Neoplasia • Vol. 6, No. 3, May/June 2004, pp. 240–247 240

www.neoplasia.com

High-Resolution Analysis of Gene Copy Number Alterations in Human Prostate Cancer Using CGH on cDNA Microarrays: Impact of Copy Number on Gene Expression[1]

Maija Wolf, Spyro Mousses†, Sampsa Hautaniemi‡, Ritva Karhu§, Pia Huusko†, Minna Allinen¶, Abdel Elkahloun¶, Outi Monni#, Yidong Chen¶, Anne Kallioniemi§ and Olli-P Kallioniemi*,†*

*Medical Biotechnology, VTT Technical Research Centre of Finland and University of Turku, Turku FIN-20520, Finland; †Translational Genomics Research Institute, Gaithersburg, MD 20878-1762, USA; ‡Institute of Signal Processing, Tampere University of Technology, Tampere FIN-33101, Finland; §Laboratory of Cancer Genetics, Tampere University Hospital and Institute of Medical Technology, University of Tampere, Tampere FIN-33520, Finland; ¶Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892, USA; #Biomedicum Biochip Center, University of Helsinki, Helsinki FIN-00014, Finland

Abstract

Identification of target genes for genetic rearrangements in prostate cancer and the impact of copy number changes on gene expression are currently not well understood. Here, we applied high-resolution comparative genomic hybridization (CGH) on cDNA microarrays for analysis of prostate cancer cell lines. CGH microarrays identified most of the alterations detected by classical chromosomal CGH, as well as a number of previously unreported alterations. Specific recurrent regions of gain (28) and loss (18) were found, and their boundaries defined with sub-megabasepair accuracy. The most common changes included copy number decreases at 13q, and gains at 1q and 5p. Refined mapping identified several sites, such as at 13q (33–44, 49–51, and 74–76 Mbp from the p-telomere), which matched with minimal regions of loss seen in extensive loss of heterozygosity mapping studies of large numbers of tumors. Previously unreported recurrent changes were found at 2p, 2q, 3p, and 17q (losses), and at 3q, 5p, and 6p (gains). Integration of genomic and transcriptomic data revealed the role of individual candidate target genes for genomic alterations as well as a highly significant ($P < .0001$) overall association between copy number levels and the percentage of differentially expressed genes. Across the genome, the overall impact of copy number on gene expression levels was, to a large extent, attributable to low-level gains and losses of copy number, corresponding to common deletions and gains of often large chromosomal regions.

Neoplasia (2004) **6**, 240–247

Keywords: Copy number alteration, prostate cancer, gene expression, cDNA microarray, CGH microarray.

Introduction

Despite the high prevalence of prostate cancer, its molecular basis continues to be poorly understood. Gene microarray technologies have provided a new means for exploring the transcriptome of prostate cancer cells. This information has been applied to distinguish patterns and gene expression fingerprints for benign prostate hyperplasia, to identify subtypes of prostate cancer, and to identify markers for diagnosis and candidate targets for therapy [1,2]. In addition to the gene expression, gene copy number changes have been shown to be involved in the development and progression of prostate cancer [3–6]. However, most of the target genes for the genetic alterations remain unknown and their impact on global gene expression profiles remains unclear [7]. Cancer cell lines are widely used as models for tumorigenesis and progression of prostate cancer, and their genomic and transcriptomic profiling is relevant to understanding the molecular mechanisms underlying the disease.

Here, we applied cDNA microarrays to survey changes in DNA copy number throughout the genomes of four cell lines (PC-3, DU 145, LNCaP, and CWR22R) established from advanced prostate cancer. Gains and losses throughout the cancer genome were examined with a cDNA microarray having an average clone spacing of 280 kb, and the results were compared and validated with data from chromosomal comparative genomic hybridization (CGH) and locus-specific fluorescence *in situ* hybridization (FISH) analyses. We also sought to define the boundaries of common copy number alterations and to determine the impact of these alterations on gene deregulation in prostate cancer.

Abbreviations: CGH, comparative genomic hybridization; FISH, fluorescence *in situ* hybridization; BAC, bacterial artificial chromosome

Address all correspondence to: Olli-P Kallioniemi, Medical Biotechnology, VTT Technical Research Centre of Finland and University of Turku, Itäinen Pitkäkatu 4, Turku FIN-20520, Finland. E-mail: olli.kallioniemi@vtt.fi

[1]This study was partly supported by the Academy of Finland, Sigrid Juselius Foundation, and Finnish Cancer Society.

Received 13 November 2003; Revised 23 December 2003; Accepted 7 January 2004.

DOI 10.1593/neo.03439

Materials and Methods

Prostate Cancer Cell Lines and Sample Preparation

PC-3, DU 145, and LNCaP cell lines were obtained from the American Type Culture Collection (ATCC, Manassas, VA). CWR22R cells were kindly provided by Dr. Jim Jacobberger's Laboratory at Case Western Reserve University (Cleveland, OH). Cell culture conditions were as recommended by the supplier. For each cell line, at least four T175 flasks were pooled to obtain enough material. PolyA-RNA was isolated from the samples using FastTrack 2.0 mRNA isolation kit (Invitrogen, Carlsbad, CA). After adding the lysis buffer, genomic DNA was spooled from the same sample by swirling a glass rod in the cell lysate. After spooling, DNA samples were subjected to phenol/chloroform purification.

cDNA Microarray Preparation

A customized 16K cDNA microarray was used for the detection of copy number and expression level changes. In addition to a genome-wide coverage of 11,500 clones, 4700 clones were included from regions at 2p23–p25, 5p, 8q, 10q21–q24, 11q12–q14, 12q13–q15, 17q11.2–q23, 20q, and Xcen-q13, representing common amplification sites in human cancers. Preparation and printing of the clones on glass slides were performed as described in Ref. [8]. On average across the genome, the 11,500 clones correspond to a resolution of 3.6 clones per megabasepair (or one clone every 280 kb) with, however, much greater density at the nine amplified regions due to the set of 4700 specific clones targeting these regions.

