Int J Clin Exp Pathol 2011;4(3):230-240
www.ijcep.com /IJCEP1102004

Original Article

Transcriptional factor typing with SOX2, HNF4aP1, and CDX2 closely relates to tumor invasion and Epstein-Barr virus status in gastric cancer

Hiroshi Uozaki, Rita Rani Barua, Sun Minhua, Tetsuo Ushiku, Rumi Hino, Aya Shinozaki, Takashi Sakatani, Masashi Fukayama

Department of Pathology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan.

Received February 20, 2011; accepted February 28, 2011; Epub March 1, 2011; published March 31, 2011

Abstract: Background: Gastric cancer (GC) is a major cancer, sometimes associated with Epstein-Barr virus (EBV). Some transcriptional factors (TFs) are specific to the digestive tract and related to the character of the tumors. Methods: We studied three TFs, SOX2, CDX2, and hepatocyte nuclear factor 4 alpha-promoter 1 (HNF4aP1) in GC. First, 255 tumors including 31 EBV-associated GC were immunohistochemically examined using tissue arrays and compared TF type and mucin phenotype. We classified them into 4 TF types: N-TF type as SOX2-/HNF4aP1- tumor, G: SOX2+/HNF4aP1-, GI: SOX2+/HNF4aP1+, and I: SOX2-/HNF4aP1+. Next, 915 GCs were intensely investigated and compared with their clinicopathological factors.Results: In the first study, 255 GCs were classified into N-TF 44%, G-TF 31%, GI-TF 3%, and I-TF 2%. The TF type did not strictly accord with the mucin phenotype, classified by MUC2/5AC/6/CD10 expression. EBV status was the only factor related to both the TF and mucin phenotype classifications ($P<0.0001$, <0.0001). TF classification is related to more factors including tumor stage, than mucin phenotype classification. The second study using 915 GCs revealed that N-TF gradually increased and I-TF decreased as GC invaded deeper. TF classification was not related to nodal involvement in each tumor stage. HNF4aP1 and CDX2 were independent factors for early stage tumor in logistic regression analysis. Conclusions: EBV-associated GC is a discriminating group in both TF and mucin phenotype. TF classification, especially the absence of HNF4aP1 and CDX2, is related to tumor invasion. TF classification is a useful marker to study the carcinogenesis of GC further.

Keywords: Gastric cancer, transcriptional factor, Epstein-Barr virus, SOX2, HNF4a

Introduction

Gastric cancer (GC) is a major cancer worldwide. In addition, 5–10% of GC is infected by Epstein-Barr virus (EBV), known as EBV-associated GC (EBVaGC). Monoclonal EBV infection is observed in EBVaGC, and EBV is thought to be closely related to the carcinogenesis of EBVaGC [1]. EBVaGC has unique characteristics including poorly differentiated histology and the preference for a null or gastric mucin phenotype.

Some transcriptional factors (TFs) are rather specific to the digestive tract and also relate to the development of organs or germ cells. The SRY-related high-mobility group (HMG)-box protein-2 (SOX2) is a member of the HMG-domain DNA-binding-protein family, which is implicated in the regulation of transcription and chromatin architecture [2], Recently, SOX2 has been identified as a stem cell marker, important in establishing IPS cells. SOX2 regulates the apoptosis and pluripotency of stem cells, is highly expressed in the neuroepithelium of the developing nerve system, and has been shown to maintain neural stem cells. It is also highly expressed in the upper gastrointestinal tract from the pharynx to the upper half of the stomach in adult mice [3], and acts as a transcription factor in the stomach [4]. SOX2 plays important roles in growth inhibition through cell-cycle arrest and apoptosis in the gastric epithelium. In gastric cancer, the expression of SOX2 is low and its loss is related to a poor prognosis [5]. The gastric phenotype is related to SOX2 expression among gastric cancers [6].

Table 1. Summary of the clinicopathological data of the examined GCs

	First study (TF and mucins)	Second detailed study (TF)
Tumors	255	915
Patients	244	854
EBVaGC/ EBV-negative GC	31/224	58/857
Gender (male/female)	165/79	619/235
Age	63.9±12.1 [a]	64.5±11.3
Tumor size (mm)	54.8±38.4 [a]	52.9±36.4
Tumor location (upper/middle/lower)	74/116/65	213/389/313
Tumor depth (M/SM/MP/SS/SE/SI) [b]	46/52/27/53/72/5	231/164/107/178/216/19
Hislology in Japanese GC classification (pap/tub1/tub2/por/sig/muc/others) [c]	6/43/77/105/22/2/0	32/191/249/335/88/18/2
Lauren's classification (Intestinal/Diffuse)	125/130	471/444
Lymphatic infiltration (p/n) [d]	134/121	424/491
Vessel infiltration (p/n)	145/110	463/452
Lymph node metastasis (p/n)	116/128	404/474

[a] Mean ± SD. [b] M: intramucosal tumor, SM: invasion to submucosal layer, MP: to muscularis propria, SS: to subserosal layer, SE: exposed to serosa, SI: direct invasion of an adjacent organ. [c] pap: papillary adenocarcinoma, tub1: well differentiated tubular adenocarcinoma, tub2: moderately differentiated tubular adenocarcinoma, por: poorly differentiated adenocarcinoma, sig: signet-ring cell carcinoma, muc: mucinous adenocarcinoma. [d] p, positive; n, negative

CDX2 is a mammalian homeobox analogue to the caudal homeobox genes of the Drosophila melanogaster fruit fly that determine organogenesis, anterior-posterior polarity, and cell identity. CDX2 regulates the transcription of intestinal specific genes and is critical in intestinal development and homeostasis [7]. In human adults, CDX2 is expressed in the colonic epithelium and also in the intestinalized epithelium of the stomach [4]. About half of all GC express CDX2 [8].

