Oncogene (2004) 23, 6830–6844

www.nature.com/onc

Comparison of gene-expression profiles between diffuse- and intestinal-type gastric cancers using a genome-wide cDNA microarray

Natini Jinawath[1], Yoichi Furukawa[1], Suguru Hasegawa[1], Meihua Li[1], Tatsuhiko Tsunoda[2], Seiji Satoh[3], Toshiharu Yamaguchi[4], Hiroshi Imamura[5], Masatomo Inoue[6], Hitoshi Shiozaki[7] and Yusuke Nakamura*[,1]

[1]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Tokyo, Japan; [2]SNP Research Center, RIKEN (Institute of Physical and Chemical Research), Tokyo, Japan; [3]Department of Gastroenterological Surgery, Kyoto University Graduate School of Medicine, Kyoto, Japan; [4]Department of Surgery, Cancer Institute Hospital, Tokyo, Japan; [5]Department of Surgery, Sakai City Hospital, Osaka, Japan; [6]Department of Surgery, Nara Hospital, Kinki University, School of Medicine, Nara, Japan; [7]Department of Surgery, Sayama Hospital, Kinki University, School of Medicine, Osaka, Japan

Gastric cancer is the fourth leading cause of cancer-related death in the world. Two histologically distinct types of gastric carcinoma, 'intestinal' and 'diffuse', have different epidemiological and pathophysiological features that suggest different mechanisms of carcinogenesis. A number of studies have investigated intestinal-type gastric cancers at the molecular level, but little is known about mechanisms involved in the diffuse type, which has a more invasive phenotype and poorer prognosis. To clarify the mechanisms that underlie its development and/or progression, we compared the expression profiles of 20 laser-microbeam-microdissected diffuse-type gastric-cancer tissues with corresponding noncancerous mucosae by means of a cDNA microarray containing 23 040 genes. We identified 153 genes that were commonly upregulated and more than 1500 that were commonly downregulated in the tumors. We also identified a number of genes related to tumor progression. Furthermore, comparison of the expression profiles of diffuse-type with those of intestinal-type gastric cancers identified 46 genes that may represent distinct molecular signatures of each histological type. The putative signature of diffuse-type cancer exhibited altered expression of genes related to cell–matrix interaction and extracellular-matrix (ECM) components, whereas that of intestinal-type cancer represented enhancement of cell growth. These data provide insight into different mechanisms underlying gastric carcinogenesis and may also serve as a starting point for identifying novel diagnostic markers and/or therapeutic targets for diffuse-type gastric cancers.
Oncogene (2004) **23,** 6830–6844. doi:10.1038/sj.onc.1207886
Published online 26 July 2004

Keywords: diffuse-type gastric cancer; cDNA microarray; laser-microbeam microdissection; expression profile

*Correspondence: Y Nakamura, Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan; E-mail: yusuke@ims.u-tokyo.ac.jp
Received 27 November 2003; revised 12 May 2004; accepted 14 May 2004; published online 26 July 2004

Introduction

Gastric cancer is the fourth most frequent cancer worldwide, accounting for 10.4% of cancer deaths in 2000; Japan has the highest age-standardized incidence of this disease (Parkin, 2001). Since the 5-year survival rate of patients who are diagnosed at an advanced stage is generally less than 10% (Peddanna *et al*., 1995), identification of sensitive diagnostic markers and development of novel therapeutic modalities other than surgery are of some urgency.

Gastric cancers are histologically classified into diffuse type (infiltrating, poorly differentiated, noncohesive cancer cells with vast fibrous stroma) and intestinal type (cohesive, glandular-like cell groups) (Lauren, 1965). These two types have different epidemiology, etiology, pathogenesis and biological behavior. The diffuse type occurs in relatively younger individuals regardless of gender, and often metastasizes to peritoneum or lymph nodes with a poorer prognosis as a result. *Helicobacter pylori*, which tends to be associated with gastric cancer (Sipponen, 1995) might not have a role in diffuse-type gastric carcinogenesis because prolonged infection leads to intestinal metaplasia, a precursor lesion of intestinal-type gastric cancer. The precursor lesion of diffuse-type gastric cancer is proposed to arise from hyperplastic neck cells, but that has not yet been proven (Ming, 1998).

Molecular investigations have identified multiple genetic alterations in gastric carcinomas including point mutations, deletions, loss of heterozygosity and microsatellite instability. Notably, mutations in *APC*, *CTNNB1* (*β*-catenin) and/or *p53* genes (Craanen *et al*., 1995; Park *et al*., 1999) are more frequently associated with the intestinal type, whereas aberrant expression of *CDH1* (encoding E-cadherin) is often a feature of diffuse-type gastric cancers; somatic mutations of *CDH1,* or methylation in its promoter region, occur in about 50% of these tumors (Becker *et al*., 1994). Additionally, a limited number of genes including *IQGAP1* (Takemoto *et al*., 2001), *TGFβ1*, *TGFβRII* (Chung *et al*., 1997; Maehara *et al*., 1999),

HGF and *HGFR* (c-met) (Lee *et al*., 2000), have been reported to show altered expression in diffuse-type gastric cancers.

Since most cancer cells in the diffuse type are scattered and accompanied by marked stromal reactions such as desmoplasia and lymphocytic infiltrates, bulk tissues would necessarily yield expression data from a mixture of cancer cells and abundant noncancerous cells. However, laser-microdissection technology now allows us to obtain precise expression patterns in a pure population of diffuse-type gastric-cancer cells. Although several studies of gene expression in diffuse-type gastric cancer have been reported (Hippo *et al*., 2001; Hippo *et al*., 2002; Leung *et al*., 2002; Boussioutas *et al*., 2003; Chen *et al*., 2003; Kim *et al*., 2003; Tay *et al*., 2003), none of the studies used laser microdissection for the analysis. Thus, our data is the first genome-wide expression profiles of pure population of diffuse-type gastric-cancer cells.

In a previous study, we carried out a genome-wide analysis of gene-expression profiles of 20 intestinal-type gastric-cancer tissues using a cDNA microarray containing 23 040 genes (Hasegawa *et al*., 2002). In those experiments, a large number of genes showed altered expression in the cancer tissues compared to corresponding noncancerous gastric mucosae. Since diffuse- and intestinal-type gastric cancers show distinctly different clinicopathological features in many aspects, we were encouraged to compare gene-expression profiles of diffuse-type tumors with those obtained earlier in intestinal-type gastric cancers.

For the present study, we used the same cDNA microarray system and microdissected cell populations from 20 diffuse-type gastric cancers. We identified genes showing altered expression in this type of tumor including some associated with tumor progression. We were also able to identify numerous genes that were differentially expressed between diffuse- and intestinal-type gastric cancers. The data presented here provide important information regarding the mechanisms of gastric carcinogenesis and may contribute to the development of type-specific diagnostic markers and/or therapeutic targets.

Results

Identification of genes commonly up- or downregulated in diffuse-type gastric cancers

Comparison of the expression profiles of 20 tumors with their corresponding noncancerous mucosae identified 153 genes (including 23 of unknown function) that were upregulated (Supplementary Material: 1) (Table 1 and http://www.ims.u-tokyo.ac.jp/nakamura/furukawa/microarray.html) and 1553 (including 395 of unknown function) that were downregulated (Supplementary Material: 2) (Table 2 and http://www.ims.u-tokyo.ac.jp/nakamura/furukawa/microarray.html) in 50% or more of the samples examined. The upregulated genes represented a variety of functions including genes associated with signal-transduction pathways (*S100A10*, *ANXA1*, *GNAI2*, *LY6E*, *RAI3* and *PLAB*), genes encoding transcription factors (*TGIF2* and *ETV4*), or genes involved in various metabolic pathways (*GPX1, BACH*, *NNMT* and *GNPI*), transport systems (*ATP1B3* and *KCNA3*), cell proliferation (*TGFBI*), apoptosis (*CARD4* and *BIRC5*), protein translation and RNA processing (*PRKDC*, *HSPCB* and *FBL)*, cell-cycle regulation (*PRC1*, *CDC25B* and *CDC20*), cell adhesion and cytoskeleton (*ZYX*, *LCP1*, *ARHGDIB* and *MSLN*), cell motility and extracellular matrix (*CD81, MMP7*, *SPARC*, *COL3A1* and *FN1*), immunity (*MIF*, *IFITM2* and *G1P2*), and other functions (*CTSB*).

Commonly downregulated elements included genes associated with various metabolic pathways (*ALDH3A1*, *GSTA1*, *FBP1*, *CA2, AKR1C3* and *CYP3A7*), small-molecule or heavy-metal transport (*ATP2A3*, *GIF*, *MT1F* and *SLC7A8)*, defense response (*TFF1*, *TFF2* and *GSTA3*), immunity (*LTF, IL1R2, FCGBP* and *C5*), signal transduction and cell-cycle regulation (*MAL* and *ERBB2IP*), cell proliferation (*PAP* and *REG1A*) or other functions (*SYTL2* and *PGC*).

To validate our microarray data, we selected eight representative genes (*TGFBI, SPARC, COL3A1*, *MSLN, FLJ20736, GW112* and two ESTs) that were commonly upregulated in diffuse-type tumors, and performed semiquantitative RT–PCR using eight pairs of the same RNA samples that had served for the microarray analysis (Figure 1). Among the 42 microarray data that showed normalized Cy3 and Cy5 signal intensities above cutoff value, 39 were consistent with those obtained by semiquantitative RT–PCR. Hence, we estimated the concordance between the microarray and semiquantitative RT–PCR data to be 92.85%, corroborating the high reliability of our microarray data.

