Oncogene (2013), 1–10

www.nature.com/onc

ORIGINAL ARTICLE

The *OCT4* pseudogene *POU5F1B* is amplified and promotes an aggressive phenotype in gastric cancer

H Hayashi[1,2], T Arao[1], Y Togashi[1], H Kato[1], Y Fujita[1], MA De Velasco[1], H Kimura[1], K Matsumoto[1], K Tanaka[2], I Okamoto[2], A Ito[3], Y Yamada[4], K Nakagawa[2] and K Nishio[1]

POU5F1B (POU domain class 5 transcription factor 1B), a processed pseudogene that is highly homologous to *OCT4*, was recently shown to be transcribed in cancer cells, but its clinical relevance and biological function have remained unclear. We now show that *POU5F1B*, which is located adjacent to *MYC* on human chromosome 8q24, is frequently amplified in gastric cancer (GC) cell lines. *POU5F1B*, but not *OCT4*, was also found to be expressed at a high level in GC cell lines and clinical specimens. In addition, the DNA copy number and mRNA abundance for *POU5F1B* showed a positive correlation in both cancer cell lines and GC specimens. Overexpression of *POU5F1B* in GC cells promoted colony formation *in vitro* as well as both tumorigenicity and tumor growth *in vivo*, and these effects were enhanced in the additional presence of *MYC* overexpression. Furthermore, knockdown of *POU5F1B* expression with a short hairpin RNA confirmed a role for the endogenous pseudogene in the promotion of cancer cell growth *in vitro* and tumor growth *in vivo*. *POU5F1B* overexpression induced upregulation of various growth factors in GC cells as well as exhibited mitogenic, angiogenic and antiapoptotic effects in GC xenografts. Finally, amplification of *POU5F1B* was detected in 17 (12%) of 145 cases of GC and was a significant predictor of poor prognosis in patients with stage IV disease. In conclusion, we found that the *POU5F1B* pseudogene is amplified and expressed at a high level in, as well as confers an aggressive phenotype on, GC, and that *POU5F1B* amplification is associated with a poor prognosis in GC patients.

Oncogene advance online publication, 23 December 2013; doi:10.1038/onc.2013.547

Keywords: POU5F1B; OCT4; pseudogene; gastric cancer; gene amplification

INTRODUCTION

The gene for octamer-binding transcription factor 4 (OCT4), a POU-domain transcription factor, is expressed specifically in cells of the early mammalian embryo, including embryonic stem cells, and it has a key role in maintenance of pluripotency and self-renewal ability.[1] OCT4 is also one of the four factors that have been found to reprogram somatic cells and thereby give rise to induced pluripotent stem cells.[2] Two isoforms of human OCT4 mRNA, designated OCT4A and OCT4B, are generated by alternative splicing,[3] with only the OCT4A protein functioning as a stem cell factor.[3,4] *OCT4* is also expressed in germ cell tumors and epithelial cell tumors,[5–7] suggesting that OCT4-positive cancer cells likely represent cancer stem cells.

Pseudogenes are genomic loci that resemble a parental gene and are often considered to be nonfunctional because of the presence of defects that either prevent transcription or result in the generation of a nonfunctional protein.[8] Microarray technology has allowed detection of pseudogene transcription at the whole-genome level and has thus revealed that numerous pseudogenes are indeed transcribed. Compilation of RNA-seq transcriptome data focusing specifically on pseudogene transcripts identified 2082 pseudogenes that are expressed in human cells.[9] Although many of these pseudogenes appeared to be expressed ubiquitously or nearly so, 218 were found to be expressed only in cancer cells and 40 were restricted to a single cancer type.

To date, six *OCT4*-related pseudogenes that are highly homologous to *OCT4A* have been identified in the human genome, with the high level of homology having given rise to some confusion in cancer research because of the difficulty in discriminating between *OCT4* and its pseudogenes.[10] Two of these *OCT4* pseudogenes were found to be transcribed in cancer cells:[11] *POU5F1B* (POU domain class 5 transcription factor 1B gene, also known as *OCT4-pg1*, *OTF3C*, *OTF3P1* and *POU5F1P1*), which has been mapped to chromosomal band 8q24, encodes a protein that is 95% homologous identical to OCT4A, whereas *POU5F1P5*, which is localized to human chromosome 10q21, is theoretically incapable of producing a protein from the identified transcripts.[12] Although *POU5F1B* is known to be expressed in some cancer tissues,[13,14] its expression status and biological function have remained largely uncharacterized in cancer and other cell types.

With the use of array-based comparative genomic hybridization at the whole-genome level, we previously found that several human gastric cancer (GC) cell lines manifest amplification of the proto-oncogene *MYC*.[15] A subsequent detailed analysis revealed that *POU5F1B* was frequently co-amplified with *MYC* in the amplicon. Given that the implications of the expression of *OCT4* and its pseudogenes in human cancer have remained unclear,[11,16,17] we have now investigated the clinical relevance and biological function of *POU5F1B*.

[1]Department of Genome Biology, Kinki University Faculty of Medicine, Osaka-Sayama, Japan; [2]Department of Medical Oncology, Kinki University Faculty of Medicine, Osaka-Sayama, Japan; [3]Department of Pathology, Kinki University Faculty of Medicine, Osaka-Sayama, Japan and [4]Department of Gastrointestinal Oncology, National Cancer Center Hospital, Tokyo, Japan. Correspondence: Professor K Nishio, Department of Genome Biology, Kinki University Faculty of Medicine, 377-2 Ohno-higashi, Osaka-Sayama 589-8511, Japan.
E-mail: knishio@med.kindai.ac.jp
Received 3 April 2013; revised 12 November 2013; accepted 15 November 2013

RESULTS

POU5F1B is co-amplified with *MYC* in GC cell lines

Array-based comparative genomic hybridization analysis of nine human GC cell lines showed that the region of chromosome 8q24 around *MYC* is amplified in the 58As1, HSC44, 44As3, HSC39 and SNU-16 lines (Figure 1a). Real-time PCR-based analysis of DNA copy number revealed amplification of *POU5F1B* together with *MYC* in four of these GC cell lines but not in SNU-16 (Figure 1b). Fluorescence *in situ* hybridization (FISH) analysis of the 58As1, 44As3, HSC39 and TU-KATOIII cell lines with break-apart probes for *MYC* (an orange probe that includes the *POU5F1B* region centromeric to *MYC* and a green probe that includes the region telomeric to *MYC*) showed directly that *POU5F1B* was markedly ($\geqslant$20 copies) and focally amplified in 58As1, 44As3 and HSC39 cells, whereas it was present in only three or four copies in TU-KATOIII cells (Figure 1c). Amplification was mostly apparent as regions of homogeneous staining. Further analysis of various cancer cell lines revealed that *POU5F1B* amplification was largely restricted to GC lines but was also apparent in some breast and colorectal cancer lines (Figure 1d).

POU5F1B is expressed in GC cell lines

Given that *OCT4* and its pseudogenes are difficult to discriminate from one another because of their high level of homology, we examined the expression of *POU5F1B* in GC cell lines with the use of reverse transcription (RT) and PCR analysis with specific primers after taking precautions to avoid genomic DNA contamination. A schematic representation of transcripts corresponding to *OCT4*, *POU5F1B*, *POU5F1P3* and *POU5F1P4* based on previous studies[10,18,19] and BLAST searches is shown in Figure 2a (see also Supplementary Table S1).

POU5F1B transcripts, as detected with *POU5F1B*-specific primer pairs (POU5F1B-1 and POU5F1B-2), were apparent at high levels in several GC cell lines but not in the OCT4A-positive human testicular embryonal carcinoma cell line NTERA-2 (Figure 2b). As expected, no obvious band was detected with HEK293 cells, which have been examined as a negative control in previous studies.[13] In contrast, *OCT4A* transcripts, as detected with an *OCT4A*-specific primer pair (OCT4FA) designed to discriminate *OCT4A* from *OCT4B* and *OCT4* pseudogenes,[3] were apparent at a high level in NTERA-2 cells but were present at only low levels in GC cell lines. A primer set (OCT4FL) designed to detect transcripts of *POU5F1B*, *OCT4A*, *POU5F1P3* and *POU5F1P4* yielded a band of the expected size in both GC and NTERA-2 cell lines. These results thus indicated that GC cell lines preferentially express *POU5F1B* rather than *OCT4A* and other *OCT4* pseudogenes, whereas NTERA-2 cells preferentially express *OCT4A*.

