[CANCER RESEARCH 64, 4744–4748, July 15, 2004]

High-Resolution Global Profiling of Genomic Alterations with Long Oligonucleotide Microarray

Cameron Brennan,[1] Yunyu Zhang,[1] Christopher Leo,[1] Bin Feng,[1] Craig Cauwels,[1] Andrew J. Aguirre,[1] Minjung Kim,[1] Alexei Protopopov,[1] and Lynda Chin[1,2]

[1]Department of Medical Oncology, Dana Farber Cancer Institute, and [2]Department of Dermatology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts

ABSTRACT

Cancer represents the phenotypic end point of multiple genetic lesions that endow cells with a full range of biological properties required for tumorigenesis. Among the hallmark features of the cancer genome are recurrent regional gains and losses that, upon detailed characterization, have provided highly productive discovery paths for new oncogenes and tumor suppressor genes. In this study, we describe the use of an oligonucleotide-based microarray platform and development of requisite assay conditions and bioinformatic mining tools that permits high-resolution genome-wide array-comparative genome hybridization profiling of human and mouse tumors. Using a commercially available 60-mer oligonucleotide microarray, we demonstrate that this platform provides sufficient sensitivity to detect single-copy difference in gene dosage of full complexity genomic DNA while offering high resolution. The commercial availability of the microarrays and associated reagents, along with the technical protocols and analytical tools described in this report, should provide investigators with the immediate capacity to perform DNA analysis of normal and diseased genomes in a global and detailed manner.

INTRODUCTION

Cancer cell genomes are typified by widespread chromosomal structural aberrations leading to regional amplifications and deletions of cancer-relevant loci. Comparative genomic hybridization (CGH) has emerged as a cornerstone technology for the identification and characterization of such chromosomal numerical aberrations (CNAs) on a genome-wide level (1–3). This approach has been enabled further by the availability of the human and mouse draft sequences and the adaptation of CGH to a microarray platform (4). Bacterial artificial chromosome (BAC)-based arrays have proven highly effective in defining the location of regional copy number changes (5). The current BAC arrays typically offer approximately 1 Mb of coverage (containing ~3000 BACs), translating into a resolution limit of 2 Mb (6–8). Using this platform, the additional delimitation of regional alterations is made possible by custom microarrays containing BAC contigs that tile across the locus of interest in an iterative locus-specific manner. Prior work has clearly documented the effectiveness of iterative BAC array-CGH profiles to identify candidate cancer genes residing in a focal amplicon (9, 10).

Several studies have documented the utility of cDNA-based microarrays for CGH profiling of human cancers (11, 12). We have gained considerable experience in the use of cDNA platforms in the analysis of more than 300 human tumors. These studies have demonstrated that commercially available cDNA array-CGH platforms are sufficiently robust to detect regional single-copy changes (13), providing that high background probes are eliminated by empirical and bioinformatic means.[3] Although highly effective, the full potential of cDNA microarrays has been constrained by currently available validated cDNA repositories.

Against the backdrop of this past experience, oligo-based microarrays hold the potential of enhanced design flexibility and eventual full-genome representation of probes capable of accurately reporting single-copy number changes. One long-standing concern has been whether the high complexity of the full genome would undermine the accurate reporting potential of short DNA substrates on a microarray. To date, two types of oligonucleotide microarrays have shown the potential to detect genomic alterations (14–16). One platform uses short oligonucleotides shown previously to be effective in the detection of single-nucleotide polymorphisms, whereas another is a photoprint array of custom-designed 70-mers. In both cases, a PCR-based genomic representation is required to reduce the complexity of the input genomic DNA by ~98% as a means to improve hybridization kinetics (14–17). Left unanswered is the extent to which PCR-based amplification biases impact the result.

In this study, we describe assay conditions and bioinformatic tools that enhance the utility of oligo-based microarray platforms in genome-wide DNA copy number analyses of human and mouse cancers. Using a commercially available 60-mer platform, we provide evidence of reliable detection of single copy number alterations in full-complexity genomic DNA. In the analyses of human and mouse cancer genomes, this high-resolution approach readily detects regional and focal CNAs that can be verified by quantitative PCR and are consistent with spectral karyotyping (SKY) data. We suggest that the methodology and analytical tools described in this report should provide investigators with an immediate opportunity to make use of available platforms to gain a detailed and global view of normal and diseased genomes in different species.

