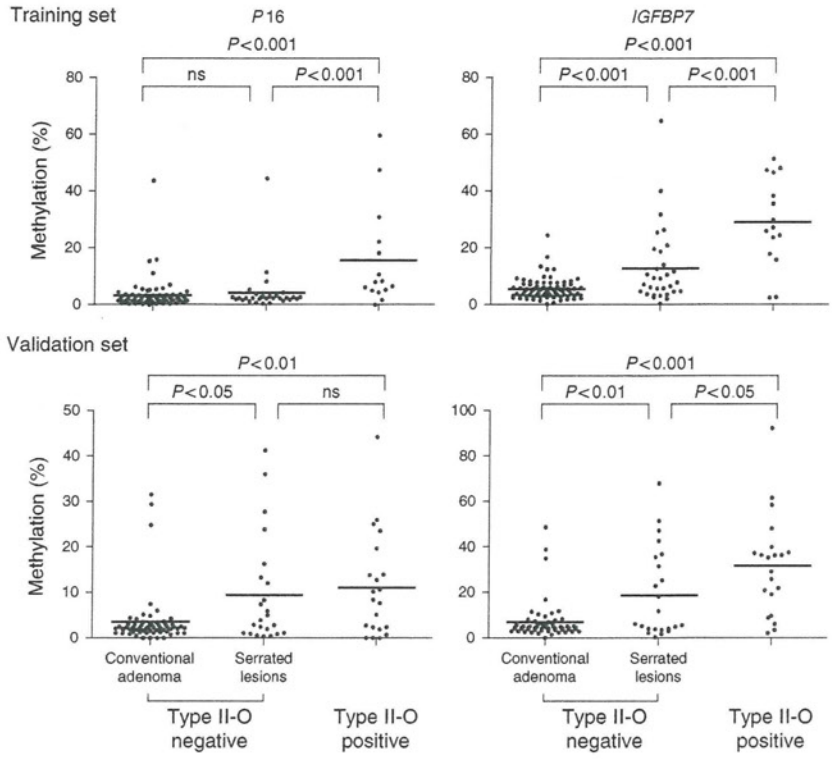

Figure 2. Increased CpG island methylation in Type II open-shape (Type II-O) pit-positive lesions. All of the Type II-O pit-positive specimens are serrated lesions. Type II-O pit-negative specimens contain serrated lesions and conventional adenomas. Methylation levels of *p16* and *IGFBP7* were obtained by bisulfite pyrosequencing. Each point represents an individual specimen; the horizontal bars represent the respective means.

show somewhat higher frequencies of *BRAF* mutation and CIMP, they never showed MSI (**Figure 3e**, **Supplementary Table 2**). Thus, Type II-O pit-positive and -negative serrated lesions appear to develop through distinct tumorigenic pathways.

To develop an efficient diagnostic method for detecting precursors of *BRAF* mutant and CIMP-positive CRCs, we next constructed a diagnostic tree to classify serrated lesions on the basis of their pit patterns (**Figure 4a**). We first excluded lesions containing only advanced pits (Type III or more) because they are unlikely to be SSAs. We next divided the lesions according to whether they were Type II-O pit-positive or -negative as the second node of the diagnostic tree. Finally, we subdivided the lesions into four groups according to the coexisting advanced pits (**Figure 4a; Supplementary Table 3**). As shown in **Figure 4a**, group 1 (Type II-O plus Type III, IV or V) exhibited significant specificity for *BRAF* mutation and CIMP-positive lesions. To evaluate the performance of the diagnostic tree in defining *BRAF*-mutant plus CIMP-positive lesions, a receiver operating characteristic curve was constructed by plotting the sensitivity over one specificity at three cut-offs for each group (group 1 vs. groups 2–4; groups 1–2 vs. groups 3–4; groups 1–3 vs. group 4). Areas under the curves in both the training set and the validation set were very high (for training set, 0.987; for validation set, 0.846).

We summarize our findings in **Figure 5**. We found that the Type II-O pit pattern is a clinically useful hallmark of SSAs. Moreover, during the progression of carcinogenesis, SSAs undergo additional molecular and histological changes, and our results demonstrate that colonoscopic observations can be used to predict these changes.

DISCUSSION

SSAs were previously reported to be precursors of CRCs with MSI (13,14). Although identification of SSAs in screening colonoscopies has important implications for the prevention and early detection of CRCs, the colonoscopic findings for SSAs have been less than definitive. Criteria for colonoscopic diagnosis of colorectal lesions were established on the basis of histological findings; however, there is no clear histological definition to distinguish SSAs from HPs, and the rate of concordance among gastrointestinal pathologists for diagnosis of SSA is low (10). Detection of SSAs also reportedly differs among endoscopists, and classification of HP and SSA differs among pathologists (12).

In the present study, we performed an integrated analysis of the genetic, epigenetic and clinical features of serrated lesions in an effort to establish more definitive criteria for the colonoscopic diagnosis of SSA. We identified an SSA-specific pit pattern Type II-O and prospectively validated its clinical utility for detecting SSA. Molecular dissection of mixed serrated lesions suggested that subcomponents with advanced pit patterns were derived from coexisting Type II-O pit-positive SSAs, and additional molecular

Table 2. Clinicopathological features of the patients

	Training Set			Validation Set		
	Type II-O			Type II-O		
	Positive	Negative	*P*-value	Positive	Negative	*P*-value
Age (mean±s.d.)	71.4±7.84	68.7±10.09	NS	65.48±8.77	64.05±10.72	NS
Sex						
F (%)	9 (60)	29 (27.1)	0.013	10 (47.62)	34 (40.96)	NS
M (%)	6 (40)	78 (72.9)		11 (52.38)	49 (59.04)	
Location						
Right (%)	14 (93.33)	43 (40.19)	<0.001	18 (85.71)	44 (53.01)	0.002
Left (%)	1 (6.67)	28 (26.17)		3 (14.46)	27 (32.53)	
Rectum (%)	0 (0)	36 (33.64)		0 (0)	12 (14.46)	
Morphology						
Flat (%)	7 (46.67)	49 (45.79)	NS	15 (71.43)	43 (51.81)	NS
Flat+protruded (%)	6 (40)	13 (12.15)		2 (9.52)	4 (5.82)	
Protruded (%)	2 (13.33)	45 (42.06)		4 (19.05)	36 (43.37)	
Size (mm, mean±s.d.)	11.00±4.02	10.45±8.56	NS	11.71±6.02	13.83±10.73	NS
Pit pattern						
Type II (%)		19 (19.92)			16 (19.28)	
Type II+Type IV (%)		13 (12.26)			6 (7.23)	
Type IV (%)		53 (50)			34 (40.96)	
Type III (%)		21 (19.81)			26 (31.33)	
Type III+Type IV (%)		0 (0)			1 (1.2)	
Type II-O (%)	6 (40.00)			16 (76.2)		
Type II-O+Type III (%)	0 (0)			2 (9.52)		
Type II-O+Type IV (%)	7 (46.67)			2 (9.52)		
Type II-O+Type V (%)	2 (13.33)			1 (4.76)		
Histology						
HP/IM (%)	0 (0)	14 (13.08)	<0.001	1(4.76)	4 (4.82)	<0.001
HP/IM+Ad-C (%)	0 (0)	0 (0)		0 (0)	2 (2.40)	
HP/IM+TSA (%)	0 (0)	9 (8.41)		1(4.76)	2 (2.41)	
TSA (%)	0 (0)	6 (5.61)		0 (0)	2 (2.41)	
Ad+se (%)	0 (0)	2 (1.87)		0 (0)	2 (2.41)	
Ad (%)	0 (0)	76 (71.03)		0 (0)	61 (73.49)	
SSA (%)	8 (53.33)	0 (0)		17 (80.95)	8 (9.64)	
SSA+Ad-C (%)	4 (26.67)	0 (0)		1 (4.76)	2 (2.41)	
SSA+HGD (%)	2 (13.33)	0 (0)		1 (4.76)	0 (0)	
SSA+TSA (%)	1 (6.67)	0 (0)		0 (0)	0 (0)	

Ad, conventional adenoma; Ad-C, adenomatous change; Ad+se, conventional adenoma with serration; HGD, high grade dysplasia; HP/IM, hyperplastic polyp/intermediate; NS, not significant; SSA, sessile serrated adenoma; TSA, traditional serrated adenoma; Type II-O, Type II open-shape.

alterations (e.g., *p16* and *MLH1* methylation) were acquired during tumorigenesis. The inactivation of cell cycle regulatory genes, including *p16* and *p53*, has an important role in CRC development by enabling tumor cells to escape oncogene-induced cellular senescence (28,29). *IGFBP7* has been shown to have a central role in oncogenic *BRAF*-induced senescence, and it is also a direct

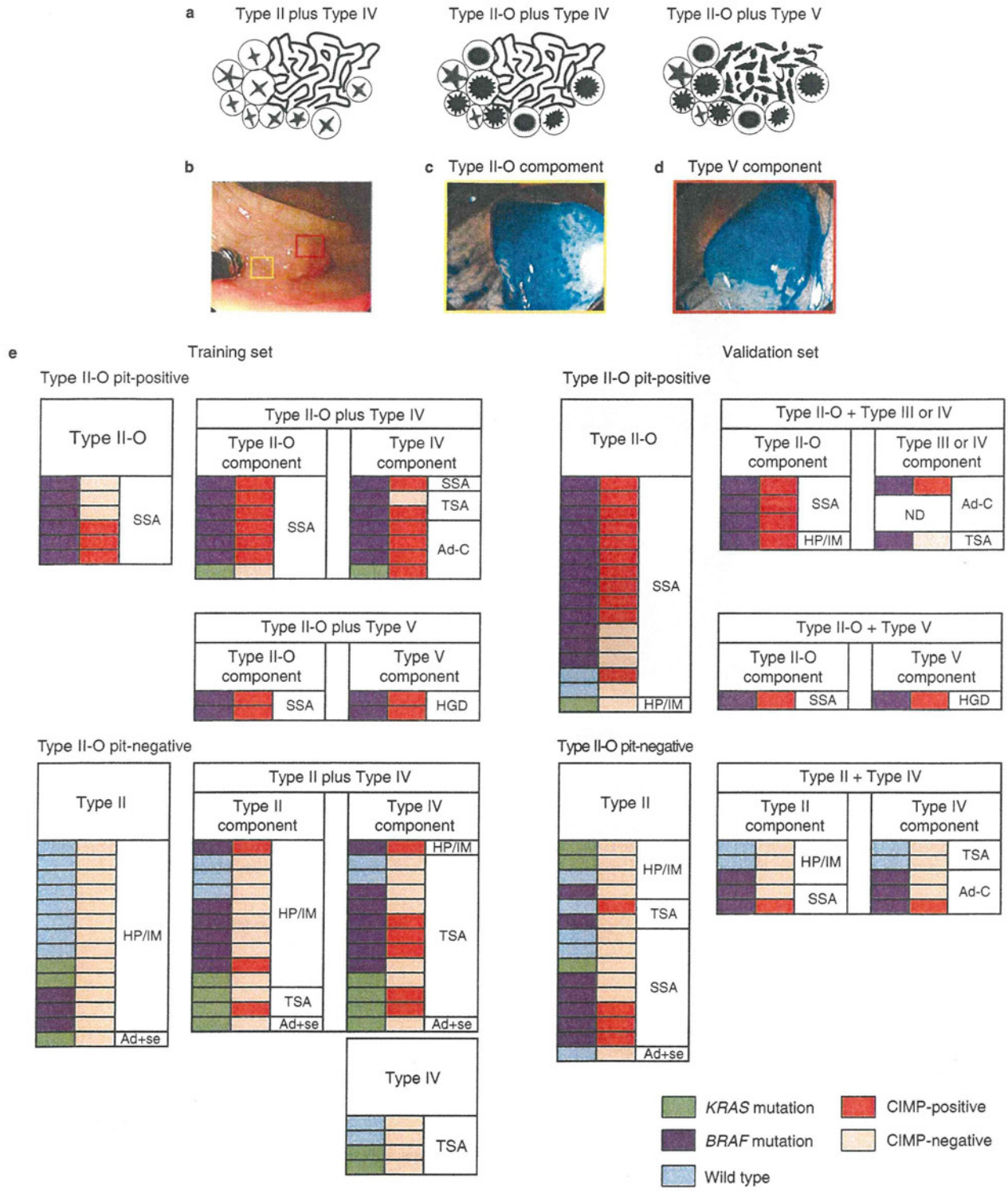

Figure 3. Morphological and molecular signatures reveal progression of sessile serrated adenomas (SSAs) with Type II open-shape (Type II-O pits). (**a**) Schematic diagram of serrated lesions with mixed pit patterns. (**b**) Colonoscopic view of a representative serrated lesion with Type II-O and Type V pits. (**c**, **d**) Magnified views of the Type II-O subcomponent (**c**) indicated by the yellow box (**b**) and the Type V subcomponent (**d**) indicated by the red box (**b**). (**e**) Summary of the molecular, colonoscopic and histological features of serrated lesions with or without Type II-O pits. Each row indicates one lesion. Results of the training set are on the left, and those of the validation set are on the right. ND, not determined.

