NEOPLASIA
www.neoplasia.com

Volume 10 Number 8 August 2008 pp. 897–907 **897**

Homozygous Deletions and Recurrent Amplifications Implicate New Genes Involved in Prostate Cancer[1,2]

Wennuan Liu[*,†], Chunmei Carol Xie[*,†], Yi Zhu[*,†], Tao Li[*,†], Jishan Sun[*,†], Yu Cheng[*,†], Charles M. Ewing[‡], Sue Dalrymple[‡], Aubrey R. Turner[*,†], Jielin Sun[*,†], John T. Isaacs[‡], Bao-Li Chang[*,†], Siqun Lilly Zheng[*,†], William B. Isaacs[‡] and Jianfeng Xu[*,†]

*Center for Cancer Genomics, Wake Forest University School of Medicine, Winston-Salem, NC, USA; †Center for Human Genomics, Wake Forest University School of Medicine, Winston-Salem, NC, USA; ‡Johns Hopkins Medical Institutions, Baltimore, MD, USA

Abstract

Prostate cancer cell lines provide ideal *in vitro* systems for the identification and analysis of prostate tumor suppressors and oncogenes. A detailed characterization of the architecture of prostate cancer cell line genomes would facilitate the study of precise roles of various genes in prostate tumorigenesis in general. To contribute to such a characterization, we used the GeneChip 500K single nucleotide polymorphic (SNP) array for analysis of genotypes and relative DNA copy number changes across the genomes of 11 cell lines derived from both normal and cancerous prostate tissues. For comparison purposes, we also examined the alterations observed in the cell lines in tumor/normal pairs of clinical samples from 72 patients. Along with genome-wide maps of DNA copy number changes and loss of heterozygosity for these cell lines, we report previously unreported homozygous deletions and recurrent amplifications in prostate cancers in this study. The homozygous deletions affected a number of biologically important genes, including *PPP2R2A* and *BNIP3L* identified in this study and *CDKN2A*/*CDKN2B* reported previously. Although most amplified genomic regions tended to be large, amplifications at 8q24.21 were of particular interest because the affected regions are relatively small, are found in multiple cell lines, are located near *MYC*, an oncogene strongly implicated in prostate tumorigenesis, and are known to harbor SNPs that are associated with inherited susceptibility for prostate cancer. The genomic alterations revealed in this study provide an important catalog of positional information relevant to efforts aimed at deciphering the molecular genetic basis of prostate cancer.

Neoplasia (2008) 10, 897–907

Introduction

Prostate cancer (PCa) cell lines have been widely used not only as homogeneous tumor genomes for the identification of genetic alterations including deletion, amplification, translocation, and single nucleotide mutation but also as *in vitro* systems for studying the expression and function of various candidate tumor suppressor genes (TSGs) and oncogenes. To fully evaluate the expression and/or function of a specific (group of) gene(s) in a tumor genome, it is important to know the relative status of the overall genetic integrity across the genome because most genes and gene products interact with others to impose a biologic effect resulting in a specific phenotype. Various cytogenetic methods, including G-banding, spectral karyotyping and fluorescent *in situ* hybridization (FISH), and comparative genomic hybridization (CGH) such as metaphase (or conventional) and

Address all correspondence to: Dr. William B. Isaacs, Marburg 115, Johns Hopkins Hospital, 600 N. Wolfe Street, Baltimore, MD 21287. E-mail: wisaacs@jhmi.edu
[1]The study was partially supported by National Cancer Institute grants CA105055, CA106523, and CA95052 and by a Department of Defense grant PC051264 to J.X. and by a grant from the Dubie H. Holleman Trust for Cancer & Heart Research to W.L.
[2]This article refers to supplementary materials, which are designated by Figures W1–W22 and WX and are available online at www.neoplasia.com.
Received 27 March 2008; Revised 21 May 2008; Accepted 24 May 2008

Copyright © 2008 Neoplasia Press, Inc. All rights reserved 1522-8002/08/$25.00
DOI 10.1593/neo.08428

BAC/oligo/cDNA/single nucleotide polymorphic (SNP; or array-based) hybridization, have been used in the characterization of various PCa cell lines, primary and metastatic tumors, resulting in the identification of a number of recurrent losses and gains in the PCa tumor genome [1–11].

Some losses and/or gains of genomic DNA may result from a generalized genomic instability combined with long-term culture and associated proliferation of cancer cell lines and may not reflect alterations of genes "driving" the cancer initiation and progression process. However, a number of TSGs and oncogenes, such as *PTEN* and *MYC*, have been identified through the analysis of deletion and amplification of genetic material in tumor genomes and have been verified through the functional analysis of the genes in appropriate cell lines or in animal model systems.

Although single-copy losses and one– to two–extra-copy gains of DNA are common, complete genetic loss of both alleles and more extensive amplification of specific DNA sequences have been only rarely observed in PCa. This may be, in part, due to the limited resolution of the methods used in the identification of these copy number alterations. Using cDNA array CGH with a resolution ~570 kb, Clark et al. [5] identified two novel regions of complete loss at 17q21.31 and 10q23.1 in the PC3 cell line. The sizes of homozygous deletions in general are much smaller than typical hemizygous deletions affecting only one allele, presumably due to the fatal effect of large homozygous losses on the survival of the cancer cells. Therefore, analysis of these homozygous deletions can make a more direct contribution to the discovery of new genes involved in tumorigenesis. As the resolution of these analyses increases, additional homozygous deletions of novel TSG are expected to be revealed.

Analyzing loss of heterozygosity (LOH) in the cancer genome has also contributed to the identification of TSGs. Loss of heterozygosity due to hemizygous deletion may or may not be accompanied by duplication of the remaining allele. An analysis of both copy number changes and genotyping to determine allele-specific alterations is necessary to differentiate these possibilities. Although a number of CGH platforms with various resolutions have been used for mapping most DNA copy number alterations in the genome of PCa cell lines, SNP array is currently the only whole-genome approach for simultaneously mapping copy number alteration and LOH in a single analysis. To date, this type of high-resolution SNP arrays has not been used to analyze DNA copy number changes and LOH simultaneously for prostate cell lines.

In addition to the deletion of TSGs and gain of oncogenes, gene fusion events between androgen-regulated genes and the ETS family of oncogenes have been recently demonstrated in most PCa [12,13]. These fusions result from balanced and more commonly nonbalanced genetic translocations and deletions [13,14]. Using a combined cytogenetic and tiling array BAC CGH with a resolution of ~78 kb, Watson et al. [10] have demonstrated that a majority of balanced translocations are actually imperfectly balanced and are associated with focal deletions and duplications in PCa cell lines. Therefore, mapping the breakpoints of these translocations, which requires very high resolution, can provide valuable information characterizing gene fusion events in PCa [15].

In this study, we simultaneously analyzed DNA copy number alteration and LOH using the GeneChip 500K SNP array with an average resolution of ~5.8 kb in 11 prostate cell lines. In this report, we uncover some previously unreported homozygous deletions and amplifications with gains of three or more extra copies of DNA in PCa, along with a comprehensive map of copy number alteration and LOH and the related copy number data of more than 500K SNP probes for each of the cell lines. These data provide important positional information relevant to efforts aimed at deciphering the underlying molecular pathogenesis of PCa.

Materials and Methods

Cell Lines and DNA Isolation

PrECs are nonimmortalized early-passage human normal prostate epithelial cells maintained in serum-free defined medium (i.e., PrEC complete media). The cell and medium were obtained from Lonza Walkerville, Inc. (Walkerville, MD). The human DU145, LNCaP, PC3, and VCaP PCa cell lines were purchased from American Type Culture Collection (Manassas, VA). LNCaP is hypotetraploid, whereas DU145, PC3, LAPC4, and VCaP are near-triploid cell lines (American Type Culture Collection) [6]. BPH1 is a nontumorigenic, SV-40-immortalized, human prostate epithelial cell line generously provided by Simon Hayward [16]. CWR22Rv1 is derived from a human prostatic carcinoma xenograft that was serially propagated in mice after castration-induced regression and relapse of the parental, androgen-dependent CWR22 xenograft and was generously provided by Thomas G. Pretlow [17]. E006AA is a hypertriploid cell line established from primary PCa cells from an African-American patient, which demonstrates androgen-sensitive growth in culture; it was generously provided by Shahriar Koochekpour [18]. 975E/hTERT is a human telomerase reverse transcriptase–immortalized cell line derived from normal prostate cells generated from a radical prostatectomy (RP) specimen of a patient with a family history of PCa [19]. All of the cell lines were cultured in the RPMI-1640 medium with 10% fetal bovine serum or as previously described [19]. PC82 is a hormone-dependent transplantable tumor, which was maintained by serial xenograft transplantation in nude mice [20]; it was generously supplied by Fritz Schroeder [21]. We isolated genomic DNA using a DNA isolation kit from Gentra (Qiagen, Valencia, CA).

Clinical Samples

The clinical samples used in this study include 67 primary tumors and 5 lymph node metastases. All primary tumors were from PCa patients undergoing RP for treatment of clinically localized disease at Johns Hopkins Hospital. For the analysis of somatic DNA copy number alterations in these tumors, we selected cases from which genomic DNA of sufficient quantity (>5 μg) and purity (>70% cancer cells for cancer specimens; no detectable cancer cells for normal samples) could be obtained by macrodissection of matched nonmalignant (hereafter referred to as *normal*) and cancer-containing areas of prostate tissue as determined by histologic evaluation of hematoxylin and eosin–stained frozen sections of snap-frozen RP specimens. These include 9 with Gleason 6 tumors, 33 with Gleason 7 tumors, 6 with Gleason 8 tumors, 17 with Gleason 9 tumors, 1 with Gleason 10 tumors, and 1 having no Gleason score available.

GeneChip Mapping 500K SNP Assay

The GeneChip Mapping 500K set comprised Nsp (~262,000 SNPs) and Sty (~238,000 SNPs) arrays. The median physical distance between SNPs is ~2.5 kb, and the average distance between SNPs is ~5.8 kb. The 500K SNP arrays were purchased from Affymetrix, Inc., Santa Clara, CA. All of the reagents used for the assay were obtained from manufacturers recommended by Affymetrix. We labeled

and hybridized the 500K SNP mapping arrays according to the manufacturer's instructions [15]. We washed and stained the arrays using an Affymetrix Fluidics Station 450 and scanned the arrays using a GeneChip Scanner 3000 7G (Affymetrix, Inc.). The genotype of each SNP was generated by the GeneChip Genotyping analysis software (GTYPE) of Affymetrix.

DNA Copy Number, LOH, and Classification of Deletions and Gains

A database containing probe intensity and genotype data sets generated from the 500K SNP array using more than 200 DNA samples from both blood and normal prostate cells was used as reference samples for DNA copy number analysis in the cell lines. Because the ploidy levels vary among different cell lines, we normalized the data of $\log_2$ ratio using most of the chromosomes as the baseline within each of the tumor genomes based on the ploidy levels of the cell lines. DNA copy number and LOH were calculated based on allele intensity using two different software packages: Copy Number Analyzer for Affymetrix GeneChip (CNAG2.0 [22]) and dChip analyzer (dChip [23]). The physical positions of the detected deletions were determined based on the Human hg17 Assembly (NCBI Build 35). Deletions and gains are deduced from the average of $\log_2$ ratio with a smoothing window of 10-SNP probes using CNAG2.0 and dChip. The breakpoint SNPs were identified based on the median $\log_2$ ratio with a smoothing window of five SNP probes.

To distinguish which allele is deleted or gained in the clinical samples, we used allele-specific analysis as described by Nannya et al. [22]. The algorithm of allele-specific analysis takes advantage of genotype information and allele-specific intensities from paired samples to estimate DNA copy numbers for each heterozygous SNP. Allele-specific analysis can also minimize the effect of normal DNA contamination from nonmalignant cells in the tumors on the identification of DNA copy number changes, because the probe intensity of the other allele can be used as an internal reference for the detection of hemizygous deletion or gain. When the allele intensities of the match nontumor DNA are used as references, the homozygous deletion or gain can also be detected by allele-specific analysis.

We used the criteria described previously to define deletions and gains [24]. Hemizygous deletion refers to the loss of one of the alleles, whereas homozygous (biallelic) deletion refers to the loss of both alleles identified by allele-specific analysis in the clinical samples. In cell lines, homozygous deletions were identified by CNAG2.0 with a copy number of the probes equal to zero in default settings. The absolute $\log_2$ ratios of one- and two-copy losses varied depending on the ploidy levels of the cell lines but can be clearly distinguished according to the criteria described previously [24]. For example, in near-triploid cell lines, we used $\log_2$ ratios <-0.2 with an average of ~ -0.25 in the 10-SNP genomic smoothed data to define one-copy loss, whereas we used $\log_2$ ratios <-0.5 with an average of ~ -0.7 to define two-copy loss and $\log_2$ ratios <-0.9 with an average of ~ -1.8 to define three-copy (complete) loss. We used the same method to define gains. For example, in near-triploid cell lines, we used $\log_2$ ratios >0.2 with an average of ~0.24 in the 10-SNP genomic smoothed data to define one extra copy gain, whereas we used $\log_2$ ratios >0.4 with an average of ~0.6 to define two–extra-copy gain and >0.8 with an average of ~1 to define three–extra-copy gain or more. Where the gains were three or more extra copies, we define these gains as amplifications. We used both quantitative polymerase chain reaction (PCR) and FISH data to validate and refine the criteria for the copy number estimations.

Quantitative PCR

On the basis of the 500K SNP array analysis, we selected two normal DNA samples, PrEC from a primary culture of prostate epithelial cells and 083 from the blood of an unaffected individual, as reference samples for quantitative PCR (qPCR). The reaction was performed using the ABI Prism 7500 Sequence Detection System (Foster City, CA). Primers were designed using PrimerQuest software and were synthesized by Integrated DNA Technologies, Inc. (Coralville, IA). Amplicons were designed against the putatively homozygous-deleted and -amplified loci in the tumor cells and a control locus of known normal DNA copy number. The sequences of the primers for *BNIP3L*, *PPP2R2A*, *CDKN2A*, *CDKN2B*, *RAB20*, *ING1*, *OTUB1*, *MARK1*, and a locus at 8q24.11 are available on request. Polymerase chain reaction kinetics at the control locus with normal DNA copy number was used to control for sample-to-sample differences in genomic DNA purity and concentration. Polymerase chain reaction conditions and determination of deletion and gain were described previously [15].

Fluorescent In Situ *Hybridization*

Fluorescent *in situ* hybridization was performed to analyze the DNA copy number of the androgen receptor gene (*AR*) in E006AA cells. CEP × (DX21) spectrum green probe (for centromere control) and the *AR* (Xq12) spectrum orange probe were purchased from Vysis, Inc. (Downers Grove, IL). Cell preparation and hybridization were carried out according to the manufacturer's protocol. Both interphase and metaphase analyses were performed to estimate DNA copy number of *AR* in relationship to centromere in E006AA and cells from a normal male as control. Fluorescent *in situ* hybridization analysis was performed using fluorescent microscopy with the appropriate filters to visualize the probes. Slides were blinded, then for each sample, a total of 100 interphase nuclei were scored and a ratio of red signals (*AR*) to green signals (Xcen) was calculated.

