Table 1. Antitumor effects of imatinib and irinotecan in TMK-1 gastric carcinoma cells

Treatment group	Body weight (g)	Tumor incidence	Tumor weight (g)	Lymph node metastasis	Peritoneal metastasis	Liver metastasis
Control	21.8 (17.5–25.2)	10/10	3.13 (2.40–3.89)	8/10	6/10	1/10
Imatinib (50 mg/kg)	20.6 (17.2–22.7)	10/10	3.34 (1.92–4.50)	7/10	3/10	0/10
Imatinib (200 mg/kg)	22.0 (19.5–24.3)	9/9	2.26 (0.68–5.50)	6/9	3/9	2/9
Irinotecan (5 mg/kg)	21.3 (19.1–25.2)	10/10	1.33 (0.66–2.16)**	5/10	5/10	2/10
Irinotecan (5 mg/kg) + imatinib (50 mg/kg)	19.0 (16.1–21.4)	9/9	1.11 (0.53–1.80)**	7/9	3/9	1/9
Irinotecan (5 mg/kg) + imatinib (200 mg/kg)	21.8 (19.1–24.5)	9/9	0.59 (0.32–1.01)**	2/9*	1/9*	0/9

TMK-1 cells (1×10^6) were injected into the gastric wall of nude mice. Two weeks later, the mice were randomly assigned to receive water daily by oral gavage and intraperitoneal injection of PBS once per week (control); oral gavage of imatinib (50 mg/kg or 200 mg/kg) daily and weekly intraperitoneal injection of PBS; irinotecan intraperitoneally once per week and daily administration of water by oral gavage; or oral gavage of imatinib (50 mg/kg or 200 mg/kg) daily and weekly intraperitoneal injection of irinotecan. Treatments continued for 28 days. Number of individual mice is shown unless otherwise indicated. Ranges are shown in parentheses.
*$p < 0.05$ *vs.* control group, values. **$p < 0.01$ *vs.* control group, values.

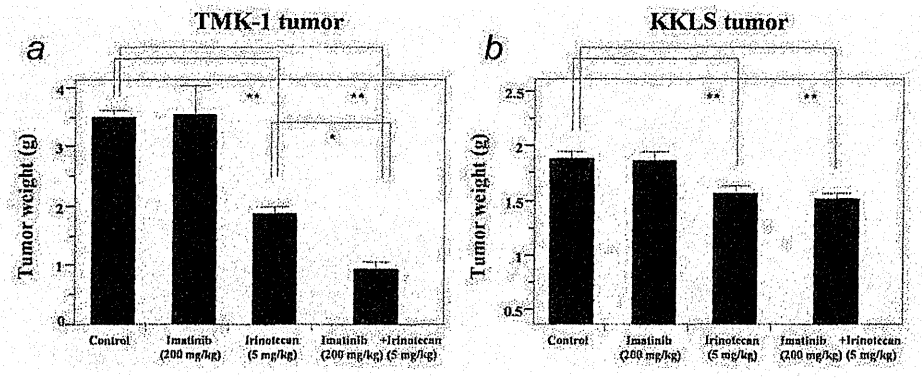

Figure 5. Antitumor effects of imatinib and irinotecan on the growth of TMK-1 and KKLS orthotopic tumors. In TMK-1 tumors, tumor growth was significantly inhibited in mice treated with irinotecan alone and imatinib in combination with irinotecan enhanced the antitumor effects of irinotecan (*a*). In contrast, imatinib did not enhance the antitumor effects of irinotecan in KKLS tumors (*b*). *, $p < 0.01$, **, $p < 0.001$; bars, SE.

Rβ was expressed mainly by stromal cells, including CAFs, pericytes and lymphatic endothelial cells. Blockade of PDGF-R signaling decreased stromal reaction, areas of vascular and lymphatic vessels and number of pericytes.

Kim *et al.* reported that inhibition of PDGF-R by imatinib enhanced the effect of chemotherapeutic reagents; however, they used an ectopic (subcutaneous) model of gastric carcinoma.[37] The organ microenvironment is known to influence the growth, vascularization, invasion and, hence, metastasis of human neoplasms.[38] Because ectopic tumors generally do not metastasize, the metastatic ability of tumor cells cannot be evaluated in an ectopic transplantation model. In addition, we reported previously that the expression of PDGF-Rβ by stromal cells in a colon cancer model was higher in tumors growing at the orthotopic site than in those growing at the ectopic site.[39] In the orthotopic site (cecal wall), the tumors induced a strong stromal reaction, whereas in the subcutis, the stromal reaction was minimal.[39] Treatment with imatinib and irinotecan significantly inhibited tumor growth in the cecum but not in the subcutis.[25] Thus, orthotopic transplantation models should be used in experimental studies to evaluate the inhibitory effect of therapeutic agents on metastasis.[38]

Recent studies have shown that tumor growth is determined not only by malignant cancer cells themselves but also by the tumor stroma.[40] Tumor stroma consists of CAFs, smooth muscle cells, inflammatory cells, microvessels and abundant ECM.[41] CAFs synthesize many of the constituents of fibrillar ECM, such as Type-I, Type-III and Type-V collagen and fibronectin.[42,43] CAFs are known to modulate tumorigenic properties of neoplastic cells, including proliferation, apoptosis and angiogenesis,[44] and they are thought to play a central role in the complex process of tumor-stroma interaction and consequent-tumorigenesis.[36] Experiments involving coinjection of CAFs and tumor cells have shown that, in contrast to normal fibroblasts, CAFs promote tumor growth.[45] Moreover, recent studies have revealed extensive

Table 2. Immunohistochemical analysis of TMK-1 human gastric carcinoma cells

Treatment group	Tumor cells		Tumor-associated endothelial cells		Lymphatic endothelial cells
	Ki-67 LI[1] (%)	TUNEL[2] (%)	MVA[3] (/× 10^3 μm^2)	Pericyte coverage[4] (%)	LVA[5] (/× 10^3 μm^2)
Control	57.3 ± 15.0	5.54 ± 1.64	38.6 ± 4.75	56.8 ± 4.65	27.8 ± 5.28
Imatinib (50 mg/kg)	50.7 ± 9.97	5.96 ± 2.45	22.2 ± 4.55**	31.5 ± 3.34*	27.0 ± 2.73
Imatinib (200 mg/kg)	41.7 ± 6.74*	4.74 ± 0.55	27.7 ± 5.34*	22.1 ± 6.16*	14.9 ± 1.51*
Irinotecan (5 mg/kg)	37.7 ± 21.1*	4.54 ± 0.36	40.8 ± 12.6	63.7 ± 10.5	21.5 ± 1.83
Irinotecan (5 mg/kg) + imatinib (50 mg/kg)	39.7 ± 7.85**	5.19 ± 1.22	26.6 ± 0.87*	45.5 ± 4.44*	20.8 ± 1.64
Irinotecan (5 mg/kg) + imatinib (200 mg/kg)	28.1 ± 14.3**	6.23 ± 3.12	18.8 ± 3.36**	10.1 ± 2.04*	16.1 ± 2.43*

[1]Number (mean ± SE) of Ki-67-positive tumor cells/field determined by measuring 10 random 0.81-mm^2 fields at 100× magnification. [2]Percentage of TUNEL-positive tumor cells (out of total number of cells) per 0.81-mm^2 fields at 100× magnification. [3]MVA was determined by measuring 10 random fields at 100× magnification. [4]LVA was determined by measuring 10 random fields at 100× magnification. [5]CD31-positive cells in direct contact with desmin-positive cells were counted in five random fields at 100× magnification. *$p < 0.05$ *vs.* control group, values. **$p < 0.01$ *vs.* control group, values.
Abbreviations: Ki-67 LI: Ki-67 labeling index; MVA: microvessel area; LVA: lymphatic vessel area.

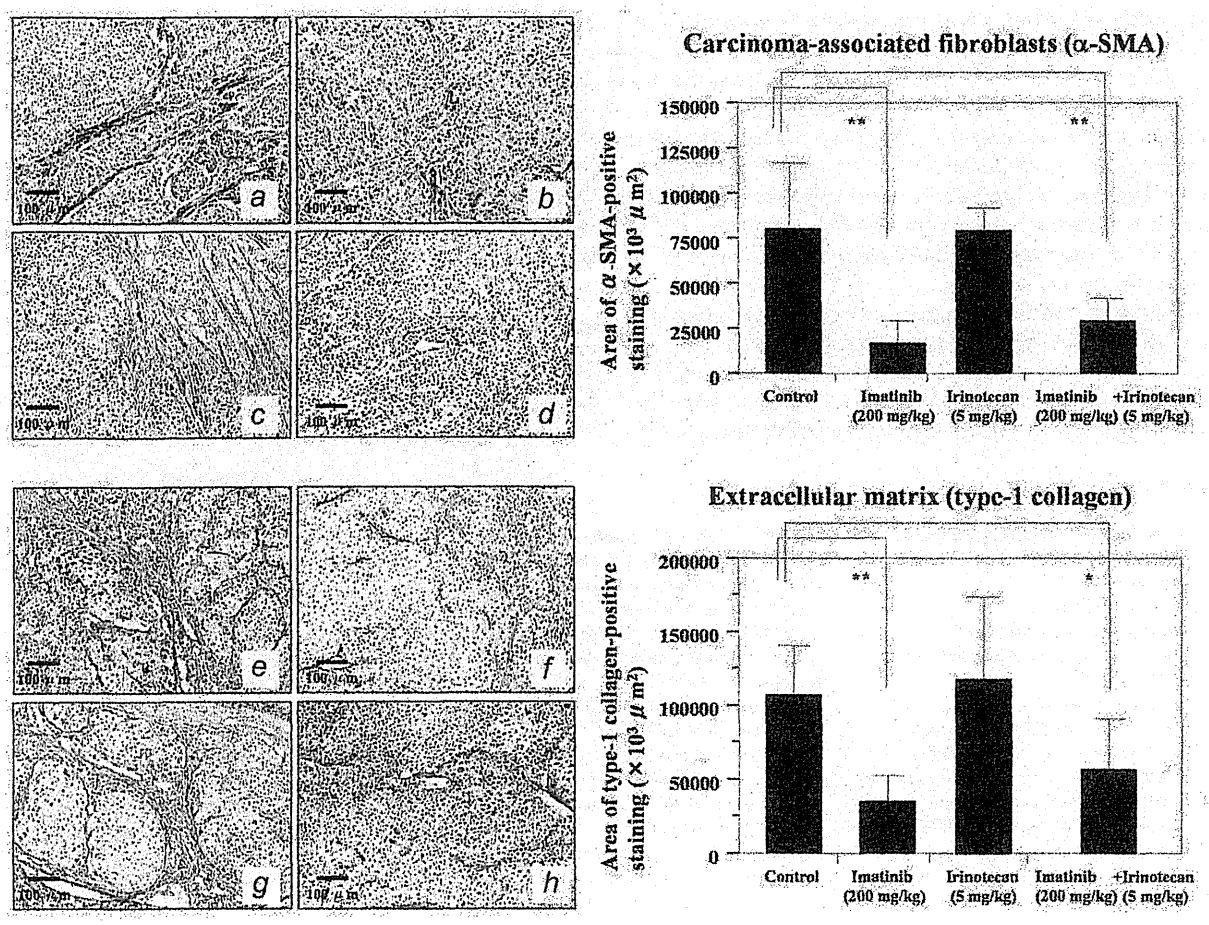

Figure 6. CAFs and ECM were quantified according to the α-SMA-positive (*a–d*) and Type-1 collagen-positive (*e–h*) areas. The areas of CAFs and ECM were reduced with treatment with imatinib alone and imatinib in combination with irinotecan. (*a* and *e*): control; (*b* and *f*): imatinib (200 mg/kg); (*c* and *g*): irinotecan (5 mg/kg); (*d* and *h*): imatinib (200 mg/kg) plus irinotecan (5 mg/kg). *, $p < 0.05$, **, $p < 0.01$; bars, SE. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

changes in the phenotype, and even the genotype, of CAFs compared with their normal counterparts.[46] Lieubeau *et al.* reported that progressive tumor growth correlates with proliferation of myofibroblasts, whereas regression of tumors is linked to the fibrous capsule, suggesting that the presence of reactive stromal cells (myofibroblasts) may contribute to the growth of tumor cells.[47] Other studies have shown that a poor prognosis in cases of colorectal carcinoma is associated with abundance of CAFs or increased expression of fibroblast-activated protein.[48,49] It has also been shown that CAFs can affect sensitivity of pancreatic carcinoma cells to chemo- or radio-therapy; such tumor cells become less sensitive to chemotherapy when co-cultured with CAFs or grown in fibroblast-conditioned medium.[50] Therefore, CAFs are considered novel targets for anti-tumor therapies.

