NPY2R, were associated with these pathological conditions in non-Hispanic, European-Caucasian families.

In this study, we investigated the relationships between the polymorphisms rs4425326 (C/T) and rs6857715 (T/C), both of which were claimed to be associated with alcohol and cocaine dependence in the report by Wetherill *et al.*,[16] and nicotine dependence in an elderly Japanese population.

MATERIALS AND METHODS

Questionnaire

Blood was collected from 2517 subjects in the clinical laboratory of Iwata City Hospital during the 5-year period from 2003 to 2008. The criteria for recruitment as subjects of this study were being ambulant, able to communicate orally and 60 years of age or older. All subjects provided written informed consent regarding participation in this study. A leaflet containing a questionnaire about life style, including alcohol consumption, smoking, diet and cancer history, was handed to each subject, and professional interviewers assisted them in filling them out and confirmed their answers. Some of the questions on smoking behavior were the same as those in the revised Fagerström Tolerance Questionnaire,[17] that is, the Fagerström Test for Nicotine Dependence (FTND),[18] which contains six of the original eight questions in the Fagerström Tolerance Questionnaire, and the Tobacco Dependence Screener (TDS) (a screening questionnaire for tobacco/nicotine dependence according to the *International Statistical Classification of Diseases and Related Health Problems* (ICD)-10, *Diagnostic and Statistical Manual of Mental Disorders* (DSM)-III-R and DSM-IV),[19] which consists of 10 questions. The questionnaire also included questions about the numbers of cigarettes smoked per day (CPD), age when the subject started smoking, how many times current smokers had tried to quit and how many times ex-smokers had tried to quit smoking before they succeeded. FTND scores were available for 1296 subjects (1220 men, 76 women (90.4 and 91.6%, respectively, of the ever-smokers)), and TDS scores were available for 1252 subjects (1183 men, 69 women (87.6 and 83.1%, respectively, of the ever smokers)).

The study design was approved by the institutional review board of Hamamatsu University School of Medicine (19-87 and 21-8).

Genotype analysis

DNA was extracted from whole blood by using a QIAamp DNA Blood Maxi kit according to the manufacturer's instructions (Qiagen, Hamburg, Germany). A 50 ng sample of each subject's DNA was used for PCR amplification with the primer sets for NPY2R polymorphisms rs4425326 and rs6857715 by using the Start One (Applied BioSystems, Carlsbad, CA, USA), and assayed by using the Custom TaqMan SNP Genotyping Assay C__26159385_10 and C__29013142_10, respectively. The locations of the *NPY2R* polymorphisms and the exons are shown in Figure 1. The distribution of both genotypes distributions differs among populations. The minor alleles of rs4425326 and rs6857715 are Ts in the European populations, but C in rs4425326 and T in rs6857715 are listed as the ancestral alleles in the SNP database. The allele frequency of the C allele at the rs4425326 in Japanese is 0.341 and at the rs6857715 is 0.557. The background information regarding rs4425326 and rs6857715 is available at http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=4425326 and at http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=6857715, respectively.

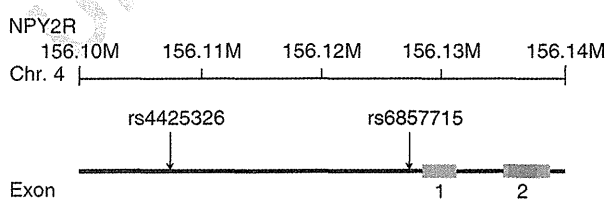

Figure 1

Statistical analysis

Genotype distributions were tested for Hardy–Weinberg equilibrium by using SPSS statistics 17.0. software (SPSS Japan, Tokyo, Japan). χ^2-Tests of each genotype or dominant model were performed for smoking status. The odds ratios were estimated by using a logistic model. The CPD values, FTND scores and TDS scores were evaluated according to smoking status and a dominant model of the polymorphisms by the Kruskal–Wallis test or Mann–Whitney *U*-test (SPSS Japan).

RESULTS

The age, sex and smoking status of the subjects are shown in Table 1. The subjects ranged in age from 60 to 94 years, and they consisted of generations born between 1910 and 1948. There were 1350 ever-smokers (current smokers and ex-smokers) (83.6%) among the 1615 male subjects and 83 (9.2%) among the 902 female subjects, and 21.5% of the men and 3.4% of the women were current smokers. These values are slightly lower than in a recent report[20] stating that the prevalence of male and female Japanese current smokers 60 years of age or more is 27.8 and 6.2%, respectively. The mean CPD values, FTND scores and TDS scores in this study were lower than those previously reported in Japanese[19,21] or American smokers (Table 1).[22] The CPD values of male ex-smokers were higher than those of male current smokers, and TDS scores of the current smokers of both sexes were higher than the TDS scores of ex-smokers of both sexes.

The drinking status and lung cancer history of the subjects are also shown in Table 1. Current drinkers were overrepresented among current smokers, and never-smokers had tendencies to be never-drinkers in both sexes (Supplementary Table 1).

Figure 2, a (male) and b (female), shows the distributions of the FTND scores, and Figure 3, a (male) and b (female), shows the distributions of the TDS scores. The FTND scores of the male smokers ranged from 0 to 10, and the mode value and the mean value were 2 and 3.59, respectively. The FTND scores of the female smokers ranged from 0 to 7, and the mode value and mean value were 1 and 2.41, respectively. The TDS scores of the male smokers ranged from 0 to 10, and the mode value and mean value were 2 and 3.06, respectively. The TDS scores of the female smokers ranged from 0 to 9, and the mode value and mean value were 0 and 2.87, respectively (Figures 2 and 3).

Spearman's rank correlation coefficient for the correlation between the FTND scores and the TDS scores of the male ever-smokers and the female ever-smokers was 0.307 and 0.347, respectively. The *P*-values of both rank correlation coefficients indicated that they were significant at the 1% level. The frequency distribution of both rs4425326 (men: χ^2=1.155, *P*=0.566; women: χ^2=1.595, *P*=0.451) and rs6857715 (men: χ^2=0.020, *P*=0.991; women: χ^2=1.035, *P*=0.593) obeyed the Hardy–Weinberg law in both sexes.

The distribution of rs4425326 of the smokers revealed that the subjects having the C allele (genotypes CC and CT) were over-represented among male current smokers (Table 2), or expressed another way, men having the C allele in this SNP tended to continue to smoke (current smokers). This relation between the rs4425326 C allele and smoking status category was not found in female smokers. There were no differences in the prevalence of the rs6857715 polymorphism in any smoking category in either sex.

Male, but not female ever-smokers who had the rs4425326 TT genotype had significantly higher FTND scores and greater CPD than those with other genotypes (Table 3). In addition, male ever-smokers having the rs6857715 TT genotype had greater CPD than those having the C allele, but not significantly. Any other relations between the

Table 1 Subject profile

Variables	*Male*	P-*value*	*Female*	P-*value*
Number of subjects	1615		902	
Mean age, years (±s.d.)	73.1 (±6.2)		73.0 (±6.4)	
Age distribution, n *(%)*				
60–64	81 (5.0)		51 (5.7)	
65–69	426 (26.4)		251 (27.8)	
70–74	455 (28.2)		240 (26.6)	
75–79	418 (25.9)		197 (21.8)	
80–84	170 (10.5)		134 (14.9)	
85–89	51 (3.2)		25 (2.8)	
90–	14 (0.9)		4 (0.4)	
Smoking status, n *(%)*				
Current smokers	348 (21.5)		31 (3.4)	
Ex-smokers	1002 (62.1)		52 (5.8)	
Never-smokers	265 (16.4)		819 (90.8)	
Mean age according to smoking status, years (±s.d.)				
Current smokers	72.1 (±6.0)	0.002[a]	71.1 (±5.2)	0.078[a]
Ex-smokers	73.4 (±6.0)		71.7 (±6.4)	
Never-smokers	73.4 (±7.0)		73.2 (±6.4)	
Mean age at start of smoking, years (±s.d.)				
Ever smokers	19.7 (±3.7)		34.5 (±13.3)	
Current smokers	19.9 (±4.3)	0.282[b]	37.5 (±15.0)	0.136[b]
Ex-smokers	19.6 (±3.5)		32.7 (±11.9)	
Mean numbers of CPD (±s.d.)				
Ever smokers	21.1 (±13.0)		13.3 (±8.1)	
Current smokers	16.6 (±9.1)	<0.001[b]	12.4 (±6.1)	0.776[b]
Ex-smokers	22.7 (±13.7)		13.8 (±9.1)	
Mean FTND score (±s.d.)				
Ever smokers	3.59 (±2.21)		2.41 (±2.00)	
Current smokers	3.63 (±2.10)	0.495[b]	2.26 (±1.81)	0.777[b]
Ex-smokers	3.57 (±2.25)		2.51 (±2.13)	
Mean TDS score (±s.d.)				
Ever smokers	3.06 (±2.49)		2.87 (±2.47)	
Current smokers	3.69 (±2.41)	<0.001[b]	3.78 (±2.50)	0.018[b]
Ex-smokers	2.84 (±2.48)		2.41 (±2.35)	
Drinking status, n *(%)*				
Current drinkers	854 (52.9)		177 (19.6)	
Ex-drinkers	319 (19.8)		50 (5.5)	
Never-drinkers	442 (27.4)		675 (74.8)	
Lung cancer history, n *(%)*				
Yes	47 (2.9)		12 (1.3)	
No	1568 (97.1)		890 (98.7)	

Abbreviations: CPD, cigarettes smoked per day; FTND, the Fagerström Test for Nicotine Dependence; s.d., standard deviation; TDS, the Tobacco Dependence Screener.
Ever smokers: current smokers and ex-smokers.
[a]Kruskal-Wallis test comparing three statuses.
[b]Mann–Whitney *U*-test comparing current smokers and ex-smokers.

rs6857715 polymorphism and the FTND scores or CPD value were not found in either sex (Table 3). Neither the rs4425326 nor the rs6857715 polymorphism was associated with the TDS scores in either sex (Table 3).

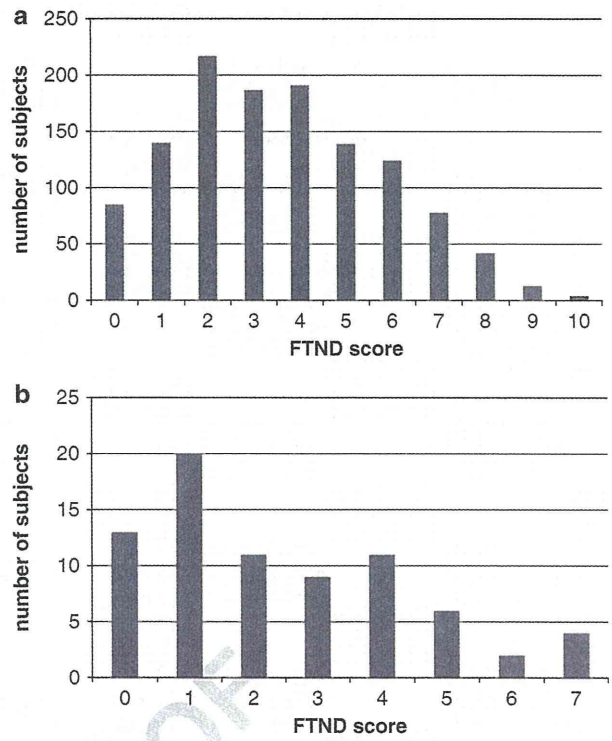

Figure 2 (**a**) Distribution of the Fagerström Test for Nicotine Dependence (male). (**b**) Distribution of the Fagerström Test for Nicotine Dependence (female).

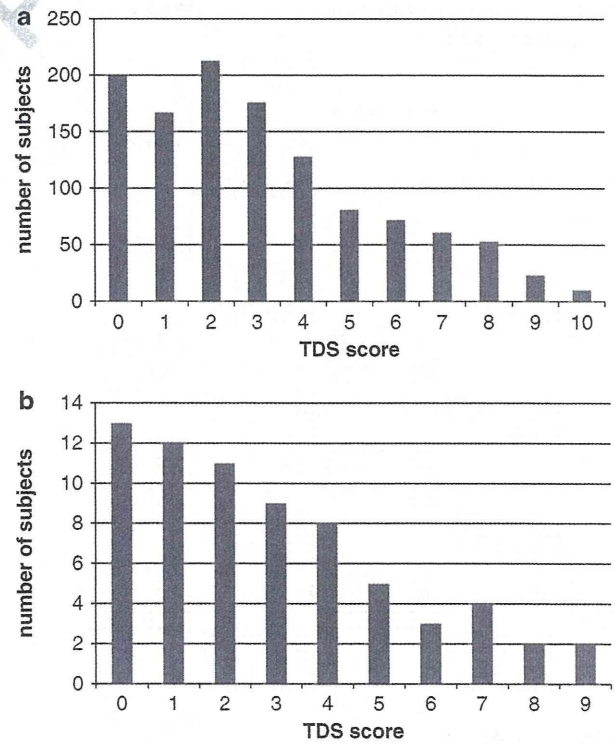

Figure 3 (**a**) Distribution of the Tobacco Dependence Screener (male). (**b**) Distribution of the Tobacco Dependence Screener (female).