MATERIALS AND METHODS

Array Hybridization. Based on experience with more than 800 labeling reactions and 400 hybridizations on both the cDNA and oligo array platforms, we have established a set of quality assurance parameters on the pre- and post-labeling products that are predictive of successful hybridization. Specifically, digested genomic DNAs should have $A_{260/280}$ ratios of 1.7 to 2.0. Dye incorporations of post-labeling products are measured by NanoDrop reading, and the parameters that predict successful hybridization include minimum Cy3 incorporation of 0.5pmol/μl and Cy5 incorporation of 0.3 pmol/μl.

In brief, genomic DNA was fragmented by *Dpn*II restriction digest before labeling. After purification with the QIAquick PCR Purification kit (Qiagen), digested DNA was visualized using the Agilent 2100 BioAnalyzer. For each labeling reaction, 2 μg of digested DNA were used. Each sample is dye-swap labeled for hybridization against normal pooled-human male reference (Promega). DNA samples (2 μg) were denatured in the presence of 740 ng/μl Cy dye-labeled Random Primer (Trilink) and Reaction Buffer (Invitrogen BioPrime Labeling kit) at 98°C for 5 min and then cooled to 2°C for 5 min. The denatured sample was incubated with Klenow fragment, dNTP mix [2.0

Received 4/7/04; revised 5/12/04; accepted 5/21/04.

Grant support: NIH Grants RO1 CA99041, U01 CA84313, and P50 CA93683; the Goldhirsh Brain Tumor Foundation Award (L. Chin); NIH Training Grant T32 CA09382 (C. Brennan); the LeBow fund for Myeloma Cure (C. Brennan and A. Protopopov). L. Chin is a Charles E. Culpeper Scholar.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Note: Supplementary data for this article can be found at Cancer Research Online (http://cancerres.aacrjournals.org).

Requests for reprints: Lynda Chin, Department of Medical Oncology, Dana Farber Cancer Institute, Boston, MA 02115; E-mail: Lynda_Chin@dfci.harvard.edu.

[3] C. Brennan and L. Chin, unpublished observations.

mM dATP dGTP dTTP, 1.0 mM dCTP in 10 mM Tris (pH 8.0), 1 mM EDTA], and Cy3 or Cy5 dCTP nucleotides (1 mM; Perkin-Elmer) for 2 h at 37°C. Reactions were terminated using 0.5 M EDTA (pH 8.0). Cy3 and Cy5 reaction pairs (labeled pair, Cy5-sample:Cy3-reference; reversed labeled pair, Cy3-sample:Cy5-reference) were pooled, precipitated, and resuspended in 18.5 μl of SDS (0.514%). After a quality assurance check with NanoDrop determination of Cy3 and Cy5 incorporation, samples were mixed with blocking solution concentrated from 50 μl of human Cot-1 DNA (1 mg/ml; Life Technologies), 20 μl of yeast tRNA (5 mg/ml; Gibco), and 4 μl of (dA)-poly(dT) (5 mg/ml; Sigma). SSC and SDS were added to final concentrations of 3.9× and 0.25%, respectively, in a final volume of 60 μl. For hybridization, samples are denatured at 98°C for 2 min and then cooled at 37°C for 30 min under light protection with foil. Labeled reactions in a volume of 45 μl were pipetted onto Agilent Human 1A oligonucleotide arrays. Hybridization was carried out for 18–20 h at 65°C using the MAUI Hybridization System (BioMicro Systems, Salt Lake City, UT). After hybridization was complete, arrays were washed in 2× SSC and 0.03% SDS at 65°C for 5 min, followed by additional 5-min wash steps in 1× SSC and then 0.2× SSC, each at room temperature. Detailed labeling and hybridization protocols are available for download.[4]

Image Acquisition and Raw Data Processing. After drying, hybridized arrays were scanned on an Axon 4000B scanner, and spot finding and flagging were accomplished using GenePix Pro software, version 3.0.[5] Alternatively, images were scanned at 10 μm resolution using Agilent scanner equipped with automatic spot finding and flagging ability in addition to reporting of Cy3 and Cy5 signal and background for each spot. Data extraction was performed using Agilent Feature Extraction Software, version 7.1. Custom tools including probe-to-chromosome mapping, ratio calculation, normalization, and visualization were used to compile the CGH profiles from these array data points. These tools are available for download.[4] Segmented profiles are also generated as described before (13).

Custom Analytical Tools. A package of analysis tools has been designed specifically for oligo array-CGH.[4] These consist of an annotation file for human and mouse arrays (which can be adapted to catalogue or custom arrays) and an analysis program. The annotation is generated for all oligo probes, including assembly alignment with BLAST-like alignment tool (BLAT) (18), GC-content, Tm (19), and minimum free energy for secondary structure with MFOLD (20).

