Allele-specific loss of heterozygosity in multiple colorectal adenomas: toward an integrated molecular cytogenetic map II

Xin Mao,*, Rifat A. Hamoudib, Ian C. Talbotc, Michael Baudisd

aSkin Tumour Unit, St. John’s Institute of Dermatology, St. Thomas’ Hospital, 4th Floor, South Wing, Block 7, Lambeth Palace Road, London SE1 7EH, UK
bCancer Gene Cloning Centre, Haddow Laboratories, Institute of Cancer Research, 15 Cotswold Road, Sutton, Surrey SM2 5NG, UK
cAcademic Department of Pathology, St. Mark’s and Northwick Park NHS Trust, Watford Road, Harrow, Middlesex HA1 3UJ, UK
dDivision of Hematology/Oncology, Department of Pediatrics, University of Florida, SW 34th Street, Gainesville, FL 32610-0296

Received 5 July 2005; accepted 25 August 2005

Abstract

Colorectal cancer (CRC) remains a significant public health challenge despite our increased understanding of the genetic defects underlying the pathogenesis of this common disease. It has been thought that multiple mechanisms lead to the malignant phenotype, with familial predisposition syndromes accounting for only a small proportion of all CRC cases. To identify additional loci likely involved in CRC and to test the hypothesis of allele-specific loss of heterozygosity (LOH) for the localization of CRC susceptibility genes, we initially conducted a genome-wide allelotyping analysis of 48 adenomas from a patient with familial adenomatous polyposis coli (FAP) and 63 adenomas from 7 patients with sporadic CRC using 79 fluorescently tagged oligonucleotide primers amplifying microsatellite loci covering the human genome. Frequent allelic losses were identified at D17S802 (41%), D7S518 (40%), D18S53 (38%), D10S249 (32%), D2S391 (29%), D16S419 (27%), D15S1005 and D15S120 (24%), D9S274 and D11S1318 (23%), D14S65 (20%), D14S274 and D17S953 (19%), D19S424 (18%), D5S346 and D15S397 (15%), and D6S468 (13%) in multiple FAP adenomas. Common LOH was also detected at D4S1584 (42%), D11S968 (31%), D17S595 (28%), D5S394, D9S286 and D10S249 (24%), D8S511 (23%), D13S158 (21%), D7S669 (20%), D18S58 (19%), D2S162 and D16S432 (16%), D2S206 (15%), D7S496 and D17S946 (14%), D6S292 (13%), D4S1586 and D8S283 (11%), and D1S2766 (10%) in multiple CRC adenomas. In addition, allele-specific LOH at D5S346, D15S1005, and D15S120 was observed in multiple FAP adenomas (P < 0.01) and at D2S206 and D16S423 in multiple CRC (P < 0.05). To compare our data to previous reports, we determined the band-specific frequency of chromosomal imbalances in CRC karyotypes reported in the Mitelman database, and from the CGH results of cases accessible through the PROGENETIX website. Furthermore, published genome-wide allelotyping analysis of CRC and other allele-specific LOH studies were compiled and collated with our LOH data. The combined results not only provide a comprehensive view of genetic losses in CRC, indicating the comparability of these different techniques, but they also reveal different novel loci in multiple adenomas from FAP and sporadic CRC patients, suggesting that they represent a distinct subtype of CRC in terms of allelic losses. Allele-specific LOH is an alternative approach for cancer gene mapping. © 2006 Elsevier Inc. All rights reserved.

1. Introduction

Colorectal cancer (CRC) represents a group of heterogeneous epithelial malignancies, of which familial adenomatous polyposis coli (FAP) and hereditary nonpolyposis colon cancer (HNPCC) are two major CRC predisposition syndromes [1–3]. FAP usually presents in the second decade and is characterized by large numbers of adenomatous polyps (usually more than 100), carpeting the large bowel. Malignant change usually takes place in one or more polyps by the age of 50 years [4–7]. Almost all FAP cases result from truncating mutations in the APC gene [5–10]. In contrast, HNPCC patients have a normal or only slightly elevated tendency to develop adenomas, but the probability and rate of progression to carcinoma is increased [11–13] and an increased risk of other carcinomas is also
a recognizable feature of this syndrome [14]. Germline defects in one or more of a group of DNA mismatch repair genes, including MLH1, MSH2, MLH3, PMS1, PMS2, MSH6, and GTBP, are associated with HNPCC [15–21]. Approximately 15% of CRC are caused by dominantly inherited predisposition to the disease [22,23]. Only 2–6% of cases have been attributed to FAP or HNPCC [8,9,12], however, suggesting the presence of additional predisposition genes [24]. A proportion of this residual risk may be due to primary predisposition to colorectal adenomas, which subsequently progress to carcinoma. Previous epidemiological studies have shown that the relatives of CRC patients have a two- to threefold risk of developing adenomas [25–27], and relatives of probands with adenomas are at a twofold risk of developing CRC [28]. Furthermore, it has been suggested that predisposition to colorectal adenoma is common in the general population and that colorectal adenomas and carcinomas may occur predominantly in susceptible individuals [22].

