Genome-wide Allelic State Analysis on Flow-Sorted Tumor Fractions Provides an Accurate Measure of Chromosomal Aberrations

Willem E. Corver, Anneke Middeldorp, Natalja T. ter Haar, Ekaterina S. Jordanova, Marjo van Puijlenbroek, Ronald van Eijk, Cees J. Cornelisse, Gert Jan Fleuren, Hans Morreau, Jan Oosting, and Tom van Wezel

Department of Pathology, Leiden University Medical Center, Leiden, the Netherlands

Abstract
Chromosomal aberrations are a common characteristic of cancer and are associated with copy number abnormalities and loss of heterozygosity (LOH). Tumor heterogeneity, low tumor cell percentage, and lack of knowledge of the DNA content impair the identification of these alterations especially in aneuploid tumors. To accurately detect allelic changes in carcinomas, we combined flow-sorting and single nucleotide polymorphism arrays. Cells derived from archival cervical and colon cancers were flow-sorted based on differential vimentin and keratin expression and DNA content and analyzed on single nucleotide polymorphism arrays. A new algorithm, the lesser allele intensity ratio, was used to generate a molecular measure of chromosomal aberrations for each case. Flow-sorting significantly improved the detection of copy number abnormalities; 31.8% showed an increase in amplitude and 23.2% were missed in the unsorted fraction, whereas 15.9% were detected but interpreted differently. Integration of the DNA index in the analysis enabled the identification of the allelic state of chromosomal aberrations, such as LOH ([A]), copy-neutral LOH ([AA]), balanced amplifications ([AABB]), and allelic imbalances ([AAB] or [AAAB], etc.). Chromosomal segments were sharply defined. Fluorescence in situ hybridization copy numbers, as well as the high similarity between the DNA index and the allelic state index, which is the average of the allelic states across the genome, validated the method. This new approach provides an individual molecular measure of chromosomal aberrations and will likely have repercussions for preoperative molecular staging, classification, and prognostic profiling of tumors, particularly for heterogeneous aneuploid tumors, and allows the study of the underlying molecular genetic mechanisms and clonal evolution of tumor subpopulations. [Cancer Res 2008;68(24):10333–40]

Introduction
Chromosomal aberrations are common characteristics of human cancer and arise early during tumorigenesis. These aberrations are believed to be one of the driving forces behind tumor progression (1, 2). The process results in aneuploid cancer cells, which can be observed by genomic copy number abnormalities (CNA), allelic imbalances, loss of heterozygosity (LOH), and abnormal DNA content. Recently, a meta-analysis showed that these aberrations are associated with a worse prognosis in colorectal cancer (3). CNAs can be detected by array CGH (aCGH) or single nucleotide polymorphism (SNP) arrays (4–6). SNP arrays are preferred over aCGH because they additionally identify copy-neutral LOH (cnLOH; ref. 7), balanced CNAs (equal multiplication of both alleles), and allelic imbalances in which one allele is duplicated or amplified. For example, Kloth and colleagues showed that 75% of the LOH events found by SNP array were unnoticed using aCGH on the same cohort of cervical cancer cell lines (8). Furthermore, we and others have shown that reliable genotypes and profiles of CNAs and LOH can be generated from the fragmented DNA derived from formalin-fixed, paraffin-embedded (FFPE) cancer tissue using SNP arrays (5, 9–11).

The generated data is usually interpreted relative to the average DNA content of a tumor. For that reason, only relative copy numbers can be detected. Because tumors often show extensive genomic CNAs with almost doubled, near-tetraploid genomes (12), the extent of CNAs could be misinterpreted. For example, without knowledge of the DNA index, the CNA profiles from near-tetraploid tumors can be difficult to distinguish from those of near-diploid tumors. Consequently, only five or more copies will be interpreted as a gain, four copies will be misinterpreted as neutral and two or three copies as a loss. To improve copy number analysis, use of the DNA index has been suggested (13). Yamamoto and colleagues (14) analyzed cell lines and acute leukemia with limited CNAs and a mostly hyperdiploid DNA content (15). A recent study from Lyng and colleagues clearly showed the necessity for measuring the DNA index for calculation of absolute copy numbers from aCGH data (16). Also, varying proportions of normal cells (inflammatory and stromal cells) impair the detection of genomic and genetic alterations in tumor samples. For example, for the detection of LOH, samples should contain at least 50% tumor cells (17–19). Furthermore, solid tumors often contain subpopulations of tumor cells that harbor different chromosomal aberrations and may differ in their DNA index. Microdissection, either manually or by laser capture, is only a partial solution to the sampling problem in solid tumors because it is not possible to select tumor cells based on ploidy. Also, microdissection has a low cell yield and is prohibitively time-consuming when the neoplastic cells are highly intermingled with normal cells.

