Different gene expression profiles in metastasizing midgut carcinoid tumors

Katarina Edfeldt, Peyman Björklund, Göran Åkerström, Gunnar Westin, Per Hellman and Peter Stålberg

Department of Surgical Sciences, Uppsala University, SE-751 85 Uppsala, Sweden
(Correspondence should be addressed to P Stålberg; Email: peter.stalberg@surgsci.uu.se)

Abstract

The genetic events leading the progression of midgut carcinoid tumors are largely unknown. The disease course varies from patient to patient, and there is a lack of reliable prognostic markers. In order to identify genes involved in tumor progression, gene expression profiling was performed on tumor specimens. Samples comprised 18 primary tumors, 17 lymph node (LN) metastases, and seven liver metastases from a total of 19 patients. Patients were grouped according to clinical data and histopathology into indolent or progressive course. RNA was subjected to a spotted oligo microarray and B-statistics were performed. Differentially expressed genes were verified using quantitative real-time PCR. Self-organizing maps demonstrated three clusters: 11 primary tumors separated in one cluster, five LN metastases in another cluster, whereas all seven liver metastases, seven primary, and 12 LN metastases formed a third cluster. There was no correlation between indolent and progressive behavior. The primary tumors with Ki67 >5%, with low frequency of the carcinoid syndrome, and a tendency toward shorter survival grouped together. Primary tumors differed in expression profile from their associated LN metastases; thus, there is evidence for genetic changes from primary tumors to metastases. ACTG2, GREM2, REG3A, TUSC2, RUNX1, TPH1, TGFBR2, and CDH6 were differentially expressed between clusters and subgroups of tumors. The expression profile that assembles tumors as being genetically similar on the RNA expression level may not be concordant with the clinical disease course. This study reveals differences in gene expression profiles and novel genes that may be of importance in midgut carcinoid tumor progression.

Endocrine-Related Cancer (2011) 18 479–489

Introduction

Intestinal carcinoid tumors originate from the neuroendocrine enterochromaffin (EC) cells of the gut. The WHO classification from 2000 divides the tumors into three groups: well-differentiated tumors, well-differentiated endocrine carcinomas, and poorly differentiated carcinomas (Kloppel et al. 2004). The older and previously more commonly used classification divides the tumors according to their embryologic origin into foregut, midgut, and hindgut tumors, which also reflects its localization in the gastrointestinal tract (Arnold 2005).

Midgut carcinoid tumors are the most common type of carcinoid tumors in the gastrointestinal tract and arise in the lower jejunum, ileum, appendix, and cecum. This malignant neoplasm grows slowly (Ki67 proliferating index is often <2%) but nevertheless has a lethal outcome; overall five-year survival is ~60% (Modlin et al. 2003). The annual incidence is about two cases per 1 00 000 persons, but the rate is increasing, probably due to increased sensitivity and awareness in the diagnostic phase. However, there is still often a malignant spread to regional lymph nodes (LNs) and the liver at the time of diagnosis, because the available biomarkers are of low specificity and the clinical course often diffuses (Schnirer et al. 2003).

The tumors are neoplasms of peptide- and amine-producing cells. Some patients will develop the carcinoid syndrome due to excessive secretion of serotonin and other metabolites, which includes flushing, diarrhea, abdominal pain, bronchospasm, and heart disease and other symptoms resulting from fibrosis (Soga et al. 1999). Symptoms and progression rate can greatly differ between individual patients and is almost impossible
to predict due to lack of reliable prognostic indicators. The biomarkers used correlate to total tumor burden, such as chromogranin A in serum and 5-hydroxyindoleacetic acid (5-HIAA) in urine. Several analyses have stated that factors leading to a bad prognosis are advanced age, LN involvement, presence of more than five liver metastases, lack of symptoms at the time of diagnosis, high levels of 5-HIAA, high levels of plasma chromogranin A or neuropeptide K, and the carcinoid syndrome (Janson et al. 1997, Soreide et al. 2000, Hellman et al. 2002, Kolby et al. 2004, Ahmed et al. 2009).

