Comparative Analyses of Chromosome Alterations in Soft-Tissue Metastases within and across Patients with Castration-Resistant Prostate Cancer

Ilona N. Holcomb, Janet M. Young, Ils a M. Coleman, Keyan Salari, Douglas I. Grove, Li Hsu, Lawrence D. True, Martine P. Roudier, Colm M. Morrissey, Celestia S. Higano, Peter S. Nelson, Robert L. Vessella, and Barbara J. Trask

Abstract
Androgen deprivation is the mainstay of therapy for progressive prostate cancer. Despite initial and dramatic tumor inhibition, most men eventually fail therapy and die of metastatic castration-resistant (CR) disease. Here, we characterize the profound degree of genomic alteration found in CR tumors using array comparative genomic hybridization (array CGH), gene expression arrays, and fluorescence in situ hybridization (FISH). By cluster analysis, we show that the similarity of the genomic profiles from primary and metastatic tumors is driven by the patient. Using data adjusted for this similarity, we identify numerous high-frequency alterations in the CR tumors, such as 8p loss and chromosome 7 and 8q gain. By integrating array CGH and expression array data, we reveal genes whose correlated values suggest they are relevant to prostate cancer biology. We find alterations that are significantly associated with the metastases of specific organ sites, and others with CR tumors versus the tumors of patients with localized prostate cancer not treated with androgen deprivation. Within the high-frequency sites of loss in CR metastases, we find an overrepresentation of genes involved in cellular lipid metabolism, including PTEN. Finally, using FISH, we verify the presence of a gene fusion between TMPRSS2 and ERG suggested by chromosome 21 deletions detected by array CGH. We find the fusion in 54% of our CR tumors, and 81% of the fusion-positive tumors contain cells with multiple copies of the fusion. Our investigation lays the foundation for a better understanding of and possible therapeutic targets for CR disease, the poorly responsive and final stage of prostate cancer. [Cancer Res 2009;69(19):7793–802]

Introduction
Genomic analyses of malignant disease are intended to distinguish the molecular features that underlie carcinogenesis and identify clinical targets. Moreover, comparing primary and metastatic tumors is an exceptionally useful way to assess the molecular alterations associated with stage or progression.

The plethora of molecular events that occur in prostate cancer, with no single ubiquitous alteration, illustrate the complex biology of this disease. This complexity is likely to result, in part, from different genetic backgrounds and environmental exposures of the patients. One way to reduce this heterogeneity when comparing primaries to metastases is to assay both tumor types from the same patient, but these matching sets are difficult to obtain. A decade or more can elapse between the resection of the primary tumor by prostatectomy, detection of overt metastases, and death from castration-resistant (CR) disease. To address this deficiency, we set out to obtain sets of primary and metastatic tumors from the same patient. Although technical limitations prevented us from including bone metastases in this study, we present here analyses of an extraordinary set of multiple soft-tissue metastases.

To block the effects of androgens on tumor growth, patients with advanced disease are often deprived of androgen by surgical or chemical castration. However, aggressive and ultimately lethal CR disease inevitably develops. Few treatment options with clinical benefits exist (1–3), and most represent palliative interventions once this CR state is achieved.

One potential diagnostic marker and treatment target, the fusion of the TMPRSS2 and ERG genes, has generated considerable interest (4–10). Deletion of the 3-Mbp region of 21q22.2 between the promoter of the androgen-regulated serine-protease TMPRSS2 and the 3’ exons of ERG, a member of the oncogenic ETS family of transcription factors, is the principal mechanism for this gene fusion (7, 10). Androgens are presumed to drive the expression of this oncogenic fusion, which is found in ∼30% of prostate cancers.

Here, we characterize the genomic changes in CR prostate cancer using matching sets of primary prostate tumors and metastases. This work reveals alterations that might have causative properties, clinical value, or site-specific consequences for this end-stage prostate cancer.

Materials and Methods
Sample acquisition. Use of human samples was approved by the Institutional Review Boards of participating institutions. Tumor samples were collected by radical prostatectomy or from autopsies performed at the University of Washington Medical Center under the rapid autopsy program as described previously (11). Available clinical data (stage, Gleason grade, treatment, etc.) are provided in Supplementary Table S1. Autopsies were done within 2 to 4 h of death on 14 patients [median (range) age at death, 67 y (47–83 y)] with clinically diagnosed CR disease. Fifty-four tumors were obtained from various organ sites (Supplementary Table S2). Radical prostatectomy specimens from 19 individuals with organ-confined (i.e., localized) prostate cancer not treated by androgen deprivation were also collected.