To determine whether a specific sample has been successfully profiled, the program measures the percentage of acceptable (good) spots and distribution (mean, median, and SD) of the average deviation of the Log_2 ratio of paired hybridizations (dye swap). To estimate noise level of the hybridization, a vector is created from the difference between Log_2 ratios of consecutive probes along the physical distance of the chromosome. Noise is estimated by taking the SD of this vector scaled by 1 over the square root of 2 (so-called "derivative SD"). The rationale for this method instead of SD of raw Log_2 ratio is based on the fact that derivative SD does not penalize profiles with chromosome changes spanning a large region. Another important quality assurance parameter is the consistency of the paired hybridizations, which is measured by linear correlation based on raw Log_2 ratios between the pairs. Unlike expression profiles, most of the probes are expected to report no significant changes in a CGH profile, because regions containing CNAs typically comprise only a small proportion of the entire genome. Thus, probes reporting no copy number alterations (defined as absolute Log_2 ratio is less than 0.2 from median filtered profiles with window size 9) are excluded from calculation of correlation. In this way, we measure the concordance of changes that are detected in a pair of dye-swap hybridizations or on multiple-day replicate hybridizations based on correlation of changed regions only.

Real-Time Quantitative PCR. Relative gene copy numbers were determined by real-time PCR (qPCR) using SYBR Green I detection chemistry and the ABI Prism 7700 Sequence Detection System (Applied Biosystems). Amplification reactions contained 1× QuantiTect SYBR Green PCR buffer (Qiagen), 5 ng of genomic DNA template, 0.25 unit of uracil-*N*-glycosylase (Applied Biosystems), and 300 nM of each primer in a final volume of 25 μl. The comparative threshold cycle (C_T) method was used to quantify target gene copy number in the tumor DNA sample relative to that of an endogenous control gene (Assay-Z) and a reference DNA sample (normal pooled-human male reference; Promega). For primer sequences used in this study, see Supplemental Table 1.

Calculation of Gene Dosage for Array-CGH and Real-Time PCR Data. Real-time PCR determines gene dosage at a region of interest relative to a control assay designed to interrogate a specific unchanged region in the genome (so-called Assay-Z). This reference region is selected based on array-CGH profiles as a region without copy number alterations. We further verify the actual copy number of the Assay-Z location for each sample by SKY (Supplemental Table 2). The Assay-Z used in this study for human cell lines is located on ATP2B4 gene (1q32, locus 200.77~200.89 Mb on UCSC hg 16 map). The gene dosage for real-time PCR is thus calculated by multiplying relative gene copy number by Assay-Z base copy. It is typical for any array-CGH platform to quantitatively underestimate copy number change in Log_2 ratio (12). To compare estimates of gene dosage between CGH and PCR directly, we built a linear regression model based on Log_2 ratio of real-time PCR to adjust segmented Log_2 ratio (Supplemental Fig. 1). Given an R^2 of 0.9527, the model showed that these two types of measurements are highly correlated and demonstrate the same trend in gene dosage change. We used this model to calculate an estimated array-CGH Log_2 ratio, and the array-CGH gene dosage is finally obtained as (Assay-Z base copy)$*2^{(\text{estimated log2 ratio})}$.

SKY. SKY was performed according to manufacturer's protocol (Applied Spectral Imaging, Carlsbad, CA).

RESULTS

Probe Selection for Genomic Hybridization. In this study, we examined the utility of long oligonucleotide (60-mer) microarrays (Agilent Technologies) for array-CGH. These ink-jet printed arrays consisting of approximately 22,500 elements were designed for expression profiling, raising questions as to whether a significant proportion of the probes would be suitable for high-complexity genomic hybridizations. In an initial pass to cull potentially problematic probes, all 22,500 sequences were subjected to BLAT (18) alignment with the latest draft of human or mouse genome sequence (University of California at Santa Cruz genome hg16 and mm4, respectively[6]) in an effort to eliminate (*a*) un-mappable probes, which were defined arbitrarily as alignment length of best hit less than 55 bp, and (*b*) un-informative probes, which were identified by having a second best hit of 95% or higher on the alignment length of the best hit. As shown in Table 1, these criteria selected 13,502 and 15,729 informative gene-specific probes for the human and mouse arrays, respectively. All subsequent analyses make sole use of these selected informative probes, which together translate into a median resolution of 69.2 kb for human and 48.6 kb for mouse genomes (Table 1).

Detection of Single-Copy Difference in Normal Genomes. To first determine whether this platform can detect single-copy difference, we performed array-CGH profiles on normal female and male genomic DNAs (pooled; Promega). As shown in Fig. 1, the CGH profile clearly detected a single-copy increase in X chromosome in the female compared with the male reference. Similarly, loss of Y is readily evident, even in the face of sparse probe density along the Y chromosome.