Table 2 Comparison of subjects distribution of current smokers and ex-smokers according to two polymorphisms of *NPY2R*

	Smoking status				*Dominant model (CC+CT vs TT)*		*Dominant model (CC+CT vs TT)*	
	Total n *(%)*	*Current smokers* n *(%)*	*Ex-smokers* n *(%)*	P-*value*[a]	*OR*[b] *(95% CI)*	P-*value*	*OR*[c] *(95% CI)*	P-*value*
Males								
rs4425326								
TT	680 (50.4)	158 (45.4)	522 (52.1)		1		1	
CT	566 (41.9)	156 (44.8)	410 (40.9)	0.055	1.308 (1.024–1.670)	0.034	1.300 (1.017–1.662)	0.036
CC	104 (7.7)	34 (9.8)	70 (7.0)					
rs6857715								
TT	319 (23.6)	82 (23.6)	237 (23.7)		1		1	
CT	679 (50.3)	164 (47.1)	515 (51.4)	0.244	1.005 (0.754–1.339)	1.000	1.013 (0.759–1.351)	0.931
CC	352 (26.1)	102 (29.3)	250 (25.0)					
Females								
rs4425326								
TT	37 (44.6)	17 (54.8)	20 (38.5)		1		1	
CT	44 (53.0)	14 (45.2)	30 (57.7)	0.281	0.515 (0.209–1.268)	0.175	0.512 (0.208–1.264)	0.147
CC	2 (2.4)	0 (0)	2 (3.8)					
rs6857715								
TT	19 (22.9)	10 (32.3)	9 (17.3)		1		1	
CT	47 (56.6)	14 (45.2)	33 (63.5)	0.205	0.440 (0.155–1.245)	0.176	0.446 (0.157–1.268)	0.130
CC	17 (20.5)	7 (22.6)	10 (19.2)					

Abbreviations: CI, confidence interval; OR, odds ratio.
Allele frequency of C in rs4425326: 0.278. Allele frequency of T in rs6857715: 0.489. The risk alleles of these two polymorphisms are designated as C.
[a]The χ^2-tests were performed based on 3×2 tables.
[b]The odds ratios were calculated for the genotypes concerned in the current smoking status.
[c]The odds ratios were adjusted for age and calculated for the genotypes concerned in the current smoking status.

Table 3 Comparison of the numbers of CPD, FTND scores and TDS scores of ever smokers according to the dominant model (CC+CT vs TT) of two polymorphisms of *NPY2R*

	CPD			*FTND*			*TDS*		
	N	*Mean* ± *s.d.*	P-*value*[a]	n	*Mean* ± *s.d.*	P-*value*[a]	n	*Mean* ± *s.d.*	P-*value*[a]
Males (CPD, n=*1348; FTND,* n=*1220; TDS,* n=*1183)*									
rs4425326									
TT	680	22.3 ± 13.3	<0.001	614	3.80 ± 2.31	0.003	598	2.95 ± 2.49	0.085
CT+CC	668	19.9 ± 12.5		606	3.37 ± 2.08		585	3.18 ± 2.51	
rs6857715									
TT	319	21.8 ± 12.6	0.054	291	3.76 ± 2.35	0.197	281	2.88 ± 2.32	0.314
CT+CC	1029	20.9 ± 13.1		929	3.53 ± 2.16		902	3.12 ± 2.55	
Females (CPD, n=*83; FTND,* n=*76; TDS,* n=*69)*									
rs4425326									
TT	37	14.0 ± 8.0	0.417	34	2.56 ± 1.99	0.480	29	3.31 ± 2.99	0.501
CT+CC	46	12.7 ± 8.2		42	2.29 ± 2.02		40	2.55 ± 2.00	
rs6857715									
TT	19	11.2 ± 6.1	0.218	17	1.88 ± 1.50	0.334	14	2.21 ± 2.29	0.264
CT+CC	64	13.9 ± 8.5		59	2.56 ± 2.10		55	3.04 ± 2.51	

Abbreviations: CPD, cigarettes smoked per day; FTND, the Fagerström Test for Nicotine Dependence; s.d., standard deviation; TDS, the Tobacco Dependence Screener.
[a]Mann–Whitney *U*-test.

DISCUSSION

A long list of genes is considered candidates for a relation to smoking behavior or nicotine dependence. Li and Burmeister,[23] recently conducted an extensive reviewed of research on genes related to addictions, and in that seminal review, they listed genes that are considered candidates for an association with at least one drug addiction (62

genes) in two large tables and genes having one or more of whose variants have been associated with addiction to at least one substance (41 genes) in supplementary tables. *NPY*, but not *NPYR*, was included in the tables, and thus far there have been few studies on associations between nicotine dependence and the NPY–NPYR axis.[23]

We attempted to determine whether the rs4425326 and rs6857715 polymorphisms of the *NPY2R* gene are related to another addictive behavior in humans, nicotine dependence, because Wetherill *et al.*[16] have recently shown an association between these polymorphisms and both alcohol dependence and cocaine addiction in humans. Dependence on nicotine is one of the addictions that has been extensively studied, and the genetic aspects of smoking behavior are being extensively investigated now.[23]

Our study on smoking behavior is the first to report a correlation between *NPY2R* gene polymorphism and nicotine dependence.

We found a significant correlation between rs4425326 C-containing allelotypes, which have been hypothesized to be high-risk allelotypes for addiction and current smoking status in male smokers. This is the first report to show that the *NPY2R* polymorphism is related to human addictive behaviors in a non-Caucasian population. This is also the first time an association between *NPY2R* and nicotine dependence has been shown. However, no relation was found between TDS scores and any of the *NPY2R* allelotypes, perhaps because the traits detected by the scores on two questionnaires (FTND and TDS) are different, and how the genetic components control these traits in establishing individual nicotine dependence is not elucidated.

The mechanistic significance of the rs4425326 locus, 0.2 Mb upstream from the first exon of *NPY2R* (Figure 1), is unknown.

Interestingly, 'addiction'-allele rs4425326 C holders (according to Wetherill *et al.*[16]) were more prevalent in the current smokers in this study than among the ex-smokers in spite of the lower FTND scores of the allele C holders, meaning that the rs4425326 C allele holders have milder nicotine dependence based on FTND scores but they do not quit smoking ~~(Supplementary Table 2)~~. Like the FTND scores, the CPD values of the rs4425326 C allele holders were lower, too. The analysis of male current smokers also showed lower CPD values and FTND scores in the rs4425326 C holders (P=0.100 and 0.098, respectively), but not significantly (Supplementary Table 3). Thus, Japanese male rs4425326 C allele holders are modest but very persistent nicotine-seekers. This paradoxical result may reflect the complex decision-making process regarding smoking in the elderly Japanese. No genetic tendencies were detected in any of the nicotine dependence scores among the female smokers in our study. The failure to find any genetic tendencies among them was mainly because of the relatively small numbers of female smokers in our study.

Smoking is a complicated personal behavior that is influenced by multiple factors, including the social, cultural and sometimes even the political environment. The role of a single genetic polymorphism must not be overestimated in explaining individual smoking behavior. The interpretations of this study have several limitations. We recruited the elderly people in a rural city, and this population has demographical and occupational characteristics different from those in urban cities or agricultural villages. We expected our study subjects have established smoking behavior. This might imply the dependence scores collected here have some biases, causing the limitation of our study. In a different point of view, the information on the established smoking behavior as a life-long habit in individuals may make a unique contribution to understand the genetic effect on whole life of humans. Anyway, it will be necessary to validate our observations by replication studies in the future. However, we think our data provide a major clue to understanding human smoking behavior.

ACKNOWLEDGEMENTS

We thank all the participants of this study, the nurses and laboratory technicians of Iwata City Hospital, and Dr Kazuhiko Nakamura for critical reading of the article and helpful comments. This work was supported by grants-in-aid from the Japanese Ministry of Health, Labour and Welfare for the Comprehensive 10-Year Strategy for Cancer Control (19-19), from the Japanese Ministry of Education, Culture, Sports, Science and Technology for priority area (20014007, 221S0001), from the Smoking Research Foundation and from the 21st Century COE program.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

1 Rose, P. M., Fernandes, P., Lynch, J. S., Frazier, S. T., Fisher, S. M., Kodukula, K. *et al.* Cloning and functional expression of a cDNA encoding a human type 2 neuropeptide Y receptor. *J. Biol. Chem.* **270,** 22661–22664 (1995).
2 Elbers, C. C., de Kovel, C. G., van der Schouw, Y. T., Meijboom, J. R., Bauer, F., Grobbee, D. E. *et al.* Variants in neuropeptide Y receptor 1 and 5 are associated with nutrient-specific food intake and are under recent selection in Europeans. *PLoS One* **4,** e7070 (2009).
3 Wang, L., Rao, F., Zhang, K., Mahata, M., Rodriguez-Flores, J. L., Fung, M. M. *et al.* Neuropeptide Y(1) receptor NPY1R discovery of naturally occurring human genetic variants governing gene expression in cella as well as pleiotropic effects on autonomic activity and blood pressure *in vivo*. *J. Am. Coll. Cardiol.* **54,** 944–954 (2009).
4 Hodges, G. J., Jackson, D. N., Mattar, L., Johnson, J. M. & Shoemaker, J. K. Neuropeptide Y and neurovascular control in skeletal muscle and skin. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **297,** R546–R555 (2009).
5 Zhou, Z., Zhu, G., Hariri, A. R., Enoch, M. A., Scott, D., Sinha, R. *et al.* Genetic variation in human NPY expression affects stress response and emotion. *Nature* **452,** 997–1001 (2008).
6 Cotton, C. H., Flint, J. & Campbell, T. G. Is there an association between NPY and neuroticism? *Nature* **458,** E6; discussion E7 (2009).
7 Okahisa, Y., Ujike, H., Kotaka, T., Morita, Y., Kodama, M., Inada, T. *et al.* Association between neuropeptide Y gene and its receptor Y1 gene and methamphetamine dependence. *Psychiatry Clin. Neurosci.* **63,** 417–422 (2009).
8 Kauhanen, J., Karvonen, M. K., Pesonen, U., Koulu, M., Tuomainen, T. P., Uusitupa, M. I. *et al.* Neuropeptide Y polymorphism and alcohol consumption in middle-aged men. *Am. J. Med. Genet.* **93,** 117–121 (2000).
9 Lappalainen, J., Kranzler, H. R., Malison, R., Price, L. H., Van Dyck, C., Rosenheck, R. A. *et al.* A functional neuropeptide Y Leu7Pro polymorphism associated with alcohol dependence in a large population sample from the United States. *Arch. Gen. Psychiatry* **59,** 825–831 (2002).
10 Ilveskoski, E., Kajander, O. A., Lehtimaki, T., Kunnas, T., Karhunen, P. J., Heinälä, P. *et al.* Association of neuropeptide y polymorphism with the occurrence of type 1 and type 2 alcoholism. *Alcohol Clin. Exp. Res.* **25,** 1420–1422 (2001).
11 Bhaskar, L. V., Thangaraj, K., Shah, A. M., Pardhasaradhi, G., Praveen Kumar, K., Reddy, A. G. *et al.* Allelic variation in the NPY gene in 14 Indian populations. *J. Hum. Genet.* **52,** 592–598 (2007).
12 Zhu, G., Pollak, L., Mottagui-Tabar, S., Wahlestedt, C., Taubman, J., Virkkunen, M. *et al.* NPY Leu7Pro and alcohol dependence in Finnish and Swedish populations. *Alcohol Clin. Exp. Res.* **27,** 19–24 (2003).
13 Mottagui-Tabar, S., Prince, J. A., Wahlestedt, C., Zhu, G., Goldman, D. & Heilig, M. A novel single nucleotide polymorphism of the neuropeptide Y (NPY) gene associated with alcohol dependence. *Alcohol Clin. Exp. Res.* **29,** 702–707 (2005).
14 Okubo, T. & Harada, S. Polymorphism of the neuropeptide Y gene: an association study with alcohol withdrawal. *Alcohol Clin. Exp. Res.* **25,** 59S–62S (2001).
15 Thiele, T. E. & Badia-Elder, N. E. A role for neuropeptide Y in alcohol intake control: evidence from human and animal research. *Physiol. Behav.* **79,** 95–101 (2003).
16 Wetherill, L., Schuckit, M. A., Hesselbrock, V., Xuei, X., Liang, T., Dick, D. M. *et al.* Neuropeptide Y receptor genes are associated with alcohol dependence, alcohol withdrawal phenotypes, and cocaine dependence. *Alcohol Clin. Exp. Res.* **32,** 2031–2040 (2008).
17 Fagerstrom, K. O. Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addict. Behav.* **3,** 235–241 (1978).
18 Heatherton, T. F., Kozlowski, L. T., Frecker, R. C. & Fagerstrom, K. O. The Fagerstrom Test for Nicotine Dependence: a revision of the Fagerstrom Tolerance Questionnaire. *Br. J. Addict.* **86,** 1119–1127 (1991).

19 Kawakami, N., Takatsuka, N., Inaba, S. & Shimizu, H. Development of a screening questionnaire for tobacco/nicotine dependence according to ICD-10, DSM-III-R, and DSM-IV. *Addict. Behav.* **24,** 155–166 (1999).
20 Japan Smoking Rate Survey [homepage on the Internet], 2009. http://www.health-net.or.jp/tobacco/product/pd090000.html, available on 21 July 2010.
21 Ito, H., Hamajima, N., Matsuo, K., Okuma, K., Sato, S., Ueda, R. *et al.* Monoamine oxidase polymorphisms and smoking behaviour in Japanese. *Pharmacogenetics* **13,** 73–79 (2003).
22 Piper, M. E., McCarthy, D. E., Bolt, D. M., Smith, S. S., Lerman, C., Benowitz, N. *et al.* Assessing dimensions of nicotine dependence: an evaluation of the Nicotine Dependence Syndrome Scale (NDSS) and the Wisconsin Inventory of Smoking Dependence Motives (WISDM). *Nicotine Tob. Res.* **10,** 1009–1020 (2008).
23 Li, M. D. & Burmeister, M. New insights into the genetics of addiction. *Nat. Rev. Genet.* **10,** 225–231 (2009).

