Universidade de São Paulo

Biblioteca Digital da Produção Intelectual - BDPI

Departamento de Psiquiatria - FM/MPS

Artigos e Materiais de Revistas Científicas - FM/MPS

2012

Identification of protein expression signatures in gastric carcinomas using clustering analysis

JOURNAL OF GASTROENTEROLOGY AND HEPATOLOGY, MALDEN, v. 27, n. 2, pp. 378-384, FEB, 2012

http://www.producao.usp.br/handle/BDPI/42596

Downloaded from: Biblioteca Digital da Produção Intelectual - BDPI, Universidade de São Paulo

Journal of Gastroenterology and Hepatology JGHF

doi:10.1111/j.1440-1746.2011.06849.x

GASTROENTEROLOGY

Identification of protein expression signatures in gastric carcinomas using clustering analysis

Maria D Begnami,* José Humberto T G Fregnani,† Helena Brentani,‡ Cesar Torres,§ Wilson Luiz Costa Jr,¶ Andre Montagnini,¶ Suely Nonogaki* and Fernando A Soares*

Departments of *Pathology and §Medical Archives and Statistics and ¶Abdominal Surgery, A C Camargo Hospital, Sao Paulo, and †Department of Gynecology, Cancer Hospital of Barretos, Barretos, and ‡Department of Psychiatry, Medical School of University of Sao Paulo, Sao Paulo, Brazil

Key words
classifier, clustering analysis, gastric carcinoma.

Accepted for publication 28 June 2011.

Correspondence
Professor Maria D Begnami, Department of Pathology, Hospital A C Camargo, Rua Professor Antonio Prudente 211, Liberdade, São Paulo 01509-900, Brazil. Email: mariadirlei@gmail.com

Abstract

Background and Aim: The identification of gastric carcinomas (GC) has traditionally been based on histomorphology. Recently, DNA microarrays have successfully been used to identify tumors through clustering of the expression profiles. Random forest clustering is widely used for tissue microarrays and other immunohistochemical data, because it handles highly-skewed tumor marker expressions well, and weighs the contribution of each marker according to its relatedness with other tumor markers. In the present study, we e identified biologically- and clinically-meaningful groups of GC by hierarchical clustering analysis of immunohistochemical protein expression.
Methods: We selected 28 proteins (p16, p27, p21, cyclin D1, cyclin A, cyclin B1, pRb, p53, c-met, c-erbB-2, vascular endothelial growth factor, transforming growth factor [TGF]-βI, TGF-βII, MutS homolog-2, bcl-2, bax, bak, bcl-x, adenomatous polyposis coli, clathrin, E-cadherin, β-catenin, mucin (MUC)1, MUC2, MUC5AC, MUC6, matrix metalloproteinase [MMP]-2, and MMP-9) to be investigated by immunohistochemistry in 482 GC. The analyses of the data were done using a random forest-clustering method.
Results: Proteins related to cell cycle, growth factor, cell motility, cell adhesion, apoptosis, and matrix remodeling were highly expressed in GC. We identified protein expressions associated with poor survival in diffuse-type GC.
Conclusions: Based on the expression analysis of 28 proteins, we identified two groups of GC that could not be explained by any clinicopathological variables, and a subgroup of long-surviving diffuse-type GC patients with a distinct molecular profile. These results provide not only a new molecular basis for understanding the biological properties of GC, but also better prediction of survival than the classic pathological grouping.

Introduction

Gastric carcinoma (GC) is caused by multiple molecular and genetic alterations that underlie the malignant transformation of gastric mucosa.[1] The main molecular and genetic alterations include microsatellite instability, inactivation of tumor suppressor genes, such as *TP53*, *p16*, adenomatous polyposis coli, Deleted in Colorectal Cancer, and fragile histidine triad, and the activation of oncogenes, such as human epidermal growth factor receptor-2, *Ras*, *c-myc*, *cyclin E*, and *K-sam*[1]. However, these genetic changes do not precisely reflect the biological nature of the tumor cells or the clinical characteristics of GC patients.

