C motif) receptor 4	1.65
12	DNAJC15	NM_013238	DnaJ (Hsp40) homologue, subfamily C	1.64
13	DUSP2	NM_004418	Dual specificity phosphatase 2	1.53
14	EBI2	NM_004951	EBV-induced gene 2	1.56
15	GZMA	NM_006144	Granzyme A	1.63
16	IFI44L	NM_006820	IFN-induced protein 44-like	1.63
17	IGH	M87789	Immunoglobulin heavy locus	1.78
18	IGHM	BC001872	Immunoglobulin heavy constant mu	2.37
19	IGJ	AV733266	Immunoglobulin J polypeptide	2.02
20	IGKC	L14458	Immunoglobulin κ constant	1.88
	IGKC	BG485135	Immunoglobulin κ constant	1.67
	IGKC	M63438	Immunoglobulin κ constant	1.6
	IGKC	BC005332	Immunoglobulin κ constant	1.59
21	IGKV1D-13	AW408194	Immunoglobulin κ variable 1D-13	1.73
22	IGL	M87790	Immunoglobulin λ locus	1.82
	IGL	X57812	Immunoglobulin λ locus	1.81
	IGL	AA680302	Immunoglobulin λ locus	1.74
	IGL	AV698647	Immunoglobulin λ locus	1.67
23	IL7R	NM_002185	Interleukin 7 receptor	1.61
24	INDO	M34455	Indoleamine-pyrrole 2,3 dioxygenase	1.64
25	ISG15	NM_005101	ISG15 ubiquitin-like modifier	1.6
26	ITK	D13720	IL2-inducible T-cell kinase	1.76
27	LOC283970	AI925734	Hypothetical protein LOC283970	1.53
28	LOC339562	M20812	Similar to Igκ chain V-I region Walker precursor	1.64
	LOC651629	AW404894	Similar to Igκ chain V-I region Walker precursor	1.91
	LOC652745	AJ408433	Similar to Igκ chain V-I region Walker precursor	2.12
29	LOC96610	X79782	Hypothetical protein similar to KIAA0187	1.57
30	LYZ	AV711904	Lysozyme (renal amyloidosis)	1.61
31	MMP9	NM_004994	Matrix metallopeptidase 9	1.86
32	PI15	NM_015886	Peptidase inhibitor 15	1.63
33	PSCDBP	L06633	Pleckstrin homology, Sec7	1.87
34	PSPHL	NM_003832	Phosphoserine phosphatase-like	4.64
35	PTPRC	NM_002838	Protein tyrosine phosphatase, receptor type, C	1.6
36	RASA4	AB011110	RAS p21 protein activator 4	1.5
37	SF1	D26121	Splicing factor 1	2.17
38	STAT1	BC002704	Signal transducer and activator of transcription 1	1.71
	STAT1	M97935_MB	Signal transducer and activator of transcription 1	1.51
39	TRBV19	M15564	T-cell receptor β variable 19	1.64
	TRBV19	AL559122	T-cell receptor β variable 19	1.57
40	TRBV21-1	AF043179	T-cell receptor β variable 21-1	1.58

(*Continued on the following page*)

Table 3. Most differently expressed tumor genes by race/ethnicity (Cont'd)

Genes lower expressed in prostate tumors of African-American patients (0.7-fold or less)

No.	Gene symbol	Genbank ID	Gene title	Fold change
1	ADRB1	NM_000684	Adrenergic receptor, β-1	0.63
2	ARL17P1	AF119889	ADP-ribosylation factor-like 17 pseudogene 1	0.67
3	EIF2S3	NM_001415	Eukaryotic translation initiation factor 2	0.65
4	LGALS8	AF342815	Lectin, galactoside-binding, soluble, 8	0.7
	LGALS8	AF342816	Lectin, galactoside-binding, soluble, 8	0.58
5	PLGLB2	BC005379	Plasminogen-like B2	0.66
6	SMA3	NM_006780	SMA3	0.7
	SMA3	X83301	SMA3	0.61
7	ZNF654	NM_018293	Zinc finger protein 654	0.7

NOTE: FDR ≤ 5% for all genes in the list.
*Fold change for annotated probesets with European-American patients as the reference.

(88%) and 5 of 5 patients from Italy into the European-American category (100%).