Fluorescence-activated cell sorting has proven to be an excellent tool for the purification of cell subpopulations from human tumors (20–23). Furthermore, since the development of robust dissociation methods in the early 1980s, flow cytometry has been widely used to determine the DNA content of solid tumors. Although these studies initially required fresh or frozen tumor samples, we developed a technique that markedly improved the yield and resolution of flow
cytometric DNA content measurements of FFPE samples (24). This technique allows simultaneous flow-sorting of tumor and stromal cells based on differential expression of vimentin and keratin, as well as DNA content, and was successfully applied to study cervical, gastric, and colon cancers (25–27).

In the present study, we show that the combined use of multiparameter DNA flow-sorting and SNP array analysis significantly improves the detection of CNAs in archival FFPE cervical and colon cancers. For analysis, we used a novel algorithm, lesser allele intensity ratio (LAIR), which is incorporated in beadarraySNP (5). LAIR integrates the DNA index in the analysis and defines the allele intensity ratio (LAIR), which is similar to that of paired alleles of the reference sample (balanced, left dotted line; two copies [AB], four copies [AABB], etc.). LAIR is 0 when no signal is found for one of the alleles in the tumor (LOH, right dotted line; one or more copies [A], [AA], and [AAA], etc.). Allelic imbalances (imbalanced) are indicated by intermediate values depending on the copy number ratio between the two alleles: [AAABB], [AAAAB], [AAAB], and [AAB] are shown equidistantly (n = number of copies).

**Materials and Methods**

**Tumor dissociation.** Cervical and colorectal tumors were obtained from the FFPE tissue bank of the Department of Pathology, Leiden University Medical Center (LUMC), Leiden, the Netherlands. Samples were handled according to the medical ethical guidelines described in the Code Proper Secondary Use of Human Tissue established by the Dutch Federation of Medical Sciences.1 Paraffin sections taken from all samples were H&E-stained and reviewed by two pathologists (G.J. Fleuren and H. Morrean). Cell suspensions were prepared as described (24) from either 6 to 10 60-μm sections or 2 4-mm tissue punches from each paraffin block.

**Antibodies.** Clone MNF116 [anti-keratin 5, 6, 8, and 17, IgG1 (DAKO)] was used at working concentrations of 2 μg/mL for 1 × 10⁶ cells and 10 μg/mL for 5 × 10⁶ cells. Clones AE1/AE3 [anti-pan-keratin, premixed 20:1, IgG1 (Chemicon)] were used at working concentrations of 5 μg/mL for 1 × 10⁶ cells and 25 μg/mL for 5 × 10⁶ cells. Clone V9-2b (anti-vimentin, IgG1) was developed in our department, was used as a diluted culture supernatant (1:5 or 1:1, depending on the cell concentration). Goat F(ab′)₂ anti-mouse IgG-FITC and goat F(ab′)₂ anti-mouse IgG2a-RPE (Southern Biotechnology Associates) were both diluted 1:100 in PBATw.

**Staining.** One million cells were incubated with 100 μL of a monoclonal antibody mixture containing clones MNF116, AE1/AE3, and V9-2b overnight at 4°C. The next day, cells were washed twice with ice-cold PBATw and centrifuged at 500 × g for 5 min at 4°C. The cells were then incubated with 100 μL of premixed FITC- or RPE-labeled secondary reagents. After 30 min on ice, cells were washed twice with ice-cold PBATw and incubated with 500 μL of DNA staining solution containing 10 μmol/L of propidium iodide (PI; Calbiochem) and 0.1% DNsase-free Rnase (Sigma) diluted in PBATw. Cells were kept at room temperature for 30 min to activate the RNase and were then incubated at 4°C overnight to allow for stoichiometric staining of the DNA.