Today, surgery is the only potential cure. To achieve prolonged survival and symptom relief, the patients are treated with cytotoxic agents, biological therapies, and tumor-targeted radionucleotides (Schnirer et al. 2003). Liver metastases can be treated with surgery, radiofrequency ablation, or liver embolization (Eriksson et al. 2008). Nevertheless many tumors are resistant to therapy and there is an urgent need to develop new therapies.

Besides the lack of diagnostic and therapeutic tools, a greater understanding of molecular arrangements in midgut carcinoid tumors is required. Loss of chromosome 18q has been found to be the most frequent genetic aberration (Zhao et al. 2000, Kytola et al. 2001, Lollgen et al. 2001, Tonnies et al. 2001, Wang et al. 2005, Kulke et al. 2008). No alterations in the tumor suppressor genes DPC4/SMAD4 and DCC, located on chromosome 18, are found indicating that other, currently unknown genes are important for pathogenesis (van Eeden & Offerhaus, 2006). Moreover, previous research has treated neuroendocrine neoplasms as a homogenous group of tumors. This is most likely insufficient, because recent studies have demonstrated the heterogeneous nature and different gene expression profiles among subgroups of tumors (Zhao et al. 2000, Kytola et al. 2001, Tonnies et al. 2001, Wang et al. 2005).

To our knowledge, only three expression microarray assays, which include midgut carcinoid tumors, have been performed until today. The gene expression profiles have only been analyzed in a limited number of tumors; in total, three primary tumors, 16 liver metastases, and none LN metastases. There are some suggestions of novel genes (PNMA2, SPOCK1, SERPINA10, GRIA2, GRP112, OR51E1, CXCL14, NKX23, NAP1L1, MAGE-D2, MTA-1, and APLP1) that may be of importance for tumorigenesis, but their role in the development of midgut carcinoid tumors is still unknown (Kidd et al. 2006, Arvidsson et al. 2008, Leja et al. 2009). In two of the microarrays, the gene expression in tumor tissues was compared with normal EC cells and/or bowel mucosa (Kidd et al. 2006, Leja et al. 2009), and in one microarray, the gene expression was compared with a cell line (GOT1; Arvidsson et al. 2008). It is difficult to extract normal EC cells from the bowel mucosa, because these cells are scattered throughout the gastrointestinal mucosa, and isolating these cells will result in contamination from surrounding tissue. Owing to these difficulties, this study has a new approach. To clarify the gene expression pattern in midgut carcinoid tumors, with emphasis on genes involved in progression and aggressiveness, an expression microarray assay was performed comparing primary midgut carcinoid tumors with metastases and comparing groups of tumors with different clinical features.

Materials and methods

Tumor material and RNA preparation

Primary and metastatic midgut carcinoid tumors were snap frozen in liquid nitrogen and kept at −70 °C. The tumors came from a cohort of patients who were diagnosed with midgut carcinoid tumor and operated at Uppsala University Hospital between 1989 and 2007. Patient records were evaluated, and totally 42 tumors from 19 patients were selected for further analysis. Clinical features are shown in Table 1; two patients had been treated with interferon α prior to surgery and ten with somatostatin analogues. According to the WHO classification, all of the tumors were well-differentiated neuroendocrine carcinomas. Endocrine tumors have often been treated as a homogenous group in the literature, but resent research has stated their heterogeneity; thus, pancreatic neuroendocrine tumors are foregut and not midgut carcinoid tumors. In 11 patients (no. 1–5, 7–10, 17, and 19), RNA was extracted from both primary tumors and LN metastases, and from five patients (no. 11–15), RNA was extracted from the primary tumor, LN, and liver metastases. From one patient (no. 6), tumor material came from only the primary tumor, and from another patient (no. 16), the tumor material came from the primary tumor and liver metastasis, and from one patient (no. 18) RNA was extracted from LN and liver metastasis. Total RNA was extracted using TriZol Reagent (Invitrogen) according to the manufacturer’s instructions. Sections from all the tumors were stained with Mayer’s hematoxylin and histopathologically evaluated, and the tissues used contained at least 80%, and in most cases more than 90%, of neoplastic cells. When needed, the tumor was macroscopically resected from excessive stromal tissue, blood vessels, bowel mucosa,
lymphoid, and liver tissue. Quality and quantity of the extracted RNA was assessed by Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). Only one LN metastasis, of the 17 LN in total, contained lymphatic tissue (< 5%; tumor no. 18).

