Gene Dosage Alterations Revealed by cDNA Microarray Analysis in Cervical Cancer: Identification of Candidate Amplified and Overexpressed Genes

Gopeshwar Narayan,1 Veronique Bourdon,2 Seeta Chaganti,3 Hugo Arias-Pulido,1,4 Subhadra V. Nandula,1 Pulivarthi H. Rao,5 Lutz Gissmann,6 Matthias Dürst,7 Achim Schneider,8 Bhavana Pothuri,9 Mahesh Mansukhani,1 Katia Basso,10 R. S. K. Chaganti,2 and Vundavalli V. Murty1,10∗

1Department of Pathology, Columbia University Medical Center, NY10032
2Cell Biology Program, Memorial Sloan-Kettering Cancer Center, NY10021
3Department of Molecular Genetics and Microbiology, University of New Mexico, Albuquerque, NM
4Department of Molecular Genetics and Microbiology, University of New Mexico, Albuquerque, NM
5Baylor College of Medicine, Texas Children’s Cancer Center, Houston, TX
6Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 242, Heidelberg, Germany
7Department of Obstetrics and Gynecology, Friedrich Schiller University, Jena, Germany
8Department of Obstetrics and Gynecology, Charité Universitätsmedizin Berlin, Hindenburgdamm 30, Berlin, Germany
9Department of Gynecologic Oncology, Columbia University Medical Center, NY10032
10Institute for Cancer Genetics, Columbia University Medical Center, NY10032

Cervical cancer (CC) cells exhibit complex karyotypic alterations, which is consistent with deregulation of numerous critical genes in its formation and progression. To characterize this karyotypic complexity at the molecular level, we used cDNA array comparative genomic hybridization (aCGH) to analyze 29 CC cases and identified a number of over represented and deleted genes. The aCGH analysis revealed at least 17 recurrent amplicons and six common regions of deletions. These regions contain several known tumor-associated genes, such as those involved in transcription, apoptosis, cytoskeletal remodeling, ion-transport, drug metabolism, and immune response. Using the fluorescence in situ hybridization (FISH) approach we demonstrated the presence of high-level amplifications at the 8q24.3, 11q22.2, and 20q13 regions in CC cell lines. To identify amplification-associated genes that correspond to focal amplicons, we examined one or more genes in each of the 17 amplicons by Affymetrix U133A expression arrays and semiquantitative reverse-transcription PCR (RT-PCR) in 31 CC tumors. This analysis has enabled the identification of focal amplicons characteristic for the CC genome and facilitated the validation of relevant genes in these amplicons. These data, thus, form an important step toward the identification of biologically relevant genes in CC pathogenesis. This article contains Supplementary Material available at http://www.interscience.wiley.com/jpages/1045-2257/suppmat.
of 3q, occur very early in the progression of CC (Heselmeyer et al., 1996; Kersemakers et al., 1999; Chatterjee et al., 2001). However, the molecular nature of the global genetic changes recognized as complex cytogenetic alterations in CC remains poorly understood. Elucidation of these changes is critical for understanding the molecular basis of CC.

In an effort to identify the molecular alterations associated with invasive CC, we performed microarray CGH (aCGH) analysis to identify gene dosage changes in CC. This analysis identified 17 amplified and six deleted chromosomal regions characteristic of CC. FISH analysis demonstrated high-level amplifications at 8q24.3, 11q22.2, and 20q13, and increased copies of 3q. Affymetrix gene expression profile and RT-PCR analyses identified overexpression of a number of genes mapped within the focal amplicons in CC and enabled the identification of relevant transcriptional targets.

**MATERIALS AND METHODS**

**Tumor Specimens and Cell Lines**

A total of 54 CC cases (9 cell lines and 45 primary tumor specimens) and 16 normal cervical tissues obtained from hysterectomy specimens as controls were used in this study. Twenty-nine tumors (21 primary tumors—6 stage IB, 8 stage IIB, 7 stage IIIB; and 8 cell lines) were used in aCGH analysis and 25 additional tumors (24 primary tumors and one cell line) were used in expression studies. The cell lines (HT-3, ME-180, CaSkii, MS751, C-4I, C-33A, SW756, HeLa, and SiHa) were obtained from American Type Culture Collection (ATCC, Manassas, VA) and grown in tissue culture as per the supplier’s specifications. The tumors were obtained from patients evaluated at the Instituto Nacional de Cancerologia (Santa Fe de Bogota, Colombia) (Pulido et al., 2000), the Department of the Obstetrics and Gynecology of Friedrich Schiller University (Jena, Germany), and Columbia University Medical Center, NY. All specimens were obtained after appropriate informed consent and approval of protocols by institutional review boards. The primary tumors were classified as FIGO stage IB (8 tumors), IIB (18 tumors) or IIIB/IV (19 tumors). Forty-two tumors were diagnosed as squamous-cell carcinoma (SCC) and three as adenocarcinoma. All tumor specimens were determined to contain at least 70% tumor cellularity by H&E staining. High molecular weight DNA and total RNA from tumor and normal tissues, and cell lines were isolated by standard methods. DNA isolated from placenta was used as a reference in aCGH analysis.

