Integrative Analysis of Genomic Aberrations Associated with Prostate Cancer Progression

Jung H. Kim,[1,5] Saravana M. Dhanasekaran,[1] Rohit Mehra,[1,6] Scott A. Tomlins,[1] Wenjuan Gu,[3] Jianjun Yu,[1,5] Chandan Kumar-Sinha,[7] Xuhong Cao,[1] Atreya Dash,[8] Lei Wang,[1] Debashis Ghosh,[3] Kerby Shedden,[4] James E. Montie,[2,6] Mark A. Rubin,[9,10] Kenneth J. Pienta,[6] Rajal B. Shah,[1,2,6] and Arul M. Chinnaiyan[1,2,5,6]

[1]Michigan Center for Translational Pathology, Department of Pathology, Departments of [2]Urology, [3]Biostatistics, and [4]Statistics, [5]Program of Bioinformatics, and [6]Comprehensive Cancer Center, University of Michigan Medical School, Ann Arbor, Michigan; [7]Advanced Center for Treatment, Research & Education in Cancer, Tata Memorial Center, Navi Mumbai, India; [8]Memorial Sloan-Kettering Cancer Center, New York, New York; and [9]Department of Pathology, Brigham and Women's Hospital, and [10]Dana-Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts

Abstract

Integrative analysis of genomic aberrations in the context of trancriptomic alterations will lead to a more comprehensive perspective on prostate cancer progression. Genome-wide copy number changes were monitored using array comparative genomic hybridization of laser-capture microdissected prostate cancer samples spanning stages of prostate cancer progression, including precursor lesions, clinically localized disease, and metastatic disease. A total of 62 specific cell populations from 38 patients were profiled. Minimal common regions (MCR) of alterations were defined for each sample type, and metastatic samples displayed the most number of alterations. Clinically localized prostate cancer samples with high Gleason grade resembled metastatic samples with respect to the size of altered regions and number of affected genes. A total of 9 out of 13 MCRs in the putative precursor lesion, high-grade prostatic intraepithelial neoplasia (PIN), showed an overlap with prostate cancer cases (amplifications in 3q29, 5q31.3-q32, 6q27, and 8q24.3 and deletions in 6q22.31, 16p12.2, 17q21.2, and 17q21.31), whereas postatrophic hyperplasia (PAH) did not exhibit this overlap. Interestingly, prostate cancers that do not overexpress ETS family members (i.e., gene fusion–negative prostate cancers) harbor differential aberrations in 1q23, 6q16, 6q21, 10q23, and 10q24. Integrative analysis with matched mRNA profiles identified genetic alterations in several proposed candidate genes implicated in prostate cancer progression. [Cancer Res 2007;67(17):8229–39]

Introduction

Chromosomal aberrations due to genome instability are a characteristic of human solid tumors (1) and are considered the primary drivers in the development and progression of cancer (2). Precise measurements of gene copy number alterations with high resolution are now possible with array comparative genomic hybridization (aCGH) done on BAC (bacterial artificial chromosome) arrays, cDNA microarrays, or oligoCGH arrays (3). Several tumors including breast, prostate, and lung cancers among others have been analyzed using aCGH technology (4, 5). More recently, studies have been documenting genome-wide copy number changes with parallel mRNA expression profiling for various cancers using microarray platforms (6–10).

aCGH analyses of human prostate cancer cell lines (11–13), xenografts (14, 15), and prostate cancer tissues (16, 17) have been reported. Although all of the above studies have used grossly dissected tissues, profiling of laser-captured, microdissected prostate cancer specimens has been shown to resolve cancer-specific genomic aberrations with higher sensitivity (17), Recently, Hughes et al. (18) reported aCGH profiling of a small set of laser-captured prostate cancer specimens. In the present study, we carried out a comprehensive characterization of cytogenetic profiles of 62 prostate cell populations using aCGH on a cDNA microarray platform as described by Pollack et al. (19). Cells from specific prostate tissue foci were isolated by laser-capture microdissection (LCM), and the samples belonged to various groups that include benign prostatic hyperplasia (BPH); stromal, atrophic epithelia; proliferative inflammatory atrophy (PIA; ref. 20); postatrophic hyperplasia (PAH; ref. 21); prostatic intraepithelial neoplasia (PIN); clinically localized prostate cancer (PCA) low-grade (L-PCA) (Gleason pattern 3); foamy, F-PCA; high-grade, H-PCA (Gleason pattern 4); and metastatic prostate cancer (MET). Our group has recently noted the enrichment of various molecular concepts in gene expression signature of prostate cancer progression using this sample set (22). In the present study, we defined the minimal common regions (MCR) corresponding to various sample groups and identified novel regions of aberrations and candidate genes that lie within. Our study also provides information on the occurrence of these MCRs in successive stages of cancer progression. Furthermore, we integrated aCGH and corresponding gene expression data (22) obtained from matched samples to evaluate genomic aberrations accompanying gene expression changes through prostate cancer progression. Notably, molecular concept map (MCM) analysis (22) of the cancer specimens identified various chromosomal regions including 6q21 that distinguish ETS-overexpressing samples from others.

Materials and Methods

Tissue specimen and genomic DNA isolation. Tissues were obtained from the radical prostatectomy series at the University of Michigan and from the Rapid Autopsy Program, which are both part of the University of Michigan Prostate Cancer Specialized Program of Research Excellence

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

J.H. Kim, S.M. Dhanasekaran, and R. Mehra contributed equally to this report.

Requests for reprints: Arul M. Chinnaiyan, Departments of Pathology and Urology, University of Michigan Medical School, 1400 E. Medical Center Drive, 5410 CCGC University of Michigan, Ann Arbor, MI 48109-0940. Phone: 734-615-4062; Fax: 734-615-4055; E-mail: arul@umich.edu.

