ABSTRACT

Current epidemiological evidence supports a pathogenetic model of gastric cancer involving intermediate stages that include chronic gastritis and intestinal metaplasia. This study explores the molecular features of gastric cancer and premalignant stages using DNA microarray-based gene expression profiling and relates these findings to clinical, pathological, and ethnic parameters. A total of 124 tumor and adjacent mucosa samples were analyzed using spotted cDNA microarrays containing 9381 nonredundant gene elements. Tumor specimens were diffuse, intestinal, or mixed gastric cancer and adjacent mucosa, which generally displayed signs of chronic gastritis or intestinal metaplasia. Expression patterns could be discerned that readily defined premalignant and tumor subtypes. Chronic gastritis exhibits a pronounced mitochondrial gene expression signature, which may be linked to Helicobacter pylori pathogenesis. Intestinal metaplasia was associated with increased expression of many intestinal differentiation genes, many of which were not overexpressed in tumors. Samples were obtained from 91 Australian and 33 Chinese patients to explore potential variation in gene expression between these populations. Despite differences in the incidence, and potentially the etiology, of gastric cancer between these ethnic groups, we found the tumors to be molecularly similar. The identification of molecular signatures that are characteristic of subtypes of gastric cancer and associated premalignant changes should enable further analysis of the steps involved in the initiation and progression of this disease.

INTRODUCTION

GC is one of the most frequent malignant tumors in Asia, Eastern Europe, and South America, with an incidence in many regions ~10-fold that seen in Western countries. It is presently unclear whether the molecular pathology of GC is similar in Western, Asian, and Eastern European countries (1–3). GC has a poor prognosis in Western countries, including Australia and United States (4), although better results are obtained in Japan, possibly attributable to aggressive approaches to early detection and surgery (5).

The molecular understanding of GC subtypes is limited, and unlike some other solid tumors, molecular features of GC are not used to individualize treatment. All existing GC classification systems are based on histology (Bormann, Mings, Goseki, WHO, and Lauren classification). The Lauren system uses cell morphology to distinguish three broad GC subtypes: IGC and DGC, which roughly correspond to well and poorly differentiated subtypes, and MGC, representing a combination of the two (6, 7). IGC and DGC may result from the transformation of different epithelial cells (8) or distinct molecular changes in a common cell type. β-catenin mutation is more common in IGC (9), whereas mutation in CDH1 is more often seen in DGC, including familial GC (10–12); however, neither mutation is exclusive to either subtype. Chronic infection by Helicobacter pylori is commonly associated with all subtypes of GC but particularly to IGC. Infection results in ChG and is linked to GC pathogenesis (13, 14) through an intermediate stage involving IM (15, 16).

High throughput methods, such as cDNA and oligonucleotide microarrays, are increasingly being used to systematically compare molecular features of individual cancers to key clinical parameters (17–20). Previous expression profiling studies of GC highlighted differences in gene expression between tumor and adjacent mucosa (21–23) but failed to identify gene expression patterns discriminating histological subtypes. There is increasing interest in correlating clinical outcomes with gene expression signatures (24, 25). Hasegawa et al. (26) studied IGC specifically and correlated findings with lymph node status of tumors, whereas Leung et al. (27) describes PLA2G2A (Phospholipase A2 group IIA) expression in relation to patient survival. An additional study has correlated gains in chromosome 17q and gene expression, prompting study into possible mechanisms of GC progression (28). Our study uses a larger number of GC tumor and adjacent mucosa samples to focus on the relationship of gene expression pattern to clinical, pathological, and ethnic parameters.

MATERIALS AND METHODS

Patients and Samples. Tumor and normal specimens were collected from patients with GC requiring curative or palliative resection in Melbourne or in China (Shenyang or Dalian). Written informed consent was obtained before collection. Table 1 summarizes the characteristics of patient samples. Tumors were resected and divided (fresh frozen in liquid nitrogen and 70% ethanol), within 25 min ex vivo. All non-neoplastic samples were isolated from patients undergoing gastrectomy for GC. Mucosa ≥2 cm from the macroscopic tumor margin but from the same anatomical region was stripped of underlying muscle and serosa. In addition, cardia, body, fundus, and antral mucosa at maximal distance from the tumor were collected from 11 patients, pooled, and used to create a single batch of reference RNA.

