

## Pfetin as a Prognostic Biomarker of Gastrointestinal Stromal Tumors Revealed by Proteomics

Yoshiyuki Suehara,<sup>1,7,8</sup> Tadashi Kondo,<sup>1</sup> Kunihiro Seki,<sup>4</sup> Tatsuhiro Shibata,<sup>2</sup> Kiyonaga Fujii,<sup>1</sup> Masahiro Gotoh,<sup>3</sup> Tadashi Hasegawa,<sup>4</sup> Yasuhiro Shimada,<sup>5</sup> Mitsuru Sasako,<sup>6</sup> Tadakazu Shimoda,<sup>4</sup> Hisashi Kurosawa,<sup>8</sup> Yasuo Beppu,<sup>7</sup> Akira Kawai,<sup>7</sup> and Setsuo Hirohashi<sup>1</sup>

**Abstract Purpose:** We aimed to develop prognostic biomarkers for gastrointestinal stromal tumors (GIST) using a proteomic approach.

**Experimental Design:** We examined the proteomic profile of GISTs using two-dimensional difference gel electrophoresis. The prognostic performance of biomarker candidates was examined using a large-scale sample set and specific antibodies.

**Results:** We identified 43 protein spots whose intensity was statistically different between GISTs with good and poor prognosis. Mass spectrometric protein identification showed that the 43 spots corresponded to 25 distinct gene products. Eight of the 43 spots derived from pfetin, a potassium channel protein, and four of the eight pfetin spots had a high discriminative power between the two groups. Western blotting and real-time PCR showed that pfetin expression and tumor metastasis were inversely related. The prognostic performance of pfetin was also examined by immunohistochemistry on 210 GIST cases. The 5-year metastasis-free survival rate was 93.9% and 36.2% for patients with pfetin-positive and pfetin-negative tumors, respectively ( $P < 0.0001$ ). Univariate and multivariate analyses revealed that pfetin expression was a powerful prognostic factor among the clinicopathologic variables examined, including risk classification and c-kit – or platelet-derived growth factor receptor A mutation status.

**Conclusions:** These results establish pfetin as a powerful prognostic marker for GISTs and may provide novel therapeutic strategies to prevent metastasis of GIST.

Gastrointestinal stromal tumors (GIST) are the most common primary mesenchymal tumors of the digestive tract, with a prevalence of 15 to 20 per 1,000,000 (1, 2). The clinical course of GISTs spans a wide spectrum from a curable disorder to a highly malignant disease that leads to metastasis and death. Thus, the molecular background of

GISTs has been extensively studied to predict the behavior of individual tumors. GISTs are characterized by the presence of mutations and overexpression of c-kit and, clinically, by their response to treatment with imatinib (3–7). Tumor location and certain molecular aberrations, including c-kit, platelet-derived growth factor (PDGFR), and p16 alterations, have been found to correlate with patient prognosis and response to treatment with imatinib (8–12). However, none of these variables has been proven to be clinically useful in improving patient outcomes yet. GISTs arise from the intestinal cells of Cajal, which are the mesenchymal pacemaker cells of the guts (13), their biological characteristics, however, remain largely obscure.

Recent comprehensive studies offered a global view of molecular aberrations associated with the malignant spectrum of GISTs. Genomic studies using fluorescence *in situ* hybridization and array-based comparative genomic hybridization identified chromosomal regions frequently amplified and target genes within these regions, the copy number status of which correlated with tumor behavior (14, 15). Global mRNA expression studies using DNA microarrays identified the genes that are involved in the signaling pathways specific to kit or PDGFR and aberrantly regulated in GISTs (16), the genes associated with histologic features denoting malignancy (17), and the genes differentially expressed based on the KIT genotype and GIST anatomic site (18). These comprehensive studies will further increase our understanding of the biology of GIST and lead to the

**Authors' Affiliations:** <sup>1</sup>Proteome Bioinformatics Project, <sup>2</sup>Cancer Genomics Project, and <sup>3</sup>Pathology Division, National Cancer Center Research Institute; <sup>4</sup>Clinical Laboratory Division, <sup>5</sup>Gastrointestinal Oncology Division, <sup>6</sup>Gastric Surgery Division, and <sup>7</sup>Orthopedic Surgery Division, National Cancer Center Hospital; and <sup>8</sup>Department of Orthopedic Surgery, Juntendo University School of Medicine, Tokyo, Japan

Received 6/15/07; revised 12/25/07; accepted 1/9/08.

**Grant support:** Ministry of Health, Labor, and Welfare and by the Program for Promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation of Japan (patent no. 2006-286087).

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

**Note:** Supplementary data for this article are available at Clinical Cancer Research Online (<http://clincancerres.aacrjournals.org/>).

Current address for T. Hasegawa: Department of Clinical Pathology, Sapporo Medical University School of Medicine.

**Requests for reprints:** Tadashi Kondo, Proteome Bioinformatics Project, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. Phone: 81-3-3542-2511, ext. 3004; Fax: 81-3-357-5298; E-mail: takondo@gan2.res.ncc.go.jp.

© 2008 American Association for Cancer Research.  
doi:10.1158/1078-0432.CCR-07-1478

**Table 1.** Clinicopathologic features of the training set samples (17 cases)

Sample No	Age	Gender	Site	Histologic types	Size (cm)	Risk classification*
1	68	M	Stomach	Spindle	19	High
2	56	M	Stomach	Spindle	38	High
3	58	M	Stomach	Spindle	13	High
4	50	F	Rectum	Spindle	4	High
5	51	M	Stomach	Mixed (spindle main)	12	High
6	34	M	Small intestine	Spindle	18	High
7	68	M	Small intestine	Mixed (spindle main)	7	High
8	72	F	Stomach	Spindle	25	High
9	64	F	Stomach	Spindle	3.5	Low
10	64	F	Stomach	Spindle	4	Low
11	54	M	Stomach	Spindle	10	Intermediate
12	68	M	Small intestine	Spindle	3.7	Low
13	77	M	Small intestine	Spindle	4	Low
14	40	F	Stomach	Spindle	10	Intermediate
15	52	M	Stomach	Spindle	7	Intermediate
16	76	M	Small intestine	Spindle	7	Intermediate
17	81	M	Stomach	Spindle	5.5	Intermediate

NOTE: PDGFR mutations: All samples lacked of PDGFR mutations. Detail data: Supplementary Table S1.

Abbreviations: NED, no evidence of disease; AWD: alive with disease; DOD, dead of disease.

\*Prognostic classification based on tumor size and MIB-1 grade (Hasegawa, T. et al. Hum Pathol. 2002, 33:669-676).

development of practical tumor markers to support individualized therapy (8). Emerging technologies that examine the overall features of the expressed proteins, namely the proteome, have identified many candidate proteins associated with early diagnosis (19), differential diagnosis (20), prognosis (21), and response to chemotherapy (22) in various diseases. Many lines of evidence have indicated that DNA copy number and mRNA expression levels do not necessarily correspond to the protein contents, and that posttranslational modifications cannot be predicted by DNA sequences (23, 24), suggesting that proteomic studies offer unique data that cannot be obtained by other approaches. The proteomic profile of GISTs has not been established yet, and a proteomic study using a large-scale clinical sample set would complement the genome and transcriptome studies.