Further evaluation of POU5F1B and OCT4A mRNA abundance in various cancer cell lines by RT and real-time PCR analysis with

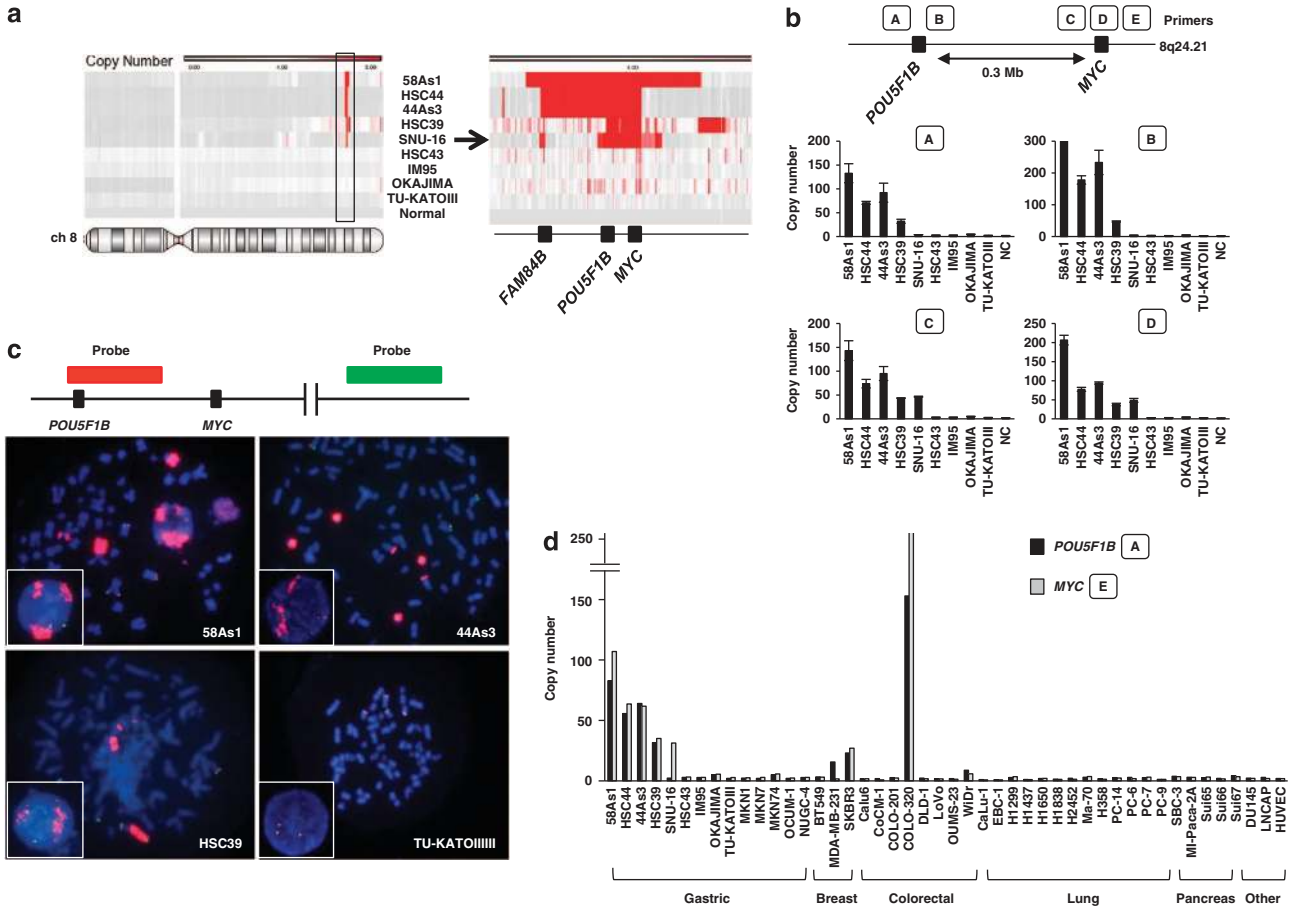

Figure 1. Amplification of *POU5F1B* and *MYC* in GC cell lines. (**a**) Array-based comparative genomic hybridization analysis of DNA copy number for chromosomal region 8q24 in nine GC cell lines. Gene amplification ($\geqslant$5 copies) is shown in red. Normal, normal human leukocytes. (**b**) Real-time PCR-based analysis of DNA copy number for *POU5F1B* and *MYC* in GC cell lines with primers targeting the indicated genomic loci A to E. NC, normal control. Data are means ± s.d. for three independent experiments. (**c**) FISH analysis of 58As1, 44As3, HSC39 and TU-KATOIII cell lines with the indicated (orange and green) break-apart probes for *MYC*. Main and inset images were metaphase and interphase FISH, respectively. (**d**) Real-time PCR-based analysis of DNA copy number for *POU5F1B* and *MYC* in 45 human cancer cell lines and human umbilical vein endothelial cells (HUVEC). Data are means of triplicates from a representative experiment.

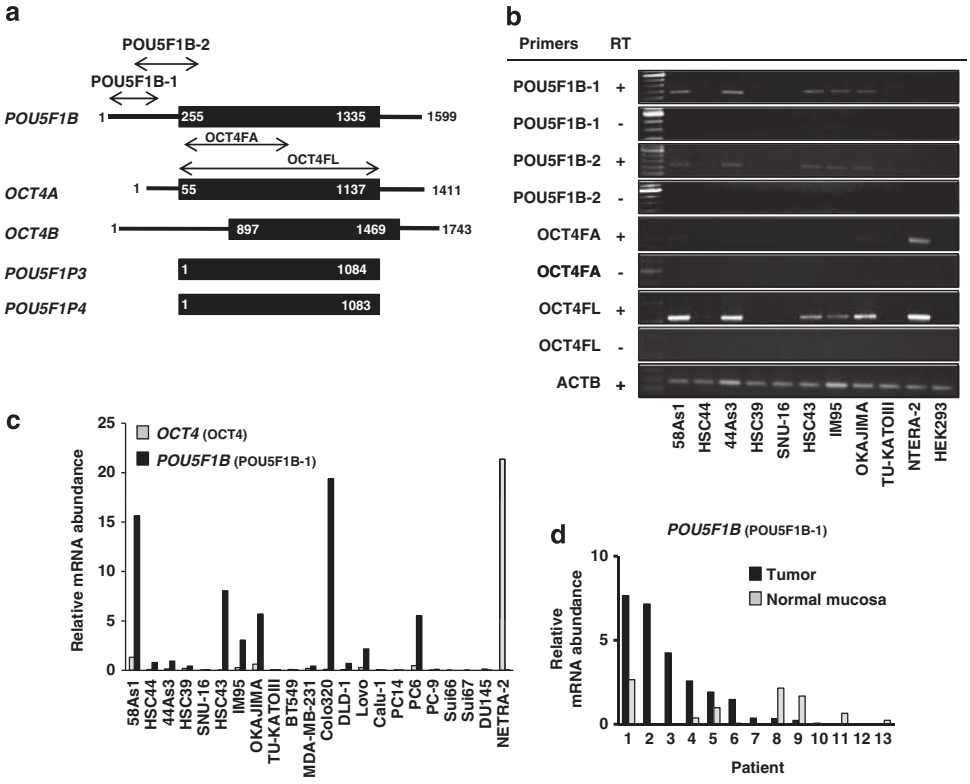

Figure 2. Expression of *POU5F1B* in cancer cell lines and clinical specimens. (**a**) Schematic representation of transcripts derived from *OCT4* and its pseudogenes. The thick and thin black bars indicate coding and noncoding regions, respectively. Regions amplified by the indicated PCR primer sets (POU5F1B-1, POU5F1B-2, OCT4FA and OCT4FL) are also shown. (**b**) RT–PCR analysis of *POU5F1B* and *OCT4* expression in GC cell lines with *POU5F1B*-specific (POU5F1B-1 and POU5F1B-2), *OCT4A*-specific (OCT4FA) and *OCT4* full-length (OCT4FL) primers. The β-actin gene (*ACTB*) was examined as an internal control. No-RT control incubations without reverse transcriptase were also performed to rule out genomic DNA contamination. NTERA-2 and HEK293 cells were examined as a positive control for *OCT4A* expression and a negative control, respectively. Lane M contains molecular size standards. (**c**, **d**) RT and real-time PCR analysis of *POU5F1B* and *OCT4* transcripts with the indicated primer sets in 22 human cancer cell lines (**c**) and 13 paired specimens of GC and noncancerous gastric mucosa (**d**). Relative mRNA abundance represents the *POU5F1B*/*GAPDH* or *OCT4*/*GAPDH* transcript ratio.

specific primers (Supplementary Table S2) revealed that *POU5F1B* is overexpressed in the 58As1 and Colo320 cell lines (Figure 2c), both of which manifested *POU5F1B* amplification (Figure 1d). Several GC or lung cancer cell lines that did not show *POU5F1B* amplification (including HSC43, IM95, OKAJIMA and PC6) were also found to overexpress *POU5F1B*, presumably as a result of transcriptional activation. *OCT4A* was overexpressed only in NTERA-2 cells. A significant correlation ($\rho = 0.56$, $P = 0.02$) was apparent between *POU5F1B* copy number (Figure 1d) and transcript abundance (Figure 2c), suggesting that *POU5F1B* amplification is responsible for *POU5F1B* overexpression in cancer cells.