Supplementary Information accompanies the paper on Journal of Human Genetics website (http://www.nature.com/jhg)

Association between neurexin 1 (*NRXN1*) polymorphisms and the smoking behavior of elderly Japanese

Naomi Sato[a,b], Shinji Kageyama[a], Renyin Chen[a], Masaya Suzuki[a], Fumihiko Tanioka[c], Takaharu Kamo[a], Kazuya Shinmura[a], Akiko Nozawa[b] and Haruhiko Sugimura[a]

Psychiatric Genetics 2010, **00**:000–000

Departments of [a]Pathology, [b]Clinical Nursing, Hamamatsu University School of Medicine, Hamamatsu, Shizuoka and [c]Division of Pathology and Laboratory Medicine, Iwata City Hospital, Iwata, Japan

Correspondence to Professor Haruhiko Sugimura, MD, PhD, Department of Pathology, Hamamatsu University School of Medicine, 1-20-1, Handayama, Higashi-ward, Hamamatsu, 431-3192, Shizuoka, Japan

Tel: +81 53 435 2220; fax: +81 53 435 2225; e-mail: hsugimur@hama-med.ac.jp

Present address: Renyin Chen, Department of Pathology, Zhengzhou University, Zhengzhou, Henan, China

Received 21 September 2009 **Revised** 16 January 2010
Accepted 7 March 2010

Neurexin 1 (*NRXN1*, MIM 600565) is a molecule that expressed in neurons and is a receptor for α-latrotoxin (Rowen *et al.*, 2002). Bierut *et al.* (2007) pinpointed the *NRXN1* gene locus as being responsible for nicotine dependence in a case–control analysis in which the cases were American and Australian smokers whose Fagerström test for nicotine dependence (FTND) score was 4 or more. The other group also documented a significant contribution of the polymorphism at the *NRXN1* locus to nicotine dependence in the European–American and African–American populations (Nussbaum *et al.*, 2008). They evaluated 21 *NRXN1* polymorphisms in European–American and African–American smokers for associations with FTND scores, number of cigarettes smoked per day (CPD), and the heaviness of smoking index, which is a combination of CPD and time before the first cigarette of the day. The results showed that two of 21 single nucleotide polymorphisms, rs2193225 and rs6721498, were associated with heaviness of smoking index and FTND scores in some American populations of European and African origin. In this letter, we report the evaluation of these polymorphisms in an elderly Japanese population (60–94 years old) according to smoking history. rs6721498 and rs2193225 were genotyped in 2516 Japanese with various smoking habits [1348 ever-smokers (current smokers and exsmokers; 83.6%) among the 1612 male participants and 82 (9.1%) among the 904 female participants, and current smokers were 21.6 and 3.4% of the male and female participants, respectively]. FTND and the tobacco dependence screener (TDS), which was developed so that the characteristics detected would be more closely correlated with those defined by the International Statistical Classification of Diseases and Related Health Problems, 10th version, *Diagnostic and Statistical Manual of Mental Disorders*, Revised Third Edition, and *Diagnostic and Statistical Manual of Mental Disorders*, Fourth Edition (Kawakami *et al.*, 1999), were applied to almost all of the ever-smokers among the participants.

Genotype frequencies of these polymorphisms were 0.168 (AA), 0.486 (AG), and 0.347 (GG) in the rs6721498 (A allele 0.411) and 0.659 (AA), 0.304 (AG), and 0.037 (GG) in the rs2193225 (G allele 0.189). Male ever-smokers with the rs2193225 GG type were more prevalent in the higher TDS score category ($P = 0.056$), but not in the higher FTND score category. The CPD was greater in the male ever-smokers with the rs2193225 GG genotype ($P = 0.06$ by analysis of variance; 0.032 and 0.025 by the Dunnett multiple test). In contrast, the rs2193225 GG type was overrepresented in the male never-smokers ($P = 0.006$). There were no differences in the distributions of the FTND scores or TDS scores according to the rs6721498 genotype. We concluded that *NRXN1* polymorphisms are associated with the smoking behavior of elderly Japanese. The contrasting associations between rs2193225 GG type and smoking initiation (ever vs. never) and smoking persistence (according to the TDS score) may imply a difference between the role of *NRXN1* in the initiation and persistence of smoking. This genotype may be more closely associated with the smoking behavior patterns detected by the TDS than that by the FTND.

Smoking is a highly complicated behavior, and no single genetic polymorphism has much influence on it. Further validation and replication in several populations are important.

Acknowledgements

This work is supported by a Grant-in-Aid for Scientific Research on priority areas (2001407) from the Japanese Ministry of Education, Culture, Sports, Science and Technology, and from the Smoking Research Foundation.

DOI: 10.1097/YPG.0b013e32833a21f9

References

Bierut LJ, Madden PA, Breslau N, Johnson EO, Hatsukami D, Pomerleau OF, *et al.* (2007). Novel genes identified in a high-density genome wide association study for nicotine dependence. *Hum Mol Genet* **16**:24–35.

Kawakami N, Takatsuka N, Inaba S, Shimizu H (1999). Development of a screening questionnaire for tobacco/nicotine dependence according to ICD-10, DSM-III-R, and DSM-IV. *Addict Behav* **24**:155–166.

Nussbaum J, Xu Q, Payne TJ, Ma JZ, Huang W, Gelernter J, *et al.* (2008). Significant association of the neurexin-1 gene (NRXN1) with nicotine dependence in European- and African-American smokers. *Hum Mol Genet* **17**:1569–1577.

Rowen L, Young J, Birditt B, Kaur A, Madan A, Philipps DL, *et al.* (2002). Analysis of the human neurexin genes: alternative splicing and the generation of protein diversity. *Genomics* **79**:587–597.

Cancer Science
The official journal of the Japanese Cancer Association

Imaging mass spectrometry of gastric carcinoma in formalin-fixed paraffin-embedded tissue microarray

Yoshifumi Morita,[1,2] Koji Ikegami,[2] Naoko Goto-Inoue,[2] Takahiro Hayasaka,[2] Nobuhiro Zaima,[2] Hiroki Tanaka,[1] Takashi Uehara,[1] Tomohiko Setoguchi,[1] Takanori Sakaguchi,[1] Hisashi Igarashi,[3] Haruhiko Sugimura,[3] Mitsutoshi Setou[2,4] and Hiroyuki Konno[1]

[1]Second Department of Surgery, [2]Departments of Molecular Anatomy, and [3]Pathology, Hamamatsu University School of Medicine, Shizuoka, Japan

(Received June 6, 2009/Revised September 16, 2009/Accepted September 20, 2009/Online publication November 24, 2009)

The popularity of imaging mass spectrometry (IMS) of tissue samples, which enables the direct scanning of tissue sections within a short time-period, has been considerably increasing in cancer proteomics. Most pathological specimens stored in medical institutes are formalin-fixed; thus, they had been regarded to be unsuitable for proteomic analyses, including IMS, until recently. Here, we report an easy-to-use screening method that enables the analysis of multiple samples in one experiment without extractions and purifications of proteins. We scanned, with an IMS technique, a tissue microarray (TMA) of formalin-fixed paraffin-embedded (FFPE) specimens. We detected a large amount of signals from trypsin-treated FFPE-TMA samples of gastric carcinoma tissues of different histological types. Of the signals detected, 54 were classified as signals specific to cancer with statistically significant differences between adenocarcinomas and normal tissues. We detected a total of 14 of the 54 signals as histological type-specific with the support of statistical analyses. Tandem MS revealed that a signal specific to poorly differentiated cancer tissue corresponded to histone H4. Finally, we verified the IMS-based finding by immunohistochemical analysis of more than 300 specimens spotted on TMAs; the immunoreactivity of histone H4 was remarkably strong in poorly differentiated cancer tissues. Thus, the application of IMS to FFPE-TMA can enable high-throughput analysis in cancer proteomics to aid in the understanding of molecular mechanisms underlying carcinogenesis, invasiveness, metastasis, and prognosis. Further, results obtained from the IMS of FFPE-TMA can be readily confirmed by commonly used immunohistochemical analyses. (*Cancer Sci* 2010; 101: 267–273)

Intensive genome-based surveys are performed on candidate biomarker transcripts relevant to cancer tissues by utilizing the advances in high-throughput microarrays.[1,2] Further, various single-nucleotide polymorphism (SNP) analyses have been performed to further understand cancer.[3,4] Recently, cancer genome resequencing has been increasingly performed to acquire specific genomic data.[5,6] To achieve the systematic understandings of cancer by systems biology, data from genome resequencing and the corresponding data from the transcriptomes should be combined with the individual metabolome and proteome data of the cancer patient.[7]

Several approaches for the investigation of global alterations in proteomics have emerged. Mass spectrometry (MS) is used in combination with 2D electrophoresis or liquid chromatography.[8,9] Further, protein microarrays offer a means of effective identification of cancer-specific protein alterations to researchers.[10]

Despite the existence of techniques for the global detection of cancer-specific alterations at the protein level, proteomic approaches continue to possess two major disadvantages. Under most circumstances, proteomic approaches only allow a limited number of samples to be analyzed in an experiment. Additionally, proteomic techniques are not adopted for the investigation of large amounts of archival specimens that are stored in hospitals and medical institutes. TMAs have been developed for the analysis of a large number of specimens by antibody labeling.[11] It allows high-throughput profiling of the molecular and pathological alterations in tissue specimens.

In recent years, IMS has emerged and developed dramatically in the field of proteomics and metabolomics.[12,13] IMS enables simultaneous analysis of thousands of proteins directly from a tissue sample without protein extraction and usage of target-specific reagents such as antibodies.[14,15]

In this study, we combine the TMA and IMS technique, and introduce a simple and easy-to-use protocol to detect, by a single experimental trial, cancer-specific or histological type-specific proteins. Further, we optimized the IMS procedure for the FFPE samples that are commonly used in hospitals and stored for long time.

Materials and Methods

Specimens. We chose gastric cancer to perform the study for the evaluation of our experimental paradigm. Gastric cancer is the fourth most frequent cancer and the second leading cause of cancer-related death in the world.[16] In Japanese hospitals, large amounts of gastric cancer specimens are stored and are, thus, readily available to perform studies.

Human gastric cancer tissues and adjacent normal tissues were provided by the Diagnostic Pathology Division, Hospital of Hamamatsu University School of Medicine, Shizuoka, Japan. The study was performed in accordance with the guidelines for pathological specimen handling, which was approved by the ethical committee of the Hamamatsu University School of Medicine. Histological classification was based on Japanese Classification of Gastric Carcinoma, 2nd English edition.[17] Further, each tissue did not contain non-tumor tissue, confirmed by two pathologists.

For IMS, we used the old specimens which had been fixed in 10% neutral formalin promptly after surgery and had been stored for up to 2 years in paraffin. Tissue blocks of three cancer tissues and one adjacent normal tissue were cored using tissue microarrayer type KIN (Azumaya, Tokyo, Japan). A cylinder, 3 mm in diameter, was taken and placed into the recipient block. Three cancer tissues and one non-tumor gastric mucosal tissue were aligned as shown in Fig. 1a.

Sample preparation. For analysis, the FFPE tissue microarray blocks were sliced into 10-μm-thick serial sections; further, for hematoxylin–eosin staining, these blocks were sliced into 1-μm-thick sections, using a microtome (Tissue-Tek, Feather Trustome; Sakura Finetek, Tokyo, Japan). The analysis samples were deposited onto indium-tin-oxide (ITO)-coated glass slides

[4]To whom correspondence should be addressed.
E-mail: setou@hama-med.ac.jp

doi: 10.1111/j.1349-7006.2009.01384.x

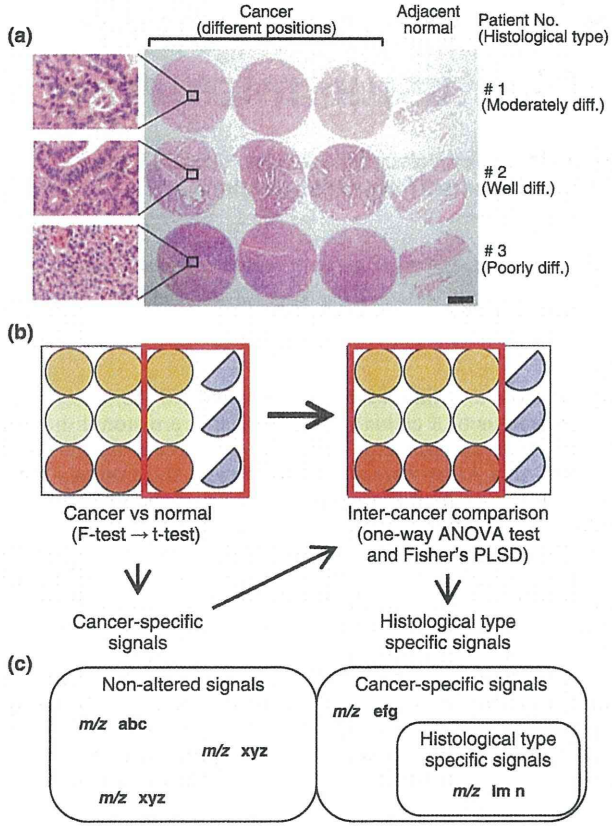

Fig. 1. Experimental paradigm design. (a) Formalin-fixed paraffin-embedded (FFPE) samples were cored with a 3-mm diameter needle and arranged in a line with three cancer tissues and one adjacent normal tissue. The histological type of the cancer of Patient 1 was moderately differentiated adenocarcinoma, that of Patient 2 was well-differentiated adenocarcinoma, and that of Patient 3 was poorly differentiated adenocarcinoma. Hematoxylin–eosin stain, ×10. Scale bar, 1 mm. Enclosed area corresponded to magnified microscopic image. Hematoxylin–eosin stain, ×400 (b) The schema of FFPE samples and the workflow of statistical analysis are shown. (c) The schema that categorizes the acquired signals is presented.

(Bruker Daltonics, Bremen, Germany), and the staining samples were loaded onto regular glass microscope slides by scooping the sections in a 50°C water bath, and then dried on an extender at 45°C. Paraffin was removed by 10-min immersion in xylene at 60°C. Subsequently, the slides were washed by stepwise immersions of 5-min duration each; this involved slide immersion in 100% ethanol twice, and once each in 90% ethanol, 80% ethanol, and 70% ethanol. After rehydration, these slides were incubated in a humid chamber at 55°C overnight.

Tryptic digestion. The sample slide was inserted into a slot on matrix-assisted laser desorption/ionization (MALDI) target plates affixed with conductive tape, and inserted into a chemical inkjet printer (CHIP-1000; Shimadzu, Kyoto, Japan). Trypsin solution was prepared by dissolving 20 μg of trypsin (Sigma, St. Louis, MO, USA) in 200 μL of 20-mM ammonium hydrogen carbonate (NH_4HCO_3). Trypsin microspotting was performed with CHIP-1000 in 5-nL droplets by five cycles of 1000 pl on each spot at a spatial interval of 250 μm. After spotting, MALDI target plates were incubated overnight at 37°C under high-humidity conditions.

