20 psi (MEHP, PA, and their internal standards for negative ion mode); nebulizer gas (N_2) pressure, 20 psi for positive ion mode and 30 psi for negative ion mode; and turbo ion spray gas (N_2) pressure, 10 psi for positive ion mode and 80 psi for negative ion mode. The ion source temperature was maintained at 650°C and the turbo ion spray voltages for positive ion mode (DEHP, DEHP-d_4) and negative ion mode (MEHP, PA, and their internal standards) were 3500 and −4500 V, respectively. DEHP and DEHP-d_4 were detected in the positive ion mode, whereas MEHP, PA, and their internal standards were detected in the negative ion mode. The product ion mass spectra of DEHP, MEHP, and PA obtained by the LC-MS/MS system are shown in Figure 1. The combinations of precursor ion and product ions were as follows: DEHP (precursor ion → product ion, m/z 391 → 149), DEHP-d_4 (m/z 395 → 153), MEHP (m/z 277 → 134), MEHP-d_4 (m/z 281 → 138), PA (m/z 165 → 121), and PA-d_4 (m/z 169 → 125). The collision gas (N_2) pressures were set at 5 units (positive ion mode) and 4 units (negative ion mode).

Method Validation

After selection of the optimum conditions for sample preparation and LC-MS/MS, the method was thoroughly evaluated using DEHP, MEHP, and PA standard solutions. The linearity of the response of this system was examined with a calibration curve obtained at six different concentrations of the standard solution containing the certain amount of internal standard. Linear regression was performed using the ratio of DEHP peak area/DEHP-d_4 (internal standard) peak area plotted against the concentration. The calibration curves for MEHP and PA were also obtained in the same way. To assess the accuracy and precision of this method, low and high quality control samples were determined by replicate analysis. Intraday precision and accuracy were determined by replicate analysis of standard solutions in one day (n = 3), and interday precision and accuracy were determined over a span of three days.

The method was applied to 5% glucose solution and HCO-60 (0.02 mg/mL) samples that were spiked with 100 ng/mL DEHP, MEHP, and PA standards and certain amounts of internal standards. Each recovery was obtained from three replicates.

Migration Test

The migration of DEHP, MEHP, and PA from PVC sheets (1 × 3 cm) into 5 mL of each extraction solvent was examined. Five percent glucose solution, HCO-60, and purified water were used as extraction solvent, and served as simulated pharmaceuticals. HCO-60 is a surfactant that is

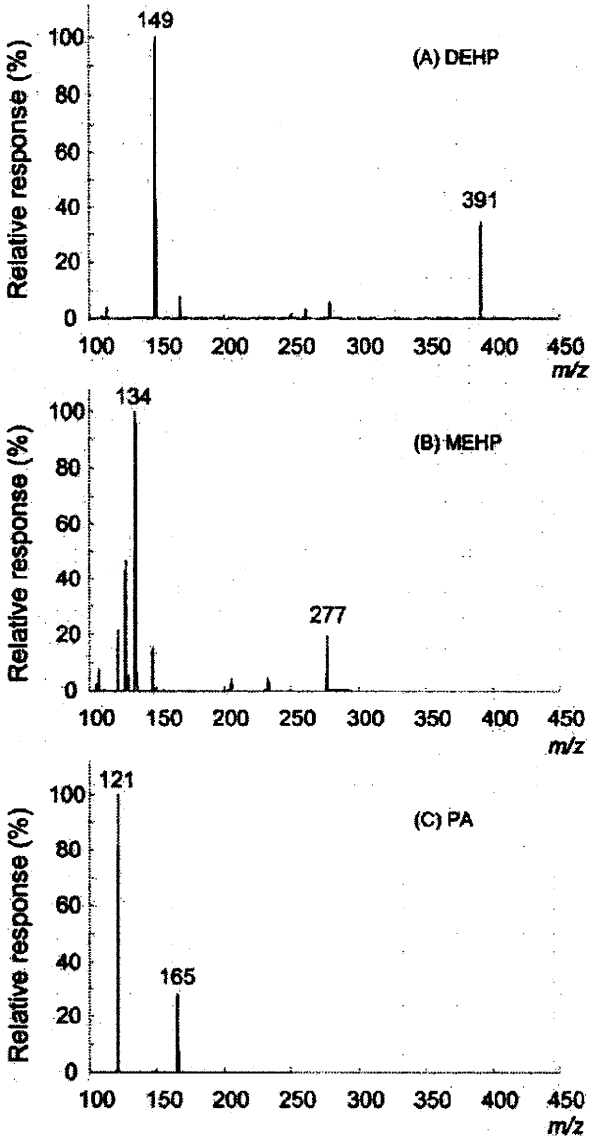


Figure 1. Product ion mass spectra of DEHP, MEHP and PA standard solutions. (A) 1 μg/mL DEHP standard solution, (B) 1 μg/mL MEHP standard solution, (C) 1 μg/mL PA standard solution. Each standard solution was infused directly into the MS system.

involved in the migration of DEHP into drugs such as Prograf®. The extent of DEHP migration was dependent on the concentration of HCO-60;[23] however, the injection of DEHP at high concentrations contaminated the MS system. Therefore, in this study, 0.02 mg/mL HCO-60 was prepared for the migration test. The samples were kept in test tubes and extraction was carried out by shaking at room temperature for 1 hr. A 1 mL aliquot of the extract was pipetted into another test tube, and DEHP-d_4, MEHP-d_4, and PA-d_4 were added. Then, the sample solution was appropriately diluted prior to LC-MS/MS analysis.

RESULTS AND DISCUSSION

Optimizing the LC-MS/MS Method

In the scan mode, DEHP, MEHP, and PA were monitored at m/z 391, 277, and 165 which were assigned to $[M + H]^+$, $[M\text{-}H]^-$, and $[M\text{-}H]^-$, respectively. Moreover, in the product ion MS/MS measurement, the selective reaction monitoring ions (SRM) of DEHP, DEHP-d_4, MEHP, MEHP-d_4, PA, and PA-d_4 were set depending upon their precursor ions. For the separation and the MS ionization, formic acid was added to purified water as the mobile phase. The optimum concentration of formic acid in purified water was 0.05% (Figure 2). In addition, the sample solution was acidified (1%) to improve separation. No interference from peaks of other compounds present in the extraction solvents was noted. The SRM chromatograms of DEHP, MEHP, and PA spiked into HCO-60 were shown in Figure 3.

