Int. J. Cancer: **123,** 817–825 (2008)

Integrated gene copy number and expression microarray analysis of gastric cancer highlights potential target genes

Samuel Myllykangas[1], Siina Junnila[2], Arto Kokkola[3], Reija Autio[4], Ilari Scheinin[1,5], Tuula Kiviluoto[3], Marja-Liisa Karjalainen-Lindsberg[1], Jaakko Hollmén[6], Sakari Knuutila[1], Pauli Puolakkainen[3] and Outi Monni[2]*

[1]*Department of Pathology, Haartman Institute and HUSLAB, University of Helsinki and Helsinki University Central Hospital, Helsinki, Finland*
[2]*Biomedicum Biochip Center and Institute of Biomedicine/Biochemistry, Genome-Scale Biology Research Program, University of Helsinki, Helsinki, Finland*
[3]*Department of Surgery, Helsinki University Central Hospital, University of Helsinki, Helsinki, Finland*
[4]*Institute of Signal Processing, Tampere University of Technology, Tampere, Finland*
[5]*Genome Informatics Unit and Department of Pathology, Biomedicum Helsinki and Haartman Institute, University of Helsinki, Helsinki, Finland*
[6]*Department of Information and Computer Science, Helsinki University of Technology, Espoo, Finland*

We performed an integrated array comparative genomic hybridization (aCGH) and expression microarray analysis of 8 normal gastric tissues and 38 primary tumors, including 25 intestinal and 13 diffuse gastric adenocarcinomas to identify genes whose expression is deregulated in association with copy number alteration. Our aim was also to identify molecular genetic alterations that are specific to particular clinicopathological characteristics of gastric cancer. Distinct molecular genetic profiles were identified for intestinal and diffuse gastric cancers and for tumors obtained from 2 different locations of the stomach. Interestingly, the *ERBB2* amplification and gains at 20q13.12-q13.33 almost exclusively discriminated intestinal cancers from the diffuse type. In addition, the 17q12-q25 gain was characteristic to cancers located in corpus and the 20q13.12-q13.13 gain was more common in the antrum. Statistical analysis was performed using integrated copy number and expression data to identify genes showing differential expression associated with a copy number alteration. Genes with the highest statistical significance included *ERBB2*, *MUC1*, *GRB7*, *PPP1R1B* and *PPARBP* with concomitant changes in copy number and expression. Immunohistochemical analysis of ERBB2 and MUC1 on a tissue microarray containing 78 independent gastric tissues showed statistically significant differences ($p < 0.05$ and <0.001) in immunopositivity in the intestinal (31 and 70%) and diffuse subtypes (14 and 41%), respectively. In conclusion, our results demonstrate that intestinal and diffuse type gastric cancers as well as cancers located in different sites of the stomach have distinct molecular profiles which may have clinical value.

Key words: gastric cancer; array CGH; gene expression; microarray

Gastric cancer is the fourth most common cancer worldwide with almost 1 million people being diagnosed annually, and it is the second most common cause of cancer-related death inflicting 700,000 annual deaths worldwide.[1] The high mortality rate of gastric cancer is explained by the fact that the majority of the tumors in the stomach are malignant gastric adenocarcinomas,[2] detected often at an advanced stage and manifested by lymph node invasion and metastasis.[3] Therapeutic interventions to treat such late stage carcinomas are often restricted to non-curative gastrectomy and lymphadenectomy. Thus, 5-year survival rates of gastric cancer patients are only 10 to 30%.[4–6] Because the early stages of gastric cancer are usually asymptomatic, it is of great clinical importance to identify molecular markers for diagnostic purposes as well as for targeted treatment.

Gastric adenocarcinoma can be divided into 2 main histological subtypes, intestinal (differentiated) and diffuse (undifferentiated).[7] Diffuse gastric cancer is characterized by non-cohesive and scattered tumor cells that infiltrate deep into the stroma and do not form any glandular structures. In contrast, intestinal gastric cancer shows distinguishable, albeit distorted, gland formation and grows by expansion rather than infiltration.[8] Chronic gastritis and intestinal metaplasia are risk factors involved in the pathogenesis of intestinal gastric cancer, whereas the diffuse type is less associated with inflammatory stress and pre-malignant lesions. In addition, the diffuse type usually has a poorer prognosis than the intestinal type.[9] Altogether, diffuse and intestinal types of gastric cancer represent different disease entities in regard to epidemiological and biological characteristics.

Gene copy number aberrations are hallmarks of various human cancers, but their impact on gene expression remains unknown to a large degree. Recently, integration of genome-wide aCGH and gene expression microarray data has provided new information about the molecular mechanisms underlying gene expression alterations. For example, up to 40–60% of the amplified genes have demonstrated simultaneous overexpression in primary tumors and cell lines[10–13] and 10–15% of all gene expression changes have been found to be directly associated with gene copy number changes.[12,14] Identification of deregulated genes that are associated with gene amplification or deletion is likely to reveal targets for copy number alterations and genes with potential functional significance in the malignant phenotype.

Here, we performed genome-wide gene copy number and expression profiling of 46 gastric tissue specimens, including tumors from 2 histological subtypes and locations of the stomach, to identify genes in which alteration in gene expression is associated with a change in gene copy number. We also studied molecular profiles of different gastric cancer subtypes to identify molecular markers with potential clinical significance. Our study is one of the first reports of integrated genome-wide analysis of clinical gastric cancer material enlightening potential target genes for further clinical and functional validation.

Material and methods

Sample processing and collection of clinical data

This research project has been reviewed and approved by the Ethical Committee of the Department of Medical Genetics and Surgery, and the Clinical Review Board of Helsinki University

Samuel Myllykangas and Siina Junnila contributed equally to this work.

This article contains supplementary material available via the Internet at http://www.interscience.wiley.com/jpages/0020-7136/suppmat.

Grant sponsors: Academy of Finland, the State Appropriations of Helsinki University Central Hospital, the Sigrid Jusélius Foundation, Finnish Cancer Organizations, Biocentrum Helsinki and Research Foundation of University of Helsinki.

***Correspondence to:** Biomedicum Biochip Center and Institute of Biomedicine/Biochemistry, P.O. Box 63 (Haarmaninkatu 8) FI-00014 University of Helsinki, Helsinki, Finland. E-mail: outi.monni@helsinki.fi

Received 28 September 2007; Accepted after revision 18 February 2008
DOI 10.1002/ijc.23574
Published online 27 May 2008 in Wiley InterScience (www.interscience.wiley.com).

Central Hospital authorized this investigation. Gastric tissue specimens were prospectively collected from patients who underwent gastric surgery or gastroscopy in the Helsinki University Central Hospital between 1999 and 2005. Samples that had an unclear or mixed histology and samples from patients that had received chemotherapeutic treatment were excluded from further analysis. Also, we prioritized samples of which both high quality RNA and DNA could be extracted from 1 tissue homogenate. These criteria left us with a total of 38 cancer and 8 normal good quality tissue samples. Normal tissue samples were taken as far away from the tumor as possible (minimum distance to tumor was 5 cm). Informed consent was obtained from each participating subject. All tumors were reviewed for invasion (T), lymph node status (N) and metastasis (M), and information about sampling location in the stomach was collected according to standard practice at the Helsinki University Central Hospital (Appendix 1, supplementary data).

Our tissue material consisted of 2 different histological gastric cancer subtypes; intestinal ($n = 25$) and diffuse ($n = 13$). Nine of the cancer samples represented T1–T2 stage tumors and 29 T3–T4 stage tumors, out of which 11 also showed metastasis (M1). Lymph node colonization was not seen in 6 tumor samples (N0) whereas 32 tumors showed colonization in 1–15 regional lymph nodes (N1–N3). The ages of the studied patients ranged from 30 to 99 years (mean age 68) and equal numbers of females and males were included in the study. Location distributions of tumor samples and normal samples were similar, because 50% of the samples in both groups are from the corpus (19 tumor samples and 4 normal samples) and 50% are from the antrum (19 tumor samples and 4 normal samples). Clinical and histological information of all the samples are publicly available at www.cangem.org.

