Carcinogenesis vol.34 no.1 pp.121–127, 2013
doi:10.1093/carcin/bgs327
Advance Access publication October 11, 2012

Lipid peroxidation-induced DNA adducts in human gastric mucosa

Tomonari Matsuda[1], Hong Tao[2], Masanori Goto[2], Hidetaka Yamada[2], Masaya Suzuki[2], Yijia Wu[3], Nong Xiao[4], Qiong He[5], Wenwen Guo[6], Zhenming Cai[6], Nobuya Kurabe[2], Keiko Ishino[2], Yoshitaka Matsushima[7], Kazuya Shinmura[2], Hiroyuki Konno[8], Masato Maekawa[9], Yaping Wang[6] and Haruhiko Sugimura[2,*]

[1]Research Center for Environmental Quality Management, Kyoto University, Otsu, Shiga 520–0811, Japan, [2]Department of Tumor Pathology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi-ku, Hamamatsu 431–3192, Japan, [3]Department of Gastroenterology, Lujiang People Hospital, 32 Wenmingzhong Road, Lujiang, Anhui 231501, China, [4]Department of Internal Medicine, Lujiang People Hospital, 32 Wenmingzhong Road, Lujiang, Anhui 231501, China, [5]Department of Pathology, Lujiang People Hospital, 32 Wenmingzhong Road, Lujiang, Anhui 231501, China, [6]Jiangsu Key Laboratory of Molecular Medicine, Nanjing University School of Medicine, 22 Hankou Road, Nanjing 210093, China, [7]Department of Human Science, Chemistry Division, Hamamatsu University School of Medicine, Hamamatsu 431–3192, Japan, [8]Department of Surgery II, Hamamatsu University School of Medicine, Hamamatsu 431–3192, Japan and [9]Department of Laboratory Medicine, Hamamatsu University School of Medicine, Hamamatsu 431–3192, Japan

*To whom correspondence should be addressed. Tel: +81-53-435-2220; Fax: +81-53-435-2225;
Email: hsugimur@hama-med.ac.jp

DNA adducts are a major cause of DNA mutation and DNA mutation-related diseases, but the simultaneous identification of multiple DNA adducts has been a challenge for a decade. An adductome approach using consecutive liquid chromatography and double mass spectrometry after micrococcal nuclease treatment has paved the way to demonstrations of numerous DNA adducts in a single experiment and is expected to contribute to the comprehensive understanding of overall environmental and endogenous exposures to possible mutagens in individuals. In this report, we applied an adductome approach to gastric mucosa samples taken at the time of a gastrectomy for gastric cancer in Lujiang, China, and in Hamamatsu, Japan. Seven lipid peroxidation-related DNA adducts [1,*N*6-etheno-2′-deoxyadenosine, butanone-etheno-2′-deoxycytidine (BεdC), butanone-etheno-2′-deoxy-5-methylcytidine, butanone-etheno-2′-deoxyadenosine (BεdA), heptanone-etheno-2′-deoxycytidine, heptanone-etheno-2′-deoxyadenosine (HεdA) and heptanone-etheno- 2′-deoxyguanosine] were identified in a total of 22 gastric mucosa samples. The levels of these adducts ranged from 0 to 30 000 per 10^9 bases. Although the presence of *Helicobacter pylori* DNA in the mucosa was not related to these adducts level, the levels of BεdC, BεdA and HεdA were higher in the Japanese gastric mucosa samples. The profiles of these 7 adduct levels among the 21 cases were capable of discriminating between the possible origins (China or Japan) of the gastric mucosa samples. Our report is the first demonstration of lipid peroxidation-related DNA adducts in the human stomach, and these observations warrant further investigation in the context of the significance of DNA adducts in human gastric carcinogenesis.

Introduction

DNA adducts are a major cause of mutation (1); thus, the recognition of adducts in target organs in individuals provides very basic information on the susceptibility to and the initiation of cancer in particular organs and particular individuals (2).

Abbreviations: BεdA, butanone-etheno-2′-deoxyadenosine; BεdC, butanone-etheno-2′-deoxycytidine; BεmedC, butanone-etheno-2′-deoxy-5-methylcytidine; εdA, 1,N6-etheno-2′-deoxyadenosine; HεdA, heptanone-etheno-2′-deoxyadenosine; HεdC, heptanone-etheno-2′-deoxycytidine.

The identification of DNA adducts has been a technical challenge for several decades (3), but recent innovations have enabled us to identify multiple DNA adducts in human tissues (4,5). This adductome approach has successfully revealed a considerable number of lipid peroxidation-related adducts in human tissues, including human lungs (6). The human stomach is another organ that is almost directly exposed to the environment and especially to dietary substances with the potential to damage the DNA of the host on a daily basis. Because dietary carcinogens and mutagens have been identified in ordinary food and/or food additives (7), DNA adducts related to such environmental carcinogens are probably to be present in human gastrointestinal tracts, especially in patients who suffer from cancers of the gastrointestinal tract. In addition, the gastric mucosa is a site where chronic inflammation may occur, especially in individuals infected with *Helicobacter pylori (H. pylori)* (8–10); thus, inflammation-related DNA adducts may exist in the gastric mucosa, possibly playing an important role in gastric carcinogenesis (11). Previously, several attempts have been made to detect a variety of DNA adducts in the stomach of experimental animals (12,13) and human beings (14). Among them, Abdul-Momen and colleagues detected five DNA adduct spots using the P^{32}-postlabeling method (15). They claimed that these five DNA adducts were specific to the gastric mucosa in gastric cancer patients and were absent in newborn stomach. This observation, however, has not been followed-up to date, probably because the P^{32}-postlabeling method cannot further delineate the molecular species in DNA from the gastric mucosa. Recently, our group established a method consisting of liquid chromatography followed by double tandem mass spectrometry to identify numerous adducts in human and animal tissues (4–6,16) during a single experiment. In contrast with the P^{32}-postlabeling method, this method directly facilitates the identification of modified DNA bases by placing the standard substances in parallel. Several adducts with different chemical formulae and mass-per-charge values have been confirmed to exist in human tissues (4,6). Recently, Marsden and colleagues (17) used liquid chromatography and accelerator mass spectrometry and showed a dose-response relationship with the quantity of a specific DNA adduct (N7-(2-hydroxyethyl)guanine, a presumed cancer-causing adduct) in an experimental model of stomach cancer; they stated that their method and calibration would be useful for the study of human tissues and epidemiology frameworks. However, they have not yet reported any information specific to human gastrointestinal tissue.

In this report, we applied the adductome approach to human stomach samples obtained during gastrectomies for the resection of gastric cancer and validly identified seven different lipid peroxidation-derived DNA adducts in the human stomach for the first time.

Materials and methods

DNA extraction from gastric mucosa

Gastric mucosa was taken from non-tumor tissue sections of stomachs resected for the treatment of gastric cancer at Lujiang Hospital, China, and at Hamamatsu University Hospital, Japan. For sample quality control, pathologists from each institution made a site-visit and confirmed that non-tumor, mucosal portions were immediately obtained after resection and were frozen in liquid nitrogen and kept at −70°C in a freezer until DNA extraction. The clinicopathological profiles of the cases are shown in Table I. For all the samples, DNA extraction was conducted according to the same protocol and by the same person (N.K.) without knowledge of the origin (Lujiang or Hamamatsu) of the tissue using a Gentra Puregene™ Tissue Kit (Qiagen, Valencia, CA). DNA extraction was undertaken according to the protocol provided by the manufacturer, with the addition of desferroxamine to all solutions to a final concentration of 0.1 mM to protect against the oxidation of the DNA during the procedure.

Sample preparation for pooled-DNA adductome analysis

To grasp the overall picture of DNA adducts in the gastric mucosa of Japanese and Chinese patients, we performed a pooled-DNA adductome analysis. DNA

Table I. Clinicopathological profiles of the 22 cases from Lujiang (C) and Hamamatsu (J)

	Origin	Age	Sex (M: male; F: female)	Cancer location[a]	Histological type[b]
1	C	45	F	U	U
2	C	54	M	L	D
3	C	60	M	M	D
4	C	52	M	U	U
5	C	58	M	U	U
6	C	60	F	M	D
7	C	55	F	L	D
8	C	57	M	U	U
9	C	58	M	L	D
10	C	61	M	U	U
11	C	60	M	U	U
12	C	58	F	U	U
13	J	57	F	M	D
14	J	64	M	U	U
15	J	60	M	L	D
16	J	52	F	U	U
17	J	53	M	U	U
18	J	64	M	L	D
19	J	57	M	M	D
20	J	60	M	U	U
21	J	64	M	M	D
22	J	60	M	M	D

[a]U, M and L represent the upper, middle and lower thirds of the stomach.
[b]D, differentiated type (intestinal type); U, undifferentiated type (diffuse type).

samples from each patient (approximately 100 μg) were mixed with 54 μl of digestion buffer (17 mM sodium succinate and 8 mM calcium chloride, pH 6.0) containing 67.5 units of micrococcal nuclease (Worthington, Lakewood, NJ) and 0.255 units of spleen phosphodiesterase (Worthington). After 3 h of incubation at 37°C, three units of alkaline phosphatase (Sigma–Aldrich, St Louis, MO), 30 μl of 0.5 M Tris–HCl (pH 8.5), 15 μl of 20 mM zinc sulfate and 101 μl of milliQ water were added; the mixture was then incubated for another 3 h at 37°C. After incubation, an 87 μl aliquot from 10 Japanese samples (Patient Nos. 13–22) was taken and pooled as the Japanese group, and an 87 μl aliquot from 10 Chinese samples (Patient Nos. 2–11) was also pooled as the Chinese group. The pooled mixtures were concentrated to approximately 100 μl using a Speed-Vac concentrator, and 500 μl of methanol was added to precipitate the protein. After centrifugation, the methanol fraction (supernatant) was transferred to a new Eppendorf tube and evaporated to dryness, then redissolved in 320 μl of 30% dimethyl sulfoxide (5).

Sample preparation for quantification of lipid peroxidation-derived DNA adducts

Seven kinds of lipid peroxidation-derived DNA adducts [1,N^6-etheno-2′-deoxyadenosine (εdA), butanone-etheno-2′-deoxycytidine (BεdC), butanone-etheno-2′-deoxy-5-methylcytidine (BεmedC), butanone-etheno-2′-deoxyadenosine (BεdA), heptanone-etheno-2′-deoxycytidine (HεdC), heptanone-etheno-2′-deoxyguanosine (HεdG) and heptanone-etheno-2′-deoxyadenosine (HεdA)] were quantified basically as described by Chou and colleagues (6). A 58 μl aliquot of DNA digest was spiked with 2.2 μl of 4 nM stable-isotope internal standard mix ([$^{15}N_5$]-derivatives of εdA, BεdC, BεdA, HεdC, HεdG and HεdA), and the volume was reduced using a Speed-Vac concentrator; then, 100 μl of methanol was added to precipitate the protein. The methanol fraction (supernatant) was transferred to a new Eppendorf tube and evaporated to dryness, and then redissolved in 22 μl of 30% dimethyl sulfoxide.

DNA adductome analysis

For the DNA adducts analysis, a Quattro Ultima Pt triple-stage quadrupole mass spectrometer equipped with an Alliance 2695 separation module and a 2487 Dual λ Absorbance Detector (Waters, Milford, MA) was used. The pooled-DNA adductome analysis was performed in a manner similar to that described by Kanaly and colleagues (5). An aliquot of the digested DNA sample (20 μl) was injected and separated using a Shim-pack XR-ODS column (3.0 × 75 mm; Shimadzu, Kyoto, Japan). The column was eluted in a linear gradient of 5–80% methanol in water from 0 to 20 min at a flow rate of 0.2 ml/min. Multireaction monitoring was performed in the positive-ion mode. The experimental conditions were set as follows: ion source temperature, 130°C; desolvation temperature, 380°C; cone voltage, 35 V; collision energy, 15 eV; desolvation gas flow rate, 700 l/h; cone gas flow rate, 35 l/h and collision gas, argon. This strategy was designed to detect the neutral loss of 2′-deoxyribose from positively ionized 2′-deoxynucleoside adducts by monitoring the samples with $[M + H]^+ \rightarrow [M + H\text{-}116]^+$ transitions. Each of the Japanese- and Chinese-pooled samples were injected 15 times to complete the monitoring of the 451 multireaction monitoring transitions over the *m/z* range from the transition *m/z* 250→134 to the transition 702→586. The transitions of normal deoxynucleosides, including 252→136 ($[dA + H]^+$) and 268→152 ($[dG + H]^+$), were not monitored in the adductome analysis.

DNA adduct quantification

The same liquid chromatography and double mass spectrometry system was used for DNA adduct quantification. An aliquot (20 μl) of each sample was injected and separated using the Shim-pack XR-ODS column, eluted in a linear gradient of 5–30% methanol in water from 0 to 27 min and then of 30–80% methanol from 27 to 35 min, then kept in 80% methanol from 35 to 40 min at a flow rate of 0.2 ml/min. The collision energies and characteristic reactions monitored for the different DNA adducts were as follows [cone voltage (V), collision energy (eV), base ion→product ion]: [U-$^{15}N_5$]-εdA (35, 14, 280.9→164.9), [U-$^{15}N_3$]-HεdC (35, 10, 367.0→251.0), [U-$^{15}N_5$]-HεdA (35, 10, 393.0→277.0), [U-$^{15}N_5$]-HεdG (35, 10, 409.0→293.0), [U-$^{15}N_3$]-BεdC (35, 10, 324.8→208.6), [U-$^{15}N_5$]-BεdA (35, 10, 351.0→234.8), εdA (35, 14, 275.9→159.9), HεdC (35, 10, 364.0→248.0), HεdA (35, 10, 388.0→272.0), HεdG (35, 10, 404.0→288.0), BεdC (35, 10, 321.8→205.6), BεmedC (35, 20, 335.9→220.0) and BεdA (35, 10, 351.0→234.8). The amount of each DNA adduct was quantified by calculating the peak area ratio of the target DNA adduct and its specific internal standard ([U-$^{15}N_3$]-BεdC was used for BεdC and BεmedC). Calibration curves were obtained using authentic standards spiked with isotope internal standards.

Histological analysis

The tissues next to the sampled portion were used for histological evaluation using hematoxylin and eosin staining following the usual formalin-fixation and paraffin-embedding steps.

Detection of H.pylori DNA in the DNA sample

DNA from each sample was tested for *H.pylori* DNA according to a previously reported method, in principle (18,19). An automatic gene analyzing system, GENECUBE™ (Toyobo, Osaka, Japan), was used to detect S (sensitive to Clarithromycin) and R (resistant to Clarithromycin) DNA fragments encoding *H.pylori* 23S ribosomal RNA with an internal control, *CYP2C19*, for the human genome (20). All the cases had a peak at *CYP2C19* for either genotype. Cases without an R or S peak were considered to be *H.pylori* negative.

Statistical analysis

For each adduct level, a Wilcoxon rank sum test was performed to compare the DNA adduct numbers/10^9 bases. We noticed one outlier case in our data set, and a cluster analysis validated the independency of this single case. After an exclusion of the outlier case, a discriminant function analysis was performed using seven adduct levels to categorize the cases based on the origin of tissue (Lujiang or Hamamatsu), the location of the cancer (upper, middle or lower part of the stomach), the histology of the cancer (differentiated or undifferentiated) and the sex (male or female). All the calculations were performed using the statistics package software JMP7™ (SAS Institute Japan Ltd, Tokyo, Japan).