Because the microarray data indicated a greater heterogeneity of *PSPHL* expression in nontumor tissue than tumor tissue, we performed qRT-PCR validation of the *PSPHL* expression data (Fig. 3). qRT-PCR results revealed a bimodal distribution of the *PSPHL* expression in prostate tumors resulting in a 160-fold difference in gene expression between African-American and European-American patients. In the surrounding nontumor tissue, this difference was 38-fold. Furthermore, *PSPHL* was found to be up-regulated 84-fold in tumors of African-Americans when compared with the surrounding nontumor tissues from this population. These fold change values should be considered estimates. Because of the low expression of *PSPHL* in prostate tissue from European-Americans, accurate determination of *PSPHL* baseline expression was difficult. Nevertheless, the qRT-PCR results unequivocally validated the microarray data by showing that *PSPHL* is significantly higher expressed in tumors of African-Americans than in those of European-Americans.

Discussion

The causes of the prostate cancer health disparity that exists between African-American and European-American men remain to be fully elucidated. Much of the research on this disparity has focused on socioeconomic factors. To date, very few studies have examined differences in tumor biology despite suggestive clinical evidence that they exist.

We analyzed the gene expression profiles of 69 tumors from 33 African-American and 36 European-American patients. Highly significant differences in the expression of genes related to immunobiology within the tumor microenvironment were identified between these two groups. This interpretation was supported by both disease association and pathway analyses. Most of the immune-related genes were higher expressed in tumors from African-American patients than in those from European-American patients. Although preliminary, these findings are novel and could have implications for cancer therapy. Clinical trials of immunotherapy for prostate cancer have been conducted (29), and these therapies may enter clinical practice in the near future. Our data suggest that African-American patients and European-American patients might respond differently to these types of therapy.

We observed that previously recognized diagnostic and prognostic marker genes for prostate tumors (23–25) were not significantly differently expressed by race/ethnicity. The findings suggest that the most common alterations in gene expression in prostate carcinogenesis do not develop any differently in African-Americans than in European-Americans. This result corroborates published data that showed a similar pattern of chromosomal alterations in prostate tumors of African-American and European-American men (8).

Several metastasis-associated genes were found to be more highly expressed in tumors of African-American patients when compared with European-American patients. This group of genes contained *AMFR, CXCR4, CCR7*, and *MMP9*, among others. To the best of our knowledge, none of those genes have been previously recognized as a prostate cancer marker gene by examination of gene expression profiles. Unfortunately, most previous gene expression profiling studies do not disclose the race/ethnic make up of the patients when describing the sample sets used (30–33). It is likely that the majority of the studied tumors were obtained from European-American patients. Thus, genes that are dysregulated in prostate tumors of African-American patients may have been overlooked until now. However, multiple studies have shown that both *CXCR4* and *MMP9* are key metastasis regulators in prostate cancer (27, 34–37), and both genes are linked through a common pathway (38). The causal role of *CXRC4* and also *CCR7* in cancer metastasis is well demonstrated (39–41), and therapeutic approaches to target these receptors are being developed (42).

From the present study, we do not know why prostate tumors from African-American patients would have a different immunologic profile than tumors from European-American patients. We hypothesize that the causes of these differences are multifactorial. The nature of the differently expressed transcripts indicates the involvement of various cell types in the generation of the gene signature, e.g., B cells, T cells, and tumor-infiltrating macrophages.