Fresh frozen gastric tissue samples were stored at −70°C. Frozen samples were embedded in Tissue-Tek OCT Compound (Sakura Finetek, Torrance, CA) and 5 μm frozen ice-sections were prepared and stained using Alcian Blue-Periodic Acid Schiff (AB-PAS) and Trypan Blue. Histology of normal tissue and gastric adenocarcinoma subtypes (intestinal and diffuse) as well as tumor cell content in the samples were evaluated by an experienced pathologist (M.-L. K.-L.).

Nucleic acid extraction

After sectioning, Tissue-Tek was removed from the tissues and frozen samples were stored at −80°C in RNAlater-ICE reagent (Ambion, Austin, TX). Because paired DNA and RNA samples were extracted from the same source, tissue samples were co-homogenized in RLT (β-mercaptoethanol) lysis buffer of the RNeasy midi kit (Qiagen, Hilden, Germany) using Ultra-Turrax homogenizer (IKA Works, Wilmington, NC) to ensure that an identical sample was taken for both the copy number and gene expression analyses. Prior to sample homogenization, we verified that the RLT (β-mercaptoethanol) lysis buffer had no effect on the quality of DNA or CGH microarray hybridization (data not shown). After homogenization, DNA extraction was continued using the DNeasy tissue extraction kit (Qiagen) and RNA was extracted using the RNeasy midi kit. DNA and RNA concentrations were measured using Eppendorf Biophotometer (Eppendorf AG, Hamburg, Germany) and nucleic acid quality was evaluated using Agilent's 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and gel electrophoresis (data not shown).

Array comparative genomic hybridization and gene expression microarray analyses

Forty-six gastric specimens were selected for the study based on clinical and histological data. Samples with a majority of tumor cells with high-quality RNA and DNA were analyzed on array comparative genomic hybridization (aCGH) and gene expression microarrays. RNA quality was analyzed using Bioanalyzer 2100 (Agilent), and high-quality total RNA had to have distinct peaks representing 18S and 28S ribosomal RNA with low background values. DNA quality was analyzed using a spectrophotometer (Eppendorf Biophotometer) and the mean 260/280 ratio for high-quality DNA was 1.88. Labeling and hybridization of the DNA samples were performed as described previously.[10,11] Briefly, 20 μg of genomic DNA from gastric samples and reference samples (pooled DNA extracted from buffy coat fractions of whole blood obtained from gender-matched healthy individuals provided by the Finnish Red Cross) was digested overnight at 37°C using *AluI* and *RsaI* restriction enzymes (Sigma–Aldrich, St. Louis, MO) and purified using phenol–chloroform extraction. Four micrograms of digested reference and sample DNA were labeled using Cy3-/Cy5-dUTP fluorescent dyes, respectively (Amersham Biosciences, Piscataway, NJ), and RadPrime DNA Labeling System (Invitrogen, Carlsbad, CA). Labeled DNA was then hybridized on Human-1 cDNA Microarray (Agilent Technologies) at 65°C for 16 h. Prior to scanning the slides were washed in 0.1% SDS (RT, 2 min), followed by 0.5× SSC, 0.01% SDS (RT, 2 min), and 0.06× SSC (RT, 2 min), and dried using centrifugation ($300 \times g$, RT, 5 min).

Gene expression experiments were performed using the Whole Human Genome Oligo Microarray kit (Agilent Technologies). In brief, 20 μg of total RNA extracted from gastric samples and reference RNA (pool of 10 cancer cell lines, non-gastric; ATCC, Manassas, MA) was labeled using the Fluorescent Direct Label kit (Agilent Technologies) and Cy5-dCTP/Cy3-dCTP dyes, respectively (Perkin–Elmer, Wellesley, MA). Hybridization was performed with Agilent's *In Situ* Hybridization Kit Plus (Agilent Technologies) and rinsed according to the manufacturer's instructions. Microarray slides (both aCGH and expression) were scanned using DNA Microarray Scanner (Agilent Technologies).

Microarray data pre-processing

Scanned images were analyzed using Feature Extraction Software 8.1 (Agilent Technologies). Intensity data were imported into the GeneSpring 7.3 program (Agilent Technologies) and aCGH and gene expression microarray normalization was carried out using the Feature Extraction data import plug-in (Agilent Technologies). Default parameters in the Feature Extraction data import procedure for 2-color data normalization were applied. Probes flagged as outliers by Feature Extraction were ignored in the normalization. Sample spot intensities were divided by control channel, and microarrays were normalized to the 50th percentile and probes to the median value. Microarray results were quality filtered using outliers determined by the Feature Extraction Software.

DNA sequence information (3′ and 5′ end reads) for the cDNA probes spotted on Human-1 cDNA microarray was obtained from Agilent Technologies and Incyte (Wilmington, DE). Genomic map positions for every available cDNA probe and Whole Human Genome Oligo microarray probes (60-mers) were identified by aligning them against the human genome (NCBI build 35) by MegaBlast analysis. Each sequence was mapped separately and all probe sequences representing a specific annotated gene or transcript were combined. If the sequence had multiple hits in the genome, it was removed from the analysis. Corresponding Ensembl gene identifiers (EnsGene IDs) were determined for the microarray probes using the genomic positions of the probe sequences. Probes mapping to exon sequences were prioritized. Secondarily, the whole gene sequence was used to match the microarray probe positions with EnsGenes. EnsGene IDs were subsequently used to collect gene annotation data from the Ensembl database (version 37). Original microarray data, including scanned images and image analysis output files, are available at www.cangem.org. Altogether, 11,848 and 37,374 features in aCGH (Appendix 2, available at www.cangem.org) and expression microarrays (Appendix 3, available at www.cangem.org), respectively, had unambiguous genomic map positions and were included in the analysis. Genomic map positions of the probes were used to integrate gene copy number data with the gene expression data. This was accom-

plished by selecting the closest copy number value within a 375-kb distance from the gene expression probe. The middle base pair position of the probes was used to measure the distance. Integration of copy number and gene expression data yielded 33,607 gene expression microarray probes that had the corresponding gene copy number probe within 375 kb (Appendix 4, available at www.cangem.org).

Gene copy number profiling

The CGH-Explorer software[15] was applied to identify gene copy number aberrations and to produce binary data (0; no gain or loss, 1; gain and −1; loss) from array CGH measurements. The ACE algorithm and false discovery rate of 0.002 were used to identify gene copy number aberrations in individual samples and to determine aberration frequencies in gastric cancer (Appendix 5, available at www.cangem.org).

To study whether any copy number alteration clustered into specific histology or location, we used receiver operating characteristic (ROC) analysis.[16] The clinical sample group comparisons were performed for intestinal ($n = 25$) *vs*. diffuse ($n = 13$) and antrum ($n = 19$) *vs*. corpus ($n = 19$) tumors. First, the ROC curve was estimated for each gene using class labels (histology or location) and the information about the gains and losses. The area under the ROC curve was used to measure which genes were significant as classifying the 2 compared sample groups. To ensure meaningful copy number patterns, at least 30 percent of the samples had to have classifying gains or losses in at least one of the compared classes.

Classification of gastric cancers based on multidimensional copy number patterns

To find a set of genes with good performance of classifying patients into classes defined by labels histology and tumor location, we ran a forward selection algorithm[17] on all of the available variables. In order to take into account the statistical variability in the forward selection procedure, we performed a 10-fold cross-validation procedure[17] repeated 50 times at each step of the algorithm. From the experiments, we can calculate the average training error and the average validation error of the resulting 500 repetitions of the experiment trained with different subsamples of the data. The number of selected variables is the number for which we achieve the minimum validation error, suggesting the best generalization ability. As a base classifier, we used the Naïve Bayes classifier. The basis assumption in the Naïve Bayes classifier is that the observed copy number changes are conditionally independent of each other given the class label. The parameters of the classifier were trained in the framework of maximum likelihood.