RT and real-time PCR analysis of 13 paired clinical specimens of GC and normal mucosa revealed that the amount of POU5F1B mRNA tended ($P = 0.098$) to be increased in the tumor tissue compared with the paired normal gastric mucosa (Figure 2d). Furthermore, analysis of paired DNA and RNA samples prepared from 16 GC specimens revealed a significant correlation ($R = 0.47$, $P < 0.05$) between *POU5F1B* copy number and transcript abundance (Supplementary Figure S1).

We also evaluated the expression of *OCT4A* and *OCT4* pseudogenes in cancer cell lines and clinical GC specimens by DNA sequencing of cloned RT–PCR products amplified with the OCT4FL primer set, as previously described.[13] We found that *POU5F1B* accounted for most such clones in the 58As1 and 44As3 GC cell lines as well as in the GC tumor specimens, whereas *OCT4A* expression was largely limited to NTERA-2 cells (Table 1). *POU5F1P3* and *POU5F1P4* accounted for most clones derived from MCF7 and MDA-MB-231 breast cancer and U251 glioma cell lines. The DNA sequencing analysis thus confirmed that *POU5F1B* is expressed predominantly in GC cell lines and clinical samples.

Table 1. Expression of *POU5F1B*, *OCT4A*, *POU5F1P3* and *POU5F1P4* in cancer cell lines and clinical specimens as determined by DNA sequencing of cloned RT–PCR products

	POU5F1B, n *(%)*	OCT4A, n *(%)*	POU5F1P3, n *(%)*	POU5F1P4, n *(%)*	*Total clones*
58As1	16 (100)	0 (0)	0 (0)	0 (0)	16
44As3	16 (84)	1 (5)	0 (0)	2 (11)	19
MCF7	3 (17)	0 (0)	3 (17)	12 (67)	18
MDA-MB-231	3 (18)	1 (6)	0 (0)	13 (76)	17
U251	1 (5)	0 (0)	1 (5)	17 (89)	19
NTERA-2	0 (0)	20 (100)	0 (0)	0 (0)	20
GC 1	11 (79)	0 (0)	0 (0)	3 (21)	14
GC 2	10 (71)	0 (0)	0 (0)	4 (29)	14

RT–PCR products generated with the OCT4FL primer set specific for *POU5F1B*, *OCT4A*, *POU5F1P3* and *POU5F1P4* were cloned and sequenced. The number and percentage of sequenced clones corresponding to each target were determined. GC 1 and 2, clinical specimens of gastric cancer.

Overexpression of *POU5F1B* promotes colony formation *in vitro*

To investigate the function of *POU5F1B* in cancer cells, we established GC cell lines that stably express the enhanced green

fluorescent protein (EGFP) gene either alone or together with *POU5F1B* (Figure 3a). For these experiments, we used TU-KATOIII and SNU-16 cells, in both of which endogenous *POU5F1B* is expressed at a relatively low level (Figure 2c). A clonogenic assay revealed that the number of colonies formed by TU-KATOIII or SNU-16 cells overexpressing *POU5F1B* (TK3/POU and SNU/POU cells, respectively) was significantly greater than that formed by the corresponding control (TK3/EGFP and SNU/EGFP, respectively) cells (Figure 3b).

POU5F1B promotes tumorigenicity and tumor growth *in vivo*

We next examined whether *POU5F1B* might affect tumorigenicity or tumor growth *in vivo*. To assess tumorigenicity, we injected TK3/EGFP or TK3/POU cells subcutaneously into the flank of nude mice. The tumorigenicity of TK3/POU cells was significantly greater than that of the control cells (Figure 3c). We also established HEK/POU and HEK/EGFP cells from HEK293 cells as a control and found that overexpression of *POU5F1B* in these cells also significantly increased their tumorigenicity (Figure 3c). In addition, an *in vivo* assay of tumor growth revealed that the volume of tumors formed by TK3/POU or SNU/POU cells was markedly greater than that of those formed by the corresponding control cells (Figure 3d). Together, these results suggested that overexpression of *POU5F1B* confers an aggressive phenotype on GC cells. To validate these findings, we introduced either a short hairpin RNA (shRNA) that targets *POU5F1B* or a control shRNA into 58As1 and Colo320 cells, in both of which *POU5F1B* is highly amplified (Figure 1d) and expressed (Figure 2c). The stable transfectants were designated 58As1/sh-POU, 58As1/sh-Cont, Colo/sh-POU and Colo/sh-Cont, respectively. Whereas RT and real-time PCR analysis with *OCT4*-specific primers revealed that expression of the *OCT4* parental gene was not affected by the POU5F1B shRNA, the expression level of *POU5F1B* in cells harboring this shRNA was reduced to 5–18% of that apparent in the corresponding control cells (Figure 3e). Knockdown of *POU5F1B* expression in 58As1 or Colo320 cells resulted in marked inhibition of both colony formation *in vitro* (Figure 3f) and tumor growth *in vivo* (Figure 3g), consistent with the results obtained with cells overexpressing *POU5F1B*.

POU5F1B promotes angiogenesis and cell proliferation and inhibits apoptosis

To gain insight into the changes in gene expression induced by *POU5F1B* overexpression, we performed microarray analysis with TK3/EGFP and TK3/POU cells. Overexpression of *POU5F1B* resulted in the upregulation of 35 genes by a factor of $\geqslant 4$, with these genes including several that contribute to tumor growth such as those for amphiregulin, interleukin-8 (IL-8) and fibroblast growth factor 21 (FGF21) (Figure 4a). RT and real-time PCR analysis, immunoblot analysis and an enzyme-linked immunosorbent assay (ELISA) confirmed these results for amphiregulin and IL-8 (Figure 4b and c). Furthermore, the clonogenic assay showed that knockdown of IL-8 or amphiregulin (Figure 4d) resulted in significant inhibition of colony formation by TK3/POU cells (Figure 4e), suggesting that the upregulation of IL-8 and amphiregulin gene expression contributes to the promotion of colony formation induced by *POU5F1B* overexpression. On the other hand, with the exception of the gene for leucine-rich repeat-containing G protein-coupled receptor 5 (LGR5), few stemness-related genes were modulated by *POU5F1B* overexpression, suggesting that *POU5F1B* might not contribute to stemness but rather may regulate other genes related to angiogenesis or cell proliferation.

To examine cell proliferation and angiogenesis in tumors formed by TK3/EGFP, SNU/EGFP, TK3/POU or SNU/POU cells *in vivo*, we performed immunohistochemical staining for Ki67 and CD31, respectively. Overexpression of *POU5F1B* markedly promoted both tumor cell proliferation (Figure 5a) and tumor angiogenesis (Figure 5b). We also evaluated the effect of *POU5F1B* overexpression on doxorubicin-elicited apoptosis in GC cells *in vitro*, given that doxorubicin was previously shown to induce apoptosis mainly via the p53–Fas–caspase-8 system.[20] The doxorubicin-induced cleavage of poly(ADP-ribose) polymerase (PARP) and caspase-3 was markedly attenuated in TK3/POU and SNU/POU cells compared with that apparent in the corresponding control cells (Figure 5c). In addition, the abundance of the cleaved forms of PARP and caspase-3 was greatly reduced in tumor xenografts formed by TK3/POU cells than in those formed by TK3/EGFP cells (Figure 5d). The terminal deoxynucleotidy transferase-mediated dUTP nick-end labeling (TUNEL) assay also revealed that apoptosis was significantly attenuated in tumors formed by TK3/POU or SNU/POU cells compared with that in those formed by the corresponding control cells (Figure 5e). These results thus suggested that POU5F1B exerts an antiapoptotic effect both *in vivo* and *in vitro*.

POU5F1B promotes colony formation and tumor growth cooperatively with *MYC*

The GC cell line SNU-16 exhibits *MYC* amplification and overexpression, whereas the TU-KATOIII cell line exhibits a low copy number and low expression level for *MYC* (Figure 1d, Supplementary Table S3). Given that demonstration of a role for *POU5F1B* as an oncogene in the 8q24 amplicon is complicated by its tight linkage with *MYC*, we investigated the relation between *POU5F1B* and *MYC* function by establishing TK3/EGFP and TK3/POU cell lines that stably overexpress *MYC* as well as corresponding mock-transfected cells (designated TK3/EGFP/MYC, TK3/EGFP/Mock, TK3/POU/MYC and TK3/POU/Mock cells). RT and real-time PCR and immunoblot analyses of these cell lines revealed that overexpression of *POU5F1B* did not interfere with that of *MYC* and vice versa (Figure 6a). The clonogenic assay revealed that colony formation by TK3/POU/Mock and TK3/POU/MYC cells was significantly enhanced compared with that by TK3/EGFP/Mock and TK3/EGFP/MYC cells, respectively ($P = 0.01$ and 0.04, respectively) (Figure 6b). Furthermore, the growth of tumors formed after subcutaneous injection of cells into nude mice increased according to the rank order TK3/EGFP/Mock, TK3/EGFP/MYC, TK3/POU/Mock and TK3/POU/MYC (Figure 6c). These results suggested that overexpression of *POU5F1B* confers an aggressive phenotype on GC cells in a cooperative manner with *MYC*.

