DNA copy number gain of FAK in wild-type KRAS colorectal cancer

Submitted for publication

Cindy Postma
Sandra Mongera
Rebecca PM Brosens
Danielle AM Heideman
Miguel A Cuesta
Alexander F Engel
Henk Verheul
Sabine Tejpar
Heike Grabsch
Philip Quirke
Bauke Ylstra
Beatriz Carvalho
Gerrit A Meijer
Abstract

Background: KRAS mutation status in colorectal cancer (CRC) is a powerful predictive marker for response to anti-EGFR monoclonal antibody therapy. However, response rates in wild-type KRAS CRC patients are still low, indicating that other factors, next to KRAS mutation status, influence the response to EGFR inhibitors. The aim of the present study was to investigate if particular DNA copy number alterations in wild-type microsatellite stable (MSS) KRAS CRC exist that could relate to the EGFR signaling pathway.

Material & Methods: DNA of 64 primary MSS CRCs was subjected to genome-wide array comparative genomic hybridization analysis and KRAS mutation analysis, and the DNA copy number profiles in CRC with or without KRAS mutations were compared. Gene dosage effect on mRNA expression was evaluated by means of expression array analysis. Candidate markers were further evaluated in an independent series of 35 MSS colorectal carcinomas.

Results: The most interesting region significantly different between CRC patients with or without mutated KRAS (P<0.03) was chromosomal region 8q23.3-q24.3, where MYC and FAK co-localize. DNA copy number gains of MYC and FAK were more frequent in the wild-type KRAS tumors in both the learning set (P=0.04 and P=0.01, respectively) and the validation set (P=0.02 and P=0.21, respectively). Of these two genes FAK showed a positive correlation between copy number gain and mRNA overexpression in both learning and validation set (both P=0.02).

Conclusion: Within the subgroup of wild-type KRAS CRC heterogeneity exists at the DNA copy number level. Especially copy number gain of FAK, which among other functions is an upstream regulator of the RAS signaling pathway, is likely to affect the response to anti-EGFR therapy.
Introduction

Colorectal cancer (CRC) is a leading cause of cancer death worldwide. New chemotherapeutic and targeted drugs have contributed to improved median overall survival of patients with metastatic CRC. One category of drugs targets the epidermal growth factor receptor (EGFR), which is a transmembrane tyrosine kinase receptor and a member of the ERBB family. EGFR can be activated by overexpression, ligand-dependent and ligand-independent mechanisms. EGFR activation results in activation of the RAS/RAF/MAPK, STAT and PI3K/AKT signaling pathways, which together control cell proliferation, angiogenesis, metastasis and survival [1]. Although anti-EGFR therapies, i.e. cetuximab and panitumumab, have improved clinical outcome for metastatic CRC, they are effective in only a subset of patients. Approximately 10% of patients achieve objective response [2-4]. Therefore, there is a clear need to identify biomarkers that can identify CRC patients who will likely respond to these EGFR-targeted therapies among the numerous patients who will not and require other therapeutic strategies. Several studies have indicated that the presence of mutant KRAS in CRC is associated with lack of response to EGFR inhibitors [5-11]. KRAS is a small G-protein and an essential downstream component of the EGFR signal transduction pathway. Mutation of the KRAS gene can yield a constitutively active protein that isolates the pathway from EGFR status, rendering EGFR inhibitors ineffective. KRAS mutations occur in approximately 40% of CRCs [5-12]. Therefore, KRAS mutation analysis is used for selecting CRC patients for anti-EGFR therapy. While mutation screening is powerful in predicting which CRC patients will not respond to these drugs, the response in wild-type KRAS CRC patients is still highly variable. Other factors, next to KRAS mutation, thus impact the response rates to anti-EGFR therapy. Therefore, additional predictive markers would be beneficial to select CRC patients with wild-type KRAS that will or will not respond to EGFR inhibitors.