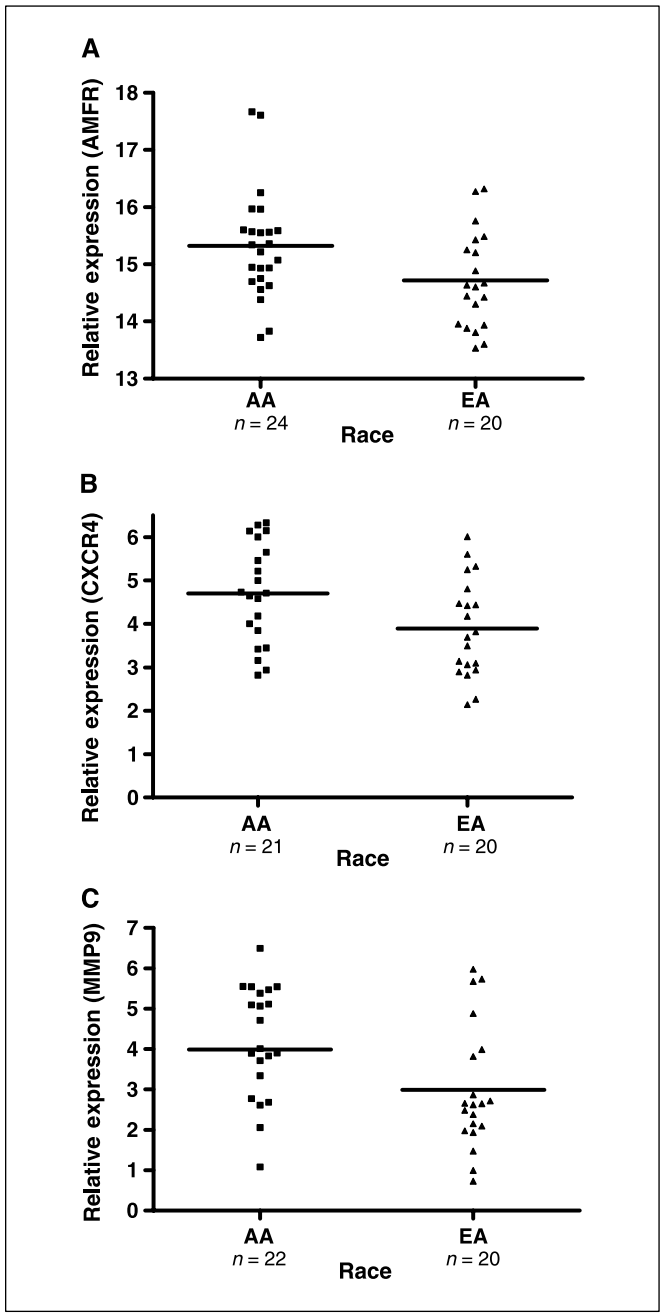

Figure 2. qRT-PCR expression analysis of *AMFR* (*A*), *CXCR4* (*B*), and *MMP9* (*C*) in prostate tumors of African-Americans and European-Americans. *Points,* relative C_t values for the individual samples; *bars,* median value for the sample set. The genes are significantly higher expressed in African-American patients than in European-American patients; fold change = 1.52, $P = 0.033$ (two-sided *t* test) for *AMFR*; fold change = 1.75, $P = 0.027$ (two-sided *t* test) for *CXCR4*; fold change = 2.00, $P = 0.044$ (two-sided Wilcoxon rank-sum test) for *MMP9*. Graphs were prepared using relative C_t values that were calculated by subtracting the 18s C_t values from the corresponding C_t values for the gene being measured (see Subjects and Methods). Fold changes were derived from the $2^{-\Delta\Delta Ct}$ values for each gene.

An association of those cells with inflammation in prostate carcinogenesis has been recognized (43). To investigate the contribution from B-cell and T-cell infiltration, we evaluated B-cell and T-cell density of 48 prostate tumors (23 African-American and 25 European-American) in a pilot study. No significant difference in tumor B-cell and T-cell infiltration was found between the two race/ethnic groups (data not shown). These preliminary data argue that the infiltrating B-cell and T-cell populations could be functionally different in the African-American and European-American tumors. Possible causes contributing to the different immunologic profiles could be environmental factors, e.g., infections, genetic factors, e.g., genetic variations in immune regulatory genes, or the interactions of both. There is evidence that infectious diseases are associated with prostate carcinogenesis (43, 44), and two genes involved in innate immunity, *MSR1* and *RNASEL*, are candidate susceptibility genes for prostate cancer (45, 46). Notably, the gene signature that was generated in this study comparing prostate tumors from African-American and European-American men revealed many up-regulated IFN-related genes in the African-American prostate tumors. This distinctive signature could be indicative of a possible viral involvement in this African-American population.

Chronic inflammation is also thought to be a causative factor in prostate carcinogenesis (43, 47). Studies have suggested that inflammation is more prevalent in nontumor prostate biopsy specimens from African-American men when compared with European-American men (48). Other investigations have shown that an immune response signature in the liver of cancer patients predicts metastasis and recurrence of hepatocellular carcinoma (49). Thus, future studies should evaluate whether the immunologic profile of prostate tumors in African-American men is a predisposing factor for tumor progression and metastasis.