Consistent Performance on Pair-Wise Dye-Swap and Replicate Hybridizations. Next, we sought to determine the performance consistency of the oligo platform by measuring correlation of detected changes in multiple hybridizations of the same samples (*i.e.*, correlation of the variance). To this end, we conducted multiple array-CGH profiling analyses of human cancer cell lines and mouse tumor cell lines (Supplemental Table 3). For each sample, a pair-wise dye-swap hybridization was carried out on the same day. In the case of human pancreatic cell line ASPC1 and mouse melanoma cell lines MK1 and MK2, we also conducted replicate dye-swap hybridizations on different days. We determined that the correlation of changes detected in a

[4] Internet address: http://genomic.dfci.harvard.edu.
[5] Internet address: http://www.axon.com/gn_GenePixSoftware.html.
[6] Internet address: http://genome.ucsc.edu/.

Table 1 *Probe selection of human and mouse oligo arrays*

The oligonucleotide sequences (60-mer) for all probes printed on both human and mouse catalogue microarrays (Agilent respectively) are aligned against their corresponding genome databases (UCSC hg16 and mm4, respectively) by BLAT. Probes with alignment length greater or equal to 55 bp are defined as mappable. "Informative probe" is further defined as mappable probes with no more than one alignment greater than 95% of the best alignment length.

Array type	Total probes	Mappable probes	Informative probes	Median resolution
Human	17,168	14,950	13,502	69.2 kb
Mouse	20,317	17,310	15,729	48.6 kb

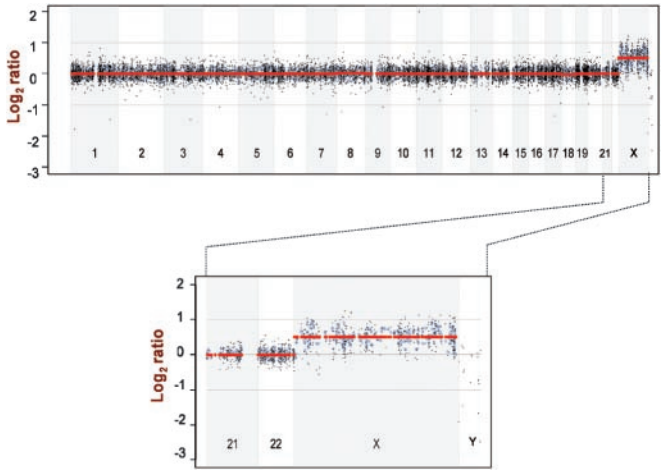

Fig. 1. Genomic profiles of female *versus* male normal genomic DNAs. Array-CGH profiles of female DNA against male DNA as reference (both pooled samples) with *X*-axis coordinates representing oligo probes ordered by genomic map positions. *Top panel*, whole-genome profile. *Bottom panel*, expanded view of the bracketed region. Segmented data are displayed in *red*, median filtered data (three nearest neighbors) in *blue*, and raw data in *black*.

pair-wise dye-swap hybridization for the same sample on the same day ranges from 0.65 to 0.9 for human and 0.67 to 0.78 for mouse samples (for example, see Supplemental Fig. 2*A*). Moreover, when the same samples were profiled repeatedly on different days, the correlation of detected changes was 0.84 to 0.92 (Supplemental Fig. 2*B*). For example, in sample ASPC1, a zoom-in view of chromosome 7 demonstrates that distinct hybridizations detected the same amplicon and amplicon structural features (Supplemental Fig. 2*B*). Together, these performance data strongly support the view that this oligo array-CGH platform and assay conditions yields consistent data sets across independent experiments.

Verification of CNAs Detected by Oligo Array-CGH. To further validate our CGH data sets and the robustness of our assay conditions and analytical tools, we used two independent and highly reliable methods to ascertain gene dosage alterations in complex cancer genomes: SKY analysis and qPCR for human cell lines (Supplemental table 3). For SKY studies, we used both published SKY data (*e.g.*, ASPC1; Ref. 21) as well as newly generated SKY profiles (*e.g.*, DanG, HUP-T4, HPAC, PANC1, and TU8902). SKY ideograms were created for visualization using the National Center for Biotechnology Information Automatic Karyotype to SKYGRAM Converter tool[7] and compared with pseudo-karyotype representations of the segmented array-CGH profiles obtained in this study. Fig. 2 illustrates two representative comparisons of the array-CGH and SKY data sets. Noteworthy in the SKY analysis of HPAC is the detection of a one-copy gain of part of chromosome 10q (Fig. 2*A*, *green* chromosome), which is translocated to chromosome 12, and the presence of four copies of chromosome 12p (Fig. 2*A*, *magenta* chromosome). These SKY features mirror those obtained in the segmented array-CGH profile of HPAC cells. Another example is cell line PANC1, which readily shows three copies of 8q (Fig. 2*B*, *orange* chromosome) by SKY and array-CGH assays, underscoring that large regional alterations revealed by SKY are readily detected by array-CGH method used here.