Results

Copy Number Alterations and LOH Revealed by the 500K SNP Array in PCa Cell Lines

DNA copy number alterations in each of the 11 genomes for each chromosome are presented in Figures W1 to WX. The deletions and gains are inferred from the 10-SNP genomic smoothed $\log_2$ ratios of the hybridization intensity for each of the 500K probes analyzed (presented in a supplementary data set for each of the cell lines, http://www1.wfubmc.edu/Genomics/Publications+and+Data). Among the cell lines analyzed, PC3 and VCaP harbor the most DNA copy number alterations. Within these two cell lines, chromosome 5 harbors the most changes (Figure W5). In contrast, there were no large DNA copy number aberrations found in PrEC, although some small copy number variations were observed, for example, on chromosomes 14q and 15q (Figures W14 and W15, *red arrows*) in these presumably normal prostate epithelial cells. We also documented deletions and gains in prostate cell lines that had not been analyzed previously by array CGH (e.g., 975E, E006AA, and PC82; Figures W1–WX).

Because the SNP array also generates genotypes at each of the 500K SNPs, it was possible to analyze LOH across the whole genome for each of the cell lines using CNAG2.0 (Figures W1–WX). As would be predicted from similar previous studies, some regions of LOH cover the entire chromosomes in some of the cell lines. These include chromosomes 3, 7, 9, 13, and 14 in E006AA, chromosome 13 in DU145, chromosomes 8, 9, 10, 16, and 22 in PC3, and chromosome

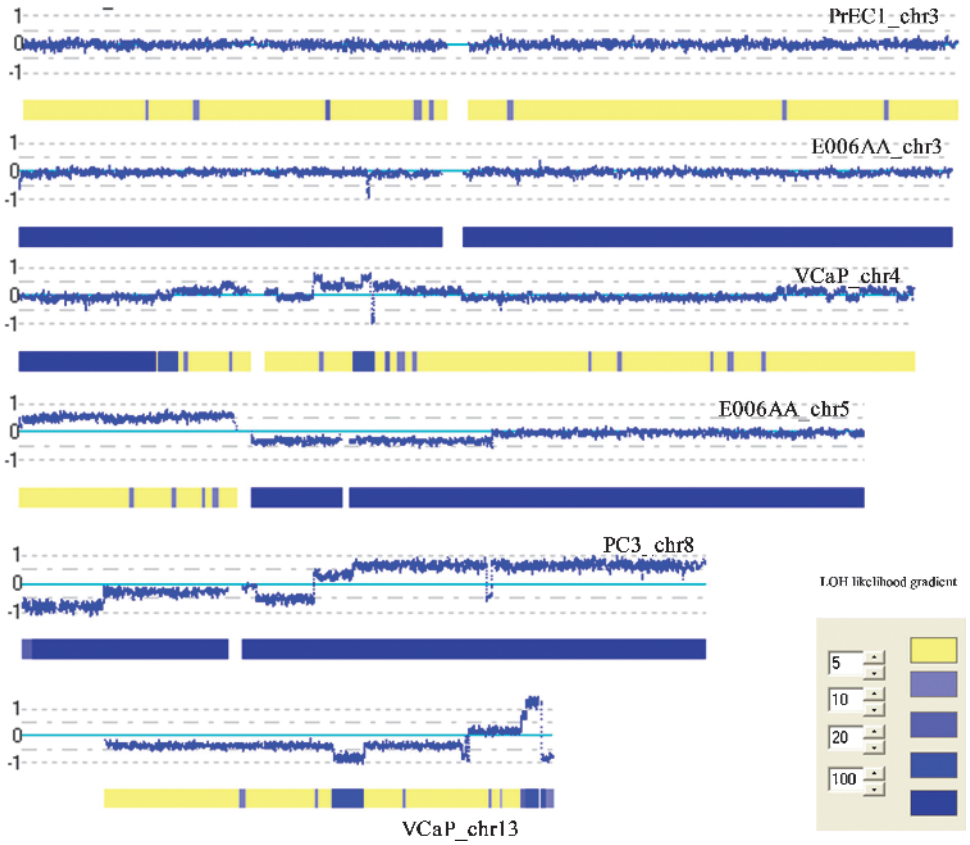


Figure 1. Examples of DNA copy number changes and LOH identified by CNAG2.0. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the likelihood gradients labeled on the left.

18 in VCaP, reducing hundreds to thousands of genes to homozygosity in these cell lines. Because some LOH events were not accompanied by copy number changes, LOH in these instances apparently occurred along with gain of the remaining chromosomal copy, either completely or partially (Figure 1). In addition, we analyzed the minimum common regions (MCRs) of recurrent LOH and listed the six that occurred in at least four cell lines in Table 1. Among the many genes affected by these recurrent alterations are known and suggested TSGs such as *TUSC1*, *PTEN*, and *KLF5* and genes involved in cellular senescence (*MORF4*), DNA repair (*HMGB2* and *EPC1*), and oxidative stress (*MGST1*).

Homozygous Deletions Implicate New Genes Involved in Prostate Cancer

We used three different approaches in the identification of homozygous deletions, including the value of log_2, SNP call (no call), and qPCR. Although, in general, it is not feasible to use default settings of CNAG2.0 to generate calls for a specific copy number alteration (other than homozygous deletion with a copy number equal to zero) in the presence of multiple levels of deletion in polyploid cells, there are obvious differences in log_2 ratios between levels of deletion within a given genome when it is analyzed using the 500K SNP array. Taking the deletions on chromosome 5 in near-triploid cell lines PC3, for example (Figure 2*a*), we were able to clearly identify one-copy deletions (*light blue arrows*), two-copy deletions (*black arrows*), and complete or homozygous deletions (*green arrows*). In the regions of homozygous deletion, we found the no-call rate increased in comparison to the regions without homozygous deletion, consistent with loss of all copies of the specific genomic interval. To validate the homozygous deletion calls, we performed qPCR and confirmed the results as shown in Figure 2*b*.

Using the 500K SNP array, we identified 12 homozygous deletions in five cell lines, with 7 apparently being novel (Figure 2, *a* and *c*). Half of these deletions occurred in a single cell line, PC3. The genes affected

Table 1. Examples of Recurrent LOH Revealed by the 500K Mapping Array in Prostate Cancer Cell Lines.

Cell Line	Chromosome	Start	End	Size (bp)	Gene
975E, CWR22Rv1, DU145, E006AA, PC3	4	171,890,559	176,165,998	4,275,439	*AK128523*, *SAP30*, *GALNT7*, *HMGB2*, *SCRG1*, *MORF4*, *FBXO8*, *HPGD*, *GLRA3*
DU145, E006AA, PC3, PC82	9	22,021,005	25,996,565	3,975,560	*AF109294*, *DMRTA1*, *ELAVL2*, *TUSC1*
BPH1, DU145, PC3, PC82	10	30,398,784	34,803,394	4,404,610	19 genes, *MAP3K8*, *TCF8*, *EPC1*
E006AA, LNCaP, PC3, PC82	10	88,889,330	90,573,072	1,683,742	11 genes, *MINPP1*, *PTEN*, *LIPF*
DU145, LAPC4, LNCaP, PC3	12	16,235,442	17,277,690	1,042,248	*MGST1*, *LMO3*
DU145, E006AA, LNCaP, PC82, VCaP	13	72,316,041	74,078,996	1,762,955	*C13orf24*, *KLF5*, *KLF12*

by these novel homozygous deletions include *CDH18*, *AK130123*, *PPP2R2A*, *BNIP3L*, *MAT1A*, *LOC143241*, *AK125908*, *C10orf58*, *BC005871*, *TSPAN14*, *SH2D4B*, *MAP4K5*, *SPG3A*, *SAV1*, *USP10*, *CR593410*, *ZDHHC7*, *KIAA0513*, *FLJ44299*, *MGC22001*, and *PPAP2C*. The cyclin-dependent kinase inhibitors *CDKN2A* and *CDKN2B* were found to be deleted in the E006AA line, the only line in this study derived from PCa arising in an African-American. Notably, each homozygous deletion was unique in that it was observed only once among the lines examined, although a subset of these deletions was observed additional times in the clinical specimens examined (see below). All of the homozygous deletions reside in larger regions of LOH (Figures W2, W5, W6, W8, W9, W10, W14, W16, W17, and W19). We confirmed five other regions that have been reported to be completely lost: on chromosome 2 in LNCaP; chromosomes 5, 10, and 17 in PC3; and on chromosome 6 in DU145, including the genes *MSH2*, *KIAA0416*, *PTEN*, *SFTPA2*, *CTNNA1*, *STAT3*, and *STAT5B*. Furthermore, using the 500K SNP array, we defined boundaries of previously reported homozygous deletions more closely and identified more affected genes, which include *KCNK12*, *LRRTM2*, *C10orf59*, *LIPL1*, *LIPF*, *LGP2*, *GCN5L2*, *HSPB9*, *RAB5C*, *KCNH4*, *HCRT*, *AK075277*, *LGP1*, *STAT5B*, *STAT5A*, *PTRF*, and *ATP6V0A1* (Figure 2*c*).

Screening for these homozygous deletions in the clinical samples from 72 patients with PCa using allele-specific analysis, we found five recurrent biallelic losses on chromosomes 8, 9, 10, and 16, with examples shown in Figure 2*d* (*green arrows*). The most common complete loss of both alleles occurred at chromosome 10q23.31 and included *PTEN*, with a frequency of 21%, whereas the frequency of homozygous deletions at 9p21.3 and 16q24.1 was 4.2% (Figure 2*c*). The regions of complete loss at 9p21.3 in the clinical samples overlap with that observed in cell line E006AA, with an MCR affecting only *CDKN2A* and *CDKN2B* from 21,968 to 22,125 kb (Figure 2*c*). The MCR of complete loss at 16q24.1 in the VCaP cell line and clinical samples ranges from 83,300 to 83,541 kb, affecting the genes *USP10* and *CRISPLD2*. We also observed single occurrences of homozygous deletions on 8p in the 72 clinical samples. The complete loss at 8p21.2 in a primary tumor affected *AK130123*, *EBF2*, *PPP2R2A*, and *BNIP3L*. We did not observe any complete losses at the other regions listed in Figure 2*c* in the clinical samples analyzed.

Amplifications with More Than Two Extra Copies

Using the same approach as described above for the analysis of multiple levels of deletion, we were able to define multiple levels of gain, as shown in Figure 3. Taking the gains on chromosomes 1 (Figure 3*a*) and 10 (Figure 2*a*) in PC3, for example, we identified gained regions with one, two, and three or more extra copies, marked by *brown*, *pink*, and *red arrows*, respectively, in this near-triploid genome. In comparison to homozygous deletions (Figure 2, *a* and *c*), the size of the amplifications, as defined by the presence of three or more extra copies, and ranging from approximately 88 kb to 7.8 Mb, was generally far larger than the homozygous deletions observed and affected a number of genes in various cell lines (Figure 3*b*). To confirm these amplifications, we selected five different regions on chromosomes 8, 11, and 13 and performed qPCR analysis of DNA copy number using *GAPDH* as the control locus. For example, in VCaP as shown Figure 4, the results of qPCR analysis revealed gains of multiple copies of *RAB20*, *ING1*, *OTUB1*, and *MARK2*. In addition, we also validated the amplification of an amplicon at 8q24.21 where several germ line PCa risk alleles have been identified in populations of both European and African descent. Whereas most of the regions with amplification were observed only once in this group of analyzed cell lines, the amplification at 8q24.21 was found in 3 of these 11 cell lines (Figure 3*b*).

Analyzing the gains in the 72 clinical samples, we did not find the amplifications that we had observed in the cell lines shown in Figure 3*b*, except at chromosome 8q (Figure 2*d*). Among these clinical samples, approximately 38% harbor gains on 8q with most of them being hemizygous and affecting the entire q arm. In the allele-specific analysis, we identified three tumors harboring biallele gains, ranging from 98,928 kb to 8qter (Figure 2*d*, *red arrows*), which overlap with the amplifications at 8q24. However, the MCR ranging from 128,776 to 136,826 kb in the clinical samples is located 3′ from the MCR in the cell lines and affected a number of genes including *MYC*.

In addition, we detected an amplification event on chromosome X, ranging from 65,893 to 66,811 kb in VCaP. This alteration harbored only the *AR* gene and was present in at least five extra copies (Figure 3). Amplification of *AR* was also observed in the cell line E006AA, although the affected region was larger, ranging from 64,829 to 76,599 kb, covering a number of genes (Figure WX). The gain of *AR* in E006AA cells was confirmed by FISH where metaphase analysis indicated two copies of the X centromere (Figure 4*i*, *green*), consistent with the previously reported karyotype [18] and what seems to be two to three or more copies of the *AR* gene per X chromosome (Figure 4*i*, *red*). Interphase analysis demonstrated multiple red signals in relation to the number of green signals that could be more easily counted than in the metaphase. The number of green signals per nucleus ranged from 1 to 7 and the number of red signals per nucleus ranged from 3 to 18 (Figure 4*h*), indicating extensive genetic heterogeneity within these cells. On average, we observed 5.5 copies of *AR* with 2.2 copies of CEP X signals per nucleus, which is consistent with the result from the SNP array analysis (Figure WX). In comparison to E006AA, VCaP apparently harbors a higher level of *AR* amplification, according to the $\log_2$ ratios and metaphase analysis of FISH (Figure WX and data not shown).

Discussion

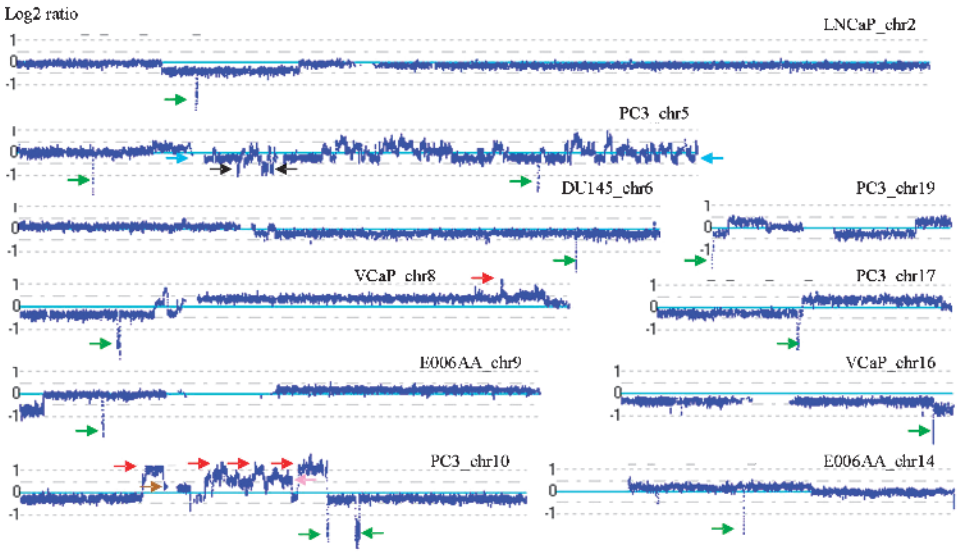
Alterations of DNA copy number in prostate cell lines, such as LNCaP, PC3, and DU145, have been systematically analyzed using cDNA microarray [5,8] and BAC tiling path array CGH [10]. Although the large-scale genomic alterations uncovered by various platforms are very similar, we find that the deletions and gains revealed by the 500K SNP array are more consistent with the results generated from the higher-resolution tiling path array than those from the cDNA arrays, which contain fewer and less evenly distributed probes. The varied results from study to study are most likely due to these factors, although variations in the cell lines themselves cannot be ruled out. Nevertheless, the comprehensive analysis and documentation of DNA copy number changes in these widely used *in vitro* model systems provides a detailed genetic background for studying PCa biology.

