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 *Sss*I 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

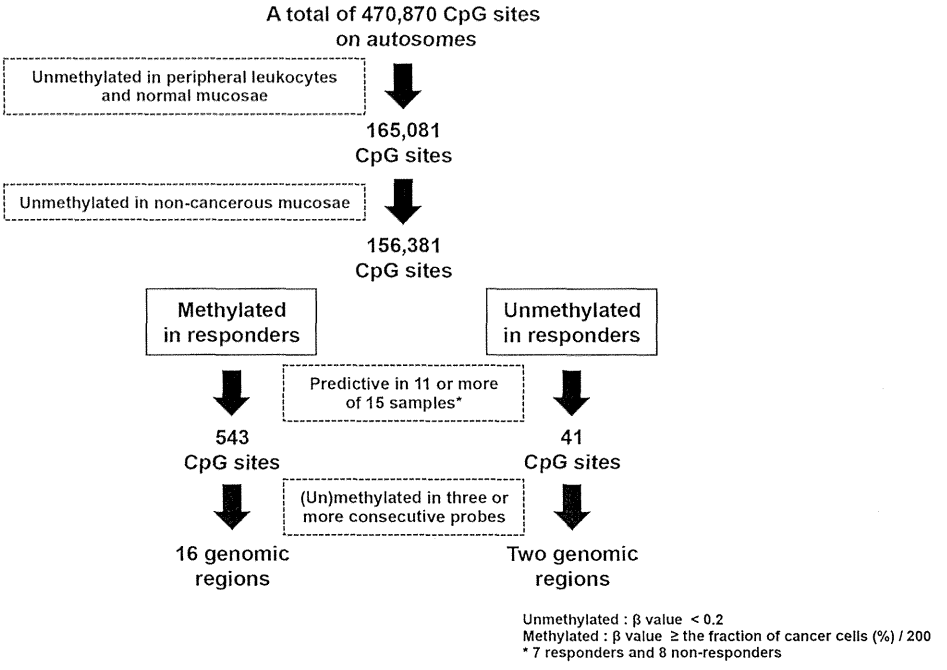

Fig. 1 Selection processes of genomic regions whose methylation statuses were associated with the response to dCRT. From the 482,421 CpG sites on the Infinium HumanMethylation450 BeadChip array, those on autosomes and unmethylated in non-cancerous mucosae were selected. As specifically methylated in 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

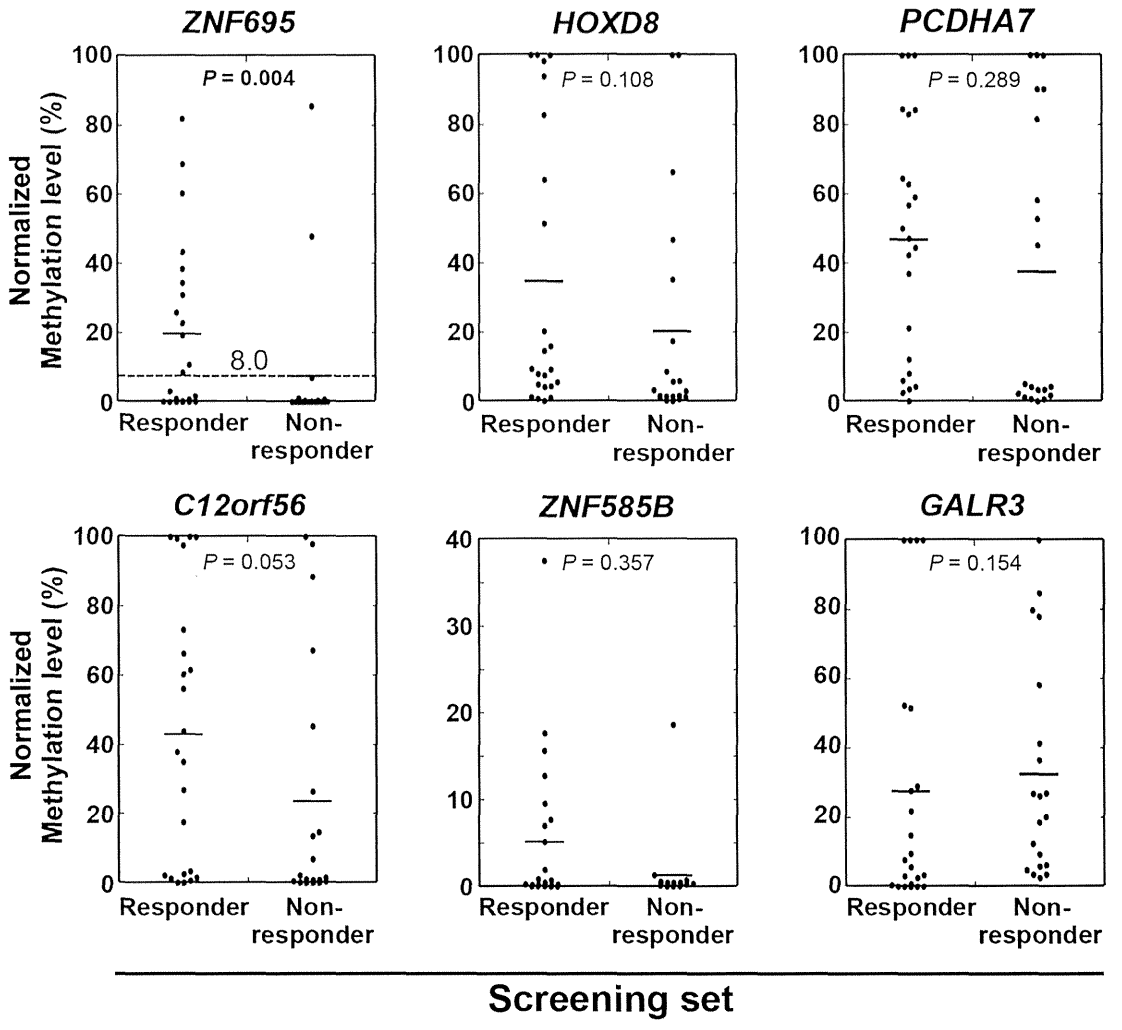

Fig. 2 Normalized methylation levels of the candidate genomic regions in the screening set. Methylation levels of six genomic regions were measured by qMSP in 23 responders and 20 non-responders in the screening set. A normalized methylation level was calculated using the fraction of cancer cells in 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.

