Clinical responses of large cell neuroendocrine carcinoma of the lung to cisplatin-based chemotherapy

Shigeo Yamazaki[a,d], Ikuo Sekine[a,*], Yoshihiro Matsuno[b], Hidefumi Takei[c], Noboru Yamamoto[a], Hideo Kunitoh[a], Yuichiro Ohe[a], Tomohide Tamura[a], Tetsuro Kodama[a], Hisao Asamura[c], Ryosuke Tsuchiya[c], Nagahiro Saijo[a]

[a] Division of Thoracic Oncology and Internal Medicine, National Cancer Center Hospital, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan
[b] Division of Clinical Laboratory, National Cancer Center Hospital, Tokyo
[c] Division of Thoracic Surgery, National Cancer Center Hospital, Tokyo
[d] Department of Surgery, Keiyu-kai Sapporo Hospital, Sapporo, Japan

Summary
Background: The efficacy of chemotherapy in patients with large cell neuroendocrine carcinoma of the lung (LCNEC) remains unclear.
Methods: Patients with LCNEC who received cisplatin-based chemotherapy were identified by reviewing 567 autopsied and 2790 surgically resected lung cancer patients. The clinical characteristics and objective responses to chemotherapy in these patients were analyzed.
Results: Overall, 20 cases of LCNEC were identified, including stage IIIA ($n = 3$), stage IIIB ($n = 6$), stage IV ($n = 6$) and postoperative recurrence ($n = 5$) cases. Six patients had received prior chemotherapy, and 14 were chemo-naive patients. The patients had received a combination of cisplatin and etoposide ($n = 9$), cisplatin, vindesine and mitomycin ($n = 6$), cisplatin and vindesine ($n = 4$), or cisplatin alone ($n = 1$). One patient showed complete response and nine showed partial response, yielding an objective response rate of 50%. The response rate did not differ between patients with the initial diagnosis of SCLC and those with the initial diagnosis of NSCLC, however, the response rate in chemo-naive patients (64%) was significantly different from that in previously treated patients (17%).
Conclusions: Our results suggest that the response rate of LCNEC to cisplatin-based chemotherapy was comparable to that of SCLC.
© 2005 Elsevier Ireland Ltd. All rights reserved.

* Corresponding author. Tel.: +81 3 3542 2511; fax: +81 3 3542 3815.
E-mail address: isekine@ncc.go.jp (I. Sekine).

1. Introduction

Pulmonary neuroendocrine tumors include a spectrum of four clinicopathological entities classified on the basis of the morphological and biological features: typical carcionoid and atypical carcinoid, which are tumors of low to intermediate grade malignancy, and large cell neuroendocrine carcinoma (LCNEC) and small cell carcinoma (SCLC), which are high-grade malignant tumors. Travis et al. proposed the term LCNEC in 1991 [1], for classifying a type of poorly differentiated high-grade carcinoma characterized by a neuroendocrine appearance under light microscopy. LCNEC exhibits more prominent cellular pleomorphism and higher mitotic activity than the atypical carcinoid (AC), and is distinguished from SCLC by the tumor cell size and chromatin morphology. Although several different terminologies and classifications have been proposed previously, and even the present classification of pulmonary neuroendocrine tumors lacks uniform definition criteria, this class of tumors could become widely accepted and included in the updated histological classification of the World Health Organization [2].

The clinical features of LCNEC have not yet been completely clarified. The prognosis of patients with surgically resected LCNEC is reported to be intermediate between that of AC and SCLC [3–5], and the same as that of resected NSCLC, except that stage I LCNEC has a poorer prognosis than stage I non-small cell lung cancer (NSCLC) [6]. To the best of our knowledge, however, there are no studies that have examined the role of chemotherapy for LCNEC and the prognosis of patients with unresectable LCNEC, even though several reports have been published on the association between response to chemotherapy and the neuroendocrine differentiation of NSCLC [7–9]. The appropriate treatment of unresectable LCNEC, therefore, remains unclear. In the present study, we attempted to investigate the effectiveness of chemotherapy with cisplatin-based regimens for LCNEC in patients with unresectable and recurrent LCNEC.

2. Materials and methods

Eighty-seven of 2790 patients with primary lung cancer who underwent tumor resection from 1982 to 1999 at the National Cancer Center Hospital were found to have tumors with the histological characteristics of LCNEC [6]. Of these, five had received cisplatin-based chemotherapy at the time of recurrence, and were enrolled as subjects of this study. In addition, 303 of 567 patients who were autopsied from 1983 to 1997 at the National Cancer Center Hospital who had the following histological diagnoses were first selected: SCLC ($n = 112$), poorly differentiated adenocarcinoma ($n = 99$), large cell carcinoma ($n = 58$), poorly differentiated squamous cell carcinoma ($n = 29$), poorly differentiated adenosquamous carcinoma ($n = 2$), LCNEC ($n = 2$), and carcinoid ($n = 1$). Of these, 161 had received cisplatin-based chemotherapy were selected for a pathological review. Finally, specimens from 17 of these 161 cases were found to have histological characteristics consistent with the diagnosis of LCNEC, and were selected as subjects of this study. We focused on cisplatin, because since the 1980s, cisplatin has been the only anticancer agent with proven efficacy against both SCLC and NSCLC [10,11]; we, therefore, considered that the effectiveness of chemotherapy for LCNEC could be reasonably evaluated if cisplatin were included in the regimen. Cases which had received adjuvant chemotherapy without evaluable lesions were excluded from the analysis.

All the available paraffin-embedded tissue sections stained with hematoxylin—eosin were reviewed. We classified LCNEC according to the histopathological criteria in the WHO classification [2]. Immunohistochemical analysis was performed to confirm the neuroendocrine features of the tumors. For this purpose, formalin-fixed paraffin sections were stained for a panel of neuroendocrine markers, including chromogranin A (CGA), synaptophysin (SYN), and neural cell adhesion molecule (NCAM), using standard methods. The intensity of immunostaining for these markers was scored as follows: +, when the proportion of stained tumor cells was >50%; ±, when 10—50% of tumor cells were stained; and —, when <10% of tumor cells were stained, as previously described [6]. One case included in this study had the typical histological features of LCNEC, but no neuroendocrine features as determined by the immunohistochemical analysis. For specimens obtained after treatment, we routinely confirmed that the histopathological and morphological features showed no changes due to treatment as compared with the pretreatment biopsy or cytologic specimens. Such cases for which no pretreatment samples were available were excluded from the study; since it has been reported that histological changes may occur after treatment in SCLC [12], we were concerned that misdiagnosis might occur if the same were also true for LCNEC.
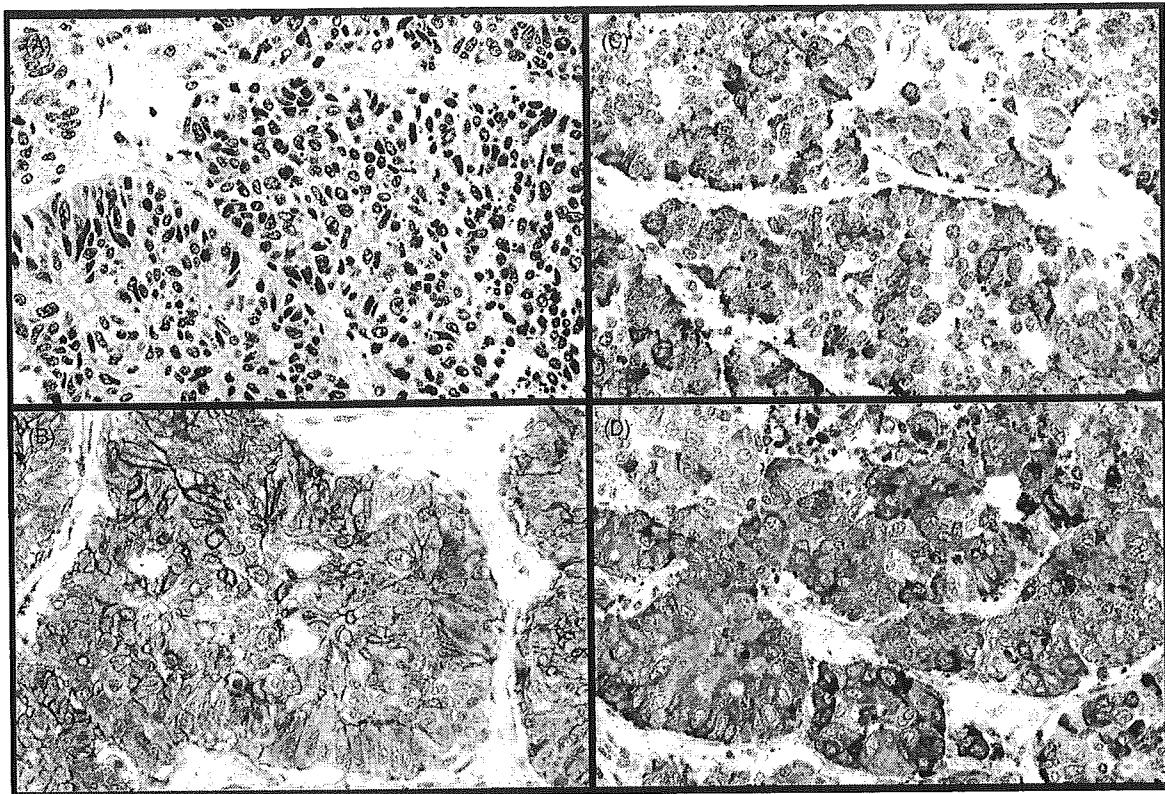
Fig. 1 Case no. 2, 57-year-old man. (A) The tumor cells which are large-sized, polygonal in shape and have a low nuclear-cytoplasmic ratio, are arranged in organoid nests and trabeculae (H&E stain, ×200). Positive staining for neural cell adhesion molecule (B), chromogranin A (C), and synaptophysin (D) (immunostain, ×400).