**aCGH Hybridization and Image Analysis**

The EST arrays generated at the Albert Einstein College of Medicine microarray facility (www.aecom.yu.edu/cancer/new/cores/microarray/default.htm#) contained 9,206 T3/T7 PCR-amplified cDNA inserts of human IMAG.E. consortium clones printed on glass slides. The slides were hybridized as previously described (Bourdon et al., 2002). Briefly, the slides were first incubated for 1–5 hr with 20 μl of prehybridization mix. A total of 5 μg of test and reference DNAs was digested with DpnII for 1 hr, purified using a PCR clean-up kit (Promega, Madison, WI), and extracted in 50 μl of water. Digested DNAs were concentrated by ultrafiltration (Microcon YM-30, Amicon; Millipore, Bedford, MA). Equal amounts of test and reference (placenta) DNAs (1.8–2.2 μg) were labeled separately in 50 μl reactions using Cy3-dUTP or Cy5-dUTP (Amersham, Piscataway, NJ), respectively. The reaction mixtures were pooled, purified, and hybridized in the presence of a blocking reagent (Pollack et al., 1999). The slides were prepared after posthybridization washes as described (Bourdon et al., 2002).

The arrays were scanned using an Axon dual color laser scanner (GenePix 4000A; Axon, Union City, CA). At the time of the scanning, the laser power was adjusted to have <5% features saturated; the digitized Cy3 and Cy5 signals were pseudocolored green and red, respectively (GenePix Pro 3.0; Axon). After gridding, each dot on the 24-bit ratio image was visually inspected and unsatisfactory dots were manually flagged if necessary. A GenePix results (*.gpr) file of the raw data (F635 median-B635 median, F532 median-B532 median) was used for further analyses.

The signals obtained after laser excitation of the dyes were digitized, and the raw data (median feature pixel intensity with the median local background intensity subtracted at each wavelength) were then subjected to statistical analysis. To correct for systematic errors introduced by the intensity-dependent dye efficiencies, the hybridization signal data from each slide were normalized using a local regression of the log-ratio variable $Y = \log_2(G/R)$ versus the log-product $X = \log_{10}(R \times G)$ / 2 (R and G represent the intensities of the Cy3 and Cy5, respectively). It was important to construct an indicator to identify ESTs that exhibited significant signal deviation from normal in a given slide. To this end, we computed the intensity-dependent
(local) variance $\sigma(X_N)^2$ from a local regression of $Y_N^2$ vs. $X_N$ after normalization ($X_N$ and $Y_N$ represent the normalized $X$ and $Y$ variables) and attributed significance to amplified/deleted ESTs according to the values of $Y_N/\sigma(X_N)$ (LR/SD), independently for each slide (Bourdon et al., 2002). With the binomial distribution it is extremely unlikely to get more than two false positive calls out of 29 samples with a $P < 0.001$. Therefore, a sequence was called amplified or deleted when the value $LR/SD$ was $\geq 3.1$ or $\leq -3.1$, respectively, in at least three tumor samples and none of the controls (three placenta versus placenta experiments). All EST clones on the array were mapped in silico using NCBI genome map viewer build 34.3 (www.ncbi.nlm.nih.gov/mapview/) and assigned to subchromosomal regions. The normalized data have been deposited in the Gene Expression Omnibus (GEO) database (Accession GSE1715) (www.ncbi.nlm.nih.gov/geo/).

**HPV Typing**

HPV types were determined as previously described (Narayan et al., 2003a).

**Semiquantitative RT-PCR Analysis**

Total RNA from normal cervix was obtained from three commercial sources (Ambion, Austin, TX; Stratagene, La Jolla, CA; BioChain, Hayward, CA). Total RNA was isolated from nine cell lines (eight used in aCGH analysis), 18 primary tumors (all SCC; nine of these also studied by aCGH), and five normal cervix were reverse transcribed using random primers and the Pro-STAR first strand RT-PCR kit (Stratagene, La Jolla, CA). A semiquantitative analysis of gene expression was performed in duplicate or triplicate experiments using 26–28 cycles of multiplex RT-PCR with $\beta$-actin ($ACTB$) as the control and gene specific primers spanning at least two exons (Supplementary Table 1; supplementary material for this article can be found at http://www.interscience.wiley.com/jpages/1045-2257/suppimat).

The PCR products were run on 1.5% agarose gels, visualized by ethidium bromide staining and quantified using the Kodak Digital Image Analysis System (Kodak, New Haven, CT). The values obtained for each gene were normalized against $ACTB$. For each gene, at least three different normal cervix RNA samples were used to calculate the mean and SD. A gene was considered upregulated if the gene/control ratio was $\geq$ mean + 2 SD of the normal cervix.

**Oligonucleotide Microarray Gene Expression Analysis**

Biotinylated cRNA preparation and hybridization to Affymetrix U133A oligonucleotide microarray (Affymetrix, Santa Clara, CA), which contains 14,500 genes was performed on 22 primary CC cases (only one of these cases was studied by aCGH), nine CC cell lines (eight were studied by aCGH), and 16 normal cervical epithelium specimens by the standard protocols supplied by the manufacturer. Arrays were subsequently developed and scanned to obtain quantitative gene expression levels. Expression values for the genes were determined using the Affymetrix GeneChip Operating Software (GCOS) and the Global Scaling option, which allows a number of experiments to be normalized to one target intensity to account for the differences in global chip intensity. To perform the supervised gene expression analysis, we used the Genes@Work software platform, which is a gene expression analysis tool based on the pattern discovery algorithm SPLASH (Structural Pattern Localization Analysis by Sequential Histograms) (Califano, 2000).

**Fluorescence In Situ Hybridization**

FISH was performed by standard methods on chromosomes prepared from eight CC cell lines. DNA prepared from human BAC clones RP11-750P5 (11q22.2), RP11-480A16 (3q29), RP11-374B7 (8q24.3), and RP11-30F23 (20q13.1) (Open Biosystems, Huntsville, Alabama) was labeled by nick-translation using spectrum red or spectrum green dUTP fluorochromes (Vysis, Downers Grove, IL). Spectrum red or spectrum green-labeled centromeric probes (CEP) were obtained from Vysis (Downers Grove, IL). Hybridization signals were scored on at least 20 metaphase spreads on DAPI counterstained slides.