Since FAP and HNPCC patients usually develop multiple independent adenomas, the presence of colorectal adenomas may offer additional support for the localization of susceptibility genes by genetic linkage analysis in families with multiple affected cases. If the underlying susceptibility gene were a tumor suppressor gene, there should be loss of heterozygosity (LOH) occurring in a substantial proportion of tumors within the vicinity of the gene, as previously shown for APC [29]. With multiple tumors occurring in the same individual, each tumor should lose the same allele inherited from the non–mutation-carrying parent. This type of allelic losses has been described as allele-specific LOH [30]. Other LOH events not related to a susceptibility gene locus, which might even occur at high frequency, would usually not be allele specific. The utilization of allele-specific LOH in individuals with multiple tumors may obviate the requirement for ascertainment of multiple cases from the same family and, hence, is applicable to susceptibility syndromes with low or variable penetrance. In addition, since allelic losses in tumors often span large chromosomal distances, the marker map used in an allele-specific LOH search for a susceptibility gene could be less dense than the 10–20 cM usually employed in conventional linkage analysis. Allele-specific LOH analysis has previously been used to identify the clonal origin and progression of several types of tumors [30–47]. To our knowledge, however, it has not been tested for cancer susceptibility gene identification.

As a model for both multistep and multipathway carcinogenesis [48,49] CRC provides paradigms of alterations of tumor suppressor genes (TSG) and oncogenes in malignant transformation [50]. These genetic changes can be detected by different techniques, including conventional cytogenetics, metaphase or array-based comparative genomic hybridization (CGH), and allelotyping. Previous cytogenetic studies have revealed chromosomal abnormalities in 30–80% of CRC, including deletions of 1p, 3p, 5p, 10p, and 17p, as well as loss of 18 [51–54]. CGH studies have shown DNA copy number losses of 5q, 10q, 11q, 17p, and 18q [55]. Molecular studies have demonstrated frequent allelic losses at 1p, 5q, 7q, and 15q [56–58]. However, none of these have put the cytogenetic data [banding, CGH, multiplex fluorescent in situ hybridization (M-FISH)] and molecular data (LOH, genomic and expression microarrays) of a specific tumor type together to create a user-friendly map in a single setting. We have constructed the integrated molecular cytogenetic maps for Sézary syndrome and breast cancer via this approach, facilitating the direct assessment of genetic alterations at chromosomal and molecular levels [59,60]. This provides a basis for the comparison between different techniques to create integrated molecular cytogenetic maps for different tumors.

To identify additional loci likely to be associated with the pathogenesis of CRC through the assessment of allele-specific LOH, and to construct the integrated molecular cytogenetic map for CRC, we initially conducted a genome-wide allelotyping analysis of 48 adenomas from 1 FAP patient and 63 adenomas from 7 patients with sporadic CRC using 79 fluorescently tagged oligonucleotide primers amplifying microsatellite (MS) loci covering the human genome. We then combined our LOH data with published cytogenetic, CGH, and allelotyping data of CRC by using dedicated karyotype parsing softwares and conventional literature searches.

2. Materials and methods

2.1. Allelotyping

2.1.1. Specimens and DNA extraction

Two sets of samples were collected for genome-wide allelotyping analysis. The first one included 48 adenomas and a normal control tissue (appendix) that were microdissected from formalin-fixed, paraffin-embedded tissue sections from a 45-year-old male patient with FAP who had 5,409 discrete adenomas including tubular, villous, and tubulovillous adenomas. The second one consisted of 63 adenomas and 7 normal control tissues that were dissected from paraffin sections from 7 patients with sporadic CRC (5 male and 2 female, 39–80 years old, average adenoma number <10). Dissected tissue samples were incubated in 10 mmol/L Tris hydrogen chloride (pH 7.5), 1 mmol/L ethylenediaminetetraacetic acid, 1% (wt/vol) sodium dodecyl sulfate, and 500 μg/mL proteinase K at 37°C for 72 hours. The mixture was then heated at 100°C for 10 minutes and directly used for polymerase chain reaction (PCR) amplification without further purification.