CGH and Expression Analysis on cDNA Microarrays

CGH on cDNA arrays was performed as described earlier in Ref. [9] with slight modifications [10]. Briefly, 20 μg of genomic DNA was digested overnight using *Alu*I and *Rsa*I (Life Technologies, Inc., Rockville, MD). Gender-matched DNA obtained from white blood cells of healthy individuals was used as a reference for CGH array hybridizations. Digested DNA samples were purified by phenol/chloroform extraction. Six micrograms of digested cell line DNA and reference DNA was labeled with Cy5-dUTP and Cy3-dUTP (Amersham Biosciences, Piscataway, NJ), respectively, in a random priming reaction using Bioprime Labeling kit (Life Technologies, Inc.). Hybridization and washes were performed as described earlier [9]. For cDNA microarray hybridizations, 3 μg of test mRNA and 3 μg of reference sample, composed of a pool from 18 different cancer cell lines, were labeled with Cy5-dUTP and Cy3-dUTP, respectively, and hybridizations were done as described in Ref. [10]. A laser confocal scanner (Agilent Technologies, Palo Alto, CA) was used to measure signal intensities from targets for both CGH and cDNA microarray hybridizations.

Microarray Data Analysis

DeArray software (Scananalytics Inc., Fairfax, VA) was used in image analysis [11]. Average intensities of the tumor samples were divided by the average intensities of the reference sample at each microarray spot after background intensity subtraction. Within-slide normalization for both cDNA and CGH data was done using ratio statistics [11], followed by mean centering across the genes. After normalization, both cDNA and CGH data were filtered using ratio quality values [12] below 0.5 as a threshold for exclusion. This quality cutpoint value has, in the past, been shown to represent less reliable cDNA microarray measurements due to either low signal intensity, high local background level, uneven distribution of the target intensity, and/or small target size.

A custom-made database was created including the genomic sequence alignment information for all available mRNA sequences according to the assembly by the University of California Santa Cruz's (UCSC) Genome Browser database (http://genome.ucsc.edu/) (December 2001 freeze), as well as the Unigene information obtained from Build 146 (http://ncbi.nih.gov/). Genomic basepair localizations of the clones were retrieved by assigning each clone to its Unigene cluster and then relating the data to the genomic alignment of the mRNA sequence according to the largest alignment size from the corresponding cluster. To visualize the whole genome coverage of copy number data, we plotted the clone-derived ratio data as a function of position along the genome sequence for each clone individually (maximal theoretical resolution ~300 kb, depending on the region and density of the clones). We also plotted moving averaged ratios of adjacent clones with unique genomic positions to improve the signal-to-background ratio. The figures shown were plotted using 30-clone moving average ratios.

Regions of gain and loss were defined as loci where at least two adjacent clones had a moving averaged CGH ratio $\geq$1.05 (gain/amplification) or <0.95 (deletion). These threshold values for altered copy number were determined based on the normal variation in the control hybridizations; 98.5% of the 30-clone averaged copy number ratios in a self-*versus*-self experiment varied between ratios 0.95 and 1.05. The breakpoints for the changes that were scored to be true genetic alterations were then defined using individual clone data or five-clone averaging. Integration of the transcriptomic data with the genomic information was achieved by annotating the gene expression data on copy number ratios using color coding—red for upregulated genes (global upper 7% of expression ratios) and green for downregulated genes (global lower 7% of expression ratios). Statistical significance of the impact of gene copy number on gene expression was determined with chi-square analysis.

Chromosomal CGH

To evaluate the sensitivity and reliability of the array-based CGH, the results were compared with the profiles obtained from chromosomal CGH [13]. Cell line DNA samples representing the same batch of DNA extraction were used for chromosomal and array CGH. The DNA samples were labeled with FITC-12-dUTP (DuPont, Boston, MA) and sex-matched normal reference DNA with Texas Red-6-dUTP (DuPont). Six hundred nanograms of labeled cell line DNA, 400 ng of reference DNA, and 10 μg of unlabeled Cot-1 DNA (Gibco BRL, Gaithersburg, MD) were hybridized to

normal lymphocyte metaphase chromosomes. After overnight hybridization and washing, the chromosomes were counterstained with 0.5 M 4′,6-diamidino-2-phenylindole (DAPI) in an antifade solution. Fluorescence intensities were quantitated using an Olympus BX 50 epifluorescence microscope (Olympus, Tokyo, Japan) and the Quips digital image analysis system (Vysis, Inc., Downers Grove, IL). Based on hybridization results from control samples (normal male hybridized against normal female), chromosomal regions with copy number ratios under 0.85 were considered as lost, and above 1.2 as gained, in the chromosomal CGH analysis.