Recently, large-scale molecular technologies, such as DNA microarrays, have enabled the global monitoring of gene expression changes.[2] Several extensive global gene expression studies have already been reported for GC.[3–6] Previous studies have described the differences between gastric tumors and normal samples,[5,6] as well as the process of gastric carcinogenesis, using gene expression profiling.[3,7] This has permitted a preliminary selection of classifiers of the subtypes[7,8] and the identification of new molecular subtypes.[9] However, such high-throughput experiments often generate hundreds of candidate genes or proteins, sometimes with futile results. The cost, complexity, and interpretation of DNA microarrays are currently unsuitable for routine use in standard clinical settings. Recently, in large-scale molecular studies, formalin-fixed, embedded fixed tissue using immunohistochemistry and clustering analysis of multiple markers have been reported to be significantly correlated with patient survival.[10,11] In addition, previous results have demonstrated that protein expressions are likely to have a more direct impact on the biological behaviors of GC than expressions at the DNA and RNA levels.[12]

In the present study, we identified biologically- and clinically-meaningful groups of GC by hierarchical analysis of 28 selected

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

protein expressions, previously described to be either putative tumor suppressors or involved in GC progression using tissue microarray (TMA) and immunohistochemistry.

Methods

Specimens

A total of 482 cases of primary GC surgically treated from 1980 to 1998 were identified in the files of the Department of Pathology, A. C. Camargo Hospital (Sao Paulo, Brazil). All the cases (mean age: 62 years, range: 26–84 years; 308 males and 174 females) failed to accept radiotherapy or chemotherapy neoadjuvant. The study items included age, sex, tumor location, tumor size, and pathological tumor lymph node metastasis (pTNM) stage. The clinical characteristics and pathological findings are summarized in Table 1. Tissue slides were reviewed for histological classification, according to Lauren's classification.[13] The series included 234 intestinal type, 166 diffuse type, 22 mixed type, and 60 cases of unclassified type. Patient clinical outcome was followed up from the date of surgery up to a period of 0.6–108.6 months (mean: 28.3 months). Patients who were lost to

Table 1 Clinical characteristics and histopathological findings of 482 gastric carcinomas

Variable	Category	No. cases (%)
Sex	Male	308 (64%)
	Female	174 (36%)
Age	< 60 years	273 (56%)
	> 60 years	209 (44%)
Tumor size	< 5 cm	198 (41%)
	> 5 cm	284 (59%)
Location	Proximal	48 (10%)
	Distal	407 (84%)
	Whole stomach	27 (6%)
Depth of infiltration	Tis-T1	27 (6%)
	T2-4	455 (94%)
Lymph node metastases	Present	358 (75%)
	Absent	119 (25%)
Histological type	Intestinal	234 (48.5%)
	Diffuse	166 (34%)
	Mixed	60 (12%)
	Unclassified	22 (4.5%)

Table 2 Antibodies used for immunohistochemistry

Antibody	Clones	Source	Dilution	Expression patterns
β-catenin	17C2	Novocastra	1:100	Membranous
APC	Polyclonal	Santa Cruz Biotechnologies	1:800	Cytoplasm
Bak	Polyclonal	Dako	1:400	Cytoplasm
Bax	Polyclonal	Dako	1:50	Cytoplasm
Bcl-2	124	Dako	1:40	Cytoplasm
Bcl-x	Polyclonal	Dako	1:50	Cytoplasm
c-erbB-2	Polyclonal	DAKO	1:500	Membranous
Clathrin	23	BD Transduction	1:2000	Cytoplasm
C-met	Polyclonal	Novocastra	1:50	Membranous
Cyclin A	Polyclonal	Santa Cruz Biotechnologies	1:40	Nuclear
Cyclin B1	V152	DAKO	1:50	Cytoplasm
Cyclin D1	RBT-14	BIO SB	Read to use	Nuclear
E-cadherin	36B5	Novocastra	1:50	Membranous
MMP-2	75-7F7	Oncogene	1:40	Cytoplasm
MMP-9	56-2A4	Oncogene	1:80	Cytoplasm
MSH-2	Polyclonal	Santa Cruz	1:25	Nuclear
MUC1	Ma695	Novocastra	1:500	Cytoplasm
MUC2	CCp58	Novocastra	1:1000	Cytoplasm
MUC5AC	CLH2	Novocastra	1:500	Cytoplasm
MUC6	CLH5	Novocastra	1:600	Cytoplasm
NOS-1	nNOS	Santa Cruz Biotechnologies	1:200	Cytoplasm
NOS-2	iNOS	Santa Cruz Biotechnologies	1:40	Cytoplasm
NOS-3	eNOS	Santa Cruz Biotechnologies	1:100	Cytoplasm
p16	C-20	MTM	1:25	Nuclear
p21	SX118	DAKO	1:30	Nuclear
p27	SX53G8	DAKO	1:200	Nuclear
p53	DO7	DAKO	1:100	Nuclear
pRb	Rb1	DAKO	1:50	Nuclear
TFG-βI	Polyclonal	Santa Cruz Biotechnologies	1:50	Cytoplasm
TFG-βII	Polyclonal	Santa Cruz Biotechnologies	1:200	Cytoplasm
VEGF	Polyclonal	Santa Cruz Biotechnologies	1:500	Cytoplasm