Apart from wild-type KRAS status, recently also increased EGFR gene copy number, as detected by fluorescent or chromogenic in situ hybridization, and high mRNA expression of EGFR ligands, particularly amphiregulin (AREG) and epiregulin (EREG), have been associated with response to cetuximab and/or panitumumab therapy [5,10,11,13-15]. Biologically, CRC is a heterogeneous disease and this has influence on clinical behavior, including the response to therapy. This biological diversity occurs for a substantial part at the chromosomal level, giving rise to DNA copy number alterations [16]. Therefore, the aim of the present study was to determine 1) if colorectal carcinomas with mutated KRAS compared to wild-type KRAS show different DNA copy number profiles, 2) if within the category of wild-type KRAS carcinomas, heterogeneity exists at the level of DNA copy number changes and consequently 3) if gene dosage
affects the expression of genes in the EGFR signaling pathway. This approach may unravel putative predictors of response to EGFR inhibitors in wild-type KRAS CRC patients.

**Materials and Methods**

**Material**

The learning set consisted of 64 primary colorectal carcinomas, known to be MSS or MSI-L as determined by MSI Analysis System (Promega Corporation, Madison, USA) according to the manufacturer’s instructions. The microsatellite instable (MSI) colorectal tumors were excluded from this study, because these tumors are known to have little or no chromosomal alterations and often carry BRAF mutations, rendering this category less relevant for the purpose of the present study. The selected series consisted of 26 snap-frozen and 38 formalin-fixed and paraffin-embedded (FFPE) colorectal carcinomas. The whole series has been previously analyzed for chromosomal abnormalities by array-comparative genomic hybridization (aCGH) and in 22 also mRNA expression was determined by microarray analysis [17]. The validation set consisted of 35 MSS/MSI-L, FFPE colorectal carcinomas, collected from the tissue archive of the department of Pathology at the Zaans medical center (NL) and Leeds General Infirmary (UK) [18]. All samples were used in compliance with the respective institutional ethical regulations for surplus material and use of material from Leeds General Infirmary (UK) was approved by the Leeds (West) Research Ethics Committee, unique identifier CA 02/014.

**DNA and RNA isolation**

DNA from microdissected FFPE material, containing at least 70% tumor cells, was isolated using the QIAamp microkit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions and DNA and RNA from snap-frozen tissues were obtained as described before [17-18]. All DNA and RNA concentrations were measured on a Nanodrop ND-1000 spectrophotometer (Isogen, IJsselstein, NL).

**KRAS mutation analysis**

KRAS mutations located in exon 1 and 2 were screened. For the learning set a two step nested PCR approach was used as described before [19,20] and for the validation set high resolution melting (HRM) PCR based-technology was used as described before [21], both followed by direct
sequencing using the BigDyeTM Terminator version 3.1 Cycle Sequencing Kit (Applied Biosystems) and the 3130 Genetic Analyzer (Applied Biosystems). Sequences of KRAS exon 2 HRM M13-tagged forward primer, reverse primer and unlabelled probe with a 3’-conjugated C3 spacer were 5’-CCTTCTCAGGATTCTACAGGAAGCAAG-3’, 5’-AGTCCTCATGTACTGGTCCCTCA-3’ and 5’-CCTCTTGACCTGCTGTGTCGAGAATAT-3’, respectively.

In all analyses, genomic DNA extracted from HCT116 of H460 cells were used as positive control for mutations in exon 1 and 2, respectively, and genomic DNA from HT29 cells was used as wild-type KRAS control.

**aCGH**

For the learning set, a 5K bacterial artificial chromosome (BAC) array platform was used as described before [18,22]. DNA from CRCs of the validation set was hybridized on an Agilent 4x44K oligonucleotide array (Agilent Technologies, Palo Alto, USA). DNA labeling and hybridization was done as described before [18]. Local background, for both BAC and Agilent array, was subtracted from the signal median intensities of both tumor and reference DNA. The log2 tumor to reference ratio was calculated for each spot, in case of BAC array the median of the triplicate spots was calculated, and tumor/reference ratios were normalized against the median of the ratios of all autosomes.