One of the hallmarks of cancers is genomic instability resulting in aneuploidy, which has been considered as a valuable predictor of clinical outcome [25]. To more specifically characterize the aneuploidy in PCa cell lines, we took advantage of the genotype data from the 500K SNP array and analyzed LOH across the entire genome using CNAG2.0. We found that some entire chromosomes or arms lost heterozygosity, although the DNA copy number either did not change or increased. This type of LOH apparently resulted from the loss of

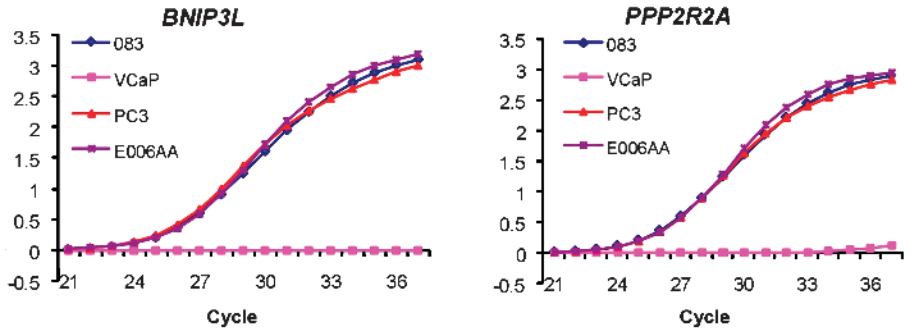
one chromosome or arm in combination with the amplification of the other chromosome or arm. It would be interesting to explore the biologic impact in processes, such as progression of PCa, of large LOH events resulting in the reduction to homozygosity of many hundreds to thousands of genes, with or without accompanying further changes through deletion and gain.

Complete losses of DNA at five chromosomal regions have been reported in PC3 and DU145 [5]. We have validated three of these

a.



b.



c.

Homozygous deletions revealed by the 500K mapping array in prostate cancer cell lines

Cell_line	Chr	Start	End	Size (bp)	Gene	P tumor (%)
LNCaP	2	47,612,563	47,868,470	255,907	*MSH2 , KCNK12*	
PC3	5	19,901,834	19,989,940	88,106	*CDH18*	
PC3	5	138,164,484	138,341,045	176,561	*CTNNA1 , LRRTM2, SIL1*	
DU145	6	148,169,277	148,334,801	165,524		
VCaP	8	25,960,305	26,416,283	455,978	*AK130123, PPP2R2A, BNIP3L*	1.4
E006AA	9	21,968,443	22,126,489	158,046	*CDKN2A , CDKN2B*	4.2
PC3	10	82,002,079	82,310,973	308,894	*MAT1A, LOC143241, AK125908, C10orf58, BC005871, TSPAN14, SH2D4B*	
PC3	10	89,664,137	90,502,286	838,149	*PTEN, C10orf59 , LIPL1, LIPF*	20.8
E006AA	14	50,039,198	50,186,809	147,611	*MAP4K5, SPG3A, SAV1*	
VCaP	16	83,300,379	83,882,068	581,689	*USP10, CR593410, ZDHHC7, KIAA0513, FLJ44299, MGC22001*	4.2
PC3	17	37,510,024	37,871,777	361,753	*LGP2, GCN5L2, HSPB9, RAB5C, KCNH4, HCRT, AK075277, LGP1, STAT5B, STAT5A, STAT3, PTRF, ATP6V0A1*	
PC3	19	212,033	252,668	40,636	*PPAP2C*	

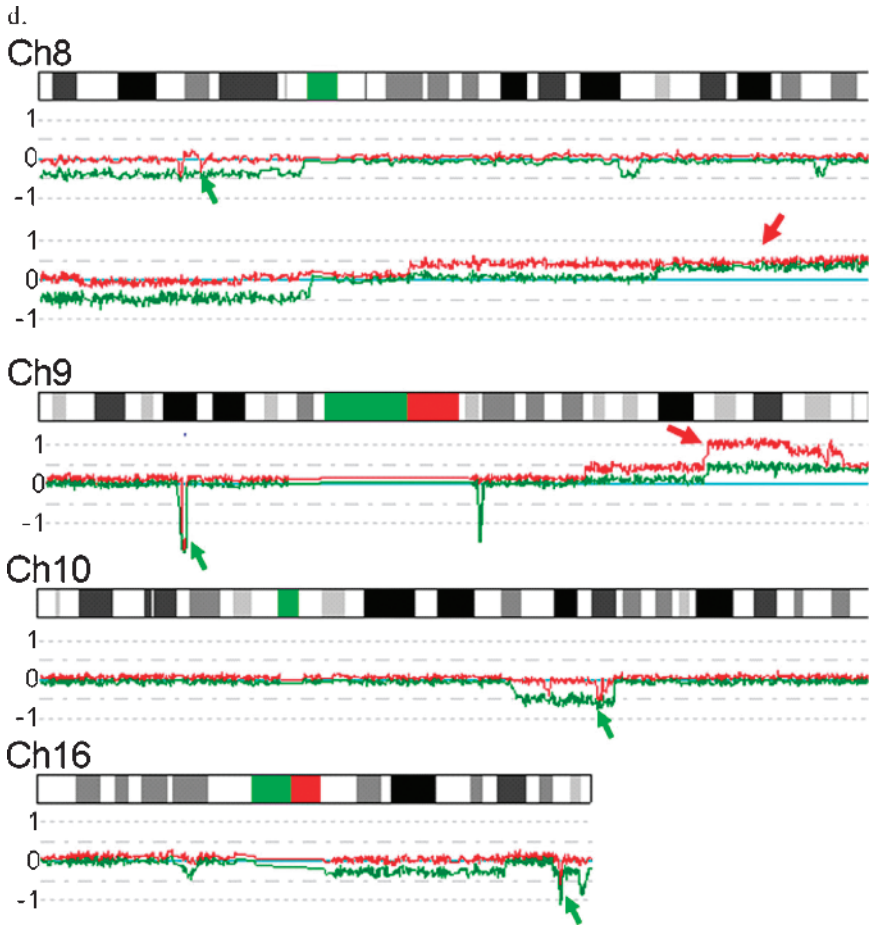


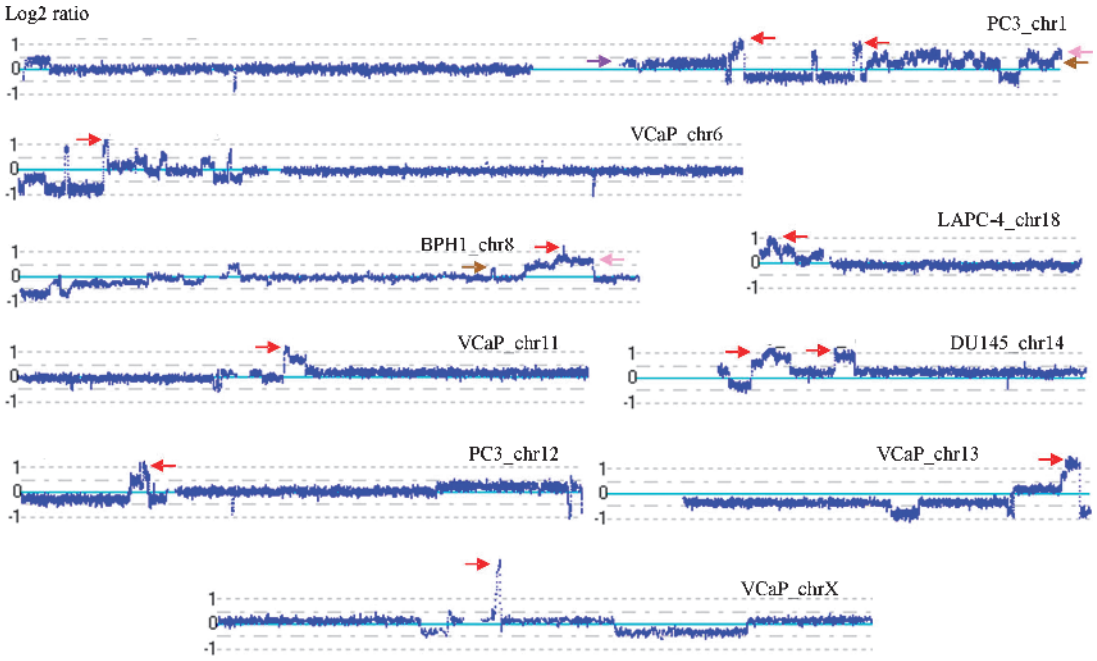
Figure 2. (*continued*)

regions on chromosomes of 5, 10, and 17. It is not surprising that the most frequent homozygous deletion in these 72 clinical samples analyzed in this study was found at 10q23.31, affecting *PTEN*. These results emphasize the significance of losing PTEN in the development of PCa. Somewhat surprisingly, occurrences of deletions between the *TMPRSS2* and *ERG* genes were less common in these cancer cell lines (Figure W21), particularly considering the high frequency of these deletions in clinical samples of prostate cancer, although the number of the cancer cell lines analyzed is too small to be conclusive. This may suggest a counterselection for this genomic alteration in PCa cells, which can grow and can be established in culture. Furthermore, these results emphasize some of the different selection pressures that exist for PCa cells in patients *versus* in culture. Although these cell lines are valuable models to study human PCa in general, some critical differences should be borne in mind when extrapolating results from *in vitro* to *in vivo* settings.

Particularly interesting genes affected by complete losses identified in PCa cell lines and the clinical samples in this study include *BNIP3L* and *PPP2R2A* at 8p21.2, *CDKN2A* and *CDKN2B* at 9p21.3, *USP10* and *CRISPLD2* at 16q24.1, and *MAP4K5* at 14q21.3. *BNIP3L* (BCL2/adenovirus E1B 19-kDa interacting protein) encodes a protein that is a functional homolog of *BNIP3* and may play a role in tumor suppression [26]. Fei et al. [27] found that the expression of *BNIP3L* was induced by p53 under hypoxia, and its knockdown promoted tumor growth *in vivo*. Down-regulation of *BNIP3L* has been reported to be correlated with liver metastases and tumor invasion [28] and

Figure 2. Homozygous deletions identified in PCa cell lines. (a) Log_2 ratios (horizontal dot lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of the 500K probes. Green arrows indicate homozygous deletion. Red arrow indicates amplifications with three or more extra copies. Light blue arrow indicates one-copy loss. Black arrow indicates two-copy loss. Brown arrow indicates one–extra-copy gain. Pink arrow indicates two–extra-copy gain. (b) Validation of homozygous deletions in PCa cell lines using qPCR. The raw signal intensities for each of the test loci were normalized against the signal intensity of control locus (*GAPDH*) with normal copy numbers for each of the cycles. The normalized relative signal intensities (*Y*-axis) are plotted for each of the PCR cycles (*X*-axis), with 083 from normal blood DNA used as a reference. No signals of *PPP2R2A* and *BNIP3* detected in VCaP indicate complete loss of these two genes. (c) Physical locations and genes involved in homozygous deletions. *Chr* indicates chromosome; *P tumor*, clinical sample. (d) Examples of recurrent homozygous deletions (green arrows) and biallelic gains (red arrows) identified in different clinical samples by allele-specific analysis. Log_2 ratios of the two alleles (red and green curves) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). G6, G7, G8, and G9 represent Gleason scores of 6, 7, 8, and 9, respectively, for different subjects.

a.



b.

Cell_line	Chr	Start	End	Size	# of gene	Known Gene
PC3	1	168,165,989	168,692,795	526,807	5	*BAT2D1, MYOC, VAMP4, KIAA0859, DNM3*
PC3	1	169,086,789	170,545,105	1,458,317	28*	*DNM3, AL035301, LOC92346, PIGC, C1orf9, ...*
PC3	1	196,478,779	197,040,038	561,260	1	*NR5A2*
PC3	1	197,393,872	198,162,732	768,861	10	*CAMSAP1L1, GPR25, FLJ10901, AB007918, AK131318, TMEM9, DKFZp434B1231, PKP1, TNNT2, LAD1*
VCaP	6	11,088,294	11,585,571	497,278	4	*ELOVL2, HERV-FRD , NEDD9, AK127678*
VCaP	6	19,948,471	20,923,239	974,769	5	*ID4, OACT1, E2F3, AF116619, CDKAL1*
BPH1	8	128,288,090	128,540,858	252,769	1	*AF268618*
VCaP	8	128,048,873	128,273,398	224,526	0	
LAPC-4	8	128,040,097	128,672,788	632,692	1	*AF268618*
PC3	10	33,454,754	38,273,379	4,818,626	9	*NRP1, PARD3, CUL2, CREM, AF465728, CX40.1, FZD8, ANKRD30A, ZNF248*
PC3	10	50,253,092	50,713,871	460,780	6	*ERCC6, PGBD3, SLC18A3, CHAT, C10orf53, OGDHL*
PC3	10	74,184,621	81,966,905	7,782,285	109*	*OIT3, PLA2G12B, P4HA1, NUDT13, HSGT1, ...*
VCaP	11	62,876,042	67,750,958	4,874,917	298*	*MARK2, OTUB1, MAP4K2, PPP2R5B, PTPRCAP, ...*
PC3	12	28,027,256	28,115,299	88,044	0	
PC3	12	28,703,465	29,449,882	746,418	3	*AF113698, MLSTD1, PTX1*
VCaP	13	107,369,152	111,787,023	4,417,872	38*	*IRS2, RAB20, ING1, COL4A1, COL4A2, ...*
DU145	14	29,798,243	36,253,365	6,455,123	66*	*KIAA1333, SCFD1, COCH, STRN3 , AP4S1, ...*
DU145	14	46,693,956	51,498,797	4,804,842	56*	*RPS29, PPIL5, POLE2, ARF6, L2HGDH, ...*
PC3	14	67,786,720	67,914,548	127,829	1	*RAD51L1*
PC3	14	72,257,827	72,961,640	703,814	7	*DPF3, WDR21A, ZFYVE1, RBM25, PSEN1, PAPLN , NUMB*
LAPC-4	18	1,749,809	4,004,959	2,255,151	10	*METTL4, KNTC2, CR627458, EMILIN2, LPIN2, MYOM1, MRCL3, MRLC2, TGIF, DLGAP1*
VCaP	X	65,873,500	66,811,243	937,744	1	*AR*

Figure 3. Examples of amplifications with three or more extra copies identified in PCa cell lines. (a) Log_2 ratios (horizontal dot lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Red arrow indicates amplification with three or more extra copies. Brown arrow indicates one–extra-copy gain. Pink arrow indicates two–extra-copy gain. (b) Physical locations and genes involved in amplification. An asterisk (*) represents a known gene item count from the UCSC Table Browser utility (University of California, Santa Cruz, CA).

poor prognosis in hepatocellular carcinoma [29]. *PPP2R2A* belongs to the phosphatase 2 regulatory subunit B family. Protein phosphatase 2 is one of the four major Ser/Thr phosphatases and is implicated in the negative control of cell growth and division. Studying the fusion of *CHEK2* and *PPP2R2A* in childhood teratoma, Jin et al. [30] reported that the deregulation of *CHEK2* and/or *PPP2R2A* is of pathogenetic importance in at least a subset of germ cell tumors. *CDKN2A* and *CDKN2B* are cyclin-dependent kinase inhibitors. *CDKN2A*, also named as multiple tumor suppressor (*MTS1*), encodes $p16^{INK4a}$ and $p14^{ARF}$ and acts as a negative regulator of the proliferation of normal cells by interacting strongly with *CDK4* and *CDK6*. Defect in *CDKN2A* is involved in tumor formation in a wide range of tissues and is the cause of familial atypical multiple mole melanoma–pancreatic carcinoma syndrome. Although hemizygous deletion of

CDKN2A has been reported in prostate tumors [31–33], homozygous deletion of this gene in prostate tumors and cancer cell lines was not observed in several studies [33–35]. However, homozygous deletion of *CDKN2A* has been observed at low frequency in prostate tumors with a high Gleason score [36]. *CDKN2B* lies adjacent to the TSG *CDKN2A*. Using a systematic multiplex reverse transcription–polymerase chain reaction, Yamamoto et al. [37] found a decrease in expression of this gene in most breast and PCa cell lines. *USP10* may function as a cofactor of the DNA-bound androgen receptor complex and is inhibited by a protein in the Ras–GTPase pathway, whereas *CRISPLD2* seems to be highly expressed in prostate. *MAP4K5* encodes a member of the serine/threonine protein kinase, which was shown to activate Jun kinase in mammalian cells, suggesting a role in stress response. This kinase is expressed at high levels in the ovary, testis, and prostate, whereas expression is decreased in PCa (ONCOMINE, http://www.oncomine.org/).