**RESULTS**

Our previous molecular cytogenetic analyses of CC have identified complex chromosome alterations that include recurrent sites of high-level amplifications, +3q, and del(2q) (Harris et al., 2003; Narayan et al., 2003b; Rao et al., 2004). To characterize this karyotypic complexity at the molecular level, we performed cDNA array CGH (aCGH) analysis of a series of 29 CC cases that included 8 cell lines and 21 primary tumor biopsies. Of these, 27 (91%) were HPV positive (20 with HPV16/18; 7 harbored other HPV types) and two were HPV negative. Among the 9,206-cDNA
sequences on the array, 445 (64.1 ± 35.2/tumor) were found to be over represented and 121 (16.8 ± 13.9/tumor) were deleted. A gene was considered either gained or lost if present in >10% of tumors based on the criteria described in the methods. Although the frequency of gene copy number gains was similar in cell lines and primary tumors, deletions were more common in cell lines than in primary tumors (28.4 ± 16.1 vs. 12.3 ± 10.1; P = 0.003). Representative examples of genomic alteration profiles of a cell line (CaSki) and primary tumor (T-982) are shown in Figures 1A and 1B. The data from the cell lines and primary tumors were combined in all subsequent analyses.

Identification of Increased Gene Dosage Profiles

The majority of the 445 cDNA clones with an increased dosage were mapped to few chromosomes, such as X (15%), 1 (13%), 5 (9%), 3 (6%), 19 (6%), and 20 (4%) (Fig. 1C) (Supplementary Table 2A; supplementary material for this article can be found at http://www.interscience.wiley.com/jpages/1045-2257/suppmat). Therefore, the distribution of amplified clones in the chromosomal complement was nonrandom. The top 124 clones amplified in five or more tumors were associated with distinct cellular phenotypes, such as transcriptional regulation, cell cycle, cytoskeleton remodeling, apoptosis, angiogenesis, mitochondrial, ribosomal, immune

![Copy number alteration profiles in cervical cancer.](image)
response, drug metabolism, and ion transport (Table 1) (Supplementary Table 2B).

**Identification of Genes with Copy Number Deletions**

A total of 121 cDNAs were under represented in the CC genome compared to normal (Fig. 1C). The under represented genes will be referred as deleted genes hereafter. The deleted cDNA clones were found to be preferentially localized to chromosomes 4 (21%), 2 (11%), 13q (8%), 8 (8%), 11 (7%), 3p (5%), and 12q (5%). This nonrandom distribution of deleted regions in the genome suggests that these chromosomal regions harbor candidate tumor suppressor genes relevant to CC.

To identify common contiguous regions of deletions on the frequently affected chromosomes, we mapped in silico all of the deleted clones to exact sequence map positions using the NCBI’s MapView genome browser (www.ncbi.nlm.nih.gov/mapview/). This analysis identified 4q13.3 (3 Mb; 4 cDNA clones), 2q33-37 (36 Mb; 10 cDNA clones), 13q14.1 (8 Mb; 5 cDNA clones), 8p23 (6.2 Mb; 3 cDNA clones), 11q11.2-14 (27.2 Mb; 4 cDNA clones).

**Table 1. Genes with Known Function Amplified in Five or More Tumors**

| Gene   | Function                                                                 | Gene   | Function                                                                 |
|--------|--------------------------------------------------------------------------|--------|--------------------------------------------------------------------------|
| TFE3   | Transcription factor binding to IGHM enhancer 3                          | ATP7A  | ATPase, Cu^{2+} transporting, α polypeptide                               |
| TFAP2C | DNA-binding protein                                                       | OCLN   | Regulation of the tight junction (TJ)                                     |
| SMARCD2| SWI/SNF related, matrix associated, actin dependent regulator of chromatin| RANBP2 | Transport factor (RAN-GTP, karyopherin)-mediated protein                  |
| SMARCC2| SWI/SNF related, matrix associated, actin-dependent regulator of chromatin| SCL9A6 | Na(+) and K(+) across the mitochondrial inner membrane                    |
| ZNF161 | Binds to the CT/CG-rich region of the interleukin-3 promoter             | MN     | Crosslinker between plasma membranes and actin-based cytoskeleton        |
| ZFR    | Zinc finger RNA binding protein                                           | MMP1   | Interstitial collagen                                                     |
| CTNNBL1| Bipartite nuclear localization signal and a leucine-isoleucine zipper     | MMP1   | Membrane protein, palmitoylated                                           |
| MYST2  | Histone acetyltransferase                                                 | ITGA2  | Receptor for laminin, collagen, collagen c-propeptides, fibronectin and E-cadherin |
| USP9X  | Preventing degradation of proteins                                        | RANBP2 | Transport factor (RAN-GTP, karyopherin)-mediated protein                  |
| Apoptosis | Induction of apoptosis                                                    | RANBP3| Transport factor (RAN-GTP, karyopherin)-mediated protein                  |
| MAP3K10| Apoptosis-related cysteine protease; promoter of apoptosis               | RDX    | Binding of the barbed end of actin filaments to the plasma membrane      |
| CASP4  | Apoptosis-related cysteine protease; promoter of apoptosis               | SCL9A6 | Na(+) and K(+) across the mitochondrial inner membrane                    |
| PHLDA2 | Implicated in Fas expression and apoptosis                               | EffB1  | Binds to the receptor tyrosine kinases                                     |
| NALP2  | Apoptosis and inflammation                                               | COL16A1| Integrity of the extracellular matrix                                     |
| ZFPM2  | Regulation of transcription from RNA polymerase II promoter              | PTTRF  | Cell–matrix and cytoskeletal rearrangments                                |
| Mitochondrial | Binding of initiation tRNA to mitochondrial 28S ribosomes            |       |                                                                          |
| MTIF2  |                                                                           |        |                                                                          |
| Apoptosis | Induction of apoptosis                                                    | MMP7   | Degrades casein, gelatins, fibronectin, and activates procollagenase     |
| Immune response | Adapter protein of the FYN and SH2-domain                              | AMPH   | Control the properties of the membrane-associated cytoskeleton           |
| FYB    | Adapter protein of the FYN and SH2-domain                                | RDX    | Binding of the barbed end of actin filaments to the plasma membrane      |
| IL7R   | Cytokine–cytokine receptor interaction                                   | SIAHBP1| Fuse-binding protein-interacting repressor                                |
| IK     | Downregulator of HLA II                                                 | SF3A3  | Binding of U2 snRNP to the branchpoint sequence in pre-mRNA              |
| IL13RA2| Protein folding and intracellular transport                              |        |                                                                          |
| IL2RG  | Protein folding and intracellular transport                              | RPS6KA3| Phosphorylates substrates of ribosomal protein S6                         |
| CR1    | Complement component C3b receptor activity                               |        |                                                                          |
| Drug metabolism | Glutathione metabolism                                               | SNRPE  | RNA binding                                                               |
| GSTT1  |                                                                           |        |                                                                          |
| GSTM2  |                                                                           |        |                                                                          |