In this report, we did a comprehensive quantitative expression study on the intact proteins of GIST clinical samples using two-dimensional difference gel electrophoresis and mass spectrometry. Proteomic studies on peptides have been used to develop tumor markers, but intact proteins have not been considered for this purpose, with a few exceptions. Two-dimensional difference gel electrophoresis, as the most advanced form of two-dimensional gel electrophoresis, has the great advantage of being able to be used to study intact proteins. We found that the expression levels of 43 proteins, including eight variants of pftin (predominantly fetal-expressed tetramerization domain; potassium channel tetramerisation domain containing protein 12), which was originally reported as a protein highly expressed in fetal cochlea and brain (25), correlated with prognosis. We verified the prognostic value of pftin expression on 210 GIST cases using immunohistochemistry. Our findings indicate that the use of pftin expression as a prognostic indicator may facilitate tailored medical care for GIST patients.

## Materials and Methods

**Patients and clinical information.** We examined the tumor tissues of 212 GIST patients who underwent surgery at the National Cancer Center Hospital consecutively from October 1977 to December 2005. All patients underwent resection with curative intent and were not treated with adjuvant chemotherapy, including treatment with imatinib, until distant metastasis was diagnosed. Histologic features of the tissues were reviewed by three board-certified pathologists (K.S., T.S., and T.H.). Diagnosis was based on the WHO classification system for soft-tissue tumors (26), including the examination of tumor size, presence of necrosis, differentiation, mitotic rate, MIB-1 index, presence of epithelioid cells, and CD34 and CD117 expression. Using this large, well-characterized single hospital-based sample set, we were able to identify proteomic features that differ significantly when examined in relation to certain clinicopathologic variables. This project was approved by the institutional review board of the National Cancer Center.

Previous reports indicated that GIST patients that were histologically classified as of being at low or intermediate risk did not develop metastases within 2 y postsurgery, whereas GIST patients histologically classified as of being at high risk developed metastases within 1 y postsurgery (27). For proteomic analysis, we grouped the GIST samples into two groups. GISTs that had metastases at diagnosis or developed metastases within 1 y postsurgery and were categorized in the high-risk group based on their histologic features were defined as poor-prognosis GISTs (P-GIST; Table 1, samples 1-8). GISTs that did not have metastases within 2 y postsurgery and were grouped in the low- or intermediate-risk group based on the histologic features were defined as good-prognosis GISTs (G-GIST; Table 1, samples 9-17). The samples listed in Supplementary Table S1 were excluded from this classification; samples 18, 19, 24, and 25 were excluded because RNA data were not available for the validation study and the other samples because they did not meet the criteria for classification either as P-GISTs or G-GISTs.

For the immunohistochemical study, we selected 210 patients who did not have distant metastases at the time of surgery.

**Protein expression profiling.** Frozen samples were crushed to powder with a CryoPress (Microtech Nichion) under cooling with liquid nitrogen. The frozen powder was then treated with urea lysis buffer

**Table 1.** Clinicopathologic features of the training set samples (17 cases) (Cont'd)

Type of KIT mutation	Metastatic site (first development)	Metastasis time after diagnosis (mo)	Follow-up time after diagnosis (mo)	Follow-up status
Wild-type	Peritoneal metastasis	8	16	DOD
EX11 deletion	Peritoneal metastasis	7	9	DOD
EX11 deletion	Peritoneal metastasis	6	11	DOD
EX11 deletion	Peritoneal metastasis	11	60	AWD
EX11 deletion	Peritoneal metastasis	5	69	AWD
EX11 deletion	Liver metastasis	At diagnosis	31	AWD
EX9 insertion	Peritoneal metastasis	At diagnosis	9	AWD
EX11 559 V-D	Peritoneal metastasis	5	8	AWD
Wild-type	—	—	68	NED
Wild-type	—	—	81	NED
EX11 560 V-G	—	—	77	NED
EX9 insertion	—	—	50	NED
Wild-type	—	—	69	NED
EX11 559 V-D	—	—	88	NED
EX11 576 L-P	—	—	62	NED
EX11 deletion	—	—	48	NED
EX11 559 V-D	—	—	43	NED

(6 mol/L urea, 2 mol/L thiourea, 3% CHAPS, and 1% Triton X-100). After centrifugation at 15,000 rpm for 30 min, the supernatant was used as the source of cellular proteins for protein expression studies.

Two-dimensional difference gel electrophoresis was done as described previously (20, 28, 29). In brief, the internal control sample was prepared by mixing a portion of all individual samples. Five micrograms of the internal control sample and of each individual sample were labeled with Cy3 and Cy5, respectively (CyDye DIGE Fluor saturation dye; GE Healthcare Biosciences) according to the manufacturer's instructions. The differently labeled protein samples were mixed and separated by two-dimensional gel electrophoresis, which was achieved using IPG DryStrip gels for the first dimension separation (length, 24 cm; isoelectric point range, between 4 and 7; GE Healthcare Biosciences) and SDS-PAGE for the second dimension separation (EttanDalt II; GE Healthcare Biosciences). The gels were scanned using laser scanners (Typhoon Trio; GE Healthcare Biosciences) at appropriate wavelengths. For all spots, the intensity of the Cy5 image was normalized by that of the Cy3 image in the identical gel so that gel-to-gel differences were compensated, using the DeCyder image software (GE Healthcare Biosciences). System reproducibility was verified by comparing the protein profiles obtained from three independent separations of the identical sample (sample 22; Supplementary Table S1). Scatter plot analysis revealed that the standardized intensity of >96% of the spots ranged within a 2.0-fold difference (Supplementary Fig. S1). Representative two-dimensional images with the numbers of the identified spots are shown in Fig. 1A and Supplementary Fig. S3.

**Data analysis.** We identified protein spots whose intensity was statistically (Wilcoxon test,  $P < 0.001$ ) different between the groups examined. Hierarchical clustering, principal component analysis, correlation matrix study, and spot ranking were done using the Expressionist software (Genedata).

