Table 1 continued

Neo-adjuvant chemotherapy, *n* (%)	
Present	54 (5.4)
Absent	952 (94.6)
Adjuvant chemotherapy, *n* (%)	
Present	67 (6.7)
Absent	939 (93.3)

[a] T, primary tumor

[b] N, regional lymph node

[c] R, residual tumor

total, 495 (49.2 %) patients were classified into the well-differentiated and 489 (48.6 %) into the poorly differentiated tumor category. Mucinous carcinoma was diagnosed in 21 (2.1 %) patients; there was 1 case of mixed endocrine and tubular adenocarcinoma. Although gastrectomy and systematic D2 regional lymph node dissection with curative intent was performed in 952 (94.6 %) patients, 4 were found to be microscopically margin positive. Splenectomy was added to total gastrectomy in patients when the tumor was located on the greater curvature of the stomach and where the tumor extended further than the submucosal layer of the stomach. Palliative resection was performed in 54 (5.4 %) patients. The median number of dissected lymph nodes and number of lymph nodes with metastasis were 37 (range 6–118) and 0 (range 0–68), respectively.

Although 127 patients died of gastric cancer, 28 patients died of other disease. Recurrence was observed in 152 patients. In the 948 patients who had undergone R0 resection, both median DSS and median RFS were longer than 5 years, and the cumulative 5-year disease-specific survival (5Y-DSS) rate and 5-year recurrence-free survival (5Y-RFS) rate were 84.5 and 83.3 %, respectively. Of the total of 1,006 patients, the 3Y-DSS and 5Y-DSS of the patients analyzed by the disease stage were as follows: 99.2 and 98.1 % in stage I patients, 91.2 and 86.5 % in stage II patients, and 68.1 and 61.0 % in stage III patients, respectively. The 3Y-DSS of the patients with stage IV disease was 23.2 %.

HER2 scores

Immunohistochemical analysis was performed to examine the expression of HER2 in all the cases (Fig. 1). The numbers of patients with an HER2 score of 0, 1+, 2+, and 3+ were 455 (45.2 %), 360 (35.8 %), 94 (9.3 %), and 97 (9.6 %), respectively. Concordance of the immunohistochemistry results for HER2 protein was evaluated between duplicate samples. The concordance rates for immunohistochemistry scores of 2+/3+ and 0/1+ between the samples were 74.5 and 94.7 %, respectively. In 94 patients

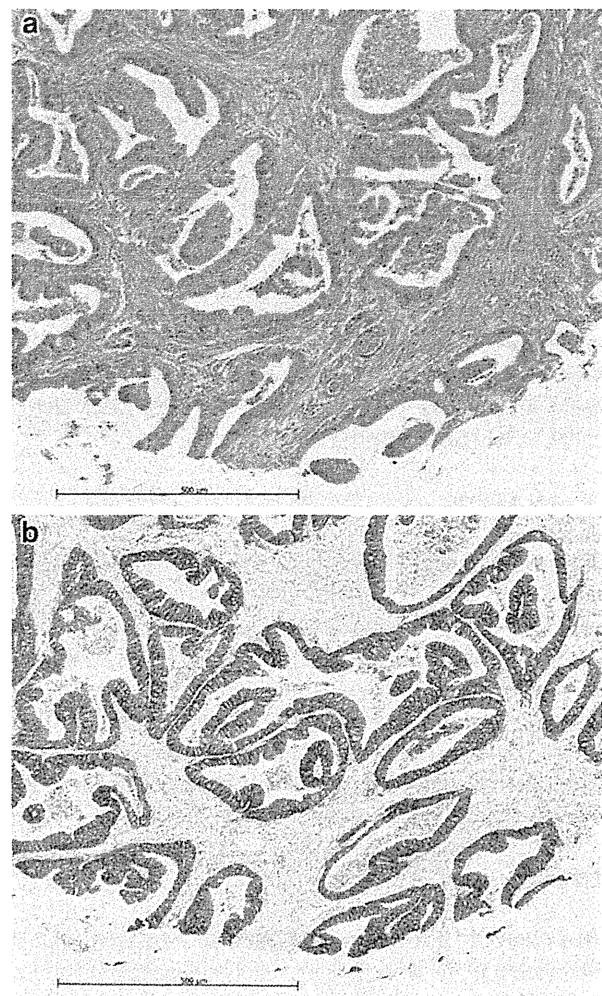

Fig. 1 Representative photomicrograph of a human epidermal growth factor receptor 2 (HER2) 3+ specimen. **a** Hematoxylin and eosin (H&E) staining shows a typical tubular arrangement of well-differentiated tumor cells. **b** Immunohistochemical staining for HER2 protein shows strong basolateral membrane staining of the tumor cells

with an immunohistochemistry score of 2+, 21 patients were also detected to have HER2 gene amplification by double-color fluorescent in situ hybridization. Finally, 118 (11.7 %) patients were defined as being positive for tumor HER2 overexpression.

Demographics, tumor-related factors and tumor HER2 overexpression

Correlations between HER2 overexpression and demographic and tumor-related factors are shown in Table 2. HER2 overexpression was more commonly observed in patients of greater age (14.1 %, $p = 0.024$), males (14.3 %, $p < 0.001$), patients with tumors with well-differentiated histology (20.2 %, $p < 0.001$), patients with

Table 2 Correlation of human epidermal growth factor receptor 2 (HER2) overexpression with demographics and tumor-related factors

Variables	HER2 negative, *n* (%)	HER2 positive, *n* (%)	*p* value
Total number	888 (88.3)	118 (11.7)	–
Age (years)			
≤65	477 (90.5)	50 (9.5)	0.024
>65	411 (85.9)	68 (14.1)	
Gender			
Male	580 (85.7)	97 (14.3)	<0.001
Female	308 (93.7)	21 (6.3)	
Tumor location			
Esophagogastric junction	28 (90.4)	3 (9.6)	0.497
Stomach	860 (88.2)	115 (11.8)	
Tumor diameter (cm), median (range)	4.0 (0.4–34.4)	4.9 (1.2–18.0)	0.219
Macroscopic tumor type			
Type 0	461 (88.5)	60 (11.5)	0.002
Type 1, 2	115 (79.3)	30 (20.7)	
Type 3, 4	296 (91.6)	27 (8.4)	
Type 5	16 (94.1)	1 (5.9)	
Grade of differentiation			
Well (tubular, papillary)	395 (79.8)	100 (20.2)	<0.001
Poorly (poorly, signet cell)	472 (96.6)	17 (3.4)	
Others	21 (95.5)	1 (4.5)	
Pathological (p) T[a] disease			
pT1	446 (88.3)	59 (11.7)	0.027
pT2	116 (91.4)	11 (8.6)	
pT3	194 (83.6)	38 (16.4)	
pT4	132 (93.0)	10 (7.0)	
Pathological N[b] disease			
pN0	550 (90.3)	59 (9.7)	0.038
pN1	116 (82.3)	25 (17.7)	
pN2	86 (83.7)	14 (16.3)	
pN3a	73 (79.5)	15 (20.5)	
pN3b	56 (92.9)	4 (7.1)	
Pathological stage			
I	487 (89.5)	57 (10.5)	0.556
II	184 (86.0)	30 (14.0)	
III	164 (87.7)	23 (12.3)	
IV	53 (86.9)	8 (13.1)	
Neo-adjuvant chemotherapy			
Absent	839 (86.7)	113 (13.3)	0.669
Present	49 (90.7)	5 (9.3)	

[a] T, primary tumor

[b] N, regional lymph node

tumors exhibiting expansive growth (macroscopic type 1 or type 2, $p = 0.002$), and patients with lymph node metastasis (N1 disease, 17.7 %; N2 disease, 16.3 %; N3a disease, 20.5 %; $p = 0.038$). Although differences in frequency of HER2 overexpression among the T stages were estimated to be significant ($p = 0.027$), no consistent gradient was observed. The presence of HER2 overexpression in the tumor was not influenced by tumor location, tumor size, pathological stage, or history of administration of neoadjuvant chemotherapy.

HER2 overexpression and the prognosis after surgery

The 5Y-DSS rates in the patients with and without HER2 overexpression who underwent R0 resection were 87.2 and 87.7 %, respectively. With respect to the 5-year RFS (5Y-RFS) rate after R0 resection, the percentage of patients with and without tumor HER2 overexpression was 81.1 and 83.5 %, respectively. HER2 overexpression in the tumors was not correlated with either DSS (log-rank test; $p = 0.14$) or RFS (log-rank test; $p = 0.718$). Furthermore, tumor HER2 overexpression was also not related to survival in the 58 inoperable patients (data not shown). Univariate analyses using the Cox proportional hazard model (Table 3) to identify the factors influencing DSS and RFS consistently failed to exhibit the prognostic significance of tumor HER2 overexpression. However, age >65 years, male gender, tumor location in the esophagogastric junction, tumor diameter ≥10 cm, tumor exhibiting infiltrative growth (macroscopic type 3 or type 4), T4 disease, and N disease were significant ominous prognostic factors, as previously known. The sites of recurrence in the patients are shown in Table 4. Although patients with tumor HER2 overexpression were predisposed to the development of solitary liver or lymph node metastasis, peritoneal seeding was more frequent in the patients without tumor HER2 overexpression. Surgery was more often performed for operable recurrences in patients with tumor HER2 overexpression, whereas best supportive care tended to be undertaken in patients not showing tumor HER2 overexpression (data not shown).