doi:10.1158/0008-5472.CAN-07-1297

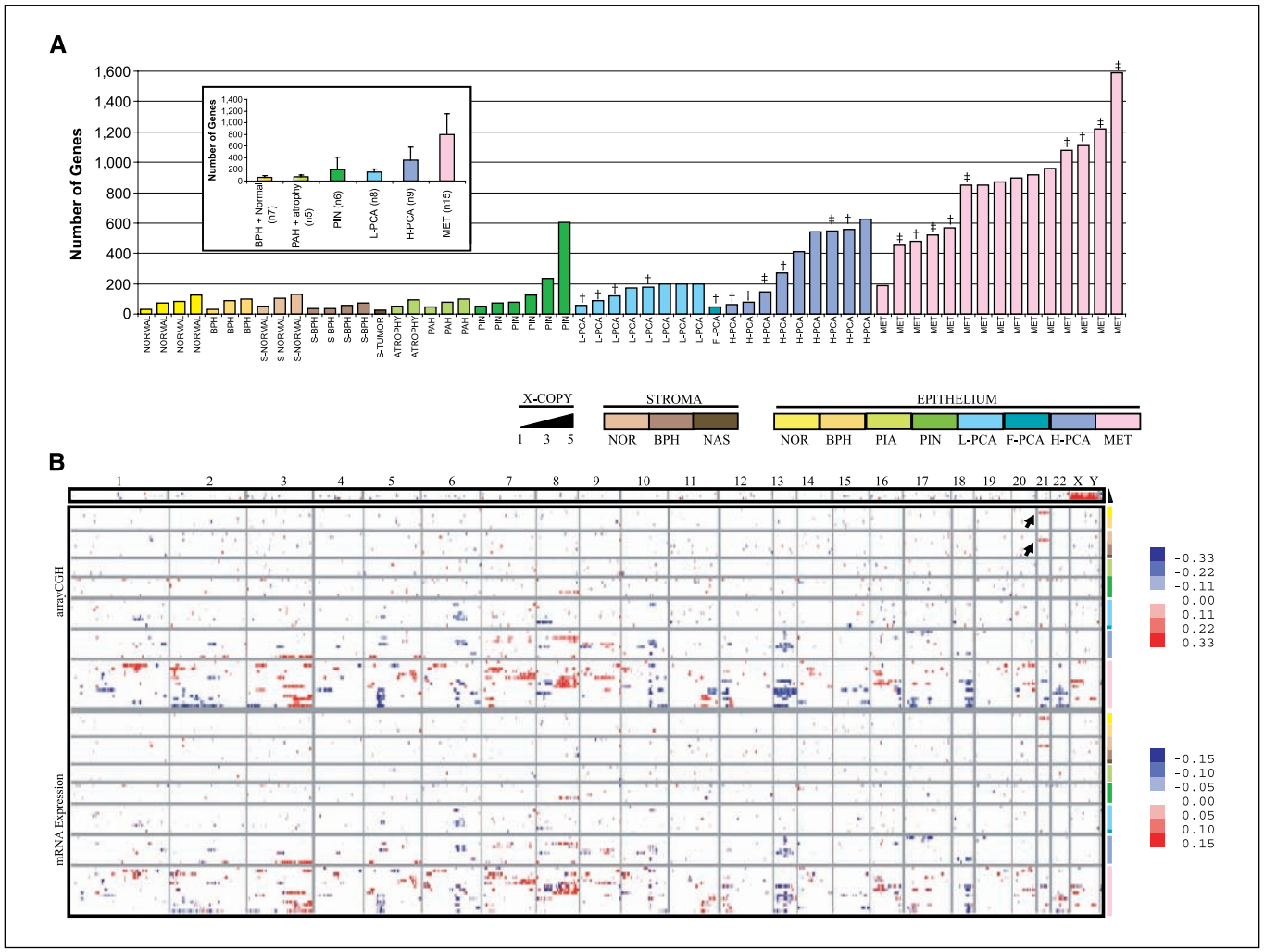

Figure 1. Genome-wide view of chromosomal alterations in prostate cancer progression. *A,* the number of altered genes from significantly altered regions (*q* value of <0.01) in each sample is represented in a bar graph. *Top left inset,* averages of the data presented. *B, top,* profiles are depicted for cell lines containing different numbers of X chromosomes. Genomic DNA isolated from laser-captured cells from various prostate tissue sections were profiled for DNA copy number changes. *Row,* tumor, benign prostate, or cell line; *column,* one of 9,550 unique genes, ordered by genome map position from chromosome 1 to Y. *Red,* fold amplification; *blue,* fold deletion; *white,* no change. *Bottom,* mRNA expression of matched samples within regions of significant genomic alteration. *NOR,* normal prostate from organ donors and patient; *BPH,* benign prostatic hyperplasia; *S,* adjacent stroma; *Atrophy,* atrophic epithelium; *L-PCA,* low-grade localized prostate cancer (Gleason pattern 3); *F-PCA,* foamy localized prostate cancer; *H-PCA,* high-grade localized prostate cancer (Gleason pattern 4); *MET,* metastatic prostate cancer. *Arrows,* single copy gain in chromosome 21 in a Down's syndrome patient. †, *ERG*-overexpressing samples. ‡, *ETV1*-overexpressing samples.

Tissue Core. All samples were collected with informed consent of the patients and prior Institutional Review Board approval. The prostate cancer samples as shown in the MIAME checklist (Supplementary Methods) obtained from a total of 38 patients/organ donors include 7 normal/BPH; 8 stromal, S; 5 PAH (2 atrophic epithelium, ATR; 3 PIA), 7 PIN; 18 localized prostate cancer (8 L-PCA, 1 F-PCA, and 9 H-PCA); and 17 metastatic prostate cancer (MET; Supplementary Table S1). A precision cut using LCM was done on frozen tissue sections (6 μm) containing a minimum of 10,000 cells placed on specially manufactured membrane slides (MMI) with the SL Microtest device (MMI) using μCUT software (MMI; ref. 22). Genomic DNA was isolated from the cells using QIAamp DNA Mini Kit (Qiagen), and DNA concentration was determined using Quant-iT DNA Assay Kit, High Sensitivity (Invitrogen). For the threshold analysis, human genomic DNA samples having varying copies (1 to 5) of the X chromosome were purchased from the National Institute of General Medical Sciences Human Genetic Mutant cell repository.[12] Normal human male and female genomic DNA was purchased from Promega Inc.

[12] http://www.nigms.nih.gov/Initiatives/HGCR/

Array-CGH on cDNA microarrays. In-house cDNA microarrays containing 20,000 spotted elements representing ~13,000 different UniGene clusters used in our previous gene expression profiling studies (22, 23) were used for the aCGH studies. One hundred nanograms of genomic DNA were amplified using OmniPlex Whole Genome Amplification (WGA) kit (Sigma-Aldrich) following the manufacturer's protocol. Amplified normal human male genomic DNA was used as reference for all the hybridizations. The amplified DNA was quantified by Quant-iT DNA Assay Kit, High Sensitivity (Invitrogen), and 4 μg of DNA from each sample was labeled using BioPrime Array CGH Genomic Labeling System (Invitrogen). Two color hybridizations were done as described earlier by Pollack et al. (19). The use of WGA was previously evaluated on frozen and formalin-fixed, paraffin-embedded Wilm's tumor specimens on aCGH platform by Little et al. (24).

Data collection and analysis. cDNA microarray slides were scanned using an Axon GenePix 4000B dual-laser scanner, and its fluorescence signal was quantified with GenePix Pro 6.0 software (Axon Instruments). Bad spots were flagged out, and data were Lowess normalized (25). Genes with multiple representations were averaged using GEPAS software (26), and $\log_2$ transformed. A total of 9,550 unique genes were analyzed as

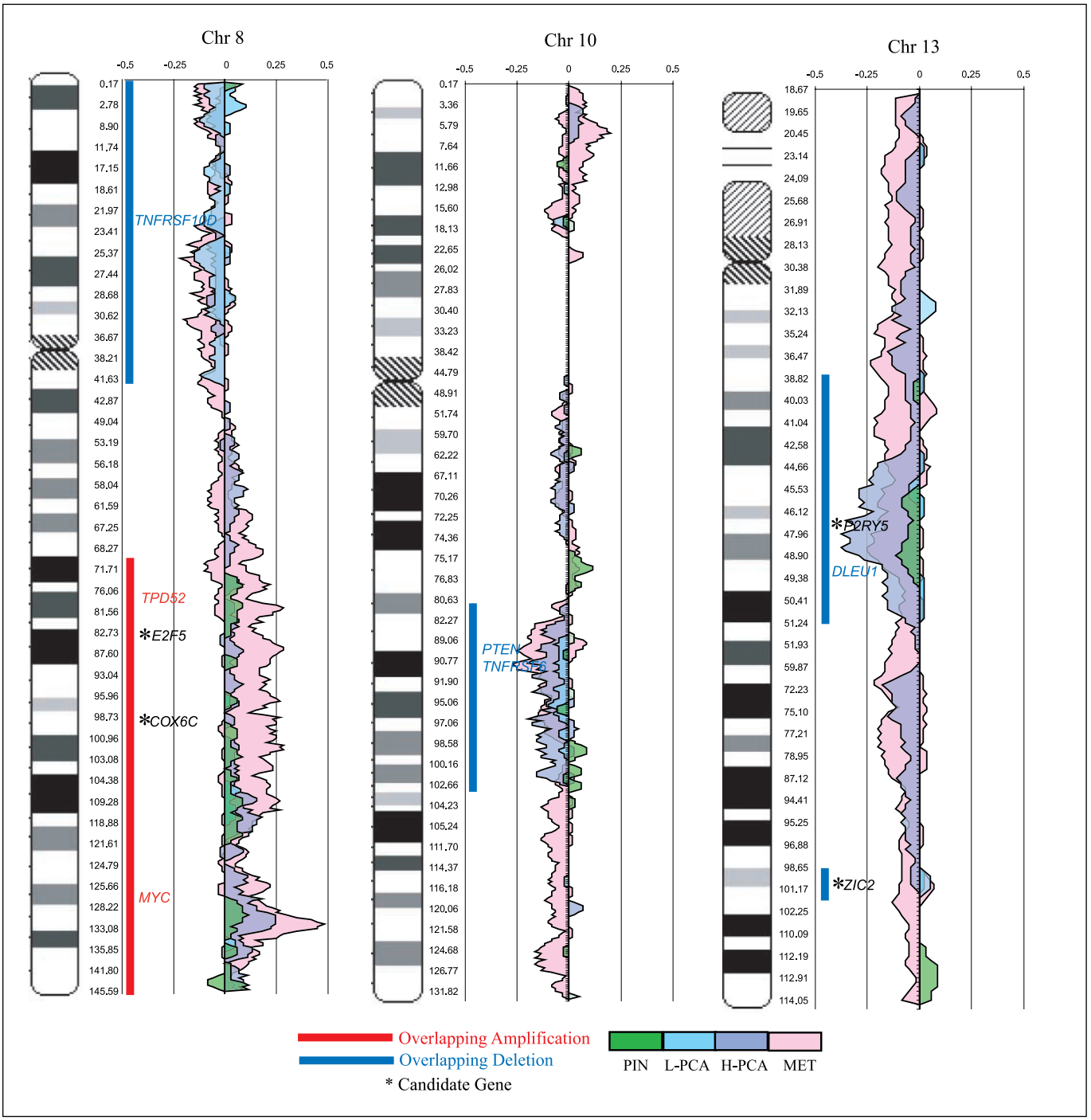

Figure 2. Representative chromosomal alterations in prostate cancer. The chromosomal aberrations observed in chromosomes 8, 10, and 13 of PIN, L-PCA, H-PCA, and MET samples are depicted, and data for all chromosomes are provided in the Supplementary Fig. S1. *Peaks moving to the right,* amplification; *peaks moving to the left,* deletion. The amplified genes, *MYC* and *TPD52*, and deleted tumor-suppressor *PTEN* are located within these frequently amplified and deleted regions, respectively. Regions harboring genes, including *E2F5, COX6C, P2RY5, MYC*, and *ZIC2*, are altered in specimens from all stages of prostate cancer.

