

Expression Profile of Malignant and Nonmalignant Lesions of Esophagus and Stomach: Differential Activity of Functional Modules Related to Inflammation and Lipid Metabolism

Luciana I. Gomes,^{1,2} Gustavo H. Esteves,³ Alex F. Carvalho,¹ Elier B. Cristo,³ Roberto Hirata, Jr.,³ Waleska K. Martins,^{1,2} Sarah M. Marques,¹ Luiz P. Camargo,² Helena Brentani,² Adriane Pelosof,² Cláudia Zitron,² Rubens A. Sallum,² André Montagnini,² Fernando A. Soares,² E. Jordão Neves,³ and Luiz F.L. Reis^{1,2}

¹Ludwig Institute for Cancer Research; ²Hospital do Câncer A.C. Camargo; and ³Instituto de Matemática e Estatística da Universidade de São Paulo, São Paulo, Brazil

Abstract

Adenocarcinomas of stomach and esophagus are frequently associated with preceding inflammatory alterations of the normal mucosa. Whereas intestinal metaplasia of the gastric mucosa is associated with higher risk of malignization, Barrett's disease is a risk factor for adenocarcinoma of the esophagus. Barrett's disease is characterized by the substitution of the squamous mucosa of the esophagus by a columnar tissue classified histopathologically as intestinal metaplasia. Using cDNA microarrays, we determined the expression profile of normal gastric and esophageal mucosa as well as intestinal metaplasia and adenocarcinomas from both organs. Data were explored to define functional alterations related to the transformation from squamous to columnar epithelium and the malignant transformation from intestinal metaplasia to adenocarcinomas. Based on their expression profile, adenocarcinomas of the esophagus showed stronger correlation with intestinal metaplasia of the stomach than with Barrett's mucosa. Second, we identified two functional modules, lipid metabolism and cytokine, as being altered with higher statistical significance. Whereas the lipid metabolism module is active in samples representing intestinal metaplasia and inactive in adenocarcinomas, the cytokine module is inactive in samples representing normal esophagus and esophagitis. Using the concept of relevance networks, we determined the changes in linear correlation of genes pertaining to these two functional modules. Exploitation of the data presented herein will help in the precise molecular characterization of adenocarcinoma from the distal esophagus, avoiding the topographical and descriptive classification that is currently adopted, and help with the proper management of patients with Barrett's disease. (Cancer Res 2005; 65(16): 7127-36)

Note: Supplementary data for this article are available at http://www.lbc.ludwig.org.br/adenogest_project/. The raw data from hybridizations and experimental conditions can be obtained at Gene Expression Omnibus (<http://www.ncbi.nlm.nih.gov/geo>) under accession no. GSE2444. For detailed description of the 4.8K array, the accession no. is GPL1930.

Requests for reprints: Luiz F.L. Reis, Ludwig Institute for Cancer Research and Hospital do Câncer A.C. Camargo, Rua Professor Antonio Prudente 109 4th Floor, São Paulo, Brazil 01509-010. Phone: 55-11-33883228; Fax: 55-11-32077001; E-mail: lreis@ludwig.org.br.

©2005 American Association for Cancer Research.
doi:10.1158/0008-5472.CAN-05-1035

Introduction

Although alterations in tumor suppressors and oncogenes underlie the cell-autonomous defects that are characteristic of cancer, cross-talk between normal and neoplastic cells is increasingly recognized to influence various stages of carcinogenesis (1). Immune cells constitute a prominent component of the stroma and a functional link between chronic inflammation and cancer has also long been suspected for the development of many human tumors, including those affecting the liver, esophagus, stomach, large intestine, and urinary bladder (2). Unresolved inflammation elicits cell turnover in an effort to restore tissue homeostasis, which, together with carcinogen or phagocyte-induced DNA damage, can eventually culminate in cellular transformation. In this process, deregulated cytokine production and aberrant cytokine signaling can lead to altered cell growth, differentiation, and apoptosis (1, 3).

Adenocarcinomas of the distal esophagus and gastroesophageal junction (GEJ) have increased in Western countries over the last two decades. They are normally detected at an advanced stage and the patients have a correspondingly poor prognosis (4). Chronic symptomatic gastroesophageal reflux and Barrett's esophagus are generally regarded to be important risk factors (5). In Barrett's esophagus, squamous epithelium damaged by reflux esophagitis is replaced by a metaplastic intestinal-type epithelium that is predisposed to malignancy (6). It is estimated that Barrett's esophagus is found in 10% to 16% of patients undergoing endoscopy for symptoms of gastroesophageal reflux disease and 1% to 3% of unselected patient populations undergoing endoscopy (7, 8). It carries a 30- to 125-fold increased risk for developing adenocarcinoma, with best estimates of cancer incidence of ~0.5% to 1.0% per year (9), highlighting the need for a clearer understanding of Barrett's metaplasia and the factors involved in tumor progression.

Epidemiologic data have shown a strong association between chronic inflammation with metaplasia, such as Barrett's esophagus and *Helicobacter pylori*-associated atrophic gastritis, and subsequent progression to neoplasia (10–12). The severe and prolonged reflux of gastric (acid, pepsin, and mucous) and duodenal (bile salts, trypsin, cholesterol, and lipase) contents can cause injury of the esophageal mucosa, initiating chronic inflammation and esophagitis. Inflammation in Barrett's tissue can lead to an increase in oxidative stress, and if the level of reactive oxygen species (ROS) exceeds the antioxidant and DNA repair capacity of the tissue, the resultant oxidative stress could lead to promutagenic

DNA damage, including DNA adducts, strand breaks, and other lesions (13).

Gastric adenocarcinoma is the second leading cause of cancer-related death in the world (14). Epidemiologic studies have associated *H. pylori* infection with peptic ulcers, non-Hodgkin's lymphoma of the stomach, gastric atrophy, and distal gastric adenocarcinoma (15). The pathway from gastritis to gastric atrophy, dysplasia, and carcinoma is thought to be a multistep process probably triggered by free radicals within the gastric epithelium and increased exposure to luminal carcinogens (12). However, the risk for gastric cancer seems to be related to *H. pylori* genetic elements, including the *cag* pathogenicity island, the *VACA* gene, and the *BABA2* gene; to the inflammatory responses governed by host genetics; and to specific interactions between host and microbial determinants (15).

A promising approach to exploit data generated by microarray is the idea that alterations in gene expression might manifest at the level of biological pathways or coregulated gene sets that share a common function rather than individual genes (16–19). A straightforward strategy for identifying the expression level of a set of genes corresponding to a specific pathway promises an integrated understanding of the process being studied, although the regulatory mechanisms of a cell are far from transparent in these data (19). Here, we describe the expression profile of normal

and diseased tissues from the stomach and the esophagus and the identification of genes with altered expression in the different tissues under investigation and determine their interactions in biological processes.