Validation of the Method

In the proposed method, the limits of detection (LODs; signal-to-noise ratio = 3) of DEHP, MEHP, and PA were 5, 0.5, and 1 ng/mL, respectively. The limits of quantification (LOQs) (signal-to-noise ratio > 10) of DEHP, MEHP, and PA were 20, 2, and 5 ng/mL, respectively. For DEHP measurement, a calibration curve was obtained by plotting the peak area ratio (DEHP/DEHP-d_4) versus DEHP concentration, and was linear over the range of 20 to 1000 ng/mL ($r = 0.999$). For MEHP measurement, a calibration curve was obtained by plotting the peak area ratio (MEHP/MEHP-d_4) versus MEHP concentration, and was linear over the range of 2 to 1000 ng/mL ($r = 0.999$). For PA measurement, a calibration curve was obtained by plotting the peak area ratio (PA/PA-d_4) versus PA concentration, and was linear over the range of 2 to 1000 ng/mL ($r = 0.999$).

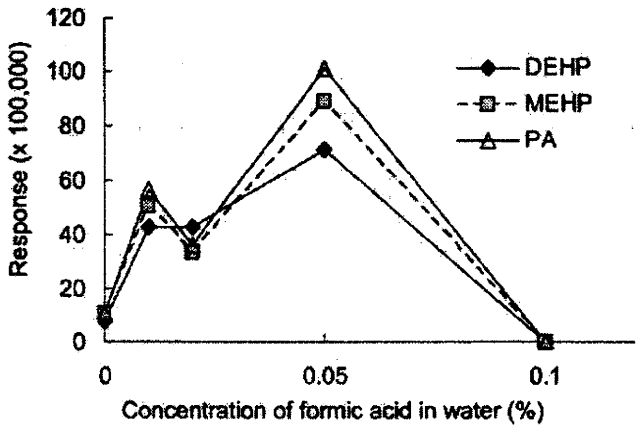


Figure 2. Effect of concentration of formic acid in purified water on response.

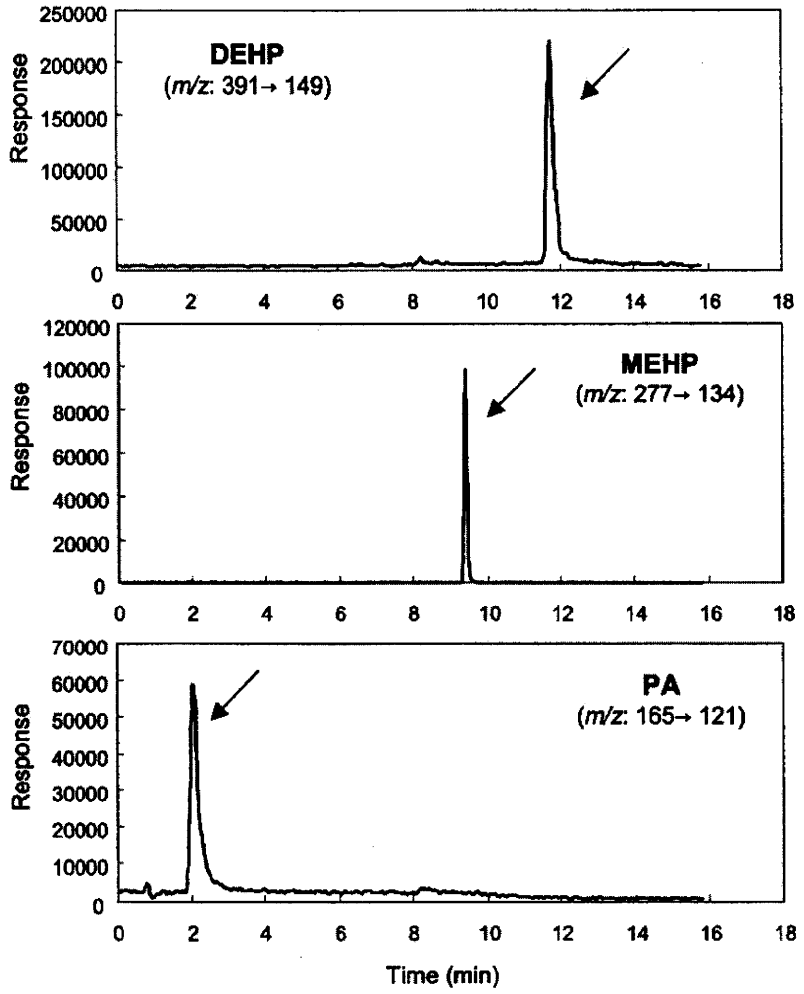


Figure 3. Chromatograms of DEHP, MEHP, PA spiked into HCO-60. 100 ng/mL of DEHP, MEHP, and PA were spiked into HCO-60 solution. SRM chromatograms were monitored as follows; DEHP (m/z 391 → 149), MEHP (m/z 277 → 134), and PA (m/z 165 → 121).

Intraday precision was expressed as relative standard deviation (RSD), which was calculated by measuring low (50 ng/mL) and high (500 ng/mL) concentrations of the standard solution three times (n = 3) in one day. Interday precision and accuracy were calculated using values measured at two concentrations (50 and 500 ng/mL) of the standard solutions over a span of three days. As Table 1 shows, all values of intra- and interday precision were less than 10%.

We also examined recovery using 5% glucose solution and HCO-60 as extraction solvents. For 5% glucose solution that was spiked with 100 ng/mL DEHP, MEHP, and PA, the average recoveries ranged from 100.4 to 104.2% (RSD <9.5%; Table 2). For HCO-60 that was spiked with 100 ng/mL DEHP, MEHP, and PA, the average recoveries ranged from 98.9 to 102.9% (RSD <10.3%; Table 2).