2.1.2. Primers, PCR, data, and statistic analysis

Our previous study showed that DNA extracted from paraffin sections could not reliably yield products greater than 200 base pairs (bp) upon PCR [61], hence MS markers
with product sizes less than 200 bp were selected for this study. A total of 79 fluorescently labeled oligonucleotide primers amplifying MS loci covering the human genome were used (Table 1). Three PCR methods were tested. First, dissected DNA extracts were directly amplified with the standard PCR protocol, and then a modified double PCR procedure was used. In the first round of nested PCR, multiple primers (one primer/locus, 0.3 μL of each primer) and 1 μL of DNA were added into 15 μL of reaction mixture consisting of 1.5 μL of 10× PCR buffer, 1.5 mmol/L MgCl₂, 1.5 μL (2 mmol/L each nucleotide) of dNTPs, 0.15 μL (10 mg/mL) of bovine serum albumin (BSA), 0.1 μL of recombinant Thermus thermophilus DNA polymerase XL (Perkin Elmer, Branchburg, NJ), and 9.25 μL of water. PCR conditions consisted of 40 cycles of denaturation at 94°C for 1 minute, annealing at the appropriate temperature (50–60°C) for 1 minute, and extension at 72°C for 1 minute. In the second round of PCR, 1 μL of the nested PCR product and primer pair (0.3 μL of each primer) were added into the same reaction mixture and amplified under the same conditions described above. Finally, DNA samples were amplified with the degenerated oligonucleotide primer (DOP) PCR method. This was conducted in 15 μL reaction mix (10.5 μL of distilled H₂O, 1.5 μL of 1× DOP PCR buffer, 1.5 μL of 0.005% (vol/vol) dNTP, 1 μL of 1.5 mmol/L MgCl₂, 1 μL of DNA extract, 0.25 μL of BSA, 0.15 μL of DOP (5’-OH CCGACCTCGAGNNNNNNATGTGGG OH-3’), and 0.1 μL of Taq DNA polymerase). The DOP-PCR conditions were as follows: 5 cycles of 94°C for 0.5 minute, 30°C for 1.30 minutes, and 30–72°C for 3 minutes, and 35 cycles of 94°C for 1 minute, 62°C for 1 minute, and 72°C for 2 minutes, as recommended by the supplier (Boehringer Ingelheim GmbH, Ingelheim, Germany). The DOP-PCR product was further amplified using the same method as the second round of PCR described above.

The PCR products were analyzed on a 29:1 (acrylamide/bis) 4.5% polyacrylamide denaturing gel premix (National Diagnostics, Hull, UK) in 1× TBE buffer using ABI 377 automated fluorescent DNA sequencer (Applied Biosystems, Foster City, CA). Two microliters of each PCR reaction were combined with 2 μL blue dye with formamide and 0.5 μL of a TAMRA fluorescent size marker (Applied Biosystems). This mix was denatured for 10 minutes at 94°C, after which 1.5 μL was loaded into each well on a prewarmed gel on a 36-cm Well-to-Read plate. The gel was run for 2.5 hours at 200 watts power, 60 amps current, 2,900 volts voltage, scan rate of 2,400 scans/hour, and 50°C temperature. While the samples were undergoing electrophoresis, fluorescence was detected in the laser-scanning region using filter set C and data were collected and stored using the GeneScan Collection Software 2.0 (Applied Biosystems). The fluorescent gel data collected during the run were automatically analyzed by GeneScan Analysis software (version 2.0.2; Applied Biosystems) at the end of the run. Each fluorescent peak was quantitated in terms of peak height and peak area. The results were then imported into Genotyper (version 1.1.1; Applied Biosystems) for further analysis.

The comparison of the ratios between tumors and their controls was made using the following two formulas for calculation: (1) \( T1/T2/N1:N2 \) and (2) \( T2/T1/N2:N1 \). In these formulas, \( T1 \) and \( N1 \) are the peak height of the smaller allele, and \( T2 \) and \( N2 \) are the peak height of the larger allele. Formula 1 was used to calculate the ratio of the smaller allele, while formula 2 was used to calculate the ratio of the larger allele. For ratios greater than 1, the reciprocal of the ratio is calculated to give a value between 0.00 and 1.00. A value of 0.25 or less was assigned as indicative of LOH [60–63].

In this study, allele-specific LOH is determined as consistent loss of one allele in more than two tumor samples from the same patients, as suggested previously [30]. To exclude the probability of loss of the same allele of each polymorphism at a specific locus in multiple tumor samples occurring as a chance event, the probability equation \( P = 1/2^n \) was used for the statistic analysis. In this formula, \( P \) represents the probability of an event taking place by chance, and \( n \) stands for the number of tumors with loss of the same allele of each polymorphism at a specific locus.

To further exclude the possibility of field effect of uninvolved tissues on the determination of LOH, multiple normal samples from different sites, including colorectal tissue of the same individuals, were tested with the MS markers described above. No LOH or abnormal band shifts (microsatellite instability) were detected, indicating that the field effect is insignificant in this study. Despite the standard PCR procedure failing to yield enough PCR products for analysis, both double PCR and DOP-PCR methods gave rise to ideal products, with consistent results.

### 2.2. Literature survey and analysis

#### 2.2.1. Cytogenetics

A total of 861 adenomatous neoplasias of the colon with available cytogenetic data were identified in a survey of the Mitelman Database of Chromosome Aberrations in Cancer (http://cgap.nci.nih.gov/Chromosomes/Mitelman) and the PROGENETIX database (http://www.progenetix.net). For the final data set, only malignant cases were included, which consisted of 659 cases reported in 53 original publications, including 603 cases of primary tumors originating in the large intestine and 56 cases of adenocarcinomas of the rectum.