Discussion

Only R0 resection with prophylactic lymph node dissection has been established as a potentially curative treatment for gastric cancer [13]. Although the beneficial effect of optimal resection based on tumor-related factors would seem to be the maximum achievable for local control and survival, a substantial number of cases eventually show locoregional

Table 3 Univariate analysis by the COX proportional hazard model to identify factors influencing disease-specific survival (DSS) and recurrence-free survival (RFS) in 948 patients who had undergone no residual cancer (R0) resection

Variables	*n*	Risk for cancer-related death			Risk of recurrence		
		HR	95 % CI	Log-rank (*p* value)	HR	95 % CI	Log-rank (*p* value)
Age (years)							
≤65	497	1.000	–	–	1.000	–	–
>65	451	2.003	1.336–3.004	0.001	1.667	1.206–2.304	0.001
Gender							
Female	311	1.000	–	–	1.000	–	–
Male	637	1.542	0.980–2.426	0.059	1.388	0.969–1.988	0.841
Tumor diameter (cm)							
<10	805	1.000	–	–	1.000	–	–
≥10	54	5.016	2.936–8.568	0.004	4.583	3.148–7.480	<0.001
Tumor location							
Stomach	917	1.000	–	–	1.000	–	–
Esophagogastric junction	31	2.774	1.346–5.715	<0.001	3.625	2.090–6.287	<0.001
Macroscopic tumor type							
Type 1 or 2	140	1.000	–	–	1.000	–	–
Type 3 or 4	273	2.037	1.277–3.247	0.002	1.498	1.004–2.235	0.048
Well-differentiated histology	481	1.000	–	–	1.000	–	–
Poorly differentiated histology	450	1.354	0.927–1.979	0.718	1.234	0.910–1.673	0.116
Pathological (p) T[a] disease							
pT1	504	1.000	–	–	1.000	–	–
pT2	126	8.021	3.010–21.37	<0.001	6.891	3.154–15.06	<0.001
pT3	222	19.71	8.406–46.21	<0.001	21.61	11.18–41.77	<0.001
pT4	96	40.63	17.09–96.60	<0.001	33.73	16.93–67.08	<0.001
Pathological (p) N[b] disease							
pN0	605	1.000	–	–	1.000	–	–
pN1	136	3.421	1.676–6.982	0.003	4.532	2.616–7.850	<0.001
pN2	89	8.419	4.451–15.92	<0.001	9.227	5.476–15.55	<0.001
pN3a	79	12.34	6.649–22.91	<0.001	18.03	11.03–29.47	<0.001
pN3b	39	38.02	20.50–70.53	<0.001	31.94	18.32–55.69	<0.001
HER2-negative	836	1.000	–	–	1.000	–	–
HER2-positive	112	0.565	0.262–1.219	0.140	1.109	0.685–1.796	0.718

[a] T, primary tumor

[b] N, regional lymph node

failure. Multimodality therapies are now being assessed, and some clinical trials have already demonstrated the efficacy and safety of adjuvant chemotherapy administered after resection [14–16]. Recently, HER2 expression has attracted attention because of the encouraging results of the ToGA trial. Actually, the results provided the first clue to the potential benefit of tumor biology-based treatment in patients with gastric cancer. Despite the increasing importance of this relationship, the relationship between HER2 overexpression and tumor biology in cases of gastric cancer still remains to be elucidated.

Based on the results of the ToGA trial, several studies have evaluated HER2 overexpression in esophagogastric cancer using validated criteria [17–22]. The frequency of HER2 overexpression has been reported to be 9–16 %, consistent with the results of two major reviews of previous studies [4, 9]. However, the criterion for diagnosing HER2 overexpression still varies among studies. We evaluated HER2-based tumor biology in 1,006 cases of gastric cancer using the following criterion for HER2 overexpression in gastric cancer: immunohistochemistry score of 3+, or immunohistochemistry score of 2+ plus gene amplification; in the present study, using this criterion, the rate of HER2 overexpression was estimated to be 11.7 % in the present study. In a subgroup analysis of the ToGA trial conducted after excluding cases with tumors showing an

Table 4 Sites of recurrence in patients with and without tumor HER2 overexpression

	HER2 negative, *n* (%)	HER2 positive, *n* (%)
Incidence of recurrence	133	19
Site of recurrence, *n* (%)		
Peritoneum	38 (28.6)	1 (5.3)
Liver	23 (17.3)	9 (47.4)
Lymph node	23 (17.3)	5 (26.3)
Lung	3 (2.3)	1 (5.3)
Other sites	16 (12.0)	1 (5.3)
Multiple	26 (19.5)	2 (10.5)
Unclear	4 (3.0)	0 (0.0)

immunohistochemistry score of 0–1+ plus gene amplification, the gain in median survival was 4.2 months in the intent-to-treat population. This gain was superior to the actual gain in the median survival of 2.7 months under the initial entry criteria [10] and was regarded as comparable to the gain of 4.8 months observed in the trial of advanced breast cancer [2, 23]. For obtaining clinical benefit, the criterion for HER2 overexpression of an immunohistochemistry score of 3+ or 2+ plus gene amplification seemed to be the most reliable, and the present study evaluated a large number of patients of gastric cancer using this criterion.

Heterogeneity of HER2 expression level in gastric cancer cases has been suggested as another cause of the discordance of HER2 overexpression among studies [24]. TMA is a useful tool for analyzing a large number of samples, but its major limitation is that tumor cores may not represent the whole tissue section. A recent study that evaluated both whole tissue sections and TMA [19] concluded that use of duplicate TMA samples for each case could minimize the discordance. The concordance rates of immunohistochemistry scores of 2+/3+ and 0/1+ between duplicate TMAs constructed from the same tissue section in a patient were 74.5 and 91.8 %, respectively. Although this was in good agreement with the previously reported results of analysis of biopsy specimens from 261 patients with invasive intestinal-type gastric cancers [25], further investigations are still required to elucidate the reliability of HER2 scoring in TMA samples in gastric cancer.

In the present study, age, gender, growth pattern, grade of differentiation of the tumor, and N disease were found to be correlated with tumor HER2 overexpression (Table 2). Similarly, in the ToGA trial, tumor with HER2 overexpression was reported to be more common in patients with a well-differentiated tumor histology [17–19, 21]. Actually, well-differentiated tumor histology was significantly related to greater age, male gender, and an expansive growth pattern of the tumor in the present study (data not shown). Therefore, age, gender, and the growth pattern of tumors were indirectly related to the likelihood of HER2 overexpression in the tumors. On the other hand, lymph node metastasis was significantly more frequent in patients with poorly differentiated tumor histology (data not shown). Several previous studies have reported the existence of a relationship between the presence of lymph node metastasis and tumor HER2 overexpression [18, 19]. This finding suggests that tumor biology in the subgroup of patients with tumor HER2 overexpression predisposes to spread via the lymphatic system. This finding suggests the potential clinical benefit of HER2-targeted therapy in the adjuvant or neo-adjuvant setting for patients with node-positive gastric cancer. However, no significant correlation between HER2 overexpression and any of the factors of tumor location, depth of tumor (T stage), or pathological stage was observed in the present study.

In the survival analysis, no significant correlation was observed between HER2 overexpression and either DSS or RFS. Actually, recent studies in which the diagnosis of HER2 overexpression was based on immunohistochemical staining using criteria suggested by Hofmann et al. [9] reported conflicting results with respect to the prognostic significance of HER2 overexpression in gastric cancer [20–22]. The inconsistent results could be attributed to multiple factors. First, even the foregoing studies have not standardized treatment of patients having tumors with immunohistochemistry scores of 0 to 2+ plus HER2 gene amplification. As a subgroup analysis of the ToGA trial suggested the indispensable requirement of a different criterion for the diagnosis of HER overexpression in gastric cancer from that in breast cancer, further analysis with the criterion for HER2 overexpression set as an immunohistochemistry scores of 3+ or 2+ plus gene amplification, which is potentially the most reliable at present, would be needed. Second, the studies were of one accord in suggesting that a well-differentiated histology and lymph node metastasis were closely related to HER2 overexpression. However, these two were known as opposing prognostic factors, which could provide controversial results among studies. Gastric cancer with a well-differentiated histology or of the intestinal type generally shows a favorable prognosis, whereas a positive nodal status has an ominous prognostic value. Third, the present study evaluated the relationship between HER2 overexpression and the prognosis in a cohort of operable cases. Fourth, analysis of HER2 overexpression using TMA still remains to be validated, as already stated.

Interestingly, in the analysis of the pattern of recurrence, differences in the sites of recurrence were observed between the two groups with and without tumor HER2 overexpression (Table 4). Solitary metastases in the lymph node or liver were more common in patients with tumor HER2 overexpression, whereas peritoneal seeding and

multiple metastases were more frequent in patients without tumor HER2 overexpression. In a previous analysis of 643 patients administered systemic chemotherapy alone, the 5-year survival rates in the patients bearing metastasis confined to the abdominal lymph nodes and liver were 10.4 and 1.7 %, respectively, which was better than that of 0 % in the patients showing peritoneal dissemination [26]. The slight trend of favorable DSS in patients with HER2 overexpression despite these same patients showing a trend toward poor RFS might be attributable to the higher incidence of recurrence in lymph nodes in these patients. This finding implied that patients with tumor HER2 overexpression possibly included quite a considerable subpopulation responding to chemotherapy and exhibiting a favorable tumor biology. In a subclass analysis of the ToGA trial, lower gain of survival time was observed in an intent-to-treat analysis in the Asian population [10]. Although differences in environmental factors and biology of gastric cancer between Asian and Western populations [27] might be a possible explanation, the large number of patients who received subsequent chemotherapy following failure of first-line chemotherapy could be a likely reason [28]. Survival of patients showing tumor HER2 overexpression in the ToGA trial could have been influenced by the second-line chemotherapies.

Although further elucidation of the HER2-based biology of gastric cancer, which seems to be different from that of breast cancer, is required, multimodality therapy with trastuzumab for operable cases could be limited to a small subgroup. Patients with extensive spread to the lymph nodes may be good candidates. Furthermore, establishment of other treatment targets in patients without tumor HER2 overexpression may be more important to improve overall survival in gastric cancer patients.

In summary, we evaluated the relationship between tumor HER2 overexpression as assessed using a validated immunohistochemistry scoring system and clinical course of the patients in 1,006 cases of gastric cancer. Tumor HER2 overexpression was correlated with a well-differentiated histology and presence of lymph node metastasis. Patients with tumor HER2 overexpression are considered to constitute a subgroup that may be expected to show favorable response to HER2-targeted therapy.

Acknowledgments This work was partially supported by grants from the Third-Term Comprehensive 10-year Strategy for Cancer Control from the Ministry of Health, Labour and Welfare.

