

# Variation in Gene Expression Patterns in Human Gastric Cancers

Xin Chen,<sup>\*†‡</sup> Suet Y. Leung,<sup>‡§</sup> Siu T. Yuen,<sup>§</sup> Kent-Man Chu,<sup>¶</sup> Jiafu Ji,<sup>||</sup>  
Rui Li,<sup>\*</sup> Annie S.Y. Chan,<sup>§</sup> Simon Law,<sup>¶</sup> Olga G. Troyanskaya,<sup>#</sup>  
John Wong,<sup>¶</sup> Samuel So,<sup>\*</sup> David Botstein,<sup>#</sup> and Patrick O. Brown<sup>†\*\*††</sup>

Departments of <sup>\*</sup>Surgery, <sup>†</sup>Biochemistry, and <sup>#</sup>Genetics, and <sup>\*\*</sup>Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, California 94305; Departments of <sup>§</sup>Pathology and <sup>¶</sup>Surgery, The University of Hong Kong, Hong Kong, China; and <sup>||</sup>Department of Surgery, Perking University School of Oncology, Beijing Cancer Hospital, Beijing, China

Submitted December 18, 2002; Revised February 24, 2003; Accepted March 17, 2003  
Monitoring Editor: Keith Yamamoto

Gastric cancer is the world's second most common cause of cancer death. We analyzed gene expression patterns in 90 primary gastric cancers, 14 metastatic gastric cancers, and 22 nonneoplastic gastric tissues, using cDNA microarrays representing ~30,300 genes. Gastric cancers were distinguished from nonneoplastic gastric tissues by characteristic differences in their gene expression patterns. We found a diversity of gene expression patterns in gastric cancer, reflecting variation in intrinsic properties of tumor and normal cells and variation in the cellular composition of these complex tissues. We identified several genes whose expression levels were significantly correlated with patient survival. The variations in gene expression patterns among cancers in different patients suggest differences in pathogenetic pathways and potential therapeutic strategies.

## INTRODUCTION

Gastric cancer is the second most common cause of cancer death worldwide (Parkin *et al.*, 1999). Environmental and genetic factors are both important in gastric carcinogenesis. Gastritis, most commonly due to *Helicobacter pylori* infection, is strongly associated with development of gastric cancer (Parsonnet *et al.*, 1991). In a small fraction of gastric cancers, the cancer cells have been found to be latently infected by Epstein-Barr virus (EBV; Shibata and Weiss, 1992). Intestinal metaplasia in response to chronic inflammation often precedes development of gastric cancer and may be an important precursor (Correa, 1988). The Lauren classification, which distinguishes tumors based on their histological appearance into "intestinal" and "diffuse" types, has proven to be clinically useful (Lauren, 1965). These two histological

types have been proposed to be related to differences in pathogenetic mechanisms (Lauren, 1965; Grabiec and Owen, 1985).

The prognosis of gastric cancer depends heavily on tumor stage at diagnosis. Surgical resection, still the mainstay of treatment, is very effective in early stage cancer. Most patients present in advanced stages, however, when the prognosis is extremely poor. Recent studies have shown that comprehensive analysis of gene expression patterns can identify new molecular criteria for classification and prognostication of diverse cancers. The molecular portraits provided by the tumors' gene expression patterns provide a basis for recognizing new subtypes of cancer that, although morphologically indistinguishable, differ in clinical behavior, molecular pathways of tumor development, and underlying genetic predisposition (Alizadeh *et al.*, 2000; Perou *et al.*, 2000; Dhanasekaran *et al.*, 2001; Garber *et al.*, 2001; Hedenfalk *et al.*, 2001; Rosenwald *et al.*, 2001; Sorlie *et al.*, 2001; Pomeroy *et al.*, 2002; Rosenwald *et al.*, 2002; van 't Veer *et al.*, 2002).

Previous studies have reported expression profiles in small numbers of gastric cancer samples, identifying a few hundred genes that distinguish tumor from normal tissue (El-Rifai *et al.*, 2001; Hippo *et al.*, 2002). However, these studies were too small to distinguish subgroups or identify any genes whose expression may be correlated with clinical outcomes. Here we describe a systematic study of gene

Article published online ahead of print. Mol. Biol. Cell 10.1091/mbc.E02-12-0833. Article and publication date are available at [www.molbiolcell.org/cgi/doi/10.1091/mbc.E02-12-0833](http://www.molbiolcell.org/cgi/doi/10.1091/mbc.E02-12-0833).

<sup>††</sup> Corresponding author. E-mail address: [pbrown@cmgm.stanford.edu](mailto:pbrown@cmgm.stanford.edu).

<sup>‡</sup>These authors contributed equally to this work.

Supplementary Information is available through the author's Web supplement site at: [http://genome-www.stanford.edu/Gastric\\_Cancer2](http://genome-www.stanford.edu/Gastric_Cancer2).

Abbreviations used: EBV: Epstein-Barr virus; *H. pylori*: *Helicobacter pylori*.

expression patterns in gastric adenocarcinomas and nonneoplastic gastric mucosa. The results define potential molecular criteria for disease classification, early detection, and prognostication as well as providing new insights into the pathogenesis and progression of gastric cancers.

## MATERIALS AND METHODS

### *Samples and RNA Preparation*

Samples of tumor and normal gastric mucosa were collected from gastrectomy specimens from Department of Surgery, Queen Mary Hospital, The University of Hong Kong. Tissues were frozen in liquid nitrogen within half an hour after they were resected. Non-neoplastic mucosa from stomach, small intestine, and colon was dissected free of muscle and histologically confirmed to be tumor free by frozen section. Total RNA was extracted using Trizol (Invitrogen, Carlsbad, CA) and mRNA was isolated by the FastTrack mRNA isolation kit (Invitrogen). This study was approved by the Ethics Committee of the University of Hong Kong and the IRB from Stanford University.