International System for Human Cytogenetic Nomenclature (ISCN 1995) karyotype annotations collected in their respective databases were converted to band-specific aberration status information using dedicated parsing algorithms developed for the PROGENETIX project [64]. Briefly, ISCN annotations were split into their information atoms describing events involving one or several chromosomes. The information atoms were then analyzed for the
Table 1
A summary of genome-wide allelotyping analysis of multiple FAP and sporadic CRC adenomas

<table>
<thead>
<tr>
<th>Loci</th>
<th>Genetic (cM)</th>
<th>Chromosome bond</th>
<th>FAP</th>
<th>Sporadic</th>
<th>Case with LOH (%)</th>
<th>FAP</th>
<th>Sporadic</th>
<th>Case with LOH (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1S508</td>
<td>18.1</td>
<td>1p36.21p36.31</td>
<td>48 (100)</td>
<td>31 (49)</td>
<td>2 (4)</td>
<td>2 (6)</td>
<td>D1S249</td>
<td>0</td>
</tr>
<tr>
<td>D1S233</td>
<td>62.3</td>
<td>1p36</td>
<td>n/a</td>
<td>61 (97)</td>
<td>n/a</td>
<td>4 (7)</td>
<td>D10S1647</td>
<td>91.4</td>
</tr>
<tr>
<td>D1S551</td>
<td>97.8</td>
<td>1p22</td>
<td>0 (0)</td>
<td>11 (17)</td>
<td>0 (0)</td>
<td>4 (36)</td>
<td>D10S219</td>
<td>105.1</td>
</tr>
<tr>
<td>D1S2766</td>
<td>100.5</td>
<td>1p22</td>
<td>48 (100)</td>
<td>63 (100)</td>
<td>0 (0)</td>
<td>6 (10)</td>
<td>D10S574</td>
<td>124.4</td>
</tr>
<tr>
<td>BCL10</td>
<td>100.5</td>
<td>1p22</td>
<td>0 (0)</td>
<td>n/a</td>
<td>0 (0)</td>
<td>n/a</td>
<td>D1S187</td>
<td>143.9</td>
</tr>
<tr>
<td>D1S1588</td>
<td>104.6</td>
<td>1q25</td>
<td>45 (94)</td>
<td>42 (67)</td>
<td>13 (29)</td>
<td>5 (12)</td>
<td>D1S269</td>
<td>58.3</td>
</tr>
<tr>
<td>D1S435</td>
<td>128.9</td>
<td>2q24</td>
<td>n/a</td>
<td>55 (87)</td>
<td>n/a</td>
<td>9 (16)</td>
<td>D1S158</td>
<td>86.9</td>
</tr>
<tr>
<td>D2S391</td>
<td>73.8</td>
<td>2p21</td>
<td>48 (100)</td>
<td>63 (100)</td>
<td>7 (15)</td>
<td>2 (3)</td>
<td>D1S2766</td>
<td>100.5</td>
</tr>
<tr>
<td>D2S138</td>
<td>191.8</td>
<td>2p24</td>
<td>0 (0)</td>
<td>41 (65)</td>
<td>0 (0)</td>
<td>5 (12)</td>
<td>D1S728</td>
<td>53.8</td>
</tr>
<tr>
<td>D2S206</td>
<td>248.6</td>
<td>2q33</td>
<td>n/a</td>
<td>47 (75)</td>
<td>n/a</td>
<td>7 (15)</td>
<td>D1S285</td>
<td>112.8</td>
</tr>
<tr>
<td>TGFBIIR</td>
<td>11</td>
<td>3q26</td>
<td>40 (80)</td>
<td>32 (51)</td>
<td>0 (0)</td>
<td>5 (16)</td>
<td>D1S724</td>
<td>53.8</td>
</tr>
</tbody>
</table>

1 n/a: not available.
occurrence and status (gain and loss) of each chromosomal band (862 bands resolution). For the transformation of the banding data, only karyotypes of the main clones and their subclones were evaluated. For cases analyzed by metaphase banding, the relative status of each chromosomal band was generated from the sum of all gains and losses involving the band. A modified online version of the main software (ISCN2matrix converter) is accessible through the PROGENETIX project’s website (http://www.progenetix.net).

For comparison of cytogenetic results and locus-specific LOH data, the maximum loss percentage on a given chromosomal arm was used.

2.2.2. LOH

Extensive literature search for genome-wide allelotyping analysis of CRC and allele-specific LOH in malignancies was conducted. Seven reports were found [65–71], three of which contained data sufficient for compilation [65–67] and were then compared with the results of this study to assess the degree of consistency among these studies. In addition, 18 studies on allele-specific LOH were compiled and compared with this study (Table 2).

3. Results

3.1. Allelotyping

As described above, there were 5,409 adenomas in the FAP patient, of which 48 were tested by genome-wide allelotyping (0.89%). This revealed allelic losses on most chromosomes. Frequent LOH were seen in a descending order of D17S942 (41%), D7S518 (40%), D18S58 (38%), D10S249 (32%), D2S391 (29%), D16S419 (27%), D15S1005 and D15S120 (24%), D9S274 and D11S1318 (23%), D14S65 (20%), D14S274 and D17S953 (19%), D19S424 (18%), D5S346 and D15S124 (17%), and D6S468 (13%) (Table 1). Of the 48 adenomas analyzed, 46 had allelic losses (96%), ranging from 0–8 per samples, with mean LOH of 3.56 and standard deviation of 2.09 (Fig. 1; Table 1). Fifty-one MS markers were used to examine multiple FAP adenomas, of which 35 were informative (69%), with 25 showing LOH (71%).