Hepatocyte nuclear factor 4 alphas (HNF4a) are essential for development of the liver, and are related to fat metabolism [9]. In humans, heterozygous mutation of HNF4a causes maturity onset diabetes of the young 1 (MODY-1). Several isoforms of HNF4a are generated by alternative promoter (P1 and P2) usage and splicing. The usage of HNF4a promoter depends on the organ. Cancers show alternative expression patterns of P1 and P2 promoter-driven HNF4a. In the small intestine and colon, both P1-driven HNF4a and P2-driven HNF4a are expressed. P2-driven HNF4a is expressed in the normal stomach. P1-driven HNF4a (HNF4aP1) is expressed in about half of all gastric cancers [9,10].

We investigated the expression of these 3 TFs, SOX2, CDX2, and HNF4aP1, in GC and explored the relation between the TF expression and clinicopathological factors. We revealed some significance of the TF expression in GC carcinogenesis.

Materials and methods

Patients and samples

Formalin-fixed, paraffin-embedded (FFPE) tissue blocks with primary GC were obtained from the archives of the Department of Pathology, The University of Tokyo Hospital, from 1990 to 2007.

First, 255 tumors, including 31 EBVaGC, were compared for their TF type and mucin (M) phenotype. Next, 915 GCs from 854 patients, including 58 EBVaGCs, were retrieved from the archives, investigated for their TF type, and compared with their clinicopathological factors.

Histological and pathological data were evaluated according to the Japanese classification of GC [11]. The Lauren classification [12] was also adopted for histological classification. The presence of EBV in carcinoma was determined by *EBER1*-in situ hybridization using FFPE sections. The clinicopathological data of the tumors are summarized in **Table 1**.

Tissue cores, 2mm in diameter, from FFPE

blocks were arrayed using a manual device (Beecher Instruments, WI, USA). Two cores were taken from each tissue block and inserted into a new paraffin block. Suitable cases were defined as tumors occupying more than 50% of the core area. The arrayed tissues were sliced at 4 μm thickness and placed on slides.

This study was approved by our Institutional Review Board (No. 710).

IHC staining

IHC staining was performed on tissue array sections using the VECTASTAIN ABC kit (Vector Laboratories, CA, USA) except for HNF4aP1 staining. Ventana XT System Discovery (Roche Diagnostics, USA) was used for HNF4aP1. The primary antibodies used were mouse anti-SOX2 antibody (Perseus PPZ0113, 1:500), mouse anti-CDX2 antibody (BioGenex CDX2-88, 1:200), mouse anti-HNF4aP1 antibody (Perseus K9218, 1:250), mouse anti-MUC5AC antibody (Novocastra CLH2, 1:50), mouse anti-MUC6 antibody (Novocastra CLH5, 1:50), mouse anti-MUC2 antibody (Novocastra Ccp58, 1:100), and mouse anti-CD10 antibody (Novocastra 56C, 1:100). The sections were autoclaved for 10 minutes in citrate buffer pH6 for antigen retrieval and incubated with primary antibodies at 4°C overnight.

Evaluation of IHC stain

The IHC results were graded into 4 or 5 groups according to the positive rate of cancer cells. The immunohistochemical stains of TFs were graded into 4 groups according to the extent of positivity of carcinomas as follows; grade 0: no positive stain, grade 1: less than 5%, grade 2: less than 50%, grade 3: more than 50% of the tumor cells showed positive stain, respectively. Stains of MUCs and CD10 were graded into 5 groups as follows; grade 0: no positive stain, grade 1: less than 10%, grade 2: less than 30%, grade 3: less than 70%, grade 3: more than 70%. After 2 cores of each tumor had been evaluated, the average of the two scores was adopted as the total grading of the tumor. The average was rounded up to an integer. After grading, positive cases were defined as grade 2 or more for each antibody.

TF type / M phenotype

GCs were classified into 4 transcriptional factor (TF) types according to SOX2 and HNF4aP1 expression: null (N)-TF type (SOX2-/HNF4aP1-), gastric (G)-TF type (SOX2 +/HNF4aP1-), intestinal (I)-TF type (SOX2 -/HNF4aP1+), and gastric and intestinal mixed (GI)-TF type (SOX2 +/HNF4aP1+).

For M phenotypes, the G-M phenotype was defined as MUC5AC and/or MUC6 positivity, and the I-M phenotype as MUC2 and/or CD10 positivity, with the GI-M phenotype positive for both gastric and intestinal phenotypes [13]. N-M phenotype was defined as a phenotype negative for both gastric and intestinal phenotype. TF and M classifications were compared with each other and with tumor stages.

Statistical analysis

The chi-square test was used to examine the distribution of two variables, that is, two or oligo-valued data. The unpaired t-test was used to test for differences in tumor size and patient age. Differences among more than two groups were tested by one-factor ANOVA analysis of variance. The effect of TF on tumor progression was also tested in logistic regression models. P values less than 0.05 were defined as significant.

Results

Expression of TFs and classification

Positive signals of TFs, SOX2, CDX2, and HNF4aP1, were found in the nuclei of non-neoplastic and neoplastic epithelial cells (**Figure 1**). CDX2 positivity was sometimes observed in cytoplasm, but only the nuclear stain was evaluated. MUC2/5AC/6/CD10 signals were found in the cytoplasm of epithelial cells. In non-neoplastic stomach mucosa, the foveolar epithelium was SOX2-positive, and intestinal metaplasia was CDX2- and HNF4aP1-positive. Two tumors were missing for HNF4aP1 stain due to deficit of the sample. The results of IHC are summarized in **Table 2**. When the expression of each TF was compared, HNF4aP1 and CDX2 were positively correlated ($P<0.0001$) and HNF4aP1 and SOX2 were negatively correlated ($P=0.016$). MUC2 was positively related to CDX2 and HNF4aP1 ($P=0.002$ and $P=0.010$, respectively), but not to SOX2. MUC5AC, MUC6, and CD10 were not related to TFs.