To assess the statistical significance of the identified variables, we compared them with randomly selected variables. The Naïve Bayes classifier was trained 10,000 times with randomly selected variables. The classification performance of the true selected variables and the performances of the randomly selected variables were calculated. An empirical *p*-value was calculated as the proportion of values in the empirical distribution representing samples of the test statistic under the assumption of the null hypothesis exceeding the classification performance of the true classifier.

Integration of gene copy number and expression data

Thirty-eight tumor samples were included in the integration analysis. Normal tissue samples were not used in the integration, because we compared the gene expression levels of tumor samples with copy number alterations to tumor samples without copy number alterations to assess if an alteration in gene copy number is reflected on the gene expression level in gastric tumors. The integration of gene copy number and gene expression data was done as previously reported by Hautaniemi *et al*.[18] Chromosomal regions of gains and losses were identified using CGH-Explorer and the ACE algorithm with a false discovery rate of 0.002. Gains and losses were treated separately in the analysis. First, expression in tumor samples with a copy number gain was compared to expression in tumor samples with a normal copy number. Second, expression in tumor samples with a copy number loss was compared to expression in tumor samples with a normal copy number. Gene copy number data was specified as a vector of zeros (normal copy number) and ones (gain or loss) indicating the copy number status of each gene across all tumor samples. Then, a weight, w_g, was calculated for expression levels of each probe using signal-to-noise statistics $w_g = (m_{g1} - m_{g0})/(\sigma_{g1} + \sigma_{g0})$, where m_{g1} and σ_{g1} represent the mean and the standard deviation of the expression levels of probes in samples showing gain or loss ($g1$) and m_{g0} and σ_{g0} mark the expression level and the standard deviation of probes in samples showing normal copy number ($g0$). Random permutations ($n = 10{,}000$) of the label vectors were then performed to calculate the statistical significance of the association between gene expression and gain or loss. α-Values (comparable with *p*-value) below 0.05 indicated a statistically significant association between gene copy number and expression. The global effect of gene copy number on expression was evaluated by calculating the proportion of probes that showed association between expression and copy number changes.

Immunohistochemistry of gastric tissues on tissue microarray

Expression of ERBB2 and MUC1 proteins in gastric cancer was validated using immunohistochemistry and an independent tissue microarray containing 78 gastric tumor specimens. The preparation of tissue microarray has been documented previously.[19] Five micrometer sections were cut from the tissue microarray block. The microarray slides were first stained with hematoxylin to evaluate the histology of the samples. In brief, tissue array slides were pre-treated with Tris-EDTA (pH 9.0) or with citrate (pH 6.0) for 24 min in a microwave oven prior to staining with ERBB2 and MUC1 antibodies, respectively. Slides were then stained using 1:500 dilution of NCL-CBII (c-erbB-2 Oncoprotein) and 1:25 dilution of NCL-MUC-1-CORE (Muc-1 Core Glycoprotein) mouse monoclonal antibodies (Novocastra Laboratories, Newcastle, UK). Immunoreactivity of the tested samples was scored as negative, weak positive, positive, or strong positive. Significance of the differences in the frequencies of immunopositive samples between intestinal and diffuse types and samples from distal (antrum) and proximal (corpus and cardia) stomach were evaluated using a *z*-test for 2 proportions.

Results

Gastric cancers of different histology and location show specific copy number alterations

Our gastric cancer samples showed a number of copy number alterations, with the most frequent gains at 17q12-q21, 20q, 8q, and X, and losses at 4q (Fig. 1). 17q12-21 was the only highly amplified region in our material and was only present at the intestinal type (Appendix 6, supplementary data). ROC analysis was applied to identify genes that classify intestinal and diffuse type gastric cancers and tumors located in antrum and corpus. We identified 125 genes with increased copy number that classified intestinal and diffuse types of gastric cancers (Appendix 7, supplementary data). These include gains of *ERBB2* at 17q12 and gains at 19q13.32, 20q11.22-q13.33 and Xp. Gains at 19q13.32 were characteristic to diffuse type gastric cancer, whereas all other gains at the specified regions were more common in the intestinal type. On the other hand, 138 genes classified gastric cancers located in antrum and corpus (Appendix 7). These genes were located at 8q24.13-q24.21, 17q21.2 and 20q13 regions. Gains at 17q21.2 were specific for corpus located gastric cancers and gains at 8q24 and 20q13 were specific for gastric cancers located in the antrum. No deletions classified gastric cancers based on histology or location.

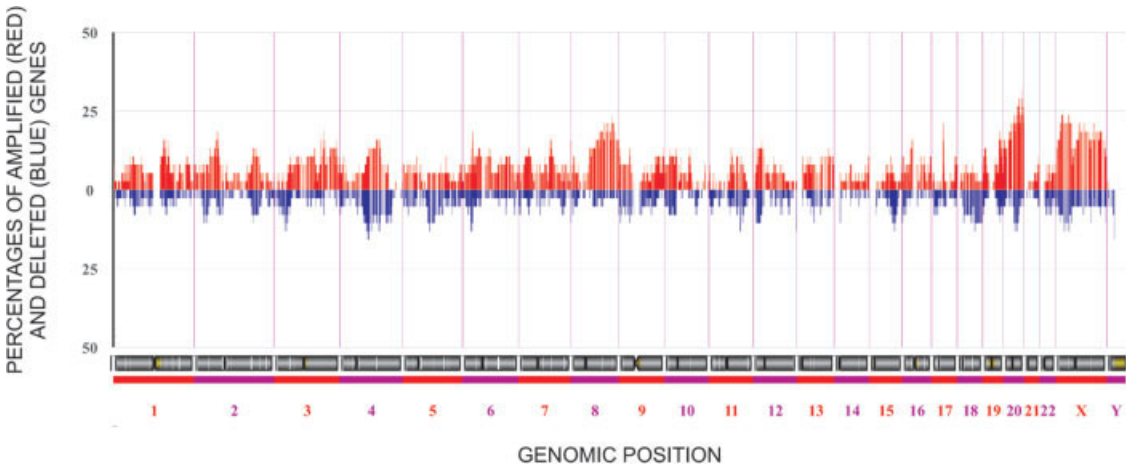

Figure 1 – Frequency plot of DNA copy number changes in 38 gastric adenocarcinomas using CGH microarrays and CGH-Explorer software. Frequencies of copy number aberrations of each clone have been plotted on the *Y*-axis. Gains are marked in red and losses in blue. Genomic positions of the probes are marked on *X*-axis.

Specific copy number patterns classify gastric cancers of different histology and location

We identified six amplified genes that discriminated intestinal from diffuse gastric cancers by Naïve Bayes model. These genes included *ERBB2*, *NAPA*, *GNAS*, *XK*, *PTGIS* and 1 unknown gene (ENSG00000104866). With these genes, the percentage of correct classification was 81.6% with a *p*-value of 0.0002 as compared to the randomly selected variables. On the other hand, a pattern of 5 amplified genes discriminated cancers located in antrum from corpus located gastric cancers. These genes included *SLC35C2*, *BCAS1*, *NEURL2*, *CYP24A1* and 1 unknown gene (ENSG00000131747). The corresponding percentage of correct classification was 73.7% with a *p*-value of 0.0003.

Gene copy number changes are associated with alterations in gene expression

Altogether, 5.9% of the probes ($n = 1980$) were deleted in at least 10% of the cancer samples, and 6.3% of these probes ($n = 125$) showed an association between copy number loss and underexpression (α-value < 0.05). On the other hand, 27.8% of the probes ($n = 9336$) were gained or amplified in at least 10% of the cancer samples. Of these probes, 9.6% ($n = 895$) showed an association between copy number gain and overexpression (α-value <0.05).