Increased IFP in solid tumors, which is mediated by an increase in stromal compartment pressure, acts as a barrier between tumor cells and normal tissues against effective distribution of anti-tumor drugs. It has also been reported that increased IFP prevents tumor transvascular transport of effector cells.[51] The PDGF/PDGF-R signaling pathway plays a critical role in the control of IFP.[24,51] A number of studies have shown that inhibition of PDGF-R signaling can decrease this pressure and, hence, enhance the effects chemotherapeutic reagents.[14] In our TMK-1 model, we observed that inhibition of tumor growth and metastasis by combination imatinib and irinotecan treatment was associated with the reduction in CAFs and ECM, suggesting that similar mechanisms may be involved in gastric carcinoma.

Treatment with imatinib influences microvessel structure in tumor tissues. Pericytes in tumor-associated vessels, but not those in vessels of normal mucosa, are enlarged and they overexpress PDGF-Rβ and p-PDGF-Rβ.[39] In agreement with earlier reports,[13,31] treatment with imatinib decreased pericyte coverage on tumor-associated endothelial cells in our experimental model. The inhibition of PDGF-R signaling may decrease pericyte recruitment and attachment to endothelial cells and destabilize the tumor vasculature.

VEGF-C and -D are well-known lymphangiogenic factors in human gastric carcinoma. Other growth factors recently reported to be lymphangiogenic are FGF-2,[52] PDGF-BB[53] and angiopoietin-2.[54] In this study, we showed that lymphatic endothelial cells express PDGF-R and that inhibition of PDGFR results in a reduction of the lymphatic vessel area. We recently found correlation between the level of PDGF-B mRNA expression in gastric carcinoma tissue and lymphatic metastasis.[55] These findings suggest that PDGF-B acts as a lymphangiogenic factor. However, it is a possible that treatment with imatinib exerts indirect anti-lymphangiogenic effects *via* inhibition of VEGF-C and -D production by tumor cells and/or stromal cells. In addition, imatinib at a high concentration can block various tyrosine kinases to different extents. Molecules other than PDGF-Rs may be involved in the regulation of lymphangiogenesis. Further studies are needed to confirm whether the PDGF/PDGF-R axis is directly involved in lymphangiogenesis in gastric carcinoma.

In summary, targeting stromal cells by PDGF-R inhibitor appears promising as a therapeutic strategy against carcinoma. Stromal compartment-rich tumors, such as diffuse-type gastric carcinomas, may be tumors in which targeting the PDGF/PDGF-R signaling pathway for enhancement of the chemotherapeutic effect is most applicable.

Acknowledgements

This work was carried out with the kind cooperation of the Analysis Center of Life Science and Institute of Laboratory Animal Science, Hiroshima University, Hiroshima, Japan, and we thank Novartis Pharma K.K. (Basal, Switzerland) for providing the imatinib.

References

1. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. *Nature* 2008;454:436–44.
2. Whiteside TL. The tumor microenvironment and its role in promoting tumor growth. *Oncogene* 2008; 27:5904–12.
3. Mueller MM, Fusenig NE. Friends or foes—bipolar effects of the tumour stroma in cancer. *Nat Rev Cancer* 2004;4:839–49.
4. Kitadai Y. Cancer-stromal cell interaction and tumor angiogenesis in gastric cancer. *Cancer Microenviron*, 2010;3:109–16.
5. Uehara H, Kim SJ, Karashima T, Shepherd DL, Fan D, Tsan R, Killion JJ, Logothetis C, Mathew P, Fidler IJ. Effects of blocking platelet-derived growth factor-receptor signaling in a mouse model of experimental prostate cancer bone metastases. *J Natl Cancer Inst* 2003;95: 458–70.
6. Antoniades HN, Galanopoulos T, Neville-Golden J, O'Hara CJ. Malignant epithelial cells in primary human lung carcinomas coexpress in vivo platelet-derived growth factor (PDGF) and PDGF receptor mRNAs and their protein products. *Proc Natl Acad Sci USA* 1992;89:3942–6.
7. Lindmark G, Sundberg C, Glimelius B, Pahlman L, Rubin K, Gerdin B. Stromal expression of platelet-derived growth factor beta-receptor and platelet-derived growth factor B-chain in colorectal cancer. *Lab Invest* 1993;69:682–9.
8. Seymour L, Dajee D, Bezwoda WR. Tissue platelet derived-growth factor (PDGF) predicts for shortened survival and treatment failure in advanced breast cancer. *Breast Cancer Res Treat* 1993;26: 247–52.
9. Yi B, Williams PJ, Niewolna M, Wang Y, Yoneda T. Tumor-derived platelet-derived growth factor-BB plays a critical role in osteosclerotic bone metastasis in an animal model of human breast cancer. *Cancer Res* 2002;62:917–23.
10. Bergsten E, Uutela M, Li X, Pietras K, Ostman A, Heldin CH, Alitalo K, Eriksson U. PDGF-D is a specific, protease-activated ligand for the PDGF beta-receptor. *Nat Cell Biol* 2001;3:512–16.
11. Ostman A. PDGF receptors-mediators of autocrine tumor growth and regulators of tumor vasculature and stroma. *Cytokine Growth Factor Rev* 2004;15: 275–86.
12. Risau W, Drexler H, Mironov V, Smits A, Siegbahn A, Funa K, Heldin CH. Platelet-derived growth factor is angiogenic in vivo. *Growth Factors* 1992;7:261–6.
13. Bergers G, Song S, Meyer-Morse N, Bergsland E, Hanahan D. Benefits of targeting both pericytes and endothelial

cells in the tumor vasculature with kinase inhibitors. *J Clin Invest* 2003;111:1287–95.
14. Pietras K. Increasing tumor uptake of anticancer drugs with imatinib. *Semin Oncol* 2004;31 (Suppl 6):18–23.
15. Druker BJ, Lydon NB. Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *J Clin Invest* 2000;105:3–7.
16. Fletcher JA. Role of KIT and platelet-derived growth factor receptors as oncoproteins. *Semin Oncol* 2004;31 (Suppl 6):4–11.
17. Day E, Waters B, Spiegel K, Alnadaf T, Manley PW, Buchdunger E, Walker C, Jarai G. Inhibition of collagen-induced discoidin domain receptor 1 and 2 activation by imatinib, nilotinib and dasatinib. *Eur J Pharmacol* 2008;599:44–53.
18. Heinrich MC, Blanke CD, Druker BJ, Corless CL. Inhibition of KIT tyrosine kinase activity: a novel molecular approach to the treatment of KIT-positive malignancies. *J Clin Oncol* 2002;20: 1692–703.
19. Beppu K, Jaboine J, Merchant MS, Mackall CL, Thiele CJ. Effect of imatinib mesylate on neuroblastoma tumorigenesis and vascular endothelial growth factor expression. *J Natl Cancer Inst* 2004;96: 46–55.
20. Kilic T, Alberta JA, Zdunek PR, Acar M, Iannarelli P, O'Reilly T, Buchdunger E, Black PM, Stiles CD. Intracranial inhibition of platelet-derived growth factor-mediated glioblastoma cell growth by an orally active kinase inhibitor of the 2-phenylaminopyrimidine class. *Cancer Res* 2000;60:5143–50.
21. Krystal GW, Honsawek S, Litz J, Buchdunger E. The selective tyrosine kinase inhibitor STI571 inhibits small cell lung cancer growth. *Clin Cancer Res* 2000; 6:3319–26.
22. Merchant MS, Woo CW, Mackall CL, Thiele CJ. Potential use of imatinib in Ewing's sarcoma: evidence for in vitro and in vivo activity. *J Natl Cancer Inst* 2002;94: 1673–9.
23. Sjöblom T, Shimizu A, O'Brien KP, Pietras K, Dal Cin P, Buchdunger E, Dumanski JP, Ostman A, Heldin CH. Growth inhibition of dermatofibrosarcoma protuberans tumors by the platelet-derived growth factor receptor antagonist STI571 through induction of apoptosis. *Cancer Res* 2001;61: 5778–83.
24. Hwang RF, Yokoi K, Bucana CD, Tsan R, Killion JJ, Evans DB, Fidler IJ. Inhibition of platelet-derived growth factor receptor phosphorylation by STI571 (Gleevec) reduces growth and metastasis of human pancreatic carcinoma in an orthotopic nude mouse model. *Clin Cancer Res* 2003; 9:6534–44.
25. Kitadai Y, Sasaki T, Kuwai T, Nakamura T, Bucana CD, Fidler IJ. Targeting the expression of platelet-derived growth factor receptor by reactive stroma inhibits growth and metastasis of human colon carcinoma. *Am J Pathol* 2006;169:2054–65.
26. Pietras K, Pahler J, Bergers G, Hanahan D. Functions of paracrine PDGF signaling in the proangiogenic tumor stroma revealed by pharmacological targeting. *PLoS Med* 2008;5:123–38.
27. Farhat FS. A general review of the role of irinotecan (CPT11) in the treatment of gastric cancer. *Med Oncol* 2007;24:137–46.
28. Apte SM, Fan D, Killion JJ, Fidler IJ. Targeting the platelet-derived growth factor receptor in antivascular therapy for human ovarian carcinoma. *Clin Cancer Res* 2004; 10:897–908.
29. Kobie K, Kawabata M, Hioki K, Tanaka A, Matsuda H, Mori T, Maruo K. The tyrosine kinase inhibitor imatinib [STI571] induces regression of xenografted canine mast cell tumors in SCID mice. *Res Vet Sci* 2007;82:239–41.
30. Wolff NC, Veach DR, Tong WP, Bornmann WG, Clarkson B, Ilaria RL, Jr. PD166326, a novel tyrosine kinase inhibitor, has greater antileukemic activity than imatinib mesylate in a murine model of chronic myeloid leukemia. *Blood* 2005; 105:3995–4003.
31. Yokoi K, Sasaki T, Bucana CD, Fan D, Baker CH, Kitadai Y, Kuwai T, Abbruzzese JL, Fidler IJ. Simultaneous inhibition of EGFR, VEGFR, and platelet-derived growth factor receptor signaling combined with gemcitabine produces therapy of human pancreatic carcinoma and prolongs survival in an orthotopic nude mouse model. *Cancer Res* 2005;65:10371–80.
32. Peng B, Hayes M, Resta D, Racine-Poon A, Druker BJ, Talpaz M, Sawyers CL, Rosamilia M, Ford J, Lloyd P, Capdeville R. Pharmacokinetics and pharmacodynamics of imatinib in a phase I trial with chronic myeloid leukemia patients. *J Clin Oncol* 2004;22:935–42.
33. Hermanson M, Funa K, Hartman M, Claesson-Welsh L, Heldin CH, Westermark B, Nistér M. Platelet-derived growth factor and its receptors in human glioma tissue: expression of messenger RNA and protein suggests the presence of autocrine and paracrine loops. *Cancer Res* 1992;52: 3213–19.
34. Smits A, Funa K, Vassbotn FS, Beausang-Linder M, af Ekenstam F, Heldin CH, Westermark B, Nister M. Expression of platelet-derived growth factor and its receptors in proliferative disorders of fibroblastic origin. *Am J Pathol* 1992;140: 639–48.
35. Forsberg K, Valyi-Nagy I, Heldin CH, Herlyn M, Westermark B. Platelet-derived growth factor (PDGF) in oncogenesis: development of a vascular connective tissue stroma in xenotransplanted human melanoma producing PDGF-BB. *Proc Natl Acad Sci USA* 1993;90:393–7.
36. Micke P, Ostman A. Tumour-stroma interaction: cancer-associated fibroblasts as novel targets in anti-cancer therapy? *Lung Cancer* 2004;45 (Suppl 2): S163–S175.
37. Kim R, Emi M, Arihiro K, Tanabe K, Uchida Y, Toge T. Chemosensitization by STI571 targeting the platelet-derived growth factor/platelet-derived growth factor receptor-signaling pathway in the tumor progression and angiogenesis of gastric carcinoma. *Cancer* 2005;103:1800–9.
38. Fidler IJ, Wilmanns C, Staroselsky A, Radinsky R, Dong Z, Fan D. Modulation of tumor cell response to chemotherapy by the organ environment. *Cancer Metastasis Rev* 1994;13:209–22.
39. Kitadai Y, Sasaki T, Kuwai T, Nakamura T, Bucana CD, Hamilton SR, Fidler IJ. Expression of activated platelet-derived growth factor receptor in stromal cells of human colon carcinomas is associated with metastatic potential. *Int J Cancer* 2006;119: 2567–74.
40. Kalluri R. Basement membranes: structure, assembly and role in tumour angiogenesis. *Nat Rev Cancer* 2003;3:422–33.
41. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70.
42. Rodemann HP, Muller GA. Characterization of human renal fibroblasts in health and disease. II. In vitro growth, differentiation, and collagen synthesis of fibroblasts from kidneys with interstitial fibrosis. *Am J Kidney Dis* 1991;17:684–6.
43. Tomasek JJ, Gabbiani G, Hinz B, Chaponnier C, Brown RA. Myofibroblasts and mechano-regulation of connective tissue remodelling. *Nat Rev Mol Cell Biol* 2002;3:349–63.
44. Micke P, Ostman A. Exploring the tumour environment: cancer-associated fibroblasts as targets in cancer therapy. *Expert Opin Ther Targets* 2005;9:1217–33.
45. Tuxhorn JA, Ayala GE, Rowley DR. Reactive stroma in prostate cancer progression. *J Urol* 2001;166:2472–83.
46. Moinfar F, Man YG, Arnould L, Bratthauer GL, Ratschek M, Tavassoli FA. Concurrent and independent genetic alterations in the stromal and epithelial cells of mammary carcinoma: implications for tumorigenesis. *Cancer Res* 2000;60: 2562–6.
47. Lieubeau B, Garrigue L, Barbieux I, Meflah K, Gregoire M. The role of transforming growth factor beta 1 in the fibroblastic reaction associated with rat colorectal tumor development. *Cancer Res* 1994;54: 6526–32.