Amplification of *POU5F1B* is associated with poor prognosis in patients with advanced GC

Finally, we examined the prognostic impact of *POU5F1B* and *MYC* amplification in a cohort of 145 patients with advanced GC. *POU5F1B* and *MYC* amplification were detected in 8 (13.3%) and 10 (16.7%) of the 60 tumors diagnosed as stage III, respectively, and in 9 (10.6%) and 14 (16.5%) of the 85 tumors diagnosed as stage IV, respectively. Kaplan–Meier analysis revealed that *MYC* amplification did not affect the overall survival of the patients with GC of stage III (mean of 2551 vs 1674 days, $P = 0.48$), whereas *POU5F1B* amplification tended to be associated with a poor overall survival in these individuals (989 vs 2040 days, $P = 0.07$) (Figure 7a). For patients with GC of stage IV, *POU5F1B* and *MYC* amplification were each significantly associated with a shorter overall survival (320 vs 870 days for *POU5F1B*, $P = 0.0003$; 499 vs 870 days for *MYC*, $P = 0.02$) (Figure 7b). Both *POU5F1B* and *MYC* amplification were thus associated with a significantly poorer prognosis in patients with GC of stage IV, and *POU5F1B* amplification seemed to be indicative of a poorer prognosis compared with *MYC* amplification.

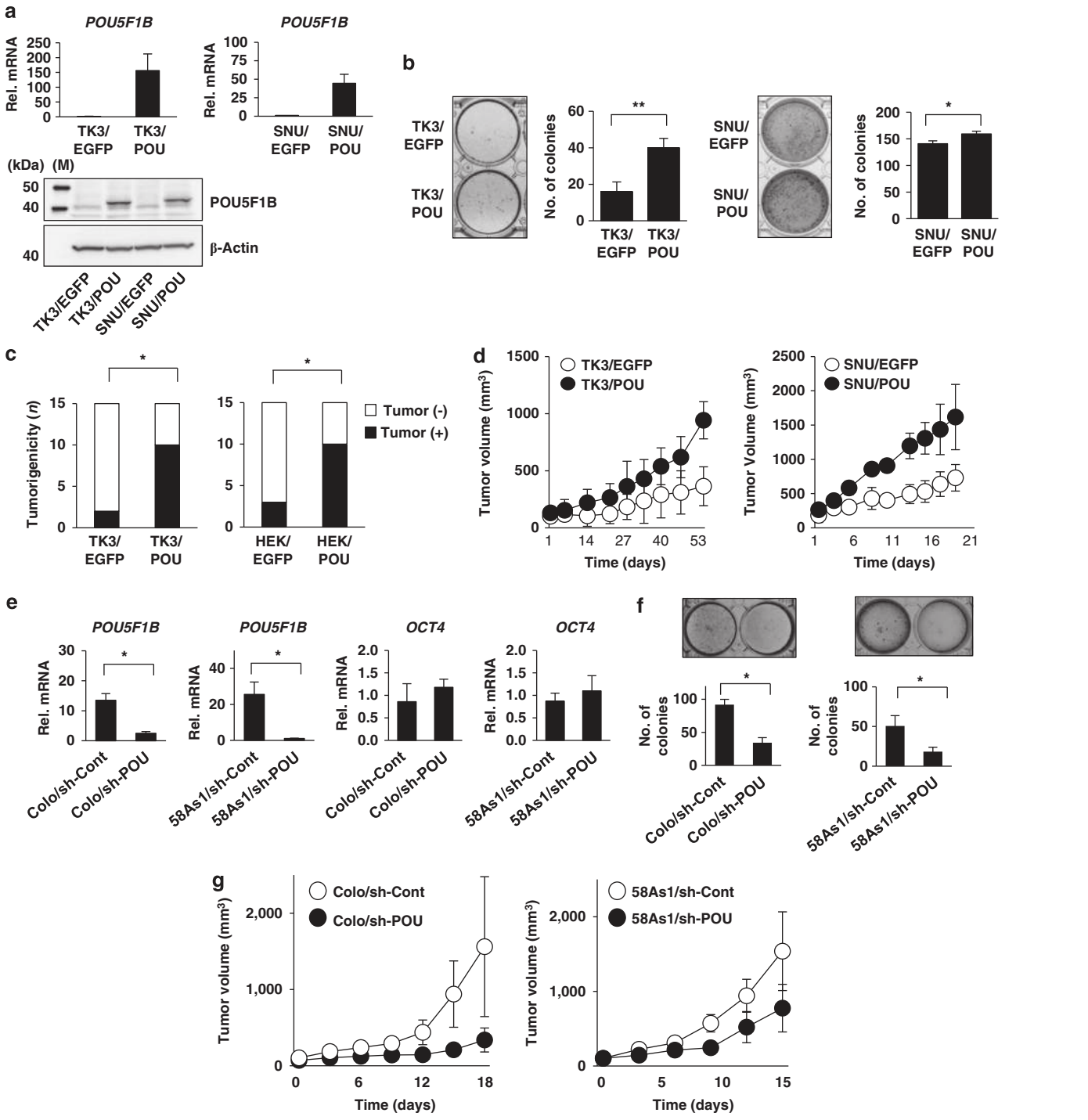

Figure 3. Effects of overexpression or depletion of POU5F1B on GC cell growth *in vitro* and *in vivo*. (**a**) TU-KATOIII or SNU-16 cells stably expressing EGFP either alone or together with POU5F1B were subjected to RT and real-time PCR analysis of POU5F1B mRNA (upper panels) or to immunoblot analysis of POU5F1B protein (with antibodies to OCT4 that recognize POU5F1B) (lower panels). Data on relative (Rel.) mRNA abundance correspond to the *POU5F1B/GAPDH* transcript ratio $\times 10^4$ and are means ± s.d. from three independent experiments. The blot was also probed for β-actin as an internal control. Lane M contains molecular size standards. (**b**) Clonogenic assay performed with TK3/EGFP, TK3/POU, SNU/EGFP and SNU/POU cells. Quantitative data are means ± s.d. from three independent experiments. $^*P<0.05$, $^{**}P<0.005$. (**c**) Tumorigenicity of TK3/EGFP, TK3/POU, HEK/EGFP and HEK/POU cells. Data represent the number of nude mice that developed tumors 4 weeks after subcutaneous injection of 2×10^5 (TK3) or 1×10^6 (HEK) cells into 15 animals for each group. $^*P<0.05$. (**d**) Growth of tumors formed after subcutaneous injection of 2×10^6 TK3/EGFP or TK3/POU cells or of 5×10^6 SNU/EGFP or SNU/POU cells into nude mice. Data are means ± s.d. for five mice of each group. Tumor volume was significantly larger for TK3/POU (942 ± 163 vs 364 ± 170 mm^3 on day 55, $P=0.001$) and SNU/POU (1618 ± 477 vs 731 ± 194 mm^3 on day 19, $P=0.01$) cells than for the corresponding control cells. (**e**) RT and real-time PCR analysis of POU5F1B and OCT4 mRNAs in Colo320 (Colo) or 58As1 cells stably expressing POU5F1B (sh-POU) or control (sh-Cont) shRNAs. Data are presented as the *POU5F1B/GAPDH* or *OCT4/GAPDH* transcript ratio and are means ± s.d. from three independent experiments. $^*P<0.05$. (**f**) Clonogenic assay performed with Colo/sh-Cont, Colo/sh-POU, 58As1/sh-Cont and 58As1/sh-POU cells. Quantitative data are means ± s.d. from three independent experiments. $^*P<0.05$. (**g**) Growth of tumors formed after subcutaneous injection of Colo/sh-Cont, Colo/sh-POU, 58As1/sh-Cont or 58As1/sh-POU cells (5×10^6) into nude mice Data are means ± s.d. for five mice of each group. Tumor volume was significantly smaller for Colo/sh-POU (338 ± 157 vs 1562 ± 917 mm^3 on day 18, $P=0.03$) and 58As1/sh-POU (776 ± 319 vs 1539 ± 528 mm^3 on day 15, $P=0.02$) cells than for the corresponding control cells.