Tumors were classified using the Lauren's classification into intestinal, diffuse, mixed and indeterminate types (Lauren, 1965). The presence of *H. pylori* in the gastrectomy specimens was determined by histological examination, supplemented by modified Giemsa staining. The presence of EBV in cancer cells was assayed by *in situ* hybridization for EBER as described previously (Yuen *et al.*, 1994). The tumor stage was defined by the General Rules for Gastric Cancer Study of the Japanese Research Society for Gastric Cancer (Nishi *et al.*, 1995).

### *Microarray Procedure and Data Analysis*

We used cDNAs microarray containing 44,500 cDNA clones, representing ~30,300 unique genes. The methods for microarray production and hybridization, and data analysis, including hierarchical clustering, closely followed previously described procedures (Alizadeh *et al.*, 2000; Perou *et al.*, 2000). In brief, 2  $\mu$ g of sample mRNA and 2  $\mu$ g of reference mRNA were used as template for synthesis of cDNA, labeled with Cy5-dUTP and Cy3-dUTP (Amersham, Piscataway, NJ), respectively, using reverse transcriptase (Invitrogen) for 2 h at 42°C. The two labeled cDNA samples were separated from unincorporated nucleotides by filtration, mixed, and hybridized to microarray at 65°C overnight. After hybridization, each microarray was washed with 2 $\times$  SSC, 0.03% SDS for 5 min at 65°C and then with 1 $\times$  SSC for 5 min and 0.1 $\times$  SSC for 5 min, both at room temperature. The array was then scanned with a GenePix 4000B microarray scanner (Axon Instruments, Union City, CA). Primary data collection and analysis were carried out using GenePix Pro3.0 (Axon Instruments). Areas of the array with obvious blemishes were manually flagged and excluded from subsequent analysis. The raw data were deposited into Stanford Microarray Database (Sherlock *et al.*, 2001) at: <http://genome-www4.stanford.edu/MicroArray/SMD/index.html>.

### *Statistical Analysis*

We selected nonflagged array elements for which the fluorescent intensity in either channel was greater than 3 times the local background. Genes for which fewer than 80% of measurements met this standard, across all the samples in this study, were excluded from the analysis. We further selected genes whose expression level differed by at least 2.5-fold, in at least three samples, from their mean expression level across all samples for hierarchical clustering. Before the statistical analysis, the missing values in the dataset were estimated by a KNN impute algorithm using 12 neighbors (Troyanskaya *et al.*, 2001). A nonparametric *t* test with p value cutoff of 0.001 or 0.002 based on 10,000 random column permutations was used to identify genes differentially expressed in gastric cancer and

nonneoplastic gastric tissue samples (Troyanskaya *et al.*, 2002). The false discovery rate (FDR) was estimated based on the number of genes that passed the p value cutoff of 0.001 or 0.002 in five datasets with randomized columns. Cox regression survival analysis was performed to identify genes whose expression levels correlated with patient survival, using the Cox regression function in the statistical package R (<http://lib.stat.cmu.edu/R/CRAN>). The p value reflecting the correlation of each gene's expression pattern with patient survival was calculated individually using the Cox function in SPSS, without adjustment for multiple testing, therefore, the p values presented in the article might overestimate statistical significance because of multiple hypothesis testing.