Materials and Methods

Patients and tissue samples. Patients were recruited at Hospital do Câncer A.C. Camargo (São Paulo, Brazil) during a 4-year period (2001-2004). All patients signed a preinformed consent and the study was approved by our institutional review board. Tissue samples were snap frozen in liquid nitrogen (when obtained by surgery) or collected in *RNAlater* (Ambion, Austin, TX). At the time of RNA extraction, diagnosis was confirmed by H&E staining. Frozen samples were hand dissected for removal of infiltrating inflammatory cells and enriched to have only intestinal metaplasia or Barrett's mucosa or, in the case of adenocarcinomas, at least 70% of tumor cells. A total of 71 samples were analyzed: 39 esophagus and GEJ samples [9 normal esophageal mucosa, 6 esophagitis mucosa, 10 Barrett's mucosa (4 for the long type and 6 for the short type), 5 adenocarcinomas of the esophagus, and 9 adenocarcinomas of the GEJ] and 32 stomach samples (6 normal body and antrum mucosa, 5 normal cardiac mucosa, 9 intestinal metaplasia mucosa, 7 samples of intestinal-type adenocarcinoma, and 5 samples of diffuse-type carcinoma). Detailed descriptions of tumor samples are presented on Tables 1 and 2. Esophagitis were obtained from patients with history of reflux and defined based on histopathology. For Barrett's samples, diagnosis required evidence of intestinal metaplasia together with

Table 1. Description of samples representing tumors from distal esophagus and adenocarcinomas of the GEJ

ID	WHO	Grade	Localization	Size (cm)	Infiltration	Lymph node	TNM	Stage	Associated Barrett's metaplasia
GH852	Adenocarcinoma	Moderately differentiated	Distal esophagus and GEJ	5.5	Fat tissue	12/38	T ₃ N ₁ M ₀	III	Absent
GH857	Adenocarcinoma	Poorly differentiated	Distal esophagus and GEJ	6.0	Muscularis propria	1/25	T ₂ N ₁ M ₀	IIB	Present
GH861	Adenocarcinoma	Moderately differentiated	Distal esophagus and GEJ	4.5	Adventitia	6/29	T ₃ N ₁ M _{1b}	III	Absent
GH865	Adenocarcinoma	Poorly differentiated	Distal esophagus and GEJ	6.0	Fat tissue	30/43	T ₃ N ₁ M _{1b}	IVB	Present
GH871	Adenocarcinoma	Moderately differentiated	GEJ	3.3	Submucosa	1/4	T ₁ N ₁ M ₀	IIB	Present
GF44	Pylorocardiac carcinoma	Moderately differentiated	Distal esophagus, GEJ, and proximal stomach	9.2	Muscularis propria	3/11	T ₂ N ₁ M ₀	IIB	Absent
HC01	Adenocarcinoma	Poorly differentiated	Middle-distal esophagus and GEJ	6.0	Muscularis propria	10/24	T ₂ N ₁ M ₀	IIB	Present
BIO465B	Adenocarcinoma	Poorly differentiated	Distal esophagus and GEJ	NA	NA	NA	NA	NA	Present
HC03	Adenocarcinoma	Poorly differentiated	Distal esophagus and GEJ	6.5	Adventitia	5/55	T ₃ N ₁ M ₀	III	Absent
HC05	Adenocarcinoma	Moderately differentiated	Distal esophagus, GEJ, and proximal stomach	6.0	Muscularis propria	0/46	T ₂ N ₀ M ₀	IIA	Absent
HC07	Adenocarcinoma	Moderately differentiated	GEJ	7.0	Fat tissue	0/82	T ₃ N ₀ M ₀	IIA	Absent
HC09	Adenocarcinoma	Well differentiated	GEJ	1.0	Muscularis propria	0/25	T ₂ N ₀ M ₀	IIA	Absent
BH01	Adenocarcinoma	Moderately differentiated	GEJ	5.5	Adventitia	3/17	T ₃ N ₁ M ₀	III	Absent
CP03	Adenocarcinoma	Moderately differentiated	GEJ	NA	NA	NA	NA	NA	NA

NOTE: TNM, tumor-node-metastasis; NA, not applicable. We used the TNM classification for esophageal carcinomas.

Table 2. Description of samples representing gastric tumor

ID	WHO	Laurén	Grade	Size (cm)	Borrmann	Infiltration	Lymph node	TNM
GF48	Signet-ring cell carcinoma	Diffuse	Poorly differentiated	7.5	IV	Serosa	0/7	T ₂ N ₀ M ₁
GF62	Signet-ring cell carcinoma	Diffuse	Poorly differentiated	3.5	III	Serosa	1/40	T ₃ N ₁ M ₀
GF68	Signet-ring cell carcinoma	Diffuse	Poorly differentiated	6.0	II	Serosa	9/40	T ₄ N ₂ M ₀
GH877	Signet-ring cell carcinoma	Diffuse	Poorly differentiated	7.5	II	Serosa	7/44	T ₃ N ₂ M ₀
GH977	Signet-ring cell carcinoma	Diffuse	Poorly differentiated	6.5	IV	Muscularis propria	13/23	T ₂ N ₂ M ₀
GF50	Tubular carcinoma	Intestinal	Moderately differentiated	10.0	III	Fat tissue	3/42	T ₄ N ₁ M ₀
GF54	Tubular carcinoma	Intestinal	Poorly differentiated	4.5	III	Fat tissue	2/22	T ₃ N ₁ M ₀
GF70	Tubular carcinoma	Intestinal	Moderately differentiated	11.8	III	Serosa	0/42	T ₃ N ₀ M ₀
GH831	Mucinous adenocarcinoma	Intestinal	Moderately differentiated	11.0	II	Fat tissue	06/18	T ₄ N ₁ M ₁
GH843	Mucinous adenocarcinoma	Intestinal	Well differentiated	8.0	II	Fat tissue	10/37	T ₃ N ₂ M ₀
GH969	Papillary adenocarcinoma	Intestinal	Well differentiated	7.0	II	Serosa	11/18	T ₄ N ₂ M ₁
GH973	Tubular adenocarcinoma	Intestinal	Moderately differentiated	14.0	II	Serosa	20/41	T ₃ N ₃ M ₀

tongues or segments of red columnar-appearing epithelium extending upwards from the GEJ. To define a lesion as adenocarcinomas of the esophagus, the bulk of the lesion was above the GEJ with associated Barrett's mucosa in the nonneoplastic tissue. Adenocarcinoma of the GEJ was defined as a tumor involving esophagus and GEJ without Barrett's mucosa. For the gastric adenocarcinoma samples, the Laurén's classification was used (20).

Extraction, amplification, and labeling of the antisense RNAs. Total RNA was extracted using TRIzol (Life Technologies, Inc., Grand Island, NY) and amplified by a T7-based protocol (21). As reference RNA, we used a pool of total RNA from disease-free fragment of esophageal mucosa. For replica hybridizations with dye swap, amplified RNA (3-8 µg) was added to synthetic antisense RNAs corresponding to internal controls and labeled with either Cy3-dTCP and Cy5-dCTP (Amersham Biosciences, Piscataway, NJ) using random primer and SuperScript II (Invitrogen, Carlsbad, CA). The quality of dye incorporation was determined as described elsewhere (22).

Hybridization and scanning of cDNA microarray. Glass arrays containing 4,800 cDNA sequences were prepared in our laboratory with the aid of the Flexys robot (Genomic Solutions, Cambridgeshire, United Kingdom; ref. 23). Prehybridization, hybridization, and washing were done as described previously (21) and slides were scanned on a laser scanner (ScanArray Express, Perkin-Elmer Life Sciences, Boston, MA). Data were extracted with ScanArray Express software using the histogram method.

Statistical analysis. For data analysis, we used R,⁴ an open source interpreted computer language for statistical computation and graphics and for tools from the Bioconductor project.⁵ After image acquisition and quantification, spots with signal lower or equal to background were excluded from the analysis. Background-subtracted spot intensities were normalized by loess using span equal to 0.4 and degree equal to 2.

For the identification of differentially expressed genes, a nonparametric test (Mann-Whitney) was applied to determine the *P* for each individual gene in each pair-wise tissue comparison. For clustering samples based on their expression profile, we applied hierarchical clustering based on correlation distance and complete linkage. Once clusters were obtained, genes were organized hierarchically, based on their correlation distances.

Next, we searched gene expression patterns associated to disease types but which are not necessarily associated to individual sample (19). For that, we defined functional modules according to the Kyoto Encyclopedia of Genes and Genomes.⁶ A gene in a sample is declared induced if its expression is >2-fold the average expression in all samples and declared

repressed if its expression is <2-fold below the same average. Next, we identify the modules with a larger than expected number of induced or repressed genes (19). For the corresponding hypothesis test, the number of induced and repressed genes within each module would, under the null hypothesis, have hypergeometric distribution. A 5% false discovery rate is used to identify those active and repressed modules.