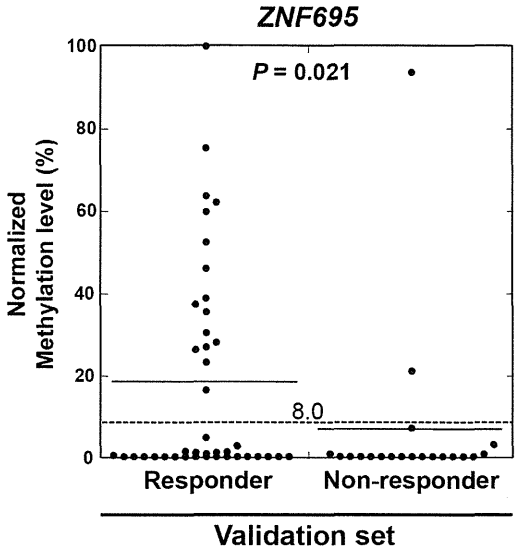

Fig. 3 Normalized methylation levels of *ZNF695* in the validation set. Methylation levels of *ZNF695* were measured by qMSP in 41 responders and 20 non-responders in the validation set. A normalized methylation level was calculated as described in the legend of Fig. 2. The *horizontal dotted line* shows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

Clinical Translational Research

Oncology

Oncology
DOI: 10.1159/000369904

Received: September 18, 2014
Accepted after revision: November 11, 2014
Published online: January 14, 2015

Comprehensive DNA Methylation and Extensive Mutation Analyses of HER2-Positive Breast Cancer

Takeshi Yamaguchi[a] Hirofumi Mukai[c] Satoshi Yamashita[b] Satoshi Fujii[d] Toshikazu Ushijima[b]

[a]Department of Medical Oncology, Musashino Red Cross Hospital, and [b]Division of Epigenomics, National Cancer Center Research Institute, Tokyo, and [c]Department of Breast and Medical Oncology, and [d]Pathology Division, Research Center for Innovative Oncology, National Cancer Center Hospital East, Chiba, Japan

Key Words
Breast cancer · Epigenome · Gene mutation · HER2 · Ki-67 · Neoadjuvant chemotherapy · Next-generation sequencer · Trastuzumab

Abstract
Objective: Resistance to trastuzumab is a problem that remains to be solved in HER2-positive breast cancer. We aimed to characterize profiles of genetic and epigenetic alterations in cancer-related pathways in HER2-positive breast cancers, using biopsy tissue samples obtained from patients enrolled in a prospective neoadjuvant clinical trial. ***Methods:*** HER2-positive breast cancer tissue samples were collected and processed with the PAXgene Tissue System. A total of 24 breast cancers were analyzed. Genetic alterations of 409 cancer-related genes were analyzed by a bench-top next-generation sequencer. DNA methylation statuses were analyzed by a bead array with 485,512 probes. ***Results:*** The WNT pathway was potentially activated by aberrant methylation of its negative regulators, such as *DKK3* and *SFRP1*, in 9 breast cancers. The AKT/mTOR pathway was activated by mutations of *PIK3CA* in 5 breast cancers. The Notch pathway was potentially activated by mutations of *NOTCH1* and *NOTCH2* in 4 breast cancers. The p53 pathway was inactivated by mutations of *TP53* in 13 breast cancers and potentially by aberrant methylation of its downstream genes in 10 breast cancers. Cell adhesion was affected by mutations of *CDH1* in 1 breast cancer. ***Conclusion:*** Genes involved in cancer-related pathways were frequently affected not only by genetic but also by epigenetic alterations in HER2-positive breast cancer.

Introduction

Overexpression of human epidermal growth factor receptor 2 (HER2) or amplification of the *HER2* gene is observed in 15–25% of breast cancers (BCs), and, independently, both are factors discriminating for poor prognosis [1]. Adding trastuzumab, which is a humanized monoclonal antibody that targets HER2, to chemotherapy improves survival in patients with HER2-positive metastatic BC [2]. Therefore, HER2 was recognized as an important therapeutic target in BC. Trastuzumab has also shown efficacy against early-stage BC as an adjuvant and neoadjuvant therapy [3]. Although trastuzumab has revolutionized the treatment of HER2-positive BC, resis-

0030–2414/15/0000–0000$39.50/0

E-Mail karger@karger.com
www.karger.com/ocl

Dr. Hirofumi Mukai
Department of Breast and Medical Oncology
National Cancer Center Hospital East
6-5-1, Kashiwanoha, Kashiwa-shi, Chiba 277-8577 (Japan)
E-Mail hrmukai@east.ncc.go.jp

tance to this substance is a problem that remains to be solved. Response rates to trastuzumab monotherapy as first- and second-line treatments for metastatic BC have been reported to be 15 and 26%, respectively [4, 5]; most patients do not respond to this agent alone (de novo resistance). The combination of chemotherapy and trastuzumab shows a high response rate in metastatic BC; however, all patients eventually show resistance (acquired resistance). Therefore, the discovery of new biomarkers and an elucidation of the mechanisms of resistance to trastuzumab are needed for the development of new treatments in the future.