48. Henry LR, Lee HO, Lee JS, Klein-Szanto A, Watts P, Ross EA, Chen WT, Cheng JD. Clinical implications of fibroblast activation protein in patients with colon cancer. *Clin Cancer Res* 2007; 13:1736–41.
49. Tsujino T, Seshimo I, Yamamoto H, Ngan CY, Ezumi K, Takemasa I, Ikeda M, Sekimoto M, Matsuura N, Monden M. Stromal myofibroblasts predict disease recurrence for colorectal cancer. *Clin Cancer Res* 2007;13:2082–90.
50. Hwang RF, Moore T, Arumugam T, Ramachandran V, Amos KD, Rivera A, Ji B, Evans DB, Logsdon CD. Cancer-associated stromal fibroblasts promote pancreatic tumor progression. *Cancer Res* 2008;68:918–26.
51. Pietras K, Ostman A, Sjoquist M, Buchdunger E, Reed RK, Heldin CH, Rubin K. Inhibition of platelet-derived growth factor receptors reduces interstitial hypertension and increases transcapillary transport in tumors. *Cancer Res* 2001;61: 2929–34.
52. Kubo H, Cao R, Brakenhielm E, Makinen T, Cao Y, Alitalo K. Blockade of vascular endothelial growth factor receptor-3 signaling inhibits fibroblast growth factor-2-induced lymphangiogenesis in mouse cornea. *Proc Natl Acad Sci USA* 2002;99: 8868–73.
53. Cao R, Bjorndahl MA, Religa P, Clasper S, Garvin S, Galter D, Meister B, Ikomi F, Tritsaris K, Dissing S, Ohhashi T, Jackson DG, et al. PDGF-BB induces intratumoral lymphangiogenesis and promotes lymphatic metastasis. *Cancer Cell* 2004;6:333–45.
54. Gale NW, Thurston G, Hackett SF, Renard R, Wang Q, McClain J, Martin C, Witte C, Witte MH, Jackson D, Suri C, Campochiaro PA, et al. Angiopoietin-2 is required for postnatal angiogenesis and lymphatic patterning, and only the latter role is rescued by angiopoietin-1. *Dev Cell* 2002;3:411–23.
55. Kodama M, Kitadai Y, Sumida T, Ohnishi M, Ohara E, Tanaka M, Shinagawa K, Tanaka S, Yasui W, Chayama K. Expression of platelet-derived growth factor (PDGF)-B and PDGF-receptor beta is associated with lymphatic metastasis in human gastric carcinoma. *Cancer Sci*, 2010: 101:1984–9.

GASTROENTEROLOGY 2011;140:892–902

A Functional Single Nucleotide Polymorphism in *Mucin 1*, at Chromosome 1q22, Determines Susceptibility to Diffuse-Type Gastric Cancer

NORIHISA SAEKI,* AKIRA SAITO,*,‡ IL JU CHOI,§ KEITARO MATSUO,‖ SUMIKO OHNAMI,* HIROHIKO TOTSUKA,*,¶ SUENORI CHIKU,*,# AYA KUCHIBA,* YEON-SU LEE,§ KYONG-AH YOON,§ MYEONG-CHERL KOOK,§ SOOK RYUN PARK,§ YOUNG-WOO KIM,§ HIDEO TANAKA,‖ KAZUO TAJIMA,‖ HIROSHI HIROSE,** FUMIHIKO TANIOKA,‡‡ YOSHIHIRO MATSUNO,§§ HARUHIKO SUGIMURA,‖‖ SHUNJI KATO,¶¶ TSUNEYA NAKAMURA,## TOMOHIRO NISHINA,*** WATARU YASUI,‡‡‡ KAZUHIKO AOYAGI,* HIROKI SASAKI,* KAZUYOSHI YANAGIHARA,§§§ HITOSHI KATAI,‖‖‖ TADAKAZU SHIMODA,¶¶¶ TERUHIKO YOSHIDA,* YUSUKE NAKAMURA,###,**** SETSUO HIROHASHI,**** and HIROMI SAKAMOTO*

**Genetics Division, National Cancer Center Research Institute, Tokyo; ‡Statistical Genetics Analysis Division, StaGen Co, Ltd, Tokyo, Japan; §Research Institute and Hospital, National Cancer Center, Gyeonggi-do, Korea; ‖Division of Epidemiology and Prevention, Aichi Cancer Center Research Institute, Nagoya; ¶Bioinformatics Group, Research and Development Center, Solution Division 4, Hitachi Government and Public Corporation System Engineering Ltd, Tokyo; #Science Solutions Division, Mizuho Information and Research Institute, Inc, Tokyo; **Department of Internal Medicine, Keio University School of Medicine, Tokyo; ‡‡Iwata City Hospital, Shizuoka; §§Department of Surgical Pathology, Hokkaido University Hospital, Sapporo; ‖‖First Department of Pathology, Hamamatsu University School of Medicine, Shizuoka; ¶¶Department of Surgery, Nippon Medical School Hospital, Tokyo; ##Department of Endoscopy, Aichi Cancer Center Hospital, Nagoya; ***Department of Internal Medicine, Shikoku Cancer Center, Ehime; ‡‡‡Department of Molecular Pathology, Hiroshima University Graduate School of Biomedical Sciences, Hiroshima; §§§Department of Life Sciences, Yasuda Women's University Faculty of Pharmacy, Hiroshima; ‖‖‖Department of Surgical Oncology, National Cancer Center Hospital, Tokyo; ¶¶¶Center for Cancer Control and Information Services, National Cancer Center, Tokyo; ###Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo; and ****National Cancer Center Research Institute, Tokyo, Japan*

BACKGROUND & AIMS: Two major types of gastric cancer, intestinal and diffuse, develop through distinct mechanisms; the diffuse type is considered to be more influenced by genetic factors, although the mechanism is unknown. Our previous genome-wide association study associated 3 single nucleotide polymorphisms (SNPs) with diffuse-type gastric cancer (DGC); 1 was a functional SNP (rs2294008) in *prostate stem cell antigen* (*PSCA*), but the loci of the other 2 were not investigated. **METHODS:** We performed high-density mapping to explore a linkage disequilibrium status of the 2 SNPs at chromosome 1q22. A DGC case-control study was conducted using DNA from 606 cases and 1264 controls (all Japanese individuals) and validated using DNA from Japanese (304 cases, 1465 controls) and Korean (452 cases, 372 controls) individuals. The effects of SNPs on function were analyzed by reporter assays and analyses of splice variants. **RESULTS:** A region of a strong linkage disequilibrium with the 2 SNPs contained *mucin 1* (*MUC1*) and other 4 genes and SNPs significantly associated with DGC (rs2070803: $P = 4.33 \times 10^{-13}$; odds ratio [OR], 1.71 by meta-analysis of the studies on the 3 panels) but not with intestinal-type gastric cancer. Functional studies demonstrated that rs4072037 ($P = 1.43 \times 10^{-11}$; OR, 1.66 by meta-analysis) in *MUC1* affects promoter activity and determines the major splicing variants of *MUC1* in the gastric epithelium. Individuals that carry both SNPs rs2294008 in *PSCA* and rs4072037 in *MUC1* have a high risk for developing DGC (OR, 8.38). **CONCLUSIONS: *MUC1* is the second major DGC susceptibility gene identified. The SNPs rs2070803 and rs4072037 in *MUC1* might be used to identify individuals at risk for this type of gastric cancer.**

Keywords: Stomach Cancer; Risk Genotype; Cancer Prevention; Genome-Wide Association Study.

Gastric cancer (GC) is the fourth most common cancer and the second most common cause of cancer death in the world.[1] More than 90% of GC are adenocarcinomas, which are classified into diffuse-type GC (DGC) and intestinal-type GC (IGC).[2] Typically, IGC arises through a sequence of pathologic changes of the gastric epithelium: chronic gastritis mainly because of *Helicobacter pylori* infection, atrophic gastritis, intestinal metaplasia, dysplasia, and adenocarcinoma.[3] On the other hand, the origin of DGC has been considered to be gastric epithelial stem cells and/or precursors present in the isthmus region of the middle portion of the epithelium (Supplementary Figure 1). Genetic and epigenetic events acting on the stem/precursor cells may cause a deviation from their normal differentiation program and lead to a DGC development,[4] although details are yet to be revealed. In contrast to the steady decline of the incidence of IGC, mainly because of the decreasing prevalence of *H pylori* infection, DGC appears to be increasing.[5] Moreover, some DGC develops to a highly malignant form,

Abbreviations used in this paper: **DGC, diffuse-type gastric cancer; GC, gastric cancer; GWAS, genome-wide association study; IGC, intestinal-type gastric cancer; kb, kilobase; LD, linkage disequilibrium; *MUC1*, mucin 1; OR, odds ratio; por, poorly differentiated adenocarcinoma; *PSCA*, *prostate stem cell antigen*; sig, signet-ring cell carcinoma; TR, tandem repeats.**

0016-5085/$36.00
doi:10.1053/j.gastro.2010.10.058

linitis plastica.[6] Identification of genetic predisposing factors and molecular pathways for the DGC development is one of the fundamentals for conceiving effective prevention, early diagnosis and therapeutic strategies.

Previously, we conducted a gene-centric genome-wide association study (GWAS) on DGC and identified 3 statistically significant single nucleotide polymorphisms (SNPs) at 2 loci after Bonferroni correction ($P < 1.8 \times 10^{-5}$) in the second stage of the screening.[7] Of the 3 SNPs, rs2976392 showed the lowest *P* value and tagged a linkage disequilibrium (LD) block at chromosome 8q24.3, in which we identified *prostate stem cell antigen* (*PSCA*) encoding prostate stem cell antigen as the novel DGC susceptibility gene. In the present study, we investigated the second genomic region of interest at chromosome 1q22, which harbors the remaining 2 SNPs, rs2075570 and rs2070803,[7] and identified *mucin 1* (*MUC1*) as the possible causal gene of the association of the region to DGC. The association between the gene and GC had been suggested also in previous reports.[8–11] However, unlike the previous candidate gene approach, we have reached the gene by performing a hypothesis-free GWAS followed by biologic studies in which a rationale of the association was obtained through the analyses of the function of a SNP rs4072037. Moreover, this study has a sufficient power as a systematic survey of genetic factors with the expected range of effect size and allele frequencies, generating a convincing level of statistical association ($P < 10^{-10}$ as compared with $P \sim 10^{-2}$ by the previous candidate gene approach[11]). The SNP rs4072037 is known to determine a splicing acceptor site in the second exon of *MUC1*.[12] In this study, we showed that the SNP is also related to major splicing variant selection in the stomach and has effect on the *MUC1* promoter activity, both of which may result in MUC1 functional difference between the individuals.

Materials and Methods

Samples

In Japan, the common type of GC is classified into 7 categories: papillary adenocarcinoma (pap), tubular adenocarcinoma (tub1 and tub2), poorly differentiated adenocarcinoma (por1 and por2), signet-ring cell carcinoma (sig), and mucinous adenocarcinoma (muc). However, a classification into 2 major categories by Lauren,[2] intestinal and diffuse types, is used worldwide especially for clinicoepidemiologic studies. A review of the classification systems is described elsewhere.[7] Basically, the DGC under the Lauren classification corresponds to por2 (nonsolid type) of poorly differentiated adenocarcinoma and sig by Japanese classification, although some investigators consider that por1 (solid type) is also included in DGC.[13]

Details of DNA samples used in the SNP typing and the association study of the chromosome 1q22 locus are as follows: In the Tokyo data set study, 610 DNA samples from patients with DGC (320 males; mean age, 55.4; 290 females; mean age, 54.0) were prepared either from methanol-fixed, paraffin-embedded tissues of noncancerous gastric mucosa or lymph nodes, or from peripheral blood, of patients with either the linitis plastica type of GC or early-stage cancer diagnosed as macroscopic type 0 IIc with histologic type of por2 and/or sig. The DGC samples in the Tokyo data set were collected at 4 institutions: the National Cancer Center Hospital in Tokyo: 360 paraffin-embedded tissues and 164 blood samples; Nippon Medical School Hospital in Tokyo: 76 blood samples; Aichi Cancer Center in Aichi: 1 blood sample; and Shikoku Cancer Center in Ehime: 9 blood samples. The control DNA samples were from peripheral blood leukocytes of 1266 volunteer individuals (male, 849; mean age, 67.2; female, 417; mean age, 59.8) with no known malignancies, who offered blood at a health check examination at Iwata City Hospital in Shizuoka and at Keio University campuses in Tokyo.