follows. Cutoff values were set at $\log_2$ ratio $\geq$0.22 for amplification and $\leq$−0.22 for deletion (97% and 3% quantiles, respectively); the cutoff values for complementary expression profiling data were set at $\log_2$ ratio of $\geq$0.4 for overexpression of genes and $\leq$−0.4 for underexpression ($\pm$4 SDs of the middle 50% quantile of data; refs. 9, 10, 27). For the genome-wide integrative analyses, the data from aCGH and gene expression microarrays were moving averaged (symmetrical five nearest neighbors) using CGH-Miner software (28). The CGH-Miner output for 9,550 unique genes was ordered according to the genome map positions from chromosome 1 to Y, and the moving averaged (symmetrical five nearest neighbors) fluorescence ratios were depicted using $\log_2$-based pseudocolor scale.

Percentage of alterations in prostate cancer. Prostate cancer data represented in global view analysis were classified into four different sample groups: PIN, L-PCA (Gleason pattern 3), H-PCA (Gleason pattern 4), and MET. From the region of chromosomal aberrations, the amplified and deleted genes were selected using the thresholds set above. Among the

Table 1. The most frequently observed chromosomal alteration sites in prostate cancer progression

Chromosomes	MET cytogenetic band	MCR recurrence (n = 12)	PCA cytogenetic band	MCR recurrence (n = 17)
Gain				
2	2p25.1-p24.3	4	2p25.1*	1
	2p24.3	4		
	2p23.3 †	5		
3	3q13.33	4	3q13.33 ‡	1
	3q21.2	3	3q21.2*	3
	3q26.32-q26.33	4	3q26.32	1
	3q26.33	4	3q26.33	1
5			5q32 †	5
	5q35.1	4	5q35.1 ‡	1
7	7p21.3	4	7p21.3-p21.2 ‡	1
	7p15.3	4	7p15.3 ‡	1
	7p15.2	4	7p15.2*	1
	7p15.2-p15.1 †	4		
	7p15.1 †	5	7p15.1 †	2
	7p14.2-p14.1	4	7p14.2*	1
	7p14.1	4	7p14.1 ‡	2
	7p13	4	7p13 ‡	1
	7q34	4	7q34	3
	7q36.1	4	7q36.1	2
8	8q21.11 †	4	8q21.11 † ‡	1
	8q21.13 †	5	8q21.13 † ‡	2
	8q21.3 †	4	8q21.3 † ‡	1
	8q22.1 †	4	8q22.1 †	2
	8q22.1-q22.2 †	4		
	8q22.2 †	5	8q22.2 † ‡	1
	8q22.2-q22.3 †	6		
	8q22.3 †	5	8q22.3 † ‡	1
	8q23.1 †	5	8q23.1 †	3
	8q23.1-q23.3 †	4	8q23.2-q23.3 †	2
	8q23.3-q24.11 †	5	8q24.11 †	4
	8q24.11 †	4	8q24.11-q24.12 †	3
	8q24.12 †	3	8q24.12 †	3
	8q24.13 †	4	8q24.13 † ‡	2
	8q24.21 †	4	8q24.21 † ‡	3
	8q24.22 †	7	8q24.21-q24.22 † ‡	3
	8q24.22-q24.23 †	4	8q24.22 † ‡	3
	8q24.3 †	4	8q24.23-q24.3 †	3
9	9q33.3 †	4	9q33.3 †	2
16	16p12.3 †	4		
20	20q13.33	4		
Loss				
1	1q23.1	1	1q23.1	3
3	3q26.33	1	3q26.33	3
5	5q13.3	2	5q13.3	4
6	6q14.1	3	6q14.1	3
	6q14.2	3	6q14.2	3
	6q14.3	3	6q14.3	4
	6q15-q16.1	4	6q15	4
	6q16.1-q16.2	3	6q16.1-q16.3	3
	6q21	2	6q21	3
	6q22.31 †	2	6q22.31 †	3
8	8p21.2	4	8p21.2	2
	8p21.1	4	8p21.1	2
	8p12	4	8p12	2
10	10q23.2-q23.31 †	4	10q23.2-q23.31 †	4
	10q23.31 †	5	10q23.31 †	4
	10q23.31-q23.32 †	3	10q23.31-q23.32 †	3

(*Continued on the following page*)

Table 1. The most frequently observed chromosomal alteration sites in prostate cancer progression (Cont'd)

Chromosomes	MET cytogenetic band	MCR recurrence (n = 12)	PCA cytogenetic band	MCR recurrence (n = 17)
	10q23.33 †	2	10q23.33 †	3
			10q23.33-q24.1	3
	10q24.1 †	1	10q24.1	4
13	13q13.3-q14.11 †	4	13q14.11-q14.12 †	3
	13q14.12 †	3	13q14.12 †	5
	13q14.13-q14.2 †	3	13q14.13-q14.2 † ‡	3
	13q14.2 †	4	13q14.2 † ‡	5
	13q14.2-q14.3 †	4	13q14.3 † ‡	3
16	16p12.2 †	2	16p12.2 †	4
	16p12.1	1	16p12.1	3
17	17q21.31 †	2	17q21.31 †	4
18	18q21.2	4	18q21.2 ‡	1
	18q21.2-q21.31	4		
	18q21.31	5	18q21.31 ‡	1
	18q21.31-q21.32	4		
	18q21.32	5	18q21.32	2
	18q21.32-q21.33	4		
	18q21.33-q22.1	4	18q21.33	2
	18q21.1	5	18q22.1 ‡	1
	18q22.2-q22.3	4	18q22.1-q22.3 ‡	1
X			Xp22.11-p21.3 ‡	3

NOTE: Using automated locus definition (Materials and Methods), the MCRs in localized and metastasized prostate cancer samples are defined. The cytogenetic bands where the located genes showing frequent amplification and deletion in MET and PCA are listed.

* Alteration only observed in L-PCA, but not in H-PCA within the cytoband.

† At least one PIN sample has alteration within the cytoband.

‡ Alteration only observed in H-PCA, but not in L-PCA within the cytoband.

samples profiled in the metastatic group, data from six hormone refractory tumors obtained from different metastatic sites from a single patient (case number 34) were averaged to reduce sample bias. Three additional samples (two MET and one PIN) excluded in the global view analysis (as transcript profiling was not available) were included here, taking the total number of samples in the MET group to 12 and PIN group to 7. The percentage of alterations for the selected genes was calculated for each group. The gene list was ordered according to the chromosomal location of each gene and was moving averaged (n = 5) for graphical representation. The residual prostate carcinoma obtained from case 34 was included in the metastatic group in all of our analyses as the gene expression analysis clustered this sample in the metastatic group. This sample had 851 alterations, which is in the range of alterations observed in metastatic samples from case 34 (456–1,588 altered genes).

MCR characterization. MCR characterization was done as described earlier (9, 10, 27), with some modifications. A Perl-based algorithm was applied to the normalized data. Genes with $\log_2$ ratios greater or less than the predefined cutoff values (as described above in Data collection and analysis), within the significantly altered regions identified by CGH miner, were considered as altered. The CGH miner output arranges genes according to their chromosomal location. Samples were grouped into six categories (PIA, PIN, L-PCA, H-PCA, PCA, MET). To identify most commonly amplified or deleted genes, a score was given to each gene based on the number of samples with alteration. We then scanned the scores to identify contiguous spans of altered genes having at least 75% of the peak alteration percentage to denote the MCRs.