To identify genes that are important for tumor progression and prognosis, the patients were classified and divided into two groups according to clinical data. A group with six patients featured a more aggressive clinical pattern and a second group with ten patients had a less aggressive clinical behavior. The clinical criteria are viewed in Table 2. Of the clinical criteria, progression was defined according to RECIST 1.1 (Eisenhauer et al. 2009), and slow progression was defined as progression at more than 12-month intervals on radiological imaging and three patients that were included in the microarray did not classify into any of the groups (patient no. 12, 14, and 18). According to the TNM staging for midgut endocrine tumors (Rindi et al. 2007), 18 patients were classified as stage IV and one patient (no. 2) was classified as IIIB (and displayed no clinically evident liver metastases).

Gene expression array

Microarrays including a human Genome Oligo Set containing 24 650 genes and 37 123 gene transcripts (version 3.0) were performed. The arrays were printed with a QArray2 (Genetix, Hampshire, UK) instrument with 48 K2805 pins (Genetix) on Ultra GAPS slides (Corning, Lowell, MA, USA). Detailed protocols and the complete gene list are available at http://www.biotech.kth.se/molbio/microarray. Neoplasm samples were labeled with Cy5 and RNA reference with Cy3.

Gene expression analysis

Image analysis was performed in GenePix 5.1 Software (Axon Instruments, Sunnyvale, CA, USA). After removal of bad quality spots (if <70% of foreground pixels were below background intensity plus 2 s.d. in both channels or if the difference between ratio of medians and regression ratio exceeded 20% in one of the channels), analysis of the gene expression data was carried out in the freely available statistical computing language R (http://www.r-project.org) using a package available from the Bioconductor project (www.bioconductor.org). The raw data was normalized using the LOWESS method (Yang et al. 2002). In order to search for differentially expressed genes between groups, an empirical Bayes moderated t test was applied, using the Limma package (Smyth 2004). To address the problem of multiple testing, the P values were adjusted using the Benjamini and
Hochberg method (Benjamini 1995). Hierarchical clustering of genes for three-dimensional visualization of expression pattern was performed by the principal component analysis (PCA). PCA is used to reduce the dimensionality of a data set consisting of a large number of variables, while retaining as much as possible the variation present in the data set. The data is transformed to a new set of variables, the principal components (PCs), which are uncorrelated. This simplifies the structure of the data and visualizes hidden structures. PCA mapping was observed in three-dimensional space in which the x-, y-, and z-axes represent the first, second, and third PC respectively (Ma & Dai 2011).

Gene ontology enrichment and gene network analysis

The web-based gene ontology (GO) analysis tool DAVID (http://david.abcc.ncifcrf.gov/; Dennis et al. 2003, Huang da et al. 2009) and Ingenuity Pathways Analysis (Ingenuity Systems, www.ingenuity.com) were used for functional annotations and to determine statistical overrepresentation of GO terms among genes that differed between groups of tumors. A fold change cutoff of 2 was set to identify genes whose expression was differentially regulated and a P value of $<0.05$ was considered significant. For all analysis, the GO terms biological process, cellular compartment, and top canonical pathways were considered. To map differentially expressed genes with genes already known to be involved in tumorigenesis, the Computational biology center web site was used (http://cbio.mskcc.org/CancerGenes/SelectAndDisplay.action; Higgins et al. 2007).

Quantitative real-time RT-PCR analysis

Genes with a fold change difference $>2$ and with known gene function, involved in different types of tumorigenesis, were selected for further analyses. Quantitative real-time PCR (RT-qPCR) was performed to verify the microarray results and relative mRNA expression was determined. DNA-free RNA was prepared using the Nucleospin RNA kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany). Successful DNase treatments were established by PCR analysis of all RNA preparations. Reverse transcription of total DNA-free RNA was performed with random hexamer primers using the First-Strand cDNA Synthesis kit (GE Healthcare, Buckinghamshire, UK) according to the manufacturer’s instructions. RT-qPCR reaction was performed on step 1 RT-PCR system (Applied Biosystems, Foster city, CA, USA). In addition to cDNA, all PCR reactions contained 200 mM of each primer and 1 µl probe and TaqMan Gene Expression Master Mix (Applied Biosystems). All samples were amplified in triplicate, and non-template controls were included. Each sample’s mean threshold value was corrected for the corresponding mean value for GAPDH rRNA, used as internal control.