Ethanol fixed tissue sections were stained with H&E and classified independently by two pathologists (S. L. and P. W.). These results were compared with the pathology record from the contributing institution. Final pathology was determined by consensus and review if necessary. Each specimen was attributed a diagnosis and scored for Lauren classification, differentiation, and percentage of tumor cells and inflammatory cells. Pathological Tumor-Node-Metastasis stage and demographic data were obtained from clinical records with patient permission. H. pylori serological testing was performed by PanBio using the PANBIO H. pylori IgG ELISA on preoperative serum collected on patients recruited by the Australian centers.

Microarray Analysis. Total RNA was isolated by phenol-chloroform extraction (TRIzol; Invitrogen) and column chromatography (RNeasy; Qiagen). Total RNA (50–70 μg) was labeled using reverse transcription with Super-
beled cDNA was cohybridized in 3.125 dCTP-Cy3 (Amersham) for reference, similar to previous methods (29). La-
script II (Invitrogen) and dCTP-Cy5 (Amersham) for test samples and with using a robotic arrayer (Virtek). Slides were washed in 0.5 to spotted cDNA arrays printed at the PMCI Microarray Core facility. These comprise Ha t4 2
Quantarray (GSI Lumonics).

AJCC, American Joint Committee on Cancer.

No data on lymph node status of one patient.
One adenosquamous and one mucinous. b
Age in years (range in years).
The order of sample processing was randomized to avoid systematic bias.

RESULTS AND DISCUSSION

Distinct Gene Expression Profiles Define Each Non-neoplastic and GC Subtype. Non-neoplastic samples were classified histologically as ChG or IM, and tumor samples were classified as IGC, DGC, and GC Subtype. G-3

Results and pathologic parameters of patient samples included for microarray analysis

<table>
<thead>
<tr>
<th>Tumor tissue parameters</th>
<th>No.</th>
<th>Pathology</th>
<th>Gender</th>
<th>Age</th>
<th>Tumor location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tumor cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>64.0 (32–83)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>73.7 (47–85)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pathology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intestinal</td>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diffuse</td>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adenosquamous</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor location</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GOI</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fundus</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greater curve-body</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lesser curve-body</td>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Body unspecified</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antrum</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Differentiation

Well: 4
Moderate: 13
Moderate to poor: 6
Poor: 20
Undifferentiated: 20
Other: 2

Depth of invasion

Mucosa: 1
Submucosa: 3
Muscularis propria: 13
Serosal/full thickness: 45

Lymph node status

N0 (<.5 Lymph nodes): 6
N0 (1–6): 10
N1 (7–15): 15
N1 (>15): 3
Unspecified: 1

AICC stage

IA: 3
IB: 9
II: 14
IIIA: 22
IIIB: 13
IV: 3
Unspecified: 1

Normal tissue parameters

Country

Australia: 41
China: 18
Pathology

Normal: 27
Chronic gastritis: 22
Squamous: 1

All tumor and normal RNA samples were competitively hybridized with the common reference RNA described above. In Immunohistochemistry, GC and premalignant tissues were ethanol fixed, paraffin embedded, and used to create tissue microarrays. Ki-67 (DAKO Corp., Carpinteria, CA; Ref. 30), E-cadherin (E-Cad; Zymed, San Francisco, CA; Ref. 31), and TFF1 (a gift from Dr A. Giraud) immunostaining was performed either using tissue arrays which were derived from the samples used for microarray analysis (Ki67) as described previously (32), or on an independent data set of samples (E-Cadherin and TFF1). Staining was revealed with a horseradish peroxidase linked secondary antibody using the DAKO LSAB+ kit (DAKO), following the manufacturer’s instructions.