*Genes, Chromosomes & Cancer DOI 10.1002/gcc*
clones), 3p14.2-21.3 (29 Mb; 4 cDNA clones) as commonly deleted sites. However, no minimal region could be derived from 12q due to dispersed deletions. A number of genes relevant to tumorigenesis maps to the minimal region of deletions. For example, the 2q33-37 regions contain at least two such genes, \textit{SMARCAL1} and \textit{COPS8}. \textit{SMARCAL1} is an SWI/SNF related matrix-associated, actin-dependent regulator of chromatin, and \textit{COPS8} plays a role in chromatin remodeling and DNA repair. The 8p23 deleted region contains CUB and Sushi multiple domains one (\textit{CSMD1}) gene. The 8p23 region containing \textit{CSMD1} has been shown to be homozygously lost in tumors derived from bladder, colon, prostate, and fallopian tube (Blaveri et al., 2005).

The 3p14-21 deletion spans the fragile histidine triad (\textit{FHIT}) gene, which is deleted in multiple cancer types including CC (Sozzi et al., 1998). Of the 121 deleted genes, the top 52 clones with decreased copy numbers in four or more tumors include tumor suppressors (\textit{FHIT}, \textit{ZDHHC2}), and genes related to apoptosis (\textit{CASP1}), transcription regulation, immune response (chemokine ligand 4-like 1; \textit{CCL4L1}), cell-cell interaction, and DNA repair (Table 2).

### Identification of Amplicons

The nonrandom clustering of the majority of over represented clones in the dataset to a few chromosomal regions prompted us to use an objective criterion to identify and define the amplicons. Toward this end, all 445 amplified clones identified by aCGH were mapped in silico to specific chromosomal sites at the sequence level using the MapView browser (http://www.ncbi.nlm.nih.gov/mapview/). A discrete locus of regional copy number increase represented by four or more clones within 15 Mb genomic region in three or more tumors was considered a potential amplicon. This algorithm was highly effective as it identified 17 amplicons (four on X chromosome, three on chromosome 1, and one each at chromosomal regions 3q27.3-29, 5p12-13, 5q31.3, 8q24.3, 11q22-23, 14q32, 17q21-22, 19p13, 19q13.3, and 20q13.1) (Table 3). A number of recurrent over represented cDNA clones, however, remained single genes at their specific chromosomal regions, which requires further confirmation by other methods.

### Expression-Array Validation of Genes in Amplicons

In the present study, we restricted the validation to the genes present within the 17 amplicons identified above. To identify expression profiles of the genes within the amplicons, we used the Affymetrix U133A array data sets derived from 16 normal

