Table III. Clinicopathological characteristics of 10 patients with *EPHA7-S*-positive lung cancer.

Patient no.	Sex	Age	Histology	Stage	BI
9	F	77	SCC	3b	0
15	M	62	SQCC	3a	1260
18	M	61	SQCC	2b	1200
24	M	69	SQCC	1a	740
31	M	39	SQCC	1b	920
34	M	66	LCC (LCNEC)	1b	1350
42	M	65	ADC	1a	1350
47	F	82	ADC	1a	0
55	F	54	SQCC	3b	720
63	F	67	SQCC	1a	0

ADC, adenocarcinoma; LCC, large cell carcinoma; LCNEC, large cell neuroendocrine carcinoma; SCC, small cell carcinoma; SQCC, squamous cell carcinoma; BI, Brinkman index.

Table IV. Associations between *EPHA7*-S and clinicopathological features.

	EPHA7-S$^+$	*EPHA7*-S$^-$	P-value
Gender			
Male	6	35	0.7924[a]
Female	4	28	
Histology			
Adenocarcinoma	2	48	0.0004[a]
Other carcinomas	8	15	
Brinkman index			
0≤BI<400	3	25	0.557[a]
400≤BI	7	29	
TNM stage			
I	6	43	0.3548[b]
II	1	4	
III	3	10	
IV	0	3	

[a]χ^2 test. [b]Cochran-Armitage trend test.

other clinicopathological factors were related to EPHA7-S expression.

Discussion

While pursuing the possibility of EPHA7 fusion partners, we incidentally isolated EPHA7-S in lung cancer cell lines and tissues. Intriguingly, EPHA7-S tends to be expressed in lung but not other cancer cell lines. Furthermore, this EPHA7-S expression in lung cancers was seen mainly in non-adenocarcinomas. These findings may facilitate diagnosing of the primary site of such cancers, especially non-adenocarcinoma types.

Lung carcinogenesis probably constitutes heterogeneous steps according to its histological subtype and/or environmental effects such as smoking. Classically, non-adenocarcinomas are believed to have a stronger association with smoking than adenocarcinoma of the lung (24), such that the greater representation of EPHA7-S in non-adenocarcinoma than adenocarcinoma may suggest this difference to reflect EPHA7-S involvement in smoking related carcinogenesis. The amount of tobacco did not, however, correlate with the EPHA7-S expressions in the tumors of our subjects. Thus, the significance of this tendency must be carefully interpreted and further accumulation of cases is warranted. Recently, expression of EPHA2, another EPHA family member, was reported to be associated with smoking history (25). EPHA7-S is possibly another member of the EPH family of genes related to all form of smoking. The prevalences of EPHA7-S in lung cancer cell lines of different histological origin did not, however, differ between adenocarcinomas and non-adenocarcinomas. Again, the significance of the apparent difference in primary lung tumors according to subtype remains an open question. There is only one report documenting EPHA7-S expression in human tumors. Dawson *et al* reported EPHA7-S in germinal center lymphocytes (23). They speculated that hypermethylation of the EPHA7 promoter and the secreted form of EphA7 interacted in lymphomatogenesis. In fact, the finding of frequent promoter methylation in *EPHA7* in colorectal cancer raised the possibility that *EPHA7* is one of the tumor suppressor genes inactivated in colorectal carcinogenesis (8). The EPHA7 expression profile in human lung cancers was quite different from those in colorectal, gastric, prostate, and brain cancers (8,13,14,26). EPHA7 was expressed mainly in the tumor portion (data not shown). In this context, EPHA7 may be an oncogene rather than a suppressor gene in lung carcinogenesis. A larger clinical study is needed to corroborate this view of EPHA7 as an oncogene. This paradoxical situation (oncogene in one and suppressor gene in an other organ) has been described in several organs and settings in which the *EPH* family of genes play roles (27). EPHA7-S is also overexpressed in some lung cancers, meaning that the role of EPHA7-S in lung cancer may not be consistent with that assumed in lymphoma, as previously reported (23). We detected EPHA7-S in two non-tumor lung tissues. We do not know the exact reason, but environmental stress may induce various forms of aberrant splicing (28). Since we also detected EPHA7 in SAEC and 16HBE14o- cell lines, both of which are from non-tumorigenic bronchogenic epithelium, the same microenvironment as in lung cancer tissue may exist in the non-tumor portion of the lung. In addition, there was an exception: HSC-39, a gastric cancer cell line, expresses EPHA7-S. The biological and pathological situations triggering EPHA7-S expression clearly require further study. In conclusion, we have, for the first time, described the secreted form of EPHA7 in a subset of human lung cancers.

Acknowledgments

This work is supported in part by a Grant-in-Aid from the Ministry of Education, Culture, Sports, Science and Technology (20014007), from the Ministry of Health, Labor, and Welfare (19-19), and the Smoking Research Foundation.

References

1. Eph nomenclature and committee: unified nomenclature for Eph family receptors and their ligands, the ephrins. Eph Nomenclature Committee. Cell 90: 403-404, 1997.
2. Pasquale EB: Eph-ephrin bidirectional signaling in physiology and disease. Cell 133: 38-52, 2008.
3. Nakamoto M and Bergemann AD: Diverse roles for the Eph family of receptor tyrosine kinases in carcinogenesis. Microsc Res Tech 59: 58-67, 2002.
4. Noren NK, Lu M, Freeman AL, Koolpe M and Pasquale EB: Interplay between EphB4 on tumor cells and vascular ephrin-B2 regulates tumor growth. Proc Natl Acad Sci USA 101: 5583-5588, 2004.
5. Campbell TN, Attwell S, Arcellana-Panlilio M and Robbins SM: Ephrin A5 expression promotes invasion and transformation of murine fibroblasts. Biochem Biophys Res Commun 350: 623-628, 2006.
6. Tanaka M, Sasaki K, Kamata R and Sakai R: The C-terminus of ephrin-B1 regulates metalloproteinase secretion and invasion of cancer cells. J Cell Sci 120: 2179-2189, 2007.
7. Nakamura R, Kataoka H, Sato N, *et al*: EPHA2/EFNA1 expression in human gastric cancer. Cancer Sci 96: 42-47, 2005.
8. Wang J, Kataoka H, Suzuki M, *et al*: Downregulation of EphA7 by hypermethylation in colorectal cancer. Oncogene 24: 5637-5647, 2005.
9. Taddei ML, Parri M, Angelucci A, *et al*: Kinase-dependent and -independent roles of EphA2 in the regulation of prostate cancer invasion and metastasis. Am J Pathol 174: 1492-1503, 2009.
10. Kataoka H, Igarashi H, Kanamori M, *et al*: Correlation of EPHA2 overexpression with high microvessel count in human primary colorectal cancer. Cancer Sci 95: 136-141, 2004.
11. Kataoka H, Tanaka M, Kanamori M, *et al*: Expression profile of EFNB1, EFNB2, two ligands of EPHB2 in human gastric cancer. J Cancer Res Clin Oncol 128: 343-348, 2002.
12. Ding L, Getz G, Wheeler DA, *et al*: Somatic mutations affect key pathways in lung adenocarcinoma. Nature 455: 1069-1075, 2008.
13. Guan M, Xu C, Zhang F and Ye C: Aberrant methylation of EphA7 in human prostate cancer and its relation to clinicopathologic features. Int J Cancer 124: 88-94, 2009.
14. Wang J, Li G, Ma H, *et al*: Differential expression of EphA7 receptor tyrosine kinase in gastric carcinoma. Hum Pathol 38: 1649-1656, 2007.
15. Foster RE, Abdulrahman M, Morris MR, *et al*: Characterization of a 3;6 translocation associated with renal cell carcinoma. Genes Chromosomes Cancer 46: 311-317, 2007.
16. Tognon C, Knezevich SR, Huntsman D, *et al*: Expression of the ETV6-NTRK3 gene fusion as a primary event in human secretory breast carcinoma. Cancer Cell 2: 367-376, 2002.
17. Tomlins SA, Laxman B, Dhanasekaran SM, *et al*: Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 448: 595-599, 2007.
18. Soda M, Choi YL, Enomoto M, *et al*: Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature 448: 561-566, 2007.
19. Takeuchi K, Choi YL, Togashi Y, *et al*: KIF5B-ALK, a novel fusion oncokinase identified by an immunohistochemistry-based diagnostic system for ALK-positive lung cancer. Clin Cancer Res 15: 3143-3149, 2009.
20. Rikova K, Guo A, Zeng Q, *et al*: Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell 131: 1190-1203, 2007.
21. Okudela K, Suzuki M, Kageyama S, *et al*: PIK3CA mutation and amplification in human lung cancer. Pathol Int 57: 664-671, 2007.
22. Kleihues Pand L and Sobin H: (eds.) Tumours of the lung, pleura, thymus and heart. IARC Press, Lyon, 2004.
23. Dawson DW, Hong JS, Shen RR, *et al*: Global DNA methylation profiling reveals silencing of a secreted form of Epha7 in mouse and human germinal center B-cell lymphomas. Oncogene 26: 4243-4252, 2007.
24. Kreyberg L: Relationship of different histological lung tumour groups to tobacco smoking. Br J Cancer 15: 51-53, 1961.
25. Brannan JM, Dong W, Prudkin L, *et al*: Expression of the receptor tyrosine kinase EphA2 is increased in smokers and predicts poor survival in non-small cell lung cancer. Clin Cancer Res 15: 4423-4430, 2009.
26. Wang LF, Fokas E, Juricko J, *et al*: Increased expression of EphA7 correlates with adverse outcome in primary and recurrent glioblastoma multiforme patients. BMC Cancer 8: 79, 2008.
27. Noren NK and Pasquale EB: Paradoxes of the EphB4 receptor in cancer. Cancer Res 67: 3994-3997, 2007.
28. Korner M and Miller LJ: Alternative splicing of pre-mRNA in cancer: focus on G protein-coupled peptide hormone receptors. Am J Pathol 175: 461-472, 2009.