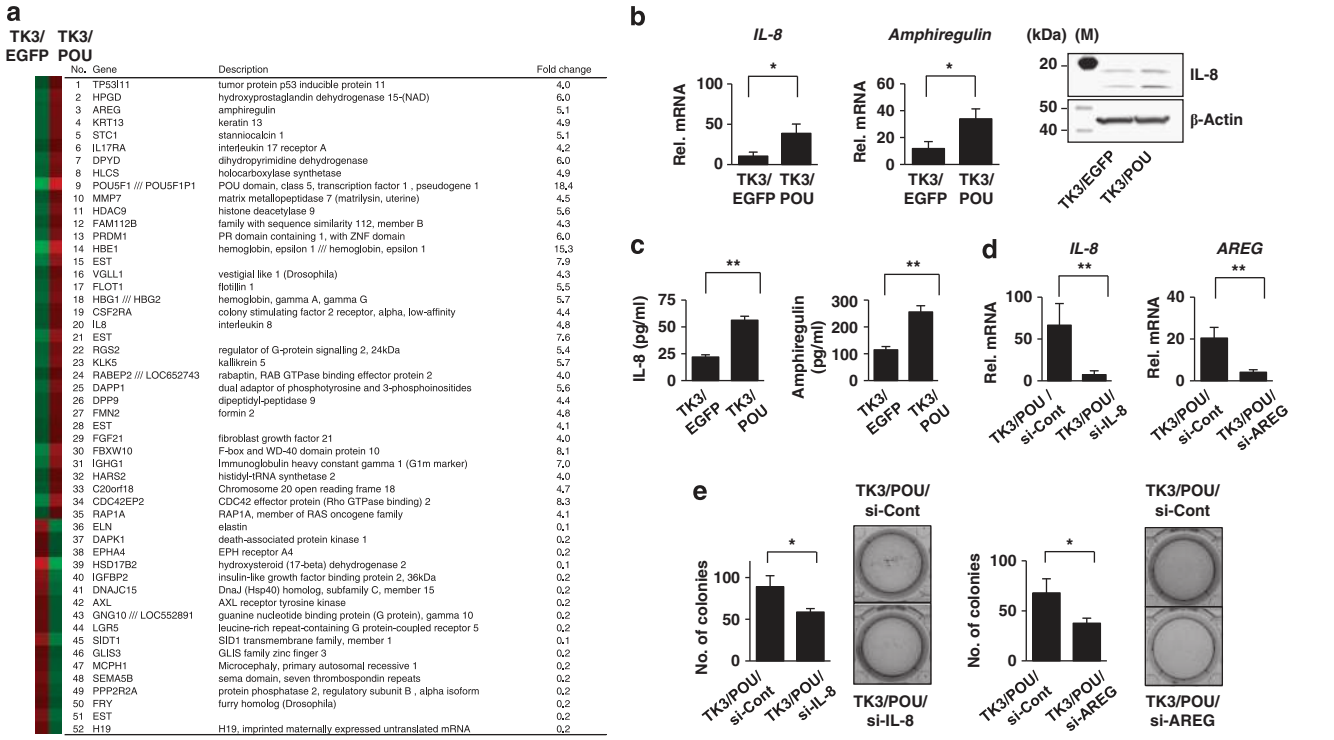

Figure 4. Effects of *POU5F1B* overexpression on gene expression in GC cells. (**a**) Microarray analysis of gene expression in TK3/EGFP and TK3/POU cells. Genes that were up- or downregulated by a factor of $\geqslant 4$ are listed. (**b**) RT and real-time PCR analysis of IL-8 and amphiregulin mRNAs as well as immunoblot analysis of IL-8 in TK3/EGFP and TK3/POU cells. Lane M contains molecular size standards. Relative mRNA abundance represents the IL-8/GAPDH or amphiregulin/GAPDH mRNA ratio $\times 10^4$ and is presented as means ± s.d. from three independent experiments. $^*P<0.05$. (**c**) Release of IL-8 and amphiregulin by TK3/EGFP or TK3/POU cells into the culture medium was determined with an ELISA. Data are means ± s.d. from three independent experiments. $^{**}P<0.005$. (**d**) RT and real-time PCR analysis of IL-8 and amphiregulin (AREG) mRNAs in TK3/POU cells transiently expressing corresponding targeted or control small interfering RNAs (siRNAs). Data are presented as the IL-8/GAPDH or amphiregulin/GAPDH mRNA ratio $\times 10^4$ and are means ± s.d. from three independent experiments. $^{**}P<0.005$. (**e**) Clonogenic assay performed with TK3/POU cells depleted of IL-8 or amphiregulin. Quantitative data are means ± s.d. from three independent experiments; $^*P<0.05$.

DISCUSSION

We have shown that amplification of *POU5F1B* was associated with a high expression level of the pseudogene in several cancer cell lines, most prominently in the 58As1 and Colo320 lines. A positive correlation was thus detected between *POU5F1B* copy number and mRNA abundance in cancer cell lines as well as in clinical specimens from GC patients. As far as we are aware, this is the first demonstration that amplification of a pseudogene gives rise to an increased abundance of the corresponding mRNA and confers a malignant phenotype in cancer cells. The amount of POU5F1B mRNA was also relatively high in the HSC43, IM95 and OKAJIMA cell lines despite the absence of *POU5F1B* amplification, suggesting that another mechanism of *POU5F1B* overexpression may be operative in these cells, possibly one involving transcriptional activation due to the presence of single-nucleotide polymorphisms or mutations in the promoter or enhancer regions.

Amplification of chromosomal region 8q24, including *MYC*, has been found to occur in ~10–30% of GCs.[21] Consistent with this finding, we detected *MYC* amplification in 17% of GCs examined in the present study. The frequency of *POU5F1B* amplification (12%) in the present study was lower than that of *MYC* amplification because of the presence of two major types of amplicon: a shorter amplicon containing *MYC* alone present in ~30% of cases, and a longer amplicon containing both *MYC* and *POU5F1B* present in the remaining cases (data not shown). The longer amplicon appeared to be associated with a poorer prognosis than the shorter one in patients with advanced GC. Several large-scale genome-wide association studies have revealed that genomic loci in the vicinity of *POU5F1B* are associated with susceptibility to various cancers including prostate, colorectal, breast and ovarian tumors,[22–28] providing further support for the malignant potential of *POU5F1B*. In addition, we found that overexpression of *POU5F1B* increased cell growth *in vitro* as well as tumorigenicity and tumor growth *in vivo*, and these effects were more pronounced in the presence of *MYC* overexpression. These results suggest that *POU5F1B* acts cooperatively with *MYC* to promote tumor formation and growth. Although the mechanism underlying the induction of an aggressive phenotype by *POU5F1B*, especially in cooperation with *MYC*, remains to be fully elucidated, our results suggest that the upregulation of various growth factors and cytokines and the consequent promotion of cell proliferation and angiogenesis and attenuation of apoptosis may have a role.

POU5F1B might be expected to possess stemness potential because of its high level of homology with *OCT4A*. However, the transcriptional activity of *POU5F1B* was previously found to be relatively low compared with that of *OCT4A*.[12] Consistent with this finding, overexpression of *POU5F1B* did not confer a stemness phenotype on cancer cells in a sphere assay nor did it affect the expression of stem cell markers in the present study (data not shown).

The biological relevance of pseudogenes remains largely uninvestigated because they are often considered to be neither transcribed nor translated. However, increasing evidence suggests a role for pseudogenes in cellular processes such as tumorigenesis.[29–31] We found that *POU5F1B* is not only transcribed but is also capable of promoting cell growth and of producing a functional protein *in vivo*. In general, pseudogenes are capable of regulating the expression of tumor suppressor genes and oncogenes by