Integrative analysis of copy number–based differential gene expression. The lists of genes that are candidates for the copy number–based differential expression were selected on the basis of three criteria: (a) the percent alteration among samples (described above), (b) the correlation between mRNA expression and aCGH, and (c) the significance of copy number change. Genes were ranked by their correlation with mRNA expression data and the degree of copy number change in either direction. The genes that ranked among the top 500 in either category were selected individually from MET and PCA groups. Among these, the genes that show aberration in more than five samples in MET and PCA groups (n = 364) were mapped and compared (n = 210) to the other prostate cancer gene expression data sets from Dhanasekaran et al. (29), LaPointe et al. (30), and Varambally et al. (31) studies, obtained from the Oncomine database.[13] For candidate gene progression analysis, a given gene must have chromosomal alteration in all stages of prostate cancer progression starting from the precursor lesions. A total of 504 (overexpressed and amplified) and 241 (underexpressed and deleted) filtered genes were mapped to the mRNA expression progression list (P value <0.05) from our matched study available in the Oncomine database (22). Genes that are ranked among the top/bottom 100 genes in mRNA expression progression list are reported in Supplementary Table S3.

Results and Discussion

Chromosomal aberrations in prostate cancer. We used LCM to isolate 62 specific cell populations from 38 patients representing a histopathologic spectrum of prostate cancer progression to perform aCGH analysis. This paper describes the results from the aCGH and its integrative analysis with gene expression data, which have been reported previously (22). Analysis of the array CGH data using CGH-Miner software identified significantly altered contiguous chromosomal regions (q value of <0.01) within each sample

[13] www.oncomine.org

(28; see Materials and Methods). The total number of genes located within these altered regions was highest (average n = 800) in metastatic samples (Fig. 1*A* and Supplementary Table S1), whereas it was lowest in the benign samples (average n = 61). On average, high-grade PIN (HGPIN) samples had 194 gene alterations, whereas 151 and 360 altered genes were found in low- and high-grade localized prostate cancer samples, respectively. PAH and atrophy had 74 alterations on average, and the range of the number of observed chromosomal aberrations in these samples was not significantly different (P value > 0.05) from benign/normal samples. However, in both HGPIN and cancer cases, the variability in the number of alterations compared with that of benign/normal samples was significantly different (HGPIN P value <0.0005; H-PCA P value <0.0003, and MET P value <0.0001).

To calibrate the resolution of our array CGH technique, we used genomic DNA samples with varying copies (1 to 5) of chromosome X hybridized against normal human male genomic DNA obtained from a commercial source. The gain in copy number was evident with increasing signal for all genes derived from chromosome X (Fig. 1*B*). Using these data, when the mean fluorescence ratios of the genes located in X chromosome from each experiment were plotted, the slope was 0.2228, with R^2 = 0.9998 (data not shown). In addition, genomic DNA from the epithelium and stroma of normal prostate tissue from a cadaveric donor, previously diagnosed with Down's syndrome, showed a single copy gain in chromosome 21 (Fig. 1*B*). Resolution of the X chromosome copy number changes, as well as the detection of a single copy gain in chromosome 21, validated the performance of our arrays and the amplification technique employed with laser-captured specimens.

The significantly altered genes from various prostate samples identified in the above analysis (Fig. 1*A*) were ordered according to human genome map position from chromosome 1 to Y, and the regions of gains and losses were displayed as a heat map (Fig. 1*B*). Benign samples, as expected, showed no significant regions of alteration (excluding the two samples with chromosome 21 amplification), whereas metastatic samples displayed the most alterations. Among the metastatic prostate cancer samples, frequent amplifications were observed in chromosomal arms 2p, 3q, 7, 8q, 9q, 16p, and 20q, and deletions were observed in 6q, 8p, 10q, 13q, and 18q (Fig. 1*B*). When the size of the altered span was considered, the H-PCA (Gleason pattern 4) samples showed more resemblance to the MET samples than to L-PCA (Gleason pattern 3), and the alteration sites tended to extend further, even encompassing entire chromosomal arms, for example, in 3q, 8q, and 13 (Fig. 1*B*). Next, we calculated the percent alteration in each group, and the alteration frequency is displayed for all the chromosomes (Supplementary Fig. S1). The aberrations observed in PIN, L-PCA, H-PCA, and MET groups in chromosomes 8, 10, and 13 revealed a distinct overlap (Fig. 2). Several known alterations, including amplifications in *TPD52* and *MYC* (8q21.13 and 8q24.22) and deletion in *PTEN* (10q23.3), are located within the highly altered regions on these chromosomes. For 8q24.22 and 10q23.3, the percent alteration in MET and H-PCA ranged from 30% to 50%, whereas 8q21.13 had a 40% alteration in MET samples. Among the PIN samples, the whole-arm amplification in 8q, which is often observed in the advanced form of the disease, was present in at least one sample. None of our other PIN samples had chromosomal arm–spanning alterations, although there were a number of smaller altered sites throughout the entire chromosome. Beheshti et al. (17) have previously reported that more extensive aberrations are observed in PCA than in PIN, and in contrast to 8q gain that was consistently observed from different tumor foci, 8q gain in PIN is not a common event. A recent prostate cancer aCGH study by Hughes et al. (18) on a BAC platform consisting of 2,400 clones identified both 8q21.11-qter gain and 8p11.23-p23.3 loss in PIN and PCA samples, with gain in more than 37.5% and 50% and loss in more than 50% of PIN and PCA samples, respectively (sample size n = 7 for PIN and 8 for PCA). Aberrations in chromosome 8 were not very common in our PIN sample cohort, with only 1 case of 8p loss and 1 case of 8q gain. However, gain of 8q and loss of 8p were frequent events (40% and 30%, respectively) in the tumor samples.

MCRs in prostate cancer. To refine the regions of alteration obtained from CGH-Miner output, we used MCR identification as described in Materials and Methods. The automated algorithm we used has been previously applied to define overlapping regions of amplification, deletion, and focal regions of recurred alterations in myeloma, pancreatic, and lung cancer aCGH studies (9, 10, 27). By using this method, we identified MCRs in PCA and MET sample groups in our prostate cancer data set (Table 1). More detailed information on all MCRs, including precise aberration sites, cytogenetic bands, number of samples altered, and the gene names within the predicted regions that meet the cutoff threshold for amplification and deletion, along with the over- and under-expressed genes from matching mRNA profiles, is provided in Supplementary Table S2.

The region 8q24.22 (132.98–134.64 Mb) had the highest percentage of alterations (50%) among the METs. Other frequently amplified regions with more than 40% of MET samples showing the alterations are 2p23.3 (24.06–24.34 Mb), 7p15.1 (29.48–31.5 Mb), 8q21.13 (31.6–81.62 Mb), 8q22.2 (99.21–99.65 Mb), 8q22.2-q22.3 (101.23–101.99 Mb), and 8q22.3 (102.57–103.29 Mb). Regions 10q23.31 (90.96–91.08 Mb), 18q21.31 (53.37–54.17 Mb), 18q21.32 (54.68–55.25 Mb), and 18q22.1 (63.32–64.54 Mb) were among the frequently deleted regions exhibited by 40% of MET samples. The region 18q21.2, which harbors *SMAD4*, is deleted in 30% of MET samples. SMAD4, a transforming growth factor-β superfamily signaling molecule, is significantly underexpressed in prostate cancer (32).

In PCA samples, regions 5q32 with 30% amplification and 13q14.12 with 30% deletion were the most frequently altered sites. Other alterations, which include amplification in 5q32 and 8q24.11 and deletion in 5q13.3, 6q14.3-q15, 10q23.2-q24.1, 13q14.12-q14.2, 16p12.2, and 17q21.31, were observed with a frequency >20%. In H-PCA samples, the deletion in 13q14.2 (47.55–48.9 Mb) was observed in more than 40% of the cases. Previously, van Dekken et al. (33) compared Gleason patterns 3 and 4 tumors obtained from the same cases by the 2,400-element BAC array, and a 34% overlap in genomic aberrations, mainly in deleted regions, was reported. The defined MCRs for H-PCA and L-PCA samples in our cohort show overlap in deleted regions 5q13.3, 6q14.2-q21, and 16p12.2; however, there were no overlapping regions of amplification between the two groups. This could be due to the identification of only four altered sites by automated MCR definition in L-PCA.