Most of the regions with amplification were large, affected a number of genes, and like homozygous deletions, were observed only once in this group of prostate cell lines. It is therefore hard to draw a connection between these amplifications and tumorigenesis because of the large number of genes affected and because most of the affected genes have little obvious relevance to the development of PCa on the basis the functional descriptions in public databases. However, the amplification at 8q24.21 was very interesting because it was found in two PCa cell lines and was located near *MYC*, an oncogene that is frequently up-regulated in a number of cancers. It is of particular interest

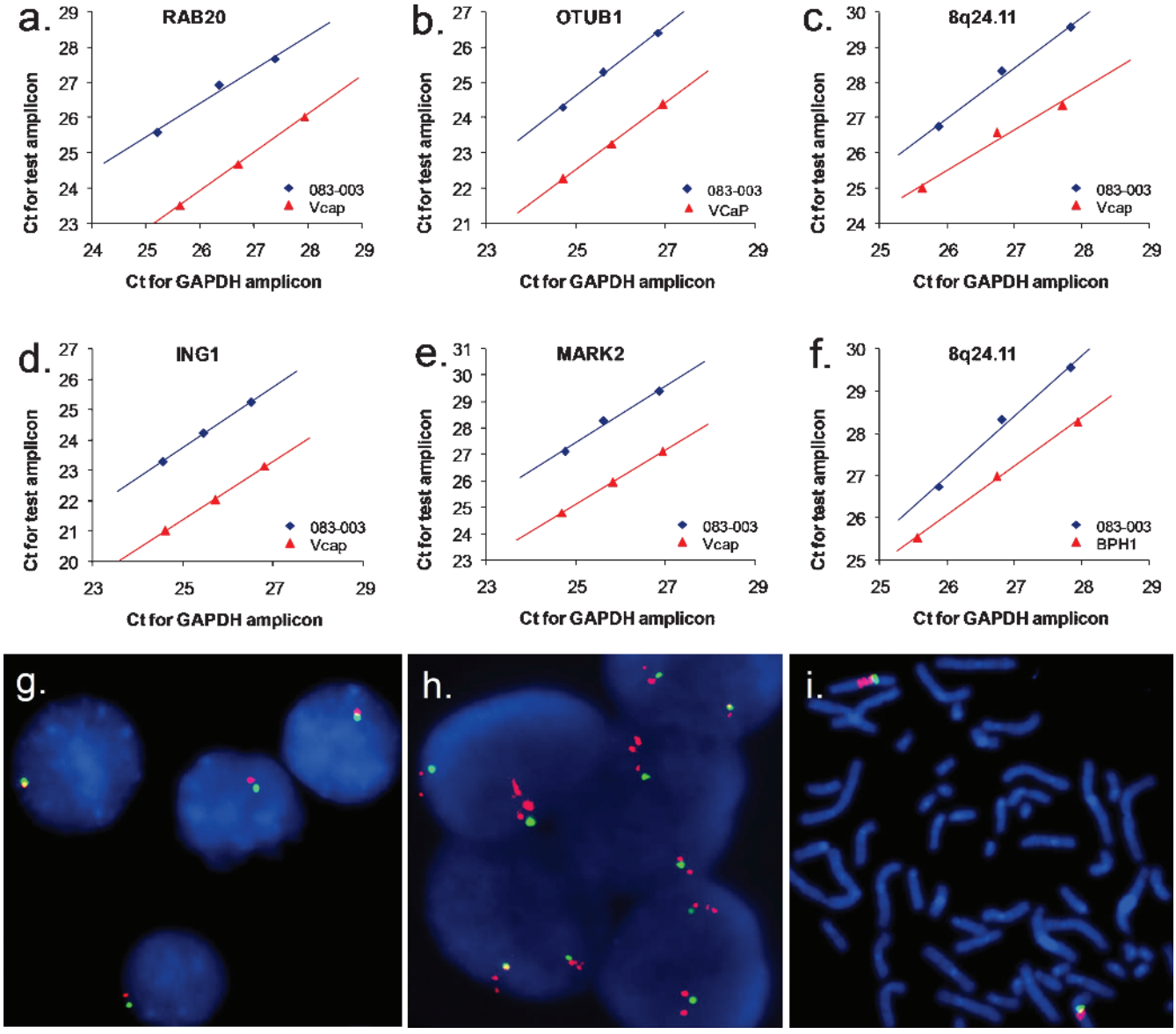


Figure 4. Validation of amplifications in PCa cell lines using qPCR and FISH. Amplifications of *RAB20* (a), *OTUB1* (b), *ING1* (d), *MARK2* (e), and a locus at 8q24 (c and f) confirmed by qPCR. Cycle numbers (C_t) of the control (*GAPDH*; *X*-axis) with normal copy number and test amplicons (*Y*-axis) for the three dilutions of each DNA sample were plotted against each other and the offset between the reference sample (083–003; blue) and DNA from cell lines (red). Fluorescent *in situ* hybridization analysis of DNA copy number of *AR* in a normal male control cell line (g) and in the PCa cell line E006AA (h and i). Green indicates the X centromere probe. Red indicates the *AR* probe at Xq12. Metaphase analysis of E006AA indicated two copies of X centromere (green) and what seem to be multiple copies of the *AR* gene (red).

that the gained region harbors a number of risk variants, including rs10086908, rs16901979, rs6983267, rs10505477, rs7017300, and rs1447295, which are associated with both prostate and colorectal cancer susceptibility in various human populations [38,39]. This amplification was also observed in BPH1, a nontumorigenic prostate epithelial cell line, indicating that the occurrence of this amplification event by itself is not sufficient for tumorigenesis. Further study is warranted to dissect the function of these DNA sequences in regulating the development of PCa.

The other recurring region of amplification was observed at Xq12, occurring in two of the PCa cell lines analyzed. In both cases, the *AR* gene was included in the amplified region, either alone or in combination with several other genes. It is well established that the androgen receptor plays a key role in the progression of PCa to hormone independence. A highly localized amplification of *AR* was apparently associated with the androgen-deprived selection of the PCa cells in the vertebral metastatic deposit from which the VCaP cell line was derived [40]. However, *AR* amplification was also observed in E006AA that was derived from a clinically localized primary prostate cancer from an African-American patient that apparently had not undergone hormone deprivation treatment. Whether the gain of *AR* sequences occurred before or after establishment of the E006AA cells in culture is unclear. In any case, these results emphasize that gain of *AR* sequences is a recurring but by no means a mandatory event associated with prostate cancer progression.

In summary, while documenting DNA copy number changes and LOH at >500,000 locations across the entire genome of 11 prostate cell lines, we have highlighted the following: 1) unreported LOH regions, including frequent monosomy with accompanying gain and loss, 2) novel homozygous deletions that implicate new genes involved in PCa, 3) gains of *AR* in two cell lines, and 4) recurrent amplification of regions that harbor a number of germ line risk SNPs that have been identified in both prostate and colorectal cancers. Further studies are needed to evaluate whether risk SNPs in the germ line play any role in the amplification of chromosome 8q including the oncogene *MYC*, and ultimately whether amplification of these SNPs in the somatic genome contributes to the formation of PCa.

Deletion and gain of a number of regions that affect numerous genes in the tumor genomes of PCa have been identified [41]. However, only a few these genes have been investigated in the functional analyses using *in vitro* and *in vivo* systems with limited knowledge on their genomic architectures. In this study, we provide the genomic architectures of 11 prostate cell lines at high resolution, which will facilitate further functional analysis of a number of candidate TSGs and oncogenes, including the novel ones identified in this study.

References

[1] Bernardino J, Bourgeois CA, Muleris M, Dutrillaux AM, Malfoy B, and Dutrillaux B (1997). Characterization of chromosome changes in two human prostatic carcinoma cell lines (PC-3 and DU145) using chromosome painting and comparative genomic hybridization. *Cancer Genet Cytogenet* **96**, 123–128.

[2] 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.

[3] Virgin JB, Hurley PM, Nahhas FA, Bebchuk KG, Mohamed AN, Sakr WA, Bright RK, and Cher ML (1999). Isochromosome 8q formation is associated with 8p loss of heterozygosity in a prostate cancer cell line. *Prostate* **41**, 49–57.

[4] Aurich-Costa J, Vannier A, Gregoire E, Nowak F, and Cherif D (2001). IPM-FISH, a new M-FISH approach using IRS-PCR painting probes: application to the analysis of seven human prostate cell lines. *Genes Chromosomes Cancer* **30**, 143–160.

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

[6] van Bokhoven A, Caires A, Maria MD, Schulte AP, Lucia MS, Nordeen SK, Miller GJ, and Varella-Garcia M (2003). Spectral karyotype (SKY) analysis of human prostate carcinoma cell lines. *Prostate* **57**, 226–244.

[7] Lieberfarb ME, Lin M, Lechpammer M, Li C, Tanenbaum DM, Febbo PG, Wright RL, Shim J, Kantoff PW, Loda M, et al. (2003). Genome-wide loss of heterozygosity analysis from laser capture microdissected prostate cancer using single nucleotide polymorphic allele (SNP) arrays and a novel bioinformatics platform dChipSNP. *Cancer Res* **63**, 4781–4785.

[8] Zhao H, Kim Y, Wang P, Lapointe J, Tibshirani R, Pollack JR, and Brooks JD (2005). Genome-wide characterization of gene expression variations and DNA copy number changes in prostate cancer cell lines. *Prostate* **63**, 187–197.

[9] Paris PL, Hofer MD, Albo G, Kuefer R, Gschwend JE, Hautmann RE, Fridyland J, Simko J, Carroll PR, Rubin MA, et al. (2006). Genomic profiling of hormone-naive lymph node metastases in patients with prostate cancer. *Neoplasia* **8**, 1083–1089.

[10] Watson SK, deLeeuw RJ, Horsman DE, Squire JA, and Lam WL (2007). Cytogenetically balanced translocations are associated with focal copy number alterations. *Hum Genet* **120**, 795–805.

[11] Kim JH, Dhanasekaran SM, Mehra R, Tomlins SA, Gu W, Yu J, Kumar-Sinha C, Cao X, Dash A, Wang L, et al. (2007). Integrative analysis of genomic aberrations associated with prostate cancer progression. *Cancer Res* **67**, 8229–8239.

[12] Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, et al. (2005). Recurrent fusion of *TMPRSS2* and *ETS* transcription factor genes in prostate cancer. *Science* **310**, 644–648.

[13] Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, et al. (2007). Distinct classes of chromosomal rearrangements create oncogenic *ETS* gene fusions in prostate cancer. *Nature* **448**, 595–599.

[14] Liu W, Chang B, Sauvageot J, Dimitrov L, Gielzak M, Li T, Yan G, Sun J, Sun J, Adams TS, et al. (2006). Comprehensive assessment of DNA copy number alterations in human prostate cancers using Affymetrix 100K SNP mapping array. *Genes Chromosomes Cancer* **45**, 1018–1032.

[15] Liu W, Ewing CM, Chang BL, Li T, Sun J, Turner AR, Dimitrov L, Zhu Y, Sun J, Kim JW, et al. (2007). Multiple genomic alterations on 21q22 predict various *TMPRSS2/ERG* fusion transcripts in human prostate cancers. *Genes Chromosomes Cancer* **46**, 972–980.

[16] Hayward SW, Dahiya R, Cunha GR, Bartek J, Despande N, and Narayan P (1995). Establishment and characterization of an immortalized but non-tumorigenic human prostate epithelial cell line: BPH-1. *In Vitro* **31A**, 14–24.

[17] Sramkoski R, Pretlow TG, Giaconia J, Pretlow TP, Schwartz S, Sy MS, Marengo S, Rhim J, Zhang D, and Jacobberger J (1999). A new human prostate carcinoma cell line, 22Rv1. *In Vitro Cell Dev Biol Anim* **35**, 403–409.

[18] Koochekpour S, Maresh GA, Katner A, Parker-Johnson K, Lee TJ, Hebert FE, Kao YS, Skinner J, and Rayford W (2004). Establishment and characterization of a primary androgen-responsive African-American prostate cancer cell line, E006AA. *Prostate* **60**, 141–152.

[19] Yasunaga Y, Nakamura K, Ewing MC, Isaacs WB, Hukku B, and Rhim JS (2001). A novel human cell culture model for the study of familial prostate cancer. *Cancer Res* **61**, 5969–5973.

[20] Hoehn W, Schroeder FH, Reimann JF, Joebsis AC, and Hermanek P (1980). Human prostatic adenocarcinoma: some characteristics of a serially transplantable line in nude mice (PC 82). *Prostate* **1**, 95–104.

[21] Romijn JC, Verkoelen CF, and Schroeder FH (1984). Determination of the growth rate of human prostatic cells in primary culture by a morphometric technique. *Cell Biol Int Rep* **8**, 363–371.

[22] Nannya Y, Sanada M, Nakazaki K, Hosoya N, Wang L, Hangaishi A, Kurokawa M, Chiba S, Bailey DK, Kennedy GC, et al. (2005). A robust algorithm for copy number detection using high-density oligonucleotide single nucleotide polymorphism genotyping arrays. *Cancer Res* **65**, 6071–6079.

[23] Lin M, Wei LJ, Sellers WR, Lieberfarb M, Wong WH, and Li C (2004). dChipSNP: significance curve and clustering of SNP-array–based loss-of-heterozygosity data. *Bioinformatics* **20**, 1233–1240.

[24] Chang BL, Liu W, Sun J, Dimitrov L, Li T, Turner AR, Zheng SL, Isaacs WB, and Xu J (2007). Integration of somatic deletion analysis of prostate cancers and germline linkage analysis of prostate cancer families reveals two small consensus regions for prostate cancer genes at 8p. *Cancer Res* **67**, 4098–4103.

[25] Borre M, Høyer M, Nerstrøm B, and Overgaard J (1998). DNA ploidy and survival of patients with clinically localized prostate cancer treated without intent to cure. *Prostate* **36**, 244–249.

[26] Mellor HR and Harris AL (2007). The role of the hypoxia-inducible BH3-only proteins BNIP3 and BNIP3L in cancer. *Cancer Metastasis Rev* **26**, 553–566.

[27] Fei P, Wang W, Kim SH, Wang S, Burns TF, Sax JK, Buzzai M, Dicker DT, McKenna WG, Bernhard EJ, et al. (2004). *Bnip3L* is induced by p53 under hypoxia, and its knockdown promotes tumor growth. *Cancer Cell* **6**, 597–609.

[28] Niedergethmann M, Alves F, Neff JK, Heidrich B, Aramin N, Li L, Pilarsky C, Grutzmann R, Allgayer H, Post S, et al. (2007). Gene expression profiling of liver metastases and tumour invasion in pancreatic cancer using an orthotopic SCID mouse model. *Br J Cancer* **97**, 1432–1440.

[29] Calvisi DF, Ladu S, Gorden A, Farina M, Lee JS, Conner EA, Schroeder I, Factor VM, and Thorgeirsson SS (2007). Mechanistic and prognostic significance of aberrant methylation in the molecular pathogenesis of human hepatocellular carcinoma. *J Clin Invest* **117**, 2713–2722.