Institutional review boards

The study protocol was approved by the institutional review boards of Hamamatsu University School of Medicine (23–91), Lujiang People's Hospital and Nanjing University.

Results

DNA adductome

A DNA adductome map of Japanese and Chinese gastric mucosa DNA is shown in Figure 1. This adductome map corresponds to a mixture of 10 Japanese samples (blue) and a mixture of 10 Chinese samples (red), so that this map reflects the average picture of the gastric mucosa in both countries. In this adductome map, peaks that might be derived from normal deoxynucleosides were omitted. We identified 141 peaks in the Japanese sample and 159 peaks in the Chinese sample. Ninety-two of these peaks were present in both the Japanese and Chinese samples. Although most of the peaks (circles) shown on this map have not been identified, we found peaks corresponding to εdA (m/z: 276, RT 12.5 min) and HεdC (m/z: 364, RT 22.8 min), which are lipid peroxidation-derived DNA adducts. These observations motivated

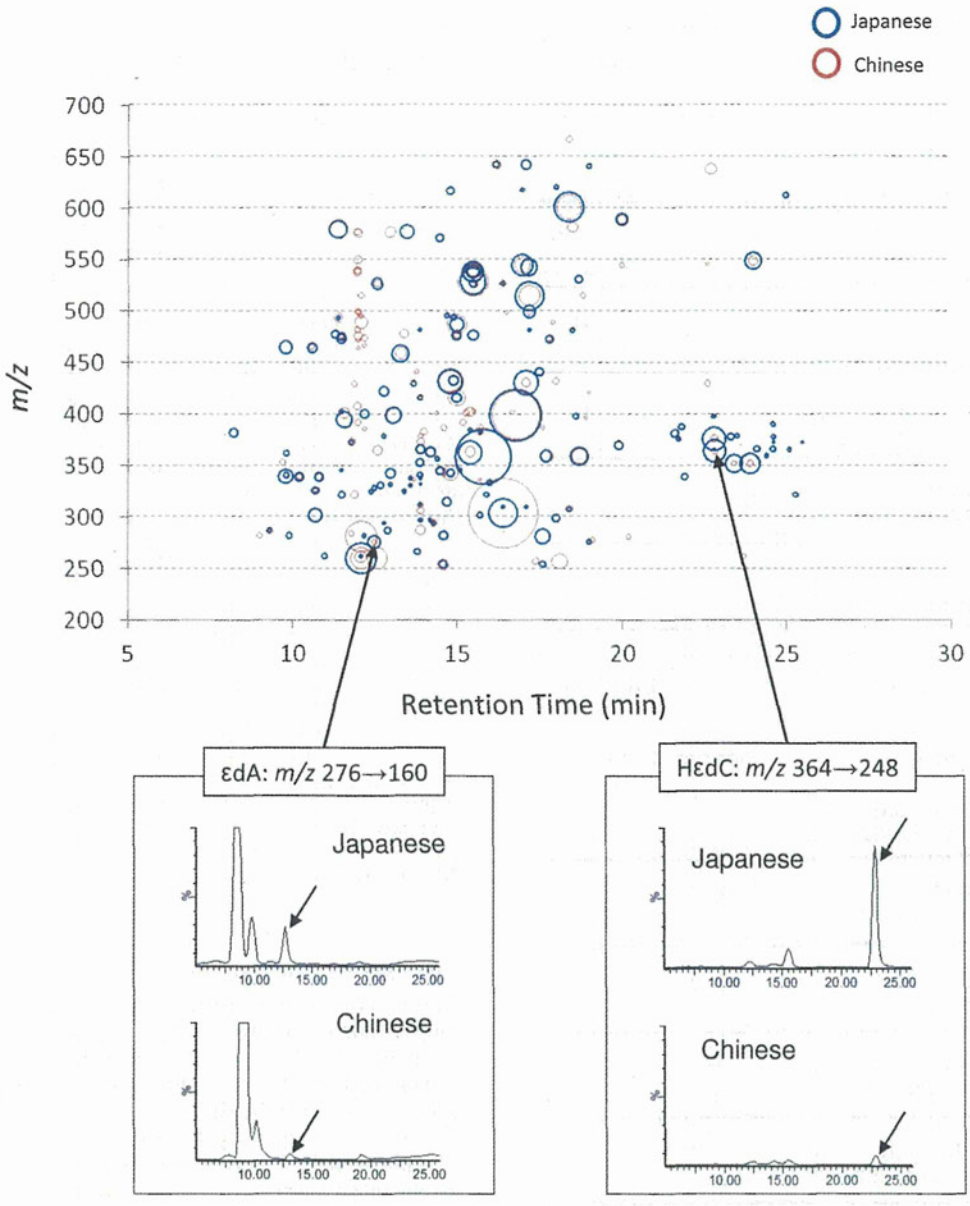

Fig. 1. Pooled-DNA adductome map of gastric mucosa samples from 10 Japanese and 10 Chinese patients. In this plot, the bubble size represents the relative peak area of MS/MS chromatogram. The liquid chromatography and double mass spectrometry peaks corresponding to the lipid peroxidation-induced DNA adducts, εdA and HεdC, are shown.

us to quantify specific lipid peroxidation-derived DNA adducts in individual samples.

Quantification of lipid peroxidation-derived DNA adducts

The representative chromatogram from one individual is shown in Figure 2. Chromatogram of 1 nM standards (left panel) and a stomach sample (right panel) were shown. The numbers of seven kinds of lipid peroxidation-derived DNA adducts per 10^9 bases in 22 cases of gastric mucosa are shown in Table II. The approximate detection limits for each DNA adduct were 1.7 per 10^9 bases (εdA, BεdC, BεmedC and BεdA), 3.3 per 10^9 bases (HεdC) and 16.5 per 10^9 bases (HεdA and HεdG). Among the seven adducts, εdA and HεdC were detected in all 22 cases, and their levels were extremely high, compared with previous reports describing their presence in other organs (6). Actually, the median value for εdA in the Chinese and Japanese samples was 3–5 adducts per 10^7 bases, and the median value for HεdC was 5–13 adducts per 10^7 bases. These values (3–13 adducts per 10^7 bases) are as high as the level of 8-hydroxydeoxyguanosine, one of the most prevalent DNA adducts currently known (6,21). Because minimal comparable data for the levels of these adducts in human gastric tissues is available in previous reports, we can only report the presently observed values. In addition, some DNA adducts other than εdA and HεdC were also detected at a high frequency, and their levels were also high compared with their levels in previous reports describing their presence in other organs (6,22). The median values for the DNA adducts were higher in the Japanese samples than in the Chinese samples. Three adducts, BεdC, BεdA and HεdA, were significantly higher in the Japanese gastric mucosa (Wilcoxon rank sum test) (Figure 3).

Cluster analysis and discriminant function analysis

A cluster analysis using the Ward method was undertaken (23). The dendrogram (Figure 4) identified one outlier case (case number 14) as a single isolated group (green). Case number 14 had extremely high values for almost all the adducts that were investigated.

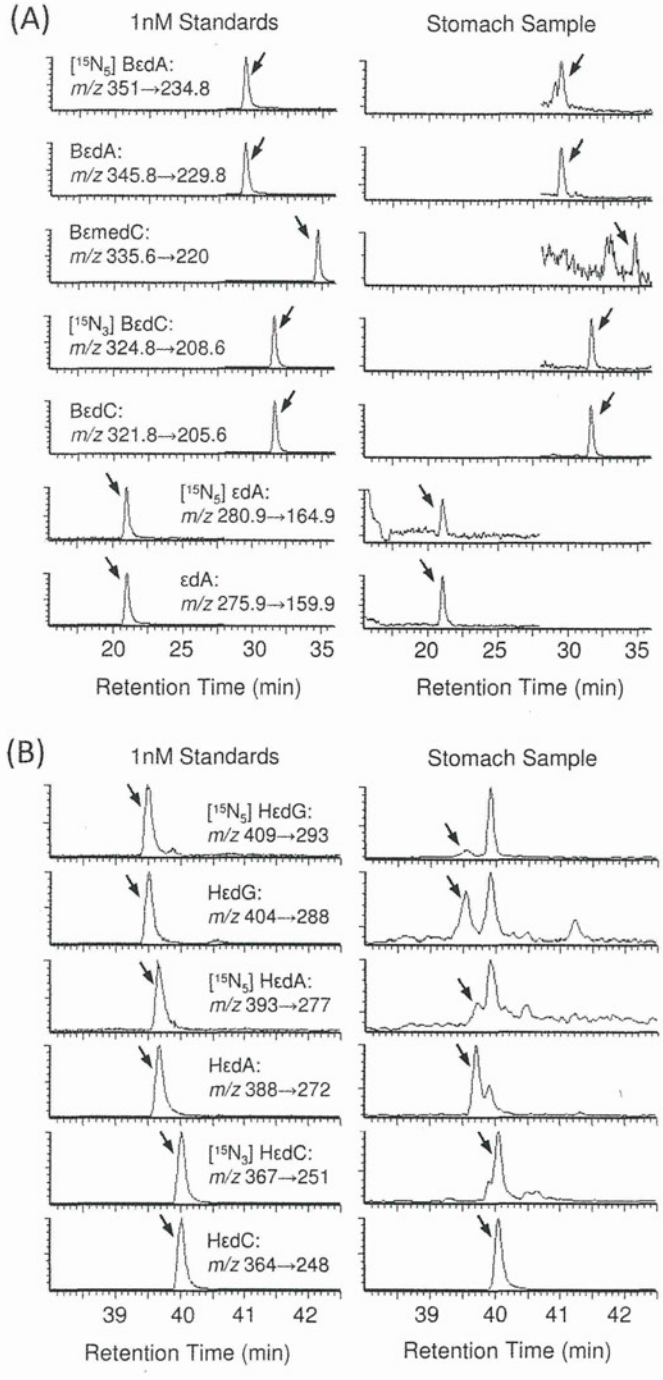

Fig. 2. Representative liquid chromatography and double mass spectrometry chromatogram of the lipid peroxidation-induced DNA adducts. Chromatogram of 1 nM standards (left panel) and a stomach sample (right panel) were shown. This sample is from the case number 21 of Table II. (**A**) Chromatogram of the εdA, BεdC, BεmedC and BεdA with their stable isotopes. (**B**) Chromatogram of the HεdC, HεdA and HεdG with their stable isotopes.

Because we were interested in whether the profiles of these seven adduct levels could identify the origin of the mucosa, we performed a discriminant function analysis. Excluding the outlier case, the discriminant function analysis generated a discriminant score in which the threshold of each category was set to 20%. When the origin of the tissue, Lujiang or Hamamatsu, was used as the dependent category, we were able to discriminate the origins of the tissues perfectly according to the seven adduct levels (Table III). The results of the discriminant function analysis were successfully transformed into a plot graph against two covariants (Figure 5).

Table II. DNA adduct level (per 10^9 bases) in Japanese and Chinese gastric mucosa samples

Case number	εdA	BεdC	BmεdC	BεdA	HεdC	HεdA	HεdG
Chinese samples							
1	281	447	ND	91	1619	305	118
2	211	34	ND	ND	112	ND	ND
3	797	73	ND	ND	149	ND	47
4	333	248	ND	ND	807	353	53
5	209	34	ND	ND	106	33	19
6	158	438	14	61	846	178	17
7	1792	1649	ND	319	3943	1491	209
8	1434	983	ND	220	3390	642	202
9	1105	1071	18	106	1662	331	108
10	190	ND	ND	ND	35	ND	ND
11	383	16	ND	ND	57	ND	ND
12	119	17	ND	ND	37	ND	18
Maximum	1792	1649	18	319	3943	1491	209
Median	307	160.5	ND	ND	478	105.5	33
Minimum	119	ND	ND	ND	35	ND	ND
Japanese samples							
13	262	94	ND	66	436	371	25
14	2763	7616	444	2261	29298	34456	482
15	627	989	ND	231	1630	594	50
16	505	697	23	206	1575	672	69
17	406	1204	ND	220	1103	266	ND
18	436	294	ND	82	869	482	73
19	763	468	ND	107	1022	427	33
20	571	1508	24	383	2523	974	110
21	964	2261	57	472	3035	1520	190
22	390	91	ND	25	203	106	ND
Maximum	2763	7616	444	2261	29298	34456	482
Median	538	843	ND	213	1339	538	59.5
Minimum	262	91	ND	25	203	106	ND

ND, not detectable.

Clinicopathological analysis

H.pylori DNA was detected in eight cases: four in Chinese gastric mucosa samples and four in Japanese gastric mucosa samples. On the other hand, genotyping for 2C19 was successful in all the cases; thus, our detection system for *H.pylori* DNA seemed sound. No differences in any of the adduct levels or in the profile of the seven adducts was seen between the *H.pylori* DNA-positive and -negative cases (data not shown).

Clinicopathological parameters including sex, age, location of cancer and histological type were compared with the adducts levels, but no correlations were obtained (data not shown). A determinant function analysis based on the parameters of sex, age, location of cancer and histological type did not generate a discriminant function capable of dividing these categories (data not shown).

Discussion

We identified lipid peroxidation-derived DNA adducts in the human stomach for the first time. Considerable numbers of seven different adducts were demonstrated. In our previous study, we quantified these DNA adducts in 68 Japanese autopsy tissues from various organs other than the stomach (lung, colon, pancreas, spleen, liver, kidney and small intestine (6)). The median values of the lipid peroxidation-derived DNA adducts in Japanese gastric mucosa observed in this study were one order of magnitude higher than those observed in the other organs. The active inflammation in the stomach may be one of the most plausible explanations for this phenomenon, but complete information on the other adducts must also be considered.