Table 1. Results of intra- and interday assays to validate proposed LC-MS/MS method

Analyte	Concentration (ng/mL)	Intraday			Interday		
		Detected average (ng/mL)	RSD (%)	Accuracy (%)	Detected average (ng/mL)	RSD (%)	Accuracy (%)
DEHP	50	54.4	4.0	106.8	50.3	8.9	100.6
	500	508.6	0.8	101.7	503.6	1.4	100.7
MEHP	50	48.9	3.5	97.8	48.6	1.1	97.1
	500	500.9	1.6	100.2	503.0	1.3	100.6
PA	50	49.9	4.8	99.9	49.4	1.6	98.9
	500	501.5	1.7	100.1	503.0	0.3	100.6

(n = 3).

Table 2. Recoveries of DEHP, MEHP and PA

Compound	Spiked amount (ng/mL)	Average recovery (%)		
		Water	5% Glucose solution	0.02 mg/mL HCO-60
DEHP	100	106.9 ± 6.9	104.2 ± 5.9	100.6 ± 10.3
MEHP	100	104.2 ± 2.4	103.4 ± 9.5	102.9 ± 5.8
PA	100	105.6 ± 7.6	100.4 ± 1.8	98.9 ± 8.2

(Mean ± SD, n = 3).

DEHP, MEHP, and PA Migration from Gamma-Ray Irradiated PVC Sheet

The proposed method was applied to the determination of DEHP, MEHP, and PA migration from the gamma-ray irradiated PVC sheet. DEHP migrated from both irradiated and unirradiated PVC sheets. The concentrations of DEHP that migrated from gamma-ray irradiated and unirradiated PVC sheets into purified water, or 5% glucose solution, were almost the same level (53.0–69.1 ng/mL). In contrast, the concentrations of DEHP that migrated from irradiated and unirradiated PVC sheets into HCO-60 were both high level (average concentration: 88.9 ng/mL) compared with the other solution. These concentrations were similar to those reported previously.[22] In our previous study, we noted that temperature and optical irradiation had an influence on DEHP release from the PVC sheet.[24] Therefore, DEHP release from the examined PVC sheet might have been influenced by temperature and/or optical irradiation, although the PVC sheet was stored in the dark at room temperature. The concentrations of MEHP that migrated from gamma-ray irradiated and unirradiated PVC sheets were also similar to those reported previously.[22] Gamma-ray irradiated PVC sheets released a high concentration of MEHP (Figure 4). In contrast, PA migration from unirradiated PVC sheets into any of the extraction solvents was not detected, whereas the gamma-ray irradiated PVC sheets released detectable levels of PA (Figure 4).

In our previous study, not only DEHP but also MEHP migrated from PVC medical devices into simulated pharmaceuticals. MEHP migration from PVC sheets was detected even though MEHP was not used as a plasticizer. In addition, MEHP was detected in gamma-ray irradiated PVC sheets but was not detected in PVC sheets sterilized by autoclaving or exposure to ethylene oxide gas.[22] The concentration of MEHP migrating from gamma-ray irradiated PVC sheets was significantly high compared with that from the unirradiated ones. Moreover, the concentration of PA migrating from gamma-ray irradiated PVC sheets was increased compared with unirradiated ones. MEHP and PA migrated from gamma-ray irradiated PVC sheets into

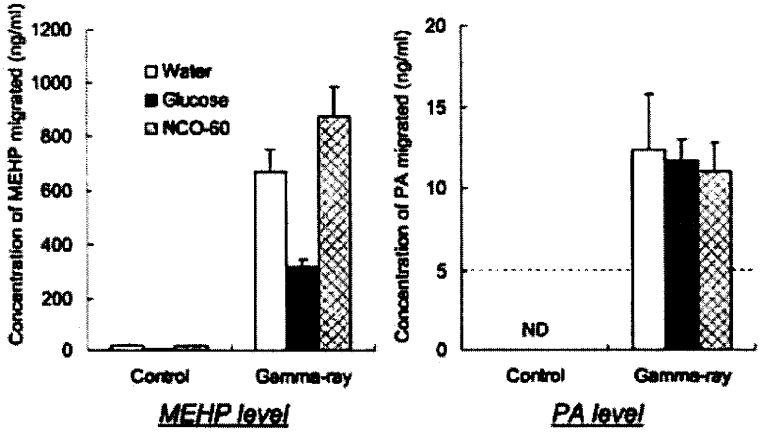


Figure 4. Concentrations of MEHP and PA migrating into various solutions from gamma-ray irradiated PVC sheets. Control is a PVC sample without gamma-ray irradiation. Each column is the mean of triplicate analysis (n = 3). Error bar represents standard deviation (S.D.). ND means "not detected." The dotted line at 5 ng/mL PA represents the limit of quantification.

both HCO-60 and purified water. We have already shown that MEHP was produced from DEHP as a breakdown product.[22] MEHP was produced by cleavage of one of two ester bonds in DEHP (Figure 5). We surmise that if MEHP was produced from DEHP by gamma-ray irradiation, PA would be produced by the same mechanism.

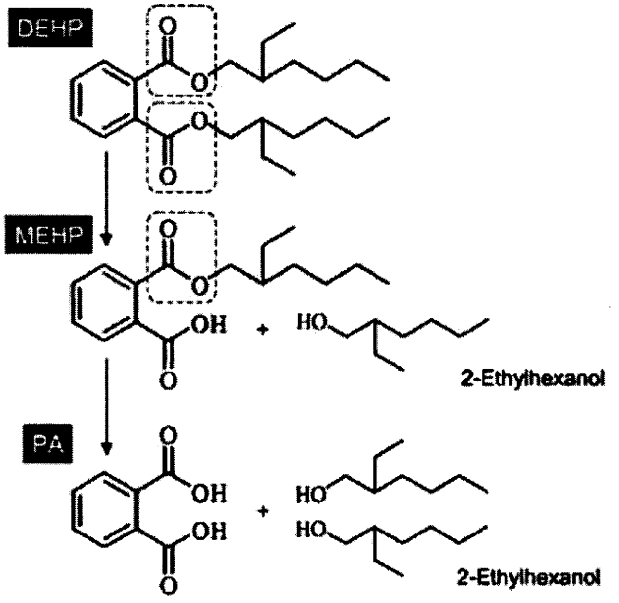


Figure 5. Chemical structures of DEHP, MEHP, and PA. The dotted circle represents the ester bond. MEHP was produced by cleavage of one of two ester bonds in DEHP. PA was produced by cleavage of two ester bonds in DEHP, and by cleavage of an ester bond in MEHP.