In the Aichi data set study, the DGC case samples were obtained from peripheral blood of 304 patients with histologic diagnosis of por1, por2, or sig (199 males; mean age, 57.3; 105 females; mean age, 56.4). Control blood samples were from 1467 volunteer individuals (1098 males; mean age, 59.8; 369 females; mean age, 57.3) with no known malignancies. Power calculations for the DGC analysis showed that the sample size of 304 cases and 1467 controls would provide the study with a power of over 98% for detecting an association of a SNP with a minor allele frequency of 0.2 or higher and per-allele odds ratio (OR) for risk allele of 1.63 or higher (estimates on rs2070803 obtained from Tokyo data set) in a 2-sided test at a significance level of .05.

In the Korea data set study, peripheral blood samples were donated from 455 patients with DGC who were diagnosed or treated at the National Cancer Center in Seoul, Korea (260 males; mean age, 52.4; 195 females; mean age, 48.5). The control subjects were 372 volunteers who participated in the National Cancer Screening Program at the National Cancer Center, Korea, and were confirmed by endoscopy not to have GC (191 males; mean age, 54.2; 181 females; mean age, 52.5). Power calculations showed that the sample size of the Korea data set would provide the study with a power of over 95% for detecting an association of rs2070803 at a significance level of .05 for the DGC study.

In the association studies (results shown in Figure 1, Table 1, and Supplementary Tables 1–4), 11 subjects (4 DGC and 2 controls from Tokyo data set, 2 controls from Aichi data set, 3 DGC cases from Korea data set) were excluded because of at least 1 missing covariate. Distributions of the covariates from subjects included in the studies are shown in Supplementary Figures 9–11.

Haplotype-based association study was performed on DNA samples from 380 DGC cases (200 males; mean age,

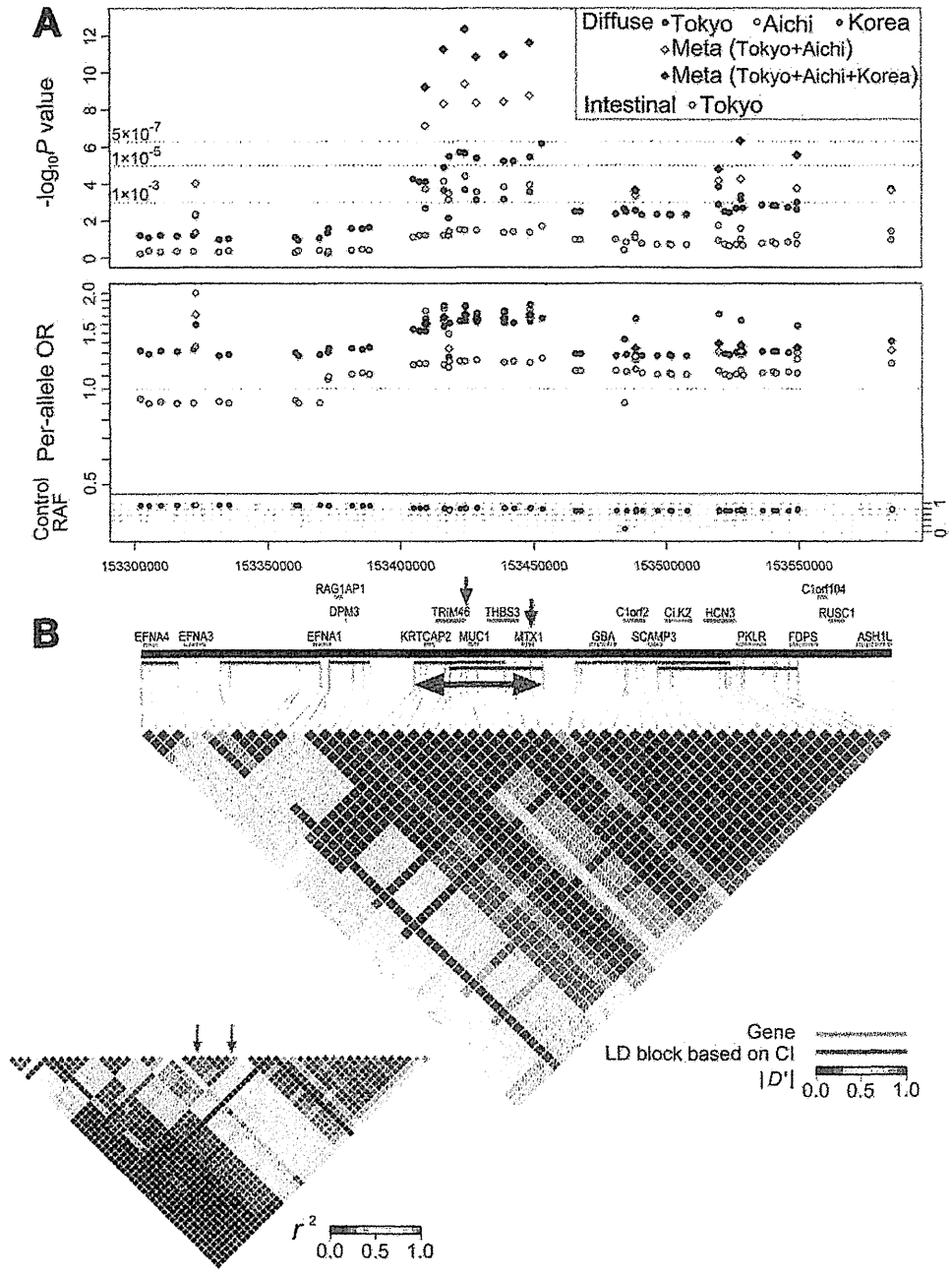

Figure 1. Association of the SNPs around rs2070803 and rs2075570 in chromosome 1q22 with GC and LD analyses of the SNPs. (*A*) The association study on DGC disclosed 8 SNPs with $P < 4 \times 10^{-5}$ within the LD block around rs2070803 and rs2075570 (*arrows* in *B*) in the Japanese population (Tokyo data set: *blue dots*). The association with DGC was replicated in another Japanese population (Aichi data set: *orange dots*) and also in the Korean population (Korea data set: *red dots*) for 4 selected SNPs: rs2070803, rs4072037, rs2066981, and rs2075570 ($P < 1 \times 10^{-3}$). Meta-analysis on these 3 data sets was also conducted (Tokyo and Aichi: *grey dots*; Tokyo, Aichi, and Korea: *black dots*). The study on IGC in the Japanese population (Tokyo data set: 599 cases, 1264 controls) showed no significant association of SNPs (*green dots*). *Upper panel* shows *P* value of each SNP in negative common logarithmic scale; *lower panel* shows OR and frequency of the risk allele (control RAF) of the SNPs. The position of the *dots* representing each SNP corresponds vertically to that in the physical map in *B*. (*B*) An LD analysis based on |*D'*| showed a strong LD around the 2 SNPs, rs2070803 and rs2075570 (*red arrows*), identified as DGC-associated SNPs in GWAS.[7] The strength of the LD is indicated by heat maps. *Blue bars* represent LD blocks defined by confidence intervals of |*D'*|.[15] Five genes—*KRTCAP2*, *TRIM46*, *MUC1*, *THBS3*, and *MTX1*—reside in the region of strong LD (*double-headed red arrow*) harboring rs2070803 and rs2075570. An LD map with r^2 is also shown in small scale. The analysis was performed with genotyping data of 1266 controls of the Tokyo data set.

Table 1. Results of Association Studies With 3 Independent Data Sets—Tokyo, Aichi, and Korea—for GC Susceptibility, on the 8 SNPs in the LD Block Containing rs2070803 and rs2075570 in Chromosome 1q22

SNPs, major/minor/risk[a]	Diffuse, Tokyo data set[b,c]	Diffuse, Aichi data set[c,d]	Diffuse, Korea data set[c,e]	Diffuse, Meta-analysis[f,g]	Diffuse, Meta-analysis[g,h]	Intestinal, Tokyo data set[c,i]
rs9426886 T/A/T	OR: 1.61 (1.32–1.97) *P*: 3.17×10^{-6} MAF (case): 0.133 MAF (control): 0.199					OR: 1.22 (1.01–1.46) *P*: 3.62×10^{-2} MAF (case): 0.169 MAF (control): 0.199
rs4971100 A/G/A	OR: 1.63 (1.33–1.99) *P*: 2.04×10^{-6} MAF (case): 0.133 MAF (control): 0.200					OR: 1.22 (1.02–1.47) *P*: 3.13×10^{-2} MAF (case): 0.169 MAF (control): 0.200
rs2070803[j] G/A/G	OR: 1.63 (1.33–1.99) *P*: 2.20×10^{-6} MAF (case): 0.133 MAF (control): 0.200	OR: 1.81 (1.36–2.40) *P*: 3.93×10^{-5} MAF (case): 0.104 MAF (control): 0.178	OR: 1.82 (1.32–2.49) *P*: 2.19×10^{-4} MAF (case): 0.103 MAF (control): 0.178	OR: 1.69 (1.43–1.99) *P*: 4.25×10^{-10}	OR: 1.71 (1.48–1.98) *P*: 4.33×10^{-13}	OR: 1.22 (1.02–1.47) *P*: 3.34×10^{-2} MAF (case): 0.168 MAF (control): 0.200
rs4072037 A/G/A	OR: 1.62 (1.32–1.99) *P*: 4.04×10^{-6} MAF (case): 0.126 MAF (control): 0.187	OR: 1.69 (1.27–2.25) *P*: 2.82×10^{-4} MAF (case): 0.099 MAF (control): 0.164	OR: 1.74 (1.26–2.39) *P*: 7.82×10^{-4} MAF (case): 0.093 MAF (control): 0.163	OR: 1.64 (1.39–1.94) *P*: 4.46×10^{-9}	OR: 1.66 (1.44–1.93) *P*: 1.43×10^{-11}	OR: 1.23 (1.02–1.48) *P*: 3.36×10^{-2} MAF (case): 0.157 MAF (control): 0.187
rs2066981 T/C/T	OR: 1.61 (1.31–1.98) *P*: 5.91×10^{-6} MAF (case): 0.125 MAF (control): 0.186	OR: 1.74 (1.31–2.32) *P*: 1.50×10^{-4} MAF (case): 0.099 MAF (control): 0.167	OR: 1.76 (1.27–2.43) *P*: 6.98×10^{-4} MAF (case): 0.092 MAF (control): 0.159	OR: 1.65 (1.40–1.95) *P*: 3.77×10^{-9}	OR: 1.67 (1.44–1.94) *P*: 1.11×10^{-11}	OR: 1.21 (1.01–1.46) *P*: 4.41×10^{-2} MAF (case): 0.158 MAF (control): 0.186
rs914615 G/A/G	OR: 1.61 (1.31–1.98) *P*: 5.83×10^{-6} MAF (case): 0.126 MAF (control): 0.187					OR: 1.22 (1.01–1.47) *P*: 3.94×10^{-2} MAF (case): 0.158 MAF (control): 0.187
rs2075570[j] A/G/A	OR: 1.63 (1.32–2.00) *P*: 3.45×10^{-6} MAF (case): 0.127 MAF (control): 0.189	OR: 1.77 (1.33–2.37) *P*: 1.12×10^{-4} MAF (case): 0.100 MAF (control): 0.170	OR: 1.84 (1.32–2.55) *P*: 2.73×10^{-4} MAF (case): 0.091 MAF (control): 0.163	OR: 1.67 (1.42–1.98) *P*: 1.73×10^{-9}	OR: 1.71 (1.47–1.98) *P*: 2.26×10^{-12}	OR: 1.21 (1.01–1.46) *P*: 4.25×10^{-2} MAF (case): 0.159 MAF (control): 0.189
rs1057941 T/C/T	OR: 1.67 (1.36–2.04) *P*: 6.59×10^{-7} MAF (case): 0.130 MAF (control): 0.200					OR: 1.25 (1.04–1.50) *P*: 1.95×10^{-2} MAF (case): 0.168 MAF (control): 0.200

NOTE. OR: per-allele odds ratio for risk allele with 95% confidence intervals in parentheses.
MAF, minor allele frequency; SNP, single nucleotide polymorphism.
[a]Major, minor, and risk alleles are common in Tokyo, Aichi, and Korea data sets.
[b]Association study on diffuse-type GC in the Japanese population (Tokyo data set, 606 cases, 1264 controls), performed with fine-mapping data.
[c]*P* values and ORs with 95% CI were calculated under an additive model using logistic regression adjusted for age, gender, and number of risk alleles of rs2294008 of the *PSCA* gene, which was associated with DGC in our previous GWAS (Study Group of Millennium Genome Project for Cancer[7]).
[d]Association study on diffuse-type GC in the Japanese population (Aichi data set: 304 cases, 1465 controls), for replication of the study on Tokyo data set.
[e]Association study on diffuse-type GC in the Korean population (Korea data set: 452 cases, 372 controls), for replication of the study on Tokyo data set.
[f]Meta-analysis on the data of the Tokyo and Aichi data sets.
[g]*P* values and ORs with 95% CI were calculated using a random effects model.
[h]Meta-analysis on the data of the Tokyo, Aichi, and Korea data sets.
[i]Association study on intestinal-type GC in the Japanese population (Tokyo data set: 599 cases, 1264 controls).
[j]SNPs identified in the GWAS (Study Group of Millennium Genome Project for Cancer[7]).