Interphase FISH Analysis

We validated copy number data from array CGH with gene-specific FISH analyses at 1q with prominent copy number alterations. Dual-color interphase FISH was done for PC-3 using bacterial artificial chromosome (BAC) clones RP11-119A2, RP11-79I7, and RP11-101I9 mapping to 1q24.3 (173 Mbp from the p-telomere), 1q25.3 (187 Mbp), and 1q32 (205 Mbp), respectively. Clones were labeled with SpectrumOrange (Vysis, Inc.) using random priming. SpectrumGreen-labeled 1p telomeric probe was used as reference. DAPI staining was performed and the mean number of red and green signals was determined.

Results

Overviews of Copy Number Alterations by Array CGH and Comparison with Chromosomal CGH

CGH microarray analysis revealed numerous recurrent copy number changes in prostate cancer (Figure 1*A*,

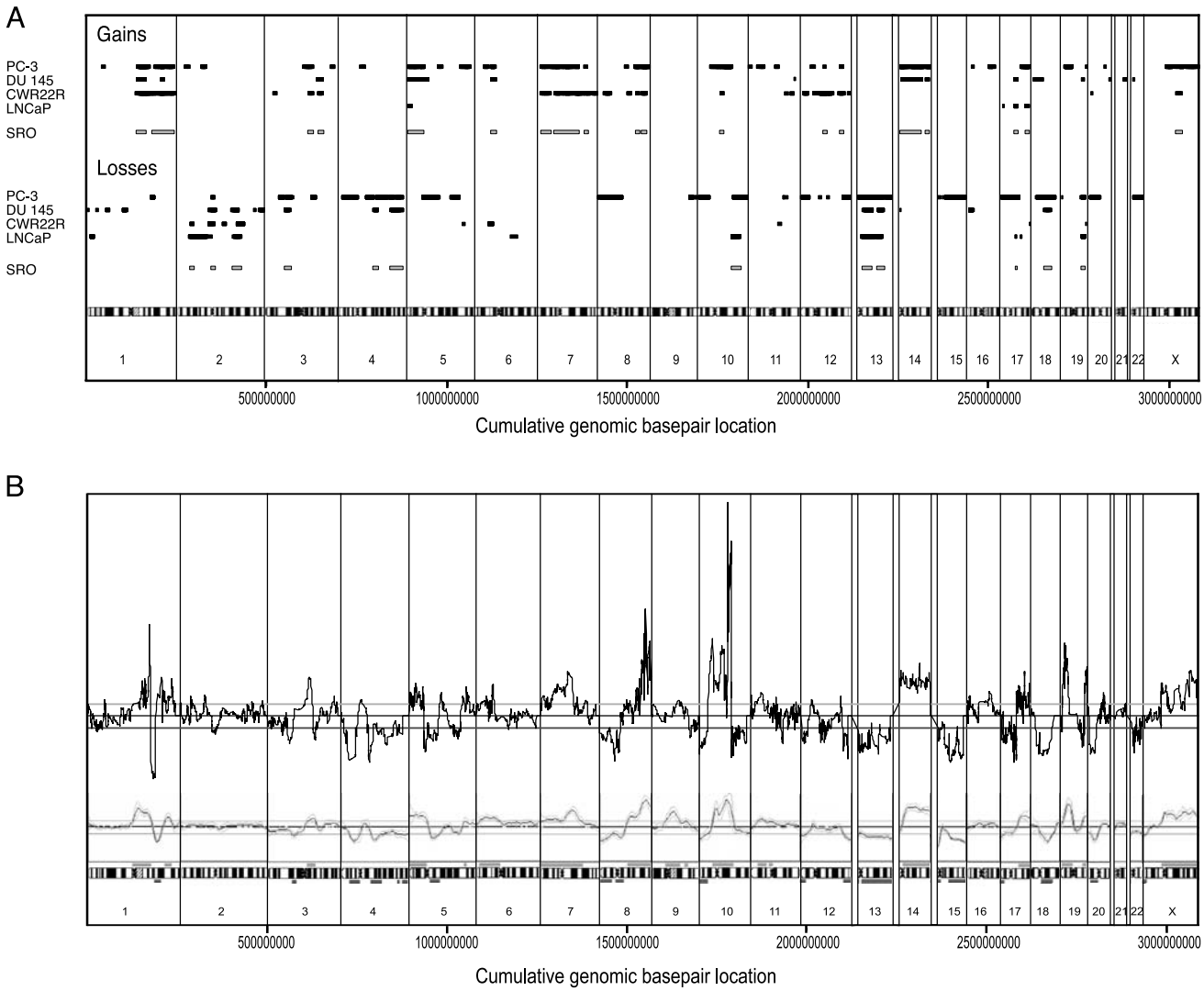

Figure 1. *Genome-wide copy number analysis of prostate cancer cell lines. (A) Copy number alterations in each cell line (black bars) and minimal common regions of overlap (SRO) for gains and losses detected in at least two samples (gray bars) are indicated. Vertical bars designate chromosomal borders. Cumulative genomic basepairs were calculated according to chromosome sizes (UCSC; December freeze 2001). Overall, hormone-refractory cell lines PC-3, DU 145, and CWR22R manifested more unbalanced chromosomal changes than the hormone-responsive cell line LNCaP. (B) Alignment of chromosomal (bottom) and microarray-based CGH (top) displaying the genome-wide copy number profile for PC-3. For array-based profile, averaged ratios from 30 adjacent clones (y-axis) were plotted as a function of the position along the human genome (x-axis).*

Table 1. Minimal Common Regions of Overlap for Gains and Losses Detected in at Least Two Samples.

