Small bowel adenocarcinoma copy number profiles are more closely related to colorectal than to gastric cancers


Departments of 1Pathology; 2Epidemiology and Biostatistics, VU University Medical Center, Amsterdam; 3Department of Mathematics, VU University Amsterdam, Amsterdam, The Netherlands; 4Section of Medicine, Surgery and Anaesthesia, Leeds Institute of Molecular Medicine, St James’s University Hospital, University of Leeds, Leeds, UK; 5Department of Gastroenterology, VU University Medical Center, Amsterdam; 6Department of Surgery, Leiden University Medical Center, Leiden, The Netherlands; 7Department of Pathology and Tumour Biology, Leeds Institute of Molecular Medicine, St James’s University Hospital, University of Leeds, Leeds, UK; 8Department of Pathology, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

Received 30 November 2010; revised 22 February 2011; accepted 28 February 2011

Background: Small bowel adenocarcinoma (SBA) is a rare cancer and consequently, the options for clinical trials are limited. As they are treated according to either a colorectal or a gastric cancer regimen and the molecular biology of a tumor is a pivotal determinant for therapy response, chromosomal copy number aberrations were compared with the colorectal and gastric adenocarcinomas.

Materials and methods: A total of 85 microsatellite stable (MSS) adenocarcinomas from the stomach, colorectum and small bowel were selected from existing array comparative genomic hybridization (aCGH) datasets. We compared the aCGH profiles of the three tumor sites by supervised analysis and hierarchical clustering.

Results: Hierarchical clustering revealed substantial overlap of 27 SBA copy number profiles with matched colorectal adenocarcinomas but less overlap with profiles of gastric adenocarcinomas. DNA copy number aberrations located at chromosomes 1p36.3-p34.3, 4p15.3-q35.2, 9p24.3-p11.1, 13q13.2-q31.3 and 17p13.3-p13.2 were the strongest features discriminating SBAs and colorectal adenocarcinomas from gastric adenocarcinomas.

Conclusions: We show that MSS SBAs are more similar to colorectal than to gastric cancer, based on the 27 genome-wide DNA copy number profiles that are currently available. These molecular similarities provide added support for treatment of MSS small bowel cancers according to colorectal cancer regimens.

*Correspondence to: Dr B. Ylstra, Department of Pathology, VU University Medical Centre, PO Box 7057, 1007 MB Amsterdam, The Netherlands. Tel: +31-20-4442495; Fax: +31-20-4442964; E-mail: B.Ylstra@vumc.nl

© The Author 2011. Published by Oxford University Press on behalf of the European Society for Medical Oncology. All rights reserved. For permissions, please email: journals.permissions@oup.com
introduction

Small bowel adenocarcinomas (SBAs) are rare and comprise only ~2% of gastrointestinal malignancies [1, 2]. Prognosis of patients with SBAs is poor with a 5-year overall survival rate of 20%–30% [3]. SBAs are usually diagnosed at an advanced stage requiring systemic drug therapy in addition to surgical resection. Clinical trials to determine optimal treatment regimens for patients with SBAs have been limited due to the low incidence. As a result, patients with SBAs are treated with either colorectal or gastric cancer drug regimens [2, 4].

The molecular biology of a tumor is a pivotal determinant for therapy response, implying that common molecular characteristics lead to comparable drug sensitivity. Relations in the molecular biology between SBAs on the one hand and colorectal and gastric cancer on the other could therefore aid in therapy selection. Associations of response to systemic chemotherapy with molecular characteristics are ample; however, they are limited for a relatively new technique such as array comparative genomic hybridization (aCGH) [5–8]. One genome-wide study demonstrates that a deletion of chromosome 11q is correlated with a better response to anthracycline-based chemotherapy in early breast cancer [8]. In another study, we demonstrated that colorectal cancer patients who do respond to systemic combination therapy with capcitabine and irinotecan have regions located on chromosome 18 frequently deleted [6]. Specific gene amplifications, such as v-myc myelocytomatosis viral-related oncogene, epidermal growth factor receptor and erythroblastic leukaemia viral oncogene homologue-2 (ERBB2), are decisive factors in therapy choice [5]. Copy number aberrations are frequently referred to as a hallmark of cancer. Indeed, by analyzing DNA copy number profiles of 373 cancers, we demonstrated that cancers cluster according to their embryological origin [9]. Others have shown that tumor subtypes that are based on copy number can have different prognosis [10]. Thus, although currently limited precedent exists for genome-wide copy number profiles in relation to therapy, copy numbers well represent the biology of a tumor.