References

1. Moasser MM. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. Oncogene. 2007;26(45):6469–87.
2. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, Fleming T, Eiermann W, Wolter J, Pegram M, Baselga J, Norton L. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. N Engl J Med. 2001;344(11):783–92.
3. Barros-Silva JD, Leitao D, Afonso L, Vieira J, Dinis-Ribeiro M, Fragoso M, Bento MJ, Santos L, Ferreira P, Rego S, Brandao C, Carneiro F, Lopes C, Schmitt F, Teixeira MR. Association of ERBB2 gene status with histopathological parameters and disease-specific survival in gastric carcinoma patients. Br J Cancer. 2009;100(3):487–93.
4. Jorgensen JT. Targeted HER2 treatment in advanced gastric cancer. Oncology. 2010;78(1):26–33.
5. Nakajima M, Sawada H, Yamada Y, Watanabe A, Tatsumi M, Yamashita J, Matsuda M, Sakaguchi T, Hirao T, Nakano H. The prognostic significance of amplification and overexpression of c-met and c-erb B-2 in human gastric carcinomas. Cancer (Phila). 1999;85(9):1894–902.
6. Park DI, Yun JW, Park JH, Oh SJ, Kim HJ, Cho YK, Sohn CI, Jeon WK, Kim BI, Yoo CH, Son BH, Cho EY, Chae SW, Kim EJ, Sohn JH, Ryu SH, Sepulveda AR. HER-2/neu amplification is an independent prognostic factor in gastric cancer. Dig Dis Sci. 2006;51(8):1371–9.
7. Tanner M, Hollmen M, Junttila TT, Kapanen AI, Tommola S, Soini Y, Helin H, Salo J, Joensuu H, Sihvo E, Elenius K, Isola J. Amplification of HER-2 in gastric carcinoma: association with topoisomerase II alpha gene amplification, intestinal type, poor prognosis and sensitivity to trastuzumab. Ann Oncol. 2005;16(2): 273–8.
8. Yu GZ, Chen Y, Wang JJ. Overexpression of Grb2/HER2 signaling in Chinese gastric cancer: their relationship with clinicopathological parameters and prognostic significance. J Cancer Res Clin Oncol. 2009;135(10):1331–9.
9. Hofmann M, Stoss O, Shi D, Buttner R, van de Vijver M, Kim W, Ochiai A, Ruschoff J, Henkel T. Assessment of a HER2 scoring system for gastric cancer: results from a validation study. Histopathology. 2008;52(7):797–805.
10. Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, Lordick F, Ohtsu A, Omuro Y, Satoh T, Aprile G, Kulikov E, Hill J, Lehle M, Ruschoff J, Kang YK. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. Lancet. 2010;376(9742):687–97.
11. Sobin LH, Gospodarowicz MK, Wittekind C, editors. TMN classification of malignant tumours. Chichester: Wiley; 2010.
12. Graham AD, Faratian D, Rae F, Thomas JS. Tissue microarray technology in the routine assessment of HER-2 status in invasive breast cancer: a prospective study of the use of immunohistochemistry and fluorescence in situ hybridization. Histopathology. 2008;52(7):847–55.
13. Maruyama K, Okabayashi K, Kinoshita T. Progress in gastric cancer surgery in Japan and its limits of radicality. World J Surg. 1987;11(4):418–25.
14. Cunningham D, Allum WH, Stenning SP, Thompson JN, Van de Velde CJ, Nicolson M, Scarffe JH, Lofts FJ, Falk SJ, Iveson TJ, Smith DB, Langley RE, Verma M, Weeden S, Chua YJ, Participants MT. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. N Engl J Med. 2006;355(1):11–20.
15. Macdonald JS, Smalley SR, Benedetti J, Hundahl SA, Estes NC, Stemmermann GN, Haller DG, Ajani JA, Gunderson LL, Jessup JM, Martenson JA. Chemoradiotherapy after surgery compared with surgery alone for adenocarcinoma of the stomach or gastroesophageal junction. N Engl J Med. 2001;345(10):725–30.
16. Sakuramoto S, Sasako M, Yamaguchi T, Kinoshita T, Fujii M, Nashimoto A, Furukawa H, Nakajima T, Ohashi Y, Imamura H,

Higashino M, Yamamura Y, Kurita A, Arai K. Adjuvant chemotherapy for gastric cancer with S-1, an oral fluoropyrimidine. N Engl J Med. 2007;357(18):1810–20.
17. Ruschoff J, Nagelmeier I, Baretton G, Dietel M, Hofler H, Schildhaus HU, Buttner R, Schlake W, Stoss O, Kreipe HH. Her2 testing in gastric cancer. What is different in comparison to breast cancer? Pathologe. 2010;31(3):208–17.
18. Im SA, Kim JW, Kim JS, Kim MA, Jordan B, Pickl M, Han SW, Oh DY, Lee HJ, Kim TY, Kim WH, Yang HK, Bang YJ. Clinicopathologic characteristics of patients with stage III/IV (M(0)) advanced gastric cancer, according to HER2 status assessed by immunohistochemistry and fluorescence in situ hybridization. Diagn Mol Pathol. 2011;20(2):94–100.
19. Kim KC, Koh YW, Chang HM, Kim TH, Yook JH, Kim BS, Jang SJ, Park YS. Evaluation of HER2 protein expression in gastric carcinomas: comparative analysis of 1,414 cases of whole-tissue sections and 595 cases of tissue microarrays. Ann Surg Oncol. 2011;18(10):2833–40.
20. Thompson SK, Sullivan TR, Davies R, Ruszkiewicz AR. HER-2/neu gene amplification in esophageal adenocarcinoma and its influence on survival. Ann Surg Oncol. 2011;18:2010–7.
21. Wu HM, Liu YH, Lin F, Xu FP, Luo DL, Zhang F, Zhuang HG, Luo XL, Wu WL, Lin XT. Association of HER2 protein expression with clinicopathologic features and prognosis in Chinese patients with gastric carcinoma. Zhonghua Bing Li Xue Za Zhi. 2011;40(5):296–9.
22. Kunz PL, Mojtahed A, Fisher GA, Ford JM, Chang DT, Balise RR, Bangs CD, Cherry AM, Pai RK. HER2 expression in gastric and gastroesophageal junction adenocarcinoma in a US population: clinicopathologic analysis with proposed approach to HER2 assessment. Appl Immunohistochem Mol Morphol. 2012;20(1): 13–24.
23. Okines AFC, Cunningham D. Trastuzumab in gastric cancer. Eur J Cancer. 2010;46(11):1949–59.
24. Grabsch H, Sivakumar S, Gray S, Gabbert HE, Muller W. HER2 expression in gastric cancer: rare, heterogeneous and of no prognostic value. Conclusions from 924 cases of two independent series. Cell Oncol. 2010;32(1–2):57–65.
25. Yano T, Doi T, Ohtsu A, Boku N, Hashizume K, Nakanishi M, Ochiai A. Comparison of HER2 gene amplification assessed by fluorescence in situ hybridization and HER2 protein expression assessed by immunohistochemistry in gastric cancer. Oncol Rep. 2006;15(1):65–71.
26. Yoshida M, Ohtsu A, Boku N, Miyata Y, Shirao K, Shimada Y, Hyodo I, Koizumi W, Kurihara M, Yoshida S, Yamamoto S. Long-term survival and prognostic factors in patients with metastatic gastric cancers treated with chemotherapy in the Japan Clinical Oncology Group (JCOG) study. Jpn J Clin Oncol. 2004;34(11):654–9.
27. Ma BB, Hui EP, Mok TS. Population-based differences in treatment outcome following anticancer drug therapies. Lancet Oncol. 2010;11(1):75–84.
28. Sawaki A, Ohashi Y, Omuro Y, Satoh T, Hamamoto Y, Boku N, Miyata Y, Takiuchi H, Yamaguchi K, Sasaki Y, Nishina T, Satoh A, Baba E, Tamura T, Abe T, Hatake K, Ohtsu A. Efficacy of trastuzumab in Japanese patients with HER2-positive advanced gastric or gastroesophageal junction cancer: a subgroup analysis of the Trastuzumab for Gastric Cancer (ToGA) study. Gastric Cancer. 2012;15:313–22.

J Cancer Res Clin Oncol (2015) 141:453–463
DOI 10.1007/s00432-014-1841-x

ORIGINAL ARTICLE – CANCER RESEARCH

ZNF695 methylation predicts a response of esophageal squamous cell carcinoma to definitive chemoradiotherapy

Takamasa Takahashi · Satoshi Yamahsita · Yasunori Matsuda · Takayoshi Kishino · Takeshi Nakajima · Ryoji Kushima · Ken Kato · Hiroyasu Igaki · Yuji Tachimori · Harushi Osugi · Masato Nagino · Toshikazu Ushijima

Received: 24 June 2014 / Accepted: 20 September 2014 / Published online: 2 October 2014
© Springer-Verlag Berlin Heidelberg 2014

Abstract
Purpose Definitive chemoradiotherapy (dCRT) is one of the standard treatments for esophageal squamous cell carcinoma. Patients with a response to dCRT have a better prognosis than those resistant to dCRT while survival benefits for patients with residual tumors are limited. Nevertheless, few molecular markers to predict the response to dCRT are currently available. Here, we aimed to establish a DNA methylation marker to predict the response to dCRT.
Methods A total of 104 patients were divided into screening ($n = 43$) and validation ($n = 61$) sets. A genome-wide DNA methylation analysis was performed using an Infinium HumanMethylation450 BeadChip array. Methylation levels were measured by quantitative methylation-specific PCR and normalized by the fraction of cancer cells in a sample.
Results The genome-wide methylation analysis of seven responders and eight non-responders identified 18 genomic regions specifically (un)methylated in the responders. Among these, methylation of the promoter CpG island of *ZNF695* was significantly associated with the response to dCRT in the screening set ($P = 0.004$), and a cutoff value was determined. In the validation set, the association was successfully validated ($P = 0.021$), and a high specificity (90 %) for the prediction of responders was obtained using the prefixed cutoff value. In addition, a multivariate analysis showed that *ZNF695* methylation was an independent predictive factor for the response to dCRT (OR 7.55, 95 % CI 2.12–26.9, $P = 0.002$).
Conclusion *ZNF695* methylation was significantly associated with the response to dCRT and is a promising predictive marker for the response to dCRT.

Keywords Epigenetics · DNA methylation · Biomarker · Esophageal squamous cell carcinoma · Chemoradiotherapy

Electronic supplementary material The online version of this article (doi:10.1007/s00432-014-1841-x) contains supplementary material, which is available to authorized users.

T. Takahashi · S. Yamahsita · Y. Matsuda · T. Kishino · T. Ushijima (✉)
Division of Epigenomics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
e-mail: tushijim@ncc.go.jp

T. Takahashi · T. Kishino · H. Igaki · Y. Tachimori
Esophageal Surgery Division, National Cancer Center Hospital, Tokyo, Japan

T. Takahashi · M. Nagino
Division of Surgical Oncology, Department of Surgery, Nagoya University Graduate School of Medicine, Nagoya, Japan

Y. Matsuda · H. Osugi
Department of Gastroenterological Surgery, Graduate School of Medicine, Osaka City University, Osaka, Japan

T. Nakajima
Endoscopy Division, National Cancer Center Hospital, Tokyo, Japan

R. Kushima
Pathology and Clinical Laboratory Division, National Cancer Center Hospital, Tokyo, Japan

K. Kato
Gastrointestinal Medical Oncology Division, National Cancer Center Hospital, Tokyo, Japan

Introduction

Esophageal squamous cell carcinoma (ESCC) is a predominant histological type of esophageal cancer worldwide and one of the most serious malignant cancers due to its rapid development and fatal prognosis (Chung et al. 2010; Pennathur et al. 2013). The prognosis of patients with advanced ESCC is still unsatisfactory although aggressive treatment strategies, such as surgery with preoperative or postoperative chemotherapy or preoperative chemoradiotherapy (CRT), and definitive CRT (dCRT) (Allum et al. 2009; Ando et al. 2003, 2012; Blum et al. 2013; Conroy et al. 2014; Kato et al. 2011, 2013; Kelsen et al. 2007; Kleinberg and Forastiere 2007; Tepper et al. 2008; van Hagen et al. 2012), have been implemented. Among these aggressive treatments, dCRT has been an important therapeutic strategy for advanced ESCC (Conroy et al. 2014; Kato et al. 2011, 2013; Kleinberg and Forastiere 2007; van Hagen et al. 2012). By dCRT, a complete response of primary tumor is seen in 58.0–70.6 % of patients (Kato et al. 2011, 2013; Tahara et al. 2005), and such patients not only have a good prognosis but also can preserve their esophagus and keep a better quality of life than patients who undergo esophagectomy. In contrast, survival benefits of dCRT for patients who have residual tumors are limited even after adequate salvage treatments (Conroy et al. 2014; Kato et al. 2011, 2013). Although preoperative chemo- or chemoradiotherapy followed by surgery may be powerful, operative mortality and morbidity are not ignored. If we can predict the response of ESCC to dCRT before starting a treatment, we can stratify patients who will benefit by it for personalized treatments.