Quantitative PCR. Nine representative genes (CDX1, BENE, MYO1A, CDH11, SFRP4, PDGFRB, DAB2, SPARCL1, and IGFBP7) were selected that were differentially expressed, based on hierarchical clustering of gene expression signatures. Quantitative expression analysis was performed with real-time PCR (Applied Biosystems, Foster City, CA), using oligonucleotide primers to known sequences. SYBR green intercalation to PCR products was estimated using an ABI PRISM 7000 (Applied Biosystems), according to the manufacturer’s specifications. TP63 and BIKE served as control genes, because these showed little variation in expression across the samples used in microarray experiments. Test genes were normalized to the average values obtained for the two controls. Sequences of primers were: TP63 forward primer, 5’-CCAAAG-GCAGGCACCCCTA-3’ and reverse, 5’-GGGAGATTAGGTGACATGAG-G-3’; BIKE forward primer, 5’-CTAAGCTGTTGA-CATTGCCT-3’ and reverse, 5’-GTGAGGTAGTCTGGCTG-3’; CDX1 forward primer 5’-CC-CTGACCTCT-GGACATG-3’ and reverse 5’-GTGAGTGAGGTGGGAGGTC-3’; BENE forward primer 5’-GCAGTGAGAGAAGTGCTG-3’ and reverse, 5’-TGTGCTTCCT-CATCTCCGC-3’; MYO1A forward primer, 5’-CTGAGGCAAGCATGTGAT-3’ and reverse, 5’-CACCCCGGTGCA-TTTTGG-3’; CDH11, forward primer, 5’-AACAGTGTGTTGTC-3’ and reverse, 5’-CTGTCGCTCTTTGTAACATG-3’; SFRP4 forward primer, 5’-CTTACAGGATGAGCCTTG-3’ and reverse, 5’-CA-TGCGCTTACATAAGTCGTC-3’; PDGFRB forward primer, 5’-TGTCCAGAAGTGAAGGCC-3’ and reverse, 5’-CTGAGCCCAAGGATGGA-3’; DAB2 forward primer, 5’-ACAAAGCTGATGACACAGACAG-3’ and reverse, 5’-CAAAGCGTGGA-AAGGGGCAG-3’; SPARCL1 forward primer, 5’-ATCATGTGCACTTCAAGGAA-AATGG-3’ and reverse, 5’-ACACGTAACCCAAAGAGGATCATC-3’; and IGFBP7 forward primer, 5’-ATGCTGGAGAATATGAGTGCCA-3’ and reverse, 5’-CTGAAGGCT-3’.

Data Analysis. Data were imported into GeneSpring (Silicon Genetics) and intensity-dependent normalization carried out using LOWESS (33). A group of 7385 genes on the satisfied filtering criteria based on the presence of significant signal in ≥80% of hybridizations. Hierarchical clustering was performed using Pearson correlation as a measure of similarity, after average linkage and median centering of values using the program Cluster, and viewed with Treeview (34). Discriminant genes and differences between groups were analyzed using two-tailed ANOVA with the Benjamini Hochberg multiple testing correction factor at P ≤ 0.01 (35), unless otherwise specified.

RESULTS AND DISCUSSION

Distinct Gene Expression Profiles Define Each Non-neoplastic and GC Subtype. Non-neoplastic samples were classified histologically as ChG or IM, and tumor samples were classified as IGC, DGC, or MGC. To systematically profile the spectrum of gene expression, we compared each sample to a common reference RNA. Independent hybridizations (192) were conducted, including 37 replicate experiments performed with separate slide batches and 31 direct comparisons of tumor to adjacent mucosa from the same patient. Hierarchical clustering revealed that all repeat samples clustered together (Supporting Information). Similarly, data from multiple clones for individual genes almost invariably culled, indicating a high degree of experimental reproducibility.

Unsupervised clustering of the data demonstrated a major division of samples into malignant and nonmalignant groups. Moreover, there was a highly structured partitioning of samples into

Internet address: http://www.ccgpm.org.

*Age in years (range in years).
*One adenosquamous and one mucinous.
*No data on lymph node status of one patient.
*AJCC, American Joint Committee on Cancer.
recognized histological groups of ChG, IM, IGC, DGC, or MGC (Fig. 1). MGC samples, which display both IGC and DGC elements histologically, fell predominantly at the interface of IGC and DGC clusters. Although expression profiling allowed good segregation of samples, several samples clustered into groups inconsistent with their histology, e.g., despite having expression signatures that overlapped with their correct group, samples ab077IM, ab156IM, ab135CG, and ab149N cosegregated in the DGC cluster. With the exception of ab156IM, the samples were derived from patients with DGC. These samples had no histologically detectable tumor cells but appear to have been misclassified because they share a strong stromal signature with DGC samples (Fig. 1, Region K, discussed below). To validate our RNA expression results, we selected nine genes and a representative cohort of samples to perform real-time PCR as a semiquantitative assessment of relative RNA abundance. Table 2 shows that there was an excellent con-