The relations between TFs and clinicopathologi-

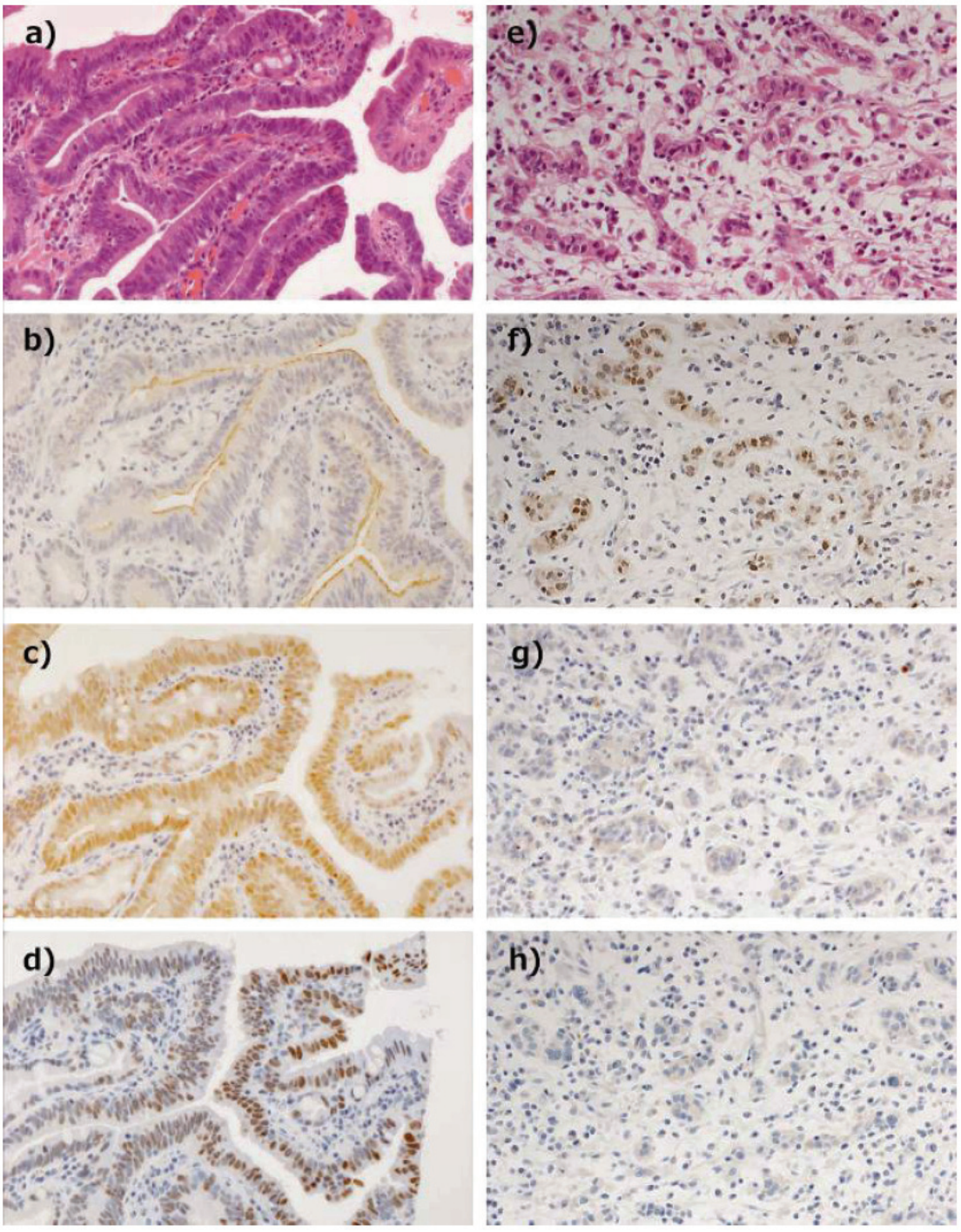


Figure 1. Examples of IHC staining and scoring. **a-d**) EBV-negative GC, an identical tumor. **e-h**) EBV-associated GC, an identical tumor. **a,e**) HE stains. **b, f**) SOX2 stains show grade 0 and 2, respectively. **c, g**) CDX2 stains show grade 3, 0. **d, h**) HNF4aP1 stains show grade 3, 0.

cal factors in GC are presented in **Table 3**. CDX2 and HNF4aP1 were more frequently expressed in early stage tumors (Tis or T1) ($P=0.019$, $P<0.0001$). Their expression was also correlated to the absence of nodal involvement and to smaller tumor size. HNF4aP1-positive tumors

Table 2. Positivity of molecules

Molecules	First study (0/1/2/3/4)	Positive cases (%)	Second detailed study (0/1/2/3)	Positive cases (%)
SOX2	144/24/41/45/-	89(33.7)	461/138/199/117	316(34.5)
CDX2	114/38/48/55/-	104(40.8)	258/220/277/160	437(47.8)
HNF4aP1	84/107/50/12/-	62(24.5)	222/412/229/52	282(30.8)
MUC5AC	74/63/64/40/12	126(45.9)		
MUC6	91/68/60/29/7	114(37.6)		
MUC2	168/49/23/12/3	59(14.9)		
CD10	72/89/59/28/7	108(36.9)		