Specimens (n = 73) were embedded in freezing media (Tissue-Tek OCT Compound, Sakura Finetek) and stored in liquid nitrogen. Cells from the 54 CR tumors, 9 localized prostate cancers (LocPC), or normal stromal of 10

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

Requests for reprints: Barbara J. Trask, Human Biology Division, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, Mail Stop C3-168, PO Box 19024, Seattle, WA 98109, Phone: 206-667-1470, Fax: 206-667-4023, E-mail: btrask@fhcrc.org.

Published Online First on September 22, 2009 as 10.1158/0008-5472.CAN-08-3810
LocPC patients were isolated by laser-capture microdissection (LCM) as described previously (12). A pathologist (L.D.T.) reviewed all LCM images.

**DNA isolation and amplification.** DNA from LCM-collected samples was isolated using the QiAamp DNA Micro kit (Qiagen, Inc.). DNA from 42 tumors (Supplementary Table S2) was amplified by ligation-mediated PCR (LMP) as described previously (13). DNA samples from 12 CR tumors and 9 LocPC tumors could not successfully be amplified by LMP and instead were amplified using a whole genome amplification (WGA) kit, WGA2 (Sigma-Aldrich). DNAs amplified from these two methods are comparable (14). Of the 10 normal samples, 5 were amplified by LMP and 5 by WGA.

Reference DNA was obtained from the peripheral blood of a single female individual isolated using the QiAamp DNA Blood Mini kit (Qiagen, Inc.). The amplification method for all reference samples was matched with the test sample.

**Array comparative genomic hybridization analysis.** The bacterial artificial chromosome (BAC) clones that make up the array, array comparative genomic hybridization (array CGH) methods, and analyses have been described previously (14, 15). Clone coordinates given refer to the May 2004 sequence assembly (Build 35).

The log2/ratio array data were normalized with a block-level Loess algorithm (16) and processed by Circular Binary Segmentation (CBS; ref. 17) to organize the output into segments of approximately equal copy number.

Thresholds for calling loss and gain in arrays of LMP- or WGA-amplified material were determined using the array results obtained from the normal-cell reference DNA.

**Genomic alteration data** were identified by performing a Pearson's correlation for the CBS-segmented log2(ratio) array data. A gene was considered to have a genomic alteration if the correlation coefficient from the CBS algorithm (similarity metric was correlation centered) using Gene Cluster software was greater than 1.0. The consistency of alterations for the tumors was evaluated using a non-parametric statistical test (FDR). For each analysis, 1,000 permutations were done.

**Significant associations between alterations and tumor state (i.e., CR primaries or CR metastases) or CR organ site (i.e., prostate, lymph-node metastasis, or liver metastasis) were identified using the Significance Analysis of Microarrays (SAM) method (19) using response formats two-class (unpaired) and multiclass, respectively. SAM is based on a modified t statistic and uses random permutations of class labels to estimate a false discovery rate (FDR). For each analysis, 1,000 permutations were done.

**Results**

Hierarchical clustering groups tumors from the same patient together. We performed hierarchical cluster analysis to evaluate relationships among the 54 CR tumors from 14 patients based on their genomic alterations. These CR tumors include those resected from the prostate (and called CR primaries here) and multiple soft-tissue metastatic sites. Tumors most often group with other tumors from the same patient, rather than by organ of origin (Fig. 1A). The tumors for 11 of the 14 (79%) patients define patient-specific clusters. The closest neighbor in the hierarchy for 33 of 54 tumors is a tumor from the same patient. Including LocPC and normal arrays did not alter the relationships observed for the CR tumors (Fig. 1B). The consistency of alterations for the tumors of a given patient is illustrated in Fig. 1C and emphasized by the separation of the distributions of pairwise correlation coefficients calculated for all intratumor pairs versus for all intertumor pairs (Supplementary Fig. S1).

**Tumor-related loci and candidate genes in the frequent alterations of CR tumors.** Every position in the genome is represented in the cumulative spectrum of changes observed in one or more of the 54 CR tumors (Fig. 2). To summarize these results, we first adjusted for the variation in number of tumors evaluated for each patient before calculating the frequency of genomic alterations across the CR patients. BAC clones encompassed by a lost or gained segment were assigned a value of −1 or 1, respectively; no change was assigned a value of 0. For each loss or gain, we first calculated the average value for each patient and then averaged the resultant fractions across all 14 patients to represent the frequency of a given deviation in this patient set. All frequencies noted in this work are these adjusted frequencies.

To define deviations of interest, we used the one-sample binomial test to calculate a threshold frequency such that for any deviation with frequency exceeding this threshold, the 99% confidence interval of its frequency did not include 0. This frequency is \( \frac{z^2}{(n + z^2)} \), where \( z \) is the critical value (Standard normal table) and \( n \) is the sample size. For \( n \) of 14 and \( P \) value of 0.05, the threshold frequency is 0.05.<ref offenders in this work are these adjusted frequencies.