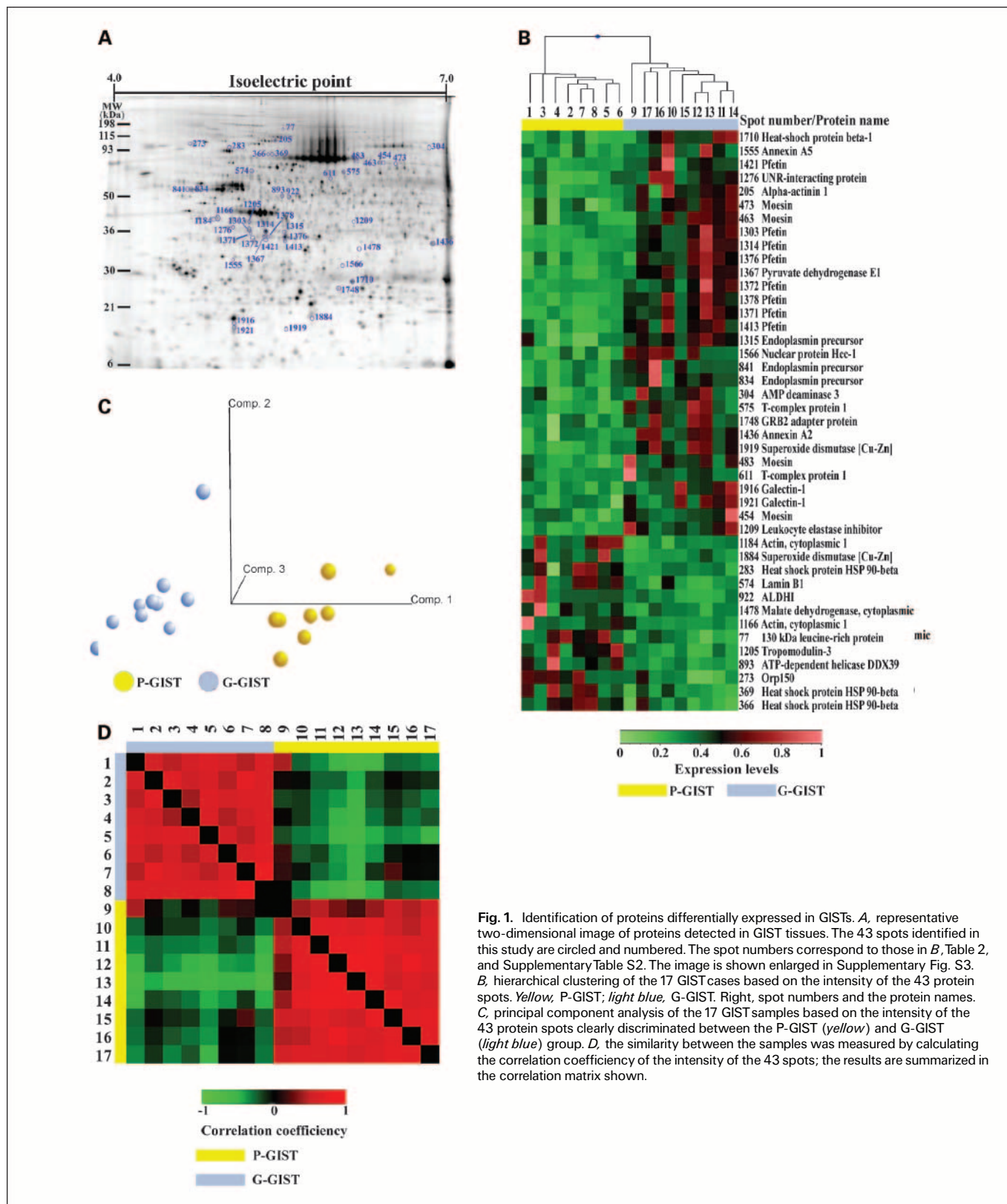
**Protein identification by mass spectrometry.** Proteins corresponding to the spots of interest were identified by mass spectrometry according to our previous report (20, 30). Cy5-labeled proteins separated by 2D-PAGE were recovered in gel plugs and digested with modified trypsin (Promega). The trypsin digests were subjected to liquid chromatography coupled with tandem mass spectrometry equipped with a nano-electrospray ion source (Paradigm MS4 dual solvent delivery system; Michrom BioResources, Inc.) for microflow high performance liquid chromatography, an HTS PAL auto sampler (CTC Analytics), and a Finnigan LTQ linear ion trap mass spectrometer (Thermo Electron Co.) equipped with a nano-electrospray ion source (AMR, Inc.). The Mascot software (version 2.1; Matrix Science) was used

to search for the mass of the peptide ion peaks against the SWISS-PROT database (Homo sapiens, 12867 sequence in Sprot\_47.8 fasta file). Proteins with a Mascot score of 35 or more were used for protein identification. When multiple proteins were identified in a single spot, the proteins with the highest number of peptides were considered as those corresponding to the spot.

**Western blotting and immunohistochemistry.** Protein samples were separated by SDS-PAGE and subsequently blotted on a nitrocellulose membrane. The membrane was incubated with rabbit polyclonal antibody against pfetin (1:1,000 dilution) kindly provided by Dr. Morton (25), and then horseradish peroxidase-conjugated secondary antibody (1:1,000 dilution; GE Healthcare Biosciences). Pftetin was detected using an enhanced chemiluminescence system (GE Healthcare Biosciences) and LA 1000 (Fuji film).

Pftetin expression was examined immunohistochemically using paraffin-embedded tissues. In brief, 4- $\mu$ m-thick tissue sections were autoclaved in 10 mmol/L citrate buffer (pH 6.0) at 121°C for 30 min and incubated with the antibody against pfetin (1:500 dilution). Immunostaining was done according to the streptavidin-biotin peroxidase method using the Strept ABC Complex/horseradish peroxidase kit (DAKO). One pathologist (K. S.) and one medical doctor (Y. S.) reviewed the sections stained with antipftetin antibody in a blinded fashion regarding clinical data (age, sex, anatomic site, and outcome). Positively stained cells were defined as those that had higher staining intensity than that of vascular endothelial cells, which served as positive controls. Cases with >20% of tumor cells stained positively with the antipftetin antibody were considered as pfetin positive, whereas cases with <20% pfetin-positive tumor cells were considered as pfetin negative. In most cases, the difference was so obvious that two reviewers had consistent results. Inconsistencies, if any, were resolved by discussion, as a usual process of pathologic diagnosis in the hospital.

**Mutation study for c-kit and PDGFRA.** We examined the c-kit and PDGFRA genes for the presence of mutations as previously described (31) in the 39 cases where DNA samples were available. In brief, DNA was extracted from the frozen tissues, and the exons including the frequent mutation sites for c-kit and PDGFRA were amplified by PCR. The PCR products were purified with 2% agarose gel electrophoresis, extracted with an QIAquick PCR Purification kit (Qiagen), and sequenced using an ABI Prism 3100 Genetic Analyzer (Applied Biosystems). The primer sets for c-kit were as follows: 5'-TCTAGTG-CATTCAAGCACAATGG-3' and 5'-CATGACTGATATGGTAGACAGAG-3' for exon 9, and 5'-CCAGAGTGGCTCTAATGACTGAGAC-3' and 5'-AAAGGTGACATGGAAAGCCCCTG-3' for exon 11. The primer set for exon 12 of the PDGFRA gene was 5'-CCTGGTCATTTATAGAAACCGAG-3'



**Fig. 1.** Identification of proteins differentially expressed in GISTs. *A*, representative two-dimensional image of proteins detected in GIST tissues. The 43 spots identified in this study are circled and numbered. The spot numbers correspond to those in *B*, Table 2, and Supplementary Table S2. The image is shown enlarged in Supplementary Fig. S3. *B*, hierarchical clustering of the 17 GIST cases based on the intensity of the 43 protein spots. *Yellow*, P-GIST; *light blue*, G-GIST. Right, spot numbers and the protein names. *C*, principal component analysis of the 17 GIST samples based on the intensity of the 43 protein spots clearly discriminated between the P-GIST (*yellow*) and G-GIST (*light blue*) group. *D*, the similarity between the samples was measured by calculating the correlation coefficient of the intensity of the 43 spots; the results are summarized in the correlation matrix shown.

and 5'-CTCCCATCTTGAGTCATAAGGCA-3'. PCR cycling variables were as follows: cycle at 96°C for 1 min; 50 cycles at 94°C for 30 s, 56°C for 30 s, and 68°C for 2.5 min; and finally 1 cycle at 72°C for 5 min.