For the identification of genes with altered correlation in their expression profile, we applied the concept of relevance networks. We made comprehensive pair-wise comparisons of genes within each functional module and determined their correlation coefficients. We then constructed a matrix in a heat map representation where green represents correlation of -1, red represents correlation of +1, and black indicates lack of correlation. To determine the statistical significance of the observed changes in correlation coefficients, when two distinct classes of samples were compared, we first apply the Fisher's *z* transformation, where the correlation distance *r* is transformed into a new variable denoted as *z* (24) and the hypothesis test on the correlation is then applied with the new *z* variable.

Results

Identification of differentially expressed genes. To identify the genes with differential expression among tissues representing the various pathologies, we did pair-wise comparisons and selected the 50 genes with lowest *P*s for each comparison leading to a nonredundant set of 958 genes. Based on the expression profile of these genes, samples were hierarchically clustered (Fig. 1A). Two main branches representing squamous and columnar tissues were obtained. The branch containing columnar tissues could be further divided into two groups, representing malignant and nonmalignant samples. The list of the 958 genes represented in Fig. 1A can be obtained as Supplementary Data.

Our data show that, at least at the molecular level, adenocarcinoma of the GEJ and adenocarcinoma of the esophagus have similar expression profile, with samples dispersed within the branch representing adenocarcinomas in the hierarchical clustering presented in Fig. 1A.

Based on the observed distinction between squamous and columnar tissues, we then identified the genes whose altered expression could best distinguish between these two groups of samples. Among the 30 genes with lowest *P*s for this comparison (Fig. 1B), we identified a group of genes whose expression pattern are well-established markers for either squamous (*KRT4*, *PPL*, and *KLK13*) or columnar (*TM4SF3*, *MUC13*, *CITED1*, and *CHK*) tissues.

⁴ <http://cran.r-project.org>

⁵ <http://www.bioconductor.org>

⁶ <http://www.kegg.org>

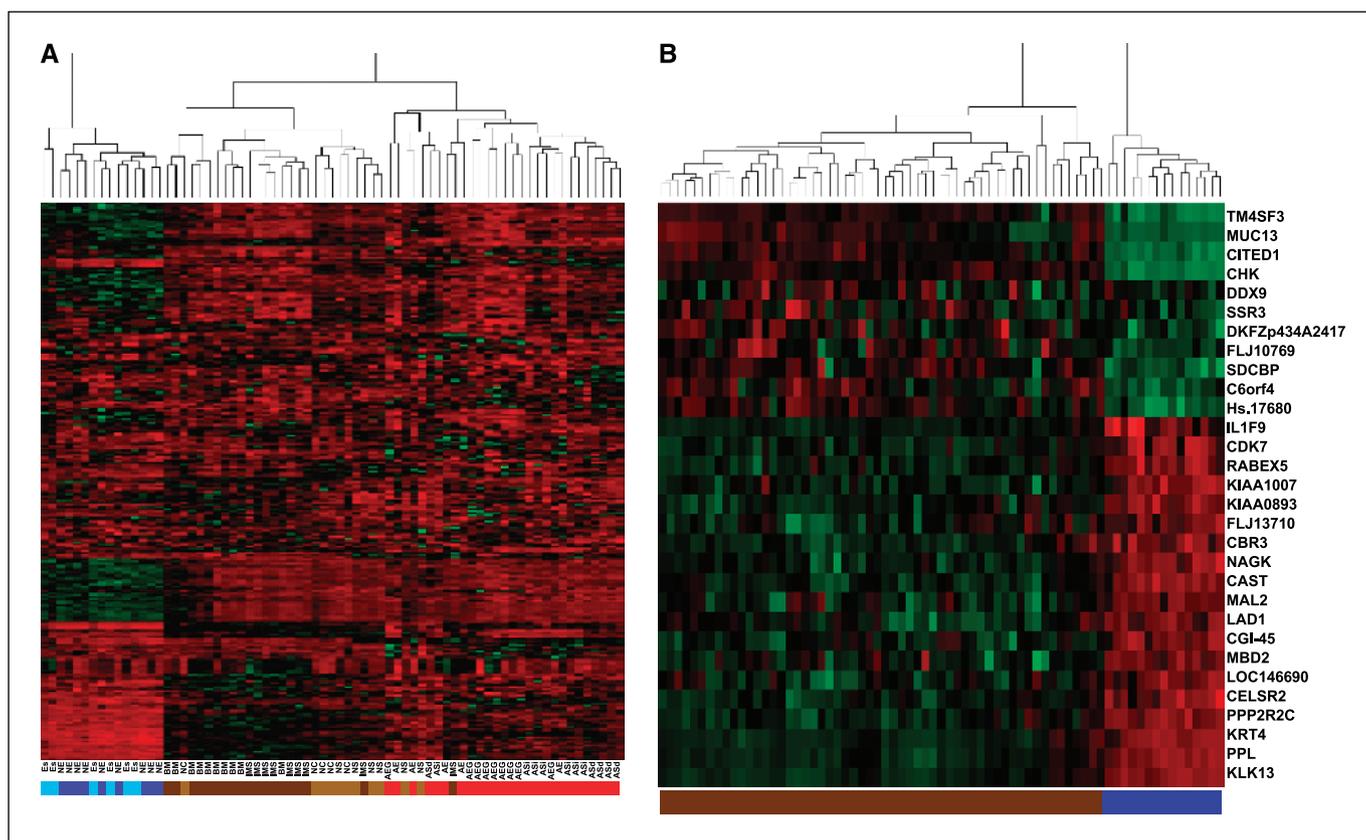


Figure 1. Hierarchical clustering of samples from stomach and esophagus. After selecting the top 50 genes with lowest P s for every pair-wise comparison, samples were clustered hierarchically using correlation distance (A). Samples are identified as normal esophagus (NE), esophagitis (Es), normal stomach (NS), normal cardia (NC), intestinal metaplasia of the stomach (IMS), Barrett's mucosa (BM), adenocarcinomas of the esophagus (AE), adenocarcinoma of the GEJ (AGEJ), adenocarcinoma of the stomach intestinal type (ASi), and adenocarcinoma of the stomach diffuse type (ASd). Throughout the article, the color code below sample identification will be used. B, the top 30 genes with lowest P s for squamous versus columnar comparison were selected and samples were clustered hierarchically using correlation distance. Blue bars, squamous tissues; brown bars, columnar tissues.

Considering the controversy regarding the nomenclature and classification of adenocarcinoma of the esophagus, we determined the linear correlation between adenocarcinomas of the esophagus with either intestinal metaplasia of the stomach or Barrett's mucosa. We observed a stronger correlation between adenocarcinomas of the esophagus and intestinal metaplasia of the stomach ($R^2 = 0.9618$) than with Barrett's mucosa ($R^2 = 0.9477$), with $P < 0.0001$ (figure available as Supplementary Data).

Analyses of functional modules. To analyze the impact of altered gene expression in functional modules, we first grouped the genes present in our array according to Kyoto Encyclopedia of Genes and Genomes and 17 modules were represented and listed as Supplementary Data. Two modules showed alterations with higher statistical significance: the process of lipid metabolism (the glycerolipid metabolism module) and the cytokine-cytokine receptor interaction module (figures available as Supplementary Data).

The glycerolipid metabolism module was active in 6 of 10 (60%) Barrett's mucosa samples and 5 of 9 (55.5%) stomach intestinal metaplasia samples. It was inactive in 7 of 14 (50%) samples of adenocarcinoma of the esophagus and GEJ and 3 of 7 (43%) intestinal type of gastric adenocarcinoma samples. The data describing the scores for each gene in each functional module as well as the status of each functional module in each sample group can be obtained at the Web site with Supplementary Data.

We used the nonsupervised K -means algorithm to group samples according to the expression profile of the genes belonging to either module. For both modules, glycerolipid (Fig. 2) or cytokine (figure available as Supplementary Data), there were a near-perfect separation of samples when we used $K = 3$, with one cluster representing normal squamous tissues (blue bar), one cluster representing nonmalignant columnar tissues (brown bar), and one cluster representing adenocarcinomas of the stomach and esophagus (red bar). In each cluster, samples and genes were ordered according to their hierarchical distance as indicated by the dendrograms.