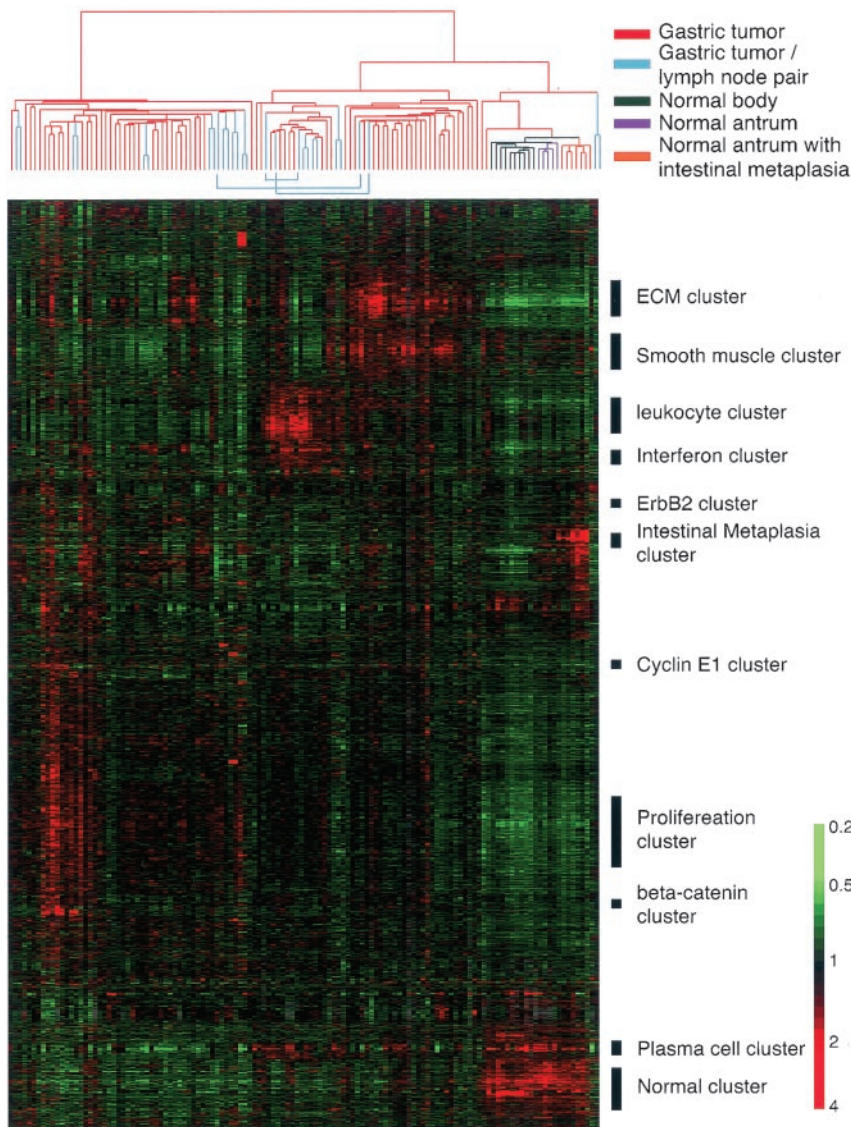
## RESULTS AND DISCUSSION

Using a cDNA microarray containing 44,500 clones, representing ~30,300 different genes, we profiled mRNA populations in 90 primary gastric adenocarcinomas, lymph node metastases from 14 of the same 90 tumors, and 22 samples of nonneoplastic gastric mucosa adjacent from resected tumors, representing both antrum (12 samples) and body (10 samples). Among the 12 antral tissues, 7 showed intestinal metaplasia. From the roughly 30,000 genes whose expression was measured in these samples, we identified a set of ~5200 genes whose transcripts varied by at least 2.5-fold from their mean abundance in the sample set, in at least 3 samples (Figure 1). We used a hierarchical clustering method to group these genes on the basis of similarity in expression across the samples and to group samples based on similarity in their gene expression patterns.

### *Heterogeneity of the Gene Expression Patterns among Gastric Cancer Tissues*

Simple hierarchical clustering analysis readily segregated nontumor samples from all the tumors and revealed complex variations in gene expression patterns among the gastric cancers. Eleven of the 14 lymph node metastases were clustered with primary tumor from which they arose, implying that the gene expression patterns in the metastatic tumor masses were more similar to the patterns in the corresponding primary tumors than to any of the other metastatic or primary tumors. The three exceptions were attributable to differences between the primary and metastatic tumors in the admixture of nontumor cells; when hierarchical clustering of the tumors was carried out based on expression of set of genes that excluded genes characteristically expressed in lymphocytes and stromal cells, all 14 pairs of primary and metastatic tumors from the same patient clustered together as nearest neighbors (see Web supplement). This result implies that each gastric cancer has a unique characteristic gene expression pattern and that the distinctive features of these molecular portraits are maintained even when tumors metastasize to the lymph node. We can use these 14 paired samples to identify a set of "intrinsic genes," whose expression levels vary widely among tumors from different patients, but are closely matched in primary and metastatic tumors from the same patient. These genes are likely to be especially useful for recognition of molecular subtypes of gastric cancer (Perou *et al.*, 2000; see Web supplement).

Genes with consistent differences in expression between the gastric cancer and the nonneoplastic gastric tissues are of particular interest. Using a nonparametric *t* test, we identi-



**Figure 1.** Hierarchical clustering of the patterns of variation in expression of 5200 genes in 90 gastric tumors, 14 paired lymph node metastasis, and 22 nonneoplastic gastric mucosa. Each row represents a separate cDNA clone on the microarray and each column a separate mRNA sample. In each gastric tissue sample, the ratio of abundance of transcripts of each gene to its mean abundance across all tissue samples is depicted according to the color scale shown at the bottom. Gray indicates missing or excluded data. The dendrogram at the top provides a measure of the relatedness of gene expression in each sample and is color coded according to the category of mRNA sample studied. The dendrogram at the top of the figure represents the hierarchical clustering of the tumors based on similarity in their pattern of expression of these genes. See supplementary information for full data.

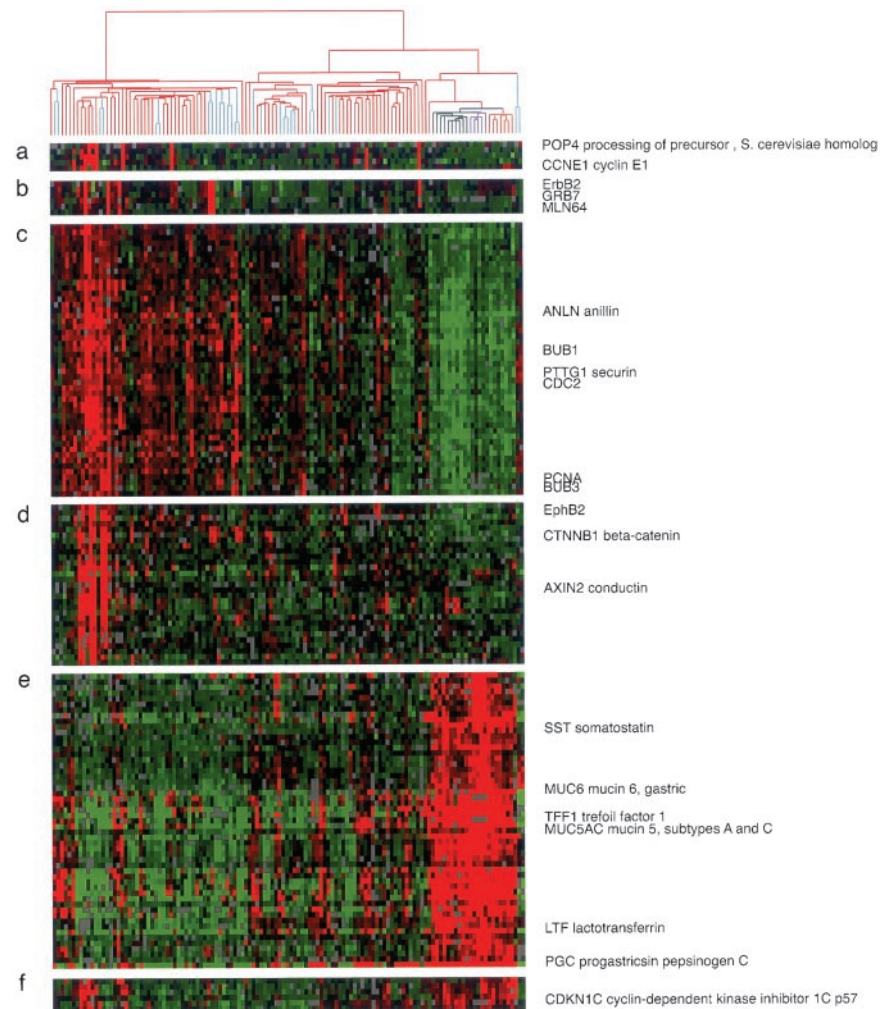
fied 2656 genes whose expression levels were significantly different between tumor and normal tissues, with  $p \leq 0.001$  and a false discovery rate of 0.13% (see Web supplement).