**Quantitative reverse transcription-PCR.** We extracted mRNA and generated cDNA using the Super Script III kit (Invitrogen) in the 39 cases where mRNA samples were available. The quantitative amplification was monitored with Taq Man Gene Expression Assays

**Table 2.** A list of identified protein

Spots no*	Accession no <sup>†</sup>	Identified protein <sup>†</sup>	Wilcoxon test <i>P</i>	Fold difference (ratio of means)	Overall rank
1378	Q96CX2	Potassium channel tetramerisation domain containing protein 12 (Pfetin)	8.23E-05	4.862	1
1314	Q96CX2	Potassium channel tetramerisation domain containing protein 12 (Pfetin)	8.23E-05	5.805	2
1,371	Q96CX2	Potassium channel tetramerisation domain containing protein 12 (Pfetin)	8.23E-05	3.781	5
1413	Q96CX2	Potassium channel tetramerisation domain containing protein 12 (Pfetin)	1.65E-04	3.198	26
575	P17987	T-complex protein 1, $\alpha$ subunit	3.29E-04	1.405	28
893	O00148	ATP-dependent helicase DDX39	6.99E-04	2.321	41
1367	P11177	Pyruvate dehydrogenase estrone component $\beta$ subunit	8.94E-04	3.669	24
1303	Q96CX2	Pfetin	9.87E-04	2.917	4
1919	P00441	Superoxide dismutase (Cu-Zn)	1.33E-03	2.547	12
611	P17987	T-complex protein 1, $\alpha$ subunit	1.56E-03	1.982	9
1184	P60709	Actin, cytoplasmic 1	1.56E-03	1.902	20
1376	Q96CX2	Pfetin	1.56E-03	2.850	38
834	P14625	Endoplasmic precursor	1.75E-03	3.487	8
483	P26038	Moesin	1.86E-03	1.558	15
369	P08238	Heat shock protein HSP 90- $\beta$	1.86E-03	1.977	16
922	P05091	Aldehyde dehydrogenase, mitochondrial precursor	2.04E-03	1.359	31
273	Q9Y4L1	150 kDa oxygen-regulated protein precursor	3.32E-03	1.384	36
1436	P07355	Annexin A2	3.32E-03	1.590	39
1,748	P62993	Growth factor receptor-bound protein 2	3.32E-03	1.545	40
205	P12814	$\alpha$ -actinin 1	3.70E-03	1.993	11
1710	P04792	Heat-shock protein $\beta$ -1	3.70E-03	2.078	13
574	P20700	Lamin B1	3.70E-03	2.575	14
1555	P08758	Annexin A5	3.70E-03	1.438	19
454	P26038	Moesin	3.70E-03	1.441	37
1315	P14625	Endoplasmic precursor	3.73E-03	2.895	23
1276	Q9Y3F4	Serine-threonine kinase receptor-associated protein	3.85E-03	1.650	29
304	Q01432	AMP deaminase 3	4.00E-03	1.691	22
1372	Q96CX2	Pfetin	4.04E-03	2.972	30
1478	P40925	Malate dehydrogenase, cytoplasmic	4.08E-03	1.534	27
1916	P09382	Galectin-1	4.21E-03	1.917	7
1209	P30740	Leukocyte elastase inhibitor	4.33E-03	2.380	18
1884	P00441	Superoxide dismutase (Cu-Zn)	5.25E-03	3.121	3
77	P42704	130 kDa leucine-rich protein	5.51E-03	1.456	21
1166	P60709	Actin, cytoplasmic 1	5.51E-03	3.971	42
1921	P09382	Galectin-1	6.22E-03	1.912	6
1421	Q96CX2	Pfetin	6.22E-03	1.665	34
366	P08238	Heat shock protein HSP 90- $\beta$	7.02E-03	1.705	25
1566	P82979	Nuclear protein Hcc-1	7.02E-03	1.526	43
473	P26038	Moesin	7.59E-03	1.323	32
1205	Q9NYL9	Tropomodulin-3	7.90E-03	1.935	10
283	P08238	Heat shock protein HSP 90- $\beta$	7.90E-03	1.300	33
463	P26038	Moesin	7.90E-03	1.524	35
841	P14625	Endoplasmic precursor	9.45E-03	2.478	17

\*Spot numbers refer to those in Fig. 1A and Supplementary Fig. S3.

<sup>†</sup> Accession numbers of protein were derived from SWISS-PROT and National Center for Biotechnology Information nonredundant databases.

<sup>‡</sup> Observed isoelectric point and molecular weight calculated according to location on the two-dimensional gel.

<sup>§</sup> Theoretical isoelectric point and molecular weight obtained from Swiss-Prot and the ExPASy database. (<http://au.expasy.org>).

<sup>||</sup> Mascot score for the identified proteins based on the peptide ions score ( $P < 0.05$ ; <http://www.matrixscience.com>).

using premade primers for pfetin and Human glyceraldehyde-3-phosphate dehydrogenase, and Taq Man Universal PCR Master Mix with 7500 Real-time PCR system (Applied Biosystems) according to the manufacturer's instructions.

**Statistical analysis.** The tumor-specific and metastasis-free survival times were calculated from the first resection of the primary tumor to death from tumor-specific causes or to first evidence of metastasis,

respectively. All time-to-event end points were computed by the Kaplan-Meier method (32). Patients who died of causes unrelated to GISTs were censored at the time of death. Potential prognostic factors were identified by univariate analysis using the log-rank test. Independent prognostic factors were evaluated using a Cox's proportional hazards regression model and a stepwise selection procedure. To arrive at a parsimonious multivariate model, covariates were selected into the

**Table 2.** A list of identified protein (Cont'd)

pI (obs) <sup>‡</sup>	pI (cal) <sup>§</sup>	MW (obs; kDa) <sup>‡</sup>	MW (cal; kD) <sup>§</sup>	Protein score <sup>  </sup>	Peptide matches	Sequence coverage (%)
6.4	5.5	35.1	35.7	406	8	20.0
5.4	5.5	44.0	35.7	338	7	19.7
6.2	5.5	35.7	35.7	146	2	8.3
6.5	5.5	34.5	35.7	306	8	23.4
5.1	5.8	73.3	60.3	267	6	12.4
5.5	5.5	51.8	49.1	109	2	3.7
6.4	6.2	35.1	39.2	174	3	9.7
5.2	5.5	44.7	35.7	236	5	14.5
6.6	5.7	15.3	15.8	60	1	7.8
4.9	5.8	67.9	60.3	667	11	22.5
5.9	5.3	46.3	41.7	276	4	11.2
6.5	5.5	35.1	35.7	396	11	23.7
4.7	4.8	57.2	92.5	977	18	20.5
5.2	6.1	84.0	67.7	208	4	6.9
4.4	5.0	89.4	83.1	263	5	7.3
5.6	6.6	50.0	56.4	323	5	10.1
4.7	5.2	113.7	111.3	102	2	2.4
6.9	7.6	34.5	38.5	262	5	13.3
6.0	5.9	25.5	25.2	79	2	7.4
4.4	5.3	114.1	103.1	593	12	12.9
6.1	6.0	27.5	22.8	491	15	54.1
4.2	5.1	75.1	66.3	1616	33	36.4
6.1	4.9	31.5	35.8	96	2	5.3
5.4	6.1	82.3	67.7	366	8	9.2
5.5	4.8	44.3	92.5	216	5	4.7
5.1	5.0	45.0	38.4	250	5	15.4
4.8	6.5	113.4	88.8	185	3	4.6
6.3	5.5	35.1	35.7	118	2	8.3
6.2	6.9	33.6	36.3	116	2	5.7
6.1	5.3	16.9	14.6	244	6	35.8
5.2	5.9	46.0	42.7	551	9	23.0
6.8	5.7	18.4	15.8	139	2	17.0
4.5	5.5	195.2	145.2	120	2	1.7
5.9	5.3	46.3	41.7	246	7	15.5
6.1	5.3	15.8	14.6	90	3	11.9
6.4	5.5	34.5	35.7	336	13	23.4
4.4	5.0	89.4	83.1	420	8	13.8
6.1	6.1	30.6	23.5	196	2	13.9
5.6	6.1	80.5	67.7	775	15	18.8
5.2	5.1	46.0	39.6	168	3	9.1
4.1	5.0	113.1	83.1	907	17	22.1
5.4	6.1	82.3	67.7	356	7	9.7
4.7	4.8	57.2	92.5	159	3	4.0

Downloaded from <http://aacrjournals.org/clinoncres/article-pdf/14/6/1707/1980614/1707.pdf> by guest on 29 November 2022

model only if they contributed significantly to the fit of the model based on the  $\chi^2$  test value. *P* value differences of <0.05 were considered to be significant. Statistical analyses were done using the SPSS statistical package (SPSS).