Table 3. Transcriptional factors and clinicopathological factors

		SOX2		CDX2		HNF4aP1	
		+/-	P value	+/-	P value	+/-	P value
Age*		61.3±13.1/ 65.9±11.3	0.004	65.7±12.6/ 63.5±11.7	0.160	65.6±10.4/ 63.9±11.6	0.039
Tumor size*		51.3±37.5/ 56.6±38.8	0.303	47.6±33.9/ 59.8±40.5	0.013	43.5±28.0/ 57.0±38.8	<0.0001
Tumor stage	Early**	52/105		49/49		41/55	
	Advanced	34/64	0.903	54/103	0.017	22/135	<0.0001
Lauren's classification	Intestinal	43/82		50/75		37/87	
	Diffuse	43/87	0.928	53/77	1	26/103	0.102
EBV	Positive	23/8		1/30	***	0/30	***
	Negative	63/161	<0.0001	102/122	<0.0001	63/160	0.002
Lymphatic infiltration	Positive	43/91		47/87		23/111	
	Negative	43/78	0.654	56/65	0.09	40/79	0.002
Vessel infiltration	Positive	49/96		53/92		23/122	
	Negative	37/73	1	50/60	0.192	40/68	0.0002
Lymph node metastasis	Positive	37/83		39/81		19/101	
	Negative	49/86	0.431	64/71	0.022	44/89	0.003

* mean±SD, ** Tis or T1. M or SM in Japanese GC classification, *** tested by Fisher's exact test

showed lower frequencies of lymphatic and vascular infiltration.

The expression of all 3 TFs, SOX2, CDX2, and HNF4aP1, was strongly related with the EBV status, either positively or negatively. SOX2 expression was associated with the presence of EBV (P<0.0001), while CDX2 and HNF4aP1 expression with the absence of EBV (P<0.0001, P=0.002).

HNF4aP1 showed a close correlation with more clinicopathological factors than CDX2; therefore, we adopted HNF4aP1 as an excellent intestinal transcriptional factor to examine the relationship with clinicopathological factors further and classified tumors using the combination of SOX2 and HNF4aP1 expression, into 4 groups; null (N)- (SOX2-/HNF4aP1-), gastric (G)- (SOX2+/HNF4aP1-), gastric and intestinal mixed (GI)-TF type (SOX2+/HNF4aP1+), and intestinal (I)- (SOX2-/HNF4aP1+) TF types.

Comparison of TF and M classifications

According to TF types, 255 cases of GC were classified into N-TF (N=112, 44%), G-TF (N=78, 31%), GI-TF (N=8, 3%), and I-TF (N=55, 22%),

Table 4. TF types/M phenotypes and clinicopathological factors among 255 tumors

		TF type		M phenotype	
		N/G/GI/I	P value	N/G/GI/I	P value
Age*		65.5±11.6/60.5±13.2/ 69.4±9.9/66.5±10.6	0.007	62.9±13.3/64.9±12.6/ 63.5±12.4/65.4±9.7	0.619
Tumor size*		65.6±41.3/53.3±38.3/ 32.5±24.0/39.8±25.8	0.0001	60.7±33.0/58.3±42.0/ 49.6±32.3/54.7±46.7	0.352
Tumor stage	Early	27/28/6/35		10/27/42/19	
	Advanced	85/50/2/20	<0.0001	33/41/51/32	0.11
Lauren's	Intestinal	49/38/5/32		17/29/49/30	
classification	Diffuse	63/40/3/23	0.299	26/39/44/21	0.164
EBV	Positive	7/23/0/0		16/9/4/2	
	Negative	105/55/8/55	<0.0001	27/59/89/49	<0.0001
Lymphatic	Positive	71/40/3/20		26/42/40/26	
Infiltration	Negative	41/38/5/35	0.008	17/26/53/25	0.077
Vessel	Positive	74/48/1/22		29/38/46/32	
Infiltration	Negative	38/30/7/33	0.0006	14/30/47/19	0.188
Lymph node	Positive	45/44/5/39		21/36/38/25	
metastasis	Negative	67/34/3/16	0.002	22/32/55/26	0.473

* mean±SD, tested by one-factor ANOVA analysis of variance.

respectively. M phenotypes consisted of N-M (N=43, 17%), G-M (N=68, 27%), GI-M (N=93, 36%), and I-M (N=51, 20%) phenotypes, respectively. The relationship between TF types and M phenotypes are presented in **Figure 2**. N- and G-TF types more frequently showed GI-M phenotype, but TF types did not show strict correspondence to M phenotypes.

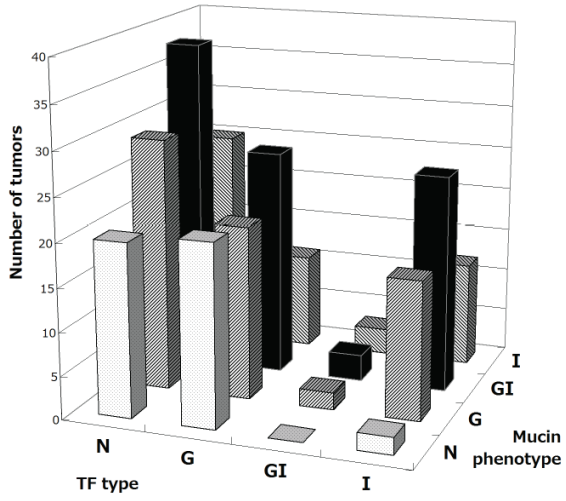


Figure 2. Distribution of each M phenotype GC among TF types. Relations between M phenotype and TF type are shown. P value was 0.003 in a chi-square test, but the distribution showed no constant gradient.