To define deviations of interest, we used the one-sample binomial test to calculate a threshold frequency such that for any deviation with frequency exceeding this threshold, the 99% confidence interval of its frequency did not include 0. This frequency is \( \frac{z^2}{(n + z^2)} \), where \( z \) is the critical value (Standard normal table) and \( n \) is the sample size. For \( n \) of 14 and \( P \) value of 0.05, the threshold frequency is 0.05. To identify genes with significant differences both in copy number and gene expression level for each gene, a modified Student's \( t \) test was applied to generate a copy-number score and a gene-expression score for the primary or metastatic samples. The two scores were summed and weighted to favor genes with substantial differences between the two tumor types. Significance was established by recalculating the scores on 1,000 random permutations of the sample labels (FDR of 5%).

**TMPRSS2:ERG fluorescent in situ hybridization.** Five-micrometer tumor sections from frozen tissue blocks were fixed in 3:1 methanol/acetate acid. Four BAC DNAs were used as probes, RP11-35C4 (probe 1), RP11-95I21 (probe 2), RP11-476D17 (probe 3), and RP11-120C17 (probe 4) (Fig. 4). BAC DNA was labeled using a nick-translation kit and either Spectrum Red-dUTP, Spectrum Green-dUTP, or Spectrum Aqua-dUTP (Abbot Molecular). Hybridization was performed as described previously (22).

Fluorescence in situ hybridization (FISH) signals were scored manually (<100 oil immersion). Fusion-positive tumors had positive nuclei (i.e., juxta-position of probe 1 and probe 3 with concurrent loss or dissociation of probe 2) in at least 20 of 50 cells analyzed. We designated the tumor as possessing multiple copies of the fusion if at least five of 20 positive nuclei showed at least two gene-fusion probe conformations. The results were confirmed in a second FISH experiment (using the same scoring criteria) testing for retention of probe 1 concurrent with dissociation or loss of probe 4.

**Results**

Hierarchical clustering groups tumors from the same patient together. We performed hierarchical cluster analysis to evaluate relationships among the 54 CR tumors from 14 patients based on their genomic alterations. These CR tumors include those resected from the prostate (and called CR primaries here) and multiple soft-tissue metastatic sites. Tumors most often group with other tumors from the same patient, rather than by organ of origin (Fig. 1A). The tumors for 11 of the 14 (~79%) patients define patient-specific clusters. The closest neighbor in the hierarchy for 33 of 54 tumors is a tumor from the same patient. Including LocPC and normal arrays did not alter the relationships observed for the CR tumors (Fig. 1B). The consistency of alterations for the tumors of a given patient is illustrated in Fig. 1C and emphasized by the separation of the distributions of pairwise correlation coefficients calculated for all intratumor pairs versus for all intertumor pairs (Supplementary Fig. S1).

**Tumor-related loci and candidate genes in the frequent alterations of CR tumors.** Every position in the genome is represented in the cumulative spectrum of changes observed in one or more of the 54 CR tumors (Fig. 2). To summarize these results, we first adjusted for the variation in number of tumors evaluated for each patient before calculating the frequency of genomic alterations across the CR patients. BAC clones encompassed by a lost or gained segment were assigned a value of ±1 or 1, respectively; no change was assigned a value of 0. For each loss or gain, we first calculated the average value for each patient and then averaged the resultant fractions across all 14 patients to represent the frequency of a given deviation in this patient set. All frequencies noted in this work are these adjusted frequencies.

To define deviations of interest, we used the one-sample binomial test to calculate a threshold frequency such that for any deviation with frequency exceeding this threshold, the 99% confidence interval of its frequency did not include 0. This frequency is \( \frac{z^2}{(n + z^2)} \), where \( z \) is the critical value (Standard normal table) and \( n \) is the sample size. For \( n \) of 14 and \( P \) value of 0.05, the threshold frequency is 0.05.

**Results**

Hierarchical clustering groups tumors from the same patient together. We performed hierarchical cluster analysis to evaluate relationships among the 54 CR tumors from 14 patients based on their genomic alterations. These CR tumors include those resected from the prostate (and called CR primaries here) and multiple soft-tissue metastatic sites. Tumors most often group with other tumors from the same patient, rather than by organ of origin (Fig. 1A). The tumors for 11 of the 14 (~79%) patients define patient-specific clusters. The closest neighbor in the hierarchy for 33 of 54 tumors is a tumor from the same patient. Including LocPC and normal arrays did not alter the relationships observed for the CR tumors (Fig. 1B). The consistency of alterations for the tumors of a given patient is illustrated in Fig. 1C and emphasized by the separation of the distributions of pairwise correlation coefficients calculated for all intratumor pairs versus for all intertumor pairs (Supplementary Fig. S1).