Our investigations identified a two-gene tumor signature that accurately differentiated between African-American and European-American patients. Little is known about the functions of the two genes, *PSPHL* and *CRYBB2*, in this classifier. *PSPHL* was first isolated from Fanconi anemia fibroblasts in a genetic screen and was identified as a homologue of a putative phosphoserine

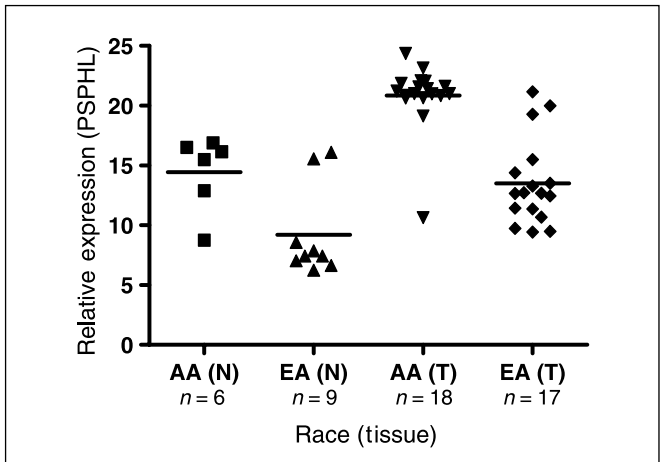

Figure 3. qRT-PCR expression analysis of *PSPHL* in prostate tumors (*T*) and surrounding nontumor tissue (*N*) of African-Americans and European-Americans. *Points,* relative C_t-values for the individual samples; *bars,* median value for the sample set. *PSPHL* is significantly higher expressed in surrounding nontumor tissue (fold change = 38, $P = 0.012$; two-sided Wilcoxon rank-sum test) and tumor tissue (fold change = 160, $P < 0.001$; two-sided Wilcoxon rank-sum test) of African-American patients when compared with European-American patients. Graphs were prepared using relative C_t values that were calculated by subtracting the 18s C_t values from the corresponding C_t values for the gene being measured (see Subjects and Methods). Fold changes were derived from the $2^{-\Delta\Delta Ct}$ values for each gene.

phosphatase (26). Interestingly, *PSPHL* is located on chromosome 7q11.2, a chromosomal region known to have gain of function related to advanced tumor stage in prostate cancer (50). Despite its chromosomal location, there are no studies linking *PSPHL* expression to cancer progression. Our data also showed that *PSPHL* is higher expressed in prostate tumors than the surrounding nontumor tissue. Thus, the possible role of *PSPHL* in contributing to prostate cancer susceptibility will have to be tested in future studies.

Whole-genome gene expression analysis experiments are prone to findings that are either unique to a selected patient population or are artificially created by the applied technology. To exclude the possibility of an artifact, three approaches were used to cross-validate our gene expression data. First, we compared our results of the differently expressed genes between tumor and surrounding nontumor tissue with those of other studies and found very good agreement (23–25, 33). Second, we validated the expression of several key genes via qRT-PCR and confirmed that these genes are higher expressed in tumors from African-Americans than from European-Americans. Lastly, we used prediction analysis on a blinded dataset and corroborated that a two-gene signature can distinguish between African-American patients and European-American patients. Despite all of the cross-validation approaches used to strengthen our results, we recognize that the findings of the current study are preliminary and will have to be further investigated, including the immunohistochemical analysis of larger sample sets.

In conclusion, the gene expression profile of prostate tumors corresponds to differences in tumor biology between African-American and European-American men. The implications of those differences in disease aggressiveness and response to therapy should be evaluated in future studies.

Acknowledgments

Received 7/10/2007; revised 10/2/2007; accepted 12/4/2007.

Grant support: Intramural Research Program of NIH National Cancer Institute Center for Cancer Research.

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 the National Cancer Institute Cooperative Prostate Cancer Tissue Resource and James Goedert (Viral Epidemiology Branch, National Cancer Institute) for providing tissue specimens and supporting data and Audrey Salabes, John Cottrel, Raymond Jones, and other personnel at University of Maryland and Baltimore Veterans Administration for their contributions.

References

1. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. CA Cancer J Clin 2007;57:43–66.