The shared regions of alterations among PIN, L-PCA, H-PCA, and MET samples were further investigated by the following method. If a given cytogenetic band shows 20% recurrent aberration within a sample type, it is mapped to the same cytogenetic band region in other sample types to assess percent recurrent aberration in that region. Identifying these shared regions is of great interest because it might shed light on the mechanism of tumor progression, especially when the alteration is detected in the precursor lesions such as PIN and is preserved or becomes more frequent in other progressive stages of the cancer. The cytogenetic bands harboring

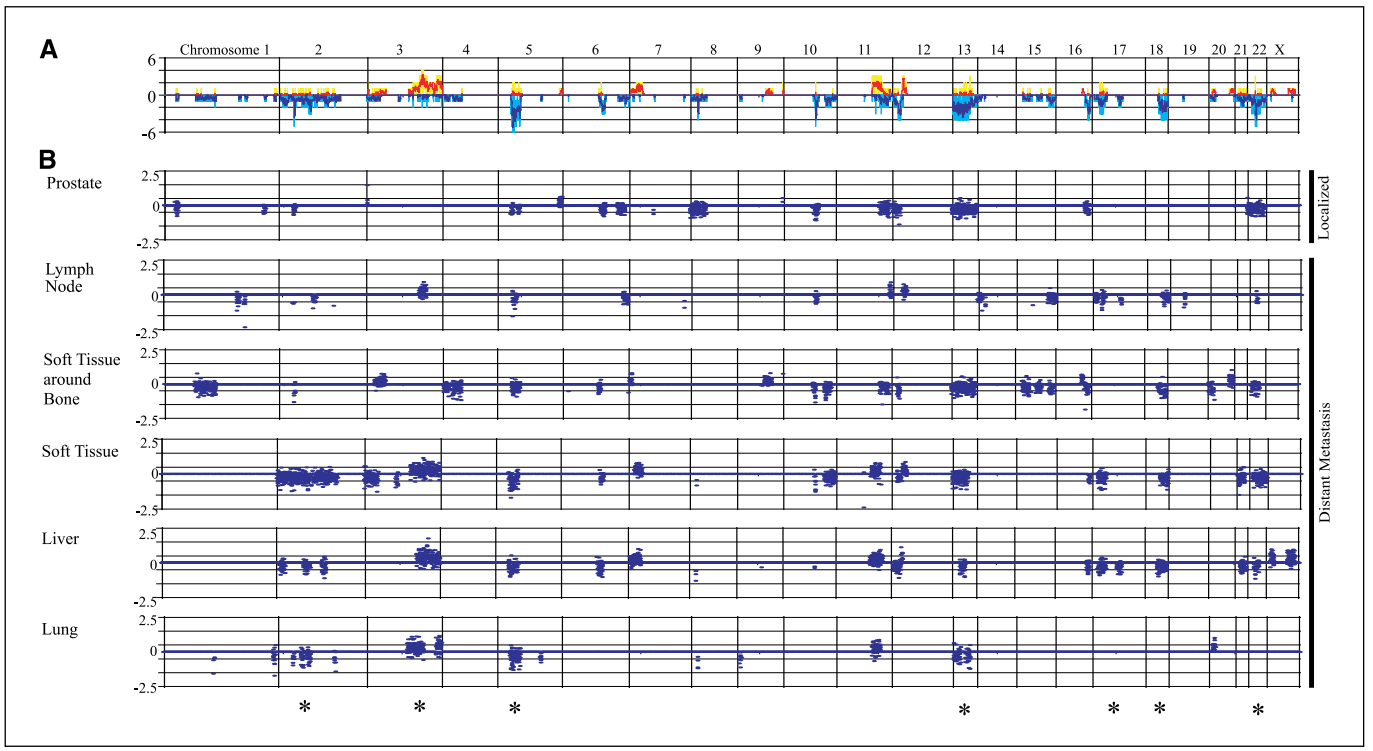

Figure 3. Chromosomal aberrations across multiple metastatic sites in a patient with lethal prostate cancer. Distribution of chromosomal alterations in tumor samples collected from five metastatic sites in addition to the residual carcinoma of prostate from a single patient. *A,* the percentage of alterations observed across all the samples. *B,* alterations observed in each individual sample. *, shared regions of alterations.

the shared amplicons in all four groups are 1q21.1-q21.3, 1q24.1-q24.2, 3q21.1-q21.3, 3q29, 6p21.33-p21.1, 7p15.1-p14.3, 8q22.2-q24.12, 11p15.4-p15.1, 11q13.1, and 12p13.32-p13.2. In most of the PCA and PIN samples, the aberration within designated cytogenetic region was seen in <20% cases. Gain in 1q or 6p denotes poor outcome in melanomas (34), suggesting that alterations in those region may be associated with poor prognosis and lower survival rate in prostate cancer as well. All groups shared deletion at 6q16.2-q22.31, 13q12.12-q32.1, 17p12-p11.2, and 18q21.1-q23, whereas the deleted region in chromosome 13 is a frequent event in both MET and H-PCA cases (>30%). Among 13 defined MCRs in PIN, 9 are shared with PCA. These regions include amplifications in 3q29, 5q31.3-q32, 5q32, 6q27, and 8q24.3 and deletions in 6q22.31, 16p12.2, 17q21.2, and 17q21.31. 17q21.31 is known to be completely lost in the PC3 cell line (35). However, among three defined MCRs in PIA samples, none are shared with PIN or PCA. Among the putative precursor lesions, PIN, but not PIA, bears a closer resemblance to prostate cancer in copy-number alterations.

Overall, the most frequent amplifications are observed in 2p, 3q, 5q, 7p, 8q, 9q, 16p, and 20q for MET and 3q, 5q, 7q, and 8q for PCA. Deletion-prone sites in MET are 6q, 8p, 10q, 13q, and 18q, and they are 1q, 3q, 5q, 6q, 10q, 13q, 16p, 17q, and Xp in PCA, where most of the alterations are obtained from H-PCA samples.

Chromosomal aberrations in distant metastases. Several metastatic specimens from a single patient (case 34) that include lymph node, lung, liver, soft tissue around bone, other soft tissue, and tissue from the residual prostate gland were profiled in this study (Supplementary Table S1). CGH-Miner output of significant alterations (q value of <0.01) for each sample was displayed for comparison (Fig. 3). Common alterations that occur in at least three samples are detected in 3q, 5q, 11q, 12p, 13q, 16q23-q24, 17p, 18q, and 22q (Fig. 3). The majority of the alteration sites are overlapping; however, there are also unshared altered sites present, likely accounted for by the heterogeneity in each clonal group as well as due to the alterations that occur after metastasis. The number of genes within the identified aberrant sites ranges from 583 to 1723, and the residual carcinoma of prostate gland and lymph node MET displayed lower number of altered genes compared with the other sites. Interestingly, in an earlier integrative mRNA and copy-number study on chromosome 16q, deletion in 16q23.1 to 16qter (a region harboring many candidate tumor suppressor genes) was reported in more than 50% of prostate cancer samples examined (36). Many other genes known to be deleted in prostate cancer, such as *CYP1B1, TNFRSF10B, ATAD1*, and *PTEN*, are located within the commonly deleted regions in these distant metastasis samples.