**aCGH analysis**

Median absolute deviation (MAD) was determined for each case as a quality control. Samples from the learning set hybridized on BAC arrays with a MAD value ≥0.2 were excluded. For the validation set, MAD value <0.3 was accepted, because these 50 tumors were hybridized on another aCGH platform with a higher resolution. The mean MAD value of these 50 carcinomas was 0.16 (range 0.0-0.33), only 1 tumor was excluded (MAD=0.33).

R package CGHcall was used for data segmentation, defining copy number gains, high-level amplifications and losses, and converting log2 ratios into ordinal data, i.e. ‘+1’ for gains, ‘+2’ for high-level amplifications, ‘-1’ for losses, and ‘0’ if no DNA copy number alterations were present [23].

To reduce the dimensions of the aCGH data and gain statistical power we used the CGHregions algorithm. This algorithm combines consecutive (in terms of genomic order) probes on the array with similar copy number status to single events. A threshold for average error rate of 0.01 was used [24].
Expression array and analysis

For the learning set, the Human Release 2.0 oligonucleotide library, containing 60-mer oligonucleotides representing 28,830 unique genes, designed by Compugen (San Jose, CA, USA) was used and analyzed as described before [17].

For the validation set, Agilent single 44K formatted expression arrays (G4112A, Agilent Technologies, Paolo Alto, California, USA), containing 41,675 60-mer oligonucleotides representing over 27,000 well-characterized full-length or partial human genes and expressed sequence tag clusters, were used as described before [18].

Unsupervised hierarchical cluster analysis

To analyze the distribution of whole genome DNA copy number profiles in these colorectal tumors with and without KRAS mutations, unsupervised hierarchical cluster analysis was performed using Weighted Clustering of Called aCGH data (WECCA) with the parameter total linkage [25].

Supervised analysis

To determine clinico-pathological differences between CRCs with and without mutated KRAS, the Mann-Whitney U test for comparing means of continuous variables between the two groups and the Chi-square test for testing differences in distribution of categorical variables were applied. Boxplots and scatterplots representing different genes were created by taking the specific probe of that gene or in case there were several probes available, the mean of 2 or more genes was calculated. If no specific probe was available, then the mean of 2 flanking probes was taken. To calculate significance of differences in frequencies of DNA copy number changes between tumors with wild-type KRAS and mutated KRAS, CGH multiarray was used including a Wilcoxon test with ties, the P-values of which were corrected for multiple testing using a permutation version of false discovery rate (FDR). To determine differences in gene copy number of specific genes between CRCs with and without mutated KRAS the Mann-Whitney U test was applied. Correlation between log2 copy number changes and mRNA expression of the different genes were evaluated by linear regression analysis. SPSS version 15.0 statistical software package (SPSS Inc., Chicago, IL) was used. P-values less than 0.05 and FDR-values below 0.20 were considered to be significant.
Results

KRAS mutation status in relation to clinicopathological data
All CRCs, either snap-frozen or FFPE material, could successfully be analysed for KRAS mutation status. Forty-four percent (28/64) of the cases from the learning and 37% (13/35) of the cases from the validation set had mutated KRAS. The most common KRAS mutations in both learning and validation set were detected in exon 1, i.e. G12V, G12D and G13D mutations (35.7% versus 46.1%, 17.9% versus 15.4% and 21.4% versus 38.5%, respectively).

In both sets, patient and tumor characteristics were mostly balanced between the carcinomas with and without mutated KRAS. However, in the learning set, tumors with KRAS mutations were significantly more often located in the proximal part of the colon, while tumors with wild-type KRAS more in the distal colon (P<0.005) (Table 1).

Copy number profiles in relation to KRAS mutation status in the learning set
In the learning set, no significant differences in the number of chromosomal gains and losses were found between CRCs with and without mutated KRAS. The mean number of total alterations, gains and losses were 8.8 (range 1-25) versus 10.6 (range 3-23), 5.8 (range 1-11) versus 4.3 (range 1-15), and 4.8 (range 0-16) versus 4.6 (range 0-10) for the CRC with and without KRAS mutations (P = 0.12, P = 0.63 and P = 0.05, respectively). Figure 1 shows the frequencies of gains and losses according to chromosomal location in the CRCs with (Fig 1A) and without (Fig 1B) mutated KRAS.