Cluster analysis of diffuse- and intestinal-type gastric cancers

To arrange the samples and genes on the basis of similarities, we carried out a two-way unsupervised hierarchical clustering algorithm using a total of 1051 genes that had passed the first filter described in Materials and methods. The cluster analysis classified the 40 tumor samples into two major classes that correctly corresponded to their histological subtypes, namely 20 diffuse- and 20 intestinal-type gastric cancers (Figure 2). The data indicated that diffuse-type cancer has a distinctly different expression profile from that of intestinal-type cancer. Regarding the gene axis, the algorithm classified the genes into several clusters that represented different expression signatures among the tumors. We identified clusters that included a number of genes associated with similar biological processes. For example, clusters A, B and C, which contained genes that were relatively upregulated in diffuse-type cancers, included a number of genes associated with cell adhesion or migration and the extracellular matrix (ECM), for example, *SPARC*, *TUBB2*, *ITGB1*, *ARHGDIA,*

Table 1 The 75 most upregulated genes in diffuse-type gastric cancer[a]

No.	*Accession no.*[c]	*Symbol*	*Title*	*Function*[b]	*Median*[d]
1	J03040	*SPARC*	Secreted protein, acidic, cysteine-rich (osteonectin)	ECM remodeling, cell motility	17.02495567
2	D49441	*MSLN*	Mesothelin	Cell adhesion	16.89094109
3	X02761	*FN1*	Fibronectin 1	Cell adhesion, migration, metastasis	16.75009195
4	AW298180	*MMP7*	Matrix metalloproteinase 7 (matrilysin, uterine)	ECM remodeling, metastasis	16.50926608
5	U18018	*ETV4*	ets variant gene 4 (E1A enhancer binding protein, E1AF)	Transcription factor, oncogenesis	16.05115993
6	M13690	*SERPING1*	Serine (or cysteine) proteinase inhibitor, clade G (C1 inhibitor), member 1, (angioedema, hereditary)	Complement activation	14.85969213
7	NM_006169	*NNMT*	Nicotinamide *N*-methyltransferase	Nicotinamide metabolism	5.200336287
8	Z68179	*LY6E*	Lymphocyte antigen 6 complex, locus E	Signal transduction	4.592021044
9	J03464	*COL1A2*	Collagen, type I, alpha 2	ECM	4.538226354
10	X03963	*COL4A1*	Collagen, type IV, alpha 1	ECM	4.514567118
11	AI684645	*HCA112*	Hepatocellular carcinoma-associated antigen 112	Tumor Ag	4.257192083
12	AV758898	MGC27165	Hypothetical protein MGC27165	Unknown	4.025887854
13	K02215	*AGT*	Angiotensinogen (serine (or cysteine) proteinase inhibitor, clade A, member 8)	Cell–cell signaling, hormone activity	3.529505325
14	AL138409		Homo sapiens mRNA; cDNA DKFZp313L231 (from clone DKFZp313L231)	Unknown	3.521028423
15	Y14735	*IGHG3*	Immunoglobulin heavy constant gamma 3 (G3 m marker)	Immune response	3.4029154
16	X14420	*COL3A1*	Collagen, type III, alpha 1 (Ehlers–Danlos syndrome type IV, autosomal dominant)	ECM	3.135499176
17	AF044588	*PRC1*	Protein regulator of cytokinesis 1	Cell-cycle regulation	2.948876913
18	AF097021	*GW112*	Differentially expressed in hematopoietic lineages	Unknown	2.873366561
19	AJ011497	*CLDN7*	Claudin 7	Membrane integrity	2.738268954
20	AK023969	*CARD4*	Caspase recruitment domain family, member 4	Apoptosis, signal transduction	2.701428204
21	AA789233	*COL1A1*	Collagen, type I, alpha 1	ECM	2.69564816
22	S55551	*TPSB2*	Tryptase beta 2	Inflammation	2.681107922
23	D78013	*DPYSL2*	Dihydropyrimidinase-like 2	Signal transduction, nucleic acid metabolism	2.519226943
24	NM_006435	*IFITM2*	Interferon-induced transmembrane protein 2 (1–8D)	Immune response	2.50940959
25	AJ002231	*GNPI*	Glucosamine-6-phosphate isomerase	Glucosamine metabolism	2.345160429
26	AA742701	*LCP1*	Lymphocyte cytosolic protein 1 (L-plastin)	Cytoskeleton	2.31821868
27	BE621666	*LAPTM4B*	Lysosomal-associated protein transmembrane 4 beta	Membrane integrity	2.300573197
28	M77349	*TGFBI*	Transforming growth factor, beta-induced, 68 kDa	Cell proliferation, cell adhesion	2.215203495
29	AF001601	*PON2*	Paraoxonase 2	Organophosphate catabolism	2.129971645
30	M12670	*TIMP1*	Tissue inhibitor of metalloproteinase 1 (erythroid potentiating activity, collagenase inhibitor)	ECM remodeling	2.092756794
31	NM_007274	*BACH*	Brain acyl-CoA hydrolase	lipid metabolism	2.073257092
32	M13755	*G1P2*	Interferon, alpha-inducible protein (clone IFI-15 K)	Immune response	1.919872602
33	AF026166	*CCT2*	Chaperonin containing TCP1, subunit 2 (beta)	Cell-cycle regulation	1.917393514
34	BF727014	*S100A10*	S100 calcium binding protein A10 (annexin II ligand, calpactin I, light polypeptide (p11))	Signal transduction, calcium binding protein	1.880840314
35	AA625198		ESTs, Weakly similar to S10889 proline-rich protein – human [H.sapiens]	Unknown	1.880031974
36	NM_000310	*PPT1*	Palmitoyl-protein thioesterase 1 (ceroid-lipofuscinosis, neuronal 1, infantile)	Lipid-modified protein catabolism	1.878755527
37	X05610	*COL4A2*	Collagen, type IV, alpha 2	ECM	1.878270024
38	D42073	*RCN1*	Reticulocalbin 1, EF-hand calcium binding domain	Calcium binding protein	1.841472524
39	NM_005787	*NOT56L*	Not56 (*D. melanogaster*)-like protein	Membrane integrity	1.834076317
40	AB042646	*TGIF2*	TGFB-induced factor 2 (TALE family homeobox)	TGFbeta-induced transcription factor	1.819070249
41	NM_000700	*ANXA1*	Annexin A1	Cell motility, signal transduction	1.814148773
42	NM_003979	*RAI3*	Retinoic acid-induced 3	Signal transduction	1.813389648
43	NM_001436	*FBL*	Fibrillarin	RNA processing	1.775163278
44	U44839	*USP11*	Ubiquitin-specific protease 11	Unknown	1.769244784
45	AA706499		Homo sapiens cDNA FLJ32269 fis, clone PROST1000526.	Unknown	1.7622519
46	M33680	*CD81*	CD81 antigen (target of antiproliferative antibody 1)	Cell motility, signal Transduction	1.736299903
47	AA536113	*TMEPAI*	Transmembrane, prostate androgen-induced RNA	TGFbeta signaling pathway	1.695098803
48	J03004	*GNAI2*	Guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 2	Oncogenesis, signal transduction	1.686220648
49	D21853	KIAA0111	KIAA0111 gene product	Unknown	1.680174948
50	X03212	*KRT7*	Keratin 7	Cytoskeleton	1.671195308
51	AI127415	*PRKDC*	Protein kinase, DNA-activated, catalytic polypeptide	Protein modification	1.598914278
52	NM_000581	*GPX1*	Glutathione peroxidase 1	Glutathione metabolism	1.59105557
53	NM_002085	*GPX4*	Glutathione peroxidase 4 (phospholipid hydroperoxidase)	Glutathione metabolism	1.589020836
54	X94991	*ZYX*	Zyxin	Cell adhesion	1.579104324
55	AK024245	FLJ20736	Hypothetical protein FLJ20736	Unknown	1.570844127
56	J04208	*IMPDH2*	IMP (inosine monophosphate) dehydrogenase 2	Purine biosynthesis	1.567535435
57	AF077350	*BIRC5*	Baculoviral IAP repeat-containing 5 (survivin)	Apoptosis inhibitor	1.544943286
58	NM_021874	*CDC25B*	Cell division cycle 25B	Cell cycle regulation, oncogenesis	1.539860782
59	AK022241	FLJ10774	Hypothetical protein FLJ10774	Unknown	1.516671984

Table 1 (*continued*)

No.	*Accession no.*[c]	*Symbol*	*Title*	*Function*[b]	*Median*[d]
60	NM_001255	*CDC20*	CDC20 cell division cycle 20 homolog (*S. cerevisiae*)	Cell cycle regulation	1.511488165
61	AA143060	FLJ22283	Hypothetical protein FLJ22283	Unknown	1.508073355
62	L20688	*ARHGDIB*	Rho GDP dissociation inhibitor (GDI) beta	Cytoskeleton	1.491020602
63	N30179	*PLAB*	Prostate differentiation factor	TGFbeta receptor signaling pathway	1.482028239
64	NM_001679	*ATP1B3*	ATPase, Na+/K+ transporting, beta 3 polypeptide	Electron transportation	1.467881262
65	AW957113		Homo sapiens mRNA full-length insert cDNA clone EUROIMAGE 1499812	Unknown	1.460990905
66	AK027260	KIAA1268	KIAA1268 protein	Unknown	1.45155437
67	M61831	*AHCY*	*S*-adenosylhomocysteine hydrolase	Carbon compound metabolism	1.448553035
68	AA143048	DKFZP564-O0463	DKFZP564O0463 protein	Unknown	1.443429831
69	AI273886	*HSPCB*	Heat-shock 90 kDa protein 1, beta	Heat shock protein	1.436606731
70	D16294	*ACAA2*	Acetyl-Coenzyme A acyltransferase 2 (mitochondrial 3-oxoacyl-Coenzyme A thiolase)	Lipid metabolism	1.430780056
71	BE793569	*NME1*	Nonmetastatic cells 1, protein (NM23A) expressed in	Nucleic acid metabolism	1.398911778
72	M85217	*KCNA3*	Potassium voltage-gated channel, shaker-related subfamily, member 3	Potassium ion transportation	1.394269542
73	Y10805	*HRMT1L2*	HMT1 hnRNP methyltransferase-like 2 (*S. cerevisiae*)	Methylation	1.380666396
74	L16510	*CTSB*	Cathepsin B	Lysosome, proteolysis	1.378887483
75	D00172	*ANXA5*	Annexin A5	Signal transduction	1.334294834

[a]The 75 representative upregulated genes were sorted by their expression ratio ranging from the highest. [b]Gene functions were summarized from literature sources or according to LocusLink in NCBI (www.ncbi.nlm.nih.gov/LocusLink). [c]GenBank Accession number. [d]Median of log 2 transformed expression ratio of diffuse-type gastric cancer cases

Table 2 The 75 most downregulated genes in diffuse-type gastric cancer[a]