Next, real-time qPCR was performed to assess the ability of array-CGH to accurately and reproducibly report focal CNAs and their complex structural features, particularly those that may not be detected by the low-resolving power of SKY. Ten randomly selected CNAs detected in a total of five different human tumor cell lines were subjected to qPCR verification. As shown in Table 2, the qPCR results were completely concordant with the array-CGH data. For example, real-time qPCR quantitation confirmed the presence of a Chr7 amplicon detected by array-CGH profile of ASPC1 tumor cell line (Fig. 3*A*). Of particular significance, real-time qPCR also confirmed the capacity of oligo array-CGH to detect homozygous deletion of CDKN2A in the 9p region of ASPC1 (Fig. 3*B*). Moreover, the fine structural complexity of this CNA as revealed by the array-CGH profile is mirrored by the real-time qPCR data points. Specifically, three consecutive features within a 20 Mb genomic distance are observed: a two-copy deletion of CDKN2A region at 21 Mb from 9p tip; followed by a focal gain spanning only 1 Mb verified by NGX6 at 35.8 Mb; and single-copy loss from 36.7 Mb. Together, these data provide strong evidence that the oligo array-CGH approach described here provides accurate and reliable representation of complex regional copy number changes in cancer genomes.

High-Density Coverage of the Oligo Platform Provides Ready Detection of Focal CNAs. It stands to reason that the increased resolution of the oligo platform compared with antecedent microarray platforms should provide a greater likelihood for the identification of highly focal copy number changes. To ascertain this possibility, we performed a comparative array-CGH study of the same mouse tumor cell line sample using the cDNA and current oligo-based microarrays. The median resolution of the mouse cDNA platform is ~100 kb, whereas the oligo platform is ~50 kb (Table 1). As shown in Fig. 4, the oligo array-CGH profile of MK1 (Fig. 4*A*, *blue*) revealed 4 distinct focal amplicons on chromosome 10 measuring 1.4, 0.13, 0.8, and 0.7 Mb, respectively. All four amplifications were confirmed by real-time qPCR (Fig. 4*C*). In contrast, the cDNA-based profile (Fig. 4*A*, *or-*

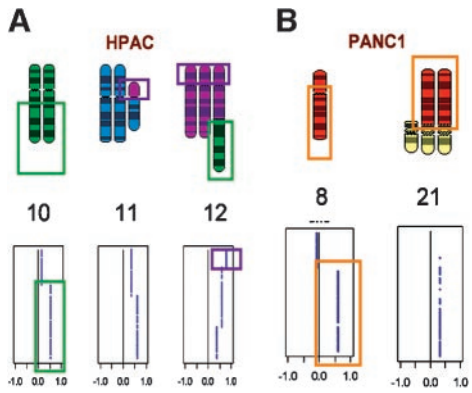

Fig. 2. SKY verification of CNAs detected by oligo array-CGH. *Top panel* shows the SKY ideogram corresponding to specific chromosomes as indicated by numbers below. Each chromosome is represented by a different pseudo-color as determined by the SKY software. *Bottom panel* is the array-CGH pseudo-karyotype based on segmented data in Log_2 ratio of the corresponding chromosomes. The Log_2 ratio of 0 represents two copies. Segmented profiles to the *right* of 0 indicate gain or amplification and to the *left* indicate loss or deletion. *A*, copy numbers of chromosome 10 (*green*), 11 (*blue*), and 12 (*magenta*) of sample HPAC are shown. The *green boxes* highlight the one-copy gain of 10q, where the third copy is translocated to the q arm of chromosome 12. The *magenta boxes* highlight the amplification of 12p, where the fourth copy is translocated to chromosome 11p. *B*, *orange boxes* outline the three copies of chromosome 8q (*orange*) in sample PANC1, where two of the copies are translocated to chromosome 21 (*yellow*).

[7] Internet address: http://www.ncbi.nlm.nih.gov/sky/.