For DNA index validation, two tissue blocks from an archival cervical carcinoma were taken and thick sections were cut at different time intervals and prepared for multiparameter DNA analysis as described. In total, nine independent measurements were performed, of which the DNA index and coefficient of variation (CV) of the G₀G₁ populations was calculated.

**Flow cytometry and sorting.** For analysis, data from 20,000 single cell events were collected using a standard FACScalibur (BD Biosciences) flow cytometer, equipped with a 15 mW Argon-ion laser (488 nm) and a 12 mW diode laser (635 nm; ref. 30). The FL-3-A versus FL-3-W pulse processor was used to enrich for single cell events during acquisition and analysis. For data analysis, DNA index, and CV calculation, the WinList 6.0 and ModFit 3.1 software packages were used (Verity Software House, Inc.). N-color compensation was used for postacquisition spectral cross-talk correction according to the manufacturer’s instructions, without the use of hyperlog transformation or log bias.

For flow-sorting, the cell concentration was increased to 5 × 10⁶ cells/mL. The PI concentration was simultaneously increased to 50 μmol/L. G₀G₁ vimentin-negative, keratin-positive tumor cells and G₀G₁ vimentin-positive, keratin-negative stromal cells were flow-sorted using a FACSaria flow-sorter at 40 psi (BD Biosciences) with a 100-μm nozzle at a frequency of ~52 kHz. The 488 laser line was used for excitation. The FACSaria purity mode was used during sorting. These settings allowed us to typically flow-sort 800 × 10⁶ cells in 5 mL Falcon tubes. The following detector and filter settings were used during sorting: FITC fluorescence, detector E, 530/30 nm BP filter; R-PE fluorescence, detector D, 575/26 nm BP filter; PI fluorescence, detector C, 610/20 nm BP filter. A detector C-Area versus detector C-Width dot plot was used to gate out doublet and aggregates during sorting. After sorting, cells were centrifuged at 4,000 × g for 10 min before DNA was extracted.

For fluorescence in situ hybridization (FISH) analysis of flow-sorted cells (20 psi, 100 μm nozzle), samples were labeled for keratin, vimentin and DNA using APC- and RPE-conjugated antibodies, and 4′,6-diamidino-2-phenylindole as DNA stain. This approach reduced background fluorescence during the examination of the interface nuclei after hybridization.

**DNA isolation.** DNA was isolated as described (31) and DNA was further purified using the Promega Protein Precipitation solution (Promega) according to the manufacturer’s instructions. DNA concentrations were determined using the PicoGreen method (Invitrogen).

**SNP array analysis.** SNP arrays were performed at the Leiden Genome Technology Center2 as described (32) with minor modifications: 1 μg of DNA was used as the input in a multi-use activation step and was subsequently dissolved in 60 μL of resuspension buffer. Genotypes and the Gene

---

1 http://www.federa.org/

2 http://www.lgtc.nl/
Call Score were extracted using GeneCall version 6.0.7 (Illumina). Only genotypes with a GCS <0.5 were used to eliminate low-quality calls from this analysis.

Reference normal sample with the highest average quality score was chosen. LOH within a tumor sample was determined by comparison to the reference sample. Informative SNPs, SNPs at which the reference sample is heterozygous, were checked for homozygosity and LOH was indicated when two neighboring informative SNPs were homozygous in the tumor sample. The beadarraySNP package (5) was adapted to combine copy number profiles, allele-specific intensities, and the DNA index.