To establish a biomarker to predict the response of ESCC to dCRT, a lot of effort has been made in the field of molecular markers, and markers using specific RNA and protein expression have been reported (Akutsu et al. 2011; Gao et al. 2013; Makuuchi et al. 2013; Okamoto et al. 2013; Zhang et al. 2013). Although these markers were associated with the response to dCRT, their utility has not been confirmed by independent studies yet. Also, their accuracy even in the screening set was not satisfactory for clinical use. Therefore, further searching for a predictive marker for dCRT is required, and validating the utility of the marker is essential for clinical application.

As a molecular marker, DNA methylation has several advantages over RNA and proteins (Goel 2010; Issa 2012; Laird 2003). First, DNA methylation status is stable, even if a cell is placed in different environments, and consistent results can be obtained even under different conditions of sample collection where gene expression profiles are affected. Secondly, DNA methylation can be analyzed using DNA, which can remain relatively intact in samples with RNA and protein degradation. Thirdly, DNA methylation has only two statuses, methylated or unmethylated, and its profile cannot be affected by a small amount of contaminating cells, different from gene expression. Also, DNA methylation status has been reported to correlate with clinicopathological features of many types of cancers (Brait et al. 2008; Kim et al. 2013; Sato et al. 2002; Yang and Park 2012). Taking these advantages, DNA methylation markers for the prediction of a response to a cancer therapy have been identified in other types of cancers (Amatu et al. 2013; Giovannetti et al. 2012; Hegi et al. 2005; Mikeska et al. 2012; Park et al. 2009; Toyota et al. 2009). However, in ESCC, few studies have demonstrated an association between DNA methylation status and response to dCRT (Brabender et al. 2009).

In the present study, we aimed to identify genomic regions whose methylation statuses are associated with the response of ESCC to dCRT by a genome-wide methylation analysis and to validate the isolated candidate predictive markers for the response to dCRT.

Methods

Samples and patient profiles

A total of 187 ESCC samples were collected from ESCC patients from January 2006 to April 2013 at the National Cancer Center Hospital and the Osaka City University Hospital. This study was approved by the Institutional Review Board of the National Cancer Center, Tokyo, Japan (Reference No. 2010-094), and the Osaka City University, Osaka, Japan (Reference No. 1500). Written informed consents were obtained from all the individuals. They had histologically confirmed ESCC and were at cStage II–IV according to the 6th edition of the TNM classification (Sobin 2002). Patients with clinical T4 and with distant organ metastasis were excluded.

A group of 128 patients underwent chemotherapy comprised of two courses of infusion of 5-fluorouracil (5-FU) (700 mg/m^2/day) on days 1–4 and 29–32 and 2-h infusion of cisplatin (CDDP) (70 mg/m^2) on days 1 and 29. A total of 60 Gy was concurrently administered in 30 fractions. A 1-week break was provided after 30 Gy irradiation, and radiotherapy was resumed on day 29 with the second chemotherapy course. The other group of 59 patients underwent chemotherapy comprised of two courses of infusion of 5-FU (1,000 mg/m^2/day) on days 1–4 and 29–32 and a 2-h infusion of CDDP (75 mg/m^2) on day 1 and 29. A total of 50.4 Gy was concurrently administered in 28 fractions. Break period of radiation was not planned. An additional two cycles of chemotherapy were conducted after completion of the radiotherapy until the tumor disappeared or progressed.

The response to dCRT was determined based on the findings of endoscopic evaluation of the primary tumor after each course of chemotherapy using the modified criteria of the 10th edition of the Japanese Society for Esophageal Diseases (Kuwano et al. 2008). A responder was defined as a patient with disappearance of the primary tumor without the presence of irregular erosion, ulceration, visible elevated lesions, acute esophagitis, or malignant cells in biopsy specimens. A non-responder was defined as a patient not diagnosed as a responder after finishing all courses of chemotherapy.

DNA extraction, estimation of the fraction of cancer cells in a DNA sample, and sample selection

The 187 ESCC samples consisted of 48 biopsy samples stored in RNAlater (Applied Biosystems, Foster City, CA, USA) at −80 °C (frozen samples), and 139 biopsy samples embedded in paraffin wax block after fixation with formalin (FFPE samples). Genomic DNA was extracted by using the phenol/chloroform method. The fraction of cancer cells in a DNA sample was assessed using a DNA methylation marker for the fraction of cancer cells (Takahashi et al. 2013).

Among the 48 frozen samples, five samples were excluded from further analysis because the fraction of cancer cells was <20 % (four samples) or the total number of DNA molecules assessed by quantitative methylation-specific PCR (qMSP) was <50 (1 sample). As a result, the 43 remaining frozen samples were used as the screening set (Table 1). Among the 139 FFPE samples, 78 samples were excluded from further analysis because the fraction of cancer cells was <20 % (two samples) or the total number of DNA molecules was <50 (76 samples). Consequently, the 61 remaining FFPE samples were used as the validation set (Table 1).

Genome-wide DNA methylation analysis

A genome-wide screening of differentially methylated CpG sites was conducted using an Infinium HumanMethylation450 BeadChip array, which covered 482,421 CpG sites (Illumina, San Diego, CA, USA) (Shigematsu et al. 2012). We excluded 11,551 CpG sites on the sex chromosomes and used the remaining 470,870 CpG sites for the analysis. To adjust probe design biases, intra-array normalization was conducted using a peak-based correction method, Beta MIxture Quantile dilation (BMIQ) (Teschendorff et al. 2013). The methylation level of each CpG site was represented by a β value, which ranged from 0 (completely unmethylated) to 1 (completely methylated).

qMSP and normalization of the methylation levels

Sample DNA was treated with sodium bisulfite and purified (Kaneda et al. 2004). qMSP was conducted by real-time PCR using 25 ng bisulfite-modified DNA and specific primers (Supplementary Table 1), SYBR Green I (BioWhittaker Molecular Applications, Rockland, MD, USA), and an MyiQ real-time PCR detection system (Bio-Rad Laboratories, CA, USA). Standard DNA for a methylated sequence was prepared by purifying PCR products with primers for the methylated sequence and genomic DNA treated with *SssI* methylase (New England Biolabs, Beverly, MA, USA). That for an unmethylated sequence was prepared by purifying PCR products with primers for the unmethylated sequence and DNA amplified twice with a GenomiPhi HY DNA amplification kit (GE Healthcare Bio-Science, Buckinghamshire, England). The numbers of methylated and unmethylated molecules in a sample were calculated by comparing its amplification with that of methylated and unmethylated standard DNA, respectively, that contained known numbers of molecules (10–10^6 molecules).

A methylation level was calculated as the fraction of methylated molecules in the total number of DNA molecules (the number of methylated molecules + the number of unmethylated molecules) in a sample (Oka et al. 2009). The measured methylation level was normalized for each gene using the fraction of cancer cells in a sample [the normalized methylation level = 100 × (the measured methylation level (%))/(the fraction of cancer cells in the sample (%))].

Statistical analysis

Fisher's exact test was used to evaluate the significant difference in relative frequency between two independent groups. Differences in the normalized methylation levels between the responders and the non-responders were evaluated by Mann–Whitney U test. In the univariate analysis, the odds ratios (ORs) and 95 % confidence interval (95 % CI) were calculated. The confounders affecting the apparent likelihood of the response to dCRT were tested using multivariate logistic regression analysis. Overall survival was calculated from the initiation of the first course of treatment to the date of death or final date of survival confirmation. Survival curves were drawn by the Kaplan–Meier method and were compared by the log-rank test. All statistical analyses were conducted by PASW statistics version 18.0.0 (SPSS Japan Inc., Tokyo, Japan).

Results

Genome-wide screening to isolate genomic regions whose methylation statuses were associated with the response to dCRT in the screening set

To isolate genomic regions specifically methylated or unmethylated in responders, a genome-wide methylation analysis was performed using an Infinium HumanMethylation450 BeadChip array. For this analysis, we used (1) seven ESCC samples from responders with 40 % or more of cancer cells, (2) eight ESCC samples from non-responders with 40 % or more of cancer cells, (3) peripheral leukocytes of one healthy volunteer, (4) a pool of normal esophageal mucosae of four healthy volunteers, and (5) a pool of non-cancerous esophageal mucosae of eight ESCC patients. The ESCC samples were selected from 43 samples in the screening set (Table 1).

From 470,870 CpG sites on autosomes, we selected 165,081 CpG sites unmethylated (β value <0.2) in the peripheral leukocytes and the pool of normal esophageal mucosae Fig. 1. From the 165,081 CpG sites, 156,381 CpG sites unmethylated (β value <0.2) in the pool of non-cancerous mucosae were selected. From the 156,381 CpG sites, we isolated 543 CpG sites (1) specifically methylated [β value $\geq$ cancer cell fraction (%)/200] in the responders and (2) by which the response to dCRT was predicted in 11 or more of the 15 ESCCs. Finally, from the 543 CpG sites,

Table 1 Clinicopathological data and the fraction of cancer cells in the screening and validation sets

Feature	Genome-wide analysis set[a]			Screening set			Validation set		
	Responder	Non-responder	*P* value[b]	Responder	Non-responder	*P* value[b]	Responder	Non-responder	*P* value[b]
Number of patients	7	8		23	20		41	20	
Age (years)[c]	66.1 ± 4.4	66.8 ± 5.9	0.827	67.4 ± 6.3	65.1 ± 8.3	0.318	66.0 ± 8.1	60.5 ± 7.7	0.014
Gender									
Male	6	5	0.569	20	15	0.44	35	19	0.258
Female	1	3		3	5		6	1	
Location									
Cervical	1	0	0.709	2	0	0.116	7	3	0.247
Upper	1	2		1	5		3	4	
Middle	4	5		17	11		14	9	
Lower	1	1		3	4		17	4	
Radiation dose (Gy)									
60	5	5	1	11	8	0.76	33	12	0.083
50.4	2	3		12	12		8	8	
cT									
1	1	1	0.529	5	2	0.016	14	0	0.001
2	1	0		6	0		7	2	
3	5	7		12	18		20	18	
cN									
0	1	2	1	3	2	1	7	2	0.377
1	6	6		20	18		34	18	
cM									
0	5	5	1	18	10	0.064	33	13	0.158
1	2	3		5	10		8	7	
cStage									
2	2	3	0.765	11	4	0.09	23	2	0.003
3	3	2		7	6		10	11	
4	2	3		5	10		8	7	
Fraction of cancer cells (%)[c]	65.6 ± 16.1	63.5 ± 14.2	0.793	59.7 ± 19.3	57.9 ± 17.6	0.756	64.1 ± 17.1	60.3 ± 17.6	0.417