| Table 2. List of 33 Genes with Known Function Deleted in Four or More Tumors |
|-----------------------------|----------------------------------|
| **Gene** | **Function** |
| Tumor suppressor genes | | |
| FHIT | Cleaves A-5′-PPP-5′A to yield AMP and ADP |
| ZDHHC2 | Mutated in hepatocellular and colorectal cancer |
| Transcription factors | | |
| SP100 | SP100 nuclear antigen |
| LDB2 | Neuronal and other development |
| NFKB1 | Transcription factor; immune response, apoptosis, and cell-growth |
| MEF2A | Transcription factor, muscle-specific |
| ZNF20 | Regulation of transcription, DNA-dependent |
| Pro-apoptosis | | |
| CASP1 | Promoter of apoptosis, interleukin 1-β specific |
| Immune response | | |
| CCL4L1 | Immunoregulatory and inflammatory processes |
| FCGR3A | FC fragment of IgG, low affinity IIa, receptor for (CD16) |
| GYPA | MN blood group receptors |
| FGB | Polymerization into fibrin and platelet aggregation |
| ESR2 | Immune and inflammatory responses |
| DNA damage and repair | | |
| COPS8 | Regulation of transcription in response to DNA damage |
| Cell-interaction | | |
| SPARC | Antiadhesive extracellular matrix property |
| NRG1 | Ligand for ERBB3 and ERBB4 tyrosine kinase receptors |
| PICAM1 | Mediating cell-adhesion to extracellular matrix |
| Miscellaneous genes | | |
| EV15 | Ecotropic viral integration site 5 |
| MCFD2 | Multiple coagulation factor deficiency 2 |
| TRAK2 | Trafficking protein, kinesin binding 2 |
| MOBP | Myelin-associated oligodendrocyte basic protein |
| GBE1 | Glycogen biosynthesis |
| GABRA2 | Mediating inhibitory neurotransmission |
| UGT2B4 | Starch and sucrose metabolism |
| SULT1E1 | Androgen and estrogen metabolism |
| METAP1 | Methionyl aminopeptidase activity |
| SNX19 | Intracellular signaling |
| GPR109B | G-protein coupled receptor signaling pathway |
| CPB2 | Carboxypeptidase activity |
| TPP2 | Tripeptidyl-peptidase |
| MPI | Fructose and mannose metabolism |
| ELAC1 | elAC homolog 1 |
| BCR | GTPase-activating protein, chronic myeloid leukemia |
| Amplicon | Size (Mb) | Amplified region (kb) | aCGH array amplified genes | Genes overexpressed on affymetrix array | No. | Selected genes of importance to tumorigenesis |
|---------|-----------|-----------------------|----------------------------|-----------------------------------------|-----|---------------------------------------------|
| 1p36    | 10.5      | 15640–25843           | MSTP9, TCEA3                | SPEN, CROCC, PADI3, KIAA0090, CaMKIIN1, EGF4G3, EPHB2, HRNRPR, ID3, AD7c-NT7, TCEB3, FUSIP1, RNUX3 | 13  |                                                                                          |
| 1p34    | 8.0       | 31576–39695           | TDE2L, COL16A1, RBBP4, ZNF258, SF3A3, PABPC4 | LCN7, PTP4A2, TMEM39B, KPN6, MARCKSL1, FLJ10276, RBBP4, YARS, AK2, CDC48, FLJ12666, PABPC4 | 12  |                                                                                          |
| 1q22–q23| 6.6       | 155815–162367         | IFI16, APCS, NNLH1, DDR2   | IFI16, AIM2, PDN2, HSPA6, FCGR3B, DUSP12 | 6   |                                                                                          |
| 3q27–q29| 11.8      | 185391–191797         | ABCF3, SFP2, RFC4           | ALG3, PSMD2, POL2R2H, MAGEF1, IMP-2, ETV5, RFC4, SIAT1, RPL39L, IFRG28, FLJ42393, ILIRAP, HRASLS, OPA1, HES1, TPI1A3A, FLJ1301, FLJ3227, MUC4, Hs.124620 | 20  |                                                                                          |
| 5p12–p13| 11.6      | 32161–43740           | NNT, FYB, NIPBL, IL7R, AGXT2, ZFR, GOLPH3 | GOLPH3, TARS, RAD1, SKP2, SLC1A3, NIPBL, NUP155, OSMR, FYB, PRKAA1, Hs.44986, PAIP1, NNT | 13  |                                                                                          |
| 5q31.3  | 1.3       | 140021–141371         | IFI16, APCS, NHLH1, DDR2   | IFI16, AIM2, PDN2, HSPA6, FCGR3B, DUSP12 | 2   |                                                                                          |
| 8q24.3  | 4.6       | 141599–146047         | EIF2C2, PTK2, SIAHBP1, DGAT1, COMM5 | EIF2C2, PTK2, LOC51337, LYE6, EXOSC4, GPAA1, CYC1, BOP1, DGAT1, GPR172A, CPSFI, SLC39A4, REQCly4, LRC14 | 14  |                                                                                          |
| 11q22.2 | 0.8       | 102137–102955         | MMP7, MMP1, MMP12, MMP13, YAP1 | BIRC2, BIRC3, PORIMIN, MMP7, MMP10, MMP1, MMP3, MMP12, MMP13 | 9   |                                                                                          |
| 14q32.33| 4.0       | 102265–106330         | RCOR1, TNFAIP2             | CDC42BPB, TNFAIP2, EIF5, KNS2, BAG5, XRC3, SIVA, AKT1, CDC4A, JAG2, BRF1, PACS1L, MTA1, CRIP1, IGHE, Hs.44901 | 16  |                                                                                          |
| 17q21–q22| 9.8      | 44024–53809           | PHB, SPOP, MYST2, ZNF161 | HOXB7, ITGA3, ABC3, CROP, NME2, MMD, AKAP1 | 7   |                                                                                          |
| 19p13.13| 5.4       | 10265–15668           | ICAMS5, KLFI1, PRKACA, CYP4F12 | FLJ11286, TYK2, FLJ12949, ILF3, SMARCA4, PRKCSH, TNPO2, RNASEH2A, DNASE2, RAD23A, Hs.293379, FLJ20244, BTBD14B, FLJ23447, ASF1B, DDX3, BRD4 | 17  |                                                                                          |
| 19q13.3 | 14.4      | 45390–59790           | ARHGEF1, GRLFI1, DHX34, LIG1, SNRP70, PRKCG, LIRLB1 | SNRPA, EGLN2, HNRNPUL1, FLJ10241, ERF, PAAHA1B3, PLAUR, KCNN4, PRKCSH, TNPO2, RNASEH2A, DNASE2, RAD23A, Hs.293379, FLJ20244, BTBD14B, FLJ23447, ASF1B, DDX3, BRD4 | 46  |                                                                                          |
| 20q13.1 | 2.6       | 42683–45242           | ADA, YWHAB, SLPI, EYA2     | ADA, YWHAB, TOMM34, STK4, KCN51, SDC4, PIGT, ZNF335, UBE2C, MUPP9, CD40, EYA2 | 12  |                                                                                          |
| Xp11.2–p11.3| 7.5   | 47188–54924          | ALAS2, SMCL1L1, PLP2, TS4, TFE3, SSX3, UXT, TIMPI | TIMP1, PGBP1, WDR45, SMCL1L1, GLN3L1 | 5   |                                                                                          |
| Xq12    | 6.8       | 67539–74307           | EFNB1, IL2RG, ITGB1BP2, OGT, SLC16A2 | KIF4A, SEPHS1 | 2   |                                                                                          |
| Xq22    | 2.2       | 100459–102691         | GLA, ALEX3, PLP1, TCEAL1, MAGEA10, NSOHL, ID3HG, PLXN3, MPP1 | HMG3, MAGEA3, MAGE6, CSAG2, TREGX2, DUSP9, SLC6A8, SLC6A10, ARHGAP4, IRAK1, G6PD, DKC1, FBA1 | 2   |                                                                                          |
| Xq28    | 5.6       | 148342–154107         | MAGEA10, NSOHL, ID3HG, PLXN3, MPP1 | TMSNB, PRKCI | 13  |                                                                                          |