Table 2 *qPCR verification of 10 CNAs detected by array-CGH from five human cancer cell lines*

Array-CGH values are adjusted based on a regression model (Supplemental Figure 1). Estimated gene dosage is calculated as described in text. For more details, see Supplemental Table 2.

			Array-CGH			Real-time PCR	
			Log_2 ratio				
Sample	Chr no.[a]	Locus (Mb)	Raw	Adjusted	Gene dosage	Log_2 ratio	Gene dosage
DanG	12	18.13–21.58	2.459	2.459	48.1	3.597	36.3
PANC1	19	43.75–45.41	2.434	2.655	40.8	3.903	29.9
TU8902	12	0.18–5.03	1.580	1.683	18.5	2.719	19.8
ASPC1	7	93.7–98.6	1.553	1.553	15.8	3.247	28.5
TU8902	12	13.62–32.85	1.006	1.109	9.1	2.120	13.1
HPAC	21	31.96–44.08	−0.449	−0.449	1.3	−0.702	1.8
DanG	8	11.86–19.89	−0.595	−0.595	1.1	−1.899	0.8
PANC1	8	0.35–19.89	−0.652	−0.431	0.9	−1.953	0.5
ASPC1	9	0.96–27.94	−1.187	−1.187	0.5	−2.229	0.6
PANC1	9	21.16–23.68	−1.438	−1.217	0.5	−13.883	0.0

[a]Chr no., chromosome number.

ange) of the same sample captured only two of the four amplicons (amplicons C and D). In the case of amplicon B, the region in question is interrogated by a single probe on the oligo microarray, located physically at 102.99 Mb on chromosome 10 and flanked by a left neighbor 30 kb away (102.96 Mb) and a right neighbor 100 kb away (103.09 Mb). Within this 130-kb region, there was no probe represented on the cDNA microarray, providing a basis for the lack of detection of this high-amplitude event on that platform (Fig. 4*B*). An observation such as this should provide impetus for the design of higher density oligo microarrays for full-genome interrogation.

DISCUSSION

This study reports the successful implementation of technical protocols and analytical tools for an array-CGH platform that uses commercially available microarrays composed of long oligonucleotide probes. We show that this platform, supported by the described analytical tools, is sufficiently sensitive to detect single-copy alterations. In addition, genome-wide oligonucleotide array-CGH profiles of human and mouse cancer genomes reveal regional and focal CNAs with high resolution that can be verified by quantitative PCR and by SKY. Furthermore, we provide evidence that increased resolution of the current oligo array-CGH platform enabled the identification of focal alterations that were missed by a lower resolution cDNA-based platform.

Although previous proof-of-principal studies established the utility of oligonucleotide microarrays in copy number analyses (14–16), an important finding of this study is that a PCR enrichment step is not necessary to reduce the genome complexity before hybridization to the microarray. Direct genome labeling and hybridization should

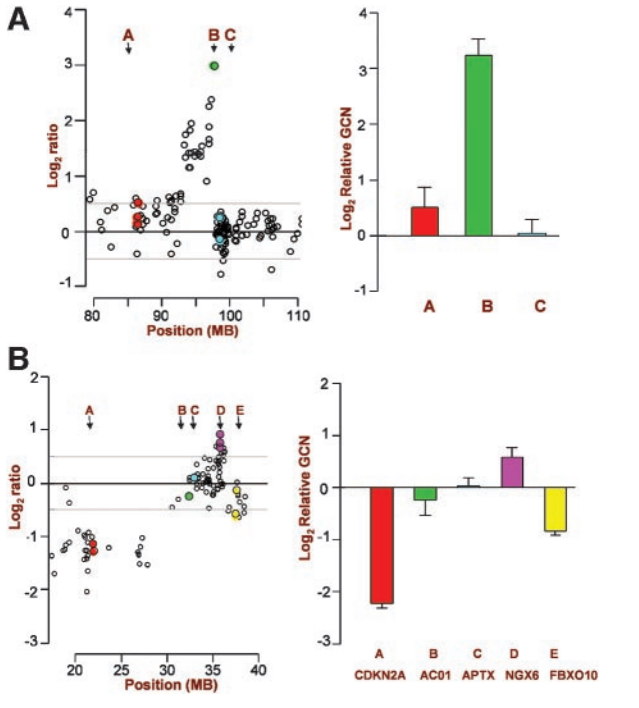

Fig. 3. Real-time qPCR verification of CNAs detected by oligo array-CGH. Comparison of gene dosages detected by array-CGH and qPCR in ASPC1 in two genomic locations. *A*, chromosome 7 from 80–110 Mb. *B*, chromosome 9 from 20–40 Mb. Raw Log_2 ratio values for each probe on the microarray are plotted in the *left panel*. *Colored dots* correspond to probes interrogating genomic regions assayed by qPCR on the *right panel*. The height of histogram bar represents the relative gene copy numbers in Log_2 scale as measured by qPCR (for copy number calculation, see Supplemental Table 2). The *bar* and *dots* of same color are measuring the gene dosage in the same genomic location.