[a] This set was derived from samples in the screening set

[b] *P* values were calculated by the chi-square test

[c] Mean ± SD

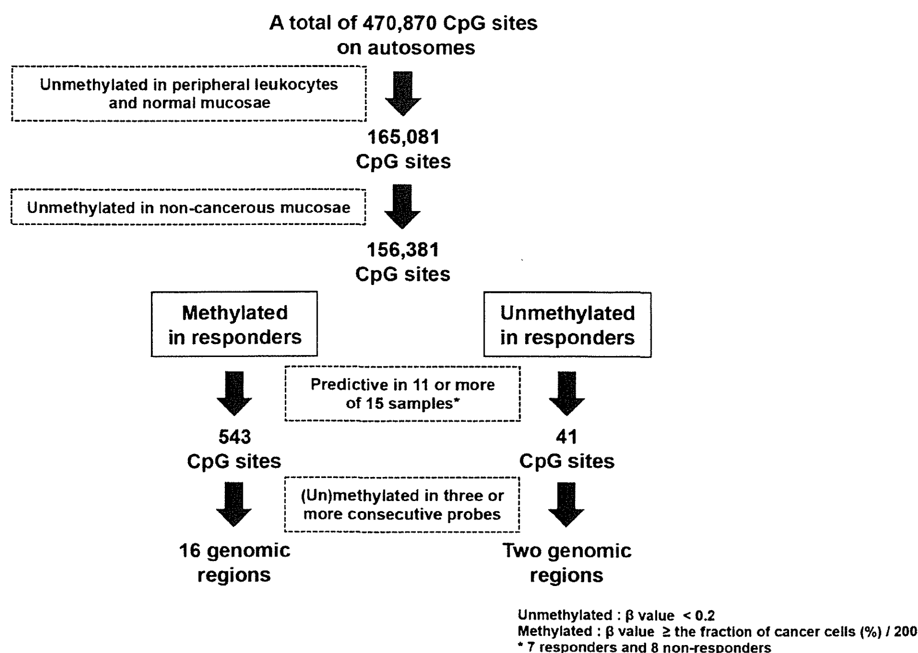

Fig. 1 Selection processes of genomic regions whose methylation statuses were associated with the response to dCRT. From the 482,421 CpG sites on the Infinium HumanMethylation450 BeadChip array, those on autosomes and unmethylated in non-cancerous mucosae were selected. As specifically methylated in responders, 543 CpG sites, resulting in 16 genomic regions, were isolated. As specifically unmethylated in responders, 41 CpG sites, resulting in two genomic regions, were isolated. An unmethylated CpG site was defined when its β value was <0.2. A methylated CpG site was defined when its β value was more than or equal to half of the fraction of cancer cells in a sample [β value $\geq$ the fraction of cancer cells in the sample (%)/200]

we isolated 16 genomic regions methylated or unmethylated in three or more consecutive probes in each sample. In the same way, we isolated two genomic regions specifically unmethylated (β value <0.2) in the responders (Table 2; a representative locus in Supplementary Fig. S1). Among the 18 genomic regions, we were able to design primers for qMSP in six regions.

Selection of genomic regions whose methylation statuses were associated with the response to dCRT in the screening set

To select genomic regions whose methylation statuses were associated with the response to dCRT, methylation levels of the six genomic regions were quantified in the screening set (Table 1) by qMSP. To exclude the effect of contamination of non-cancerous cells in a sample, a normalized methylation level was calculated using the fraction of cancer cells in a sample assessed using a DNA methylation marker (Takahashi et al. 2013) [the normalized methylation level = 100 × (the measured methylation level (%))/(the fraction of cancer cells in the sample (%))]. Among the six genomic regions, the normalized methylation levels of the promoter CpG island of *ZNF695* in the responders were significantly higher than those in the non-responders ($P = 0.004$; Fig. 2).

For *ZNF695*, a cutoff value of 8.0 was established, so that the Youden index (sensitivity + specificity − 1) would be maximized. Also, we implemented a criterion that a cutoff value should be larger than 2.0 because methylation levels in this range were experimentally reliable. Using this cutoff value, the incidence of *ZNF695* methylation was significantly higher in the responders than in the non-responders ($P = 0.004$).

Validation of the association between *ZNF695* methylation and the response to dCRT

To validate the association between *ZNF695* methylation and the response to dCRT, its methylation levels were quantified in an independent sample set (validation set, Table 1) and were normalized by the fraction of cancer cells in a sample. The normalized methylation level of *ZNF695* was significantly higher in the responders than in the non-responders ($P = 0.021$; Fig. 3). Using the cutoff value of 8.0, which had been prefixed in the screening set, the incidence of *ZNF695* methylation in the responders was significantly higher than that in the non-responders ($P = 0.036$). These results demonstrated that the association between *ZNF695* methylation and the response to dCRT was valid even in an independent sample set and that *ZNF695* methylation was a candidate predictive marker for the response to dCRT. The response to dCRT was detected with a sensitivity of 39 %, specificity of 90 %, and positive predictive value of 89 %.

Association between *ZNF695* methylation and clinicopathological characteristics

The predictive power of *ZNF695* methylation was compared with other clinical factors. Univariate analyses showed that age, clinical T stage, clinical M stage, and

Table 2 Genomic regions identified by genome-wide methylation analysis

No.	Gene symbol	Location		Position against a gene	Relation to a CpG island	Mean β value		Incidence of methylation		*P* value[a]	
		Chr	nt. number			Responder ($n = 7$)	Non-responder ($n = 8$)	Responder ($n = 7$)	Non-responder ($n = 8$)	Screening	Validation
1	*AGRN*	1	976168–976227	Body	Island	0.30	0.04	4	0	ND	ND
2	*ZNF695**	1	247171252–247171572	1stExon-TSS1500	Island-S_Shore	0.29	0.02	4	0	0.004	0.021
3	*HOXD8**	2	176994665–176995088	5′UTR	Island	0.33	0.10	4	1	0.108	ND
4	*TRIM71*	3	32859377–32859445	TSS200	Island	0.24	0.04	3	0	ND	ND
5	*PCDHA7**	5	140306181–140306213	Body	Island	0.32	0.11	4	1	0.289	ND
6	*KIAA1949*	6	30653659–30653736	Body	N_Shore	0.35	0.15	4	1	ND	ND
7	*LEMD2*	6	33739406–33739607	3′UTR	Island	0.31	0.08	4	0	ND	ND
8	*PCLO*	7	82792105–82792284	5′UTR-TSS200		0.24	0.06	3	0	ND	ND
9	*C12orf56**	12	64784252–64784626	1stExon-TSS1500	Island	0.40	0.16	6	2	0.053	ND
10	*MEIS2*	15	37390176–37390326	1stExon	Island	0.49	0.27	6	3	ND	ND
11	*CCDC64B*	16	3079708–3079953	Body	Island	0.28	0.06	3	0	ND	ND
12	*FAM38A*	16	88850218–88850534	Body	Island	0.36	0.10	4	0	ND	ND
13	*ZNF585B**	19	37701550–37701642	TSS200		0.25	0.01	3	0	0.357	ND
14	*ZNF155*	19	44488121–44488181	TSS200		0.27	0.03	3	0	ND	ND
15	*ZNF350*	19	52490223–52490339	TSS200-TSS1500		0.22	0.05	4	0	ND	ND
16	*GALR3**	22	38221009–38221187	Body	Island	0.37	0.11	4	1	0.154	ND
17	*CXCL5*	4	74864165–74864313	1stExon	Island	0.36	0.59	4	8	ND	ND
18	*PKD2L2*	5	137224904–137225191	TSS1500-1stExon	N_Shore-Island	0.07	0.29	1	4	ND	ND

Chr, chromosome number; nt. number, nucleotide number in the NCBI database (NCBI37/hg19); TSS200, within 200 bp upstream of the transcription start site (TSS); TSS1500, 200–1,500 bp upstream of a TSS; body, within exons or introns of a gene; 3′UTR, 3′ untranslated region of a gene, 5′UTR; 5′ untranslated region of a gene; Island, CpG island; S_Shore, within 2,000 bp from the end of a CpG island; N_Shore, within 2,000 bp of the start of a CpG island; and ND, not done

* Primers for qMSP were successfully designed

[a] Difference was evaluated by Mann–Whitney *U* test

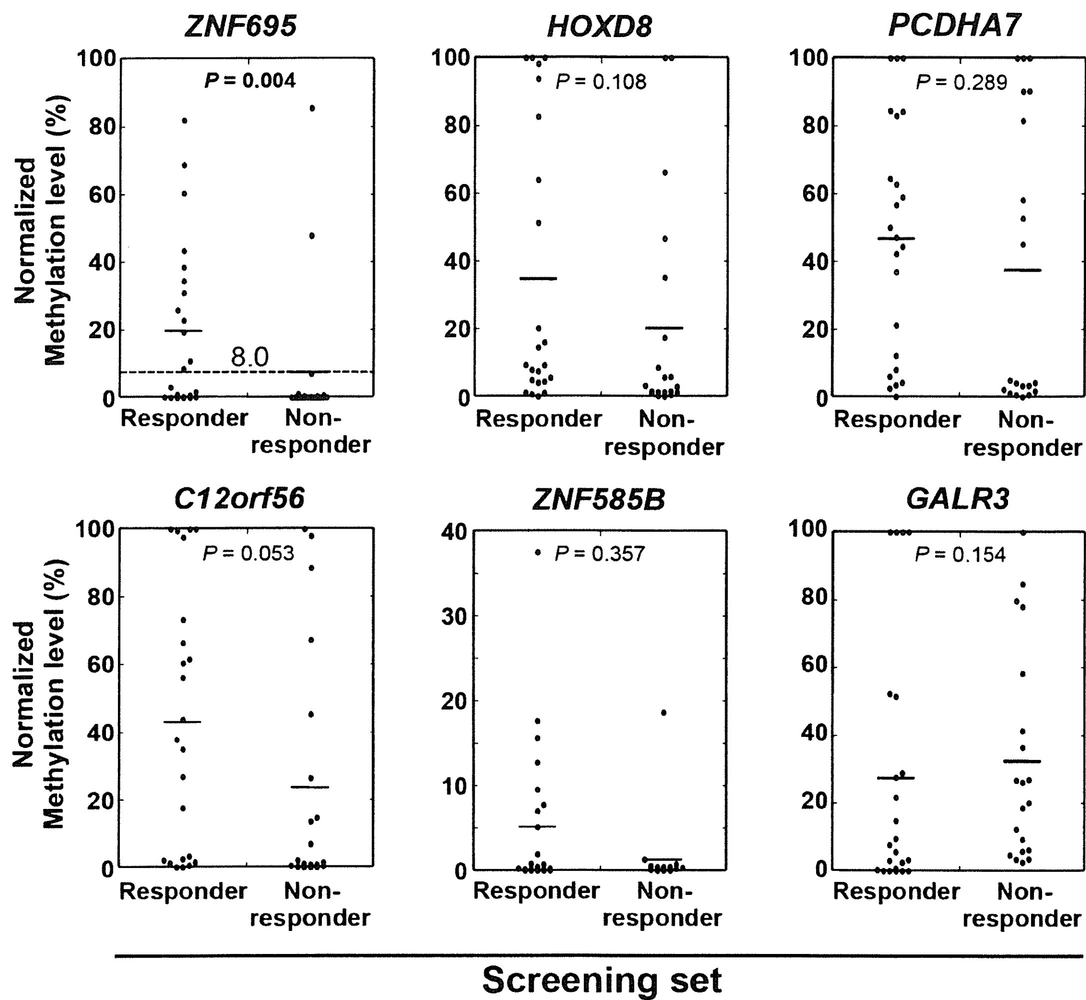

Fig. 2 Normalized methylation levels of the candidate genomic regions in the screening set. Methylation levels of six genomic regions were measured by qMSP in 23 responders and 20 non-responders in the screening set. A normalized methylation level was calculated using the fraction of cancer cells in a sample [the normalized methylation level = 100 × (the measured methylation level (%))/(the fraction of cancer cells in the sample (%))]. The *horizontal dotted line* shows a cutoff value of 8.0 adopted for *ZNF695*

ZNF695 methylation were significantly associated with the response to dCRT (Table 3). A multivariate logistic regression analysis using the age, sex, radiation dose, clinical T stage, clinical M stage, and *ZNF695* methylation showed that, in addition to the clinical T stage, *ZNF695* methylation was an independent predictive factor for the response to dCRT (OR 7.55, 95 % CI 2.12–26.9, $P = 0.002$; Table 4).