APC, matrix metalloproteinase; MMP, matrix metalloproteinase; MSH-2, MutS homolog-2; MUC, mucin; NOS, nitric oxide synthase; TGF, transforming growth factor; VEGF, vascular endothelial growth factor.

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

follow up or who died from any other cause than GC were regarded as censored data for the analysis of the survival rates. No hereditary tumors or tumors associated with specific genetic mutations were included. The institutional board for ethical studies at A. C. Camargo Hospital approved this study.

TMA methods

Three array blocks, containing a total of 482 cases, were prepared, as described previously. Using a tissue microarrayer (Beecher Instruments, Silver Spring, MD, USA), core tissue biopsies (0.6 mm diameter) were taken from individual paraffin-embedded GC (donor blocks) and transferred to a recipient paraffin block (tissue array blocks). The area of interest in the donor block was cored twice, and the cores ($n = 964$) were arrayed onto three separate blocks that contained two tumor cores from each case. Subsequently, 4-μm-thick sections were cut from each tissue array block, deparaffinized, and dehydrated.

Immunohistochemistry

A standard peroxidase-conjugated streptavidin–biotin method was used to detect the staining reaction against tumor-associated gene products. Commercially-available antibodies were tested using a stomach cancer control slide. After the test procedure, 28 antibodies, which were properly stained in each positive and negative control, were selected for this study. The primary antibodies studied are described in Table 2. The immunohistochemical staining was evaluated by means of light microscopic examination, and interpreted by two independent pathologists who were blinded to clinical information. The final consensus was discussed and determined in a common session. For the statistical analysis, the cases were considered positive when the tumor cells showed a dark brown color in more than 10% of the neoplastic cells.[11] Immunohistochemistry was done in two slides of the three blocks of TMA for each antibody.

Cluster analyses

Cluster analyses were used to identify the similarity profile of variations with our data. In this study, hierarchical cluster analyses were done using a random forest-clustering method (TMEV, http://www.tm4.org/mev.html), and results were displayed using TreeView.[14] This method is an unsupervised learning method that aims to find molecular classifications with distinct global expression profiles blinded to clinicopathological covariates.

Statistical analyses

The survival curves were estimated using the Kaplan–Meier product-limit method, and the significant differences between the survival curves were determined using the log–rank test. Multivariate survival analysis was performed using the Cox proportional hazards model. The correlation coefficients between the immunohistochemical expression status and clinicopathological findings, as well as between every antibody expression status, were estimated by Pearson's correlation. The χ^2-test or Fisher's exact test (two sided) was performed to determine the correlation between the antibodies expression and clinicopathological parameters. The results were considered to be statistically significant at $P < 0.05$. All statistical analyses were conducted using the SPSS 10.0 statistical software program (SPSS, Chicago, IL, USA).