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).

Preparation of the recombinant MUTYH proteins

E. coli BL21-CodonPlus (DE3)-RP competent cells (Stratagene) were transformed with the MUTYH-pET25b vector and cultured at 37°C until an A_{600} of 0.6. After incubation with 0.1 mM IPTG at 15°C for 12 h, MUTYH-His_6 protein was purified with TALON metal affinity resins (Clontech, Palo Alto, CA) and a TALON 2-ml disposable gravity column (Clontech). The protein was then dialyzed against buffer containing 10 mM sodium phosphate (pH 7.6), 50 mM NaCl, 0.5 mM DTT, 0.1 mM EDTA, 0.5 mM PMSF, 2 μg/ml pepstatin, 2 μg/ml leupeptin, 50 μM chymostatin, and 10% glycerol. The quality and concentration of MUTYH proteins were determined by using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA) and Image J software (National Institutes of Health, Bethesda, MD).

Western blot analysis

Purified recombinant protein was mixed with an equal volume of 2x SDS sample buffer and boiled. A 2 μg protein was subjected to SDS-polyacrylamide gel electrophoresis (PAGE) and electrophoretically transferred to a polyvinylidene difluoride membrane (GE Healthcare Bio-Science Corp., Piscataway, NJ). The membrane was blocked with non-fat milk and incubated with an anti-MUTYH polyclonal antibody (Ohtsubo et al., 2000). After washing, the membrane was incubated with an anti-rabbit HRP-conjugated secondary antibody (GE Healthcare Bio-Science Corp.). The membrane was then washed, and immunoreactivity was visualized with an ECL Plus chemiluminescence system (GE Healthcare Bio-Science Corp.).

DNA cleavage activity assay

30-mer oligonucleotides containing and not containing a single 8-OHG (5'-CTG GTG GCC TGA C[8-OHG or T]C ATT CCC CAA CTA GTG-3') were chemically synthesized and purified by PAGE (Japan Bio Services, Saitama, Japan). Complementary oligonucleotides containing an adenine opposite the 8-OHG or T were ^{32}P-labeled at the 5' terminus with a MEGALABEL kit (Takara, Osaka, Japan) and a [γ-^{32}P]ATP (PerkinElmer, Tokyo, Japan), and then annealed to oligonucleotides containing a single 8-OHG or T. The reaction mixture containing 20 mM sodium phosphate (pH 7.6), 100 mM NaCl, 0.5 mM DTT, 0.5 mM EDTA, 5 μM $ZnCl_2$, 1.5% glycerol, 2.5 nM labeled oligonucleotide, 50 μg/ml BSA, and purified MUTYH protein was incubated at 37°C, and the mixture was treated with 0.1 M NaOH at 95°C for 4 min. After adding denaturing formamide dye to the mixture, it was heated at 95°C for 3 min, and subjected to 20% PAGE. A ^{32}P-labeled marker oligonucleotide was used as a size marker for the cleavage products. The radioactivity of intact and cleaved oligonucleotides was quantified by using an FLA-3000 fluoroimage analyzer (Fuji Film, Tokyo, Japan) and ImageGauge software (Fuji Film) (Goto et al., 2009).

Active site titration and evaluation of the rate constant k_f

The active site titration and evaluation of the rate constant were performed as described previously (Fersht, 1985; Kundu et al., 2009). A 100 ng amount of total proteins was incubated at 37°C for 0 - 30 min with 5 nM 8-OHG containing substrate and the cleavage products were monitored. To determine the amplitude of the burst (A_0), which is proportional to the active protein fraction concentration, the data were fitted to equation (1):

$$[P]_t = A_0[1\text{-}\exp(\text{-}k_b t)] + k_l t \qquad (1)$$

where $[P]_t$ is the cleavage product concentration at time t and k_b and k_l are the rate constants of the burst phase and the linear phase, respectively.

Rate constant, k_f, was evaluated under single-turnover conditions. A 10 nM concentration of active MUTYH enzymes was incubated at 37°C for 0 - 15 min with 2.5 nM 8-OHG containing substrate and the cleavage products were monitored. To estimate the k_f, the data were fitted to equation (2):

$$[P]_t = A_0[1\text{-}\exp(\text{-}k_f t)] \qquad (2)$$

Mutation nomenclature and reference sequence

Mutation nomenclature is according to den Dunnen and Antonarakis (2000) and den Dunnen and Paalman (2003). The reference sequence for the *MUTYH* gene encoding type 2 protein is accession number NM_001048174.1. The nucleotide numbering system uses the A of the ATG translation initiation start site as nucleotide +1.

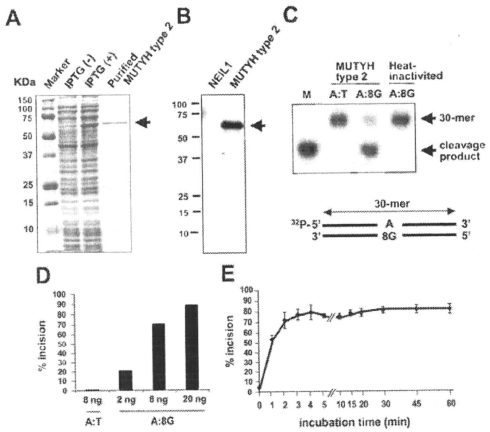


Figure 1. Measurement of the adenine DNA glycosylase activity of wild-type MUTYH type 2 protein. (**A**) Purification of wild-type MUTYH type 2 protein resolved by SDS-polyacrylamide gel electrophoresis (PAGE) and stained with Coomassie Brilliant Blue. Lysates of *E. coli* culture without or with IPTG induction and purified MUTYH type 2 protein are shown. The arrow points to the MUTYH-His_6 protein band. (**B**) Western blot of purified wild-type MUTYH type 2 protein tagged with His_6. MUTYH-His_6 protein is indicated by the arrow. Purified recombinant NEIL1 (MIM# 608844)-His_6 protein, which was prepared by using pET25b(+) vector (Novagen) and *E. coli* BL21-CodonPlus (DE3)-RP cells (Stratagene) previously (Shinmura et al., 2004), was included as a negative control. (**C**) The DNA glycosylase activity of MUTYH type 2 protein on double-stranded DNA containing an A:8-hydroxyguanine (8-OHG). The MUTYH type 2 protein and a ^{32}P-labeled double-stranded oligonucleotides containing or not containing a single 8-OHG mispair were incubated and subjected to 20% PAGE. The intact 30-mer oligonucleotides and cleavage products are indicated by the arrows. Heat-inactivation of the MUTYH protein was accomplished by heating the protein at 100°C for 5 min. 8G means 8-hydroxyguanine. (**D**) Protein concentration dependency of cleavage of DNA containing an A:8-OHG by MUTYH type 2 protein. The MUTYH protein (2, 8, and 20 ng) was incubated at 37°C for 15 min with a 30-mer oligonucleotide containing an A:8-OHG or A:T (50 fmole). The amount of cleavage products as a proportion of total oligonucleotides was calculated as % incision. (**E**) Time-course assay of cleavage of DNA containing an A:8-OHG by MUTYH type 2 protein. The 8 ng amount of MUTYH type 2 protein was incubated at 37°C for 0 - 60 min with double-stranded oligonucleotide containing an A:8-OHG (50 fmole). The % incision valves are means ± standard errors of data from three independent experiments.