Clinical information about the cases was obtained from the medical records. The clinical disease staging was reassessed according to the latest International Union Against Cancer (UICC) staging criteria [13]. The response to chemotherapy and overall survival rate were assessed retrospectively. The objective tumor response was evaluated according to the WHO criteria published in 1979 (WHO, 1979) [14]. The survival time was measured from the date of start of chemotherapy with a cisplatin-containing regimen. Survival curves were drawn using the Kaplan—Meier method [15]. Drug toxicity could not be assessed as the study was a retrospective one and records were often incomplete.

3. Results

Overall, 22 cases were recognized as having tumors with histological characteristics consistent with LC-NEC among the autopsied and surgically resected

cases of primary lung cancer that had received cisplatin-based chemotherapy and had evaluable lesions; of these 17 were autopsied cases and five were surgically resected cases. Two of the autopsied cases were excluded, because no pretreatment pathological or cytological samples were available. The typical microscopic appearance of the tumor specimens is shown in Fig. 1A. The specimen sources for the prechemotherapy-diagnosis included surgically resected specimens ($n = 5$), biopsy specimens ($n = 9$), and cytology specimens ($n = 6$). The histological and cytological findings in the specimens obtained before chemotherapy were consistent with those in the specimens obtained after chemotherapy. We therefore finally enrolled 20 cases in this study. The initial pathologic diagnoses in these patients were as follows: small cell carcinoma ($n = 10$), poorly differentiated adenocarcinoma ($n = 6$), large cell carcinoma ($n = 2$), undifferentiated carcinoma ($n = 1$), and poorly differentiated carcinoma ($n = 1$) (Table 1). None of the cases had been labeled as LCNEC at the time of initial diagnosis, probably because the concept of LCNEC

Table 1 Patient characteristics

Characteristics	N	%
No. of patients	20	
Sex		
Male	18	90
Female	2	10
Age, median (range)	58 (37—74)	
Smoking history		
Yes	19	95
No	1	5
Performance status		
1—2	19	95
>2	1	5
Initial pathological diagnosis		
Small cell carcinoma	10	50
Adenocarcinoma	6	30
Large cell carcinoma	2	10
Others	2	10
Clinical stage at the start of chemotherapy		
IIIA	3	15
IIIB	6	30
IV	6	30
Postoperative recurrence	5	25
Prior treatment		
None	10	50
Surgery	4	20
Radiotherapy	2	10
Chemotherapy without cisplatin	6	30

Table 2 Staining for neuroendocrine markers in 20 LCNECs

Case	NCAM	CGA	SYN
1	+	+	+
2	+	+	+
3	+	+	+
4	±	+	+
5	+	+	+
6	+	+	+
7	−	+	−
8	+	−	−
9	−	−	−
10	−	+	±
11	+	−	+
12	+	+	+
13	+	+	+
14	+	−	±
15	+	+	+
16	+	−	NA
17	+	−	+
18	+	−	NA
19	+	−	+
20	−	+	+

NCAM, neural cell adhesion molecule; CGA, chromogranin A; SYN, synaptophysin; NA, not assessed.

was not completely accepted at our hospital at that time.

The results of the immunohistochemical staining are shown in Table 2, and a typical case showing positive staining is shown in Fig. 1B and D. Of the 20 LCNECs, 19 expressed at least one of the three general neuroendocrine markers, namely CGA, SYN, and NCAM. Sixteen of the 20 LCNECs exhibited positive staining for NCAM, while one showed equivocal staining. Twelve of the 20 LCNECs showed positive staining for CGA. Thirteen LCNECs showed positive staining for SYN and three showed equivocal staining. Only one case was negative for all the three general neuroendocrine markers, however, this case exhibited the typical histological features of LCNEC on light microscopy.

The clinical characteristics of the patients are summarized in Table 1. The extremely high predominance of men and smokers in this study was comparable to the demographic features of our LCNEC patients treated by surgical resection [6]. Previous chemotherapy was given in six patients: nedaplatin in one and cyclophosphamide-based regimen in five

patients. The chemotherapy regimens used were as follows: cisplatin ($80\,mg/m^2$, day 1) and etoposide ($100\,mg/m^2$, days 1—3) ($n = 9$), cisplatin ($80\,mg/m^2$, day 1), vindesine ($3\,mg/m^2$, days 1 and 8) and mitomycin ($8\,mg/m^2$, day 1) ($n = 6$), cisplatin ($80\,mg/m^2$, day 1) and vindesine ($3\,mg/m^2$, days 1 and 8) ($n = 4$), or cisplatin ($100\,mg/m^2$, day 1) alone ($n = 1$). The median (range) number chemotherapy cycles were 2 (1—6). Of the 20 patients, one showed CR and nine showed PR, yielding an overall response rate of 50% (95% confidence interval, 27.2—72.8%). One CR and four PRs were observed among the cases treated with cisplatin and etoposide, two PRs were found among those treated with cisplatin, vindesine and mitomycin, and three PRs were found among those treated with cisplatin and vindesine. Seven patients showed NC, and three showed progressive disease. While the response rate did not differ between patients with an initial diagnosis of SCLC and those patients with an initial diagnosis of NSCLC, previous chemotherapy affected the response to cisplatin: the response rate in chemo-naive patients was 64%, whereas that in previously treated patients was 17%. The median progression-free survival in the 20 patients was 103 days, median survival was 239 days, 1-year survival rate was 35%, and 2-year survival rate was 15%.

4. Discussion

In this extensive review of over 3000 lung cancer patients, we found considerable difficulty in evaluating the response of LCNEC to systemic chemotherapy. The pathological diagnosis of LCNEC was established in 87 (3.1%) of 2790 patients treated by surgical resection. This low incidence of LCNEC in surgically treated lung cancer patients is comparable to that in other previously published reports: 2.4% (50/2070), 2.9% (22/766), and 3.6% (53/1530) [16—18]. Of the 87 patients, only five who had received cisplatin-based chemotherapy for recurrent tumor that was evaluable for the response. While LCNEC is difficult to diagnose prior to the start of treatment on the basis of the findings in biopsy or cytological specimens, the architectural neuroendocrine features may, more or less, be reflected in these small samples [19,20]. We, therefore, conducted a review of 567 autopsy cases of lung cancer, and identified 15 cases of LCNEC who had received cisplatin-based chemotherapy. We obtained a response rate to cisplatin-based chemotherapy of 50% in these 20 patients with LCNEC, however, the clinical characteristics of patients with medically treatable advanced LCNEC would still remain to be clarified, because autopsy is conducted only in highly selective cases.

Travis et al. suggested that immunohistochemical or electron-microscopic evidence of neuroendocrine features were important to diagnose LCNEC [1]. We assessed the neuroendocrine marker expression by immunohistochemical staining for CGA, SYN, and NCAM. Our cases included one that was negative for all the three neuroendocrine markers examined, but showed the typical histological features of LCNEC, which could be attributable to technical staining problems. Immunohistochemical staining for neuroendocrine tumors is generally recognized as only a supplementary diagnostic tool. In addition, the post-surgical survival rate did not differ between histologically diagnosed cases of LCNEC with neuroendocrine differentiation in marker expression as assessed by immunohistochemical staining and large cell carcinoma with neuroendocrine morphology where the neuroendocrine markers were negative (data not shown). Thus, we decided to include the case with negative staining as LCNEC on the basis of its typical neuroendocrine morphology.

To the best of our knowledge, only one study on the efficacy of chemotherapy in patients with LCNEC has been reported previously. In the study, 13 patients with LCNEC received chemotherapy when relapse was noted after surgical resection, and two (20%) of 10 evaluable patients showed an objective response. The evaluable lesion in these patients, however, was the brain in seven, liver in two, and bone in one patient [21]. Thus, the relatively low response rate in the report may be due to the site of the evaluable lesion. In addition, reports on the correlation between response to chemotherapy and neuroendocrine differentiation of NSCLC may be helpful. Graziano et al. reported that the proportion of NSCLC positive for neuroendocrine markers was higher in responders than in non-responders among 52 NSCLC patients treated by chemotherapy, and that the result suggested a correlation between positivity for neuroendocrine marker expression and the likelihood of response to chemotherapy [7]. On the other hand, others have reported the absence of any correlation between the presence of neuroendocrine differentiation and the response to chemotherapy [8,9]. The neuroendocrine differentiation in NSCLCs in the aforementioned studies was confirmed only by immunohistochemical staining and not on the basis of the morphological definition of LCNEC. Therefore, these groups might have potentially included heterogeneous subtypes of lung carcinoma, such as adenocarcinoma or squamous cell carcinoma, with components of neuroendocrine differentiation. The conflicting conclusions of these studies may, therefore, reflect differences in the biological characteristics of the tumors included in the analysis. Since the definition of LCNEC is based on morphological criteria as well as positivity for neuroendocrine marker expression, LCNEC is may be considered to be a clinically homogeneous group. Therefore, our study of LCNEC may endorse the former reports about the relationship between neuroendocrine differentiation and the sensitivity to chemotherapy.

Objective response to chemotherapy can be observed in only 15—30% of NSCLCs, even when they are treated with regimens containing cisplatin [10]. In SCLC, however, effective combination regimens yield objective response rates in the range of 80—90% [11]. Our study showed an overall response rate of LCNEC of 50% to cisplatin-based chemotherapy, and a response rate of 64% in chemo-naive patients, which seemed to be higher than the response rate of NSCLC to chemotherapy. Considered together, these results suggest that the chemosensitivity of LCNEC is intermediate between that of NSCLC and SCLC, although we were unable to obtain firm evidence from this retrospective study, which included only a small cohort of patients.

Since LCNEC is a relatively rare subtype of lung cancer, a prospective study is difficult to perform, and may only be possible as a multicenter study.

For this purpose, it is an urgent task to establish diagnostic criteria for LCNEC based on examination of biopsy or cytologic specimens. Although the histological definition of LCNEC in surgically resected specimens proposed by Travis et al. is commonly accepted, its diagnostic reproducibility is not satisfactory [22]. It is also difficult to apply the definition to biopsy specimens, in which artifacts can easily be produced and detailed examination may be difficult due to insufficient specimen size. Thus, definitive diagnostic criteria also applicable to biopsy and cytologic specimens are required.