**Tumor-related loci and candidate genes in the frequent alterations of CR tumors.** Every position in the genome is represented in the cumulative spectrum of changes observed in one or more of the 54 CR tumors (Fig. 2). To summarize these results, we first adjusted for the variation in number of tumors evaluated for each patient before calculating the frequency of genomic alterations across the CR patients. BAC clones encompassed by a lost or gained segment were assigned a value of ±1 or 1, respectively; no change was assigned a value of 0. For each loss or gain, we first calculated the average value for each patient and then averaged the resultant fractions across all 14 patients to represent the frequency of a given deviation in this patient set. All frequencies noted in this work are these adjusted frequencies.

To define deviations of interest, we used the one-sample binomial test to calculate a threshold frequency such that for any deviation with frequency exceeding this threshold, the 99% confidence interval of its frequency did not include 0. This frequency is \( \frac{z^2}{(n + z^2)} \), where \( z \) is the critical value (Standard normal table) and \( n \) is the sample size. For \( n \) of 14 and \( P \) value of 0.05, the threshold frequency is 0.05.
Figure 1. Similarity of the tumors of a given patient. Trees from hierarchical clustering of segmented data from (A) the CR tumors and (B) the CR tumors, LocPCs, and normal prostate stromal tissue. In both trees, the patient number precedes the organ description, and numerical suffixes indicate multiple tumors from the same organ. Each color in A indicates tumors from a different patient; in B, the three sample types are each shaded with a different color. The third panel (C) shows heat maps by chromosome of the CBS segment data for each of the seven tumors from patient 10. Red, negative segment values (regions of copy number loss); blue, positive segment values (regions of copy number gain). Note the expected relative loss of X-chromosome sequences and gain of Y chromosome in this male/female comparison. For each chromosome, the vertical black line within each box indicates centromere position, and the large gaps without data indicate unsurveyed repetitive regions. The Y-axis indicates the sample: A, adrenal metastasis; Li, liver metastasis; 1 to 4, the four lymph-node metastases; and S, spleen metastasis. The color bar below the figure indicates the range of colors representing the segment values shown.
<0.01, i.e., $z = 2.56$, the threshold frequency is $\sim 0.3$. The dashed gray lines in Fig. 2 mark this threshold. High-frequency sites of loss or gain exceeding this threshold are listed in Table 1.

Table 1 also provides the locations of the minimally overlapping regions (MOR), which are the most frequently deviant subregions within each set of overlapping deviations. We have identified novel MORs on the two chromosome arms that are thought to play an important role in the biology of prostate cancer, 8p and 8q. Of the four MORs identified for 8p, two have been frequently identified in prostate cancer (23–27), but the more distal MORs in 8p23 have not. We identified four previously reported MORs on 8q (between 8q21.3-q24.22; ref. 23) and three MORs (between 8q11.1-8q21.12) that are heretofore unreported for prostate cancer.

We also examined single BAC clones with extreme log ratio values (i.e., those in the upper and lower 95th percentile) before smoothing by the CBS method. These singleton deviations, when high, might represent focal sites of gene amplification or, when low, homozygous deletions or loss of multiple copies from polyploid tumor cells. BAC clones whose normalized log ratio value was outside the 95 percentile in two or more normal control arrays were excluded from this analysis. Supplementary Table S3 lists the singleton deviations that encompass known cancer-related genes (AR, CDH13, MMP16, MYC, PTEN, and PTK2) and androgen-regulated genes (AR, MMP16, MYC, NDRG1, and TSC22D1).

Genomic profiles of CR tumors and untreated localized primaries are significantly different. To distinguish sites that relate to metastasis or CR disease, we identified those alterations that occur in significantly greater numbers in the CR tumors ($n = 54$) versus the LocPCs ($n = 9$) and vice versa. We found 26 losses spanning a total of 306 Mbp and 8 gains spanning a total of 255 Mbp that were significantly (Fisher’s exact $P < 0.05$) associated with the CR tumors (Table 1). One of the gains encompasses the androgen receptor, an amplification specifically associated with CR disease (28). In the converse analysis, we found only three losses totaling 20 Mbp (5q22.2-q23.1, 11q14.1, and 20q11.23-q13.12) and six gains spanning 69 Mbp (1p36.33-p34.3, 3p21.31-p21.1, 10q21.3-q22.1, 12q13.11-q14.1, 15q15.1-q15.3, and 16q22.1-q22.2) in significantly greater numbers of LocPC than CR tumors.

To verify our findings, we compared our results to a recent survey summarizing 41 CGH studies examining 872 prostate tumors (23). All of the MORs encompassed by the seven alterations that this survey called common (>10%) to primary prostate cancer (loss at 5q15, 6q15, 8p21.3, 13q21.33, 16q22.1, and 18q21.33-22.1 and gain at 8q22.2) were observed in our LocPCs and were in high frequency in our CR tumors ($\geq 30\%$). Of these, our analysis found that the MORs on 16q and 8q were significantly more associated with CR tumors than with LocPCs (Table 1).