Analysis of array CGH data and mRNA expression profiles on matched LCM prostate specimens. The aim of our integrative analysis was to identify candidate regions with genetic alterations that accompany corresponding transcriptomic changes. Transcript expression patterns of genes located in the chromosomal regions with significant aberrations in identical samples were compared (Fig. 1*B*). An association between mRNA overexpression and chromosomal gain was observed, such that among highly amplified genes ($\log_2$ ratio $\geq$0.5), there was a 26% and 20% concordance in high-level mRNA expression ($\log_2$ ratio $\geq$1) for MET and PCA samples, respectively, whereas at moderate-level mRNA expression ($\log_2$ ratio $\geq$0.4), a 42% and 22% concordance was observed for MET and PCA tissues. Among all the amplified genes ($\log_2$ ratio $\geq$0.22), a moderate-level overexpression ($\log_2$ ratio $\geq$0.4) was observed in 38% of MET and 20% of PCA cases, and a high-level overexpression ($\log_2$ ratio $\geq$1) was observed in 23% MET and 15%

of PCA cases. A previous breast cancer aCGH and coupled gene expression study estimated a 62% (representing 54 unique genes) association between 117 highly amplified genes and transcript overexpression (8). Hyman et al. (37) have reported 44% highly amplified genes associated with overexpression and 10.5% of highly overexpressed genes to be amplified in breast cancer. In pancreatic cancer, as many as 60% of the genes located within highly amplified regions were reported overexpressed (38). Similar high-level concordance between DNA copy number change and mRNA expression level was also observed in our prostate study, in which the increased dosage in the gene copy number most likely plays a major role in their transcriptional up-regulation. The MCM analysis (22) from the Oncomine database revealed the enrichment of overexpressed genes in chromosomal arms 8q, 1q, 7p, 9q, 16p, 10p, and 3q (P value <0.05), where 8q, 7p, 9q, 16p, and 3q are among the top chromosomal alteration sites in our MCR analysis (Table 1). Integrative analysis of our aCGH and gene expression data allows a direct comparison between the change in copy number and transcript expression levels, and genes within regions of significant genomic alterations show concordance at the mRNA expression level.

Integrative analysis of genomic and transcriptomic profiles associated with prostate cancer progression. To identify the top altered genes that are associated with a change in expression level, we selected the candidate genes based on three criteria mentioned under "Integrative analysis of copy number–based differential gene expression" in Materials and Methods. These significantly altered genes are located within the commonly observed regions of chromosomal aberrations and are accompanied with the altered mRNA expression in a correlated manner. The chosen genes from PCA and MET samples are likely to play a role in mRNA expression level, and they are cross-indexed with three independent gene expression data sets available in the public domain (Fig. 4). The gene expression data were obtained from grossly dissected localized and metastatic prostate tumor tissues from previously

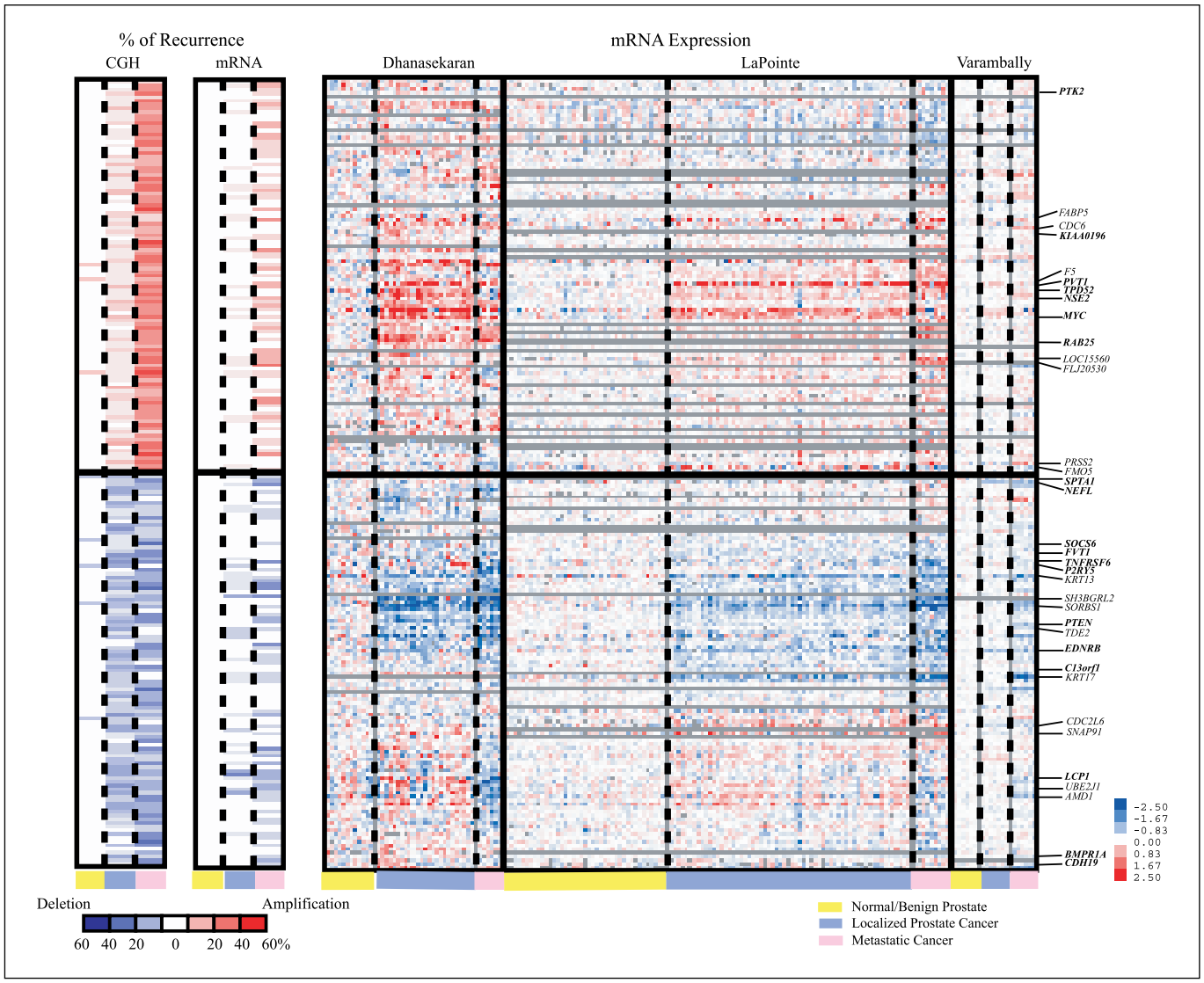

Figure 4. Concordantly altered candidate genes in various prostate cancer studies. The proposed candidate amplified/deleted genes that are correlated with matched mRNA expression data with high percentage of alterations are mapped to Dhanasekaran et al., Lapointe et al., and Varambally et al. data sets available from www.oncomine.org and are displayed. *Left,* percentage of alterations in CGH (*blue,* percentage of deleted samples; *red,* percentage of amplified samples) and the percentage of samples with overexpressed/underexpressed genes in matched mRNA expression data.