There were four MS markers (16%), D15S1005, D15S120, D5S346 (APC), and D16S419, demonstrating specifically loss of allele 2, namely allele-specific LOH, although the frequency of loss at these loci was less than that of D7S518, D17S942, D18S53, and D10S249 (Fig. 1). Seven adenomas had LOH of allele 2 at the D5S346 (APC) locus, the P value of which was 1/27. The probability of this event occurring by chance was 1 out of 128 (P < 0.01), which was negligible. The same is true for D16S419 because there were seven adenomas with loss of allele 2 at this locus. Again, the P value for D15S120 was 1/2 or 1:512 (P < 0.01), and for D15S1005 it was 1/211 or 1:2048 (P < 0.001). Thus, it is unlikely that loss of allele 2 at these loci was randomly distributed in the genome of multiple FAP adenomas.

A total of 166 allelic losses were identified in 63 multiple adenomas from 7 CRC patients, with the mean LOH of 2.36 per sample ranging from 1.3 to 4 and a standard deviation of 1.77, which affected almost all chromosomes in the human genome (Table 1). Frequent allelic losses present in more than 5 adenomas were detected at 19 loci. This included D4S1584 (42%), D11S968 (31%), D17S953 (28%), D5S346, D9S286 and D10S249 (24%), D8S511 (23%), D13S158 (21%), D7S669 (20%), D18S58 (19%), D2S162 and D16S432 (16%), D2S206 (15%), D7S496 and D17S946 (14%), D6S292 (13%), D4S1586 and

Table 2
A summary of allele-specific LOH identified in different types of tumours

<table>
<thead>
<tr>
<th>Tumour type</th>
<th>Loci showing allele-specific LOH</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple CRC from FAP patient</td>
<td>E5.55, L5.71, E5.57 (franking APC)</td>
<td>Miki et al 1992</td>
</tr>
<tr>
<td>Mice multiple intestinal neoplasia</td>
<td>apc+</td>
<td>Levy et al 1994</td>
</tr>
<tr>
<td>Non-small cell lung cancer</td>
<td>IFNA, D9S171</td>
<td>Kishimoto et al 1995</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>D3S1228, D3S1029, D3S1038</td>
<td>Hung et al 1995</td>
</tr>
<tr>
<td>Endemic gallbladder carcinoma</td>
<td>D9S171, TP53</td>
<td>Wistuba et al 1995</td>
</tr>
<tr>
<td>Cylindromas</td>
<td>D16S419, D16S408 (franking CYLD1)</td>
<td>Biggs et al 1995 and 1996</td>
</tr>
<tr>
<td>Peutz-Jegher’s syndrome</td>
<td>D19S886, D19S883 (franking LKB1)</td>
<td>Hemminki et al 1997</td>
</tr>
<tr>
<td>Cervical carcinoma</td>
<td>D3S1597, D3S1244, D3S2432, D4S1565</td>
<td>Larson et al 1997</td>
</tr>
<tr>
<td>Cylindromas</td>
<td>CYLD1</td>
<td>Verhoef et al 1998</td>
</tr>
<tr>
<td>Rat brain tumour</td>
<td>BDIV allele</td>
<td>Kindler-Rohrborn et al 1999</td>
</tr>
<tr>
<td>Cylindromas</td>
<td>CYLD1</td>
<td>Thomson et al 1999</td>
</tr>
<tr>
<td>Squamous cell lung carcinoma</td>
<td>D3S1447, D8S277, D8S1130, D8S1106, D8S602, D8S254, D8S261, LPL-GZ, D8S136, D9S1748, D11S1391</td>
<td>Wistuba et al 1999a and b</td>
</tr>
<tr>
<td>Cylindromas</td>
<td>CYLD1</td>
<td>Takahashi et al 2000</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>DAL-1/4.1B (EPB41L3)</td>
<td>Kintiymya et al 2004</td>
</tr>
<tr>
<td>Breast cancer</td>
<td>LKB1/STK11</td>
<td>Nakamishi et al 2004</td>
</tr>
<tr>
<td>Melanoma</td>
<td>HLA-B</td>
<td>Rodriguez et al 2004</td>
</tr>
<tr>
<td>Multiple FAP adenomas</td>
<td>D5S346 (APC), D15S1005, D15S120, D16S419</td>
<td>This study</td>
</tr>
<tr>
<td>Multiple sporadic CRC adenomas</td>
<td>D2S206, D16S423</td>
<td>This study</td>
</tr>
</tbody>
</table>

[30–47]
D8S283 (11%), and D1S2766 (10%) (Fig. 2; Table 1). Less common LOH (>4 samples) was also observed at D18S452 (26%), TGFBIIR (16%), D6S468 and D17S 802 (15%), as well as D2S391 and D2S138 (12%) (Table 1).

There were seven adenomas showing allelic loss at D2S206, with a P value of 1.128 (P < 0.01; Fig. 2). Five adenomas demonstrated allele-specific LOH at D16S432, with a P value of 1.32 (P < 0.05; Fig. 2), which was slightly higher than the average of P values of the above loci in multiple FAP adenomas. This might be due to the fact that this MS locus was non-informative in two patients (Fig. 2).