A cluster of  $\sim 2000$  genes was expressed at higher levels in most of gastric cancers compared with nontumor tissues (Figure 1). Among these 2000 genes were a group of genes associated with cell cycle progression, the “proliferation cluster” (Iyer *et al.*, 1999; Whitfield *et al.*, 2002), whose transcripts were generally much more abundant in the tumors than in nonneoplastic tissues, though their levels varied considerably among the tumor samples (Figure 2c).

Several clusters of genes that appear to reflect primary events important for gastric carcinogenesis, such as chromosomal amplifications or alterations in transcriptional regulators or signal transduction pathways, were specifically associated with subsets of these cancers. For example, cyclinE1, POP4, RMP, UQCRFS1, and DKFZP762D096 were coexpressed at greatly elevated levels in  $\sim 10\%$  of gastric tumors

(Figure 2a). All these genes are located on chromosome 19q12-13, suggesting amplification of a locus containing CyclinE1 in this group of tumors. Similarly, in  $\sim 10\%$  of tumors, a cluster of genes including ErbB2 and two neighboring genes on chromosome 17q12-21 (Grb7 and MLN64), showed markedly elevated expression, strongly suggesting amplification of the ErbB2 locus in these tumors (Figure 2b).

Another cluster of genes jointly expressed at elevated levels in some gastric cancers included AXIN2 and beta-catenin (Figure 2d). AXIN2 is induced by beta-catenin/TCF as part of an auto-regulatory pathway in Wnt signaling. Colorectal and liver tumors with activation of the Wnt signaling pathway have been shown consistently to express a high level of AXIN2 (Lustig *et al.*, 2002). Activation of this pathway by mutations in APC or beta-catenin leads to increased levels of beta-catenin and its nuclear localization (Korinek *et al.*, 1997). Indeed, immunohistochemical staining revealed nuclear beta-catenin in many of the gastric tumors



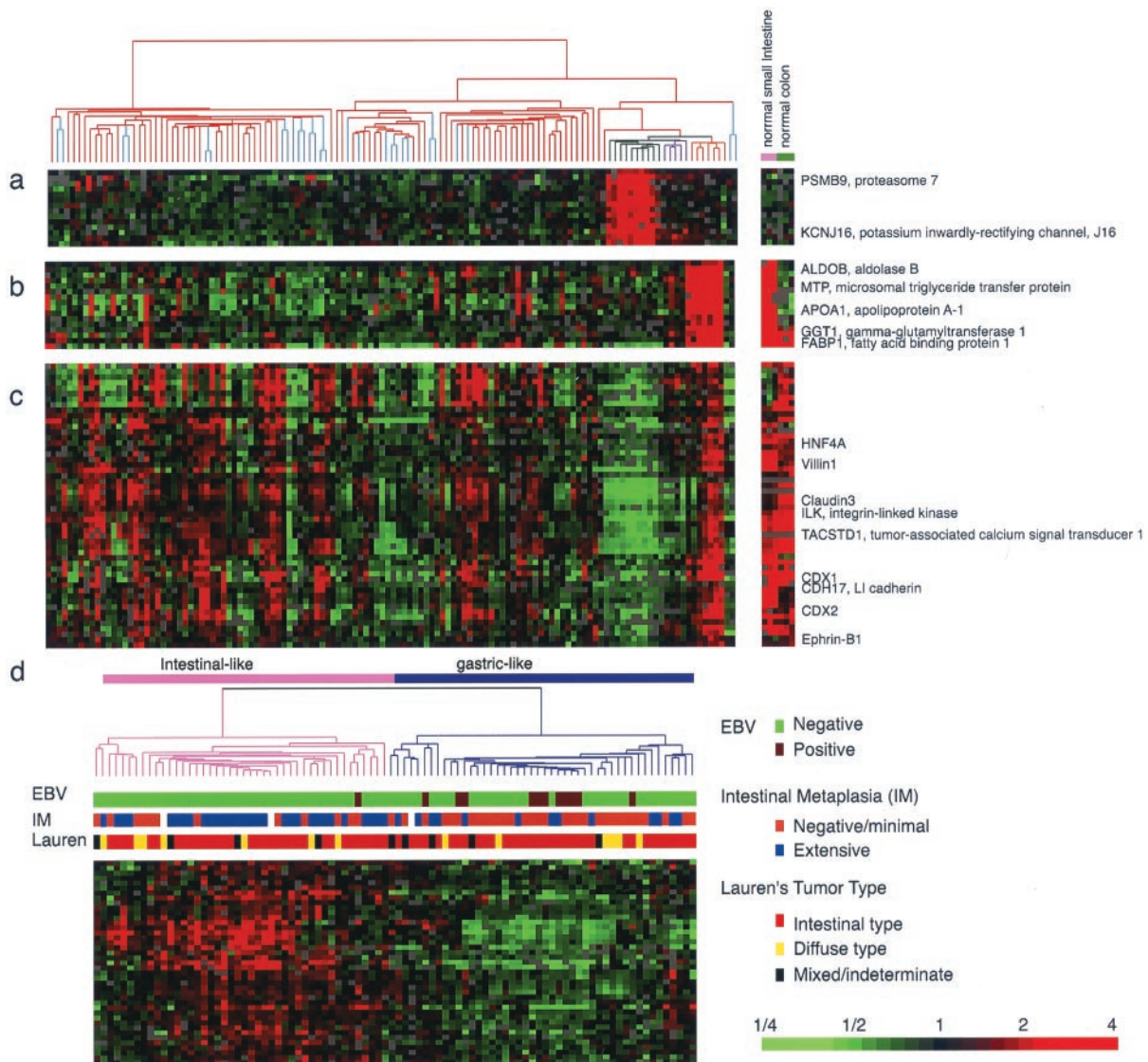
**Figure 2.** Features of the variation in gene expression patterns can be related to specific physiological and genotypic variations in the tumor and nontumor tissue samples. Data are the same as in Figure 1. Selected gene clusters (a–d) illustrate coordinated variation in expression of genes in tumor samples, reflecting apparent variation in signaling pathways, chromosomal amplification, or cell proliferation. (a) Cyclin E cluster; (b) ErbB2 cluster; (c) proliferation cluster; (d) AXIN cluster. Genes depicted in the normal cluster (e and f) reflect loss of specialized gastric functional component as well as reduced expression of putative tumor suppressor gene or cell cycle inhibitor in the tumor samples.