First, a segmentation procedure is applied to find genomic regions that have the same copy number (33). Secondly, LAIR, a measure of the contribution of the two original alleles, is calculated for all informative SNPs. This value is close to 1 when the contribution of both alleles of an SNP to the total intensity in the tumor is similar to that of the reference sample. The value is close to 0 when there is no signal for either of the alleles in the tumor (LOH). Allelic imbalances will show intermediate values (Fig. 1). By using the assumption that regions with LAIR close to 1 should have an even copy number and that the total calculated DNA index should be similar to the measured DNA index obtained by flow cytometry, it is possible to determine the allelic copy number in each genomic region. We refer to this as the allelic state. The following allelic states [modified from Nancarrow and colleagues (34)] can be distinguished: (a) [AB], normal; (b) [A], abnormalities with copy number 1, called LOH, resulting from the loss of either the A or B allele; (c) [AA], a diploid abnormality with either the genotype AA or BB, which is referred to as cnLOH; (d) [AAA] or [AAAA], etc., or amplified LOH, in which only the A or B allele is present at the locus and is present in three or more copies; (e) [AABB] or [AAABB], etc., or balanced amplifications resulting in a 4n, 6n or higher genomic region with equal amounts of both alleles; and (f) allelic imbalances for [AABB], [AAABB], or [AAAAAB], etc., for which the copy number at the locus is three or higher and both alleles are present at unequal amounts.

Generally, all SNPs on the array are equidistantly spaced across the genome. Thus, the average copy number of the allelic states of all SNPs, the allelic state index, was calculated and found to be an accurate measure of the copy number of the tumor. The allelic state index closely matched the DNA index measured by flow cytometry.

Interphase FISH analysis of flow-sorted cells. Interphase FISH analysis was performed as previously described (33) on 500 to 2,000 cells. The following probes were used: aliphid satellite centromeric probes for chromosome 4 (PYAM 11.39, kindly provided by Dr. A. von Bergh, Department of Clinical Genetics, Erasmus MC, Rotterdam, the Netherlands), chromosome 6 (p308), chromosome 8 (D8Z2), and chromosome 18 (L184; kindly provided by Dr. K. Szuhai, Department of Molecular Cell Biology, LUMC, Leiden, the Netherlands). The probes were biotin-16-dUTP–labeled (Roche Diagnostics) by standard nick translation. The BAC probes 149A7 (4q), 86C11 (6p), 10G10 (8q), 536K17 (8q), 748M14 (18q), and 154H12 (18q; gift from Dr. K. Szuhai) were similarly labeled with digoxigenin-12-dUTP (Roche Diagnostics).

The centromere CEP17 Alpha SpectrumGreen and Vysis LSI TP53 SO SpectrumOrange (17p) probes were purchased from Abbot Molecular, Inc. Hybridization and immunodetection were performed as advised by the manufacturer with an additional denaturation step for 8 min at 80°C.

Results

Flow cytometry and sorting. Flow cytometric analysis of six solid tumors showed DNA histograms with two G0G1 fractions, which shows the presence of an aneuploid population in all cases. Representative DNA histograms and dot plots are shown in Fig. 2. After gating of the vimentin-positive, keratin-negative populations, a single population of tumor stromal cells remained in all cases (CV, 4.48 ± 0.78%). In the vimentin-negative, keratin-positive cell fraction of five cases, a single DNA aneuploid population was found. The DNA index ranged from 1.30 to 1.91 (Supplementary Table S1). The vimentin-negative, keratin-positive population of sample 5 showed two distinct populations with a near-diploid DNA index of 0.97 and an aneuploid fraction with a DNA index of 1.87. Cell suspensions were simultaneously stained for keratin (epithelial cells, FITC), vimentin (stromal cells, R-PE fluorescence), and DNA (PI fluorescence). Sorting was restricted to the G0G1 populations (vertical bars, right).

Figure 2. Flow-sorting of tumor cell subpopulations from cervical and colorectal carcinoma FFPE tissue. A, cervical cancer sample shows a unimodal DNA histogram after gating of the vimentin-negative, keratin-positive cell fraction. B, colorectal sample shows a bimodal DNA histogram after keratin gating, containing two DNA fractions: a near-diploid fraction with a DNA index of 0.97 and an aneuploid fraction with a DNA index of 1.87. Cell suspensions were simultaneously stained for keratin (epithelial cells, FITC), vimentin (stromal cells, R-PE fluorescence), and DNA (PI fluorescence). Sorting was restricted to the G0G1 populations (vertical bars, right).

