Molecular discrimination of sessile rectal adenomas from carcinomas for a better treatment choice: integration of chromosomal instability patterns and expression array analysis

Lips, E.H.

CHAPTER 2

Reliable high-throughput genotyping and loss-of-heterozygosity detection in formalin-fixed, paraffin-embedded tumors using single nucleotide polymorphism arrays

Esther H. Lips, Jan Willem F. Dierssen, Ronald van Eijk, Jan Oosting, Paul H.C. Eilers, Rob A.E.M. Tollenaar, Eelco J. de Graaf, Ruben van 't Slot, Cisca Wijmenga, Hans Morreau and Tom van Wezel.

Cancer Res. 2005 Nov 15; 65(22):10188-91

Abstract

Most human cancers show genetic instabilities leading to allelic imbalances, including loss of heterozygosity (LOH). Single nucleotide polymorphism (SNP) arrays can be used to detect LOH. Currently these arrays require intact genomic DNA as obtained from frozen tissue, however for most cancer cases only low quality DNA from formalin-fixed paraffin embedded (FFPE) tissue is available. In this study we tested Illumina BeadArrays to genotype FFPE tissue and detect LOH/allelic imbalances in matched colorectal tumor and normal tissue. Genotypes were compared between leukocyte and FFPE normal tissue as well as between frozen and FFPE tumor tissue. Identical genotypes and LOH profiles were obtained from normal and tumor isolates. LOH was mainly observed on chromosomes 4, 5q, 12q, 14q, 15q, 17p, 18, and 20p, which are commonly detected regions in colorectal cancer. LOH profiles of the BeadArrays were compared with profiles obtained by Affymetrix GeneChip 10K arrays, showing identical LOH patterns. These data show that genome-wide genotyping of FFPE tissue with the BeadArray gives reliable results and is a powerful technique for LOH analysis.

Introduction

Colorectal cancers, like most human cancers, are characterized by genetic instabilities, indicated by allelic imbalances (AI) and loss of heterozygosity (LOH). Such alterations can lead to inactivation of tumor suppressor genes. Recently introduced single nucleotide polymorphism (SNP) arrays can be used for high resolution genome-wide genotyping and LOH detection (1-5). These arrays typically require high quality DNA from fresh frozen tissue and leukocytes, which often is not available for large retrospective series of cancer cases. In contrast, tissue archives contain vast amounts of formalin-fixed, paraffin-embedded (FFPE) tissue of large clinical series with follow-up data. However, DNA is degraded during formalin fixation, and unsuitable for most SNP array platforms. Recently, BeadArrays with a SNP linkage mapping panel, consisting of 5861 SNPs distributed over the human genome, have been introduced (6). These arrays allow allelic discrimination directly on short genomic segments of ~40 bp surrounding the SNPs of interest, thus overcoming the need for high quality integer DNA. Using two oligonucleotides that recognize the SNP and flanking sequences a synthetic PCR template is generated by allele specific primer extension and subsequent ligation (7). In this study we show that these BeadArrays can be used to obtain reliable genotyping and genome wide LOH profiles from FFPE normal and tumor tissue, respectively, and compared them with normal leukocyte DNA and frozen tumors and with LOH profiles obtained by GeneChip arrays.

Material and Methods

Samples

Colorectal tumor tissue and corresponding normal tissue from four patients were used, following medical ethical guidelines (www.fmwv.nl/gedragscodes/goedgebruik/CodeProperSecondaryUseOfHumanTissue.pdf). From one additional case only normal tissue was used. The samples included one rectal adenoma and three right-sided carcinomas (one Dukes B and two Dukes C) (Table 1). Ploidy status of the tumors was previously assessed by flow cytometry[1]. A pathologist (H.M.) assessed the normal and tumor areas and percentage of tumor cells based on H&E slides.