Genes shown in bold are commonly found in both aCGH and Affymetrix arrays.
cervical epithelium samples and 31 CC cases (unpublished data). Only nine cases (one primary tumor and eight cell lines) were commonly studied by both aCGH and Affymetrix expression profiling studies due to limitations in obtaining good quality RNA. First, we identified all of the probe sets present in the U133A array within the genomic intervals of all 17 amplicons using Affymetrix-UCSC Genome Browser Query Tool (www.affymetrix.com/analysis/netaffx/query_uosc.aaffx?mapping). As expected, this analysis identified a large number of genes in each of the amplicons (data not shown). These probe sets were examined in the normalized expression profiles derived from normal cervix and invasive CC to identify over expressed genes. A gene was considered over expressed if the expression levels exceeded mean + 2SD of normal in >10% tumors. This analysis identified a number of overexpressed genes in each of the amplicons: 1p36, 13; 1p34, 12; 1q22-23, 6; 3q27-29, 21; 5p12-13, 13; 5q31.3, 2; 8q24.3, 14; 11q22.2, 9; 14q32.33, 16; 17q21-22, 7; 19q13.3, 46; 20q13.1, 12; Xp11.2-11.3, 5; Xq12, 2; Xq22, 2; and Xq28, 13 (Table 3; Supplementary Table 3A; supplementary material for this article can be found at http://www.interscience.wiley.com/jpages/1045-2257/suppmat).

RT-PCR Validation of Genes Overexpressed by Array Expression Analysis

To further validate the genes that showed evidence of over expression by Affymetrix expression profiles, we chose one to three genes functionally relevant in tumor development in each of the focal amplicons and analyzed them by semiquantitative RT-PCR. Thus, a total of 23 genes that mapped to 17 amplicons were examined by RT-PCR in 8 normal cervical epithelium, 9 CC cell lines, and 10 primary tumors. These analyses showed that all genes tested, except PTPA2 at 1p34 and HMGB3 at Xq28, showed similar levels of increased expression in tumors compared to the corresponding normal cervix (Supplementary Table 3B). Thus, in a subset of tumors, the overexpression of genes within amplicons was confirmed by both Affymetrix expression and RT-PCR analyses, which showed a similar fold increase (Fig. 2, and supplementary Table 3B). RT-PCR analysis of CDCA8, AIM2, ABCG3, RECQL4, SMARCA4, and CSAG2 genes showed no detectable levels of expression in normal cervix and the fold increase for these genes in tumor specimens was considered 100% (Fig. 2). Thus, using two different validation methods, we showed overexpression of a number of genes mapped within the amplicons identified by aCGH array.
DISCUSSION

Like many other epithelial cancers, invasive CC exhibits complex chromosomal changes (Atkin, 1997; Harris et al., 2003; Rao et al., 2004). The molecular consequence of this genomic complexity is poorly understood. Extensive genome-wide LOH studies have shown allelic deletions of chromosome arms 2q, 3p, 4, 5p, 6p, and 11q (Mitra et al., 1994a; Mullokandov et al., 1996; Narayan et al., 2003b). A number of studies have provided evidence for gain or amplification of chromosomal regions or genes (Mitra et al., 1994b; Heselmeyer et al., 1996; Narayan et al., 2003b; Rao et al., 2004). Gene dosage changes play a major role in tumor formation and progression (Albertson et al., 2003). The analyses presented here identified several such gene dosage alterations in CC. The copy number changes identified by aCGH showed a near concordance with the previously reported chromosomal CGH (cCGH) data on the same panel of tumors validating the present data (Harris et al., 2003; Narayan et al., 2003b; Rao et al., 2004). Comparing the recurrent increased copy number of one or more cDNA clones by aCGH in the present study with those of the cCGH data showed common amplifications that correspond to the regions 1p31, 2q32, 7q22, 8q21-q24, 10q23, 11q22, 16q23-q24, 20q11.2, 20q13.1, and Xp (Fig. 4A). Analysis of these data sets also showed similar concordance of chromosomal gains at 1p, 3q, 5p, 9q, 14q, 17q, and X (Fig. 4A). In addition, analysis of the data on deletions showed a similar correlation between the cCGH and aCGH with common regions of deletions at 2q33-37, 3p, 4p, 6q, 8p, 10p, 11q22-25, 13q, and 18q (Fig. 4B). However, the chromosomal amplifications at 7p11.2, 10q21, 11q13, and 12q15 regions revealed by cCGH could not be confirmed by aCGH (Fig. 4A). This discrepancy between cCGH and aCGH data may be due to the differences in coverage by each of the techniques. Although cCGH will identify all of the genomic changes at a resolution of megabase level, the array
Figure 4. Ideogram showing correlation of chromosome copy number alterations identified by cCGH and aCGH in 29 cervical cancer specimens. (A) Copy number gains and amplifications. cCGH data are shown in green vertical lines on left of the ideogram. Thin vertical lines indicate gains. Thick vertical lines indicate high-level amplifications. The aCGH data on increased copy number of clones is shown on the right of the ideogram in circles. Each small circle represents one tumor. Large circles represent 10 tumors. Green circles represent increased copies of multiple clones within the chromosomal sub-band, while the red circles represent a single clone. (B) Copy number deletions. Deletions identified by cCGH are shown in red vertical lines on left of the ideogram. The aCGH data on decreased copy number of clones is shown on the right of ideogram in circles. Each circle represents one tumor. Large circles represent 10 tumors. Green circles represent deletion of multiple clones within the chromosomal sub-band, while the red circles represent a single clone.
we used for aCGH has only an average coverage of 300 kb. Since the cDNA array used by us had low genomic representations in certain regions of the genome, we assume that genomic regions of some of the amplicons identified by cCGH are under represented in the cDNA array. Second, the criteria that we applied to identify amplifications at aCGH analysis in the present study will be eliminated amplifications present in less than three tumors.