Identifying genes responsible for activation/inactivation of functional modules. As described by Segal et al. (19), only a subset of genes belonging to a given functional module may contribute to its activation or inactivation. Hence, the identification of active/inactive status of a given module was based on the fraction of genes with high score and low P as described in Materials and Methods. For the cytokine-cytokine receptor interaction module, these genes were *IL1R2*, *CCL20*, *CCL18*, *INHBA*, *IL4R*, and *IFNAR2* (Fig. 3A), and for the glycerolipid metabolism module, they were *AKR1B10*, *ALDH3A2*, *ADH1B*, *CDS1*, and *DGKQ* (Fig. 3B).

The expression of *IL1R2* and *IFNAR2* showed a diminished expression for normal esophageal and esophagitis tissue when compared with intestinal metaplasia of the stomach and Barrett's disease ($P < 0.0001$, for both genes, Mann-Whitney). Likewise, adenocarcinomas of stomach or esophagus, including those from

GEJ, showed lower expression of *IL1R2* and *IFNAR2* when compared with tissues representing intestinal metaplasia ($P < 0.0001$). *CCL20*, *CCL18*, and *ILAR* genes showed lower expression in normal esophageal and esophagitis when compared with tissues representing intestinal metaplasia ($P = 0.001$, 0.002 , and 0.007 , respectively). The *INHBA* gene showed a distinct profile with overexpression only for adenocarcinomas tissues ($P < 0.0001$, compared with intestinal metaplasia tissues; Fig. 3A).

The genes responsible for activation of the glycerolipid metabolism module (Fig. 3B) showed lower expression in samples representing adenocarcinomas of esophagus and GEJ or intestinal-type gastric adenocarcinoma when compared with intestinal metaplasia tissues ($P < 0.0001$ for *AKR1B10*, *ALDH3A2*, *ADH1B*, and *CDS1* and $P = 0.001$ for *DGKQ*).

Relevance networks within the cytokine and glycerolipid modules. Finally, we determined the linear correlation for all pairs of genes within the cytokine and glycerolipid modules and statistical significance of the change in the linear correlation when samples representing intestinal metaplasias were compared with adenocarcinomas. In Fig. 4, we represent in a heat map, the linear correlation for genes belonging to the glycerolipid module for Barrett's mucosa (left) and for adenocarcinomas of the esophagus and GEJ (right). By following a given line, one can identify all genes that have a positive correlation (red squares) or negative correlation (green squares) with that particular gene. In Fig. 5, the heat maps represent the statistical significance of changes in correlation for the glycerolipid module when intestinal metaplasia and adenocarcinoma of the stomach of the intestinal type (right) or Barrett's disease and adenocarcinomas of the esophagus and GEJ (left) are compared. The figure representing these changes for the cytokine module as well as the raw data for linear correlation for all pairs of genes for each condition and the P s for the represented comparisons can be obtained at the Web site as Supplementary Data.

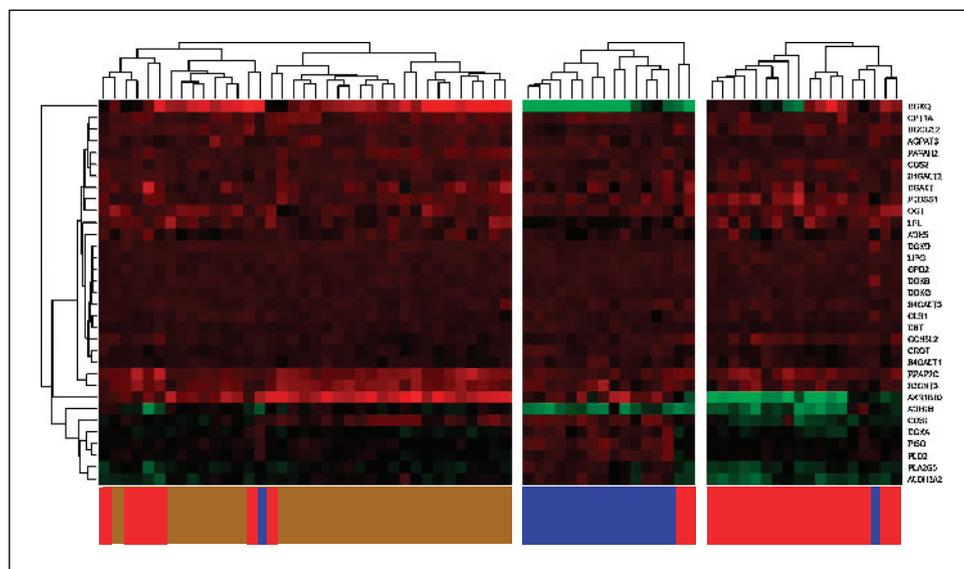
To illustrate how the data represented in Fig. 5 could be used to generate data-oriented hypothesis, we represented in Fig. 6 the changes in linear correlation of two pairs of genes pertaining to the glycerolipid metabolism when intestinal metaplasia and adenocarcinomas of the stomach (*CDS1* \times *PLD2*; Fig. 6A) or Barrett's mucosa and adenocarcinomas of the esophagus or GEJ (*CPT1A* \times *GCN5L2*; Fig. 6B) were compared.

Discussion

Intestinal metaplasia of the stomach or the distal esophagus, named Barrett's disease, are both considered as lesions with higher risk of malignization to adenocarcinoma and also both arise as a consequence of inflammatory stimuli (10–12). From a collection of 4,595 cDNA sequences representing full-length human genes, we selected a nonredundant set of 958 genes representing those with lowest P s for pair-wise comparisons. Based on the expression profiles of these genes, it was possible to discriminate two main groups of samples representing squamous and columnar tissues. Samples representing normal esophageal mucosa and esophagitis showing very similar molecular profiles were all grouped in a single branch (Fig. 1A). Recently, we described molecular signatures for gastric samples (25), and no distinction between normal gastric mucosa and gastritis also could be observed, suggesting that, apart from the presence of inflammatory cells that are removed before RNA extraction, no major changes in gene expression are triggered by the inflammatory stimuli also in the esophagus mucosa.

Classification of adenocarcinoma of the GEJ is controversial, with some authors classifying them as gastric tumors, others as esophageal tumors, and yet a third group that suggests a separate classification based on the location of the main portion of the lesion (26). Because the outcome for adenocarcinoma of the GEJ does not seem to be related to the affected area (27), we would like to suggest that the critical issue is to define its origin, which could help in defining better strategies for follow-up of patients with Barrett's disease, diagnosis, and treatment. As shown in the far right branch of Fig. 1A, all but one adenocarcinoma of the GEJ cluster together with adenocarcinomas of the stomach, mainly of the intestinal type. In contrast, adenocarcinomas classified as from the esophagus (those with the majority of the tumor mass above the GEJ and with Barrett's mucosa at its upper margin) were dispersed among the malignant diseases, indicating the absence of a common expression profile. Considering the fact that, as in the classification proposed by Rudiger et al. (27), the vast majority of adenocarcinomas affecting the distal esophagus also affect the GEJ and that, rarely, one can detect tumor-free Barrett's mucosa between the GEJ and adenocarcinomas of the esophagus, it is tempting to speculate that adenocarcinomas of

Figure 2. Nonsupervised cluster of samples based on the expression of genes pertaining to altered function modules. Using expression data from genes pertaining to the glycerolipid metabolism module, samples were clustered by the nonsupervised algorithm K -means using $K = 3$. Once clusters were obtained, samples and genes were ordered hierarchically based on correlation distance and complete linkage. Blue bars, normal esophagus or esophagitis; brown bars, columnar nonmalignant samples; red bars, adenocarcinomas.