Differences in these adduct levels between the Chinese and Japanese gastric mucosa samples were not, at first glance, apparent, but a discriminant analysis disclosed detectable differences according to the origin of the gastric mucosa. When examined separately, three of the seven adducts (BεdC, BεdA and HεdA) were

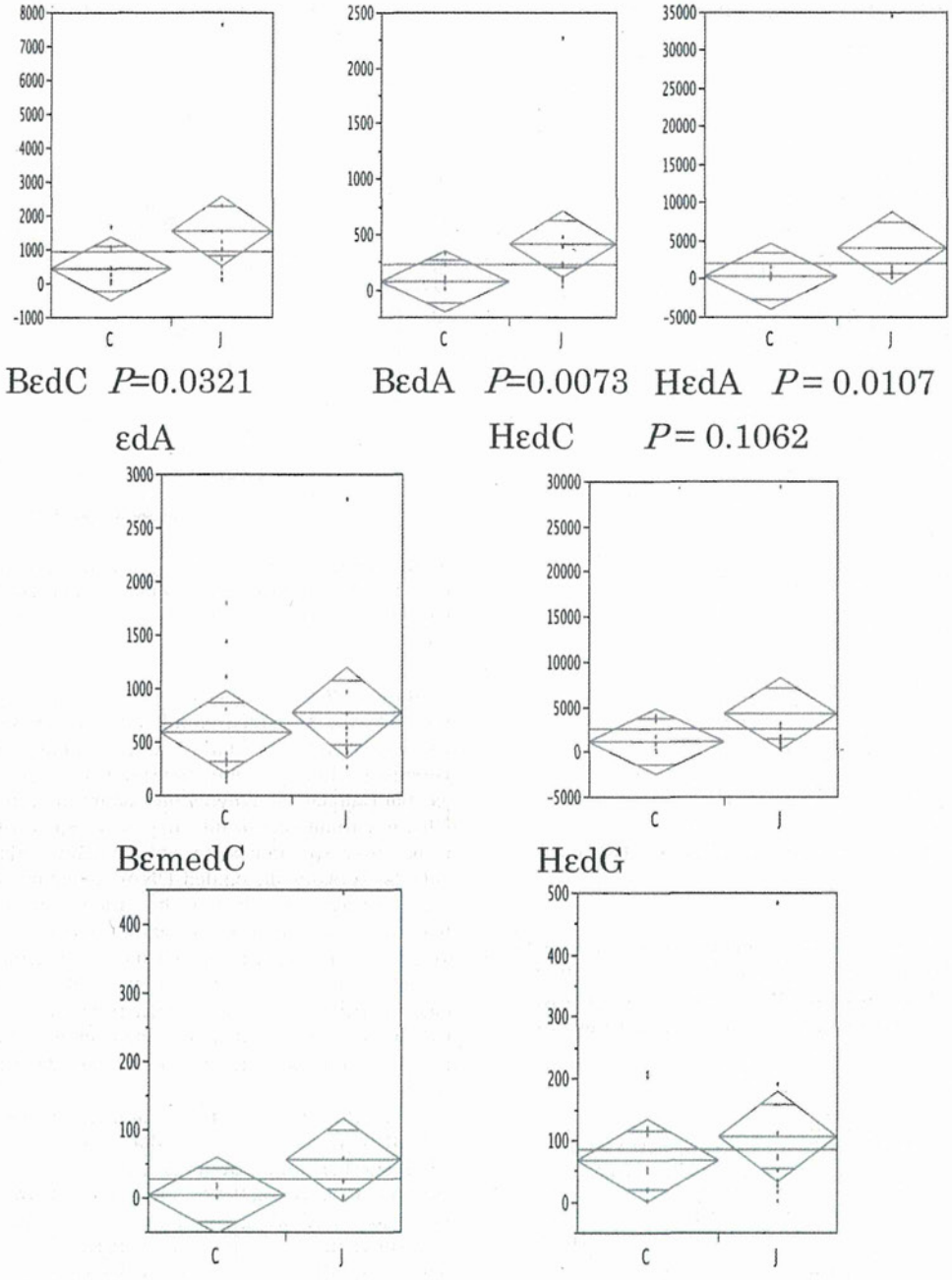

Fig. 3. Comparison of each adduct level between the Lujiang (C) and Hamamatsu (J) cases. The levels of BεdC, BεdA and HεdA were significantly higher for the Hamamatsu cases ($P < 0.05$ according to a Wilcoxon rank sum test). The horizontal lines in the diamonds indicate the upper quadrant, the mean and the lower quadrant. The axis indicates the numbers of adducts per bases.

significantly higher in the Japanese stomach specimens than in the Chinese stomach specimens. The significance of this observation remains unknown.

Lujiang county is known as an area of China where gastric cancer is endemic (24); thus, we expected higher adduct levels in this group than in the Japanese group. This analysis, however, showed that the Japanese stomach mucosa samples had higher levels of these lipid peroxidation-derived adducts. This observation may reflect the fact that the presently evaluated markers were limited to inflammation-related markers. Actually, our adductome map (Figure 1, (5,6)) for these specimens had many other un-annotated spots, some of which were more prevalent in the Chinese stomach specimens.

In our previous papers, Chou and colleagues argued that the adducts investigated in this study are formed by exposure to 4-ONE [4-oxo-2(*E*)-nonenal] and 4-OHE [4-oxo-2(*E*)-hexenal], which can be made from ω-6 and ω-3 polyunsaturated fatty acids endogenously. They cited a paper by Blair's group and hypothesized that COX2 and lipoxygenase contributed to the formation of these adducts (25). On the other hand, continued inflammation, or chronic atrophic gastritis, has long been thought to predispose an individual to gastric cancer (26,27). Hence, our evidence of the presence of oxidation-related DNA adducts may strengthen the idea of the inflammation-mediated pathogenesis of gastric cancer (11).

We showed the usefulness of the adductome profile for indicating the origin of a specimen, but whether these adducts actually indicate a gastric cancer predisposition warrants the further investigation of gastric mucosa samples from subjects without gastric cancer. Previously, inflammation-mediated carcinogenesis has been discussed in the context of dynamic changes of cellular machinery such

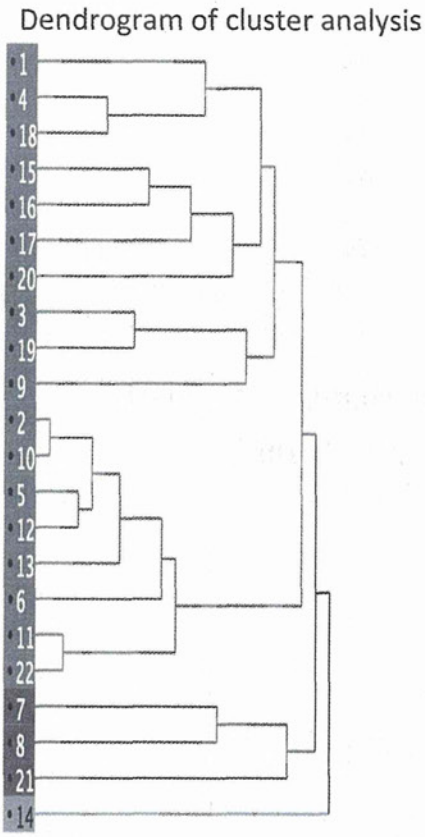

Fig. 4. Cluster analysis of the 22 cases. Case No. 14 was identified as an outlier.

as microRNA, not of formation of DNA adducts (28), but some of the adducts described in this study are known to be mutagenic (29–31); thus, these lipid peroxidation-derived DNA adducts may, at least in part, be responsible for the carcinogenic origins of the human gastric cancers in this series.

Table III. Results of discriminate analysis using seven adduct levels as the discriminating score to predict the origin of the specimen in 21 cases

Real origin	Square distance	Probability	Predicted probability	Predicted category
C	8.91653	0.9986	0.9986	C
C	1.41919	0.8686	0.8686	C
C	8.60098	0.9444	0.9444	C
C	6.09944	0.998	0.998	C
C	1.36231	0.9175	0.9175	C
C	7.82509	0.9769	0.9769	C
C	14.87972	0.6381	0.6381	C
C	12.39559	0.9982	0.9982	C
C	11.20971	0.9999	0.9999	C
C	1.71078	0.7945	0.7945	C
C	2.57192	0.7579	0.7579	C
C	2.10673	0.9117	0.9117	C
J	2.73173	0.947	0.947	J
J	2.77308	0.9958	0.9958	J
J	5.78756	0.9818	0.9818	J
J	11.33611	0.9832	0.9832	J
J	5.8705	0.5033	0.5033	J
J	3.35645	0.8798	0.8798	J
J	5.57043	0.9998	0.9998	J
J	13.27304	0.9983	0.9983	J
J	3.2031	0.575	0.575	J

C indicates Lujiang Hospital and J indicates Hamamatsu University Hospital.

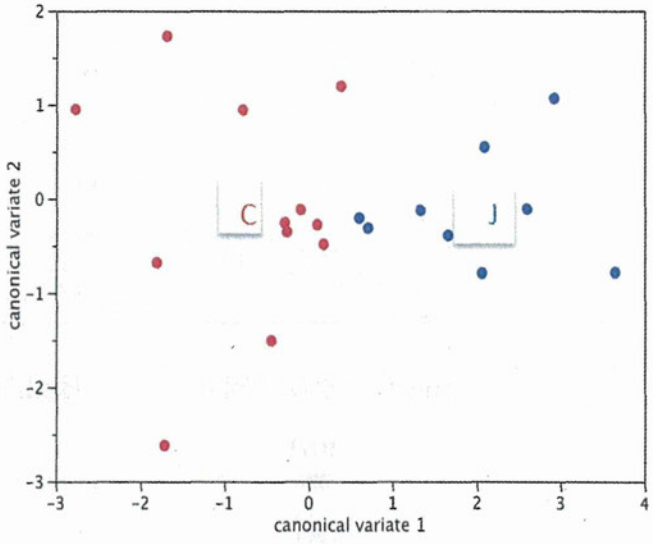

Fig. 5. A scattered diagram showing a discriminant analysis of the 21 cases (excluding the one outlier case) with seven variables. The red and blue dots represent the cases from Lujiang (C) and Hamamatsu (J), respectively.

From another perspective, our data may suggest that a pathway other than the peroxidation-inflammatory pathway may contribute to carcinogenesis. The Chinese gastric mucosa samples derived from patients in Lujiang county had lower levels of oxidative DNA damage, but Lujiang has a higher prevalence of gastric cancer than Japan. Other environmental insults may be revealed with further annotation of the observed adducts, possibly including alkylating agent-related adducts. Actually, the pooled-DNA adductome map, which was used as a screening procedure in this study, contains many other peaks that may not be lipid peroxidation related. It is assumed that there are many subjects who do not have detectable lipid peroxidation-related adducts. The continued effort of identifications of the other adducts will be necessary to comprehensive understanding of gastric carcinogenesis, and adductome approach, though at the burgeoning stage, may become one of the important omics in the field of carcinogenesis.

In conclusion, we first demonstrated the existence of lipid peroxidation-related DNA adducts in the human stomach and addressed their implications in the assessment of the environmental and endogenous exposure of human beings to these possible mutagens. In addition, considering that gastric cells have a battery of repair genes that respond to or repair DNA damage, the presently reported results may promote understanding of the role of repair genes in gastric carcinogenesis, a topic that has recently attracted enthusiastic interest in the field of carcinogenesis research (32–36).

Funding

This project was approved by the government of Jiangsu Province, China, under the project title, 'Study on the relationship between the epi-germline mutation of tumor suppressor genes and the pathogenesis of gastrointestinal cancers' (the International Science and Technology Cooperation Project of Jiangsu, grant number: BZ2008055). This work was also supported by Grants-in-Aids (Research on International Cooperation in Medical Science, Grants-in-Aids for Cancer Research, 21-1) from the Ministry of Health, Labour and Welfare; the Japan Society for the Promotion of Science (22590356 and 22790378); the Ministry of Education, Culture, Sports, Science and Technology (221S0001); the Princess Takamatsu Cancer Research Foundation and the Smoking Research Foundation of Japan.

Conflict of Interest Statement: None declared.

References

1. Weinberg,R.A. (2007) The Biology of Cancer. Garland Science, New York, NY.
2. Clayson,D.B. (1985) International Commission for Protection Against Environmental Mutagens and Carcinogens. ICPEMC Working Paper No. 2. Diet, mutation and cancer. *Mutat. Res.*, **154**, 205–217.
3. Shields,P.G. *et al.* (1991) Molecular epidemiology and the genetics of environmental cancer. *JAMA*, **266**, 681–687.
4. Kanaly,R.A. *et al.* (2006) Development of the adductome approach to detect DNA damage in humans. *Antioxid. Redox Signal.*, **8**, 993–1001.
5. Kanaly,R.A. *et al.* (2007) Application of the adductome approach to assess intertissue DNA damage variations in human lung and esophagus. *Mutat. Res.*, **625**, 83–93.
6. Chou,P.H. *et al.* (2010) Detection of lipid peroxidation-induced DNA adducts caused by 4-oxo-2(E)-nonenal and 4-oxo-2(E)-hexenal in human autopsy tissues. *Chem. Res. Toxicol.*, **23**, 1442–1448.
7. Nagao,M. (1999) A new approach to risk estimation of food-borne carcinogens–heterocyclic amines–based on molecular information. *Mutat. Res.*, **431**, 3–12.
8. De Koster,E. *et al.* (1994) *Helicobacter pylori*: the link with gastric cancer. *Eur. J. Cancer Prev.*, **3**, 247–257.
9. Farinati,F. *et al.* (1998) Oxidative DNA damage accumulation in gastric carcinogenesis. *Gut*, **42**, 351–356.
10. Kawanishi,S. *et al.* (2006) Oxidative and nitrative DNA damage as biomarker for carcinogenesis with special reference to inflammation. *Antioxid. Redox Signal.*, **8**, 1047–1058.
11. Piazuelo,M.B. *et al.* (2010) Gastric cancer: an infectious disease. *Infect. Dis. Clin. North Am.*, **24**, 853–69, vii.
12. Tan,S.L. *et al.* (2011) Is the tissue persistence of O(6)-methyl-2'-deoxyguanosine an indicator of tumour formation in the gastrointestinal tract? *Mutat. Res.*, **721**, 119–126.
13. Terasaki,M. *et al.* (2008) Detection of endogenous DNA adducts, O-carboxymethyl-2'-deoxyguanosine and 3-ethanesulfonic acid-2'-deoxycytidine, in the rat stomach after duodenal reflux. *Cancer Sci.*, **99**, 1741–1746.
14. Spigelman,A.D. *et al.* (1991) DNA adducts, detected by 32P-postlabelling, in the foregut of patients with familial adenomatous polyposis and in unaffected controls. *Carcinogenesis*, **12**, 1727–1732.
15. Abdul-Momen,M. *et al.* (2003) DNA adducts detected in human gastric mucosa. *Cancer Detect. Prev.*, **27**, 209–215.
16. Nagayoshi,H. *et al.* (2009) Increased formation of gastric N(2)-ethylidene-2'-deoxyguanosine DNA adducts in aldehyde dehydrogenase-2 knockout mice treated with ethanol. *Mutat. Res.*, **673**, 74–77.
17. Marsden,D.A. *et al.* (2009) Dose-response relationships for N7-(2-hydroxyethyl)guanine induced by low-dose [14C]ethylene oxide: evidence for a novel mechanism of endogenous adduct formation. *Cancer Res.*, **69**, 3052–3059.
18. Nakamura,A. *et al.* (2007) Determination of mutations of the 23S rRNA gene of *Helicobacter pylori* by allele specific primer-polymerase chain reaction method. *J. Gastroenterol. Hepatol.*, **22**, 1057–1063.
19. Furuta,T. *et al.* (2005) Influence of CYP2C19 polymorphism and *Helicobacter pylori* genotype determined from gastric tissue samples on response to triple therapy for *H pylori* infection. *Clin. Gastroenterol. Hepatol.*, **3**, 564–573.
20. Matsumoto,K. *et al.* (2009) Correlation between voriconazole trough plasma concentration and hepatotoxicity in patients with different CYP2C19 genotypes. *Int. J. Antimicrob. Agents*, **34**, 91–94.
21. Kim,H.N. *et al.* (2001) Changes in DNA 8-hydroxyguanine levels, 8-hydroxyguanine repair activity, and hOGG1 and hMTH1 mRNA expression in human lung alveolar epithelial cells induced by crocidolite asbestos. *Carcinogenesis*, **22**, 265–269.
22. Kawai,K. *et al.* (2010) DNA modifications by the omega-3 lipid peroxidation-derived mutagen 4-oxo-2-hexenal *in vitro* and their analysis in mouse and human DNA. *Chem. Res. Toxicol.*, **23**, 630–636.
23. Ward,J.H. (1963) Hierarchical grouping to optimize an objective function. *J. Am. Statist. Assoc.* **58**, 236–244.
24. Yang,L. (2006) Incidence and mortality of gastric cancer in China. *World J. Gastroenterol.*, **12**, 17–20.
25. Speed,N. *et al.* (2011) Cyclooxygenase- and lipoxygenase-mediated DNA damage. *Cancer Metastasis Rev.*, **30**, 437–447.
26. Correa,P. (1995) The role of antioxidants in gastric carcinogenesis. *Crit. Rev. Food Sci. Nutr.*, **35**, 59–64.
27. Correa,P. *et al.* (2006) Etiopathogenesis of gastric cancer. *Scand. J. Surg.*, **95**, 218–224.
28. Schetter,A.J. *et al.* (2010) Inflammation and cancer: interweaving microRNA, free radical, cytokine and p53 pathways. *Carcinogenesis*, **31**, 37–49.
29. Pollack,M. *et al.* (2006) Translesion DNA Synthesis across the heptanone–etheno-2'-deoxycytidine adduct in cells. *Chem. Res. Toxicol.*, **19**, 1074–1079.
30. Levine,R.L. *et al.* (2000) Mutagenesis induced by a single 1,N6-ethenodeoxyadenosine adduct in human cells. *Cancer Res.*, **60**, 4098–4104.
31. Moriya,M. *et al.* (1994) Mutagenic potency of exocyclic DNA adducts: marked differences between *Escherichia coli* and simian kidney cells. *Proc. Natl. Acad. Sci. U.S.A.*, **91**, 11899–11903.
32. Li,W.Q. *et al.* (2009) Association between genetic polymorphisms of DNA base excision repair genes and evolution of precancerous gastric lesions in a Chinese population. *Carcinogenesis*, **30**, 500–505.
33. Goto,M. *et al.* (2009) Altered expression of the human base excision repair gene NTH1 in gastric cancer. *Carcinogenesis*, **30**, 1345–1352.
34. Goto,M. *et al.* (2010) Three novel NEIL1 promoter polymorphisms in gastric cancer patients. *World J. Gastrointest. Oncol.*, **2**, 117–120.
35. Shinmura,K. *et al.* (2004) Inactivating mutations of the human base excision repair gene NEIL1 in gastric cancer. *Carcinogenesis*, **25**, 2311–2317.
36. Shinmura,K. *et al.* (2011) Reduced expression of MUTYH with suppressive activity against mutations caused by 8-hydroxyguanine is a novel predictor of a poor prognosis in human gastric cancer. *J. Pathol.*, **225**, 414–423.