CONCLUSIONS

In this study, a method for the simultaneous determination of DEHP, MEHP, and PA was developed. The method had sufficient precision and accuracy to determine the concentrations of DEHP and its breakdown products migrating from PVC medical devices. Using the developed method, not only MEHP but also PA was found to be the breakdown product of DEHP. MEHP is thought to be more toxic than DEHP. The assessment of DEHP exposure in high risk patients is necessary to determine exposure to MEHP and PA, although PA is not as toxic as MEHP.

ACKNOWLEDGMENTS

This study was partly supported by Health Sciences Research Grants from the Ministry of Health, Labour, and Welfare of Japan. This study was partly supported by Science/Technology Frontier Research Base from the Ministry of Education, Science, Sports, and Culture of Japan. This study was partly supported by Grant-in-Aid for Young Scientists (B).

REFERENCES

1. Earls, A.O.; Axford, I.P.; Braybrook, J.H. J. Chromatogr. A **2003**, *983*, 237–246.
2. Inoue, K.; Kawaguchi, M.; Okada, F.; Yoshimura, Y.; Nakazawa, H. Anal. Bioanal. Chem. **2003**, *375*, 527–533.
3. Takatori, S.; Kitagawa, Y.; Kitagawa, M.; Nakazawa, H.; Hori, S. J. Chromatogr. B **2004**, *804*, 397–401.
4. Ito, R.; Seshimo, F.; Miura, N.; Kawaguchi, M.; Saito, K.; Nakazawa, H. J. Pharm. Biomed. Anal. **2005**, *39*, 1036–1041.
5. Koizumi, M.; Ema, M.; Hirose, A.; Hasegawa, R. Jpn. J. Food Chem. **2000**, *7*, 65–73.
6. Koizumi, M.; Ema, M.; Hirose, A.; Kurokawa, Y.; Hasegawa, R. Jpn. J. Food Chem. **2001**, *8*, 1–10.
7. Tickner, J.A.; Schettler, T.; Guidotti, T.; McCally, M.; Rossi, M. Am. J. Ind. Med. **2001**, *39*, 100–111.
8. Yakubovich, M.; Vienken, J. Med. Device Technol. **2000**, *11*, 18–21.
9. Hill, S.; Shaw, B.; Wu, A. Clin. Chim. Acta **2001**, *304*, 1–8.
10. Thomas, J.A.; Northup, S.J. Toxicol. Environ. Health **1982**, *9*, 141–152.
11. Lake, B.G.; Phillips, J.C.; Linnel, J.C.; Gangolli, S.D. Toxicol. Appl. Pharmacol. **1977**, *39*, 239–248.
12. Albro, P.W.; Thomas, R.O. Biochim. Biophys. Acta **1973**, *306*, 380–390.
13. Albro, P.W.; Lavenhar, S.R. Drug Metab. Rev. **1989**, *21*, 13–34.
14. Heindel, J.J.; Chapin, R.E. Toxicol. Appl. Pharmacol. **1989**, *97*, 377–385.
15. Grasso, P.; Heindel, J.J.; Powell, C.J.; Reichert, L.E. Biol. Reprod. **1993**, *48*, 454–459.
16. Richburg, J.H.; Boekelheide, K. Toxicol. Appl. Pharmacol. **1996**, *137*, 42–50.

17. Lee, J.; Richburg, J.H.; Shipp, E.B.; Meistrich, M.L.; Boekelheide, K. Endocrinology **1999**, *140*, 852–858.
18. Richburg, J.H.; Nanex, A.; Williams, L.R.; Embree, M.E.; Boekelheide, K. Endocrinology **2000**, *141*, 787–793.
19. Davis, B.J.; Maronpot, R.R.; Heindel, J.J. Toxicol. Appl. Pharmacol. **1994**, *128*, 216–223.
20. Lovekamp, T.N.; Davis, B.J. Toxicol. Appl. Pharmacol. **2001**, *172*, 217–224.
21. Shintani, H. Chromatographia **2000**, *52*, 721–726.
22. Ito, R.; Seshimo, F.; Miura, N.; Kawaguchi, M.; Saito, K.; Nakazawa, H. J. Pharm. Biomed. Anal. **2006**, *41*, 455–460.
23. Haishima, Y.; Seshimo, F.; Higuchi, T.; Yamazaki, H.; Hasegawa, C.; Izumi, S.; Makino, T.; Nakahashi, K.; Ito, R.; Inoue, K.; Yoshimura, Y.; Saito, K.; Yagami, T.; Tsuchiya, T.; Nakazawa, H. Int. J. Pharm. **2005**, *298*, 126–142.
24. Ito, R.; Seshimo, F.; Haishima, Y.; Hasegawa, C.; Isama, K.; Yagami, T.; Nakahashi, K.; Yamazaki, H.; Inoue, K.; Yoshimura, Y.; Saito, K.; Tsuchiya, T.; Nakazawa, H. Int. J. Pharm. **2005**, *303*, 104–112.