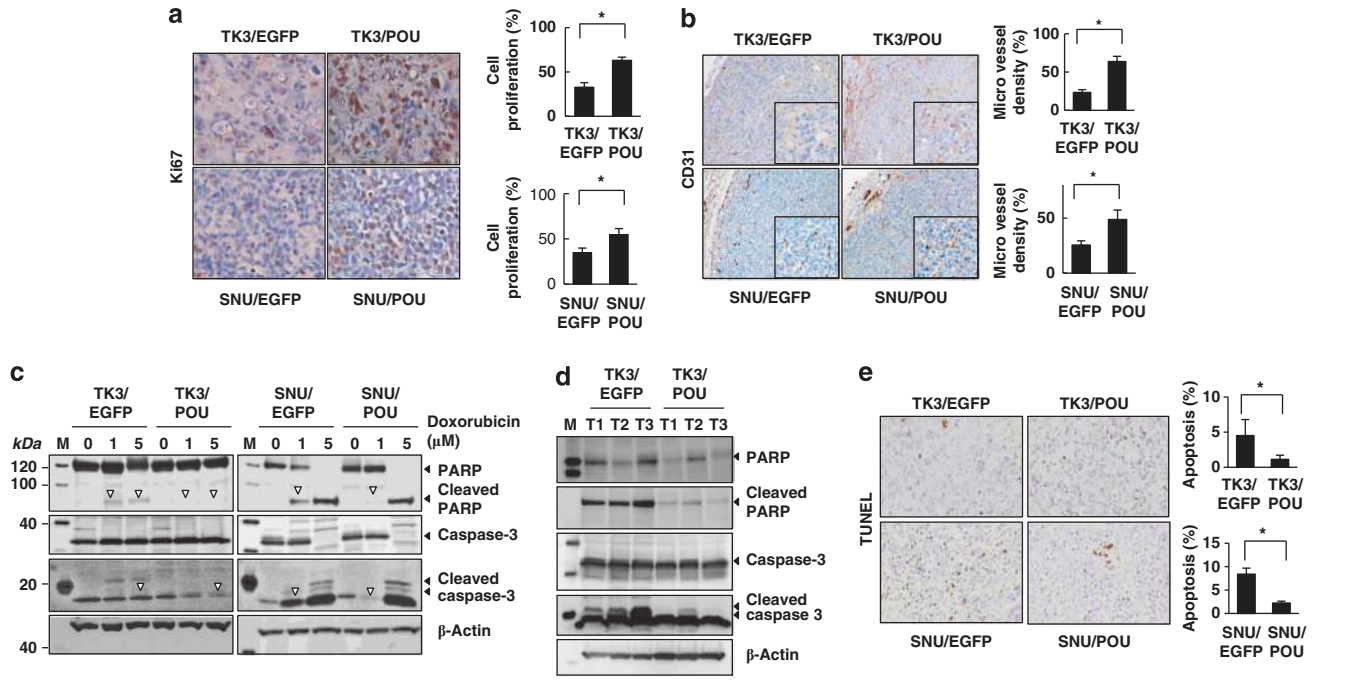

Figure 5. Overexpression of *POU5F1B* promotes cell proliferation and angiogenesis and inhibits apoptosis. (**a**) Immunohistochemical staining of Ki67 in tumors formed by TK3/EGFP, TK3/POU, SNU/EGFP or SNU/POU cells in nude mice. Cell proliferation was quantitated as the percentage of Ki67-positive tumor cells and is presented as means ± s.d. for three tumors of each group. *$P<0.05$. (**b**) Immunohistochemical staining of CD31 in tumors formed by TK3/EGFP, TK3/POU, SNU/EGFP or SNU/POU cells in nude mice. Insets show higher magnification images. Microvessel density was determined on the basis of the number of CD31-positive endothelial cells in the tumor specimens and is presented as means ± s.d. for three tumors of each group. *$P<0.05$. (**c**) TK3/EGFP, TK3/POU, SNU/EGFP or SNU/POU cells were exposed to the indicated concentrations of doxorubicin for 24 h *in vitro* and then subjected to immunoblot analysis with antibodies to intact or cleaved forms of PARP or caspase-3. Open arrowheads indicate marked differences in the amounts of the cleaved forms of PARP or caspase-3 between the cells with or without *POU5F1B* overexpression. Lanes M contain molecular size standards. (**d**) Immunoblot analysis of PARP and caspase-3 in individual (T1–T3) tumor xenografts formed by TK3/EGFP or TK3/POU cells in nude mice. (**e**) TUNEL staining of tumor xenografts formed by the indicated cell lines in nude mice. The percentage of apoptotic cells was determined as means ± s.d. for three tumors of each group; *$P<0.05$.

acting as microRNA decoys.[32] Given the relatively low expression level of the parental gene *OCT4A* in cancer cells, however, *POU5F1B* likely does have play a substantial role as a microRNA decoy in regulation of *OCT4A* expression. Our findings suggest that the POU5F1B protein promotes tumor formation and growth, at least in GC cells.

In conclusion, *POU5F1B* is amplified and overexpressed in GC, and it promotes tumorigenicity and tumor growth *in vivo*. Furthermore, *POU5F1B* amplification was identified as a novel prognostic factor for patients with advanced GC.

MATERIALS AND METHODS

Cell lines and culture

The HSC39, HSC43, HSC44, OKAJIMA, SNU-16, TU-KATOIII, 44As3 and 58As1 cell lines were cultured in RPMI 1640 medium (Sigma, St Louis, MO, USA), whereas IM95, NTERA-2 and HEK293 cells were cultured in Dulbecco's modified Eagle's medium (Sigma). Both media were supplemented with 10% heat-inactivated fetal bovine serum (Life Technologies BRL, Carlsbad, CA, USA), and all cells were maintained under a humidified atmosphere of 5% CO_2 at 37 °C.

Array-based comparative genomic hybridization

Genomic DNA (250 ng) from GC cell lines was digested with both *Nsp*I and *Sty*I or with *Nsp*I alone in independent parallel reactions for comparative genomic hybridization with the Genome-wide Human SNP Array 6.0 or the GeneChip Human Mapping 250K Nsp Array (Affymetrix, Santa Clara, CA, USA), respectively. The digested DNA was ligated to the adapter and amplified by PCR with a universal primer and TITANIUM Taq DNA Polymerase (Clontech, Mountain View, CA, USA). The PCR products were then quantified, fragmented, end-labeled and hybridized with the arrays. The arrays were then washed and stained in a Fluidics Station 450 (Affymetrix) before scanning to generate .CEL files with the use of a GeneChip Scanner 3000 and GeneChip Operating Software version 1.4 (Affymetrix).

Assay of DNA copy number

The DNA copy numbers for *MYC* and *POU5F1B* were determined with the use of TaqMan Copy Number Assays (Applied Biosystems, Foster City, CA, USA), as previously described,[33] and with the primers Hs06213526_cn and Hs05028821_cn for *POU5F1B* and Hs03693129_cn, Hs03660964_cn and Hs04346708_cn for *MYC*. The *TERT* locus was assayed as an internal reference control. Human genomic DNA (TaKaRa, Otsu, Japan) and formalin-fixed paraffin-embedded (FFPE) normal tissue samples were examined as normal controls. The criterion for gene amplification (copy number of $\geqslant 5$) was based on that adopted in previous studies.[33,34]

FISH analysis

Amplification of the 8q24 region was examined by FISH with the use of the LSI (locus-specific identifier) *MYC* BAP (break-apart probe) set (Abbott/Vysis, Wiesbaden, Germany), which consists both of a spectrum orange-labeled 5′ LSI *MYC* probe that begins 119 kb upstream of the 5′ end of *MYC* and extends 277 kb toward the centromere (including the region containing *POU5F1B*) as well as a spectrum green-labeled 3′ LSI *MYC* probe that begins at ~1.5 Mb downstream of *MYC* and extends toward the telomere for 407 kb. The hybridization protocols and evaluation criteria were described previously.[35]

RT–PCR analysis

Total RNA was extracted from cell lines with the use of an RNeasy Kit (Qiagen, Hilden, Germany) and was treated with DNase I to remove

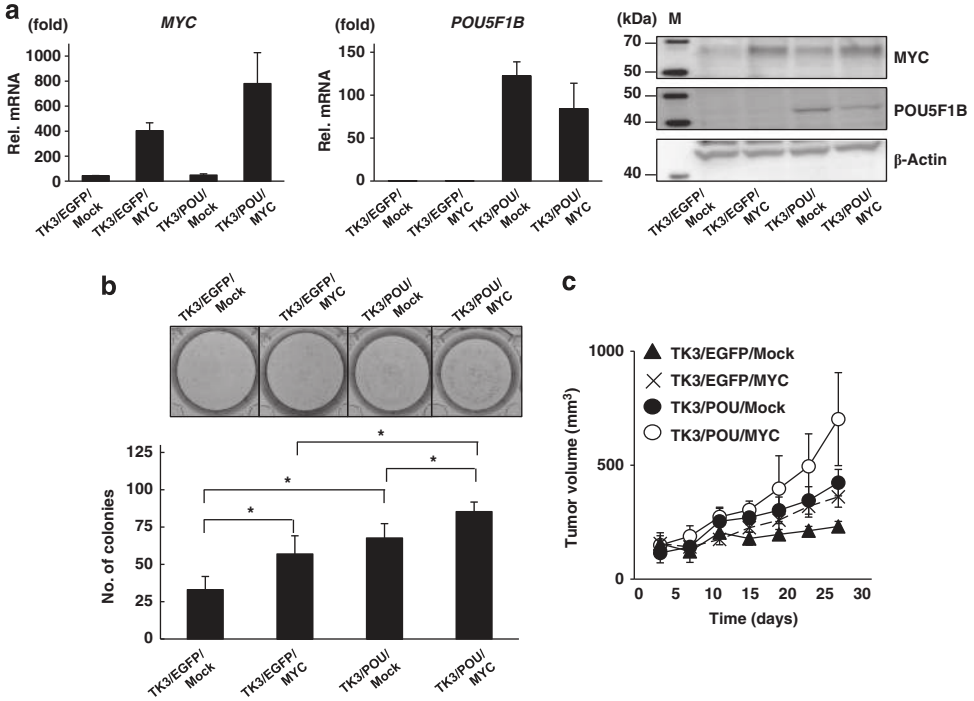

Figure 6. Overexpression of *POU5F1B* promotes colony formation and tumor growth in cooperation with overexpression of *MYC*. (**a**) RT and real-time PCR and immunblot analyses of *POU5F1B* and *MYC* expression in TK3/EGFP/Mock, TK3/EGFP/MYC, TK3/POU/Mock and TK3/POU/MYC cells. Relative mRNA abundance represents the *POU5F1B/GAPDH* transcript ratio $\times 10^4$ or the *MYC/GAPDH* transcript ratio $\times 10^3$ and is presented as means ± s.d. from three independent experiments. The blot was probed with antibodies to MYC and to OCT4 (for detection of POU5F1B). Lane M contains molecular size standards. (**b**) Clonogenic assay for TK3/EGFP/Mock, TK3/EGFP/MYC, TK3/POU/Mock and TK3/POU/MYC cells. Data are means ± s.d. from three independent experiments; $*P<0.05$. (**c**) Growth of tumors formed in nude mice after subcutaneous injection of 5×10^6 cells of each line. Data are means ± s.d. for five mice of each group.