Matrix deposition. The matrix solution was prepared by dissolving 50 mg of 2, 5-dihydroxybenzoic acid (DHB; Bruker Daltonics) in 1 mL of 70% methanol/0.1% trifluoroacetic acid. DHB is a widely used matrix for lower molecules including peptides.[18] A thin matrix layer was applied to the surface of the plates using a 0.2-mm nozzle caliber airbrush (Procon Boy FWA Platinum; Mr. Hobby, Tokyo, Japan). The spraying distance was maintained at 15 cm from the tissue surface. The total amount of the matrix solution on each slide was 2–3 mL. The spraying technique enabled full matrix coverage over the entire tissue surface and facilitated co-crystallization of matrix and bio-molecules. A desalting process such as ethanol wash was not performed, since the process does not significantly improve DHB-assisted imaging mass spectrometry (IMS) and NH_4HCO_3 is a highly volatile buffer.[18,19]

Direct analysis of tissue sections by MALDI mass spectrometry. Mass spectra were acquired using the QSTAR XL (Applied Biosystems, Foster City, CA, USA), a hybrid quadrupole/time-of-flight mass spectrometer equipped with an orthogonal MALDI source and a pulsed YAG laser that was operated at a repetition rate of 100 Hz, and a power modulator. Spectra were acquired in positive ion mode. Spectra were acquired in the range of m/z 500–2000. Representative mass spectra were acquired using random laser irradiation-sections. The number of laser shots was 150. An alignment of the mass spectra was performed to compare the datasets using SpecAlign software (http://physchem.ox.ac.uk/~jwong/specalign/). The peak intensity value of the spectra was normalized by dividing them with the total ion current (TIC) as previously described.[20,21]

Imaging of tissue section by mass spectrometry. IMS was performed using orthogonal MALDI (oMALDI) server software by defining a region of interest around the tissue slice. The mechanical resolution, which is the value that refers to the length of the stepwise movement of the laser beam on the sample stage, was 300 μm × 300 μm, and the accumulation time per spot was about 2 s. The acquired mass spectra were visualized using BioMap software (http://www.maldi-msi.org). Molecular images were captured using this software by applying baseline correction to the spectra and integrating these spectra over the peak of interest. Alignment of these mass spectra was performed using SpecAlign software.

Tissue protein identification. The quadrupole ion trap time-of-flight mass spectrometer, namely, AXIMA-QIT (Shimadzu), was used to perform MS/MS analysis. In the MS/MS operation, the data acquisition conditions (i.e. the laser power, collision energy, and the number of laser irradiations) were adjusted to obtain good-quality mass spectra with high intensity and signal-to-noise ratios (S/N) in the fragmented peaks. MS/MS spectra were processed using the Mascot search engine (http://www.matrixscience.com) using the National Center for Biotechnology Information (NCBI)/basic local alignment search tool (BLAST) protein database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) with a taxonomy filter for humans, and the peptide and MS/MS tolerance at 0.3 Da. The search criteria was allowed to consider up to one missed cleavage and variable modifications including protein N-terminus acetylation, histidine/tryptophan oxidation, and methionine oxidation.

Statistical analysis. All statistical analyses were performed with StatView software version 5.0 (SAS Institute, Cary, NC, USA). First, statistical analyses were performed on adjoining cancer and normal tissue. The Student's t-test ($\alpha = 0.05$) was performed between peak intensity means of cancer and normal tissue samples on the basis of equal variance. Welch's test for unequal variance ($\alpha = 0.05$) was performed between peak intensity means of cancer and normal tissues. The corresponding P-value, i.e. P (T t), was reported as a measure of significant statistical variability between conditions.

We extracted signals that showed significantly higher intensity in cancer than in normal tissue in the form of cancer-specific

doi: 10.1111/j.1349-7006.2009.01384.x

peaks. To determine specific peaks related to the degree of differentiation, analysis of variance (ANOVA) and Fisher's protected least significant difference (PLSD) were performed as post-hoc tests on three histological types of three different cancer tissue samples. Fig. 1b shows the statistical workflow.

Immunohistochemical staining for histone H4. For immunohistochemistry (IHC), we used the formalin-fixed tissue microarray specimens containing a wide range of preservation time, up to 30 years after being embedded in paraffin. Each TMA was composed of 50 primary gastric tumors such as adenocarcinoma of various grades of differentiation. Three micrometer sections were cut from the TMA blocks. Immunostaining was performed by the Dako Autostainer System (Dako Japan, Tokyo, Japan) according to the manufacturer-recommended procedure. In brief, paraffin was removed by immersing the TMA slides in xylene for 5 min twice. Subsequently, the tissue sections were rehydrated by immersing the slides in 100% ethanol for 10 min twice, followed by two-time 10-min immersion in 95% ethanol. After rehydration, these slides were incubated at 96°C for 40 min in 10 mM sodium citrate buffer pH 6.0 and then cooled on the bench top for 20 min. Then, these sections were incubated in 3% hydrogen peroxide for 5 min and washed in Tris-buffered saline. A monoclonal antibody against histone H4 (L64C1; Cell Signaling Technology, Danvers, MA, USA) was used as the primary antibody at a dilution of 1:300. N-Histofine® Simple Stain MAX-PO (Multi) (Nichirei Biosciences, Tokyo, Japan) was used as the secondary antibody. After removing the secondary antibody, the sections were exposed to diaminobenzidine for 5 min, and then washed with distilled water. Counterstaining was performed with hematoxylin for 10 s.

IHC evaluation. The IHC evaluation was carried out in two independent ways by two of the authors (Y.M. and H.S.). An evaluation was performed by visual inspection, where IHC staining was classified into four ranks (0, negative; 1, slightly positive; 2, positive; 3, strongly positive) for common types of gastric carcinoma. Papillary adenocarcinoma was interpreted as well-differentiated adenocarcinoma and signet-ring cell carcinoma as poorly differentiated adenocarcinoma. Assignment of mucinous carcinoma category was made according to the other predominant elements. Special types of gastric carcinoma and other tumors were excluded. In total, 169 specimens were evaluated; they included 42 well-differentiated, 38 moderately differentiated, and 89 poorly differentiated adenocarcinomas. As described,[22] Steel–Dwass' test was performed by using free software available on a web site, MEPHAS (http://www.gen-info.osaka-u.ac.jp/testdocs/tomocom/). Another evaluation was carried out by quantifying the signal intensity of IHC staining with Scion image software version 4.0.3.2 (Scion, Frederick, MD, USA). We analyzed additional 170 specimens containing 43 well-differentiated, 40 moderately differentiated, and 87 poorly differentiated adenocarcinomas. The data were represented as the mean value of intensity ± SD. ANOVA and Fisher's PSLD were performed as post-hoc test among three histological types (well, moderately, and poorly differentiated).

Results

Experimental paradigm design. To statistically detect cancer-specific signals, we placed three cancer tissues and one normal tissue from three patients in a TMA. The histological type of cancer differed among the three patients (Fig. 1a). The sample from Patient 1 was moderately differentiated adenocarcinoma, that from Patient 2 was well-differentiated adenocarcinoma, and that from Patient 3 was poorly differentiated adenocarcinoma. We acquired tissue samples from three different regions of each patient for further statistical analysis to detect signals specific to cancer-differentiation status (Fig. 1a). During the first screening, we compared signals of three individual cancers with those of normal tissues (Fig. 1b). Thereafter, we screened the detected cancer-specific peaks by multiple comparisons of three different cancer regions belonging to three histological types (Fig. 1b). We categorize the obtained results on the basis of the statistical workflow in Fig. 1c.

Acquisition of mass spectra from FFPE-TMA samples. A previous report showed relatively weak signal intensities and low S/N ratio was obtained with FFPE tissue samples compared to freshly frozen ones.[23] We first examined if peptide signals could be sufficiently detected with FFPE-TMA. We employed chemical inkjet technology to equalize the quantity and the interval of trypsin solution application.[14] We detected vast quantities of signals that were sufficient to generate imaging data from FFPE-TMA samples. Figure 2 shows representative spectra obtained from three individual cancer tissues and normal tissues. The peaks obtained were mainly concentrated below m/z 2000 and could hardly be detected over m/z 2000. Thus, using SpecAlign, we performed signal-intensity normalization.

Detection of cancer specifically increased signals in IMS of digested FFPE tissue microarray. Subsequently, we performed IMS of FFPE-TMA on the samples and obtained mass spectra. We setup a spatial interval of 300 μm to prevent repeated laser irradiation, as the irradiated laser diameter was 200 μm. We completed the scanning of the TMA samples with 12 spots in approximately 1 h. We detected a total of 72 signals with FFPE-TMA samples. Fig. S1 shows the obtained array images. To perform statistical data analysis, we quantified the signal intensities of m/z peaks. The first statistical screening (Fig. 1b) revealed 54 signals, the intensity of which was detected to be significantly increased in cancer tissues (Fig. 3). We examined the reliability of this screening by performing two independent trials with sibling arrays. Forty of the 54 signals were detected in the two independent trials (Fig. 3). Figure 4(a) shows the representative array result of the signal significantly increased in cancer. In contrast, Fig. 4b shows that another signal has no significant difference between cancer and normal tissues, i.e. it shows an even distribution pattern.

Detection of histological type-specific increased signals in IMS of digested FFPE tissue microarray. We further analyzed the quantified signal intensities to examine whether such histological type-specific signals could be detected with our experimental paradigm. To this end, we compared signals detected in cancer tissues among the well-differentiated, moderately differentiated, and poorly differentiated tissues (Fig. 1b). To detect specific signals, we conducted statistical analyses with one-way ANOVA followed by Fisher's PLSD. Certain signals demonstrating a cancer-specific pattern appeared to demonstrate uneven signal intensities in different degrees of cancer differentiation (Fig. 3, Fig. 5a–c). Of the detected signals, peaks having m/z 537.2, 1168.4, 1387.6, 1475.8 were reproducibly detected in another experimental trial. Other signals were detected only once in two experimental trials. Owing to the two-step screening, it is probable that the detection of histological type-specific signals shows worse reproducibility than the simple detection of cancer-specific signals.

Identification of protein-specific increase in poorly differentiated cancer tissues by MS/MS analysis. We attempted to identify the signals that were specifically detected in poorly differentiated cancer tissue. We performed MS/MS analyses on the FFPE-TMA, and analyzed the resultant data with the Mascot search engine. We identified one signal with an m/z 1325.6 as histone H4 that is specific to poorly differentiated cancers (Fig. 6a), and identified a protein that demonstrated non-specific expression and had an m/z 976.4 corresponding to that of actin (Fig. 6b). Other peaks could not be detected due to their weak intensity.

IHC staining for histone H4 using another TMA specimen of larger numbers of the cases. Finally, we examined whether

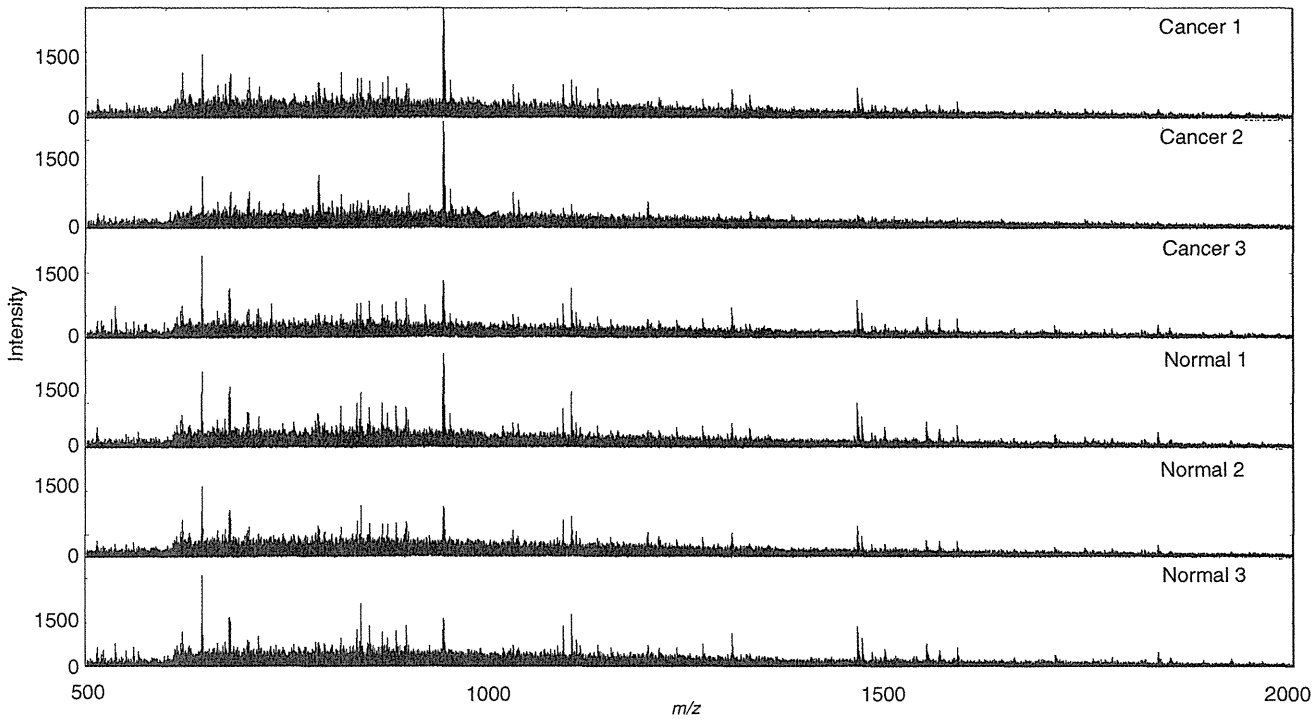

Fig. 2. Acquired mass spectra from formalin-fixed paraffin-embedded (FFPE) tissue microarray samples (TMA) by random laser irradiation. Acquired mass spectra from each TMA spot of adjoining cancer and normal tissue are shown as representative spectra. *m/z* refers to mass per charge ratio.