that express high levels of AXIN2 (see Web supplement). Abnormal activation of the Wnt pathway may therefore play a role in the pathogenesis of this distinct subgroup of gastric tumors. This gene cluster also included EphB2, a member of the Eph receptor tyrosine kinases. Ephrin signaling has been shown to be important in embryonic development, including axon guidance and angiogenesis (Kullander and Klein, 2002). This pathway has also been implicated in promoting tumorigenesis (Dodelet and Pasquale, 2000). A recent study showed that the activation of beta-catenin led to high expression of EphB receptors and that expression of EphB is crucial for the correct positioning of epithelial cells in gut (Batlle *et al.*, 2002). It is therefore interesting to note that EphB2 coclustered with beta-catenin and AXIN2 because of its patterns of expression in these gastric cancer tissues. The relationship of expression of EphB and Wnt signaling pathways in gastric cancer development deserves further investigation.

Clusters of genes that were generally more highly expressed in normal gastric mucosa than in gastric cancers are also of considerable interest (Figure 2, e and f). Among them, many are known to have role in the digestive function of the stomach or the integrity of the mucosa, such as pepsinogen C, carbonic

anhydrase II, specific subtypes of gastric mucins (MUC5AC and MUC6) and somatostatin. Some putative tumor suppressor genes or genes with known growth inhibitory effects in the gastrointestinal tract can also be found in this cluster. These include TFF1—mice lacking TFF1 develop gastric adenomas and intramucosal carcinomas (Lefebvre *et al.*, 1996)—and CDKN1C, the gene encoding the p57KIP2 cyclin-dependent kinase inhibitor, whose promoter has been found to be methylated, accompanied by transcriptional silencing, in some gastric cancer cell lines (Shin *et al.*, 2000).

Gastric cancer is a histologically complex tissue. The reciprocal interactions between tumor cells and stromal cells play important roles in tumor biology. A cluster of genes, including many that encode extracellular matrix components, appears to reflect variation in the proportion of stromal cells in the tumor. The genes in this cluster tend to be significantly more highly expressed in tumors of the diffuse histological type than in tumors of the intestinal type (see Web supplement), consistent with greater propensity of this group of tumors for invasive growth, often provoking a dense fibrous reaction. The expression patterns of two clusters of genes, which we term the “leukocyte” and “interferon” clusters, respectively, appear to



**Figure 3.** Expanded view of the gastric body-specific gene cluster (a), intestinal metabolism gene cluster (b), and intestinal differentiation gene cluster (c). Expression patterns of the same genes in small intestine and colon are depicted in the right panel for comparison. (d) Discovery of gastric cancer subtypes by expression of intestinal differentiation gene cluster. Hierarchical clustering of gastric cancer based on the genes of the intestinal differentiation signature shown in c. Two GC subgroups, “intestinal-like” (pink) and “gastric-like” (purple) were defined by this process. Associations with histological features or etiological agents are depicted by color-coded box below the dendrogram.

reflect the density, composition, and physiology of infiltrating leukocytes. The genes in these two clusters were generally significantly more highly expressed in the EBV-associated tumors (see Web supplement), consistent with previous reports that these EBV-associated gastric cancers are often, though not invariably, accompanied by intense lymphoid infiltrate. A cluster of genes that included immunoglobulin genes was expressed at high levels in gastric mucosa and in some tumors, and their expression correlated well with the number of plasma cells as determined by histological examination (see Web supplement). Similarly, a cluster enriched in genes characteristically expressed by smooth muscle appears to reflect

adventitious smooth muscle cells in the samples (see Web supplement).

### *A Gene Expression Pattern Associated with Intestinal Metaplasia Defines a Distinct Subset of Gastric Cancers*

Gastric cancers frequently arise in a background of chronic gastritis with intestinal metaplasia. As a step toward understanding this relationship, we explored the variations in gene expression patterns in different anatomical regions of the normal stomach, in metaplastic gastric mucosa, and in

samples of normal mucosa from small intestine and colon (displayed to the right of the main cluster; Figure 3). Based on their global gene expression patterns, the samples of nonneoplastic gastric mucosa were segregated into three branches, one corresponding to the gastric body and the other two corresponding to antral mucosa with and without intestinal metaplasia, respectively.