Received June 22, 2007
Accepted July 11, 2007
Manuscript 6161

ANALYTICAL SCIENCES NOVEMBER 2008, VOL. 24

Notes

Measurement of Benzophenones in Human Urine Samples by Stir Bar Sorptive Extraction and Thermal Desorption-Gas Chromatography-Mass Spectrometry

Migaku KAWAGUCHI,* Rie ITO,* Hidehiro HONDA,* Naoyuki ENDO,* Noriya OKANOUCHI,* Koichi SAITO,* Yasuo SETO, and Hiroyuki NAKAZAWA*†**

**Department of Analytical Chemistry, Faculty of Pharmaceutical Sciences, Hoshi University, 2-4-41 Ebara, Shinagawa, Tokyo 142-8501, Japan*
***National Research Institute of Police Science, 6-3-1 Kashiwanoha, Kashiwa, Chiba 277-0882, Japan*

Determination of benzophenones (BPs) in human urine samples by stir bar sorptive extraction (SBSE) and thermal desorption (TD)-gas chromatography-mass spectrometry (GC-MS) is described. As analytes, BP, its metabolites benzhydrol (BP-OH) and 2-hydroxybenzophenone (2OH-BP), and its derivatives 2-hydroxy-4-methoxybenzophenone (BP-3) and 2-hydroxy-4-methoxy-4′-methylbenzophenone (BP-10) were selected. After enzymatic hydrolysis, a polydimethylsiloxane (PDMS) stir bar was placed in a urine sample diluted 1:1 with water and stirred for 60 min at room temperature. The limit of quantification (LOQ) of BPs is 0.2 - 0.5 ng ml^{-1} (ppb). The method showed linearity over the calibration range (0.2 - 10 or 0.5 - 10 ng ml^{-1}), and the correlation coefficients were equal to or higher than 0.993 for all of the analytes. The average recoveries of BPs were equal to or higher than 98.7% (RSD: 1.5 - 4.8%, $n = 6$).

(Received July 3, 2008; Accepted September 2, 2008; Published November 10, 2008)

Introduction

Benzophenone (BP) and its derivatives are the most commonly used sunscreen agents in cosmetics. However, various studies have revealed the estrogenic activity of BPs.[1-3] The migration of BPs from multilayer plastic-paper material intended for food packaging and the contamination of food samples with BPs have recently been reported.[4,5] The fact that healthy humans may be exposed to BPs *via* a variety of daily activities indicates that the assessment of human exposure to BPs is an important issue.

It has been reported that when a human ingests BP, it is excreted in urine as a metabolite, such as benzhydrol (BP-OH) glucuronide.[6] Thus, it is thought that human exposure can be presumed by measuring these compounds in human urine samples. However, because the concentration of BPs in human urine is at the sub ng ml^{-1} level, a method with both high sensitivity and high accuracy is required.

Several analytical methods for the determination of BPs in human urine samples have been reported, including liquid chromatography (LC) with diode array detection (DAD),[7] and tandem mass spectrometry (MS-MS).[8,9] On the other hand, gas chromatography-mass spectrometry (GC-MS) is useful for the determination of BPs in human urine samples.[10]

The sample preparation method known as solid-phase micro extraction (SPME) has been achieved the determination of BPs in human urine samples.[10] However, because the limit of detection (LOD) of 2-hydroxy-4-methoxybenzophenone (BP-3) was 5 ng ml^{-1}, the sensitivity of the above SPME method remains low. In 1999, a sorptive extraction technique that uses a stir bar coated with polydimethylsiloxane (PDMS) was developed,[11] and is known as stir bar sorptive extraction (SBSE). Its main advantages compared to a general sample preparation technique are high sensitivity and a wide application range that includes volatile aromatics, halogenated solvents, polycyclic aromatic hydrocarbons (PAHs), polychlorinated biphenyls (PCBs), pesticides, preservatives, odor compounds and organotin compounds.[12] Rodil and Moeder have reported that the determination of UV filters contains BP-3 in a water sample by SBSE method.[13] In addition, we have reported that the trace analysis of BPs in river water samples by the SBSE method,[14,15] and liquid-phase microextraction (LPME) method.[16]

The aim of this study was to determine trace amounts of BPs in human urine samples by the SBSE and thermal desorption (TD)-GC-MS methods. Emphasis was placed on the determination of BP, its metabolites BP-OH and 2-hydroxybenzophenone (2OH-BP), and its derivatives BP-3 and 2-hydroxy-4-methoxy-4′-methylbenzophenone (BP-10). The developed method was applied to human urine samples.

Experimental

Materials and reagents

Benzophenone (BP) and benzophenone-d_{10} (BP-d_{10}) as internal standard were purchased from Kanto Chemical Inc. (Tokyo, Japan). Benzhydrol (BP-OH) and 2-hydroxybenzophenone (2OH-BP) were purchased from Wako Pure Chemical, Inc. (Osaka, Japan). 2-Hydroxy-4-methoxybenzophenone (oxybenzone, BP-3) was purchased from Sigma-Aldrich Co. (St. Louis, MO). 2-Hydroxy-4-methoxy-4′-methylbenzophenone (BP-10) was purchased from Lancaster Synthesis (Morecambe, England). The chemical structures are shown in Fig. 1. *E. coli*

† To whom correspondence should be addressed.
E-mail: nakazawa@hoshi.ac.jp
M. K. present address: Bio-Medical Standard Section, National Metrology Institute of Japan (NMIJ), National Institute of Advanced Industrial Science and Technology (AIST), 1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, Japan.

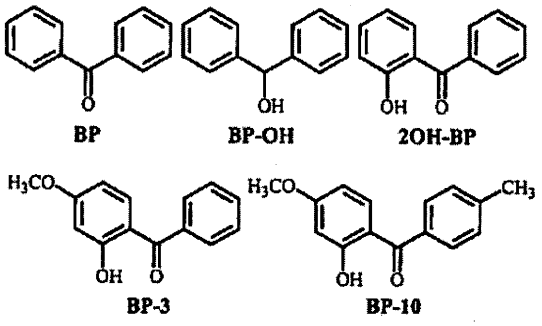


Fig. 1 Chemical structures of BPs.

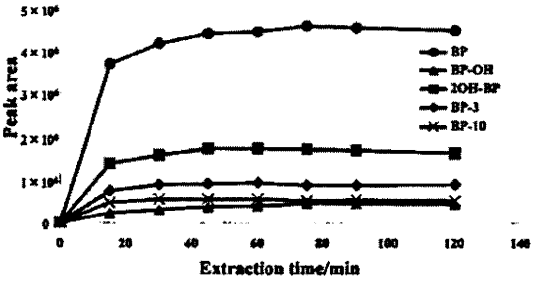


Fig. 2 Extraction time profiles of BPs in human urine sample subjected to the SBSE method.