[30] Jin Y, Mertens F, Kullendorff CM, and Panagopoulos I (2006). Fusion of the tumor-suppressor gene *CHEK2* and the gene for the regulatory subunit B of protein phosphatase 2 *PPP2R2A* in childhood teratoma. *Neoplasia* **8**, 413–418.

[31] Jarrard DF, Bova GS, Ewing CM, Pin SS, Nguyen SH, Baylin SB, Cairns P, Sidransky D, Herman JG, and Isaacs WB (1997). Deletional, mutational, and methylation analyses of *CDKN2* (*p16/MTS1*) in primary and metastatic prostate cancer. *Genes Chromosomes Cancer* **19**, 90–96.

[32] Mangold KA, Takahashi H, Brandigi C, Wada T, Wakui S, Furusato M, Boyd J, Chandler FW, and Allsbrook WC Jr (1997). *p16* (*CDKN2/MTS1*) gene deletions are rare in prostatic carcinomas in the United States and Japan. *J Urol* **157**, 1117–1120.

[33] Heidenreich B, Heidenreich A, Sesterhenn A, Srivastava S, Moul JW, and Sesterhenn IA (2000). Aneuploidy of chromosome 9 and the tumor suppressor genes *p16(INK4)* and *p15(INK4B)* detected by *in situ* hybridization in locally advanced prostate cancer. *Eur Urol* **38**, 475–482.

[34] Komiya A, Suzuki H, Aida S, Yatani R, and Shimazaki J (1995). Mutational analysis of *CDKN2* (*CDK4I/MTS1*) gene in tissues and cell lines of human prostate cancer. *Jpn J Cancer Res* **86**, 622–625.

[35] Chen W, Weghorst CM, Sabourin CL, Wang Y, Wang D, Bostwick DG, and Stoner GD (1996). Absence of *p16/MTS1* gene mutations in human prostate cancer. *Carcinogenesis* **17**, 2603–2607.

[36] Konishi N, Nakamura M, Kishi M, Nishimine M, Ishida E, and Shimada K (2002). Heterogeneous methylation and deletion patterns of the INK4a/ARF locus within prostate carcinomas. *Am J Pathol* **160**, 1207–1214.

[37] Yamamoto M, Metoki R, and Yamamoto F (2004). Systematic multiplex polymerase chain reaction and reverse transcription–polymerase chain reaction analyses of changes in copy number and expression of proto-oncogenes and tumor suppressor genes in cancer tissues and cell lines. *Electrophoresis* **25**, 3349–3356.

[38] Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A, Farrington SM, Prendergast J, Olschwang S, Chiang T, Crowdy E, et al. (2007). Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* **39**, 989–994.

[39] Zheng SL, Sun J, Cheng Y, Li G, Hsu FC, Zhu Y, Chang BL, Liu W, Kim JW, Turner AR, et al. (2007). Association between two unlinked loci at 8q24 and prostate cancer risk among European Americans. *J Natl Cancer Inst* **99**, 1525–1533.

[40] Korenchuk S, Lehr JE, MClean L, Lee YG, Whitney S, Vessella R, Lin DL, and Pienta KJ (2001). VCaP, a cell-based model system of human prostate cancer. *In Vivo* **15**, 163–168.

[41] Sun J, Liu W, Adams TS, Sun J, Li X, Turner AR, Chang B, Kim JW, Zheng SL, Isaacs WB, et al. (2007). DNA copy number alterations in prostate cancers: a combined analysis of published CGH studies. *Prostate* **67**, 692–700.