Our study did not include any cases labeled as LCNEC at the time of initial diagnosis. One half of the cases was originally diagnosed as SCLC and the other half as NSCLC, including poorly differentiated adenocarcinoma and large cell carcinoma. This was attributed to the fact that the concept of LCNEC was not clearly defined prior to its being proposed by Travis et al. [1]. Thus, it is possible that patients with LCNEC were included in earlier clinical trials for NSCLC or SCLC. If LCNEC shares the poor prognosis of NSCLC, the reported results of chemotherapy for NSCLC may have been worse in studies in which cases of LCNEC were included. Similarly, the results of clinical studies of SCLC to study their objective response to chemotherapy may also have been worse because of the confounding effects of the inclusion of LCNECs among the cases.

In conclusion, our results suggest that the response rate of LCNEC to cisplatin-based chemotherapy was comparable to that of SCLC. However, because of the retrospective nature of this study and the small sample size, we could not arrive at any definitive conclusion; we, therefore, propose to conduct a prospective study in the future aimed at elucidating the efficacy of chemotherapy for LCNEC. To that end, firm diagnostic criteria for LCNEC need to be established, even when the diagnosis must be based only on examination of biopsy and cytology specimens.

Acknowledgment

References

[1] Travis WD, Linnoila RI, Tsokos MG, Hitchcock CL, Cutler Jr GB, Nieman L, et al. Neuroendocrine tumors of the lung with proposed criteria for large-cell neuroendocrine carcinoma. An ultrastructural, immunohistochemi-cal, and flow cytometric study of 35 cases. Am J Surg Pathol 1991;15:529—53.

[2] Travis W, Corrin B, Shimosato Y, Brambilla E. Histological typing of lung and pleural tumours. Berlin: Springer-Verlag; 1999.

[3] Rusch VW, Klimstra DS, Venkatraman ES. Molecular markers help characterize neuroendocrine lung tumors. Ann Thorac Surg 1996;62:798—809 [discussions 09—10].

[4] Dresler CM, Ritter JH, Patterson GA, Ross E, Bailey MS, Wick MR. Clinical-pathologic analysis of 40 patients with large cell neuroendocrine carcinoma of the lung. Ann Thorac Surg 1997;63:180—5.

[5] Travis WD, Rush W, Flieder DB, Falk R, Fleming MV, Gal AA, et al. Survival analysis of 200 pulmonary neuroendocrine tumors with clarification of criteria for atypical carcinoid and its separation from typical carcinoid. Am J Surg Pathol 1998;22:934—44.

[6] Takei H, Asamura H, Maeshima A, Suzuki K, Kondo H, Niki T, et al. Large cell neuroendocrine carcinoma of the lung: a clinicopathologic study of eighty-seven cases. J Thorac Cardiovasc Surg 2002;124:285—92.

[7] Graziano SL, Mazid R, Newman N, Tatum A, Oler A, Mortimer JA, et al. The use of neuroendocrine immunoperoxidase markers to predict chemotherapy response in patients with non-small-cell lung cancer. J Clin Oncol 1989;7:1398—406.

[8] Schleusener JT, Tazelaar HD, Jung SH, Cha SS, Cera PJ, Myers JL, et al. Neuroendocrine differentiation is an independent prognostic factor in chemotherapy-treated nonsmall cell lung carcinoma. Cancer 1996;77:1284—91.

[9] Carles J, Rosell R, Ariza A, Pellicer I, Sanchez JJ, Fernandez-Vasalo G, et al. Neuroendocrine differentiation as a prognostic factor in non-small cell lung cancer. Lung Cancer 1993;10:209—19.

[10] Ginsberg RJ, Vokes EE, Rosenzweig K. Non-small cell lung cancer. In: DeVita VT, Hellman S, Rosenberg SA, editors. Cancer: principles & practice of oncology. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 925—83.

[11] Murren J, Glatstein E, Pass HI. Small cell lung cancer. In: DeVita VT, Hellman S, Rosenberg SA, editors. Cancer: principles & practice of oncology. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2001. p. 983—1018.

[12] Fushimi H, Kikui M, Morino H, Yamamoto S, Tateishi R, Wada A, et al. Histologic changes in small cell lung carcinoma after treatment. Cancer 1996;77:278—83.

[13] Sobin LH, Wittenkind C. TNM classification of malignant tumors. New York: Wiley-Liss; 1997.

[14] Handbook for reporting results of cancer treatment. Geneva: World Health Organization (WHO); 1979.

[15] Kaplan ELMP. Non-parametric estimation from incomplete observations. J Am Stat Assoc 1958;53:457—81.

[16] Jiang SX, Kameya T, Shoji M, Dobashi Y, Shinada J, Yoshimura H. Large cell neuroendocrine carcinoma of the lung: a histologic and immunohistochemical study of 22 cases. Am J Surg Pathol 1998;22:526—37.

[17] Iyoda A, Hiroshima K, Toyozaki T, Haga Y, Fujisawa T, Ohwada H. Clinical characterization of pulmonary large cell neuroendocrine carcinoma and large cell carcinoma with neuroendocrine morphology. Cancer 2001;91:1992—2000.

[18] Paci M, Cavazza A, Annessi V, Putrino I, Ferrari G, De Franco S, et al. Large cell neuroendocrine carcinoma of the lung: a 10-year clinicopathologic retrospective study. Ann Thorac Surg 2004;77:1163—7.

[19] Yang YJ, Steele CT, Ou XL, Snyder KP, Kohman LJ. Diagnosis of high-grade pulmonary neuroendocrine carcinoma by fine-needle aspiration biopsy: nonsmall-cell or small-cell type? Diagn Cytopathol 2001;25:292—300.

[20] Wiatrowska BA, Krol J, Zakowski MF. Large-cell neuroendocrine carcinoma of the lung: proposed criteria for cytologic diagnosis. Diagn Cytopathol 2001;24:58—64.

[21] Mazieres J, Daste G, Molinier L, Berjaud J, Dahan M, Delsol M, et al. Large cell neuroendocrine carcinoma of the lung: pathological study and clinical outcome of 18 resected cases. Lung Cancer 2002;37:287—92.

[22] Travis WD, Gal AA, Colby TV, Klimstra DS, Falk R, Koss MN. Reproducibility of neuroendocrine lung tumor classification. Hum Pathol 1998;29:272—9.

SNP Communication

Genetic Variations and Haplotypes of UGT1A4 in a Japanese Population

Mayumi SAEKI[1], Yoshiro SAITO[1,2,*], Hideto JINNO[1,3], Kimie SAI[1,4], Akiko HACHISUKA[2],
Nahoko KANIWA[1,5], Shogo OZAWA[1,6], Manabu KAWAMOTO[7], Naoyuki KAMATANI[7],
Kuniaki SHIRAO[8], Hironobu MINAMI[9], Atsushi OHTSU[10], Teruhiko YOSHIDA[11],
Nagahiro SAIJO[12], Kazuo KOMAMURA[13,14], Takeshi KOTAKE[15], Hideki MORISHITA[15],
Shiro KAMAKURA[13], Masafumi KITAKAZE[13], Hitonobu TOMOIKE[13] and Jun-ichi SAWADA[1,2]

[1]*Project Team for Pharmacogenetics,* [2]*Division of Biochemistry and Immunochemistry,*
[3]*Division of Environmental Chemistry,* [4]*Division of Xenobiotic Metabolism and Disposition,*
[5]*Division of Medicinal Safety Science,* [6]*Division of Pharmacology, National Institute of Health Sciences,*
Tokyo, Japan [7]*Division of Genomic Medicine, Department of Advanced Biomedical Engineering and*
Science, Tokyo Women's Medical University, Tokyo, Japan [8]*Division of Internal Medicine,*
National Cancer Center Hospital, [11]*Genetics Division, National Cancer Center Research Institute,*
National Cancer Center, Tokyo, Japan [9]*Division of Oncology/Hematology,*
[10]*Division of GI Oncology/Digestive Endoscopy,* [12]*Deputy Director, National Cancer Center Hospital East,*
Chiba, Japan [13]*Division of Cardiology,* [14]*Department of Cardiovascular Dynamics Research Institute,*
[15]*Department of Pharmacy, National Cardiovascular Center, Osaka, Japan*

Full text of this paper is available at http://www.jssx.org

Summary: Nineteen genetic variations, including 11 novel ones, were found in exon 1 and its flanking region of the UDP-glucuronosyltransferase (UGT) 1A4 gene from 256 Japanese subjects, consisting of 60 healthy volunteers, 88 cancer patients and 108 arrhythmic patients. These variations include $-217T>G$ and $-36G>A$ in the 5'-flanking region, $30G>A$ (P10P), 127delA (43fsX22; frame-shift from codon 43 resulting in the termination at the 22nd codon, codon 65), 175delG (59fsX6), $271C>T$ (R91C), $325A>G$ (R109G), and $357T>C$ (N119N) in exon 1, and IVS1+1G>T, IVS1+98A>G and IVS1+101G>T in the following intron. Among them, 127delA and 175delG can confer early termination of translation, resulting in an immature protein that probably lacks enzymatic activity. Variation IVS1+1G>T is located at a splice donor site and thus may lead to aberrant splicing. Since we did not find any significant differences in the frequencies of all the variations among the three subject groups, the data were analyzed as one group. The allele frequencies of the novel variations were 0.006 for IVS1+101G>T, 0.004 for $30G>A$ (P10P) and $357T>C$ (N119N), and 0.002 for the 8 other variations. In addition, the two known nonsynonymous single nucleotide polymorphisms (SNPs), $31C>T$ (R11W) and $142T>G$ (L48V), were found at 0.012 and 0.129 frequencies, respectively. The SNP $70C>A$ (P24T), mostly linked with $142T>G$ (L48V) in German Caucasians, was not detected in this study. Sixteen haplotypes were identified or inferred, and some haplotypes were confirmed by cloning and sequencing. It was shown that most of $142T>G$ (L48V) was linked with $-219C>T$, $-163G>A$, $448T>C$ (L150L), $804G>A$ (P268P), and IVS1+43C>T, comprising haplotype *3a*; haplotype *4a* harbors $31C>T$ (R11W); 127delA (43fsX22) and $142T>G$ (L48V) were linked (haplotype *5a*); 175delG (59fsX6) was linked with $325A>G$ (R109G) (*6a* haplotype); and $-219C>T$, $-163G>A$, $142T>G$ (L48V), $271C>T$ (R91C), $448T>C$ (L150L), $804G>A$ (P268P), and IVS1+43C>T comprised haplotype *7a*. Our results provide fundamental and useful information for genotyping *UGT1A4* in the Japanese and probably Asian populations.