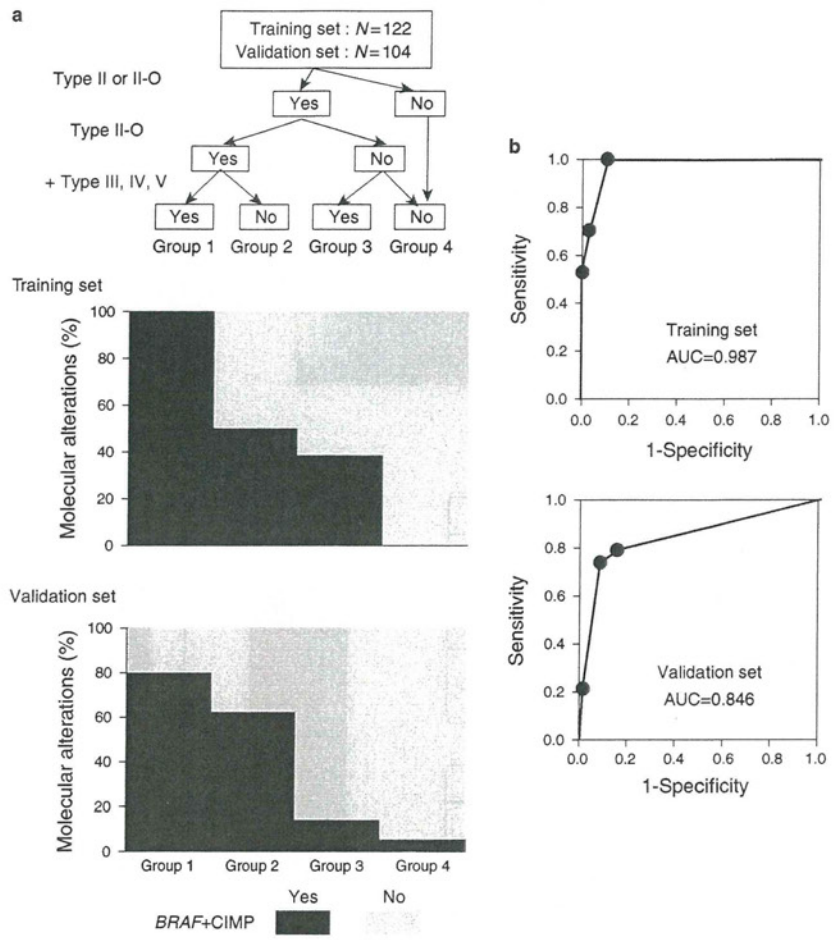

Figure 4. Clinical usefulness of the Type II open-shape (Type II-O) pit pattern for defining sessile serrated adenomas (SSAs). (**a**) A diagnostic tree to identify SSAs with *BRAF* mutation and CpG island methylator phenotype (CIMP). Serrated lesions are classified into four groups according to their pit patterns. The ratios of *BRAF*-mutant plus CIMP-positive lesions in the respective groups of the training and validation sets are shown below. (**b**) Receiver operating characteristic (ROC) curve analysis of the performance of the diagnostic tree. Sensitivities and specificities for definition of *BRAF*-mutant plus CIMP-positive lesions were calculated at three cutoffs for each group (group 1 vs. groups 2–4; groups 1–2 vs. groups 3–4; groups 1–3 vs. group 4).

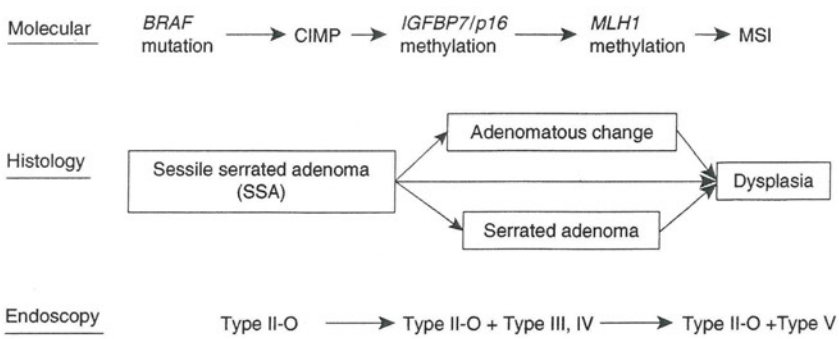

Figure 5. Progression of the molecular, histological and colonoscopic features in the serrated neoplastic pathway.

target of p53 (27,30). In CRCs, methylation of *IGFBP7* is strongly associated with *BRAF* mutation, the lack of *p53* mutation and the presence of MSI and CIMP (27). In the present study, we showed that levels of *IGFBP7* methylation are elevated in Type II-O pit-positive serrated lesions, suggesting the tumor suppressor function of p53 is attenuated at an early stage of the MSI pathway.

The mechanism underlying the strong relationship between Type II-O pits and SSAs remains unclear. SSAs exhibit a variety of histological features, including exaggerated serration, boot-shaped crypts, and the branching and dilatation of the crypts (8,15,31). These features are usually observed near the base of the crypts, making it difficult to endoscopically discriminate SSAs from HPs. SSAs are often covered by abundant mucus production, and accumulation of the mucin within crypts may lead to their dilatation (31). It is thus conceivable that overproduction of mucin may be associated with the Type II-O pit pattern in SSAs.

Identification of SSAs in screening colonoscopies has great significance for the prevention and surveillance of CRCs, as it is generally accepted that colonoscopy and polypectomy reduce the incidence of CRCs (32–35). Although SSAs are usually treated as conventional adenomas, it is still uncertain whether all SSAs should be endoscopically resected (16,36). In the present study, a large majority of Type II-O-positive lesions were histologically SSAs, and were tightly associated with *BRAF* mutation and CIMP, suggesting these lesions are appropriate targets for endoscopic resection. Moreover, as serrated lesions with Type II-O plus Type III, IV or V pits exhibit additional malignant potential, they too should be targets of early treatment. On the other hand, we noted that approximately half of the serrated lesions with only conventional Type II pits in our validation set were diagnosed as SSAs, though all Type II lesions in the training set were HP/IMs. This difference reflects the extremely high specificity and relatively low sensitivity of Type II-O pits for distinguishing SSAs from HP/IMs (**Table 1**).

There are several possible explanations for the low sensitivity of Type II-O pits for definition of SSAs. First, SSAs with only conventional Type II pits in our validation set may have been at a very early stage of development, where Type II-O pits were not yet established. Alternatively, it may simply reflect the difficulty of histologically distinguishing SSAs from HP/IMs. Of these two possibilities, the first is supported by the fact that although *BRAF* was frequently mutated in these lesions, aberrant DNA methylation was infrequent and only a limited number of the specimens were CIMP-positive. Further prospective study will be needed to unravel the time course of the emergence of Type II-O pits in the serrated pathway.

There are several potential limitations to the current study. First, our diagnostic system is based on endoscopic observations, so it may be affected by the skills of the endoscopists. As described above, among the ten endoscopists participating in this study, four had experience with more than 1,000 endoscopy cases, whereas the remaining six were less experienced. We therefore divided the data into two parts according the endoscopists' experience and compared the diagnostic accuracy of Type II-O pits for defining SSAs. As shown in **Supplementary Table 4**, the sensitivities and specificities were similar between the two groups, suggesting the intra-observer variability in our diagnostic system is relatively limited. Second, there are several biases in the selection of tumor specimens in this study. For instance, the average size of the lesions is relatively large (>10 mm). This is because these specimens were collected through endoscopic resection, and smaller lesions, which are usually subject to follow-up observation, were less likely to be included in this study. In addition, during the collection of the training set specimens, we treated as many possible serrated lesions as possible, which might have resulted in a somewhat high frequency of TSAs in our training set. Moreover, the identification of Type II-O pits enabled us to distinguish non-neoplastic HP/IM from neoplastic serrated lesions, which might have reduced the frequency of HP/IM in the validation set. Thus, our findings should be validated in an independent multicenter study that includes a larger number of samples. Third, our study does not provide effective criteria to endoscopically define TSAs. Further study may enable us to find new clues to refine the endoscopic diagnosis of serrated lesions.

In summary, we have identified a novel surface microstructure that is specific to SSAs. Recent studies have shown that the presence of SSAs is associated with an increased risk of synchronous CRCs (37–40). In the context of those findings, our observations indicate detection of Type II-O pits could be predictive of CRC risk. Thus, more intensive and frequent colonoscopic surveillance may be appropriate for patients in whom Type II-O-positive lesions were once detected. Not only could these findings contribute to the prevention of MSI-positive CRCs, they also demonstrate that integrative analysis of molecular and endoscopic characteristics greatly improves our understanding of the pathogenesis of CRC and the quality of colonoscopic surveillance.

ACKNOWLEDGMENT

The authors thank Dr William F. Goldman for editing the manuscript.

CONFLICT OF INTEREST

Guarantor of the article: Hiromu Suzuki, MD, PhD.

Specific author contributions: T. Kimura, E. Yamamoto, H. Yamano and M. Toyota conceived the study; E. Yamamoto, H. Suzuki, K. Imai, Y. Shinomura and M. Toyota designed the study; E. Yamamoto and H. Suzuki wrote the manuscript; T. Kimura, H. Yamano, K. Yoshikawa, R. Takagi, R. Kato, T. Harada and R. Suzuki carried out material support; T. Sugai performed histological analysis; M. Nojima performed statistical analysis; S. Kamimae, T. Sawada, M. Ashida, R. Maruyama, M. Kai and T. Sugai carried out the experiments; E. Yamamoto, H. Suzuki, M. Nojima performed data analysis.

Financial support: This study was supported in part by Grants-in-Aid for Scientific Research in Priority Areas (M. Toyota and K. Imai), a Grant-in-Aid for the Third-term Comprehensive 10-year Strategy for Cancer Control (M. Toyota), a Grant-in-Aid for Cancer Research from the Ministry of Health, Labor and Welfare, Japan (M. Toyota), and a Research Grant from the Princess Takamatsu Cancer Research Fund 09-24119 (M. Toyota).

Potential competing interest: None.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Recent evidence suggests that sessile serrated adenomas (SSAs) are precursor lesions for colorectal cancers (CRCs) with microsatellite instability (MSI).
- ✓ Pit pattern analysis using magnifying colonoscopy is an effective method of distinguishing malignant from benign lesions.

WHAT IS NEW HERE

- ✓ The Type II open-shape pit pattern (Type II-O) is a novel surface microstructure specific to SSAs.
- ✓ The presence of the Type II-O pit pattern is a hallmark of the premalignant stage of MSI and CpG island methylator phenotype (CIMP)-positive CRCs.
- ✓ Our findings will improve the efficacy of colonoscopic surveillance to prevent CRCs, especially those with MSI and CIMP.