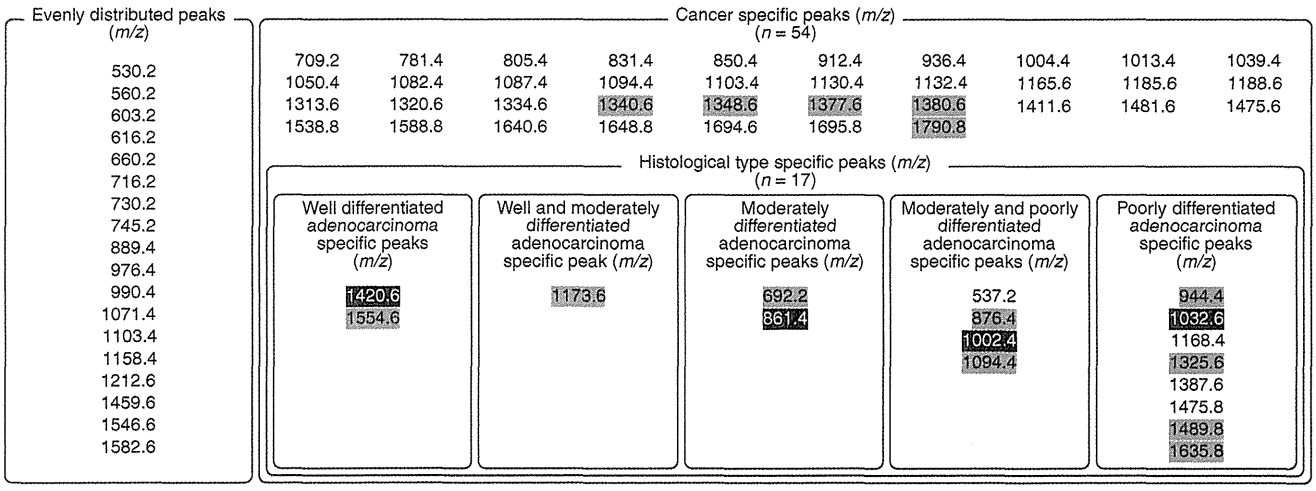

Fig. 3. The peak list acquired from mass spectra. The shaded values represent signals that showed significantly increased intensity in cancer tissues in one independent trial ($P < 0.05$). The white-on-black values represent signals that showed significant difference among three cancers in one independent trial ($P < 0.05$).

histone H4 was specifically strongly detected in poorly differentiated cancers. In total, we stained 400 gastric tumors including adenocarcinoma, squamous carcinoma, neuroendocrine carcinoma, metastatic carcinoma, malignant lymphoma, and adenoma. We excluded 61 specimens such as a special type of gastric carcinoma, malignant lymphoma, and benign lesion. We examined 339 gastric carcinomas composed of 85 well-differentiated carcinomas, 78 moderately differentiated carcinomas, and 176 poorly differentiated carcinomas. Figure 7(a) shows representative photomicrographs of each histological type. We evaluated the result of IHC in two approaches. First, we determined the staining appearance according to four ranks (0, negative; 1, slightly positive; 2, positive; 3, strongly positive). Slightly positive and positive staining reached a high rate in well- and moderately differentiated carcinoma. In contrast, poorly differentiated adenocarcinomas were categorized into much more positive staining such as rank 2 or rank 3 (Fig. 7b). Second, we also performed more quantitative analysis. We

doi: 10.1111/j.1349-7006.2009.01384.x

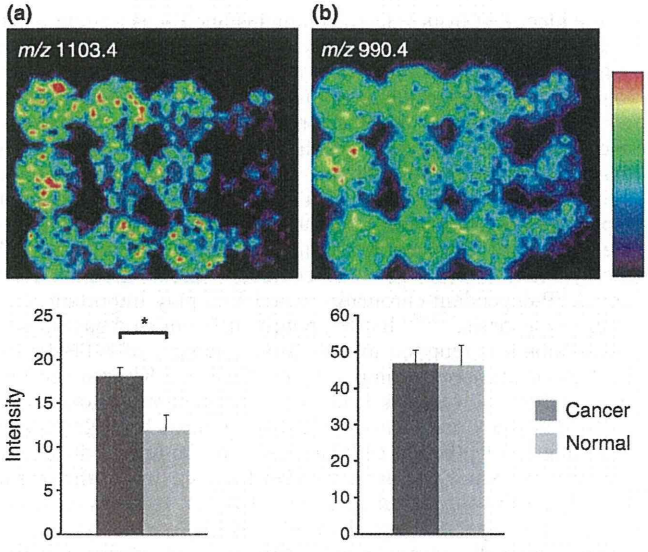

Fig. 4. Cancer-specific signal increase and even distribution of signals in imaging mass spectrometry (IMS) of digested formalin-fixed paraffin-embedded (FFPE) tissue microarrays (TMA). (a) Significantly strong peak intensity was detected at m/z 1103.4. (b) No significant difference was observed between cancer and normal tissues at an m/z of 990.4. Values are represented as mean ± SD (n = 3). *P < 0.05.

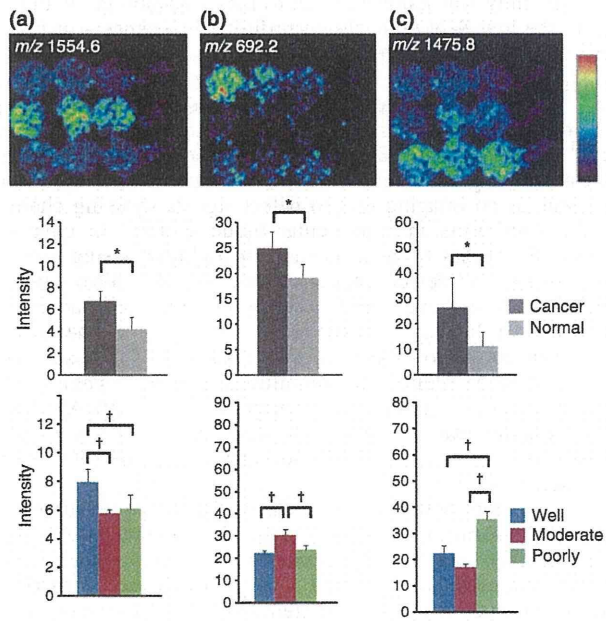

Fig. 5. Histological type-specific signal increase in imaging mass spectrometry (IMS) of digested formalin-fixed paraffin-embedded (FFPE) tissue microarrays (TMA). (a) Ion imaging revealed a peak with significantly strong signal intensity in well-differentiated adenocarcinoma at an m/z of 1554.6. (b) Ion imaging revealed a peak with significantly strong signal intensity in moderately differentiated adenocarcinoma at an m/z of 692.2. (c) Ion imaging revealed a peak with significantly strong signal intensity in poorly differentiated adenocarcinoma at m/z 1475.8. Values are represented as mean ± SD (n = 3). *P < 0.05, †P < 0.05.

quantified the signal intensity using Scion image software. In agreement with our visual inspection, poorly differentiated carcinoma showed significantly higher value compared to well-

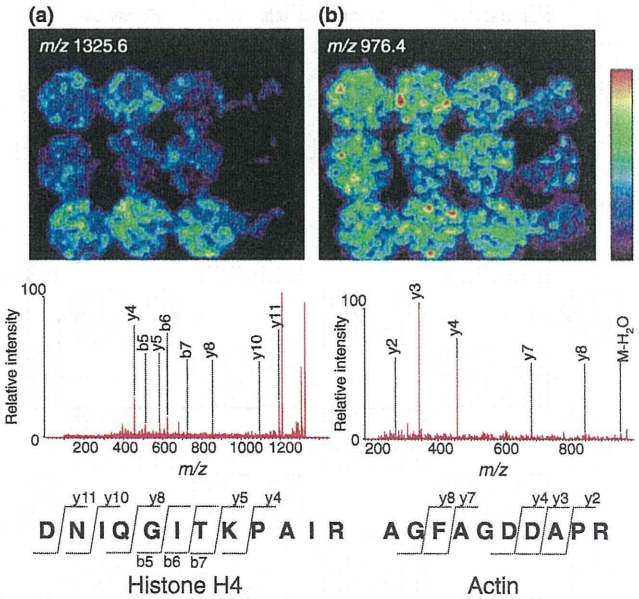

Fig. 6. MS/MS analysis of digested peptide and protein identification. (a) The biomolecule of an m/z 1325.6 was identified as histone H4. DNIQGITKPAIR: abbreviation of the amino-acid sequence aspartic acid–asparagine–isoleucine–glutamine–glycine–isoleucine–threonine–lysine–proline–alanine–isoleucine–arginine. y4, y5, y8, y10, y11 represent each fragment ion, which includes the C-terminal domain. b5, b6, b7 represent each fragment ion, which includes the N-terminal domain. (b) The biomolecule with an m/z 976.4 was identified as actin. AGFAGDDAPR: abbreviation for the amino-acid sequence alanine–glycine–phenylalanine–alanine–glycine–aspartic acid–aspartic acid–alanine–proline–arginine. y2, y3, y4, y7, y8 refer to each fragment ion, which includes the C-terminal domain.

differentiated or moderately differentiated carcinomas. There was no significant difference between well- and moderately differentiated carcinoma (Fig. 7c).

Discussion

In this report, we presented a simple and easy-to-use method for the detection and identification of cancer-specific proteins, i.e. strong candidates for biomarkers or drug targets, with high reliability in an experimental trial (Fig. 1). We succeeded in detecting cancer-specific signals with 75% (40 per 54) reliability in two independent experiments (Fig. 3, Fig. 4). Furthermore, we detected signals that were specific for each status of cancer differentiation (Fig. 5). Finally, we successfully identified one of the signals that was specifically increased in the poorly differentiated cancer tissue as histone H4 (Fig. 6).

We analyzed 12 different tissue samples within 1 h. The TMA-IMS technique has prominent advantages when compared to existing proteomic techniques. This technique enables the analysis of multiple proteins in multiple tissue samples in just one experiment. The existing proteomic techniques lack either multiprocessing property with respect to the analysis of samples or detection of proteins. Proteomic techniques employing 2D electrophoresis-MS or protein microarrays can analyze only single or double samples in one experiment; however, they can detect and identify multiple proteins in one experiment. In contrast, the limitation of TMA is that it enables that analysis of only one or two proteins in one experiment; however, it enables simultaneous analysis of multiple tissue samples. Thus, the TMA-IMS technique has two advantages compared to existing techniques. Moreover, the TMA-IMS technique does not require

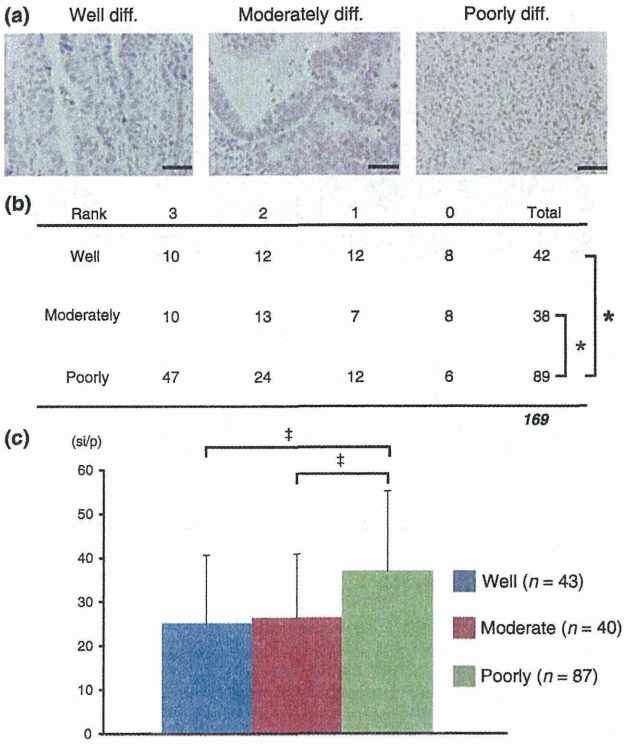

Fig. 7. Immunohistochemical (IHC) staining for histone H4. (a) Representative photomicrograph of IHC for histone H4 protein, ×400. Scale bar, 50 μm. (b) Evaluation of IHC according to four ranks (0, negative; 1, slightly positive; 2, positive; 3, strongly positive). The 169 gastric carcinomas comprised 42 well-differentiated carcinomas, 38 moderately differentiated carcinomas, 89 poorly differentiated carcinomas. *$P < 0.05$ by Steel–Dwass' test. (c) Quantitative analysis of IHC signal intensity. The 170 gastric carcinomas comprised 43 well-differentiated carcinomas, 40 moderately differentiated carcinomas, and 87 poorly differentiated carcinomas. Values are represented as mean ± SD. si/p, signal intensity per pixel. ‡$P < 0.01$.

the complicated sample-preparation steps which are required in 2-DE-MS-based proteomics.

The high-intensity signals detected from cancer tissues account for approximately two-thirds (54 of 72) of all signals detected. We failed to detect signals specific to normal tissues. This could be explained by the heterogeneity underlying normal tissues. The adjacent normal tissues consisted of varied types of tissues, such as mucosa, fatty tissue, and muscle. Acquired signal intensities in normal tissues were the average of whole spectrum derived from each of the tissues. Hence, tissue type-specific peaks were totally obscured, resulting in the failure to detect adjacent normal tissue-specific peaks.

We detected 17 histological type-specific peaks in two independent experiments. (Fig. 3) The reproducibility of signal detection from normal tissues was lower (4 per 17; 24%) than that of signal-detection from cancer tissues (75%). This could be explained by our experimental paradigm under which we performed two-step screening. The high severity of the first screening step excluded 9 of 13 peaks in either trial due to high variance among cancer tissues. In this technique, the two-step screening is essential as we cannot rule out the possibility of detecting false positive signals that are specific merely for certain patients. This problem could be addressed by spotting more samples under a variety of cancer tissue conditions in a TMA and performing multiple direct comparisons.