Twenty-seven CRCs had high-level amplifications, which were mainly on chromosome regions 8q24.21, 10q22.3, 12p13.33-p13.32, 13q12.11-q12.3, 20q13.12-q13.13 and 20q13.31-q13.32. The mean number of high-level amplifications for CRC with mutated KRAS and wild-type KRAS was 0.54 (range 0-3) versus 0.69 (range 0-3), respectively (P = 0.40).
Table 1. Clinical and pathologic characteristics of colorectal tumors in relation to KRAS mutation status.

<table>
<thead>
<tr>
<th></th>
<th>All Mutant KRAS</th>
<th>Wild-Type KRAS</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Learning set:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcinomas-no.</td>
<td>64</td>
<td>28</td>
<td>36</td>
</tr>
<tr>
<td>Age-yr (Mean±SD)</td>
<td>68±9.5</td>
<td>69.9±9.0</td>
<td>66.5±9.8</td>
</tr>
<tr>
<td>Sex-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>28 (43.7%)</td>
<td>11 (39.3%)</td>
<td>17 (47.2%)</td>
</tr>
<tr>
<td>Female</td>
<td>35 (54.7%)</td>
<td>16 (57.1%)</td>
<td>19 (52.8%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (1.6%)</td>
<td>1 (3.6%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Site of tumor-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td>16 (25.0%)</td>
<td>12 (42.8%)</td>
<td>4 (11.1%)</td>
</tr>
<tr>
<td>Distal</td>
<td>47 (73.4%)</td>
<td>15 (53.6%)</td>
<td>32 (88.9%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (1.6%)</td>
<td>1 (3.6%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Tumor grade-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td>16 (25.0%)</td>
<td>7 (25.0%)</td>
<td>9 (25.0%)</td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>42 (65.6%)</td>
<td>18 (64.3%)</td>
<td>24 (66.7%)</td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>6 (9.4%)</td>
<td>3 (10.7%)</td>
<td>3 (8.3%)</td>
</tr>
<tr>
<td>Stage of disease-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>34 (53.1%)</td>
<td>14 (50.0%)</td>
<td>20 (55.6%)</td>
</tr>
<tr>
<td>II</td>
<td>12 (18.7%)</td>
<td>6 (21.4%)</td>
<td>6 (16.7%)</td>
</tr>
<tr>
<td>III</td>
<td>14 (21.9%)</td>
<td>6 (21.4%)</td>
<td>8 (22.2%)</td>
</tr>
<tr>
<td>IV</td>
<td>3 (4.7%)</td>
<td>1 (3.6%)</td>
<td>2 (5.5%)</td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (1.6%)</td>
<td>1 (3.6%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td><strong>Validation set:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carcinomas-no.</td>
<td>35</td>
<td>13</td>
<td>22</td>
</tr>
<tr>
<td>Age-yr (Mean±SD)</td>
<td>73.7±10</td>
<td>73.5±8.1</td>
<td>73.9±11.1</td>
</tr>
<tr>
<td>Sex-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>16 (45.7%)</td>
<td>7 (53.8%)</td>
<td>9 (40.9%)</td>
</tr>
<tr>
<td>Female</td>
<td>19 (54.3%)</td>
<td>6 (46.2%)</td>
<td>13 (59.1%)</td>
</tr>
<tr>
<td>Site of tumor-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proximal</td>
<td>17 (48.6%)</td>
<td>5 (38.5%)</td>
<td>12 (54.5%)</td>
</tr>
<tr>
<td>Distal</td>
<td>18 (51.4%)</td>
<td>8 (61.5%)</td>
<td>10 (45.5%)</td>
</tr>
<tr>
<td>Tumor grade-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well differentiated</td>
<td>2 (5.7%)</td>
<td>0 (0%)</td>
<td>2 (9.1%)</td>
</tr>
<tr>
<td>Moderately differentiated</td>
<td>30 (85.7%)</td>
<td>13 (100%)</td>
<td>17 (77.3%)</td>
</tr>
<tr>
<td>Poorly differentiated</td>
<td>3 (8.6%)</td>
<td>0 (0%)</td>
<td>3 (13.6%)</td>
</tr>
<tr>
<td>Stage of disease-no. (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>3 (8.6%)</td>
<td>1 (7.7%)</td>
<td>2 (9.1%)</td>
</tr>
<tr>
<td>II</td>
<td>32 (91.4%)</td>
<td>12 (92.3%)</td>
<td>20 (90.9%)</td>
</tr>
</tbody>
</table>
Unsupervised hierarchical clustering on DNA copy number data using WECCA revealed three clusters. However, KRAS mutation status was not correlated with cluster membership (data not shown).