The amplification of oncogenes is a known genetic mechanism underlying the development of a number of tumor types. Our previous studies suggested that gene amplification is a common event in CC (Mitra et al., 1994b; Harris et al., 2003; Narayan et al., 2003b). The present analysis identified increased copy number of cDNA clones on the entire X chromosome suggesting gain of this chromosome (Rao et al., 2004). The gain of 3q26-29 has been commonly reported in invasive CC and was shown to occur during the progression from low- to high-grade cervical intraepithelial neoplasia (CIN) (Heselmeyer et al., 1996; Heselmeyer-Haddad et al., 2003; Hidalgo et al., 2005; Narayan et al., 2003b; Rao et al., 2004; Fitzpatrick et al., 2006). Here we identified an amplicon spanning 11.8 Mb in 3q27.3-29. Gain of distal 3q is commonly seen in many other tumor types, such as head and neck squamous-cell carcinomas, and lung and ovarian cancer. Potential target oncogenes at 3q26-29 such as PIK3CA, TP73L, CCNL1, and EIF5A2 have been reported (Redon et al., 2002). Previous studies have implicated PIK3CA and TERC as target genes in CC (Ma et al., 2000; Sugita et al., 2000; Heselmeyer-Haddad et al., 2003). Mapped to 3q26.3, these genes are, however, 10 Mb proximal to the 3q amplicon identified in the present study. The 3q27.3-29 region contains several genes of relevance to cancer (Table 3). We showed here a 2.5 to 10.9-fold increased expression of three genes (RFC4, MUC4, and HRASLS) by both microarray expression profiles and RT-PCR (Fig. 2). RFC4 (replication factor 4) plays a critical role in DNA damage checkpoint pathways (Ellison and Stillman, 2003). Mucin 4 (MUC4) secreted by epithelial surfaces including cervix is implicated in renewal and differentiation of these cells. MUC4 has been reported to be overexpressed in pancreatic cancer and cervical dysplasias, and acts as ligand for ERBB2 and a target for the TGFB pathway (Lopez-Ferrer et al., 2001; Jonckheere et al., 2004; Singh et al., 2004). The mouse homologue of HRASLS encodes a ras-responsive gene, which modulates the HRAS-mediated signaling pathway.

Amplicons at 8q24.3 and 20q13.1 have been found in many tumor types, including CC (Zhang et al., 2002; Hodgson et al., 2003). The 8q24 region harbors a number of genes including MYC. In the present study, two genes, PTK2 and RECQL4, mapped to this amplicon were examined and shown to exhibit 3 to >9.7-fold increased expression in CC. The protein tyrosine kinase 2 (PTK2) gene, which encodes a cytoplasmic protein tyrosine kinase, is implicated in signaling pathways involved in cell motility, proliferation, and apoptosis (McLean et al., 2005). The RecQ protein-like 4 (RECQL4) encodes a DNA helicase involved in the maintenance of genomic integrity (Hickson, 2003). No report of RECQL4 amplification and/or over expression in human cancer is known thus far and it remains to be seen whether the over expression of RECQL4 has any functional role in CC tumorigenesis or represents a bystander affect. The 20q13.1 region, known to be amplified in diverse tumor types, harbors several genes implicated in tumorigenesis, such as ABI1, BTAK, and PTPN1. Our Affymetrix gene expression profiles identified increased expression of 12 genes, including UBE2C, within the 20q13.1 amplicon. Of these, the overexpression of UBE2C gene was further confirmed by RT-PCR analysis (Fig. 2). Ubiquitin-conjugating enzyme E2C (UBE2C) encodes a member of the E2 ubiquitin-conjugating enzyme family, which is essential for destruction of mitotic cyclins and for cell cycle progression. The ubiquitin-conjugate family genes are amplified and overexpressed in many human tumors, including CC (Wagner et al., 2004; Santin et al., 2005). The present study also identified a number of previously uncharacterized amplifications, which could include genes relevant to CC that may be revealed by positional approaches. For instance, the 820 kb 11q22.2 amplicon contains a number of matrix metalloproteinase (MMPs 1, 3, 12, and 13) genes, which are known to be overexpressed in many tumor types and that promote tumor growth, cell proliferation, and migration (Overall and Lopez-Otin, 2002).

This work represents the first high-resolution aCGH analysis of CC, which forms a basis for further studies on a subset of candidate genes in delineating the molecular mechanisms involved in its development. Identification of tumor-specific gene dosage profiles has important potential diagnostic and therapeutic implications. The distinct genetic losses and gains seen in the present study may be characteristic of CC as some of these changes (e.g., gain of 5p, 5q, and loss of 2q, 4p, 4q) are not commonly seen in
other epithelial cancers. Detailed characterization of the amplified and deleted regions may facilitate the identification and functional characterization of genes involved in CC development. In addition, examination of these changes in CIN lesions may provide new insights into the role of these genes in the progression of CC and thus in the diagnostic identification of lesions at high-risk for progression into invasive cancer.