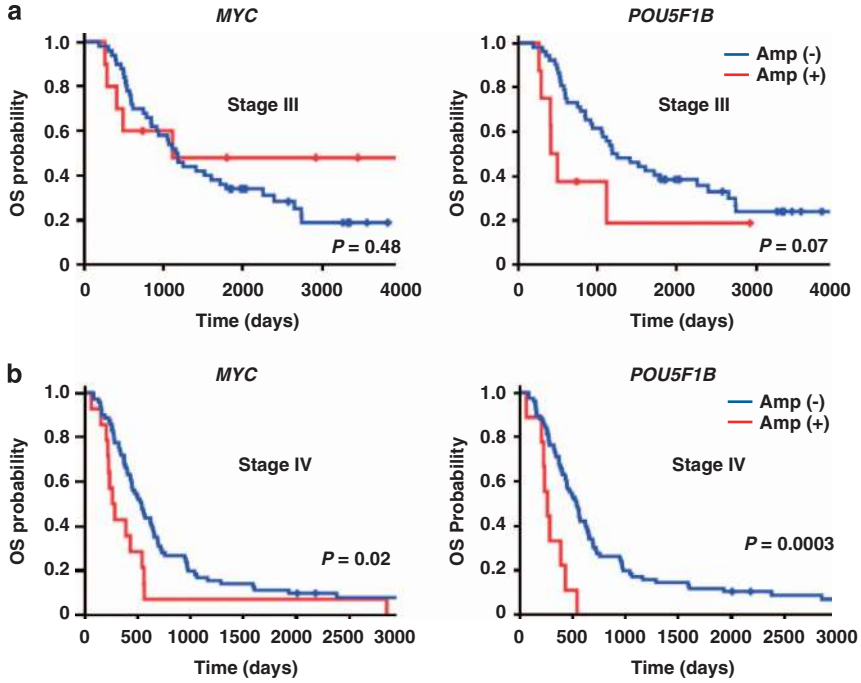

Figure 7. Kaplan–Meier plots of overall survival for GC patients according to amplification of *MYC* or *POU5F1B*. (**a**) Plots for 60 patients with GC of stage III. (**b**) Plots for 85 patients with GC of stage IV. Gene amplification (Amp) was defined as the presence of ⩾5 copies. The *P*-values were obtained with the log-rank test.

contaminating genomic DNA with the use of an RNase-Free DNase Set (Qiagen). RT and PCR were performed as previously described.[36] In brief, total RNA (1 μg) was converted to cDNA and PCR was then performed with the use of a GeneAmp RNA-PCR Core Kit (Applied Biosystems). β-actin (ACTB) mRNA was examined as an internal control. The primer sequences are shown in Supplementary Table S1. The reaction mixtures were first incubated at 94 °C for 5 min and then subjected to the following thermal profiles: 35 cycles of 95 °C for 30 s, 58 °C for 30 s and 72 °C for 30 s (POU5F1B-1, POU5F1B-2 and ACTB); 35 cycles of 95 °C for 30 s, 60 °C for 30 s and 72 °C for 35 s (OCT4FA); or 35 cycles of 94 °C for 60 s, 64 °C for 45 s and 72 °C for 60 s (OCT4FL). All mixtures were then subjected to a final extension at 72 °C for 7 min. The PCR products were separated by electrophoresis on a 1.0% agarose gel, stained with ethidium bromide and visualized under ultraviolet light.

RT and real-time PCR and immunoblot analyses

RT and real-time PCR as well as immunblot analyses were performed as described previously.[37] With the exception of POU5F1B-1, the primers used for real-time PCR were obtained from TaKaRa (Supplementary Table S4). For immunoblot analysis, antibodies to OCT4 and to β-actin were obtained from Santa Cruz Biotechnology (Santa Cruz, CA, USA), those to intact or cleaved forms of PARP or caspase-3 and to c-MYC were from Cell Signaling Technology (Danvers, MA, USA), and those to IL-8 were from Abcam (Cambridge, UK).

Cloning and sequencing analysis

RT–PCR products were extracted from the agarose gel with the use of a MiniElute Gel Extraction Kit (Qiagen) and were then cloned into the pCR-2.1-TOPO vector (Invitrogen, Carlsbad, CA, USA). The resulting constructs were introduced into DH5α-competent cells (TOYOBO, Osaka, Japan) and plasmids with an insert were subsequently sequenced in both directions with the use of M13 forward and reverse primers.

Plasmid construction, virus production and cell transfection

Plasmid construction, virus production and cell transfection were performed as previously described.[37] A cDNA encoding full-length human POU5F1B was isolated by PCR with Prime STAR HS DNA polymerase (TaKaRa) and with the sense and antisense primers 5′-CGGGATCCCGACCATGGATGGCGGGA CACCTGGCTT-3′ and 5′-CGGAATTCCGTCAGTTTGAATGCATGGGAGAG-3′, respectively. The vector pMX-MYC (plasmid 17220) was obtained from Addgene (Cambridge, MA, USA). Stable cotransfectants of TU-KATOIII cells expressing MYC together with EGFP or POU5F1B were established from bulk transfected cells rather than from cloned single cells. Vectors for a shRNA targeting *POU5F1B* and for a control shRNA were constructed from oligonucleotides encoding the corresponding small interfering RNAs (siRNAs), 5′-CATACGGTCACAGAGCTTTTT-3′ and 5′-AACTAGTATAGATCGATT CTT-3′, respectively. These sequences were cloned into the RNAi-Ready pSIREN-RetroQZsGreen vector (Clontech).

Clonogenic assay of cell growth

Cells were isolated by exposure to trypsin to yield a single-cell suspension and were then seeded in six-well plates at a density of 1000 cells per well and cultured for 10 days. The resulting colonies were stained with crystal violet, and those containing at least 50 cells were counted.

ELISA

Cells (5×10^5 per well of a six-well plate) were cultured in complete medium for 12 h and then exposed to serum-free medium for 12 h. The latter culture supernatants were then collected, centrifuged to remove floating cells and assayed for amphiregulin and IL-8 with the use of a DuoSet ELISA Development Kit (R&D Systems, Minneapolis, MN, USA).

Xenograft studies

Xenograft studies were performed as described previously.[38] Six-week-old female nude (BALB/c nu/nu) mice were obtained from CLEA Japan (Tokyo, Japan) and were cared for in accordance with recommendations in the Handling of Laboratory Animals for Biomedical Research compiled by the Committee on Safety and Ethical Handling Regulations for Laboratory Animal Experiments, Kinki University. For assessment of tumorigenicity and tumor growth, cell suspensions in 0.1 ml of phosphate-buffered saline were injected subcutaneously into the left or right flank of the mice.

Microarray analysis

Microarray analysis was performed according to Affymetrix protocols with BRB Array Tools software, version 3.3.0, developed by Richard Simon and Amy Peng, as described previously.[15,37]

Patients and specimens

Amplification of *POU5F1B* and *MYC* in DNA samples from 145 GC patients was examined as previously described.[33] In brief, macrodissection of FFPE tissue specimens was performed to select a cancer region, which was marked by a pathologist after paraffin depletion. Genomic DNA was extracted from the cancer tissue with the use of a QIAamp DNA Micro Kit (Qiagen). Total RNA was also isolated from 16 of the 145 GC specimens in order to investigate the relation between expression and amplification of *POU5F1B*. The RNA was extracted with the use of an RNeasy FFPE Kit (Qiagen), treated with DNase I to remove contaminating genomic DNA and subjected to RT and real-time PCR analysis as described previously.[22] In addition, cDNA samples from 13 paired GC and noncancerous gastric mucosa specimens were described previously.[39] These analyses were approved by the local institutional review board, and informed consent was obtained from all patients.

Immunohistochemical analysis

Paraffin sections were subjected to immunohistochemical staining with antibodies to Ki67 (MIB-1, 1/50 dilution; Dako, Glostrup, Denmark) and to CD31 (NCH-38, 1:200 dilution; Dako) according to standard procedures.[36] Random microscopic fields of tumor tissue were identified, digitally captured and saved for quantification with ImageJ software (http://rsb.info.nih.gov/ij). The stained slides were independently examined and scored by two investigators. For evaluation of Ki67, the percentage of positively stained malignant nuclei was determined for five fields of each slide. Cancer cells undergoing apoptosis were detected by the TUNEL method, using the *in situ* apoptosis detection TUNEL kit (Takara) according to the manufacturer's recommended protocol.

Statistical analysis

Quantitative data are presented as means ± s.d. and were analyzed with Student's *t*-test as performed with SPSS version 14.0 software (SPSS, Chicago, IL, USA). The relation between copy number and expression for *POU5F1B* was evaluated with Spearman's rank correlation coefficient. Survival analysis was performed by the Kaplan–Meier method, and the curves were compared with the log-rank test. A *P*-value of <0.05 was considered statistically significant.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

ACKNOWLEDGEMENTS

This study was supported by the Third-Term Comprehensive 10-Year Strategy for Cancer Control and a Grant-in-Aid for Cancer Research from the Ministry of Health, Labor and Welfare of Japan. We thank Shinji Kurashimo, Yoshihiro Mine, Ayaka Kurumatani and Tomoko Kitayama for technical assistance.