Finally, by Kaplan–Meier analysis, patients with the response to dCRT had significantly better overall survival than those without the response to dCRT ($P < 0.001$; Supplementary Fig. S2a). However, *ZNF695* methylation (methylated vs. unmethylated) and clinical T stage (cT1 or T2 vs. cT3) were not significantly associated with overall survival (methylated vs. unmethylated, $P = 0.469$; cT1 or T2 vs. cT3, $P = 0.244$; Supplementary Fig. S2b and S2c). Also, by Kaplan–Meier analysis stratified by clinical T stage, *ZNF695* methylation (methylated vs. unmethylated) was not significantly associated with overall survival (cT1 or T2, $P = 0.559$; cT3, $P = 0.593$; Supplementary Fig. S3a and S3b).

Discussion

In the present study, a candidate predictive marker for the response of ESCC to dCRT, *ZNF695* methylation, was successfully identified by a genome-wide methylation analysis. Because any genome-wide screening has a risk of obtaining "too good" results due to overfitting, a biomarker isolated by such a screening needs to be validated

using a different set of samples. Here, the predictive power of *ZNF695* methylation was confirmed in an independent set of samples ($P = 0.021$), and the concern of overfitting was resolved.

The specificity and sensitivity in the validation set were 90 and 39 %, respectively, using the cutoff value established in the screening set. The multivariate analysis showed that *ZNF695* methylation was an independent predictive factor. These results indicated that a patient with *ZNF695* methylation can achieve a complete response of a primary tumor with high probability. Since the risk of failing to respond to dCRT is the major reason why some clinicians are reluctant to recommend dCRT to patients, this characteristic of *ZNF695* methylation may become a great advantage in the selection of appropriate treatments. In contrast, even among the patients within a specific clinical tumor stage, *ZNF695* methylation was not significantly associated with the response to dCRT, due to the low sensitivity of *ZNF695* methylation. Therefore, if a patient does not have *ZNF695* methylation, we cannot conclude that the patient will not benefit by dCRT. To overcome the low sensitivity, a combination with other high specificity markers might be useful.

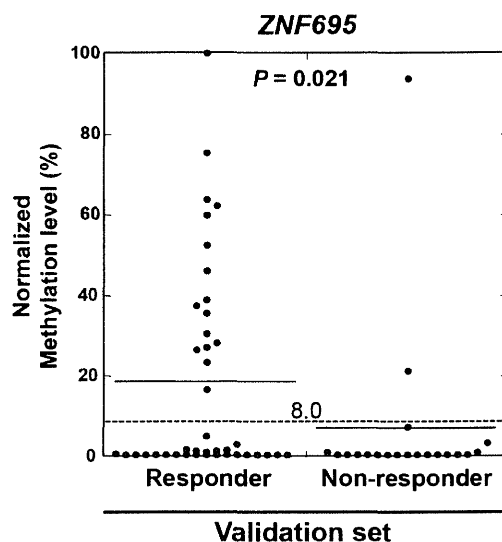

Fig. 3 Normalized methylation levels of *ZNF695* in the validation set. Methylation levels of *ZNF695* were measured by qMSP in 41 responders and 20 non-responders in the validation set. A normalized methylation level was calculated as described in the legend of Fig. 2. The *horizontal dotted line* shows the cutoff value of 8.0 prefixed in the screening set

In general, patients with a complete disappearance of a primary tumor have better prognosis that those with a residual primary tumor (Tahara et al. 2005). Also in this study, the responders showed significantly better survival than the non-responders (Supplementary Fig. S2a). It is therefore expected that a biomarker to predict the response to dCRT could also predict the prognosis of patients. However, *ZNF695* methylation was not associated with better overall survival. Among the patients with unmethylated *ZNF695*, more than half of the patients were responders due to the low sensitivity of this marker. This was considered as a reason why we were not able to observe a significant difference in overall survival between patients with methylated and unmethylated *ZNF695*.

Table 3 Univariate analysis of the response to dCRT

Features	Categories	No. of cases	Responder	Non-responder	OR	95 % CI	*P* value[a]
Age	≥60	83	57	26	4.84	1.58–12.2	0.005
	<60	21	7	14			
Gender	Female	15	9	6	1.08	0.35–3.30	1.000
	Male	89	55	34			
Location	Cervical, upper	25	13	12	1.68	0.68–4.18	0.346
	Middle, lower	79	51	28			
Clinical T stage	T1, T2	36	32	4	9.00	2.87–28.2	<0.001
	T3	68	32	36			
Clinical N stage	N0	14	10	4	1.67	0.49–5.74	0.558
	N1	90	54	36			
Clinical M stage	M0	64	51	23	2.90	1.21–6.95	0.025
	M1	40	13	17			
Radiation dose (Gy)	60	69	45	24	1.58	0.69–3.62	0.294
	50.4	35	19	16			
ZNF695 methylation	Methylated	32	28	4	7.00	2.23–22.0	<0.001
	Unmethylated	72	36	36			

[a] *P* values were calculated by Fisher's exact test

Table 4 Multivariate logistic regression analysis to identify independent factors

	No. of cases	OR	95 % CI	*P* value
Clinical T category (cT1 or cT2 vs. cT3)	36/68	7.58	2.07–27.8	0.002
ZNF695 methylation (methylated vs. unmethylated)	32/72	7.55	2.12–26.9	0.002
Age (≥60 vs. <60)	83/21	2.25	0.66–7.68	0.196
Clinical M category (cM0 vs. cM1)	64/40	2.18	0.75–6.32	0.153
Radiation dose (60 vs. 50.4)	69/35	1.96	0.68–5.61	0.213
Gender (male vs. female)	89/15	0.73	0.19–2.82	0.647

One advantage of DNA methylation marker is that DNA methylation can be assessed using DNA, which is chemically stable. In our study, we used samples fixed with formalin and embedded in paraffin wax. Even using the samples with which analysis of RNA and most proteins is difficult, we were able to assess a methylation level of *ZNF695* and to use the samples for the analysis. However, we had to exclude samples in which the total number of DNA molecules assessed by qMSP was <50. This resulted in the selection of 61 FFPE samples from the initial 139 FFPE samples. It is known that storage in formalin for a prolonged period results in the degradation of DNA (Bonin et al. 2010; Funabashi et al. 2012; Turashvili et al. 2012), and use of samples with shorter storage is considered desirable.

Contamination of non-cancerous cells in a tumor DNA sample can affect the results of any DNA methylation analysis (Loh et al. 2010). To overcome this issue, we excluded samples with low cancer cell content (<20 %) from the analysis and normalized the measured methylation level by the fraction of cancer cells in a sample assessed by a DNA methylation marker (Takahashi et al. 2013). The normalized methylation level was considered to well reflect the methylation status of cancer cells in a sample, compared with the actual methylation level measured. By comparison of the normalized methylation levels between the responders and non-responders, we were able to analyze subtle difference in methylation status between the two groups.

The biological basis of why *ZNF695* methylation was associated with the response to dCRT needs to be established. The methylated genomic region was located in a CpG island in the promoter region, and its methylation is considered to be critical for the regulation of gene expression. Because *ZNF695* is a zinc finger protein and likely to be a transcription factor, one possible mechanism is that *ZNF695* protein regulates the expression of genes involved in DNA repair introduced by therapy, including radiation, cisplatin, or 5-FU. Another possible mechanism is that a mechanism that alters methylation of multiple genes, such as the CpG island methylator phenotype, causes *ZNF695* methylation, and the mechanism is critical for cancer cells surviving dCRT. Recently, splice variants of *ZNF695* have been reported to be highly expressed in ovarian cancer (Juarez-Mendez et al. 2013). Although it is still unknown whether such splice variants are associated with aggressive behavior of ovarian cancer, there is a possibility that by loss of *ZNF695* function, cancer cells acquire a high proliferation ability, resulting in resistance to therapy.

In conclusion, we identified that methylation of the promoter CpG island of *ZNF695* was associated with the response of ESCC to dCRT with a high specificity of 90 %. *ZNF695* methylation is a promising predictive marker for dCRT, and a large-scale analysis is warranted.

Acknowledgments We thank the National Cancer Center Research Core Facility for preparation of sections from paraffin-embedded biopsy specimens in this study. This study was supported by Project for Development of Innovative Research on Cancer Therapeutics (P-DIRECT) from the Ministry of Education, Culture Sports, Science, and Technology, Japan, and by the Third-term Comprehensive Cancer Control Strategy from the Ministry of Health, Labour, and Welfare, Japan (H22-general-002). T.T. and Y.M. are recipients of a Research Resident Fellowship from the Foundation for Promotion of Cancer Research. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Conflict of interest None.