Downloaded from cancerres.aacrjournals.org on July 20, 2017. © 2008 American Association for Cancer Research.
Figure 3. The allelic state of the autosomes for all cases provides a detailed measure of chromosomal aberrations. Cervical carcinomas, cases 1 to 3 (A); colon carcinomas, cases 4 to 6 (B). For each chromosome, the number of bar(s) to the right of the ideogram indicates the copy number. The bars are only depicted for informative SNPs (heterozygous in normal tissue). Red, the A allele; blue, B allele. Orange and cyan bars, clonal heterogeneity of this locus. For example; case 1, chromosome 7—a mixture of [AB] and [AABB] allelic states; see also Supplementary Fig. S2. Note the striking clonal relation between the near-diploid and near-tetraploid fractions of case 5. Besides endoreduplication of most chromosomes, including LOH ([AA]) on chromosome 5, the breakpoints on chromosomes 5, 8, 17, and 20 are identical.
Signal amplitudes and improved detection of chromosomal alterations. Using the beadarraySNP package, LOH profiles and relative copy numbers were generated for all cancers. Vimentin-positive, keratin-negative stromal cells from the archival tissue were used as a reference because these showed a high genotypic concordance (>98.8%) with paired normal fractions (lymph nodes or endometrium) and contained normal diploid genomes.

To evaluate the effect of flow-sorting for the macrodissected cervical tumors, the unsorted and flow-sorted tumor fractions were compared (Supplementary Table S2). Flow-sorting considerably improved signal amplitudes and identification of chromosomal segments with LOH (Supplementary Fig. S1). In total, 119 segments were identified: 24 with gains, 33 with physical losses; 12 showed cnLOH, and 50 showed retention. Of these segments, 92 were identical in the sorted and unsorted fractions, although 22 of the CNAs (31.8%) showed an increase in amplitude. Of all of the alterations.

Intratumor heterogeneity. In case 5 (colon cancer), two vimentin-negative, keratin-positive tumor fractions with a different DNA index were clearly distinguished (0.97 and 1.86 for the near-diploid and aneuploid fraction, respectively; Fig. 2). To study their clonal relation, both fractions were analyzed for genomic aberrations. Fifteen chromosomes (1–5p, 7, 9–12, 14–16, 18, 20–22) were duplicated in the aneuploid fraction compared with the near-diploid fraction (Fig. 3; Supplementary Fig. S2). For example, the allelic states of chromosomes 1, 2, and 3 were duplicated from [AB] to a balanced gain [AABB] in the aneuploid fraction. Similarly, LOH [A] at 8p, 15, and 18 in the near-diploid fraction was duplicated to cnLOH [AA] in the aneuploid fraction. FISH confirmed the chromosome 18 copy numbers (Fig. 4B and C).

<table>
<thead>
<tr>
<th>Locus</th>
<th>Probe</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5, near-diploid fraction</th>
<th>5, aneuploid fraction</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>4q28.3</td>
<td>149A7</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>2/2 85%, 4/4 13%</td>
<td>4/4</td>
<td>—</td>
</tr>
<tr>
<td>6p25.1</td>
<td>86C11</td>
<td>3/3</td>
<td>31%, 4/4 67%</td>
<td>—</td>
<td>—</td>
<td>2/2</td>
<td>4/4 34%, 4/2 57%</td>
<td>4/4</td>
</tr>
<tr>
<td>8q22</td>
<td>10G10</td>
<td>2/2 33%, 3/3 63%</td>
<td>—</td>
<td>—</td>
<td>2/2 52%, 4/4 31%, &gt;4/&gt;4 12% (52)</td>
<td>&gt;4/&gt;4</td>
<td>4/4 75%, 4/3 15% (20)</td>
<td>—</td>
</tr>
<tr>
<td>8q23.3</td>
<td>536K17</td>
<td>2/2 42%, 3/3 52%</td>
<td>—</td>
<td>—</td>
<td>2/2 86%, &gt;4/&gt;4 10%</td>
<td>&gt;4/&gt;4</td>
<td>4/4 85%, 4/3 10% (20)</td>
<td>—</td>
</tr>
<tr>
<td>Smad2</td>
<td>748M14</td>
<td>3/3 38%, 4/4 50%</td>
<td>—</td>
<td>—</td>
<td>2/2 36%, 1/1 61%</td>
<td>2/2</td>
<td>2/2</td>
<td>—</td>
</tr>
<tr>
<td>18qter</td>
<td>154H12</td>
<td>3/3 42%, 4/4 58%</td>
<td>—</td>
<td>—</td>
<td>2/2 50%, 1/1 43%</td>
<td>2/2</td>
<td>2/2</td>
<td>—</td>
</tr>
<tr>
<td>DNA index</td>
<td>1.84</td>
<td>1.91</td>
<td>1.82</td>
<td>1.30</td>
<td>0.97</td>
<td>1.86</td>
<td>1.54</td>
<td>—</td>
</tr>
<tr>
<td>Allelic state index</td>
<td>1.86</td>
<td>1.88</td>
<td>1.88</td>
<td>1.32</td>
<td>0.96</td>
<td>1.95</td>
<td>1.54</td>
<td>—</td>
</tr>
</tbody>
</table>