## Results

We compared the protein expression profiles between nine G-GISTs and eight P-GISTs using two-dimensional difference gel electrophoresis. We selected 1,513 protein spots that appeared in at least 75% of the images of the Cy3-labeled internal control sample to decrease irrelevant expression data.

The G- and P-GIST samples were not classified into their respective groups based on the overall protein expression features of the samples (Supplementary Fig. S2). However, we found that 43 protein spots had significantly different intensity between the two groups (*P* < 0.01). The localization of the 43 spots on the two-dimensional image is shown in Fig. 1A (enlarged image in Supplementary Fig. S3). Hierarchical clustering and principal component analysis accurately classified the 17 GIST samples into either the G- or the P-GIST group based on the intensity of the 43 selected spots (Fig. 1B and C). The profiles of the 43 spots were similar between samples of the same group and different between

the two groups (Fig. 1D). Mass spectrometric protein identification revealed that the 43 protein spots corresponded to 25 distinct gene products (Fig. 1B; Table 2; Supplementary Table S2).

We aimed to prioritize the protein spots according to their discriminative power for the two groups. We created a classifier based on a support vector machine algorithm that used the intensity of the 43 spots and ranked the 43 spots according to their contribution to the classification using support vector machine weight algorithm (Table 2). We found that 4 of the 8 identified pfetin spots were ranked within the top 10 protein spots whose intensity was different between the groups (Table 2), and that pfetin spots appeared 8 times in the list of the 43 protein spots (Fig. 1B; Table 2).

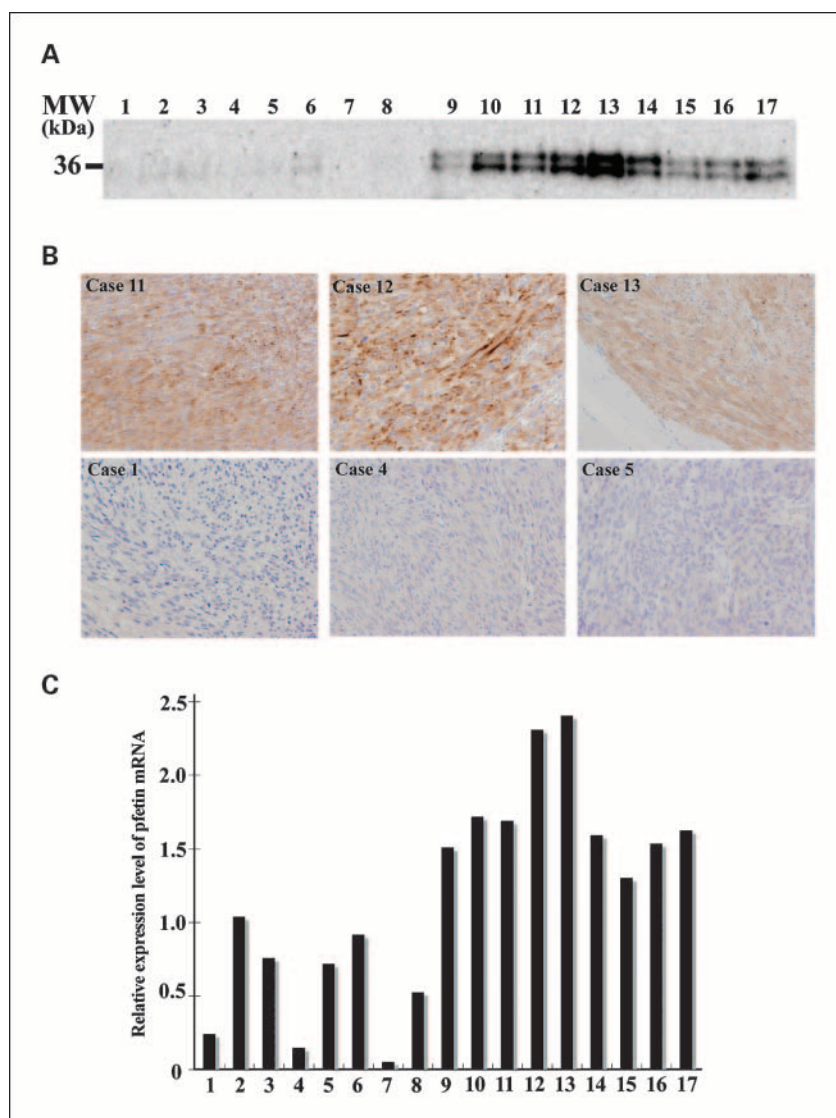
Pfetin is highly expressed in fetal cochlea and brain (25), consistent with the fact that the origin of GIST is Cajal cells, neuronal cells in the gut. Thus, we further validated the relationship of the expression of pfetin with the malignant potential of GISTs. SDS-PAGE/Western blotting showed that the expression of pfetin was lower in the P-GIST compared with the G-GIST group (Fig. 2A). Two bands on SDS-PAGE/

Western blotting corresponded to the location of protein spots for pfetin variants on the two-dimensional image (Fig. 1A; Supplementary Fig. S3). These results were further validated in an additional four GIST samples that were not included in the initial proteome study (Supplementary Fig. S4A). Pfetin expression was not observed in the 6 liver metastases examined or in primary high-risk GISTs that developed metastases between 13 and 30 months postsurgery (Supplementary Fig. S4B; Supplementary Table S1).

We used immunohistochemistry to evaluate pfetin expression *in situ*. Positive cells were diffusely stained with antipfetin antibody in membrane and cytoplasm. All nine G-GISTs expressed pfetin, whereas none of the eight P-GISTs did (Fig. 2B). Pfetin expression was not observed in neighboring host cells or in the intestinal cells of Cajal (Supplementary Fig. S5).

Real-time RT-PCR revealed pfetin mRNA levels were higher in G-GISTs than in P-GISTs (Fig. 2C). However, the difference between the P- and G- GIST group was less obvious at the mRNA than at the protein level, suggesting that pfetin expression is partially regulated at the mRNA level, and that

**Fig. 2.** Validation of the differential expression of pfetin. G-GISTs expressed pfetin at significantly higher levels than P-GISTs. *A*, Western blotting. Case numbers correspond to those in Fig. 1. *B*, immunohistochemistry; pfetin is overexpressed in G-GISTs (*top*), whereas it is not expressed in P-GISTs (*bottom*). *C*, the pfetin mRNA expression levels detected in the 17 GIST samples examined. P-GIST, 1;8; G-GIST, 9;17.



postranscriptional regulation may also play an important role in pftetin expression.

As pftetin expression has been reported to correlate with c-kit mutation status (33), we examined 39 primary GISTs for the presence of c-kit and PDGFRA gene mutations and monitored their pftetin expression levels by Western blotting (Supplementary Tables S1 and S3). Overexpression of pftetin was observed in 12 of 29 c-kit mutation positive cases and in 6 of 10 negative cases ( $P = 0.389$ ; Supplementary Table S3). PDGFRA mutations were not detected in the series. We observed no significant correlation between pftetin expression and c-kit or PDGFR mutation status.