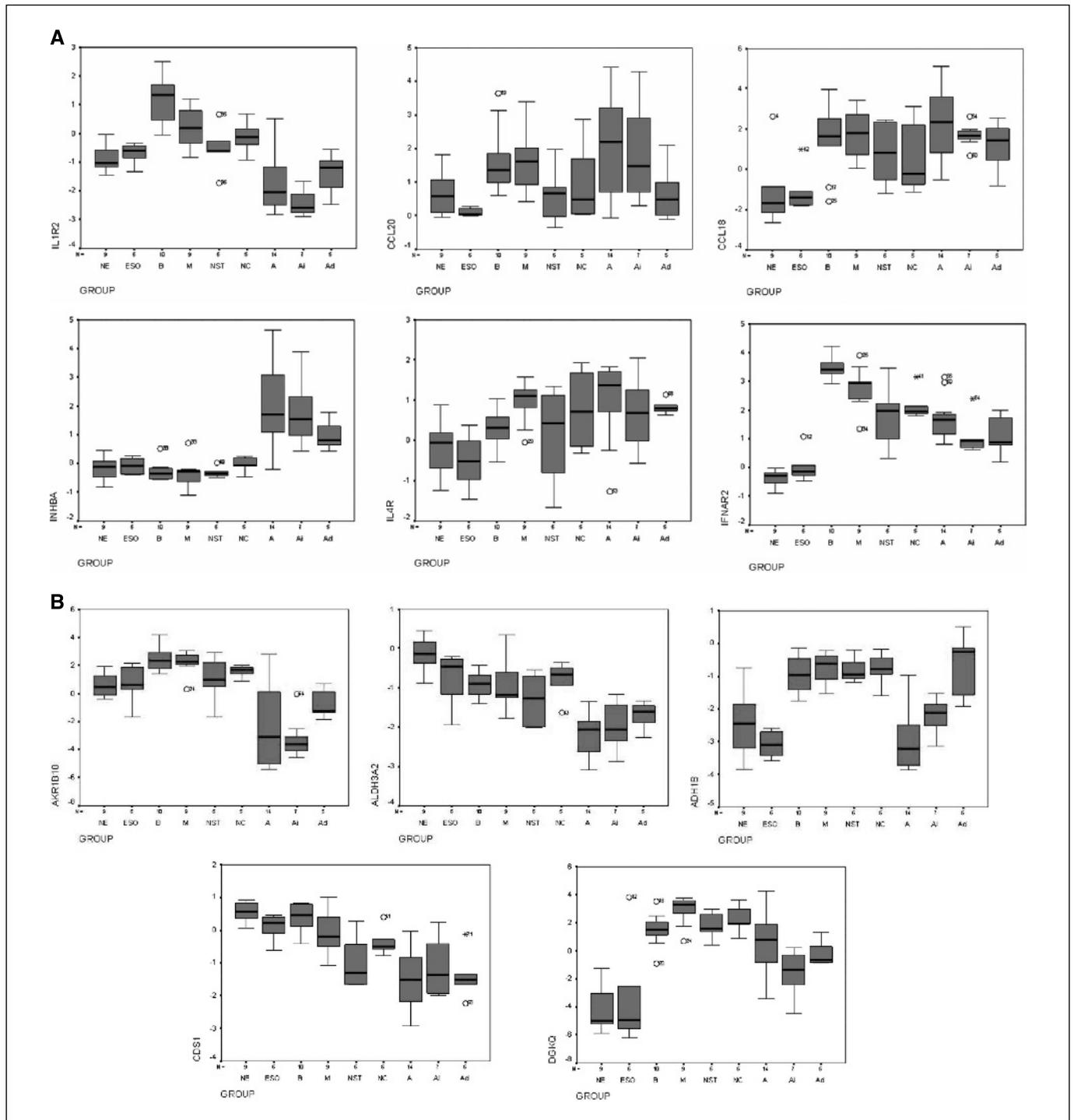


Figure 3. Identification of individual genes contributing to alterations of functional modules and their expression levels in samples of stomach and esophagus. *A*, expression profile of the six genes with highest score and lower *P*s for the cytokine module along all sample groups. *B*, expression profile of the five genes with highest score and lower *P*s for the glycerolipid metabolism module along all sample groups. Statistical significance for the observed differences was determined by Mann-Whitney.

the esophagus could originate at the cardia invading the esophagus from below, taking advantage of the rich lymphatic drainage that goes from the cardia upwards. This hypothesis is further supported by data demonstrating a stronger correlation between adenocarcinomas of the esophagus and intestinal metaplasia of the stomach than with Barrett's mucosa ($P < 0.0001$; figure available as Supplementary Data).

Considering the transition from squamous to columnar tissue, it would be helpful to determine the molecular changes that are associated with this event. As depicted in Fig. 1*B*, squamous and columnar tissues could be precisely grouped based on their differential expression profile. Among the genes, *LADI*, *KRT4*, and *PPL* showed increased expression in squamous tissues. The protein ladinin, encoded by the gene *LADI*, is a component of

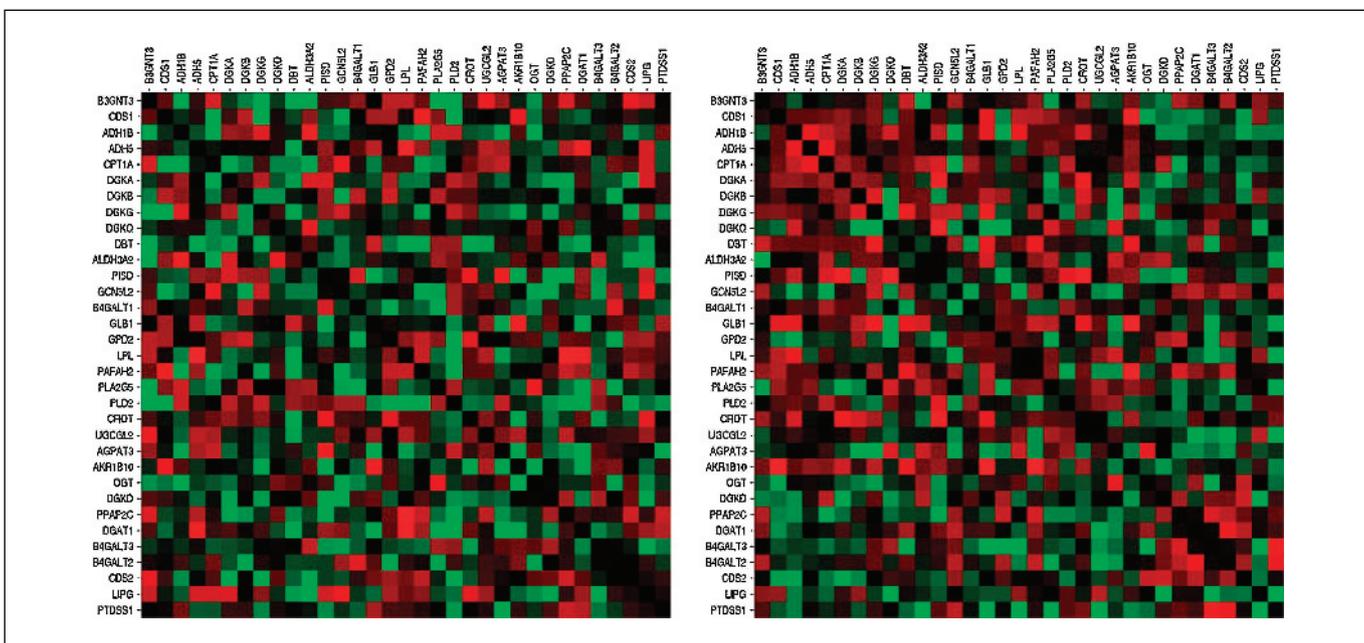


Figure 4. Heat map of relevance networks for genes pertaining to glycerolipid metabolism. The linear correlation of all pairs of genes pertaining to the functional module representing glycerolipid metabolism was determined for Barrett's mucosa (*left*) and adenocarcinomas of the esophagus and GEJ (*right*). *Red squares*, pairs of genes with positive correlation; *green squares*, pairs of genes with negative correlation; *brighter colors*, stronger correlations.

the basement membranes and may function in contributing to the stability of the association of the epithelial layers with the underlying mesenchyme (28). Cytokeratin 4, encoded by the gene *KRT4*, is expressed by the suprabasal layer of nonkeratinized squamous epithelia, such as esophagus (29). Finally, periplaklin,

encoded by the gene *PPL*, is a component of the desmosomes of keratinocytes (30).