Received August 2, 2012; revised September 4, 2012; accepted October 7, 2012

nature genetics

A genome-wide association study identifies two new susceptibility loci for lung adenocarcinoma in the Japanese population

Kouya Shiraishi[1], Hideo Kunitoh[2,13], Yataro Daigo[3,4], Atsushi Takahashi[5], Koichi Goto[6], Hiromi Sakamoto[7], Sumiko Ohnami[7], Yoko Shimada[1], Kyota Ashikawa[8], Akira Saito[9], Shun-ichi Watanabe[10], Koji Tsuta[11], Naoyuki Kamatani[5], Teruhiko Yoshida[7], Yusuke Nakamura[4], Jun Yokota[12], Michiaki Kubo[8] & Takashi Kohno[1]

Lung adenocarcinoma is the most common histological type of lung cancer, and its incidence is increasing worldwide. To identify genetic factors influencing risk of lung adenocarcinoma, we conducted a genome-wide association study and two validation studies in the Japanese population comprising a total of 6,029 individuals with lung adenocarcinoma (cases) and 13,535 controls. We confirmed two previously reported risk loci, 5p15.33 (rs2853677, $P_{combined} = 2.8 \times 10^{-40}$, odds ratio (OR) = 1.41) and 3q28 (rs10937405, $P_{combined} = 6.9 \times 10^{-17}$, OR = 1.25), and identified two new susceptibility loci, 17q24.3 (rs7216064, $P_{combined} = 7.4 \times 10^{-11}$, OR = 1.20) and 6p21.3 (rs3817963, $P_{combined} = 2.7 \times 10^{-10}$, OR = 1.18). These data provide further evidence supporting a role for genetic susceptibility in the development of lung adenocarcinoma.

Lung cancer is the leading cause of cancer-related death in most countries[1]. Lung cancer consists of three major histological types: adenocarcinoma, squamous-cell carcinoma and small-cell carcinoma[1–3]. Adenocarcinoma is the most common type, comprising ~40% of all cases of lung cancer, and its incidence is increasing in both Asian and Western countries. The development of lung adenocarcinoma is more weakly associated with smoking than are the developments of squamous and small-cell carcinomas, indicating that the mechanisms of carcinogenesis differ among these histological types. A better understanding of the genetic factors underlying the development of lung adenocarcinoma is strongly needed to elucidate the etiology of disease and identify high-risk individuals for targeted screening and/or prevention. In particular, the proportion of females and never smokers among patients with lung adenocarcinoma is considerably higher in Asians than in Europeans[2,3], suggesting that genetic factors contribute differently to disease in the two populations.

Genome-wide association studies (GWAS) of lung cancer with a full range of histological types have been conducted in European populations, and associations at 15q25.1, 5p15.33 and 6p21.33 have been identified[4–8]. Variants at these regions have been defined in European populations by a meta-analysis of GWAS according to histological types, and rs2736100 in *TERT* at 5p15.33 was found to be associated with risk of lung adenocarcinoma[9]. However, no additional loci reached genome-wide significance in the study; therefore, GWAS focusing on lung adenocarcinoma were greatly needed[9]. A recent GWAS on lung adenocarcinoma risk in the Japanese and Korean populations identified a new locus, 3q28 (*TP63*)[10]. Subsequently, a significant but weaker association of 3q28 variations with lung adenocarcinoma risk was validated in Europeans[11]. Notably, the association of this locus with cancer risk was supported by a recent GWAS on lung cancer with a full range of different histological types in the Chinese population[12]. These results indicate that there may be differences in the magnitude of the contribution of these loci to lung cancer susceptibility by ethnicity. Here, to further elucidate the genetic factors contributing to the development of lung adenocarcinoma, we performed a GWAS focusing on lung adenocarcinoma in the Japanese population and expanded the scale of our previous study in terms of both sample size and SNP coverage[10].

Using Illumina Omni1-Quad and OmniExpress chips, we genotyped 1,722 cases and 5,846 controls for 709,857 SNPs (**Supplementary Table 1**). Based on the results of a stringent quality-control analysis, we chose 538,166 autosomal SNPs, 1,695 cases and 5,333 control subjects for our GWAS analyses (Online Methods and

[1]Division of Genome Biology, National Cancer Center Research Institute, Tokyo, Japan. [2]Division of Thoracic Oncology, National Cancer Center Hospital, Tokyo, Japan. [3]Department of Medical Oncology, Shiga University of Medical Science, Shiga, Japan. [4]Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo, Tokyo, Japan. [5]Laboratory for Statistical Analysis, Center for Genomic Medicine, RIKEN, Kanagawa, Japan. [6]Thoracic Oncology Division, National Cancer Center Hospital East, Chiba, Japan. [7]Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan. [8]Laboratory for Genotyping Development, Center for Genomic Medicine, RIKEN, Kanagawa, Japan. [9]Statistical Genetics Analysis Division, StaGen Co. Ltd, Tokyo, Japan. [10]Division of Thoracic Surgery, National Cancer Center Hospital, Tokyo, Japan. [11]Division of Pathology and Clinical Laboratories, National Cancer Center Hospital, Tokyo, Japan. [12]Division of Multistep Carcinogenesis, National Cancer Center Research Institute, Tokyo, Japan. [13]Present address: Department of Respiratory Medicine, Mitsui Memorial Hospital, Tokyo, Japan. Correspondence should be addressed to T.K. (tkkohno@ncc.go.jp).

Received 21 February; accepted 18 June; published online 15 July 2012; doi:10.1038/ng.2353

Table 1 Summary of the GWAS and validation studies and the combined analyses

dbSNP locus	Gene	Allele [risk allele]	Stage	Cases Total	Cases RAF	Controls Total	Controls RAF	P^a	OR (95% CI)	P_{het}
rs2853677	*TERT*	T/C	GWAS	1,695	0.384	5,333	0.308	8.66×10^{-17}	1.41 (1.30–1.53)	
5p15.33	intron 2	[C]	First validation	2,955	0.374	7,036	0.297	8.62×10^{-21}	1.43 (1.32–1.54)	
			Second validation	1,373	0.360	1,132	0.290	5.88×10^{-6}	1.35 (1.19–1.54)	
			Combined validation[b]	4,328	0.370	8,168	0.296	3.90×10^{-25}	1.42 (1.32–1.50)	0.49
			Combined all[b]	6,023	0.374	13,501	0.300	2.80×10^{-40}	1.41 (1.32–1.50)	0.79
rs2736100	*TERT*	T/G	GWAS	1,695	0.458	5,329	0.391	7.31×10^{-12}	1.32 (1.22–1.42)	
5p15.33	intron 2	[G]	First validation	2,954	0.458	7,036	0.385	2.13×10^{-19}	1.39 (1.29–1.49)	
			Second validation	1,343	0.432	1,166	0.368	1.79×10^{-4}	1.27 (1.12–1.44)	
			Combined validation[b]	4,297	0.450	8,202	0.383	3.97×10^{-22}	1.36 (1.28–1.44)	0.22
			Combined all[b]	5,992	0.452	13,531	0.386	2.50×10^{-32}	1.34 (1.28–1.41)	0.39
rs10937405	*TP63*	C/T	GWAS	1,695	0.728	5,333	0.677	1.10×10^{-8}	1.29 (1.18–1.40)	
3q28	intron 1	[C]	First validation	2,953	0.714	7,036	0.663	9.22×10^{-10}	1.27 (1.18–1.37)	
			Second validation	1,375	0.704	1,166	0.682	1.22×10^{-1}	1.11 (0.97–1.26)	
			Combined validation[b]	4,328	0.711	8,202	0.666	8.17×10^{-10}	1.23 (1.15–1.31)	0.076
			Combined all[b]	6,023	0.715	13,535	0.670	6.92×10^{-17}	1.25 (1.19–1.32)	0.15
rs7216064	*BPTF*	A/G	GWAS	1,695	0.747	5,333	0.706	1.07×10^{-5}	1.22 (1.12–1.34)	
17q24.3	intron 9	[A]	First validation	2,955	0.736	7,036	0.708	7.72×10^{-5}	1.17 (1.08–1.27)	
			Second validation	1,376	0.744	1,166	0.708	4.70×10^{-3}	1.21 (1.06–1.39)	
			Combined validation[b]	4,331	0.739	8,202	0.708	1.34×10^{-6}	1.18 (1.10–1.26)	0.65
			Combined all[b]	6,026	0.741	13,535	0.707	7.40×10^{-11}	1.20 (1.13–1.26)	0.76
rs3817963	*BTNL2*	A/G	GWAS	1,695	0.363	5,331	0.327	5.54×10^{-5}	1.18 (1.09–1.28)	
6p21.3	intron 4	[G]	First validation	2,951	0.347	7,028	0.310	1.59×10^{-5}	1.18 (1.09–1.27)	
			Second validation	1,376	0.358	1,166	0.329	2.41×10^{-2}	1.16 (1.02–1.32)	
			Combined validation[b]	4,327	0.350	8,194	0.313	1.14×10^{-6}	1.17 (1.10–1.25)	0.86
			Combined all[b]	6,022	0.354	13,525	0.318	2.69×10^{-10}	1.18 (1.12–1.24)	0.97

RAF, risk allele frequency; P_{het}, *P* value for heterogeneity.
[a]Adjusted for age and gender. [b]The combined meta-analysis was performed using a fixed effect model.

Supplementary Fig. 1). We generated a quantile-quantile plot using the results of a logistic regression trend test (**Supplementary Fig. 1d**). The genomic inflation factor ($\lambda_{1,000}$)[13] was 1.021, indicating a low possibility of false-positive associations resulting from population stratification or genotype misclassification (**Supplementary Fig. 2**).

In the GWAS, two loci reached genome-wide significance for association ($P < 5 \times 10^{-8}$; **Supplementary Fig. 1e**); these two loci have been reported in previous GWAS (rs2736100 at 5p13.33 and rs10937405 at 3q28)[9,10]. We also identified a significant association for a SNP (rs2853677 at 5p13.33) that was not examined in our previous GWAS (**Table 1**). In addition, we examined associations of other previously reported loci with lung cancer risk (**Supplementary Table 2**). We found one locus (rs2131877 at 3q29)[14] to be associated with lung adenocarcinoma risk, but we could not confirm the associations between lung adenocarcinoma risk and the other loci identified in a recent GWAS of the European and Han Chinese populations[12]. These results are probably the result of the lower statistical power in our GWAS than in the previous GWAS (**Supplementary Table 2**). In addition, most of the earlier GWAS were performed in lung cancer representing a full range of histological types and in subjects of European descent. Therefore, differences in genetic modifiers and/or environmental factors in different histological types and populations might have contributed to the differing results.

To investigate additional susceptibility loci, we conducted a validation study using two independent sample sets consisting of 2,955 cases and 7,036 controls (first validation cohort) and 1,379 cases and 1,166 controls (second validation cohort) (**Supplementary Table 1**). Among 125 SNPs with a logistic regression trend of $P < 1 \times 10^{-4}$ in our GWAS, we selected 78 SNPs, excluding 38 SNPs within the same locus ($r^2 > 0.8$) and nine SNPs located at the previously reported loci, 5p13.33 and 3q28. We successfully genotyped all 78 SNPs in the first validation set using the multiplex PCR-based Invader assay, and 8 SNPs had ORs with a significance of $P < 0.05$ in the same direction of association (**Supplementary Table 3**). We then subjected these eight SNPs to the second validation set analysis. When we combined the results of both validation sets using a fixed effects model, two SNPs, rs7216064 at 17q24.3 and rs3817963 at 6p21.3, showed significant associations after Bonferroni correction ($P < 6.4 \times 10^{-4}$, calculated as 0.05/78) in addition to three SNPs at the two known loci described above (**Table 1**). When we combined the results of the GWAS and the validation study, both of the newly discovered loci reached genome-wide significance (rs7216064, $P = 7.4 \times 10^{-11}$, OR = 1.20; rs3817963, $P = 2.7 \times 10^{-10}$, OR = 1.18) (**Table 1**). The ORs were similar between the GWAS and the validation study, with no heterogeneity (**Table 1**). The strengths of the associations remained similar after adjustment for smoking (**Supplementary Table 4**). In a subgroup analysis (**Supplementary Table 5**), there was no clear association between the two newly discovered loci and gender or smoking behavior, and there was also no such association for the two known loci[10].