Statistical analysis of the dependencies in the aCGH and gene expression microarray measurements revealed 657 individual genes that showed an association between overexpression and copy number gain and 95 genes that showed an association between underexpression and copy number loss (Appendix 8, available at www.cangem.org). Natural logarithm of the ratio of mean expression levels between samples labeled with 1 (denoting amplification or deletion) and 0 (denoting no copy number change) ($\ln[m_{g1}/m_{g0}]$) were plotted to visualize the genes that had an association between the gene copy number aberration and altered gene expression in gastric adenocarcinoma (Fig. 2). Only those genes that had at least a 2-fold change in the gene expression ratio ($\ln[m_{g1}/m_{g0}] > 0.7$ or < -0.7) between the samples with amplification or deletion and normal copy number are shown in Figure 2 and Table I. Integrated analysis of copy number and expression pinpointed several known gastric cancer genes such as *PPP1R1B*, *STARD3*, *PNMT*, *PERLD1*, *ERBB2*, *GRB7*, *JUP* (17q12-21),[20–24] *MUC1* (1q22),[25] *VEGF* (7p21.1)[26] and *MYC* (8q24)[27] (Table I). Novel genes included *PPARBP*, *SMURF1*, *HOXA9*, *HOXA10*, *MYO6*, *MYST4*, *SMAD1*, *NARG1*, *FGFR3*, *VDAC2*, *GSDML*, *KRT23* and *MALT1* with no previous association with gastric cancer (Table I).

ERBB2 and MUC1 protein overexpression is associated with intestinal cancers

Protein expression levels of ERBB2 (HGNC Id: 3430) and MUC1 (HGNC Id: 7508) were investigated on a tissue microarray to identify differences in the protein expression between the different histological subtypes or tumor location (Fig. 3; Appendix 9, supplementary data). The analysis was successful for 64 and 68 samples stained with MUC1 and ERBB2, respectively. According to the *z*-test for 2 proportions, both ERBB2 (*z*-score 1.98, $p < 0.05$) and MUC1 (*z*-score 3.06, $p < 0.001$) showed positive staining more frequently in the intestinal type gastric cancer (31% and 70%, respectively) than in the diffuse type (14% and 41%, respectively). Furthermore, ERBB2 and MUC1 were not differentially expressed in tumors located in the proximal (corpus and cardia) and the distal (antrum) sections of the stomach. These data suggest that *ERBB2* and *MUC1* overexpression may have clinical value in the intestinal type gastric cancers, regardless of the location of the tumor in the stomach.

Discussion

Several studies have provided insights into the importance of specific copy number alterations in the development of solid tumors, showing that these alterations may lead to the altered expression of cancer critical genes.[11,28] Our aim was to integrate the genome-wide gene copy number and expression microarray data of 46 surgical stomach specimens to identify genes whose expression is deregulated due to altered copy number. Using aCGH, gene expression and tissue microarrays as well as custom-developed bioinformatics tools, we pinpointed genes whose altered expression was associated with increased or decreased copy number. Our data also showed distinct genomic profiles in gastric cancers depending on histology (intestinal and diffuse) or site (antrum and corpus), implicating genes with specific biological or therapeutic roles in gastric cancer.

Different carcinogenetic pathways have been suggested for intestinal and diffuse gastric cancer and also our data revealed differences in molecular genetic alterations in 2 histological subtypes. Well-differentiated intestinal adenocarcinoma is proposed to arise through chronic gastritis–intestinal metaplasia-carcinoma sequence whereas genetic risk factors are thought to be associated with the diffuse type.[9] According to our results, intestinal type gastric cancers could be differentiated by gains of *ERBB2* at 17q12, as well as gains at 20q11.22-q13.33 and Xp, whereas gains at 19q13.32 were more common in the diffuse type. We and others have previously reported the higher frequency of 17q12 and 20q

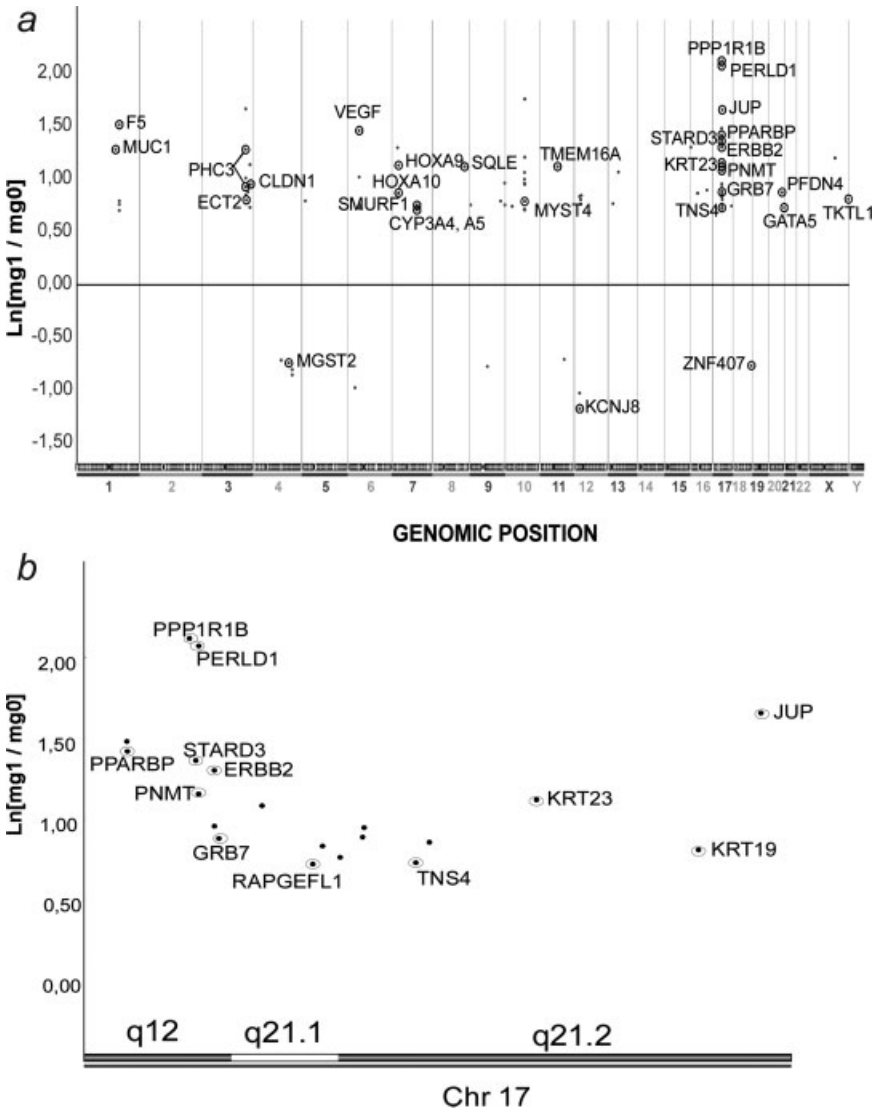

FIGURE 2 – Genes showing highest statistical association between gene copy number and expression in gastric cancer. Gene expression and copy number data sets were integrated using genome positions of the arrayed probes. Genes with consistent copy number and expression values were then analyzed using signal-to-noise statistics and random permutation test. Genes with an α-value <0.05, altered copy number in at least 10% of the cancer samples, and at least a 2-fold change in gene expression ratio are shown. Genomic position of each probe is plotted on the *X*-axis and the natural logarithm of mean gene expression ratios of samples with copy number alteration to samples without alterations are shown in the *Y*-axis. (*a*) Genome-wide image and (*b*) a zoomed-in image of the 17q12-q21 region from position 34,653 to 37,277 kb.

amplifications in intestinal gastric cancer.[23,29,30] Moreover, we identified 2 sets of genes with copy number gains that classified gastric cancers based on histology and location. *ERBB2* was included in the gene set that separated intestinal from the diffuse type gastric cancers. Additionally, ERBB2 was shown to be specifically expressed in the intestinal type by immunohistochemistry and tissue microarray. The differences in copy number of the diffuse and intestinal tumor components were recently investigated in mixed types of gastric tumors by Carvalho *et al*.[31] In their study, no differences in copy number profiles were identified between the intestinal and diffuse components of the mixed type gastric cancers. These findings support the idea that the mixed type is a third histological subtype in gastric cancer with specific phenotype and genomic changes and that the intestinal and diffuse types are separate entities. The identified patterns of copy number gains could be valuable biomarkers in diagnostics and in selecting treatment modalities for different gastric cancer subtypes.