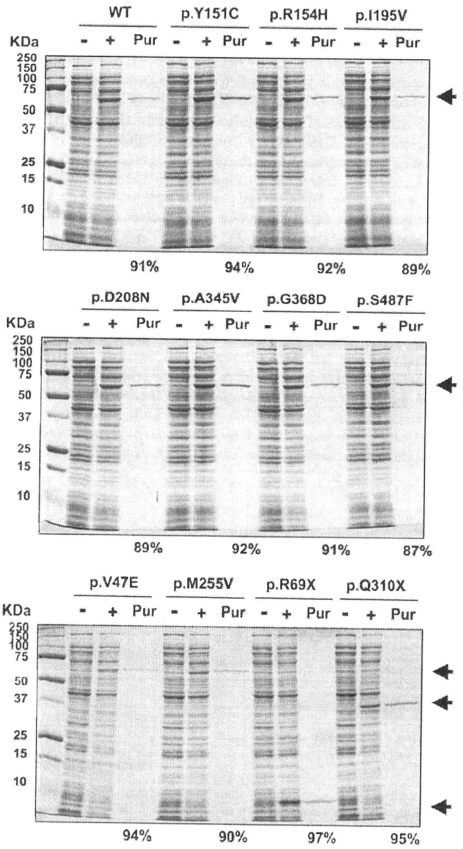


Figure 2. Purification of variant-type MUTYH type 2 recombinant proteins. Each protein was overexpressed and purified under conditions essentially the same as used for the wild-type (WT) protein. Representative results of expression and purification of MUTYH proteins resolved by SDS-PAGE and stained with Coomassie Brilliant Blue are shown. '–' and '+' mean absence and presence, respectively, of IPTG induction, and 'Pur' means purified MUTYH type 2 proteins. The arrow points to the MUTYH-His_6 protein band. The purification level is indicated below the SDS-PAGE panels.

RESULTS

To overcome the difficulty of preparing highly purified recombinant MUTYH proteins, in this study, we used a pET25b(+) expression vector and BL21-CodonPlus (DE3)-RP *E. coli* host cells for induction of MUTYH expression and a TALON metal affinity resin and gravity column for purification of the MUTYH proteins. Wild-type human MUTYH type 2 protein tagged with His_6 at its C-terminus was successfully overexpressed in *E. coli* and purified to approximately 90% homogeneity (Figure 1A). The specificity of the purified MUTYH protein was confirmed by Western blotting with anti-MUTYH polyclonal antibody (Figure 1B). Their molecular size of approximately 61 kDa was determined by SDS-PAGE / Coomassie Brilliant Blue (CBB) staining and Western blotting, and it corresponded to their size calculated from the cDNA sequence. The DNA glycosylase activity of the wild-type MUTYH protein was tested by determining its capacity to cleave a double-stranded oligonucleotide containing an adenine mispaired with 8-OHG. The cleavage products were analyzed on a denaturing polyacrylamide gel, and their mobility was compared with that of a marker oligonucleotide. No clear cleavage products were detected when oligonucleotide containing an unmodified A:T base pair was exposed to the MUTYH protein, but cleavage products having the same mobility as the marker oligonucleotide were detected when MUTYH proteins were allowed to react with oligonucleotide containing an A:8-OHG base pair (Figure 1C). No cleavage was detected when allowed to react after heat-inactivation of the MUTYH protein (Figure 1C). The amount of cleavage products was calculated as percent of total oligonucleotides and expressed as % incision, and the % incision of protein substrate containing an A:8-OHG mispair increased in a protein-concentration–dependent manner (Figure 1D). The time-course assay of the cleavage activity of MUTYH protein on substrate containing an A:8-OHG mispair indicated that the amount of cleavage products peaked within 5 min and was almost constant from 5 min to 60 min (Figure 1E). The results showed that highly purified wild-type MUTYH type 2 protein had been obtained with our expression and purification system, and the adenine DNA glycosylase activity of the protein on an A:8-OHG mispair was satisfactorily detected by our assay. We therefore decided to apply our experimental systems to assessment of the adenine DNA glycosylase activity of various MUTYH variant proteins.

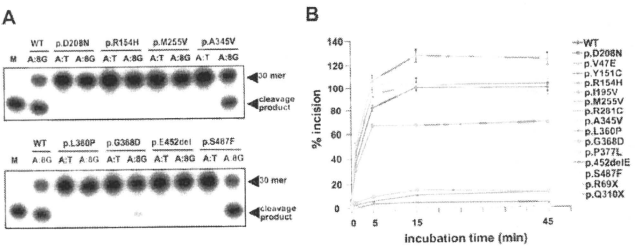


Figure 3. Measurement of the adenine DNA glycosylase activity of variant MUTYH type 2 proteins. **(A)** Representative results of DNA cleavage assays of MUTYH variant proteins are shown. MUTYH proteins (130 fmole) were allowed to act on double-stranded oligonucleotide containing a single A:8-OHG (8G) mispair at 37°C for 15 min. The reaction mixture was analyzed by 20% PAGE. A ^{32}P-labeled marker oligonucleotide was used as a size marker for the cleavage products. The intact 30-mer oligonucleotides and cleavage products are indicated by the arrows. **(B)** Time-course assay of cleavage of DNA containing an A:8-OHG by MUTYH type 2 protein. MUTYH type 2 proteins (130 fmole) were incubated at 37°C for 0 - 45 min with double-stranded oligonucleotide containing an A:8-OHG (50 fmole). The amount of cleavage products as a proportion of total oligonucleotides was calculated as % incision, and the % incision of each variant-type MUTYH protein is shown relative to that of wild-type (WT) MUTYH protein, which has been set equal to 100. The % incision valves are means ± standard errors of data obtained from three independent experiments in which three independently prepared MUTYH proteins were used.

Table 1. DNA glycosylase activity of 14 variants of MUTYH type 2 protein on DNA containing adenine mispaired with 8-hydroxyguanine

MUTYH type 2 protein	Type of mutation[a]	Purified protein yield (ng) per 10 ml culture[b]	Relative % incision[c]
WT		4365, 3562, 3454	100
p.D208N	negative control	2073, 1826, 1808	extremely severely defective
p.V47E	c.140T>A, missense	1018, 1893, 1916	128.1±5.20
p.Y151C	c.452A>G, missense	5485, 4779, 2835	4.5±0.21
p.R154H	c.461G>A, missense	894, 1501, 2109	extremely severely defective
p.I195V	c.583A>G, missense	1733, 1102, 671	66.9±0.35
p.M255V	c.763A>G, missense	2446, 1726, 2488	10.7±0.47
p.R281C	c.841C>T, missense	912, 934, 447	103.0±1.39
p.A345V	c.1034C>T, missense	2564, 2435, 3864	103.5±5.43
p.L360P	c.1079T>C, missense	152, 303, 120	extremely severely defective
p.G368D	c.1103G>A, missense	2486, 1246, 2270	15.2±0.71
p.P377L	c.1130C>T, missense	885, 626, 430	extremely severely defective
p.452delE	c.1353_1355delGGA, inframe deletion	822, 614, 364	extremely severely defective
p.S487F	c.1460C>T, missense	1099, 984, 706	102.8±2.31
p.R69X	c.205C>T, nonsense	6289, 5245, 6289	extremely severely defective
p.Q310X	c.928C>T, nonsense	918, 5726, 2771	extremely severely defective

[a]The reference sequence for the *MUTYH* gene encoding type 2 protein is accession number NM_001048174.1.

[b]Amount of MUTYH proteins purified from 10 ml of *E. coli* culture expressing MUTYH type 2 protein. Each protein was purified three times and has been listed in this table.

[c]The DNA cleavage activity of MUTYH protein was measured under conditions of 37°C for 15 min. The amount of cleavage products as a proportion of total oligonucleotides was calculated as % incision, and the % incision of each variant-type MUTYH protein is shown relative to that of wild-type (WT) MUTYH protein, which has been set equal to 100. Values are means ± standard errors of data obtained from three independent experiments in which three independently prepared MUTYH proteins were used.