We next performed imputation analyses using the Japanese in Tokyo (JPT) and Han Chinese in Beijing (CHB) reference sets from the 1000 Genomes Project database (June 2010 release) (Online Methods), and we examined the associations between 1,665 putative SNPs and lung adenocarcinoma risk. We found a series of signals in high linkage disequilibrium (LD) with a marker SNP at 17q24.3 (rs7216064), and we observed significant associations with lung adenocarcinoma risk for 33 of the imputed SNPs (**Fig. 1a** and **Supplementary Table 6**). However, none of the SNPs in LD at 6p21.3 reached the *P* value of our marker SNP (**Fig. 1b**).

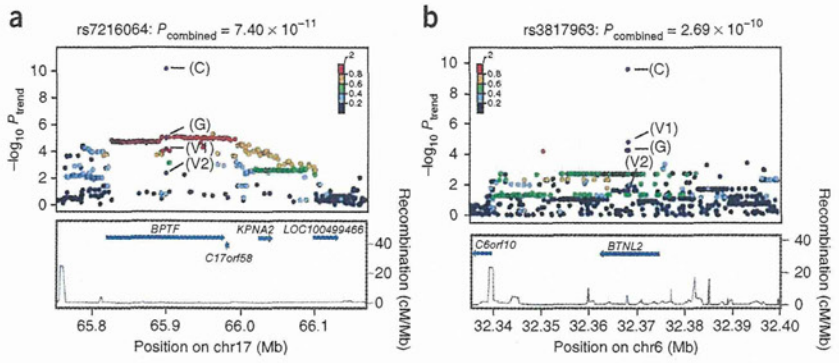

Figure 1 Regional plots of the identified marker SNPs. (**a**) rs7216064 at 17q24.3. (**b**) rs3817963 at 6p21.3. The marker SNP is shown in purple, and the r^2 values for the other SNPs are indicated by different colors. The correlations were estimated using data from the 1000 Genomes Project. The genes within the region of interest are annotated and are indicated by arrows. The blue lines indicate the recombination rates in cM per Mb. The $-\log_{10}$ P_{trend} values of the marker SNPs are shown for the GWAS (G), the first validation study (V1), the second validation study (V2) and the combined study (C).

SNP rs7216064 resides within intron 9 of *BPTF* (encoding bromodomain PHD finger transcription factor) at 17q24.3. Other imputed SNPs in this locus showing similarly significant associations were also synonymous (not resulting in amino acid changes in translated proteins). Based on the regional plot and recombination rates, we found that rs7216064 represented an LD region that includes three genes: *BPTF*, *C17orf58* (encoding a protein without known domains) and *KPNA2* (encoding karyopherin α 2) (**Fig. 1a**). Thus, to address the biological importance of 17q24.3 variants, we examined the mRNA expression levels of these three genes in 314 noncancerous lung tissues by real-time quantitative PCR (**Supplementary Note**). We detected expression of *BPTF*, but not of *C17orf58* or *KPNA2*, in these lung tissues. The expression of *BPTF* was marginally different depending on the genotype of the rs7216064 SNP ($P = 0.02$), implying low expression from the risk (G) allele (**Supplementary Table 7**). *BPTF* encodes a chromatin remodeling factor that regulates transcription through the specific recognition of methylated histone proteins[15]. Recently, chromatin remodeling genes have been implicated as tumor suppressors in lung[16] and other cancers[17]. Therefore, a low level of *BPTF* mRNA being associated with the risk allele might lead to an elevated risk for lung adenocarcinoma through decreased transcriptional regulation. However, further studies are needed to conclude whether *BPTF* is responsible for lung adenocarcinoma susceptibility.

SNP rs3817963 is located in intron 4 of *BTNL2* (encoding butyrophilin-like 2) at 6p21.3 (**Fig. 1b**). Based on the regional plot and recombination rates, rs3817963 represents an LD region that includes only a single gene, *BTNL2*. The top ten SNPs (genotyped or imputed), including rs3817963, were synonymous. The effects of the SNPs on the expression of *BTNL2* could not be assessed because of the low or absent expression of this gene in noncancerous lung tissues. *BTNL2* encodes a T cell co-stimulatory molecule, and associations between *BTNL2* SNPs and risk have been reported in several immune-related diseases, including asthma[18], vitiligo[19] and ulcerative colitis[20,21]. Therefore, *BTNL2* might affect lung adenocarcinoma risk by affecting immune responses against tumor cells. However, 6p21.3 is a part of the extended major histocompatibility complex (MHC) region, whose association with lung cancer risk has previously been reported[5]. The previously identified marker SNPs, rs3117582 and rs3131379, located 700 kb from the *BTNL2* locus, were monomorphic in our study populations. Therefore, it is possible that the association at 6p21.3 identified in the present study is not new, and further studies are warranted.

We here provide further evidence for the existence of genetic susceptibility in the development of lung adenocarcinoma through the identification of two candidate susceptibility loci, 17q24.3 and 6p21.3, at genome-wide significance. rs7216064 at 17q24.3 showed a tendency of association in the same direction as lung cancer risk in Europeans, although this association did not reach statistical significance, whereas rs3135353 at 6p21.3, which is in LD with rs3817963, showed a statistically significant association with lung cancer risk in European and American populations (**Supplementary Table 8**)[7,9]. Therefore, these loci might be involved in lung cancer risk in individuals of European descent. Further studies of these loci in multiple populations, including those with other histological types of lung cancers, will help to elucidate the etiology of lung adenocarcinoma.

URLs. The BioBank Japan project, http://biobankjp.org/; R, http://cran.r-project.org/; PLINK statistical software v1.06, http://pngu.mgh.harvard.edu/~purcell/plink/; Primer3 v0.3.0, http://frodo.wi.mit.edu/primer3/; UCSC Genome Browser, http://genome.ucsc.edu/; LocusZoom, http://csg.sph.umich.edu/locuszoom/; a catalog of genome-wide association studies, http://www.genome.gov/gwastudies/; SNPinfo Web Server, http://manticore.niehs.nih.gov/index.html; Illumina's IconDB resource, http://www.illumina.com/science/icontroldb.ilmn.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We thank all of the subjects for participating in the study, and we also thank the collaborating physicians for assisting with sample collection. We are grateful to the members of BioBank Japan, the National Cancer Center Biobank and the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for supporting our study. We thank Y. Aoi, T. Odaka, M. Okuyama, H. Totsuka, S. Chiku, A. Kuchiba and the technical staff of the Center for Genome Medicine, National Cancer Center Research Institute, for providing technical and methodological assistance. We also thank H. Hirose of Health Center, Keio University and D. Saito of National Cancer Center Hospital (present affiliation: Nihonbashi Daizo Clinic) for DNA samples of control subjects. This work was supported in part by Grants-in-Aid from the Ministry of Health, Labor and Welfare for Research on Applying Health Technology and for the 3rd-term Comprehensive 10-year Strategy for Cancer Control; from the Ministry of Education, Culture, Sports, Science and Technology of Japan for Scientific Research on Innovative Areas (22131006); from the Japan Society for the Promotion of Science for Research Activity Start-up (23800073) and for Young Scientists (B) (24790340); and by the National Cancer Center Research and Development Fund. This work was also conducted as a part of the BioBank Japan Project supported by the Ministry of Education, Culture, Sports, Science and Technology, Japan. The National Cancer Center Biobank is supported by the National Cancer Center Research and Development Fund, Japan.

AUTHOR CONTRIBUTIONS

K.S., J.Y., M.K. and T.K. designed the study. A.T., K.A., S.O., N.K. and A.S. analyzed the GWAS and replication study. H.S., Y.S., T.Y. and K.S. performed the genotyping for the GWAS and the replication study. H.K., K.G., S.W. and K.T. recruited

subjects and participated in diagnostic evaluations. K.S. and T.K. wrote the manuscript. M.K., Y.D., T.Y. and Y.N. contributed to the overall GWAS design.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2353.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Colvy, T.V. *et al.* Adenocarcinoma. in *World Health Organization Classification of Tumors: Pathology and Genetics, Tumours of Lung, Pleura, Thymus and Heart* (eds. Travis, W.D., Brambilla, E., Muller-Hermelink, H.K. & Harris, C.C.) 35–44 (IARC Press, Lyon, France, 2004).
2. Subramanian, J. & Govindan, R. Lung cancer in never smokers: a review. *J. Clin. Oncol.* **25**, 561–570 (2007).
3. Sun, S., Schiller, J.H. & Gazdar, A.F. Lung cancer in never smokers—a different disease. *Nat. Rev. Cancer* **7**, 778–790 (2007).
4. Broderick, P. *et al.* Deciphering the impact of common genetic variation on lung cancer risk: a genome-wide association study. *Cancer Res.* **69**, 6633–6641 (2009).
5. Wang, Y. *et al.* Common 5p15.33 and 6p21.33 variants influence lung cancer risk. *Nat. Genet.* **40**, 1407–1409 (2008).
6. McKay, J.D. *et al.* Lung cancer susceptibility locus at 5p15.33. *Nat. Genet.* **40**, 1404–1406 (2008).
7. Hung, R.J. *et al.* A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. *Nature* **452**, 633–637 (2008).
8. Amos, C.I. *et al.* Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. *Nat. Genet.* **40**, 616–622 (2008).
9. Landi, M.T. *et al.* A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk for adenocarcinoma. *Am. J. Hum. Genet.* **85**, 679–691 (2009).
10. Miki, D. *et al.* Variation in *TP63* is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. *Nat. Genet.* **42**, 893–896 (2010).
11. Wang, Y. *et al.* Variation in *TP63* is associated with lung adenocarcinoma in the UK population. *Cancer Epidemiol. Biomarkers Prev.* **20**, 1453–1462 (2011).
12. Hu, Z. *et al.* A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. *Nat. Genet.* **43**, 792–796 (2011).
13. Freedman, M.L. *et al.* Assessing the impact of population stratification on genetic association studies. *Nat. Genet.* **36**, 388–393 (2004).
14. Yoon, K.A. *et al.* A genome-wide association study reveals susceptibility variants for non-small cell lung cancer in the Korean population. *Hum. Mol. Genet.* **19**, 4948–4954 (2010).
15. Ruthenburg, A.J. *et al.* Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145**, 692–706 (2011).
16. Medina, P.P. & Sanchez-Cespedes, M. Involvement of the chromatin-remodeling factor BRG1/SMARCA4 in human cancer. *Epigenetics* **3**, 64–68 (2008).
17. Wilson, B.G. & Roberts, C.W. SWI/SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer* **11**, 481–492 (2011).
18. Hirota, T. *et al.* Genome-wide association study identifies three new susceptibility loci for adult asthma in the Japanese population. *Nat. Genet.* **43**, 893–896 (2011).
19. Jin, Y. *et al.* Variant of *TYR* and autoimmunity susceptibility loci in generalized vitiligo. *N. Engl. J. Med.* **362**, 1686–1697 (2010).
20. Asano, K. *et al.* A genome-wide association study identifies three new susceptibility loci for ulcerative colitis in the Japanese population. *Nat. Genet.* **41**, 1325–1329 (2009).
21. Anderson, C.A. *et al.* Meta-analysis identifies 29 additional ulcerative colitis risk loci, increasing the number of confirmed associations to 47. *Nat. Genet.* **43**, 246–252 (2011).

ONLINE METHODS

Study design and subjects. We performed a three-stage GWAS of lung adenocarcinoma in the Japanese population using independent samples. The characteristics of each case-control group are shown in **Supplementary Table 1**. The discovery GWAS samples consisted of 1,722 cases from the National Cancer Center Hospital (NCCH) and 5,846 controls from the BioBank Japan project[22], Osaka-Midosuji Rotary Club (MRC) and the Pharma SNP consortium (PSC). The BioBank Japan project (see URLs) was begun in 2003 for the collection of genomic DNA, serum and clinical information from 300,000 individuals diagnosed with any of 47 diseases by a collaboration network of 66 hospitals in all areas of Japan[22]. The subjects from MRC were 1,018 healthy volunteers, and the subjects from PSC were 906 Japanese healthy volunteers from whom immortalized B lymphoblast cell lines were established by the PSC. The cell lines were obtained from the Japan Health Sciences Foundation (JHSF)/Health Science Research Resources Bank (HSRRB). Individuals with any cancer were excluded from the control group.

The validation study consisted of two independent cohorts. The first validation cohort included 2,955 cases with lung adenocarcinoma and 7,036 controls. The cases included 1,747 subjects from the BioBank Japan project and 1,208 subjects from NCCH. All control subjects were from the BioBank Japan project. Individuals with any cancer were excluded from the control group. The second validation cohort included 1,379 cases with lung adenocarcinoma from the NCCH and 1,166 controls (cancer-free volunteers) from the NCCH and the Keio University in Tokyo[23]. Individuals with any cancer were excluded from the control group.

All of the cases with lung adenocarcinoma were diagnosed by cytological and/or histological examination according to the World Health Organization classification[24]. The cases from NCCH in the GWAS and first validation sets consisted of cases for which enough DNA for large-scale SNP analyses were available, and the cases in thee second validation set consisted of cases for which less DNA were available, for example, because only small noncancerous tissues were available for DNA extraction. Controls from the BioBank Japan project were genotyped previously. Genome-wide genotyping data obtained before and after 2010 were used as the controls for the GWAS and the first validation study described below, respectively, and thus there was no specified rationale for the selection of control subjects other than time of genotyping. Eight controls and 1,529 cases with lung adenocarcinoma from BioBank Japan and 906 control subjects from the Osaka-MRC were used in the previous GWAS[10]. All the participants provided written informed consent. This project was approved by the ethical committees of each participating institution.

Sample preparation and genotyping. Genomic DNA was extracted from peripheral blood leukocytes or noncancerous lung tissues using standard methods.

In the GWAS, we genotyped 1,722 cases with lung adenocarcinoma from the NCCH using the Illumina HumanOmni1-Quad Chip. For the controls in the GWAS, we used genome-wide data from 5,846 individuals with cerebral aneurysms, chronic obstructive pulmonary disease or glaucoma, which were genotyped using the Illumina Human OmniExpress Genotyping BeadChip.

In the study of the first validation cohort, we genotyped 2,955 cases with lung adenocarcinoma using the multiplex PCR-based Invader assay (Third Wave Technologies), as previously described[10]. The control group consisted of genome-wide data from 7,036 individuals with epilepsy, nephrosis syndrome, atopic dermatitis, urinary tract stone disease or Basedow's (Graves') disease, which were genotyped using the Illumina Human OmniExpress Genotyping BeadChip. The same quality control criteria were applied as in the GWAS (see below) to confirm that no unexpected duplicates or probable relatives were present between the control subjects of the GWAS and those of the first validation study.

In the study of the second validation cohort, we genotyped 1,379 cases with lung adenocarcinoma and 1,166 controls using the TaqMan method, according to the protocol for the ABI PRISM 7900HT Sequence Detection System (Applied Biosystems).