No.	*Accession no.*[c]	*Symbol*	*Title*	*Function*[b]	*Median*[d]
1	AW008222	*PGC*	Progastricsin (pepsinogen C)	Enzyme (precursor of pepsin)	−20.14449893
2	AB039886	*AMP18*	18 kDa antrum mucosa protein	Unknown	−19.76923519
3	M63154	*GIF*	Gastric intrinsic factor (vitamin B synthesis)	Ion transportation	−19.40546803
4	M84337	*PAP*	Pancreatitis-associated protein	Cell proliferation, cell adhesion	−18.78530191
5	U07643	*LTF*	Lactotransferrin	Humoral immune response	−18.55183966
6	NM_002371	*MAL*	Mal, T-cell differentiation protein	Signal transduction	−18.13945316
7	NM_001869	*CPA2*	Carboxypeptidase A2 (pancreatic)	Proteolysis, vacuolar protein catabolism	−17.97663684
8	AA992910	*CTXL*	Cortical thymocyte receptor (*X. laevis* CTX) like	Membrane integrity	−17.08322999
9	BE719697	LOC130576	Hypothetical protein LOC130576	Unknown	−16.7612826
10	N51406	FLJ14503	Hypothetical protein FLJ14503	Unknown	−16.14088801
11	AV704982	*GSTA1*	Glutathione *S*-transferase A1	Carcinogen and toxic metabolism	−16.0913734
12	M57951	*UGT1A4*	UDP glycosyltransferase 1 family, polypeptide A4	Carbohydrate metabolism	−16.00923816
13	AI340056		ESTs	Unknown	−16.00834108
14	L76465	*HPGD*	Hydroxyprostaglandin dehydrogenase 15-(NAD)	Prostaglandin metabolism	−15.95505956
15	AA400080		ESTs	Unknown	−15.67843787
16	AK025909		Homo sapiens cDNA: FLJ22256 fis, clone HRC02860.	Unknown	−15.12654316
17	AA788924	*C5*	Complement component 5	Immune response	−14.84904567
18	T95199	*MT1F*	Metallothionein 1F (functional)	Metal ion binding	−13.91008053
19	AF043498	*PSCA*	Prostate stem cell antigen	Tumor Ag	−10.26053898
20	AA741431	*TFF2*	Trefoil factor 2 (spasmolytic protein 1)	Defense response	−10.10803184
21	AA522445	*SYTL2*	Synaptotagmin-like 2	Neurotransmission regulator	−9.563551832
22	AI247475	*AKR1C2*	Aldo–keto reductase family 1, member C2	Electron transportation	−7.91861773
23	AI627636		Homo sapiens cDNA FLJ32260 fis, clone PROST1000334.	Unknown	−7.73673358
24	U37100	*AKR1B10*	Aldo–keto reductase family 1, member B10 (aldose reductase)	Bile acid biosynthesis	−6.988414261
25	D00408	*CYP3A7*	Cytochrome *P*450, family 3, subfamily A, polypeptide 7	Drug metabolism	−6.943817124
26	NM_001275	*CHGA*	Chromogranin A (parathyroid secretory protein 1)	Neuroendocrine negative modulator	−6.398707092
27	AB018283	*RHOBTB1*	Rho-related BTB domain containing 1	Signal transduction, protein binding	−6.197027765
28	BE970031	*CLPS*	Colipase, pancreatic	Enzyme, lipid catabolism	−6.144322204
29	BE868254		ESTs	Unknown	−5.941929445
30	U57961	13CDNA73	Hypothetical protein CG003	Unknown	−5.910096103
31	BF973589	*MT1X*	Metallothionein 1X	Heavy metal transportation	−5.888563301
32	AA564381	*KRT20*	Keratin 20	Cytoskeleton	−5.804261199
33	AK024409	DKFZP586-A0522	DKFZP586A0522 protein	Unknown	−5.801520054
34	AV702143	*GSTA3*	Glutathione *S*-transferase A3	Defense response	−5.757357962

Table 2 (*continued*)

No.	*Accession no.*[c]	*Symbol*	*Title*	*Function*[b]	*Median*[d]
35	M18963	*REG1A*	Regenerating islet-derived 1 alpha (pancreatic stone protein, pancreatic thread protein)	Cell proliferation	−5.714284913
36	AI970797		ESTs	Unknown	−5.577186347
37	AA868345	*RNASE1*	Ribonuclease, RNase A family, 1 (pancreatic)	RNA binding protein	−5.575601794
38	AF034803	*PPFIBP2*	PTPRF interacting protein, binding protein 2 (liprin beta 2)	DNA integration	−5.181314162
39	AI278995		Homo sapiens cDNA FLJ37907 fis, clone CO-LON2009337.	Unknown	−5.157417787
40	AA476677	MGC11324	Hypothetical protein MGC11324	Unknown	−5.108701604
41	NM_005036	*PPARA*	Peroxisome proliferative-activated receptor, alpha	Transcription factor, fatty acid metabolism	−5.061877465
42	AL080177	*UBL3*	Ubiquitin-like 3	Ubiquitination	−5.008658426
43	BE675124	*JFC1*	NADPH oxidase-related, C2 domain-containing protein	Membranetrafficking	−4.918163001
44	M77477	*ALDH3A1*	Aldehyde dehydrogenase 3 family, memberA1	Aldehyde metabolism	−4.904053163
45	BF106962	*FAM3B*	Chromosome 21 open reading frame 11	Cytokine activity	−4.671527904
46	AA628346		ESTs	Unknown	−4.590544114
47	J03037	*CA2*	Carbonic anhydrase II	Carbondioxide metabolism	−4.494765303
48	AK000188	*ACAS2*	Acetyl-Coenzyme A synthetase 2 (ADP forming)	Lipid metabolism	−4.367872099
49	AW954291	CGI-40	CGI-40 protein	Unknown	−4.314783694
50	AA478113	LOC283385	Morn	Unknown	−4.27957966
51	NM_005518	*HMGCS2*	3-Hydroxy-3-methylglutaryl-Coenzyme A synthase 2 (mitochondrial)	Cholesterol biosynthesis	−4.274855877
52	X59770	*IL1R2*	Interleukin 1 receptor, type II	Immune response	−4.034032544
53	AV682257	*FABP1*	Fatty acid binding protein 1, liver	Fatty acid metabolism	−4.030921493
54	BF971884	*MT2A*	Metallothionein 2A	Heavy metal transportation	−3.965939867
55	Z69881	*ATP2A3*	ATPase, Ca^{++} transporting, ubiquitous	Electron transportation	−3.909188751
56	NM_012244	*SLC7A8*	Solute carrier family 7 (cationic amino-acid transporter, y+ system), member 8	Small molecule transportation	−3.902622755
57	D17793	*AKR1C3*	Aldo–keto reductase family 1, member C3	Lipid metabolism	−3.899483839
58	S68287	*AKR1C4*	Aldo–keto reductase family 1, member C4	Lipid metabolism	−3.872209776
59	AA573793	*TFF1*	Trefoil factor 1 (breast cancer, estrogen-inducible sequence expressed in)	Defense response	−3.777161641
60	AL122066	*FKBP5*	FK506 binding protein 5	Protein folding	−3.763801026
61	AI633883	LOC92558	Hypothetical protein LOC92558	Unknown	−3.710260693
62	AA259255		ESTs	Unknown	−3.682262717
63	AW014155		ESTs, Moderately similar to hypothetical protein FLJ20378 [Homo sapiens] [H.sapiens]	Unknown	−3.656475021
64	AB018263	KIAA0720	KIAA0720 protein	Unknown	−3.64599766
65	AI985579	*SH3BGRL2*	SH3 domain binding glutamic acid-rich protein like 2	Signal transduction	−3.637604549
66	L10320	*FBP1*	Fructose-1,6-bisphosphatase 1	Fructose metabolism	−3.583676033
67	D84239	*FCGBP*	Fc fragment of IgG binding protein	Immune response	−3.519272848
68	AI433935	*REG1B*	Regenerating islet-derived 1 beta (pancreatic stone protein, pancreatic thread protein)	Islet cell regeneration	−3.50240627
69	AF227899	*BCAA*	RBP1-like protein	Tumor-associated Ag	−3.501043543
70	AK027077	*ERBB2IP*	Erbb2 interacting protein	Signal transduction	−3.443668435
71	AW242997		ESTs, Weakly similar to S10590 cysteine proteinase (EC 3.4.22.-) – human [H.sapiens]	Unknown	−3.412788641
72	AI356283	FLJ34443	Hypothetical protein FLJ34443	Unknown	−3.412680583
73	AK024449	PP2135	PP2135 protein	Unknown	−3.400982961
74	BF448529	*B3GALT4*	UDP-Gal:betaGlcNAc beta 1,3-galactosyltransferase, polypeptide 4	Glycolipid biosynthesis	−3.383794412
75	X04299	*ADH1C*	Alcohol dehydrogenase 1C (class I), gamma polypeptide	Alcohol metabolism	−3.383024368

[a]The 75 representative downregulated genes were sorted by their expression ratio ranging from the lowest. [b]Gene functions were summarized from literature sources or according to LocusLink in NCBI (www.ncbi.nlm.nih.gov/LocusLink). [c]GenBank Accession number. [d]Median of log 2-transformed expression ratio of diffuse-type gastric cancer cases

COL1A1 and *COL1A2*, and genes associated with immune response or metabolism (*IGHG3, IGKC, C3, APOA1, GPX1* and *LDHA*). On the other hand, genes in groups D and F, which were relatively upregulated in intestinal-type cancers, contained genes associated with growth-factor receptors such as *GRB10*, *IRS2* and *EPS15R*, and genes related to cellular proliferation or mitochondrial function such as *EGFR*, *EGR1*, *PCNA*, *CDK2*, *PSORT*, *TIMM10* and *BCAT2*. These data appear to reflect the different natures of the two types of tumor. Notably, clones corresponding to the same gene that had been spotted at different positions on the microarray slides were present in the same branch of the gene-cluster dendrogram, indicating a high degree of reliability for our microarray data.

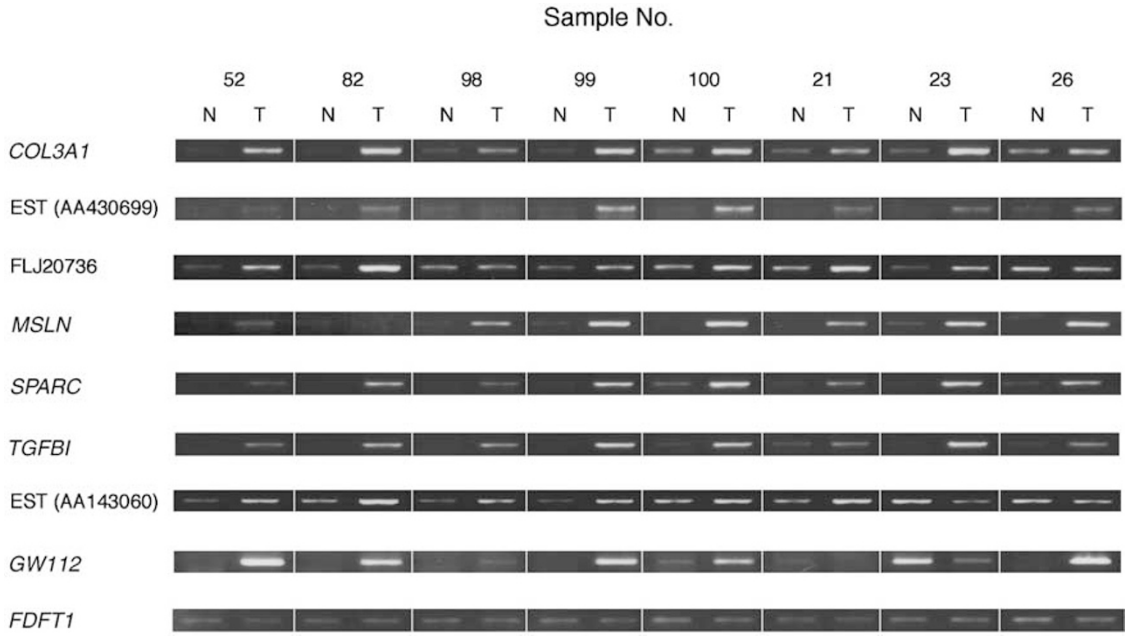

Figure 1 Re-evaluation of elevated expression of eight genes in diffuse-type gastric cancers (T) and corresponding noncancerous epithelia (N) by semiquantitative RT–PCR using eight pairs of the aRNAs utilized in microarray analysis. Expression of *FDFT1* served as an internal control