REFERENCES

1. Vogelstein B, Fearon ER, Hamilton SR *et al.* Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525–32.
2. Jones PA, Baylin SB. The epigenomics of cancer. Cell 2007;128:683–92.
3. Toyota M, Ahuja N, Ohe-Toyota M *et al.* CpG island methylator phenotype in colorectal cancer. Proc Natl Acad Sci USA 1999;96:8681–6.
4. Toyota M, Ohe-Toyota M, Ahuja N *et al.* Distinct genetic profiles in colorectal tumors with or without the CpG island methylator phenotype. Proc Natl Acad Sci USA 2000;97:710–5.
5. Weisenberger D, Siegmund K, Campan M *et al.* CpG island methylator phenotype underlies sporadic microsatellite instability and is tightly associated with BRAF mutation in colorectal cancer. Nat Genet 2006;38:787.
6. Shen L, Toyota M, Kondo Y *et al.* Integrated genetic and epigenetic analysis identifies three different subclasses of colon cancer. Proc Natl Acad Sci USA 2007;104:18654.
7. Longacre TA, Fenoglio-Preiser CM. Mixed hyperplastic adenomatous polyps/serrated adenomas. A distinct form of colorectal neoplasia. Am J Surg Pathol 1990;14:524–37.
8. Torlakovic E, Skovlund E, Snover D *et al.* Morphologic reappraisal of serrated colorectal polyps. Am J Surg Pathol 2003;27:65–81.
9. Snover D, Jass J, Fenoglio-Preiser C *et al.* Serrated polyps of the large intestine: a morphologic and molecular review of an evolving concept. Am J Clin Pathol 2005;124:380–91.
10. Farris A, Misdraji J, Srivastava A *et al.* Sessile serrated adenoma: challenging discrimination from other serrated colonic polyps. Am J Surg Pathol 2008;32:30–5.
11. Torlakovic E, Gomez J, Driman D *et al.* Sessile serrated adenoma (SSA) vs. traditional serrated adenoma (TSA). Am J Surg Pathol 2008;32:21–9.
12. Hetzel J, Huang C, Coukos J *et al.* Variation in the detection of serrated polyps in an average risk colorectal cancer screening cohort. Am J Gastroenterol 2010;105:2656–64.
13. Kambara T, Simms LA, Whitehall VL *et al.* BRAF mutation is associated with DNA methylation in serrated polyps and cancers of the colorectum. Gut 2004;53:1137–44.
14. Spring KJ, Zhao ZZ, Karamatic R *et al.* High prevalence of sessile serrated adenomas with BRAF mutations: a prospective study of patients undergoing colonoscopy. Gastroenterology 2006;131:1400–7.
15. Jass JR, Baker K, Zlobec I *et al.* Advanced colorectal polyps with the molecular and morphological features of serrated polyps and adenomas: concept of a 'fusion' pathway to colorectal cancer. Histopathology 2006;49:121–31.
16. Leggett B, Whitehall V. Role of the serrated pathway in colorectal cancer pathogenesis. Gastroenterology 2010;138:2088–100.
17. Boparai KS, van den Broek FJ, van Eeden S *et al.* Hyperplastic polyposis syndrome: a pilot study for the differentiation of polyps by using high-resolution endoscopy, autofluorescence imaging, and narrow-band imaging. Gastrointest Endosc 2009;70:947–55.
18. Kudo S, Tamura S, Nakajima T *et al.* Diagnosis of colorectal tumorous lesions by magnifying endoscopy. Gastrointest Endosc 1996;44:8–14.
19. Hurlstone DP, Cross SS, Adam I *et al.* Efficacy of high magnification chromoscopic colonoscopy for the diagnosis of neoplasia in flat and depressed lesions of the colorectum: a prospective analysis. Gut 2004;53:284–90.
20. Kudo S, Lambert R, Allen J *et al.* Nonpolypoid neoplastic lesions of the colorectal mucosa. Gastrointest Endosc 2008;68:S3–47.
21. Matsuda T, Fujii T, Saito Y *et al.* Efficacy of the invasive/non-invasive pattern by magnifying chromoendoscopy to estimate the depth of invasion of early colorectal neoplasms. Am J Gastroenterol 2008;103:2700–6.
22. Morita T, Tamura S, Miyazaki J *et al.* Evaluation of endoscopic and histopathological features of serrated adenoma of the colon. Endoscopy 2001;33:761–5.
23. Oka S, Tanaka S, Hiyama T *et al.* Clinicopathologic and endoscopic features of colorectal serrated adenoma: differences between polypoid and superficial types. Gastrointest Endosc 2004;59:213–9.
24. Toyota M, Suzuki H, Sasaki Y *et al.* Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. Cancer Res 2008;68:4123–32.
25. Sugai T, Habano W, Endoh M *et al.* Molecular analysis of gastric differentiated-type intramucosal and submucosal cancers. Int J Cancer 2010;127:2500–9.
26. Boland CR, Thibodeau SN, Hamilton SR *et al.* A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998;58:5248–57.
27. Suzuki H, Igarashi S, Nojima M *et al.* IGFBP7 is a p53-responsive gene specifically silenced in colorectal cancer with CpG island methylator phenotype. Carcinogenesis 2010;31:342–9.
28. Serrano M, Lin AW, McCurrach ME *et al.* Oncogenic ras provokes premature cell senescence associated with accumulation of p53 and p16INK4a. Cell 1997;88:593–602.
29. Collado M, Serrano M. The power and the promise of oncogene-induced senescence markers. Nat Rev Cancer 2006;6:472–6.
30. Wajapeyee N, Serra RW, Zhu X *et al.* Oncogenic BRAF induces senescence and apoptosis through pathways mediated by the secreted protein IGFBP7. Cell 2008;132:363–74.
31. Higuchi T, Sugihara K, Jass JR. Demographic and pathological characteristics of serrated polyps of colorectum. Histopathology 2005;47:32–40.
32. Winawer SJ, Zauber AG, Ho MN *et al.* Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. N Engl J Med 1993;329:1977–81.
33. Citarda F, Tomaselli G, Capocaccia R *et al.* Efficacy in standard clinical practice of colonoscopic polypectomy in reducing colorectal cancer incidence. Gut 2001;48:812–5.
34. Brenner H, Chang-Claude J, Seiler CM *et al.* Does a negative screening colonoscopy ever need to be repeated? Gut 2006;55:1145–50.
35. Winawer S, Zauber A, Fletcher R *et al.* Guidelines for colonoscopy surveillance after polypectomy: a consensus update by the US Multi-Society Task Force on Colorectal Cancer and the American Cancer Society. Gastroenterology 2006;130:1872–85.
36. Jass J. Gastrointestinal polyposes: clinical, pathological and molecular features. Gastroenterol Clin North Am 2007;36:927 46, viii.
37. Li D, Jin C, McCulloch C *et al.* Association of large serrated polyps with synchronous advanced colorectal neoplasia. Am J Gastroenterol 2009;104:695.
38. Hiraoka S, Kato J, Fujiki S *et al.* The presence of large serrated polyps increases risk for colorectal cancer. Gastroenterology 2010;139: 1503,10, 1510.e1-3.
39. Schreiner MA, Weiss DG, Lieberman DA. Proximal and large hyperplastic and nondysplastic serrated polyps detected by colonoscopy are associated with neoplasia. Gastroenterology 2010;139:1497–502.
40. Boparai K, Mathus-Vliegen EMH, Koornstra J *et al.* Increased colorectal cancer risk during follow-up in patients with hyperplastic polyposis syndrome: a multicentre cohort study. Gut 2010;59:1094–100.

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

Molecular and Cellular Pathobiology

Cancer Research

Upregulation of miR-196a and *HOTAIR* Drive Malignant Character in Gastrointestinal Stromal Tumors

Takeshi Niinuma[1], Hiromu Suzuki[1,3], Masanori Nojima[4], Katsuhiko Nosho[1], Hiroyuki Yamamoto[1], Hiroyuki Takamaru[1], Eiichiro Yamamoto[3], Reo Maruyama[3], Takayuki Nobuoka[2], Yasuaki Miyazaki[9], Toshirou Nishida[9,10], Takeo Bamba[11], Tatsuo Kanda[11], Yoichi Ajioka[12], Takahiro Taguchi[13], Satoshi Okahara[7], Hiroaki Takahashi[7], Yasunori Nishida[8], Masao Hosokawa[8], Tadashi Hasegawa[5], Takashi Tokino[6], Koichi Hirata[2], Kohzoh Imai[14], Minoru Toyota[3], and Yasuhisa Shinomura[1]

Abstract

Large intergenic noncoding RNAs (lincRNA) have been less studied than miRNAs in cancer, although both offer considerable theranostic potential. In this study, we identified frequent upregulation of miR-196a and lincRNA *HOTAIR* in high-risk gastrointestinal stromal tumors (GIST). Overexpression of miR-196a was associated with high-risk grade, metastasis and poor survival among GIST specimens. miR-196a genes are located within the *HOX* gene clusters and microarray expression analysis revealed that the *HOXC* and *HOTAIR* gene were also coordinately upregulated in GISTs which overexpress miR-196a. In like manner, overexpression of *HOTAIR* was also strongly associated with high-risk grade and metastasis among GIST specimens. RNA interference–mediated knockdown of *HOTAIR* altered the expression of reported *HOTAIR* target genes and suppressed GIST cell invasiveness. These findings reveal concurrent overexpression of *HOX* genes with noncoding RNAs in human cancer in this setting, revealing miR-196a and *HOTAIR* as potentially useful biomarkers and therapeutic targets in malignant GISTs. *Cancer Res; 72(5); 1126–36. ©2012 AACR.*

Introduction

Gastrointestinal stromal tumors (GIST) are the most common mesenchymal tumors of the gastrointestinal tract (1–3). GISTs arise predominantly in the stomach (60%) and small intestine (25%) but also occur in colon and rectum (5%), esophagus (2%), and other organs (3). Immunohistochemically, GISTs are positive for KIT and CD34 and are negative or variably positive for other neural and smooth muscle cell markers. The expression of KIT and CD34 is a characteristic feature of the intestinal cells of Cajal (ICC), which are located in the intestinal wall and regulate gastrointestinal motility. GISTs are thus thought to originate from ICCs or ICC precursors. Activating *KIT* mutations have been identified in 80% to 90% of GISTs, and mutation of the platelet-derived growth factor receptor alpha gene (*PDGFRA*) is observed in approximately 5% of GISTs (1–3). In that context, imatinib mesylate (formerly STI571) was developed as a tyrosine kinase inhibitor and has been shown to inhibit the activities of BCR-ABL, KIT, and PDGFR. Imatinib mesylate is currently being used for the treatment of both chronic myeloid leukemia and metastatic GISTs.

Predicting the biologic potential of GISTs is often difficult, and considerable effort has been made to define the variables that could enable more accurate identification of tumors with malignant potential. In most classification systems, the key prognostic factors for estimating malignant potential are tumor size and mitotic rate, and to a more variable degree, the proliferation index or tumor site (4). Other potential and promising markers of GIST malignancy are molecular alterations. As mentioned, a large majority of GISTs exhibit activating *KIT* or *PDGFRA* mutations. By itself, however, mutation status does not fully explain the diverse biology of GISTs, and it is believed that additional molecular alterations are required for the progression of high-risk GISTs. For instance, expression of CD26 (encoded by *DPP4*) is strongly associated with poor survival among patients with gastric GISTs, suggesting its involvement in the malignant progression of the disease (5).

Authors' Affiliations: First Departments of [1]Internal Medicine and [2]Surgery, Departments of [3]Molecular Biology, [4]Public Health, [5]Surgical Pathology, [6]Medical Genome Science, Research Institute for Frontier Medicine, Sapporo Medical University School of Medicine; Departments of [7]Gastroenterology and [8]Surgery, Keiyukai Sapporo Hospital, Sapporo; [9]Department of Surgery, Osaka University Graduate School of Medicine; [10]Department of Surgery, Osaka Police Hospital, Osaka; Divisions of [11]Digestive and General Surgery and [12]Molecular and Diagnostic Pathology, Niigata University Graduate School of Medical and Dental Sciences, Niigata; [13]Division of Human Health and Medical Science, Graduate School of Kuroshio Science, Kochi University, Nankoku; and [14]Division of Novel Therapy for Cancer, The Advanced Clinical Research Center, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan

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

T. Niinuma and H. Suzuki contributed equally to this work.

Corresponding Authors: Hiromu Suzuki, Department of Molecular Biology, Sapporo Medical University; S1, W17, Chuo-Ku, Sapporo 060-8556, Japan. Phone: 81-11-611-2111; Fax: 81-11-622-1918; E-mail: hsuzuki@sapmed.ac.jp; and Yasuhisa Shinomura, First Department of Internal Medicine, Sapporo Medical University; S1, W16, Chuo-Ku, Sapporo 060-8543, Japan. Fax: 81-11-611-2282; E-mail: shinomura@sapmed.ac.jp

doi: 10.1158/0008-5472.CAN-11-1803

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

In addition, we recently showed that hypomethylation of repetitive DNA elements is predominantly observed in malignant GISTs, and that global hypomethylation correlates with increased chromosomal aberration (6).

miRNAs are a class of small noncoding RNAs that regulate gene expression by inducing translational inhibition or direct degradation of target mRNAs through base pairing to partially complementary sites (7). miRNAs are highly conserved among species and play critical roles in a variety of biologic processes, including development, differentiation, cell proliferation, and apoptosis. Consistent with their role in these processes, a number of studies have shown widespread alteration of miRNA expression patterns in cancer (8, 9). It has also been shown that in cancer global miRNA expression profiles, as well as expression of specific miRNAs, correlate with disease prognosis and clinical outcome (10). To date, however, only a few groups have studied miRNA expression in GISTs (11, 12), and no specific miRNAs that could serve as prognostic markers have yet been identified.

In this study, we investigated the global pattern of miRNA expression in GISTs. Our aim was to evaluate the contribution made by miRNAs to the malignant potential of GISTs and to identify predictive biomarkers. We determined that upregulation of miR-196a is strongly associated with high risk and poor prognosis in GIST patients. Furthermore, we provide evidence that overexpression of miR-196a is accompanied by upregulation of *HOXC* cluster genes and a metastasis-associated noncoding RNA in GISTs.

Materials and Methods

Tumor samples

A total of 56 fresh frozen GIST specimens were obtained from Sapporo Medical University Hospital, Keiyukai Sapporo Hospital, and Osaka University Hospital, as described (6). In addition, formalin-fixed paraffin-embedded (FFPE) tissue sections of 100 GIST specimens were obtained from Niigata University Hospital. Informed consent was obtained from all patients before collection of the specimens, and this study was approved by the respective Institutional Review Boards. Risk grade was assessed according to the risk definition system proposed by Fletcher and colleagues (4). Tumors that were less than 2 cm in diameter with a mitotic count of less than 5/50 high-power fields (HPF) were categorized as very low risk. Tumors that were 2 to 5 cm in diameter with a mitotic count of less than 5/50 HPF were considered to be low risk. Tumors that were less than 5 cm in diameter with a mitotic count of 6 to 10/50 HPF, or were 5 to 10 cm with a mitotic count of less than 5/50 HPF were considered to be intermediate risk. Tumors that were more than 5 cm in diameter with a mitotic count of more than 5/50 HPF, more than 10 cm in diameter with any mitotic count, or any size with a mitotic count of more than 10/50 HPF were considered to be high risk. Total RNA was extracted from fresh frozen tissue specimens using a mirVana miRNA Isolation Kit (Ambion). Total RNA was extracted from FFPE tissue specimens using a RecoverAll Total Nucleic Acid Isolation Kit for FFPE (Ambion). Tumor tissues were reviewed by pathologists and were macrodissected; laser capture microdissection was not carried out in this study.

miRNA microarray analysis

One-color microarray-based miRNA expression analysis was carried out according to the manufacturer's instructions (Agilent Technologies). Briefly, 100 ng of total RNA from fresh frozen GIST tissues was labeled using miRNA Labeling Reagent (Agilent Technologies), after which the labeled RNA was hybridized to a Human miRNA Microarray V3 (Rel 12.0, G4470C; Agilent Technologies), which covers 859 human miRNAs and 80 viral miRNAs. The microarray data were analyzed using GeneSpring GX version 11 (Agilent Technologies). The normalized microarray data were then compared with the TaqMan assay results using GraphPad PRISM version 5 (GraphPad Software Inc.). The Gene Expression Omnibus accession number for the miRNA microarray data is GSE31741.