Genetic and epigenetic alterations are involved in cancer development and progression by activating growth-promoting pathways and inactivating tumor-suppressive pathways. Various studies have been conducted in vivo and in vitro, and several mechanisms of resistance to HER2-targeted therapy have been proposed. However, many of these studies are retrospective [6–9]. Until recently, these genetic and epigenetic alterations have been assessed individually, because technologies for their comprehensive analysis have not been available at a reasonable cost. Now, point mutations and gene amplifications of a large number of target genes can be analyzed by bench-top next-generation sequencers [10]. A comprehensive DNA methylation profile can be analyzed using a bead array [11]. The process of the combination of comprehensive DNA methylation and mutation analyses as well as pathway analysis using these data was established in our previous studies of gastric cancer [12, 13].

In this study, we aimed to establish an integrated profile of genetic and epigenetic alterations in HER2-positive BC by using tissue samples obtained from a neoadjuvant clinical trial.

Methods

Study Design of the Neoadjuvant Clinical Trial

We have conducted a trial to determine if neoadjuvant chemotherapy can be optimized by using the Ki-67 index changes during neoadjuvant chemotherapy in HER2-positive BC. This trial was registered in the UMIN Clinical Trials Registry (registration No. UMIN000007074). The details of the trial have been described in a previous paper [14]. In brief, this was a randomized phase II trial in which women aged 20–75 years with histologically confirmed HER2-positive BC stage II–III were eligible patients. HER2 positivity was defined as overexpression by immunohistochemistry (3+) or gene amplification by fluorescence in situ hybridization.

The patients were required to have a good performance status, sufficient organ functions, a normal left ventricular ejection fraction as well as no prior endocrine therapy or chemotherapy for BC. They were randomly assigned to weekly paclitaxel (80 mg/m^2 a week) and trastuzumab (a loading dose of 4 mg/kg followed by 2 mg/kg a week) for a total of 12 doses or to Ki-67 index-guided treatment as a preoperative treatment. In Ki-67 index-guided treatment, the patients initially received weekly paclitaxel and trastuzumab as in the reference arm, and a primary tumor was biopsied during days 15–21 for estimation of the Ki-67 index. The subsequent chemotherapy regimen was adjusted according to changes in Ki-67 index from baseline.

Between December 2011 and December 2013, 133 patients were registered, with the aim of reaching 200 patients; registration is currently still in progress (as of April 2014). The study protocol was approved by the National Cancer Center Ethics Committee (approval No. 2010-250). The trial met the Ethical Guidelines for Clinical Studies of the Japanese Ministry of Health, Labor and Welfare and was conducted in compliance with the Declaration of Helsinki. The patients provided written informed consent.

Samples

From among the patients enrolled in the neoadjuvant clinical trial so far, samples obtained from 24 patients were used. One or two core needle biopsy samples of the primary tumors were collected at baseline and days 15–21. An experienced pathologist (S.F.) evaluated whether the components of the tumor cells in the samples were sufficient enough for molecular analysis. In the present study, 24 HER2-positive BC tissue samples obtained at baseline were analyzed (22 invasive ductal carcinoma samples, 1 invasive lobular carcinoma sample and 1 invasive micropapillary carcinoma sample); 8 BCs were estrogen receptor positive, and 8 were progesterone receptor positive.

Tissue samples were fixed and stored with the PAXgene Tissue System according to the manufacturer's instructions (Qiagen, Germany). Genomic DNA was extracted with a PaxGene DNA Kit (Qiagen), and extracted DNA was quantified using a Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, USA). Total RNA was extracted with a PaxGene RNA Kit (Qiagen). The PAXgene Tissue System is a recently developed tissue stabilization technology. So far, researchers have reported that RNA and DNA purified from PAXgene-fixed tissues are of high integrity and that they perform as well as those from fresh frozen tissue [15–17].

Analysis of Somatic Mutations

Four independent multiplex PCRs amplifying a total of 15,991 regions in 409 cancer-related genes were performed on genomic DNA, using the Ion AmpliSeq Library Kit 2.0 with the Comprehensive Cancer Panel (Life Technologies). The amplified products were mixed equally and then uniquely barcoded with an Ion Xpress Barcode Adapters 1-96 Kit (Life Technologies); 4–8 barcoded libraries were pooled, and an emulsion PCR was performed with an Ion OneTouch 2 device with an Ion PI Template OT2 200 Kit v2 (Life Technologies). Template-positive Ion Sphere particles were concentrated with an Ion OneTouch ES (Life Technologies) and loaded onto an Ion PI Chip v2 (Life Technologies). Sequencing was performed with a bench-top next-generation sequencer, an Ion Proton sequencer with an Ion PI Sequencing 200 Kit v2 (Life Technologies). Primer sequences were removed from the obtained sequences, and the trimmed sequences were aligned onto the human reference genome hg19 with Torrent Suite (Life Technologies). By using CLC Genomics Workbench 6.0 (CLC bio, Denmark), a variation was identified as a functional mutation only

if (1) its frequency was >10%, (2) its read count was >40, (3) it was found in both at least 5 forward and 5 reverse reads, (4) it was not present in >10% of the reads in any normal samples and (5) it caused amino acid changes or splicing defects. All the mutations identified by computational analysis were validated by manual inspection of the alignment data.

Analysis of Gene Amplifications

Gene amplifications of 409 genes were analyzed using the alignment data; the average reading depth of each target region (amplicon) in each sample was obtained by CLC Genomics Workbench. To evaluate copy number variations in a gene, the relative reading depth to the reference (RRDR) of an individual gene was calculated as following: RRDR = (average of the reading depths of the target regions in an individual gene/average of the reading depths of all regions in the panel)$_{cancer}$/(average of the reading depths of the target regions in an individual gene/average of the reading depths of all regions in the panel)$_{normal}$, where 'cancer' and 'normal' mean a cancerous sample and its matched normal tissue (blood), respectively. Regions in which the average reading depth of all the samples was >50 were employed for the analysis. Genes whose RRDRs were larger (two-fold or more) than those of the other genes were defined as amplified genes.