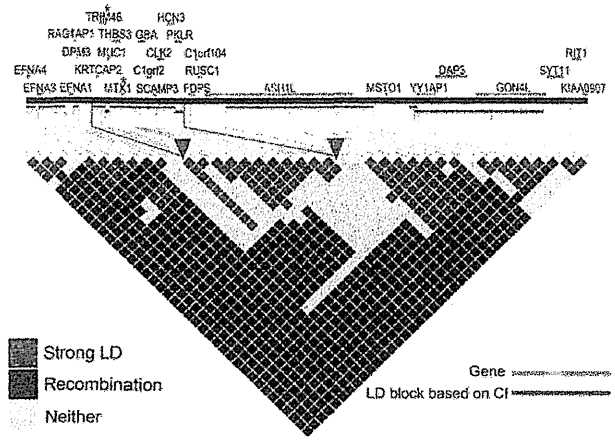

Figure 2. LD analyses on the SNPs in chromosome 1q22 using Gabriel's criteria[15] based on genotyping data of 680 Japanese controls. LD blocks with the criteria are in *blue horizontal lines*. *Red asterisks* indicate positions of rs2075570 and rs2070803. The LD block (both ends indicated by *red triangles*), in which the 2 SNPs reside, contains 12 SNPs and 11 genes. CI, confidence interval.

56.1; 178 females; mean age, 54.6; 2 missing gender information; 372 of which were the same as DGC cases in the Tokyo data set) and 309 controls (151 males; mean age, 49.1; 158 females; mean age, 46.1; 306 of which were included in the Tokyo data set). DGC samples were collected at 2 institutions as follows: 318 paraffin-embedded tissues at the National Cancer Center Hospital, and 62 blood samples at Nippon Medical School Hospital. Control DNA samples were from Keio University campuses.

The Japanese part of the study was approved by the ethics committees of the participating institutions in accordance with the Ethics Guidelines For Human Genome/Gene Analysis Research in Japan. The Korean side of the GC case control study was approved by the Ethics Committee of the National Cancer Center, Korea. Informed consent was obtained from all living subjects, including opt-out consent for the paraffin block archival samples.

LD Analysis

The LD map of chromosome 1q22 (Figure 2) was constructed based on the genotype data of 41 SNPs (Supplementary Table 10) obtained from 680 Japanese controls (436 males; mean age, 43.7; 242 females; mean age, 43.7; 2 missing gender information; 371 of which were from controls in the Tokyo data set) genotyped by Illumina Human610-Quad BeadChip (Illumina, San Diego, CA). The LD map shown in Figure 1 was constructed based on the fine mapping data of the 52 SNPs (Supplementary Table 10) from the Single Nucleotide Polymorphism database (http://www.ncbi.nlm.nih.gov/projects/SNP/) on the 1266 Japanese controls (the same as controls in the Tokyo data set). The pattern of LD was analyzed using 2 parameters, r^2 and $|D'|$,[14] and the confidence interval of the $|D'|$ was also utilized.[15]

Statistical Analyses

Statistical significance of the association was evaluated for each SNP by logistic regression. *P* values under an additive model adjusted for 3 age categories (≤39, 40–59, and ≥60 years), gender, and the risk alleles at rs2294008 in *PSCA* (Supplementary Figures 9–11). The significance level was set to .05 by Bonferroni correction for multiple testing, meaning $P = 8.9 \times 10^{-4}$ before correction for the Tokyo data set. Meta-analyses of the Tokyo data set and the Aichi and Korea replication data sets were performed using a random effects model.[16] Haplotype-based association was tested by Fisher exact test. Haplotype phases in each individual were inferred by fastPHASE software.[17] Other statistical analyses were carried out using the R suite (http://www.r-project.org/) and the StatXact 8 (Cytel Inc, Cambridge, MA). Population stratification of the Tokyo data set was examined previously by the STRUCTURE software,[18] the Genomic Control, and mixture model methods,[19,20] and no significant subpopulation was detected.[7]

In the association studies using 2-locus genotype data of rs4072037 and rs2294008, the biologic effect of the SNPs' risk allele was assumed to be recessive (rs4072037) or dominant (rs2294008), ie, the risk genotype for rs4072037 is AA and, for rs2294008, TT and TC (Figure 3). Risk factor variables consist of 4 categories based on the genotypes of rs4072037 and rs2294008. *P* value and OR and its 95% confidence interval (CI) for each category was obtained by logistic regression adjusted for age and gender.

Other Analyses

The materials and methods used in IGC association studies, genotyping, resequencing, and functional studies are described in Supplementary Materials and Methods and Supplementary Tables 9–11.

Results

Identification of the Susceptibility Region in Chromosome 1q22

Initially, we analyzed LD ($|D'|$) around the 2 marker SNPs based on the genotyping data on 680 control subjects. The criteria based on a confidence interval of $|D'|$[15] was applied to find an LD block containing 12 SNPs (including rs2075570 and rs2070803) and 11 genes (Figure 2). The second analysis of high-density genotyping around this block was performed on 610 cases of DGC and 1266 controls (Tokyo data set) for 52 SNPs selected from the Single Nucleotide Polymorphism database. A solid 49kb-LD block was identified spanning 13 SNPs including rs2075570 and rs2070803. Eight SNPs in the block showed strong associations ($P < 1.0 \times 10^{-5}$)

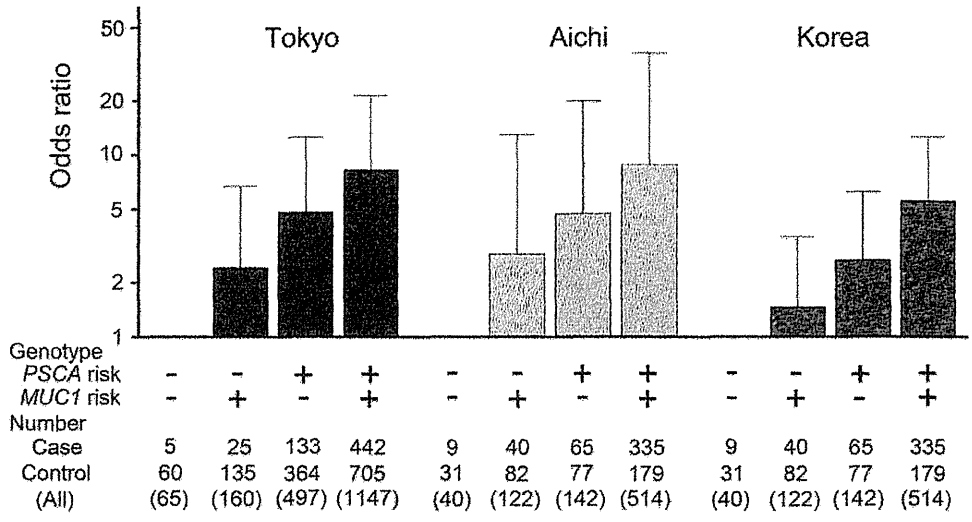

Figure 3. Association studies for DGC using 2-locus genotype data of rs4072037 in *MUC1* and rs2294008 in *PSCA*. The association studies were performed with a distinct model for each risk allele's effect, recessive for rs4072037 and dominant for rs2294008. *Bar*, upper bound of 95% confidence interval.

with DGC (Figure 1 and Table 1). Of the 8 SNPs, 4 were selected and genotyped on independent case-control sets in Japan (Aichi data set: 304 cases, 1467 controls) and in Korea (Korea data set: 455 cases, 372 controls), and the association was replicated in both data sets (Figure 1 and Table 1). A meta-analysis of the 3 case-control studies also showed significant correlation of the region: $P = 2.26 \times 10^{-12}$; OR, 1.71 for rs2075570 and $P = 4.33 \times 10^{-13}$; OR, 1.71 for rs2070803. Because this 1q22 region was originally identified by the GWAS on DGC,[7] an association with IGC was examined on the 3 populations: 601, 274, and 415 cases from Tokyo, Aichi, and Korea, respectively, using the same control subjects analyzed for DGC. The 1q22 region was not significantly associated with IGC (Figure 1, Supplementary Figure 2). Full data of the association studies are shown in Supplementary Tables 1–4.

Polymorphisms in MUC1 *Gene and Haplotype-Based Association Study*

The 49-kilobase (kb) block contained 5 genes encoding keratinocyte associated protein 2 (*KRTCAP2*), tripartile motif protein 46 (*TRIM46*), mucin 1 (*MUC1*), thrombospondin 3 (*THBS3*), and metaxin 1 (*MTX1*). Based on their expression patterns and gene annotations, we prioritized *MUC1* for further analyses because *MUC1* is expressed in pit cells in the pit region, mucous neck cells in the neck region, chief (zymogenic) cells in the base region, and parietal cells in the neck and base regions of the gastric epithelium (Figure 4*A*, Supplementary Figure 1).[21] Moreover, previous studies based on a candidate-gene approach reported an association between its polymorphisms and GC.[8–11]

The resequencing of the *MUC1* gene identified a total of 7 polymorphisms in 48 Japanese individuals: 4 SNPs without rs numbers (numbers 1, 2, 4, and 7), 1 indel (No. 6, rs66597679), rs12411216 (No. 3), and rs4072037 (No. 5) (Figure 4*B* and Supplementary Table 5). The 7 were genotyped on 380 Japanese cases and 309 controls (Table 2), and, in a subsequent haplotype analysis, SNPs numbers 1 and 7 were removed from analysis because they were monomorphic in the 689 Japanese individuals. The remaining 5 SNPs were used for a haplotype-based association study, which revealed 3 major haplotypes, numbers 1–3, with ORs of 1.32, 0.90, and 0.65, respectively, and 1 minor haplotype, No. 4, with minor allele frequency of 0.0105 in cases and 0.0032 in controls (Table 3).

Functional Analyses of MUC1 *SNPs*

Seven transcriptional variants are registered as *MUC1* messenger RNA in the National Center for Biotechnology Information database (http://www.ncbi.nlm.nih.gov/) (Supplementary Figure 3), and the rs4072037 SNP ($P = 1.43 \times 10^{-11}$ and OR of 1.66 by meta-analysis of the 2 Japanese and 1 Korea data sets, Table 1) located in exon 2 of *MUC1* had been found to be related to the splicing site selection in the exon.[12] To identify the variants expressed in the stomach, we conducted RNA ligase-mediated rapid amplification of the 5′ complementary DNA end procedure.

Our results showed that the major transcripts in the stomach are variants 2 and 3 (Figure 4*B* and Supplementary Figure 4) and that all the examined clones of variant 2 possessed G allele at rs4072037, in contrast to those of the variant 3 possessing the A allele, as reported previously (Supplementary Figure 5).[12] This suggests that rs4072037 is significantly involved in the splicing regulation of the second exon. In other words, it is likely that the SNP directly determines the relative dominance of the 2 major *MUC1* splicing variants, the variants 2 and 3, in the gastric epithelium.