To determine putative specific chromosomal regions related to KRAS mutation status, univariate analysis with CGHtest was performed yielding 15 chromosomal regions to be significantly different between CRCs with and without mutated KRAS (Table 2). Three regions were located on chromosome 1q23.1-q44, nine on chromosome 8q23.3-q24.3, two on chromosome 11q23.3.
and one region on chromosome 13q21.33-q22.1. Regions 1q23.1-q44 and 13q21.33-q22.1 showed significantly more gains in the mutated KRAS tumors, while regions 8q23.3-q24.3 and 11q23.3 showed significantly more gains in the wild-type KRAS tumors.

**Table 2.** Chromosomal regions which are significantly different between tumors that have mutated KRAS and wild-type KRAS

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Size (Mb)</th>
<th>BAC clones (N)</th>
<th>Mutated KRAS (N=28) % gain</th>
<th>Wild-type KRAS (N=36) % gain</th>
<th>p-value</th>
<th>FDR</th>
<th>Genes (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1q23.1-q25.3</td>
<td>25.72</td>
<td>29</td>
<td>25.0</td>
<td>0</td>
<td>0.003</td>
<td>0.03</td>
<td>345</td>
</tr>
<tr>
<td>1q31.1-q31.3</td>
<td>11.74</td>
<td>15</td>
<td>21.4</td>
<td>0</td>
<td>0.01</td>
<td>0.04</td>
<td>38</td>
</tr>
<tr>
<td>1q32.1-q44</td>
<td>48.50</td>
<td>47</td>
<td>21.4</td>
<td>0</td>
<td>0.01</td>
<td>0.04</td>
<td>591</td>
</tr>
<tr>
<td>8q23.3-q24.11</td>
<td>0.92</td>
<td>5</td>
<td>3.6</td>
<td>25</td>
<td>0.03</td>
<td>0.11</td>
<td>8</td>
</tr>
<tr>
<td>8q24.11-q24.12</td>
<td>4.22</td>
<td>8</td>
<td>3.6</td>
<td>27.8</td>
<td>0.02</td>
<td>0.07</td>
<td>24</td>
</tr>
<tr>
<td>8q24.13-q24.21</td>
<td>5.78</td>
<td>20</td>
<td>7.1</td>
<td>30.6</td>
<td>0.03</td>
<td>0.11</td>
<td>63</td>
</tr>
<tr>
<td>8q24.21</td>
<td>0.16</td>
<td>3</td>
<td>3.6</td>
<td>30.6</td>
<td>0.01</td>
<td>0.04</td>
<td>0</td>
</tr>
<tr>
<td>8q24.21</td>
<td>0.34</td>
<td>5</td>
<td>7.1</td>
<td>30.6</td>
<td>0.03</td>
<td>0.11</td>
<td>1</td>
</tr>
<tr>
<td>8q24.22</td>
<td>2.13</td>
<td>9</td>
<td>3.6</td>
<td>30.6</td>
<td>0.01</td>
<td>0.04</td>
<td>11</td>
</tr>
<tr>
<td>8q24.22</td>
<td>2.18</td>
<td>6</td>
<td>0</td>
<td>27.8</td>
<td>0.003</td>
<td>0.03</td>
<td>19</td>
</tr>
<tr>
<td>8q24.23</td>
<td>1.35</td>
<td>6</td>
<td>0</td>
<td>27.8</td>
<td>0.003</td>
<td>0.03</td>
<td>2</td>
</tr>
<tr>
<td>8q24.3</td>
<td>3.80</td>
<td>5</td>
<td>3.6</td>
<td>33.3</td>
<td>0.03</td>
<td>0.09</td>
<td>103</td>
</tr>
<tr>
<td>11q23.3</td>
<td>2.40</td>
<td>22</td>
<td>0</td>
<td>27.8</td>
<td>0.003</td>
<td>0.03</td>
<td>72</td>
</tr>
<tr>
<td>11q23.3</td>
<td>0.39</td>
<td>2</td>
<td>0</td>
<td>19.4</td>
<td>0.03</td>
<td>0.09</td>
<td>9</td>
</tr>
<tr>
<td>13q21.33-q22.1</td>
<td>2.11</td>
<td>10</td>
<td>67.9</td>
<td>38.9</td>
<td>0.04</td>
<td>0.12</td>
<td>17</td>
</tr>
</tbody>
</table>