REFERENCES

1 Scholer HR, Ruppert S, Suzuki N, Chowdhury K, Gruss P. New type of POU domain in germ line-specific protein Oct-4. *Nature* 1990; **344**: 435–439.
2 Takahashi K, Yamanaka S. Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 2006; **126**: 663–676.
3 Atlasi Y, Mowla SJ, Ziaee SA, Gokhale PJ, Andrews PW. OCT4 spliced variants are differentially expressed in human pluripotent and nonpluripotent cells. *Stem Cells* 2008; **26**: 3068–3074.
4 Cauffman G, Liebaers I, Van Steirteghem A, Van de Velde H. POU5F1 isoforms show different expression patterns in human embryonic stem cells and preimplantation embryos. *Stem Cells* 2006; **24**: 2685–2691.
5 Atlasi Y, Mowla SJ, Ziaee SA, Bahrami AR. OCT-4, an embryonic stem cell marker, is highly expressed in bladder cancer. *Int J Cancer* 2007; **120**: 1598–1602.
6 Monk M, Holding C. Human embryonic genes re-expressed in cancer cells. *Oncogene* 2001; **20**: 8085–8091.

7 Wu CL, Shieh GS, Chang CC, Yo YT, Su CH, Chang MY *et al*. Tumor-selective replication of an oncolytic adenovirus carrying oct-3/4 response elements in murine metastatic bladder cancer models. *Clin Cancer Res* 2008; **14**: 1228–1238.
8 Pink RC, Wicks K, Caley DP, Punch EK, Jacobs L, Carter DR. Pseudogenes: pseudo-functional or key regulators in health and disease? *RNA* 2011; **17**: 792–798.
9 Kalyana-Sundaram S, Kumar-Sinha C, Shankar S, Robinson DR, Wu YM, Cao X *et al.* Expressed pseudogenes in the transcriptional landscape of human cancers. *Cell* 2012; **149**: 1622–1634.
10 Liedtke S, Enczmann J, Waclawczyk S, Wernet P, Kogler G. Oct4 and its pseudogenes confuse stem cell research. *Cell Stem Cell* 2007; **1**: 364–366.
11 Suo G, Han J, Wang X, Zhang J, Zhao Y, Dai J. Oct4 pseudogenes are transcribed in cancers. *Biochem Biophys Res Commun* 2005; **337**: 1047–1051.
12 Panagopoulos I, Moller E, Collin A, Mertens F. The POU5F1P1 pseudogene encodes a putative protein similar to POU5F1 isoform 1. *Oncol Rep* 2008; **20**: 1029–1033.
13 Zhao S, Yuan Q, Hao H, Guo Y, Liu S, Zhang Y *et al*. Expression of OCT4 pseudogenes in human tumours: lessons from glioma and breast carcinoma. *J Pathol* 2011; **223**: 672–682.
14 Kastler S, Honold L, Luedeke M, Kuefer R, Moller P, Hoegel J *et al.* POU5F1P1, a putative cancer susceptibility gene, is overexpressed in prostatic carcinoma. *Prostate* **70**: 666–674.
15 Furuta K, Arao T, Sakai K, Kimura H, Nagai T, Tamura D *et al*. Integrated analysis of whole genome exon array and array-comparative genomic hybridization in gastric and colorectal cancer cells. *Cancer Sci* 2012; **103**: 221–227.
16 Schoenhals M, Kassambara A, De Vos J, Hose D, Moreaux J, Klein B. Embryonic stem cell markers expression in cancers. *Biochem Biophys Res Commun* 2009; **383**: 157–162.
17 Cantz T, Key G, Bleidissel M, Gentile L, Han DW, Brenne A *et al.* Absence of OCT4 expression in somatic tumor cell lines. *Stem Cells* 2008; **26**: 692–697.
18 Wang X, Dai J. Concise review: isoforms of OCT4 contribute to the confusing diversity in stem cell biology. *Stem Cells* 2010; **28**: 885–893.
19 Pain D, Chirn GW, Strassel C, Kemp DM. Multiple retropseudogenes from pluripotent cell-specific gene expression indicates a potential signature for novel gene identification. *J Biol Chem* 2005; **280**: 6265–6268.
20 Friesen C, Herr I, Krammer PH, Debatin KM. Involvement of the CD95 (APO-1/FAS) receptor/ligand system in drug-induced apoptosis in leukemia cells. *Nat Med* 1996; **2**: 574–577.
21 Calcagno DQ, Leal MF, Assumpcao PP, Smith MA, Burbano RR. MYC and gastric adenocarcinoma carcinogenesis. *World J Gastroenterol* 2008; **14**: 5962–5968.
22 Fletcher O, Johnson N, Gibson L, Coupland B, Fraser A, Leonard A *et al.* Association of genetic variants at 8q24 with breast cancer risk. *Cancer Epidemiol Biomarkers Prev* 2008; **17**: 702–705.
23 Gudmundsson J, Sulem P, Manolescu A, Amundadottir LT, Gudbjartsson D, Helgason A *et al.* Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24. *Nat Genet* 2007; **39**: 631–637.
24 Haiman CA, Le Marchand L, Yamamato J, Stram DO, Sheng X, Kolonel LN *et al.* A common genetic risk factor for colorectal and prostate cancer. *Nat Genet* 2007; **39**: 954–956.
25 Schumacher FR, Feigelson HS, Cox DG, Haiman CA, Albanes D, Buring J *et al.* A common 8q24 variant in prostate and breast cancer from a large nested case-control study. *Cancer Res* 2007; **67**: 2951–2956.
26 Yeager M, Orr N, Hayes RB, Jacobs KB, Kraft P, Wacholder S *et al.* Genome-wide association study of prostate cancer identifies a second risk locus at 8q24. *Nat Genet* 2007; **39**: 645–649.
27 Zanke BW, Greenwood CM, Rangrej J, Kustra R, Tenesa A, Farrington SM *et al.* Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. *Nat Genet* 2007; **39**: 989–994.
28 Amundadottir LT, Sulem P, Gudmundsson J, Helgason A, Baker A, Agnarsson BA *et al.* A common variant associated with prostate cancer in European and African populations. *Nat Genet* 2006; **38**: 652–658.
29 Bier A, Oviedo-Landaverde I, Zhao J, Mamane Y, Kandouz M, Batist G. Connexin43 pseudogene in breast cancer cells offers a novel therapeutic target. *Mol Cancer Ther* 2009; **8**: 786–793.
30 Yu H, Zhao X, Su B, Li D, Xu Y, Luo S *et al.* Expression of NF1 pseudogenes. *Hum Mutat* 2005; **26**: 487–488.
31 Zhang J, Wang X, Li M, Han J, Chen B, Wang B *et al.* NANOGP8 is a retrogene expressed in cancers. *FEBS J* 2006; **273**: 1723–1730.
32 Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP. A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature* 2010; **465**: 1033–1038.
33 Matsumoto K, Arao T, Hamaguchi T, Shimada Y, Kato K, Oda I *et al.* FGFR2 gene amplification and clinicopathological features in gastric cancer. *Br J Cancer* 2012; **106**: 727–732.
34 Graziano F, Galluccio N, Lorenzini P, Ruzzo A, Canestrari E, D'Emidio S *et al.* Genetic activation of the MET pathway and prognosis of patients with high-risk, radically resected gastric cancer. *J Clin Oncol* 2011; **29**: 4789–4795.
35 Klapper W, Stoecklein H, Zeynalova S, Ott G, Kosari F, Rosenwald A *et al.* Structural aberrations affecting the MYC locus indicate a poor prognosis independent of clinical risk factors in diffuse large B-cell lymphomas treated within randomized trials of the German High-Grade Non-Hodgkin's Lymphoma Study Group (DSHNHL). *Leukemia* 2008; **22**: 2226–2229.
36 Kaneda H, Arao T, Tanaka K, Tamura D, Aomatsu K, Kudo K *et al.* FOXQ1 is overexpressed in colorectal cancer and enhances tumorigenicity and tumor growth. *Cancer Res* 2010; **70**: 2053–2063.
37 Tanaka K, Arao T, Maegawa M, Matsumoto K, Kaneda H, Kudo K *et al.* SRPX2 is overexpressed in gastric cancer and promotes cellular migration and adhesion. *Int J Cancer* 2009; **124**: 1072–1080.
38 Takeda M, Arao T, Yokote H, Komatsu T, Yanagihara K, Sasaki H *et al.* AZD2171 shows potent antitumor activity against gastric cancer over-expressing fibroblast growth factor receptor 2/keratinocyte growth factor receptor. *Clin Cancer Res* 2007; **13**: 3051–3057.
39 Yamada Y, Arao T, Gotoda T, Taniguchi H, Oda I, Shirao K *et al.* Identification of prognostic biomarkers in gastric cancer using endoscopic biopsy samples. *Cancer Sci* 2008; **99**: 2193–2199.

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