Table 3 Expression of the 28 proteins and association with histological subtype of gastric carcinomas

Antibody	Category	Histological subtype			*P*-value
		Intestinal	Diffuse	Total	
E-cadherin	Positive	87 (39%)	11 (7%)	98 (100%)	#0.001
	Negative	137 (61%)	144 (93%)	281 (100%)	
MUC2	Negative	212 (93%)	128 (82%)	340 (100%)	0.001
	Positive	14 (7%)	29 (18%)	43 (100%)	
Clathrin	Negative	56 (25%)	86 (54%)	142 (100%)	< 0.001
	Positive	172 (75%)	73 (46%)	245 (100%)	
MMP-2	Negative	123 (54%)	101 (66%)	224 (100%)	0.019
	Positive	104 (46%)	51 (34%)	155 (100%)	
Cyclin B1	Negative	119 (52%)	64 (41%)	183 (100%)	0.029
	Positive	107 (48%)	92 (59%)	199 (100%)	
pRb	Negative	86 (39%)	36 (23%)	122 (100%)	0.002
	Positive	138 (61%)	121 (77%)	259 (100%)	
p21	Negative	189 (84%)	144 (93%)	333 (100%)	0.007
	Positive	37 (16%)	11 (7%)	48 (100%)	
p53	Negative	146 (65%)	118 (76%)	264 (100%)	0.001
	Positive	80 (35%)	36 (24%)	116 (100%)	
TGF-βI	Negative	31 (14%)	40 (25%)	71 (100%)	0.005
	Positive	197 (86%)	117 (75%)	314 (100%)	
TGF-βII	Negative	3 (2%)	9 (6%)	12 (100%)	0.018
	Positive	223 (98%)	147 (94%)	370 (100%)	
c-erbB-2	Negative	177 (78%)	154 (97%)	331 (100%)	< .001
	Positive	49 (22%)	5 (3%)	54 (100%)	
c-Met	Negative	17 (8%)	25 (16%)	42 (100%)	0.012
	Positive	208 (92%)	129 (84%)	333 (100%)	
VEGF	Negative	10 (8%)	23 (15%)	33 (100%)	< 0.001
	Positive	207 (92%)	126 (85%)	333 (100%)	
MSH-2	Negative	6 (3%)	24 (15%)	30 (100%)	< 0.001
	Positive	220 (97%)	136 (85%)	356 (100%)	
NOS-3	Negative	2 (1%)	11 (8%)	13 (100%)	0.002
	Positive	215 (99%)	137 (92%)	352 (100%)	

MMP-2, matrix metalloproteinase-2; MSH-2, MutS homolog-2; MUC2, mucin2; NOS-3, nitric oxide synthase-3; TGF, transforming growth factor; VEGF, vascular endothelial growth factor.

Results

Immunohistochemical findings

The staining results of the 28 antibodies are summarized in Table 3. For most of the antibodies tested, there was a significant difference in immunostaining between the diffuse- and intestinal-type GC. Of the 28 proteins studied, 13 were more often immunopositive in intestinal-type GC: clathrin ($P < 0.001$), E-cadherin ($P = 0.001$), matrix metalloproteinase-2 ($P = 0.019$), p21 ($P = 0.007$), p53 ($P = 0.001$), transforming growth factor (TGF)-βI ($P = 0.005$), TGF-βII ($P = 0.018$), c-erbB-2 ($P < 0.001$), vascular endothelial growth factor (VEGF; $P < 0.001$), MutS homolog-2 ($P < 0.001$), and nitric oxide-3 ($P = 0.002$). However,

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

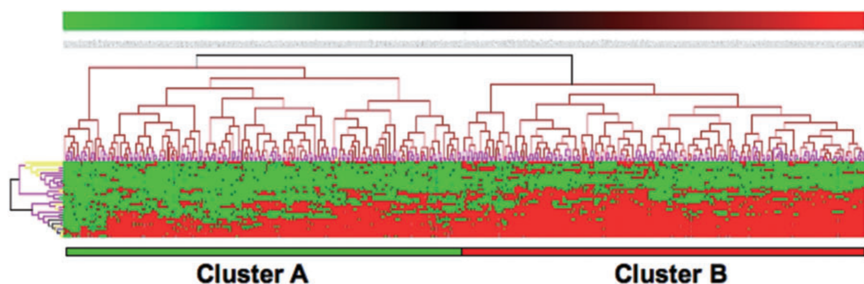

Figure 1 Two-way hierarchical cluster analysis of 482 gastric carcinomas according to the expression of 28 proteins. Tumors were grouped into three clusters, based on the protein expression profile. Rows represent proteins ordered according to their hierarchical distances. Colors in columns represent expression levels: red indicates positive staining, and green represents the absence of staining for each of the antibodies studied. Within each cluster, samples were ordered on the basis of their correlation distances.