Figure W1. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

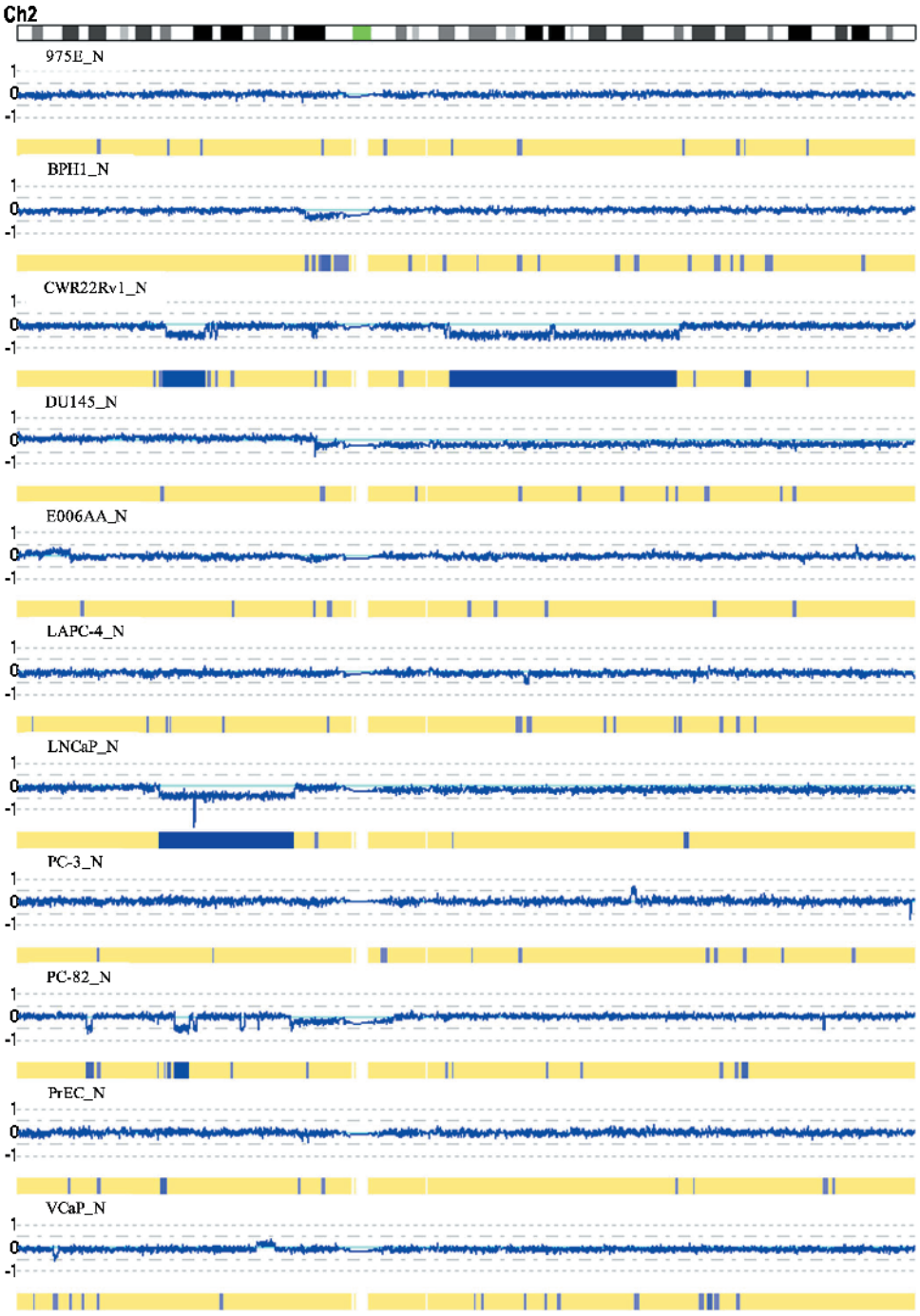


Figure W2. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

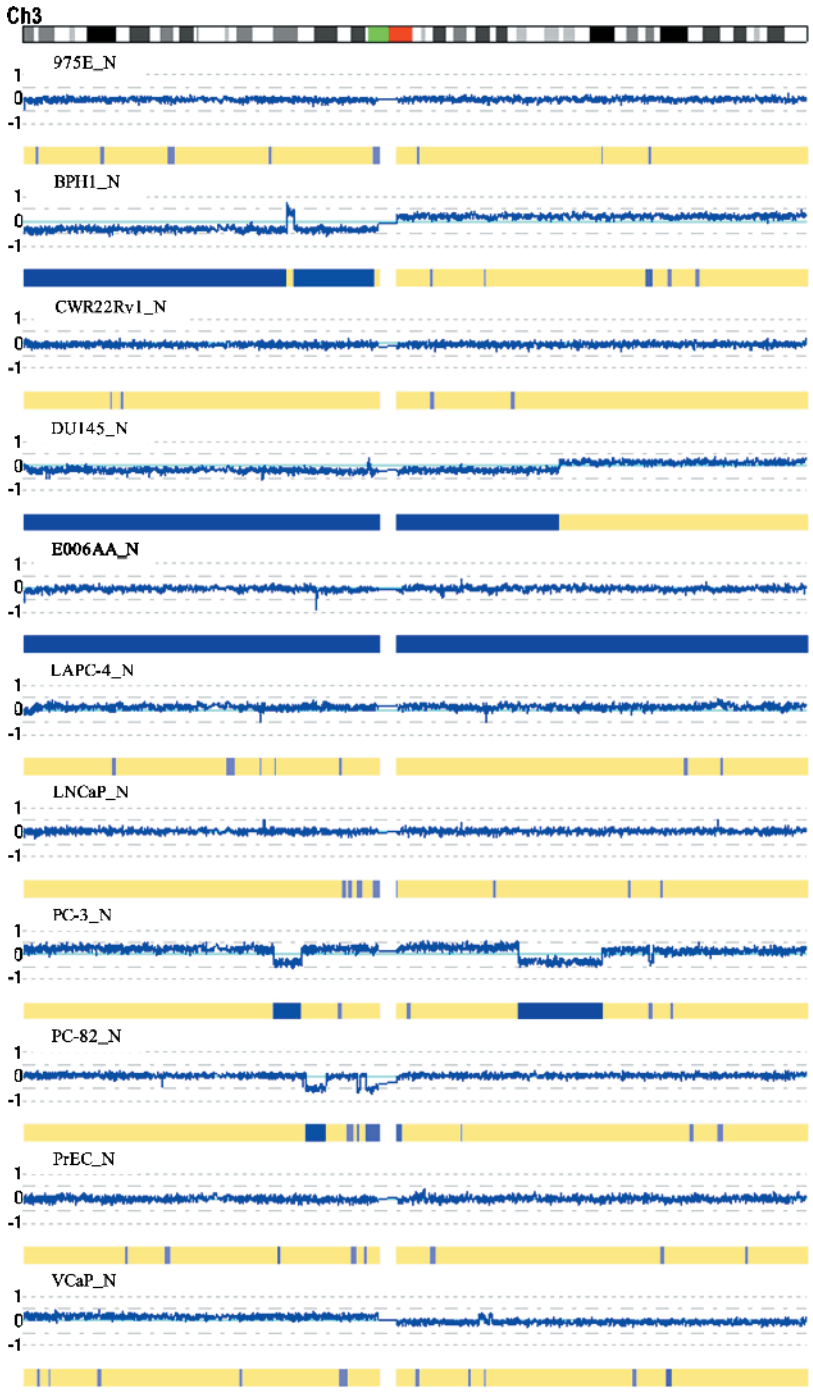


Figure W3. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

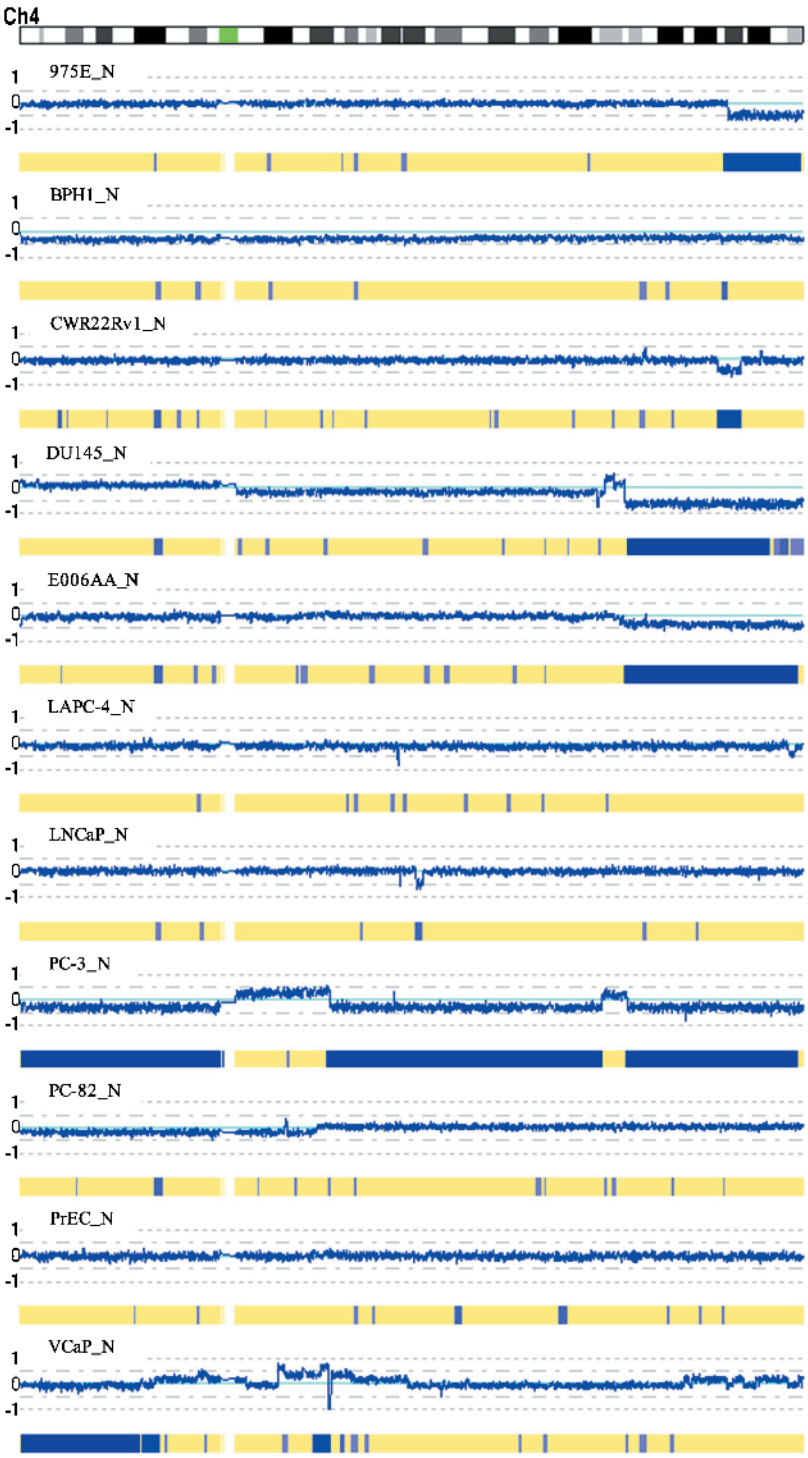


Figure W4. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

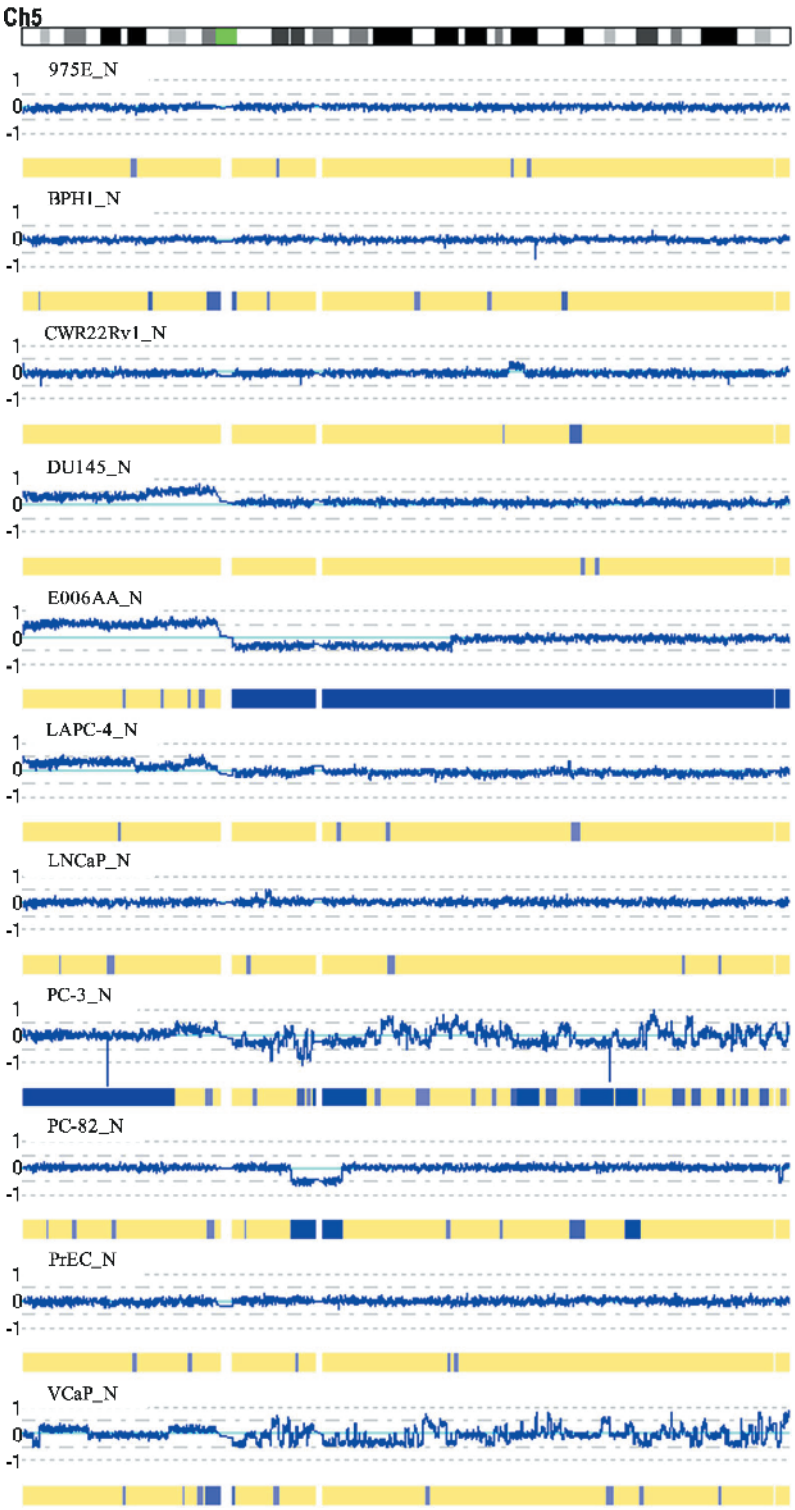


Figure W5. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

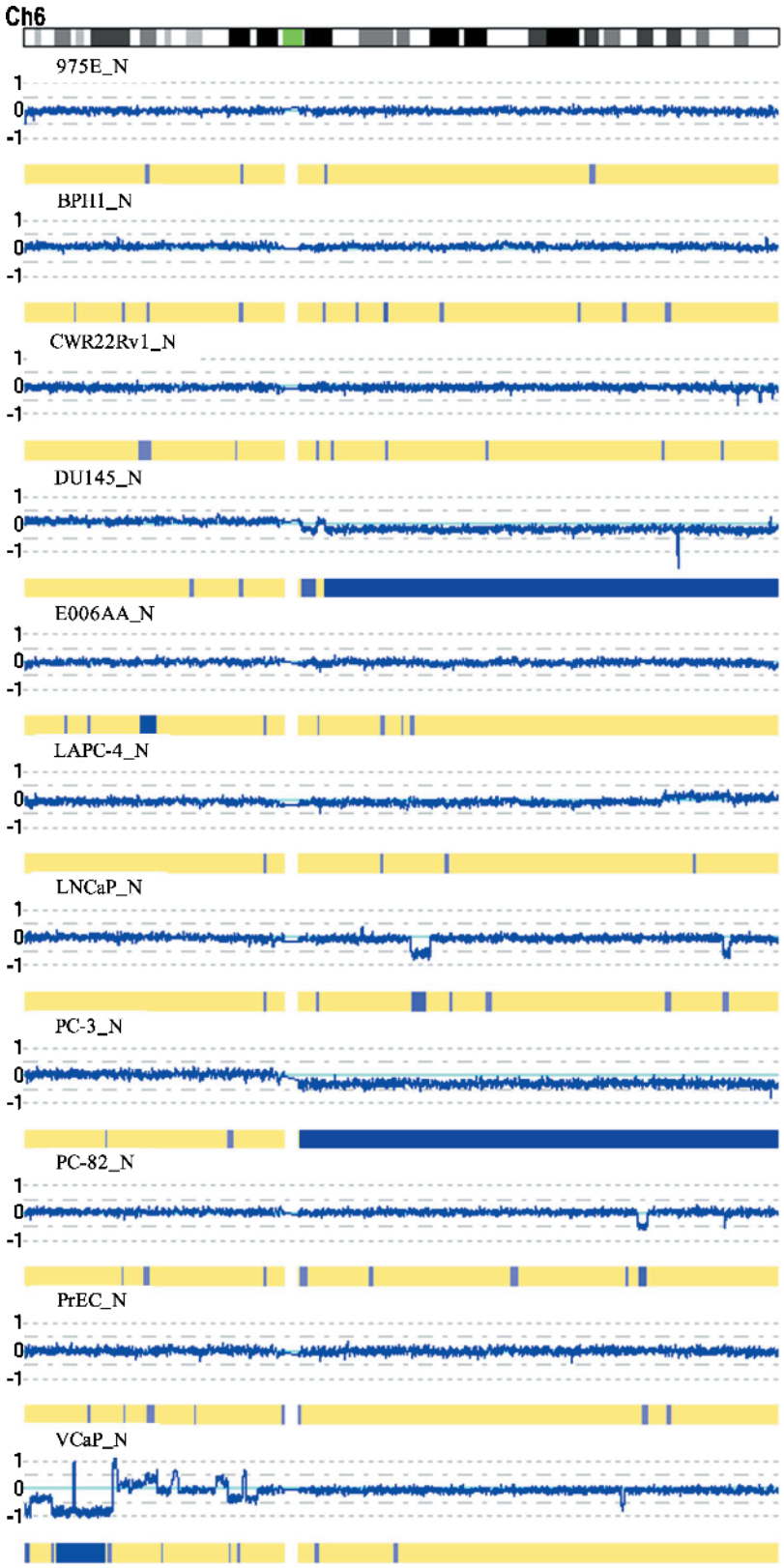


Figure W6. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

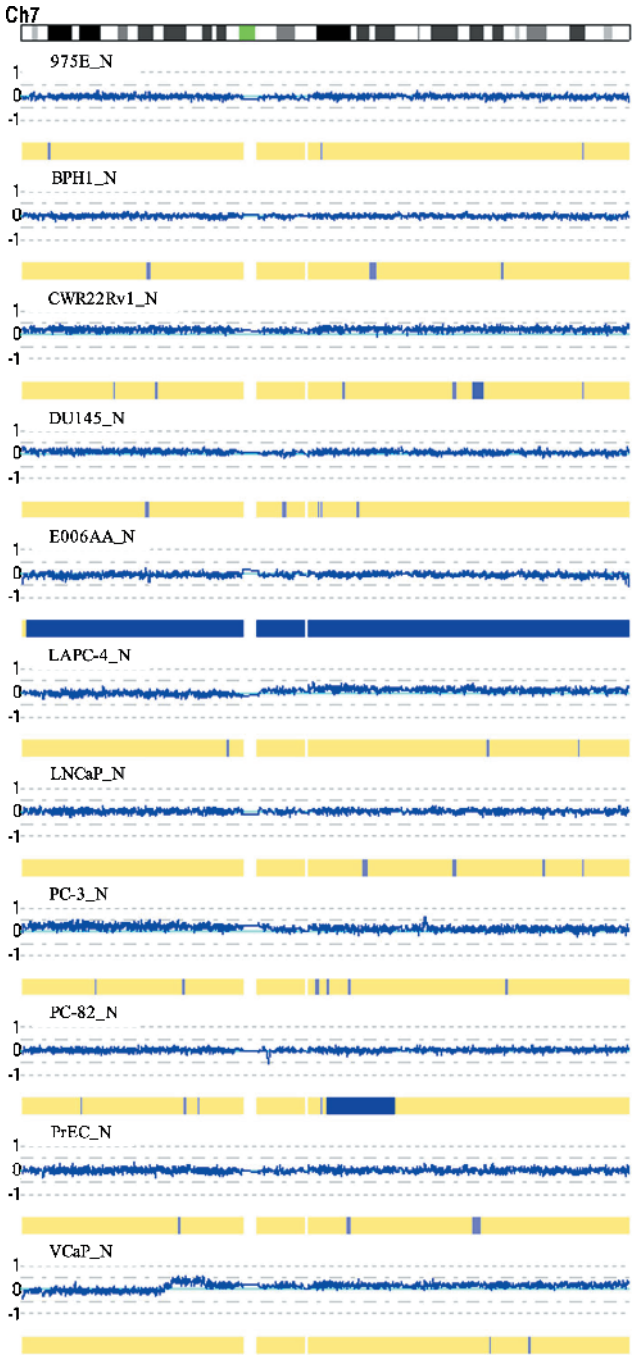


Figure W7. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

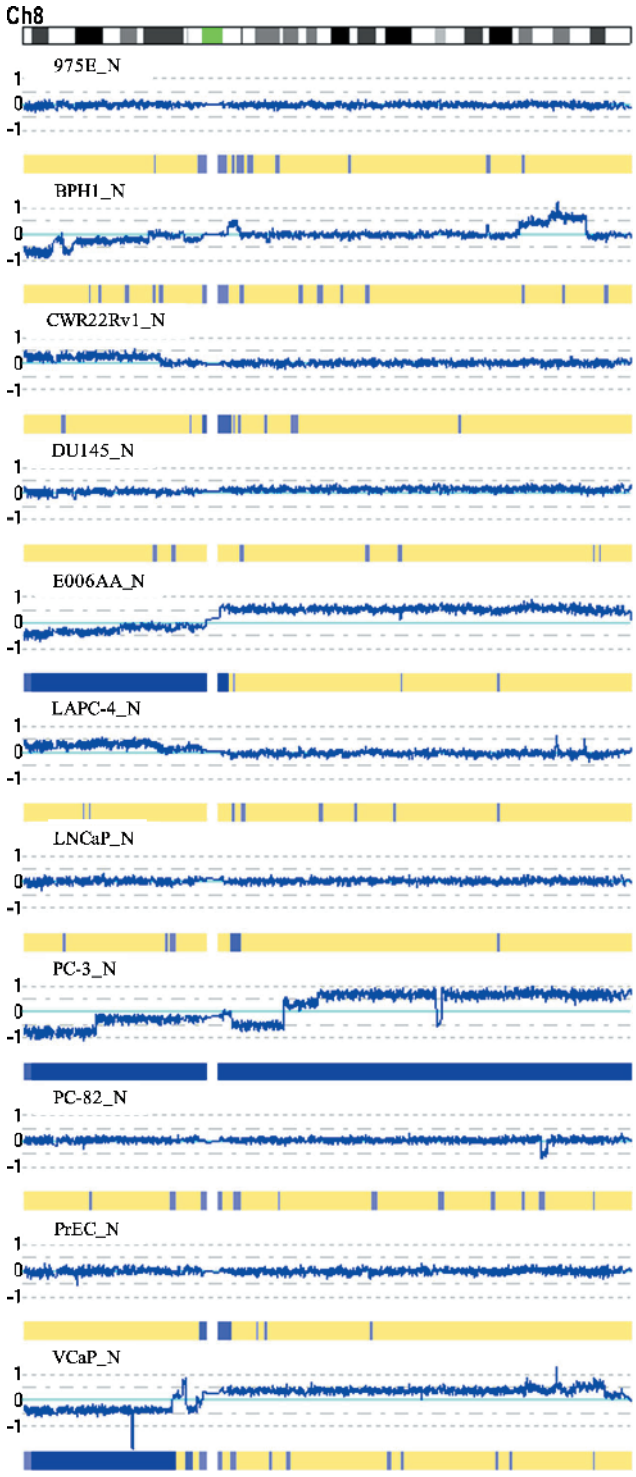


Figure W8. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

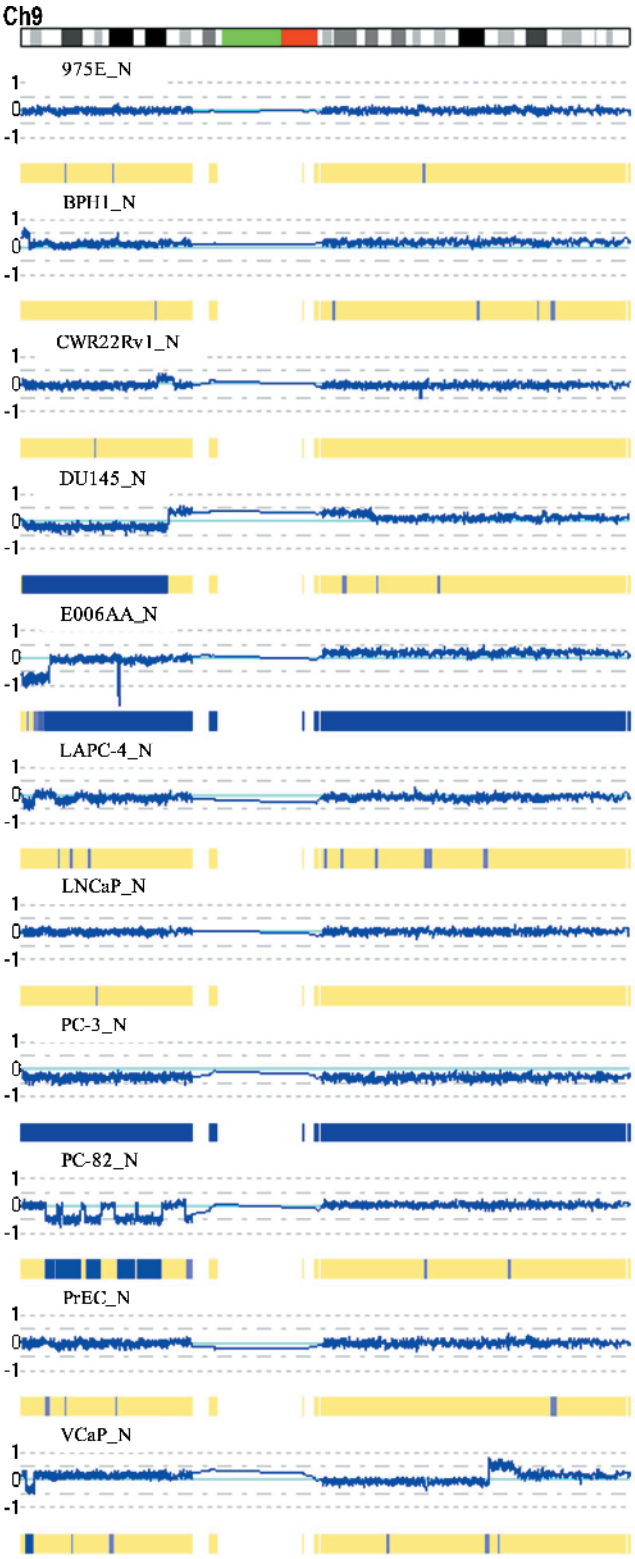


Figure W9. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

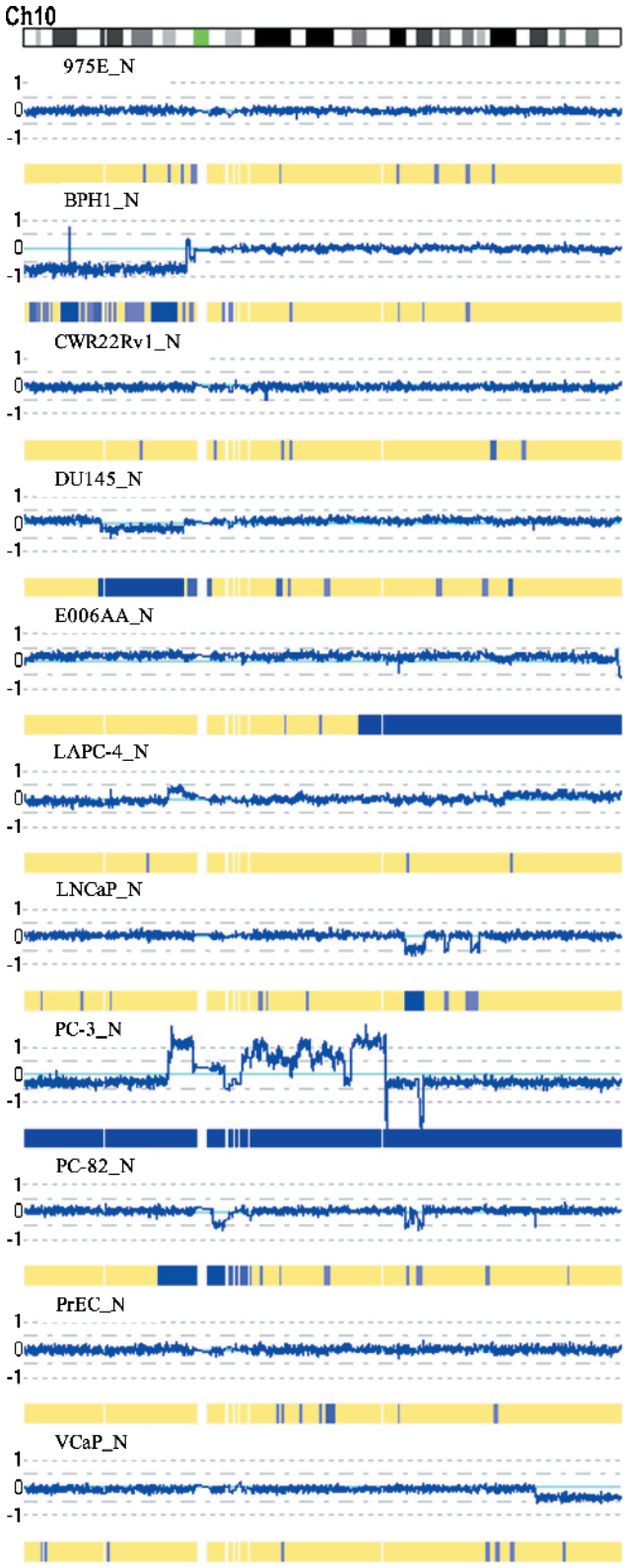