NOTE: The DNA index and the allelic state index are given for each fraction (—, not determined).

* One hundred nuclei were counted, unless otherwise noted in parentheses.
† Copy numbers are given as n/n representing centromere signal/probe signal.
‡ >4 = more than four copies.
Figure 4. FISH confirmation of allelic state copy number and tumor heterogeneity. A, case 6, allelic state analysis of chromosomes 8 and 18 and interphase FISH for chromosome 18q on vimentin-positive, keratin-negative and vimentin-negative, keratin-positive nuclei. Two centromere 18 signals (red) and two signals (green) in the SMAD2 region are visible. B, case 5, near diploid (DNA index, 0.97) fraction. Copy numbers of the allelic states at 8q ([AAAB]) and 18 ([A]) are confirmed by FISH. On chromosome 8, four centromere signals (red) and four 8q signals (green) are visible, and on chromosome 18, one centromere signal and one 18q signal. Inset, vimentin-positive, keratin-negative interphase nucleus showing two centromere signals and two 18q signals. C, case 5, aneuploid fraction (DNA index, 1.86) and FISH for chromosomes 8 and 18 of the aneuploid fraction. Chromosome 8 shows seven centromere signals and seven 8q signals. Chromosome 18 shows two centromere signals and two 18q signals. Note the striking similarity between the chromosomal aberrations of different colon tumor fractions. Chromosomes 8 and 18 are from different OPA panels, causing the small difference in the level of the red segmentation line. D, intratumor heterogeneity observed by FISH. The allelic state of chromosome 6p of case 2 was [AAAA] according to LAIR analysis (a low LAIR score; red bars). FISH analysis of flow-sorted vimentin-negative, keratin-positive tumor cells showed that this population is composed of a mixture of two fractions: one fraction (67%) containing four copies of 6p25/4 centromeric copies and one fraction (33%) containing three copies of 6p25/3 centromeric copies. The vimentin-positive, keratin-negative fraction was shown to be normal (2/2, [AB]). [A], [AAB], etc., indicate the allelic state; black dots, normalized copy number with a red segmentation line for all SNPs; horizontal blue dashes, LAIR (calculated on informative SNPs) scale from 0 to 1; vertical bars; green, LAIR = 1 (retention); red, LAIR = 0 (LOH); blue, intermediate LAIR (-0.2 to -0.8, allelic imbalance). Probes: centromere 6, p308 (red); 6p, 86C11 (green); centromere 8, D8Z2 (red); 8q, 536K17 (green); centromere 18, L1.84 (red); 18q, 748M14 (green); 18qter, 154H12 (green).
Allelic State Analysis of Chromosomal Aberrations

(data not shown); however, FISH revealed a nearly equal mixture of 2n and 4n abnormalities on chromosome 6p in the aneuploid fraction (Table 1). These abnormalities may be a mixture of either [AA] and [AABB] or [AB] and [AAABB], leading to the detected [AAB]. Similarly, in case 2 (cervical squamous cell carcinoma), the allelic state estimate of chromosome 6p is [AAAA] (Fig. 3; Supplementary Fig. S2), whereas FISH analysis of the flow-sorted G2G1 aneuploid vimentin-negative, keratin-positive tumor fraction clearly revealed a mixture of a 3n and 4n population, with allelic states [AAA] in 51% and [AAAA] in 67% of the nuclei, respectively (Fig. 4D).