Most gastric cancers occur in the distal third of the stomach, the antrum. Other common gastric cancer sites are the cardia (upper third) and corpus (middle third). In our study, all the samples were obtained either from the corpus or antrum to ensure exclusion of esophageal tumors expanding to gastric cardia. Our results show that 17q12.2 region is gained in gastric cancers of the corpus whereas gains at 8q24 and 20q13 were more frequent in antrum-derived gastric tumors. Interestingly, proximal gastrointestinal tumors (e.g., Barrett's adenocarcinoma) also have amplifications at 17q, whereas distal gastrointestinal tumors (*e.g*., colorectal cancer) have frequent amplifications at 20q.[32] We hypothesize that their location in the gastro-intestinal tract and a similar cell-of-origin[33] might explain comparable molecular genetic alteration for these cancer types.

Although gene expression and copy number profiles specific to gastric cancer and its histological subtypes have been identified,[34–37] gene expression or copy number profiling alone is not

TABLE I – PUTATIVE GASTRIC CANCER TARGET GENES IN 38 PRIMARY GASTRIC ADENOCARCINOMAS IDENTIFIED USING CGH AND EXPRESSION MICROARRAY ANALYSES

Probe name	Gene symbol	Chromosome band	α	$\ln[m_{g1}/m_{g0}]$	Copy number aberration frequency	Examples of ssociated cancers	PubMed ID
A_23_P137856	*MUC1*	1q22	0.0107	1.28	0.11	Gastric, ovarian, breast, colorectal	16403482, 16880776, 16846534, 16857798
A_23_P148768	*F5*	1q23	0.0027	1.51	0.13	–	–
A_23_P409988	*PHC3*	3q26.2	0.0061	1.28	0.13	Osteosarcoma	17001316
A_23_P9574	*ECT2*	3q26.31	0.0248	0.80	0.13	Esophagus	15761962
A_24_P165949	*CLDN1*	3q28	0.0022	0.95	0.11	Colorectal	16253248
A_23_P110167	*MGST2*	4q28.3	0.0026	−0.75	0.13	–	–
A_24_P12401	*VEGF*	6p21.1	0.0001	1.46	0.11	Gastric, ovarian, breast, osteosarcoma	16761623, 16835828, 16596190, 16528367
A_23_P500998	*HOXA9*	7p15.2	0.0057	1.13	0.13	ALL	11314021
A_23_P253368	*HOXA10*	7p15.2	0.0071	0.87	0.13	Endometrium, breast	15627890, 15044858
A_23_P353316	*SMURF1*	7q22.1	0.0055	0.75	0.13	Pancreatic	16036106
A_32_P223327	*CYP3A4, CYP3A5*	7q22.1	0.0284	0.70	0.13	Prostate	15548719
A_23_P146284	*SQLE*	8q24.1	0.0027	1.11	0.21	Lung	17316888
A_24_P941586	*MYST4*	10q22.2	0.0077	0.78	0.11	AML	15147375, 16626284
A_24_P87036	*TMEM16A*	11q13.3	0.0376	1.12	0.11	Breast, pancreatic, gastric, parathyroid, head and neck	12739008
A_23_P64879	*KCNJ8*	12p11.23	0.0000	−1.18	0.11	–	–
A_23_P425704	*PPARBP*	17q12	0.001	1.42	0.11	Breast, gastric	10485914, 12430186, 11691784
A_23_P129835	*PPP1R1B*	17q12	0.0007	2.12	0.18	Gastric, UGC	12124342, 14991576, 16825313, 15188007
A_24_P224538	*STARD3*	17q12	0.0025	1.36	0.18	Breast, UGC	16708353, 1682531
A_23_P100642	*PNMT*	17q12	0.016	1.15	0.18	Gastric, breast, adrenal	15010812, 12727839, 16047163
A_24_P275828	*PERLD1*	17q12	0.0002	2.07	0.18	Gastric, breast	12739007, 15010812, 14578197
A_23_P89249	*ERBB2*	17q12	0.0035	1.30	0.21	Gastric, breast, lung	16718853, 14991576, 16622439, 16847810
A_23_P163992	*GRB7*	17q12	0.0002	0.88	0.18	Gastric, pancreatic, TGCT	11980659, 16595785, 16354586
A_23_P207850	*TNS4*	17q21.2	0.0008	0.73	0.18	Lung	12711115
A_23_P78248	*KRT23*	17q21.2	0.0011	1.12	0.13	Pancreatic	11135429
A_23_P501822	*JUP*	17q21.2	0.0001	1.65	0.13	Gastric, TGCT	11980659, 11956097, 14612934
A_23_P380954	*ZNF407*	18q23	0.0057	−0.73	0.11	–	–
A_23_P166023	*PFDN4*	20q13.2	0.0032	0.87	0.21	Breast	11381030
A_23_P371835	*GATA5*	20q13.33	0.0192	0.73	0.24	Gastric, colorectal, ovarian, lung	14612389, 16337738, 15585625
A_23_P259901	*TKTL1*	Xq28	0.0454	0.81	0.11	Colon, urethelial	16465194

α Value reports the significance of the association between gene copy number alteration and expression. $\ln[mg_1/mg_0]$ corresponds to natural logarithm of mean expression ratio of samples with copy number alterations (mg_1) to samples with no copy number alterations (mg_0). Positive values correspond to a gain or amplification related overexpression and negative values to loss related underexpression. The examples of associated cancers indicate the cancer types in which mutations, altered RNA expression, protein expression, or copy number have been reported with corresponding PubMed IDs.

SCC, squamous cell carcinoma; ALL, acute lymphoid leukemia; MM, malignant mesothelioma; UGC, upper gastrointestinal cancers (gastric and esophageal); AML, acute myeloid leukemia; TGCT, testicular germ cell tumor.

optimal in identification of specific targets for therapeutic applications. We integrated both the copy number and gene expression information to define the genes whose altered gene copy number is associated with gene expression change and which might therefore be critical in gastric cancer development. We identified 752 putative gastric cancer target genes with altered copy number (gain or loss) present in more than 10% of the cancer samples and whose expression was attributable to copy number alteration (α-value $<$ 0.05). Overall, the impact of gene copy number on gene expression in gastric cancer was in line with previous reports of other solid tumors[10–13] as our data showed that 10% and 6% of the genes within the regions of copy number gains or losses were over- or underexpressed, respectively. For example, *PFDN4* and *GATA5* were amplified and overexpressed at 20q13.2, which is one of the most commonly amplified regions in gastric cancer, and *MUC1* copy number had the most significant impact on gene expression at 1q22. We then compared our results to a recent copy number and gene expression microarray study of 30 gastric tumors.[38] The overlapping genes included *GATA6*, *ANKRD11*, *CDC6*, *MAZ*, *MFHAS1*, *GATA4, PPARBP, GRB7* and *SEC22L3*, of which *GRB7* (18.4 %), *CDC6* (15.8%), *MAZ* (13.2%) and *PPARBP* (10.5%) were the most frequently amplified genes in our material. The frequency of copy number alterations in the other genes ranged from 5.3 to 7.9 % in our study and from 3.3 to 6.7 % in the study by Yang et al.[38] Although the frequencies reported by Yang et al. were relatively low, these results might point to the biological role of these genes in gastric cancer.