Three prominent gene clusters distinguished these three different types of gastric mucosa. A group of genes were expressed selectively in gastric body, but not in antrum, small intestine or colon, nor, to any significant degree, in gastric tumors (Figure 3a). This group included genes involved in transport (KCNJ16) and proteolysis (PSMB9), presumably reflecting functional specialization of the mucosa of the gastric body. A second cluster of genes, expressed at high level in metaplastic mucosa, but very little in nonmetaplastic mucosa and most gastric tumors (Figure 3b), included genes with roles in lipid (MTP, FABP1, APOA1), amino acid (GGT1), or glucose (ALDOB) transport or metabolism. Most of these genes are also highly expressed in small intestine and to a lesser extent in the colon. The third cluster of genes, also highly expressed in metaplastic antral mucosa compared with nonmetaplastic gastric mucosa, appears to be related to the control of intestinal differentiation (Figure 3c). Genes in this cluster include transcription factors that are involved in inducing and maintaining differentiation of the small and large intestine (CDX1, CDX2, and HNF4A) as well as integral membrane proteins and adhesion molecules that are characteristic of intestinal epithelium (CLDN3, CLDN4, CDH17, and VIL1). The genes in this cluster are expressed in small intestine and colon mucosa, supporting their association with intestinal differentiation. Surprisingly, this group of genes was also expressed at high levels in many gastric cancers. On the basis of expression of 41 genes in the cluster, the 90 gastric cancer samples can be segregated into two groups, which for the purpose of this discussion, we labeled "intestinal-like" and "gastric-like" (Figure 3d). We hypothesized that tumors showing intestinal-like gene expression patterns were derived from gastric mucosa cells that have undergone intestinal metaplasia.

Gastric cancer is traditionally classified into two subtypes based on their histological architecture (Lauren, 1965). The "intestinal type," characterized by cohesive growth of cells in quasi-glandular structures, has been presumed to arise secondarily from intestinal metaplasia, whereas the "diffuse type," in which tumor cells were dispersed and generally not organized into glandular structures, has been thought to arise directly from nonmetaplastic foveolar or mucous neck cells (Grabiec and Owen, 1985). We therefore explored the relationship between expression of this intestinal differentiation gene cluster and the two histologically defined cancer subtypes. To our surprise, the intestinal-like gene expression pattern was seen with equal frequency in the histologically defined intestinal and diffuse groups (Figure 3d and Web supplement). In contrast, there was a strong association between the intestinal-like gene expression pattern and the presence of intestinal metaplasia at the tumor edge ( $p < 0.001$ ; Figure 3d and Web supplement). Our data therefore suggest the possibility that the intestinal-like gastric cancers, as defined based on their gene expression patterns, may arise from gastric mucosal cells that have previously undergone intestinal metaplasia and thus retain the molecular

signature of the intestinal enterocyte, whereas the gastric-like cancers may arise by alternative pathway that does not involve intestinal metaplasia as an intermediate step. It will be important in following up these results to focus for the attention on the progression of events involved in metaplasia and gastric cancer development.

Our identification of a transcriptional program related to intestinal differentiation in gastric cancer provides one potential basis for extending classification of gastric cancer beyond morphological distinctions. Antibodies against proteins encoded by genes in the "intestinal differentiation" cluster may provide a convenient adjunct to conventional histopathology in classifying gastric cancers. Because the distinct molecular phenotypes of the intestinal-like and gastric-like gastric cancers are likely to have counterparts in the physiology and behavior of the tumors, recognition of these differences may be of significant value in evaluating clinical behavior and treatment trials.

### *Differential Expression of the Molecular Targets of Potential Therapeutic Agents*

Specific features of the gene expression programs that distinguish gastric cancers may not only provide clues to differences in their regulation but may provide critical information for individualized therapy using molecularly targeted treatment modalities. For example, TACSTD1, which encodes a homophilic calcium-independent cell adhesion molecule, is among the genes characteristically expressed in the intestinal-like subset of gastric cancers. An mAb (CO17-1A) against this protein has shown considerable promise in treatment of colorectal cancer in phase II/III randomized clinical trials (Riethmuller *et al.*, 1998). The differential expression of this gene among gastric cancers suggests its potential selective application in treatment of patients with intestinal-like gastric cancer. Similarly, the elevated expression of the EGF receptor in ~15% of gastric cancers suggests that these patients may benefit from compounds (both small molecular inhibitors and monoclonal antibodies) that target this protein, which have shown promise in clinical trials against a variety of cancers, and the highly elevated expression of ErbB2 in a small but significant fraction of gastric cancers raises the possibility that this group of patients may benefit from treatment with Herceptin.

### *Gene Expression Patterns and Infectious Etiologies*

Despite the strong association between *H. pylori* infection and gastric cancer, we were unable to identify genes whose expression levels were significantly associated with *H. pylori* infection. Although the intestinal-like gene expression pattern was somewhat more common than the gastric-like pattern in cancers associated with *H. pylori* infection, the difference was not statistically significant (see Web supplement). The critical role that *H. pylori* infection can play in the pathogenesis of gastric cancer may precede the development of overt cancers by many years. Therefore, histological evidence of contemporaneous *H. pylori* infection in the surgical samples may not reflect the role this infection played in the development of a tumor. It will be of great interest in future studies to compare gene expression patterns in gastric can-

cer with a more comprehensive assessment of the patient's past history of *H. pylori* infection.