**Discussion**

We showed the feasibility of combining high-throughput SNP arrays and flow-sorting of tumor cell subpopulations from different formalin-fixed archival samples. Our approach significantly improved the simultaneous detection of numerical and structural chromosomal aberrations, allowing clear discrimination between allelic imbalances and LOH, and definition of aberrant chromosomal segments. Signal amplitudes of the losses and gains were generally higher when compared with those of the unsorted samples. Furthermore, <20% of the chromosomal abnormalities were missed in the unsorted cervical tumor fractions, whereas 16% were interpreted differently, either due to intratumor heterogeneity or to low tumor percentage. In the sorted samples, absolute LOH ([A]) or cnLOH ([AA]) was frequently identified, creating a sharp definition of a chromosomal segment.

We determined the allelic states of flow-sorted tumor fractions and confirmed copy numbers by FISH for several loci. The concordance between the DNA index and the allelic state index verified the method. In line with Nancarrow and colleagues (34), we propose the use of the following nomenclature: I, LOH for the loss of one allele, resulting in the allelic state [A]; II, cnLOH for the allelic state [AA]; III, amplified LOH for the allelic states [AAA], [AAAA], etc.; IV, balanced amplification for 4n [AABB] or 6n [AAABB], etc.; and V, allelic imbalances for [AB], [AAABB], [AAAAA], etc. We prefer allelic imbalance over amplification because allelic imbalance is more neutral and not suggestive of the molecular mechanism that was responsible for the observed abnormality. For example, an [AAB] status can be caused by a single gain of an A allele within an overall diploid genomic background; however, for a highly aneuploid tumor, this status can arise from endoreduplication of [AB] to an [AABB] status, followed by an additional loss of a B allele during tumor progression.

For accurate estimates of the allelic states of CNAs in FFPE aneuploid tumors, we showed that integration of the DNA index in conjunction with SNP arrays (8, 27) is crucial. Otherwise, most, if not all, CNAs will be misinterpreted, which is also supported by the findings of Lyng and coworkers on aCGH (16). For example, the patterns of CNAs on chromosomes 8 and 18 of the near-diploid and the aneuploid fractions of case 5 seem to be identical, but their allelic states are shown to differ after integration of the DNA index (Fig. 4). For near-diploid tumors, the allelic state of a tumor fraction will likely be correctly interpreted (14). Determination of the DNA index alone is feasible in most laboratories. However, for highly intermingled tumors (infiltrate and stroma) or for tumors containing different tumor subpopulations, flow-sorting seems to be mandatory.

An important potential of flow cytometry is the identification and sorting of multiple clones based on differences in DNA content. This advancement has not yet been achieved by any other enrichment or purification method. This approach allows for the study of intratumor heterogeneity, chromosomal aberrations which develop during tumor progression and clonal relationships between tumor subpopulations (20, 36). A clear example is case 5, in which the near-diploid and aneuploid tumor fractions probably originated from a common hypothetical near-diploid precursor fraction. Most chromosomes were endoreduplicated from [AB] in the near-diploid fraction to [AABB] in the aneuploid fraction (Fig. 3). Subsequent parallel divergence of the two clones resulted in the current fractions. For example, on chromosome 13, physical loss and mitotic recombination could have led from [AB] in the common precursor to cnLOH [AA] in the near-diploid fraction. The [AAAAA] status of chromosome 13 of the aneuploid fraction could have resulted from endoreduplication with subsequent loss of the B allele, and doubling of the A alleles.