In contrast, the gene *MUC13* was shown to be expressed at highest levels in columnar epithelia, predominantly in the gastrointestinal track with the protein localized at the apical

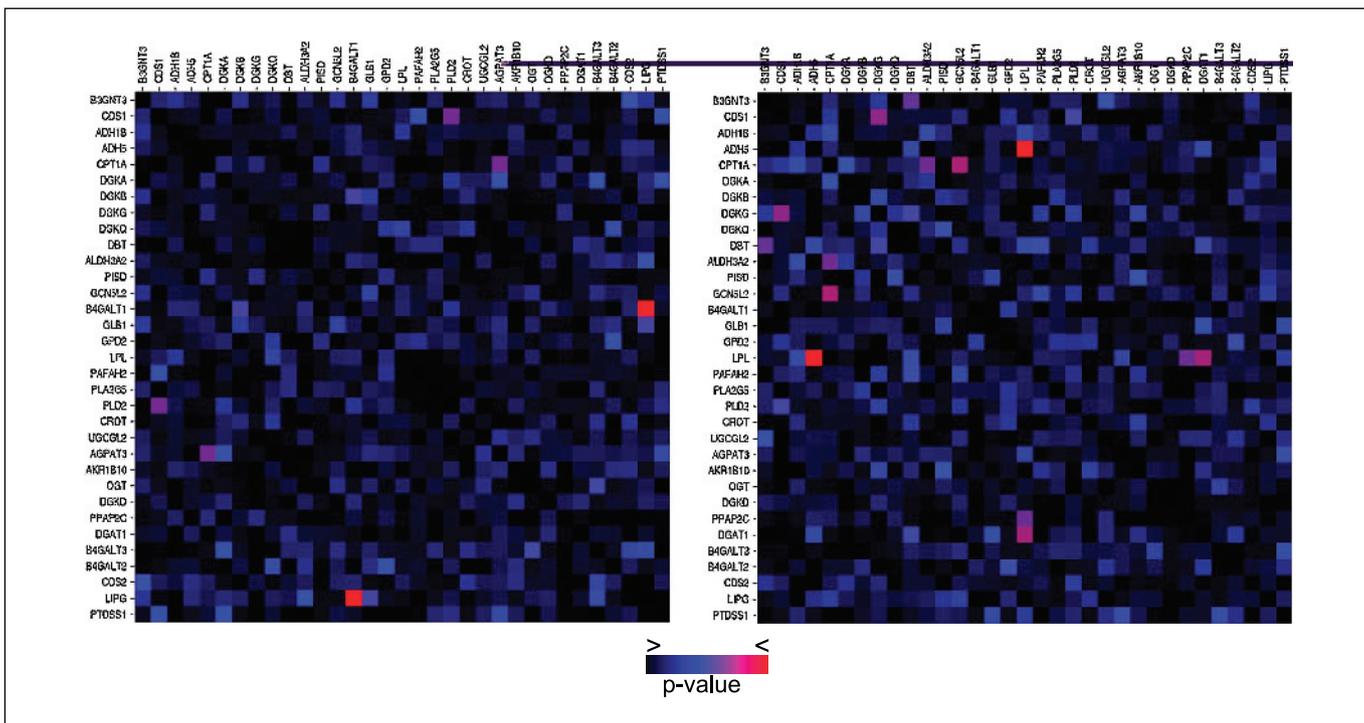


Figure 5. Heat map of relevance networks for genes with statistically significant changes in correlation when nonmalignant and malignant samples were compared. The changes in linear correlation for genes pertaining to glycerolipid metabolism were determined as described in Materials and Methods. Comparison for intestinal metaplasia of the stomach and gastric adenocarcinomas of the intestinal type is on the *left* and Barrett's mucosa and adenocarcinomas of the esophagus and GEJ on the *right*. *Black squares*, changes with lowest significance; *red squares*, changes with highest significance.

membrane of columnar cells (31). Hence, the changes in the expression pattern of the remaining genes listed in Fig. 1B deserve further investigation and might help in the understanding of the changes related to the appearance of Barrett's mucosa.

Recently, Segal et al. (19) described a method of global analysis of gene expression that allows for construction of a map representing the status of functional modules based on existing biological knowledge. Functional modules are represented by gene sets whose functions are related and their status as active or inactive is determined by the expression of that particular gene set.

Using gene sets from the Kyoto Encyclopedia of Genes and Genomes, we defined 17 functional modules represented in our array, and based on the expression profile of 71 samples tested here, we constructed heat maps representing the functional modules as being active or inactive in each individual sample as well as in sample groups. Two modules showed changes with higher statistical significance: the glycerolipid metabolism module and the cytokine-cytokine receptor interaction module.

Confirming the significance of changes in the expression of genes pertaining to these two modules, we obtained a near-perfect separation of squamous, nonmalignant columnar and malignant columnar tissues when we applied a nonsupervised algorithm (*K*-means, with *K* = 3) for clustering samples based on the expression profile of these genes (Fig. 2; Supplementary Data). The genes that gave significance to the changes observed in these two modules are listed as Supplementary Data and their expression levels in the various pathologies are presented in Fig. 3A and B.

It has to be mentioned that activation of cell cycle and integrin-mediated cell adhesion modules was observed in adenocarcinomas of the GEJ and diffuse type of gastric adenocarcinoma, respectively. The hallmark of diffuse-type gastric carcinoma is the lack of glandular organization with spreading of tumor cells through the parenchyma (32–34), implying the need for extracellular matrix destruction. We (25) and others (35) showed that diffuse-type adenocarcinomas of the stomach have an increased expression of matrix metalloproteinase-2 as well as other tissue remodeling proteins.

There are several possible mechanisms by which chronic inflammation may lead to metaplasia and neoplastic change in epithelial cells (36). ROS have been identified at high levels in ulcerated gastroesophageal (36) and *H. pylori* gastritis mucosa (12). When ROS are generated in inflamed tissues, they lead to damage of macromolecules, such as DNA, proteins, and lipids, and this effect is exacerbated in rapidly dividing cells as, for example, during tissue repair (37). Moreover, ROS attack to lipids induces lipid peroxidation and subsequent alteration of mitochondrial respiratory chain. This event in turn reinforces ROS production and the generation of aldehyde 4-hydroxy-2-nonenal, which is highly reactive and stimulates additional ROS formation and DNA damage. Hence, the regulation of lipid peroxidation needs to be tightly regulated and that involves the participation of pro-oxidative and reductive enzymes (38). Thus, it is not unexpected that the glycerolipid metabolism is active in both intestinal metaplasia and Barrett's mucosa with higher expression of *AKR1B10*, *ALDH3A2*, and *ADH1B* compared with adenocarcinomas of both stomach and esophagus. Whereas the first two enzymes play a key role in detoxification of reactive aldehydes generated by food digestion (39) or during alcohol metabolism and lipid peroxidation (40), *ADH1B* is necessary for the metabolism of ethanol, retinol, and products of lipid peroxidation (41).