β-glucuronidase (25000 units/0.4 ml, 62500 units ml^{-1}) and *H. pomatia sulfatase* (3540 units ml^{-1}) were purchased from Sigma-Aldrich Co. Prior to use, the β-glucuronidase was added to 0.1 M ammonium acetate to make a total concentration of 10000 units ml^{-1}. Other reagents were purchased from Wako Pure Chemical, Inc. The water-purification system used was a Milli-Q gradient A 10 with an EDS polisher (Millipore, Bedford, MA).

Standard solutions

Stock solutions (1.0 mg ml^{-1}) of BP, BP-OH, 2OH-BP, BP-3 and BP-10 standards were dissolved in methanol. Standard solutions used for calibration were prepared by the addition of purified water. The calibration curve was performed daily for all samples containing an internal standard (5 ng ml^{-1}) using the SBSE and TD-GC-MS methods. For quantification, the internal standard method was used.

SBSE tools

Stir bars coated with a 0.5-mm-thick PDMS layer (24 µl; Twister™) were obtained from Gerstel (Mülheim an der Ruhr, Germany). The stir bars were conditioned for 1 h at 300°C in a flow of helium. Then, the stir bars were stored in new 2 ml vials until immediately prior to use. The stir bars could be used more than 50 times with appropriate reconditioning. For extraction, a 10 ml headspace vial from Agilent Technologies (Palo Alto, CA) was used.

Instrumentation

TD was performed with a Gerstel TDS 2 thermodesorption system equipped with a Gerstel TDS A autosampler and a Gerstel Cooled Injection System (CIS) 4 programmable temperature vaporization (PTV) inlet. GC-MS was performed with an Agilent 6890N gas chromatograph equipped with a 5973N mass-selective detector with an ultra ion source (Agilent Technologies).

The TDS 2 temperature was programmed from 20°C (held for 1 min) to 250°C (held for 5 min) at 60°C min^{-1}. The desorbed compounds were cryofocused in CIS 4 at −150°C. Then, the CIS 4 temperature was programmed from −150 to 300°C (held for 10 min) at 12°C s^{-1} to inject trapped compounds into the analytical column. Injection was performed in the solvent vent mode. Separations were conducted on a DB-5ms fused silica column (30 m × 0.25 mm i.d., 0.25 µm film thickness, J & W Scientific, Agilent Technologies). The oven temperature was programmed from 40°C (held for 1 min) to 190°C at 5°C min^{-1} and from 190°C to 280°C (held for 3 min) at 15°C min^{-1}. Helium was used as the carrier gas at a flow rate of 1.2 ml min^{-1}. The mass spectrometer was operated in the selected ion-monitoring (SIM) mode with electron ionization (EI) at 70 eV.

Table 1 Figures of merit of SBSE and TD-GC-MS

Compound	SIM[a] (*m/z*)	LOD[b]/ ng ml^{-1}	LOQ[c]/ ng ml^{-1}	Range/ ng ml^{-1}	Correlation coefficient, *r*
BP	<u>182</u>, 105	0.1	0.5	0.5 - 10	0.999
BP-OH	<u>184</u>, 105	0.05	0.2	0.2 - 10	0.998
2OH-BP	<u>197</u>, 121	0.1	0.5	0.5 - 10	0.997
BP-3	<u>227</u>, 151	0.1	0.5	0.5 - 10	0.993
BP-10	<u>241</u>, 227	0.1	0.5	0.5 - 10	0.994

a. The underlined number is the *m/z* of the ion used for quantification.
b. LOD: limit of detection (*S/N* = 3).
c. LOQ: limit of quantification (*S/N* > 10).

The ions selected for ion monitoring are given in Table 1.

Urine samples

Urine was collected from six healthy volunteers (I, II, III, IV, V, and VI). All samples were stored at 4°C prior to use.

Sample preparation

One milliliter of a urine sample spiked with BP-d_{10} (5 ng ml^{-1}) was pipetted into a 10-ml vial. Then, 1.0 M ammonium acetate (100 µl, pH 6.8) was added. After adding β-glucuronidase (10 µl, 10000 units ml^{-1}) and sulfatase (10 µl, 3540 units ml^{-1}), the sample was gently mixed. A hydrolysis process to release free BPs was accomplished by incubating at 37°C for 3 h. One milliliter of water was added to the sample and SBSE was performed at room temperature for 60 min while stirring at 500 rpm. After extraction, the PDMS stir bar was thermally desorbed, and subjected to GC-MS thereafter.

Results and Discussion

Optimization of extraction time

One important parameter affecting SBSE was the extraction time. To optimize the extraction time, 1 ml of human urine sample (5 ng ml^{-1} BPs standard solutions) and 1 ml of purified water were mixed and used. The extraction time profiles (0 - 180 min) of BPs in the human urine sample that was subjected to SBSE were determined by TD-GC-MS, and are shown in Fig. 2. BPs reached equilibrium after approximately 60 min. Therefore, this condition was used for the determination of BPs in human urine samples.

Figures of merit of SBSE and TD-GC-MS for the determination of BPs

The calculated LODs of BPs were 0.05 to 0.1 ng ml^{-1} for

Table 2 Recoveries of BPs in human urine samples

Compound	Amount spiked			
	1 ng ml^{-1}		10 ng ml^{-1}	
	Recovery, %	RSD, %[a]	Recovery, %	RSD, %[a]
BP	101.5	2.4	98.7	1.5
BP-OH	99.1	4.3	99.4	4.8
2OH-BP	101.7	3.0	100.0	3.4
BP-3	99.7	3.2	100.7	2.2
BP-10	99.3	4.0	100.5	2.3

a. The recoveries and precision were also examined by replicate analysis (n = 6) of human urine samples.