Stratification by CR tumor location identifies organ-specific alterations. We stratified the CR tumors to look for differences that might relate to particular states (i.e., CR primaries or CR metastases) or sites (i.e., CR primaries, lymph-node, or liver metastases). We found that the CR primaries possessed significantly fewer gains and losses (average of 14 gains and 18 losses; Student’s $P$ values = 0.008 and 0.004, respectively) than the CR metastases (average of 23 gains and 24 losses).

We used SAM (19) to identify alterations with significantly different frequencies between the CR primaries and metastases. Only two alterations were found, both in significantly more metastases, deletions of 10p15.2-10p15.1 (FDR, 0%, 20% of primaries, and 56% of metastases) and 22q12.1-22q12.3 (FDR, 0%, 27% of primaries, and 80% of metastases). These deletions were also found in significantly fewer LocPCs (11% and 33%, respectively) compared with the entire set of CR tumors. Among the genes encompassed by these loci are the tumor suppressors PRKCQ (10p15.1), MYO18B (20q12.1), and SEZ6L (20q12.1).

11 http://hgsv.washington.edu/
We next asked what differences might exist between the genomic profiles of the CR tumors from the three organ sites that make up 80% of our CR tumors, i.e., prostate (n = 15 from 12 patients), lymph node (n = 19 from 11 patients), and liver (n = 9 from 8 patients). Figure 3A shows the frequencies of alterations for each set. We found that the prostate tumors had significantly fewer gains (average 14; Student’s P = 0.03 and 0.005, respectively) and losses (average 18; Student’s P = 0.02 and 0.04, respectively) than lymph-node (averages 21 and 22, respectively) or liver (averages 30 and 24, respectively) metastases. Lymph-node metastases had significantly fewer gains (Student’s P = 0.04) but not losses compared with liver metastases.

Using SAM, we identified five alterations in significantly more liver metastases, three in more lymph-node metastases, and none in the primary tumors (Fig. 3B). The sites for liver metastases were gains at 5q31.2-5q31.3, 5q35.2-5q35.3, 8p11.21, 11q13.2, and 16p13.3. The gain at 8p encompasses a single gene ANK1. The sites for lymph-node metastases were gains at 6p21.32-6p21.2 and 6p21.1 and loss at 22q12.1. Among the genes encompassed by the gains were the oncogene EV77 (6p21.31) and the tumor-related genes PTK7 (6p21.1; ref. 29) and VEGF (6p21.1; ref. 30).

**Merged copy-number and expression results reveal candidate genes.** Genomic alterations in cancer can encompass many genes. To identify those that might relate to tumor phenotypes, we integrated our array CGH data with microarray analyses of expression levels. We were able to obtain copy-number and expression data for 51 of our CR tumors.

We identified 131 genes that showed correlated array CGH segment and microarray expression values at a significance level of 0.05 (Supplementary Table S4; tumor-related and androgen-regulated genes are highlighted). Supplementary Fig. S2 illustrates the correlation between copy number and expression for the following eight genes. The gene with the third highest correlation value was TMPRSS2. A significant subset of tumors exhibited negative CBS segment values and lower expression (n = 15) or positive CBS segment values and higher expression (n = 19) of TMPRSS2. Other genes identified were the tumor suppressor retinoblastoma 1 (RB1, deleted in 79% of the CR tumors), MYC binding protein 2 (MYCBP2, deleted in 60%), mucin 1 (MUC1, gained in 57%), PBX1 (gained in 49%), PARP1 (gained in 33%), LSM1 (gained in 31%), and TP53 binding protein 2 (TP53BP2, gained in 30%).

To identify genes with significant differences in both copy number and gene expression between primary and metastatic prostate cancer specimens, we used the DNA/RNA-SAMs (DR-SAM) method of DR-Integrator. At an FDR of 5%, we found 19 genes with higher copy number and expression in CR metastases versus CR primaries and 6 with higher copy number and expression in CR primaries versus CR metastases (Supplementary Table S5). Supplementary Fig. S3 illustrates the correlation between selected genes.

**Overrepresentation of ontologically related genes in regions frequently altered in CR tumors.** To further prioritize possible genes of interest, we looked for overrepresentation of genes of particular gene-ontology categories in regions frequently (≥30%, adjusted frequencies) altered in the CR tumors. Lost and gained segments were evaluated separately. Gene ontology categories and genes are given in Supplementary Table S6. The P value “SimPValue” is a statistic that corrects for gene ontology categories that might be overrepresented due to gene clustering. Categories with a SimPValue of <0.05 should be considered with caution.