Table 4 Associations between clusters and clinicopathological features in gastric carcinomas

Variable	Category	Cluster A	Cluster B	Total	*P*-value
Sex	Male	138 (48%)	154 (52%)	292 (100%)	NS
	Female	90 (54%)	78 (46%)	168 (100%)	
Age	< 60 years	129 (50%)	130 (50%)	259 (100%)	NS
	> 60 years	99 (49%)	102 (51%)	201 (100%)	
Tumor size	< 5 cm	102 (54%)	86 (46%)	188 (100%)	NS
	> 5 cm	126 (46%)	146 (54%)	272 (100%)	
Dept of infiltration	T2-4	221 (50%)	218 (50%)	340 (100%)	NS
	Tis-T1	7 (33%)	14 (67%)	21 (100%)	
Lymph node metastases	Present	175 (51%)	167 (49%)	342 (100%)	NS
	Absent	52 (46%)	61 (54%)	113 (100%)	
Location	Distal	192 (50%)	194 (50%)	386 (100%)	NS
	Proximal	23 (48%)	25 (52%)	48 (100%)	
	Whole stomach	13 (52%)	12 (48%)	25 (100%)	
Histological type	Intestinal	100 (44%)	126 (56%)	226 (100%)	NS
	Diffuse	88 (56%)	68 (44%)	156 (100%)	
	Mixed	28 (48%)	30 (52%)	58 (100%)	
	Unclassified	12 (60%)	8 (40%)	20 (100%)	
Clinical stage	I and II	135 (53%)	121 (47%)	256 (100%)	0.046
	III and IV	61 (42%)	83 (58%)	144 (100%)	

NS, not significant.

the following were more often immunopositive in diffuse-type GC: mucin2 (MUC2) ($P = 0.001$), cyclin B1 ($P = 0.029$), and pRb ($P = 0.002$).

Hierarchical cluster analysis

The overall expression patterns for the 460 samples of GC were analyzed by hierarchical clustering after excluding those with values missing in more than 20% of the columns. The combined protein expression patterns defined two clusters: cluster A (228 cases) and cluster B (232 cases; Fig. 1). Cluster A consisted of tumors frequently negative for proteins related to cell-cycle control and apoptosis (e.g. Bcl-x, Bak, Bcl-2, cyclins), whereas cluster B consisted of tumors more often positive for suppressor proteins (e.g. p21, p16, p53, E-cadherin).

Significant associations between clusters and clinicopathological findings, such as sex, age, location, and histological tumor type, were not observed. However, cluster B was significantly correlated with advanced pTNM stage ($P = 0.046$, χ^2). All of the data are summarized in Table 4.

Previous studies have indicated distinct genetic changes and an expression pattern of a subset of genes and proteins between the histological subtypes of GC. To identify novel genes and proteins associated with the histological subtypes, we performed a separate clustering analysis of the protein expression in the two major types (intestinal and diffuse). A total of 116 diffuse-type GC and 344 intestinal-type GC were studied. The tumors were separated into three main clusters (Figs 2a,3a).

To determine whether these clusters might represent clinically distinct subgroups of patients, a univariate survival analysis was carried out. In the group of diffuse-type tumors, cluster A was associated with poorer prognosis, when compared with clusters B and C. The rate of overall survival time was 50.8% for patients in cluster C, 36.1% for patients in cluster B, and 12.3% for patients in cluster A. These differences were statistically significant ($P = 0.001$, Fig. 2b). Cluster C was characterized by tumors