Identification of genes that were differentially expressed between diffuse- and intestinal-type gastric cancers

We carried out random permutation tests to identify genes that could discriminate diffuse-type cancer from intestinal-type cancer. The tests identified a total of 46 genes with *P*-values of less than 0.01. We further performed an additional, supervised hierarchical clustering analysis using the data for the 46 genes that had clearly separated the two classes of tumors (Figure 3). Of these 46 genes, 14 (including three of unknown function) showed relatively higher expression levels, and 32 (including eight of unknown function) showed relatively lower expression levels in diffuse-type cancer than intestinal-type gastric cancer. The 14 upregulated elements included genes encoding chaperones (*CCT3* and *TOR1B*) and genes associated with cell motility and cytoskeleton (*CD81* and *TUBA3*), glycosylation (*RPN2, MGAT1* and *MPI*), or other functions *(SDCCAG8* and *HRMT1L2)*. The 32 genes with relatively lower expression levels in diffuse-type cancers included some involved in signal transduction and transcriptional regulation *(RHBDL, SFRS8, MLL5, LDB3* and *GFRA2)*, nuclear transportation (*KPNB2* and *NUP133*), cell adhesion (*PSK-1, ITGB5, SRPX* and *IBSP*), or other functions (*HRG* and *TG737*) (Table 3). We also carried out quantitative RT–PCR experiments with three randomly selected discriminating genes (*SLC25A4, GFRA2, HSPCB*) to verify the microarray data. The results of RT–PCR experiments using 16 pairs of RNA from eight intestinal-type and eight diffuse-type cancers consistently showed significantly different levels of expression between these two types of cancer, supporting the reliability of our strategy (Figure 4).

Identification of genes related to LN metastasis, venous invasion, or lymphatic-vessel involvement in diffuse-type gastric cancer

Additional random permutation tests identified 13 genes (including three of unknown function) associated with venous invasion and 11 (including three of unknown function) associated with lymphatic vessel invasion. Discriminators in both invasion groups included genes involved in differentiation (*S100P*), cell adhesion (*SDCBP*), transport (*SNX2, HCN3* and *ATP6V0A1*), transcription (*FOXD1, ZFP36*), or signal transduction (*SYK, PPP3CA* and *PTPRJ*). We also identified 31 genes that were differentially expressed between tumors with and without LN metastasis; 13 (including six of unknown function) showed elevated expression and 18 (including two of unknown function) showed reduced expression in tumors with LN metastasis. Discriminators in the LN metastasis group included genes involved in differentiation (*NOTCH2*), cytoskeleton (*TUBB2*), metabolism (*FACL5* and *ACAA1*), signal transduction (*EPS15*), transcription, or protein synthesis (*ZFP103, RPS23, RPS10, RPL31* and *EIF4G2*) (Figure 5).

Discussion

The advent of laser-microdissection technology has brought about a great improvement in our ability to isolate cancer cells from interstitial tissues. The proportion of contaminated cells using this method is estimated to be less than 0.3% (Yanagawa *et al*., 2001) and 0.29% (Nakamura *et al*., 2004) in our earlier studies, which

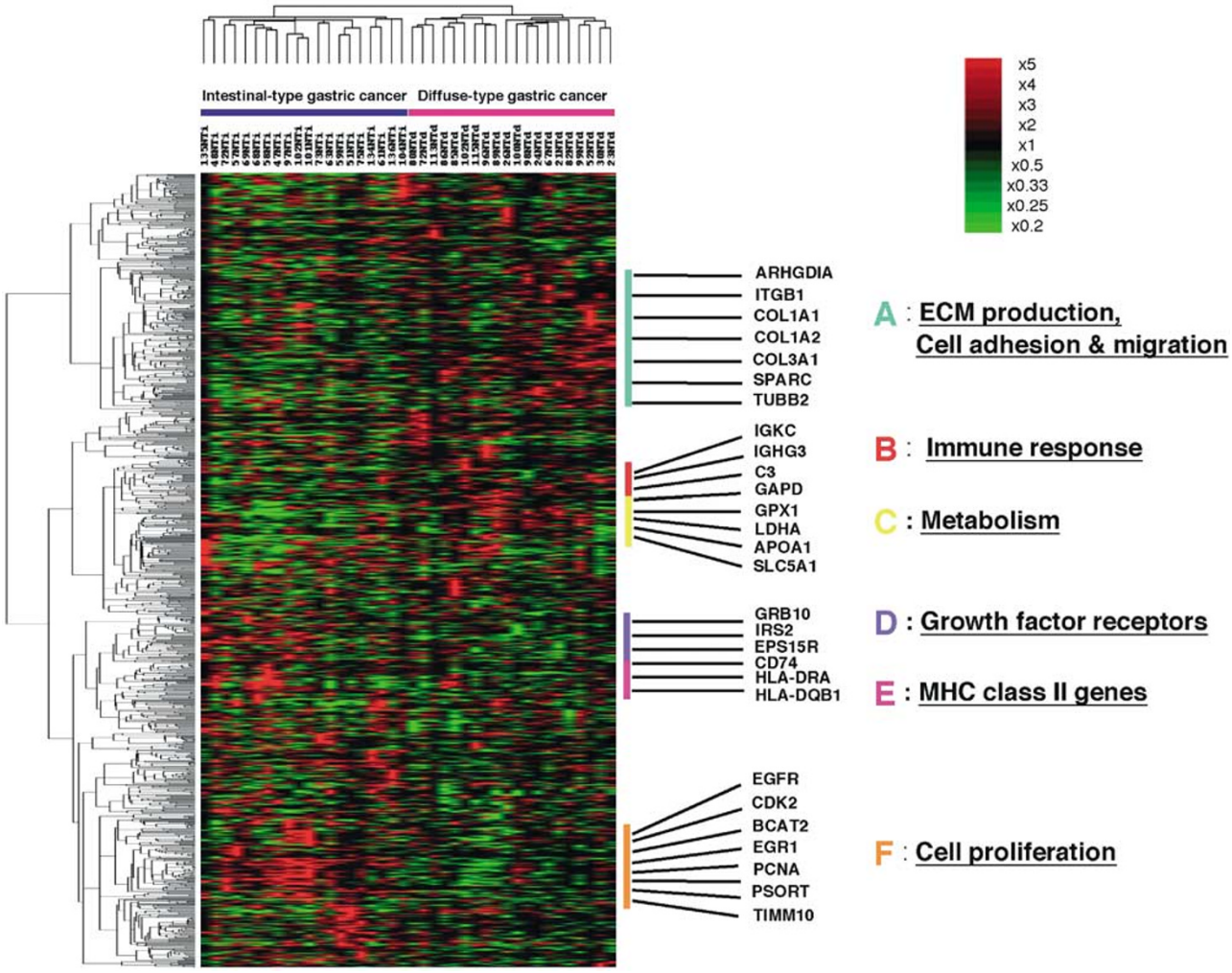

Figure 2 An unsupervised two-way hierarchical clustering analysis of 1051 genes across 40 gastric-cancer samples. In the horizontal axis, the 40 tumors were classified into two major trunks, while in the vertical axis, the 1051 genes were clustered according to similarities in relative expression ratios. The color of each cell in the matrix represents the expression level of a single gene, with red and green indicating expression levels, respectively, above and below the median derived from log-transformed relative expression ratios for that gene across all samples. Black represents unchanged expression, gray indicates no or insignificant expression (intensities of both Cy3 and Cy5 under the cutoff value)

suggests that our data should represent the expression profile of a highly pure population of diffuse-type gastric-cancer cells.

Some of the genes that were commonly upregulated in our experiments reflected the distinct nature of diffuse-type gastric cancer. A number of genes involved in the TGF-β signaling pathway, such as *TGFBI* and *PLAB,* were upregulated in this type of gastric cancer. *TGFBI*, first isolated from human lung-adenocarcinoma cell line A549 (Skonier *et al*., 1992), has a recognition site for integrin, a key molecule for adhesion and migration of cancer cells. *PLAB* helps regulate tissue differentiation and maintenance, and also inactivates macrophages by suppressing production of TNF-α (Bootcov *et al*., 1997). Since previous analyses of gene expression in intestinal-type gastric cancers (Hasegawa *et al*., 2002), colorectal cancers (Kitahara *et al*., 2001) and pancreatic cancers (Iacobuzio-Donahue *et al*., 2003) also showed elevated expression of *TGFBI* or *PLAB*, enhanced activity of the TGF-β signaling pathway seems to play an important carcinogenetic role in a wide range of human tissues.

We showed also that *MMP-7* and *TIMP1*, genes associated with cell–ECM interaction, were highly expressed in diffuse-type but not in intestinal-type gastric cancers (Hasegawa *et al*., 2002). As MMPs catalyse extracellular matrices including collagen and fibronectin, they are thought to be involved in invasion and/or metastasis of cancer cells. Enhanced expression of *MMP7*, a uterine matrilysin, has been reported in cancers of the pancreas (Crnogorac-Jurcevic *et al*., 2002), ovary (Schwartz *et al*., 2003), bladder (Sumi *et al*., 2003) and colon (Zeng *et al*., 2002). Immunohistochemical staining has demonstrated marked accumulation of *MMP-7* in the tumor-front areas of metastatic