Fig. 1. Gastric cancer and adjacent non-neoplastic tissues readily segregate by gene expression patterns. A total of 124 experimental samples (Key) were sorted by hierarchical clustering of expression measurements made on 7383 genes and depicted as a dendrogram with all experiment numbers (A) and their corresponding pathological diagnosis (D, diffuse GC; I, intestinal GC; M, mixed GC; AS, adenosquamous cancer; N, normal; SN, normal squamous mucosa) and a gene expression matrix (B). Major clusters (A–K) are discussed in the text and subsequent figures. Arrow, non-neoplastic samples clustering in DGC. Normal (Key) indicates Chinese adjacent mucosa samples for which no detailed histological information was available, and these are denoted “normal” to indicate this fact. The scale bar indicates fold-difference relative to median level across all samples.
cordance of the results of real-time PCR analysis and microarray findings.

A clear distinction of tumor and adjacent mucosa has been observed in other GC expression profiling studies (22), but we believe this is the first to distinguish between DGC and IGC. The number of genes and specimens in our data set has allowed a high-resolution analysis of GC and its reported premalignant conditions.

GC is highly endemic in Asia and Eastern Europe, presumably because of differences in the balance of predisposing genes and/or environmental risk factors relative to Western countries. We were interested to determine whether the gene expression pattern of GCs arising in Chinese nationals was similar to those in Australians. We did not observe clear segregation of GC samples based on ethnicity using unsupervised clustering (Fig. 1) or ANOVA (data not shown), indicating that the molecular features of GC are similar in Chinese and Australian samples. Although we were able to find statistically significant differences in gene expression in non-neoplastic samples based on country of origin (Supporting Information), this analysis was confounded by unequal distribution of IM and ChG samples from each country.

ChG Is Characterized by a Mitochondrial Gene Expression Signature. ChG was defined by expression patterns of groups of genes sharing similar biological function, including metallothionein or elongation factor gene clusters (Figs. 1 and 2, Region B). In particular, there was a preponderance of nuclear genes encoding mitochondrial proteins, especially those involved in the oxidative production of ATP, including enzymes that degrade fatty acids and other oxidative substrates (Fig. 2, and Supporting Information). These included a large number of genes associated with the respiratory complex, such as subunits of cytochrome c oxidase (COX), NADH dehydrogenase.

### Table 2: Quantitative real-time PCR values and cDNA ratios for selected genes

<table>
<thead>
<tr>
<th>Samples</th>
<th>CDX1</th>
<th>BENE</th>
<th>MYO1A</th>
<th>CDH11</th>
<th>DAB2</th>
<th>SPARCL1</th>
<th>SFRP4</th>
<th>PDGFRB</th>
<th>IGFBP7</th>
</tr>
</thead>
<tbody>
<tr>
<td>ab094-ChG</td>
<td>0.01</td>
<td>0.32</td>
<td>0.06</td>
<td>0.28</td>
<td>0.04</td>
<td>0.38</td>
<td>0.27</td>
<td>0.64</td>
<td>0.03</td>
</tr>
<tr>
<td>ab169-ChG</td>
<td>0.00</td>
<td>0.29</td>
<td>0.02</td>
<td>0.29</td>
<td>0.02</td>
<td>0.54</td>
<td>0.04</td>
<td>0.67</td>
<td>0.13</td>
</tr>
<tr>
<td>ab069-IM</td>
<td>1.51</td>
<td>1.68</td>
<td>3.27</td>
<td>3.03</td>
<td>4.28</td>
<td>3.01</td>
<td>0.64</td>
<td>0.71</td>
<td>1.03</td>
</tr>
<tr>
<td>ab171-IM</td>
<td>0.89</td>
<td>1.24</td>
<td>2.44</td>
<td>2.60</td>
<td>3.92</td>
<td>3.63</td>
<td>0.26</td>
<td>0.78</td>
<td>0.53</td>
</tr>
<tr>
<td>ab076-DGC</td>
<td>0.08</td>
<td>0.33</td>
<td>0.41</td>
<td>0.47</td>
<td>0.31</td>
<td>0.38</td>
<td>1.49</td>
<td>1.18</td>
<td>1.61</td>
</tr>
<tr>
<td>ab121-DGC</td>
<td>0.34</td>
<td>0.41</td>
<td>0.81</td>
<td>0.60</td>
<td>0.56</td>
<td>0.52</td>
<td>3.17</td>
<td>1.62</td>
<td>2.49</td>
</tr>
<tr>
<td>ab173-DGC</td>
<td>0.30</td>
<td>0.55</td>
<td>0.85</td>
<td>0.75</td>
<td>0.70</td>
<td>0.56</td>
<td>1.43</td>
<td>1.87</td>
<td>1.24</td>
</tr>
<tr>
<td>ab172-IGC</td>
<td>0.18</td>
<td>0.32</td>
<td>0.67</td>
<td>0.47</td>
<td>0.07</td>
<td>0.25</td>
<td>0.73</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>ab059-IGC</td>
<td>0.94</td>
<td>2.22</td>
<td>0.24</td>
<td>0.55</td>
<td>0.09</td>
<td>0.37</td>
<td>0.40</td>
<td>0.60</td>
<td>0.46</td>
</tr>
<tr>
<td>ab108-IGC</td>
<td>5.75</td>
<td>2.65</td>
<td>1.23</td>
<td>0.96</td>
<td>0.04</td>
<td>0.35</td>
<td>1.56</td>
<td>1.07</td>
<td>1.32</td>
</tr>
<tr>
<td>ab172-IGC</td>
<td>0.18</td>
<td>0.32</td>
<td>0.67</td>
<td>0.47</td>
<td>0.07</td>
<td>0.25</td>
<td>0.73</td>
<td>0.85</td>
<td>0.85</td>
</tr>
</tbody>
</table>