Mb, Megabase; BAC, Bacterial artificial chromosome; FDR, False discovery rate.

Of special interest are the genes MYC (128.8 Mb) and FAK (141.7 Mb) at chromosome 8q23.3-q24.3. MYC is a downstream target of KRAS and FAK is a focal adhesion kinase, which can activate the RAS signaling pathway. Figure 2A and B shows that CRC with wild-type KRAS had significantly higher copy number ratios of MYC and FAK compared to the CRC with mutated KRAS (P=0.04 and P=0.01, respectively).

To determine gene dosage effect of FAK copy number status on mRNA expression, combined data of 22 snap-frozen carcinomas was available, which revealed a positive correlation between copy number ratio and mRNA expression for FAK (P=0.02, r=0.49; Figure 3A). Gene dosage effects for MYC could not be evaluated since there was no MYC oligonucleotide probe present on the expression array.
DNA copy number gain of FAK in wild-type KRAS colorectal cancer

| Figure 2. | Boxplots of DNA copy number ratios of MYC (A; C) and FAK (B; D) in CRCs with and without KRAS mutation. Boxplots A and B show CRCs of the learning set (n=64), C and D of CRCs of the validation set (n=35). In the learning set, both genes had significantly higher copy number ratios in the wild-type KRAS tumors.

Differences in copy number profiles within carcinomas with wild-type KRAS

To determine variation in DNA copy number profiles within 36 CRCs of the learning set with wild-type KRAS, we performed unsupervised hierarchical cluster analysis. This yielded 2 clusters, one cluster (“cluster 1”) containing 13 carcinomas and the other (“cluster 2”) 23 carcinomas. In cluster 1, gain of 7p11.2 that harbors the EGFR locus, occurred more often than in cluster 2 (61.5% versus 30.4% respectively, P=0.09), while in cluster 2 gain of 8q24, which harbors MYC and/or FAK, was more frequent than in cluster 1 (43.5% versus 7.7%, P=0.03 and 43.5% versus 15.4%, P=0.14,
respectively). Copy number gains of \textit{EGFR} and \textit{MYC}/\textit{FAK} co-occurred in only a few tumor samples (for \textit{MYC} and \textit{FAK}, 5.6% and 8.3%, respectively) while most of the cases showed gain of either locus (36.1% and 25.0% for \textit{EGFR} and \textit{MYC}, respectively and 33.3% and 25.0% for \textit{EGFR} and \textit{FAK}, respectively).