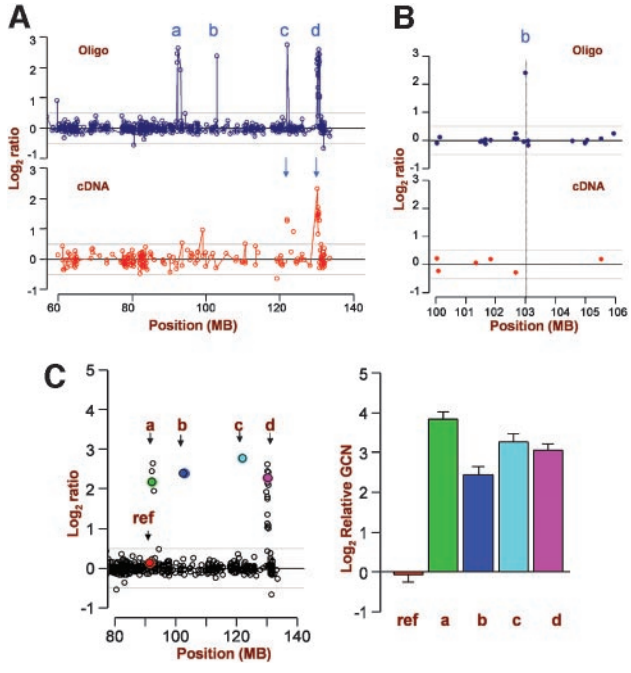

Fig. 4. Comparison of array-CGH profiles on oligo *versus* cDNA microarray platforms. *A*, chromosome 10 array-CGH profile of mouse sample MK1 generated on both oligo (*blue tracing*) and cDNA (*orange tracing*) array platforms. The *X* axis represents physical position of chromosome 10 in Mb. Four amplification events (*a–d*) are detected in the oligo profile, whereas only two (*c* and *d; arrows*) are detected in the cDNA profile. *B*, enlarged view of the profiles in region containing amplicon b. Probes on each platform are as indicated, showing a lack of probe on cDNA microarray corresponding to region of focal amplicon b (*dashed line*). *C*, qPCR validation of all four amplification events. As in Fig. 3, *bar* and *dots* in same color are measuring the gene dosage in the same genomic location. "Ref" (*red*) marks a flanking control region found unchanged on array-CGH.

serve to eliminate prevailing concerns of PCR amplification bias that may not capture highly focal CNAs or provide fine structural complexity within a given CNA to facilitate mechanistic studies of amplification or deletion. There exists a need to determine the extent to which PCR-based procedure may influence copy number profiles, because amplification may be needed in cases in which clinical materials are limiting. The approach described in this study now makes this possible.

Although current available BAC-based array-CGH platforms typically provide genome-wide coverage at ~1-Mb resolution, a recent report has described the construction of tiling BAC microarrays possessing 32,433 BACs spotted in triplicate on two separate glass slides (22). These BAC microarrays offer an approximate resolution of 80 kb across the human genome, comparable with the resolution provided by the commercially available expression oligo arrays used in this study. Furthermore, it is reassuring that the array-CGH performance characteristics of the reported tiling BAC arrays and the oligo arrays are comparable with respect to magnitude of signal in changed regions and noise in unchanged regions (22; this study). On the practical level, it is worth noting that the resolution of oligo array-CGH is not limited by the currently available oligo-arrays designed for expression profiling, rather by availability of genome sequences of any species. Furthermore, such genomic arrays can provide not only complete gene-specific representation with oligos targeted to gene-coding regions, but also representation of unique intragenic sequences, with oligos targeted to noncoding DNAs that may represent important and critical regulatory regions including *cis*-regulatory elements and microRNAs or mammalian interspersed repeats (MIR) sequences.

In summary, the experimental merits of oligonucleotide-based microarrays include (*a*) flexibility to design genome-wide or locus-specific custom microarrays with probes targeting coding and noncoding regulatory regions; (*b*) capacity to provide full genome coverage of known and predicted genes present on the latest draft of genome sequences for virtually any species; and (*c*) ease of quality control with respect to probe annotation. Availability of a commercial source obviates the need for in-house microarray printing infrastructure and provides increased access across the research community. Finally, the superior performance of this higher-resolution platform is clearly documented over antecedent cDNA platforms. These data raise the possibility that cancer genomes may harbor many focal CNAs that have eluded detection and justify continued efforts to build high-density genomic arrays that permit detailed interrogation of the entire genome.