Selection of Genes of Cancer-Related Pathways

A total of 64 genes involved in 9 cancer-related pathways (MAPK, WNT, AKT/mTOR, Notch, Hedgehog, cell cycle regulation, mismatch repair, p53 and cell adhesion) were selected according to our previous study [13]. Regarding the signaling pathways activated in BCs, their negative regulators were selected. Regarding the pathways inactivated in BCs, their positive regulators and downstream effectors were selected.

Analysis of DNA Methylation

DNA methylation levels of 485,512 probes (482,421 probes for CpG sites and 3,091 probes for non-CpG sites) were obtained using an Infinium HumanMethylation450 BeadChip array, as previously described [18]. To adjust for probe design biases, intra-array normalization was performed using a peak-based correction method, BMIQ (Beta Mixture Quantile dilation) [19]. The methylation level of each CpG site was represented by a β-value that ranged from 0 (unmethylated) to 1 (fully methylated).

DNA methylation of a CpG island (CGI) in a promoter region, especially in the 200-bp upstream region from a transcription start site (TSS200), is known to consistently silence its downstream gene, while that of downstream exons is weakly associated with increased expression [20–23]. Therefore, we paid as much attention as possible to analyze the DNA methylation of a CGI in a TSS200. To achieve this, probes for CpG sites were assembled into 296,494 genomic blocks <500 bp. Of these blocks, 59,757 were located in CGIs, and 11,307 of them were located in TSS200s. The selection of genomic blocks to analyze genes of cancer-related pathways was performed according to our previous study [13]. The DNA methylation level of a genomic block was evaluated using the mean β-value of all the probes within the genomic block. The methylation levels were corrected using cancer cell contents pathologically analyzed by an experienced pathologist (S.F.). The methylation status of the genomic blocks was classified into unmethylated (corrected β-value: 0–0.2), partially methylated (corrected β-value: 0.2–0.8) and heavily methylated (corrected β-value: 0.8–1.0).

Analysis of Gene Expression

Gene expression in normal human mammary epithelial cells (HMECs) was analyzed with the GeneChip Human Genome U133 Plus 2.0 microarray (Affymetrix, USA). Genes with signal intensities of ≥250 were defined as expressed genes [24].

Results

Methylation Silenced Genes in HER2-Positive BC

Comprehensive DNA methylation analysis was performed on 24 HER2-positive BCs. The analysis was conducted using 7,103 TSS200 CGIs unmethylated in normal mammary epithelial cells (genes unmethylated in HMECs), because a TSS200 CGI is known to play a critical role in methylation silencing [25]. The number of aberrantly methylated genes ranged from 9 to 629. Next, we focused on TSS200 CGIs of genes with positive expression in normal cells but aberrantly methylated in cancer cells, because this group of genes is known to frequently contain driver genes in carcinogenesis [26]. Using 292 TSS200 CGIs whose downstream genes were expressed in normal mammary epithelial cells and aberrantly methylated in ≥1 HER2-positive BCs (methylation-silenced genes; online suppl. table S1; for all online suppl. material, see www.karger.com/doi/10.1159/000369904), the number ranged from 2 to 132. These results showed that the number of aberrantly methylated genes was highly variable even among individual HER2-positive BCs.

Point Mutations and Gene Amplifications in HER2-Positive BC

Of the 24 HER2-positive BCs analyzed for mutations of the 409 cancer-related genes, 23 had 80 somatic mutations (online suppl. table S2). The *PIK3CA* oncogene was mutated in 20.8% of the HER2-positive BCs. *TP53* was most frequently mutated (54.2% of the HER2-positive BCs), and *PIK3CA*, *NOTCH2*, *RNF213*, *ADAMTS20*, *PKHD1*, *ROS1*, *SETD2*, *ZNF521* and *TAF1L* were mutated in ≥2 HER2-positive BCs (fig. 1).

Gene amplification was analyzed as well for the 409 cancer-related genes in the 24 HER2-positive BCs, and all BCs had 169 amplifications in 80 genes (online suppl. table S3). *ERBB2* amplification was detected in 22 of the 24 HER2-positive BCs. The remaining 2 samples (12D, 25D) showed 1.2- and 1.3-fold RRDRs, respectively. *MAF*, *PLAG1*, *EXT*, *CSMD2* and *MYC* were also amplified in ≥3 HER2-positive BCs, except for genes on chromosome 17q (close to HER2).

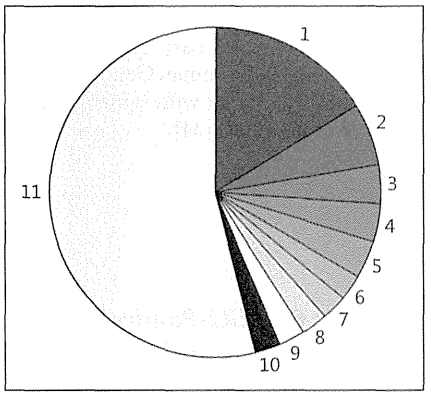

Fig. 1. Gene mutations in 24 HER2-positive BCs. A total of 80 somatic mutations found in these BCs are shown in the circle graph on the left side. Genes mutated in ≥2 HER2-positive BCs are listed on the right side.

Growth-Promoting Pathways Affected by Epigenetic and Genetic Alterations

Aberrant DNA methylation of the 64 genes involved in the 9 cancer-related pathways was combined with genetic alterations in the 24 HER2-positive BCs. First, the potential activation of growth-promoting pathways by aberrant methylation of their negative regulators, in addition to activating genetic alterations (point mutations and gene amplifications), were analyzed (fig. 2).