For DNA isolation of fresh frozen tissue, twenty 30 μm-thick sections were cut from each tumor. DNA was isolated with the Genomic Wizard kit, according to the manufacturers' protocol (Promega, Madison, WI). Leukocyte DNA was obtained by salting out precipitation as described (8). DNA from FFPE tissue was extracted with the Chelex extraction method, as described previously (9). Briefly, three tissue punches (diameter, 0.6

[1] J.W. Dierssen et al. High-resolution analysis of HLA class I alterations in colorectal cancer. BMC Cancer. 2006 Oct 2;6:233.

mm) were obtained by a tissue microarrayer (Beecher Instruments, Sun Prairie, WI), and DNA was isolated with Chelex and proteinase K. FFPE DNA was subsequently cleaned up using the Genomic Wizard kit (Promega). DNA concentrations were measured with the picogreen method (Invitrogen-Molecular Probes, Carlsbad, CA), and DNA quality was checked on a 1% agarose gel. For each cell isolate, 1µg of DNA was used for the BeadArrays, whereas 250 ng DNA was used for the GeneChip arrays.

Single Nucleotide Polymorphism array analysis

Illumina BeadArrays in combination with the linkage mapping panel version 4 (Illumina, San Diego, CA) were used, containing 5861 SNP markers distributed evenly over the genome with a physical distance of on average 482 kb. Samples were prepared according to the Goldengate assay (7). Gene calls were extracted using the gene calling programs GeneCall and GTS Reports (Illumina). A quality score, the gene call score (GCS), was automatically assigned to each genotype call. A minimal GCS can be selected that will be used to zero all genotypes below this value. Removing genotypes below this level reduced the effective call rate and removed the lowest quality genotypes. To compare genotypes of different tissue isolates a GCS cutoff was used, to eliminate wrongly genotyped SNPs. In the LOH analyses all SNPs independent of their GCS were used.

In parallel, DNA from leukocytes and fresh frozen tumors was hybridized to GeneChip Mapping 10K 2.0 arrays (Affymetrix, Santa Clara, CA) at the Leiden Genome Technology Centre (www.lgtc.nl), as previously described (10). Genotypes were scored with the GDAS software (Affymetrix).

For both platforms, LOH was determined by comparing the genotypes from tumor and their corresponding normal tissue. To this end gene call data were imported into an Access database (Microsoft, Redmond, WA). Using a SQL query, LOH was called for stretches of two or more SNPs scoring heterozygous (AB) in the normal sample and homozygous (AA or BB) in the tumor sample. These data were visualized with a chromosome visualization tool (11) in Spotfire DecisionSite (Spotfire, Sommerville, MA).

DNA sequencing

DNA sequencing was used to validate the results of 10 SNPs on chromosomes 1, 2 and 22, both from regions with and without LOH. Primers sequences are available upon request. Sequencing was done at the Leiden Genome Technology Centre (www.lgtc.nl) and analyzed with the Mutation Surveyor software package (Softgenetics, State College, PA).

Results and Discussion

To genotype FFPE tissue and to determine genome-wide LOH with BeadArrays, we used four colorectal cancer samples, from which normal leukocytes, fresh frozen tumor tissue and both normal and tumor FFPE tissue were available. Genotypes between different cell isolates were compared and subsequently genome-wide LOH was detected and confirmed with GeneChip arrays.

The suitability and reliability of the BeadArrays for genotyping FFPE tissue was tested between DNA from leukocytes and normal FFPE tissue for two cases and between frozen tumor and tumor FFPE tissue for four cases. When comparing genotypes between normal leukocytes and normal FFPE tissue, the concordance was 99.5% and 99.3% for cases 124 and 514 respectively (Table 1). Applying a GCS cutoff value of 0.5, resulted in 100 % identical genotypes for sample 124 and 99.7 % for sample 514, indicating that the minor differences were due to bad genotype calls. Setting this GCS cutoff resulted in a decline in total number of SNPs called (4370 (sample 124) and 4650 (sample 514) instead of 5861). The fact that 100% identity was not reached for 514 can be explained by genetic abnormalities in histologically normal tissue that might already be present to some extent in normal tissue adjacent to tumor tissue. Nevertheless, the percentage discrepant SNPs was very small (0.3%). When FFPE tumor samples were compared with their corresponding frozen samples, percentages identical SNPs reached 99,9 % with a GCS cutoff value of 0.5. Tumor heterogeneity possibly accounted for the minor genotyping differences because frozen and FFPE sections were derived from different parts of the same tumor and slightly varied in tumor cell percentage due to stromal and infiltrate components. Moreover, the BeadArray and the GeneChip array have 99 SNPs in common and comparing the leukocytes and frozen tumor samples for these corresponding SNPs showed identical

Table 1. Sample characteristics and concordance of genotype calls between different isolates.