Since aCGH can be carried out using DNA extracted from formalin-fixed paraffin-embedded (FFPE) archival tissue [11], it is feasible to create a series of rare tumors such as SBAs. The aim of our current study was to investigate whether SBA copy number profiles are more similar to those of colorectal or gastric adenocarcinomas. The result of such a comparison may provide further insights into the biology of SBAs and establish some evidence for choosing the optimal treatment regimen for these patients.

materials and methods

data collection

The series of samples in the present study was composed of three previously published series [12–16]. In these studies, two different aCGH platforms were used: 5K bacterial artificial chromosome (BAC) arrays [17] and 30K oligonucleotide arrays [18]. All data were generated using DNA isolated from FFPE tumor tissues and had >70% tumor cells. Obviously, the number of available SBA samples that could be included was the rate-limiting factor. Microsatellite instable (MSI) adenocarcinomas were excluded since these represent a different biological entity [19]. The 27 microsatellite stable (MSS) samples [14] were matched with gastric and colorectal adenocarcinomas based on TNM (tumor–node–metastasis) stage [20], gender and age. If multiple samples matched equally well, one sample was randomly selected (supplemental Table S1, available at Annals of Oncology online).

Gastric adenocarcinomas were selected from a set of 206 samples analyzed with BAC aCGH (Buffart et al. submitted for publication) including 122 randomly selected samples from the Dutch D1/D2 trial [15] and 84 randomly selected samples from the archives of the University of Leeds (Leeds; General Infirmary, UK). Of these samples, 139 were MSS all of intestinal type.

Colorectal adenocarcinomas were selected from an initial aCGH dataset of a panel of 74 colorectal tumors, carried out on BAC aCGH [11]. Of these 74 colorectal adenocarcinomas, 64 are MSS. All aCGH data are publicly available under accession number GSE3418 (http://www.ncbi.nlm.nih.gov/geo/ and supplemental Table S1, available at Annals of Oncology online).

cross-platform calibration

Since the original DNA copy number profiles were generated using two different aCGH platforms, before clustering and classification, resolution and deflection differences of the copy number profiles were calibrated. To facilitate the calibration procedures, two small bowel, two gastric and two colorectal adenocarcinomas were analyzed on both the BAC and oligonucleotide aCGH platform. The results from this experiment were used to establish the required preprocessing steps in order to obtain platform-independent profiles and copy number calls. Preprocessing methods as well as meta-analysis procedures were carried out using R, version 2.6.1. Processing procedures are described in the supplemental Materials and Methods (available at Annals of Oncology online). The success of the platform calibration was assessed by hierarchical clustering of the six samples, which showed that the same samples clustered in pairs independent of the aCGH platform (supplemental figure S1, available at Annals of Oncology online). For clustering a dedicated R package, (WECCA [21]) designed for clustering copy number profiles rather than expression profiles, was used with probabilities as described by Smeets et al. [22] with settings: ‘ordinal’, ‘all equal’ and ‘ward linkage’.

meta-analysis procedures

The entire dataset was preprocessed, normalized and calibrated as described for the calibration samples. DNA copy number data were determined by ‘CGHcall’, calling probabilities of 0.5 or more. The accuracy of the normalization, segmentation and calling was verified by visual inspection. Regions of gains and losses were used for the supervised and unsupervised analysis [23].