Quality control. Systemic quality control was performed on the raw genotyping data from 709,857 SNPs in DNA samples obtained from 7,568 subjects, consisting of 1,722 cases and 5,846 controls, using PLINK (v 1.06)[25]. Forty-four subjects were excluded because they showed gender discrepancies based on their X chromosome genotypes (7,524 subjects remained). Next, SNPs were excluded according to the following criteria: (i) 19,993 SNPs were not mapped on autosomal chromosomes; (ii) 102,929 SNPs had a minor allele frequency <0.01; and (iii) a total of 48,769 SNPs had a call rate <0.99 and genotype distributions that clearly deviated from those expected by Hardy-Weinberg equilibrium ($P < 1.0 \times 10^{-6}$). Together, 538,166 SNPs in autosomal chromosomes passed the quality-control filters and were used for the GWAS.

Next, an additional 75 unexpected duplicates or probable relatives in the GWAS were excluded based on pairwise identity by state according to their PI_HAT values in PLINK (all PI_HAT > 0.25) (7,449 subjects remained). Heterozygosity rates were calculated using PLINK, and more than 6 s.d. from the mean was used as the exclusion criterion. A principal components analysis (PCA) was performed on the genotype data from the samples, along with European (CEU), African (YRI) and east Asian (Japanese (JPT)) and Han Chinese (CHB)) individuals obtained from the phase II HapMap database using smartpca[26]. The PCA revealed no evident population substructure (**Supplementary Fig. 1a**) and identified seven outliers for exclusion (7,442 subjects remained). Most subjects fell into a known main cluster (Hondo) of the Japanese population (**Supplementary Fig. 1b**), and 414 subjects that fell far from the Hondo cluster[27] were excluded. The remaining 7,028 subjects, consisting of 1,695 cases and 5,333 control subjects, were used for the GWAS. The lack of population substructure between the cases and controls was validated by PCA of the subjects in the Hondo cluster (**Supplementary Fig. 1c**).

Statistical analyses. In the GWAS and the validation study, the statistical significance of the association with each SNP was assessed using a logistic regression trend test in the R program. Age and gender were included as covariates. Heterogeneity across the two stages was examined using the Breslow-Day test[28].

Imputation. We performed SNP imputation for each individual in the GWAS using the IMPUTE7 program[29]. The 1000 Genomes Project database (June 2010 release) was used as a reference panel. After excluding imputed SNPs with a low genotype information content (<0.5), posterior probability score (<0.90), call rate (<0.90), minor allele frequency (<0.01) or Hardy-Weinberg equilibrium ($P < 1.0 \times 10^{-7}$), imputed SNPs with $r^2 > 0.3$, residing 200–500 kb upstream or downstream of the two newly identified marker SNPs, were subjected to association analyses.

Software. For general statistical analyses, we used the R statistical environment version 2.6.1 or PLINK1.06 (ref. 25). We used LocusZoom to plot regional association plots[30].

22. Nakamura, Y. The BioBank Japan Project. *Clin. Adv. Hematol. Oncol.* **5**, 696–697 (2007).
23. Kohno, T. *et al.* Individuals susceptible to lung adenocarcinoma defined by combined *HLA-DQA1* and *TERT* genotypes. *Carcinogenesis* **31**, 834–841 (2010).
24. Travis, W.D. *et al. World Health Organization. International Histological Classification of Tumors: Histological Typing of Lung and Pleural Tumors* (eds. Travis, W.D., Colby, T.V., Corrin, B., Shimosato, Y. & Brambilla, E.) (Springer-Verlag, Heidelberg, Germany, 1999).
25. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
26. Price, A.L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
27. Yamaguchi-Kabata, Y. *et al.* Japanese population structure, based on SNP genotypes from 7003 individuals compared to other ethnic groups: effects on population-based association studies. *Am. J. Hum. Genet.* **83**, 445–456 (2008).
28. Breslow, N.E. & Day, N.E. *Statistical Methods in Cancer Research. Volume II—The Design and Analysis of Cohort Studies* 2–333 (IARC Scientific Publications, Lyon, France, 1987).
29. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
30. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).

doi:10.1038/ng.2353

Original Article
http://dx.doi.org/10.5115/acb.2012.45.3.149
pISSN 2093-3665 eISSN 2093-3673
Anatomy & Cell Biology acb

Prostate stem cell antigen gene is expressed in islets of pancreas

Hiroe Ono*, Kazuyoshi Yanagihara, Hiromi Sakamoto, Teruhiko Yoshida, Norihisa Saeki
Division of Genetics, National Cancer Center Research Institute, Tokyo, Japan

Abstract: Prostate stem cell antigen (PSCA) is a glycosylphosphatidylinositol-anchored cell surface antigen with an organ-dependent expression pattern in cancers; e.g., up-regulated in prostate cancer and down-regulated in gastric cancer. Previously it was reported that PSCA is not expressed in the normal pancreas but aberrantly expressed in pancreatic cancer. In this present study, we identified PSCA expression in islets of the pancreas by immunohistochemistry, which was co-localized with four islet-cell markers: insulin, glucagon, somatostatin and pancreatic polypeptide. In our investigation of the transcription start site of *PSCA*, we found a non-coding splicing variant of *PSCA* as well as authentic *PSCA* transcripts in mRNA samples from a normal pancreas. Both the transcripts were also identified in several pancreatic cancer cell lines. We previously reported that PSCA expression is correlated to the methylation status of the enhancer region in gastric and gallbladder cancer cell lines but not in pancreatic cancer cell lines, suggesting that PSCA expression is regulated in a different mode in pancreatic cancer from that in gastric and gallbladder cancers.

Key words: GPI-anchored protein, Islet cells, Pancreatic cancer, Splicing variant, Immunohistochemistry

Received June 19, 2012; Revised July 19, 2012; Accepted August 14, 2012

Introduction

The prostate stem cell antigen (PSCA) gene encodes a glycosylphosphatidylinositol (GPI)-anchored membrane protein with an unknown biological function [1]. As with other GPI-anchored proteins, it is thought that PSCA locates in a lipid raft on the outer surface of the cell membrane, a special microdomain enriched in glycosphingolipids, cholesterol and other lipidated proteins, and has some functional involvement in subcellular signal transduction [2]. *PSCA* was originally identified as a gene up-regulated in prostate cancer [3]. It is also up-regulated in other tumors including urinary bladder cancer, renal cell carcinoma, hydatidiform mole and ovarian mucinous tumor, where PSCA is thought to abet tumor progression [1]. In contrast, down-regulation of the gene was reported in gastric and gallbladder cancers, where it may act as a tumor suppressor [4, 5]. Moreover, there is another type of expression pattern: no expression in normal tissues but expressed in malignant counterparts, which is seen in pancreatic cancer [6], non-small cell lung cancer [7] and glioma [8]. It was reported that PSCA expression was not observed in the normal pancreas by northern blotting or immunohistochemistry and that its transcript could be a practical biomarker in detecting pancreatic cancer [6]. In this study, we identified PSCA expression in islets of the normal pancreas, which harbor endocrine cells, by immunohistochemistry using a mouse monoclonal anti-PSCA antibody.

Corresponding author:
Norihisa Saeki
Division of Genetics, National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan
Tel: +81-3-3542-2511 (ext3145), Fax: +81-3-3248-1631, E-mail: nsaeki@ncc.go.jp

*Present address: Neurovirology project, Tokyo Metropolitan Institute of Medical Science, 2-1-6 Kamikitazawa, Setagaya-ku, Tokyo 156-8506, Japan

Materials and Methods

Immunohistochemistry

Embedded normal pancreas specimens on glass were purchased from BioChain Institute (Newark, CA, USA). Immunohistochemistry (Fig. 1A) was performed by the same procedure as described in the previous studies [4, 5] with mouse monoclonal anti-PSCA antibody whose specificity had been proved previously [4], rabbit anti-proliferating cell nuclear antigen (PCNA) antibody (sc-7907, Santa Cruz Biotechnology, Santa Cruz, CA, USA). The sections were incubated at 4°C overnight with both antibodies simultaneously, and then with alkaline phosphatase conjugated anti-mouse IgG antibody and peroxidase conjugated anti-rabbit IgG antibody (NICHIREI Biosciences Inc., Tokyo, Japan) for an hour at room temperature. PSCA protein was visualized by Alkaline Phosphatase Substrate Kit III (Vector Laboratories, Burlingame, CA, USA) and PCNA by Vector NovaRED Substrate Kit for Peroxidase (Vector Laboratories). The fluorescent immunohistochemistry (Fig. 1B) was conducted using anti-PSCA, rabbit anti-insulin (sc-9186, Santa Cruz Biotechnology), rabbit anti-glucagon (ab18461, Abcam Japan, Tokyo, Japan), rabbit anti-somatostatin (A0566, Dako Japan, Tokyo, Japan) and goat anti-pancreatic polypeptide (NB100-1793, Novus Biologicals, Littleton, CO, USA) antibodies for primary antibodies, and anti-mouse IgG antibody-Rhodamin, anti-rabbit IgG-FITC and anti-goat IgG-FITC conjugates (sc-2300, sc-2090, sc-2024, respectively, Santa Cruz Biotechnology) for secondary antibodies. The fluorescent signals were obtained by Zeiss LSM510 META Confocal Imaging System (Carl Zeiss MicroImaging, LLC, Thornwood, NY, USA).

Cell lines

The cell lines analysed in this study were described elsewhere; pancreatic cancer cell lines [6], gastric cancer cell lines [4], and gallbladder cancer cell lines [5].

RNA ligase-mediated rapid amplification of 5′ cDNA end (RLM-5′ RACE)

The RLM-5' RACE was conducted on commercially available total RNA isolated from normal human pancreas (BioChain, Newark, CA, USA), and on total RNA isolated from the pancreatic cancer cell lines, using a GeneRacer kit (Invitrogen, Tokyo, Japan). A primer for the 1st strand synthesis (gene specific primer 1) was designed on the sequence in the 2nd exon, and nested PCR was performed with two sets of primers; a GeneRacer 5' primer and the gene specific primer 1, and a GeneRacer 5' nested primer and a primer (gene specific primer 2) containing a sequence partially overlapped with the gene specific primer 1 (Table 1). The PCR products were cloned into pCR4-TOPO vector (Invitrogen) and sequenced.

Quantitative reverse transcription polymerase chain reaction (RT-PCR)

The templates were synthesized from commercially available total RNA isolated from normal human pancreas, stomach and gallbladder (BioChain), or total RNA isolated from the pancreatic, gastric and bladder cancer cell lines, using the ThermoScript RT-PCR System (Invitrogen). Quantitative RT-PCR in Figs. 2A and 3A was performed by TaqMan Gene Expression Assay (Life Technologies Japan, Tokyo, Japan; Applied Biosystems assay ID: Hs00194665_m1 for PSCA, Applied Biosystems Part No. 4326317E for glyceraldehyde-3-phosphate dehydrogenase [GAPDH]), which was conducted for 40 cycles under a condition of 2 steps of temperature: 95°C for 15 seconds and 60°C for 60 seconds, by ABI PRISM 7900HT Sequence Detection System (Life Technologies Japan). Quantitative RT-PCR in Fig. 2C was performed with gene expression assay using SYBR Premix Ex Taq II (Takara Bio Inc., Shiga, Japan), conducted in 40 cycles under a condition of 2 steps of temperature: 95°C for 5 seconds and 60°C for 30 seconds, by the ABI PRISM 7900HT Sequence Detection System. The relative transcript level was calculated using the Ct value of *GAPDH* transcript as reference.

DNA methylation analysis

The DNA methylation status of the *PSCA* enhancer in the pancreatic cancer cell lines was analysed by bisulfite-Pyrosequencing procedure in the same manner as described in the previous report [5]. DNA samples were isolated from the pancreatic cancer cell lines with FlexiGene DNA Kit (Qiagen, Tokyo, Japan) and treated with EpiTect Bisulfite Kit (Qiagen) according to the manufacturer's instructions. DNA methylation of the *PSCA* gene (GenBank accession No. NG_011722) was investigated by bisulfite-Pyrosequencing at *PSCA* enhancer region (position, -3,090 to -2,880) which is in the *PSCA* enhancer region located at -2.7 to -3 kb from the transcription initiation site [5]. PCR was performed in triplicate under a condition of 45 cycles of 3 steps of

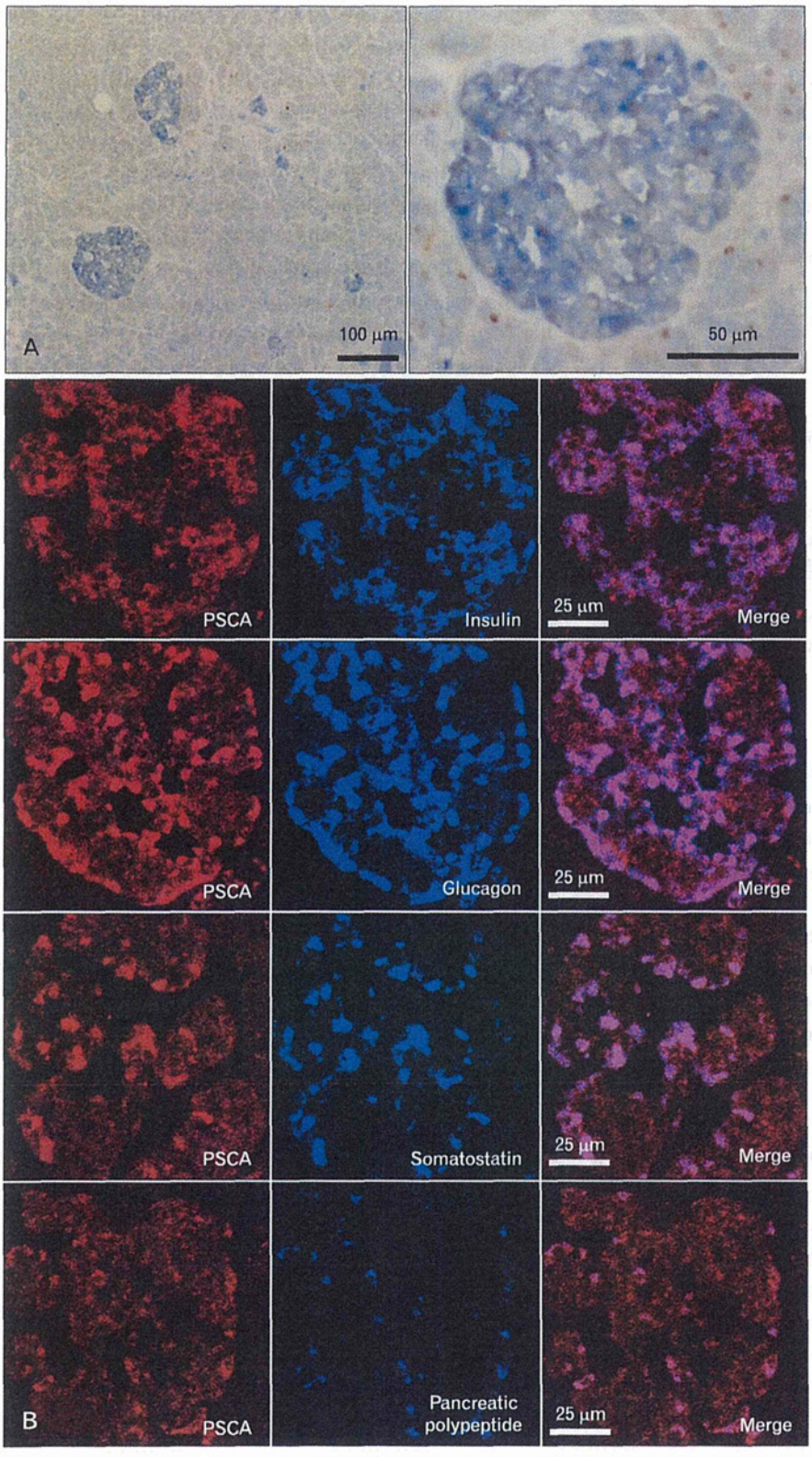

Fig. 1. Prostate stem cell antigen (PSCA) is expressed in the islet of human normal pancreas (immunohistochemistry). (A) Double staining with anti-PSCA (blue) and anti-proliferating cell nuclear (brown) antibodies. The right panel is magnification of the islet shown in the left column. (B) Fluorescent double staining of anti-PSCA antibody (red) and four islet-cell markers (blue), insulin (β cell), glucagon (α cell), somatostatin (δ cell) or pancreatic polypeptide (PP cell). Each series of the image was obtained by focusing on the islet-cell markers.