Chromosome	Distance from p-Telomere (Mbp)	Cytogenetic Region	PC-3	CWR22R	LNCaP	DU 145
(A) Minimal common areas of loss						
2	38–44	2p22.2–p21		X	X	
2	93–95	2q11.1–q11.2		X		X
2	156–160	2q24.1–q24.2			X	X
2	162–174	2q24.2–q31.1		X	X	
3	57–89	3p21.1–p11.1	X			X
4	155–157	4q31.1–q32.1	X			X
4	174–188	4q34.1–q35.2	X			X
10	93–95	10q23.2–q23.31	X		X	
10	107–108	10q24.2–q24.31	X		X	
13	33–44	13q13.3–q14.11	X		X	X
13	44–47	13q14.11–q14.2	X		X	
13	49–51	13q14.2–q14.3	X		X	X
13	74–76	13q22.1–q22.2	X		X	X
13	78–94	13q22.3–q32.1	X			X
17	36–37	17q12–q21.1	X		X	
18	36–47	18q12.2–q21.1	X			X
19	57–62	19q13.2–q13.32	X		X	
19	65–66	19q13.33	X		X	
(B) Minimal common areas of gain						
1	147–154	1q21.2–q22	X	X		X
1	156–161	1q23.1–q23.2	X	X		X
1	161–174	1q23.3–q24.3	X	X		
1	201–206	1q31.3–q32.1	X	X		
1	207–213	1q32.1–q32.2	X	X		
1	222–239	1q32.3–q42.13	X	X		
3	123–127	3q13.33–q21.1	X	X		
3	151–158	3q24 –q25.2		X		X
5	1–4	5p15.33	X		X	X
5	8–11	5p15.31–p15.2	X		X	X
5	17–29	5p15.1–p13.3	X			X
5	37–40	5p13.1–p12	X			X
6	47–53	6p21.1–p12.3	X			X
7	8–11	7p21.3	X	X		
7	18–25	7p21.1–p15.3	X	X		
7	46–92	7p13–p21.13	X	X		
7	101–104	7q22.1	X	X		
8	104–105	8q22.3	X	X		
8	123–125	8q24.12–q24.13	X	X		
10	59–60	10q21.1	X	X		
12	62	12q14.1	X	X		
12	108–115	12q23.2–q24.11	X	X		
14	25–51	14q12–q22.1	X			X
14	53–77	14q22.2–q24.3	X			X
14	93–101	14q32.13–q32.31	X			X
17	35	17q12			X	X
17	65–66	17q23.2–q23.3	X		X	
X	99–104	Xq22.2–q23	X	X		

Table 1). Of the four cell lines, PC-3 (hormone-refractory cell line), manifested the most copy number changes (total of 97, with 46 gains and 51 losses) including several high-level amplifications and deletions (Figure 1*B*), while LNCaP (hormone-responsive cell line) had the least number of alterations (4 gains and 14 losses). Comparison of the copy number profiles showed that 92% of the amplifications and 82% of the deletions detected by chromosomal CGH were also detected by array-based CGH. In addition, multiple alterations not detected using chromosomal CGH were identified by the microarray-based method. For example, deletions at chromosomes 5 (119–138 Mbp from the p-telomere, at 5q23.1–q31.1), 10 (89–142 Mbp, 10q23.1–q26.3), 17 (8–43 Mbp, 17p12–q21.2), and 19 (59–66 Mbp, 19q13.2–q13.33) in the PC-3 cell line were only detectable using high-resolution array CGH (Figure 1*B*). Aberrations close to telomeres and pericentromeric regions were also significantly better visualized by array-based, rather than conventional, CGH analysis.

Defining Critical Regions for Losses and Gains at High Resolution by Array CGH

The most commonly deleted region in the cell lines was 13q. Three separate minimal regions of deletions were mapped along 13q. These sites mapped to 33 to 44 Mbp (q13.3–q14.11), 49 to 51 Mbp (q14.2–q14.3), and 74 to 76 Mbp (q22.1–q22.2) from the p-telomere (Figure 2). These regions match with the previously found minimal loci for loss of heterozygosity (LOH) obtained from extensive analyses of large numbers of primary tumors [14–16]. This illustrates that analysis of a few informative cancer cell line

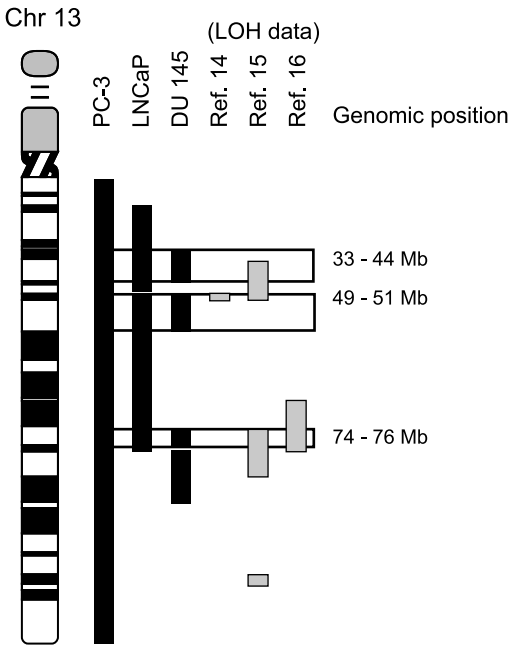

Figure 2. *Smallest region of overlap for deletions at 13q detected by CGH on cDNA microarrays. Three distinct deletions common for PC-3, LNCaP, and DU 145 cell lines were detected at 13q13.3–q14.11, 13q14.2–q14.3, and 13q22.1–q22.2. Genomic positions are indicated as megabasepairs from the p-telomere. Concordance with previous results from prostate cancer LOH studies [14–16] is shown.*