Quantitative RT-PCR of miRNA

miR-196a expression was analyzed using TaqMan microRNA Assays (Applied Biosystems). Briefly, 5 ng of total RNA were reverse transcribed using specific stem-loop RT primers, after which they were amplified and detected using PCR with specific primers and TaqMan probes. The PCR was run in triplicate using a 7500 Fast Real-Time PCR System (Applied Biosystems), and SDS v1.4 software (Applied Biosystems) was used for comparative ΔC_t analysis. U6 snRNA (RNU6B; Applied Biosystems) served as an endogenous control.

Gene expression microarray analysis

One-color microarray-based gene expression analysis was carried out according to the manufacturer's instructions (Agilent Technologies). Briefly, 700 ng of total RNA were amplified and labeled using a Quick Amp Labeling Kit One-Color (Agilent Technologies), after which the synthesized cRNA was hybridized to the Whole Human Genome Oligo DNA microarray, which includes 41,000 probe sets covering 19,416 genes (G4112F; Agilent Technologies). The microarray data were analyzed using GeneSpring GX version 11 (Agilent Technologies). The Gene Expression Omnibus accession numbers for the microarray data are GSE31802 and GSE32064.

Quantitative RT-PCR of *HOTAIR*

Single-stranded cDNA was prepared using SuperScript III reverse transcriptase (Invitrogen). Quantitative reverse transcriptase PCR (RT-PCR) of *HOTAIR* was carried out using a TaqMan Gene Expression Assay (Assay ID, Hs03296631_m1; Applied Biosystems) and a 7500 Fast Real-Time PCR System (Applied Biosystems). *GAPDH* (Assay ID, Hs99999905_m1; Applied Biosystems) served as an endogenous control.

DNA copy number and chromatin signature analysis

DNA copy number was analyzed using array-based comparative genome hybridization (CGH) as described previously (6). Trimethylated Histone H3 lysine 4 (H3K4me3) was analyzed using chromatin immunoprecipitation (ChIP) as described previously (13, 14). Details of the experimental procedures are provided in the Supplementary Methods.

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

Transfection of miRNA inhibitors and siRNA molecules

GIST-T1 cells were described previously (15). For inhibition of miR-196a, cells (3×10^5 cells in 6-well plates) were transfected with 100 pmol of Anti-miR miRNA Inhibitors (Ambion) or Anti-miR miRNA Inhibitors Negative Control #1 (Ambion) using Lipofectamine2000 (Invitrogen). For RNA interference (RNAi)-mediated knockdown of *HOTAIR*, 3 different Stealth siRNAs against *HOTAIR* were generated by Invitrogen, after which a mixture of the 3 was used for transfection. Cells (3×10^5 cells in 6-well plates) were transfected with 100 pmol of siRNA or with Stealth RNAi Negative Control Medium GC (Invitrogen) using Lipofectamine2000 (Invitrogen). Total RNA extraction, cell viability assays, and Matrigel invasion assays were carried out 48 hours after transfection as described in the Supplementary Methods.

Statistical analysis

All gene expression levels were log transformed for subsequent statistical analysis because the distribution of expression data seemed to follow a log-normal distribution. Geometric means were therefore calculated as summary statistics for expression levels. Comparisons of continuous variables were made using *t* tests or one-way ANOVA with post hoc multiple comparisons (Games–Howell test). Pearson's correlation coefficients were calculated to describe the strength of the correlation between 2 variables. Comparisons of categorical variables were made using Fisher exact test. To assess the association between prognostic factors and gene expression levels, logistic or Cox regression analyses were carried out. For these regression analyses, the most optimal cutoff points were employed to calculate ORs and HRs, with or without adjustment for clinical factors. Kaplan–Meier curves were plotted to compare 2 groups stratified by gene expression status. All statistical analyses were done using SPSS Statistics 18 (IBM Corporation).

Results

Detection of upregulated miR-196a expression in high-risk GISTs

To examine the miRNA expression signature in GISTs, we carried out miRNA microarray analysis with 32 fresh frozen GIST specimens (10 low-risk, 8 intermediate-risk, and 14 high-risk GISTs). The clinicopathologic features of the 32 patients are listed in Supplementary Table S1. Of 939 probe sets, 470 were excluded because of the absence of a detectable signal in any of the samples tested. Unsupervised hierarchical clustering using the remaining 469 probe sets revealed that GISTs in which there was abundant expression of miRNAs encoded on chromosome 14q32.31 form a separate cluster (Supplementary Figs. S1 and S2). Moreover, by comparing the miRNA expression profiles with array CGH results, we found that this cluster is enriched in tumors without 14q loss. These results are consistent with recent reports showing an inverse relationship between 14q loss and expression of miRNAs located on 14q in GISTs (11, 12), which is indicative of the reliability of our microarray analysis. We next carried out a scatter plot analysis and found that miR-196a is markedly upregulated in high-risk GISTs, as compared with low- or intermediate-risk GISTs (Fig. 1A). As shown in Fig. 1B, miR-196a was undetectable in all but one of the low- and intermediate-risk GISTs tested, whereas it was upregulated in more than half of the high-risk tumors. The elevated expression of miR-196a was observed in both gastric and small intestinal GISTs (Supplementary Table S1).

Upregulation of miR-196a is associated with GIST malignancy

To assess the clinical importance of miR-196a upregulation in GISTs, we next carried out TaqMan assay with 56 fresh frozen GIST specimens (discovery cohort), including the 32 specimens initially analyzed by microarray. The clinicopathologic features of the patients are summarized in Table 1. The TaqMan assay results were highly consistent with the microarray data, though the TaqMan assay did reveal low levels of miR-196a expression in samples in which there was no detectable signal from the microarray (Supplementary Fig. S3).

Also consistent with the microarray results was the finding that miR-196a was markedly upregulated in high-risk GISTs, as compared with the other groups ($P = 0.004$, one-way ANOVA; Fig. 1C, Supplementary Table S2). In addition, logistic regression analysis revealed that the association between miR-196a upregulation and the high-risk category was highest when we employed a cutoff value of miR-196a/U6 0.4 or more (OR = 13.7; 95% CI: 3.4–54.6; $P < 0.001$; Supplementary Table S3). Survival data were obtained for 32 patients, and Cox hazard analysis revealed the highest HR for patients with elevated miR-196a expression when a cutoff value of 1.4 was employed (Table 2). Kaplan–Meier analysis showed poor overall survival among patients with GISTs expressing high levels of miR-196a, though the effect was not statistically significant (Fig. 1D).

We next used TaqMan assay to analyze an independent validation cohort consisting of 100 FFPE GIST specimens (Table 1). Consistent with the findings summarized above, we observed that upregulation of miR-196a was associated with high-risk GISTs (Fig. 1C, Supplementary Tables S2 and S3). By using the same cutoff value (miR-196a/U6 $\geq$ 1.4), Cox hazard analysis revealed an elevated HR for patients exhibiting high levels of miR-196a expression (Table 2), and Kaplan–Meier analysis showed shorter survival times for the same patients (Fig. 1D). These results confirmed the prognostic value of miR-196a expression in both fresh frozen and FFPE GIST specimens.

Finally, we combined the GIST samples in the discovery and validation cohorts to examine the clinicopathologic significance of miR-196a. Expression of miR-196a correlated positively with high-risk grade (Fig. 1C, Supplementary Tables S2 and S3), poor clinical outcome (Fig. 1D, Table 2), tumor size, mitotic count, and metastasis (Table 3). Interestingly, although expression of miR-196a was not associated with age or gender, it was strongly associated with tumor location (Table 3). The median level of miR-196a expression was lowest in specimens from esophageal GISTs and then increased as the GIST site moved from the oral side toward the anal side of the gastrointestinal tract ($P < 0.001$; Table 3, Supplementary Fig. S4). Importantly, although the average level of miR-196a expression was higher in small intestine than in stomach, it was positively

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

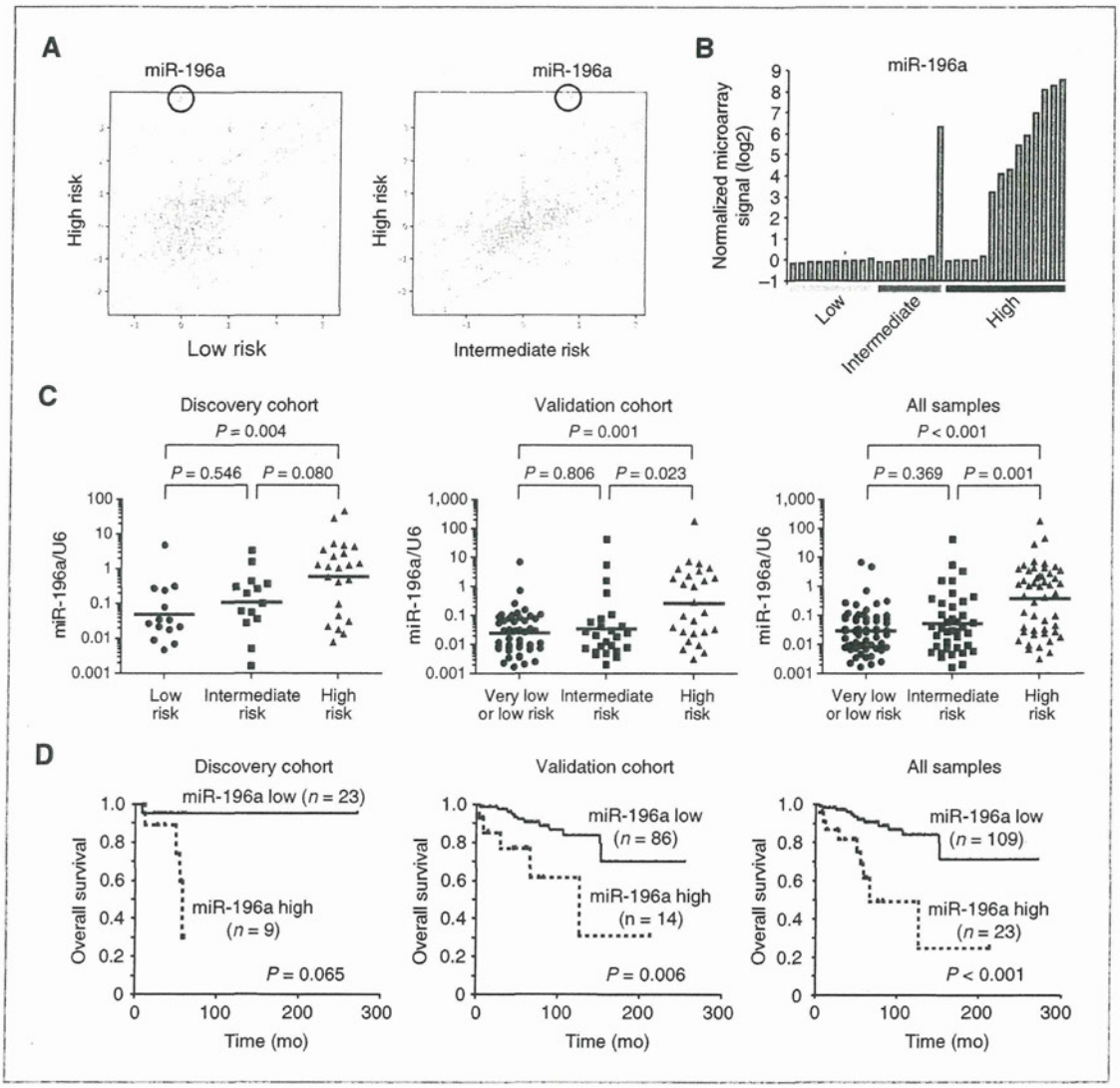

Figure 1. Upregulation of miR-196a expression in GISTs is associated with a high-risk grade and poor prognosis. A, scatter plots analyses: plotting low-risk ($n = 10$) versus high-risk GISTs ($n = 14$; left) and intermediate-risk ($n = 8$) versus high-risk GISTs ($n = 14$; right) revealed overexpression of miR-196a in high-risk GISTs. Microarray data are normalized and log transformed (base 2). Expression of miR-196a is highlighted by a circle. B, levels of miR-196a expression obtained from microarray analysis of 32 GIST specimens. Risk categories are indicated below. C, comparison of miR-196a expression using TaqMan assay in low- ($n = 14$), intermediate- ($n = 14$), and high-risk GISTs ($n = 23$) in a discovery cohort (left), very low- or low-risk ($n = 46$), intermediate-risk ($n = 25$), and high-risk GISTs ($n = 26$) in a validation cohort (middle) and all GIST specimens (right). Results are normalized to internal U6 snRNA. D, Kaplan–Meier curves showing the effect of miR-196a expression (high, miR-196a/U6 $\geq$ 1.4; low, miR-196a/U6 < 1.4) on overall survival in the discovery cohort (left), validation cohort (middle), and all GIST patients (right).

correlated with high-risk grade in both organs (Supplementary Table S4).