* Normalized to control gene and mean centered for reverse transcription-PCR and cDNA arrays.
* Italics signify <0.7 fold.
* Bold signifies >1.4 fold.
Infection with *H. pylori* is almost an invariant feature of ChG in patients with GC (36), suggesting that the ChG signature may be linked to this microbe. *H. pylori* infection was determined by histology of pathological samples and serological testing of preoperative blood samples collected from Australian centers. We found that 80% of all Australian patients were found to be positive for *H. pylori*. Strains of *H. pylori* are diverse, and several genetic loci, including the cag pathogenicity island, *babA2*, and *VacA* genes, have been implicated in the pathogenesis of GC (37–40). These genes produce a number of important pathogenic factors, including CagA, which undergoes tyrosine phosphorylation after type IV secretion into the epithelial cell (41) and deregulates SHP-2 (42) and VacA, which induces mitochondrial damage when applied to gastric epithelial cells (43). The p34 fragment of VacA localizes specifically to mitochondria, causing release of cytochrome c and apoptosis (44). Ultrastructural studies have also demonstrated that mitochondrial damage and loss are notable features of ChG (45). Our findings are consistent with loss or damage of mitochondria in ChG, leading to a high level of mitochondrial biogenesis. *H. pylori* is a neutrophile (46), and there may be an adaptive advantage in attenuating ATP production required by the proton pump for acid secretion or in triggering apoptosis of the acid-producing parietal cells through the release of cytochrome c. Mitochondrial damage may be important for *H. pylori* establishment and growth, because secretion of urease by *H. pylori* appears to be insufficient to adequately neutralize the local environment (46). At least two recent studies have investigated acute *H. pylori* infection of AGS cells *in vitro*. Although interesting from the point of view of gene expression changes in immortalized cultured cells, these studies acknowledge the need for *in vivo* experiments (47, 48). Our findings reflect chronic infection with *H. pylori* in an *in vivo* setting, rather than acute infection over hours as described by these studies. Although our findings highlight an impact on mitochondria, among the many postulated effects of *H. pylori* infection (49), expression analysis of normal and gastric mucosa from infected and uninfected individuals is needed to determine this conclusively.