\textbf{Figure 3.} Scatterplots of DNA copy number ratios and mRNA expression of \textit{FAK} and \textit{MYC} in 22 CRCs of the learning set (A) and 16 CRCs of the validation set (B, C). In the learning and validation sets, for DNA copy number profiling 5K BAC arrays and 44k Agilent oligo arrays were used, respectively. For mRNA profiling, 30K inhouse spotted Compugen and commercially available 40K Agilent arrays were used, respectively.

\textbf{EGFR gene copy number gain and gene dosage effect of EGFR ligands}

\textit{EGFR} copy number has been associated with response to \textit{EGFR}-targeted therapies. The aCGH platform used in the present study contained a 9.18 kb BAC probe, RP5-1091E12 at 7p11.2, which is part of a coding region of the \textit{EGFR} gene. Gain of \textit{EGFR} was observed in 39% (25/64) of the CRCs and \textit{EGFR} copy number gains were equally distributed over the tumors with or without mutated
KRAS (Figure 4). One cancer in the wild-type KRAS group showed a high-level amplification of the EGFR locus spanning 0.62 Mb.

Also high expression levels of the EGFR ligands ephiregulin (EREG) and amphiregulin (AREG) have been associated with response to cetuximab. In the present series, 2 of 64 (3.1%) cases showed copy number gain of 4q13.3 where both genes co-localize. To determine gene dosage effect of EREG and AREG copy number status on mRNA expression, linear regression analysis of CGH and EREG and AREG mRNA expression data of 22 snap-frozen wild-type KRAS carcinomas revealed a positive correlation between copy number ratio and mRNA expression for EREG (P=0.04 and r=0.43) but not for AREG (P=0.24 and r=0.26), as shown in Figure 5.

Figure 4. Boxplot of DNA copy number ratios of EGFR in 64 colorectal carcinomas with and without KRAS mutation from the learning set. EGFR copy number alterations were not significant different between the two groups.

**Validation of results with an independent set of CRC**

We tested the hypotheses of 1) anti-correlation of EGFR and MYC/FAK copy number gain in wild-type KRAS CRC, 2) gene dosage effect of FAK gain, and 3) gene dosage effect of EREG and AREG in a separate series of 35 MSS CRCs.

Also in this series copy number gains of MYC and FAK occurred more frequently in the wild-type KRAS CRCs than in CRCs with mutant KRAS, however, this difference was not significant for FAK
In the 22 wild-type CRCs the anti-correlation of EGFR and MYC was not as strong as in the learning set. Gain of the chromosomal regions harboring these genes co-occurred in 36.4% (8/22) of CRC samples. However, EGFR and FAK co-occurred only in 18.2% of the cases, while 36.4% of CRC samples had EGFR copy number gain without gain of FAK.

For the analysis of gene dosage effect on mRNA expression, data of only 16 tumors was available. Despite the relative small sample size, a positive correlation between copy number gain and expression level for FAK (P=0.02, r=0.56; Figure 3B) was found. In the learning set, gene dosage effect of MYC could not be analyzed due to missing probes. The expression array used for the validation set did contain two MYC oligonucleotide probes, A_23_P215956 and A_24_P178011, but no correlation between copy number and expression was found (P=0.16, r=0.37; Figure 3C). This series did not contain any copy number gains of chromosome 4q13.3 to allow gene dosage analysis of EREG and AREG.

Figure 5. Scatterplots of DNA copy number ratios and mRNA expression of the genes EREG and AREG in the learning set (n=22). Despite the low frequency of 4q13.3 gain, where both genes co-localize, a positive correlation for EREG (p=0.04 and r=0.43) was found. AREG did show a trend, but no significant correlation was found (p=0.2 and r=0.26).