Regarding the MAPK pathway, in addition to amplification of *ERBB2*, 15 of the 24 BCs had heavy aberrant methylation of its 1 negative regulator: *RASSF1* (fig. 2a). As for the AKT/mTOR pathway, only 1 of the 24 BCs had heavy aberrant methylation of its 4 negative regulators, and 5 BCs had point mutations of *PIK3CA* or *PTPN11* (fig. 2b).

Regarding the WNT pathway, all the BCs had heavy or partially aberrant methylation of ≥1 of its 16 negative regulators, such as *DKK3* and *SFRP1* (fig. 2c). To exclude concerns that we analyzed the methylation of genes that had little expression in normal epithelial cells and thus were susceptible to methylation [24], we confirmed that 7 of the 16 negative regulators were moderately or abundantly expressed (signal intensity >250) in normal epithelial cells. Even when limited to these 7 genes, *DKK3* or *SFRP1* were heavily methylated in 9 BCs. In contrast, no BCs had point mutations of *CTNNB1*.

With regard to the Notch pathway, no BCs had heavy aberrant methylation of its negative regulators, and 4 BCs had a mutation of *NOTCH1* or *NOTCH2* (fig. 2d). Regarding the Hedgehog pathway, only 1 BC had heavy aberrant methylation of its 1 negative regulator, and no BCs had a mutation of *PTCH1* (fig. 2e).

Tumor-Suppressive Pathways Affected by Epigenetic and Genetic Alterations

Tumor-suppressive pathways inactivated in BCs were also analyzed. Regarding cell cycle regulation, none of the 24 BCs had heavy aberrant methylation of *CDKN2A* and/or *CHFR*, and none of the 24 BCs had point mutations of *CDKN2A* (fig. 3a). Concerning mismatch repair, none of the 24 BCs had heavy aberrant methylation of *MLH1*, and none of the BCs had a point mutation (fig. 3b).

As for the p53 pathway, 13 BCs had point mutations of *TP53*. Twenty-four downstream genes had promoter CGIs, and 10 BCs had heavy aberrant methylation of ≥1 of the 24 genes (fig. 3c). Among the 24 genes, *IGFBP7* was abundantly expressed in normal mammary epithelial cells, and 6 BCs had heavy aberrant methylation.

Regarding cell adhesion, 13 BCs had partial aberrant methylation of *CDH1*, but none of the 24 BCs had heavy aberrant methylation. At the same time, 1 BC had its point mutations (fig. 3d). Taken together, these results demonstrate that genes in cancer-related pathways were frequently affected by epigenetic alterations even in HER2-positive BCs.

Discussion

In this study, we performed genetic and epigenetic profiling of 24 HER2-positive BCs and showed that genes in cancer-related pathways were frequently affected by not only genetic but also epigenetic alterations in HER2-positive BCs. When genetic and epigenetic alterations were combined, almost all of the 24 HER2-positive BCs had alterations in cancer-related pathways. The alterations in these pathways are potential targets for therapy, and they possibly have a role in the mechanism of resistance to trastuzumab.

In addition to amplification of HER2, 15 of the 24 HER2-positive BCs had heavy aberrant methylation of *RASSF1*. Since *RASSF1* is a negative regulator of the MAPK pathway, silencing of *RASSF1* by DNA methylation is equivalent to releasing the brakes restraining the activation of the MAPK pathway by HER2 amplification. Methylation of *RASSF1* may play an important role in the activity of the MAPK pathway in HER2-positive BC and might also be involved in the therapeutic response.

Among the 24 HER2-positive BCs, some had mutations and amplifications of target genes for molecular tar-