**IM Reflects Expression of Intestinal Genes and Transition toward a Transformed Phenotype.** Consistent with the acquisition of an intestinal phenotype, the IM group showed prominent expression of genes characteristic of intestine (Figs. 1 and 2, Regions E and F). These include *CDX1* (a homeobox gene that regulates gut morphogenesis and differentiation and a recognized marker of IM), *MYO1A* (encodes a brush border myosin found in intestinal epithelium), *MTP or microsomal triglyceride transfer protein* (encodes an enzyme involved in lipid metabolism), *cholerectokinin* (encodes a gut hormone), and *Villin1* (a major component of intestinal microvilli; Fig. 2, Region F). Interestingly, many of the markers of mature intestinal cell function are lost or show diminished expression in IGC, possibly reflecting a less differentiated state of tumor compared with metaplastic tissue. Other genes of interest that define the IM group include *FAT* (Fig. 2, Region F), a putative tumor suppressor, and *TFF1* (Fig. 2, Region E). There was a marked attenuation of *TFF1* expression in more than half of all DGC and IGC. Somatic mutations in *TFF1* are implicated in gastric lesions (50), and *TFF1* knockout mice develop GC (51). We stained tissue microarrays with TFF1 in all DGC and IGC. Region F exemplifies genes that are up-regulated in both IM and IGC. PSG proteins have been shown to interact with Kruppel-like factor 4 in placental development, and KLF-4 showed similar expression with PSG in our samples. *TGFα* was persistently up-regulated in IM and IGC (Supporting Information, Fig. 1 Region E). Given its role as a mitogen (52), *TGFα* may play a role in the progression from IM to IGC. It remains to be addressed whether genes in region E, such as *TGFα*, represent markers of terminal differentiation of intestinal epithelium that retain expression after transformation or are early participants in the transformation process.

In Fig. 3, we used ANOVA with the Benjamini Hochberg multiple testing correction factor (*P* < 0.01) as an alternative algorithm to identify genes that are most significantly differentially expressed between ChG and IM. Of 189 genes discovered by ANOVA, almost all were present in the regions B, E, and F in Fig. 2, providing validation of our approach to identifying differentially expressed genes between ChG and IM.

**Intestinal Type GC Was Characterized by Markers of Proliferation.** A notable feature of the IGC samples was the relatively strong expression of proliferative markers (Figs. 1, Regions D1–D3 and 4). These included a large number of genes required for G2–M transition, DNA replication, spindle assembly, and chromosome segregation. The list includes genes shown previously to be overexpressed in GC and other cancers, such as *topoisomerase* (TOP2A; Ref. 53) and *CDC25B* (54–56). Although it is expected that tumor cells should express genes involved in cell division, it was noteworthy that a number of the DGC samples did not show a similar elevation of these markers (Figs. 1 and 4). In principle, the infiltrative nature of DGC tumors and consequently lower propor-
tion of tumor cells in DGC samples could result in dilution of proliferative markers. However, staining of tissue arrays containing the corresponding samples with anti-Ki-67 antibodies showed that the IGC group generally had a higher proliferation score compared with DGC (Figs. 4 and 5).

Rapidly dividing cells might also be expected to up-regulate protein synthetic machinery to meet the demands of cell growth; however, examination of Fig. 1 suggests that is not the case. Region C comprised a tightly coregulated group of ribosomal protein genes, which were more strongly expressed in DGC and ChG than in IGC (Fig. 4). The expression patterns of a largely overlapping set of ribosomal subunit genes have been found to vary substantially between different tumor types (57), indicating that not all tumors strongly express genes involved in protein synthesis. Furthermore, expression profiling of germinal center B cells, which are among the most rapidly proliferating cells in the body, has revealed that high level expression of cell cycle markers was associated with relatively low level expression of many of the same ribosomal subunit genes present in region C (58). It appears there are circumstances where proliferation occurs at the expense of cell growth, such as in germinal center B cells and IGC.

Our findings accord with the high nuclear:cytoplasmic ratio characteristic of malignant cells and typical of IGC.

The DGC Expression Signature Reflected Active Extracellular Matrix Production and Remodeling and Complex Signaling by Positive and Negative Regulators of Cell Growth. DGC is an infiltrating cancer characterized by a marked stromal reaction (6). This was reflected in the expression profile, which included genes encoding components of the extracellular matrix (collagens, biglycan, osteoglycin, proteoglycan, matrix metalloproteinases, cadherin 11, Thy-1 SERPINS, and fibrillin), as well as smooth muscle and cell adhesion molecules (Figs. 1 and 4, Region K; also in Supporting Information Regions K1, K2, L, and M). A number of genes encoding extracellular proteins was up-regulated in DGC that are linked to positive and negative regulation of cell proliferation, including secreted frizzled-related protein 4, NMB, DAB2, SPARC, Wnt5a, DLG5, DKK3, FDZ1, and IGFBP7. Increased SPARC expression has been linked to more aggressive tumor types in many tissues (59, 60) but has not been implicated previously in GC. One study shows abrogation of tumorigenicity by inhibiting SPARC function in human melanoma cells (61). Wnt proteins are
involved in cell transformation and signal transduction via frizzled and frizzled-related receptors. Consistent with a previous report (62), \textit{Wnt5a} was up-regulated in both DGC and IGC (Fig. 4). Surprisingly, \textit{dickkopf 3} (\textit{DKK3}), a putative negative regulator of \textit{Wnt} signaling through the \textbeta-catenin pathway (63–66), and the \textit{Wnt}-signaling protein frizzled homologue 1 (\textit{FDZ1}) clustered together. \textit{Discs large homologue 5} (\textit{DLG5}), a putative negative regulator of cell growth (Fig. 4) cosegregated with \textit{DKK3} and \textit{FDZ1}. \textit{DAB2} is also implicated as a negative regulator of growth (67). The co-overexpression of both positive and negative regulators of cell growth in DGC samples suggests a complex interplay between their stromal and epithelial elements. There was a remarkable concordance between a cluster of genes expressed in Region K and a gene cluster associated with \textit{SPARC} in breast cancer (57). It seems likely that this group of genes reflects a characteristic signature associated with the reaction of stroma to infiltrating tumor cells.