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

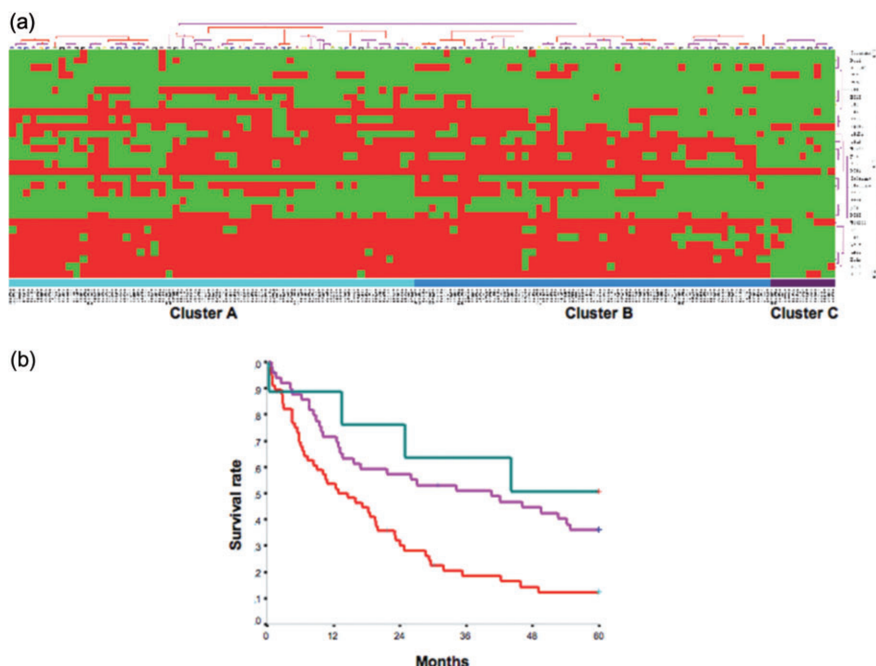

Figure 2 Classification of diffuse gastric carcinomas based on the expression of the 28 proteins. (a) Matrix format presenting the whole data. In the data matrix, a row corresponds to a single protein, and each column corresponds to a single tumor. Tumors were grouped into three clusters, based on the protein expression profile. Red denotes positive staining, and green denotes the absence of staining. (b) Univariate survival analysis by Kaplan–Meier method. $P = 0.001$. (b) ——, Cluster A; ——, Cluster B; ——, Cluster C.

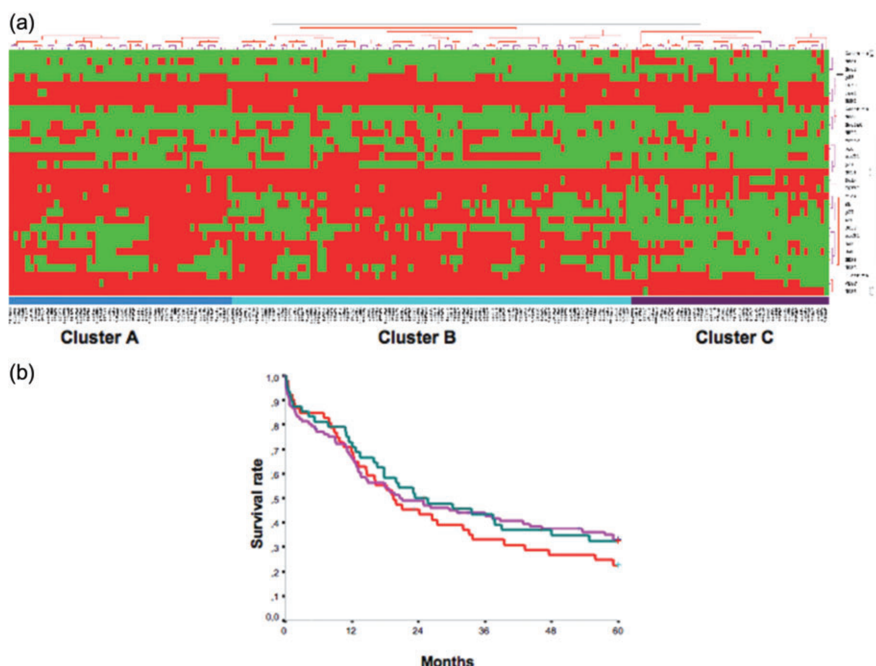

Figure 3 Classification of intestinal gastric carcinomas based on the expression of the 28 proteins. (a) Tumors were grouped into three clusters, based on the protein expression profile. Red denotes positive staining, and green denotes the absence of staining. (b) Univariate survival analysis by Kaplan–Meier method. $P = 0.610$. (b) ——, Cluster A; ——, Cluster B; ——, Cluster C.

frequently negative for most of the proteins studied. Interestingly, they were positive only for MUC2, TGF-βII, and cyclin B1. Kaplan–Meier curves did not show any differences between the clusters in the intestinal tumor of GC ($P = 0.328$, Fig. 3b).