samples by array CGH can substantially narrow down the regions of interest that are also relevant in clinical tumor specimens. Other common copy number decreases seen in half or more of the samples were mapped to specific regions at 2p, 2q, 3p, 4q, 10q, 17q, 18q, and 19q (Figure 1*A*, Table 1). Of these, previously unreported subtle alterations included losses at 2p (38–44 Mbp, p22.2–p21), 2q (93–95 Mbp, q11.1–q11.2), 3p (57–89 Mbp, p21.2–p11.1), and 17q (36–37 Mbp, q12–q21.1).

The most common copy number increases were found at 1q (147–154 and 156–161 Mbp from pter, at 1q21.2–q22 and 1q23.1–q23.2) and 5p (1–4 and 8–11 Mbp from pter, at p15.33 and p15.31–p15.2). Other recurrent copy number increases were mapped at high resolution to specific regions at 3q, 6p, 7p, 7q, 8q, 10q, 12q, 14q, 17q, and Xq (Figure 1*A*, Table 1). Novel findings include gains at 3q (123–127 Mbp, q13.33–q21.1), 5p (1–4, 8–11, and 17–29 Mbp at p15.1–p13.3 and 37–40 Mbp at p13.1–p12), and 6p (47–53 Mbp, p21.1–p12.3).

In the PC-3 cell line, two distinct amplification sites were located on chromosome 1 at 170 to 175 Mbp (1q24.2–q25.1) and 204 to 206 Mbp (1q32.1) from the p-telomere, with a deleted region in between (Figure 3). These changes were also seen by chromosomal CGH; however, the demarcation of the amplified and deleted regions could not be visualized well. We used dual-color interphase FISH to validate the copy number changes with BAC probes specific to the three regions and a 1pter reference probe. The average CGH and FISH ratios (test *vs* 1pter reference probe count) agreed well with one another.

Impact of Copy Number Alterations on Gene Expression Levels

The application of cDNA microarrays as a platform for CGH makes it possible to directly integrate information on gene copy number (DNA) and expression (RNA) levels. Across the genome, gene expression levels were significantly influenced by copy number changes ($P < .0001$) (Figure 4*A*). For these comparisons, we considered genes to be overexpressed, if their normalized cDNA ratios exceeded the 93rd percentile of all expression ratios, and downregulated, if their cDNA ratios were below the lower seventh percentile. Using 1.5 as a copy number cutpoint for DNA amplification, 19.6% of the amplified genes showed a high expression ratio. Because 6.9% of the genes with normal copy number were upregulated, increased copy number led to a 2.8-fold higher number of highly expressed genes than expected. Using 0.8 as a copy number cutpoint for a deletion, 12.1% of the genes with low copy number were significantly downregulated, corresponding to 1.8-fold higher percentage of genes than expected.

Integration of copy number and gene expression data also facilitated the direct identification of candidate target

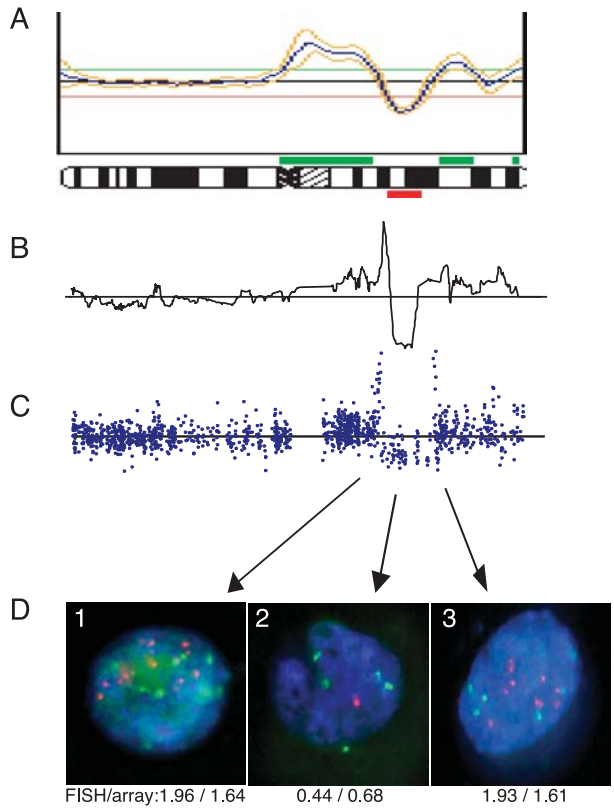

Figure 3. *Validation of the copy number changes at 1q in PC-3. (A) Copy number profile for chromosome 1 by chromosomal CGH displaying areas of gain (green vertical bars) and loss (red vertical bar) in the long arm. (B and C) Copy number profiles by CGH microarray, with moving average ratios from 30 clones (B) and actual ratios from individual clones (C). (D) Interphase FISH was performed to validate the copy number changes with BAC probes specific for q24.3 (RP11-119A2, image 1), q25.3 (RP11-79I7, image 2), and q32 (RP11-101I9, image 3) (red). Copy number ratios were determined using a reference probe from 1pter (green). The average CGH microarray ratios of cDNA clones matching the genomic localization of the BAC probes and the corresponding FISH ratios are displayed under the FISH images.*

genes for the genetic alterations. For example, several genes in the amplicon at 10q22.1–q23.1 (77–89 Mbp from the p-telomere), including annexins A7 and A11 (*ANXA7* and *ANXA11*), urokinase plasminogen activator (*PLAU*), and vinculin (*VCL*), were significantly overexpressed (Figure 4*B*). Similarly, genes with elevated expression levels could be identified from an amplified region at 14q12–q22.1 (at 27–51 Mbp from the p-telomere) in DU 145. These genes included thyroid transcription factor 1 (*TITF1*) and mitogen-activated protein kinase kinase kinase kinase 5 (*MAP4K5*) (Figure 4*C*).