Key words: UGT1A4; amino acid alteration; frameshift; splice donor site; drug metabolism

Table 1. Primers utilized for *UGT1A4* amplification and sequencing

	Direction	Primer Name	Sequences	Location[a]
Amplification	forward	UGT1A4-1stF	TTAACAAAGTAGAAGGCAGTG	135092
	reverse	UGT1A4-1stR	TGAAAACTTGAAATACACTAGGC	136460
Sequencing	forward	UGT1A4-1stF	TTAACAAAGTAGAAGGCAGTG	135092
	forward	UGT1A4seqF2	GGGCTGAGAGTGGAAAGGT	135502
	forward	UGT1A4seqF3	TCCTTCCTCCTATATTCCTAAGTT	135995
	reverse	UGT1A4seqR1-2	ATCAAATTCCTTCTGGGTCC	135698
	reverse	UGT1A4seqR2	AAGGGGCAGAAAAAGTATGG	136119
	reverse	UGT1A4-1stR	TGAAAACTTGAAATACACTAGGC	136460

[a]The 5'-end of each primer on AF297093.1.

Introduction

As phase II enzymes, the UDP-glucuronosyltransferase enzymes (UGTs) play crucial roles in the detoxification and elimination of a large number of endogenous and exogenous compounds.[1] Of the UGT1 and UGT2 subfamilies expressed in humans, the genes encoding UGT1As have a unique genetic structure consisting of at least 13 different exon 1's, including four inactive ones, and the common exons 2 to 5 clustered on chromosome 2q37.[2] One of the exon 1's can be spliced on to the common exons. The N-terminal domains (encoded by the exon 1's) of the UGT1A proteins determine their substrate-binding specificity, and the common C-terminal domain (encoded by exons 2 to 5) is important for UDP-glucuronic acid binding.[3]

UGT1A4 is expressed in the liver, bile ducts, colon, small intestine, and pancreas.[1,4,5] UGT1A4 catalyzes the conjugation of exogenous amines and alcohols, including nicotine, sapogenins, imipramine, trifluoperazine, and tamoxifen.[1,6-9] In addition, many androgens and progestins are reported as endogenous substrates of UGT1A4.[6] Several genetic polymorphisms of *UGT1A4* were reported in the public databases. Among them, two nonsynonymous single nucleotide polymorphisms (SNPs), 70C>A (P24T) and 142T>G (L48V), were found in German Caucasians, and they were shown to be closely associated.[10] The variant enzymes (24T and

48V) had reduced *in vitro* activities for β-naphtylamine, benzidine, *trans*-androsterone, and dihydrotestosterone in a substrate-specific manner.[10]

In spite of the clinical importance of UGT1A4, there is no report on the comprehensive sequencing analysis for the genetic polymorphisms of *UGT1A4* in Asian populations, including the Japanese. In the present study, *UGT1A4* exon 1 was sequenced from 256 Japanese subjects. Eleven novel genetic variations were identified, including 4 nonsynonymous ones.

Materials and Methods

Human genomic DNA samples: DNA was obtained from the blood leukocytes of 88 Japanese cancer patients and 108 Japanese arrhythmic patients. Written informed consent was obtained from all participating patients. DNA was also extracted from Epstein-Barr virus-transformed lymphoblastoid cells, for which blood samples were collected from 60 healthy Japanese volunteers at the Tokyo Women's Medical University under the auspices of the Pharma SNP Consortium (Tokyo, Japan). Informed consent was also obtained from all healthy subjects. The ethical review boards of all the participating organizations approved this study.

PCR conditions for DNA sequencing: First, exon 1 of *UGT1A4* was amplified from genomic DNA (100 ng) using 0.625 units of *Ex*-Taq (Takara Bio. Inc., Shiga, Japan) with 0.2 μM of amplification primers designed in the introns (**Table 1**). The PCR conditions were 94°C for 5 min, followed by 30 cycles of 94°C for 30 sec, 55°C for 1 min, and 72°C for 2 min, and then a final extension at 72°C for 7 min. These PCR products were then treated with a PCR Product Pre-Sequencing Kit (USB Co., Cleveland, OH, USA) and were directly sequenced on both strands using an ABI Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) (see **Table 1** for sequencing primers). The excess dye was removed by a DyeEx96 kit (Qiagen, Hilden, Germany). The eluates were analyzed on an ABI Prism 3700 DNA Analyzer (Applied Biosystems). All variations were confirmed by repeating

Mayumi Saeki, *et al.*

Table 2. Summary of *UGT1A4* polymorphisms detected in a Japanese population

| This Study	SNP ID dbSNP-NCBI database	SNP ID JSNP database	PharmGKB database[b]	Location	Position AF297093.1	Position From the translational initiation site or from the end of exon 1 (IVS1+)	Nucleotide change and flanking sequences (5′ to 3′)	Amino acid change	Wild-type	Hetero-zygote	Homo-zygote	Total (n=256)	Healthy volunteers (n=60)	Cancer patients (n=88)	Arrhythmic patients (n=108)
MP16_UIA081	rs3732219	IMS-JST085729	O	5′-flanking	135210	−219	GGGTCAGATGAGC/TTTTTCAAGATAG		195	54	7	0.133	0.133	0.142	0.125
MP16_UIA082[a]				5′-flanking	135212	−217	GTCAGATGAGCTT/GTTCAAGATAGGC		255	1	0	0.002	0.000	0.000	0.005
MP16_UIA083	rs3732218	IMS-JST085728	O	5′-flanking	135266	−163	TAACGAAAGGCAG/ATTATAGATTAAT		195	54	7	0.133	0.133	0.142	0.125
MP16_UIA084[a]				5′-flanking	135393	−36	CAGGCACAGCGTG/AGGGTGGACAGTC		255	1	0	0.002	0.000	0.006	0.000
MP16_UIA085[a]				Exon 1	135458	30	GGTTCCCCTGCCG/ACGGCTGGCCACA	P10P	254	2	0	0.004	0.000	0.000	0.009
MP16_UIA086	rs3892221			Exon 1	135459	31	GTTCCCCTGCCGC/TGGCTGGCCACAG	R11W	250	6	0	0.012	0.025	0.011	0.005
MP16_UIA087[a]				Exon 1	135555	127	AGCCCCTGGCTCA/-GCATGCGGGAGG	43fsX22	255	1	0	0.002	0.000	0.000	0.005
MP16_UIA088	rs2011425		O	Exon 1	135570	142	ATGCGGGAGGCCT/GTGCGGGAGCTCC	L48V	197	52	7	0.129	0.133	0.148	0.111
MP16_UIA089[a]				Exon 1	135603	175	GGCCACCAGGCGG/-TGGTCCTCACCC	59fsX6	255	1	0	0.002	0.000	0.000	0.005
MP16_UIA090[a]				Exon 1	135699	271	AAGGAATTTGATC/TGCGTTACGCTGG	R91C	255	1	0	0.002	0.000	0.000	0.005
MP16_UIA091[a]				Exon 1	135753	325	CATCTTCTGAAGA/GGATATTCTAGAA	R109G	255	1	0	0.002	0.000	0.000	0.005
MP16_UIA092[a]				Exon 1	135785	357	AATTATGAACAAT/CGTATCTTTGGCC	N119N	254	2	0	0.004	0.008	0.006	0.000
MP16_UIA093	rs12468274		O	Exon 1	135876	448	TTTGATGTGGTTT/CTAACAGACCCCG	L150L	195	54	7	0.133	0.133	0.142	0.125
MP16_UIA094	rs2011404		O	Exon 1	135899	471	CGTTAACCTCTGC/T[c]GGGGCCGGTGCTG	C157C	251	5	0	0.010	0.008	0.011	0.009
MP16_UIA095	rs3732217	IMS-JST085727	O	Exon 1	136232	804	CTACCCAGGCCG/AATCATGCCCAAC	P268P	195	54	7	0.133	0.133	0.142	0.125
MP16_UIA096[a]				Intron 1	136296	IVS1+1	CCACTATCTCAGG/TTCTGTATTGGTG		255	1	0	0.002	0.000	0.000	0.005
MP16_UIA097	rs2011219	IMS-JST085726	O	Intron 1	136338	IVS1+43	TTCCAGGCAAAAC/TACTTTTTAAAAA		195	54	7	0.133	0.133	0.142	0.125
MP16_UIA098[a]				Intron 1	136393	IVS1+98	ACTTATCTTTCCA/GAAGATTTTATTT		255	1	0	0.002	0.000	0.006	0.000
MP16_UIA099[a]				Intron 1	136396	IVS1+101	TATCTTTCCAAAG/TATTTTATTTTGG		253	3	0	0.006	0.008	0.006	0.005

[a] Novel variations detected in this study.
[b] The SNPs included in the PharmGKB database was shown as "O".
[c] T in the reference sequence.
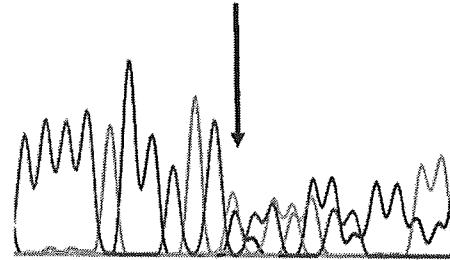
A   127delA (43 fsX 22) (sense)