http://dx.doi.org/10.5115/acb.2012.45.3.149

Table 1. Primers

Primers for RLM-5' RACE	
Gene specific primer 1	5'-GCAGTTCAAGCTGCAGCCTTTGC-3'
Gene specific primer 2	5'-AGCTGCAGCCTTTGCTGATGACGGTC-3'
Primers for quantitative RT-PCR using SYBR Premix Ex Taq II	
PSCA variant 1	5'-CAGTGACCACGAAGGCTGTG-3'
	5'-CCTGGGCTTTGCAGGAGTAG-3'
PSCA variant 2	5'-AGCTGGGGTCCAATCCATAC-3'
	5'-CAGTCCTCGTTGCTCACCTG-3'
GAPDH	5'-ACAGTCAGCCGCATCTTCTT-3'
	5'-GTTAAAAGCAGCCCTGGTGA-3'
PCR primers for bisulfite-pyrosequencing methylation assay (position, -3090 to -2880)	
	5'-TAACAACCCCTATCCCATCC-3'
	5'-GCTGCTCCGGTTCATAGATTTAACAACCCC TATCCCATCC-3'
5'-biotinylated universal primer	5'-GCTGCTCCGGTTCATAGATT-3'
Sequence primers for bisulfite-pyrosequencing methylation assay (position, -3090 to -2880)	
	5'-TATTATTAAATATTTTTTGG-3'
	5'-ATTTATAAATATTTGAGG-3'
	5'-TATAGAGATAGATAAGAAAA-3'
	5'-GGGAGTGGTTAGAGT-3'

RLM-5' RACE, RNA ligase-mediated rapid amplification of 5' cDNA end; RT-PCR, reverse transcription polymerase chain reaction; *PSCA*, prostate stem cell antigen; *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase.

temperature, 95°C for 1 minute, 50°C for 1 minute, and 72°C for 1 minute, after incubation at 95°C for 15 minutes with HotStarTaq (Qiagen). Biotinylated PCR product was purified and proceeded to the Pyrosequencing reaction, which was performed using the Pyrosequencing 96 Sample Prep Tool (Qiagen) and PSQ96MA (Biotage, Uppsala, Sweden). The data of the methylation status for the CpG site in bisulfite-treated sequences were obtained using the primers shown in Table 1, and the average frequency of methylated CpG sites was used as the methylation status index (%5mC).

Results

It was previously reported that PSCA was not expressed in the normal pancreas, but we identified PSCA expression in the islets of the normal human pancreas by immunohistochemistry, and double-staining for 4 islet-cell markers revealed that PSCA is expressed in all of the 4 islet cells: α, β, δ, and PP cells (Fig. 1). Pancreatic expression was confirmed by quantitative RT-PCR (Fig. 2A). To identify the regulatory region for PSCA expression in the pancreas, we first performed RLM-5' RACE to identify the transcription start site (TSS) of the *PSCA* transcripts, which revealed two splicing variants in the pancreas: one is the authentic transcript (variant 1, NCBI accession No. NM_005672) generally detected in other tissues; the other is a non-coding type (variant 2, NR_033343) transcribed at the TSS about 1 kb upstream of that of the authentic one, suggesting that the regulatory region of variant 2 is different from that of the authentic (Fig. 2B). As the amount of variant 2 was small, it may be expressed in the islets. We also examined the expression of variant 2 in the normal stomach and gallbladder and in cell lines derived from their cancers, in addition to pancreatic cancer cell lines, by RT-PCR and found the transcript in all the normal tissues and the cell lines except MIA Paca2 (Fig. 2C). Variant 2 expression was dominant over variant 1 expression in the normal pancreas and the pancreatic cancer cell lines, although it codes no protein. We previously reported that variant 1 expression in gastric and gallbladder cancer cell lines is associated with the DNA methylation level in the enhancer region: the cell lines with relatively high PSCA expression tend to have a lower methylation level [5], but the relation was not observed in pancreatic cancer cell lines (Fig. 3).

Discussion

Although its function is unknown, in the normal prostate epithelium, PSCA is expressed in late intermediate cells in the differentiation process but not in mature secretary cells [9]. In the stomach, PSCA is mainly expressed in the epithelium of the middle portion of the gastric gland, which harbors a pre-pit cell, a transit amplifying precursor of a mucous-secreting pit cell [5]. In these organs, PSCA seems to have a role in the cells differentiating, rather than in differentiated mature cells. It was reported that pancreatic islets contain multipotent stem cells and that the islet cells continue to proliferate slowly for turning over [10, 11]. In this sense, the islet cells possess a property of immature cells, and PSCA may have a function related to the differentiation and/or proliferation of the cells.

In this study, we performed immunohistochemistry on the normal human pancreas, which revealed PSCA expression in islet cells. The expression was confirmed by RT-PCR and, in the process of identification of the TSS in the pancreas, we observed two types of *PSCA* transcript in the normal pancreas as well as in pancreatic cancer cell lines. It is important to note the fact that PSCA is expressed in the normal pancreas, because there are several reports suggesting that PSCA can be utilized as a biomarker of pancreatic cancer [12-14] and it is

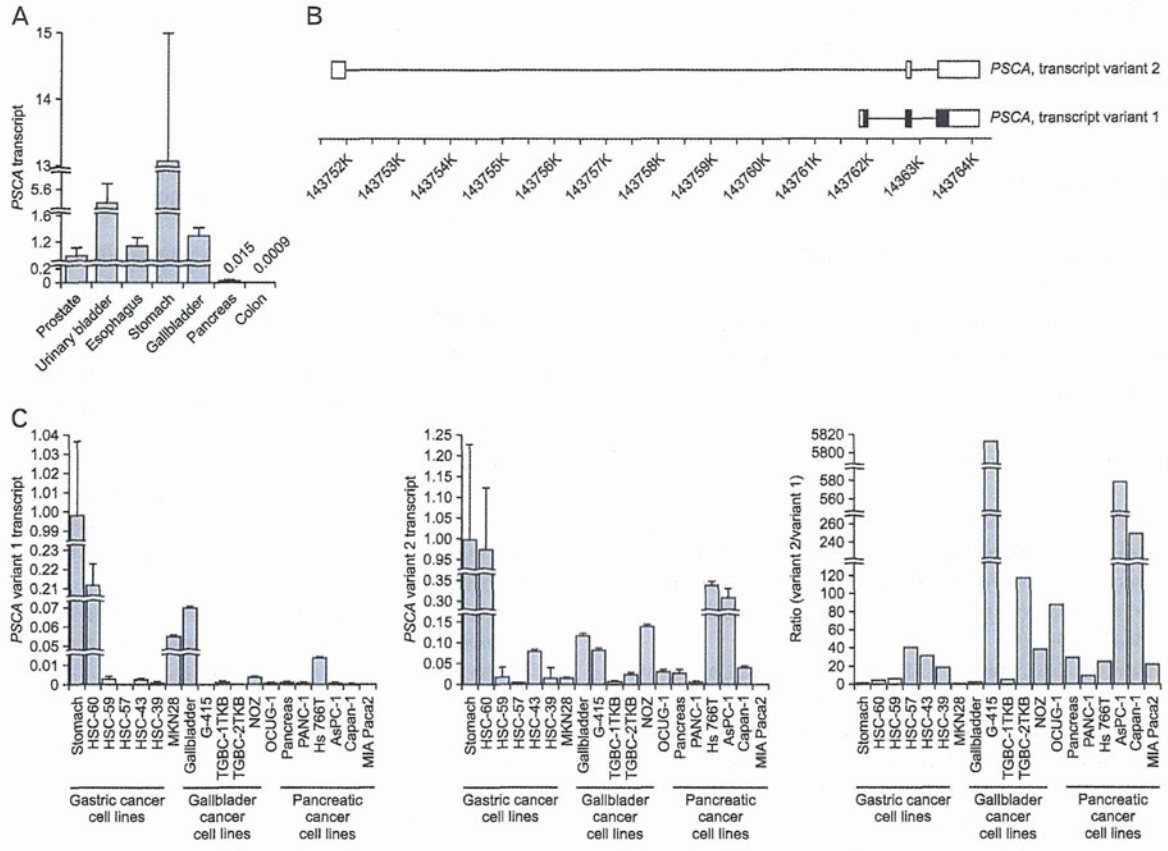

Fig. 2. Two splicing variants of the prostate stem cell antigen (*PSCA*) transcript are expressed in normal pancreatic, gastric and gallbladder tissues and in cancer cell lines derived from their cancers. (A) Quantitative reverse transcription polymerase chain reaction (RT-PCR) revealed *PSCA* transcript (variant 1) in human normal pancreas. Colon is negative control as PSCA is not expressed in that organ. (B) Schematic representation of the structure of the two variants. Position in chromosome 8q24 is based on NCBI Build 37.3. (C) Quantitative RT-PCR revealed expression of variants 1 and 2 in the normal tissues and the cancer cell lines, and variant 2 is dominant in the pancreas and pancreatic cancer cell lines.

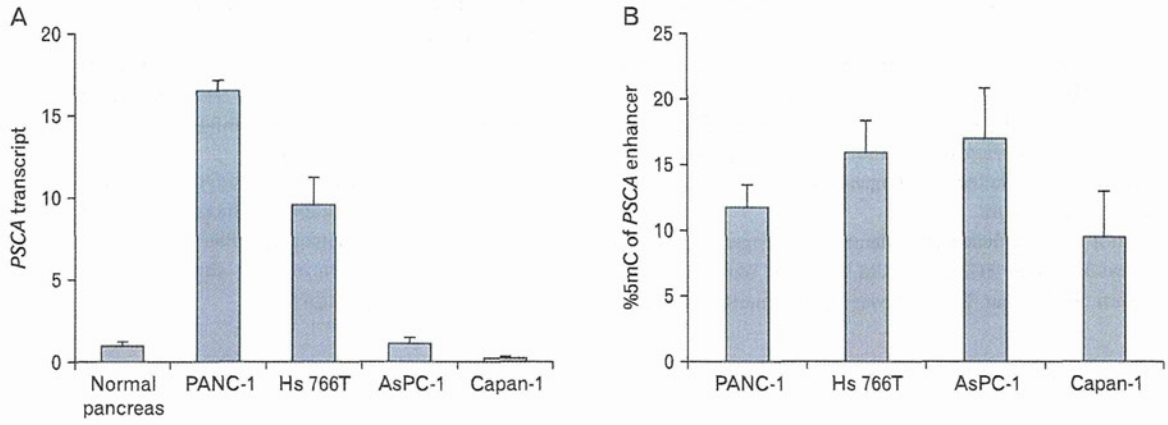

Fig. 3. Bisulfite-pyrosequencing revealed that the DNA methylation status of the prostate stem cell antigen (*PSCA*) enhancer (B) is not correlated to expression level (A) in the pancreatic cancer cell lines.

http://dx.doi.org/10.5115/acb.2012.45.3.149

considered as a target molecule in the treatment of pancreatic and prostate cancers [15, 16]. Clinical applications targeting PSCA have a potential risk of leading to misdiagnosis and adverse events in medical practice.

Acknowledgements

This study was supported by a research grant from the Ministry of Education, Culture, Sports, Science and Technology, Japan (JST grant) and by a Grants-in-Aid for Scientific Research (KAKENHI) by the Japan Society for the Promotion of Science (No. 23501327).

References

1. Saeki N, Gu J, Yoshida T, Wu X. Prostate stem cell antigen: a Jekyll and Hyde molecule? Clin Cancer Res 2010;16:3533-8.
2. Sharom FJ, Radeva G. GPI-anchored protein cleavage in the regulation of transmembrane signals. Subcell Biochem 2004;37:285-315.
3. Reiter RE, Gu Z, Watabe T, Thomas G, Szigeti K, Davis E, Wahl M, Nisitani S, Yamashiro J, Le Beau MM, Loda M, Witte ON. Prostate stem cell antigen: a cell surface marker overexpressed in prostate cancer. Proc Natl Acad Sci U S A 1998;95:1735-40.
4. Study Group of Millennium Genome Project for Cancer, Sakamoto H, Yoshimura K, Saeki N, Katai H, Shimoda T, Matsuno Y, Saito D, Sugimura H, Tanioka F, Kato S, Matsukura N, Matsuda N, Nakamura T, Hyodo I, Nishina T, Yasui W, Hirose H, Hayashi M, Toshiro E, Ohnami S, Sekine A, Sato Y, Totsuka H, Ando M, Takemura R, Takahashi Y, Ohdaira M, Aoki K, Honmyo I, Chiku S, Aoyagi K, Sasaki H, Yanagihara K, Yoon KA, Kook MC, Lee YS, Park SR, Kim CG, Choi IJ, Yoshida T, Nakamura Y, Hirohashi S. Genetic variation in *PSCA* is associated with susceptibility to diffuse-type gastric cancer. Nat Genet 2008;40:730-40.
5. Ono H, Hiraoka N, Lee YS, Woo SM, Lee WJ, Choi IJ, Saito A, Yanagihara K, Kanai Y, Ohnami S, Chiwaki F, Sasaki H, Sakamoto H, Yoshida T, Saeki N. Prostate stem cell antigen, a presumable organ-dependent tumor suppressor gene, is down-regulated in gallbladder carcinogenesis. Genes Chromosomes Cancer 2012;51:30-41.
6. Argani P, Rosty C, Reiter RE, Wilentz RE, Murugesan SR, Leach SD, Ryu B, Skinner HG, Goggins M, Jaffee EM, Yeo CJ, Cameron JL, Kern SE, Hruban RH. Discovery of new markers of cancer through serial analysis of gene expression: prostate stem cell antigen is overexpressed in pancreatic adenocarcinoma. Cancer Res 2001;61:4320-4.
7. Kawaguchi T, Sho M, Tojo T, Yamato I, Nomi T, Hotta K, Hamada K, Suzaki Y, Sugiura S, Kushibe K, Nakajima Y, Taniguchi S. Clinical significance of prostate stem cell antigen expression in non-small cell lung cancer. Jpn J Clin Oncol 2010;40:319-26.
8. Geiger KD, Hendruschk S, Rieber EP, Morgenroth A, Weigle B, Juratli T, Senner V, Schackert G, Temme A. The prostate stem cell antigen represents a novel glioma-associated antigen. Oncol Rep 2011;26:13-21.
9. Tran CP, Lin C, Yamashiro J, Reiter RE. Prostate stem cell antigen is a marker of late intermediate prostate epithelial cells. Mol Cancer Res 2002;1:113-21.
10. Smukler SR, Arntfield ME, Razavi R, Bikopoulos G, Karpowicz P, Seaberg R, Dai F, Lee S, Ahrens R, Fraser PE, Wheeler MB, van der Kooy D. The adult mouse and human pancreas contain rare multipotent stem cells that express insulin. Cell Stem Cell 2011;8:281-93.
11. Desgraz R, Herrera PL. Pancreatic neurogenin 3-expressing cells are unipotent islet precursors. Development 2009;136:3567-74.
12. Grubbs EG, Abdel-Wahab Z, Tyler DS, Pruitt SK. Utilizing quantitative polymerase chain reaction to evaluate prostate stem cell antigen as a tumor marker in pancreatic cancer. Ann Surg Oncol 2006;13:1645-54.
13. McCarthy DM, Maitra A, Argani P, Rader AE, Faigel DO, Van Heek NT, Hruban RH, Wilentz RE. Novel markers of pancreatic adenocarcinoma in fine-needle aspiration: mesothelin and prostate stem cell antigen labeling increases accuracy in cytologically borderline cases. Appl Immunohistochem Mol Morphol 2003;11:238-43.
14. Tanaka M, Komatsu N, Terakawa N, Yanagimoto Y, Oka M, Sasada T, Mine T, Gouhara S, Shichijo S, Okuda S, Itoh K. Increased levels of IgG antibodies against peptides of the prostate stem cell antigen in the plasma of pancreatic cancer patients. Oncol Rep 2007;18:161-6.
15. Wente MN, Jain A, Kono E, Berberat PO, Giese T, Reber HA, Friess H, Büchler MW, Reiter RE, Hines OJ. Prostate stem cell antigen is a putative target for immunotherapy in pancreatic cancer. Pancreas 2005;31:119-25.
16. Morris MJ, Eisenberger MA, Pili R, Denmeade SR, Rathkopf D, Slovin SF, Farrelly J, Chudow JJ, Vincent M, Scher HI, Carducci MA. A phase I/IIA study of AGS-PSCA for castration-resistant prostate cancer. Ann Oncol 2012 May 2 [Epub]. http://dx.doi.org/10.1093/annonc/mds078.