Fourteen MUTYH variants that had previously been identified in patients with colorectal polyposis and/or colorectal cancer were selected, and their expression vectors were prepared by site-directed mutagenesis. Since the Asp222 in MUTYH type 1 is the active site, and the p.D222N mutant is known not to have DNA glycosylase activity (Wooden et al., 2004), we prepared a type 2-p.D208N construct corresponding to type 1-p.D222N as a negative control. A total of 15 MUTYH type 2 proteins were successfully expressed and purified to a high level of homogeneity (Figure 2). Their molecular sizes determined by SDS-PAGE / CBB staining corresponded with their sizes calculated from their cDNA sequences. Each MUTYH variant protein was almost always reproducibly obtained in the three independent protein preparation (Table 1). Next, we compared the DNA glycosylase activity of each variant protein on oligonucleotide containing an A:8-OHG mispair with that of the wild-type MUTYH type 2 protein (Figure 3 and Table 1). As expected, no clear cleavage products were detected when any of the variant proteins were allowed to act on oligonucleotide containing an unmodified A:T base pair (Figure 3). The adenine DNA glycosylase activity of the MUTYH variant proteins on oligonucleotide containing an A:8-OHG mispair varied (Figure 3 and Table 1). The adenine DNA glycosylase activity of the p.V47E, p.R281C, p.A345V, and p.S487F proteins on the A:8-OHG substrate under conditions of 37°C for 15 min were similar to that of the wild-type protein or only slightly different (102.8% - 128.1%, with activity of the wild-type protein set equal to 100%), whereas p.I195V protein exhibited slightly lower glycosylase activity (66.9%). The p.Y151C protein, p.M255V protein and p.G368D protein exhibited only 4.5%, 10.7%, and 15.2%, respectively, of the glycosylase activity of the wild-type protein, and the adenine DNA 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 protein was almost at the background

level. We also attempted to determine whether the recombinant type 1 variant proteins had a level of repair activity that was similar to that of the corresponding type 2 proteins. We randomly chose type 1-p.R168H and type 1-p.S501F, which correspond to type 2-p.R154H and type 2-p.S487F, respectively, and found that the activity level of the type 1 and type 2 proteins of at least these two MUTYH variants in comparison with the wild-type protein is similar (Supp. Figure S1A-E). We also estimated the rate constant k_f of some type 2 proteins on the A:8-OHG substrate after correction for the active enzyme fraction as described previously (Fersht, 1985; Kundu et al., 2009). The rate constant for adenine excision by wild-type protein was 0.524, and similar to the constant for p.R281C (k_f = 0.489) (Table 2). However, the k_f value of p.M255V was 0.024 and more than 20-fold lower than that of the wild-type protein, indicating that the glycosylase activity of p.M255V was severely reduced. The above findings indicate that the adenine DNA glycosylase activity of nine of the 14 MUTYH type 2 variant proteins tested is severely impaired.

Table 2. Active yield and rate constant k_f evaluated for MUTYH type 2 protein on DNA containing adenine mispaired with 8-hydroxyguanine

MUTYH type 2 protein	Active protein yield (μg) per 1L culture[a]	k_f (min^{-1})[b]
WT	2706, 1248, 2045	0.524±0.034
p.M255V	58, 72, 75	0.024±0.004
p.R281C	944, 1128, 751	0.489±0.044

[a]Values were obtained by incubating 100 ng of total proteins with 5 nM of substrate. Results obtained with three separate protein preparations are shown.

[b]Values were obtained by incubating 10 nM of active MUTYH enzyme and 2.5 nM of substrate. Values are means ± standard errors of data obtained from three independent experiments using independently prepared proteins.

DISCUSSION

The cumulative results of recent screenings of colorectal polyposis patients for *MUTYH* mutations have revealed many *MUTYH* gene variants (reviewed in Cheadle and Sampson, 2007; Vogt et al., 2009), but the repair activity of the type 2 protein of most of the variants has never been tested. In this study we improved the method of expressing and purifying the recombinant MUTYH type 2 proteins and assessed the adenine DNA glycosylase activity of various type 2 proteins. The results revealed that the p.V47E, p.R281C, p.A345V, and p.S487F proteins largely retain adenine removing activity, but that the adenine removing activity of the p.I195V protein is mildly impaired and the activity of the p.Y151C, p.R154H, p.M255V, p.L360P, p.G368D, p.P377L, p.452delE, p.R69X, and p.Q310X proteins is severely impaired. This information should be of great help in accurately diagnosing MAP and managing MAP patients.

The adenine DNA glycosylase activity of the type 1 or type 2 MUTYH variants p.Y151C, p.G368D, p.P377L, p.452delE, and p.S487F has been assessed previously (Wooden et al., 2004; Ali et al., 2008; Kundu et al., 2009; Forsbring et al., 2009; Molatore et al., 2010). The results showed that the repair activity of p.Y151C, p.G368D, p.P377L, and p.452delE was impaired while that of p.S487F was retained, findings that were consistent with our own, and the consistency confirms the reliability of the results of our study. Interestingly, a slight difference (approximately 5% vs. approximately 15%) in repair activity was observed between p.Y151C and p.G368D in our study, and the same difference was reported in several previous papers. Nielsen et al. (2009) recently reported finding that MAP patients homozygous for a p.G368D allele have a milder clinical phenotype than MAP patients homozygous for a p.Y151C allele. The difference in adenine removing activity between the p.Y151C protein and p.G368D protein may be related to the difference in clinical phenotype.

Retention of DNA glycosylase activity by the human MUTYH proteins p.V47E, p.R281C, and p.A345V as well as mild impairment of the activity of p.I195V and severe impairment of the activity of p.R154H, p.M255V, p.L360P, p.R69X, and p.Q310X were documented for the first time in this study. Although no analyses of crystal

structure of the MUTYH has been reported, based on the cumulative results of previous studies of the biochemistry of MUTYH (reviewed in Cheadle and Sampson, 2007; Ali et al., 2008), p.R154, p.M255, and p.L360 are located in regions suspected of being important to catalytic activity or substrate recognition. Moreover, the p.R154, p.M255, and p.L360 in MUTYH protein are conserved among *E. coli*, *Mus musculus*, *Rattus norvegicus*, *Pan troglodytes*, *Canis familiaris*, and *Homo sapiens*. Substitution of any of these amino acids appears to result in a functional abnormality, resulting in severe reduction of adenine DNA glycosylase activity on DNA containing an A:8-OHG mispair. The loss of large parts of the MUTYH protein in the frameshift-type variant proteins p.R69X and p.Q310X may be responsible for the severe impairment of their adenine removing activity. An amino acid substitution in p.V47E, p.R281C, p.A345V, and p.S487F did not greatly affect their adenine DNA glycosylase activity. These results are consistent with the prediction of a possible impact of an amino acid substitution on the structure and function of MUTYH type 2 protein by the PolyPhen-2 program (http://genetics.bwh.harvard.edu/pph2/index.shtml) (Adzhubei et al., 2010). Since p.V47E, p.R281C, p.A345V, and p.S487F are not located in the region suspected of being important to catalytic activity or in a well-conserved position of the substrate recognition region, the amino acid localization may be one of the reasons for the retention of enzymatic activity. In the future, a crystal structure of the MUTYH protein alone and covalently complexed with DNA, in conjunction with the present findings on MUTYH variants, should contribute to establishing further correlations between the structure and repair function of the MUTYH protein.

Highly homogeneous MUTYH type 2 recombinant protein was prepared by using the method described in this study. As far as we have been able to determine in a review of the literature, the level of purification of the proteins in our study was higher than level of purification estimated from the SDS-PAGE images of MUTYH type 1 or type 2 proteins in previous papers and purification levels described in the other papers. Use of *E. coli* BL21-CodonPlus (DE3)-RP cells, which contain extra copies of the genes that encode the tRNAs of rare *E. coli* codons, is thought to be responsible for the increase in level of MUTYH protein expression, because the *MUTYH* gene contains many rare codons. The expression conditions (temperature and IPTG concentration), combination of expression vector and competent cells, and metal-affinity purification conditions are also thought to possibly have contributed to the improvement in the level of purification in this study. Since genetic screening of colorectal polyposis patients for *MUTYH* mutations will continue to be performed worldwide, the same as genetic screening for *APC* mutations (reviewed in Lynch et al., 2008), our method of expression and purification of human MUTYH protein should be useful for assessing MUTYH variants newly identified by genetic screening as well as MUTYH variants that have not been examined.