				NL vs FFPE normal			FFT vs. FFPE tumor		
Patient ID	Stage	Ploidy	Tumor %	% identical SNPs*	discrepancies at all GCS†	discrepancies at GCS >0.5†	% identical SNPs*	discrepancies at all GCS†	discrepancies at GCS >0.5†
124	C1	aneuploid	70	99.5 (100)	0.5% (29/5861)	0 (0/4370)			
514	adenoma	aneuploid	60	99.3 (99.7)	0.7% (41/5861)	0.3%(14/4650)	99.4 (99.9)	0.6% (35/5861)	0.1% (5/4644)
44	B2	aneuploid	50				97.4 (98.6)	2.6% (152/5861)	1.4% (57/4074)
106	C2	multiploid	90				98.6 (99.6)	1.4% (82/5861)	0.4% (17/4347)
108	C2	multiploid	80				97.2 (98.9)	2.8% (164/5681)	1.1% (48/4322)

Abbreviations: NL, normal leukocytes, FFT, fresh frozen tumor.

* Percentage identical genotypes in brackets with a GCS>0.5

† Percentage discrepant SNPs, number discrepant SNPs and total number of SNPs in brackets.

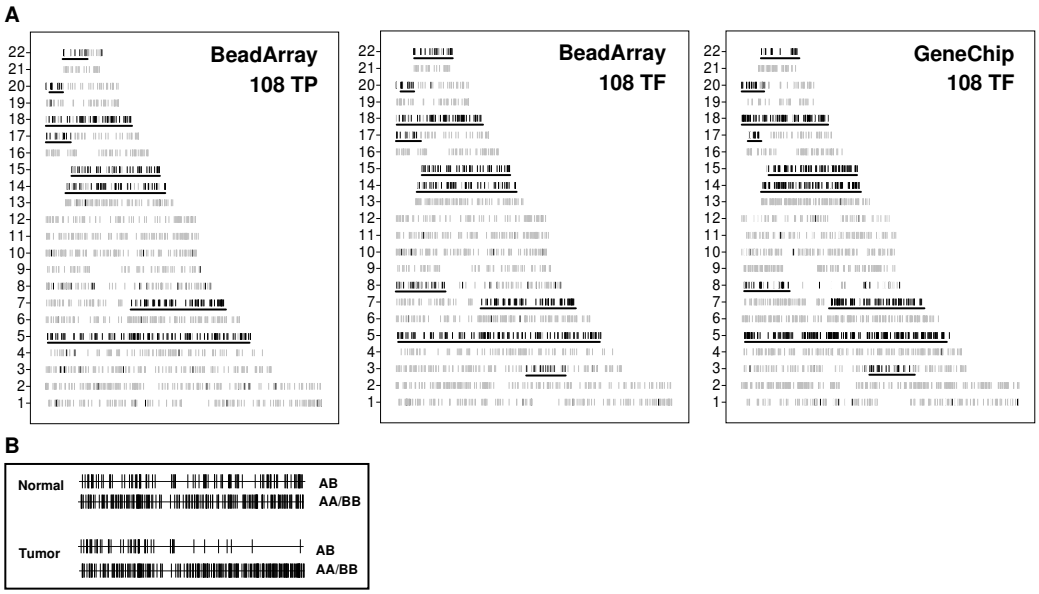

Figure 1. LOH profiles of tumor 108. A, all SNPs that are heterozygous in the normal leukocytes are aligned along the autosomes, determined by BeadArrays and GeneChip arrays. SNPs that are heterozygous in the tumor tissue (gray), and homozygous (black) indicating LOH, are shown. Only stretches of more than 2 SNPs were called as LOH (underlined with a black bar). TP, FFPE tumor tissue; TF, tumor frozen tissue. B, alignment of all SNPs for chromosome 7 as typed with the BeadArray. All heterozygous (AB) and homozygous (AA/BB) SNPs for the normal leukocytes and the FFPE tumor sample. The remarkable drop in heterozygous SNPs in the tumor alone indicates loss of chromosome 7q in this tumor.

genotyping (data not shown). From these data can be concluded that genotyping FFPE tissue using bead arrays is highly reliable and identical to leukocyte DNA or frozen tumor.