Table 3 Concentrations of BPs in human urine samples

Compound	Human urine sample/ng ml^{-1}					
	I	II	III	IV	V	VI
BP	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
BP-OH	3.7	0.5	0.4	<LOQ	0.4	<LOQ
2OH-BP	<LOD	<LOD	<LOD	<LOD	<LOD	<LOD
BP-3	1.2	0.5	<LOQ	<LOQ	<LOD	<LOD
BP-10	<LOD	<LOQ	<LOD	<LOD	1.5	<LOQ

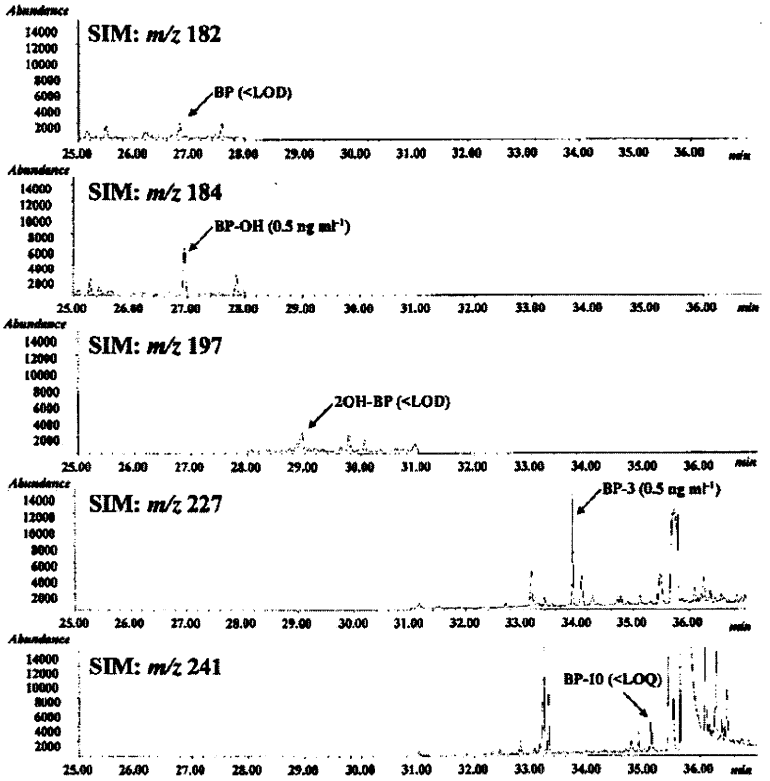


Fig. 3 SIM chromatograms of BPs in human urine sample (II).

SBSE and TD-GC-MS, with the signal to noise (S/N) ratio being 3. In addition, the limits of quantification (LOQs) of BPs when S/N > 10 were 0.2 to 0.5 ng ml^{-1}. The method showed good linearity and the correlation coefficients (r) were equal to or higher than 0.993 for all of the analytes. The figures of merit of the present method are summarized in Table 1. A comparison of the SBSE method with the SPME method,[10] used in previous studies revealed that the SBSE method was superior to the SPME method in terms of sensitivity.

The recovery and precision of the method were assessed by replicate analysis (n = 6) of human urine samples spiked at 1 and 10 ng ml^{-1} levels. Non-spiked and spiked samples were subjected to SBSE and TD-GC-MS. The recovery was calculated by subtracting the results for non-spiked samples from those for spiked samples. The results were obtained by using calibration curves acquired from standard solutions cotaining the internal standard. The recovery and the precision were 98.7 to 101.7% (RSD: 1.5 to 4.8%, n = 6) for human urine samples (Table 2). Therefore, the method enables the precise determination of standards, and may be applicable to the determination of trace amounts of BPs in human urine samples.

Determination of BPs in human urine samples

A total of six human urine samples were analyzed for BPs using the presented method, the results are given in Table 3. BP-OH (<LOQ ~ 3.7 ng ml^{-1}), BP-3 (<LOD ~ 1.2 ng ml^{-1}), and BP-10 (<LOD ~ 1.5 ng ml^{-1}) were detected in the samples with the present method. On the other hand, BP and 2OH-BP were not detected in the urine samples examined. Typical SIM chromatograms of the human urine sample (II) are shown in Fig. 3. SBSE and TD-GC-MS enabled the successful determination of trace amounts of BPs in the urine sample. The BP levels in the

urine samples were very low and could not be quantified by SPME-GC-MS.[11]

Conclusions

The determination of trace amounts of BPs in human urine samples using SBSE and TD-GC-MS was described. The proposed method has many practical advantages, including a small sample volume (1 ml) and simplicity of extraction; it is also solvent-free and has high sensitivity. The LODs of BPs were of sub ng ml^{-1} level. In addition, the present method showed good linearity and high correlation coefficients using the internal standard. The recovery was 98.7 to 101.7% and the precision was RSD: 1.5 to 4.8% ($n = 6$) for human urine samples spiked at 1 and 10 ng ml^{-1} levels.

Acknowledgements

This study was supported by "Health Sciences Research Grants from the Ministry of Health, Labour and Welfare of Japan", "Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists", "Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology" and "The Hoshi University Otani Research Grant".

References

1. D. Miller, B. B. Wheals, N. Beresford, and J. P. Sumpter, *Environ. Health Perspect.*, **2001**, *109*, 133.
2. M. Schlumpf, B. Cotton, M. Conscience, V. Haller, B. Steinmann, and W. Lichtensteiger, *Environ. Health Perspect.*, **2001**, *109*, 239.
3. S. Takatori, Y. Kitagawa, H. Oda, G. Miwa, J. Nishikawa, T. Nishihara, H. Nakazawa, and S. Hori, *J. Health Sci.*, **2003**, *49*, 91.
4. C. Nerín and E. Asensio, *Anal. Bioanal. Chem.*, **2007**, *389*, 589.
5. L. Castle, C. P. Offen, M. J. Baxter, and J. Gilbert, *Food Addit. Contam.*, **1997**, *14*, 35.
6. Y. Nakagawa, T. Suzuki, and S. Tayama, *Toxicology*, **2000**, *156*, 27.
7. V. Sarveiya, S. Risk, and H. A. E. Benson, *J. Chromatogr., B*, **2004**, *803*, 225.
8. X. Ye, Z. Kuklenyik, L. L. Needham, and A. M. Calafat, *Anal. Chem.*, **2005**, *77*, 5407.
9. X. Ye, Z. Kuklenyik, L. L. Needham, and A. M. Calafat, *Anal. Bioanal. Chem.*, **2005**, *383*, 638.
10. T. Felix, B. J. Hall, and J. S. Brodbelt, *Anal. Chim. Acta*, **1998**, *371*, 195.
11. E. Baltussen, P. Sandra, F. David, and C. Cramers, *J. Microcol. Sep.*, **1999**, *11*, 737.
12. M. Kawaguchi, R. Ito, K. Saito, and H. Nakazawa, *J. Pharm. Biomed. Anal.*, **2006**, *40*, 500.
13. R. Rodil and M. Moeder, *J. Chromatogr., A*, **2008**, *1179*, 81.
14. M. Kawaguchi, R. Ito, N. Endo, N. Sakui, N. Okanouchi, K. Saito, N. Sato, T. Shiozaki, and H. Nakazawa, *Anal. Chim. Acta*, **2006**, *557*, 272.
15. M. Kawaguchi, R. Ito, H. Honda, N. Endo, N. Okanouchi, K. Saito, Y. Seto, and H. Nakazawa, *J. Chromatogr., A*, **2008**, *1200*, 260.
16. N. Okanouchi, H. Honda, R. Ito, M. Kawaguchi, K. Saito, and H. Nakazawa, *Anal. Sci.*, **2008**, *24*, 627.