As shown in Supp. Figure S1E, the activity level of the type 1 and type 2 proteins of two MUTYH variants in comparison with the wild-type protein was similar. Although comparisons were not made for the other variants, it may not always be necessarily to study the type 2 form when analyzing MUTYH protein. However, since type 2, and not type 1, is a nuclear form (Takao et al., 1999; Ohtsubo et al., 2000), and somatic *APC* mutations occur in the nuclear DNA of MAP tumors (Al-Tassan et al., 2002), we think that evaluation of type 2 variants is likely to be more preferable when we investigate the possible pathogenic role of MUTYH in MAP. Assessment of the glycosylase activity level of the wild-type type 1 and type 2 proteins showed that the activity of type 2 was greater than that of type 1 (Figure 1 and Supp. Figure S1), a finding that is consistent with a previous report (Shinmura et al., 2000). The reason for the difference in the catalysis of adenine excision between the type 1 and type 2 proteins is unclear. One point that requires caution is that there is evidence that the type 1 protein is processed during mitochondrial transport in cells and the mature form of the protein never been identified, and thus the repair activity of the full-length type 1 may not necessarily reflect the activity *in vivo*. With regard to the method of evaluating MUTYH variant proteins, since MUTYH possesses suppressive activity against G:C to T:A mutations caused by 8-OHG (Yamane et al., 2003) and biallelic *MUTYH* inactivation leads to somatic *APC* mutation in MAP tumors (Al-Tassan et al., 2002), analysis of the mutation rate of the *APC* gene in MUTYH variant-expressing cells may be an alternative way of evaluating the level of repair activity of a MUTYH variant. A combination of mutation rate analysis and DNA glycosylase analysis would provide more definitive proof of the pathogenicity of a MUTYH variant.

ACKNOWLEDGMENTS

Contract grant sponsor: 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 (19790286, 22590356, 22790378), the Ministry of Education, Culture, Sports, Science and Technology for priority area (18014009, 20014007, 221S0001), and the Smoking Research Foundation.

REFERENCES

Aceto G, Cristina Curia M, Veschi S, De Lellis L, Mammarella S, Catalano T, Stuppia L, Palka G, Valanzano R, Tonelli F, Casale V, Stigliano V, Cetta F, Battista P, Mariani-Costantini R, Cama A. 2005. Mutations of APC and MYH in unrelated Italian patients with adenomatous polyposis coli. Hum Mutat 26: 394.

Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. 2010. A method and server for predicting damaging missense mutations. Nat Methods 7:248-9.

Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, Hodges AK, Davies DR, David SS, Sampson JR, Cheadle JP. 2002. Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. Nat Genet 30: 227-32.

Ali M, Kim H, Cleary S, Cupples C, Gallinger S, Bristow R. 2008. Characterization of mutant MUTYH proteins associated with familial colorectal cancer. Gastroenterology 135: 499-507.

Aretz S, Uhlhaas S, Goergens H, Siberg K, Vogel M, Pagenstecher C, Mangold E, Caspari R, Propping P, Friedl W. 2006. MUTYH-associated polyposis: 70 of 71 patients with biallelic mutations present with an attenuated or atypical phenotype. Int J Cancer 119: 807-14.

Bai H, Jones S, Guan X, Wilson TM, Sampson JR, Cheadle JP, Lu AL. 2005. Functional characterization of two human MutY homolog (hMYH) missense mutations (R227W and V232F) that lie within the putative hMSH6 binding domain and are associated with hMYH polyposis. Nucleic Acids Res 33: 597-604.

Bai H, Grist S, Gardner J, Suthers G, Wilson TM, Lu AL. 2007. Functional characterization of human MutY homolog (hMYH) missense mutation (R231L) that is linked with hMYH-associated polyposis. Cancer Lett 250: 74-81.

Cheadle JP, Sampson JR. 2007. MUTYH-associated polyposis--from defect in base excision repair to clinical genetic testing. DNA Repair (Amst) 6: 274-9.

den Dunnen JT, Antonarakis SE. 2000. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum Mutat 15:7-12.

den Dunnen JT, Paalman MH. 2003. Standardizing mutation nomenclature: why bother? Hum Mutat 22(3):181-182.

Fersht A. 1985. Measurement and magnitude of enzymatic rate constants. In: Enzyme Structure and Mechanism (2nd edn.). Freeman WH, New York, p 121– 154.

Forsbring M, Vik ES, Dalhus B, Karlsen TH, Bergquist A, Schrumpf E, Bjørås M, Boberg KM, Alseth I. 2009. Catalytically impaired hMYH and NEIL1 mutant proteins identified in patients with primary sclerosing cholangitis and cholangiocarcinoma. Carcinogenesis 30: 1147-54.

Goto M, Shinmura K, Igarashi H, Kobayashi M, Konno H, Yamada H, Iwaizumi M, Kageyama S, Tsuneyoshi T, Tsugane S, Sugimura H. 2009. Altered expression of the human base excision repair gene NTH1 in gastric cancer. Carcinogenesis 30: 1345-52.

Halford SE, Rowan AJ, Lipton L, Sieber OM, Pack K, Thomas HJ, Hodgson SV, Bodmer WF, Tomlinson IP. 2003. Germline mutations but not somatic changes at the MYH locus contribute to the pathogenesis of unselected colorectal cancers. Am J Pathol 162: 1545-8.

Jones S, Emmerson P, Maynard J, Best JM, Jordan S, Williams GT, Sampson JR, Cheadle JP. 2002. Biallelic germline mutations in MYH predispose to multiple colorectal adenoma and somatic G:C-->T:A mutations. Hum Mol Genet 11: 2961-7.

Kanter-Smoler G, Björk J, Fritzell K, Engwall Y, Hallberg B, Karlsson G, Grönberg H, Karlsson P, Wallgren A, Wahlström J, Hultcrantz R, Nordling M. 2006. Novel findings in Swedish patients with MYH-associated polyposis: mutation detection and clinical characterization. Clin Gastroenterol Hepatol 4: 499-506.

Kasai H, Nishimura S. 1991. Formation of 8-hydroxyguanosine in DNA by oxygen radicals and its biological significance. In Sies H. (ed.), Oxidative Stress: Oxidants and Antioxidants. Academic Press, London, UK, p 99-116.

Kundu S, Brinkmeyer MK, Livingston AL, David SS. 2009. Adenine removal activity and bacterial complementation with the human MutY homologue (MUTYH) and Y165C, G382D, P391L and Q324R variants associated with colorectal cancer. DNA Repair (Amst) 8: 1400-10.

Lejeune S, Guillemot F, Triboulet JP, Cattan S, Mouton C; PAFNORD Group, Porchet N, Manouvrier S, Buisine MP. 2006. Low frequency of AXIN2 mutations and high frequency of MUTYH mutations in patients with multiple polyposis. Hum Mutat 27: 1064.

Lynch HT, Lynch JF, Lynch PM, Attard T. 2008. Hereditary colorectal cancer syndromes: molecular genetics, genetic counseling, diagnosis and management. Fam Cancer 7: 27-39.

Molatore S, Russo MT, D'Agostino VG, Barone F, Matsumoto Y, Albertini AM, Minoprio A, Degan P, Mazzei F, Bignami M, Ranzani GN. 2010. MUTYH mutations associated with familial adenomatous polyposis: functional characterization by a mammalian cell-based assay. Hum Mutat 31: 159-66.

Nielsen M, Joerink-van de Beld MC, Jones N, Vogt S, Tops CM, Vasen HF, Sampson JR, Aretz S, Hes FJ. 2009. Analysis of MUTYH genotypes and colorectal phenotypes in patients With MUTYH-associated polyposis. Gastroenterology 136: 471-6.

Ohtsubo T, Nishioka K, Imaiso Y, Iwai S, Shimokawa H, Oda H, Fujiwara T, Nakabeppu Y. 2000. Identification of human MutY homolog (hMYH) as a repair enzyme for 2-hydroxyadenine in DNA and detection of multiple forms of hMYH located in nuclei and mitochondria. Nucleic Acids Res 28: 1355-64.

Peterlongo P, Mitra N, Sanchez de Abajo A, de la Hoya M, Bassi C, Bertario L, Radice P, Glogowski E, Nafa K, Caldes T, Offit K, Ellis NA. 2006. Increased frequency of disease-causing MYH mutations in colon cancer families. Carcinogenesis 27: 2243-9.

Russell AM, Zhang J, Luz J, Hutter P, Chappuis PO, Berthod CR, Maillet P, Mueller H, Heinimann K. 2006. Prevalence of MYH germline mutations in Swiss APC mutation-negative polyposis patients. Int J Cancer 118: 1937-40.

Sampson JR, Dolwani S, Jones S, Eccles D, Ellis A, Evans DG, Frayling I, Jordan S, Maher ER, Mak T, Maynard J, Pigatto F, Shaw J, Cheadle JP. 2003. Autosomal recessive colorectal adenomatous polyposis due to inherited mutations of MYH. Lancet 362: 39-41.

Shibutani S, Takeshita M, Grollman AP. 1991. Insertion of specific bases during DNA synthesis past the oxidation-damaged base 8-oxodG. Nature 349: 431-4.