Concurrent upregulation of miR-196a and *HOX* cluster genes in GISTs

To analyze the relationship between miR-196a upregulation and the global gene expression profiles in GISTs, we selected age-, gender-, and tumor location–matched GIST specimens showing either low ($n = 7$; average miR-196a/U6 = 0.1) or high miR-196a expression ($n = 7$; average miR-196a/U6 = 15.7) and subjected them to gene expression microarray analysis (Supplementary Table S5). We found that for 4,947 probe sets (corresponding to 3,206 unique genes), there was more than a 2-fold difference in expression between GISTs with miR-196a

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

Table 1. Clinical features of the GIST samples used in this study

Discovery cohort	
Age (y, median ± SD)	68.0 ± 15.2
Gender	
Male	32 (57.1%)
Female	24 (42.9%)
Tumor location	
Stomach	40 (71.4%)
Small intestine	14 (25.0%)
Omentum	1 (1.8%)
Colorectum	1 (1.8%)
Risk category ($n = 51$)	
Low	14 (27.5%)
Intermediate	14 (27.5%)
High	23 (45.0%)
Validation cohort	
Age (y, median ± SD)	64.0 ± 12.4
Gender	
Male	44 (44.0%)
Female	56 (56.0%)
Tumor location	
Esophagus	5 (5.2%)
Stomach	84 (84.0%)
Small intestine	8 (8.0%)
Colorectum	3 (3.0%)
Risk category ($n = 97$)	
Very low	1 (1.0%)
Low	45 (46.4%)
Intermediate	25 (25.8%)
High	26 (26.8%)

overexpression and those without it. Hierarchical clustering analysis using the 4,947 probe sets clearly distinguished between tumors on the basis of their miR-196a expression status (Fig. 2A), and Gene Ontology analysis suggested that genes related to "immune system," "plasma membrane," and "cell communication" are strongly overrepresented among the differentially expressed genes (Supplementary Table S6).

To further characterize the differentially expressed genes, we carried out a gene set analysis and obtained the highest enrichment score with the *HOX* gene set (Supplementary Table S7). We found miR-196a to be encoded at 2 paralogous loci, *miR-196a-1* and *miR-196a-2*, which are located within the *HOXB* and *HOXC* clusters, respectively (Fig. 2B; ref. 16). Hierarchical clustering analysis using the expression data for *HOXC* genes clearly differentiated the GIST samples into 2 groups, and we observed perfect correspondence between higher expression of multiple *HOXC* genes and upregulation of miR-196a (Fig. 2C). By contrast, genes in other *HOX* clusters did not show such obvious correlations with miR-196a (Fig. 2C, Supplementary Fig. S5). We next compared the microarray signal for each *HOX* gene with the miR-196a expression and found strong positive correlations between the expression levels of a number of *HOXC* genes and those of miR-196a (Fig. 3D, Supplementary Fig. S6). Notably, we also found that expression of *HOTAIR*, which encodes a large intergenic noncoding RNA (lincRNA) and is located in an antisense orientation relative to the *HOXC* genes, is concurrently upregulated with miR-196a (Fig. 2C and D). Levels of miR-196a expression also correlated moderately with those of the *HOXB* genes neighboring *miR-196a-1* (*HOXB13* and *HOXB9*), but the correlations were less significant than those with *HOXC* genes (Supplementary Fig. S7).

The similarity between the expression patterns of *HOXC* genes and those of the noncoding RNAs encoded in the same locus is indicative of a common regulatory mechanism involved in the activation of these genes in GISTs. However, array CGH analysis of 27 GIST specimens failed to detect either gain or loss in any *HOX* loci, irrespective of *miR-196a* or *HOX* gene expression, which makes it unlikely that genomic amplification is the cause of their overexpression (Supplementary Fig. S8).

Upregulation of *HOTAIR* is associated with GIST malignancy

A recent study showed that *HOTAIR* is overexpressed in primary breast cancer and is associated with metastasis (17). To examine its clinical significance in GISTs, we carried out TaqMan assays of *HOTAIR* with the discovery cohort samples. We found that the microarray signals for *HOTAIR* were highly

Table 2. miR-196a expression is associated with poor clinical outcome in GIST patients

		Outcome					
	miR-196a/U6	Survival	Death	HR (95% CI)	*P*	HR[a] (95% CI)	*P*
Discovery cohort	<1.4	22	1				
	≥1.4	5	4	6.3 (0.7–57.5)	0.104	32.9 (2.0–551.3)	0.015
Validation cohort	<1.4	73	13				
	≥1.4	9	5	3.9 (1.4–11.1)	0.011	8.4 (2.6–26.9)	<0.001
All samples	<1.4	95	14				
	≥1.4	14	9	4.9 (2.1–11.7)	<0.001	9.1 (3.5–23.7)	<0.001

[a]Age and gender adjusted HR.

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

Table 3. Correlation between miR-196a expression and clinicopathologic features of GISTs

		miR-196a/U6				
	n	Geometric mean	95% CI	*P*		
Age (y)						
<65	76	0.093	0.051–0.168			
≥65	81	0.074	0.044–0.127	0.581		
Gender						
M	76	0.104	0.059–0.186			
F	80	0.069	0.040–0.119	0.297		
Location						
Esophagus	5	0.019	0.001–0.417	Ref		
Stomach	124	0.061	0.039–0.094	0.741	Ref	
Small intestine	22	0.395	0.163–0.957	0.161	0.002	Ref
Colorectum	4	4.936	2.564–9.502	0.023	<0.001	<0.001
				<0.001		
Tumor size (cm)						
≤5.0	70	0.045	0.026–0.075			
>5.0	81	0.118	0.066–0.210	0.016		
Mitotic count (/50 HPF)						
≤5	105	0.036	0.025–0.053			
>5	35	0.539	0.215–1.353	<0.001		
Metastasis						
Yes	28	0.747	0.307–1.819			
No	108	0.041	0.027–0.063	<0.001		

consistent with the TaqMan assay results (Supplementary Fig. S9). *HOTAIR* was upregulated exclusively in high-risk GISTs, as compared with low- or intermediate-risk GISTs ($P = 0.018$; Fig. 3A), and its expression correlated positively with the expression of miR-196a (Fig. 3B) and *HOXC* genes (Fig. 3C, Supplementary Fig. S10). In addition, logistic regression analysis revealed that high levels of *HOTAIR* expression in GISTs (*HOTAIR*/*GAPDH* ≥ 0.0002) were strongly associated with metastasis (age and gender adjusted OR = 8.2; 95% CI: 1.4–48.4; $P = 0.021$). Cox hazard analysis suggested an elevated HR for patients with high *HOTAIR* expression (Table 4), and Kaplan–Meier analysis showed poor overall survival for the same patients, though the effect was not statistically significant (Fig. 3D). We also tried to analyze *HOTAIR* expression in the FFPE specimens; however, we failed to detect expression of either *HOTAIR* or *GAPDH* in these samples, most likely due to an inadequate quality of the RNA.

Reduced expression of miR-196a and *HOTAIR* target genes in GISTs

To examine the functional role of miR-196a in GISTs, we interrogated our gene expression microarray data for miR-196a target genes computationally predicted by TargetScan. Of the 2,248 genes whose expression was reduced in GISTs overexpressing miR-196a, 95 corresponded to predicted targets (Supplementary Fig. S11, Supplementary Table S8). This gene list included *ANXA1* (Annexin A1), which is an experimentally validated miR-196a target gene (18). In addition, expression of several *HOX* genes, including *HOXB8*, was reduced in GISTs overexpressing miR-196a, which is consistent with an earlier finding of miR-196a–directed cleavage of *HOXB8* mRNA (Supplementary Fig. S11; ref. 19).

In normal human fibroblasts, *HOTAIR* represses *HOXD* gene expression by interacting with polycomb repressive complex 2 (PRC2; ref. 20). In breast cancer cells, overexpression of *HOTAIR* was shown to recruit PRC2 to more than 800 gene promoters, leading to histone H3K27 methylation and epigenetic silencing of the target genes (17). We therefore examined our microarray data for the reported *HOTAIR*-induced PRC2 target genes. Among 14 GISTs analyzed with the microarray, all 7 tumors strongly expressing miR-196a showed elevated *HOTAIR* expression (average *HOTAIR*/*GAPDH* = 0.00254), whereas all tumors only weakly expressing miR-196a showed little or no *HOTAIR* expression (average *HOTAIR*/*GAPDH* = 0.00001). We found that expression of 144 *HOTAIR* target genes was reduced in GISTs overexpressing *HOTAIR* (Supplementary Fig. S11, Supplementary Table S9). These results indicated that overexpression of miR-196a and *HOTAIR* may contribute to the malignant progression of GISTs by modulating expression of their target genes.

Inhibition of *miR-196a* and *HOTAIR* suppresses GIST cell invasion

We next utilized a cultured GIST cell line to determine whether upregulation of miR-196a and *HOTAIR* is responsible for the malignant potential of GISTs. We found that both

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

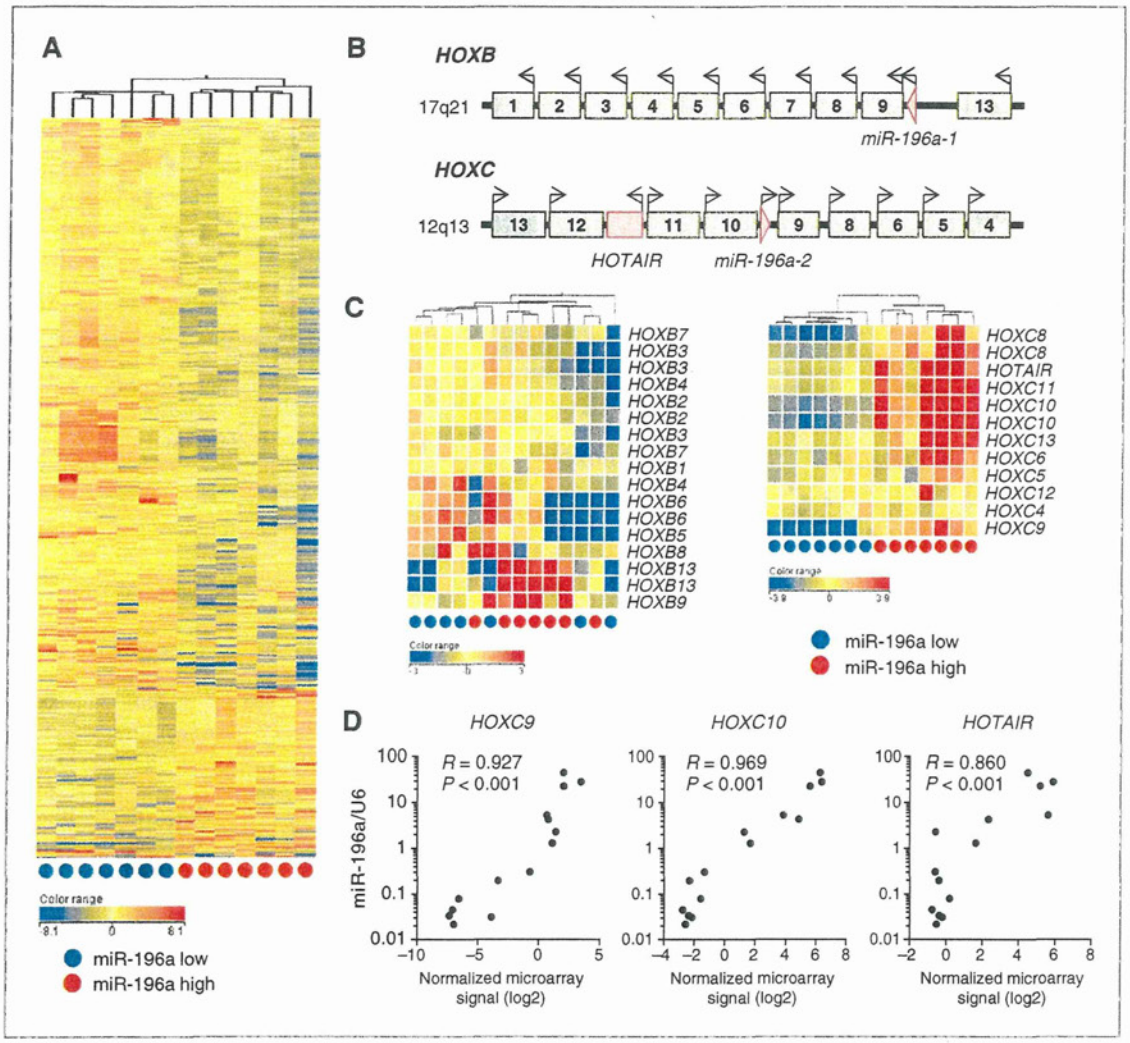

Figure 2. *GIST* gene expression signatures reveal a strong correlation between the *miR-196a* and *HOXC* genes. A, heat map of the gene expression signatures correlated with miR-196a expression. Rows represent probe sets and columns represent patients. A total of 4,947 probe sets differentially expressed (>2-fold change) between GISTs with ($n = 7$) and without ($n = 7$) miR-196a overexpression were selected, after which hierarchical clustering was carried out. The miR-196a expression status is indicated below. B, schematic representations of the *miR-196a* family locations within the *HOX* gene clusters. C, hierarchical clustering analysis using *HOXB* (left) and *HOXC* (right) expression data. miR-196a expression status is indicated below. D, correlations between the expression levels of miR-196a and *HOXC* genes or *HOTAIR*. Expression of miR-196a was analyzed using TaqMan assay and was normalized to internal U6 snRNA. Microarray data for *HOXC* and *HOTAIR* were normalized and log transformed (base 2). The Pearson correlation coefficients and *P* values are shown.