The immunohistochemical study of 210 GISTs revealed a strong correlation between pftetin expression and a number of clinicopathologic variables including the tumor size, mitotic index, MIB-1 index, degree of differentiation, and risk

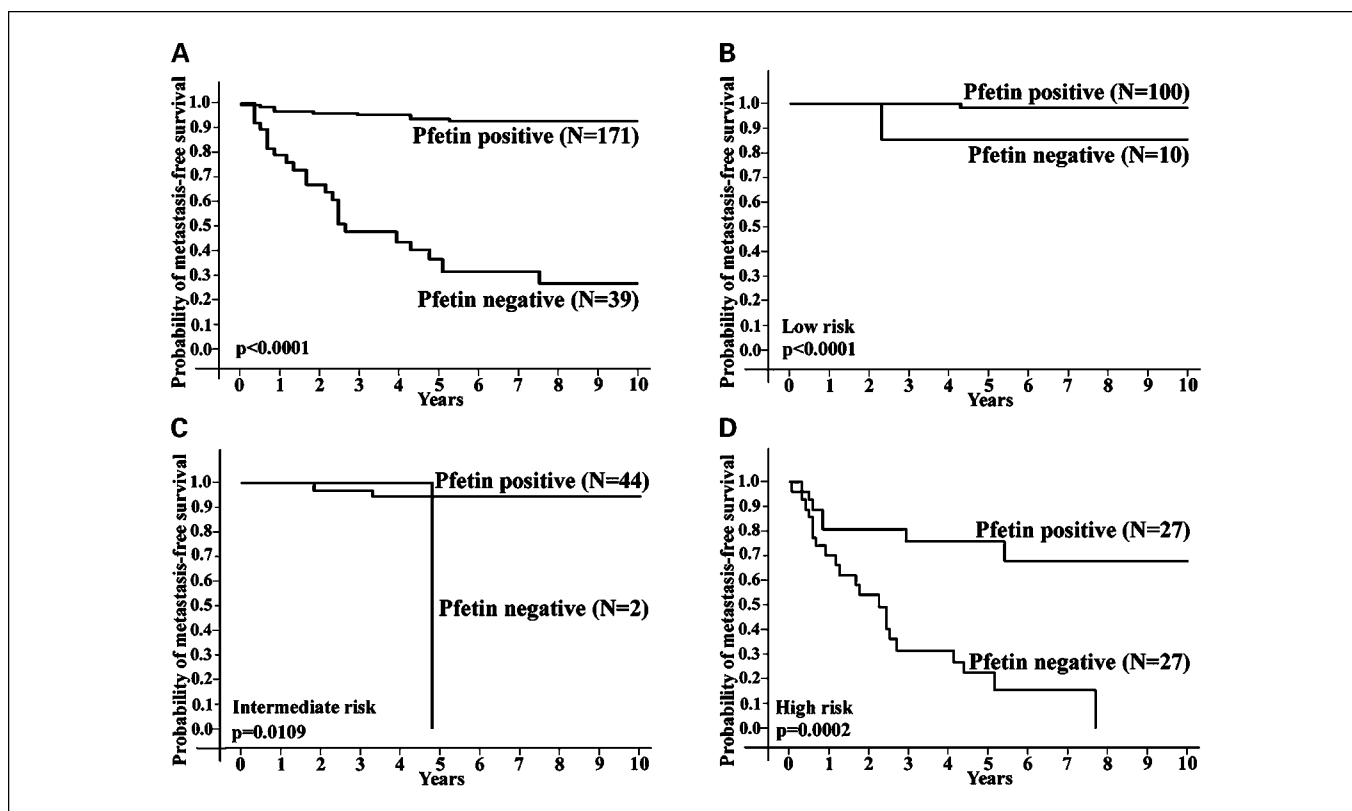
classification ( $P < 0.0001$ ; Table 3). Moreover, distant metastasis was observed in a significantly higher proportion of patients with pftetin-negative tumors compared with those with pftetin-positive tumors (24 of 39 versus 12 of 171 cases;  $P < 0.0001$ ), with a median follow-up period of 73 months. The 5-year metastasis-free survival rate was significantly higher in the pftetin-positive than in the negative group overall (93.9% versus 36.2%;  $P < 0.0001$ ; Fig. 3A; Table 3) as well as within each risk group (Fig. 3B-D). Multivariate analysis revealed that pftetin expression was a powerful predictor of disease-specific survival (Table 3). Note that high-risk cases were divided into two groups, the pftetin-positive and the pftetin-negative group, the latter having a worse prognosis. Furthermore, tumor-specific survival was statistically significantly longer in the pftetin-positive compared with the pftetin-negative group ( $P < 0.0001$ ; Table 3; Supplementary Fig. S6). These data

**Table 3.** Univariate and multivariate analysis of prognostic factor and the relationship between clinicopathologic variables and pftetin expression

Variable	Number of cases	Metastasis-free survival		Tumor-specific survival		Multivariate analysis of metastasis free survival by Cox regression			Pftetin positive (no. cases)	Pftetin negative (no. cases)	Correlation (pftetin) $\chi^2$ (P)
		5 y (%)	Log-rank (P)	5-y (%)	Log-rank (P)	P	Relative risk	95% CI			
Age			0.3290		0.8350						0.5220
<60	112	85.8 ± 3.7		91.8 ± 3.0					93	19	
60<	98	80.7 ± 4.2		95.6 ± 2.2					78	20	
Sex			0.4420		0.0393						0.8909
F	99	86.5 ± 3.7		96.1 ± 2.2					81	18	
M	111	80.4 ± 4.2		91.6 ± 2.8					90	21	
Site			0.0001		0.1655						0.4776
Stomach	170	87.8 ± 2.7		93.9 ± 2.0					140	30	
Nonstomach	40	60.7 ± 9.2		92.2 ± 5.3		0.0270	2.21	1.09-4.49	31	9	
Histology			0.1003		0.2068						0.5153
Spindle	189	85.4 ± 2.8		93.6 ± 2.0					155	34	
Epithelioid	21	67.7 ± 10.9		95.2 ± 4.6					16	5	
Size			<0.0001		<0.0001						<0.0001
<5 cm	128	92.4 ± 2.6		98.1 ± 1.3					112	16	
5-10 cm	63	76.6 ± 0.4		93.7 ± 3.5					51	12	
15 cm<	19	45.0 ± 11.9		60.5 ± 13.0		0.0070	2.05	1.22-3.44	8	11	
Necrosis			<0.0001		0.0034						0.0070
+	19	43.0 ± 12.7		76.9 ± 11.7					10	30	
-	191	86.9 ± 2.7		95.1 ± 1.7					161	9	
Miosis			<0.0001		<0.0001						<0.0001
<5/50HPF*	148	95.5 ± 2.0		98.2 ± 1.3					136	12	
5-10/50HPF	33	80.4 ± 7.2		96.8 ± 3.2					26	7	
5/50HPF<	29	29.9 ± 9.0		68.2 ± 9.4					9	20	
MIB-1			<0.0001		<0.0001						<0.0001
<9%	164	96.0 ± 1.8		98.4 ± 1.1					152	12	
10-29%	19	51.4 ± 12.6		85.6 ± 9.7					11	8	
<30%	27	32.5 ± 9.6		70.8 ± 9.4					8	19	
Differentiation			<0.0001		<0.0001						<0.0001
Score 1	161	95.9 ± 1.8		98.4 ± 1.1					149	12	
Score 2	49	43.5 ± 7.7		78.1 ± 6.6		<0.0001	10.40	3.68-29.45	22	27	
Risk classification			<0.0001		<0.0001						<0.0001
Low	110	97.7 ± 1.6		98.9 ± 1.1					100	10	
Intermediate	46	90.6 ± 5.2		97.0 ± 3.0					44	2	
High	54	48.5 ± 7.4		80.6 ± 6.0					27	27	
Pftetin			<0.0001		<0.0001						
Positive	171	93.9 ± 2.0		97.2 ± 1.4							
Negative	39	36.2 ± 8.7		76.5 ± 7.9		0.0020	3.75	1.60-8.81			

Abbreviation: 95% CI, 95% confidence interval.  
\*HPF, high-power field.