Two classifications and clinicopathological factors

The relations of TF/M types with clinicopathological factors are summarized in **Table 4**. TF types showed significant differences in many factors, such as gender, age, tumor size, tumor stage, and EBV infection, suggesting that the classification of TF types identifies significant subgroups. On the other hand, the classification based on M phenotypes showed differences only in EBV status.

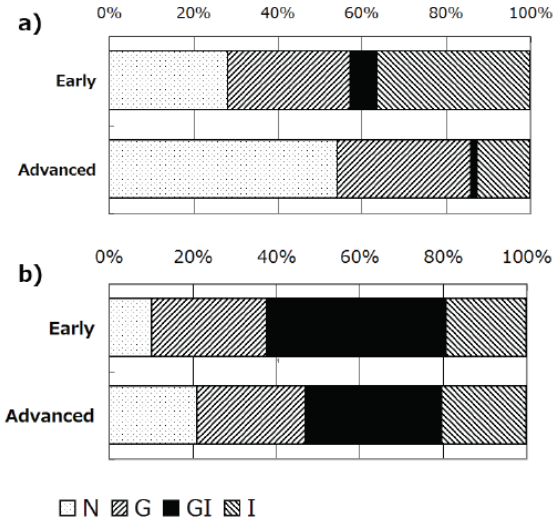
Early GC tumors are defined as tumors invading the submucosal layer or less, and advanced tumors as invading to muscular layer or deeper, according to the Japanese classification of GC [11]. When the distribution of TF/M types was compared in early and advanced carcinomas, the TF classification showed prominent differences: I-TF type predominated in early GC, and N-TF type in advanced GC (**Figure 3a**).

As for lymph node metastasis, one of the important indicators predicting the clinical behavior of GC, N-TF type showed significantly high frequency of involvement (59%), compared to other types (44% in G-TF, 38% in GI-TF and 16% in I-TF types). On the other hand, no significant difference was observed in M phenotype classification.

Table 5. TF phenotypes and clinicopathological factors among 915 tumors

		TF type with HNF4aP1 N/G/GI/I	P value	TF type with CDX2 N/G/GI/I	P value
Age*		66.3±11.5/60.6±13.1/ 67.3±12.6/63.8±10.4	0.001	65.1±10.4/61.4±12.3/ 64.3±11.2/66.1±11.3	0.0002
Tumor size*		60.2±40.3/52.5±36.3/ 42.7±29.4/43.7±27.8	<0.0001	60.9±38.6/52.6±37.1/ 48.8±33.2/46.3±33.5	<0.0001
Tumor depth	M	71/63/22/75		51/37/48/95	
	SM	52/47/14/51		43/33/28/60	
	MP	50/32/2/23		41/18/16/32	
	SS	77/50/9/42		68/32/27/51	
	SE	110/65/8/33	<0.0001,	98/44/29/45	
	SI	12/4/0/3	Fig 5a	11/2/2/4	<0.0001
Tumor	U	90/76/7/40		80/52/31/50	
Location	M	139/115/29/106	0.002,	113/71/73/132	
	L	143/70/19/81	Fig 4a	119/43/46/105	0.002
Lauren's	Intestinal	173/110/34/154	<0.0001,	153/68/76/174	
classification	Diffuse	199/151/21/73	Fig 4a	159/98/74/113	0.0005
Histology in	pap	20/5/3/4		17/5/3/7	
Japanese GC	tub1	66/35/11/79		58/21/25/87	
classification	tub2	87/71/20/71		78/43/48/80	
	por1	65/49/2/15		59/38/13/21	
	por2	98/59/11/36		77/32/38/57	
	sig	24/37/6/21		21/25/18/24	
	muc	12/3/2/1	<0.0001,	2/0/5/11	
	others	0/2/0/0	Fig 4c	0/2/0/0	<0.0001
EBV	Positive	15/40/1/2		16/37/4/1	
	Negative	357/221/54/225	<0.0001	296/129/146/286	<0.0001
Lymphatic	Positive	204/116/21/83		176/81/56/111	
Infiltration	Negative	168/145/34/144	<0.0001	136/85/94/176	<0.0001
Vessel infiltration	Positive	212/138/19/94		190/88/69/116	
	Negative	160/123/36/133	0.0002	122/78/81/171	<0.0001
Lymph node metastasis	Positive	200/106/19/79		180/69/56/99	
	Negative	158/145/32/139	<0.0001	129/93/84/170	<0.0001

* mean±SD, tested by one-factor ANOVA analysis of variance.



TF type and clinicopathological factors in a larger study

We applied this TF classification further to a series of 915 tumors (**Table 1**), including 58 cases of EBVaGC, to analyze the significance of TF types. The distribution of the four types was almost same as in the first series; 41, 29, 6 and 25% in N, G, GI, and I types, respectively.

The relationships between the TF type and clinicopathological factors were analyzed (**Table 5**).

Figure 3. TF/M classification and tumor stages. **a**) Relation between TF type and tumor stages (P<0.0001), **b**) M phenotype and tumor stages (P=0.110).