We identified from IMS screening histone H4 as a protein that is specifically increased in poorly differentiated adenocarcinoma (Fig. 6a). To validate the IMS result, we performed IHC for histone H4 protein using large amount of archival TMA specimens composed of various cellular density. In both visual inspection and quantitative analysis of IHC, histone H4 was strongly detected in poorly differentiated carcinoma (Fig. 7b,c). Similar strong detection of histone H4 in a cancer tissue has been reported by a recent study on a mouse model of brain tumor analyzed by IMS.[24] Dynamic chromatin remodeling such as DNA methylation, histone variants, covalent histone modifications, and ATP-dependent chromatin remodeling play important roles in carcinogenesis.[25,26] Indeed, poorly differentiated gastric adenocarcinoma is reported to lose Brm, a subunit of ATP-dependent chromatin-remodeling complex.[27] It can be assumed that these epigenetic changes lead to the chromatin-unfolding state and allow ready access to core histone protein. It might be also plausible that the higher cellular density of the poorly differentiated cancer tissues compared to other tissues explains the reason underlying the successful detection of a vast majority of histological type-specific signals from IMS. Conversely, signals specific to well- or moderately differentiated adenocarcinoma reflect the more-significant changes among the three histological types.

IMS technique was originally applied for the analysis of frozen tissue sections. FFPE samples are unsuitable for performing IMS due to the presence of cross-linkage between proteins and the inefficiency of enzyme digestion. Thus, few studies have reported the performance of IMS on FFPE samples.[28–30] Further, our study was hampered due to the disadvantage of FFPE. Due to the low S/N ratio, the identification of cancer- or histological type-specific proteins was rendered difficult.

TMA was originally used for IHC or *in situ* hybridization.[11,31] While this study was being conducted, another group reported the IMS of lung tumor biopsy in FFPE-TMA samples.[32] Thus, the FFPE-TMA-IMS has now emerged as the newest imaging technique. Other researchers have used this technique as an imaging tool to detect signals showing characteristic distribution in a particular tissue spot.[32] In contrast, we used the technique as a scanner for multiprocessing proteomics to readily detect cancer-specific signals. Two studies along these lines have reported highly different approaches. In this work, we loaded 12 tissue spots with 3-mm diameter in a TMA. Due to improvements in IMS resolution, the tissue spot size can now be reduced to submillimeter scales, enabling the loading of hundreds of tissue samples in one TMA. Thus, this technique can be applied for the analysis of a greater number of samples for high-throughput analysis of cancer characteristics.

Once a patient develops cancer, he or she should be subjected to medical treatments, including surgical operation and/or chemotherapy. To enable early detection of cancers, specific and sensitive biomarkers are desired. Using 2-DE based on MALDI mass spectrometry, potential proteins related to carcinogenesis have been discovered.[33,34] Further, in cases of far advanced or recurrent gastric carcinoma, chemotherapy prolongs the survival of the patient at a certain rate.[35] Proteomic analysis yielded an antidrug resistance agent,[36] while multidrug resistances were observed in certain cases. For retrospective evaluation and prospective searches of chemotherapy-related markers, high-throughput pathological evaluation methods are essential. IMS using TMA in FFPE, as we show here, will be one of the most promising gadgets in the surgical pathology laboratory.

In conclusion, we performed IMS of FFPE-TMA samples of gastric carcinoma, and successfully identified histone H4 as a signal specific to poorly differentiated cancer tissues. Moreover, the IMS-based finding was confirmed by IHC analyses of a large amount of TMAs. IMS of FFPE samples is a currently emerging

doi: 10.1111/j.1349-7006.2009.01384.x

technique and our experience represents an important step in the early phase of development. IMS of FFPE-TMA can offer fast and easy screening of cancer or tissue type-specific signals from a large amount of samples. The results of IMS-based screening can be readily verified by IHC analysis with other sets of FFPE-TMA samples. Combined with IHC confirmation, IMS of FFPE-TMA samples may be a further powerful tool in cancer proteomics.

Acknowledgments

This work was supported by a SENTAN step-up grant from JST to M.S. and by Grants-in-Aid from the Ministry of Health, Labour and Welfare for the Comprehensive 10-Year Strategy for Cancer Control (19-19) and the Third Term Comprehensive Control Research for Cancer; the Japan Society for the Promotion of Science for Scientific Research (no. 19790286), and the Smoking Research Foundation.

References

1 DeRisi J, Penland L, Brown PO *et al.* Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat Genet* 1996 Dec; **14** (4): 457–60.
2 Ramsay G. DNA chips: state-of-the art. *Nat Biotechnol* 1998 Jan; **16** (1): 40–4.
3 Zhao X, Li C, Paez JG *et al.* An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. *Cancer Res* 2004 May 1; **64** (9): 3060–71.
4 Engle LJ, Simpson CL, Landers JE. Using high-throughput SNP technologies to study cancer. *Oncogene* 2006 Mar 13; **25** (11): 1594–601.
5 McLendon R, Friedman A, Bigner D *et al.* Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* 2008 Oct 23; **455** (7216): 1061–8.
6 Ley TJ, Mardis ER, Ding L *et al.* DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. *Nature* 2008 Nov 6; **456** (7218): 66–72.
7 Hood L, Heath JR, Phelps ME, Lin B. Systems biology and new technologies enable predictive and preventative medicine. *Science* 2004 Oct 22; **306** (5696): 640–3.
8 Aebersold R, Mann M. Mass spectrometry-based proteomics. *Nature* 2003 Mar 13; **422** (6928): 198–207.
9 Diamandis EP. Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. *Mol Cell Proteomics* 2004 Apr; **3** (4): 367–78.
10 Sreekumar A, Nyati MK, Varambally S *et al.* Profiling of cancer cells using protein microarrays: discovery of novel radiation-regulated proteins. *Cancer Res* 2001 Oct 15; **61** (20): 7585–93.
11 Kononen J, Bubendorf L, Kallioniemi A *et al.* Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med* 1998 Jul; **4** (7): 844–7.
12 Stoeckli M, Chaurand P, Hallahan DE, Caprioli RM. Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat Med* 2001 Apr; **7** (4): 493–6.
13 Shimma S, Sugiura Y, Hayasaka T, Zaima N, Matsumoto M, Setou M. Mass imaging and identification of biomolecules with MALDI-QIT-TOF-based system. *Anal Chem* 2008 Feb 1; **80** (3): 878–85.
14 Shimma S, Furuta M, Ichimura K, Yoshida Y, Setou M. Direct MS/MS analysis in mammalian tissue sections using MALDI-QIT-TOFMS and chemical inkjet technology. *Surf Interface Anal* 2006; **38** (2): 1712–14.
15 Groseclose MR, Andersson M, Hardesty WM, Caprioli RM. Identification of proteins directly from tissue: *in situ* tryptic digestions coupled with imaging mass spectrometry. *J Mass Spectrom* 2007 Feb; **42** (2): 254–62.
16 Crew KD, Neugut AI. Epidemiology of gastric cancer. *World J Gastroenterol* 2006 Jan 21; **12** (3): 354–62.
17 Japanese Gastric Cancer A. Japanese classification of gastric carcinoma – 2nd English edn. *Gastric Cancer* 1998 Dec; **1** (1): 10–24.
18 Schwartz SA, Reyzer ML, Caprioli RM. Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation. *J Mass Spectrom* 2003 Jul; **38** (7): 699–708.
19 Espada A, Rivera-Sagredo A. Ammonium hydrogencarbonate, an excellent buffer for the analysis of basic drugs by liquid chromatography-mass spectrometry at high pH. *J Chromatogr A* 2003 Feb 14; **987** (1–2): 211–20.
20 Wong JW, Cagney G, Cartwright HM. SpecAlign – processing and alignment of mass spectra datasets. *Bioinformatics* 2005 May 1; **21** (9): 2088–90.
21 Whistler T, Rollin D, Vernon SD. A method for improving SELDI-TOF mass spectrometry data quality. *Proteome Sci* 2007; **5**: 14.
22 Yanaga Y, Awai K, Nakaura T *et al.* Optimal contrast dose for depiction of hypervascular hepatocellular carcinoma at dynamic CT using 64-MDCT. *AJR Am J Roentgenol* 2008 Apr; **190** (4): 1003–9.
23 Lemaire R, Desmons A, Tabet JC, Day R, Salzet M, Fournier I. Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *J Proteome Res* 2007 Apr; **6** (4): 1295–305.
24 Seeley EH, Caprioli RM. Molecular imaging of proteins in tissues by mass spectrometry. *Proc Natl Acad Sci USA* 2008 Nov 25; **105** (47): 18126–31.
25 Wang GG, Allis CD, Chi P. Chromatin remodeling and cancer, part I: covalent histone modifications. *Trends Mol Med* 2007 Sep; **13** (9): 363–72.
26 Wang GG, Allis CD, Chi P. Chromatin remodeling and cancer, part II: ATP-dependent chromatin remodeling. *Trends Mol Med* 2007 Sep; **13** (9): 373–80.
27 Yamamichi N, Inada K, Ichinose M *et al.* Frequent loss of Brm expression in gastric cancer correlates with histologic features and differentiation state. *Cancer Res* 2007 Nov 15; **67** (22): 10727–35.
28 Aoki Y, Toyama A, Shimada T *et al.* A novel method for analyzing formalin-fixed paraffin embedded (FFPE) tissue sections by mass spectrometry imaging. *Proc Jpn Acad Ser B* 2007; **83** (2): 205–14.
29 Stauber J, Lemaire R, Franck J *et al.* MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. *J Proteome Res* 2008 Mar; **7** (3): 969–78.
30 Ronci M, Bonanno E, Colantoni A *et al.* Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: application to MALDI-TOF imaging MS investigations. *Proteomics* 2008 Sep; **8** (18): 3702–14.
31 Sugimura H. Detection of chromosome changes in pathology archives: an application of microwave-assisted fluorescence *in situ* hybridization to human carcinogenesis studies. *Carcinogenesis* 2008 Apr; **29** (4): 681–7.
32 Groseclose MR, Massion PP, Chaurand P, Caprioli RM. High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. *Proteomics* 2008 Sep; **8** (18): 3715–24.
33 Yoshihara T, Kadota Y, Yoshimura Y *et al.* Proteomic alteration in gastic adenocarcinomas from Japanese patients. *Mol Cancer* 2006; **5**: 75.
34 Cheng Y, Zhang J, Li Y, Wang Y, Gong J Proteome analysis of human gastric cardia adenocarcinoma by laser capture microdissection. *BMC Cancer* 2007; **7**: 191.
35 Ohtsu A. Chemotherapy for metastatic gastric cancer: past, present, and future. *J Gastroenterol* 2008; **43** (4): 256–64.
36 Wang X, Lu Y, Yang J *et al.* Identification of triosephosphate isomerase as an anti-drug resistance agent in human gastric cancer cells using functional proteomic analysis. *J Cancer Res Clin Oncol* 2008 Sep; **134** (9): 995–1003.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Fig. S1. The array images obtained with imaging mass spectrometry (IMS) are shown. There are 18 evenly distributed images and 37 cancer specific ones, two well-differentiated adenocarcinoma-specific images, one well- and moderately differentiated adenocarcinoma-specific image, two moderately differentiated adenocarcinoma-specific images, four moderately and poorly differentiated adenocarcinoma-specific images, and eight poorly differentiated adenocarcinoma-specific images. Values are represented as mean ± SD ($n = 3$). $^*P < 0.05$. $^\dagger P < 0.05$.

WJGO World Journal of Gastrointestinal Oncology

Online Submissions: http://www.wjgnet.com/1948-5204office
wjgo@wjgnet.com
doi:10.4251/wjgo.v2.i2.117

World J Gastrointest Oncol 2010 February 15; 2(2): 117-120
ISSN 1948-5204 (online)

BRIEF ARTICLE

Three novel *NEIL1* promoter polymorphisms in gastric cancer patients

Masanori Goto, Kazuya Shinmura, Hong Tao, Shoichiro Tsugane, Haruhiko Sugimura

Masanori Goto, Kazuya Shinmura, Hong Tao, Haruhiko Sugimura, First Department of Pathology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi Ward, Hamamatsu, Shizuoka 431-3192, Japan
Shoichiro Tsugane, Epidemiology and Prevention Division, Research Center for Cancer Prevention and Screening, National Cancer Center, Tokyo 104-0045, Japan
Author contributions: Goto M performed the majority of the experiments; Shinmura K and Sugimura H designed the study and wrote the manuscript; Tao H performed the statistical analysis; Tsugane S coordinated the collection of all of the human materials and provided them.
Supported by Grants-in-Aid from Ministry of Health, Labour and Welfare for the Comprehensive 10-Year Strategy for Cancer Control (19-19); Japan Society for the Promotion of Science for Scientific Research, No. 19790286; Ministry of Education, Culture, Sports, Science and Technology for priority area, No. 20014007; and the 21st century COE program
Correspondence to: Haruhiko Sugimura, PhD, First Department of Pathology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi Ward, Hamamatsu, Shizuoka 431-3192, Japan. hsugimur@hama-med.ac.jp
Telephone: +81-53-4352220 Fax: +81-53-4352225
Received: March 14, 2009 Revised: August 5, 2009
Accepted: August 12, 2009
Published online: February 15, 2010

Abstract

AIM: To identify genetic polymorphisms in the promoter region of the human base excision repair gene *NEIL1* in gastric cancer patients.

METHODS: The *NEIL1* promoter region in DNA from 80 Japanese patients with gastric cancer was searched for genetic polymorphisms by polymerase chain reaction-single-strand conformation polymorphism and subsequent sequencing analyses.

RESULTS: Three novel genetic polymorphisms, i.e. c.-3769C>T, c.-3170T>G, and c.-2681TA[8], were identified in the *NEIL1* promoter region at an allele frequency of 0.6%, 9.4%, and 4.4%, respectively, in Japanese gastric cancer patients.

CONCLUSION: Three *NEIL1* promoter polymorphisms detected in this study may be of importance in gastric carcinogenesis.