The following labeled TaqMan probes were used for step 1 RT-PCR System: GAPDH (4333764F), actin gamma 2 isoform (ACTG2; Hs00242273_m1), gremlin 2 (GREM 2; Hs01934663_s1), regenerating islet-derived 3 alpha (REG3A; Hs00170171_m1), tumor suppressor candidate 2 (TUSC2; Hs00200725_m1), runt-related transcription factor 1 (RUNX1; Hs01021971_m1), tryptophan hydroxylase 1 (TPH1; Hs00188220_m1), transforming growth factor-β receptor II (TGFBRII; Hs00559661_m1), and cadherin 6 (CDH6; Hs00191832_m1).

Non-parametric Mann–Whitney U test or Student’s t test was used to test significance between expression levels for RT-qPCR and Kaplan–Meier to test differences in survival among clusters. Statistics were calculated using SPSS (Chicago, IL, USA) and differences were considered significant if $P < 0.05$.

Table 2 Clinical parameters for expression array comparisons (patient no. 12, 14, and 18 did not fit this classification and were exempt from this comparison)

<table>
<thead>
<tr>
<th>More aggressive ≥ 3 criteria Patient no. 1, 3, 8, 9, 15, and 42</th>
<th>Less aggressive ≥ 3 criteria Patient no. 2, 4, 5, 6, 7, 10, 11, 13, 16, and 17</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt;5 liver metastases</td>
<td>&lt;5 liver metastases</td>
</tr>
<tr>
<td>&gt;70 years old at diagnosis</td>
<td>&lt;70 years old at diagnosis</td>
</tr>
<tr>
<td>Fast progression</td>
<td>Slow progression</td>
</tr>
<tr>
<td>Carcinoid syndrome</td>
<td>No carcinoid syndrome</td>
</tr>
<tr>
<td>Severe diarrhea</td>
<td>Mild or no diarrhea</td>
</tr>
<tr>
<td>Ki67 &gt; 5%</td>
<td>Ki67 &lt; 5%</td>
</tr>
<tr>
<td>No radical surgery</td>
<td>Radical surgery</td>
</tr>
</tbody>
</table>

$^a$Slow progression defined as progression according to RECIST with more than 12-month intervals.
$^b$More than 5 per day.
Tumors with undetectable mRNA levels of the gene of interest were arbitrarily set to a relative value of 0.0001.

**Ethical permission**

This study was approved by the local ethical committee, Uppsala, and informed consent was obtained from all patients.

**Results**

**Comparison of gene expression profiles of primary tumors and metastases**

In order to identify differentially expressed genes, the mRNA expression levels of primary tumors and metastases were globally analyzed. Of the 37,123 transcripts, 7,130 were significantly ($P < 0.05$) altered between primary tumors and LN metastases. With a two-time fold change and known gene name, 289 genes differed in expression level. The number of genes with a significant and at least a 1.5-time fold change difference among the tumor subgroups is summarized in Table 3, with detailed data in Supplementary Table 1, see section on supplementary data given at the end of this article. A comparison of expression levels between primary tumors in the group defined as less aggressive (number of tumors $\geq 6$; Table 2) and the more aggressive group (number of tumors $= 10$) revealed no significant differences (Table 3). Three patients (no. 12, 14, and 18) could not be classified according to the set criteria and were therefore excluded from this comparison. The results from the microarray analysis have been deposited on NCBI’s GEO (accession number GSE27162).