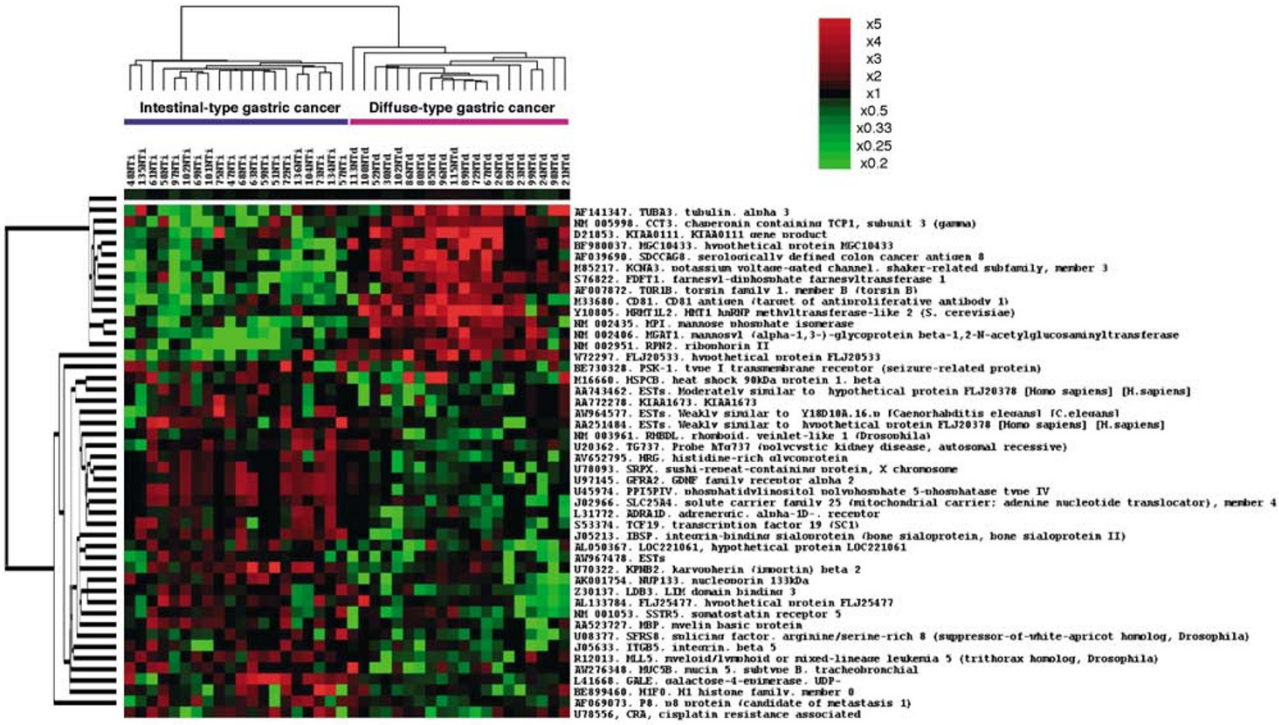

Figure 3 A supervised two-way hierarchical clustering analysis using 46 discriminating genes. The color of each cell in the matrix represents the expression level of a single gene in an individual sample, with red and green indicating expression levels, respectively, above and below the median derived from log-transformed relative expression ratios for that gene across all samples. Black represents unchanged expression

colorectal-cancer cells, suggesting involvement of *MMP-7* activation in metastasis (Zeng *et al*., 2002). Despite the fact that *TIMP-1* was initially identified as a gene encoding an inhibitor of metalloproteinases, recent studies have pointed out its paradoxical effect in carcinogenesis through regulation of cell growth, apoptosis and angiogenesis. The upregulation of *TIMP-1* in our data agrees with earlier findings that elevated expression of this gene is associated with poor prognosis among patients with several types of human tumors (Jiang *et al*., 2002), since the diffuse type of gastric cancer often shows poorer prognosis than the intestinal type.

We also observed enhanced expression of genes involved in cell migration or production of ECM components (e.g. *SPARC*, *FN1*, *TUBB2, ITGB1, COL1A1* and *COL1A2)* in diffuse-type cancers but not in intestinal-type cancers. Upregulation of *SPARC* has been implicated in invasive and/or metastatic phenotypes (De *et al*., 2003). The importance of genes encoding ECM components was emphasized in a recent study that identified *COL1A1* and *COL1A2* as members of a metastasis-associated gene signature (Ramaswamy *et al*., 2003). Although collagens were reported to be expressed specifically in tumor endothelial cells (St Croix *et al*., 2000), *COL1A1* and *COL1A2* were abundantly expressed in MKN45 and St-4 diffuse-type gastric cancer cell lines when we analysed expression profiles of 39 human cancer cell lines using the same microarray system (Dan *et al*., 2002). Therefore, augmented expression of these collagens should not result from contaminated stromal cells, but from elevated expression in cancer cells. Our data strongly suggest that elevated expression of genes encoding ECM components is in fact a characteristic feature of diffuse-type gastric-cancer cells.

Genes involved in glycosylation (*MGAT1, MPI* and *RPN2*) were significantly enhanced in diffuse-type cancers, but not in intestinal-type cancers. Increased activity of these genes may confer properties of invasion and/or metastasis to diffuse-type gastric-cancer cells, since invasion and metastasis are highly dependent on alterations in the ECM and cell–cell interactions that, in part, involve structural changes in cell-surface components including glycosylation (Couldrey and Green, 2000). Our data imply that dysregulated cell–matrix interaction is a common feature of diffuse-type gastric cancers. Interestingly, *SDCCAG8*, a tumor antigen whose elevated expression was associated with poor survival in gastric-cancer patients (Chen *et al*., 2003), was also upregulated in our diffuse-type cancers. Hence, overexpression of *SDCCAG8* in diffuse-type cancer may contribute to its poor-prognosis characteristic and may be used for prediction of prognosis in patients.

The majority of downregulated elements in our experiments represent a set of genes that function in various metabolic pathways and transport systems. Among them were several genes with specific functions in gastric epithelium, such as *PGC* and *GIF*, implying that dedifferentiation is a common feature of carcino-

Table 3 Genes differentially expressed between diffuse- and intestinal-type gastric cancers

No.	*Accession no.*[b]	*Gene symbol*	*Title*	*Function*[a]	P-*value*[c]
Genes that have significantly higher expression level in diffuse- than intestinal-type gastric cancer					
1	NM_005998	*CCT3*	Chaperonin-containing TCP1, subunit 3 (gamma)	Actin and tubulin folding	0.001186538
2	AF141347	*TUBA3*	Tubulin, alpha 3	Cytoskeleton	0.005992473
3	W72297	FLJ20533	Hypothetical protein FLJ20533	Unknown	0.000504459
4	D21853	KIAA0111	KIAA0111 gene product	Unknown	0.00028293
5	Y10805	*HRMT1L2*	HMT1 hnRNP methyltransferase-like 2 (*S. cerevisiae*)	Methylation	0.002430301
6	AF007872	*TOR1B*	Torsin family 1, member B (torsin B)	Chaperone, heat-shock protein activity	6.6376E-05
7	S76822	*FDFT1*	Farnesyl-diphosphate farnesyltransferase 1	Lipid metabolism	1.50303E-05
8	AF039690	*SDCCAG8*	Serologically defined colon cancer antigen 8	Tumor Ag	7.96073E-05
9	M85217	*KCNA3*	Potassium voltage-gated channel, shaker-related subfamily, member 3	Potassium ion transportation	1.02035E-06
10	M33680	*CD81*	CD81 antigen (target of antiproliferative antibody 1)	Signal transduction, cell motility	0.00455433
11	BF980037	MGC10433	Hypothetical protein MGC10433	Unknown	0.000447196
12	NM_002435	*MPI*	Mannose phosphate isomerase	Glycosylation	0.000549555
13	NM_002406	*MGAT1*	Mannosyl (alpha-1,3-)-glycoprotein beta-1,2-*N*-acetylglucosaminyltransferase	*N*-glycan biosynthesis, membrane integrity	0.000168686
14	NM_002951	*RPN2*	Ribophorin II	*N*-glycan biosynthesis, membrane integrity	7.47625E-05
Genes that have significantly higher expression level in intestinal- than diffuse-type gastric cancer					
1	BE730328	*PSK-1*	Type I transmembrane receptor (seizure-related protein)	Adhesion protein	0.001107428
2	AW276348	*MUC5B*	Mucin 5, subtype B, tracheobronchial	Salivary mucin synthesis	5.82751E-05
3	AA743462		ESTs, Moderately similar to hypothetical protein FLJ20378 [Homo sapiens] [H.sapiens]	Unknown	0.000156482
4	AA772278	KIAA1673	KIAA1673	Unknown	0.000260902
5	AW964577		ESTs, Weakly similar to Y18D10A.16.p [*Caenorhabditis elegans*] [*C. elegans*]	Unknown	0.002677221
6	J05633	*ITGB5*	Integrin, beta 5	Cell–matrix adhesion, signal transduction	5.23719E-07
7	M16660	*HSPCB*	Heat-shock 90 kDa protein 1, beta	Heat-shock protein activity	0.00191956
8	AA251484		ESTs, weakly similar to hypothetical protein FLJ20378 [*Homo sapiens*] [*H. sapiens*]	Unknown	1.42436E-08
9	NM_003961	*RHBDL*	Rhomboid, veinlet-like 1 (Drosophila)	Signal transduction	7.34458E-09
10	U78093	*SRPX*	Sushi-repeat-containing protein, X chromosome	Cell adhesion	1.33652E-13
11	U97145	*GFRA2*	GDNF family receptor alpha 2	Receptor, signal transduction	6.75626E-15
12	U45974	*PPI5PIV*	Phosphatidylinositol polyphosphate 5-phosphatase type IV	Catalytic enzyme activity	3.77908E-14
13	J02966	*SLC25A4*	Solute carrier family 25 (mitochondrial carrier; adenine nucleotide translocator), member 4	Mitochondrial genome maintenance	1.33423E-13
14	L31772	*ADRA1D*	Adrenergic, alpha-1D-, receptor	G-protein signaling pathway	1.2948E-09
15	S53374	*TCF19*	Transcription factor 19 (SC1)	Transcription factor, cell cycle regulation	1.07056E-09
16	J05213	*IBSP*	Integrin-binding sialoprotein (bone sialoprotein, bone sialoprotein II)	Cell adhesion	5.21228E-13
17	AV652795	*HRG*	Histidine-rich glycoprotein	Apoptosis cell clearance, anticoagulant	1.37735E-09
18	U20362	*TG737*	Probe hTg737 (polycystic kidney disease, autosomal recessive)	Excretion	4.65524E-07
19	AL050367	LOC221061	Hypothetical protein LOC221061	Unknown	1.63492E-10
20	AW967478		ESTs	Unknown	3.68484E-06
21	U70322	*KPNB2*	Karyopherin (importin) beta 2	Protein-nucleus import	1.42923E-11
22	AK001754	*NUP133*	Nucleoporin 133 kDa	mRNA-nucleus export	1.59442E-14
23	AA523727	*MBP*	Myelin basic protein	Immune response, neurotransmission	1.60146E-08
24	NM_001053	*SSTR5*	Somatostatin receptor 5	Receptor, G-protein signaling pathway	6.25063E-09
25	Z30137	*LDB3*	LIM domain binding 3	Signal transduction	6.91651E-10
26	AL133784	FLJ25477	Hypothetical protein FLJ25477	Unknown	0.00017911
27	U08377	*SFRS8*	Splicing factor, arginine/serine-rich 8 (suppressor-of-white-apricot homolog, Drosophila)	Transcriptional regulation	2.48584E-08
28	R12013	*MLL5*	Myeloid/lymphoid or mixed-lineage leukemia 5 (trithorax homolog, Drosophila)	Transcriptional regulation	6.13237E-12
29	L41668	*GALE*	Galactose-4-epimerase, UDP-	Carbohydrate metabolism	6.19831E-10
30	BE899460	*H1F0*	H1 histone family, member 0	DNA binding, nucleosome assembly	0.001187699
31	AF069073	*P8*	p8 protein (candidate of metastasis 1)	Nuclear DNA binding protein	0.001379776
32	U78556	*CRA*	Cisplatin resistance associated	Unknown	0.009855345