Key words: Gastric cancer; *NEIL1*; Base excision repair; Genetic polymorphism

Peer reviewer: Tatsuo Kanda, MD, PhD, Division of Digestive and General Surgery, Graduate School of Medical and Dental Sciences, Niigata University, Niigata City 951-8510, Japan

Goto M, Shinmura K, Tao H, Tsugane S, Sugimura H. Three novel *NEIL1* promoter polymorphisms in gastric cancer patients. *World J Gastrointest Oncol* 2010; 2(2): 117-120 Available from: URL: http://www.wjgnet.com/1948-5204/full/v2/i2/117.htm DOI: http://dx.doi.org/10.4251/wjgo.v2.i2.117

INTRODUCTION

Stomach tissue is exposed to oxidative stress, including inflammation induced by *Helicobacter pylori* infection, sodium chloride, and smoking[1-5]. Since severe oxidative stress leads to accumulation of huge amounts of damaged bases[6-9], maintenance of a system to repair damaged bases in the stomach is thought to be important. The base excision repair protein NEIL1 has activity that is capable of removing oxidatively damaged bases, including thymine glycol, 5-hydroxyuracil, urea, formamidopyrimidine-A, and formamidopyrimidine-G, which have been shown to cause mutagenesis and cell death[10-14]. We have recently demonstrated somatic inactivating *NEIL1* mutations and reduced NEIL1 expression in a subset of gastric cancers, suggesting that reduced NEIL1 activity is involved in gastric carcinogenesis[15]. In a recent investigation of the NEIL1 expression system

an approximately 1.2 kb sequence upstream of the transcriptional initiation site of the *NEIL1* gene was shown to have promoter activity by a luciferase reporter assay in human cells[16]. However, since no genetic polymorphisms have been reported in the promoter region thus far and genetic polymorphisms in the region may be of importance in gastric carcinogenesis, we tried searching DNA extracted from the blood of 80 gastric cancer patients for *NEIL1* promoter polymorphisms.

MATERIALS AND METHODS

Samples

Blood samples from 80 gastric cancer patients were obtained from hospitals in Nagano Prefecture, Japan, and genomic DNA was extracted from them with a DNA Extractor WB Kit (Wako, Osaka, Japan)[17]. The baseline characteristics of the patients have been described previously[17]. This study was approved by the Institutional Review Boards of Hamamatsu University School of Medicine and the National Cancer Center.

Polymerase chain reaction (PCR)-single-strand conformation polymorphism (SSCP) and sequencing analyses

PCR-SSCP analysis was used to examine the DNA samples for genetic polymorphisms in the *NEIL1* promoter region. An approximately 1.2 kb 5' upstream sequence that was shown to have promoter activity in a previous study[16] was divided into 8 regions (Figure 1A), and each region was amplified by PCR with HotStarTaq DNA polymerase (QIAGEN, Valencia, CA, USA). The primer sets used were: 5'-CAAATATTGCAGTCTGA AAGGGG-3' and 5'-GAAACTGATCAAGACAGGG GC-3' for region 1, 5'-GTTTTCTAATGCAGAGGTC TGG-3' and 5'-TACAGGGATAAGCCACTA CGC-3' for region 2, 5'-CCTCCTGATATGATGCAATTC-3' and 5'-CACTCCCAGCTGATTTTTG-3' for region 3, 5'-ATGGTGAAACCCCGTCTCTAC-3' and 5'-T GCTGGGAATTAGATCTAAAGGC-3' for region 4, 5'-AGCACCTAGGAAGTATCCCTG-3' and 5'-GTCTC AGCCAGTTGTTGTTTG-3' for region 5, 5'-CAAATT GAGAATGTGATGCAGC-3' and 5'-CAGATTTCCCC AATTGTCCC-3' for region 6, 5'-TGACCCATGATTG TAGCCTG-3' and 5'-GAGGTTTCGCCT TGTTGG-3' for region 7, and 5'-GAGGCGGGCAGATTACTT G-3' and 5'-CTCACTGCAGCC TCCACTTC-3' for region 8. The PCR products of regions 4 and 6 were digested with restriction enzymes *Mva*I (New England Biolabs, Beverly, MA, USA) and *Ava*I (New England Biolabs), respectively, in order to adjust their size to < 230 bp before SSCP. The PCR products of all regions were diluted with two volumes of loading solution, and after applying them to 8% polyacrylamide gels in the presence or absence of 5% glycerol, the products were electrophoresed at room temperature and 4℃ and detected by silver staining. PCR products exhibiting an abnormally shifted band in the SSCP analysis were directly sequenced with a BigDye Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Tokyo, Japan) and an ABI 3100 Genetic Analyzer (Applied Biosystems). A PCR product of region 7 was also sequenced after subcloning into a pGEM-T Easy vector (Promega, Madison, WI, USA). The reference nucleotide sequence is accession number NM_024608. Deviation of the genotype distribution from Hardy-Weinberg equilibrium (HWE) was tested by using SNPAlyze software (Dynacom, Yokohama, Japan).

RESULTS

We searched for genetic polymorphisms in the region containing *NEIL1* promoter activity by PCR-SSCP analysis using blood samples derived from 80 gastric cancer patients. Three genetic polymorphisms, c.-3769C>T, c.-3170T>G, and c.-2681TA[8], were identified in the *NEIL1* promoter region at an allele frequency of 0.6%, 9.4%, and 4.4%, respectively (Figure 1B). The distribution of the genotypes of these polymorphisms was in HWE. Examination of the frequency of the polymorphisms revealed a homozygote for the variant allele of only one of the three polymorphisms, c.-3170T>G, and in only one patient, indicating that the three polymorphisms in the *NEIL1* promoter are rare genetic polymorphisms.

DISCUSSION

In this study, we found three novel promoter polymorphisms, c.-3769C>T, c.-3170T>G, and c.-2681TA[8]. None of these polymorphisms has previously been reported or registered in the database of the single nucleotide polymorphism (dbSNP) homepage of the National Center for Biotechnology Information web site (web site : http://www.ncbi.nlm.nih.gov/SNP/) or the database of the Japanese single nucleotide polymorphism homepage (http://snp.ims.u-tokyo.ac.jp/), indicating that they are novel genetic polymorphisms.

Interestingly, when we used Genomatix software (http://www.genomatix.de/matinspector.html) to search for transcription factors that putatively bind to the sequence containing these polymorphism sites, a sequence containing c.-3170T was predicted to bind to GATA binding factors and a sequence containing c.-2681TA[7] was predicted to bind to a GZF1, a TATA-binding protein and LIM homeodomain factors. The change from c.-3170T to c.-3170G eliminates the binding site for GATA binding factors. On the other hand, although the change from c.-2681TA[7] to c.-2681TA[8] would appear to retain the sequence of binding sites for the GZF1, TATA-binding protein, and LIM homeodomain factors, there are examples of a change in the number of repetitive sequences in a promoter being associated with a difference in the expression level[18]. Thus, these nucleotide changes may be associated with a difference in the NEIL1 expression level. Moreover, since some factors involved in the regulation of the transcription level in human cells remain unknown,

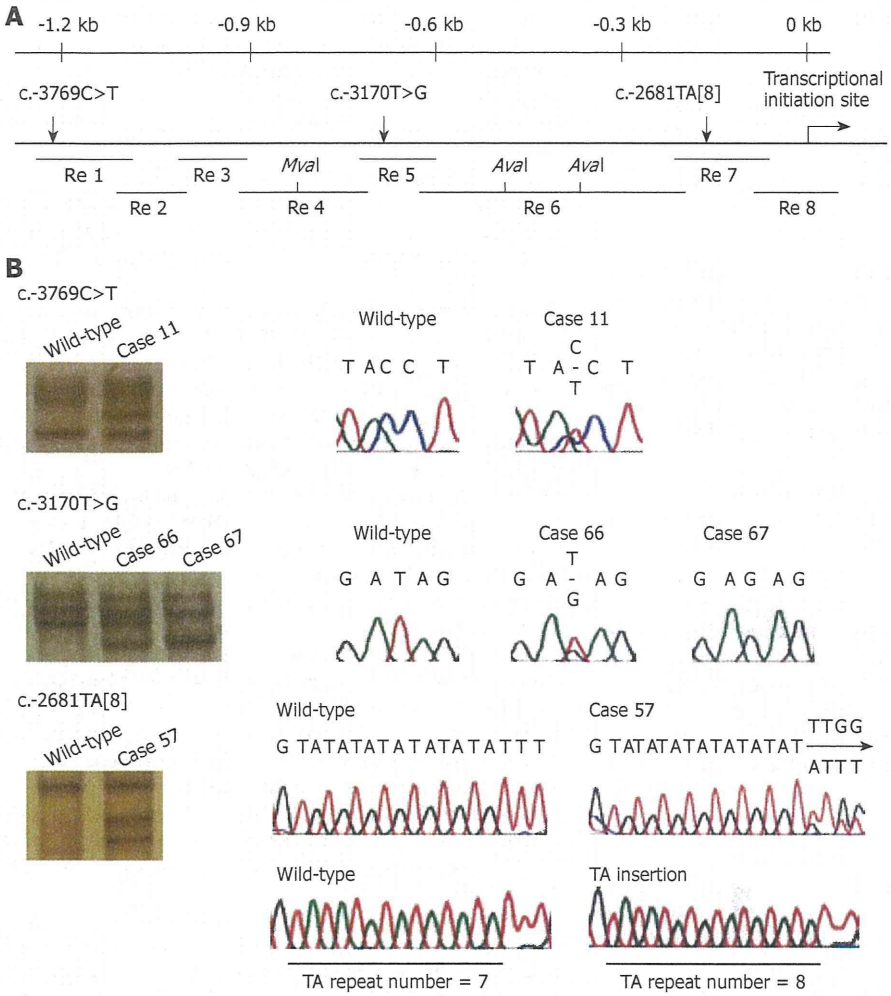

Figure 1 Identification of novel *NEIL1* promoter polymorphisms by PCR-SSCP and subsequent sequencing analyses. A: Schematic map of the *NEIL1* promoter region. The three genetic polymorphisms identified in this study are mapped, and the 8 PCR-amplified regions (Re 1-Re 8) and restriction enzyme sites are shown; B: Identification of c.-3769C>T, c.-3170T>G, and c.-2681TA[8] polymorphisms in the *NEIL1* promoter. The panels on the left show representative results of the PCR-SSCP analysis. The electropherograms on the right show the results of the sequencing analysis. Only the lowermost panels show the results of sequencing after subcloning the PCR product; the others show the results of direct sequencing.

the three *NEIL1* promoter polymorphisms may be associated with differences in the NEIL1 expression level by binding to factors that have yet to be identified.

Most common genetic polymorphisms have been registered in various genetic polymorphism databases, such as dbSNP. However, as shown in this study, there appear to be many genetic polymorphisms that still have not been registered in databases, the reason being that many unregistered genetic polymorphisms are rare. Since finding rare and novel polymorphisms requires many human samples and repeating this kind of study, the *NEIL1* data presented in this study are very valuable for future studies, such as searches for alleles that increase the risk of diseases and allele-specific expression analyses. Furthermore, NEIL1 protein plays a very important role in excision repair of oxidatively damaged bases, which have been implicated in a wide variety of human cancer. Promoter polymorphisms in some DNA repair genes, including *XRCC1*, *MLH1*, and *MSH2*, have recently been reported to be associated with an increased risk of cancer[19-21]. Like these examples, the novel *NEIL1* promoter polymorphisms identified by screening gastric cancer patients in this study may be associated with increased risk of gastric cancer. If so, this information should be of value in management to prevent the development of gastric cancer in individuals with the risk allele. We are therefore planning to examine the *NEIL1* promoter polymorphisms in the framework of a gastric cancer case-control study in the future, and we are also planning to investigate the effect of the polymorphisms on *NEIL1* promoter activity.

ACKNOWLEDGMENTS

The authors gratefully acknowledge Dr. Shusuke Natsukawa (Saku Cent Hosp), Dr. Kozo Shaura (Hokushin Gen Hosp), Dr. Yoichi Koizumi (Shinonoi Gen Hosp), and Dr. Yoshio Kasuga (Nagano Matsushiro Gen Hosp) for collecting the gastric cancer samples. The authors also gratefully acknowledge the generous assistance of the staff of each hospital and of the Agri Tech Inst of the Nagano Farmers' Federation, Ms. Aoki, Mr. Ueki, Ms. Kimijima, Ms. Komatsu, Mr. Shimazaki, Ms. Horano, and Mr. Yajima. M.G. is a COE research assistant.

COMMENTS

Background

The human base excision repair protein NEIL1 has activity that is capable of removing oxidatively damaged bases, such as thymine glycol and 5-hydroxyuracil. We have recently demonstrated somatic inactivating *NEIL1* mutations and reduced NEIL1 expression in a subset of gastric cancers, suggesting that reduced NEIL1 activity is involved in gastric carcinogenesis. In the present study, we searched for genetic polymorphisms in the promoter

region of the human *NEIL1* gene in gastric cancer patients and succeeded in identifying three novel genetic polymorphisms.

Research frontiers

Oxidized-DNA-base lesions have been implicated in carcinogenesis, and base excision repair proteins are involved in the repair of such lesions. The research frontier in the area of studying the relationship between the base excision repair genes and carcinogenesis lies in the discovery of genetic variants in the genes that are associated with increased cancer risk.

Innovations and breakthroughs

Stomach tissue is exposed to oxidative stress, including inflammation induced by *Helicobacter pylori* infection, sodium chloride, and smoking, and promoter polymorphisms in some DNA repair genes have been reported to be associated with increased cancer risk. However, there have been no reports of studies that have examined associations between *NEIL1* promoter polymorphisms and gastric cancer risk. The identification of three novel *NEIL1* promoter polymorphisms in this study should be of value for future research in this field.

Applications

The *NEIL1* data presented in this study will be useful for various future studies, such as studies that evaluate the effects of *NEIL1* polymorphisms on the risk of disease, haplotype analyses, and allele-specific expression analyses.

Peer review

This study investigated genetic polymorphisms in the human *NEIL1* gene and identified three polymorphisms in the promoter region of the gene. Although the study did not determine whether the polymorphisms are specific for gastric cancer patients, and no functional analysis of the polymorphisms was performed, the polymorphisms identified are indeed novel, and the results may facilitate future research in this field. In conclusion, the results of this study are somewhat valuable, and the paper appears to be worth publishing as a brief communication.