In our study, the only frequent high-level amplification occurred at 17q12-21 (Appendix 6). A number of genes are previously suggested as 17q12-21 targets in gastric cancer.[19,24] In our study, *PPARBP*, *PPP1R1B*, *PERLD1*, *STARD3*, *PNMT*, *ERBB2*, *GRB7* and *JUP* showed the most significant overexpression associated with copy number gain, suggesting that the 17q12-21 gain and its reflection on gene expression is instrumental in gastric carcinogenesis. *PPP1R1B* is a previously suggested target gene in the 17q amplicon in gastric cancer.[19] The amplification of 17q12-21 has been previously studied in upper gastrointestinal adenocarcinomas and the genomic region from *PPP1R1B* to *GRB7* showed the highest correlation between amplification and gene expression.[23] The analysis was carried out by FISH using a BAC clone, which spans

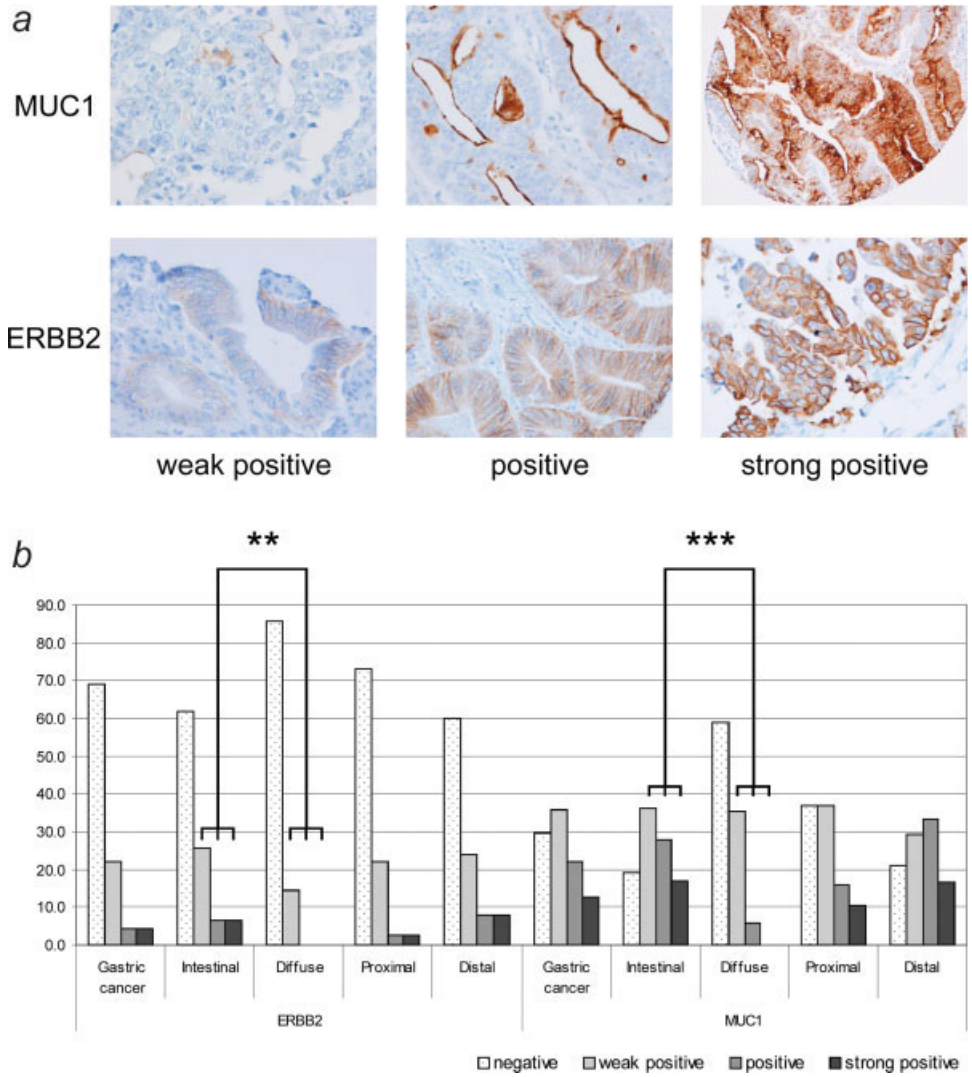

FIGURE 3 – Immunohistochemistry for ERBB2 and MUC1 proteins on a tissue microarray. (*a*) Scoring of the ERBB2 and MUC1 staining. Monoclonal antibodies and tissue microarray were applied to determine the expression of ERBB2 and MUC1 in gastric cancer. Immunoreactivity was scored as weak positive, positive, or strong positive. (*b*) Analysis of ERBB2 and MUC1 immunohistochemistry on tissue microarray. Proximal refers to corpus and cardia, whereas distal denotes antrum. ** indicates statistically significant difference ($p < 0.05$) and *** denotes statistically highly significant difference ($p < 0.001$) according to the z-test test of 2 proportions.

the whole region, and therefore, the copy number ratios of the individual genes could not be studied in great detail. In our data, *PPP1R1B*, *PERLD1* and *JUP* showed the highest copy number and gene expression levels (Fig. 2). Amplification and overexpression of *PPARBP*, located proximal to *PPP1R1B-GRB7* locus, is a novel finding in gastric cancer but it has previously been shown to be amplified and overexpressed in breast cancer.[39] Additionally, frameshift mutations of *PPARBP* have been observed in gastric cancer with a high-frequency of microsatellite instability.[40] A number of genes in the 17q12-q21 amplicon play pivotal roles in cell cycle regulation, cell signaling, and cell–cell interactions, which all act as integral parts in tumorigenesis.[20,41–44] Our results show that the distribution of amplifications appears to be non-random and confined to specific genomic sites. The most commonly amplified region was at chromosome 17, which is gene-rich (>15 genes/Mb) and abundant in open chromatin fibers.[45] The open conformation of chromatin observed in chromosome 17 might render it susceptible to DNA double-strand breaks. Amplifications at 17q12-q21 could be explained by elevated frequency of DNA breaks leading to initiation of the amplification mechanism.[46]

To determine the clinical relevance of the genes that showed an association between copy number and expression, we studied 78 clinical tumor samples using a tissue microarray. The expression levels of MUC1 and ERBB2 proteins were shown to associate with the intestinal gastric cancer more than with the diffuse type. MUC1 immunoreactivity was positive in the majority of gastric tumors. It was remarkably higher than would be expected based on genomic profiling analysis, suggesting that, in addition to gene copy number gain, MUC1 expression is increased also due to other mechanisms. Interestingly, erbB receptor tyrosine kinases have been shown to phosphorylate MUC1,[47,48] which then interacts with β-catenin.[49] Furthermore, *H. pylori* has been shown to bind tandem repeat domain of MUC1 on the surface of gastric epithelium[50] and activate β-catenin in gastric epithelial cells.[51] β-Catenin has been shown to enhance invasiveness and promote proliferation in gastric cancer cell lines.[52] Based on these data we speculate that MUC1 might be involved in gastric carcinogenesis by promoting proliferation and invasion through interaction with β-catenin after being activated by copy number gain, ERBB2 kinase or *H. pylori* infection. Moreover, *MUC1* has been shown to inhibit cell–cell and cell–stroma interactions and facilitate cell

spreading and metastasis[53] and *MUC1* expression has been used as a marker of circulating tumor cells in the peripheral blood of gastric cancer patients.[25] 31% of the analyzed gastric cancers were positive for immunohistochemical ERBB2 staining, which is in line with the results obtained from gene expression microarrays. A clear indication of *ERBB2* involvement in gastric cancer was established when trastuzumab (ERBB2 tyrosine kinase domain binding monoclonal antibody) treatment was shown to inhibit tumor growth in gastric cancer cell lines and in a patient with *ERBB2* amplification and overexpression.[43,44] Strong positive or positive staining of ERBB2 was only detected in the intestinal gastric cancer. Furthermore, *ERBB2* was the only gene at 17q12 amplicon discriminating intestinal and diffuse cancers supporting the importance of ERBB2 in the intestinal type. Altogether, the immunohistochemical analysis of MUC1 and ERBB2 confirmed the results from the gene copy number profiling, integrated copy number and gene expression analysis and illustrated that gene overexpression associated with a copy number gain also leads to higher protein expression.