As reported previously on the Caucasian population,[12] no polymorphisms other than rs4072037 were found in the region spanning from exon 1 to 2, which might affect the splicing of the second exon, by our resequencing of the

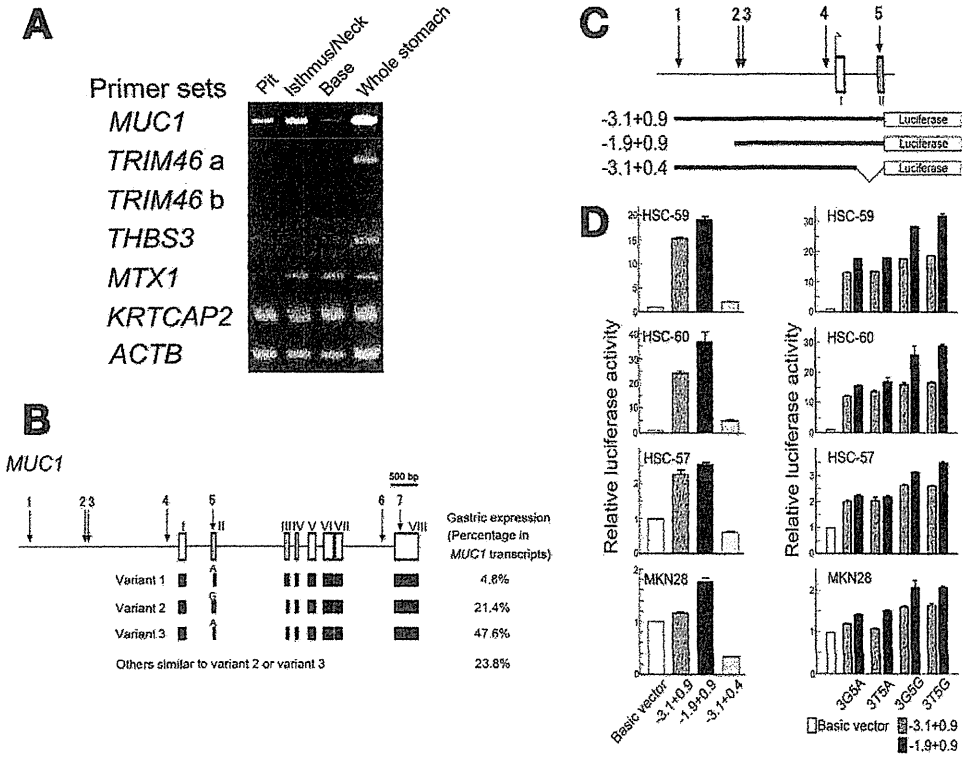

Figure 4. Functional analyses of MUC1 and its SNPs. (*A*) Expression analysis on the 5 genes in the LD block associated with DGC using microdissected gastric samples (reverse-transcription polymerase chain reaction). (*B*) Seven polymorphisms in the *MUC1* gene identified by resequencing of 48 Japanese controls and *MUC1* transcriptional variants detected in RNAs from the gastric mucosa by RNA ligase-mediated rapid amplification of the 5′ complementary DNA (cDNA) end procedure (5′RACE). SNP positions are indicated by *arrows* with numbers corresponding to Table 2 and Supplementary Table 5. The 5′RACE was conducted on a pooled gastric RNA sample from 21 Caucasian individuals. All the variant 2 clones contained exclusively the G allele of SNP No. 5 and all the variants 1 and 3 the A allele, without exception. Complete result of the 5′RACE is presented in Supplementary Figure 4. (*C*) *MUC1* genomic fragments examined in reporter assays. (*D*) Effect of SNP No.5 (rs4072037) on the transcriptional activity of the *MUC1* promoter (reporter assay). The transcriptional activity of the −3.1 to +0.9 region was significantly reduced by truncating the +0.4 to +0.9 region (*−3.1+0.4*) in gastric cancer cell lines HSC-59, HSC-60, HSC-57, and MKN28. The genomic fragments with G allele in the SNP (*3G5G* and *3T5G*) showed higher reporter activity than that with A allele (*3G5A* and *3T5A*).

MUC1 gene on 48 Japanese individuals (Figure 4*B* and Supplementary Table 5). As regards the splicing variation, the risk allele A at rs4072037, which is found in variants 1 and 3 (Supplementary Figure 5), causes a 9-amino acid deletion in the second exon and consequently modifies both the signal peptide and N-terminal amino acid of the mature protein by changing the signal-peptide cleavage site.[12] This may change the intracellular trafficking and glycosylation and folding of the protein, leading to alteration in the function of the mature protein.

To examine the difference in the MUC1 function among the variants, we evaluated in vitro functions of the

Table 2. Seven SNPs and Their Association With DGC Based on the Genotype Data of 380 Japanese Cases and 309 Controls

SNP no.	rs number	Major allele	Minor allele	Risk allele	MAF (case)	MAF (control)	OR[a]	95% CI	*P* value[b]
1		T	C		0.0000	0.0000			
2		C	T	C	0.1882	0.2039	1.11	0.85–1.44	.4945
3	12411216	G	T	G	0.1289	0.1851	1.53	1.14–2.06	.004344
4		G	A	A	0.0105	0.0032	3.28	0.69–15.49	.1998
5	4072037	A	G	A	0.1289	0.1857	1.54	1.15–2.07	.004275
6	66597679	AC	—	AC	0.1308	0.1869	1.53	1.14–2.05	.005225
7		C	T		0.0000	0.0000			

CI, confidence interval; MAF, minor allele frequency.
[a]Odds ratio for risk allele.
[b]*P* values obtained by Fisher exact test.

Table 3. Four Major Haplotypes Inferred and Their Association With DGC Based on the Genotype Data of 380 Japanese Cases and 309 Controls

	SNP no.									*P* values	
	2	3	4	5	6	Case[a]	Control[b]	OR	95% CI	Fisher[c]	Permutation[d]
Haplotype no.											
1	C	G	G	A	AC	0.6733	0.6084	1.32	1.06–1.65	.01501	.0483
2	T	G	G	A	AC	0.1884	0.2039	0.90	0.69–1.18	.4945	.9027
3	C	T	G	G	—	0.1278	0.1845	0.65	0.48–0.87	.004200	.0127
4	C	G	A	A	AC	0.0105	0.0032	3.28	0.69–15.49	.1998	.4058

CI, confidence interval.
[a]Frequency of case.
[b]Frequency of control.
[c]*P* values obtained by Fisher exact test.
[d]*P* values obtained by permutation test (100,000 permutations performed).

2 major variants expressed in the gastric epithelium: variants 2 and 3. Because the full-length product of *MUC1* is well-known for its cell growth-promoting activity in cancer cells,[22,23] we individually transfected a cytomegalovirus promoter-driven complementary DNA of *MUC1* variant 2 or 3 to the MKN28 cells, which express *MUC1* at an undetectable level (data not shown). Examination of their cell growth by both cell counting and colorimetric methods suggested that variant 2 is more potent in growth-promoting activity than variant 3 (Supplementary Figures 6 and 7). Although the observed difference seems to be small, this level of normal range of individual variation is generally expected for a common genetic variant influencing a common disease susceptibility and is probably because of the difference in the signal peptide or the N-terminal structure of the mature protein encoded by each variant because the other portion of the amino acid sequence is common between the 2 variants.

Next, we investigated the function of the *MUC1* SNPs in the context of the haplotypes. We selected haplotype No. 1 as the major risk haplotype and haplotype No. 3 as the most protective haplotype and analyzed the functions of SNPs numbers 3 and 5, excluding SNP No. 6 from our functional analyses because its location in the intron 7 made it unlikely to be involved in the transcriptional regulation and/or alternative splicing of the gene. SNP numbers 2 and 4 were also excluded because they were found on both the risk and the protective haplotypes. Because the remaining 2 SNPs, numbers 3 and 5, do not change amino acid, we first examined, by a reporter assay, their effect on the transcriptional regulation of *MUC1*; the region spanning −1.9 to 0.9 kb relative to the *MUC1* transcription start site had a transcriptional activity (Figure 4*C* and *D*). The reporter assay on base-substituted constructs showed that, in all the gastric carcinoma cell lines examined, the fragments containing the G allele at rs4072037 (SNP No. 5), which is present only in the protective haplotype, has a higher transcriptional activity than that with an A allele present in the risk haplotype (Figure 4*D*, *right panel*). The assay on truncated constructs showed that a removal of a +0.4 to +0.9-kb region, which contains rs4072037, significantly diminishes the transcriptional activity (Figure 4*D*, *left panel*). We also confirmed in the reporter assay that the T allele of SNP No. 2, which is unique to haplotype No. 2 showing OR of 0.9 but no significant *P* value, has no effect (Supplementary Figure 8).

In sum, the results in this study and from previous reports by other investigators suggest that rs4072037 SNP has at least 2 functions: (1) regulation of the alternative splicing at the second exon and (2) modification of the transcriptional activity of the promoter. The association study in the context of LD and the functional study strongly implicate rs4072037 as a functional cause of the association between the 1q22 region and DGC susceptibility.

Association Studies for DGC Using 2-Locus Genotype Data of rs4072037 in MUC1 *and rs2294008 in* PSCA

Finally, we examined the effect of 2 DGC susceptibility SNPs identified by our GWAS, rs4072037 in *MUC1* and rs2294008 in *PSCA*, both of which are functional, using Tokyo, Aichi, and Korea data sets. When a genetic model is tentatively selected for each locus by simply comparing *P* value, a recessive and dominant model was applied for rs4072037 and for rs2294008, respectively; the individuals possessing the risk genotype of both SNPs showed significant risk for developing DGC (eg, OR, 8.38 in Tokyo data set, Figure 3). Notably, individuals with protective alleles of both SNPs were observed only in controls (Supplementary Table 6).

Discussion

At chromosome 1q22, we focused on the region with strong LD around rs2075570 and rs2070803 using Gabriel et al's criteria.[15] The region contains 5 genes: *TRIM46*, *THBS3*, *MTX1*, *KRTCAP2*, and *MUC1* (Figure 1). We prioritized the genes for the subject of

further studies by the first criterion (whether the gene is expressed in the gastric epithelial cells) and then by the second one (whether annotated function suggests its involvement in carcinogenesis). We observed transcripts of 4 of the 5 genes in microdissected samples of the gastric epithelium by reverse-transcription polymerase chain reaction, but no transcript of *TRIM46* was detectable there (Figure 4*A*). *THBS3* encoding a multifunctional extracellular matrix glycoprotein is expressed in multiple human tissues including the stomach,[24] and no evidence of a causal relation to carcinogenesis has been obtained. *MTX1* encodes a component of a preprotein import complex in the outer membrane of the mammalian mitochondrion.[25] If it is involved in carcinogenesis, the effect of its SNP would be reflected in many types of cancer, yet no such involvement has surfaced. *KRTCAP2* encodes a protein possessing transmembrane domain, showing multitissue expression.[26] Its function is unknown, and its relation to carcinogenesis has not been demonstrated.

In this study, we considered *MUC1* as a strong candidate for the gene responsible for the association of 1q22 with DGC because, in addition to several previous candidate gene analyses showing an association with *MUC1* polymorphisms and GC,[8–11] MUC1 has been considered to possess an oncogenic property as described below.

Mucin family members are classified into 2 types, secreted or membranous, based on their localization, and MUC1 is a transmembrane mucin.[27] MUC1 is a multifunctional protein involved in mucosal lubrication, protection from pathogens, signal transduction, and cell-cell interaction.[27] MUC1 was over-expressed in breast, ovarian, lung, pancreatic, and prostate cancers and was a marker of poor prognosis in gastric cancer.[21,28] Several in vivo studies have provided evidence supporting its function in carcinogenesis. MUC1 has a role in cell growth, anchorage independence, cell migration, antiapoptotic property, and drug resistance of cancer cells,[22,29–33] all of which are accomplished through interaction with several signaling pathways,[34] although these lines of biologic evidence were obtained on the standard molecule containing tandem repeats (TR). Because the MUC1 expressed in the gastric epithelium has no TR, it is possible that its function in gastric epithelial cells is different from that of the TR-containing product in other epithelial cells.[35–37]

Recently, however, it is supposed that MUC1 has a protective function against environmental insults and acts against tumorigenesis in normal epithelial cells, which keep maintaining their cell polarity. In contrast, once the cells lose cell polarity in consequence of prolonged inflammation, MUC1 promotes cell growth and acts for tumorigenesis.[38] It was also reported that MUC1 functions as a growth factor receptor in human embryonic stem cells.[39] It is presumable that MUC1 is involved in growth regulation of gastric stem cells and progenitors, which are considered to be the origin of DGC. Like the function of PSCA, which is up-regulated in prostate and urinary bladder cancers but suppressed in gastric cancer,[7] the MUC1 function may differ between cell types, normal or malignant, and among different tissues. In the same manner as PSCA, MUC1 is down-regulated in intestinal metaplasia of the gastric epithelium from which IGC arises.[40] If MUC1 has some protective function in carcinogenesis, this down-regulation makes the stem and progenitor cells more susceptible to carcinogenic events. In any case, further research is needed to explore the pleiotropic functions of MUC1.

In this study, we demonstrated that rs4072037 has a role in transcriptional regulation and also in splicing site selection leading to the dominant variant determination of *MUC1* transcripts in gastric epithelial cells. If variant 3 is less functional in protection against DGC than variant 2, the possession of the A allele in the genome confers both quantitatively and qualitatively unfavorable consequences to MUC1 function, which may result in additive risk for DGC susceptibility.

In our GWAS on DGC susceptibility, the 2 loci showing the highest statistical significance directed us to the 2 functional SNPs: rs4072037 in *MUC1* and rs2294008 in *PSCA*.[7] It is noteworthy that both SNPs in the 2 genes appear to have dual functions: transcriptional regulation and signal-peptide modification.[7] Further investigation is required to validate the role of *MUC1* in DGC susceptibility and details of the mechanism that links the risk haplotype tagged by rs4072037 to DGC development.

The risk allele A of rs4072037 is in strong LD with the small allele of the variable numbers of tandem repeats in the second intron of the *MUC1* both in Europeans and Japanese; more than 90% of chromosomes have a nonrecombinant haplotype in both populations (Ng et al[12] and Supplementary Table 7), and the small allele was associated with GC in the European population.[8,9,12] However, the variable numbers of tandem repeats is unlikely to be the causal polymorphism for DGC susceptibility because the TRs are translated neither in normal nor malignant gastric epithelial cells (Supplementary Figure 4 and Supplementary Table 8).