For supervised analysis, a two-sample Wilcoxon test using 10 000 permutations was carried out to calculate the significance of DNA copy number differences between small bowel and gastric adenocarcinomas and between small bowel and colorectal adenocarcinomas for each region [24]. Two separate tests were carried
out to compare frequencies of gains and frequencies of losses between the groups. $P$ values were corrected for multiple testing using permutation-based false discovery rate (FDR). Regions with FDR < 0.05 were taken into account to calculate the difference in number of regions and length of the genome between small bowel and gastric adenocarcinomas and between small bowel and colorectal adenocarcinomas. Similarities between small bowel and gastric adenocarcinomas and between small bowel and colorectal adenocarcinomas were defined by the overlap of aberrant regions occurring in >33% per organ of origin, as previously described [25]. Additionally, to determine similarities between small bowel and colorectal and between small bowel and gastric adenocarcinomas, a Pearson’s correlation was calculated based on the frequencies of the gains and losses of the regions. Per region, the maximum frequency of either loss or gain was calculated. The frequencies of losses were multiplied by −1.

Unsupervised clustering was carried out as described for the control samples. To determine the number of most stable clusters, consensus clustering [26] was implemented. The stability of —two to five clusters was measured by running the clustering 1000 times leaving out 20% of the samples each time.

To evaluate whether the samples cluster by site of origin, the chi-square test was used. To find the most discriminative features between clusters, the distance between all regions was defined as the symmetric Kullback–Leibler divergence [27]. To obtain the list of University of California Santa Cruz (UCSC) genes located on the most discrimination regions, the UCSC table browser data retrieval tool was used [28]. Genes overlapping with the Cancer Census [29] list were identified (Table 1).

### results

#### frequency of DNA copy number aberrations in selected gastrointestinal adenocarcinomas

The overall pattern of DNA copy number changes in MSS adenocarcinomas of the small bowel, stomach and colorectum found in the current study is consistent with frequencies of deletions and gains reported in literature (Figure 1). Most common aberrant regions in colorectal adenocarcinomas were gains of chromosomes 7, 8q, 13, and 20q and losses of chromosomes 4, 8p, 14, 15, 17p and 18 [30, 31]. Most common aberrant regions in SBAs were gains of chromosomes 5p, 7, 8q, 13, 16 and 20 and losses of chromosomes 4, 5q, 8p, 17p, and 18 [32, 33]. Most common aberrant regions in gastric adenocarcinomas were gains of chromosomes 1p, 7p, 8q, 17q, 20 and losses of chromosomes 3p, 4, 5q, 6q, 9p, 12q, 13 and 18q [31, 34–36]. Thus, the selected sets of samples from each tumor site were considered to be representative.

#### differences and similarities between gastrointestinal adenocarcinomas by supervised analyses

Significant differences as well as similarities are observed between the three tumor sites (Figure 1). Similarities were determined by the overlap of the genome that was aberrant in >33% of the tumors per site. Copy number aberrations that were observed in >33% of the 27 SBAs encompassed 32% of the genome. Similarly, 30% and 36% of the genome was aberrant with frequencies of gains or losses >33% in colorectal and gastric adenocarcinomas, respectively. Sixty-four percent of these aberrant regions in SBAs were also aberrant in colorectal adenocarcinomas. The overlap of aberrant regions was slightly smaller between gastric adenocarcinomas and SBAs, namely 50% (Figure 1). Furthermore, Pearson’s correlations were calculated to estimate the similarities in frequency of aberrant regions between the different tumor types. Frequencies of the small bowel were more correlated with colorectal than with gastric adenocarcinomas (Pearson’s correlations were 0.75 and 0.63, respectively).

The differences in number and length of aberrant regions were larger between SBA and gastric adenocarcinomas (50 regions with FDR <0.05 over 499 Mb) compared with SBA and colorectal adenocarcinomas (18 regions with FDR <0.05 over 153 Mb). These results indicate that the DNA copy number aberrations might be organ specific and that SBAs are molecularly more similar to colorectal than to gastric adenocarcinomas. Differences between SBA versus gastric adenocarcinomas and SBA versus colorectal adenocarcinomas might be somewhat skewed due to platform differences.