LOH was determined by studying SNPs, which were heterozygous in the normal leukocytes and homozygous in the matching tumor tissue, both in FFPE and in fresh frozen tumor tissue (Figure 1A). Stretches of homozygous SNPs were considered as LOH, like on chromosomes 5 and 7q. Single homozygous SNPs (e.g. on chromosomes 1 and 2) were considered as noise and not taken into account. In sample 108, all large LOH regions observed were identical between DNA from frozen or FFPE tumor tissue. SNP determination in the FFPE tissue on two smaller regions, chromosomes 3 and 8, was not completely consistent with the frozen tissue (Figure 1A). The fact that this tumor is multiploid, consisting of diploid and aneuploid populations, might have caused the inconsistency between these tumor samples.

An example of a region of LOH for chromosome 7 is shown in Figure 1B. It should be noted that the LOH region on 7q displays heterozygous SNP markers in the normal that

become homozygous in the tumor sample. In fact, even without the corresponding normal tissue, the large stretch of homozygous SNPs is indicative for LOH in this tumor sample. In this way large homozygous regions can be detected as LOH in samples without an accompanying normal tissue. The few heterozygous SNPs in the LOH region can be explained by low GCSs, and some mixing of stroma or infiltrate in the tumor sample.

We confirmed our findings by typing LOH for all our samples on corresponding normal leukocytes and frozen tumors using GeneChip arrays and by sequencing 10 individual SNPs both with and without LOH. All regions of LOH were confirmed using the GeneChip arrays (Figure 1A). Moreover, SNP sequencing validated both the genotyping and LOH detection results (Figure 2).

Table 2 summarizes the regions of LOH in all samples. In four tumors many LOH regions were detected, especially on chromosomes 4, 5q, 12q, 14q, 15q, 17p, 18, and 20p. LOH at 17p and 18q was observed in all these tumors. These loci contain the tumor suppressor genes *p53* and *SMAD4* respectively, which play a role at early stages of colorectal cancer and often show LOH (12-16). LOH at 5q, containing the *APC* gene, is another common event in colorectal cancer and was observed in three tumors. These data clearly show that LOH analysis on FFPE tissue using the BeadArray can identify biologically significant regions of allelic imbalance.

Genotyping FFPE tissue can offer many possibilities in the near future, because vast sources of samples will be available for large retrospective tumor genetics studies. In addition, familial cases, either normal or tumor, from which often only FFPE tissue is available, can be genotyped. Future applications will include the use of the intensity values

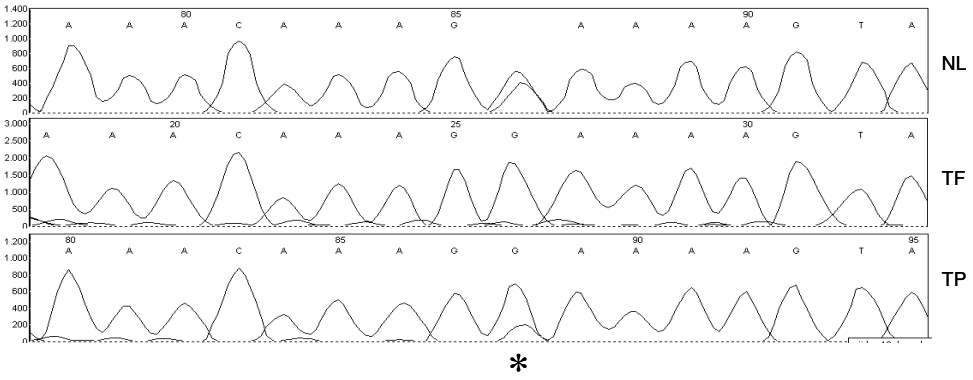

Figure 2. Sequencing result for SNP rs 2899251 on chromosome 22. Results are shown for sample 514 normal (NL), tumor frozen (TF) and tumor FFPE (TP) tissue. *, lost allele in both tumor tissues: CG in the normal becomes G in the tumor.

of the SNP callings. Upon proper normalization, these intensity values could provide quantitative measurements to each SNP, enabling the detection of genome wide gains or losses in FFPE tumors that underlie LOH patterns. This method was already applied to other platforms on frozen tissue and accompanying software tools were developed (4, 17).