This study has not only replicated the association of the *MUC1* SNP with GC in the Japanese and Korean populations, in addition to the previous reports on the Chinese and Caucasian population GCs,[8–11] but it has also disclosed that the association appears specific to DGC. Following discovery of the DGC-specific association of the *PSCA* polymorphism, this study has offered another piece of evidence to support distinct mechanisms for DGC and IGC development.

Although there was no significant interaction between the *MUC1* and *PSCA* SNPs for the DGC risk in our study (Figure 3, Supplementary Table 6, and data not shown), it is estimated that individuals with the double risk

genotype are the majority in Japanese (56%) and Korean (49%) populations with a significant OR, 8.38, in Japanese, as compared with the lowest risk category. GWAS and other emerging genome analysis tools may unveil a number of polymorphisms showing a significant statistical association, but it is important to identify functional SNPs potentially related to carcinogenesis. The accumulation of information on the functional SNPs, environmental factors, and their interactions, all of which are truly related to DGC susceptibility, will make the genotyping a more practical tool for evaluating the individual risk for DGC and offer effective prevention strategies in the future.

Supplementary Material

Note: To access the supplementary material accompanying this article, visit the online version of *Gastroenterology* at www.gastrojournal.org, and at doi: 10.1053/j.gastro.2010.10.058.

References

1. Brenner H, Rothenbacher D, Arndt V. Epidemiology of stomach cancer. In: Verma M, ed. Methods of molecular biology: cancer epidemiology. Volume 472. New Jersey: Humana, 2009:467–477.
2. Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. Acta Pathol Microbiol Scand 1965; 64:31–49.
3. Chiba T, Marusawa H, Seno H, et al. Mechanism for gastric cancer development by *Helicobacter pylori* infection. J Gastroenterol Hepatol 2008;23:1175–1181.
4. Schier S, Wright NA. Stem cell relationships and the origin of gastrointestinal cancer. Oncology 2005;69(Suppl 1):9–13.
5. Henson DE, Dittus C, Younes M, et al. Differential trend in the intestinal and diffuse types of gastric carcinoma in the United States, 1973–2000—increase in the signet ring cell type. Arch Pathol Lab Med 2004;128:765–770.
6. Rosai J. Rosai and Ackerman's surgical pathology. Edinburgh: Mosby, 2004.
7. Study Group of Millennium Genome Project for Cancer. Genetic variation in *PSCA* is associated with susceptibility to diffuse-type gastric cancer. Nat Genet 2008;40:730–740.
8. Carvalho F, Seruca R, David L, et al. *MUC1* gene polymorphism and gastric cancer—an epidemiological study. Glycoconj J 1997; 14:107–111.
9. Silva F, Carvalho F, Peixoto A, et al. *MUC1* gene polymorphism in the gastric carcinogenesis pathway. Eur J Hum Genet 2001;9: 548–552.
10. Xu Q, Yuan Y, Sun LP, et al. Risk of gastric cancer is associated with the *MUC1* 568 A/G polymorphism. Int J Oncol 2009;35: 1313–1320.
11. Jia Y, Persson C, Hou L, et al. A comprehensive analysis of common genetic variation in *MUC1, MUC5AC, MUC6* genes and risk of stomach cancer. Cancer Causes Control 2010;21:313–321.
12. Ng W, Loh AX, Teixeira AS, et al. Genetic regulation of *MUC1* alternative splicing in human tissues. Br J Cancer 2008;99:978–985.
13. Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma. 2nd English edition. Gastric Cancer 1998;1: 10–24.
14. Nei M. Molecular evolutionary genetics. New York: Columbia University Press, 1987.
15. Gabriel SB, Schaffner SF, Nguyen H, et al. The structure of haplotype blocks in human genome. Science 2002;296:2225–2229.
16. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–188.
17. Scheet P, Stephens M. A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. Am J Hum Genet 2006; 78:629–644.
18. Falush D, Stephens M, Pritchard JK. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. Genetics 2003;164:1567–1587.
19. Bacanu SA, Devlin B, Roeder K. The power of genomic control. Am J Hum Genet 2000;66:1933–1944.
20. Zhu X, Zhang S, Zhao H, et al. Association mapping, using a mixture model for complex traits. Genet Epidemiol 2002;23: 181–196.
21. Senapati S, Sharma P, Bafna S, et al. The MUC gene family: their role in the diagnosis and prognosis of gastric cancer. Histol Histopathol 2008;23:1541–1552.
22. Li Y, Liu D, Chen D, et al. Human DF3/MUC1 carcinoma-associated protein functions as an oncogene. Oncogene 2003;4: 6107–6110.
23. Tsutsumida H, Swanson BJ, Singh PK, et al. RNA interference suppression of MUC1 reduces the growth rate and metastatic phenotype of human pancreatic cancer cells. Clin Cancer Res 2006;15:2976–2987.
24. Adolph KW. Relative abundance of Thrombospondin 2 and Thrombospondin 3 mRNAs in human tissues. Biochem Biophys Res Commun 1999;258:792–796.
25. Armstrong LC, Saenz AJ, Bornstein P. Metaxin 1 interacts with metaxin 2, a novel related protein associated with the mammalian mitochondrial outer membrane. J Cell Biochem 1999;74: 11–22.
26. Bonkobara M, Das A, Takao J, et al. Identification of novel genes for secreted and membrane-anchored proteins in human keratinocytes. Br J Dermatol 2003;148:654–664.
27. Gendler SJ. MUC1, the renaissance molecule. J Mammary Gland Biol Neoplasia 2001;6:339–353.
28. Taylor-Papadimitriou J, Burchell J, Miles DW, et al. MUC1 and cancer. Biochim Biophys Acta 1999;1455:301–313.
29. Suwa T, Hinoda Y, Makiguchi Y, et al. Increased invasiveness of *MUC1* cDNA-transfected human gastric cancer MKN74 cells. Int J Cancer 1998;76:377–382.
30. Raina D, Kharbanda S, Kufe D. The MUC1 oncoprotein activates the anti-apoptotic phosphoinositide 3-kinase/Akt and Bcl-xL pathways in rat 3Y1 fibroblasts. J Biol Chemi 2004;279:20607–20612.
31. Ren J, Agata N, Chen D, et al. Human MUC1 carcinoma-associated protein confers resistance to genotoxic anticancer agents. Cancer Cell 2004;5:163–175.
32. Huang L, Chen D, Liu D, et al. MUC1 oncoprotein blocks glycogen synthase kinase 3β-mediated phosphorylation and degradation of β-catenin. Cancer Res 2005;65:10413–10422.
33. Wei X, Xu H, Kufe D. Human mucin 1 oncoprotein represses transcription of the *p53* tumor suppressor gene. Cancer Res 2007;67:1853–1858.
34. Singh PK, Hollingsworth MA. Cell surface-associated mucins in signal transduction. Trends Cell Biol 2006;16:467–476.
35. Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, et al. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. J Biol Chem 1990;265:15286–15293.

36. Zrihan-Licht S, Vos HL, Baruch A, et al. Characterization and molecular cloning of a novel MUC1 protein, devoid of tandem repeats, expressed in human breast cancer tissue. Eur J Biochem 1994;224:787–795.
37. Baruch A, Hartmann M, Zrihan-Licht S, et al. Preferential expression of novel MUC1 tumor antigen isoforms in human epithelial tumors and their tumor-potentiating function. Int J Cancer 1997; 71:741–749.
38. Kufe DW. Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer 2009;9:874–885.
39. Hikita ST, Kosik KS, Clegg DO, et al. MUC1* mediates the growth of human pluripotent stem cells. PLoS One 2008;3:e3312.
40. Silva E, Teixeira A, David L, et al. Mucins as key molecules for the classification of intestinal metaplasia of the stomach. Virchows Arch 2002;440:311–317.

Received May 8, 2010. Accepted October 26, 2010.

Reprint requests

Address requests for reprints to: Teruhiko Yoshida, MD, Genetics Division, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan. e-mail: tyoshida@ncc.go.jp; fax: (81) 3-3541-2685.

Acknowledgments

The authors thank Sachiyo Mimaki, Mineko Ushiama, Chie Naito, Yoko Odaka, Misuzu Okuyama, Miki Watanabe, and Hidemi Ito for their technical contributions and Seiji Ito for the clinical data collection of the Aichi data set.

Conflicts of interest

The authors disclose no conflicts.

Funding

Supported in Japan by the program for promotion of Fundamental Studies in Health Sciences of the National Institute of Biomedical Innovation (NiBio); in Aichi, supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science, Sports, Culture and Technology of Japan and by a Grant-in-Aid for the Third Term Comprehensive 10-Year Strategy for Cancer Control from the Ministry of Health, Labour and Welfare of Japan; and, in Korea, part of the study was supported by grant 0710340 from the National Cancer Center, Korea.

Published OnlineFirst on March 23, 2011; DOI:10.1158/1078-0432.CCR-11-0063

Clinical Cancer Research

Pulmonary Inflammatory Myofibroblastic Tumor Expressing a Novel Fusion, PPFIBP1 –ALK: Reappraisal of Anti-ALK Immunohistochemistry as a Tool for Novel ALK Fusion Identification

Kengo Takeuchi, Manabu Soda, Yuki Togashi, et al.

Clin Cancer Res Published OnlineFirst March 23, 2011.

Published OnlineFirst on March 23, 2011; DOI:10.1158/1078-0432.CCR-11-0063

Clinical Cancer Research

Imaging, Diagnosis, Prognosis

Pulmonary Inflammatory Myofibroblastic Tumor Expressing a Novel Fusion, PPFIBP1–ALK: Reappraisal of Anti-ALK Immunohistochemistry as a Tool for Novel ALK Fusion Identification

Kengo Takeuchi[1,2], Manabu Soda[4], Yuki Togashi[1,2], Emiko Sugawara[1,5], Satoko Hatano[1,2], Reimi Asaka[1,2], Sakae Okumura[3], Ken Nakagawa[3], Hiroyuki Mano[4,6], and Yuichi Ishikawa[2]

Abstract

Purpose: The anaplastic lymphoma kinase (ALK) inhibitor crizotinib has been used in patients with lung cancer or inflammatory myofibroblastic tumor (IMT), both types harboring ALK fusions. However, detection of some ALK fusions is problematic with conventional anti-ALK immunohistochemistry because of their low expression. By using sensitive immunohistochemistry, therefore, we reassessed "ALK-negative" IMT cases defined with conventional immunohistochemistry (approximately 50% of all examined cases).

Experimental Design: Two cases of ALK-negative IMT defined with conventional anti-ALK immunohistochemistry were further analyzed with sensitive immunohistochemistry [the intercalated antibody-enhanced polymer (iAEP) method].

Results: The two "ALK-negative" IMTs were found positive for anti-ALK immunohistochemistry with the iAEP method. 5′-rapid amplification of cDNA ends identified a novel partner of ALK fusion, protein-tyrosine phosphatase, receptor-type, F polypeptide-interacting protein-binding protein 1 (PPFIBP1) in one case. The presence of PPFIBP1–ALK fusion was confirmed with reverse transcriptase PCR, genomic PCR, and FISH. We confirmed the transforming activities of PPFIBP1–ALK with a focus formation assay and an *in vivo* tumorigenicity assay by using 3T3 fibroblasts infected with a recombinant retrovirus encoding PPFIBP1–ALK. Surprisingly, the fusion was also detected by FISH in the other case.

Conclusions: Sensitive immunohistochemical methods such as iAEP will broaden the potential value of immunohistochemistry. The current ALK positivity rate in IMT should be reassessed with a more highly sensitive method such as iAEP to accurately identify those patients who might benefit from ALK-inhibitor therapies. Novel ALK fusions are being identified in various tumors in addition to IMT, and thus a reassessment of other "ALK-negative" cancers may be required in the forthcoming era of ALK-inhibitor therapy. *Clin Cancer Res; 17(10); 1–8. ©2011 AACR.*

Introduction

Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase that was discovered in anaplastic large cell lymphoma (ALCL) in the form of a fusion protein, NPM–ALK. (1, 2). In addition to ALCL (fused to NPM, TPM3, TPM4, ATIC, TFG, CLTC, MSN, MYH9, or ALO17; refs. 1–10), ALK has further been found to generate fusions in inflammatory myofibroblastic tumor (IMT; TPM3, TPM4, CLTC, CARS, RANBP2, ATIC, or SEC31L1; refs. 10–15), ALK-positive large B-cell lymphoma (CLTC, NPM, SEC31L1, or SQSTM1; 16–19), lung cancer (EML4 or KIF5B; refs. 20, 21), and ALK-positive histiocytosis (TPM3; ref. 22) Besides, some ALK fusions have been reported without showing histopathologic evidence: TPM4–ALK in esophageal squamous cell carcinoma (23, 24), TFG–ALK in lung adenocarcinoma (25), and EML4–ALK in colon and breast carcinomas (26). The wild-type ALK is mainly expressed in the developing nervous system, and is usually not expressed in other normal tissues (27). A fusion protein formation with a partner through chromosomal translocations is the most common mechanism of ALK overexpression and ALK kinase domain activation. These features render ALK fusion oncokinase an ideal molecular target.