The expression levels of 326 genes showed a significant association with EBV infection ( $p < 0.002$ , and FDR = 5%; see Web supplement). Some of these genes appeared to be related to the lymphoid infiltrate often observed in EBV+ gastric cancers. Among the 135 genes whose expression tended to be significantly higher in the EBV+ cancers was p53, consistent with previous reports of elevated p53 expression by a nonmutational mechanism in these tumors (Leung *et al.*, 1998; Wu *et al.*, 2000). Interestingly, we found that most (11/12) EBV-associated gastric cancers had the gastric-like gene expression phenotype ( $p = 0.007$ ) (Figure 3d and Web supplement), although most of these tumors would be classified as intestinal by histopathology. The role of EBV in the pathogenesis of this distinct type of gastric cancer will be an interesting question for future investigation.

### **Expression Patterns Significantly Correlated with Patient Survival**

We identified several genes whose expression levels in the gastric cancer samples were significantly associated with patient survival (see Web supplement). These genes define molecular markers that may be immediately useful for prognosis and treatment stratification. Clearly, it will be important to corroborate the association between the expression of these genes and patient survival in an independent sample set. Some of these genes may play important roles in gastric tumor progression. For example, a high level of expression of IGF-2 was associated with poor patient survival ( $p < 0.002$  by Cox regression assay). Overexpression of IGF-2 has been implicated in progression and metastasis of many types of cancers, including colon cancer, breast cancer, Wilms' tumor, and neuroblastoma (Toretzky and Helman, 1996). The potential involvement of IGF-2 in the progression of gastric cancer clearly warrants further investigation. Patients whose gastric cancers expressed PLA2G2A, the gene encoding the group IIa secreted phospholipase A2, at high levels, had a highly significant survival advantage ( $p < 0.002$ ). The murine ortholog of PLA2G2A, has been shown genetically to limit the severity of intestinal neoplasia in the *Apc<sup>min1/+</sup>* mouse, a mouse model of familial adenomatous polyposis, suggesting that PLA2G2A may play a similar role in suppressing progression of human gastric cancer (MacPhee *et al.*, 1995; Cormier *et al.*, 1997). A more extensive analysis and discussion of the association between PLA2G2A and gastric cancer progression is reported separately (Leung *et al.*, 2002).

In summary, a systematic study of gene expression patterns in human gastric cancers has revealed inherent molecular heterogeneity in gastric cancer, and identified multidimensional variations in gene expression patterns that distinguish gastric carcinoma and gastric mucosa in various preneoplastic stages. By using both supervised and unsupervised methods to search for relationships and systematic patterns in the data and by comparison of the expression patterns in the cancers with those seen in normal tissue of various histological types, we can begin to place features of this molecular variation into an interpretative framework that provides insights into the biological features and potential clues to the pathogenesis of these cancers. Differences among cancers in expression of specific genes may provide

strong predictors of prognosis and suggest the potential utility of specific, molecularly targeted therapies.

### **ACKNOWLEDGMENTS**

We thank the Stanford Functional Genomic Center, Stanford Microarray database, and Stanford Asian Liver Center for their support. This work was supported by the Howard Hughes Medical Institute (P.O.B.) and by grants from National Cancer Institute (P.O.B. and D.B.), the H.M. Lui Foundation (X.C., R.L. and S. S.), Research Grants Council of the Hong Kong Special Administrative Region, HKU 7264/01M (S.Y.L., S.T.Y., and K.M.C.) and 973-G1998051230 and 863-2001AA227101 of China (J.J.). O.T.G. is a Howard Hughes Medical Institute Predoctoral Fellow and a Stanford Graduate Fellow. P.O.B. is an Investigator of the Howard Hughes Medical Institute.

### **REFERENCES**

- Alizadeh, A. A. *et al.* (2000). Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* **403**, 503–511.
- Batlle, E., Henderson, J.T., Beghtel, H., van de Born, M.M.W., Sancho, E., Huls, G., Pawson, T., and Clevers, H. (2002). beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/EphrinB. *Cell* **111**, 251–263.
- Cormier, R.T., Hong, K.H., Halberg, R.B., Hawkins, T.L., Richardson, P., Mulherkar, R., Dove, W.F., and Lander, E.S. (1997). Secretory phospholipase Pla2g2a confers resistance to intestinal tumorigenesis. *Nat. Genet.* **17**, 88–91.
- Correa, P. (1988). A human model of gastric carcinogenesis. *Cancer Res.* **48**, 3554–3560.
- Dhanasekaran, S.M., Barrette, T.R., Ghosh, D., Shah, R., Varambally, S., Kurachi, K., Pienta, K.J., Rubin, M.A., and Chinnaiyan, A.M. (2001). Delineation of prognostic biomarkers in prostate cancer. *Nature* **412**, 822–826.
- Dodelet, V.C., and Pasquale, E.B. (2000). Eph receptors and ephrin ligands: embryogenesis to tumorigenesis. *Oncogene* **19**, 5614–5619.
- El-Rifai, W., Frierson, H.F., Jr., Harper, J.C., Powell, S.M., and Knuttila, S. (2001). Expression profiling of gastric adenocarcinoma using cDNA array. *Int. J. Cancer* **92**, 832–838.
- Garber, M.E. *et al.* (2001). Diversity of gene expression in adenocarcinoma of the lung. *Proc. Natl. Acad. Sci. USA* **98**, 13784–13789.
- Grabiec, J., and Owen, D.A. (1985). Carcinoma of the stomach in young persons. *Cancer* **56**, 388–396.
- Hedenfalk, I. *et al.* (2001). Gene-expression profiles in hereditary breast cancer. *N. Engl. J. Med.* **344**, 539–548.
- Hippo, Y., Taniguchi, H., Tsutsumi, S., Machida, N., Chong, J.M., Fukayama, M., Kodama, T., and Aburatani, H. (2002). Global gene expression analysis of gastric cancer by oligonucleotide microarrays. *Cancer Res.* **62**, 233–240.
- Iyer, V.R. *et al.* (1999). The transcriptional program in the response of human fibroblasts to serum. *Science* **283**, 83–87.
- Korinek, V., Barker, N., Morin, P.J., van Wichen, D., de Weger, R., Kinzler, K.W., Vogelstein, B., and Clevers, H. (1997). Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science* **275**, 1784–1787.
- Kullander, K., and Klein, R. (2002). Mechanisms and functions of Eph and ephrin signalling. *Nat. Rev. Mol. Cell Biol.* **3**, 475–486.
- Lauren, P. (1965). The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at