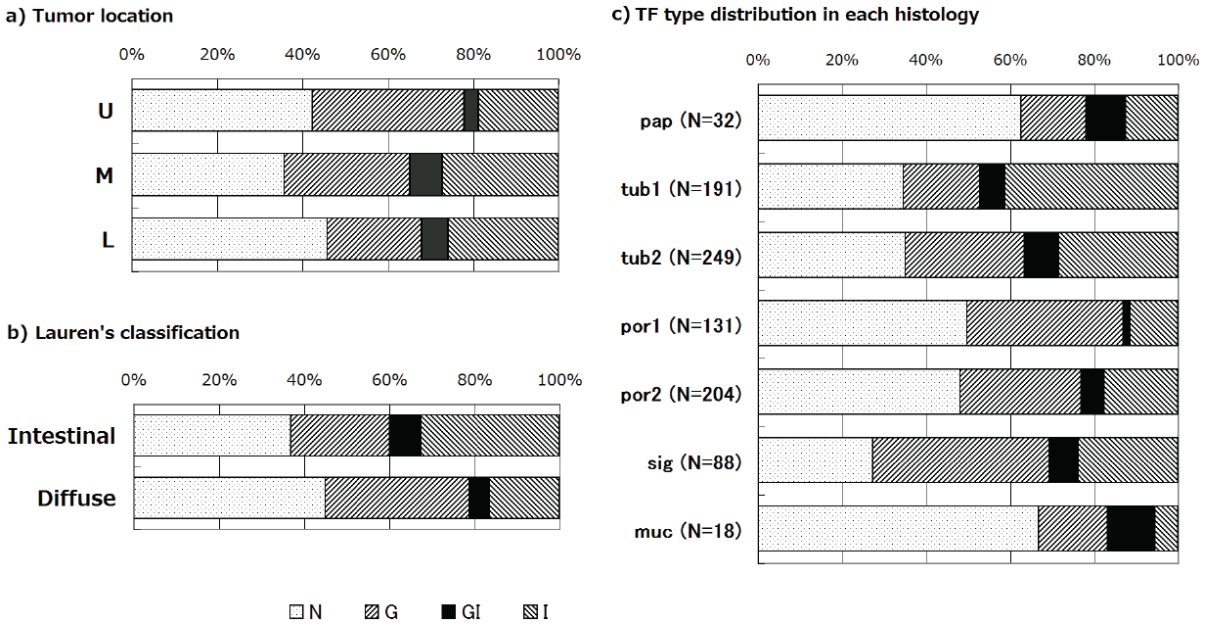


Figure 4. TF types and clinicopathological factors. **a**) Tumor location and TF types. G-TF type is more frequent among upper location GC. I-TF type GC is more frequent among lower and middle location GC (P=0.003). **b**) Lauren's histological type and TF type. Diffuse GC showed more frequent G-TF type and less I-TF type (P<0.0001). **c**) Japanese histological classification of GC and TF types. Pap: papillary adenocarcinoma, tub1: well differentiated tubular adenocarcinoma, tub2: moderately differentiated tubular adenocarcinoma, por1: poorly differentiated adenocarcinoma, solid type, por2: poorly differentiated adenocarcinoma, non-solid type, sig: signet-ring cell carcinoma, muc: mucinous adenocarcinoma. Papillary adenocarcinoma predominantly showed N-TF type; poorly differentiated adenocarcinoma, N/G -TF type; signet-ring cell carcinoma, G-TF type; mucinous adenocarcinoma, N-TF type.

The tendency was almost the same as in the previous 255 tumor series and also same when CDX2 is used as an intestinal marker. The distribution of TF types in each location is shown in **Figure 4a.** G-TF type is more frequent in upper location than in middle or lower locations. I-TF type is more frequent in middle and lower locations than in upper location.

The TF type and Lauren's histological classification were compared (**Figure 4b**). G-TF type was slightly more frequently observed in the diffuse type and the I-TF type was slightly more frequent in the intestinal type (P<0.0001). The TF type in each histological subtype of Japanese classification of GC was analyzed and is shown in **Figure 4c**. Papillary adenocarcinoma showed the N-TF type most frequently. Poorly differentiated adenocarcinoma was mainly N-TF or G-TF type. Signet ring cell carcinoma was mainly N-TF or G-TF type. Mucinous adenocarcinoma is usually N-TF type.

The relation between TF type and tumor progression (depth of tumor invasion) was also investigated in detail. The depth of tumor invasion was determined according to the Japanese classification of GC [11]. As shown in **Figure 5a**, N-TF gradually increased and GI/I-TF decreased in the proportion of each class of invasion. The frequency of G-TF showed a slight change. Next we examined the rate of lymph nodal involvement of each TF type tumor at 6 depths, respectively (**Figure 5b**). The rate of nodal involvement was similar at each invasion depth regardless of the TF type.

We tested the significance of TF in tumor progression to advanced stage using logistic regression models. Various factors were tested, including EBV status, tumor size, Lauren's intestinal type, gender, SOX2 positivity, and HNF4aP1 positivity in early/advanced stage tumor. HNF4aP1 positivity was an independent favorable factor (odds ratio 0.335 (0.230-0.487) P<0.0001, **Table 6a**). CDX2 is also an independent favorable factor when it is used as an intestinal marker (**Table 6b**).