Recently, the ALK inhibitor crizotinib has been used in patients with lung cancer or IMT, both types harboring ALK fusions (28, 29). The compound showed a 57% response rate in lung cancers (28), and a strong response for several months in IMT (29). Crizotinib and other ALK inhibitors

Authors' Affiliations: [1]Pathology Project for Molecular Targets; [2]Division of Pathology; and [3]Department of Thoracic Surgical Oncology, Thoracic Center, Cancer Institute Hospital, Japanese Foundation for Cancer Research, Tokyo; [4]Division of Functional Genomics, Jichi Medical University, Tochigi; [5]Department of Comprehensive Pathology, Graduate School, Tokyo Medical and Dental University; and [6]Department of Medical Genomics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan

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

Corresponding Author: Kengo Takeuchi, Pathology Project for Molecular Targets, The Cancer Institute, Japanese Foundation for Cancer Research, Tokyo 135-8550, Japan. Phone: +8133-520-0111; Fax: +8133-570-0230; E-mail: kentakeuchi-tky@umin.net

doi: 10.1158/1078-0432.CCR-11-0063

Published OnlineFirst on March 23, 2011; DOI:10.1158/1078-0432.CCR-11-0063

Translational Relevance

Anaplastic lymphoma kinase (ALK) inhibitors have become one of the most promising groups of molecularly targeted drugs. Therefore, ALK is no longer a mere research target or simply a diagnostic marker, but is directly linked to the therapeutic benefit of patients harboring the fusions.

Pathologic diagnoses for ALK fusion-positive tumors have been made reliably with anti-ALK immunohistochemistry. Since the discovery of EML4–ALK, however, an unexpected problem in anti-ALK immunohistochemistry has become apparent, that is, the inability to detect a low level of EML4–ALK expression. To overcome this, we developed the intercalated antibody-enhanced polymer immunohistochemistry, which successfully detected EML4–ALK.

In other words, this indicates that unknown ALK fusions, particularly those expressed at a low level, may wait to be discovered in "ALK-negative" tumors defined with conventional immunohistochemistry. In the forthcoming era of ALK-inhibitor therapy, "ALK-negative" tumors should be reassessed with a high sensitive immunohistochemistry and, if positive, be further examined with appropriate molecular method(s).

have thus become one of the most promising groups of molecularly targeted drugs. Therefore, the sensitive and accurate identification of ALK fusion in tumors has also become clinically relevant, because it is no longer a mere research target or simply a diagnostic marker, but is directly linked to the therapeutic benefit of patients harboring the fusions.

Identification of such ALK fusions, especially within ALCL, has been prompted by the immunohistochemical staining pattern with antibodies to ALK. In ALCL, the most common ALK fusion is NPM–ALK (comprising approximately 80% of all cases), and its immunohistochemical staining pattern is both nuclear and cytoplasmic. NPM has a nuclear localization signal in the C-terminal region, and therefore the heterodimers of wild-type NPM with NPM–ALK fusion protein are transported to the nucleus whereas NPM–ALK homodimers remain within the cytoplasm (30). In contrast, other fusions do not localize in the nucleus and do not show a nuclear staining pattern in anti-ALK immunohistochemistry. Interestingly, each ALK fusion usually has its own characteristic anti-ALK immunohistochemical staining pattern, because the subcellular localization of ALK fusions is dependent on the corresponding fusion partners. Anti-ALK immunohistochemistry has thus become a highly useful tool for both research and diagnostic purposes.

Since the discovery of EML4–ALK fusion in lung cancer (20), however, an unexpected problem in anti-ALK immunohistochemistry has become apparent, that is, the inability to detect a low level of fusion expression. To overcome this, we developed the intercalated antibody-enhanced polymer (iAEP) method, which moderately raises sensitivity in the immunohistochemical detection system (21). With this very simple method, anti-ALK immunohistochemistry has become a potent weapon in the diagnosis of EML4–ALK-positive lung cancer (21, 31–33). Other researchers used an anti-ALK rabbit monoclonal antibody, which is usually more sensitive than mouse monoclonal antibody, which can stain EML4–ALK (34). However, most EML4–ALK-positive lung cancer tissues do not stain well with conventional anti-ALK immunohistochemical methods because of the low message/protein level of EML4–ALK (21, 35). The expression level of a fusion gene depends on the promoter activity of the 5′-side gene, and that of EML4 is likely to be lower than that of the other ALK fusion partner genes, which may explain why EML4–ALK had not been discovered until 12 years after the development of the first anti-ALK antibody became available for immunohistochemistry (36). In other words, a tumor that immunostains for ALK only by a sensitive immunohistochemistry method may harbor a novel ALK fusion. Interestingly, in this study, we detected 2 IMT cases positive for ALK immunohistochemistry only when stained by iAEP method (21), and successfully identified a novel fusion gene, protein-tyrosine phosphatase, receptor-type, F polypeptide-interacting protein-binding protein 1 (PPFIBP1)–ALK.

Materials and Methods

Materials

Pathologic specimens from 2 pulmonary IMT cases, originally diagnosed as fibrous histiocytoma (1988: case 1, 45-year-old male; 1998: case 2, 34-year-old female), were reassessed morphologically and immunohistochemically. Surgically removed tumor specimens were routinely fixed in 20% neutralized formalin and embedded in paraffin for conventional histopathologic examination. For case 2, total RNA was extracted from the corresponding snap-frozen specimen and purified with the use of an RNeasy Mini kit (Qiagen). The study was approved by the institutional review board of the Japanese Foundation for Cancer Research.

Immunohistochemistry

Formalin-fixed, paraffin-embedded tissue was sliced at a thickness of 4 μm, and the sections were placed on silane-coated slides. For antigen retrieval, the slides were heated for 40 min at 97°C in Target Retrieval Solution (pH 9.0; Dako). For the conventional staining procedure, the slides were incubated at room temperature with Protein Block Serum-free Ready-to-Use solution (Dako) for 10 minutes and then with primary antibodies against ALK (5A4), smooth muscle actin, muscle-specific actin (HHF35), CD34, cytokeratins (AE1/AE3), S100, or desmin for 30 minutes. The immune complexes were then detected with dextran polymer reagent (EnVision + DAB system; Dako) and an AutoStainer instrument (Dako). The iAEP method was also used for the sensitive detection of ALK, as described previously (21).

Published OnlineFirst on March 23, 2011; DOI:10.1158/1078-0432.CCR-11-0063

Isolation of PPFIBP1–ALK fusion

To obtain cDNA fragments corresponding to a novel *ALK* fusion gene, we used a 5′-RACE method with the SMART RACE cDNA Amplification Kit (Clontech) according to the manufacturer's instructions, with a minor modification: the ALK2458R primer (5′-GTAGTTGGGGTTGTAGTCGGT-CATGATGGT-3′) was used as the gene-specific reverse primer.

From the oligo(dT)-primed cDNA obtained from case 2 RNA, a 471bp cDNA fragment containing the fusion point was specifically amplified with the primers PPFIBP1-592F (5′-AGAGACACAGAGGGGCTGATT-3′) and ALK3078RR (5′-ATCCAGTTCGTCCTGTTCAGAGC-3′).

PCR analysis of genomic DNA for *PPFIBP1–ALK* in case 2 was carried out with a pair of primers flanking the putative fusion point, PPFIBP1-607F (5′-CTGATTCAGGAGATCA-ATGATTTGAGGT-3′) and Fusion-RT-AS (5′-TCTTGCCAG-CAAAGCAGTAGTTGG-3′).

From the cDNA, a full-length cDNA for *PPFIBP1–ALK* was amplified by PCR with the PA-w-cDNA-in-S primer (5′-TATCTGGGTTGGAATTTGCCCCTG-3′) and the KA-w-cDNA-in-AS primer (5′-TGAGTGTGCGACCGAGCTCAGG-3′) and PrimeSTAR HS DNA polymerase (TakaraBio).

FISH

FISH analysis of gene fusion was carried out with bacterial artificial chromosome (BAC) clone-derived DNA probes for *ALK* and *PPFIBP1*. Unstained sections (4 μm thick) were subjected to hybridization with an *ALK*-split probe set (Abbott) or BAC clone-derived probes for *ALK* (RP11–984I21, RP11–62B19) and *PPFIBP1* (RP11–1060J15). Hybridized slides were then stained with DAPI and examined with the fluorescence microscope BX51 (Olympus).

Transformation assay for ALK fusion proteins

Analysis of the transforming activity of PPFIBP1–ALK was carried out as described previously (20, 37, 38). Briefly, the pMXS-based expression plasmid for PPFIBP1–ALK, EML4–ALK variant 1, or NPM–ALK was used to generate recombinant ecotropic retrovirus, followed by individual infection of mouse 3T3 fibroblasts (39). Formation of the transformed foci was evaluated after culturing the cells for 14 days. The same set of 3T3 cells was subcutaneously injected into nu/nu mice, and tumor formation was examined after 20 days. The animal experiments were approved by the animal ethics committee of Jichi Medical University.

Results

Morphology and immunophenotype of PPFIBP1–ALK-positive IMT

Histopathologic analysis of the 2 IMT cases revealed a marked proliferation of cells composed of somewhat histiocytoid spindle cells showing a fascicular or storiform pattern. The tumor cells were uniform and had pale eosinophilic cytoplasm and an oval vesicular nucleus, within which a small nucleolus was centrally located. Mild inflammatory infiltrate containing lymphocytes, plasma cells, foamy histiocytes, and multinucleated giant cells was observed (Fig. 1A and 1D). The immunophenotype of the 2 cases was negative for smooth muscle actin,

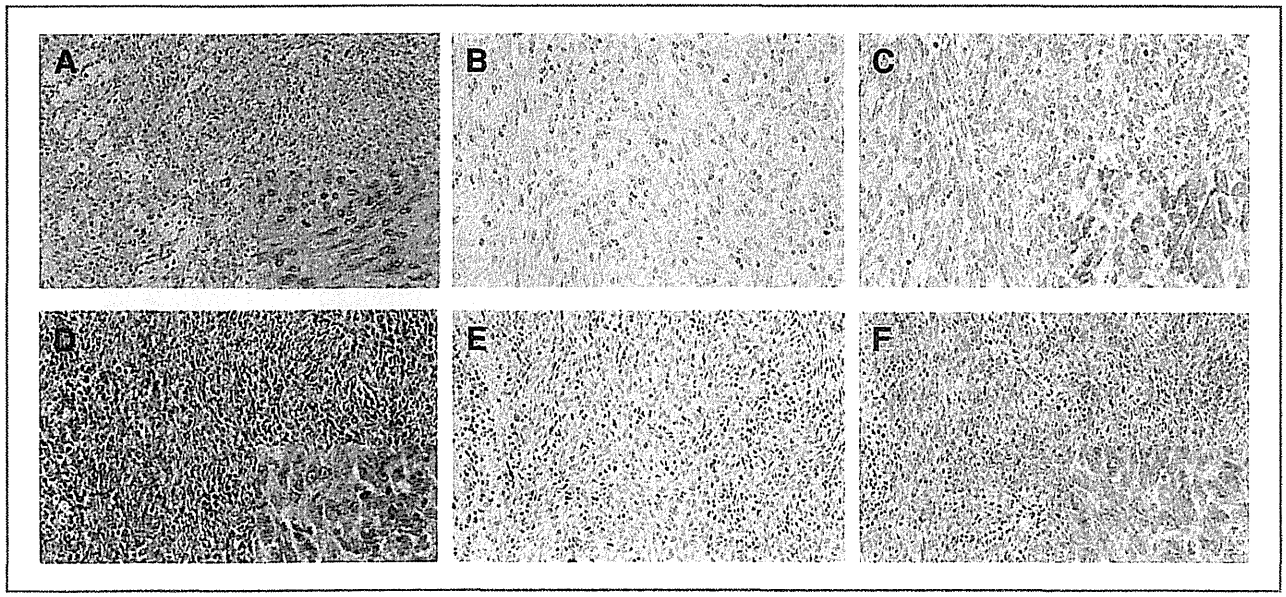

Figure 1. Histopathology of PPFIBP1–ALK-positive IMT. Diffuse proliferation of histiocytoid spindle cells showing a fascicular or storiform pattern. The tumor cells were uniform and had pale eosinophilic cytoplasm and an oval vesicular nucleus, within which a small nucleolus was centrally located. Mild inflammatory infiltrate containing lymphocytes, plasma cells, and foamy histiocytes is observed (A and D). The tumor cells were negative for ALK with conventional anti-ALK immunohistochemistry (B and E), but were clearly positive for ALK when the iAEP method was used. The staining pattern is diffuse cytoplasmic (C and F). Case 1 (A–C), Case 2 (D–F).