Figure W10. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

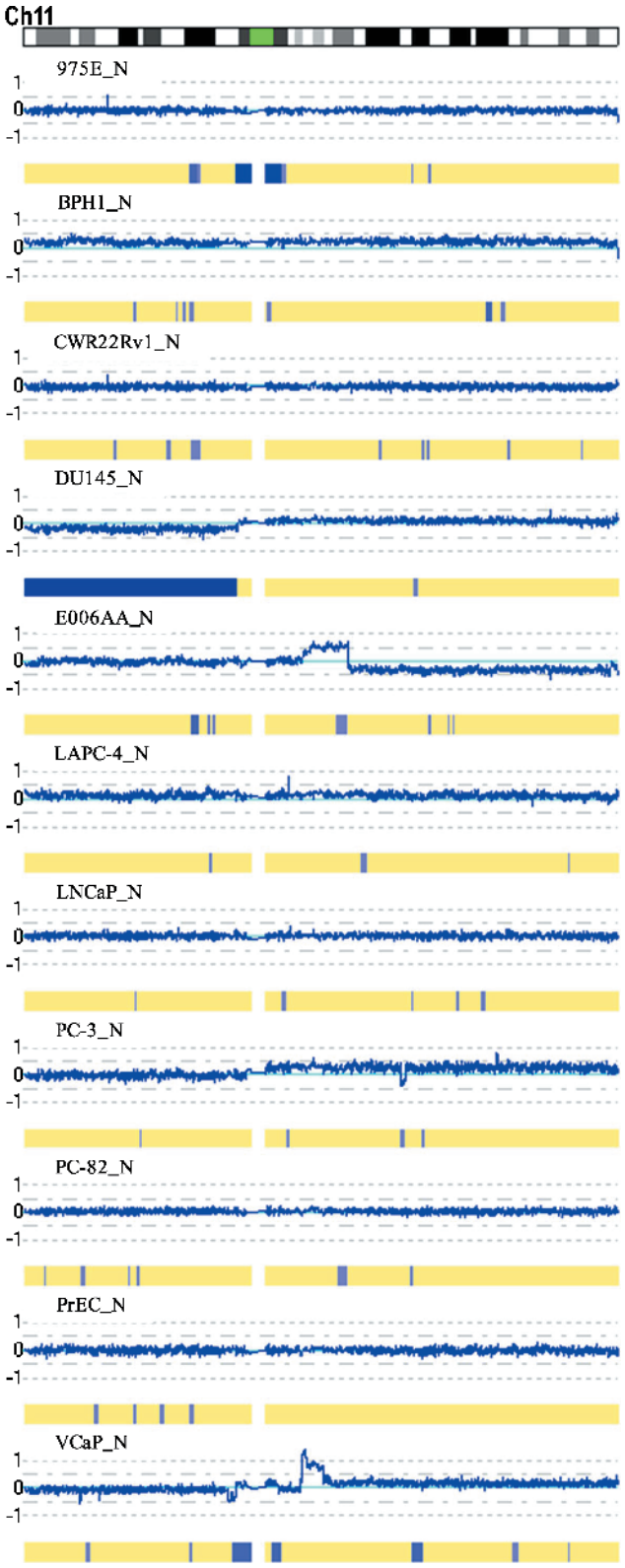


Figure W11. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

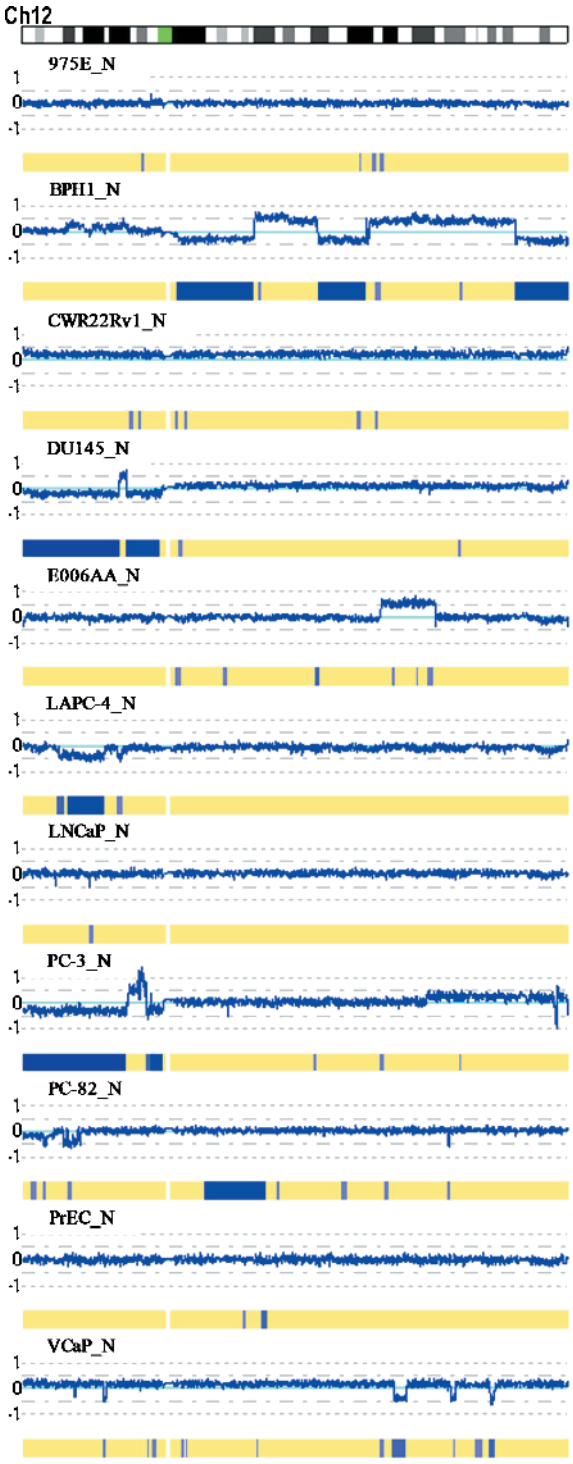


Figure W12. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

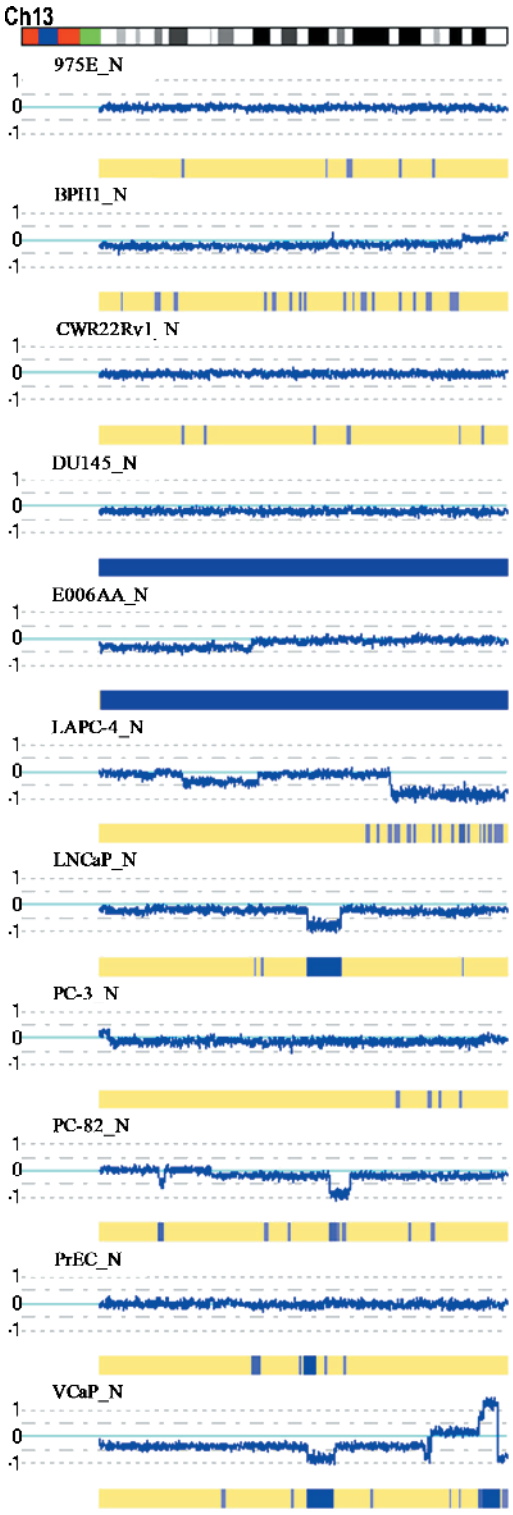


Figure W13. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

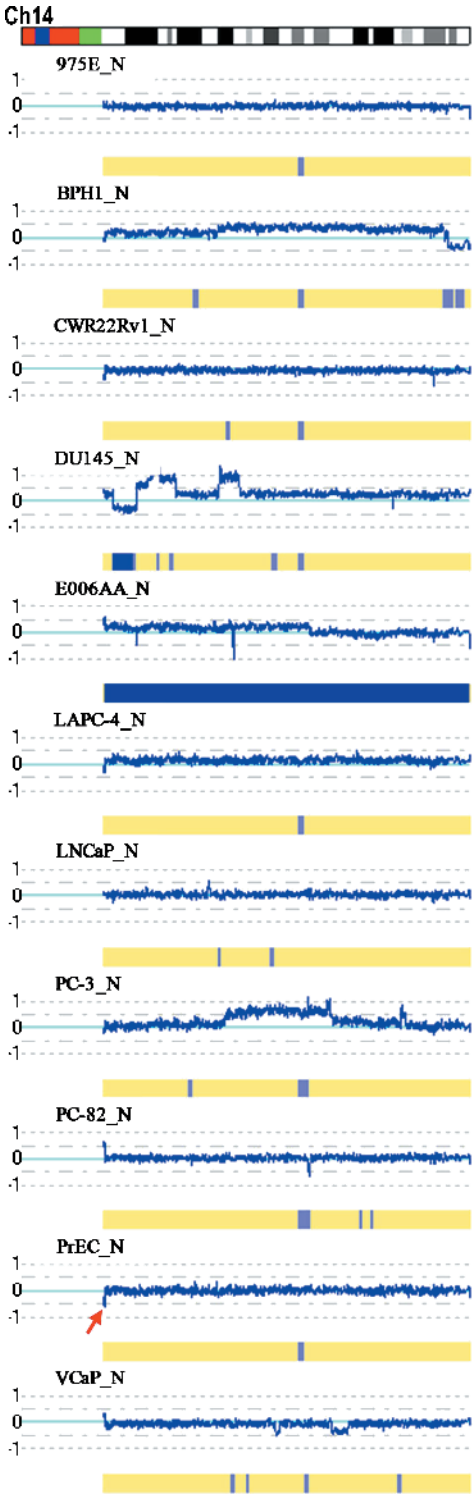


Figure W14. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

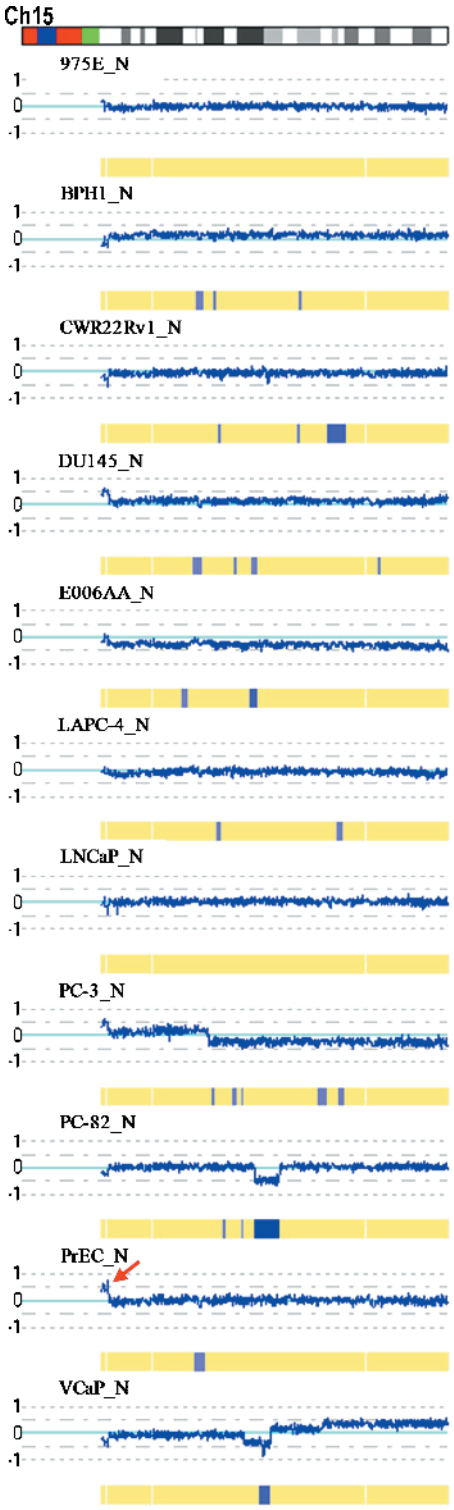


Figure W15. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

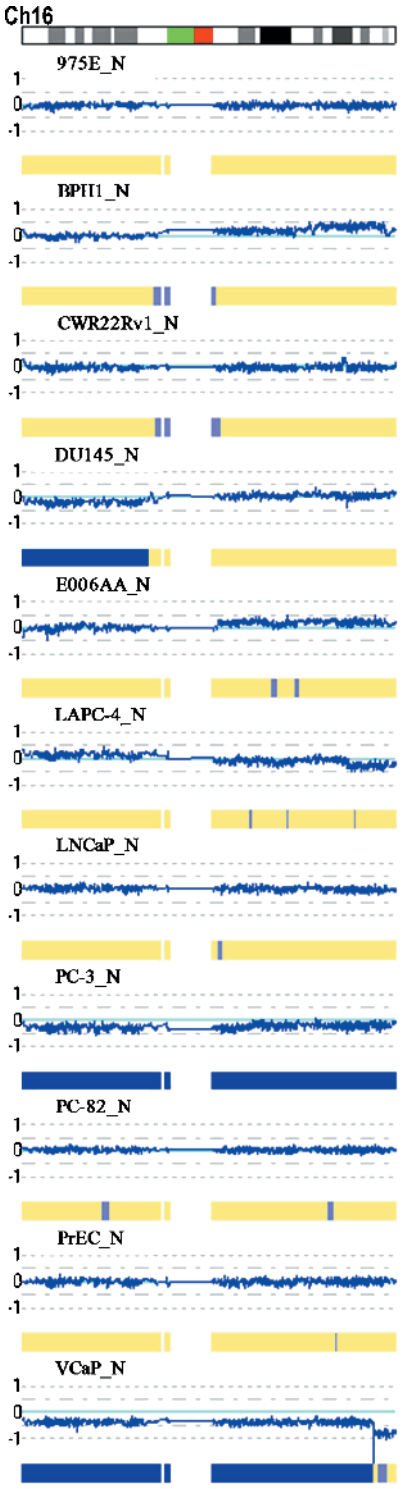


Figure W16. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

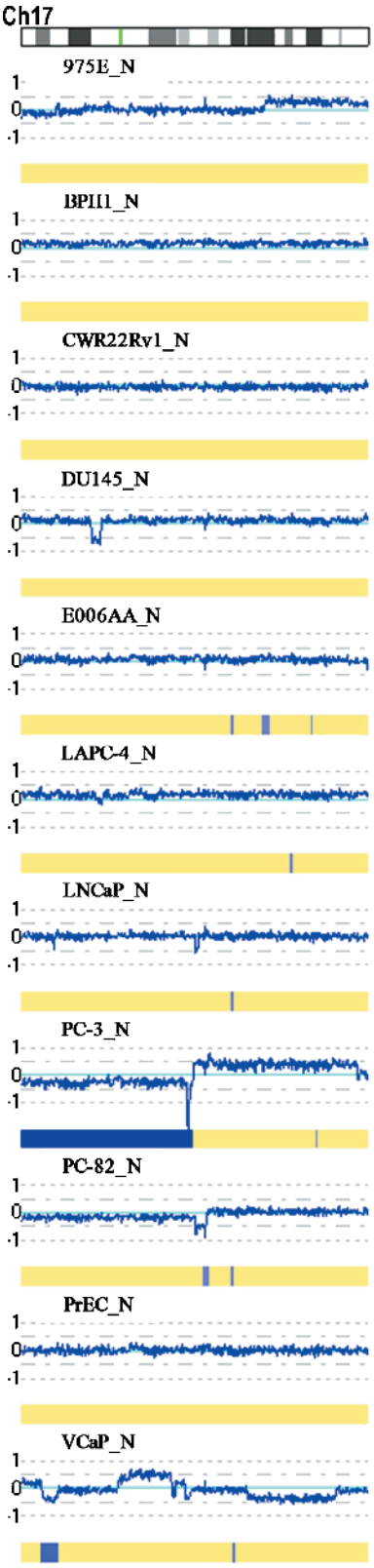


Figure W17. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

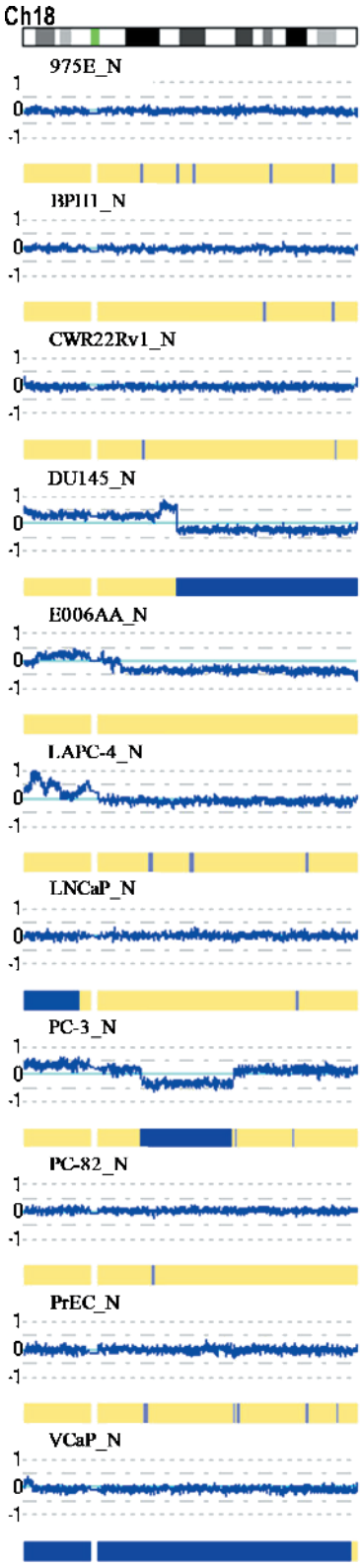


Figure W18. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

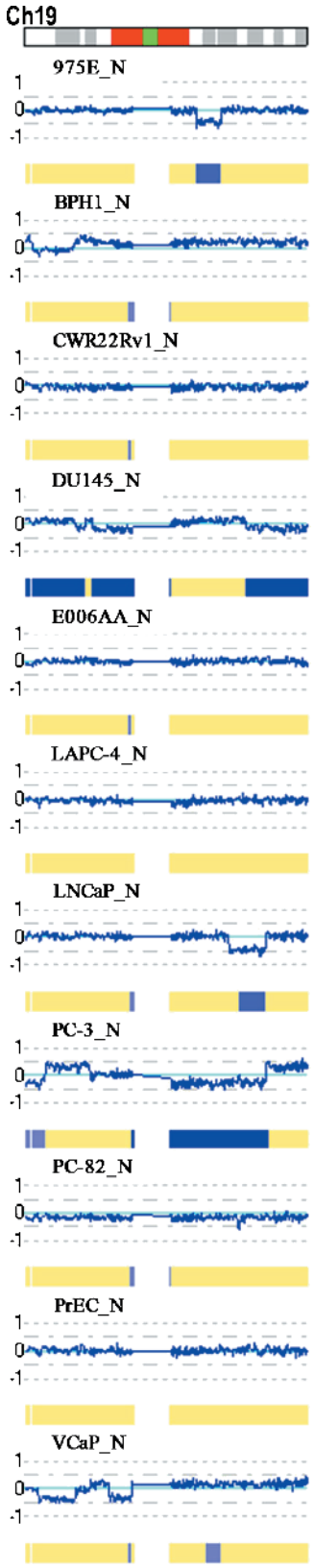