miR-196a and *HOTAIR* are expressed in GIST-T1 cells (Supplementary Fig. S12). We then carried out cell viability and Matrigel invasion assays after transfecting GIST-T1 cells with an anti–miR-196a inhibitor molecule. Gene expression microarray analysis revealed that a number of predicted miR-196a target genes, including *ANXA1* and *HOXA5*, were upregulated by inhibition of miR-196a (Supplementary Table S10), and although we observed no effects on cell viability, inhibition of miR-196a moderately suppressed cell invasion (Supplementary Fig. S13). We next disrupted *HOTAIR* expression by transfecting the cells with siRNAs targeting it (Fig. 3E). Although knockdown of *HOTAIR* did not significantly affect cell viability, it suppressed the invasiveness of GIST-T1 cells (Fig. 3E and F). Moreover, gene expression microarray analysis revealed that a number of reported *HOTAIR* target genes, including *PCHD10*, *SEMA6A*, and *STK17B*, were upregulated upon knockdown of *HOTAIR* (Supplementary Table S11). In all, we found that 1,424 genes were upregulated by siHOT (>2-fold), and Gene Ontology analysis revealed enrichment of genes related to "nucleus," "chromosome," and "membrane-bounded

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

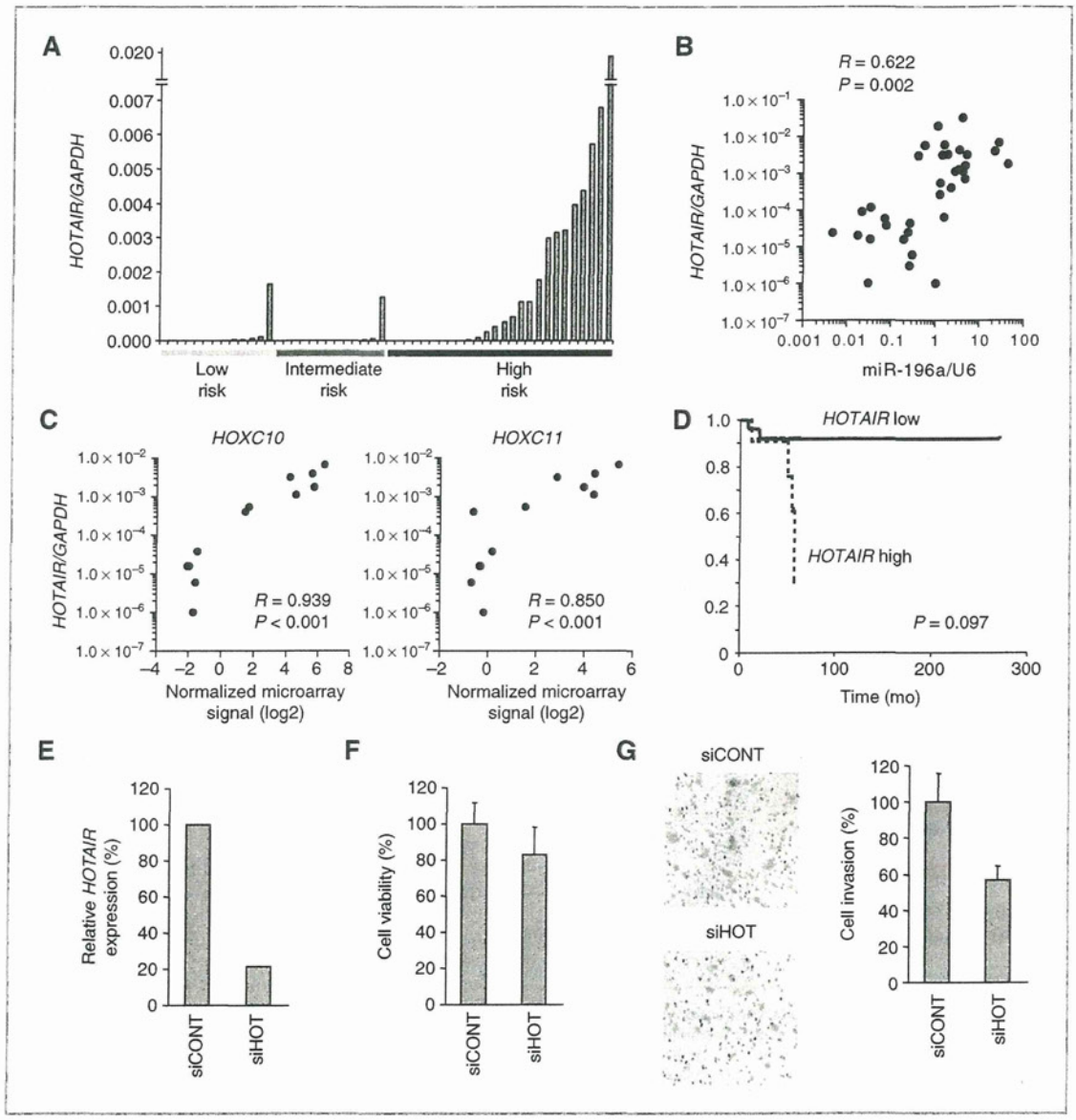

Figure 3. Upregulation of *HOTAIR* in malignant GISTs. A, TaqMan assay of *HOTAIR* in a panel of GIST specimens ($n = 52$). Results are normalized to internal *GAPDH* expression. Risk categories are indicated below. B, correlation between levels of *HOTAIR* and miR-196a expression detected using TaqMan assay. The Pearson correlation coefficient and *P* value are shown. C, correlations between levels of *HOTAIR* expression detected using TaqMan assay and those of *HOXC* genes detected from microarrays. The Pearson correlation coefficients and *P* values are shown. D, Kaplan–Meier curves showing the effect of *HOTAIR* expression (high, *HOTAIR/GAPDH* $\geq$ 0.0002; low, *HOTAIR/GAPDH* < 0.0002) on overall survival among GIST patients. E, TaqMan assay for *HOTAIR* in GIST-T1 cells transfected with control siRNA (siCONT) or siRNA targeting *HOTAIR* (siHOT). F, cell viability assay using GIST-T1 cells transfected with siCONT or siHOT. Cell viabilities were determined 48 hours after transfection. Shown are the means of 8 replications; error bars represent SDs. G, Matrigel invasion assay using GIST-T1 cells transfected with siCONT or siHOT. Shown on the right are the means of 3 random microscopic fields per membrane; error bars represent SDs.

organelle" (Supplementary Tables S12 and S13). These results suggested that *HOTAIR* may modulate transcription of a large number of genes and may have previously unidentified roles in GIST cells.

Finally, we sought to clarify the biologic relationship between miR-196a, *HOTAIR*, and *HOXC* genes. We first tested whether upregulation of miR-196a is a downstream effect of *HOTAIR* dysregulation, or vice versa. We found that inhibition

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

Table 4. HOTAIR expression is associated with poor clinical outcome in GIST patients

	Outcome					
HOTAIR/GAPDH	Survival	Death	HR (95% CI)	*P*	HR[a] (95% CI)	*P*
<0.0002	26	2				
≥0.0002	7	4	3.8 (0.7–21.2)	0.123	9.0 (1.2–68.9)	0.034

[a]Age and gender adjusted HR.

of miR-196a had no effect on *HOTAIR* expression in GIST-T1 cells, nor did knockdown of *HOTAIR* affect miR-196a expression. This suggested that overexpression of miR-196a or *HOTAIR* is not a simple downstream effect of their dysregulation (Supplementary Fig. S12). By contrast, analysis of the chromatin status in GIST-T1 cells using ChIP-PCR revealed enrichment of trimethylated histone H3 lysine 4 (H3K4me3), a hallmark of active gene transcription, at the transcription start sites of multiple *HOXC* genes and *HOTAIR* (Supplementary Fig. S14). Moreover, we found concurrent overexpression of miR-196a, *HOTAIR*, and *HOXC* genes in other cancer cells, including the KatoIII gastric cancer cell line. By carrying out high-resolution ChIP-seq analysis with the KatoIII cells, we observed significant enrichment of H3K4me3 over a wide range (more than 50 kb) of the *HOXC* cluster, which suggested that an epigenetic mechanism is involved in the dysregulation of this genomic region (Supplementary Fig. S15).

Discussion

Although the results of recent studies suggest that the gene expression signatures of GISTs are predictive of malignancy and drug sensitivity of the tumors (5, 21), the clinical significance of the miRNA expression signature is not yet fully understood. In this study, we found that upregulation of miR-196a is strongly associated with a high-risk grade, metastasis, and poor prognosis in GIST patients. Furthermore, overexpression of miR-196a is accompanied by upregulation of multiple *HOXC* genes and the metastasis-related lincRNA *HOTAIR*. To our knowledge, this is the first article to show concurrent overexpression of collinear *HOX* genes and noncoding RNAs in human malignancy.

A number of studies have implicated miR-196a in malignancy, but its role may differ among tumor types. Upregulation of miR-196a is observed in esophageal adenocarcinomas and their precancerous lesions, Barrett's esophagus and dysplasia, which suggests miR-196a is a potential marker of the malignant progression of Barrett's esophagus (22). Strong expression of miR-196a is also associated with a poor prognosis in pancreatic adenocarcinoma and glioblastoma patients (23, 24). In addition, functional analysis showed that expression of miR-196a in esophageal, breast, and endometrial cancer cells promotes proliferation and suppresses apoptosis through downregulation of *ANXA1* (18). These results suggest that miR-196a contributes to oncogenesis in cancer. On the other hand, miR-196a is significantly downregulated in melanoma, and its reexpression inhibited the invasive behavior of melanoma cells by targeting HOXC8 (25). Similarly, miR-196a suppressed HOXC8 and inhibited invasion and metastasis by breast cancer cells (26). Thus miR-196a seems to exert opposite effects in tumors of different origins.

The *HOX* genes are a highly conserved subgroup of the homeobox superfamily, and they play essential roles in a variety of biologic processes, including development, differentiation, apoptosis, and angiogenesis (27). In humans, 4 *HOX* clusters containing 39 *HOX* genes have been identified, and dysregulation of their expression is observed in various malignancies. Although the role of HOXs in cancer is not fully understood, its aberrant expression is thought to affect pathways that promote tumorigenesis and metastasis (27). For instance, *HOXC8* mRNA is overexpressed in prostate cancer cells and is associated with tumor cell proliferation and metastasis (28–30). In addition, *HOXC5* and *HOXC8* mRNAs are upregulated in cervical cancer cells (31), and one recent study suggested HOXC10 plays a key role of in the progression and invasion in cervical cancer (32).

An association between miR-196a and *HOX* expression in cancer has also been reported. Reduced expression of miR-196a in malignant melanoma cells leads to upregulation of HOXB7 and, in turn, activation of BMP4, a major modulator of migration (33). As mentioned above, miR-196a also inhibits invasion and metastasis by downregulating HOXC8 in melanoma and breast cancer cells (26, 34). Taken together, these results suggest that miR-196a acts as a tumor suppressor by targeting *HOX* genes in these tumor types. By contrast, we show in this study that both the miR-196a and *HOXC* genes are concurrently upregulated in malignant GISTs. Our findings are reminiscent of an earlier report showing that the expression patterns of miRNAs embedded in *HOX* clusters are very similar to those of *HOX* genes during mammalian embryogenesis (35). Global gene expression analysis revealed that expression of multiple putative miR-196a targets, including *ANXA1*, is diminished in GISTs overexpressing miR-196a, whereas their expression is enhanced upon inhibition of miR-196a in cultured GIST-T1 cells. In addition, inhibition of miR-196a in GIST cells overexpressing that miRNA moderately suppressed cell invasion. Taken together, our results indicate that upregulation of *HOXC* genes along with miR-196a may contribute to the malignant potential of GIST.