Down-regulation of \textit{E-cadherin} (\textit{CDH1}), which participates in cell-cell contact through the formation of tight junctions, is a recognized feature of sporadic DGC (10). In addition, germ-line mutations in \textit{CDH1} are associated with high penetrance familial DGC (11). Consistent with this, we found a significant attenuation of \textit{CDH1} expression in DGC relative to other tissues (Fig. 4) but not in IM. Immunohistochemical staining of tissue microarrays with \textit{E-cadherin} antibodies confirmed the pattern of expression seen in our microarray experiments (Fig. 5). This finding agrees with previous studies (68). \textit{Occludin}, a major membrane component of tight junctions, and \textit{actin-related protein 2/3}, coclustered with \textit{CDH1} and were similarly down-regulated in DGC. It is recognized that a loss of \textit{Occludin} and \textit{carcinoembryonic antigen 1} has been associated with loss of cell polarity in prostate cancer and could be involved in tumorigenesis (69). \textit{Occludin} is expressed in well-differentiated intestinal epithelium (70), consistent with our findings in IM (Fig. 4, \textit{E-Cad Region}). It will be interesting to test whether these genes or other components of tight junctions are mutated in GC and result in function analogous to \textit{CDH1} mutations.

A group of genes that showed pronounced differential expression between samples but independent of histological subtype were genes involved in inflammation (Fig. 1, Regions H–J). These regions included a comprehensive set of B- and T-cell markers, pro-inflammatory cytokines, and mediators of interferon signaling and interferon-induced genes (Fig. 1, and see Fig. 4 for a subregion of H). These expression changes were consistent with infiltration of plasma cells histologically (data not shown).

In summary, we found that expression profiling of a large set of GC and non-neoplastic specimens was able to distinguish premalignant and malignant subtypes of GC, and the structure of the major gene clusters reflect identifiable biological functions. It is noteworthy that the structure of the cluster reflects a characteristic signature associated with the reaction of stroma to infiltrating tumor cells.

Fig. 5. Representative tissue sections stained with antibodies to Ki-67, E-cadherin, and TFF1. \textit{Panels} show cell-specific expression of gene products reflecting the results of microarray experiments. Ki67 antibody staining shows high levels of expression in IGC, relative to other stages (×20 magnification). Down-regulation of E-cadherin and TFF1 in DGC was consistent with the RNA expression pattern in Figs. 4 and 3, respectively (×40 magnification). Scale bars, 50 μm.
GENE EXPRESSION IN GASTRIC CANCER AND ITS PRECURSORS

ACKNOWLEDGMENTS

We thank members of the Bowtell Lab in assistance with manuscript preparation and Dr. A. Giraud for providing the anti-TFF1 antibody.

REFERENCES


GENE EXPRESSION IN GASTRIC CANCER AND ITS PRECURSORS


Distinctive Patterns of Gene Expression in Premalignant Gastric Mucosa and Gastric Cancer

Alex Boussioutas, Hong Li, Jia Liu, et al.


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/63/10/2569

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2006/06/08/63.10.2569.DC1

Cited articles
This article cites 68 articles, 28 of which you can access for free at:
http://cancerres.aacrjournals.org/content/63/10/2569.full.html#ref-list-1

Citing articles
This article has been cited by 21 HighWire-hosted articles. Access the articles at:
/content/63/10/2569.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.