[a]Gene functions were summarized from literature sources or according to LocusLink in NCBI (www.ncbi.nlm.nih.gov/LocusLink). [b]GenBank Accession number. [c]Permutational *P*-values calculated as described previously (Materials and methods) were indicated

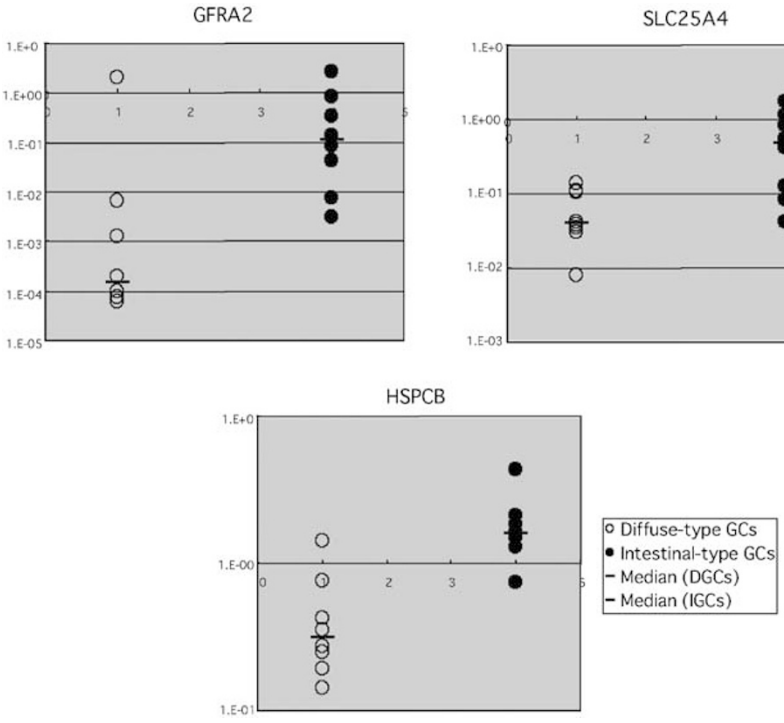

Figure 4 Quantitative RT–PCR analysis of genes differentially expressed between intestinal- and diffuse-type gastric cancers. Each dot represents a log 2-transformed expression ratio (tumor : normal) of the three selected genes in the eight intestinal-type and eight diffuse-type tumors. Short horizontal lines represent the median ratios for each type of tumor

genesis. The trefoil factor family 1 gene (*TFF1*), which participates in stabilizing the mucous gel overlying the gastrointestinal mucosa, was among the downregulated genes in both diffuse- and intestinal-type gastric cancers (Hasegawa *et al*., 2002). Since *TFF1*-knockout mice develop multiple gastric adenomas and carcinomas, and somatic mutations in *TFF1* have been reported in human gastric cancers (Park *et al*., 2000), the reduced expression of *TFF1* in our populations of gastric-cancer cells is in line with a gastric-specific tumor-suppressive role. These findings are consistent with genes depicted in the normal gastric-tissue cluster compared to tumor cluster previously reported (Chen *et al*., 2003).

One of the mechanisms that tumor cells use to escape from the immune response is to compromise the antigen-presenting functions of infiltrating dendritic cells by secreting tumor-derived factors. One of the upregulated genes on our list, *G1P2* (or *ISG15*), encodes a key cytokine that is secreted by melanoma cells to induce E-cadherin expression on dendritic cells and thereby impair migratory behavior (Padovan *et al*., 2002). This finding may reflect a novel survival mechanism of diffuse-type gastric-cancer cells. Furthermore, in our experiments MHC class II genes such as *CD74*, *HLA-DRA* and *HLA-DQB1* were expressed to a lesser degree in diffuse-type tumors than intestinal-type tumors (Figure 2, group E). Since these genes tend to be expressed abundantly in the new molecular subtypes of gastric cancer that show better prognosis (Tay *et al*., 2003), our data support the notion that the ability to escape from immunological surveillance confers aggressive properties on diffuse-type tumors.

Since diffuse-type gastric cancer has more invasive and metastasizing potential and poorer prognosis than intestinal-type cancer, it is in our interest to clarify the mechanisms underlying progression of this tumor. Therefore, we attempted to identify genes specifically associated with venous invasion, lymphatic vessel involvement or LN metastasis in diffuse-type cancers. Among such genes, *SYK*, a nonreceptor type of protein tyrosine kinase, showed lower expression in tumors having lymphatic vessel involvement than in tumors without that feature. Since reduced expression of wild-type *SYK* has been correlated with poor prognosis and with distant metastases in breast cancers (Toyama *et al*., 2003), and its overexpression suppresses motility and invasiveness of breast-cancer cells, *SYK* is regarded as a candidate metastasis-suppressor gene (Mahabeleshwar and Kundu, 2003). Hence, its decreased expression in gastric-cancer cells may also play a role in metastasis through invasion of lymphatic vessels. Expression of *CCL25*, encoding a chemokine that functions as an effector of lymphocyte migration, was higher in tumors with venous invasion than in the noninvasion group. Tumor-associated chemokines are thought to play roles in invasion and/or metastasis including control of the movement of tumor cells themselves (Balkwill, 2003). Since *CCL25* mediates lymphocyte entry into the small intestine by promoting their rapid adhesion to vascular endothelium or transmigration into lamina propria (Campbell and Butcher, 2002), cancer cells with

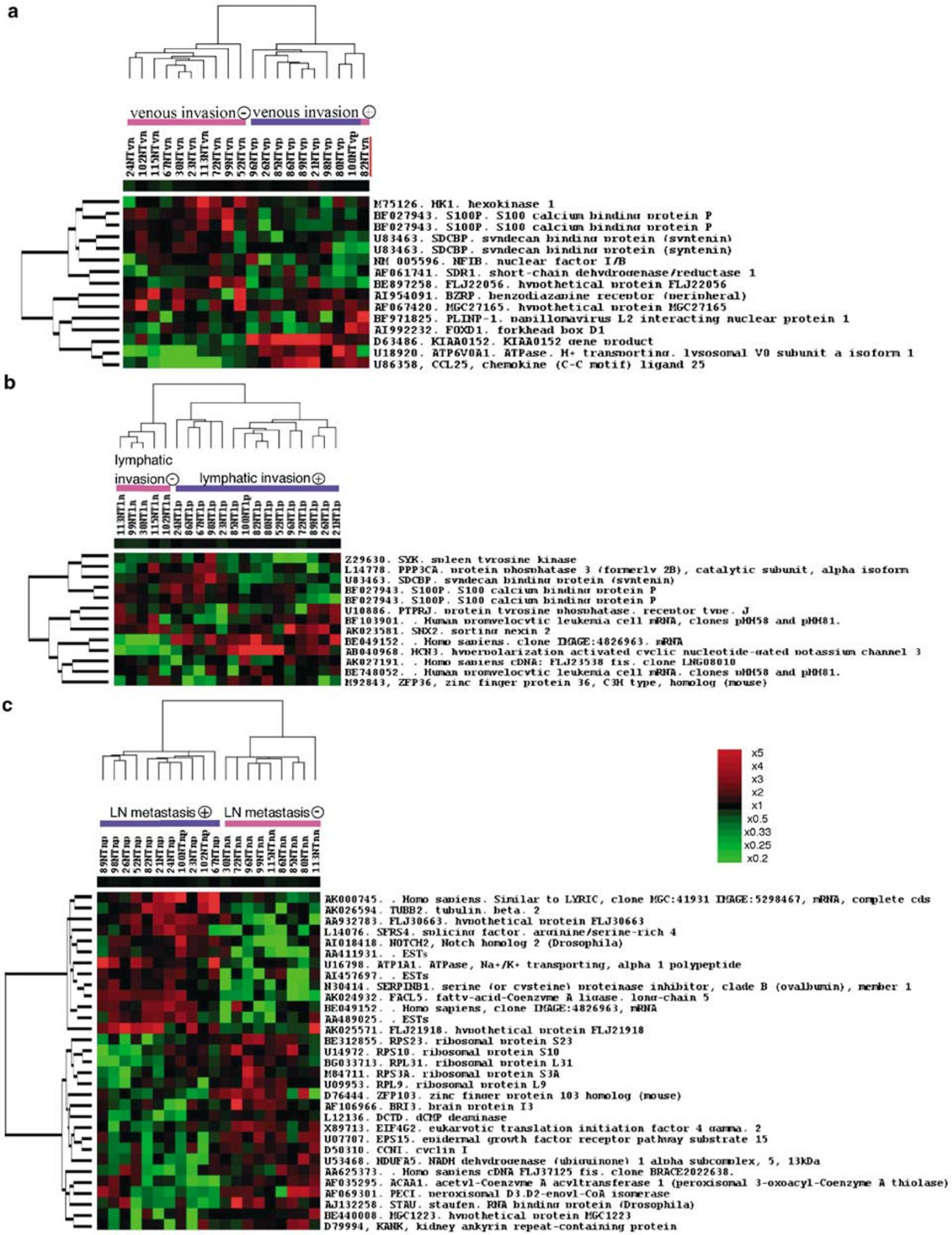

Figure 5 A supervised hierarchical clustering analysis of expression profiles associated with (**a**) venous invasion, (**b**) lymphatic vessel involvement, or (**c**) LN metastasis in diffuse-type gastric cancers. In the horizontal axis, the 20 diffuse-type tumors were correctly classified into two major trunks corresponding to the histopathological data. In the vertical axis, the genes were clustered according to their similarities in expression patterns. The color of each cell in the matrix represents the expression level of a given gene in an individual sample: red indicates elevated expression, green reduced expression and black unchanged expression. Color standards are indicated on the right of the cluster

upregulated *CCL25* may promote hematogenous metastasis by the same mechanism.

In this study, we identified many genes that were commonly altered in diffuse-type gastric cancer, as well as genes that were differentially expressed between diffuse- and intestinal-type tumors. This large body of information has not only clarified mechanisms involved in diffuse-type gastric cancer, but has revealed distinct molecular signatures of both types of gastric cancer. Further characterization of the genes identified in this study may lead to a more profound understanding of gastric carcinogenesis and facilitate the development of novel diagnostic markers and more effective therapeutic modalities for each histological type of gastric cancer.