Discussion

We showed that CGH analysis on cDNA microarrays makes it possible to achieve reliable high-resolution mapping of unbalanced genetic alterations in prostate cancer, including both high-level amplifications, but also low-level gains and losses as validated here by comparisons with chromosomal CGH and interphase FISH. As compared to chromosomal CGH, the cDNA microarray-based CGH improved the sensitivity and resolution of copy number detection so that more than two-fold higher number of alterations could be found. For example, complex copy number alterations over short distances were often only detectable by array CGH, including deletions that were located in close proximity to amplified regions, such as deletions at 5q (119–138 Mbp from p-telomere), 10q (89–142 Mbp), 17 (8–43 Mbp), and 19q (59–66 Mbp) in the PC-3 cell line (Figure 1*B*). Besides an overall low resolution (10–20 Mbp), chromosomal CGH is particularly error-prone in detecting alterations in the pericentromeric and telomeric regions of the genome. CGH

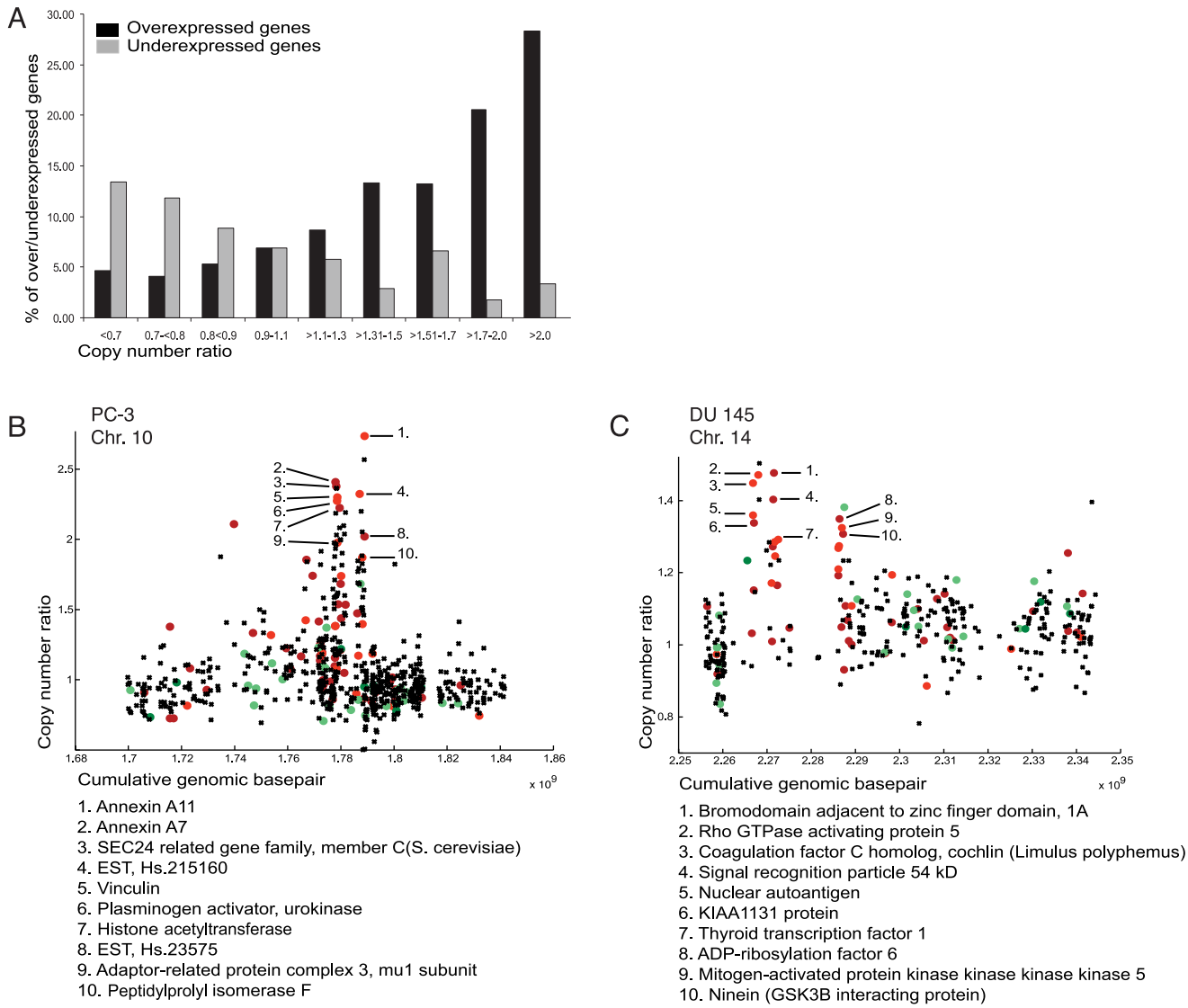

Figure 4. *Impact of copy number on gene expression in prostate cancer cell lines. (A) Impact of gene copy number (x-axis) on the percentage of genes (y-axis) whose expression was high (black bars) or low (gray bars) as a function of copy number alteration. High-level expression was defined as the top seventh percentile of all expression ratios, whereas low-level expression was defined as the bottom seventh percentile. (B and C) Expression-annotated copy number profiles of amplicons at 10q22–q23.1 in PC-3 (B) and 14q13.2–q21.3 in DU 145 (C). Color coding was used to indicate the level of expression as follows: bright red dots indicate the upper two percentiles and dark red dots indicate the next five percentiles of the gene expression ratios in the analyzed cells (overexpressed genes), with green color referencing to downregulated genes. Ten overexpressed and amplified genes are highlighted in both amplicons.*

microarray does not suffer from these limitations, and enables accurate detection of copy number at such regions.

CGH microarray provided an ability to accurately identify and map the boundaries of copy number changes. For example, at 13q, we found three distinct regions of loss, common to PC-3, LNCaP, and DU 145, residing at 33 to 44 Mbp (q13.3–q14.11), 49 to 51 Mbp (q14.2–q14.3), and 74 to 76 Mbp (q22.1–q22.2) from the p-telomere. These regions have been detected previously and narrowed down with extensive allelic loss measurements, analyzing one locus at a time [14–16]. This highlights the power of the CGH microarray technology in rapidly identifying and narrowing down regions of importance for tumor progression. Other similar recurrent deletions that deserve further study in large panels of tumors were found at specific regions at 2p (38–44 Mbp), 2q (93–95 Mbp), 3p (57–89 Mbp), and 17q (36–37 Mbp).

The resolution of the array CGH is limited only by the density of the clones on the array, which in our study was one clone every 280 kb. This is substantially higher than in most BAC clone–based CGH microarray studies. However, using cDNA and other small clones as a target, results from individual clones may not be accurate in detecting low copy number gains and losses. Therefore, analysis of moving mean ratios from adjacent clones and establishment of threshold values from control experiments were necessary to reliably score copy number alterations.