#### unsupervised hierarchical clustering of gastrointestinal adenocarcinomas

Unsupervised hierarchical cluster analysis of colorectal and gastric adenocarcinomas revealed two distinct clusters (Figure 2A): one containing primarily gastric adenocarcinomas ($n = 24$) and only two colorectal adenocarcinomas, the other one containing primarily colorectal adenocarcinomas ($n = 27$) and only five gastric adenocarcinomas ($P < 0.001$). In addition, we identified that this dataset consists of two stable clusters by consensus clustering [26] (data not shown), confirming that these two clusters are robust.

Hierarchical clustering of gastric adenocarcinomas, colorectal adenocarcinomas and SBAs also revealed two distinct clusters (Figure 2B), confirmed by consensus clustering (data not shown). The first cluster contained the same 24 gastric

<table>
<thead>
<tr>
<th>Region</th>
<th>Genes Cancer Census list</th>
<th>Feature ‘GC cluster’ or ‘CRC cluster’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1p36.3-p34.3</td>
<td>LCK, MDS2, PAX7, PRDM16, RPL22, SDHB</td>
<td>More frequently gained in ‘GC cluster’</td>
</tr>
<tr>
<td>4p15.3-q35.2</td>
<td>CHIC2, FBXW7, FIP1L1, IL2, KIT, PDGFRα, PHOX2B, RAP1GDS1</td>
<td>More frequently lost in ‘GC cluster’</td>
</tr>
<tr>
<td>9p24.3-p11.1</td>
<td>FANCG, JAK2, MLLT3, PX5</td>
<td>More frequently lost in ‘GC cluster’</td>
</tr>
<tr>
<td>13q13.2-q31.3</td>
<td>LCP1, LHFP, RB1</td>
<td>More frequently gained in ‘CRC cluster’ and more frequently lost in ‘GC cluster’</td>
</tr>
<tr>
<td>17p13.3-p13.2</td>
<td>USP6</td>
<td>More frequently lost in ‘CRC cluster’</td>
</tr>
</tbody>
</table>

CRC, colorectal adenocarcinoma; GC, gastric adenocarcinoma.

---

**Table 1.** Overview of the most discriminating chromosomal regions between the ‘colorectal, CRC, cluster’ and the ‘gastric, GC, cluster’ (in order to limit the length of the table, only Cancer Census genes in the regions are listed)
Figure 1. Frequencies of aberrations based on called data for colorectal, small bowel and gastric adenocarcinomas. On the horizontal axis, the 2000 sampled positions are plotted from chromosome 1 to chromosome 22. The X and Y chromosomes are not shown. On the vertical axis, frequencies of gains and losses are depicted. Gains are above zero and losses below zero. To determine $P$ values, two separate tests were carried out to compare frequencies of gains and losses. Bars between the frequency plots illustrate the minimum FDR of the two tests per position in gray scale such that black are the most significantly different regions and white the least significantly different regions between the groups. FDR, false discovery rate.
Figure 2. Dendrograms of unsupervised hierarchical clustering. (A) Colorectal adenocarcinomas (CRC) light gray and gastric adenocarcinomas (GC) black, (B) of all samples, small bowel adenocarcinomas (SBA) dark gray. Samples on the horizontal axis and chromosomal regions on the vertical axis, ordered by chromosomal position. Odd chromosomes are depicted in black and even chromosomes in gray. Within the heatmap, black blocks depict normal copy number, dark gray blocks loss and light gray blocks gain.
Our study demonstrates that SBAs are more like colorectal than gastric adenocarcinomas based on genome-wide DNA copy number aberrations. This is in line with previous studies that investigated other molecular characteristics of these tumor entities. For example, the DNA hypermethylation patterns of a selected panel of genes also indicate such an association [37]. Other studies have focused on similarities between colorectal adenocarcinomas and SBAs alone. These studies indicate that genes mutated in colorectal cancer, i.e. mothers against decapentaplegic homolog 4, adenomatous polyposis coli (APC), v-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (KRAS) and β-catenin were also detected in SBAs although at different frequencies [32, 38]. Methylation and mutation of the APC gene, which plays a key role in the initiation of colorectal neoplasia, were detected in both colorectal adenocarcinomas and SBAs, although less frequently in the latter. This indicates that APC may not be pivotal for the initiation of SBAs [14]. The resulting small bowel and colorectal adenocarcinoma copy number alterations, which occur at a later stage of progression, still show high similarities, likewise the DNA hypermethylation.