In conclusion, we carried out genome-wide gene copy number and expression microarray profiling of gastric adenocarcinoma and applied statistical analysis to identify putative target genes that are deregulated in association with a copy number alteration. Intestinal and diffuse types of gastric adenocarcinoma with distinctive clinicopathological features showed different molecular genetic profiles, which may have a diagnostic value. Furthermore, tumors located in the antrum and corpus were clearly 2 separate disease entities in regard to their molecular genetic characteristics. Integration of genome-wide information by CGH and expression microarrays allowed us to identify genes that are likely to be involved in gastric carcinogenesis and to discover potential therapeutic and clinical targets for gastric cancer. To further validate the clinical and biological significance of the identified targets, subsequent studies of these genes should include investigation of protein level alterations and functional properties.

Acknowledgements

The authors thank Dr. Sampsa Hautaniemi (Systems Biology Laboratory, Institute of Biomedicine, Genome-Scale Biology Research Program, University of Helsinki, Helsinki, Finland) for critical reading of the manuscript. They are also grateful to Mrs. Ulla Kiiski (HUSLAB, Helsinki University Central Hospital, Helsinki, Finland) for preparing the tissue sections and Ms. Mia Kero (HUSLAB, Helsinki University Central Hospital, Helsinki, Finland) for technical help in the immunohistochemical analyses.

References

1. Ferlay J, Bray F, Pisani P, Parkin D. GLOBOCAN 2002: cancer incidence, mortality and prevalence worldwide. IARC cancerbase no. 5 version 2.0. Lyon: IARC Press, 2004.
2. Schwartz GK. Invasion and metastases in gastric cancer: in vitro and in vivo models with clinical correlations. Semin Oncol 1996;23:316–24.
3. Hundahl SA, Phillips JL, Menck HR. The National Cancer Data Base Report on poor survival of U.S. gastric carcinoma patients treated with gastrectomy: fifth edition American joint committee on cancer staging, proximal disease, and the "different disease" hypothesis. Cancer 2000;88:921–32.
4. Green D, Ponce de Leon S, Leon-Rodriguez E, Sosa-Sanchez R. Adenocarcinoma of the stomach: univariate and multivariate analysis of factors associated with survival. Am J Clin Oncol 2002;25:84–9.
5. Msika S, Benhamiche AM, Jouve JL, Rat P, Faivre J. Prognostic factors after curative resection for gastric cancer. A population-based study. Eur J Cancer 2000;36:390–6.
6. Harrison LE, Karpeh MS, Brennan MF. Extended lymphadenectomy is associated with a survival benefit for node-negative gastric cancer. J Gastrointest Surg 1998;2:126–31.
7. Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. Acta Pathol Microbiol Scand 1965;64:31–49.
8. Werner M, Becker KF, Keller G, Hofler H. Gastric adenocarcinoma: pathomorphology and molecular pathology. J Cancer Res Clin Oncol 2001;127:207–16.
9. Fenoglio-Preiser C, Carneiro F, Correa P, Guilford P, Lambert R, Megraud F, Muñoz N, Powell SM, Rugge M, Sasako M, Stolte M, Watanabe H. Gastric carcinoma. In: Hamilton SR, Aaltonen LA, eds. Pathology and genetics. Tumors of the digestive system. Lyon: IARC Press, 2000:39–52.
10. Wolf M, Mousses S, Hautaniemi S, Karhu R, Huusko P, Allinen M, Elkahloun A, Monni O, Chen Y, Kallioniemi A, Kallioniemi OP. High-resolution analysis of gene copy number alterations in human prostate cancer using CGH on cDNA microarrays: impact of copy number on gene expression. Neoplasia 2004;6:240–7.
11. Hyman E, Kauraniemi P, Hautaniemi S, Wolf M, Mousses S, Rozenblum E, Ringner M, Sauter G, Monni O, Elkahloun A, Kallioniemi OP, Kallioniemi A. Impact of DNA amplification on gene expression patterns in breast cancer. Cancer Res 2002;62:6240–5.
12. Pollack JR, Perou CM, Alizadeh AA, Eisen MB, Pergamenschikov A, Williams CF, Jeffrey SS, Botstein D, Brown PO. Genome-wide analysis of DNA copy-number changes using cDNA microarrays. Nat Genet 1999;23:41–6.
13. Jarvinen AK, Autio R, Haapa-Paananen S, Wolf M, Saarela M, Grenman R, Leivo I, Kallioniemi O, Makitie AA, Monni O. Identification of target genes in laryngeal squamous cell carcinoma by high-resolution copy number and gene expression microarray analyses. Oncogene 2006;25:6997–7008.
14. Chin K, DeVries S, Fridlyand J, Spellman PT, Roydasgupta R, Kuo WL, Lapuk A, Neve RM, Qian Z, Ryder T, Chen F, Feiler H, et al. Genomic and transcriptional aberrations linked to breast cancer pathophysiologies. Cancer Cell 2006;10:529–41.
15. Lingjaerde OC, Baumbusch LO, Liestol K, Glad IK, Borresen-Dale AL. CGH-explorer: a program for analysis of array-CGH data. Bioinformatics 2005;21:821–2.
16. Swets JA. Measuring the accuracy of diagnostic systems. Science 1998;240:1285–93.
17. Duda RO, Hart PE, Stork DG. Pattern Classification, 2nd edn. New york: John Wiley & Sons, 2001.
18. Hautaniemi S, Ringnèr M, Kauraniemi P, Autio R, Edgren H, Yli-Harja O, Astola J, Kallioniemi A, Kallioniemi O-P. A strategy for identifying putative causes of gene expression variation in human cancers. J Franklin Institute 2004;341:77–88.
19. Varis A, Zaika A, Puolakkainen P, Nagy B, Madrigal I, Kokkola A, Vayrynen A, Karkkainen P, Moskaluk C, El-Rifai W, Knuutila S. Coamplified and overexpressed genes at ERBB2 locus in gastric cancer. Int J Cancer 2004;109:548–53.
20. Belkhiri A, Zaika A, Pidkovka N, Knuutila S, Moskaluk C, El-Rifai W. Darpp-32: a novel antiapoptotic gene in upper gastrointestinal carcinomas. Cancer Res 2005;65:6583–92.
21. Katoh M, Katoh M. Evolutionary recombination hotspot around GSDML-GSDM locus is closely linked to the oncogenomic recombination hotspot around the PPP1R1B-ERBB2-GRB7 amplicon. Int J Oncol 2004;24:757–63.
22. Tanner M, Hollmen M, Junttila TT, Kapanen AI, Tommola S, Soini Y, Helin H, Salo J, Joensuu H, Sihvo E, Elenius K, Isola J. Amplification of HER-2 in gastric carcinoma: association with Topoisomerase IIα gene amplification, intestinal type, poor prognosis and sensitivity to trastuzumab. Ann Oncol 2005;16:273–8.
23. Maqani N, Belkhiri A, Moskaluk C, Knuutila S, Dar AA, El-Rifai W. Molecular dissection of 17q12 amplicon in upper gastrointestinal adenocarcinomas. Mol Cancer Res 2006;4:449–55.
24. Varis A, Wolf M, Monni O, Vakkari ML, Kokkola A, Moskaluk C, Frierson H, Jr., Powell SM, Knuutila S, Kallioniemi A, El-Rifai W. Targets of gene amplification and overexpression at 17q in gastric cancer. Cancer Res 2002;62:2625–9.
25. Uen YH, Lin SR, Wu CH, Hsieh JS, Lu CY, Yu FJ, Huang TJ, Wang JY. Clinical significance of MUC1 and c-Met RT-PCR detection of circulating tumor cells in patients with gastric carcinoma. Clin Chim Acta 2006;367:55–61.
26. Ozdemir F, Akdogan R, Aydin F, Reis A, Kavgaci H, Gul S, Akdogan E. The effects of VEGF and VEGFR-2 on survival in patients with gastric cancer. J Exp Clin Cancer Res 2006;25:83–8.
27. Burbano RR, Assumpcao PP, Leal MF, Calcagno DQ, Guimaraes AC, Khayat AS, Takeno SS, Chen ES, De Arruda Cardoso Smith M. C-MYC locus amplification as metastasis predictor in intestinal-type gastric adenocarcinomas: CGH study in Brazil. Anticancer Res 2006;26:2909–14.
28. Knuutila S, Bjorkqvist AM, Autio K, Tarkkanen M, Wolf M, Monni O, Szymanska J, Larramendy ML, Tapper J, Pere H, El-Rifai W, Hemmer S, et al. DNA copy number amplifications in human neoplasms: review of comparative genomic hybridization studies. Am J Pathol 1998;152:1107–23.
29. Kokkola A, Monni O, Puolakkainen P, Larramendy ML, Victorzon M, Nordling S, Haapiainen R, Kivilaakso E, Knuutila S. 17q12-21