REFERENCES

Albertson DG, Collins C, McCormick F, Gray JW. 2003. Chromosome alterations in solid tumors. Nat Genet 34:369–376.

Atkin NB. 1997. Cytogenetics of carcinoma of the cervix uteri: a review. Cancer Genet Cytogenet 95:33–39.

Blaven E, Brewer JL, Roydasgupta R, Fridlyand J, DeVries S, Koppe T, Pejaver S, Mehta K, Carroll P, Simko JP, Waldman FM. 2005. Bladder cancer stage and outcome by array-based comparative genomic hybridization, Clin Cancer Res 11:7012–7022.

Bourdon V, Naef F, Rao PH, Reuter V, Mok SC, Boul GJ, Koul M, Murphy VK. Kochergapts BS, Ghaghani RS. 2002. Genomic and expression analysis of the 12p11-12p12 region using EST arrays identifies two novel amplified and overexpressed genes. Cancer Res 62:6218–6223.

Califano JA. 2000. SPLASH: Structural pattern localization analysis by sequential histograms. Bioinformatics 16:341–357.

Cappellen D, De Oliveira C, Ricol D, de Medina S, Bourdin J, Sastre-Garau X, Choppin D, Thiery JP, Radavanyi F. 1999. Frequent activating mutations of FGFR3 in human bladder and cervical carcinomas. Nat Genet 23:18–20.

Chatterjee A, Pulido HA, Koul S, Beleno N, Perilla A, Posso H, Manusukhami M, Murty VV. 2001. Mapping the sites of putative tumor suppressor genes at 6q25 and 6q21.3 in cervical carcinoma: Occurrence of allelic deletions in precancerous lesions. Cancer Res 61:2119–2123.

Ellison V, Stillman B. 2003. Biochemical characterization of DNA damage checkpoint complexes: Clamp loader and clamp complexes with specificity for 53 recessed DNA. PLoS Biol 1:E33.

Fitzpatrick MA, Funk MC, Gius D, Huetten C, Zhang Z, Bidder M, Ma D, Powell MA, Rader JS. 2006. Identification of chromosomal alterations important in the development of cervical intraepithelial neoplasia and invasive carcinoma using alignment of DNA microarray data. Gynecol Oncol 103:458–462.

Harris CR, Lu XY, Narayan G, Singh B, Murty VV, Rao PH. 2003. Comprehensive molecular cytogenetic characterization of cervical cancer cell lines. Genes Chromosomes Cancer 36:233–241.

Heselmeyer K, Schrock E, du Manoir S, Blegen H, Shah K, Steinbeck A, Schneider A, Terry MB, Murty VV, Manusukhami M, Rao PH, Murty VV. 2003b. Genetic analysis identifies putative tumor suppressor sites at 2q35-3q11 and 7q36.3-3q1, involved in cervical cancer progression. Oncogene 22:3489–3499.

Overall CM, Lopez-Otin C. 2002. Strategies for MMP inhibition in cancer: Innovations for the post-trial era. Nat Rev Cancer 2:657–668.

Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Ross DT, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. 1999. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 23:41–46.

Rao PH, Atkin NB, Lu XY, Harris CR, Varella-Garcia M. 2000. Chromosomal amplifications, 3q gain and deletions of 2q33-q37 are the frequent genetic changes in cervical carcinoma. Oncogene 19:2657–2662.

Rao PH, Varella-Garcia M, Lassonde M, Fink EA, O'Brien TJ, Xu X, Cannon MJ, Shaughnessy J, Jr., Pecorelli S. 2005. Gene expression profiles of HPV16- and HPV18-infected early stage cervical cancers and normal cervical epithelium: Identification of novel candidate molecular markers for cervical cancer diagnosis and therapy. Virology 331:269–2691.

Rao PH, Varella-Garcia M, Lassonde M, Fink EA, O'Brien TJ, Xu X, Cannon MJ, Shaughnessy J, Jr., Pecorelli S. 2005. Gene expression profiles of HPV16- and HPV18-infected early stage cervical cancers and normal cervical epithelium: Identification of novel candidate molecular markers for cervical cancer diagnosis and therapy. Virology 331:269–2691.

Singh AP, Moniaux N, Chauhan SG, Meza J, Batra SK. 2004. Inhibition of MUC4 expression suppresses pancreatic tumor cell growth and metastasis. Cancer Res 64:622–630.

Sozzi G, Huescher K, Croce CM. 1998. FHIT in human cancer. Adv Cancer Res 74:141–166.

Sugita M, Tanaka N, Davidson S, Sekiya S, Savela-Garcia M, West J, Drabkin HA, Gemmill RM. 2000. Molecular definition of a small amplification domain within 3q26 in tumors of cervix, ovary, and lung. Cancer Genet Cytogenet 117:9–18.

Wagner KW, Simpson LM, El-Rifai W, Frizzera H, Butz N, Mestan J, Hofmann E, Deveraux QL, Hampton GM. 2004. Overexpression, genomic amplification and therapeutic potential of inhibiting the UbcH10 ubiquitin conjugase in human carcinomas of diverse anatomic origin. Oncogene 23:5621–5630.

Zhang A, Anser S, Betz R, Angstrom T, Stendahl U, Bergman F, Zetterberg A, Wallin KL. 2002. Genetic alterations in cervical carcinomas: Frequent low-level amplifications of oncogenes are associated with human papillomavirus infection. Int J Cancer 101:427–433.