Discussion

GC is associated with various genetic alterations, and no single genetic marker can predict the biology and prognosis of patients with GC. The potential use of combinations of biomarkers instead of a single marker or histological feature has been previously demonstrated.[11,15] Previous studies have applied tissue array methods for the molecular classifications of various tumors, such as endometrial, breast, and brain tumors,[10,16,17] and protein expression profiling has been found to be clinically useful for the prognostic classification of tumors.[11,15] In this study, we have shown that a molecular classification of GC can be accomplished based on a hierarchical cluster analysis of the immunohistochemical profiles of tumor-associated biomarkers using TMA sections.

Studies on the molecular pathology of GC have largely been focused on the subtypes of GC: intestinal and diffuse.[1,18,19] Several regulatory genes and molecular pathways, including regulation of

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

the cell cycle, DNA repair, cell adhesion, angiogenesis, apoptosis, and matrix remodeling, have been extensively studied in GC.[6,20] Boussioutas *et al*. found distinct expression profiles between intestinal- and diffuse-type GC[3]. The upregulated genes in the intestinal-type GC were observed to be mediators of cell proliferation, including genes active in cell-cycle control, DNA replication, and arrangement of chromosomes, while genes associated with epithelial differentiation were downregulated.[3] Our results, based on immunohistochemistry, demonstrated various proteins related to these molecular mechanisms more frequently positive in the intestinal-type than in diffuse-type GC, such as TGF-βI, TGF-βII, c-erbB-2, p21, p53, and VEGF. In contrast, diffuse-type GC was characterized by the expression of MUC2, cyclin D1, and pRb. In addition, losses of adhesion molecules, such as E-cadherin, were more frequent in the diffuse-type than in intestinal-type GC. In line with our results, it has been shown that diffuse-type GC gene expression changes were found in genes of the extracellular matrix components (collagens and proteoglycan), smooth muscles, and their extracellular stroma.[3]

Hierarchical clustering analysis is a method of dealing with complex sets of data and offers a possible systematic approach to complex immunophenotypes. Clustering analysis has been successfully applied to gene expression data, based on the expression of thousands of genes, and has been remarkably successful in its ability to group tumors according to their primary site based on gene expression profile.[21–23] The clustering of tumors based on immunoreactivity is expected to be less defined than the gene expression array studies. This application of clustering analysis allows for a more objective interpretation of immunoprofiles based on staining with multiple antibodies, and holds great promise for the immunohistochemical classification of tumors.[16]

The overall protein expression patterns defined two main clusters, that is, cluster A (228 cases) and cluster B (232 cases). Although there was no survival-associated cluster, they demonstrated specific protein profiles that could be important for using target therapies. For example, patients in cluster A will not benefit from using drugs that have anti-apoptotic or cell-cycle inhibitor activities, because they were frequently negative for the proteins related to these molecular mechanisms. Moreover, we found that cluster B mainly comprised advanced tumors (stages III and IV), and the principal expression proteins were associated with gene suppressor. Although our results are promising, further validation is needed concerning the use of these classifiers for screening tests for treatment selection.

Using a hierarchical cluster analysis of the immunohistochemical profiles of the diffuse-type ($n = 116$) and intestinal-type ($n = 344$) GC, we found three distinct subgroups of the tumors. Patients with cluster C diffuse-type GC were found to have the best prognosis among the three clusters ($P = 0.001$). They were characterized by immunoexpression of cyclin D1, MUC2, and TGF-βII. However, patients with intestinal-type GC did not show any differences related to the prognosis, which might probably be due to the more heterogeneous nature of this type of GC. The fact that the random forest method was able to create clinically meaningful subgroups of tumors using protein expression signature provides indirect evidence that the method works well on real data. Using this method, we were able to discover novel, molecularly-defined patients group who might not have been identified using traditional clinicopathological features.