amplicon, a novel recurrent genetic change in intestinal type of gastric carcinoma: a comparative genomic hybridization study. Genes Chromosomes Cancer 1997;20:38–43.
30. Vauhkonen H, Vauhkonen M, Sajantila A, Sipponen P, Knuutila S. DNA copy number aberrations in intestinal-type gastric cancer revealed by array-based comparative genomic hybridization. Cancer Genet Cytogenet 2006;167:150–4.
31. Carvalho B, Buffart TE, Reis RM, Mons T, Moutinho C, Silva P, van Grieken NC, Grabsch H, van de Velde CJ, Ylstra B, Meijer GA, Carneiro F. Mixed gastric carcinomas show similar chromosomal aberrations in both their diffuse and glandular components. Cell Oncol 2006;28:283–94.
32. Myllykangas S, Himberg J, Bohling T, Nagy B, Hollmen J, Knuutila S. DNA copy number amplification profiling of human neoplasms. Oncogene 2006;25:7324–32.
33. Sell S. Stem cell origin of cancer and differentiation therapy. Crit Rev Oncol Hematol 2004;51:1–28.
34. Norsett KG, Laegreid A, Midelfart H, Yadetie F, Erlandsen SE, Falkmer S, Gronbech JE, Waldum HL, Komorowski J, Sandvik AK. Gene expression based classification of gastric carcinoma. Cancer Lett 2004;210:227–37.
35. Jinawath N, Furukawa Y, Hasegawa S, Li M, Tsunoda T, Satoh S, Yamaguchi T, Imamura H, Inoue M, Shiozaki H, Nakamura Y. Comparison of gene-expression profiles between diffuse- and intestinal-type gastric cancers using a genome-wide cDNA microarray. Oncogene 2004;23:6830–44.
36. Kim B, Bang S, Lee S, Kim S, Jung Y, Lee C, Choi K, Lee SG, Lee K, Lee Y, Kim SS, Yeom YI, et al. Expression profiling and subtype-specific expression of stomach cancer. Cancer Res 2003; 63:8248–55.
37. Weiss MM, Kuipers EJ, Postma C, Snijders AM, Pinkel D, Meuwissen SG, Albertson D, Meijer GA. Genomic alterations in primary gastric adenocarcinomas correlate with clinicopathological characteristics and survival. Cell Oncol 2004;26:307–17.
38. Yang S, Jeung HC, Jeong HJ, Choi YH, Kim JE, Jung JJ, Rha SY, Yang WI, Chung HC. Identification of genes with correlated patterns of variations in DNA copy number and gene expression level in gastric cancer. Genomics 2007;89:451–9.
39. Zhu Y, Qi C, Jain S, Le Beau MM, Espinosa R, III, Atkins GB, Lazar MA, Yeldandi AV, Rao MS, Reddy JK. Amplification and overexpression of peroxisome proliferator-activated receptor binding protein (PBP/PPARBP) gene in breast cancer. Proc Natl Acad Sci U S A 1999;96:10848–53.
40. Yamada T, Koyama T, Ohwada S, Tago K, Sakamoto I, Yoshimura S, Hamada K, Takeyoshi I, Morishita Y. Frameshift mutations in the MBD4/MED1 gene in primary gastric cancer with high-frequency microsatellite instability. Cancer Lett 2002;181:115–20.
41. Shen TL, Guan JL. Grb7 in intracellular signaling and its role in cell regulation. Front Biosci 2004;9:192–200.
42. Zhurinsky J, Shtutman M, Ben-Ze'ev A. Plakoglobin and β-catenin: protein interactions, regulation and biological roles. J Cell Sci 2000; 113:3127–39.
43. Gong SJ, Jin CJ, Rha SY, Chung HC. Growth inhibitory effects of trastuzumab and chemotherapeutic drugs in gastric cancer cell lines. Cancer Lett 2004;214:215–24.
44. Rebischung C, Barnoud R, Stefani L, Faucheron JL, Mousseau M. The effectiveness of trastuzumab (Herceptin) combined with chemotherapy for gastric carcinoma with overexpression of the c-erbB-2 protein. Gastric Cancer 2005;8:249–52.
45. Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell 2004;118:555–66.
46. Schwab M. Amplification of oncogenes in human cancer cells. Bioessays 1998;20:473–9.
47. Li Y, Ren J, Yu W, Li Q, Kuwahara H, Yin L, Carraway KL, III, Kufe D. The epidermal growth factor receptor regulates interaction of the human DF3/MUC1 carcinoma antigen with c-Src and β-catenin. J Biol Chem 2001;276:35239–42.
48. Schroeder JA, Thompson MC, Gardner MM, Gendler SJ. Transgenic MUC1 interacts with epidermal growth factor receptor and correlates with mitogen-activated protein kinase activation in the mouse mammary gland. J Biol Chem 2001;276:13057–64.
49. Yamamoto M, Bharti A, Li Y, Kufe D. Interaction of the DF3/MUC1 breast carcinoma-associated antigen and β-catenin in cell adhesion. J Biol Chem 1997;272:12492–4.
50. Vinall LE, King M, Novelli M, Green CA, Daniels G, Hilkens J, Sarner M, Swallow DM. Altered expression and allelic association of the hypervariable membrane mucin MUC1 in *Helicobacter pylori* gastritis. Gastroenterology 2002;123:41–9.
51. Franco AT, Israel DA, Washington MK, Krishna U, Fox JG, Rogers AB, Neish AS, Collier-Hyams L, Perez-Perez GI, Hatakeyama M, Whitehead R, Gaus K, et al. Activation of β-catenin by carcinogenic *Helicobacter pylori*. Proc Natl Acad Sci U S A 2005;102:10646–51.
52. Lowy AM, Clements WM, Bishop J, Kong L, Bonney T, Sisco K, Aronow B, Fenoglio-Preiser C, Groden J. β-Catenin/Wnt signaling regulates expression of the membrane type 3 matrix metalloproteinase in gastric cancer. Cancer Res 2006;66:4734–41.
53. Levi E, Klimstra DS, Andea A, Basturk O, Adsay NV. MUC1 and MUC2 in pancreatic neoplasia. J Clin Pathol 2004;57:456–62.