To further analyze gene expression patterns, we performed unsupervised hierarchal cluster analysis using the PCA method (Ma & Dai 2011). The expression data was reduced to three PCs that visualized three clusters (Fig. 1). Cluster 1 contained 11 primary tumors; cluster 2 contained five LN metastases, and cluster 3 contained seven primary tumors, 12 LN metastases, and all seven liver metastases. When comparing primary tumors with their respective associated LN metastases, 13 of 16 matched pairs were situated in different clusters, thus expressing a different gene expression profile. Only three of 11 patients with primary tumors in cluster 1 suffered from the carcinoid syndrome compared with the rest of the primary tumors (in cluster 3) where five of seven patients had the carcinoid syndrome; three tumors displayed a high proliferation rate with Ki67 $> 5\%$ in hot spots and all of these tumors were gathered in cluster 1 and seven patients eventually succumbed from their carcinoid disease and interestingly all of them resided in cluster 1. However, when analyzing cumulative survival between patients from the two clusters of primary tumors, there was no statistically significant difference ($P = 0.15$; Fig. 2). A tendency toward shorter survival for patients in cluster 1 was observed, but a larger patient cohort is needed to investigate this properly. There was no correlation between radical primary surgery and survival.

**Quantitative real-time RT-PCR**

RT-qPCR was performed to verify expression differences of eight genes selected from the identified clusters of the PCA and from the comparison of all primary tumors ($n = 18$) to all LN metastases ($n = 17$). Gene selection was based on both fold-change differences and known gene function. ACTG2, GREM2, and REG3A were selected from the comparison of all primary tumors to LN metastases. TUSC2 was chosen from the comparison of primary tumors from clusters 1 and 3, and finally RUNXI, TGFBR2, TPHI, and CDH6 were chosen from the comparison of LN metastases from clusters 2 and 3. All eight genes were verified as significantly ($P < 0.05$) differentially expressed (Fig. 3).

**GO enrichment**

To characterize the functional significance of genes with a high expression in primary tumors and LN

### Table 3

<table>
<thead>
<tr>
<th></th>
<th>No. of genes*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinically more aggressive group ($n = 6$)</td>
<td>0</td>
</tr>
<tr>
<td>Primary tumors ($n = 18$)</td>
<td>289</td>
</tr>
<tr>
<td>Primary tumors ($n = 16$)</td>
<td>324</td>
</tr>
<tr>
<td>Primary tumors ($n = 18$)</td>
<td>266</td>
</tr>
<tr>
<td>Lymph node metastases ($n = 17$)</td>
<td>11**</td>
</tr>
<tr>
<td>Interferon-treated tumors ($n = 2$)</td>
<td>7</td>
</tr>
</tbody>
</table>

*P $< 0.05$, fold change 2; **fold change 1.5.
metastases, and of significantly differently expressed genes in clusters, GO term analysis was performed and GO terms were identified. GO term analysis facilitates the interpretation of data providing biological and functional descriptions of gene products (Fig. 4). Comparing the expression of genes in all of the primary tumors with all of the LN metastases, GO terms that are unique and enriched in primary tumors were related to ‘the extracellular region’ and genes with a high expression in LN metastases were associated with ‘the organelle and cell cycle’. LEF1, FOS, and MYC genes, known to be involved in colorectal cancer, were highly expressed in the LN metastases. GO enrichment analysis also revealed that genes highly expressed in the primary tumors in cluster 1 were significantly associated with ‘the plasma membrane and cell adhesion’ in contrast to primary tumors in cluster 3, which were associated with ‘chromosomal structure, cell cycle, and mitosis’. The four genes involved in the p53 signaling pathway was highly expressed in cluster 3 primary tumors (GTSE1, CCNB2, CCNE2, and RRM2), but not in cluster 1. Both the serotonin receptor signaling pathway and the regulation of hormone levels were both associated with LN metastases in cluster 3.

Cell adhesion molecules (CAMs) were differentially expressed in all three clusters. In cluster 2, 45 CAMs were highly expressed compared with nine different CAMs in the LN metastases in cluster 3. The primary tumors in cluster 1 had a high expression of 21 CAMs compared with none in the primary tumors in cluster 3. Gene lists of differentially expressed CAMs are presented in Supplementary Table 2, see section on supplementary data given at the end of this article.

When using the Computational Biology center website and mapping the gene lists from the expression microarray against known cancer genes, several genes were identified in the clusters. Gene lists are shown in Table 4.