HOTAIR is located within the *HOXC* cluster and encodes a lincRNA known to repress its target genes by directly interacting with histone modification complexes. Epigenetic gene

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

regulation is closely associated with histone modifications in which di- or trimethylation of histone H3 lysine 4 (H3K4me2 or H3K4me3) is enriched within active gene promoters. In addition, trimethylation of histone H3 lysine 27 (H3K27me3) is a marker of gene silencing. In normal adult fibroblasts, *HOTAIR* suppresses the *HOXD* locus by recruiting the PRC2 complex, which consists of the histone H3K27 methylase EZH2, SUZ12, and EED (20). It was also recently shown that *HOTAIR* serves as a scaffold for multiple repressor complexes, including PRC2 and LSD1/CoREST/REST (36). LSD1 is a demethylase that specifically mediates demethylation of H3K4, leading to repression of the target genes. *HOTAIR* is also strongly implicated in cancer metastasis. In breast cancer cells, *HOTAIR* induces retargeting of the PRC2 complex throughout the genome, which leads to the silencing of multiple tumor suppressor and metastasis suppressor genes (17). Overexpression of *HOTAIR* is also predictive of recurrence in hepatocellular carcinoma patients after liver transplantation (37). We observed that upregulation of *HOTAIR* is closely associated with GIST aggressiveness and metastasis. In addition, functional analysis using GIST-T1 cells showed that RNAi-mediated knockdown of *HOTAIR* suppressed cell invasion. These results strongly suggest that upregulation of *HOTAIR* is one of the mechanisms that promote aggressiveness in GISTs. Interestingly, depletion of *HOTAIR* induced a significant change in the gene expression profile in GIST cells, suggesting that *HOTAIR* may regulate a spectrum of genes other than the previously reported target genes. Further studies, including genome-wide histone modification analysis, may reveal as yet unidentified roles played by *HOTAIR* in the malignant progression of GISTs.

The mechanism underlying upregulation of *HOX* cluster genes and noncoding RNAs in GISTs is intriguing. Our array CGH analysis did not detect chromosomal aberrations in any *HOX* loci, making it unlikely that gene amplification is the cause of their overexpression. However, we found that the transcription start sites of multiple genes in the *HOXC* cluster are marked by an active histone mark, H3K4me3, in GIST-T1 cells. Moreover, high-resolution ChIP-seq analysis revealed that, in cancer cells, the entire region is significantly enriched with H3K4me3, leading to overexpression of the affected genes. Our results are reminiscent of the recent finding that rearrangement of *MLL* in leukemia induces active histone modifications at the promoters of *HOXA* genes and miR-196b, resulting to their overexpression (38–40). Although such rearrangements are not known in GISTs, further study to clarify the involvement of epigenetic modifiers in malignant GISTs may lead to identification of new therapeutic targets.

Overall, our study has shown that noncoding RNAs encoded in the *HOXC* cluster could be useful predictive markers as well as novel therapeutic targets in malignant GISTs. As miRNAs are well preserved in FFPE specimens (41), miR-196a could be a reliable biomarker for risk assessment. We also provide evidence that *HOTAIR* is significantly upregulated in high-risk GISTs, indicating that this lincRNA also could be a useful biomarker, as well as a novel therapeutic target. Further study of the causes and functions of *HOXC* locus activation in GISTs may provide new strategies for the treatment of GIST patients.

Disclosure of Potential Conflicts of Interest

T. Nishida has received a research grant from Novartis Pharma K.K. The remaining authors disclose no conflicts of interest.

Acknowledgments

The authors thank Dr. William F. Goldman for editing the manuscript and M. Ashida for technical assistance.

Grant Support

This study was supported in part by grants-in-aid for Scientific Research (B) from the Japan Society for Promotion of Science (Y. Shinomura), A3 foresight program from the Japan Society for Promotion of Science (H. Suzuki), a grant-in-aid for the Third-term Comprehensive 10-year Strategy for Cancer Control (M. Toyota, H. Suzuki), a grant-in-aid for Cancer Research from the Ministry of Health, Labor, and Welfare, Japan (M. Toyota, H. Suzuki), and the Takeda Science Foundation (H. Suzuki).

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.

Received May 31, 2011; revised December 19, 2011; accepted January 6, 2012; published OnlineFirst January 18, 2012.

References

1. Shinomura Y, Kinoshita K, Tsutsui S, Hirota S. Pathophysiology, diagnosis, and treatment of gastrointestinal stromal tumors. J Gastroenterol 2005;40:775–80.
2. Rubin BP, Heinrich MC, Corless CL. Gastrointestinal stromal tumour. Lancet 2007;369:1731–41.
3. Corless CL, Fletcher JA, Heinrich MC. Biology of gastrointestinal stromal tumors. J Clin Oncol 2004;22:3813–25.
4. Fletcher CD, Berman JJ, Corless C, Gorstein F, Lasota J, Longley BJ, et al. Diagnosis of gastrointestinal stromal tumors: A consensus approach. Hum Pathol 2002;33:459–65.
5. Yamaguchi U, Nakayama R, Honda K, Ichikawa H, Hasegawa T, Shitashige M, et al. Distinct gene expression-defined classes of gastrointestinal stromal tumor. J Clin Oncol 2008;26:4100–8.
6. Igarashi S, Suzuki H, Niinuma T, Shimizu H, Nojima M, Iwaki H, et al. A novel correlation between LINE-1 hypomethylation and the malignancy of gastrointestinal stromal tumors. Clin Cancer Res 2010;16:5114–23.
7. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet 2004;5:522–31.
8. Esquela-Kerscher A, Slack FJ. Oncomirs–microRNAs with a role in cancer. Nat Rev Cancer 2006;6:259–69.
9. Croce CM. Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet 2009;10:704–14.
10. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857–66.
11. Choi HJ, Lee H, Kim H, Kwon JE, Kang HJ, You KT, et al. MicroRNA expression profile of gastrointestinal stromal tumors is distinguished by 14q loss and anatomic site. Int J Cancer 2010;126:1640–50.
12. Haller F, von Heydebreck A, Zhang JD, Gunawan B, Langer C, Ramadori G, et al. Localization- and mutation-dependent microRNA (miRNA) expression signatures in gastrointestinal stromal tumours (GISTs), with a cluster of co-expressed miRNAs located at 14q32.31. J Pathol 2010;220:71–86.
13. Maruyama R, Choudhury S, Kowalczyk A, Bessarabova M, Beresford-Smith B, Conway T, et al. Epigenetic regulation of cell type-specific expression patterns in the human mammary epithelium. PLoS Genet 2011;7:e1001369.
14. Suzuki H, Takatsuka S, Akashi H, Yamamoto E, Nojima M, Maruyama R, et al. Genome-wide profiling of chromatin signatures reveals epigenetic regulation of microRNA genes in colorectal cancer. Cancer Res 2011;71:5646–58.

Published OnlineFirst January 18, 2012; DOI:10.1158/0008-5472.CAN-11-1803

15. Taguchi T, Sonobe H, Toyonaga S, Yamasaki I, Shuin T, Takano A, et al. Conventional and molecular cytogenetic characterization of a new human cell line, GIST-T1, established from gastrointestinal stromal tumor. Lab Invest 2002;82:663–5.
16. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 2004;101:2999–3004.
17. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature 2010;464:1071–6.
18. Luthra R, Singh RR, Luthra MG, Li YX, Hannah C, Romans AM, et al. MicroRNA-196a targets annexin A1: a microRNA-mediated mechanism of annexin A1 downregulation in cancers. Oncogene 2008;27:6667–78.
19. Yekta S, Shih IH, Bartel DP. MicroRNA-directed cleavage of HOXB8 mRNA. Science 2004;304:594–6.
20. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. Cell 2007;129:1311–23.
21. Rink L, Skorobogatko Y, Kossenkov AV, Belinsky MG, Pajak T, Heinrich MC, et al. Gene expression signatures and response to imatinib mesylate in gastrointestinal stromal tumor. Mol Cancer Ther 2009;8:2172–82.
22. Maru DM, Singh RR, Hannah C, Albarracin CT, Li YX, Abraham R, et al. MicroRNA-196a is a potential marker of progression during Barrett's metaplasia-dysplasia-invasive adenocarcinoma sequence in esophagus. Am J Pathol 2009;174:1940–8.
23. Bloomston M, Frankel WL, Petrocca F, Volinia S, Alder H, Hagan JP, et al. MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. JAMA 2007;297:1901–8.
24. Guan Y, Mizoguchi M, Yoshimoto K, Hata N, Shono T, Suzuki SO, et al. MiRNA-196 is upregulated in glioblastoma but not in anaplastic astrocytoma and has prognostic significance. Clin Cancer Res 2010;16:4289–97.
25. Mueller DW, Bosserhoff AK. MicroRNA miR-196a controls melanoma-associated genes by regulating HOX-C8 expression. Int J Cancer 2011;129:1064–74.
26. Li Y, Zhang M, Chen H, Dong Z, Ganapathy V, Thangaraju M, et al. Ratio of miR-196s to HOXC8 messenger RNA correlates with breast cancer cell migration and metastasis. Cancer Res 2010;70:7894–904.
27. Shah N, Sukumar S. The Hox genes and their roles in oncogenesis. Nat Rev Cancer 2010;10:361–71.
28. Waltregny D, Alami Y, Clausse N, de Leval J, Castronovo V. Overexpression of the homeobox gene HOXC8 in human prostate cancer correlates with loss of tumor differentiation. Prostate 2002;50:162–9.
29. Miller GJ, Miller HL, van Bokhoven A, Lambert JR, Werahera PN, Schirripa O, et al. Aberrant HOXC expression accompanies the malignant phenotype in human prostate. Cancer Res 2003;63:5879–88.
30. Kikugawa T, Kinugasa Y, Shiraishi K, Nanba D, Nakashiro K, Tanji N, et al. PLZF regulates Pbx1 transcription and Pbx1-HoxC8 complex leads to androgen-independent prostate cancer proliferation. Prostate 2006;66:1092–9.
31. Alami Y, Castronovo V, Belotti D, Flagiello D, Clausse N. HOXC5 and HOXC8 expression are selectively turned on in human cervical cancer cells compared to normal keratinocytes. Biochem Biophys Res Commun 1999;257:738–45.
32. Zhai Y, Kuick R, Nan B, Ota I, Weiss SJ, Trimble CL, et al. Gene expression analysis of preinvasive and invasive cervical squamous cell carcinomas identifies HOXC10 as a key mediator of invasion. Cancer Res 2007;67:10163–72.
33. Braig S, Mueller DW, Rothhammer T, Bosserhoff AK. MicroRNA miR-196a is a central regulator of HOX-B7 and BMP4 expression in malignant melanoma. Cell Mol Life Sci 2010;67:3535–48.
34. Mueller DW, Bosserhoff AK. MicroRNA miR-196a controls melanoma-associated genes by regulating HOX-C8 expression. Int J Cancer 2011;129:1064–74.
35. Mansfield JH, Harfe BD, Nissen R, Obenauer J, Srineel J, Chaudhuri A, et al. MicroRNA-responsive `sensor' transgenes uncover Hox-like and other developmentally regulated patterns of vertebrate microRNA expression. Nat Genet 2004;36:1079–83.
36. Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, et al. Long noncoding RNA as modular scaffold of histone modification complexes. Science 2010;329:689–93.
37. Yang Z, Zhou L, Wu LM, Lai MC, Xie HY, Zhang F, et al. Overexpression of long non-coding RNA HOTAIR predicts tumor recurrence in hepatocellular carcinoma patients following liver transplantation. Ann Surg Oncol 2011;18:1243–50.
38. Okada Y, Feng Q, Lin Y, Jiang Q, Li Y, Coffield VM, et al. hDOT1L links histone methylation to leukemogenesis. Cell 2005;121:167–78.
39. Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, Sinha AU, et al. H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 2008;14:355–68.
40. Popovic R, Riesbeck LE, Velu CS, Chaubey A, Zhang J, Achille NJ, et al. Regulation of mir-196b by MLL and its overexpression by MLL fusions contributes to immortalization. Blood 2009;113:3314–22.
41. Hui AB, Shi W, Boutros PC, Miller N, Pintilie M, Fyles T, et al. Robust global micro-RNA profiling with formalin-fixed paraffin-embedded breast cancer tissues. Lab Invest 2009;89:597–606.