REFERENCES

1 **Baik SC**, Youn HS, Chung MH, Lee WK, Cho MJ, Ko GH, Park CK, Kasai H, Rhee KH. Increased oxidative DNA damage in Helicobacter pylori-infected human gastric mucosa. *Cancer Res* 1996; **56**: 1279-1282

2 **Kanada R**, Uchida T, Tsukamoto Y, Nguyen LT, Hijiya N, Matsuura K, Kodama M, Okimoto T, Murakami K, Fujioka T, Yanagisawa S, Moriyama M. Genotyping of the cagA gene of Helicobacter pylori on immunohistochemistry with East Asian CagA-specific antibody. *Pathol Int* 2008; **58**: 218-225

3 **Wang XQ**, Terry PD, Yan H. Review of salt consumption and stomach cancer risk: epidemiological and biological evidence. *World J Gastroenterol* 2009; **15**: 2204-2213

4 **Farinati F**, Cardin R, Degan P, Rugge M, Mario FD, Bonvicini P, Naccarato R. Oxidative DNA damage accumulation in gastric carcinogenesis. *Gut* 1998; **42**: 351-356

5 **Trédaniel J**, Boffetta P, Buiatti E, Saracci R, Hirsch A. Tobacco smoking and gastric cancer: review and meta-analysis. *Int J Cancer* 1997; **72**: 565-573

6 **Ernst P**. Review article: the role of inflammation in the pathogenesis of gastric cancer. *Aliment Pharmacol Ther* 1999; **13** Suppl 1: 13-18

7 **Goto M**, Shinmura K, Igarashi H, Kobayashi M, Konno H, Yamada H, Iwaizumi M, Kageyama S, Tsuneyoshi T, Tsugane S, Sugimura H. Altered expression of the human base excision repair gene NTH1 in gastric cancer. *Carcinogenesis* 2009; **30**: 1345-1352

8 **Federico A**, Morgillo F, Tuccillo C, Ciardiello F, Loguercio C. Chronic inflammation and oxidative stress in human carcinogenesis. *Int J Cancer* 2007; **121**: 2381-2386

9 **Mena S**, Ortega A, Estrela JM. Oxidative stress in environmental-induced carcinogenesis. *Mutat Res* 2009; **674**: 36-44

10 **Hazra TK**, Izumi T, Boldogh I, Imhoff B, Kow YW, Jaruga P, Dizdaroglu M, Mitra S. Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc Natl Acad Sci USA* 2002; **99**: 3523-3528

11 **Dou H**, Mitra S, Hazra TK. Repair of oxidized bases in DNA bubble structures by human DNA glycosylases NEIL1 and NEIL2. *J Biol Chem* 2003; **278**: 49679-49684

12 **Rosenquist TA**, Zaika E, Fernandes AS, Zharkov DO, Miller H, Grollman AP. The novel DNA glycosylase, NEIL1, protects mammalian cells from radiation-mediated cell death. *DNA Repair* (Amst) 2003; **2**: 581-591

13 **Miller H**, Fernandes AS, Zaika E, McTigue MM, Torres MC, Wente M, Iden CR, Grollman AP. Stereoselective excision of thymine glycol from oxidatively damaged DNA. *Nucleic Acids Res* 2004; **32**: 338-345

14 **Katafuchi A**, Nakano T, Masaoka A, Terato H, Iwai S, Hanaoka F, Ide H. Differential specificity of human and Escherichia coli endonuclease III and VIII homologues for oxidative base lesions. *J Biol Chem* 2004; **279**: 14464-14471

15 **Shinmura K**, Tao H, Goto M, Igarashi H, Taniguchi T, Maekawa M, Takezaki T, Sugimura H. Inactivating mutations of the human base excision repair gene NEIL1 in gastric cancer. *Carcinogenesis* 2004; **25**: 2311-2317

16 **Das A**, Hazra TK, Boldogh I, Mitra S, Bhakat KK. Induction of the human oxidized base-specific DNA glycosylase NEIL1 by reactive oxygen species. *J Biol Chem* 2005; **280**: 35272-35280

17 **Tsukino H**, Hanaoka T, Otani T, Iwasaki M, Kobayashi M, Hara M, Natsukawa S, Shaura K, Koizumi Y, Kasuga Y, Tsugane S. hOGG1 Ser326Cys polymorphism, interaction with environmental exposures, and gastric cancer risk in Japanese populations. *Cancer Sci* 2004; **95**: 977-983

18 **Guillemette C**, Millikan RC, Newman B, Housman DE. Genetic polymorphisms in uridine diphospho-glucuronosyltransferase 1A1 and association with breast cancer among African Americans. *Cancer Res* 2000; **60**: 950-956

19 **Hu Z**, Ma H, Lu D, Zhou J, Chen Y, Xu L, Zhu J, Huo X, Qian J, Wei Q, Shen H. A promoter polymorphism (-77T>C) of DNA repair gene XRCC1 is associated with risk of lung cancer in relation to tobacco smoking. *Pharmacogenet Genomics* 2005; **15**: 457-463

20 **Raptis S**, Mrkonjic M, Green RC, Pethe VV, Monga N, Chan YM, Daftary D, Dicks E, Younghusband BH, Parfrey PS, Gallinger SS, McLaughlin JR, Knight JA, Bapat B. MLH1 -93G>A promoter polymorphism and the risk of microsatellite-unstable colorectal cancer. *J Natl Cancer Inst* 2007; **99**: 463-474

21 **Mrkonjic M**, Raptis S, Green RC, Monga N, Daftary D, Dicks E, Younghusband HB, Parfrey PS, Gallinger SS, McLaughlin JR, Knight JA, Bapat B. MSH2 118T>C and MSH6 159C>T promoter polymorphisms and the risk of colorectal cancer. *Carcinogenesis* 2007; **28**: 2575-2580

S- Editor Li LF **L- Editor** Lalor PF **E- Editor** Lin YP

HUMAN MUTATION Mutation in Brief 31: E1861-E1874 (2010) Online

MUTATION IN BRIEF

HUMAN MUTATION

Adenine DNA Glycosylase Activity of 14 Human MutY Homolog (MUTYH) Variant Proteins Found in Patients with Colorectal Polyposis and Cancer

Masanori Goto[1], Kazuya Shinmura[1*], Yusaku Nakabeppu[2], Hong Tao[1], Hidetaka Yamada[1], Toshihiro Tsuneyoshi[3], and Haruhiko Sugimura[1*]

[1]First Department of Pathology, Hamamatsu University School of Medicine, Japan [2]Division of Neurofunctional Genomics, Department of Immunobiology and Neuroscience, Medical Institute of Bioregulation, Kyushu University, Japan [3]Department of Materials and Life Science, Shizuoka Institute of Science and Technology, Japan

*Correspondence to Haruhiko Sugimura, Professor, First Department of Pathology, Hamamatsu University School of Medicine, 1-20-1 Handayama, Higashi Ward, Hamamatsu, Shizuoka 431-3192, Japan. E-mail: hsugimur@hama-med.ac.jp or Kazuya Shinmura, Associate Professor, First Department of Pathology, Hamamatsu University School of Medicine, Hamamatsu, Shizuoka 431-3192, Japan. E-mail: kzshinmu@hama-med.ac.jp

Communicated by Riccardo Fodde

ABSTRACT: Biallelic inactivating germline mutations in the base excision repair *MUTYH (MYH)* gene have been shown to predispose to MUTYH-associated polyposis (MAP), which is characterized by multiple colorectal adenomas and carcinomas. In this study, we successfully prepared highly homogeneous human MUTYH type 2 recombinant proteins and compared the DNA glycosylase activity of the wild-type protein and fourteen variant-type proteins on adenine mispaired with 8-hydroxyguanine, an oxidized form of guanine. The adenine DNA glycosylase activity of the p.I195V protein, p.G368D protein, p.M255V protein, and p.Y151C protein was 66.9%, 15.2%, 10.7%, and 4.5%, respectively, of that of the wild-type protein, and the glycosylase activity of the p.R154H, p.L360P, p.P377L, p.452delE, p.R69X, and p.Q310X proteins as well as of the p.D208N negative control form was extremely severely impaired. The glycosylase activity of the p.V47E, p.R281C, p.A345V, and p.S487F proteins, on the other hand, was almost the same as that of the wild-type protein. These results should be of great value in accurately diagnosing MAP and in fully understanding the mechanism by which MUTYH repairs DNA in which adenine is mispaired with 8-hydroxyguanine. ©2010 Wiley-Liss, Inc.

KEY WORDS: base excision repair, 8-hydroxyguanine, MUTYH, MUTYH-associated polyposis, MAP, DNA glycosylase, colorectal cancer

INTRODUCTION

8-Hydroxyguanine (8-OHG) is an oxidized form of guanine (Kasai and Nishimura, 1991), and because 8-OHG can pair with adenine as well as cytosine, formation of 8-OHG in DNA causes a G:C to T:A transversion mutation (Shibutani et al., 1991). MUTYH protein (MIM# 604933), also known as MYH protein, is a DNA glycosylase that

Received 15 March 2010; accepted revised manuscript 1 September 2010.

DOI: 10.1002/humu.21363

catalyzes the removal of adenine mispaired with 8-OHG in double-stranded DNA (Slupska et al., 1999; Shinmura et al., 2000; Tao et al., 2008). Two major MUTYH proteins, i.e., type 1 and type 2, are expressed in human cells as a result of the presence of multiple transcription initiation sites and alternative splicing of mRNA transcripts (Takao et al., 1999; Ohtsubo et al., 2000). Type 1 is composed of 535 amino acids, and because it contains a mitochondrial targeting signal (MTS) in its N-terminal, it is localized in the mitochondria. Type 2 is composed of only 521 amino acid, because it lacks the N-terminal 14 amino acids of type 1, which contain the MTS, and as a result type 2 is localized in the nucleus (Takao et al., 1999; Ohtsubo et al., 2000). The excisional repair activity of the type 2 protein is greater than that of the type 1 protein under certain conditions (Shinmura et al., 2000).

Biallelic inactivating germline mutations in the *MUTYH* gene predispose to MUTYH-associated polyposis (MAP; MIM# 608456), an autosomal recessive disorder characterized by multiple colorectal adenomas and carcinomas (Al-Tassan et al., 2002; Jones et al., 2002; Sampson et al., 2003; Sieber et al., 2003). Since the diagnosis of MAP depends on the level of repair activity of the MUTYH variants encoded in the two *MUTYH* alleles of the patient and the presence of the clinical phenotype characteristic of MAP, even when *MUTYH* gene variations are present in a patient, information on the level of repair activity of the MUTYH variants is indispensable to making the diagnosis of MAP. However, even though more than 80 MUTYH variants have been described in the *MUTYH* gene in colorectal polyposis and colorectal cancer patients (reviewed in Cheadle and Sampson, 2007; Vogt et al., 2009), the effect of only a small number of variations on human MUTYH protein activity has been investigated (Wooden et al., 2004; Bai et al., 2005; Bai et al., 2007; Ali et al., 2008; Kundu et al., 2009; Forsbring et al., 2009; Molatore et al., 2010). One of the reasons for not investigating the effect of more variations is that human MUTYH recombinant proteins cannot be efficiently overexpressed and purified in *Escherichia coli (E. coli)* and baculovirus cultures or in a cell-free system. Thus, even in previous investigations of variant MUTYH proteins, the purified protein fraction also contained multiple other proteins, judging from the photographs of the SDS-PAGE gels. The authors of one paper (Bai et al., 2005) estimated that the purity of the GST-MUTYH fusion proteins used in the analysis was approximately 15%. Too much amount of contamination by other proteins can interfere with accurate determination of the repair activity of variant MUTYH proteins. Thus, improvement of the production and purification system is needed to enable accurate evaluation of the repair activity of variant MUTYH proteins. Moreover, since somatic *APC* (MIM# 611731) mutations occur in the nuclear DNA of a high proportion of MAP tumors (Al-Tassan et al., 2002), it is preferable to evaluate the repair activity of the type 2 protein localized in the nucleus, not the type 1 mitochondrial protein. However, except for the study by Molatole et al. (2010), the repair activity of variants of the type 1 mitochondrial MUTYH form, not the type 2 nuclear form, has been studied in previous studies. Therefore, in the present study we established an experimental system for the purification of MUTYH type 2 recombinant proteins and evaluated 14 type 2 variants, i.e., p.V47E, p.Y151C, p.R154H, p.I195V, p.M255V, p.R281C, p.A345V, p.L360P, p.G368D, p.P377L, p.452delE, p.S487F, p.R69X, and p.Q310X, which correspond to type 1 proteins p.V61E, p.Y165C, p.R168H, p.I201V, p.M269V, p.R295C, p.A359V, p.L374P, p.G382D, p.P391L, p.466delE, p.S501F, p.R83X, and p.Q324X, respectively. All of the above are MUTYH variants that have been identified in patients with colorectal polyposis and/or with colorectal cancer (Halford et al., 2003; Sieber et al., 2003; Aceto et al., 2005; Aretz et al., 2006; Kanter-Smoler et al., 2006; Lejeune et al., 2006; Peterlongo et al., 2006; Russell et al., 2006; Yanaru-Fujisawa et al., 2008). This study assessed the adenine excisional activity of a larger number of MUTYH variants than in previous studies, and the repair activity of the type 2 protein of 11 of the 14 MUTYH variants (p.V47E, p.R154H, p.I195V, p.M255V, p.R281C, p.A345V, p.L360P, p.P377L, p.S487F, p.R69X, and p.Q310X) was examined for the first in this study.

MATERIALS AND METHODS

Plasmid construction

The human MUTYH type 2 cDNA sequence was inserted into a pET25b(+) expression vector (Novagen, Darmstadt, Germany). The expression vector for 13 missense-type variants was generated by site-directed mutagenesis with a QuikChange Site-directed Mutagenesis kit (Stratagene, La Jolla, CA). The expression vector for p.R69X and p.Q310X types were constructed by inserting MUTYH cDNA sequence (nucleotides 1-204 and 1-927, respectively) into the pET25b(+) expression vector. All vectors were confirmed by DNA sequencing with a BigDye Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Tokyo, Japan) and an ABI 3100 Genetic Analyzer (Applied Biosystems).