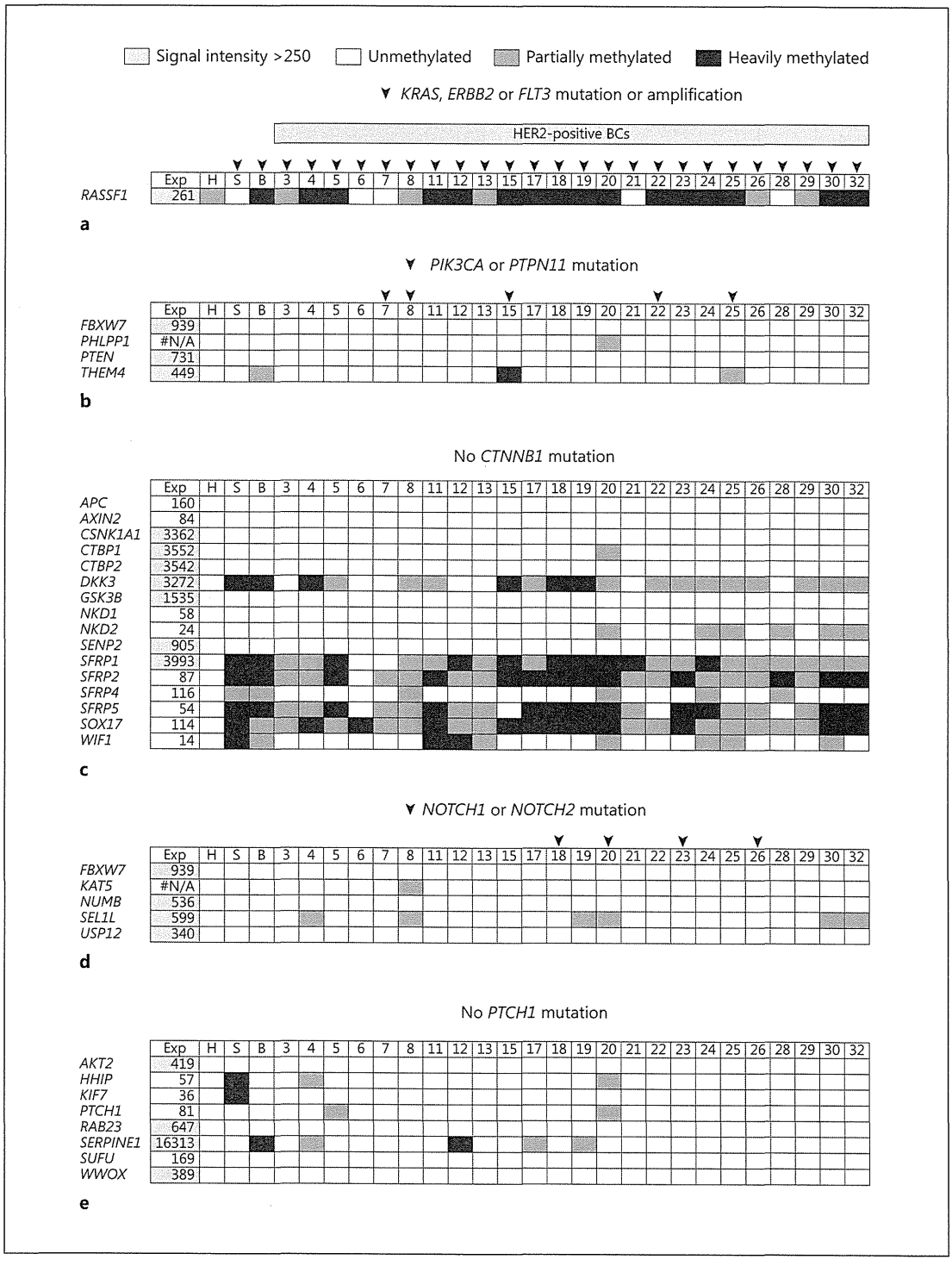

Fig. 2. Genetic and epigenetic alterations in 5 growth-promoting pathways. **a** In the MAPK pathway, 15 of the 24 HER2-positive BCs had heavy aberrant methylation of *RASSF1*. **b** In the AKT/mTOR pathway, 4 BCs had point mutations of *PIK3CA* or *PTPN11* (shown by arrowheads). In contrast, none of the 24 BCs had heavy aberrant methylation of negative regulators of the AKT/mTOR pathway. **c** WNT pathway. **d** Notch pathway. **e** Hedgehog pathway. Exp = Expression levels in HMECs; H = HMECs (normal cell lines); S = SK-BR-3 (HER2-positive BC cell lines); B = BT474 (HER2-positive BC cell lines).