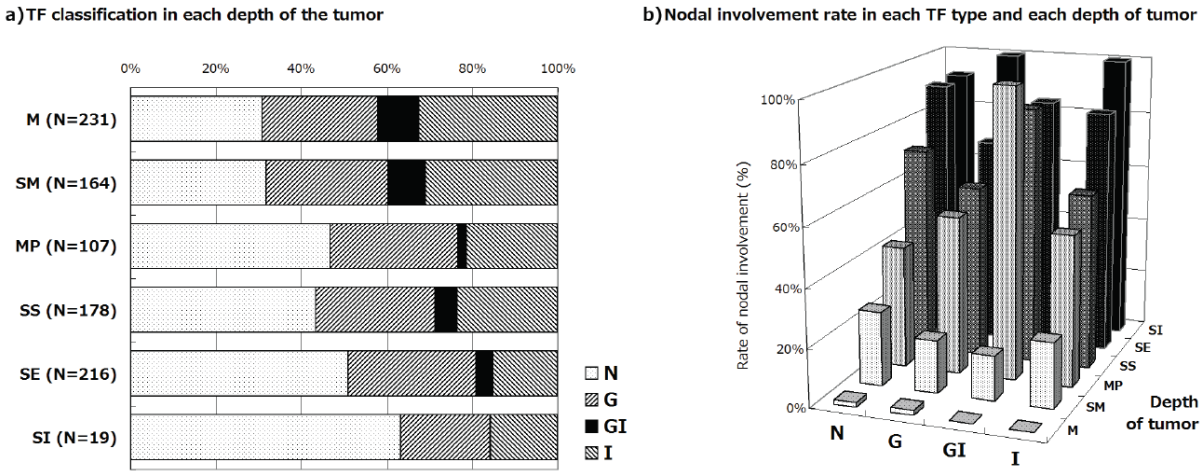


Figure 5. TF types and tumor progression. **a**) TF type distribution at each depth of tumor. M/SM/MP/SS/SE means that a tumor reached the intramucosa/submucosal layer/muscularis propria/subserosal layer/serosa, respectively. SI means that a tumor invaded an adjacent organ directly. N-TF type gradually increased and I-TF type gradually decreased with tumor progression. The frequency of G/GI-TF type did not significantly change. **b**) Nodal involvement rate in each TF type and each depth of tumor. Nodal involvement was greatly influenced only by tumor stage. TF type did not apparently influence the frequency of nodal involvement.

Table 6. TFs in logistic regression models

a) HNF4aP1

Factor	Odds ratio	P value
EBV	0.261 (0.158-0.432)	0.167
Size (mm)	1.053 (1.044-1.061)	<0.0001
Lauren's intestinal type	0.642(0.461-0.893)	0.009
Gender (male)	0.920(0.638-1.327)	0.654
SOX2	0.730(0.517-1.030)	0.073
HNF4aP1	0.335(0.230-0.487)	<0.0001

b) CDX2

Factor	Odds ratio	P value
EBV	0.249 (0.149-0.416)	0.153
Size (mm)	1.052 (1.044-1.061)	<0.0001
Lauren's intestinal type	0.594(0.430-0.821)	0.002
Gender (male)	0.958(0.667-1.376)	0.817
SOX2	0.864(0.617-1.211)	0.396
CDX2	0.574(0.416-0.793)	0.001

Discussion

We examined the expression of a gastric transcriptional factor, SOX2, and two intestinal transcriptional factors, CDX2 and HNF4aP1, in GC and investigated their relation to the clinicopathological factors and mucin phenotype. We preliminarily examined HNF4a expression in 72 GC. All tumors showed HNF4aP2 expression and some tumors showed HNF4aP1 expression, as reported previously [9]. In this study, we investigated only HNF4aP1 expression among HNF4a proteins.

The positivities of CDX2 and HNF4aP1 were significantly correlated. HNF4aP1 was correlated inversely with SOX2 and correlated to more factors than CDX2 (**Table 3**). We tried to reveal the significance of TF expression in carcinogenesis in this study, so we adopted HNF4aP1 as an intestinal factor when deciding the TF type in this study. We also tested CDX2-based TF type classification in similar analyses and obtained similar results to the results in this HNF4aP1-based study (**Table 5, 6**). Although the molecular relation between CDX2 and HNF4aP1 have not been reported, they might be associated with each other.

TF type and M phenotype

SOX2 is closely related to the gastric phenotype and CDX2 to the intestinal phenotype [4]. SOX2 is reported to upregulate MUC5AC and pepsino-

gen [14,15]. In this study, SOX2 expression did not correlate with MUC5AC or the gastric mucin phenotype.

We tested the relation between TF types and M phenotypes (**Figure 2**). Each TF type distributed broadly among M phenotypes. TF type is a different classification from M phenotype, so TF types and M phenotypes should be studied separately.

Two classifications based on TF type and M phenotype were compared with the clinicopathological factors (**Table 4**). The EBV status of GC is the only consistent factor related to the two classifications. EBV-associated GC is a discriminating group not only of EBV infection but also of cellular phenotypes.

Tumor progression and phenotype

The distribution of each TF/M type in GC of early and advanced stages was analyzed. The GI/I-TF type was found markedly less in advanced stage (**Figure 3a**). We further studied the TF type using more GC samples, 915 tumors. The tumor stage was finely categorized into 6 stages according to the Japanese classification of GC [11]. The dominant TF type revealed a striking and gradual shift from I-TF toward N-TF with tumor progression (**Figure 5a**). There are two hypotheses; first, GC changes the TF type with tumor progression. Second, a tumor with N-TF is more aggressive than one with I-TF. We examined the rate of nodal involvement of each TF type tumor in the 6 stages. Nodal involvement was influenced by tumor stage but not by TF type (**Figure 5b**). Each TF type shows almost identical metastatic potential in each tumor stage.

In multivariate models, HNF4aP1/CDX2 were independently related to early stage tumor (**Table 6**). In the medical care of GC, the tumor stage is mostly decided via endoscopic observation, so HNF4aP1/CDX2 were less useful in routine medicine; however HNF4aP1/CDX2 were revealed to be related to tumor invasion and should be paid more attention in GC studies.

TF type and other factors

The G-TF type was correlated with upper GC, which is increasing in western countries in association with hyperpepcia, and with less association of *H. pylori* infection [16]; therefore, further investigation to clarify the significance of SOX2 expression in the histogenesis of GC might prevent the increase of upper GC.

The I-TF type correlated with middle and lower GC, where chronic atrophic gastritis and intestinal metaplasia are likely to occur. In the nontumorous state, the gastric foveolar epithelium expresses SOX2. Some inflammation, including *H. pylori* gastritis, induces intestinal metaplasia in the lower portion of the stomach. The intestinalized epithelium expresses CDX2 and HNF4aP1. TF expression in GC might be greatly influenced by the precancerous state of the gastric epithelium.