For the genes belonging to the module representing cytokine-cytokine receptor interaction, we observed high levels of mRNA in

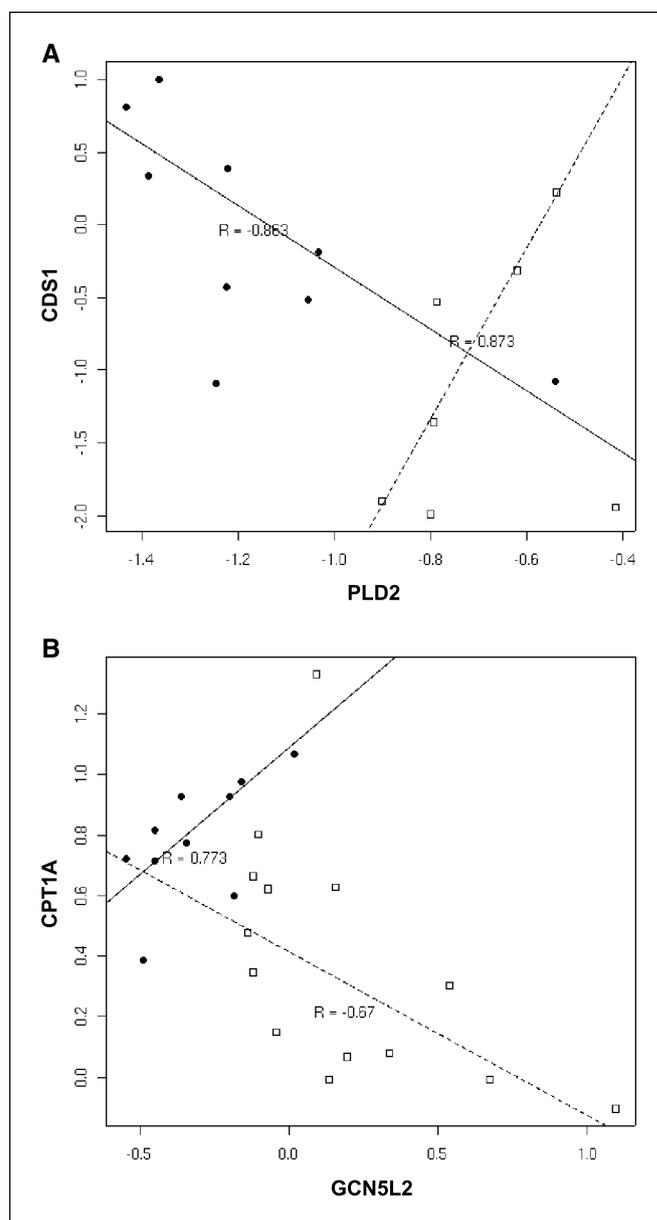


Figure 6. Changes in correlation for two pairs of genes pertaining to the glycerolipid metabolism. **A**, changes in linear correlation for *CPT1A* and *GCN5L2* in samples of Barrett's mucosa (●) and adenocarcinomas of the esophagus and GEJ (□). For Barrett's mucosa, the positive correlation is represented by the *black line* ($R = 0.773$), and for adenocarcinomas of the esophagus and GEJ, the negative correlation is represented by the *dashed line* ($R = -0.670$). **B**, changes in linear correlation for *CDS1* and *PLD2* in samples of intestinal metaplasia of the stomach (●) and intestinal-type adenocarcinomas of the stomach (□). For intestinal metaplasia, the negative correlation is represented by the *black line* ($R = -0.863$), and for adenocarcinomas of the stomach, the positive correlation is represented by the *dashed line* ($R = 0.873$).

samples representing intestinal metaplasia and Barrett's mucosa with the exception of *INHBA*, which was overexpressed only in samples representing adenocarcinomas. This gene was previously reported to be overexpressed in pancreatic tumors (42) and plays a key role during development (43).

IL-1 is a proinflammatory cytokine whose expression level is associated with the risk of adenocarcinomas of the stomach. A polymorphism in the *IL1B* gene or in the *IL1RN* gene that encodes the IL-1 receptor antagonist was associated with higher IL-B

secretion leading to hypochloridria, increased gastrin secretion, atrophic gastritis, and gastric adenocarcinomas. IL-1 binds to two types of receptors on the cell membrane, the type I (IL-1R1) and type II (IL-1R2) receptors. The IL-1R2 is a natural antagonist of IL-1 because it competes with IL-1R1 for the ligand but does not trigger a functional signaling. On binding of IL-1, the IL-1R2 interacts with IL1RAcP subtracting the coreceptor molecule from the signaling complex triggered by IL-1R1 (44). Hence, the overexpression of IL-1R2 observed in samples of intestinal metaplasia seems to be a defense mechanism to compensate the increased expression of IL-1 triggered by the inflammatory stimuli; conversely, the diminished expression of IL-1R2 observed in samples representing adenocarcinomas would represent a failure in this protective mechanism, allowing for the deleterious effects of higher IL-1.

Finally, we applied the concept of relevance networks to determine statistically significant alterations in the linear correlation of genes pertaining to the two functional modules when tissue samples with intestinal metaplasia and adenocarcinomas were compared. Associations weaker than threshold strength were flagged in the heat map leaving networks of highly correlated genes. From data presented on Fig. 4 (and data presented as Supplementary Data), one can now select a small set of genes that are linked above threshold and define new hypotheses for the malignant transformation of intestinal metaplasia to adenocarcinomas.

In the examples presented in Fig. 6A and B, two pairs of genes belonging to the representing glycerolipid metabolism functional module are depicted. As shown by Obici et al. (45), *CPT1* is a key enzyme in controlling lipid oxidation, which in turn signals nutrient availability to the hypothalamus. On the other hand, it has been shown that *GCN5L2* have histone acetyltransferase activity (46). Considering the fact that tumors cells have augmented

energy consumption, reduced *CPT1* activity could be compensated by increased transcriptional activity favored by higher levels of *GCN5L2*. This kind of analysis was proven useful to identify genes that change their expression behavior when distinct biological conditions are compared regardless of the statistical differences in their level of expression. More importantly, it allows for the discovery of gene sets that might function in an orchestrated manner that would not be revealed otherwise. These changes in correlation when associated to a given biological process, such as resistance to chemotherapy (47, 48) or oncogenic transformation, as in our case, could bring new insights to the mechanism of malignant transformation of intestinal metaplasia into adenocarcinomas.

Nevertheless, specific hypotheses linking genes to a specific biological phenomenon still suffer from the lack of biological information and precise annotation of metabolic pathways compared with the amount of information regarding gene expression, and overfitting cannot be ruled out without performing the necessary specific biological experiments.

Acknowledgments

Received 3/29/2005; revised 6/8/2005; accepted 6/20/2005.

Grant support: Conselho Nacional de Desenvolvimento Científico e Tecnológico (L.F.L. Reis and E.J. Neves) and Fundação de Amparo à Pesquisa do Estado de São Paulo grants 98/14335-2 (CEPID) and 99/11962-9 and 99/07390-0 (E.J. Neves). L.I. Gomes was a predoctoral fellow from Fundação de Amparo à Pesquisa do Estado de São Paulo.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

We thank Mariana Santos and Chamberlein Neto for technical assistance, all the members of our laboratories for helpful discussions, and Dr. Ricardo Brentani for critical reading of the article. L.F.L. Reis thanks Alexandre Koch for fruitful discussions regarding pattern recognition.