Shinmura K, Tao H, Goto M, Igarashi H, Taniguchi T, Maekawa M, Takezaki T, Sugimura H. 2004. Inactivating mutations of the human base excision repair gene NEIL1 in gastric cancer. Carcinogenesis 25: 2311-7.

Shinmura K, Yamaguchi S, Saitoh T, Takeuchi-Sasaki M, Kim SR, Nohmi T, Yokota J. 2000. Adenine excisional repair function of MYH protein on the adenine:8-hydroxyguanine base pair in double-stranded DNA. Nucleic Acids Res 28: 4912-8.

Sieber OM, Lipton L, Crabtree M, Heinimann K, Fidalgo P, Phillips RK, Bisgaard ML, Orntoft TF, Aaltonen LA, Hodgson SV, Thomas HJ, Tomlinson IP. 2003. Multiple colorectal adenomas, classic adenomatous polyposis, and germ-line mutations in MYH. N Engl J Med 348: 791-9.

Slupska MM, Luther WM, Chiang JH, Yang H, Miller JH. 1999. Functional expression of hMYH, a human homolog of the *Escherichia coli* MutY protein. J Bacteriol 181: 6210–3.

Takao M, Zhang QM, Yonei S, Yasui A. 1999. Differential subcellular localization of human MutY homolog (hMYH) and the functional activity of adenine:8-oxoguanine DNA glycosylase. Nucleic Acids Res 27: 3638-44.

Tao H, Shinmura K, Suzuki M, Kono S, Mibu R, Tanaka M, Kakeji Y, Maehara Y, Okamura T, Ikejiri K, Futami K, Yasunami Y, Maekawa T, Takenaka K, Ichimiya H, Imaizumi N, Sugimura H. 2008. Association between genetic polymorphisms of the base excision repair gene MUTYH and increased colorectal cancer risk in a Japanese population. Cancer Sci 99: 355-60.

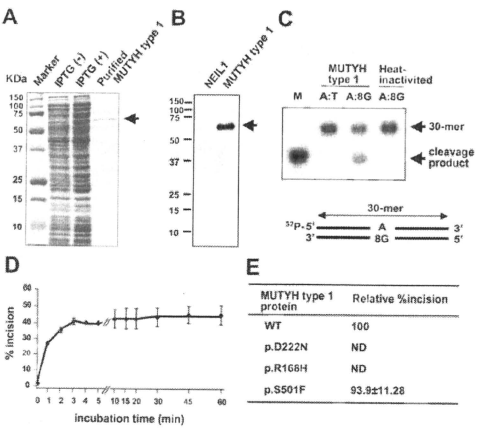
Vogt S, Jones N, Christian D, Engel C, Nielsen M, Kaufmann A, Steinke V, Vasen HF, Propping P, Sampson JR, Hes FJ, Aretz S. 2009. Expanded extracolonic tumor spectrum in MUTYH-associated polyposis. Gastroenterology 137: 1976-85.

Wooden SH, Bassett HM, Wood TG, McCullough AK. 2004. Identification of critical residues required for the mutation avoidance function of human MutY (hMYH) and implications in colorectal cancer. Cancer Lett 205: 89-95.

Yamane A, Shinmura K, Sunaga N, Saitoh T, Yamaguchi S, Shinmura Y, Yoshimura K, Murakami H, Nojima Y, Kohno T, Yokota J. 2003. Suppressive activities of OGG1 and MYH proteins against G:C to T:A mutations caused by 8-hydroxyguanine but not by benzo[a]pyrene diol epoxide in human cells in vivo. Carcinogenesis 24: 1031-7.

Yanaru-Fujisawa R, Matsumoto T, Ushijima Y, Esaki M, Hirahashi M, Gushima M, Yao T, Nakabeppu Y, Iida M. 2008. Genomic and functional analyses of MUTYH in Japanese patients with adenomatous polyposis. Clin Genet 73: 545-53.

SUPPORTING INFORMATION



MUTYH type 1 protein	Relative %incision
WT	100
p.D222N	ND
p.R168H	ND
p.S501F	93.9±11.28

Supp. Figure S1. Measurement of DNA glycosylase activity of MUTYH type 1 protein. (**A**) Purification of wild-type MUTYH type 1 protein resolved by SDS-polyacrylamide gel electrophoresis (PAGE) and stained with Coomassie Brilliant Blue (CBB). The human MUTYH type 1 cDNA sequence was inserted into a pET25b(+) expression vector (Novagen, Darmstadt, Germany). *E. coli* BL21-CodonPlus (DE3)-RP competent cells (Stratagene, La Jolla, CA) were transformed with the MUTYH-pET25b vector and cultured at 37°C until an A_{600} of 0.6. After incubation with 0.1 mM IPTG at 20°C for 12 h, MUTYH-His_6 protein was purified with TALON metal affinity resins (Clontech, Palo Alto, CA) and a TALON 2-ml disposable gravity column (Clontech). The protein was then dialyzed against buffer containing 10 mM sodium phosphate (pH 7.6), 50 mM NaCl, 0.5 mM DTT, 0.1 mM EDTA, 0.5 mM PMSF, 2 μg/ml pepstatin, 2 μg/ml leupeptin, 50 μM chymostatin, and 10% glycerol. Lysates of *E. coli* culture without or with IPTG induction and purified MUTYH type 1 protein are shown. The arrow points to the MUTYH-His_6 protein band. (**B**) Western blot of purified wild-type MUTYH type 1 protein tagged with His_6. Purified recombinant NEIL1-His_6 protein, which was prepared by using pET25b(+) vector (Novagen) and *E. coli* BL21-CodonPlus (DE3)-RP cells (Stratagene) previously (Shinmura et al., 2004), was included as a negative control. Purified recombinant protein was mixed with an equal volume of 2x SDS sample buffer and boiled. A 2 μg protein was subjected to SDS-PAGE and electrophoretically transferred to a polyvinylidene difluoride membrane (GE Healthcare Bio-Science Corp., Piscataway, NJ). The membrane was blocked with non-fat milk and incubated with an anti-MUTYH polyclonal antibody (Ohtsubo et al., 2000). After washing, the membrane was incubated with an anti-rabbit HRP-conjugated secondary antibody (GE Healthcare Bio-Science Corp.). The membrane was then washed, and immunoreactivity was visualized with an ECL Plus chemiluminescence system (GE Healthcare Bio-Science Corp.). MUTYH-His_6 protein is indicated by the arrow. (**C**) The DNA glycosylase activity of wild-type MUTYH type 1 protein on double-stranded DNA containing an A:8-hydroxyguanine (8-OHG). 30-mer oligonucleotides containing and not containing a single 8-OHG (5'-CTG GTG GCC TGA C[8-OHG or T]C ATT CCC CAA CTA GTG-3') were chemically synthesized and purified by PAGE (Japan Bio Services, Saitama, Japan). Complementary oligonucleotides containing an adenine opposite the 8-OHG or T were ^{32}P-labeled at the 5' terminus with a MEGALABEL kit (Takara, Osaka, Japan) and a [γ-^{32}P]ATP (PerkinElmer, Tokyo, Japan), and then annealed to oligonucleotides containing a single 8-OHG or T. The reaction mixture containing 20 mM sodium phosphate (pH 7.6), 100 mM NaCl, 0.5 mM DTT, 0.5 mM EDTA, 5 μM $ZnCl_2$, 1.5% glycerol, 2.5 nM labeled oligonucleotide, 50 μg/ml BSA, and purified MUTYH protein was incubated at 37°C, and the mixture was treated with 0.1 M

NaOH at 95°C for 4 min. After adding denaturing formamide dye to the mixture, it was heated at 95°C for 3 min, and subjected to 20% PAGE. A ^{32}P-labeled marker oligonucleotide was used as a size marker for the cleavage products. The radioactivity of intact and cleaved oligonucleotides was quantified by using an FLA-3000 fluoroimage analyzer (Fuji Film, Tokyo, Japan) and ImageGauge software (Fuji Film) (Goto et al., 2009). The intact 30-mer oligonucleotides and cleavage products are indicated by the arrows. Heat-inactivation of the MUTYH protein was accomplished by heating the protein at 100°C for 5 min. 8G means 8-hydroxyguanine. (**D**) Time-course assay of cleavage of DNA containing an A:8-OHG by wild-type MUTYH type 1 protein. The MUTYH type 1 protein (260 fmole) was incubated at 37°C for 0 - 60 min with double-stranded oligonucleotide containing an A:8-OHG (50 fmole). The amount of cleavage products as a proportion of total oligonucleotides was calculated as % incision. The % incision values are shown as means ± standard errors of data from three independent experiments. (**E**) DNA glycosylase activities of wild-type MUTYH type 1 protein and their variant proteins on an A:8-OHG substrate. DNA cleavage activities of MUTYH type 1 proteins were measured at 37°C for 15 min. The amount of cleavage products as a proportion of total oligonucleotides was calculated as % incision, and the % incision of each variant-type MUTYH protein is shown relative to that of wild-type (WT) MUTYH protein, which has been set equal to 100. Values are means ± standard errors of data from three independent experiments. ND, not detected.