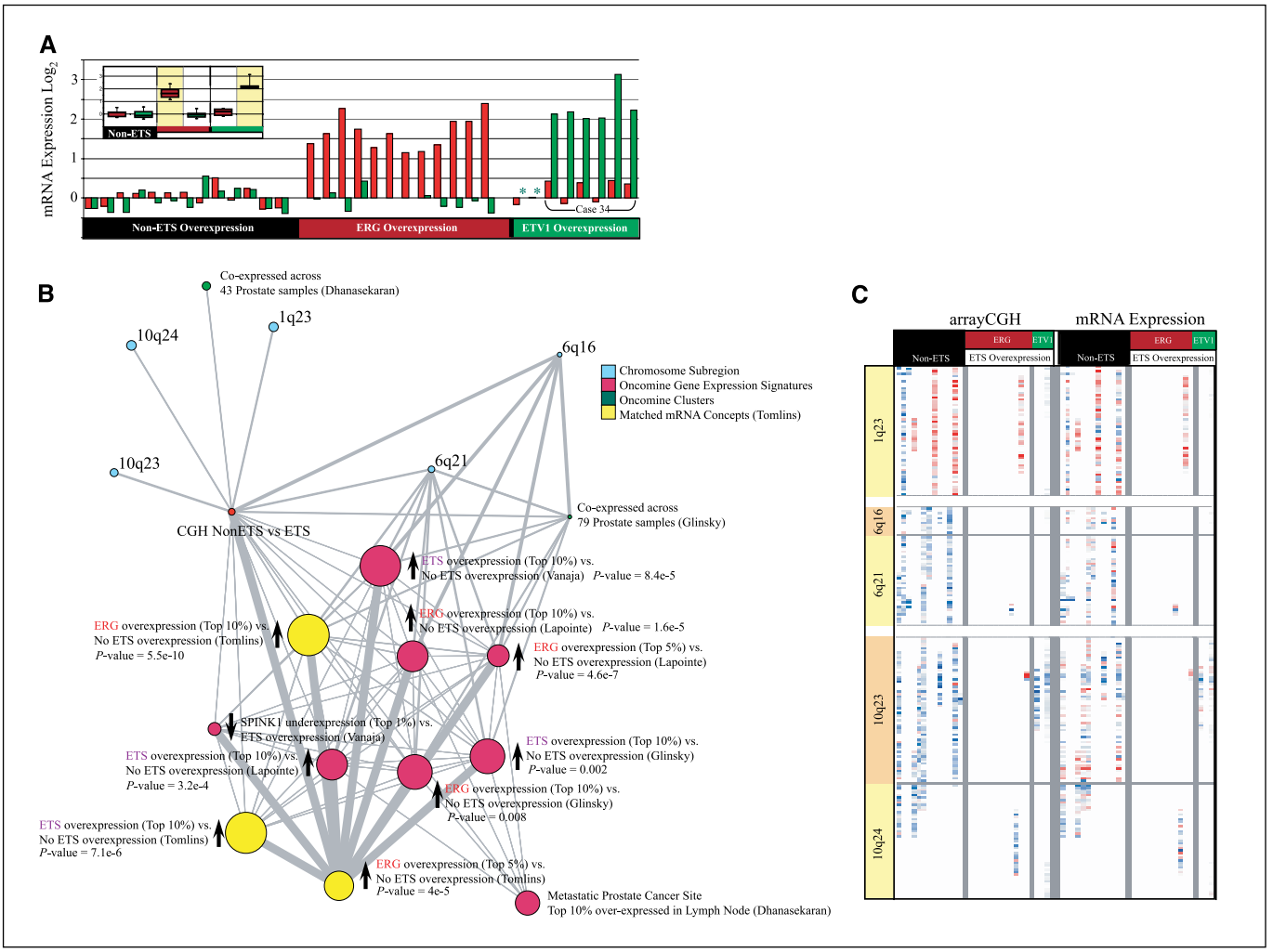

Figure 5. Genetic alterations in non-ETS versus ETS samples in prostate cancer. The genomic aberration differences as well as the enriched concepts of the genes located within differentially altered regions in non-ETS and ETS samples were analyzed using the MCM (22). *A,* mRNA expression of non-ETS and ETS overexpressing prostate cancer samples. *Red bar,* for *ERG* expression values; *green bar,* for *ETV1* expression values. *Top left inset,* average of the data presented. *B,* network map showing enrichment in chromosomal subregions and gene expression signatures that define non-ETS and ETS samples. "CGH non-ETS versus ETS" represents the data gathered from this study. *C,* heat map of differentially aberrant genomic regions between non-ETS and ETS samples. *, data not available (*ETV1* expression is confirmed from an independent sample obtained from the same case).

reported studies from our group (29, 39) and others (30). The amplified gene section was enriched with transcript overexpression, and the deleted section was enriched with mRNA downregulation. These differentially expressed genes are located in either PCA and MET or both. Well-known amplified genes such as *MYC* and *TPD52* (40–43) are among the top genes that show overexpression pattern in PCA and MET samples in various data sets. The tumor suppressor *PTEN* and suppressor of cytokine signaling, *SOCS6*, that are known to be deleted in various cancers are also seen as underexpressed. Some of the other previously described gains are *PTK2, KIAA0196, PVT1, NSE2*, and *RAB25*, and some of the earlier reported losses include *SPTA1, NEFL, FVT1, TNFRSF6, EDNRB, C13ORF1, LCP1, BMPR1A*, and *CDH19*. Some of the novel amplified and deleted genes identified in this study include *DDEF1, LCHN, F5, DDX56, P2RY5, ATAD1, ZNF532, RAB27B*, and *PPIL6*, which merit further characterization. In addition, a progressive gene signature was identified by the transcriptome analysis done on identical samples studied here. These genes showed a robust progression signature whose expression increased or decreased during the progression from benign epithelium to PIN to PCA to MET (22). We used the aCGH data for the corresponding samples to look for possible underlying genetic alterations involving the proposed candidate genes. The data are presented in Supplementary Table S3.

Genetic alterations in ETS versus non-ETS samples. ETS transcription factors that include *ERG, ETV1,* and *ETV4* were identified as outliers in prostate cancer gene expression data set and are shown to be involved in recurrent gene fusion (44). Two recent studies, one using single nucleotide polymorphism arrays on human prostate cancer tissues and the other using BAC arrays on prostate cancer xenografts (14), propose interstitial deletions as a mechanism of *TMPRSS2:ERG* gene fusion on chromosome 21 (44). These gene fusions seem to be one of the earliest events involving prostate cancer and lead to the overexpression of the fused ETS gene in an androgen-regulated manner (45). We previously characterized ETS expression in the cancer samples used in this study (22). Accordingly, the localized and metastatic prostate cancer samples were either grouped into ETS (*ERG* or *ETV1* overexpressing) or non-ETS

samples (Fig. 5*A*). A significance test was done to identify regions that distinguish between these two sample groups from the CGH miner output (q value of <0.01). A total of 50 genes passed the cutoff of P value <0.05, and this list was analyzed using MCM. MCM identified chromosome subregion concepts like 1q23 (P value <4.1e−4), 6q16 (P value <1.4e−9), 6q21 (P value <1.5e−5), 10q23 (P value <7.5e−7), and 10q24 (P value <2.1e−4). Importantly, various oncomine gene expression signature concepts that define ETS-positive versus non-ETS samples from the matched mRNA data set (22) and other independent data set like Lapointe et al. (30) and Glinsky et al. (46) were enriched in this analysis (Fig. 5*B*). The aberration summary and accompanying gene expression pattern in these chromosomal subregions are presented as a heat map (Fig. 5*C*). Gene expression analysis by Tomlins et al. showed differential enrichment in chromosome subarm 6q21 between ETS and non-ETS samples. We speculated either amplification of 6q21 in ETS or loss in non-ETS tumors. Our aCGH data showed several non-ETS samples with loss of 6q21 region (>45%), suggesting that underexpression of genes from this region could be due to deletions in a subset of non-ETS samples. Several groups have previously identified the loss of 6q21 in localized prostate cancers (47), and here we show this phenomenon to be mainly associated with non-ETS samples. *FOXO3A* (48) and *CCNC* (47) that have been proposed to participate in prostate carcinogenesis are located in this region. These alterations on 6q21 and others identified in this analysis may collectively play a role in tumor development in the non-ETS group, and further molecular characterization of these alterations is required to understand its importance in prostate cancer. This observation was also validated in an independent, grossly dissected prostate cancer aCGH data set (data not shown).

In conclusion, aCGH analysis of laser-capture–microdissected prostate cancer samples detected a multitude of chromosomally altered regions through the various stages of prostate cancer progression. Samples like PAH and PIA were characterized for the first time by aCGH in this study. MCRs were identified, and the percentage of alterations in various prostate cancer stages that reflect the entire spectrum of the disease progression was calculated. This generated a list of altered regions and candidate genes that might play a role in cancer progression. The prostate cancer precursor lesion PIN resembled PCA in its genetic alterations. The direct relationship between copy-number change and mRNA expression levels was investigated using a parallel transcriptomic study, where more than 40% of the highly altered genes were associated with elevated mRNA expression level. This study also has identified some novel regions of aberrations and candidate genes in prostate cancer. Lastly, MCM analysis of the cancer specimens identified chromosomal regions including 6q21 and gene expression concepts that distinguish ETS overexpressing samples from non-ETS samples.

Acknowledgments

Received 4/10/2007; revised 6/1/2007; accepted 6/11/2007.

Grant support: A.M. Chinnaiyan is supported by a Burroughs Wellcome Foundation Award in Clinical Translational Research. K.J. Pienta is supported by an American Cancer Society Award. S.A. Tomlins is supported by the Medical Scientist Training Program and a Rackham Pre-doctoral Award. Supported in part by NIH grants RO1 CA97063 (A.M. Chinnaiyan and D. Ghosh), RO1 CA102872 (K.J. Pienta), U01 CA111275 (A.M. Chinnaiyan and D. Ghosh), and P50 CA69568 (K.J. Pienta, A.M. Chinnaiyan, and D. Ghosh); Department of Defense PC040517 (R. Mehra) and PC051081 (A.M. Chinnaiyan); and the Ralph Wilson Medical Research Foundation Grant (K.J. Pienta).

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 A. Menon for help with microarray production; D. Rhodes, B. Laxman, and S. Kalyana-Sundaram, for helpful discussions; S. Bhagavathula, J. Siddiqui, and R. Varambally for tissue database help.