2. Hsing AW, Tsao L, Devesa SS. International trends and patterns of prostate cancer incidence and mortality. Int J Cancer 2000;85:60–7.

3. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2006. CA Cancer J Clin 2006;56:106–30.

4. Powell IJ. Epidemiology and pathophysiology of prostate cancer in African-American men. J Urol 2007;177:444–9.

5. Reddy S, Shapiro M, Morton R, Jr., Brawley OW. Prostate cancer in Black and White Americans. Cancer Metastasis Rev 2003;22:83–6.

6. Hoffman RM, Gilliland FD, Eley JW, et al. Racial and ethnic differences in advanced-stage prostate cancer: the Prostate Cancer Outcomes Study. J Natl Cancer Inst 2001;93:388–95.

7. Thompson I, Tangen C, Tolcher A, Crawford E, Eisenberger M, Moinpour C. Association of African-American ethnic background with survival in men with metastatic prostate cancer. J Natl Cancer Inst 2001;93:219–25.

8. Cher ML, Lewis PE, Banerjee M, et al. A similar pattern of chromosomal alterations in prostate cancers from African-Americans and Caucasian Americans. Clin Cancer Res 1998;4:1273–8.

9. Shibata A, Whittemore AS. Genetic predisposition to prostate cancer: possible explanations for ethnic differences in risk. Prostate 1997;32:65–72.

10. Shuch B, Mikhail M, Satagopan J, et al. Racial disparity of epidermal growth factor receptor expression in prostate cancer. J Clin Oncol 2004;22:4673–7.

11. Guo Y, Sigman DB, Borkowski A, Kyprianou N. Racial differences in prostate cancer growth: apoptosis and cell proliferation in Caucasian and African-American patients. Prostate 2000;42:130–6.

12. Haiman CA, Patterson N, Freedman ML, et al. Multiple regions within 8q24 independently affect risk for prostate cancer. Nat Genet 2007;39:638–41.

13. Yeager M, Orr N, Hayes RB, et al. Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. Nat Genet 2007;39:645–9.

14. Strickler HD, Burk R, Shah K, et al. A multifaceted study of human papillomavirus and prostate carcinoma. Cancer 1998;82:1118–25.

15. Gentleman RC, Carey VJ, Bates DM, et al. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol 2004;5:R80.

16. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98:5116–21.

17. Storey JD, Tibshirani R. Statistical significance for genomewide studies. Proc Natl Acad Sci U S A 2003;100:9440–5.

18. 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.

19. Yi M, Horton JD, Cohen JC, Hobbs HH, Stephens RM. WholePathwayScope: a comprehensive pathway-based analysis tool for high-throughput data. BMC Bioinformatics 2006;7:30.

20. Becker KG, Barnes KC, Bright TJ, Wang SA. The genetic association database. Nat Genet 2004;36:431–2.

21. Bookout AL, Mangelsdorf DJ. Quantitative real-time PCR protocol for analysis of nuclear receptor signaling pathways. Nucl Recept Signal 2003;1:e012.

22. Pettaway CA, Troncoso P, Ramirez EI, Johnston DA, Steelhammer L, Babaian RJ. Prostate specific antigen and pathological features of prostate cancer in Black and White patients: a comparative study based on radical prostatectomy specimens. J Urol 1998;160:437–42.

23. Rhodes DR, Barrette TR, Rubin MA, Ghosh D, Chinnaiyan AM. Meta-analysis of microarrays: interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. Cancer Res 2002;62:4427–33.

24. Singh D, Febbo PG, Ross K, et al. Gene expression correlates of clinical prostate cancer behavior. Cancer Cell 2002;1:203–9.

25. 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.

26. Planitzer SA, Machl AW, Rueckels M, Kubbies M. Identification of a novel c-DNA overexpressed in Fanconi's anemia fibroblasts partially homologous to a putative L-3-phosphoserine-phosphatase. Gene 1998;210:297–306.

27. Taichman RS, Cooper C, Keller ET, Pienta KJ, Taichman NS, McCauley LK. Use of the stromal cell-derived factor-1/CXCR4 pathway in prostate cancer metastasis to bone. Cancer Res 2002;62:1832–7.

28. Muller A, Homey B, Soto H, et al. Involvement of chemokine receptors in breast cancer metastasis. Nature 2001;410:50–6.