In addition, the aneuploid fraction showed a mixture of two subpopulations for chromosome 6p, which were only identified after FISH analysis. At this locus, the near-diploid fraction showed a normal allelic state [AB]. Given the high DNA index (1.86), an average of three copies, an intermediate LAIR-score ([AAB]) and FISH analysis, these aneuploid subpopulations likely contained allelic states [AA] (cnLOH) and [AABB], respectively, clearly demonstrating heterogeneity within the tumor fraction. Similar intratumor heterogeneity within tumor cell fractions was shown for case 2 (Fig. 4D). These delicate differences in chromosomal copy numbers might indicate the ongoing generation of tumor subclones due to chromosomal instability (37). These differences also show that intratumor heterogeneity is more extensive than is generally observed by ploidy and LAIR analysis because both techniques register the dominant clone(s). Intratumor heterogeneity has also been identified by aCGH (16). However, LAIR analysis has increased value relative to classical methods. It allows for a more accurate estimate of the true allelic state of a chromosome or regions of chromosomes, which, for example, could result in discrimination between three chromosomal copies [AABB] and a balanced mixture of two allelic states, [AB] and [AABB].

Integration of copy number and ploidy analysis creates a detailed view of chromosomal aberrations during tumor progression. In all three aneuploid cervical carcinoma samples, LAIR analysis revealed amplified LOH, [AAAA], of the HLA region on chromosome 6p. This region is known to be frequently targeted by LOH in cervical cancer (8, 38, 39). From these events, it might be concluded that LOH on chromosome 6p occurred in a near-diploid precursor fraction, followed by mitotic recombination and endoreduplication, leading to aneuploidy. The identification of the chromosomal break points of the amplified LOH on 6p also indicates the possibility of accurately studying the smallest regions of overlap.

The consequences of allelic state analysis for the classification of tumors and profiling, as well as for allelic dosage determination of cancer-related genes, remain to be established in a large cohort of flow-sorted tumors. For example, in a near-diploid tumor with predominantly diploid chromosomes, a locus with the allelic imbalance [AAB] will be interpreted as a gain of the A allele, whereas in a tumor with a predominantly near-tetraploid [AABB] genomic background, [AAB] will be interpreted as an additional loss of a B allele after endoreduplication. One could speculate on the relation between the allelic dosage and RNA expression of the genes on segments with CNAs (40). For the [AAB] segment, the effect might be the opposite in a near-diploid background as compared with an aneuploid background.
In conclusion, we have explicitly shown that a combined approach of flow-sorting and high-throughput SNP array profiling with DNA index integration (a) significantly improves the detection of numerical and structural chromosomal aberrations in formalin-fixed tumor samples, (b) allows the definition of the allelic state of complex CNAs and, (c) provides an individual molecular measure of chromosomal aberrations. Interestingly, a recent meta-analysis showed that chromosomal aberrations are associated with a worse prognosis in colorectal cancer (3) and should be evaluated as a prognostic marker in clinical trials. Our new procedure could be implemented to define chromosomal aberrations and the allelic state of regions that are identified in whole genome association studies and which are associated with a cancer risk (41). Finally, our combined method may provide further insight into the dynamics of genomic instability during clonal evolution in heterogeneous human carcinoma samples and can be implemented in large-scale retrospective studies and studies in which detailed genome-wide information is required.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

Received 7/11/2008; revised 9/11/2008; accepted 10/14/2008.

Grant support: Dutch Cancer Society grant UU.2005-3247 (T. van Wezel and H. Morreau).

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.

We thank Yavuz Ariyurek for assistance with SNP arrays, Esther Lips for unsorted colorectal tumor samples, and Karoly Szunai and Anne von Bergh for providing FISH probes.

References


Genome-wide Allelic State Analysis on Flow-Sorted Tumor Fractions Provides an Accurate Measure of Chromosomal Aberrations

Willem E. Corver, Anneke Middeldorp, Natalja T. ter Haar, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/68/24/10333

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2008/12/12/68.24.10333.DC1

Cited articles
This article cites 41 articles, 18 of which you can access for free at:
http://cancerres.aacrjournals.org/content/68/24/10333.full#ref-list-1

Citing articles
This article has been cited by 1 HighWire-hosted articles. Access the articles at:
http://cancerres.aacrjournals.org/content/68/24/10333.full#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.