Application of cDNA microarrays in CGH is particularly valuable in that it makes it possible to readily integrate copy number and gene expression data across the genome. We showed how this can be used to directly identify genes that are activated or silenced by specific genetic alterations. We found several amplified and highly upregulated genes that may be critical for tumor progression. For example, at 10q22.1–q23.1, there were many candidate genes that were both amplified and highly upregulated, including annexins A7 and A11 (calcium-dependent phospholipid-binding proteins previously implicated as downregulated in cancer) [17,18], urokinase plasminogen activator (a protein known to play a role in invasion and metastasis) [19–21], as well as vinculin (a cytoskeletal protein). Furthermore, *TITF1* and *MAP4K5* were overexpressed in the amplicon at 14q12–22.1. These genes provide prioritized candidate genes for further evaluation from clinical samples.

Across the genome, the impact of gene copy number on expression levels was statistically highly significant ($P < .0001$), both for regions with increased and decreased copy numbers. This validates the importance of DNA copy number changes in the development and progression of prostate cancer, as shown previously in breast cancer [22,23]. Even low-level gains and losses of copy number, corresponding to common deletions and gains of chromosomal regions, had a highly significant impact on gene expression levels. Across the genome, on average, ~12% of the genes examined in the prostate cancer cell lines were involved in unbalanced chromosomal alterations. Quantitatively, 63% of the impact of copy number on global gene expression deregulation (number of genes affected multiplied by the fold change in gene expression) was attributable to such simple gains and losses. These expression changes were often small in magnitude, but the large number of genes affected in these large regions made their contribution quantitatively highly significant. This indicates that changes in gene dosage caused by common chromosomal losses and gains may have a more significant influence on genome-wide gene expression levels in cancer than previously thought. Such gene expression changes may directly contribute to the selection of common chromosomal alterations in cancer [7,24,25].

In summary, our cDNA microarray-based genomic and transcriptomic surveys provided a comprehensive integrated view on gene copy number changes in prostate cancer, mapped recurrent regions of interest at high resolution, implicated candidate genes therein, as well as provided a global view on the impact of such changes on gene expression profiles in prostate cancer.

References

[1] Luo J, Dunn T, Ewing C, Sauvageot J, Chen Y, Trent J, and Isaacs W (2002). Gene expression signature of benign prostatic hyperplasia revealed by cDNA microarray analysis. *Prostate* **51,** 189–200.

[2] Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K, Pienta KJ, Rubin MA, and Chinnaiyan AM (2001). Delineation of prognostic biomarkers in prostate cancer. *Nature* **412,** 822–826.

[3] El Gedaily A, Bubendorf L, Willi N, Fu W, Richter J, Moch H, Mihatsch MJ, Sauter G, and Gasser TC (2001). Discovery of new DNA amplification loci in prostate cancer by comparative genomic hybridization. *Prostate* **46,** 184–190.

[4] Nupponen NN, Hyytinen ER, Kallioniemi AH, and Visakorpi T (1998). Genetic alterations in prostate cancer cell lines detected by comparative genomic hybridization. *Cancer Genet Cytogenet* **101,** 53–57.