Compared with multiple FAP adenomas, the mean LOH of multiple sporadic CRC adenomas was slightly lower. Allelic losses at D2S391, D10S249, D17S953, and D17S802, however, were consistently present in multiple adenomas from both FAP and sporadic CRC patients with a normal allele at the locus of DCC (Table 1). In addition, discordance was clearly seen between multiple FAP and sporadic CRC patients because there were six MS loci (D7S518, D9S274, D11S1318, D14S65, D15S120, and D16S419) showing frequent LOH in multiple FAP adenomas but a normal allele in multiple sporadic CRC adenomas (Table 1). The opposite is true for multiple sporadic CRC, as six other loci (D1S2766, D2S138, TGFBIIR, D8S511, D13S158 and D17S946) revealed frequent allelic losses with a normal allele in multiple FAP adenomas (Table 1). Moreover, allele-specific LOH detected in multiple FAP adenomas was absent in multiple sporadic CRC and vice versa (Figs. 1 and 2).

3.2. Analysis of cytogenetic literature in CRC

Since the first report of banding analysis in colonic polyps by Mitelman et al. [72] more than 30 years ago, there have been a large number of publications describing karyotypes of adenomatous lesions of the colon and rectum (http://www.ncbi.nih.gov/pubmed). In recent years, the molecular cytogenetic screening technique of comparative genomic hybridization (CGH) has been applied to the analysis of the virtually full spectrum of human neoplasias, including CRC (http://www.progenetix.net).

For this study, we selected 344 malignant adenocarcinomas of colon and rectum (CRC), as reported in the Mitelman database (http://www.cgap.nci.nih.gov/Chromosomes/Mitelman). We included an additional 315 CRC cases collected from the PROGENETIX database (http://www.progenetix.net). Premalignant or benign lesions were excluded from the analysis.

For detecting imbalanced regions along the genome, ISCN karyotypes were processed using the ISCN2 matrix.
As previously discussed [60], for the cases analyzed by metaphase banding, parsing of the karyotypes using a high-filter stringency (only completely annotated cases, no unresolved marker chromosomes or questionable bands) resulted in loss in the majority of cases (135 of 344 cases remaining) and selected for cases with low karyotype complexity. In the relaxed analysis method chosen thereof (all parsable bands, acceptance of ‘?’ marked annotations), a richer aberration pattern with clear delineation of hot spot regions could be observed (Fig. 5). In contrast, cases analyzed by CGH are unambiguous in the reported ‘‘rev ish’’ ISCN annotations by virtue of the technique. In total, 659 cases from 53 publications were assessed, of which 19.2% of chromosomal bands were not in a balanced state, involving on average 7 chromosomes per case. Gains and losses were roughly evenly distributed regarding their frequency, but with strong disposition toward different chromosomal regions. Fig. 3 shows the summary of chromosomal gains and losses as a percentage of affected cases. Most frequent losses (>10% of cases) with discernable peaks involved regions 18q21q22 (41.4%), 8p22 (32.2%), 17p12 (28.8%), and 14q24q31 (14.6%). Diffuse losses could be found for chromosome 4 (up to 20.2%), 1p with a maximum at 1p36 (16.5%), as well as chromosomes 15 (up to 15.5%), 22 (up to 14.4%), 5 (up to 14.3%), and 10 (up to 11.5%).

3.3. Survey of LOH studies in CRC

Since first reported by Vogelstein et al. [65] more than 15 years ago, there have been only 7 studies describing genome-wide LOH in 425 colorectal cancers [65–71]. However, LOH analysis has been widely and selectively...
applied to a limited number of subchromosomal regions [73–99]. Of the seven reports, three contained data sufficient for compilation, which were then selected for comparison with this study [65–67]. In these studies, a general concordance can be found for frequent allelic losses (mean >20%) at 17p (28%), 18p (26%), 18q and 5q (25%), 9p and 15q (24%), 8p (23%), and 14q (20%) (Figs. 4 and 5). This pattern of LOH has also been described in two other studies [68,71]. All of these allelic losses lie within regions of chromosomal losses detected by G-banded karyotyping and CGH assay (Figs. 3–5).

Discrepancy was also noted, however, as the most common LOH regions in HNPCC were at 11p and 11q [67] (Fig. 4). In contrast, Ezaki et al. [69] observed highly frequent LOH at 6q in ulcerative colitis-associated CRC, and Laiho et al. [70] described a twofold difference of allelic losses at 20q between Finnish familial and sporadic CRC cases. In contrast, both banding analysis and CGH studies revealed low-level chromosomal losses at these LOH regions (Figs. 3–5). A median LOH was also generated (Fig. 5), which was based on part of this and three previous studies [65–67] and revealed similar patterns of allelic losses as described in Fig. 4. Intriguingly, discrepancy between median LOH and cytogenetic data was seen for 5q at the APC locus, 9p at the INK4a locus, and at 11q and 13q, in which more allelic losses were detected by allelotyping, whereas for chromosomes 4, 8p, and 18q, CGH was the more sensitive method (Fig. 5).