JrnlID 13277_ArtID 378_Proof# 1 - 19/03/2012

Tumor Biol.
DOI 10.1007/s13277-012-0378-3

RESEARCH ARTICLE

Genome-wide analysis of DNA methylation identifies novel cancer-related genes in hepatocellular carcinoma

Masahiro Shitani · Shigeru Sasaki · Noriyuki Akutsu · Hideyasu Takagi · Hiromu Suzuki · Masanori Nojima · Hiroyuki Yamamoto · Takashi Tokino · Koichi Hirata · Minoru Toyota · Yasuhisa Shinomura · Kohzoh Imai

Received: 24 January 2012 / Accepted: 11 March 2012

Abstract Aberrant DNA methylation has been implicated in the development of hepatocellular carcinoma (HCC). Our aim was to clarify its molecular mechanism and to identify useful biomarkers by screening for DNA methylation in HCC. Methylated CpG island amplification coupled with CpG island microarray (MCAM) analysis was carried out to screen for methylated genes in primary HCC specimens [hepatitis B virus (HBV)-positive, $n=4$; hepatitis C virus (HCV)-positive, $n=5$; HBV/HCV-negative, $n=7$]. Bisulfite pyrosequencing was used to analyze the methylation of selected genes and long interspersed nuclear element (LINE)-1 in HCC tissue ($n=57$) and noncancerous liver tissue ($n=50$) from HCC patients and in HCC cell lines ($n=10$). MCAM analysis identified 332, 342, and 259 genes that were methylated in HBV-positive, HCV-positive, and HBV/HCV-negative HCC tissues, respectively. Among these genes, methylation of *KLHL35*, *PAX5*, *PENK*, and *SPDYA* was significantly higher in HCC tissue than in noncancerous liver tissue, irrespective of the hepatitis virus status. LINE-1 hypomethylation was also prevalent in HCC and correlated positively with *KLHL35* and *SPDYA* methylation. Receiver operating characteristic curve analysis revealed that methylation of the four genes and LINE-1 strongly discriminated between HCC tissue and noncancerous liver tissue. Our data suggest that aberrant hyper- and hypomethylation may contribute to a common pathogenesis mechanism in HCC. Hypermethylation of *KLHL35*, *PAX*, *PENK*, and *SDPYA* and hypomethylation of LINE-1 could be useful biomarkers for the detection of HCC.

Keywords Hepatocellular carcinoma · DNA methylation · CpG island · LINE-1 · Biomarker

Electronic supplementary material The online version of this article (doi:10.1007/s13277-012-0378-3) contains supplementary material, which is available to authorized users.

M. Shitani · S. Sasaki (✉) · N. Akutsu · H. Takagi · H. Suzuki · H. Yamamoto · Y. Shinomura (✉)
First Department of Internal Medicine,
Sapporo Medical University,
S1, W16, Chuo-Ku,
Sapporo 060-8543, Japan
e-mail: ssasaki@sapmed.ac.jp
e-mail: shinomura@sapmed.ac.jp

H. Suzuki · M. Toyota
Department of Molecular Biology, Sapporo Medical University,
Sapporo, Japan

M. Nojima
Department of Public Health, Sapporo Medical University,
Sapporo, Japan

T. Tokino
Medical Genome Science, Research Institute for Frontier Medicine, Sapporo Medical University School of Medicine,
Sapporo, Japan

K. Hirata
First Department of Surgery, Sapporo Medical University,
Sapporo, Japan

K. Imai
Division of Novel Therapy for Cancer, The Advanced Clinical Research Center, The Institute of Medical Science,
The University of Tokyo,
Tokyo, Japan

Introduction

Hepatocellular carcinoma (HCC) is one of the most common human malignancies, worldwide [1]. Chronic infection by hepatitis B virus (HBV) and hepatitis C virus (HCV) are well-documented risk factors for the development of HCC, while chronic alcoholism and various environmental factors, including aflatoxin B1, are also believed to be important risk

factors [2, 3]. The development and progression of HCC is often a complex, multistep process entailing the evolution of normal liver through chronic hepatitis and cirrhosis to HCC, but HCC can also arise in a noncirrhotic liver. In either case, the process is influenced by multiple genetic changes, including allelic deletions, chromosomal losses and gains, DNA rearrangements, and gene mutations [4]. In addition, a growing body of evidence suggests that epigenetic changes such as DNA methylation and histone modification also play crucial roles in hepatocarcinogenesis.

Two seemingly contradictory epigenetic events coexist in cancer: global hypomethylation, which is mainly observed in repetitive sequences throughout the genome, and regional hypermethylation, which is frequently associated with CpG islands within gene promoters [5]. Hypermethylation of CpG islands is a common feature of cancer and is associated with gene silencing. Although the classical two-hit theory posits that tumor suppressor genes are inactivated by gene mutation or deletion, it is now recognized that DNA hypermethylation is a third mechanism by which inactivation of tumor suppressor genes occurs, and that it plays a significant role in tumorigenesis. In contrast to the CpG islands, repetitive DNA elements are normally heavily methylated in somatic tissues. About 45 % of the human genome is composed of repetitive sequences, including long interspersed nuclear elements (LINEs) and short interspersed nuclear element [6], and studies have shown that methylation of such repetitive elements can serve as a surrogate for the global methylcytosine content [7]. In that regard, LINE-1 hypomethylation is known to occur during the development of various human malignancies, including HCC [8, 9].

HCC is generally diagnosed at an advanced stage of tumor progression, and a large fraction of HCC cases are fatal. Thus, a better understanding of the underlying molecular mechanisms and identification of genes critical for early detection of HCC and therapeutic intervention would be highly desirable. Although a number of hyper- or hypomethylated loci have been identified in HCC [10–12], only a few studies have been conducted to unravel the genome-wide methylation status [13–15]. In the present study, we carried out genome-wide CpG island methylation analysis in a set of primary HCC specimens, with and without hepatitis virus infection. We also evaluated the hypomethylation of LINE-1 and assessed its association with aberrant CpG island hypermethylation in HCC.

Materials and methods

Tissue samples and cell lines

A total of 57 primary HCC specimens (HBV-positive, n=21; HCV-positive, n=21; HBV/HCV-negative, n=15) were obtained through surgical resection or needle biopsy at Sapporo Medical University Hospital. Corresponding samples of noncancerous liver tissue were also obtained from 50 patients. HBV surface (HBs) antigen and anti-HCV antibody were measured serologically. An informed consent was obtained from all patients before collection of the specimens. The ten liver cancer cell lines (HT17, PLC/PRF/5, Li-7, huH-1, HuH-7, HepG2, Hep3B, HLE, HLF, and JHH-4) used have been described previously [11]. To analyze restoration of gene expression, cells were treated with 2.0 μM 5-aza-2′-deoxycytidine (5-aza-dC) (Sigma, St Louis, MO, USA) for 72 h, replacing the drug and medium every 24 h. Genomic DNA was extracted using the standard phenol–chloroform procedure. Total RNA was extracted using TRIZOL reagent (Invitrogen, Carlsbad, CA, USA) and then treated with a DNA-free kit (Ambion, Austin, TX, USA). Genomic DNA and total RNA from normal liver tissue from a healthy individual were purchased from BioChain (Hayward, CA, USA).

Methylated CpG island amplification coupled with CpG island microarray

Methylated CpG island amplification (MCA) was performed as described previously [13]. Briefly, 500 ng of genomic DNA was digested with the methylation-sensitive restriction endonuclease *Sma*I (New England Biolabs, Ipswich, MA, USA), after which it was digested with the methylation-insensitive restriction endonuclease *Xma*I. The adaptors were prepared by addition of the oligonucleotides RMCA12 (5′-CCGGGCAGAAAG-3′) and RMCA24 (5′-CCACCGCCATCCGAGCCTTTCTGC-3′). After ligation of the digested DNA to the adaptors, PCR amplification was carried out. Using a BioPrime Plus Array CGH Genomic Labeling System (Invitrogen), MCA amplicons from the HCC samples were labeled with Alexa Fluor 647, while amplicons from a normal liver sample was labeled with Alexa Fluor 555. The labeled MCA amplicons were then hybridized to a custom human CpG island microarray containing 15,134 probes covering 6,157 unique genes (G4497A; Agilent Technologies, Santa Clara, CA, USA) [16]. After washing, the array was scanned using an Agilent DNA Microarray Scanner (Agilent technologies), and the data were processed using Feature Extraction software ver. 10.7 (Agilent Technologies). The data were then analyzed using GeneSpring GX ver. 11 (Agilent Technologies).

Methylation-specific PCR

Genomic DNA (1 μg) was modified with sodium bisulfite using an EpiTect Bisulfite Kit (Qiagen, Hilden, Germany), and methylation-specific PCR (MSP) was performed as described previously [17]. Briefly, PCR was run in a 25-μl

volume containing 50 ng of bisulfite-treated DNA, 1× MSP buffer [67 mM Tris–HCl (pH 8.8), 16.6 mM $(NH4)_2SO_4$, 6.7 mM $MgCl_2$, and 10 mM 2-mercaptoethanol], 1.25 mM dNTP, 0.4 μM each primer, and 0.5 U of JumpStart REDTaq DNA Polymerase (Sigma). The PCR protocol for MSP entailed 5 min at 95°C; 35 cycles of 30 s at 95°C, 30 s at 60°C, and 30 s at 72°C; and a 7 min final extension at 72°C. Primer sequences and PCR product sizes are shown in Supplementary Table 1.

Bisulfite pyrosequencing analysis

Bisulfite pyrosequencing analysis was performed as described previously [17]. The PCR protocol entailed 5 min at 95°C; 45 cycles of 1 min at 95°C, 1 min at 60°C, and 1 min at 72°C; and a 7-min final extension at 72°C. PCR products were then bound to Streptavidin Sepharose beads HP (Amersham Biosciences, Piscataway, NJ); after which, the beads containing the immobilized PCR product were purified, washed, and denatured using a 0.2 M NaOH solution. After addition of 0.3 μM sequencing primer to the purified PCR product, pyrosequencing was carried out using a PSQ96MA system (Qiagen, Hilden, Germany) and Pyro Q-CpG software (Qiagen). Primer sequences and PCR product sizes are shown in Supplementary Table 1.

Quantitative RT-PCR

Single-stranded cDNA was prepared using SuperScript III reverse transcriptase (Invitrogen). Quantitative RT-PCR was carried out using TaqMan Gene Expression Assays (*KLHL35*, Hs00400533_m1; *PAX5*, Hs00172003_m1; *PENK*, Hs00175049_m1; *SPDYA*, Hs00736925_m1; *GAPDH*, Hs99999905_m1; Applied Biosystems, Foster City, CA, USA) and a 7500 Fast Real-Time PCR System (Applied Biosystems) according to the manufacturer's instructions. SDS1.4 software (Applied Biosystems) was used for comparative delta Ct analysis, and *GAPDH* served as an endogenous control.

Statistical analysis

To compare differences in continuous variables between groups, *t* tests or ANOVA with post hoc Tukey's tests were performed. Fisher's exact test or chi-squared test was used for analysis of categorical data. Receiver operator characteristic (ROC) curves were constructed based on the levels of methylation. Values of $P<0.05$ (two-sided) were considered statistically significant. Statistical analyses were carried out using SPSS statistics 18 (IBM Corporation, Somers, NY, USA) and GraphPad Prism ver. 5.0.2 (GraphPad Software, La Jolla, CA, USA).

Results

Genome-wide CpG island methylation analysis in HCC

To screen for CpG island hypermethylation in HCC, we carried out methylated CpG island amplification coupled with CpG island microarray (MCAM) analysis using a set of HCC tissue specimens (HBV-positive, $n=4$; HCV-positive, $n=5$; HBV/HCV-negative, $n=7$). As in an earlier study in which the same array system was used, we utilized a signal ratio (Cy5/Cy3) of >2.0 as the criterion for a methylation-positive probe [13]. The average number of methylated probe sets in the HCC specimens was 566 (range 159–846). To assess the association between hepatitis virus infection and methylation status, we categorized the HCC specimens according to their viral status. The average numbers of methylated probe sets in HBV-positive, HCV-positive, and the HBV/HCV-negative HCC specimens were 574, 598, and 539, respectively, which did not significantly differ ($P=0.840$). Interestingly, however, the numbers of methylated probe sets were more varied among HBV/HCV-negative HCCs, which is indicative of their varied pathological backgrounds (Fig. 1a).

To identify commonly methylated genes in HCC, we selected genes that were methylated in at least two tumors in each group. Among the HBV-positive HCCs, 443 probe sets (corresponding to 332 unique genes) satisfied this criterion. Among the HCV-positive HCCs, 476 probe sets (342 unique genes) satisfied the criterion, and among the HBV/HCV-negative HCCs, 348 probe sets (259 unique genes) satisfied the criterion. Collectively, 714 probes (514 unique genes) were selected as commonly methylated genes. Of those, 137, 146, and 47 probe sets were methylated in only HBV-positive, HCV-positive, or HBV/HCV-negative HCC tissues, respectively (Fig. 1b). By contrast, a large number of genes were methylated in multiple categories, and 169 probe sets were methylated in all three groups (Fig. 1b). Consistent with the above results, unsupervised hierarchical clustering analysis demonstrated that some genes were methylated irrespective of the hepatitis virus status, and that HCV-positive HCCs exhibited the largest number of methylated genes (Fig. 1c, Supplementary Fig. 1). Gene ontology analysis of the commonly methylated genes revealed that genes related to "multicellular organismal process," "developmental process," and "system development" are significantly enriched among the methylated genes (Supplementary Table 2). In addition, pathway analysis suggested that some of the methylated genes are involved in differentiation and development (Supplementary Fig. 2).

Identification of novel genes methylated in HCC

Our MCAM analysis suggested that some genes were methylated in a hepatitis virus-specific manner, but a larger