References

Akutsu Y et al (2011) COX2 expression predicts resistance to chemoradiotherapy in esophageal squamous cell carcinoma. Ann Surg Oncol 18:2946–2951. doi:10.1245/s10434-011-1645-z

Allum WH, Stenning SP, Bancewicz J, Clark PI, Langley RE (2009) Long-term results of a randomized trial of surgery with or without preoperative chemotherapy in esophageal cancer. J Clin Oncol 27:5062–5067. doi:10.1200/JCO.2009.22.2083

Amatu A et al (2013) Promoter CpG island hypermethylation of the DNA repair enzyme MGMT predicts clinical response to dacarbazine in a phase II study for metastatic colorectal cancer. Clin Cancer Res 19:2265–2272. doi:10.1158/1078-0432.CCR-12-3518

Ando N et al (2003) Surgery plus chemotherapy compared with surgery alone for localized squamous cell carcinoma of the thoracic esophagus: a Japan Clinical Oncology Group Study—JCOG9204. J Clin Oncol 21:4592–4596. doi:10.1200/JCO.2003.12.095

Ando N et al (2012) A randomized trial comparing postoperative adjuvant chemotherapy with cisplatin and 5-fluorouracil versus preoperative chemotherapy for localized advanced squamous cell carcinoma of the thoracic esophagus (JCOG9907). Ann Surg Oncol 19:68–74. doi:10.1245/s10434-011-2049-9

Blum MA, Taketa T, Sudo K, Wadhwa R, Skinner HD, Ajani JA (2013) Chemoradiation for esophageal cancer. Thorac Surg Clin 23:551–558. doi:10.1016/j.thorsurg.2013.07.006

Bonin S et al (2010) Multicentre validation study of nucleic acids extraction from FFPE tissues. Virchows Arch 457:309–317. doi:10.1007/s00428-010-0917-5

Brabender J et al (2009) Death-associated protein kinase (DAPK) promoter methylation and response to neoadjuvant radiochemotherapy in esophageal cancer. Ann Surg Oncol 16:1378–1383. doi:10.1245/s10434-009-0356-1

Brait M et al (2008) Aberrant promoter methylation of multiple genes during pathogenesis of bladder cancer. Cancer Epidemiol Biomarkers Prev 17:2786–2794. doi:10.1158/1055-9965.EPI-08-0192

Chung CS, Lee YC, Wang CP, Ko JY, Wang WL, Wu MS, Wang HP (2010) Secondary prevention of esophageal squamous cell carcinoma in areas where smoking, alcohol, and betel quid chewing are prevalent. J Formos Med Assoc 109:408–421. doi:10.1016/s0929-6646(10)60072-1

Conroy T et al (2014) Definitive chemoradiotherapy with FOLFOX versus fluorouracil and cisplatin in patients with oesophageal cancer (PRODIGE5/ACCORD17): final results of a randomised, phase 2/3 trial. Lancet Oncol. doi:10.1016/s1470-2045(14)70028-2

Funabashi KS, Barcelos D, Visona I, e Silva MS, e Sousa ML, de Franco MF, Iwamura ES (2012) DNA extraction and molecular analysis of non-tumoral liver, spleen and brain from autopsy samples: the effect of formalin fixation and paraffin embedding. Pathol Res Pract 208:584–591. doi:10.1016/j.prp.2012.07.001

Gao Y et al (2013) BRCA1 mRNA expression as a predictive and prognostic marker in advanced esophageal squamous cell carcinoma treated with cisplatin- or docetaxel-based chemotherapy/chemoradiotherapy. PLoS One 8:e52589. doi:10.1371/journal.pone.0052589

Giovannetti E, Codacci-Pisanelli G, Peters GJ (2012) TFAP2E-DKK4 and chemoresistance in colorectal cancer. N Engl J Med 366:966. doi:10.1056/NEJMc1201170#SA1

Goel A (2010) DNA methylation-based fecal biomarkers for the noninvasive screening of GI cancers. Future Oncol 6:333–336. doi:10.2217/fon.10.9

Hegi ME et al (2005) MGMT gene silencing and benefit from temozolomide in glioblastoma. N Engl J Med 352:997–1003. doi:10.1056/NEJMoa043331

Issa JP (2012) DNA methylation as a clinical marker in oncology. J Clin Oncol 30:2566–2568. doi:10.1200/jco.2012.42.1016

Juarez-Mendez S et al (2013) Splice variants of zinc finger protein 695 mRNA associated to ovarian cancer. J Ovarian Res 6:61. doi:10.1186/1757-2215-6-61

Kaneda A, Kaminishi M, Sugimura T, Ushijima T (2004) Decreased expression of the seven ARP2/3 complex genes in human gastric cancers. Cancer Lett 212:203–210. doi:10.1016/j.canlet.2004.03.020

Kato K et al (2011) Phase II study of chemoradiotherapy with 5-fluorouracil and cisplatin for stage II–III esophageal squamous cell carcinoma: JCOG trial (JCOG 9906). Int J Radiat Oncol Biol Phys 81:684–690. doi:10.1016/j.ijrobp.2010.06.033

Kato K et al (2013) Phase II study of concurrent chemoradiotherapy at the dose of 50.4 Gy with elective nodal irradiation for stage II–III esophageal carcinoma. Jpn J Clin Oncol 43:608–615. doi:10.1093/jjco/hyt048

Kelsen DP et al (2007) Long-term results of RTOG trial 8911 (USA Intergroup 113): a random assignment trial comparison of chemotherapy followed by surgery compared with surgery alone for esophageal cancer. J Clin Oncol 25:3719–3725. doi:10.1200/JCO.2006.10.4760

Kim KJ, Lee TH, Cho NY, Yang HK, Kim WH, Kang GH (2013) Differential clinicopathologic features in microsatellite-unstable gastric cancers with and without MLH1 methylation. Hum Pathol 44:1055–1064. doi:10.1016/j.humpath.2012.09.009

Kleinberg L, Forastiere AA (2007) Chemoradiation in the management of esophageal cancer. J Clin Oncol 25:4110–4117. doi:10.1200/JCO.2007.12.0881

Kuwano H et al (2008) Guidelines for diagnosis and treatment of carcinoma of the esophagus April 2007 edition: part I edited by the Japan Esophageal Society. Esophagus 5:61–73. doi:10.1007/s10388-008-0151-2

Laird PW (2003) The power and the promise of DNA methylation markers. Nat Rev Cancer 3:253–266. doi:10.1038/nrc1045

Loh M et al (2010) Impact of sample heterogeneity on methylation analysis. Diagn Mol Pathol 19:243–247. doi:10.1097/PDM.0b013e3181de4396

Makuuchi Y et al (2013) Soluble interleukin-6 receptor is a serum biomarker for the response of esophageal carcinoma to neoadjuvant chemoradiotherapy. Cancer Sci 104:1045–1051. doi:10.1111/cas.12187

Mikeska T, Bock C, Do H, Dobrovic A (2012) DNA methylation biomarkers in cancer: progress towards clinical implementation. Expert Rev Mol Diagn 12:473–487. doi:10.1586/erm.12.45

Oka D, Yamashita S, Tomioka T, Nakanishi Y, Kato H, Kaminishi M, Ushijima T (2009) The presence of aberrant DNA methylation in noncancerous esophageal mucosae in association with smoking history: a target for risk diagnosis and prevention of esophageal cancers. Cancer 115:3412–3426. doi:10.1002/cncr.24394

Okamoto H et al (2013) Murine double minute 2 and its association with chemoradioresistance of esophageal squamous cell carcinoma. Anticancer Res 33:1463–1471

Park CK et al (2009) Methylation status of the MGMT gene promoter fails to predict the clinical outcome of glioblastoma patients treated with ACNU plus cisplatin. Neuropathology 29:443–449. doi:10.1111/j.1440-1789.2008.00998.x

Pennathur A, Gibson MK, Jobe BA, Luketich JD (2013) Oesophageal carcinoma. Lancet 381:400–412. doi:10.1016/s0140-6736(12)60643-6

Sato F et al (2002) Aberrant methylation of the HPP1 gene in ulcerative colitis-associated colorectal carcinoma. Cancer Res 62:6820–6822

Shigematsu Y et al (2012) Identification of a DNA methylation marker that detects the presence of lymph node metastases of gastric cancers. Oncol Lett 4:268–274. doi:10.3892/ol.2012.708

Sobin CW LH (2002) TNM classification of malignant tumours, 6th edn. Wiley, New York

Tahara M et al (2005) Clinical impact of criteria for complete response (CR) of primary site to treatment of esophageal cancer. Jpn J Clin Oncol 35:316–323. doi:10.1093/jjco/hyi095

Takahashi T et al (2013) Estimation of the fraction of cancer cells in a tumor DNA sample using DNA methylation. PLoS One 8:e82302. doi:10.1371/journal.pone.0082302

Tepper J et al (2008) Phase III trial of trimodality therapy with cisplatin, fluorouracil, radiotherapy, and surgery compared with surgery alone for esophageal cancer: CALGB 9781. J Clin Oncol 26:1086–1092. doi:10.1200/JCO.2007.12.9593

Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S (2013) A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data Bioinformatics 29:189-196 doi:10.1093/bioinformatics/bts680

Toyota M, Suzuki H, Yamashita T, Hirata K, Imai K, Tokino T, Shinomura Y (2009) Cancer epigenomics: implications of DNA methylation in personalized cancer therapy. Cancer Sci 100:787–791. doi:10.1111/j.1349-7006.2009.01095.x

Turashvili G et al (2012) Nucleic acid quantity and quality from paraffin blocks: defining optimal fixation, processing and DNA/RNA extraction techniques. Exp Mol Pathol 92:33–43. doi:10.1016/j.yexmp.2011.09.013

van Hagen P.et al (2012) Preoperative chemoradiotherapy for esophageal or junctional cancer. N Engl J Med 366:2074–2084. doi:10.1056/NEJMoa1112088

Yang M, Park JY (2012) DNA methylation in promoter region as biomarkers in prostate cancer. Methods Mol Biol 863:67–109. doi:10.1007/978-1-61779-612-8_5

Zhang JX et al (2013) PITX2: a promising predictive biomarker of patients' prognosis and chemoradioresistance in esophageal squamous cell carcinoma. Int J Cancer 132:2567–2577. doi:10.1002/ijc.27930

Gastric Cancer
DOI 10.1007/s10120-015-0475-2

ORIGINAL ARTICLE

Establishment of a DNA methylation marker to evaluate cancer cell fraction in gastric cancer

Liang Zong · Naoko Hattori · Yukie Yoda · Satoshi Yamashita · Hideyuki Takeshima · Takamasa Takahashi · Masahiro Maeda · Hitoshi Katai · Sohachi Nanjo · Takayuki Ando · Yasuyuki Seto · Toshikazu Ushijima

Received: 1 December 2014 / Accepted: 1 February 2015
© The International Gastric Cancer Association and The Japanese Gastric Cancer Association 2015

Abstract

Background Tumor samples are unavoidably contaminated with coexisting normal cells. Here, we aimed to establish a DNA methylation marker to estimate the fraction of gastric cancer (GC) cells in any DNA sample by isolating genomic regions specifically methylated in GC cells.

Methods Genome-wide and gene-specific methylation analyses were conducted with an Infinium HumanMethylation450 BeadChip array and by quantitative methylation-specific PCR, respectively. Purified cancer and noncancer cells were prepared by laser-capture microdissection. *TP53* mutation data were obtained from our previous study using next-generation target sequencing.