The G-TF type was correlated with Lauren's diffuse type and the I-TF type with the intestinal type; however, each histology of the Japanese classification showed a characteristic TF type. Papillary adenocarcinoma tends to be N-TF type. Papillary adenocarcinoma shows a worse outcome than other differentiated GC [17]. This aggressiveness might be related to N-TF type, which is frequent among advanced stage GC. Among undifferentiated adenocarcinoma, poorly differentiated adenocarcinoma and signet-ring cell carcinoma were N/G-TF type, and mucinous adenocarcinoma was N-TF type. The TF type is a useful marker to study the carcinogenesis of GC further.

Acknowledgments

This study is supported by grants-in-aid for scientific research on priority areas (20249022, 20590337) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. We declare that we have no conflict of interest.

Please address correspondence to: Hiroshi Uozaki, MD, PhD. Department of Pathology, Graduate School of Medicine, The University of Tokyo. 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033 Japan, FAX: +81-3-3815-8379; E-mail: uozaki-tky@umin.ac.jp

References

[1] Black MM, Opler SR and Speer FD. Microscopic structure of gastric carcinomas and their regional lymph nodes in relation to survival. Surg Gynecol Obstet 1954;98:725-734.

[2] Pevny LH and Lovell Badge R. Sox genes find their feet. Curr Opin Genet Dev 1997;7:338-344.

[3] Que J, Okubo T, Goldenring JR, Nam KT, Kuro-

tani R, Morrisey EE, Taranova O, Pevny LH and Hogan BL. Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. Development 2007; 134:2521-2531.

[4] Tsukamoto T, Inada K, Tanaka H, Mizoshita T, Mihara M, Ushijima T, Yamamura Y, Nakamura S and Tatematsu M. Down-regulation of a gastric transcription factor, Sox2, and ectopic expression of intestinal homeobox genes, Cdx1 and Cdx2: inverse correlation during progression from gastric/intestinal-mixed to complete intestinal metaplasia. J Cancer Res Clin Oncol 2004; 130:135-145.

[5] Otsubo T, Akiyama Y, Yanagihara K and Yuasa Y. SOX2 is frequently downregulated in gastric cancers and inhibits cell growth through cell-cycle arrest and apoptosis. Br J Cancer 2008; 98:824-831.

[6] Tsukamoto T, Mizoshita T and Tatematsu M. Gastric-and-intestinal mixed-type intestinal metaplasia: aberrant expression of transcription factors and stem cell intestinalization. Gastric Cancer 2006; 9:156-166.

[7] Silberg DG, Swain GP, Suh ER and Traber PG. Cdx1 and cdx2 expression during intestinal development. Gastroenterol 2000; 119:961-971.

[8] Mizoshita T, Tsukamoto T, Nakanishi H, Inada K, Ogasawara N, Joh T, Itoh M, Yamamura Y and Tatematsu M. Expression of Cdx2 and the phenotype of advanced gastric cancers: relationship with prognosis. J Cancer Res Clin Oncol 2003;129:727-734.

[9] Tanaka T, Jiang S, Hotta H, Takano K, Iwanari H, Sumi K, Daigo K, Ohashi R, Sugai M, Ikegame C, Umezu H, Hirayama Y, Midorikawa Y, Hippo Y, Watanabe A, Uchiyama Y, Hasegawa G, Reid P, Aburatani H, Hamakubo T, Sakai J, Naito M and Kodama T. Dysregulated expression of P1 and P2 promoter-driven hepatocyte nuclear factor-4alpha in the pathogenesis of human cancer. J Pathol 2006; 208:662-672.

[10] Kojima K, Kishimoto T, Nagai Y, Tanizawa T, Nakatani Y, Miyazaki M and Ishikura H. The expression of hepatocyte nuclear factor-4alpha, a developmental regulator of visceral endoderm, correlates with the intestinal phenotype of gastric adenocarcinomas. Pathol 2006; 38:548-554.

[11] Japanese Gastric Cancer Association. Japanese Classification of Gastric Carcinoma - 2nd English Edition -. Gastric Cancer. 1998; 1:10-24.

[12] Lauren P. The two histological main types of gastric carcinoma. Diffuse and so-called intestinal type carcinoma. An attempt at histoclinical classification Acta Pathol Microbiol Scand, 1965;64:31-49.

[13] Barua RR, Uozaki H, Chong JM, Ushiku T, Hino R, Chang MS, Nagai H and Fukayama M. Phenotype analysis by MUC2, MUC5AC, MUC6, and CD10 expression in Epstein-Barr virus-associated gastric carcinoma. J Gastroenterol 2006;41:733-739.

[14] Park ET, Gum JR, Kakar S, Kwon SW, Deng G and Kim YS. Aberrant expression of SOX2 upregulates MUC5AC gastric foveolar mucin in mucinous cancers of the colorectum and related lesions. Int J Cancer 2008; 122:1253-1260.

[15] Tani Y, Akiyama Y, Fukamachi H, Yanagihara K and Yuasa Y. Transcription factor SOX2 upregulates stomach-specific pepsinogen A gene expression. J Cancer Res Clin Oncol 2007; 133:263-269.

[16] Gallo A and Cha C. Updates on esophageal and gastric cancers. World J Gastroenterol 2006; 12:3237-3242.

[17] Adachi Y, Yasuda K, Inomata M, Sato K, Shiraishi N and Kitano S. Pathology and prognosis of gastric carcinoma: well versus poorly differentiated type. Cancer 2000; 89:1418-1424.