Wild-type
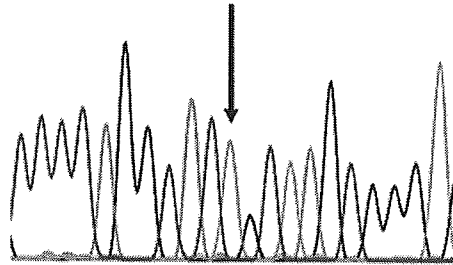
Variant

CCCCTGGCTCAGCATGCGGGA
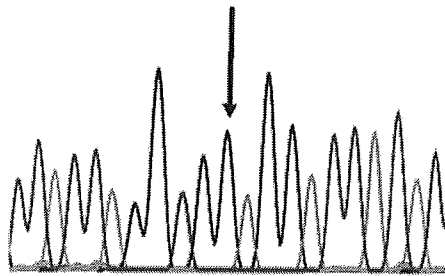
CCCCTGGCTCAGCATGCGGGA
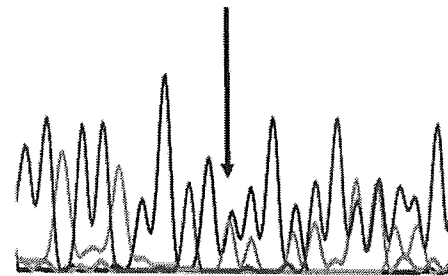   GCATGCGGGAG
(A deletion)

B   175delG (59 fsX 6) (sense)

Wild-type

Variant

CCACCAGGCGGTGGTCCTCAC
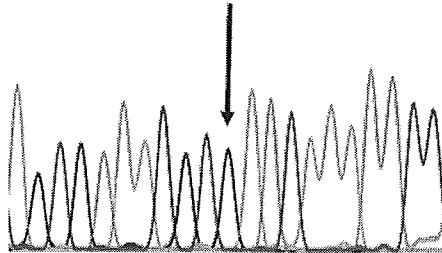
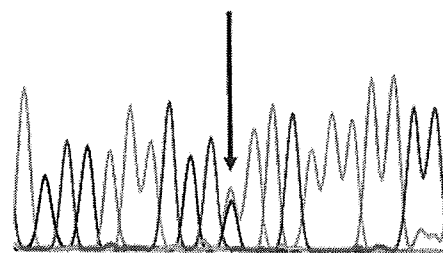CCACCAGGCGGTGGTCCTCAC
            TGGTCCTCACC
(G deletion)

C   271C>T (Arg 91 Cys) (antisense)

Wild-type

Variant

AGCGTAACGCGATCAAATTCC
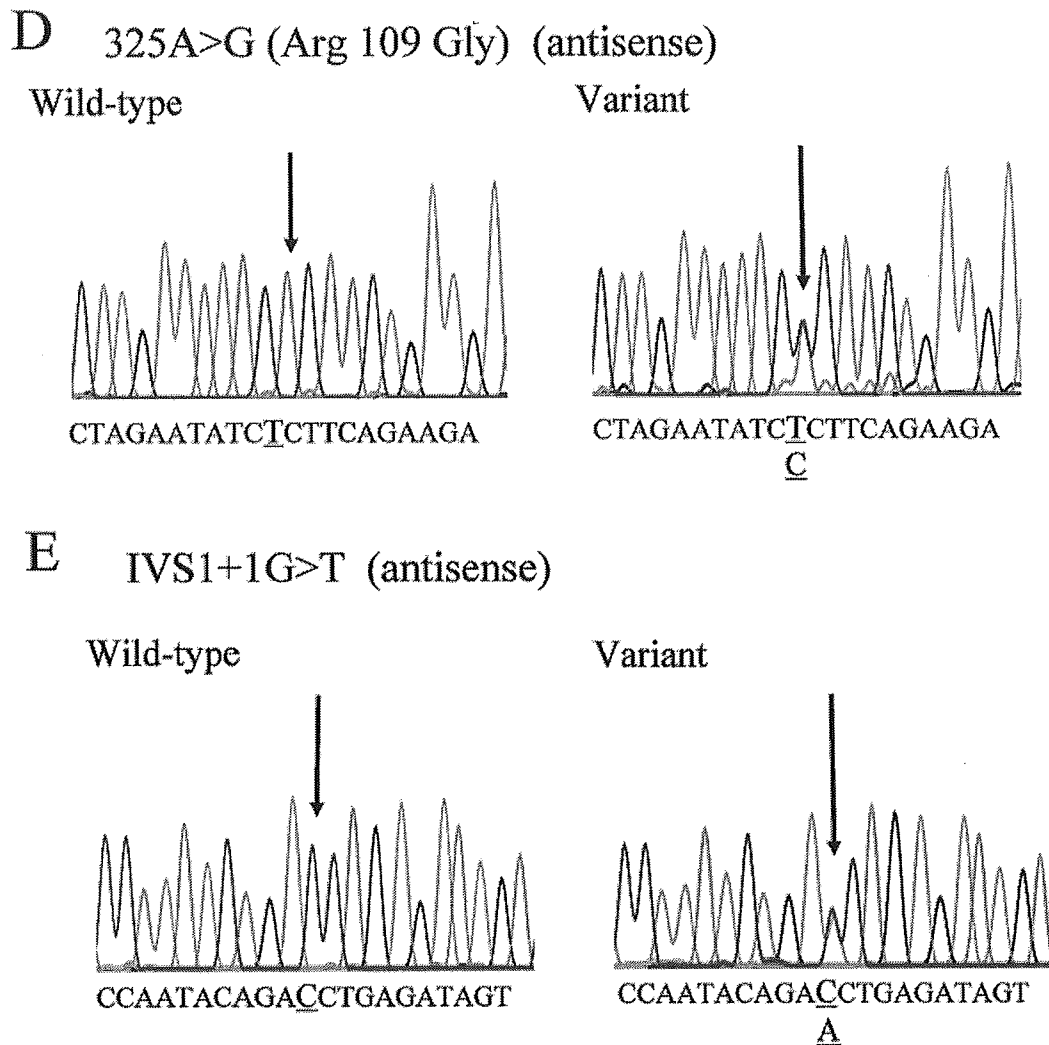
AGCGTAACGCGATCAAATTCC
          A

Fig. 1

D   325A>G (Arg 109 Gly) (antisense)

Wild-type   Variant

CTAGAATATCTCTTCAGAAGA   CTAGAATATCTCTTCAGAAGA
                                              C

E   IVS1+1G>T (antisense)

Wild-type   Variant

CCAATACAGACCTGAGATAGT   CCAATACAGACCTGAGATAGT
                                              A

Fig. 1. The 4 novel genetic variations with amino acid substitutions and 1 splice donor site variation of human *UGT1A4*. (A) MPJ6_U1A087 (wild-type, 127A/A; variant, 127A/-). (B) MPJ6_U1A089 (wild-type, 175G/G; variant, 175G/-). (C) MPJ6_U1A090 (wild-type, 271C/C; variant, 271C/T). (D) MPJ6_U1A091 (wild-type, 325A/A; variant, 325A/G). (E) MPJ6_U1A096 (wild-type, IVS1+1G/G; variant, IVS1+1G/T). Arrows indicate the positions of the nucleotide changes.

the PCR on genomic DNA and sequencing the newly generated PCR products. Furthermore, the rare variations found in only one subject were confirmed by sequencing the PCR fragments produced by amplification with a high fidelity DNA polymerase KOD-Plus-(TOYOBO, Tokyo, Japan).

Linkage disequilibrium (LD) and haplotype analysis: Hardy-Weinberg equilibrium analysis and LD analysis were performed by SNPAlyze software (Dynacom Co., Yokohama, Japan). Pairwise LDs were shown in rho square (r^2) values. Some of the haplotypes were unambiguous from the subjects with homozygous SNPs at all sites or a heterozygous SNP at only one site. Separately, the diplotype configurations (a combination of haplotypes) were inferred by LDSUPPORT software, which

determines the posterior probability distribution of the diplotype configuration for each subject based on the estimated haplotype frequencies.[11] The haplotypes were described as a number plus a small alphabetical letter.

Results and Discussion

UGT1A4 exon 1 and its flanking regions (from −286 bases upstream of the translational start site to 112 bases downstream of the end of exon 1) were sequenced from 256 Japanese subjects. Genbank accession number AF297093.1 was utilized for the reference sequence. Nineteen polymorphisms were detected, including 11 novel ones (2 were in the 5′-flanking region, 6 in exon 1, and 3 in the following intron) (**Table 2**). All of the allelic frequencies were in Hardy-Weinberg equilibrium (*p* =

0.13 or over). Since we did not find any significant differences in the frequencies of all the variations among three subject groups ($p > 0.25$ by χ^2 test) and between two of the three groups ($p > 0.13$ by χ^2 test or Fisher's exact test), the data for all subjects were analyzed as one group.

We found two novel nonsynonymous variations, 271C > T (R91C) and 325A > G (R109G), and two novel deletions, 127delA (43fsX22) and 175delG (59fsX6), as individual heterozygotes at a 0.002 frequency. Among them, 127delA (43fsX22) and 175delG (59fsX6) are the frameshift variations starting from codon 43 and 59, respectively, resulting in early stop codons at the 22nd (*i.e.* codon 65) and the 6th (*i.e.* codon 65) codons, respectively. It is most likely that these variations generate an immature protein that probably has null activity. The functional significance of 271C > T (R91C) and 325A > G (R109G) is currently unknown. Additionally, IVS1 + 1G > T, which was found at a frequency of 0.002, was located at a splice donor site and thus may lead to aberrant splicing (**Fig. 1**).

We also detected two known nonsynonymous SNPs, 31C > T (R11W) and 142T > G (L48V), at 0.012 and 0.129 frequencies, respectively. The frequency of 142T > G (L48V) was almost comparable to that of German Caucasians (0.09).[10] L48V was reported to lead to a partial decrease in glucuronidation of β-naphthylamine and benzidine, a marked decrease in the activity to *trans*-androsterone, and no activity toward dihydrotestosterone *in vitro*.[10] The functional significance of SNP 31C > T (R11W) has not been reported yet.