**Discussion**

To enhance our understanding of pathophysiological arrangements in midgut carcinoid tumors, this study aimed at identifying clinically relevant genes involved in tumor progression and aggressiveness. Comparing pathological tissue with its normal counterpart is by its very nature a difficult task because the EC cells are scattered throughout the gastrointestinal mucosa and every attempt at isolating these cells will result in contamination from surrounding tissue, thus making interpretation difficult. Therefore, the goal in this study was to identify genes involved in tumor progression, which would mirror the problem faced by the clinician in selecting the patients requiring aggressive clinical treatment from the large group of patients whose disease will follow an indolent course, despite having similar presentation and tumor burden at diagnosis. In an attempt to identify distinguishing genes of interest, we characterized two groups of patients with aggressive ($n=6$) and non-aggressive disease ($n=10$). RNA was isolated from primary tumors, LN metastases, and liver metastases from these groups (and from another three patients) and an oligo microarray analysis was performed. Surprisingly, there were no differentially expressed genes in the primary tumors from these two groups of patients when sorted according to
well-recognized clinical criteria for more or less aggressive disease. Analyzing primary tumors against LN metastases identified 289 genes that were differentially expressed. Some of these genes could be of relevance in tumor progression and to verify the microarray data and identify genes that warrant further study, RT-qPCR was performed on ACTG2, GREM2, and REG3A. ACTG2 is located on chromosome 2p13 and has been implicated in cisplatin resistance (Watson et al. 2007) and cell cycle arrest (Li et al. 2008). GREM2 is located on chromosome 1q43 and belongs to the cystine knot superfamily. GREM2 has been shown to interact with bone morphogenetic protein (BMP) in the cross talk with the WNT/β-catenin pathway (Kosinski et al. 2007) and to have an increased expression in colorectal cancer (Segditsas et al. 2008). REG3A, located on chromosome 2p12, is involved in liver and pancreatic regeneration and proliferation, and is a target of β-catenin signaling in hepatoma cells (Cavard et al. 2006). These genes may have a protective function, because the expression levels are high in primary tumors and greatly reduced or absent in metastases.

ACTG2, GREM2, and REG3A are novel genes not earlier known to be involved in carcinogenesis in endocrine tumors. These genes have not, to our knowledge, been detected in other microarray studies. Two microarray studies have compared the gene expression in midgut carcinoid tumors with normal mucosa and/or normal EC cells (Kidd et al. 2006, Leja et al. 2009) and one with a cell line from a liver metastasis (GOT1; Arvidsson et al. 2008). In the work by Leja et al. (2009), NKX23 had a lower expression level in liver metastases compared with primary tumors, and this is in accordance with the findings in this study; however, this has not been verified by RT-qPCR. None of the expression array has compared the gene expression between primary tumors and LN metastases with emphasis on tumor progression and level of aggressiveness of the disease. Thus, it is of no

Figure 3 Relative gene expression. RT-qPCR data on selected genes with comparison between primary tumors (n=18) and lymph node metastases (n=17). Relative gene expression ratios are shown adjusted for GAPDH. A, ACTG2; B, GREM2; C, REG3A. *, outlier; filled circle, mean.

Figure 4 Relative gene expression. RT-qPCR data on selected genes with comparison between A, clusters 1 (n=11) and 3 (n=7) of primary tumors, TUSC2. B–E, clusters 2 (n=5) and 3 (12) of lymph node metastases; B, RUNXI; C, TPH1; D, TGFB2, and E, CDH6. *, outlier; filled circle, mean.
relevance to compare the results from this microarray with the other, above-mentioned, microarray studies.

To find differentially expressed genes, the data was analyzed and compared between groups of tumors and hierarchal cluster analysis was performed using PCA, which surprisingly revealed three clusters. Primary tumors in PCA of cluster 1 displayed a tendency toward shorter survival, had fewer patients with the carcinoid syndrome, and showed high proliferation rate (Ki67 > 5%) implying different pathophysiological mechanisms or different stages of differentiation. Of the seven patients, four that succumbed from the disease had received radical primary surgery (Table 1); hence, the survival analysis is unlikely to depend on the surgical results. In the comparison among clusters of primary tumors, TUSC2 located on chromosome 3p21.3 was identified. TUSC2 is frequently deleted in multiple cancers and has an anti-tumor effect (Ito et al. 2004). RT-qPCR verified low expression in primary tumors in cluster 1, which reflects the tendency toward high mortality and a high proliferation rate in these tumors.