From this study we conclude that FFPE tissue can reliably be genotyped and genome-wide LOH profiles can be determined using the BeadArray. This enables the high-throughput genome wide genotyping of archived FFPE tissue blocks not only for LOH analysis, but also for linkage studies and other applications.

Table 2. LOH regions observed in four colorectal tumors.

Sample ID	LOH
44	4, **5q**, 14q, **17p**, **18**, **20**, 21q
106	**5q**, 8p, 9, 12, 13q, 15q, **17p**, **18**, **20p**
108	3q, **5,** 7q, 14q, 15q, **17p**, **18**, **20p**, 22
514	1p, 4q, 12q, **17p**, **18q**, **20p**, 22q

Note: Chromosomes in boldface are regions that are detected in three or more cases.

Acknowledgments

We thank Illumina, for providing the BeadArrays and linkage panel. The Academic Biomedical Centre, Utrecht University, supported the genotyping facility.

References

1. Bignell GR, Huang J, Greshock J, et al.: High-resolution analysis of DNA copy number using oligonucleotide microarrays. Genome Res 14:287-295, 2004
2. Huang J, Wei W, Zhang J, et al.: Whole genome DNA copy number changes identified by high density oligonucleotide arrays. Hum Genomics 1:287-299, 2004
3. Rauch A, Ruschendorf F, Huang J, et al.: Molecular karyotyping using an SNP array for genomewide genotyping. J Med Genet 41:916-922, 2004
4. Zhao X, Li C, Paez JG, et al.: An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 64:3060-3071, 2004
5. Koed K, Wiuf C, Christensen LL, et al.: High-density single nucleotide polymorphism array defines novel stage and location-dependent allelic imbalances in human bladder tumors. Cancer Res 65:34-45, 2005
6. Murray SS, Oliphant A, Shen R, et al.: A highly informative SNP linkage panel for human genetic studies. Nat Methods 1:113-117, 2004
7. Fan JB, Oliphant A, Shen R, et al.: Highly parallel SNP genotyping. Cold Spring Harb Symp Quant Biol 68:69-78, 2003
8. Miller SA, Dykes DD, Polesky HF: A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res 16:1215, 1988
9. De Jong AE, Van Puijenbroek M, Hendriks Y, et al.: Microsatellite instability, immunohistochemistry, and additional PMS2 staining in suspected hereditary nonpolyposis colorectal cancer. Clin Cancer Res 10:972-980, 2004
10. Matsuzaki H, Loi H, Dong S, et al.: Parallel genotyping of over 10,000 SNPs using a one-primer assay on a high-density oligonucleotide array. Genome Res 14:414-425, 2004
11. van Eijk R, Oosting J, Sieben N, et al.: Visualization of regional gene expression biases by microarray data sorting. Biotechniques 36:592-4, 596, 2004
12. Douglas EJ, Fiegler H, Rowan A, et al.: Array comparative genomic hybridization analysis of colorectal cancer cell lines and primary carcinomas. Cancer Res 64:4817-4825, 2004
13. Jass JR, Walsh MD, Barker M, et al.: Distinction between familial and sporadic forms of colorectal cancer showing DNA microsatellite instability. Eur J Cancer 38:858-866, 2002
14. Meijer GA, Hermsen MA, Baak JP, et al.: Progression from colorectal adenoma to carcinoma is associated with non-random chromosomal gains as detected by comparative genomic hybridisation. J Clin Pathol 51:901-909, 1998
15. Risio M, Casorzo L, Chiecchio L, et al.: Deletions of 17p are associated with transition from early to advanced colorectal cancer. Cancer Genet Cytogenet 147:44-49, 2003
16. Sugai T, Takahashi H, Habano W, et al.: Analysis of genetic alterations, classified according to their DNA ploidy pattern, in the progression of colorectal adenomas and early colorectal carcinomas. J Pathol 200:168-176, 2003
17. Lin M, Wei LJ, Sellers WR, et al.: dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. Bioinformatics 20:1233-1240, 2004