Pathology

Pathology International 2010; **60**: 543–550

doi:10.1111/j.1440-1827.2010.02561.x

Review Article

Fluorescence *in situ* hybridization analysis with a tissue microarray: 'FISH and chips' analysis of pathology archives

Haruhiko Sugimura,[1] Hiroki Mori,[1] Kiyoko Nagura,[1] Shin-ichiro Kiyose,[1] Tao Hong,[1] Masaru Isozaki,[2] Hisaki Igarashi,[1] Kazuya Shinmura,[1] Akio Hasegawa,[2] Yasuhiko Kitayama[3] and Fumihiko Tanioka[4]

[1]*Department of Pathology, Hamamamatsu University School of Medicine, Higashi-ward, Hamamatsu,* [2]*Department of Diagnostic Pathology and Laboratory Medicine, Odawara Municipal Hospital, Odawara,* [3]*Department of Pathology, International University of Health and Welfare Mita Hospital, Tokyo* [4]*Division of Pathology and Laboratory Medicine, Iwata City Hospital, Iwata, Japan*

Practicing pathologists expect major somatic genetic changes in cancers, because the morphological deviations in the cancers they diagnose are so great that the somatic genetic changes to direct these phenotypes of tumors are supposed to be correspondingly tremendous. Several lines of evidence, especially lines generated by high-throughput genomic sequencing and genome-wide analyses of cancer DNAs are verifying their preoccupations. This article reviews a comprehensive morphological approach to pathology archives that consists of fluorescence *in situ* hybridization with bacterial artificial chromosome (BAC) probes and screening with tissue microarrays to detect structural changes in chromosomes (copy number alterations and rearrangements) in specimens of human solid tumors. The potential of this approach in the attempt to provide individually tailored medical practice, especially in terms of cancer therapy, is discussed.

Key words: copy number alteration, fluorescence *in situ* hybridization (FISH), formalin-fixed paraffin- embedded (FFPE) tissue, pathology archives, tissue microarray (TMA)

INTRODUCTION

Extreme copy number alterations (aneuploidy) are the norm in human solid tumors.[1–3] Karyotyping solid tumors is so labourious[4] that only limited information on chromosomal abnormalities in human solid tumors *in situ* was available until recently. The latest methodologies that involve the use of human genome information, however, have provided us techniques that make it possible to identify any locus-specific chromosomal changes in a tumor. Several examples of applications of these state-of-the-art methodologies are essential diagnostic tools in diagnostic laboratories to, for example, identify translocation in certain solid tumors.[5–7]

New information is being obtained every day in genetic research on human solid tumors (especially carcinomas). The high-throughput, 'genome-wide' approach to genetic changes in human tumors has been widely adopted in every branch of medicine, and it is now known that there are extensive somatic changes, including multiple point mutations,[8,9] copy number alterations,[10,11] and further complex rearrangements[12] in every kind of tumor. Since most of these somatic changes have been identified in the analysis of the DNAs of advanced primary tumors and tumor cell lines, questions about when and where these genetic changes occur during cancer development in the human body remain to be answered by pathologists. Human pathology archives contain specimens of human tumors in various stages of development, from the incipient stage to the metastatic stage, and they are a treasure trove in the post-human-genome-sequencing era. The know-hows of two methods are important, especially for diagnostic pathologists: intensive application of bacterial artificial chromosome (BAC) clones as probes that have exact 'addresses' in the whole genome and construction of tissue microarrays (TMAs) which consist of hundreds of tissue specimens on a single slide. Using a combination of these two know-hows is a strategy that facilitates identification of changes at any genomic locus in several hundreds of tissue samples at once.

Correspondence: Haruhiko Sugimura, MD. PhD, Department of Pathology, Hamamamatsu University School of Medicine, 1-20-1, Handayama, Higashi-ward, Hamamatsu 431-3192, Japan. Email: hsugimur@hama-med.ac.jp

Declaration of conflicts of interest to declare: Shinichiro Kiyose is an employee of Jokoh Inc.

Received 23 March 2010. Accepted for publication 7 April 2010.

Use of some of the specific BAC probes has already acquired a niche in routine examinations in diagnostic laboratories as a means of verifying a diagnosis, selecting subjects for particular molecularly targeted therapies, and for predicting recurrence.[13–20] Use of BAC probes by diagnostic pathologists, however, is still not widespread because of the difficulty of accessing and making the BAC probes for interests of their own. In this article we review the various facets of the latest advances in the application of BAC probes to diagnostic pathology and describe some of our own experiences with using many BAC probes to investigate pathology archives. We think that using numerous BAC probes will soon become a popular diagnostic practice, the same as the current use of monoclonal antibodies.

Actually, several ambitious pathology laboratories around the world that possess these methods in their arsenals, have started to propose an agenda of TMA-FISH ('Fish and chips') approaches to tumor DNA analysis.[21–32] The recent observation of repositioning of chromosomal loci during carcinogenesis has further encouraged the analysis of human tumor specimen in various clinicopathological settings.[31,32]

APPLICATIONS OF FISH TO DETECTION OF COPY NUMBER ALTERATIONS IN HUMAN TUMORS IN PATHOLOGY ARCHIVES

The development and modifications of the FISH procedure, especially for use in formalin-fixed-paraffin-embedded (FFPE) tissues have been extensively reviewed.[30] Equivalent hybridization efficiency of probes for the arrayed pieces of tissue after different fixation times and storage methods is necessary to correctly evaluate copy number amplification. In many studies, the FISH procedure has been performed as a means of validation, that is, to verify amplification data generated by other methodologies, such as by quantitative PCR, array-based comparative genomic hybridization (aCGH), and single nucleotide polymorphism (SNP) arrays,[33] and comparisons between methods and the interpretations of the results obtained by each method have sometimes been a matter of controversy.[34,35] FISH analysis, especially of FFPE tissues, is often technically demanding, and standardized quality control, which is very important in practical settings, has just begun. There are large inconsistencies between the prevalence of amplification of well-known and familiar genes that we consider clinically useful and that are routinely used in practice without rigorous quality control guidelines.[34,35] Thousands of BAC clones are commercially available, and, in theory, any of them can be used as FISH probes. The BAC clones or labeled probes can be ordered from at least two Japanese companies (Advanced GenoTechs Co., Tsukuba, Japan; GSP laboratory, Kawasaki, Japan). When we use these BAC clones for FISH procedures in paraffin-embedded tissue sections, several steps must be carefully performed including labeling and hybridizing them to DNA. The BAC clone must be confirmed to be the correct one, because assignments of BAC clones often change to reflect the daily process of refining the human genome database. The information on exact location of each BAC probe according to the most recent Build (Build 37 in March, 2010) of the human genome is necessary. Although the reason is usually unclear, some BAC clones hybridize with multiple sites (more than 4) in normal interphase cells, and logically they cannot be used to evaluate human tumors. Thus, commercial BAC probes must be tested to determine whether they are hybridized to the two corresponding sites (or two pairs of the signals on the sister chromatids) in the metaphase chromosome spread before they are applied to human tissues containing cancer cells (Fig. 1). Sequencing of part of the BAC probes is of some help in further confirming the correctness of the BAC probes.

In addition to the above-mentioned hurdles to obtaining the right BAC probes, there is another stumbling block to completion of a FISH procedure: the labeling step. Several labeling methods are available, and some are commercially available and packaged in the form of a kit. Sufficiently efficient labeling is sometimes achieved only in an heuristic manner.

The following limitations in interpretation must be considered when using a FISH procedure to enumerate chromosomes in paraffin-embedded tissue sections. The signals can be weak for many reasons. Clinical practice has been standardized only for the system for detection of HER2 amplification in breast cancer cases.[36] The merits of protease treatment, microwave treatment, heating, and other treatments such as using various detergents have been debated. Some 'pre-treatment' kits are commercially available, but retrieval efficiency usually depends on the condition of the specimen, and individual adjustments must be made each time in each laboratory. For example, the recommended pre-treatment to augment signal strength in the two kits available, the Hercep test (Abbott, Tokyo, Japan) and the HISTRA (Jokoh, Tokyo, Japan) are different.[37] Based on our own experience, one technical tip for generating stable, sensitive signals in pathology archives that have been fixed by various methods and stored for a long period is appropriate, careful pretreatment with protease.