- a histo-clinical classification. *Acta Pathology Microbiology Scand.* 64, 31–49.
- Lefebvre, O., Chenard, M.P., Masson, R., Linares, J., Dierich, A., LeMeur, M., Wendling, C., Tomasetto, C., Chambon, P., and Rio, M.C. (1996). Gastric mucosa abnormalities and tumorigenesis in mice lacking the p52 trefoil protein. *Science* 274, 259–262.
- Leung, S.Y., Chau, K.Y., Yuen, S.T., Chu, K.M., Branicki, F.J., and Chung, L.P. (1998). p53 overexpression is different in Epstein-Barr virus-associated and Epstein-Barr virus-negative carcinoma. *Histopathology* 33, 311–317.
- Leung, S.Y. *et al.* (2002). Phospholipase A2 group IIA expression in gastric adenocarcinoma is associated with prolonged survival and less frequent metastasis. *Proc. Natl. Acad. Sci. USA* 99, 16203–16209.
- Lustig, B. *et al.* (2002). Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. *Mol. Cell. Biol.* 22, 1184–1193.
- MacPhee, M., Chepenik, K.P., Liddell, R.A., Nelson, K.K., Siracusa, L.D., and Buchberg, A.M. (1995). The secretory phospholipase A2 gene is a candidate for the Mom1 locus, a major modifier of Apc-Min-induced intestinal neoplasia. *Cell* 81, 957–966.
- Nishi M, Omori. Y., Miwa, K., and editors, J. R. S. f. G. C. (1995). Japanese Classification of Gastric Carcinoma. Tokyo: Kanehara.
- Parkin, D.M., Pisani, P., and Ferlay, J. (1999). Estimates of the worldwide incidence of 25 major cancers in 1990. *Int. J. Cancer* 80, 827–841.
- Parsonnet, J., Friedman, G.D., Vandersteen, D.P., Chang, Y., Vogelstein, J.H., Orentreich, N., and Sibley, R.K. (1991). Helicobacter pylori infection and the risk of gastric carcinoma. *N. Engl. J. Med.* 325, 1127–1131.
- Perou, C.M. *et al.* (2000). Molecular portraits of human breast tumours. *Nature* 406, 747–752.
- Pomeroy, S.L. *et al.* (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* 415, 436–442.
- Riethmuller, G. *et al.* (1998). Monoclonal antibody therapy for resected Dukes' C colorectal cancer: seven-year outcome of a multicenter randomized trial. *J. Clin. Oncol.* 16, 1788–1794.
- Rosenwald, A. *et al.* (2001). Relation of gene expression phenotype to immunoglobulin mutation genotype in B cell chronic lymphocytic leukemia. *J. Exp. Med.* 194, 1639–1647.
- Rosenwald, A. *et al.* (2002). The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N. Engl. J. Med.* 346, 1937–1947.
- Sherlock, G. *et al.* (2001). The Stanford Microarray Database. *Nucleic Acids Res.* 29, 152–155.
- Shibata, D., and Weiss, L.M. (1992). Epstein-Barr virus-associated gastric adenocarcinoma. *Am. J. Pathol.* 140, 769–774.
- Shin, J.Y., Kim, H.S., Park, J., Park, J.B., and Lee, J.Y. (2000). Mechanism for inactivation of the KIP family cyclin-dependent kinase inhibitor genes in gastric cancer cells. *Cancer Res.* 60, 262–265.
- Sorlie, T. *et al.* (2001). Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc. Natl. Acad. Sci. USA* 98, 10869–10874.
- Toretzky, J.A., and Helman, L.J. (1996). Involvement of IGF-II in human cancer. *J. Endocrinol.* 149, 367–372.
- Troyanskaya, O., Cantor, M., Sherlock, G., Brown, P., Hastie, T., Tibshirani, R., Botstein, D., and Altman, R.B. (2001). Missing value estimation methods for DNA microarrays. *Bioinformatics* 17, 520–525.
- Troyanskaya, O.G., Garber, M.E., Brown, P.O., Botstein, D., and Altman, R.B. (2002). Nonparametric methods for identifying differentially expressed genes in microarray data. *Bioinformatics* 18, 1454–1461.
- van 't Veer, L.J. *et al.* (2002). Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 415, 530–536.
- Whitfield, M.L. *et al.* (2002). Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell* 13, 1977–2000.
- Wu, M.S., Shun, C.T., Wu, C.C., Hsu, T.Y., Lin, M.T., Chang, M.C., Wang, H.P., and Lin, J.T. (2000). Epstein-Barr virus-associated gastric carcinomas: relation to H. pylori infection and genetic alterations. *Gastroenterology* 118, 1031–1038.
- Yuen, S.T., Chung, L.P., Leung, S.Y., Luk, I.S., Chan, S.Y., and Ho, J. (1994). In situ detection of Epstein-Barr virus in gastric and colorectal adenocarcinomas. *Am. J. Surg. Pathol.* 18, 1158–1163.