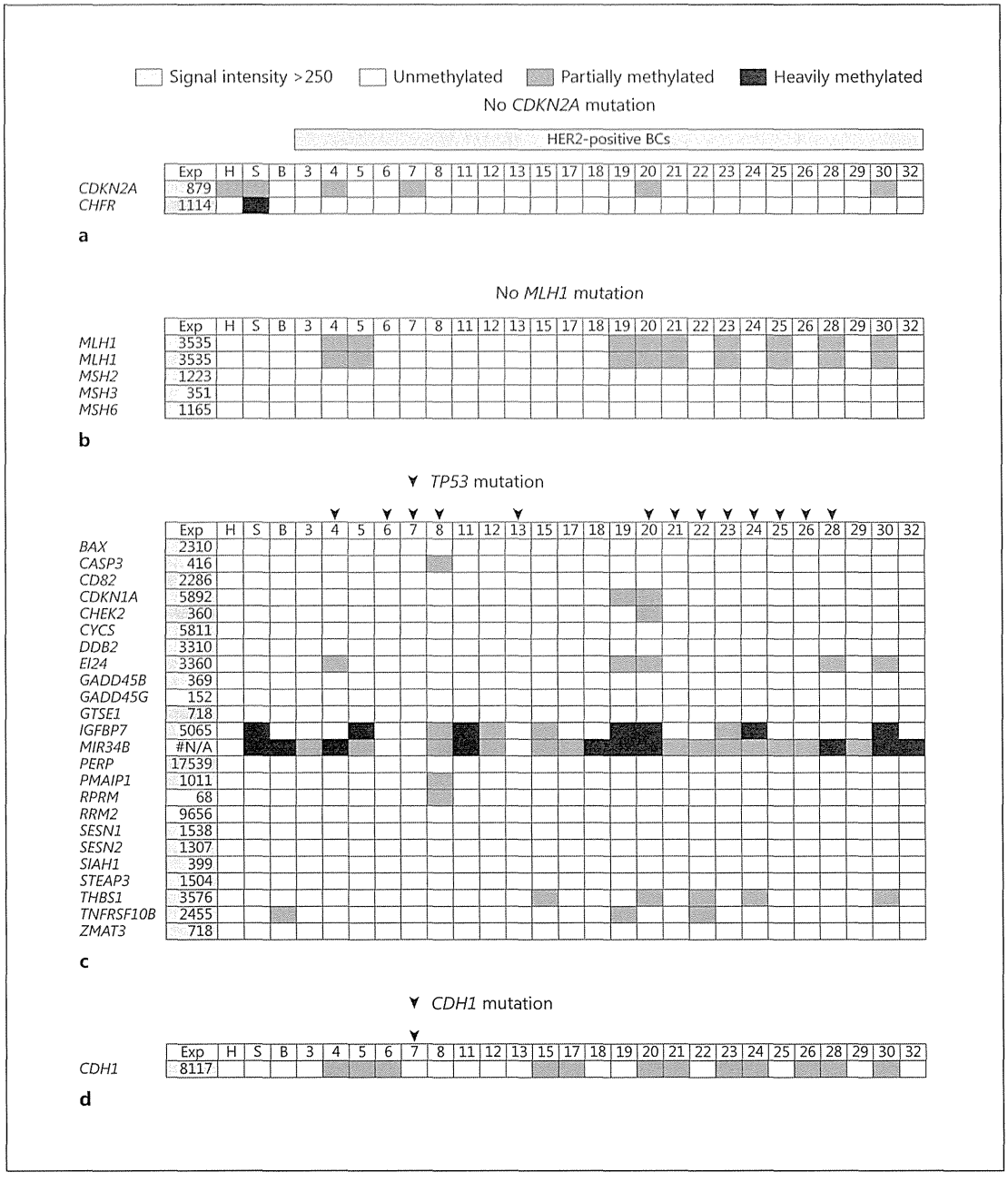

Fig. 3. Genetic and epigenetic alterations in 4 tumor-suppressor pathways. **a** Cell cycle regulation. None of the BCs had point mutations of *CDKN2A* or heavy aberrant methylation of *CDKN2A* and/or *CHFR*. **b** Mismatch repair. For *MLH1*, 2 genomic blocks in its 2 TSS200s were analyzed. **c** p53 pathway. Thirteen BCs had point mutations of *TP53* (shown by arrowheads), and 6 BCs had heavy aberrant methylation of *IGFBP7*. **d** Cell adhesion. One BC had a mutation of *CDH1* (shown by an arrowhead), but none of the 24 BCs had heavy aberrant methylation of *CDH1*.

geted therapy additionally to ERBB2 targeting. In this study, 21% of the BCs had point mutations of the *PIK3CA* gene, which is involved in the AKT/mTOR pathway. This is concordant with a study reporting that 20% of HER2-positive BCs have point mutations of this gene [11]. Activation of this pathway has been reported to be involved in therapeutic efficacy in HER2-positive BC. However, this study suggests that the possibility of activation by aberrant DNA methylation is low.

Aberrant DNA methylation of the negative regulator of the WNT pathway was observed in almost all HER2-positive BCs. Mutations of the Notch pathway-related genes were observed in 4 of the 24 BCs. As for the p53 pathway, mutations of *TP53* and aberrant methylation of the genes of the p53 pathway were observed in a complementary manner. When aberrant methylation and genetic abnormalities were combined, alterations to the p53 pathway were present in almost all HER2-positive BCs. On the other hand, the frequency of alterations in the Hedgehog pathway, cell cycle regulation, mismatch repair and cell adhesion was low both genetically and epigenetically. These results suggest the importance of the WNT and p53 pathways in HER2-positive BC. *DKK3* and *SFRP1* (involved in the WNT pathway) and *IGFBP7* (involved in the p53 pathway) were expressed in normal mammary epithelial cells and frequently methylated in HER2-positive BCs. It is known that downregulation of *DKK3* is correlated with tumor progression [27] and that *IGFBP7* can inhibit cell growth and induce apoptosis [28]. These results support that aberrant methylation of *DKK3*, *SFRP1* and *IGFBP7* was involved in the carcinogenesis of HER2-positive BCs.

ERBB2 amplification was detected with a bench-top next-generation sequencer in 22 of the 24 HER2-positive BCs. This discrepancy can be explained by the low cancer cell contents and low-level amplification of *ERBB2*. In this case, the relative copy number of *ERBB2* is low in DNA samples including both cancer and non-cancer cells.

In conclusion, an integrated profile of genetic and epigenetic alterations to cancer-related pathways in HER2-positive BCs was obtained using a bench-top next-generation sequencer and a bead array. The profile is expected to play an important role in the development of new treatments and the identification of biomarkers.

Acknowledgements

We appreciate the help in the next-generation sequencing analysis provided by Toru Niwa and Reiko Nagano.

Disclosure Statement

The authors declare that no competing interests exist.

References

1 Desmedt C, Haibe-Kains B, Wirapati P, Buyse M, Larsimont D, Bontempi G, Delorenzi M, Piccart M, Sotiriou C: Biological processes associated with breast cancer clinical outcome depend on the molecular subtypes. Clin Cancer Res 2008;14:5158–5165.

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;344:783–792.

3 Moja L, Tagliabue L, Balduzzi S, Parmelli E, Pistotti V, Guarneri V, D'Amico R: Trastuzumab containing regimens for early breast cancer. Cochrane Database Syst Rev 2012; 4:CD006243.

4 Cobleigh MA, Vogel CL, Tripathy D, Robert NJ, Scholl S, Fehrenbacher L, Wolter JM, Paton V, Shak S, Lieberman G, Slamon DJ: Multinational study of the efficacy and safety of humanized anti-HER2 monoclonal antibody in women who have HER2-overexpressing metastatic breast cancer that has progressed after chemotherapy for metastatic disease. J Clin Oncol 1999;17:2639–2648.

5 Vogel CL, Cobleigh MA, Tripathy D, Gutheil JC, Harris LN, Fehrenbacher L, Slamon DJ, Murphy M, Novotny WF, Burchmore M, Shak S, Stewart SJ, Press M: Efficacy and safety of trastuzumab as a single agent in first-line treatment of HER2-overexpressing metastatic breast cancer. J Clin Oncol 2002;20:719–726.

6 Perez EA, Dueck AC, McCullough AE, Chen B, Geiger XJ, Jenkins RB, Lingle WL, Davidson NE, Martino S, Kaufman PA, Kutteh LA, Sledge GW, Harris LN, Gralow JR, Reinholz MM: Impact of PTEN protein expression on benefit from adjuvant trastuzumab in early-stage human epidermal growth factor receptor 2-positive breast cancer in the North Central Cancer Treatment Group N9831 trial. J Clin Oncol 2013;31:2115–2122.

7 Jensen JD, Knoop A, Laenkholm AV, Grauslund M, Jensen MB, Santoni-Rugiu E, Anderson M, Ewertz M: *PIK3CA* mutations, PTEN, and pHER2 expression and impact on outcome in HER2-positive early-stage breast cancer patients treated with adjuvant chemotherapy and trastuzumab. Ann Oncol 2012; 23:2034–2042.