In this study, there were five common regions of allelic losses detected in multiple FAP adenomas [i.e., 17p (41%), 18p (38%), 10p (32%), 2p (29%), and 16q (27%)]. Six different LOH regions were found in sporadic CRC adenomas, including 4q (42%), 1p (36%), 11q (31%), 17p (28%), 18p (26%), and 1q (25%) (Table 1). Overall, the pattern of allelic losses present in multiple sporadic CRC was consistent with those described in previous allelotyping, banding, and CGH studies (Fig. 4).

Allele-specific LOH was first described at loci flanking the APC gene in multiple CRC from FAP patients by Miki et al. [30] more than 12 years ago. Since then, this genetic phenomenon has also been observed at other loci in different types of tumors (Table 2). In this study, allele-specific LOH at the APC locus (DSS346) was initially confirmed in multiple adenomas from the FAP patient and subsequently seen at D15S1005, D15S120, and D16S419 (Fig. 1; Table 2). In addition, allele-specific LOH at
D2S206 and D16S423 was identified in multiple sporadic CRC (Fig. 2; Table 2). Overall, only a small proportion of loci showed allele-specific LOH (15% in multiple FAP adenoma and 4% in multiple sporadic CRC; Figs. 1 and 2).

4. Discussion

To identify additional CRC susceptibility loci and to test the hypothesis of allele-specific LOH for gene mapping, we performed genome-wide allelotyping analysis in multiple FAP and sporadic CRC adenomas, and compared the results with previous cytogenetic and LOH studies. Allelic losses at 1q, 2p, 5q, 6q, 7q, 9p, 10p, 11q, 15q, 16q, 17p, 17q, 18p, and 19p were detected in more than 10% of FAP adenomas. In addition, 1 in 10 sporadic CRC adenomas also had LOH at 1p, 2p, 4q, 5q, 6q, 7p, 7q, 8p, 9p, 10p, 11p, 13q, 16p, 17p, 17q, or 18p. Overall, these allelic losses are generally consistent with previous cytogenetic and molecular studies of CRC [51–58]. Discrepancies were present between FAP and sporadic CRC adenomas, however, suggesting the presence of different subtypes of CRC. Furthermore, there were four loci in FAP and two loci in sporadic CRC showing allele-specific LOH, one of which was the APC locus at 5q and the probability that this could occur by chance was negligible. This observation is in concordance with previous studies [30,31], validating the efficiency of allele-specific LOH for cancer susceptibility gene identification and pointing to novel CRC susceptibility loci at 2p, 15q, 16p, and 16q.

APC was identified more than 10 years ago by use of conventional cytogenetics and genetic linkage analysis [100–103]. Despite the advent of novel and powerful techniques of cancer genome research, such as genomic, expression and single nucleotide polymorphism (SNP) microarrays [104–107], chromosome-based analysis methods remain valid. However, there are inherent limitations for these two gene mapping methods. Cytogenetic metaphase analysis requires cultured tumor cells, which may be difficult to obtain, especially while avoiding in vitro selection. On the other hand, linkage analysis requires large numbers of families with multiple affected cases, which makes it difficult to apply on rare cancer types. Cancer susceptibility genes may be difficult to study by linkage analysis because of frequent phenocopies and failure to develop clinical symptoms despite the development of premalignant lesions or early malignancies. Moreover, analysis is further complicated if the predisposition is caused by multiple genes acting either additively or multiplicatively.

Since the multiplicity of tumors is one of the clinical features of cancer predisposition syndromes, we assume that individuals with multiple adenomas might carry a mutated allele of a colorectal adenoma predisposition gene. The utilization of allele-specific LOH in this context is predicted to be useful because individuals who have many tumors can generate substantial evidence in favor of
a susceptibility locus and thus reduce the problems of genetic heterogeneity, multiplicatively acting genes, sporadic occurrence, and lack of clinical symptoms. In addition, the rate of LOH at a susceptibility locus is usually much higher in tumors that arise due to a predisposing gene mutation than in sporadic neoplasms of the same type. Even if multiple genes are acting in a single individual and some tumors do not develop as a result of a particular susceptibility gene mutation or if one or more of several adenomas arise due to chance rather than due to that particular susceptibility gene, evidence against that susceptibility locus will arise due to chance rather than due to that particular susceptibility locus will be diluted because loss of the allele linked to the susceptibility gene mutation in the sporadic tumors is less likely to occur.

