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 Chin1,2

1Department of Medical Oncology, Dana Farber Cancer Institute, and 2Department 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 demonstrated 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. 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-mer. 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 A260/280 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 0.5 pmol/μl and Cy5 incorporation of 0.3 pmol/μl.

In brief, genomic DNA was fragmented by DpnII 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 ROI 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.

C. Brennan and L. Chin, unpublished observations.
nm 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 mM EDTA (pH 8.0). Cy3 and Cy5 reaction pairs (labeled 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 3.9X 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 2X SSC and 0.03% SDS at 65°C for 5 min, followed by additional 5-min wash steps in 1X SSC and then 0.2X SSC, each at room temperature. Detailed labeling and hybridization protocols are available for download.

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. 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. Segmented profiles are also generated as described before.

Custom Analytical Tools. A package of analysis tools has been designed specifically for oligo array-CGH. 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, which together translate into a median resolution of 69.2 kb.

RESULTS

Detection of Single-Copy Difference in Normal Genomes. To first determine whether this platform can detect single-copy changes in human and mouse 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
Table 1  Probe selection of human and mouse oligo arrays

<table>
<thead>
<tr>
<th>Array type</th>
<th>Total probes</th>
<th>Mappable probes</th>
<th>Informative probes</th>
<th>Median resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human</td>
<td>17,168</td>
<td>14,950</td>
<td>13,502</td>
<td>69.2 kb</td>
</tr>
<tr>
<td>Mouse</td>
<td>20,317</td>
<td>17,310</td>
<td>15,729</td>
<td>48.6 kb</td>
</tr>
</tbody>
</table>

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. 2A). Moreover, when the same samples were profiled repeatedly on different days, the correlation of detected changes was 0.84 to 0.92 (Supplemental Fig. 2B). 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. 2B). 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 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. 2A, green chromosome), which is translocated to chromosome 12, and the presence of four copies of chromosome 12p (Fig. 2A, 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. 2B, 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. 3A). 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. 3B). 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. 4A, 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. 4C). In contrast, the cDNA-based profile (Fig. 4A, or-

---

Log 2 ratio values for each probe on the microarray are plotted in Table 2.

...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 detection of this high-amplitude event on that platform (Fig. 4) presented on the cDNA microarray, providing a basis for the lack of detection of this high-amplitude event on that platform (Fig. 4). 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...
serve to eliminate prevailing concerns of PCR amplification bias that may not capture highly focal CNAs or provide fine structural complexity within a given DNA 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 non-coding 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

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

Cameron Brennan, Yunyu Zhang, Christopher Leo, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/64/14/4744

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2004/07/29/64.14.4744.DC1

Cited articles
This article cites 20 articles, 8 of which you can access for free at:
http://cancerres.aacrjournals.org/content/64/14/4744.full.html#ref-list-1

Citing articles
This article has been cited by 32 HighWire-hosted articles. Access the articles at:
/content/64/14/4744.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.