**Discussion**

Genomic alterations within CR tumors might reveal important biological insights into this ultimately lethal stage of prostate cancer. Our cluster analysis shows that CR tumors from a given patient are more similar to each other than they are to tumors from matching organ sites of other patients. These results argue for a monoclonal origin of metastases, consistent with a recent publication by Liu and colleagues (31) who conclude that most, if not all, metastatic prostate cancers have monoclonal origins. As in our study, Liu and colleagues (31) found that metastatic tumors possess genomic profiles that reflect that of the originating tumor cell.

Androgen deprivation also undoubtedly played a part in generating the similarity of tumors of a given patient. Abrogation of androgen, a hormone with profound effects on tumor biology, places a strong selective pressure on the malignant cell population likely increasing the homogeneity of the tumor population. However, this homogenizing force seems insufficient to generate a
### Table 1. High-frequency deviations and MORs in CR tumors

<table>
<thead>
<tr>
<th>A. Losses</th>
<th>Deviation</th>
<th>MOR</th>
<th>CR-associated</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Band</td>
<td>Start (Mbp)</td>
<td>Size (Mbp)</td>
</tr>
<tr>
<td></td>
<td>1p36.23-p35.1</td>
<td>7.2</td>
<td>25.2</td>
</tr>
<tr>
<td></td>
<td>1p33-p32.3</td>
<td>48.4</td>
<td>6.8</td>
</tr>
<tr>
<td></td>
<td>2q37.3</td>
<td>238.6</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>3p21.31</td>
<td>46.3</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>4p16.3-p16.1</td>
<td>0.1</td>
<td>10.6</td>
</tr>
<tr>
<td></td>
<td>4q33-q34.3</td>
<td>171.7</td>
<td>9.9</td>
</tr>
<tr>
<td></td>
<td>4q34.3-q35.2</td>
<td>182.2</td>
<td>8.2</td>
</tr>
<tr>
<td></td>
<td>5q11.2-q13.2</td>
<td>55.1</td>
<td>15.9</td>
</tr>
<tr>
<td></td>
<td>5q14.1-q22.2</td>
<td>81.7</td>
<td>31.0</td>
</tr>
<tr>
<td></td>
<td>5q23.1-q23.2</td>
<td>115.4</td>
<td>11.8</td>
</tr>
<tr>
<td></td>
<td>5q34-q35.2</td>
<td>160.7</td>
<td>15.6</td>
</tr>
<tr>
<td></td>
<td>6q13-q22.3</td>
<td>75.2</td>
<td>43.7</td>
</tr>
<tr>
<td></td>
<td>8p23.3-p11.21</td>
<td>0.3</td>
<td>40.3</td>
</tr>
<tr>
<td></td>
<td>10p15.3-q26.3</td>
<td>0.2</td>
<td>135.1</td>
</tr>
<tr>
<td></td>
<td>11p15.4</td>
<td>4.2</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>13q11-q34</td>
<td>18.4</td>
<td>95.0</td>
</tr>
<tr>
<td></td>
<td>15q24.3-q25.1</td>
<td>75.1</td>
<td>1.9</td>
</tr>
<tr>
<td></td>
<td>15q25.1-q26.3</td>
<td>78.7</td>
<td>21.3</td>
</tr>
<tr>
<td></td>
<td>16q11.2-q24.3</td>
<td>45.3</td>
<td>43.3</td>
</tr>
<tr>
<td></td>
<td>17p13.3-p11.2</td>
<td>0.1</td>
<td>19.5</td>
</tr>
<tr>
<td></td>
<td>17q24.2-q25.3</td>
<td>64.6</td>
<td>23.2</td>
</tr>
<tr>
<td></td>
<td>17q25.3</td>
<td>74.4</td>
<td>1.6</td>
</tr>
<tr>
<td></td>
<td>18q12.1-q23</td>
<td>28.2</td>
<td>47.9</td>
</tr>
<tr>
<td></td>
<td>19p12-q13.11</td>
<td>19.9</td>
<td>20.3</td>
</tr>
<tr>
<td></td>
<td>19q13.2-q13.31</td>
<td>46.1</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>21q22.2-q23.2</td>
<td>38.7</td>
<td>4.5</td>
</tr>
<tr>
<td></td>
<td>22q11.1-q13.33</td>
<td>16.2</td>
<td>33.2</td>
</tr>
<tr>
<td></td>
<td>22q13.33-q23.3</td>
<td>16.2</td>
<td>33.2</td>
</tr>
</tbody>
</table>

(Continued on the following page)
single CR genomic signature, as the tumor sets are sufficiently distinct to cluster in patient-specific groups. These differences must reflect the uniqueness of each of our CR patients in terms of his genetic risk factors, environmental exposures, history of therapeutic modalities, and chance events.

Across patients, we find a multitude of high-frequency alterations, some encompassing candidate genes that might relate to prostate cancer, metastases, and CR disease. Within these alterations were several cancer-related genes with correlated DNA and mRNA values, including \(RB1\), the earliest recognized tumor suppressor, and \(MUC1\), a marker of prostate cancer progression and a novel therapeutic target (32).