Figure W19. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

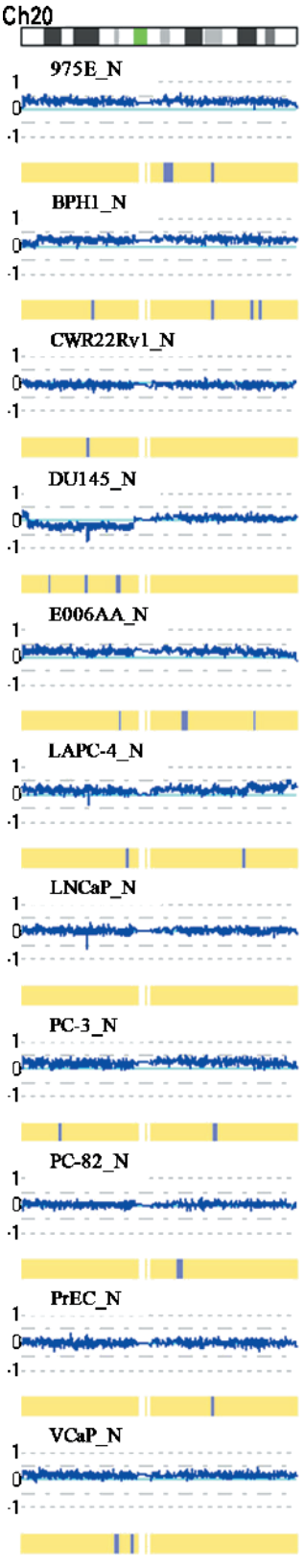


Figure W20. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

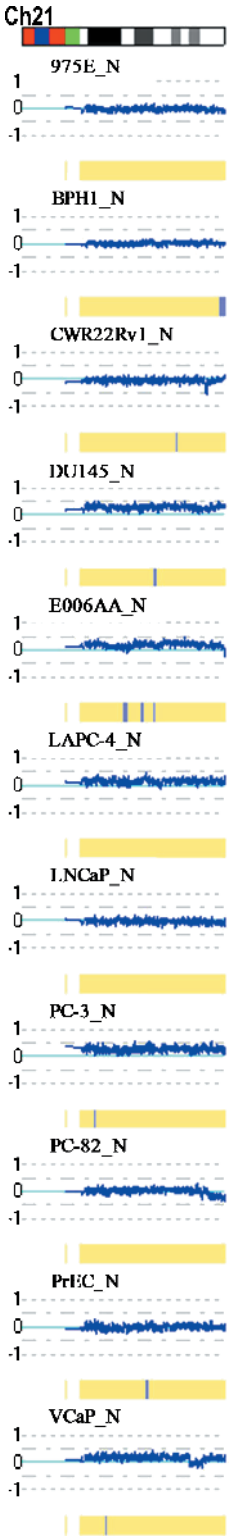


Figure W21. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

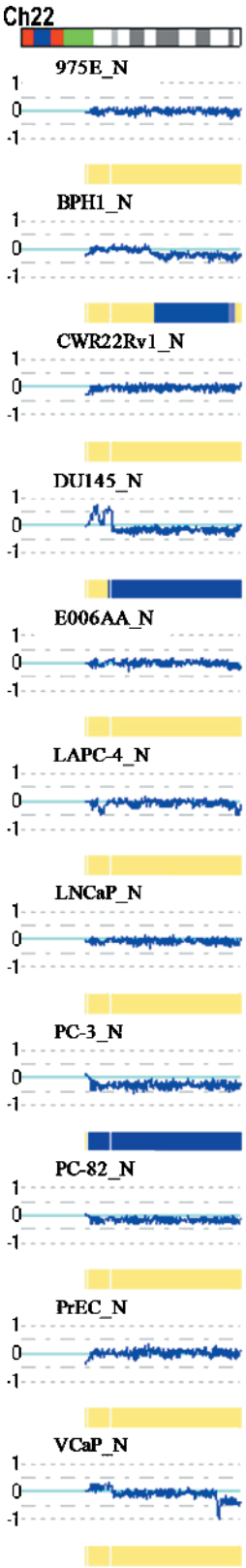


Figure W22. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.

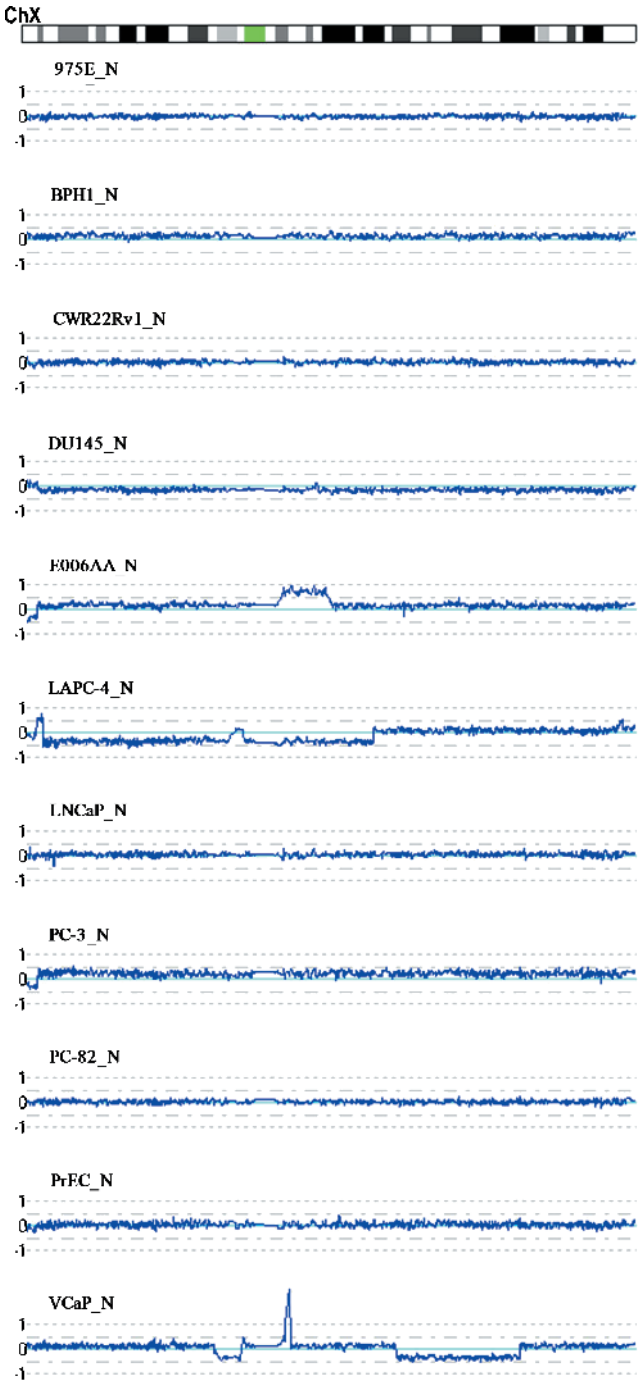


Figure WX. Maps of DNA copy number alterations and LOH generated by CNAG2.0 for chromosomes 1 (W1) to X (WX) in prostate cell lines. Log_2 ratios (horizontal dotted lines) are labeled for each of the chromosomes on the left as 1, 0, and −1, with 0 being the baseline (light blue). Dark blue curve indicates 10-SNP genomic smoothed log_2 ratios of Nsp probes. Blue bars superimposed on the yellow bars represent LOH likelihood with the gradients as shown in Figure 1. Red arrows indicate DNA copy number variations.