Results Genome-wide DNA methylation analysis of 12 GC cell lines, 30 GCs, six normal gastric mucosae, one sample of peripheral leukocytes, and four noncancerous gastric mucosae identified *OSR2*, *PPFIA3*, and *VAV3* as barely methylated in normal cells and highly methylated in cancer cells. Quantitative methylation-specific PCR using 26 independent GCs validated that one or more of them was highly methylated in all of the GCs. Using four pairs of purified cells, we confirmed the three genes were highly methylated (85 % or more) in cancer cells and barely methylated (5 % or less) in noncancer cells. The cancer cell fraction assessed by the panel of the three genes showed good correlation with that assessed by the *TP53* mutant allele frequency in 13 GCs ($r = 0.77$). After correction of the GC cell fraction, unsupervised clustering analysis of the genome-wide DNA methylation profiles yielded clearer clustering.

Conclusions A DNA methylation marker—namely, the panel of the three genes—is useful to estimate the cancer cell fraction in GCs.

Keywords Gastric cancer · Cancer cell fraction · DNA methylation · Epigenetics

Electronic supplementary material The online version of this article (doi:10.1007/s10120-015-0475-2) contains supplementary material, which is available to authorized users.

L. Zong · N. Hattori · Y. Yoda · S. Yamashita · H. Takeshima · T. Takahashi · M. Maeda · T. Ushijima (✉)
Division of Epigenomics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
e-mail: tushijim@ncc.go.jp

L. Zong · Y. Seto
Department of Gastrointestinal Surgery, Graduate School of Medicine, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-0033, Japan

H. Katai
Gastric Surgery Division, National Cancer Center Hospital, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan

S. Nanjo · T. Ando
Third Department of Internal Medicine, University of Toyama, 2630 Sugitani, Toyama 930-0194, Japan

Introduction

Extensive genomic and epigenomic analyses of a variety of human cancers, including gastric cancers (GCs), have been and are being conducted [1–4]. However, these analyses are almost always affected by contamination from coexisting normal cells in primary cancer samples. Although genomic analyses are designed to detect mutations even in a small fraction of cells, they still fail to detect gene mutations in samples with a low fraction of cancer cells [5]. Moreover, epigenomic and gene expression analyses are heavily affected by the fraction of cancer cells [6]. To

Published online: 13 February 2015

overcome the contamination from normal cells, laser-capture microdissection (LCM) is conducted [7, 8]. However, LCM is labor-intensive and time-consuming, and practically impossible for diffuse-type GCs.

Without purification of cancer cells, if a fraction of cancer cells in a sample can be assessed, a sample with an extremely low fraction of cancer cells can be excluded from subsequent analyses, or the data obtained may be corrected by the fraction of cancer cells. Such assessment has been generally conducted by an expert pathologist, which is time-consuming and almost impossible for diffuse-type GCs and a large number of samples. To overcome this limitation, efforts have been made to develop molecular markers. For example, cancer-cell-specific mutations identified by a single-nucleotide polymorphism microarray and next-generation sequencing can be used to assess the fraction of cancer cells [9, 10]. However, identification of such mutations must be conducted for each sample, and there is a sizable research cost for this approach.

To overcome these issues, in our recent study, we successfully isolated CpG islands specifically methylated in esophageal squamous cell carcinoma (ESCC) cells [11]. Three genes were methylated in almost all ESCC cells, but were not methylated or were barely methylated in normal esophageal mucosae, and at least one of the three genes was methylated in virtually all of 28 ESCC cases analyzed. Therefore, a panel of the three genes was considered to be a DNA methylation marker for the fraction of cancer cells. Using the marker, we were able to correct the fraction of ESCC cells, and showed that tumor-suppressor genes were methylated in almost all cancer cells.

In this study, for GCs, we aimed to isolate a DNA methylation marker that can be used to assess the fraction of cancer cells. Different from the esophagus, isolation of such a marker is far more difficult because gastric mucosae can have very high levels of DNA methylation owing to *Helicobacter pylori* infection [12–15], and GC samples are contaminated with such gastric mucosae. Therefore, we paid special attention to isolation of marker genes not influenced by *H. pylori* infection.

Materials and methods

GC cell lines and tissue samples

Cell lines KATOIII, MKN45, NUGC3, MKN74, and MKN7 were purchased from the Japanese Collection of Research Bioresources (Tokyo, Japan), and the AGS cell line was purchased from the American Type Culture Collection (Manassas, VA, USA). Cell lines HSC39, HSC57, 44As3, and 58As9 were gifted by K. Yanagihara from the National Cancer Center, the TMK1 cell line was gifted by W. Yasui from Hiroshima University, and the GC2 cell line was established by M. Tatematsu at Aichi Cancer Center Research Institute.

A total of 56 primary GC samples (32 intestinal type and 24 diffuse type) were collected from surgical specimens of patients who had undergone gastrectomy, and 30 of the samples were used for our previous studies [1, 16]. Genome-wide DNA methylation and *TP53* mutation data of the 30 GCs were obtained from one of the studies [1]. Peripheral leukocyte samples were collected from five healthy volunteers by a centrifugation method. Gastric mucosae were collected by endoscopic biopsy from 17 healthy volunteers (11 without and six with present *H. pylori* infection) and from noncancerous gastric mucosae of 27 GC patients. Among the 27 noncancerous gastric mucosae, 23 (nine without and 14 with present *H. pylori* infection) were used for our previous study [17]. *H. pylori* infection status was analyzed by a serum anti-*H. pylori* IgG antibody test (SRL, Tokyo, Japan), rapid urease test (Otsuka, Tokushima, Japan), or culture test (Eiken, Tokyo, Japan).

All of the samples, except for those used for LCM, were stored in RNAlater (Applied Biosystems, Foster City, CA, USA), and genomic DNA was extracted by the phenol–chloroform method. LCM was performed using formalin-fixed paraffin-embedded primary GCs by a Leica LMD7000 system [7, 18]. This study was conducted with the approval of the Institutional Review Board of the National Cancer Center. Written informed consent was obtained from all individuals.

Genome-wide DNA methylation analysis

Genome-wide DNA methylation analysis was performed using an Infinium HumanMethylation450 BeadChip array (Illumina, San Diego, CA, USA), which assessed the degree of methylation of 485,512 CpG sites. The methylation level of each CpG site was obtained as a β value, which ranged from 0 (completely unmethylated) to 1 (completely methylated). We excluded 11,551 CpG sites on the sex chromosomes, and the remaining 473,961 CpG sites were used for the analysis. Genomic blocks were defined as collections of CpG sites classified by their locations against transcription start sites and CpG islands [1].

Gene-specific DNA methylation analysis

Gene-specific DNA methylation levels were analyzed by quantitative methylation-specific PCR (qMSP). For DNA from surgical specimens in RNAlater, 1 μg was digested with *Bam*HI, treated with bisulfite, purified, and suspended in 40 μl of Tris (hydroxymethyl) aminomethane–EDTA buffer, as described in [19, 20]. For formalin-fixed paraffin-

embedded samples collected by LCM, DNA extraction and bisulfite treatment was conducted with an EpiTect Plus bisulfite kit (Qiagen, Hilden, Germany). Quantitative methylation-specific PCR (qMSP) was performed by real-time PCR using primers specific to methylated or unmethylated DNA (Table S1), the bisulfite-treated DNA, and SYBR Green I (BioWhittaker Molecular Applications, Rockland, ME, USA). The number of molecules in a sample was determined by comparing its amplification with that of standard DNA samples that contained known numbers of molecules (10^1–10^6 molecules). On basis of the numbers of methylated and unmethylated molecules, a methylation level was calculated as the fraction of methylated molecules in the total number of DNA molecules (number of methylated molecules plus number of unmethylated molecules). As a fully methylated control, blood genomic DNA treated with *Sss*I methylase (New England Biolabs, Beverly, MA, USA) was used. As a fully unmethylated control, blood genomic DNA amplified twice with Genomiphi (GE Healthcare, Piscataway, NJ, USA) was used [21].

Gene expression analysis

Complementary DNA was synthesized from 1 µg of total RNA using SuperScript III (Invitrogen, Carlsbad, CA, USA). Quantitative reverse transcription PCR was performed using SYBR Green I and an iCycler thermal cycler. The measured number of complementary DNA molecules was normalized to that of *GAPDH*. The primers and PCR conditions are shown in Table S1.

Genomic DNA copy number analysis

Copy number alteration (CNA) of a specific genomic region was analyzed by quantitative real-time PCR using an iCycler thermal cycler and SYBR Green I. *RPPH1* was used as a control gene located on a chromosomal region with infrequent CNA [22]. The number of DNA molecules in a sample was measured for the control gene and three regions flanking the target gene (Table S1). The number of DNA molecules of the target gene was normalized to that of the control gene, and the normalized number of DNA molecules in a sample was compared with that in human leukocyte DNA to obtain the CNA. All the analysis was conducted in duplicate. A CNA (gain or loss) was defined as a twofold or greater increase or a 0.5-fold or smaller decrease.

Mutations of *TP53* and mutant frequency

The *TP53* mutation status and mutant frequency were obtained from our previous study [1]. Briefly, the mutation was analyzed by target sequencing using an Ion AmpliSeq cancer panel kit (Life Technologies, Carlsbad, CA, USA) and an Ion PGM next-generation sequencer.

Statistical analyses

The correlation was analyzed using Pearson's product-moment correlation coefficients, and its P value was obtained by the parametric hypothesis test. A difference in the mean DNA methylation level was analyzed by Student's t test. A result was considered significant when the P value was less than 0.05 by a two-sided test.

Results

Selection of regions specifically methylated in GCs by a genome-wide screening

To screen specific regions not methylated in normal cells and fully methylated in GC cells using the Infinium HumanMethylation450 BeadChip array, we searched for CpG sites (1) with $\beta \leq 0.2$ in six samples of normal gastric mucosae, one sample of peripheral leukocytes, and four samples of noncancerous mucosae, and (2) with $\beta \geq 0.8$ in at least six of 12 GC cell lines. A total of 1,006 CpG sites were isolated from 473,961 informative CpG sites on autosomes. Then, to screen regions frequently methylated in primary GCs, CpG sites for which $\beta \geq 0.3$ in 20 or more of 30 primary GCs [1] were searched (Fig. 1a). From the 1,006 CpG sites, 18 CpG sites derived from 16 genomic regions were isolated (Table S2). From the 16 genomic regions, *PRDM16* was excluded because its gene amplification was known [23], and five other regions were also excluded because they did not have neighboring CpG islands or known genes.

For the remaining ten regions, we attempted to design primers for qMSP, and primers for both methylated and unmethylated DNA were successfully designed for five regions of five genes (*OSR2*, *VAV3*, *PPFIA3*, *LTB4R2*, and *DIDO1*) (Fig. 1b). To confirm the genome-wide DNA methylation data obtained by the bead array, qMSP was conducted using the 12 GC cell lines mentioned in "GC cell lines and tissue samples" and one sample of peripheral leukocytes. *DIDO1* had slight methylation in the peripheral leukocytes, and was excluded from further analysis. The methylation levels of the other four genes (*LTB4R2*, *OSR2*, *VAV3*, and *PPFIA3*) obtained by qMSP were in good accordance with the bead array data (Fig. S1).

Isolation of genes not influenced by *H. pylori* infection

Gastric mucosae with *H. pylori* infection are known to have very high DNA methylation levels [12, 13]. To