ACKNOWLEDGMENTS

We thank Drs. Kornelia Polyak, Matthew Meyerson, and Ron DePinho for critical comments on the manuscript and Tali Muller for superb technical support. Array-CGH profiles were performed at the Arthur and Rochelle Belfer Cancer Genomic Center at Dana-Farber Cancer Institute.

REFERENCES

1. du Manoir S, Speicher MR, Joos S, et al. Detection of complete and partial chromosome gains and losses by comparative genomic in situ hybridization. Hum Genet 1993;90:590–610.
2. Kallioniemi A, Kallioniemi OP, Sudar D, et al. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 1992;258:818–21.
3. Joos S, Scherthan H, Speicher MR, Schlegel J, Cremer T, Lichter P. Detection of amplified DNA sequences by reverse chromosome painting using genomic tumor DNA as probe. Hum Genet 1993;90:584–9.
4. Solinas-Toldo S, Lampel S, Stilgenbauer S, et al. Matrix-based comparative genomic hybridization: biochips to screen for genomic imbalances. Genes Chromosomes Cancer 1997;20:399–407.
5. Pinkel D, Segraves R, Sudar D, et al. High resolution analysis of DNA copy number variation using comparative genomic hybridization to microarrays. Nat Genet 1998; 20:207–11.
6. Snijders AM, Nowak N, Segraves R, et al. Assembly of microarrays for genome-wide measurement of DNA copy number. Nat Genet 2001;29:263–4.
7. Fiegler H, Carr P, Douglas EJ, et al. DNA microarrays for comparative genomic hybridization based on DOP-PCR amplification of BAC and PAC clones. Genes Chromosomes Cancer 2003;36:361–74.
8. Chung YJ, Jonkers J, Kitson H, et al. A whole-genome mouse BAC microarray with 1-Mb resolution for analysis of DNA copy number changes by array comparative genomic hybridization. Genome Res 2004;14:188–96.
9. Albertson DG, Ylstra B, Segraves R, et al. Quantitative mapping of amplicon structure by array CGH identifies CYP24 as a candidate oncogene. Nat Genet 2000;25:144–6.
10. Zafarana G, Grygalewicz B, Gillis AJ, et al. 12p-amplicon structure analysis in testicular germ cell tumors of adolescents and adults by array CGH. Oncogene 2003;22:7695–701.
11. Heiskanen MA, Bittner ML, Chen Y, et al. Detection of gene amplification by genomic hybridization to cDNA microarrays. Cancer Res 2000;60:799–802.
12. Pollack JR, Perou CM, Alizadeh AA, et al. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23:41–6.
13. Aguirre AJ, Brennan C, Bailey G, et al. High-resolution characterization of the pancreatic adenocarcinoma genome. Proc Natl Acad Sci USA. In press, 2004.
14. Lucito R, Healy J, Alexander J, et al. Representational oligonucleotide microarray analysis: a high-resolution method to detect genome copy number variation. Genome Res 2003;13:2291–305.
15. Bignell GR, Huang J, Greshock J, et al. High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res 2004;14:287–95.
16. Zhao X, Li C, Paez JG, et al. An integrated view f copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 2004;64:3060–71.
17. Lucito R, Nakimura M, West JA, et al. Genetic analysis using genomic representations. Proc Natl Acad Sci USA 1998;95:4487–92.
18. Kent WJ. BLAT: the BLAST-like alignment tool. Genome Res 2002;12:656–64.
19. Wahl GM, Allen V, Delbruck S, et al. Analysis of CAD gene amplification using a combined approach of molecular genetics and cytogenetics. Adv Exp Med Biol 1984;172:319–45.
20. Zuker M, Mathews DH, Turner DH. Algorithms and thermodynamics for RNA secondary structure prediction. In: Barciszewski J, Clark BFC, editors. A practical guide in RNA biochemistry and biotechnology. Dordrecht, The Netherlands: Kluwer Academic Publishers, 1999.
21. Ghadimi BM, Schrock E, Walker RL, et al. Specific chromosomal aberrations and amplification of the AIB1 nuclear receptor coactivator gene in pancreatic carcinomas. Am J Pathol 1999;154:525–36.
22. Ishkanian AS, Malloff CA, Watson SK, et al. A tiling resolution DNA microarray with complete coverage of the human genome. Nat Genet 2004;36:299–303.