References

1. Dranoff G. Cytokines in cancer pathogenesis and cancer therapy. *Nat Rev Cancer* 2004;4:11–22.
2. Coussens LM, Werb Z. Inflammation and cancer. *Nature* 2002;420:860–67.
3. Balkwill F, Coussens LM. Cancer: an inflammatory link. *Nature* 2004;431:405–6.
4. Mueller J, Werner M, Stolte M. Barrett's esophagus: histopathologic definitions and diagnostic criteria. *World J Surg* 2004;28:148–54.
5. Forman D. Review article: oesophago-gastric adenocarcinoma—an epidemiological perspective. *Aliment Pharmacol Ther* 2004;20:55–60.
6. Souza RF, Morales CP, Spechler SJ. Review article: a conceptual approach to understanding the molecular mechanisms of cancer development in Barrett's oesophagus. *Aliment Pharmacol Ther* 2001;15:1087–100.
7. Falk GW. Barrett's esophagus. *Gastroenterology* 2002; 122:1569–91.
8. Caygill CP, Watson A, Lao-Sirieix P, Fitzgerald RC. Barrett's oesophagus and adenocarcinoma. *World J Surg Oncol* 2004;2:12.
9. Wild CP, Hardie LJ. Reflux, Barrett's oesophagus and adenocarcinoma: burning questions. *Nat Rev Cancer* 2003;3:676–84.
10. Jankowski JA, Harrison RF, Perry I, Balkwill F, Tselepis C. Barrett's metaplasia. *Lancet* 2000;356:2079–85.
11. Jankowski JA, Anderson M. Review article: management of oesophageal adenocarcinoma—control of acid, bile and inflammation in intervention strategies for Barrett's oesophagus. *Aliment Pharmacol Ther* 2004;20: 71–80.
12. Farthing MJ, Fitzgerald R, Zhang ZW. Acid, *Helicobacter* and immunity: a new paradigm for oesophago-gastric cancer. *J Physiol Paris* 2001;95:423–7.
13. Olliver JR, Hardie LJ, Dexter S, Chalmers D, Wild CP. DNA damage levels are raised in Barrett's oesophageal mucosa relative to the squamous epithelium of the oesophagus. *Biomarkers* 2003;8:509–21.
14. Plummer M, Franceschi S, Munoz N. Epidemiology of gastric cancer. *IARC Sci Publ* 2004;311–26.
15. Peek RM Jr, Blaser MJ. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nat Rev Cancer* 2002;2:28–37.
16. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proc Natl Acad Sci U S A* 1998;95:14863–8.
17. Mootha VK, Lindgren CM, Eriksson KF, et al. PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat Genet* 2003;34:267–73.
18. Segal E, Shapira M, Regev A, et al. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genet* 2003;34:166–76.
19. Segal E, Friedman N, Koller D, Regev A. A module map showing conditional activity of expression modules in cancer. *Nat Genet* 2004;36:1090–8.
20. Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at histo-clinical classification. *Acta Pathol Microbiol Scand* 1965;64:31–49.
21. Gomes LI, Silva RL, Stolf BS, et al. Comparative analysis of amplified and nonamplified RNA for hybridization in cDNA microarray. *Anal Biochem* 2003; 321:244–51.
22. Carvalho AF, Neves EJ, Hirata JR, Reis LF. Glass slide based agarose gel electrophoresis for determining the efficiency of RNA amplification and cyanine dye incorporation during RNA labeling. *Biotechniques* 2003;35:688–90, 692.
23. Brentani RR, Carraro DM, Verjovski-Almeida S, et al. Gene expression arrays in cancer research: methods and applications. *Crit Rev Oncol Hematol* 2005;54:95–105.
24. Zar J. *Biostatistical analysis*. 2nd ed. Englewood Cliff, NJ:Prentice Hall;1998.
25. Meireles SI, Cristo EB, Carvalho AF, et al. Molecular classifiers for gastric cancer and nonmalignant diseases of the gastric mucosa. *Cancer Res* 2004;64:1255–65.
26. Abdalla EK, Pisters PW. Staging and preoperative evaluation of upper gastrointestinal malignancies. *Semin Oncol* 2004;31:513–29.
27. Rudiger SJ, Feith M, Werner M, Stein HJ. Adenocarcinoma of the esophagogastric junction: results of surgical therapy based on anatomical/topographic classification in 1,002 consecutive patients. *Ann Surg* 2000;232:353–61.
28. Motoki K, Megahed M, LaForgia S, Uitto J. Cloning and chromosomal mapping of mouse laminin, a novel basement membrane zone component. *Genomics* 1997; 39:323–30.
29. Leube RE, Bader BL, Bosch FX, Zimbelmann R, Achtstaetter T, Franke WW. Molecular characterization and expression of the stratification-related cytokeratins 4 and 15. *J Cell Biol* 1988;106:1249–61.
30. Ruhrberg C, Hajibagheri MA, Parry DA, Watt FM. Perioplakin, a novel component of cornified envelopes and desmosomes that belongs to the plakin family and forms complexes with envoplakin. *J Cell Biol* 1997;139: 1835–49.

31. Williams SJ, Wreschner DH, Tran M, Eyre HJ, Sutherland GR, McGuckin MA. Muc13, a novel human cell surface mucin expressed by epithelial and hemopoietic cells. *J Biol Chem* 2001;276:18327–36.
32. Shibata T, Ochiai A, Gotoh M, Machinami R, Hirohashi S. Simultaneous expression of cadherin-11 in signet-ring cell carcinoma and stromal cells of diffuse-type gastric cancer. *Cancer Lett* 1996;99:147–53.
33. Chen HC, Chu RY, Hsu PN, et al. Loss of E-cadherin expression correlates with poor differentiation and invasion into adjacent organs in gastric adenocarcinomas. *Cancer Lett* 2003;201:97–106.
34. Oliveira C, de Bruin J, Nabais S, et al. Intragenic deletion of CDH1 as the inactivating mechanism of the wild-type allele in an HDGC tumour. *Oncogene* 2004;23:2236–40.
35. Boussioutas A, Li H, Liu J, et al. Distinctive patterns of gene expression in premalignant gastric mucosa and gastric cancer. *Cancer Res* 2003;63:2569–77.
36. Eksteen JA, Scott PA, Perry I, Jankowski JA. Inflammation promotes Barrett's metaplasia and cancer: a unique role for TNF α . *Eur J Cancer Prev* 2001;10:163–6.
37. Bartsch H, Nair J. Oxidative stress and lipid peroxidation-derived DNA-lesions in inflammation driven carcinogenesis. *Cancer Detect Prev* 2004;28:385–91.
38. Cejas P, Casado E, Belda-Iniesta C, et al. Implications of oxidative stress and cell membrane lipid peroxidation in human cancer (Spain). *Cancer Causes Control* 2004;15:707–19.
39. Cao D, Fan ST, Chung SS. Identification and characterization of a novel human aldose reductase-like gene. *J Biol Chem* 1998;273:11429–35.
40. Rizzo WB, Lin Z, Carney G. Fatty aldehyde dehydrogenase: genomic structure, expression and mutation analysis in Sjogren-Larsson syndrome. *Chem Biol Interact* 2001;132:297–307.
41. Duester G, Farres J, Felder MR, et al. Recommended nomenclature for the vertebrate alcohol dehydrogenase gene family. *Biochem Pharmacol* 1999;58:389–95.
42. Cao D, Hustinx SR, Sui G, et al. Identification of novel highly expressed genes in pancreatic ductal adenocarcinomas through a Bioinformatics Analysis of Expressed Sequence Tags. *Cancer Biol Ther* 2004;3:1081–9.
43. Brown CW, Li L, Houston-Hawkins DE, Matzuk MM. Actins are critical modulators of growth and survival. *Mol Endocrinol* 2003;17:2404–17.
44. Lang D, Knop J, Wesche H, Raffetseder U, Kurrle R, Boraschi D, Martin MU. The type II IL-1 receptor interacts with the IL-1 receptor accessory protein: a novel mechanism of regulation of IL-1 responsiveness. *J Immunol* 1998;161:6871–7.
45. Obici S, Feng Z, Arduini A, Conti R, Rossetti L. Inhibition of hypothalamic carnitine palmitoyltransferase-1 decreases food intake and glucose production. *Nat Med* 2003;9:756–61.
46. Wang L, Mizzen C, Ying C, et al. Histone acetyltransferase activity is conserved between yeast and human GCN5 and is required for complementation of growth and transcriptional activation. *Mol Cell Biol* 1997;17:519–27.
47. Butte AJ, Tamayo P, Slonim D, Golub TR, Kohane IS. Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks. *Proc Natl Acad Sci U S A* 2000;97:12182–6.
48. Moriyama M, Hoshida Y, Otsuka M, et al. Relevance network between chemosensitivity and transcriptome in human hepatoma cells. *Mol Cancer Ther* 2003;2:199–205.