The phenomenon of allele-specific LOH has been reported in several studies (see the list of references in Table 2). In this study, there were seven FAP adenomas demonstrating LOH of allele 2 at this locus with a $P$ value of 1:128. In terms of the score of logarithmic odds (LOD) commonly used in conventional genetic linkage analysis (http://www.3-search.com/Score/lo%20score.php), the $P$ value was approximately equivalent to an LOD score of 2. The original report on linkage analysis of APC in FAP families revealed an LOD score of 3.28 [100], which was higher than the roughly estimated LOD score above. As discussed above, however, previous conventional linkage studies are based on allelotyping and statistic analysis of large numbers of families with multiple affected cases. In this study, multiple adenoma samples from just one FAP patient were tested, and the results were supported by previous studies [30,31]. Therefore, it is not only possible but also much simpler to use allele-specific LOH for mapping APC. This observation has been further consolidated with the findings of allele-specific LOH at $D15S1005$ and $D15S120$ on 15q and $D16S419$ on 16q in FAP adenomas, which showed much smaller $P$ values or likely higher LOD scores compared with that of APC. This is also true with $D2S206$ on 2q and $D16S423$ on 16p in sporadic CRC adenomas, although there was a slightly large $P$ value due to the non-informative of $D2S206$ and $D16S423$ in some patients. Thus, it is likely that these MS loci might contain novel CRC suppressor or modifier genes. A previous study has described the colorectal adenoma and carcinoma susceptibility locus (CRAC1) lying at 15q14q22 [58]. CRAC1, however, is unlikely to be a candidate in this context because it has a recombination distance of 30 cM to $D15S1005$ and 50cM to $D15S120$, and has also shown no significant association with LOH in early-onset sporadic CRC [99]. On the other hand, in the compilation of 315 CRC cases analyzed by the PROGENETIX CGH database (http://www.progenetix.net), copy number losses were detected at 15q23q24 ($D15S1005$) in 18.4%, 2q33q37 ($D2S206$) in 10.5%, 16p13.3 ($D16S423$) in 5.7%, 16q12.2q22.1 ($D16S419$) in 5.1%, and 15q26 ($D15S120$) in 14.3%. These chromosomal regions contain a variety of genes such as $TSC2$, $PKM2$, $STAT1$, $DECR2$, $RBL2$, and $AGC1$ (http://www.ensembl.org/), and reduced expression of $DECR2$ and $RBL2$ has also been noted in CRC by Affymetrix expression microarray (http://www.genome.ucsc.edu/index.html?org=Human&db=hg17&hsid=39289999). Further study is therefore required to establish if these genes are associated with the susceptibility of CRC.

In this study, non-allele-specific LOH was more commonly seen in both FAP and sporadic CRC adenomas.
This may be explained as a consequence of multiple-locus chromosomal events such as deletion, nondisjunctional chromosome loss with or without reduplication, or a locus-restricted event such as gene conversion or point mutation [108]. If the tested loci were distal to a tumor suppressor gene, however, non–allele-specific LOH is likely to emerge due to the outcome of mitotic recombination between the loci of tumor suppressor genes and tested genetic markers.

There are several explanations for the presence of different patterns of LOH in FAP and sporadic CRC adenomas. First, it may be due to the heterogeneity of the tumor samples, because all FAP adenomas came from one patient while all sporadic CRC adenomas were from seven patients. Second, it may reflect the multiclonal origin of adenomas of CRC, as suggested previously [109]. Third, it is also likely that FAP and sporadic CRC represent two different subtypes of CRC, which has been suggested by previous studies [67,68,70].

The second part of this study has been the construction of an integrated molecular cytogenetic map for CRC based on extensive literature searches with the assistance of the ISCN2matrix software. As stated above, because of its incidence, CRC is one of the genetically best-characterized malignancies, and a large number of published cytogenetic and molecular studies of CRC have been published. To date, however, those reports have not been combined to evaluate the consistency of results derived by a large number of observers using different techniques. We have attempted to draw integrated molecular cytogenetic maps for Sézary syndrome and breast cancers through this approach [59,60]. This has enabled us to directly and easily compare chromosomal aberrations in these tumors at cellular and molecular levels, and has provided a basis for the systematic comparison between different techniques to create integrated molecular cytogenetic maps for different cancers in terms of large sample size. As we had observed for breast cancer [60], results from metaphase banding and CGH showed concordance in genomic hot spot detection while differing in absolute values per locus to a certain degree. Interestingly, in this study, the compiled LOH analyses showed a higher sensitivity for the detection of losses on 5q, the map locus of the APC gene.

The concepts of LOH [110,111] and the two-hit mutation model [112] have facilitated TSG research. Despite the thousands of LOH studies in a variety of cancers, to our knowledge, no TSG has been identified by LOH analysis alone. This failure raises concern on the validity and efficiency of allelotyping for mapping TSG and as a cancer genetic research tool in general [113]. This study, however, has revealed not only the overall consistency of genetic losses detected by banding and CGH analysis and genome-wide allelotyping in CRC, but also the presence of allele-specific LOH in multiple CRC adenomas. This indicates that allelotyping remains a valid technique for the assessment of genetic losses in malignancies. In addition, the degree of similarity between banding and CGH data summary profiles should point to the ability of both techniques to correctly identify genomic imbalance hot spots when applied to a large number of cases, at least for CRC.

A recent shift toward new molecular analysis techniques can be observed (array-CGH, expression, and SNP microarrays), and indeed most future studies of genomic abnormalities in human malignancies may be based on array technologies. A recent study has compiled and analyzed all expression microarray data available at that point, revealing common and distinct gene expression patterns and clusters of signaling pathways in different types of cancers [114]. Thus, the combination of large-scale locus-based genomic aberration data, as presented here with gene expression compilations, could be a powerful tool for gaining further insights into the genetic pathways leading to cancer development. In addition, the application of SNP microarray for analysis of multiple tumor samples from individual patients would shed further light on the molecular basis or mechanisms underlying the occurrence of allele-specific LOH in multiple cancer syndromes.

Acknowledgment

This work was supported by the Peel Medical Research Trust.

References