High linkage disequilibrium ($r^2 \geq 0.89$) was observed among −219C > T, −163G > A, 142T > G (L48V), 448T > C (L150L), 804G > A (P268P), and IVS1 + 43C > T. A perfect linkage ($r^2 = 1$) was found between 175delG and 325A > G (R109G), but found in only one subject. The r^2 values were below 0.014 between other pairs of polymorphisms. The SNP 70C > A (P24T), mostly linked with 142T > G (L48V) in German Caucasians,[10] was not detected in this study. Thus, it must be clarified whether the differences in the linkage of those SNPs may lead to the ethnic differences in the enzymatic activities of UGT1A4. A similar kind of ethnic difference has been found in the *1B haplotype, which harbors the three linked SNPs in the 3′-untranslated region of UGT1A common exon 5 found in a Japanese population.[12] In Caucasian and African-American populations, this linkage of the 3 SNPs was not complete, especially in African-Americans.[13]

Using the detected SNPs, haplotype analysis was then performed (**Table 3**). Since *UGT1A4*2* [70C > A (P24T)] and *3* [142T > G (L48V)] were defined in AF465196 and AF465197 (Genbank accession numbers), respectively, the novel haplotypes with amino acid changes, frameshift variations, or splice donor site

Table 3. *UGT1A4* haplotypes in a Japanese population

[a] A of the translational start codon of *UGT1A4* is numbered 1. AF297093.1 was used as the reference sequence.
[b] The haplotypes were described as a number plus a small alphabetical letter.

Mayumi SAEKI, et al.

variation, were assigned as haplotypes *4 to *8. Several haplotypes were first unambiguously assigned by homozygous SNPs at all sites (*1a and *3a) or a heterozygous SNP at only one site (*1b, *1d-*1i, *3b, *4a, and *8a). Separately, we estimated the diplotype configuration (a combination of haplotypes) for each subject by LDSUPPORT software. The diplotype configurations of 256 subjects were inferred with probabilities (certainty) of 0.9998 or over, except for one subject. The additionally inferred haplotypes were *1c, *5a, *6a, and *7a. As for one subject with a low probability (who had heterozygous SNPs of −219C>T, −163 G>A, 31C>T, 142T>G, 448T>C, 804G>A, and IVS1+43C>T), the diplotype was determined by the cloning and sequencing of DNA fragments. One chromosome had haplotype *3a (consisting of −219C>T, −163 G>A, 142T>G, 448T>C, 804G>A, and IVS1+43C>T) and the other had haplotype *4a (31C>T). Moreover, the data obtained by cloning and sequencing analysis confirmed the presence of haplotypes *5a [127delA (43fsX22) and 142T>G (L48V)], *6a [175delG (59fsX6) and 325A>G (R109G)], and *7a [−219C>T, −163G>A, 142T>G (L48V), 271C>T (R91C), 448T>C (L150L), 804G>A (P268P), and IVS1+43C>T] (Table 3). The most frequent haplotype was *1a (frequency: 0.818), followed by *3a (0.123), *4a (0.012) and *1b (0.010). The frequencies of the other haplotypes were less than 0.01. Since 325A>G (R109G) was linked with 175delG (59fsX6), the enzymatic activity of this haplotype (*6a) is probably null. The other SNP, 271C>T confers the R91C substitution. In human UGT1A4, eight cysteine residues were located in the lumenal domain.[3,14] Though the disulfide-bond formation and its significance are not clear in the UGT1A4, it has been reported that the reduction of disulfide-bonds of rat UGT1A6 with dithiothreitol increases its enzymatic activity in the liver microsomes.[15] On the other hand, the alterations of several lumenal cysteines into serine residues seem to reduce the UGT1A6 activity when the mutant enzymes were expressed in COS cells.[15] The effect of additional cysteine residue at codon 91 in the UGT1A4 should be determined in the future.

In conclusion, we detected 19 polymorphisms, including 11 novel ones, in UGT1A4 from a Japanese population. Using the detected polymorphisms, 16 haplotypes were identified. Our results provide fundamental and useful information for genotyping UGT1A4 in the Japanese, and probably Asian populations.

References

1) Tukey, R. H. and Strassburg, C. P.: Human UDP-glucuronosyltransferase: Metabolism, Expression, and Disease. *Annu. Rev. Pharmacol. Toxicol.*, **40**: 581–616 (2000).

2) Mackenzie, P. I., Owens, I. S., Burchell, B., Bock, K. W., Bairoch, A., Belanger, A., Fournel-Gigleux, S., Green, M., Hum, D. W., Iyanagi, T., Lancet, D., Louisot, P., Magdalou, J., Chowdhury, J. R., Ritter, J. K., Schachter, H., Tephly, T. R., Tipton, K. F. and Nebert, D. W.: The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics*, **7**: 255–269 (1997).

3) Radominska-Pandya, A., Czernik, P. J. and Little, J. M.: Structural and functional studies of UDP-glucuronosyltransferases. *Drug Metab. Rev.*, **31**: 817–899 (1999).

4) Strassburg, C. P., Kneip, S., Topp, J., Obermayer-Straub, P., Barut, A., Tukey, R. H. and Manns, M. P.: Polymorphic gene regulation and interindividual variation of UDP-glucuronosyltransferase activity in human small intestine. *J. Biol. Chem.*, **275**: 36164–36171 (2000).

5) Ockenga, J., Vogel, A., Teich, N., Keim, V., Manns, M. P. and Strassburg, C. P.: UDP glucuronosyltransferase (UGT1A7) gene polymorphisms increase the risk of chronic pancreatitis and pancreatic cancer. *Gastroenterology*, **124**: 1802–1808 (2003).

6) Green, M. D. and Tephly, T. R.: Glucuronidation of amines and hydroxylated xenobiotics and endobiotics catalyzed by expressed human UGT1.4 protein. *Drug Metab. Dispos.*, **24**: 356–363 (1996).

7) Nakajima, M., Tanaka, E., Kobayashi, T., Ohashi, N., Kume, T. and Yokoi, T.: Imipramine *N*-glucuronidation in human liver microsomes: biphasic kinetics and characterization of UDP-glucuronosyltransferase isoforms. *Drug Metab. Dispos.*, **30**: 636–642 (2002).

8) Nakajima, M., Tanaka, E., Kwon, J. T. and Yokoi, T.: Characterization of nicotine and cotinine *N*-glucuronidations in human liver microsomes. *Drug Metab. Dispos.*, **30**: 1484–1490 (2002).

9) Kaku, T., Ogura, K., Nishiyama, T., Ohnuma, T., Muro, K. and Hiratsuka, A.: Quaternary ammonium-linked glucuronidation of tamoxifen by human liver microsomes and UDP-glucuronosyltransferase 1A4. *Biochem. Pharmacol.*, **67**: 2093–2102 (2004).

10) Ehmer, U., Vogel, A., Schutte, J. K., Krone, B., Manns, M. P. and Strassburg, C. P.: Variation of hepatic glucuronidation: Novel functional polymorphisms of the UDP-glucuronosyltransferase UGT1A4. *Hepatology*, **39**: 970–977 (2004).

11) Kitamura, Y., Moriguchi, M., Kaneko, H., Morisaki, H., Morisaki, T., Toyama, K. and Kamatani, N.: Determination of probability distribution of diplotype configuration (diplotype distribution) for each subject from genotypic data using the EM algorithm. *Ann. Hum. Genet.*, **66**: 183–193 (2002).

12) Sai, K., Saeki, M., Saito, Y., Ozawa, S., Katori, N.,

Jinno, H., Hasegawa, R., Kaniwa, N., Sawada, J., Komamura, K., Ueno, K., Kamakura, S., Kitakaze, M., Kitamura, Y., Kamatani, N., Minami, H., Ohtsu, A., Shirao, K., Yoshida, T. and Saijo, N.: UGT1A1 haplotypes associated with reduced glucuronidation and increased serum bilirubin in irinotecan-administered Japanese patients with cancer. *Clin. Pharmacol. Ther.*, **75**: 501–515 (2004).

13) Kaniwa, N., Kurose, K., Jinno, H., Tanaka-Kagawa, T., Saito, Y., Saeki, M., Sawada, J., Tohkin, M. and Hasegawa, R.: Racial variability in haplotype frequen-cies of UGT1A1 and glucuronidation activity of a novel SNP 686C > T (P229L) found in an African-American. *Drug Metab. Dispos.*, in press.

14) Ritter, J. K., Crawford, J. M. and Owens, I. S.: Cloning of two human liver bilirubin UDP-glucuronosyltran-sferase cDNAs with expression in COS-1 cells. *J. Biol. Chem.*, **266**: 1043–1047 (1991).

15) Ikushiro, S., Emi, Y. and Iyanagi, T.: Activation of glucuronidation through reduction of a disulfide bond in rat UDP-glucuronosyltransferase 1A6. *Biochemistry*, **41**: 12813–12820 (2002).

SNP Communication

Genetic Polymorphisms of UGT1A6 in a Japanese Population

Mayumi Saeki[1], Yoshiro Saito[1,2,*], Hideto Jinno[1,3], Kimie Sai[1,4],

Nahoko Kaniwa[1,5], Shogo Ozawa[1,6], Kazuo Komamura[7,8], Takeshi Kotake[9],

Hideki Morishita[9], Shiro Kamakura[7], Masafumi Kitakaze[7], Hitonobu Tomoike[7],

Kuniaki Shirao[10], Hironobu Minami[11], Atsushi Ohtsu[12], Teruhiko Yoshida[13],

Nagahiro Saijo[14], Naoyuki Kamatani[15] and Jun-ichi Sawada[1,2]

[1]*Project Team for Pharmacogenetics,* [2]*Division of Biochemistry and Immunochemistry,*
[3]*Division of Environmental Chemistry,* [4]*Division of Xenobiotic Metabolism and Disposition,*
[5]*Division of Medicinal Safety Science,* [6]*Division of Pharmacology,*
National Institute of Health Sciences, Tokyo, Japan
[7]*Division of Cardiology,* [8]*Department of Cardiovascular Dynamics Research Institute,*
[9]*Department of Pharmacy, National Cardiovascular Center, Osaka, Japan*
[10]*Division of Internal Medicine, National Cancer Center Hospital,* [13]*Genetics Division,*
National Cancer Center Research Institute, National Cancer Center, Tokyo, Japan
[11]*Division of Oncology/Hematology,* [12]*Division of GI Oncology/Digestive Endoscopy,*
[14]*Deputy Director, National Cancer Center Hospital East, Chiba, Japan*
[15]*Division of Genomic Medicine, Department of Advanced Biomedical Engineering and Science,*
Tokyo Women's Medical University, Tokyo, Japan

Full text of this paper is available at http://www.jssx.org

Summary: Thirteen single nucleotide polymorphisms (SNPs), including 6 novel ones, were found in exon 1 and its flanking region of UDP-glucuronosyltransferase (UGT) *1A6* from 195 Japanese subjects. Several novel SNPs were identified, including 269G > A (R90H), 279A > G (S93S), and 308C > A (S103X) in exon 1, and IVS1 + 109C > T, IVS1 + 120A > G, and IVS1 + 142C > T in the intron downstream of exon 1. Among these SNPs, 308C > A confers termination of translation at codon 103, resulting in the production of an immature protein that probably lacks enzymatic activity. The allele frequencies were 0.003 for all the 6 SNPs. In addition, the 3 known nonsynonymous SNPs were detected: 19T > G (S7A), 541A > G (T181A), and 552A > C (R184S) with frequencies of 0.226, 0.218, and 0.226, respectively. High linkage disequilibrium was observed among 19T > G (S7A), 315A > G (L105L), 541A > G (T181A), 552A > C (R184S), and IVS1 + 130G > T, as reported in Caucasian and African-American populations. Then, 11 haplotypes in *UGT1A6* were estimated. The novel nonsynonymous variant, 269A or 308A, was shown to be located on the same DNA strand together with 19G, 315G, 541G, 552C, and IVS1 + 130T. Our results provide fundamental and useful information for genotyping *UGT1A6* in the Japanese, and probably Asian populations.