References

1. Albertson DG, Collins C, McCormick F, Gray JW. Chromosome aberrations in solid tumors. Nat Genet 2003;34:369–76.
2. Cahill DP, Kinzler KW, Vogelstein B, Lengauer C. Genetic instability and darwinian selection in tumours. Trends Cell Biol 1999;9:M57–60.
3. Ylstra B, van den Ijssel P, Carvalho B, Brakenhoff RH, Meijer GA. BAC to the future! or oligonucleotides: a perspective for micro array comparative genomic hybridization (array CGH). Nucleic Acids Res 2006;34:445–50.
4. Pinkel D, Albertson DG. Array comparative genomic hybridization and its applications in cancer. Nat Genet 2005;37 Suppl:S11–7.
5. Cho EK, Tchinda J, Freeman JL, Chung YJ, Cai WW, Lee C. Array-based comparative genomic hybridization and copy number variation in cancer research. Cytogenet Genome Res 2006;115:262–72.
6. Tsafrir D, Bacolod M, Selvanayagam Z, et al. Relationship of gene expression and chromosomal abnormalities in colorectal cancer. Cancer Res 2006;66:2129–37.
7. Fritz B, Schubert F, Wrobel G, et al. Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. Cancer Res 2002;62:2993–8.
8. Pollack JR, Sorlie T, Perou CM, et al. Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. Proc Natl Acad Sci U S A 2002;99:12963–8.
9. Tonon G, Wong KK, Maulik G, et al. High-resolution genomic profiles of human lung cancer. Proc Natl Acad Sci U S A 2005;102:9625–30.
10. Aguirre AJ, Brennan C, Bailey G, et al. High-resolution characterization of the pancreatic adenocarcinoma genome. Proc Natl Acad Sci U S A 2004;101:9067–72.
11. Chaudhary J, Schmidt M. The impact of genomic alterations on the transcriptome: a prostate cancer cell line case study. Chromosome Res 2006;14:567–86.
12. Zhao H, Kim Y, Wang P, et al. Genome-wide characterization of gene expression variations and DNA copy number changes in prostate cancer cell lines. Prostate 2005;63:187–97.
13. Wolf M, Mousses S, Hautaniemi S, et al. High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression. Neoplasia 2004;6:240–7.
14. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM, Trapman J. TMPRSS2:ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. Cancer Res 2006;66:10658–63.
15. Saramaki OR, Porkka KP, Vessella RL, Visakorpi T. Genetic aberrations in prostate cancer by microarray analysis. Int J Cancer 2006;119:1322–9.
16. Paris PL, Andaya A, Fridlyand J, et al. Whole genome scanning identifies genotypes associated with recurrence and metastasis in prostate tumors. Hum Mol Genet 2004;13:1303–13.
17. Beheshti B, Vukovic B, Marrano P, Squire JA, Park PC. Resolution of genotypic heterogeneity in prostate tumors using polymerase chain reaction and comparative genomic hybridization on microdissected carcinoma and prostatic intraepithelial neoplasia foci. Cancer Genet Cytogenet 2002;137:15–22.
18. Hughes S, Yoshimoto M, Beheshti B, Houlston RS, Squire JA, Evans A. The use of whole genome amplification to study chromosomal changes in prostate cancer: insights into genome-wide signature of preneoplasia associated with cancer progression. BMC Genomics 2006;7:65.
19. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23:41–6.
20. 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.
21. Shah R, Mucci NR, Amin A, Macoska JA, Rubin MA. Postatrophic hyperplasia of the prostate gland: neoplastic precursor or innocent bystander? Am J Pathol 2001;158:1767–73.
22. Tomlins SA, Mehra R, Rhodes DR, et al. Integrative molecular concept modeling of prostate cancer progression. Nat Genet 2007;39:41–51.
23. Dhanasekaran SM, Dash A, Yu J, et al. Molecular profiling of human prostate tissues: insights into gene expression patterns of prostate development during puberty. FASEB J 2005;19:243–5.
24. Little SE, Vuononvirta R, Reis-Filho JS, et al. Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA. Genomics 2006;87:298–306.
25. Yang YH, Dudoit S, Luu P, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 2002;30:e15.
26. Herrero J, Al-Shahrour F, Diaz-Uriarte R, et al. GEPAS: A web-based resource for microarray gene expression data analysis. Nucleic Acids Res 2003;31:3461–7.

27. Carrasco DR, Tonon G, Huang Y, et al. High-resolution genomic profiles define distinct clinicopathogenetic subgroups of multiple myeloma patients. Cancer Cell 2006;9:313–25.

28. Wang P, Kim Y, Pollack J, Narasimhan B, Tibshirani R. A method for calling gains and losses in array CGH data. Biostatistics 2005;6:45–58.

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

30. 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.

31. Varambally S, Yu J, Laxman B, et al. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. Cancer Cell 2005;8:393–406.

32. Horvath LG, Henshall SM, Kench JG, et al. Loss of BMP2, Smad8, and Smad4 expression in prostate cancer progression. Prostate 2004;59:234–42.

33. van Dekken H, Paris PL, Albertson DG, et al. Evaluation of genetic patterns in different tumor areas of intermediate-grade prostatic adenocarcinomas by high-resolution genomic array analysis. Genes Chromosomes Cancer 2004;39:249–56.

34. Namiki T, Yanagawa S, Izumo T, et al. Genomic alterations in primary cutaneous melanomas detected by metaphase comparative genomic hybridization with laser capture or manual microdissection: 6p gains may predict poor outcome. Cancer Genet Cytogenet 2005;157:1–11.

35. Clark J, Edwards S, Feber A, et al. Genome-wide screening for complete genetic loss in prostate cancer by comparative hybridization onto cDNA microarrays. Oncogene 2003;22:1247–52.

36. Watson JE, Doggett NA, Albertson DG, et al. Integration of high-resolution array comparative genomic hybridization analysis of chromosome 16q with expression array data refines common regions of loss at 16q23-qter and identifies underlying candidate tumor suppressor genes in prostate cancer. Oncogene 2004;23:3487–94.

37. Hyman E, Kauraniemi P, Hautaniemi S, et al. Impact of DNA amplification on gene expression patterns in breast cancer. Cancer Res 2002;62:6240–5.

38. Heidenblad M, Lindgren D, Veltman JA, et al. Microarray analyses reveal strong influence of DNA copy number alterations on the transcriptional patterns in pancreatic cancer: implications for the interpretation of genomic amplifications. Oncogene 2005;24:1794–801.

39. Varambally S, Dhanasekaran SM, Zhou M, et al. The polycomb group protein EZH2 is involved in progression of prostate cancer. Nature 2002;419:624–9.

40. Rubin MA, Varambally S, Beroukhim R, et al. Overexpression, amplification, and androgen regulation of TPD52 in prostate cancer. Cancer Res 2004;64:3814–22.

41. Boutros R, Fanayan S, Shehata M, Byrne JA. The tumor protein D52 family: many pieces, many puzzles. Biochem Biophys Res Commun 2004;325:1115–21.

42. Byrne JA, Balleine RL, Schoenberg Fejzo M, et al. Tumor protein D52 (TPD52) is overexpressed and a gene amplification target in ovarian cancer. Int J Cancer 2005;117:1049–54.

43. Wang R, Xu J, Saramaki O, et al. PrLZ, a novel prostate-specific and androgen-responsive gene of the TPD52 family, amplified in chromosome 8q21.1 and overexpressed in human prostate cancer. Cancer Res 2004;64:1589–94.

44. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644–8.

45. Rubin MA, Chinnaiyan AM. Bioinformatics approach leads to the discovery of the TMPRSS2:ETS gene fusion in prostate cancer. Lab Invest 2006;86:1099–102.

46. Glinsky GV, Glinskii AB, Stephenson AJ, Hoffman RM, Gerald WL. Gene expression profiling predicts clinical outcome of prostate cancer. J Clin Invest 2004;113:913–23.

47. Konishi N, Shimada K, Ishida E, Nakamura M. Molecular pathology of prostate cancer. Pathol Int 2005;55:531–9.

48. Trotman LC, Alimonti A, Scaglioni PP, Koutcher JA, Cordon-Cardo C, Pandolfi PP. Identification of a tumour suppressor network opposing nuclear Akt function. Nature 2006;441:523–7.

Integrative Analysis of Genomic Aberrations Associated with Prostate Cancer Progression

Jung H. Kim, Saravana M. Dhanasekaran, Rohit Mehra, et al.

Cancer Res 2007;67:8229-8239.