[5] Visakorpi T, Kallioniemi AH, Syvänen AC, Hyytinen ER, Karhu R, Tammela T, Isola JJ, and Kallioniemi OP (1995). Genetic changes in primary and recurrent prostate cancer by comparative genomic hybridization. *Cancer Res* **55,** 342–347.

[6] Clark J, Edwards S, Feber A, Flohr P, John M, Giddings I, Crossland S, Stratton MR, Wooster R, Campbell C, and Cooper CS (2003). Genome-wide screening for complete genetic loss in prostate cancer by comparative hybridization onto cDNA microarrays. *Oncogene* **22,** 1247–1252.

[7] Phillips JL, Hayward SW, Wang Y, Vasselli J, Pavlovich C, Padilla-Nash H, Pezullo JR, Ghadimi BM, Grossfeld GD, Rivera A, Linehan WM, Cunha GR, and Ried T (2001). The consequences of chromosomal aneuploidy on gene expression profiles in a cell line model for prostate carcinogenesis. *Cancer Res* **61,** 8143–8149.

[8] DeRisi J, Penland L, Brown PO, Bittner ML, Meltzer PS, Ray M, Chen Y, Su YA, and Trent JM (1996). Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* **14,** 457–460.

[9] Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, and Brown PO (1999). Genome-wide analysis of DNA copy-number changes using cDNA microarrays. *Nat Genet* **23,** 41–46.

[10] Monni O, Bärlund M, Mousses S, Kononen J, Sauter G, Heiskanen M, Paavola P, Avela K, Chen Y, Bittner ML, and Kallioniemi A (2001). Comprehensive copy number and gene expression profiling of the 17q23 amplicon in human breast cancer. *Proc Natl Acad Sci USA* **98,** 5711–5716.

[11] Chen Y, Dougherty ER, and Bittner ML (1997). Ratio-based decisions and the quantitative analysis of cDNA microarray images. *J Biomed Opt* **2,** 364–374.

[12] Chen Y, Kamat V, Dougherty ER, Bittner ML, Meltzer PS, and Trent JM (2002). Ratio statistics of gene expression levels and applications to microarray data analysis. *Bioinformatics* **18,** 1207–1215.

[13] Kallioniemi A, Kallioniemi OP, Piper J, Tanner M, Stokke T, Chen L, Smith HS, Pinkel D, Gray JW, and Waldman FM (1994). Detection and

mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc Natl Acad Sci USA* **91,** 2156–2160.

[14] Chen C, Frierson HF Jr, Haggerty PF, Theodorescu D, Gregory CW, and Dong J-T (2001). An 800-kb region of deletion at 13q14 in human prostate and other carcinomas. *Genomics* **77,** 135–144.

[15] Hyytinen ER, Frierson HF Jr, Boyd JC, Chung LW, and Dong JT (1999). Three distinct regions of allelic loss at 13q14, 13q21–22, and 13q33 in prostate cancer. *Genes Chromosomes Cancer* **25,** 108–141.

[16] Dong JT, Chen C, Stultz BG, Isaacs JT, and Frierson HF Jr (2000). Deletion at 13q21 is associated with aggressive prostate cancers. *Cancer Res* **60,** 3880–3883.

[17] Srivastava M, Bubendorf L, Srikantan V, Fossom L, Nolan L, Glasman M, Leighton X, Fehrle W, Pittaluga S, Raffeld M, Koivisto P, Willi N, Gasser TC, Kononen J, Sauter G, Kallioniemi OP, Srivastava S, and Pollard HB (2001). *ANX7*, a candidate tumor suppressor gene for prostate cancer. *Proc Natl Acad Sci USA* **98,** 4575–4580.

[18] Xin W, Rhodes DR, Ingold C, Chinnaiyan AM, and Rubin MA (2003). Dysregulation of the annexin family protein family is associated with prostate cancer progression. *Am J Pathol* **162,** 255–261.

[19] Miyake H, Hara I, Yamanaka K, Arakawa S, and Kamidono S (1999). Elevation of urokinase-type plasminogen activator and its receptor densities as new predictors of disease progression and prognosis in men with prostate cancer. *Int J Oncol* **14,** 535–541.

[20] Linja MJ, Savinainen KJ, Saramäki OR, Tammela TLJ, Vessella RL, and Visakorpi T (2001). Amplification and overexpression of androgen receptor gene in hormone-refractory prostate cancer. *Cancer Res* **61,** 3550–3555.

[21] Helenius MA, Saramäki OR, Linja MJ, Tammela TLJ, and Visakorpi T (2001). Amplification of urokinase gene in prostate cancer. *Cancer Res* **61,** 5340–5344.

[22] Pollack JR, Sørlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Børresen-Dale AL, and Brown PO (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99,** 12963–12968.

[23] Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E, Ringnér M, Sauter G, Monni O, Elkahloun A, Kallioniemi OP, and Kallioniemi A (2002). Impact of DNA amplification on gene expression patterns in breast cancer. *Cancer Res* **62,** 6240–6245.

[24] Hughes TR, Roberts CJ, Dai H, Jones AR, Meyer MR, Slade D, Burchard J, Dow S, Ward TR, Kidd MJ, Friend SH, and Marton MJ (2000). Widespread aneuploidy revealed by DNA microarray expression profiling. *Nat Genet* **25,** 333–337.

[25] Virtaneva K, Wright FA, Tanner SM, Yuan B, Lemon WJ, Caligiuri MA, Bloomfield CD, de La Chapelle A, and Krahe R (2001). Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proc Natl Acad Sci USA* **98,** 1124–1129.