Key words: *UGT1A6*; single nucleotide polymorphisms; amino acid alteration; nonsense alteration

Table 1. Primers utilized for *UGT1A6* amplification and sequencing

	Direction	Primer Name	Sequences	Position[a]
1st Amplification	forward	UGT1A6ZF	TTCTGTAGGGACTTCTGGGACTA	107550
	reverse	UGT1A6ZR	TCAGGAGGGCATCTGTAAACACT	111117
2nd Amplification	forward	UGT1A6-1stF	TAACTTTTCAGAGAGGGAGAAGC	109328
	reverse	UGT1A6-1stR	ACTTCAGCCTCAGGTCTCCTAT	110642
Sequencing	forward	UGT1A6-1stF	TAACTTTTCAGAGAGGGAGAAGC	109328
	forward	UGT1A6seqF2	TGACAAGCTGCTGGTGGTC	109687
	forward	UGT1A6seqF3-2	CCCAGACCCTGTGTCCTACAT	110164
	reverse	UGT1A6seqR1	AAAGCAAATTAACTTCAGGCA	109808
	reverse	UGT1A6seqR2	ACAAATTAACAAGGAAGTTGGC	110258
	reverse	UGT1A6-1stR	ACTTCAGCCTCAGGTCTCCTAT	110642

[a] The position of the 5' end of each primer on AF297093.1

Introduction

The UDP-glucuronosyltransferase enzymes (UGTs) are members of a superfamily of membrane-anchored enzymes located in the endoplasmic reticulum that catalyze the conjugation of glucuronic acid to a nucleophilic substrate.[1] Glucuronidation is important for the detoxification and elimination of a large number of endogenous and exogenous compounds. So far, two *UGT* gene subfamilies have been identified in humans, *UGT1* and *UGT2*.[2] The human *UGT1A* gene cluster is located on chromosome 2q37 and consists of at least 13 different exon 1's, including four inactive exon 1's (1A2P, 1A11P, 1A12P, and 1A13P), and common exons 2 to 5. One of the exon 1's can be spliced onto the other common exons.[3] The N-terminal domains (encoded by the exon 1's) of the *UGT1A* proteins determine their substrate-binding specificity, and the C-terminal domain (encoded by the exons 2 to 5) is important for binding to UDP-glucuronic acid.[4]

UGT1A6 is expressed in liver, bile ducts, stomach, colon, kidney, and brain.[1,5,6] *UGT1A6* plays important roles in the elimination of phenols and amines, including 1-naphthol, irinotecan, acetaminophen, and β-blockers.[1,7-10] In addition, 5-hydroxytryptamine (serotonin) was reported as an endogenous substrate of *UGT1A6*.[6] Several genetic polymorphisms with functional changes were reported for *UGT1A6*. Firstly, Ciotti *et al.* reported two nonsynonymous single nucleotide polymorphisms (SNPs) 541A>G (T181A) and 552A>C (R184S), both of which are in high (but not complete) linkage disequilibrium with each other.[10] These variations are located in the putative endoplasmic reticulum-localization signal.[11] The variant enzyme with the two amino acid alterations had reduced activities *in vitro* for 4-nitrophenol, methylsalicylate, 3-*O*-methyl-dopa, and β-blockers, such as propranolol, whereas it had almost comparable activities for 3-iodophenol compared to the wild-type enzyme. On the other hand, a recent study showed that the other SNP, 19T>G (S7A), located in the N-terminal signal sequence, was also highly linked with 541 A>G (T181A) and 552A>C (R184S).[12] The homozygous variant enzyme with all 3 SNPs (corresponding to the homozygote) glucuronidated 4-nitrophenol approximately two-fold higher than the wild-type *in vitro*. Interestingly, concomitant expression of both the wild-type and the variant with these 3 SNPs (corresponding to the heterozygote) showed decreased activity *in vitro*.

Though UGT1A6 is important for detoxification of many compounds, comprehensive sequence analysis for the genetic polymorphisms of *UGT1A6* in Asian populations, which includes the Japanese, is currently lacking. In this study, exon 1 of *UGT1A6* was sequenced from 195 Japanese subjects. Six novel SNPs, including two nonsynonymous ones, were identified from this sequence analysis.

Materials and Methods

Human genomic DNA samples: All 195 subjects in this study were Japanese patients with various solid cancers (88 subjects), who were administered irinotecan, or patients with arrhythmia (107 subjects), who were administered anti-arrhythmic drugs and β-blockers. Genomic DNA was extracted from blood leukocytes and used as a template for the polymerase chain reaction (PCR). All of the ethics committees of the National Cancer Institute, the National Cardiovascular Center, and the National Institute of Health Sciences approved this study. Written informed consent was obtained from all participating subjects.

PCR conditions for DNA sequencing: First, exon 1 of *UGT1A6* was amplified from genomic DNA (150 ng) using 2.5 units of *Z*-Taq (Takara Bio. Inc., Shiga, Japan) with 0.2 μM of the 1st amplification primers (see **Table 1**). The first PCR conditions consisted of 30 cycles of 98°C for 5 sec, 55°C for 5 sec, and 72°C for

Table 2. Summary of *UGT1A6* SNPs detected in a Japanese population

This Study	dbSNP (NCBI)	Pharm GKB[c]	Refe-rence	Location	AF 297093.1	From the translational initiation site or from the end of exon 1 (IVS1+)	Nucleotide change and flanking sequences (5′ to 3′)	Amino acid change	Wild-type	Hetero-zygote	Homo-zygote	Frequency
MPJ6_U1A068	rs6759892	o	12)	Exon 1	109628	19	TGCCTCCTTCGCT/GCATTTCAGAGAA	S7A	117	68	10	0.226
MPJ6_U1A069		o		Exon 1	109714	105	GGTCCCTCAGGAC/TGGAAGCCACTGG	D35D	178	17	0	0.044
MPJ6_U1A070[a]				Exon 1	109878	269	AGCTGAAGAACCG/ATTACCAATCATT	R90H	194	1	0	0.003
MPJ6_U1A071[a]				Exon 1	109888	279	CCGTTACCAATCA/GTTTGGAAACAAT	S93S	194	1	0	0.003
MPJ6_U1A072[a]				Exon 1	109917	308	TTGCTGAGCGATC/AATTCCTAACTGC	S103X	194	1	0	0.003
MPJ6_U1A073	rs4365456	o	12)	Exon 1	109924	315	GCGATCATTCCTA/GACTGCTCCTCAG	L105L	117	68	10	0.226
MPJ6_U1A074	rs2070959[b]	o	10)	Exon 1	110150	541	TCCCTGGAGCATA/GCATTCAGCAGAA	T181A	120	65	10	0.218
MPJ6_U1A075	rs4365457	o	10)	Exon 1	110161	552	TACATTCAGCAGA/CAGCCCAGACCCT	R184S	117	68	10	0.226
MPJ6_U1A076				Exon 1	110236	627	TTCCAACGAGTG/TGCCAACTTCCTT	V209V	192	3	0	0.008
MPJ6_U1A077[a]				Intron 1	110579	IVS1+109	TTCCTGGAGAAAC/TGGTGGGGGAAG		194	1	0	0.003
MPJ6_U1A078[a]				Intron 1	110590	IVS1+120	ACGGTGGGGGGAA/GGTGATACCCGGC		194	1	0	0.003
MPJ6_U1A079	rs7592281	o		Intron 1	110600	IVS1+130	GAAGTGATACCCG/TGCTCGGAGCAGC		124	61	10	0.208
MPJ6_U1A080[a]				Intron 1	110612	IVS1+142	GGCTCGGAGCAGC/TGGGAACACATAG		194	1	0	0.003

[a] Novel variations detected in this study.
[b] Also included in the JSNP database as IMS-JST006083.
[c] SNPs publicized in the PharmGKB database were shown as "o".

190 sec. Then, the PCR products were amplified by *Ex*-Taq (0.625 units) (Takara Bio. Inc.) with the 2nd amplification primers (0.2 μM) designed in the introns (see **Table 1**, 2nd Amplification). The second round of PCR was 94°C for 5 min, followed by 30 cycles of 94°C for 30 sec, 55°C for 1 min, and 72°C for 2 min. These PCR products were then treated with a PCR Product Pre-Sequencing Kit (USB Co., Cleveland, OH, USA) and were directly sequenced on both strands using an ABI Big Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) (see **Table 1**, Sequencing). The excess dye was removed by a DyeEx96 kit (Qiagen, Hilden, Germany). The eluates were analyzed on an ABI Prism 3700 DNA Analyzer (Applied Biosystems). All the SNPs were confirmed by repeating the PCR on genomic DNA and sequencing the newly generated PCR products.