Materials and methods

Patients and tissue samples

Primary gastric cancers and corresponding noncancerous gastric mucosae were obtained with informed consent from 20 patients who underwent gastrectomy at Kyoto University Hospital, Kitano Hospital, Cancer Institute Hospital, Sakai City Hospital, and Hospitals of Kinki University, Japan. Patient profiles were obtained from medical records. Histopathological classification of each tumor, performed according to Lauren's classification (Lauren, 1965), diagnosed all samples as diffuse-type gastric adenocarcinomas. Clinical stage was determined according to the UICC TNM classification. The 20 cases analysed consisted of 19 advanced (T2–T4) and one early (T1) tumors. All cancer tissues were dissected from the margin of tumor mass, while their corresponding noncancerous tissues were collected from mucosa in the same region of the stomach. Clinicopathological data for these 20 samples and for the 20 intestinal-type gastric-cancer samples analysed previously are summarized in Table 4. No significant differences were observed in terms of gender, age of patients, size of tumor, depth of invasion, vessel invasion, or node involvement between the two groups. All samples were immediately frozen and embedded in TissueTek OCT medium (Sakura, Tokyo, Japan) and stored at $-80^{\circ}C$ until used for microarray analysis.

Laser-microbeam microdissection, extraction of RNA, and T7-based RNA amplification

Frozen tissues were prepared in 8-μm sections, as described previously (Kitahara *et al*., 2001) except for elimination of the final dehydration in xylene. A total of 30 000–40 000 cancer cells or noncancerous gastric epithelial cells were collected selectively using the EZ cut system (SL Microtest GmbH, Germany) according to the manufacturer's protocol. Extraction of total RNA, T7-based amplification, and labeling of probes were performed as described previously (Kitahara *et al*., 2001). A measure of 2.5-μg aliquots of twice-amplified RNA (aRNA) from each cancerous and noncancerous tissue were then labeled, respectively, with Cy3-dCTP or Cy5-dCTP (Amersham Biosciences).

cDNA microarray and analysis of data

We used the same cDNA microarray system containing 23 040 cDNAs in duplicate and the same procedures for hybridization, washing, photometric quantification of signal intensities of each spot, and normalization of data as in our previous analysis of 20 intestinal-type gastric cancers (Hasegawa *et al*., 2002). To filter less reliable data derived from low signal intensities, we determined the cutoff value as the intensity of spot whose S/N (signal to noise) ratio was 3. Genes were categorized into three groups according to their expression ratios (Cy3/Cy5): upregulated (ratio equal to or greater than 2.0), downregulated (ratio equal to or less than 0.5), and unchanged expression (ratios between 0.5 and 2.0). Genes with Cy3/Cy5 ratios greater than 2.0 or less than 0.5 in more than 50% of the cases examined were defined, respectively, as commonly up- or downregulated genes.

Semiquantitative RT–PCR

We selected eight representative genes (*TGFBI, SPARC, COL3A1, MSLN,* FLJ20736, *GW112* and two ESTs) that were commonly upregulated in diffuse-type tumors and examined their expression levels by semiquantitative RT–PCR. A housekeeping gene, *FDFT1*, served as an internal control because it had shown the smallest Cy3/Cy5 fluctuation in our experiments. Reverse transcription was carried out as described previously (Kitahara *et al*., 2001). The PCR reaction was preceded by 95°C for 2 min, then underwent 25 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 30 s followed by 72°C for 5 min. The primer sequences were as follows: *FDFT1* forward, 5′-TGTGTGGCTGGGACCTTTAGGAA-3′ and reverse, 5′-TCATTCTAGCCAGGATCATACTAAG-3′; *TGFBI* forward, 5′-TCCCTGGAAAAGGAGCTTCAGTA-3′ and reverse, 5′-ACACCATGGCTCTGTCACAATAG-3′; *SPARC* forward, 5′-CAAGAGTGAGATGTAGAAAGTTGT-3′ and reverse, 5′-CTTCACATCATGGTGAGAGTTTG-3′; *COL3A1* forward, 5′-AGACGCATGTTATGGTGCTAATGTA-3′ and reverse, 5′-GATCAACAACCACATACAAGCTTAC-3′; *MSLN* forward, 5′-AACGGCTACCTGGTCCTAGAC-3′ and reverse, 5′-GTTTACTGAGCGCGAGTTCTCT-3′; FLJ20736 forward, 5′-ATATGAGCAGGAACTCTGGGAG-3′ and reverse, 5′-CTCAGGATAGAGGGCAAAGAGA-3′; *GW112* forward, 5′-GAAAATCTGATGGCAGTGACAA-3′ and reverse, 5′-AAGGTTTCCAACTACTGCACTGA-3′; EST (GenBank Accession No. AA143060) forward, 5′-TGTTGCTCTTCCTTGTGGAGCT-3′ and reverse, 5′-GCAAATCCTACTTTCAACTGCAC-3′; EST (GenBank Accession No. AA430699) forward, 5′-TTTAACGCTGGTGGGCAGCA-3′ and reverse, 5′-ATAAACAGAACCCATCCCAAAG-3′.

Cluster analysis and identification of genes able to discriminate between diffuse- and intestinal-type gastric cancers

We undertook an unsupervised cluster analysis to compare the expression profiles of 20 diffuse-type gastric cancers with profiles of the 20 intestinal-type gastric tumors analysed previously (Hasegawa *et al*., 2002). For this procedure, we selected 1051 genes that had given valid values in more than 50% of all samples and standard deviations of the values greater than 2. After log-transforming and median-centering the data, we performed average-linkage hierarchical clustering and viewed the results using web-available software packages 'Cluster' and 'TreeView' (http://genome-www5.stanford.edu/MicroArray/SMD/restech.html) (Eisen *et al*., 1998). We further carried out random permutation tests using log-transformed data of genes showing Cy3 or Cy5 intensities above cutoff values, and estimated the ability of each gene to distinguish diffuse- from intestinal-type

Table 4 Histological data of the 40 gastric cancer clinical samples used for cDNA microarray analysis

No.	*Patient ID*	*Age*	*Sex*	*Japanese classification*[a]	*Depth*[b]	*T*	*N*	*M*	*Lymphatic invasion*	*Venous invasion*	*Anatomical site*[c]	*Normal mucosa*[d]
Diffuse-type gastric cancers (diagnosed according to Lauren's classification)												
1	FU010052	41	M	por2	ss	2	1	0	1	0	M	N
2	FU010067	70	M	por2	ss	2	3	0	2	0	AMC	IM
3	FU010082	46	M	por1	ss	2	1	0	1	0	M	N
4	FU010098	37	M	por2	se	3	1	0	2	1	M	N
5	FU010099	57	F	por2	ss	2	0	0	0	0	M	AG
6	FU010100	69	M	por2	se	3	2	0	2	1	AM	N
7	FU010113	57	M	tub2-por1	mp	2	0	0	0	0	M	IM
8	FU010115	60	M	por1	se	3	0	0	0	0	AM	N
9	FU020021	54	M	por1	se	3	3	0	3	3	CM	N
10	FU020023	57	F	sig	mp	2	1	0	1	0	M	N
11	FU020024	67	F	por1	ss	2	2	0	3	0	M	IM
12	FU020026	81	F	sig	se	3	3	0	1	1	A	AG
13	FU020030	71	M	por1	m	1	0	0	0	0	C	N
14	FU020072	72	M	por2	ss	1	0	0	1	0	C	N
15	FU020080	57	F	por1	se	3	0	0	3	3	C	N
16	FU020085	49	F	por1	ss	2	0	0	1	3	C	N
17	FU020086	59	M	por2	se	3	0	0	1	1	AM	IM
18	FU020089	63	M	por1	se	3	1	0	1	2	A	N
19	FU020096	72	M	tub2-por1	mp	2	0	0	2	1	M	IM
20	FU020100	58	F	sig	se	3	1	0	0	0	M	AG
Intestinal-type gastric cancers (diagnosed according to Lauren's classification)												
1	FU001047	74	M	tub2	ss	2	0	0	1	0	C	ND
2	FU001048	77	M	tub2	ss	2	2	0	2	2	A	ND
3	FU001051	59	M	tub1	se	3	2	0	3	0	AM	ND
4	FU010057	56	M	tub2	mp	2	0	0	0	0	C	ND
5	FU010058	55	M	tub2	si	4	3	0	2	1	A	ND
6	FU010059	69	M	tub2	se	3	2	0	3	2	A	ND
7	FU010061	69	F	tub1	se	3	1	0	2	2	C	ND
8	FU010063	73	M	tub2	ss	2	1	0	2	1	M	ND
9	FU010068	64	M	tub1	m	1	0	0	1	0	CM	ND
10	FU010069	70	M	tub2	m	1	0	0	0	0	A	ND
11	FU010072	65	M	tub2	se	3	2	0	2	1	CE	ND
12	FU010073	84	M	tub2	ss	2	0	0	1	1	CE	ND
13	FU010075	41	M	tub2	ss	2	2	0	1	1	A	ND
14	FU010097	56	M	tub2	ss	2	0	0	0	0	MA	ND
15	FU010101	55	F	tub2	se	3	0	0	2	1	M	ND
16	FU010102	48	M	tub1	mp	2	0	0	0	0	M	ND
17	FU010104	70	M	tub1	si	4	2	0	2	0	A	ND
18	FU010134	63	M	tub1	si	4	0	0	1	1	CMA	ND
19	FU010135	53	F	tub1	ss	2	0	0	0	3	C	ND
20	FU010136	59	M	tub1	ss	2	0	0	1	1	C	ND

[a]Japanese classification: tub1, well-differentiated tubular adenocarcinoma; tub2, moderately differentiated tubular adenocarcinoma; por1, poorly differentiated (solid) adenocarcinoma; por2, poorly differentiated (nonsolid) adenocarcinoma; sig, signet ring cell carcinoma. [b]Depth indicators: m, mucosa; mp, muscularis propria; sm, submucosa; ss, subserosa; se, serosa; si, serosal invasion. [c]Anatomical sites: A, antrum; C, cardia; E, esophagus; M, mid-body. [d]Normal mucosa: N, normal; IM, intestinal metaplasia; AG, atrophic gastritis; ND, not defined

tumors. Samples were randomly permuted 10 000 times between the two classes to calculate a *P*-value for each gene. Genes were considered significant when signal intensities were higher than cutoff values in more than 50% of the cases, *P*-values were less than 0.01, and $|Med_d\text{-}Med_i|$ was greater than or equal to 0.5, where Med indicates the median derived from log-transformed relative expression ratios in diffuse- or intestinal type. A total of 46 genes passed this filter and were selected for an additional, supervised clustering algorithm.