In addition, SBAs more often show colorectal cancer characteristic cytokeratin protein staining patterns with CK20 positive and CK7 negative [39, 40]. Under the assumption that tumor biology drives phenotypic characteristics, including drug sensitivity, mutation, methylation and our aCGH data would all support treatment of MSS SBA according to a colorectal cancer regimen.

Patients with advanced colorectal cancer benefit from 5-fluorouracil-based chemotherapy in combination with platinum compounds or irinotecan [41]. Several retrospective SBA studies also demonstrate a survival benefit with such regimens [4, 42, 43]. In addition, one of the few SBA prospective phase II clinical trials treating patients with capecitabine in combination with oxaliplatin, showed a significant increase in overall survival [44]. In this context, it appears reasonable to assume that targeted agents like bevacizumab or cetuximab could also be considered for systemic treatment of KRAS wild-type SBAs since they are also beneficial for metastatic colorectal cancer [40, 45]. As Galsky et al. [46] however point out, similar targets in different tumor types may differ in their behavior as a result of up- and downstream signaling pathways of the therapeutic targets. The genome-wide copy number signature similarities between small bowel and colorectal cancers may suggest that there are also more similarities with up- and downstream pathways.

Not all entities cluster entirely apart from each other; five SBAs clustered together with gastrics. This raises the question whether SBAs with a DNA copy number aberration profile more similar to gastric cancer originate from the duodenum and should be treated like gastric cancer. However, only one of those five SBAs originated from duodenum making a relationship between DNA copy number profile and location within the small bowel unlikely. Similarly, no significant relation was found with TNM status (supplemental Table S1, available at Annals of Oncology online) or amplifications. We specifically looked at the amplification of ERBB2, known to occur frequently in gastric cancer and the target of trastuzumab [47]. In our dataset, in four gastric adenocarcinomas and one SBA, ERBB2 was amplified. Those samples do not cluster together. One of the gastrics and the small intestinal adenocarcinomas cluster with the colorectal cancers. Subtypes of colorectal and gastric adenocarcinomas that have previously been characterized by classical pathological features such as cardia and rectum do not cluster separately based on copy number profiles, neither they influence the clustering of the small bowel, either with colorectal or gastric adenocarcinomas.

SBA is a very rare disease, in contrast to colorectal cancer, and several explanations for the low incidence of SBAs have
been proposed including rapid turnover of small intestinal cells, lower exposure to carcinogenic components because of rapid transit time, lack of bacterial degradation and relatively dilute alkaline environment and higher activity of detoxifying enzymes like glutathione S-transferase [48]. Interestingly, the substantial difference in incidence between SBAs and colorectal adenocarcinomas is not reflected in different DNA copy number patterns observed in these two tumor types.

In this study, we measure chromosomal aberrations and therefore included MSS tumors only. MSI tumors are chromosomally more stable and characterized by loss of DNA mismatch repair [19], which cannot be measured by aCGH. MSI tumors were thus excluded since we study relationships based on copy number here.

In summary, we show that genome-wide copy number profiles of most SBAs overlap primarily with colorectal adenocarcinomas. Despite limited evidence for predictive copy number profiles in general, the molecular similarities of this limited set of SBAs with colorectal cancers indicate that more patients would benefit from the colorectal-specific drug regimen rather than from a gastric-specific regimen. These similarities might also provide a further rationale for the exploration of ‘colorectal cancer type’ drug regimens for systemic treatment of MSS SBAs [4, 14, 48]. Notwithstanding, the fact that this is the only SBA aCGH dataset we know of, validation is still necessary to reinforce these findings.

acknowledgements

We would like to thank Mariette Labots and Begona Diosdado for their clinical input.

funding

This study was supported by the Dutch Cancer Society (KWF VU 2007-3832 JCH, KWF 2004-3051 TEB) and The British Society of Gastroenterology (BSG) for financial support in identifying the patients with SBA via its Research Unit and the Pathological Society of England and Ireland.

disclosure

The authors declare no conflict of interest.

references