Although preliminary, these results not only provide a new molecular basis for understanding the biological properties of GC, but also better prediction of survival than the classical pathological grouping. Further validation across different institutions and technological platforms is needed concerning the use of the above classifiers for routine pathological diagnoses.

Acknowledgment

This study was supported by a Fundação de Amparo à Pesquisa do Estado de São Paulo grant (no. 98/14335-2).

References

1 Zheng L, Wang L, Ajani J, Xie K. Molecular basis of gastric cancer development and progression. *Gastric Cancer* 2004; **7**: 61–77.
2 Schena M, Shalon D, Davis RW, Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 1995; **270**: 467–70.
3 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.
4 Becker KF, Keller G, Hoefler H. The use of molecular biology in diagnosis and prognosis of gastric cancer. *Surg. Oncol.* 2000; **9**: 5–11.
5 Hasegawa S, Furukawa Y, Li M *et al*. Genome-wide analysis of gene expression in intestinal-type gastric cancers using a complementary DNA microarray representing 23 040 genes. *Cancer Res.* 2002; **62**: 7012–17.
6 Hippo Y, Taniguchi H, Tsutsumi S *et al*. Global gene expression analysis of gastric cancer by oligonucleotide microarrays. *Cancer Res.* 2002; **62**: 233–40.
7 Chen X, Leung SY, Yuen ST *et al*. Variation in gene expression patterns in human gastric cancers. *Mol. Biol. Cell* 2003; **14**: 3208–15.
8 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.
9 Kim B, Bang S, Lee S *et al*. Expression profiling and subtype-specific expression of stomach cancer. *Cancer Res.* 2003; **63**: 8248–55.
10 Jacquemier J, Ginestier C, Rougemont J *et al*. Protein expression profiling identifies subclasses of breast cancer and predicts prognosis. *Cancer Res.* 2005; **65**: 767–79.
11 Lee HS, Cho SB, Lee HE *et al*. Protein expression profiling and molecular classification of gastric cancer by the tissue array method. *Clin. Cancer Res.* 2007; **13**: 4154–63.
12 Kim JH, Kim MA, Lee HS, Kim WH. Comparative analysis of protein expressions in primary and metastatic gastric carcinomas. *Hum. Pathol.* 2009; **40**: 314–22.
13 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.
14 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.
15 Wang YY, Ye ZY, Zhao ZS, Tao HQ, Li SG. Systems biology approach to identification of biomarkers for metastatic progression in gastric cancer. *J. Cancer Res. Clin. Oncol.* 2010; **136**: 135–41.
16 Alkushi A, Irving J, Hsu F *et al*. Immunoprofile of cervical and endometrial adenocarcinomas using a tissue microarray. *Virchows Arch.* 2003; **442**: 271–7.

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd

17 Ikota H, Kinjo S, Yokoo H, Nakazato Y. Systematic immunohistochemical profiling of 378 brain tumors with 37 antibodies using tissue microarray technology. *Acta Neuropathol.* 2006; **111**: 475–82.
18 Fiocca R, Luinetti O, Villani L *et al*. Molecular mechanisms involved in the pathogenesis of gastric carcinoma: interactions between genetic alterations, cellular phenotype and cancer histotype. *Hepatogastroenterology* 2001; **48**: 1523–30.
19 Tahara E. Genetic pathways of two types of gastric cancer. *IARC Sci. Publ.* 2004; **157**: 327–49.
20 Stock M, Otto F. Gene deregulation in gastric cancer. *Gene* 2005; **360**: 1–19.
21 Alizadeh AA, Eisen MB, Davis RE *et al*. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 2000; **403**: 503–11.
22 Perou CM, Sorlie T, Eisen MB *et al*. Molecular portraits of human breast tumours. *Nature* 2000; **406**: 747–52.
23 Sorlie T, Perou CM, Tibshirani R *et al*. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. U.S.A.* 2001; **98**: 10869–74.

© 2011 Journal of Gastroenterology and Hepatology Foundation and Blackwell Publishing Asia Pty Ltd