\(PARP1\) was also identified; it is involved in DNA repair and inhibitors of it have received considerable attention as novel therapy of breast cancer (33).

**Table 1. High-frequency deviations and MORs in CR tumors (Cont’d)**

<table>
<thead>
<tr>
<th>Deviation</th>
<th>MOR (%)</th>
<th>CR-associated</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Band</td>
<td>Start (Mbp)</td>
<td>Size (Mbp)</td>
<td>Start (Mbp)</td>
<td>Size (Mbp)</td>
</tr>
<tr>
<td>1p12-q43</td>
<td>120.0</td>
<td>117.7</td>
<td>65</td>
<td>147.9</td>
</tr>
<tr>
<td>1q11-43</td>
<td>215.6</td>
<td>22.1</td>
<td>47</td>
<td>5.5</td>
</tr>
<tr>
<td>2p25.1</td>
<td>9.6</td>
<td>1.8</td>
<td>45</td>
<td>26.8</td>
</tr>
<tr>
<td>5q31.3</td>
<td>140.1</td>
<td>0.8</td>
<td>47</td>
<td>62.0</td>
</tr>
<tr>
<td>7p22.3-q36.3</td>
<td>0.1</td>
<td>158.5</td>
<td>55</td>
<td>72.6</td>
</tr>
<tr>
<td>8p12-q24.3</td>
<td>36.5</td>
<td>109.6</td>
<td>64</td>
<td>43.3</td>
</tr>
<tr>
<td>9q33.2</td>
<td>122.4</td>
<td>1.1</td>
<td>69</td>
<td>57.2</td>
</tr>
<tr>
<td>9q32.2-q33.3</td>
<td>124.3</td>
<td>3.8</td>
<td>81</td>
<td>79.3</td>
</tr>
<tr>
<td>9q34.11</td>
<td>129.5</td>
<td>0.8</td>
<td>82</td>
<td>93.3</td>
</tr>
<tr>
<td>11q12.1-q24.1</td>
<td>59.4</td>
<td>63.0</td>
<td>79</td>
<td>102.3</td>
</tr>
<tr>
<td>11q23.3-q24.1</td>
<td>119.0</td>
<td>3.4</td>
<td>78</td>
<td>115.1</td>
</tr>
<tr>
<td>12q24.31-q24.32</td>
<td>119.8</td>
<td>4.8</td>
<td>73</td>
<td>131.1</td>
</tr>
<tr>
<td>16p13.3</td>
<td>0.9</td>
<td>2.4</td>
<td>119.0</td>
<td>3.4</td>
</tr>
<tr>
<td>16p12.1-p11.1</td>
<td>27.6</td>
<td>6.9</td>
<td>32</td>
<td>123.1</td>
</tr>
<tr>
<td>18p11.32</td>
<td>0.2</td>
<td>2.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xp22.31-p22.12</td>
<td>7.5</td>
<td>11.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xp11.22-q13.1</td>
<td>51.5</td>
<td>16.5</td>
<td>50</td>
<td>65.1</td>
</tr>
</tbody>
</table>

**NOTE:** Losses (A) and gains (B) are listed separately. High frequency is defined as observed with an adjusted frequency of \(\geq 30\). The start position, size, and any constituent MORs observed at higher frequency than the larger deviation are given for each deviation. Mbp positions are rounded to nearest 0.1 Mbp. MORs observed at \(\geq 50\) are in bold. The start site and size of CR-associated alterations are given. Abbreviation: MOR (%), the peak frequency that defines the MOR.
We found that majority of our CR tumors possessed a fusion between TMPRSS2 and ERG as a result of the deletion. The high frequency of CR tumors with multiple copies of the for TMPRSS2:ERG fusion is further evidence of the association of amplification of the fusion gene and poor clinical outcome (5). These findings support the idea that it is a promising target for therapeutic interventions (34, 35). Given that the deletion that generates the TMPRSS2:ERG fusion was found to encompass the majority of the TMPRSS2 gene, it was not surprising that we found a significant association between the fusion and negative array CGH and expression values for TMPRSS2. However, a subset of fusion-positive tumors showed higher expression of TMPRSS2. This finding may indicate that expression from an intact copy of this androgen-regulated gene might be biologically relevant for some prostate tumors.