All the LN metastases in cluster 2 are from patients who are still alive and thus the gene expression could eventually be a marker for prognosis. No other clinical

<table>
<thead>
<tr>
<th>Primary tumors compared to lymph node metastases</th>
<th>Lymph node metastases compared to primary tumors</th>
<th>Cluster 1 compared to primary tumors in cluster 3</th>
<th>Primary tumors in cluster 3 compared to cluster 1</th>
<th>Cluster 2 compared to lymph node metastases in cluster 3</th>
<th>Lymph node metastases in cluster 3 compared to cluster 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTA2</td>
<td>APOC1</td>
<td>BANP</td>
<td>ABCC4</td>
<td>BIRC3</td>
<td>ELF3</td>
</tr>
<tr>
<td>ADCA4</td>
<td>BCL2A1</td>
<td>CDK5R2</td>
<td>EME1</td>
<td>CCR7</td>
<td>ERBB3</td>
</tr>
<tr>
<td>ADCA5</td>
<td>BTG2</td>
<td>EIF4G1</td>
<td>KIAA0101</td>
<td>CD3E</td>
<td>F5</td>
</tr>
<tr>
<td>BANP</td>
<td>CA9</td>
<td>ELMO1</td>
<td>MAL</td>
<td>CXCR4</td>
<td>FOXA1</td>
</tr>
<tr>
<td>CDH17</td>
<td>CCNB2</td>
<td>FN1</td>
<td>MTSS1</td>
<td>IL8</td>
<td>MNX1</td>
</tr>
<tr>
<td>CLSN1</td>
<td>CTLA4</td>
<td>GUCY2C</td>
<td>MYB</td>
<td>LCP1</td>
<td>PKIB</td>
</tr>
<tr>
<td>CRYAB</td>
<td>DUSP1</td>
<td>HOPX</td>
<td>MYBL2</td>
<td>LYZ</td>
<td>PTGPRN</td>
</tr>
<tr>
<td>DIRAS3</td>
<td>DUSP2</td>
<td>MUC20</td>
<td>NQO1</td>
<td>MMP9</td>
<td>RAB2B</td>
</tr>
<tr>
<td>DLG2</td>
<td>E2F2</td>
<td>PAK3</td>
<td>PAH</td>
<td>PTPRC</td>
<td>RAB3C</td>
</tr>
<tr>
<td>DMD</td>
<td>EGR1</td>
<td>PRKCG</td>
<td>SERPINA3</td>
<td>TACSTD1</td>
<td></td>
</tr>
<tr>
<td>EPHA7</td>
<td>FOS</td>
<td>SYNPO2</td>
<td>SGK1</td>
<td>TPH1</td>
<td></td>
</tr>
<tr>
<td>FHL1</td>
<td>FOSE</td>
<td></td>
<td>SSX1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FLNA</td>
<td>FOBS</td>
<td></td>
<td>TUSC2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GUCY2C</td>
<td>GADD45B</td>
<td></td>
<td>ULK4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOPX</td>
<td>GAS1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITIHS</td>
<td>IER3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LPP</td>
<td>KIAA0101</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MAOB</td>
<td>KRT17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MUC20</td>
<td>LEF1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYH11</td>
<td>LOX</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPN13</td>
<td>MAL</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTPRN2</td>
<td>MYBL2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAGLN</td>
<td>MYC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNC</td>
<td>NGFR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNXB</td>
<td>NPM1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPSAB1</td>
<td>NR4A1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPSD1</td>
<td>PAH</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TTBK1</td>
<td>PKMYT1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TYRIP1</td>
<td>PTPRC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WFCDC1</td>
<td>SERPINA3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ZBTB16</td>
<td>SGK1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aFold change > 2.
bFold change > 3.
cFold change > 8.
features were correlated to this cluster, which possibly suggests that gene expression levels are not concordant with clinical features in metastases. In the comparison between clusters of LN metastases, RUNX1, TGFBR2, TPH1, and CDH6 were verified using RT-qPCR. RUNX1, located on chromosome 21q22, is a well-known oncogene in leukemia and other cancers (Cohen 2009, Hoi et al. 2010, Kumano & Kurokawa 2010) and can directly promote proliferation (Hoi et al. 2010). GO analysis identified ‘proliferation’ as one of the more significant cellular events in the LN metastases in cluster 2, which also had a high expression of RUNX1. The TGF-β signaling pathway has been considered both as a tumor suppressor pathway and as a promoter of tumor progression and invasion. Somatic mutations in TGFBR2, located on chromosome 3p22, have been shown in several different tumor types. Notably, LN metastases in cluster 3 displayed low TGFBR2 expression. Midgut carcinoid tumors produce serotonin and the gene TPH1, located on chromosome 11p15, codes for the rate limiting enzyme in the biosynthesis of serotonin. Serotonin deficiency is associated with slower tumor growth and increased expression with breast tumor progression (Pai et al. 2009). Interestingly, TPH1 had a higher expression in the LN metastases in cluster 3 compared with those in cluster 2. The GO analysis revealed that in LN metastases in cluster 3, highly expressed genes were associated with the ‘serotonin receptor signaling pathway and regulation of hormone levels’, which may be related to a high expression of TPH1. It is interesting to correlate the expression levels of TPH1 with 5-HIAA levels. Patient levels of 5-HIAA were unfortunately missing in four patients and therefore no correlation between 5-HIAA levels and TPH1 expression was detected (Table 1). If this lack of correlation reflects, the missing data remains an open question. CAMs play essential roles in development and progression of cancer and show clinical significance as prognostic biomarkers or as potential therapeutic targets in a variety of malignancies (Okegawa et al. 2004). In lymph node metastases in cluster 3, 31 CAMs were highly expressed and some of these molecules may be potential biomarkers and predictors of prognosis, for example, CDH6, which is located on chromosome 5 and is involved in cell–cell signaling. Over-expression of CDH6 has been detected in small cell lung cancers (Yokoi et al. 2002) and is implicated as a prognostic marker in ovarian carcinomas (Kobel et al. 2008).