29. Small EJ, Schellhammer PF, Higano CS, et al. Placebo-controlled phase III trial of immunologic therapy with sipuleucel-T (APC8015) in patients with metastatic, asymptomatic hormone refractory prostate cancer. J Clin Oncol 2006;24:3089–94.

30. Luo J, Duggan DJ, Chen Y, et al. Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. Cancer Res 2001;61:4683–8.

31. Magee JA, Araki T, Patil S, et al. Expression profiling reveals hepsin overexpression in prostate cancer. Cancer Res 2001;61:5692–6.

32. Welsh JB, Sapinoso LM, Su AI, et al. Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. Cancer Res 2001;61:5974–8.

33. Dhanasekaran SM, Barrette TR, Ghosh D, et al. Delineation of prognostic biomarkers in prostate cancer. Nature 2001;412:822–6.

34. Sehgal G, Hua J, Bernhard EJ, Sehgal I, Thompson TC, Muschel RJ. Requirement for matrix metalloproteinase-9 (gelatinase B) expression in metastasis by murine prostate carcinoma. Am J Pathol 1998;152:591–6.

35. Dong Z, Bonfil RD, Chinni S, et al. Matrix metalloproteinase activity and osteoclasts in experimental prostate cancer bone metastasis tissue. Am J Pathol 2005;166:1173–86.

36. Sun YX, Wang J, Shelburne CE, et al. Expression of CXCR4 and CXCL12 (SDF-1) in human prostate cancers (PCa) *in vivo*. J Cell Biochem 2003;89:462–73.

37. Sun YX, Fang M, Wang J, Cooper CR, Pienta KJ, Taichman RS. Expression and activation of $\alpha(v)\beta(3)$ integrins by SDF-1/CXC12 increases the aggressiveness of prostate cancer cells. Prostate 2007;67:61–73.

38. Chinni SR, Sivalogan S, Dong Z, et al. CXCL12/CXCR4 signaling activates Akt-1 and MMP-9 expression in prostate cancer cells: the role of bone microenvironment-associated CXCL12. Prostate 2006;66:32–48.

39. Orimo A, Gupta PB, Sgroi DC, et al. Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis

through elevated SDF-1/CXCL12 secretion. Cell 2005; 121:335–48.
40. Kang Y, Siegel PM, Shu W, et al. A multigenic program mediating breast cancer metastasis to bone. Cancer Cell 2003;3:537–49.
41. Wiley HE, Gonzalez EB, Maki W, Wu MT, Hwang ST. Expression of CC chemokine receptor-7 and regional lymph node metastasis of B16 murine melanoma. J Natl Cancer Inst 2001;93:1638–43.
42. Vaday GG, Hua SB, Peehl DM, et al. CXCR4 and CXCL12 (SDF-1) in prostate cancer: inhibitory effects of human single chain Fv antibodies. Clin Cancer Res 2004; 10:5630–9.
43. De Marzo AM, Platz EA, Sutcliffe S, et al. Inflammation in prostate carcinogenesis. Nat Rev Cancer 2007;7: 256–69.
44. Hayes RB, Pottern LM, Strickler H, et al. Sexual behaviour, STDs and risks for prostate cancer. Br J Cancer 2000;82:718–25.
45. Xu J, Zheng SL, Komiya A, et al. Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk. Nat Genet 2002;32:321–5.
46. Carpten J, Nupponen N, Isaacs S, et al. Germline mutations in the ribonuclease L gene in families showing linkage with HPC1. Nat Genet 2002;30:181–4.
47. De Marzo AM, Marchi VL, Epstein JI, Nelson WG. Proliferative inflammatory atrophy of the prostate: implications for prostatic carcinogenesis. Am J Pathol 1999;155:1985–92.
48. Eastham JA, May RA, Whatley T, Crow A, Venable DD, Sartor O. Clinical characteristics and biopsy specimen features in African-American and White men without prostate cancer. J Natl Cancer Inst 1998;90:756–60.
49. de Visser KE, Eichten A, Coussens LM. Paradoxical roles of the immune system during cancer development. Nat Rev Cancer 2006;6:24–37.
50. Alers JC, Rochat J, Krijtenburg PJ, et al. Identification of genetic markers for prostatic cancer progression. Lab Invest 2000;80:931–42.