Our analysis of high-frequency alterations in CR tumors helps refine prostate cancer-related loci and narrow in on additional candidate genes. No consensus has yet been reached about the critical locus (loci) affected by the two most common alterations seen in prostate cancer, loss at 8p and gain at 8q (28). Our study identifies

![Figure 3. Deviations in CR tumors stratified by organ of origin.](image)
four distinct MORs within 8p and seven MORs within 8q. Of the less well-characterized MORs on 8p, one of them encompasses a single gene, CSMD1, a candidate suppressor of multiple cancer types (36). EGER (a.k.a. Lyric and MTDH) is among the genes encompassed by an MOR at 8q21.3-q22.2. AEG1 is overexpressed in breast, brain, and prostate cancers (37–39) and is thought to promote tumor progression (38–40). The novel MOR at 8q21.12 contains a single gene, PKIA, an extremely potent competitive inhibitor of cyclic AMP–dependent protein kinase activity (41). The role of PKIA in prostate tumorigenesis merits exploration.

The MORs in our CR data set might help identify other genes relevant to tumorigenesis. The most frequent MOR in our CR tumors was loss at 10q23.31 (86%), which encompasses only PTEN, the well-characterized tumor suppressor. Twelve known genes are encompassed by the MOR at 13q14.13-q14.2 (85%), including the ITM2B gene, a tumor suppressor (42). Loss at 16q23.3 (82%) encompasses only CDH13 whose reduced expression in primary prostate tumors is associated with an increased risk of biochemical failure (43). Methylation of CDH13, assessed in primary prostate cancer, is generally considered the primary mechanism of gene silencing (44, 45). Our results suggest that methylation in premetastatic states might precede deletion in later stages or that deletion is an alternate mechanism of silencing.

Our gene ontology analysis of the high-frequency alterations in CR tumors also provides insight into candidate genes. The genes that overlapped between our gene ontology analysis and the integration of copy number and expression warrant particular attention. One of these genes encompassed by CR deletion encodes BNIP3, a Bcl-2 family member that can promote apoptosis (46). RNF6 and USP10, which modulate androgen receptor function (47, 48), showed correlated genomic loss and lower expression in a subset of tumors and gain and higher expression in others. This finding underscores the complexity of the relationship between CR disease and tumor processes controlled by the androgen receptor.

Our analysis of CR tumors provides evidence of accumulation of genomic change with disease progression and outgrowth toward CR disease. We found significantly fewer alterations in CR primaries than in CR metastases and no alterations significantly more associated with CR primaries relative to the matched set of lymph-node and liver metastases. However, we did find eight alterations significantly more associated with those metastatic sites (three and five, respectively). Moreover, we identified several alterations significantly more associated with CR tumors versus LocPCs.

To expand on the significance of our analysis, future work will need to include bone metastasis, a clinically important entity in prostate cancer, and validation of candidate genes. For example, it would be interesting to see the effect of PARP1 inhibitors on models of prostate cancer. The androgen-dependent LNCaP cells and the androgen-independent derivative line (49) would be particularly useful in assessing the role of genes encompassed by CR-associated alterations identified in our study. Direct injection methods for studying prostate cancer cell–bone interactions and the effects that drugs have on these interactions could be used to validate candidate genes and their therapeutic potential (50).

We have shown that CR tumors possess a profound degree of genomic change encompassing many regions that could contain therapeutic targets for metastatic prostate cancer, or at least illuminate the biology of this lethal disease. By combining the results of array CGH and expression microarrays, we have identified numerous candidate regions and genes. Moreover, we have verified that the TMPRSS2:ERG fusion, a promising target of cancer therapeutics, is highly prevalent in CR tumors. This extensive and in-depth investigation of the alterations found in

---

**Figure 4.** Experimental designs to detect the presence of the TMPRSS2:ERG fusion by FISH. In experiment 1, three probes were used to detect the fusion: a probe 5’ of TMPRSS2 (blue, probe 1), one encompassing the 5’ exons of ERG (red, probe 2), and one encompassing the 3’ exons of ERG (green, probe 3). In the normal configuration, the signals of all three probes overlap. The fusion is indicated by overlapping signals of probes 1 and 3 with loss or dissociation of probe 2. A CR tumor nucleus with two fusions and one normal probe configuration is shown to the right. In experiment 2, we confirmed the presence of fusion with probe 1 and a second probe encompassing the 3’ exons of TMPRSS2 (red, probe 4) in adjacent tissue sections. Lone probe 1 signals indicate a deletion consistent with a TMPRSS2:ERG fusion. Right, a positive nucleus from a section adjacent to the section used to capture the nucleus shown for experiment 1. Gray, 4’,6-diamidino-2-phenylindole staining; the hybridization signals are pseudocolored to correspond to the experimental schematics.
CR disease lays the foundation for a better understanding of this final stage of prostate cancer.

Disclosure of Potential Conflicts of Interest

The authors have no potential conflicts of interest.

Acknowledgments

References


Comparative Analyses of Chromosome Alterations in Soft-Tissue Metastases within and across Patients with Castration-Resistant Prostate Cancer

Ilona N. Holcomb, Janet M. Young, Ilsa M. Coleman, et al.

Cancer Res  Published OnlineFirst September 22, 2009.