8 Dave B, Migliaccio I, Gutierrez MC, Wu MF, Chamness GC, Wong H, Narasanna A, Chakrabarty A, Hilsenbeck SG, Huang J, Rimawi M, Schiff R, Arteaga C, Osborne CK, Chang JC: Loss of phosphatase and tensin homolog or phosphoinositol-3 kinase activation and response to trastuzumab or lapatinib in human epidermal growth factor receptor 2-overexpressing locally advanced breast cancers. J Clin Oncol 2011;29:166–173.

9 Scaltriti M, Rojo F, Ocana A, Anido J, Guzman M, Cortes J, Di Cosimo S, Matias-Guiu X, Ramon y Cajal S, Arribas J, Baselga J: Expression of p95HER2, a truncated form of the HER2 receptor, and response to anti-HER2 therapies in breast cancer. J Natl Cancer Inst 2007;99:628–638.

10 Gullapalli RR, Lyons-Weiler M, Petrosko P, Dhir R, Becich MJ, LaFramboise WA: Clinical integration of next-generation sequencing technology. Clin Lab Med 2012;32:585–599.
11 Dedeurwaerder S, Defrance M, Calonne E, Denis H, Sotiriou C, Fuks F: Evaluation of the Infinium Methylation 450K technology. Epigenomics 2011;3:771–784.
12 Kim JG, Takeshima H, Niwa T, Rehnberg E, Shigematsu Y, Yoda Y, Yamashita S, Kushima R, Maekita T, Ichinose M, Katai H, Park WS, Hong YS, Park CH, Ushijima T: Comprehensive DNA methylation and extensive mutation analyses reveal an association between the CpG island methylator phenotype and oncogenic mutations in gastric cancers. Cancer Lett 2013;330:33–40.
13 Yoda Y, Takeshima H, Niwa T, Kim JG, Ando T, Kushima R, Sugiyama T, Katai H, Noshiro H, Ushijima T: Integrated analysis of cancer-related pathways affected by genetic and epigenetic alterations in gastric cancer. Gastric Cancer 2015;18:65–76.
14 Yamaguchi T, Mukai H: Ki-67 index guided selection of preoperative chemotherapy for HER2-positive breast cancer: a randomized phase II trial. Jpn J Clin Oncol 2012;42:1211–1214.
15 Belloni B, Lambertini C, Nuciforo P, Phillips J, Bruening E, Wong S, Dummer R: Will PAXgene substitute formalin? A morphological and molecular comparative study using a new fixative system. J Clin Pathol 2013;66:124–135.
16 Groelz D, Sobin L, Branton P, Compton C, Wyrich R, Rainen L: Non-formalin fixative versus formalin-fixed tissue: a comparison of histology and RNA quality. Exp Mol Pathol 2013;94:188–194.
17 Staff S, Kujala P, Karhu R, Rokman A, Ilvesaro J, Kares S, Isola J: Preservation of nucleic acids and tissue morphology in paraffin-embedded clinical samples: comparison of five molecular fixatives. J Clin Pathol 2013;66:807–810.
18 Shigematsu Y, Niwa T, Yamashita S, Taniguchi H, Kushima R, Katai H, Ito S, Tsukamoto T, Ichinose M, Ushijima T: Identification of a DNA methylation marker that detects the presence of lymph node metastases of gastric cancers. Oncol Lett 2012;4:268–274.
19 Teschendorff AE, Marabita F, Lechner M, Bartlett T, Tegner J, Gomez-Cabrero D, Beck S: A beta-mixture quantile normalization method for correcting probe design bias in Illumina Infinium 450 k DNA methylation data. Bioinformatics 2013;29:189–196.
20 Ushijima T: Detection and interpretation of altered methylation patterns in cancer cells. Nat Rev Cancer 2005;5:223–231.
21 Jones PA, Baylin SB: The epigenomics of cancer. Cell 2007;128:683–692.
22 Ball MP, Li JB, Gao Y, Lee JH, LeProust EM, Park IH, Xie B, Daley GQ, Church GM: Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. Nat Biotechnol 2009;27:361–368.
23 Yamashita S, Hosoya K, Gyobu K, Takeshima H, Ushijima T: Development of a novel output value for quantitative assessment in methylated DNA immunoprecipitation-CpG island microarray analysis. DNA Res 2009;16:275–286.
24 Takeshima H, Yamashita S, Shimazu T, Niwa T, Ushijima T: The presence of RNA polymerase II, active or stalled, predicts epigenetic fate of promoter CpG islands. Genome Res 2009;19:1974–1982.
25 Lin JC, Jeong S, Liang G, Takai D, Fatemi M, Tsai YC, Egger G, Gal-Yam EN, Jones PA: Role of nucleosomal occupancy in the epigenetic silencing of the *MLH1* CpG island. Cancer Cell 2007;12:432–444.
26 Kikuyama M, Takeshima H, Kinoshita T, Okochi-Takada E, Wakabayashi M, Akashi-Tanaka S, Ogawa T, Seto Y, Ushijima T: Development of a novel approach, the epigenome-based outlier approach, to identify tumor-suppressor genes silenced by aberrant DNA methylation. Cancer Lett 2012;322:204–212.
27 Yue W, Sun Q, Dacic S, Landreneau RJ, Siegfried JM, Yu J, Zhang L: Downregulation of Dkk3 activates β-catenin/Tcf-4 signaling in lung cancer. Carcinogenesis 2008;29:84–92.
28 Ruan W, Xu E, Xu F, Ma Y, Deng H, Huang Q, Lv B, Hu H, Lin J, Cui J, Di M, Dong J, Lai M: *IGFBP7* plays a potential tumor suppressor role in colorectal carcinogenesis. Cancer Biol Ther 2007;6:354–359.

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

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