Discussion

In the present study we demonstrated that CRCs with wild-type KRAS, harbor significantly more gains of chromosomal regions 8q23.3-q24.3 and 11q23.3 in comparison to CRCs with mutated KRAS. Vice versa, in mutant KRAS CRCs, gains of the chromosomal regions 1q23.1-q44 and
DNA copy number gain of FAK in wild-type KRAS colorectal cancer | 165

13q21.33-q22.1 were significantly more frequent. Of special interest is chromosome 8q24 since MYC as well as FAK (also known as PTK2) localize in that region. Both genes are related to the RAS/RAF/MAPK signaling pathway; MYC as a downstream target [26] and FAK as an upstream regulator [27-28]. Increased copy number and overexpression of MYC have frequently been observed in primary and metastasized CRC [29-33]. However, in the small series in the present study no significant correlation between MYC DNA copy number and mRNA expression was observed, which is consistent with other previous reports [30,34]. This could indicate that gain of 8q24 actually perturbs a gene dosage effect of other genes, one of which could be FAK, for which copy number did correlate with mRNA expression. This association has previously been reported for primary lung, breast and colon cancer cell cultures, and for primary breast cancer samples [34-35]. FAK is a key molecule involved in tumor initiation, growth and metastasis by stimulating cell survival, proliferation, migration, invasion and angiogenesis [36]. The FAK protein has been reported to be frequently overexpressed in a range of human tumors, i.e. colon, breast, thyroid, prostate, oral cavity, brain and ovary [37-39], and increased FAK protein expression and activity have been found to be correlated with poor patient outcome [40-42]. Against the background of the present study, i.e. evaluation of DNA copy number changes specific for wild-type KRAS CRC with the purpose of identifying putative predictors of response to anti-EGFR therapy in wild-type KRAS CRC patients, the function of FAK is highly relevant since overexpression of its gene has been found to enhance fibronectin-stimulated RAS-dependent MAPK signaling [43]. So, continuously activated RAS signaling can be induced by mutated KRAS, but also by FAK overexpression, which again may well be the result of a gene dosage effect. This then leads to the hypothesis that CRC patients whose wild-type KRAS tumor contains FAK copy number gain may fail to benefit from anti-EGFR therapy, like CRC patients whose tumors carry a KRAS mutation. Consequently, alternative drugs that better suit the genetic make-up of their tumor may have more benefit to these patients than anti-EGFR therapy. Interestingly, FAK inhibitors do exist and currently, a FAK specific inhibitor is being tested for several solid tumors in a phase I trial with promising results [44].

When analyzing the wild-type KRAS CRC samples from the learning and validation set, we observed that copy number gains of EGFR and FAK co-occurred in a small percentage of cases (8.3% and 18.2%, respectively), while CRC samples with either EGFR or FAK copy number gain occurred more frequent (33.3% versus 25.0% and 36.4% versus 13.6%, in learning set and validation set, respectively). Conflicting data have been reported on the association of EGFR copy number gain with response to anti-EGFR therapy, where some authors reported a positive correlation [5,11,13,14], others did not [4,10]. These contradicting results could be due to the fact that those studies
did not differentiate between wild-type and mutated KRAS tumors. Another explanation could be a confounding effect of DNA copy number status of FAK on EGFR copy number gain. High expression levels of EREG and AREG, which are EGFR ligands, have also been reported as possible indicators of response to cetuximab [10,15]. We therefore aimed to investigate DNA copy number gain and gene dosage effects for these genes. DNA copy number gain of 4q13.3, where EREG and AREG co-localize, appeared to be quite rare in both learning and validation set (3% and 0%, respectively), nonetheless increased copy numbers of the EREG/AREG locus was found associated with increased mRNA expression of EREG. This could indicate that also copy number gain of the EREG/AREG locus could be a marker of response to anti-EGFR therapy.

In conclusion, we demonstrated that within the subgroup of wild-type KRAS CRC heterogeneity exists at the DNA copy number level. These copy number differences may be related to the variable response to anti-EGFR therapy among wild-type KRAS CRC patients. Specifically DNA copy number gain of FAK or the EREG/AREG locus may be relevant next to copy number status of EGFR itself. This hypothesis warrants a large scale validation study to prove whether the markers presented in this study could be of use in the selection of patients for EGFR-targeted therapy.

Acknowledgements

This study was financially supported by the Dutch Cancer Society KWF-2002-2618, KWF-2007-3974 and CTMM DeCoDe project.

References