**Fig. 3.** Pfetin expression in the primary tumor samples was predictive of the metastasis-free survival period. Statistically significant differences in the metastasis-free survival period were observed between the pfetin-positive and pfetin-negative groups ( $P < 0.0001$ ) both for the MO GIST patients overall (A) and within each risk group of patients (B-D).

clearly indicate that prognosis relying solely on the established risk classification system is not sufficiently accurate to determine the post-operative therapeutic strategy for GIST patients, and the use of pfetin expression may further refine the prognostic criteria so as to identify patients who may benefit from additional therapy.

## Discussion

Employing proteomics tools, we identified 43 protein variants corresponding to 25 distinct gene products that distinguished GIST patients according to their clinical outcome. The discriminating power of this set of proteins may be developed for use in a clinical setting. However, albeit useful to describe complex clinical variables, cutting-edge proteomic technologies cannot be transferred easily to a hospital setting, considering the high installation costs and labor intensity, the low throughput, and the required operational skills. A smaller number of proteins, measurable by simpler techniques, may be preferable for use in practice. With this notion, we showed that pfetin expression can be examined by SDS-PAGE/Western blotting, immunohistochemistry, and quantitative RT-PCR. Moreover, using a large-scale sample set, we showed that the expression levels of pfetin, as evaluated by immunohistochemistry, are predictive of patient outcome. Therefore, evaluation of pfetin expression can be applied in a clinical setting using the existing examination protocol and may allow the identification of a high-risk patient group that

may benefit from adjuvant therapy, such as treatment with imatinib, while it may also help spare low-risk patients unnecessary treatment. As mass spectrometric global surveys for phosphorylated proteins identified pfetin as a phosphorylated protein (34, 35), the multiple protein spots of pfetin may correspond to the different phosphorylation variants.

Recently, Kang et al. (33) did a proteomic study on 12 GIST samples using two-dimensional gel electrophoresis and reported that pfetin overexpression (C13orf2 in their report) correlated with histologic grading and the presence of c-kit mutations. In contrast, our results indicated that pfetin expression is inversely correlated with histologic grading (Figs. 1 and 2), and that pfetin expression levels are not associated with c-kit mutation status (Supplementary Tables S1 and S3). Moreover, the proteins reported to correlate with histologic malignancy by Kang et al. (33), including Annexin V, HMGB1, glutamate dehydrogenase 1, and fibrinogen  $\beta$  chain RoXaN (33), were not identified as such in our study, whereas 24 gene products identified in this study were not listed in their report (33). As these discrepancies may be due to differences in patient populations and proteomic modalities used, an international project to integrate all reported proteomic data in a common proteomic platform is needed to elucidate the molecular background of GISTs.

Although pfetin is known to contain a voltage-gated potassium ( $K^+$ ) channel tetramerization domain (25), its function in the process of cancer development and progression

is unknown. Although GISTs originate from Cajal cells, immunohistochemistry revealed that pftetin was absent in Cajal cells (Supplementary Fig. S5). Proteomic analysis of pftetin-associated proteins may provide clues to understanding the role of pftetin in GIST development and progression.

Our study has the limitation of not detecting proteins expressed in low levels. We did not observe overexpression of the kit (36–41) or PDGFRA gene products, or loss of CD44 (42) or p16 (11, 12). In addition, we did not detect CD34 (18) or connexin 43 (43) expression, reported to be commonly up-regulated in stomach and small intestinal GISTs, respectively. Aberrant regulation of these gene products was initially detected at the mRNA level and was later confirmed at the protein level using specific antibodies. Presently, any global approach to protein expression cannot uncover the whole proteome in a quantitative and reproducible way. The continuing efforts to improve the sensitivity of proteomic modalities have enabled the uncovering of several thousands of proteins with posttranslational modifications (24, 44, 45). We believe that such efforts will overcome some

of the inherent limitations of proteomics and lead to a more detailed understanding of the disease mechanisms and to novel therapeutic strategies in the near future.

In conclusion, we identified a possible correlation of 43 protein variants corresponding to 25 distinct gene products with variables of clinical interest in GIST and validated pftetin expression using a specific antibody. From this study, pftetin expression is predictive of metastasis and survival of patients with GISTs and, as such, may be used in clinical practice to improve existing therapeutic strategies. Assessment of the prognostic power of the combined use of pftetin and the other 24 proteins as well as more extensive validation of pftetin using additional samples are now under way in our laboratory.

## Acknowledgments

We thank Kiyooki Nomoto, Chizu Kina, and Sachiko Miura for their excellent technical support in the immunohistochemical study, and Kano Nishiyama and Yukiko Fujie in electrophoresis.