Quantitative RT–PCR

We selected three of the discriminating genes (*SLC25A4*, *GFRA2*, *HSPCB*) and examined their expression levels by means of real-time PCR experiments (Taqman PCR; Applied Biosystems, Foster City, CA, USA). *QARS* served as an internal control because it had shown the smallest Cy3/Cy5 fluctuation in our data for both diffuse- and intestinal-type cancers. The Taqman assay was carried out according to the manufacturer's protocol, with the same aRNAs that had been used for microarray analysis. The PCR reaction involved initial denaturation at 95°C for 10 min, followed by 45 cycles at 95°C for 15 s and 60°C for 1 min. The sequences of primers and probes were as follows: *QARS* forward, 5′-GGTGGATGCAGCATTAG TGGA-3′ and reverse, 5′-AAGACGCTCAAACTGGAA CTTGTC-3′; probe, 5′-VIC-CTCTGTGGCCCTGGCAAAA CCCTT-TAMRA-3′; *SLC25A4* forward, 5′-AGTTGACTG CTGGAGGAAGATTG-3′ and reverse, 5′-ATTGGACCA GGCACCTTTGA-3′; probe, 5′-FAM-CTTGGCTCCTTCG TCTTTT-TAMRA-3′; *GFRA2* forward, 5′-GTGGCGAGG CATTAAAACTTG-3′ and reverse, 5′-GGACCGTTTCTCT CTGACTTCAA-3′; probe, 5′-FAM-TTCTGCCACCGAGA AAGAA-TAMRA-3′; *HSPCB* forward, 5′-CCCCTGCT GGTGTCTAGTGTTT-3′ and reverse, 5′-CCAATCCTGC

TGTCAAGAGTAGAG-3′; probe, 5′-FAM-ACACCCTTAG TTTACTGCCT-TAMRA-3′.

Identification of gene-expression profiles related to histological data

We divided the 20 samples into three groups of two opposing characteristics according to histopathological data, namely (1) vessel invasion-positive or -negative, (2) lymphatic vessel-involvement-positive or -negative, and (3) lymph node-metastasis-positive or -negative, and performed random-permutation analyses. Genes having signal intensities higher than the cutoff value in more than 50% of the cases, *P*-values less than 0.01, and $|Med_d\text{-}Med_i| \geqslant 1.3$, 1.2 and 0.5 were selected as discriminating genes for venous invasion, lymphatic vessel-involvement and LN metastasis, respectively. These procedures selected for subsequent cluster analysis 13 genes that distinguished group 1, 11 for group 2, and 31 for group 3.

Acknowledgements
We thank Ms Tae Makino for technical assistance, Ms Emi Ichihashi for data analysis, Dr Naoki Oyaizu for reviewing the pathological data of samples, Dr Arimichi Takabayashi, Dr Koichi Matsuo, Dr Akira Togashi, Dr Kazutaka Obama, Dr Takefumi Kikuchi, Dr Toshihiko Nishidate and Dr Soji Kakiuchi for helpful discussions and Dr Meiko Takahashi for preparation of the manuscript. This work was supported in part by Research for the Future Program Grant (00L01402) from the Japan Society for the Promotion of Science.

References

Balkwill F. (2003). *Semin. Immunol.*, **15,** 49–55.
Becker KF, Atkinson MJ, Reich U, Becker I, Nekarda H, Siewert JR and Hofler H. (1994). *Cancer Res.*, **54,** 3845–3852.
Bootcov MR, Bauskin AR, Valenzuela SM, Moore AG, Bansal M, He XY, Zhang HP, Donnellan M, Mahler S, Pryor K, Walsh BJ, Nicholson RC, Fairlie WD, Por SB, Robbins JM and Breit SN. (1997). *Proc. Natl. Acad. Sci. USA*, **94,** 11514–11519.
Boussioutas A, Li H, Liu J, Waring P, Lade S, Holloway AJ, Taupin D, Gorringe K, Haviv I, Desmond PV and Bowtell DD. (2003). *Cancer Res.*, **63,** 2569–2577.
Campbell DJ and Butcher EC. (2002). *J. Clin. Invest.*, **110,** 1079–1081.
Chen X, Leung SY, Yuen ST, Chu KM, Ji J, Li R, Chan AS, Law S, Troyanskaya OG, Wong J, So S, Botstein D and Brown PO. (2003). *Mol. Biol. Cell*, **14,** 3208–3215.
Chung YJ, Park SW, Song JM, Lee KY, Seo EJ, Choi SW and Rhyu MG. (1997). *Oncogene*, **15,** 1719–1726.
Couldrey C and Green JE. (2000). *Breast Cancer Res.*, **2,** 321–323.
Craanen ME, Blok P, Dekker W, Offerhaus GJ and Tytgat GN. (1995). *Gut*, **36,** 848–852.
Crnogorac-Jurcevic T, Efthimiou E, Nielsen T, Loader J, Terris B, Stamp G, Baron A, Scarpa A and Lemoine NR. (2002). *Oncogene*, **21,** 4587–4594.
Dan S, Tsunoda T, Kitahara O, Yanagawa R, Zembutsu H, Katagiri T, Yamazaki K, Nakamura Y and Yamori T. (2002). *Cancer Res.*, **62,** 1139–1147.
De S, Chen J, Narizhneva NV, Heston W, Brainard J, Sage EH and Byzova TV. (2003). *J. Biol. Chem.*, **278,** 39044–39050.
Eisen MB, Spellman PT, Brown PO and Botstein D. (1998). *Proc. Natl. Acad. Sci. USA*, **95,** 14863–14868.
Hasegawa S, Furukawa Y, Li M, Satoh S, Kato T, Watanabe T, Katagiri T, Tsunoda T, Yamaoka Y and Nakamura Y. (2002). *Cancer Res.*, **62,** 7012–7017.
Hippo Y, Taniguchi H, Tsutsumi S, Machida N, Chong JM, Fukayama M, Kodama T and Aburatani H. (2002). *Cancer Res.*, **62,** 233–240.
Hippo Y, Yashiro M, Ishii M, Taniguchi H, Tsutsumi S, Hirakawa K, Kodama T and Aburatani H. (2001). *Cancer Res.*, **61,** 889–895.
Jiang Y, Goldberg ID and Shi YE. (2002). *Oncogene*, **21,** 2245–2252.
Kim B, Bang S, Lee S, Kim S, Jung Y, Lee C, Choi K, Lee SG, Lee K, Lee Y, Kim SS, Yeom YI, Kim YS, Yoo HS, Song K and Lee I. (2003). *Cancer Res.*, **63,** 8248–8255.
Kitahara O, Furukawa Y, Tanaka T, Kihara C, Ono K, Yanagawa R, Nita ME, Takagi T, Nakamura Y and Tsunoda T. (2001). *Cancer Res.*, **61,** 3544–3549.
lacobuzio-Donahue CA, Maitra A, Olsen M, Lowe AW, Van Heek NT, Rosty C, Walter K, Sato N, Parker A, Ashfaq R, Jaffee E, Ryu B, Jones J, Eshleman JR, Yeo CJ, Cameron JL, Kern SE, Hruban RH, Brown PO and Goggins M. (2003). *Am. J. Pathol.*, **162,** 1151–1162.
Lauren P. (1965). *Acta Pathol. Microbiol. Scand.*, **64,** 31–49.
Lee JH, Han SU, Cho H, Jennings B, Gerrard B, Dean M, Schmidt L, Zbar B and Vande Woude GF. (2000). *Oncogene*, **19,** 4947–4953.
Leung SY, Chen X, Chu KM, Yuen ST, Mathy J, Ji J, Chan AS, Li R, Law S, Troyanskaya OG, Tu IP, Wong J, So S, Botstein D and Brown PO. (2002). *Proc. Natl. Acad. Sci. USA*, **99,** 16203–16208.
Maehara Y, Kakeji Y, Kabashima A, Emi Y, Watanabe A, Akazawa K, Baba H, Kohnoe S and Sugimachi K. (1999). *J. Clin. Oncol.*, **17,** 607–614.
Mahabeleshwar GH and Kundu GC. (2003). *J. Biol. Chem.*, **278,** 6209–6221.
Ming SC. (1998). *Gastric Cancer*, **1,** 31–50.
Nakamura T, Furukawa Y, Nakagawa H, Tsunoda T, Ohigashi H, Murata K, Ishikawa O, Ohgaki K, Kashimura N, Miyamoto M, Hirano S, Kondo S, Katoh H, Nakamura Y and Katagiri T. (2004). *Oncogene*, **23,** 2385–2400.
Padovan E, Terracciano L, Certa U, Jacobs B, Reschner A, Bolli M, Spagnoli GC, Borden EC and Heberer M. (2002). *Cancer Res.*, **62,** 3453–3458.
Park WS, Oh RR, Park JY, Lee JH, Shin MS, Kim HS, Lee HK, Kim YS, Kim SY, Lee SH, Yoo NJ and Lee JY. (2000). *Gastroenterology*, **119,** 691–698.
Park WS, Oh RR, Park JY, Lee SH, Shin MS, Kim YS, Kim SY, Lee HK, Kim PJ, Oh ST, Yoo NJ and Lee JY. (1999). *Cancer Res.*, **59,** 4257–4260.
Parkin DM. (2001). *Lancet Oncol.*, **2,** 533–543.
Peddanna N, Holt S and Verma RS. (1995). *Anticancer Res.*, **15,** 2055–2064.
Ramaswamy S, Ross KN, Lander ES and Golub TR. (2003). *Nat. Genet.*, **33,** 49–54.
Schwartz DR, Wu R, Kardia SL, Levin AM, Huang CC, Shedden KA, Kuick R, Misek DE, Hanash SM, Taylor JM,

Reed H, Hendrix N, Zhai Y, Fearon ER and Cho KR. (2003). *Cancer Res.*, **63,** 2913–2922.

Sipponen P. (1995). *Am. J. Surg. Pathol.*, **19** (Suppl 1)**,** S30–S36.

Skonier J, Neubauer M, Madisen L, Bennett K, Plowman GD and Purchio AF. (1992). *DNA Cell Biol.*, **11,** 511–522.

St Croix B, Rago C, Velculescu V, Traverso G, Romans KE, Montgomery E, Lal A, Riggins GJ, Lengauer C, Vogelstein B and Kinzler KW. (2000). *Science*, **289,** 1197–1202.

Sumi T, Yoshida H, Hyun Y, Yasui T, Matsumoto Y, Hattori K, Sugimura K, Kawashima H, Nakatani T and Ishiko O. (2003). *Oncol. Rep.*, **10,** 345–349.

Takemoto H, Doki Y, Shiozaki H, Imamura H, Utsunomiya T, Miyata H, Yano M, Inoue M, Fujiwara Y and Monden M. (2001). *Int. J. Cancer*, **91,** 783–788.

Tay ST, Leong SH, Yu K, Aggarwal A, Tan SY, Lee CH, Wong K, Visvanathan J, Lim D, Wong WK, Soo KC, Kon OL and Tan P. (2003). *Cancer Res.*, **63,** 3309–3316.

Toyama T, Iwase H, Yamashita H, Hara Y, Omoto Y, Sugiura H, Zhang Z and Fujii Y. (2003). *Cancer Lett.*, **189,** 97–102.

Yanagawa R, Furukawa Y, Tsunoda T, Kitahara O, Kameyama M, Murata K, Ishikawa O and Nakamura Y. (2001). *Neoplasia*, **3,** 395–401.

Zeng ZS, Shu WP, Cohen AM and Guillem JG. (2002). *Clin. Cancer Res.*, **8,** 144–148.

Supplementary Information accompanies the paper on Oncogene website (http://www.nature.com/onc)