When comparing primary tumors with their respective associated LN metastases, different expression profiles were detected. Of 16 primary tumors, 13 separated from their respective LN metastases and resided in a different cluster. The tumor content is similar in all samples and no LN tissue (except <5% in one sample) was present in the metastases. This pattern demonstrates that the expression profile changes greatly through tumor progression and metastases development.

Interestingly, only 11 genes were expressed with a significant difference between LN metastases and liver metastases, indicating that a greater genetic change occurs in the development from primary tumors to metastases than between local and distant metastases.

There are obvious drawbacks when examining a limited number of tumors; however, this is to date one of the most comprehensive array studies on midgut carcinoid tumors. To provide solid evidence on the described gene involvement in carcinoid tumors, these findings will have to be validated in a larger independent tumor set.

This study reveals that midgut carcinoid tumors fall into three clusters based on their gene expression profiles; thus, there might be different molecular subtypes of midgut carcinoid tumors. Some primary tumors as well as some LN metastases form clusters, with interesting clinical associations. There is evidence for genetic changes from primary tumors to metastases, because most primary tumors displayed a different expression profile than their associated LN metastases. Novel genes, ACTG2, GREM2, REG3A, TUSC2, RUNX1, TGFBR2, TPH1, and CDH6, not previously known to be involved in midgut carcinoid tumorigenesis, displayed different expression levels in tumor clusters and may be of importance for tumor progression.

**Supplementary data**

This is linked to the online version of the paper at http://dx.doi.org/10.1530/ERC-10-0256.

**Declaration of interest**

The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

**Funding**

This study was supported by Lions Cancer Foundation, Swedish Cancer Society and Medical Research Council. We are grateful to Hanna Göransson Kultima for statistical and bioinformatic assistance. P B is a fellow of Swedish Association for Medical Research and supported by Åke Wibergs Foundation.
References


Received in final form 5 May 2011
Accepted 2 June 2011
Made available online as an Accepted Preprint 2 June 2011