Linkage disequilibrium (LD) and haplotype analysis: LD analysis was performed by SNPAlyze software (Dynacom Co., Yokohama, Japan), and a pairwise LD between SNPs was obtained for the rho square (r^2) values. Some of the haplotypes were unambiguous from subjects with homozygous SNPs at all sites or a heterozygous SNP at only one site. Separately, the diplotype configurations (a combination of haplotypes) were inferred by LDSUPPORT software, which determines the posterior probability distribution of the diplotype configuration for each subject based on the estimated haplotype frequencies.[13] The diplotype configurations of all subjects had a probability (certainty) of 1. The haplotypes inferred in only one chromosome are described with the haplotype name and a question mark in **Table 3**, since the predictability for these very rare haplotypes is known to be low in some cases.

Results and Discussion

UGT1A6 exon 1 and its flanking region were sequenced from 195 Japanese subjects. Genbank accession number AF297093.1 was utilized for the reference sequence. Thirteen SNPs, including 6 novel ones [3 were in the exon 1 and 3 in the following intron], were detected (see **Table 2**). All of the detected SNPs were found in Hardy-Weinberg equilibrium. Since we did not find any significant differences in the SNP frequencies between the two disease types (by χ^2 test or Fisher's exact test, $p > 0.31$), the data for all subjects were analyzed as one group.

Two novel nonsynonymous SNPs, 269G>A (R90H) and 308C>A (S103X), were found as individual heterozygotes at a 0.003 frequency (**Fig. 1**). Among them, 308C>A confers the termination of translation at codon 103, resulting in the production of an immature protein most probably with null activity, since it lacks 81% of the structure, including the C-terminal domain important for UDP-glucuronic acid binding.

Table 3. UGT1A6 haplotypes in a Japanese population

Haplotypes[b]	19 T>G (S7A)	105 C>T (D35D)	269 G>A (R90H)	279 A>G (S93S)	308 C>A (S103X)	315 A>G (L105L)	541 A>G (T181A)	552 A>C (R184S)	627 G>T (V209V)	IVS1 +109 C>T	IVS1 +120 A>G	IVS1 +130 G>T	IVS1 +142 C>T	Frequency	Frequency (group)	Haplotypes in ref 12)
*1a														0.726	0.774	*1
*1b		■												0.044		
*1c											■			0.003		
*1d?[c]				■										0.003		
*2a	■					■	■	■				■		0.197	0.213	*2
*2b	■					■	■	■				■		0.010		
*2c	■					■	■	■				■		0.003		
*2d?[c]	■					■	■	■				■		0.003		
*4a	■					■	■	■	■			■		0.008	0.008	*4
*5a	■				■	■	■	■				■		0.003		
*6a	■		■			■	■	■				■		0.003		

[a] A of the translational start codon of UGT1A6 is numbered 1. AF297093.1 was used as the reference sequence.
[b] The haplotypes were described as a number plus a small alphabetical letter.
[c] The haplotypes inferred in only one chromosome are described with the haplotype name and a question mark.

The functional significance of the other SNP, 269G>A (R90H), is currently unknown. Further functional analysis using a heterologous expression system should be pursued. Moreover, further study is necessary to evaluate the real frequencies of the very rare SNPs found in only one chromosome.

Also, the 3 known nonsynonymous SNPs, 19T>G (S7A), 541A>G (T181A), and 552A>C (R184S), were detected at 0.226, 0.218, and 0.226 frequencies, respectively. The frequencies for 541A>G (T181A) and 552A>C (R184S) were comparable to those of Asians in the previous report.[14] High linkage disequilibrium ($r^2 \geq 0.90$) was observed among 19T>G (S7A), 315A>G (L105L), 541A>G (T181A), 552A>C (R184S), and IVS1+130G>T, as reported in Caucasian and African-American populations.[12] The r^2 values were below 0.035 for the other combinations of SNPs.

Using the detected SNPs, haplotype analysis was then performed (Table 3). We basically followed Nagar et al.[12] for the haplotype numbering of UGT1A6, since they included the highly linked four SNPs, 19T>G (S7A), 315A>G (L105L), 541A>G (T181A), and 552A>C (R184S). The subtype numbering followed their frequencies (from high to low frequencies). Several haplotypes were first unambiguously assigned by homozygous SNPs at all sites (*1a and *2a) or a heterozygous SNP at only one site (*1b, *1c, *2c, and *5a). Note that the subject with novel heterozygous 308C>A (S103X) had homozygous 19T>G (S7A), 315A>G (L105L), 541A>G (T181A), 552A>C (R184S), and IVS1+130G>T (*2a haplotype), indicating that 308C>A (S103X) is linked with these SNPs (forming the *5a haplotype). Separately, the diplotype configuration (a combination of haplotypes) for each subject was estimated by LDSUPPORT software. The additionally inferred haplotypes were one *1 subtype (*1d), two *2 subtypes (*2b and *2d), *4a, and *6a. As for *6a, cloning and sequencing of DNA fragments obtained from the subject with heterozygous 269G>A (R90H) revealed that 269A is located on the same DNA strand together with 19G, 315G, 541G, 552C, and IVS1 +130T. The determined/inferred haplotypes were summarized in Table 3. We did not detect *3 haplotype, which consists of only SNP 19T>G (S7A). The most frequent haplotype was *1a (frequency: 0.726), followed by *2a (0.197), *1b (0.044), and *2b (0.010). The frequencies of the other haplotypes were less than 0.01. The total frequency of the *2 haplotypes (0.213) was almost comparable to those in Caucasians (0.274) and African-Americans (0.243).[12] On the other hand, the *4 frequency (0.008) was significantly lower ($p<0.01$ by Fisher's exact test) than those in Caucasians (0.052) and African-Americans (0.047).[12]
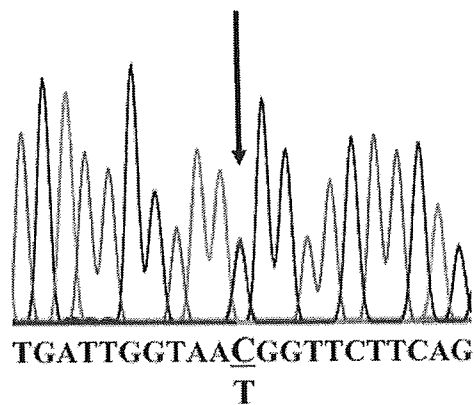
In conclusion, 13 SNPs were detected, including 6 novel ones, in UGT1A6 from the Japanese population.
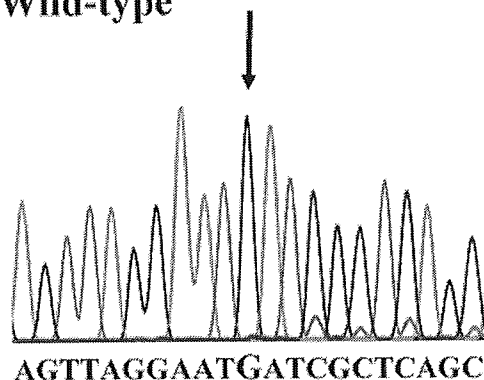
A 269G>A (Arg 90 His) (antisense)

Wild-type

Variant

TGATTGGTAACGGTTCTTCAG
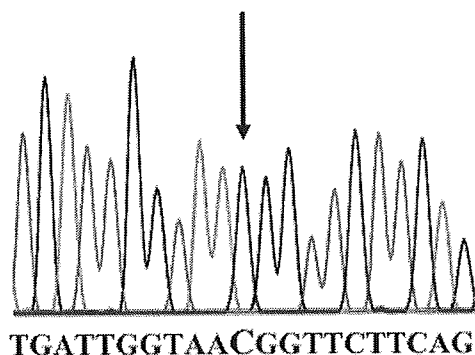
TGATTGGTAACGGTTCTTCAG
T

B 308C>A (Ser 103 X) (antisense)

Wild-type

Variant
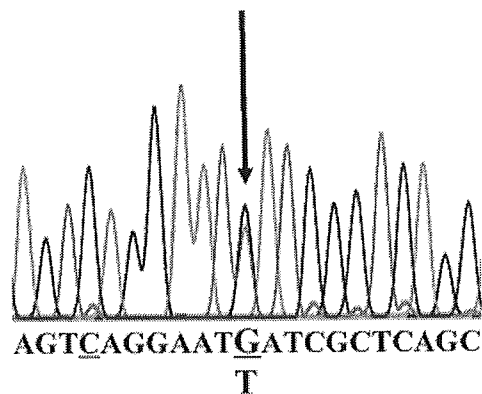
AGTTAGGAATGATCGCTCAGC

AGTCAGGAATGATCGCTCAGC
T

Fig. 1. The novel nonsynonymous SNPs of human *UGT1A6*. (A) MPJ6_U1A070 (wild-type 269G/G; variant 269G/A). (B) MPJ6_U1A072 (wild-type 308C/C; variant 308C/A). Arrows indicate the positions of the nucleotide changes. Note that the patient with heterozygous 308C>A also had homozygous 315A>G (L105L) alterations (shown in blue).

Using the detected SNPs, 11 haplotypes were determined and/or inferred. Our results provide fundamental and useful information for genotyping *UGT1A6* in the Japanese, and probably Asian populations.

References

1) Tukey, R. H. and Strassburg, C. P.: Human UDP-glucuronosyltransferase: Metabolism, Expression, and Disease. *Annu. Rev. Pharmacol. Toxicol.*, **40**: 581–616 (2000).

2) Mackenzie, P. I., Owens, I. S., Burchell, B., Bock, K. W., Bairoch, A., Belanger, A., Fournel-Gigleux, S., Green, M., Hum, D. W., Iyanagi, T., Lancet, D., Louisot, P., Magdalou, J., Chowdhury, J. R., Ritter, J. K., Schachter, H., Tephly, T. R., Tipton, K. F. and Nebert, D. W.: The UDP glycosyltransferase gene superfamily: recommended nomenclature update based on evolutionary divergence. *Pharmacogenetics*, **7**: 255–269 (1997).

3) Gong, Q.-H., Cho, J. W., Huang, T., Potter, C., Gholami, N., Basu, N. K., Kubota, S., Carvalho, S.,