## References

- Joensuu H, Kindblom LG. Gastrointestinal stromal tumors—a review. *Acta Orthop Scand Suppl* 2004;75: 62–71.
- Miettinen M, El-Rifai W, HL Sobin L, Lasota J. Evaluation of malignancy and prognosis of gastrointestinal stromal tumors: a review. *Hum Pathol* 2002;33: 478–83.
- Corless CL, Fletcher JA, Heinrich MC. Biology of gastrointestinal stromal tumors. *J Clin Oncol* 2004; 22:3813–25.
- Heinrich MC, Rubin BP, Longley BJ, Fletcher JA. Biology and genetic aspects of gastrointestinal stromal tumors: KIT activation and cytogenetic alterations. *Hum Pathol* 2002;33:484–95.
- Miettinen M, Majidi M, Lasota J. Pathology and diagnostic criteria of gastrointestinal stromal tumors (GISTs): a review. *Eur J Cancer* 2002;38 Suppl 5: S39–51.
- Demetri GD, von Mehren M, Blanke CD, et al. Efficacy and safety of imatinib mesylate in advanced gastrointestinal stromal tumors. *N Engl J Med* 2002;347: 472–80.
- van Oosterom AT, Judson I, Verweij J, et al. Safety and efficacy of imatinib (STI571) in metastatic gastrointestinal stromal tumours: a phase I study. *Lancet* 2001;358:1421–3.
- Heinrich MC, Corless CL, Blanke CD, et al. Molecular correlates of imatinib resistance in gastrointestinal stromal tumors. *J Clin Oncol* 2006;24:4764–74.
- Debiec-Rychter M, Sciot R, Le Cesne A, et al. KIT mutations and dose selection for imatinib in patients with advanced gastrointestinal stromal tumours. *Eur J Cancer* 2006;42:1093–103.
- Debiec-Rychter M, Dumez H, Judson I, et al. Use of c-KIT/PDGFRA mutational analysis to predict the clinical response to imatinib in patients with advanced gastrointestinal stromal tumours entered on phase I and II studies of the EORTC Soft Tissue and Bone Sarcoma Group. *Eur J Cancer* 2004;40: 689–95.
- Schneider-Stock R, Boltze C, Lasota J, et al. Loss of p16 protein defines high-risk patients with gastrointestinal stromal tumors: a tissue microarray study. *Clin Cancer Res* 2005;11:638–45.
- Schneider-Stock R, Boltze C, Lasota J, et al. High prognostic value of p16INK4 alterations in gastrointestinal stromal tumors. *J Clin Oncol* 2003;21: 1688–97.
- Rumessen JJ, Peters S, Thuneberg L. Light- and electron microscopic studies of interstitial cells of Cajal and muscle cells at the submucosal border of human colon. *Lab Invest* 1993;68:481–95.
- Tornillo L, Duchini G, Carafa V, et al. Patterns of gene amplification in gastrointestinal stromal tumors (GIST). *Lab Invest* 2005;85:921–31.
- El-Rifai W, Sarlomo-Rikala M, Andersson LC, Knuutila S, Miettinen M. DNA sequence copy number changes in gastrointestinal stromal tumors: tumor progression and prognostic significance. *Cancer Res* 2000;60:3899–903.
- Kang HJ, Nam SW, Kim H, et al. Correlation of KIT and platelet-derived growth factor receptor  $\alpha$  mutations with gene activation and expression profiles in gastrointestinal stromal tumors. *Oncogene* 2005;24: 1066–74.
- Koon N, Schneider-Stock R, Sarlomo-Rikala M, et al. Molecular targets for tumour progression in gastrointestinal stromal tumours. *Gut* 2004;53:235–40.
- Antonescu CR, Viale A, Sarran L, et al. Gene expression in gastrointestinal stromal tumors is distinguished by KIT genotype and anatomic site. *Clin Cancer Res* 2004;10:3282–90.
- Petricoin EF, Ardekani AM, Hitt BA, et al. Use of proteomic patterns in serum to identify ovarian cancer. *Lancet* 2002;359:572–7.
- Suehara Y, Kondo T, Fujii K, et al. Proteomic signatures corresponding to histological classification and grading of soft-tissue sarcomas. *Proteomics* 2006;6: 4402–9.
- Chen G, Gharib TG, Wang H, et al. Protein profiles associated with survival in lung adenocarcinoma. *Proc Natl Acad Sci U S A* 2003;100:13537–42.
- Okano T, Kondo T, Fujii K, et al. Proteomic signature corresponding to the response to gefitinib (Iressa, ZD1839), an epidermal growth factor receptor (EGFR) tyrosine kinase inhibitor, and mutation in EGFR in lung adenocarcinoma. *Clin Cancer Res* 2007;13:799–805.
- Orntoft TF, Thykjaer T, Waldman FM, Wolf H, Celis JE. Genome-wide study of gene copy numbers, transcripts, and protein levels in pairs of non-invasive and invasive human transitional cell carcinomas. *Mol Cell Proteomics* 2002;1:37–45.
- Varambally S, Yu J, Laxman B, et al. Integrative genomic and proteomic analysis of prostate cancer reveals signatures of metastatic progression. *Cancer Cell* 2005;8:393–406.
- Resendes BL, Kuo SF, Robertson NG, et al. Isolation from cochlea of a novel human intronless gene with predominant fetal expression. *J Assoc Res Otolaryngol* 2004;5:185–202.
- Stankey RH LA. Pathology and Genetics of Tumours of the Digestive System. Lyon: IARC Press; 2000.
- Hasegawa T, Matsuno Y, Shimoda T, Hirohashi S. Gastrointestinal stromal tumor: consistent CD117 immunostaining for diagnosis, and prognostic classification based on tumor size and MIB-1 grade. *Hum Pathol* 2002;33:669–76.
- Unlu M, Morgan ME, Minden JS. Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 1997;18:2071–7.
- Yokoo H, Kondo T, Fujii K, Yamada T, Todo S, Hirohashi S. Proteomic signature corresponding to  $\alpha$  fetoprotein expression in liver cancer cells. *Hepatology* 2004;40:609–17.
- Okano T, Kondo T, Kakisaka T, et al. Plasma proteomics of lung cancer by a linkage of multi-dimensional liquid chromatography and two-dimensional difference gel electrophoresis (2D-DIGE). *Proteomics* 2006;6:3938–48.
- Yoshida Y, Shibata T, Kokubu A, et al. Mutations of the epidermal growth factor receptor gene in atypical adenomatous hyperplasia and bronchioloalveolar carcinoma of the lung. *Lung Cancer* 2005;50:1–8.
- Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc* 1958;53: 457–81.
- Kang HJ, Koh KH, Yang E, et al. Differentially expressed proteins in gastrointestinal stromal tumors with KIT and PDGFRA mutations. *Proteomics* 2006; 6:1151–7.
- Molina H, Horn DM, Tang N, Mathivanan S, Pandey A. Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc Natl Acad Sci U S A* 2007;104:2199–204.
- Matsuoka S, Ballif BA, Smogorzewska A, et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* 2007;316:1160–6.
- Lasota J, Jasinski M, Sarlomo-Rikala M, Miettinen M. Mutations in exon 11 of c-Kit occur preferentially in malignant versus benign gastrointestinal stromal tumors and do not occur in leiomyomas or leiomyosarcomas. *Am J Pathol* 1999;154:53–60.
- Ernst SI, Hubbs AE, Przygodzki RM, Emory TS, Sobin LH, O'Leary TJ. KIT mutation portends poor

- prognosis in gastrointestinal stromal/smooth muscle tumors. *Lab Invest* 1998;78:1633–6.
38. Hirota S, Isozaki K, Nishida T, Kitamura Y. Effects of loss-of-function and gain-of-function mutations of c-kit on the gastrointestinal tract. *J Gastroenterol* 2000;35 Suppl 12:75–9.
39. Wardelmann E, Losen I, Hans V, et al. Deletion of Trp-557 and Lys-558 in the juxtamembrane domain of the c-kit protooncogene is associated with metastatic behavior of gastrointestinal stromal tumors. *Int J Cancer* 2003;106:887–95.
40. Lasota J, Dansonka-Mieszkowska A, Stachura T, et al. Gastrointestinal stromal tumors with internal tandem duplications in 3' end of KIT juxtamembrane domain occur predominantly in stomach and generally seem to have a favorable course. *Mod Pathol* 2003;16:1257–64.
41. Antonescu CR, Sommer G, Sarraf L, et al. Association of KIT exon 9 mutations with nongastric primary site and aggressive behavior: KIT mutation analysis and clinical correlates of 120 gastrointestinal stromal tumors. *Clin Cancer Res* 2003;9:3329–37.
42. Montgomery E, Abraham SC, Fisher C, et al. CD44 loss in gastric stromal tumors as a prognostic marker. *Am J Surg Pathol* 2004;28:168–77.
43. Nishitani A, Hirota S, Nishida T, et al. Differential expression of connexin 43 in gastrointestinal stromal tumours of gastric and small intestinal origin. *J Pathol* 2005;206:377–82.
44. Klose J, Nock C, Herrmann M, et al. Genetic analysis of the mouse brain proteome. *Nat Genet* 2002;30:385–93.
45. Kondo T, Hirohashi S. Application of highly sensitive fluorescent dyes (CyDye DIGE Fluor saturation dyes) to laser microdissection and two-dimensional difference gel electrophoresis (2D-DIGE) for cancer proteomics. *Nat Protoc* 2006;1:2940–56.