Since overlapping cells and cells whose nuclei are partially cut cause miscounting of the numbers of signals, cut-off values must be set based on preliminary evaluation of the signals in several non-tumorigenic tissues.[38,39] Several quality controls are necessary before applying the new probes to clinical uses the same as for the HER2 probe.

MERITS OF TMAs FOR SCREENING BY FISH

The preparation of FISH probes is a tedious task that includes several hurdles described in the previous section

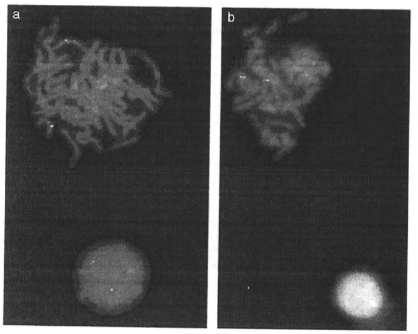


Figure 1 A metaphase spread (top) for testing a bacterial artificial chromosome (BAC) probe. (**a**) Two signals (green) with the corresponding centromere probe (red signals) are seen in the same chromosome. (**b**) Red and green signals are seen in different chromosomes, although they were supposed to be in the same chromosome according to the information in the database. Interphase cells exhibit two (pairs of) signals each (bottom).

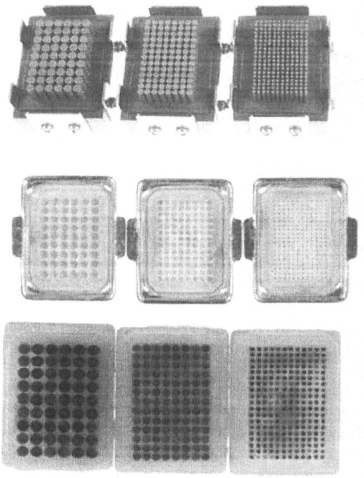


Figure 2 Tissue microarray gauges, prefabricated recipient blocks with holes, commercially available, and embedded blocks from top to bottom. The core diameters are 3 mm, 2 mm, and 1 mm in diameter (left to right).

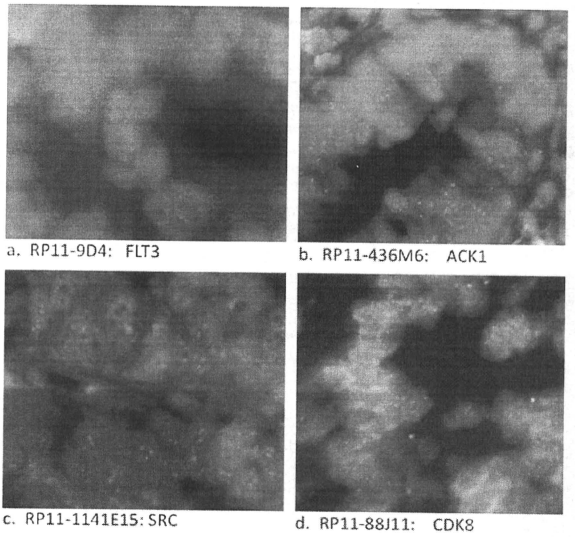


Figure 3 Amplification of kinase loci detected in FFPE tissues from the undifferentiated carcinoma of the stomach. Symbol genes are FMS related tyrosine kinase 3 (FLT3) (**a**), Activated p21CDC42 kinase(ACK1) (**b**), V-SRC avian sarcoma (Schimidt-Ruppin A-2) viral oncogene (SRC) (**c**), and Cyclin dependent kinase 8 (CDK8) (**d**). The probes were labeled with Spectrum Orange (Abbott, Abbott Park, IL, USA), and the nuclei were stained with 4, 6-diamino-2-phenyl indole dihydrochloride (DAPI, Abbot). The method is described in detail in the previous literature.

that must be overcome. Many investigators have constructed tissue microarrays for efficient use of probes they had laboriously prepared, especially in retrospective studies. The idea of embedding many pieces in a single block existed in the early days of anatomical pathology, but several embedding instruments for this purpose recently became popular, and technical refinements are under way. One well-circulating brand of microarray instruments is Beecher Instruments (Beecher Instruments, Inc. Sun Prairie, WI, USA). Their models have 0.6 mm, 1 mm, 2 mm cylinders, and the Azumaya model KIN-1 and model FIN-3 (Azumaya Cooperation, Tokyo) have wider cores that are 2 mm, 3 mm, 5 mm, and 7 mm in diameter. There are pros and cons in regard to using the smaller cores, and several problems encountered in using the instruments with various sized-cores are addressed in the instructions included with each of the instruments. A validation study in regard to possible sampling error when small core specimens are collected was performed and the results were published.[40] Very recently, donor blocks containing multiple slots and an apparatus for making them have become commercially available (Fig. 2) (Patent Application 2009-028167), and many other variations will become available commercially. In addition to genomic and immunohistochemical studies, a proteomics approach by imaging mass spectrometry on a TMA platform is also feasible.[41]

APPLICATION OF TMAS TO BIO-BANKS AND ETHICS CONSIDERATION

Preparations of TMAs and requests to prepare them will become more frequent in both investigative and diagnostic pathology laboratories, and as members of institutional review boards (IRBs) pathologists are sometimes responsible for appropriate control of these TMA bioresources. The categories of pathology specimens are described in several documents and on several websites,[42–44] and IRBs are required to facilitate research proposals of making or using TMAs to implement the research smoothly and ethically.

TMAs are a major component of tissue banks,[45] which are tissue resources for future personalized medicine and national and international bio-bank systems are now being established (websites: http://www.stn.org.sg, http://www.ukbiobank.ac.uk, http://www.bbmri.eu, and http://www.src.riken.go.jp/english/project/person/index.html).

COPY NUMBER ALTERATIONS DETECTED BY aCGH AND SNP ARRAY: USEFUL DATA FOR FISH ANALYSIS

Data on copy number alterations in solid tumors deposited in databases and publications have rapidly accumulated since the introduction of aCGH led to the discovery of many tumor-specific and stage-specific gains or losses of particular regions of chromosomes.[46–49] Much of the information generated by aCGH itself is used as a diagnostic or prognostic tool in pathology laboratories.[50–52] Information on genome-wide genetic changes in cancer DNA are now viewed as academic knowledge that is only useful to the graduate students and researchers, but sooner or later it will be an essential tool of the diagnostic pathologist facing daily challenges in diagnosis and management. There are many issues in conventional pathology research and practice to which human genome data can be applied.[53] Sano *et al.* conducted a chromosome-wide survey to the archives of adenomatous hyperplasia of the lung[38] and proposed 'adenocarcinoma in adenomatous hyperplasia' as an early stage of carcinogenesis of lung adenocarcinoma. Although the tools were genetic, the story they told was morphological. Very recently, a more powerful system, an SNP array platform containing more than 500 000 SNP sites has come into widespread use, and copy number estimation by several algorithms has facilitated identification of copy number changes, such as loss of heterozygosity, uniparental disomy, and amplification, in many clinical tumors. Midorikawa *et al.*[54] integrated the data based on pathological examination of 'nodule in nodule' in resected liver tissue with the results of a comprehensive copy number survey with the Affymetrix SNP array that were confirmed by FISH, and succeeded in clarifying genetic process in human hepatocarcinogenesis in detail.

Research on structural changes and balanced translocation of chromosomes in solid epithelial tumors is also a cutting edge area of research today.[7,51,55–57] The numbers of candidate probes that should be investigated for clinical significance seem huge. Several points need to be addressed when interpreting the results generated when an aCGH array and SNP array are used to analyze a human tumor genome. The first point is that many platforms are available to analyze copy number alterations, and a few papers on the characteristics of each platform have been published.[58–60] Furthermore, since many algorithms are available to enumerate copy numbers on the same platform, the characteristics of the platforms themselves and the benchmarks of the algorithms need to be known. Most algorithms for estimating the copy number of loci set the reference dosage of tumor autosomes as 2 (diplotype), but this reference number is not valid for most common epithelial malignant tumors. Ng *et al.*[61] recently refined the protocol for ploidy-specific copy number estimation, and obtained a better threshold for detecting CNA in cell lines, and Suzuki *et al.* performed a benchmark test of two widely used algorithms and extensively characterized the features of the algorithms in terms of different formulas for setting the gain or loss thresholds of genetic loci.[62] Because of the intrinsic limitations of each method, two or three methods need to be used simultaneously for the same tumor.