Cancer Science
The official journal of the Japanese Cancer Association

Association of dietary and genetic factors related to one-carbon metabolism with global methylation level of leukocyte DNA

Hiroe Ono,[1,2] Motoki Iwasaki,[3,9] Aya Kuchiba,[4] Yoshio Kasuga,[5] Shiro Yokoyama,[6] Hiroshi Onuma,[6] Hideki Nishimura,[7] Ritsu Kusama,[8] Sumiko Ohnami,[1] Hiromi Sakamoto,[1] Teruhiko Yoshida[1] and Shoichiro Tsugane[3]

[1]Division of Genetics, National Cancer Center Research Institute, Tokyo; [2]Biomedical Science PhD Program, Tokyo Medical and Dental University, Tokyo; [3]Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo, Japan; [4]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA; [5]Department of Surgery, Nagano Matsushiro General Hospital, Nagano; [6]Department of Breast and Thyroid Surgery, Nagano Red Cross Hospital, Nagano; [7]Department of Surgery, Nagano Municipal Hospital, Nagano; [8]Department of Surgery, Nagano Hokushin General Hospital, Nagano, Japan

(Received June 12, 2012/Revised August 21, 2012/Accepted August 30, 2012/Accepted manuscript online September 7, 2012/Article first published online October 17, 2012)

Global hypomethylation of leukocyte DNA has been associated with an increased risk of cancer. As dietary and genetic factors related to one-carbon metabolism may influence both the methylation and synthesis of DNA, we investigated associations between these factors and the global methylation level of peripheral blood leukocyte DNA based on a cross-sectional study of 384 Japanese women. Dietary intake of folate and vitamins B2, B6, and B12 was assessed with a validated semiquantitative food frequency questionnaire. Five polymorphisms in methylenetetrahydrofolate reductase (*MTHFR*) (rs1801133 and rs1801131), methionine synthase (*MTR*) (rs1805087), and methionine synthase reductase (*MTRR*) (rs10380 and rs162049) were genotyped. Global DNA methylation of leukocyte DNA was quantified using Luminometric Methylation Assay. A linear trend of association between methylation and dietary and genetic factors was evaluated by regression coefficients in a multivariable linear regression model. Mean global methylation level (standard deviation) was 70.2% (3.4) and range was from 59.0% to 81.2%. Global methylation level significantly decreased by 0.36% (95% confidence interval, 0.03–0.69) per quartile category for folate level. Subgroup analysis suggested that alcohol drinking modified the association between folate intake and global methylation level ($P_{interaction}$ = 0.01). However, no statistically significant association was observed for intake of vitamins B2, B6, and B12, alcohol consumption, or five single nucleotide polymorphisms of *MTHFR*, *MTR*, and *MTRR*. We found that higher folate intake was significantly associated with a lower level of global methylation of leukocyte DNA in a group of healthy Japanese females. (*Cancer Sci* 2012; 103: 2159–2164)

DNA methylation plays an important role in the epigenetic mechanism of gene regulation[1,2] and cellular differentiation.[3] Aberrant genomic DNA methylation, both in specific genes and in the genome overall, is widely recognized to be associated with cancer.[4] For example, hypermethylation at promoter CpG islands in tumor suppressor genes is an important means of silencing transcription in carcinogenesis.[4] Global DNA hypomethylation in normally methylated regions is thought to contribute to carcinogenesis through the induction of genomic instability.[4] In addition, some previous evidence suggests that DNA hypomethylation could lead to the activation of oncogenes, and global DNA hypomethylation has been linked to hypomethylation in multiple promoter CpG islands.[4,5] Although many studies have investigated aberrant DNA methylation at the tissue level, there is great interest in epigenetic markers in peripheral blood and several epidemiological studies have found that hypomethylation of global peripheral blood cell DNA is associated with an increased cancer risk.[6–9] However, determinants of global methylation level among healthy individuals remain largely unexplored.

Folate and vitamin Bs in one-carbon metabolism are cofactors and cosubstrates for methylation and nucleic acid synthesis and also function as regulatory molecules of these processes.[10] Accumulating epidemiological evidence has suggested that folate intake is associated with a decreased risk of some sites of cancer such as esophagus, colorectum, and pancreas,[11] which implies that folate is associated with cancer risk through the mechanisms of DNA methylation and DNA synthesis. As folate is a universal methyl donor, which is necessary in DNA methylation, it is considered to be a potential determinant of the global methylation level of leukocyte DNA.[12] Intervention studies have suggested that folate might alter DNA methylation levels, but findings have been inconsistent.[12–16] Only a few of the previous observational studies examined associations of dietary and genetic factors related to one-carbon metabolism with global methylation level of leukocyte DNA, and their overall findings showed no association.[6–8] These inconsistent findings might be explained by differences in exposure level of nutrients related to one-carbon metabolism, differences in assay methods of global methylation level, and difference in the distribution of genetic factors related to one-carbon metabolism, either alone or in combination. In particular, no study has investigated the interaction of genetic factors such as SNPs and nutrient intake related to one-carbon metabolism with DNA methylation level.

Here, we used the well-characterized control group of a breast cancer case–control study in Nagano, Japan, to carry out a cross-sectional study to evaluate the associations of dietary and genetic factors related to one-carbon metabolism with the global methylation level of peripheral blood leukocyte DNA among Japanese women.

Materials and Methods

Study subjects. Subjects were the control group in a multicenter, hospital-based case–control study of breast cancer carried out from May 2001 to September 2005 at four hospitals in Nagano Prefecture, Japan. Details of this study have been described previously.[17,18] The study protocol was

[9]To whom correspondence should be addressed.
E-mail: moiwasak@ncc.go.jp

doi: 10.1111/cas.12013

approved by the institutional review board of the National Cancer Center, Tokyo, Japan.

Briefly, healthy female individuals were selected from medical check-up examinees in two of the hospitals and confirmed to not have any cancer. Each subject was recruited as a control and matched for each case by age (within 3 years) and residential area during the study period; the cases were a consecutive series of 405 women aged 20–74 years with newly diagnosed, histologically confirmed invasive breast cancer who were admitted to one of the four hospitals during the survey period. Among potential control subjects, one declined to participate and two refused to provide a blood sample. Consequently, written informed consent was obtained from 405 matched pairs.

Data collection. Participants were asked to complete a self-administered questionnaire that included questions on demographic characteristics, anthropometric factors, smoking habit, family history of cancer, physical activity, medical history, and menstrual and reproductive history. Dietary habits were investigated using a 136-item semiquantitative FFQ that was developed and validated in a Japanese population.[19,20] In the FFQ, participants were questioned on how often they consumed the individual food items (frequency of consumption), as well as relative sizes compared to standard portions. Daily food intake was calculated by multiplying the frequency of each food item in the FFQ by its standard portion and relative size. Daily intakes of nutrients were calculated using the 5th revised and enlarged edition of the Standard Tables of Food Composition in Japan.[21] The validity of nutrient intakes estimated from the FFQ was evaluated in a subsample of the Japan Public Health Center-based Prospective Study, which includes Nagano as one its study areas. Estimated intake according to the FFQ was compared to that in four 7-day dietary records, one carried out in each of the four seasons. Spearman's correlation coefficients between energy-adjusted intakes estimated from the FFQ and from dietary records were 0.35–0.50 for folate, 0.34–0.45 for vitamin B2, 0.36–0.47 for vitamin B6, and 0.27–0.35 for vitamin B12.[18,19]

Participants provided blood samples at the time they returned their self-administered questionnaire. Whole blood in a 7-mL EDTA-2Na Vacutainer (Terumo, Tokyo, Japan) and serum samples were stored at −80°C until analyzed.

Laboratory analysis. Genomic DNA was extracted from the whole blood using a Qiagen FlexiGene DNA Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol.

Global DNA methylation was quantified by LUMA,[22,23] consequent to the findings of our in-house testing of several methods which found LUMA to be relatively reliable and unbiased in assessing small differences in global methylation levels of peripheral blood leukocytes. Three hundred nanograms of genomic DNA was cleaved with *Hap*II + *Eco*RI or *Msp*I + *Eco*RI in two separate 20-μL reaction tubes containing 2 μL of 10× T buffer (330 mM Tris-acetate, 100 mM Mg-acetate, 660 mM K-acetate, 5 mM DTT), 2 μL of 0.1% BSA, and 5 units of each of the restriction enzymes. The reactions were set up in a PSQ96 Plate Low (Qiagen) and incubated at 37°C for 1 h. Then 20 μL annealing buffer containing 200 mM Tris-acetate and 50 mM Mg-acetate (pH 7.6) was added to the cleavage reactions, and samples were assayed using a PSQ96 MA system (Biotage, Uppsala, Sweden). The instrument was programmed to add dNTPs in six steps, consisting of: step 1, dATPαS; step 2, mixture of dGTP + dCTP; step 3, dTTP; step 4, mixture of dGTP + dCTP; step 5, water; and step 6, dATP. Peak heights were calculated using the PSQ96 software. The *Hap*II/*Eco*RI and *Msp*I/*Eco*RI ratios were calculated as (dGTP + dCTP)/dATP for each reaction. The *Hap*II/*Msp*I ratio was then calculated as (*Hap*II/*Eco*RI)/(*Msp*I/*Eco*RI), which corresponds to the proportion of unmethylated CCGG. Restriction enzymes (*Hap*II, *Msp*I, and *Eco*RI) were purchased from Takara Bio (1053A, 1150A and 1040A, respectively; Shiga, Japan). PyroMark Gold Q96 Reagents for pyrosequencing were purchased from Qiagen (972804). DNA quantification was carried out using a Quan-iT PicoGreen dsDNA Reagent and kit (P7581; Invitrogen, Carlsbad, CA, USA). Intra-assay CV was 6.4% at a mean methylation level of 74% (n = 20).

In the present study, we focused on three genes, *MTHFR*, *MTR*, and *MTRR*, which are closely related to DNA methylation in one-carbon metabolism, and selected SNPs in consideration of the availability of functional information. Five polymorphisms in *MTHFR* (rs1801133 and rs1801131), *MTR* (rs1805087), and *MTRR* (rs10380 and rs162049) genes were genotyped by TaqMan SNP Genotyping Assays developed by Applied Biosystems (Foster City, CA, USA). Confirmation that the genotype frequencies were in Hardy–Weinberg equilibrium was done using a χ^2-test as quality control (all P values >0.05).

Statistical analysis. Nutrient intake (folate, vitamin B2, B6, and B12 intake) was adjusted for total energy intake using the residual method[24,25] and divided into quartile categories. Adjusted mean global methylation levels of leukocyte DNA were calculated according to nutrient intake and SNPs related to one-carbon metabolism using a multivariable linear regression model. To test linear trends for mean folate intake levels, regression coefficients (β) were calculated in the multivariable linear regression model using categories of each folate intake level as ordinal variables. The following variables were used for adjustment: age (continuous); BMI (continuous); smoking (never smokers, past smokers, current smokers); alcohol drinking (non-drinkers, occasional drinkers, regular drinkers of <150 g ethanol/week, regular drinkers of ≥150 g ethanol/week); and physical activity in the past 5 years (no, ≤2 days/week, ≥3 days/week). To investigate potential effect modification, subgroup analyses were carried out by nutrient intake and SNPs related to one-carbon metabolism, and tests for interaction were carried out. All reported P-values are two-sided, and significance level was set at $P < 0.05$. All statistical analyses were done with SAS software version 9.1 (SAS Institute, Cary, NC, USA).

Results

After exclusion of subjects who reported extremely low or high total energy intake (<500 or ≥4000 kcal, respectively) or had no DNA sample, 384 healthy Japanese women were included in the present analyses. Mean age and total calorie intake of women in the present study was 54.1 years and 1947.5 kcal, respectively. Mean global methylation level (SD) was 70.2% (3.4) and range was from 59.0% to 81.2%. Table 1 shows global methylation level according to age, BMI, smoking status, and physical activity, which were used for adjustment in Tables 2 and 3. None of these factors was associated with the level of global methylation.

Global methylation levels according to five dietary factors and five SNPs of genes for folate metabolic enzymes are shown in Table 2. We found a statistically significant association between folate intake level and the global methylation level of leukocyte DNA (P = 0.030). Global methylation level decreased by 0.36% (95% CI, 0.03–0.69) per quartile category for folate intake. No associations were found for vitamin B2, B6, or B12 intake, alcohol drinking, or five SNPs of *MTHFR*, *MTR*, and *MTRR*.

Association between mean global methylation level of leukocyte DNA and folate intake by factors related to one-carbon metabolism are shown in Table 3. Subgroup analyses revealed that alcohol drinking modified the association between folate intake and global methylation level ($P_{interaction}$ = 0.01). The

doi: 10.1111/cas.12013