DNA methylation profile of tissue-dependent and differentially methylated regions (T-DMRs) in mouse promoter regions demonstrating tissue-specific gene expression

Shintaro Yagi,[1] Keiji Hirabayashi,[1] Shinya Sato,[1] Wei Li,[2] Yoko Takahashi,[1] Tsutomu Hirakawa,[1] Guoying Wu,[1] Naoko Hattori,[1] Naka Hattori,[1] Jun Ohgane,[1] Satoshi Tanaka,[1] X. Shirley Liu,[3] and Kunio Shiota[1,4,5]

[1]*Laboratory of Cellular Biochemistry, Department of Animal Resource Sciences/Veterinary Medical Sciences, The University of Tokyo, Tokyo 113-8657, Japan;* [2]*Division of Biostatistics, Dan L. Duncan Cancer Center, Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, Texas 77030, USA;* [3]*Department of Biostatistics and Computational Biology, Dana-Farber Cancer Institute, Harvard School of Public Health, Boston, Massachusetts 02115, USA;* [4]*National Institute of Advanced Industrial Science and Technology, Tsukuba, Ibaraki 305-8561, Japan*

DNA methylation constitutes an important epigenetic regulation mechanism in many eukaryotes, although the extent of DNA methylation in the regulation of gene expression in the mammalian genome is poorly understood. We developed D-REAM, a genome-wide DNA methylation analysis method for tissue-dependent and differentially methylated region (T-DMR) profiling with restriction tag-mediated amplification in mouse tissues and cells. Using a mouse promoter tiling array covering a region from −6 to 2.5 kb (~30,000 transcription start sites), we found that over 3000 T-DMRs are hypomethylated in liver compared to cerebrum. The DNA methylation profile of liver was distinct from that of kidney and spleen. This hypomethylation profile marked genes that are specifically expressed in liver, including key transcription factors such as *Hnf1a* and *Hnf4a*. Genes with T-DMRs, especially those lacking CpG islands and those with HNF-1A binding motifs in their promoters, showed good correlation between their tissue-specific expression and liver hypomethylation status. T-DMRs located downstream from their transcription start sites also showed tissue-specific gene expression. These data indicate that multilayered regulation of tissue-specific gene function could be elucidated by DNA methylation tissue profiling.

[Supplemental material is available online at www.genome.org and at http://www.vm.a.u-tokyo.ac.jp/seika/D-REAM/. The array data from this study have been submitted to ArrayExpress (http://www.ebi.ac.uk/microarray-as/ae/) under accession no. E-TABM-551.]

In multicellular organisms, cells, and tissues form as a result of differentiation of a single fertilized egg, and phenotypes are inherited over several cell generations without alteration in the DNA sequences. Epigenetic systems are recognized as memory systems for these inheritable gene functions and, in mammals, they comprise DNA methylation and histone modifications of chromatin. DNA methylation in tissue-dependent and differentially methylated regions (T-DMRs) is involved in expression of tissue-specific genes as well as expression of key transcription factors that constitute transcription networks governing tissue or cell specificity (Shen and Maniatis 1980; Cho et al. 2001; Imamura et al. 2001; Hattori et al. 2004b, 2007; Nishino et al. 2004). Abnormal methylation of T-DMRs has been implicated in the pathogenesis of certain diseases (Jones 2002; Ushijima 2005).

DNA methylation occurs at the cytosine residue of CpG dinucleotides, which are unevenly distributed within the mammalian genome (Bird 1980). CpG islands (CGIs) have been identified as CpG-rich regions that are associated with ~50% of the promoter regions in the mouse genome (Bird et al. 1985; Gardiner-Garden and Frommer 1987). Previous genome-wide DNA methylation analyses, focusing on CGIs, have indicated that every cell and tissue type has a unique DNA methylation profile, comprising at least hundreds of T-DMRs (Ohgane et al. 1998; Shiota et al. 2002; Strichman-Almashanu et al. 2002), and these data suggested that a methylation profile could be used to identify cell types (Shiota 2004).

To identify genes with differentially methylated regions, several microarray technologies have been developed (Lieb et al. 2006), and microarray technology has been applied to identify aberrantly methylated regions in cancer cells and characterize cell lines such as human embryonic stem cells (Hatada et al. 2006; Keshet et al. 2006; Ordway et al. 2006; Rauch et al. 2006; Shen et al. 2006). However, the DNA methylation profiles obtained have not been directly related to gene function (Ching et al. 2005; Eckhardt et al. 2006; Khulan et al. 2006). The limited number of loci or regions available for genome-wide analysis of normal cells or tissues and the existence of method biases can affect the implementation of methylated profiles (for review, see Khulan et al. 2006).

[5]Corresponding author.
E-mail ashiota@mail.ecc.u-tokyo.ac.jp; fax 81-3-5841-8189.
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.074070.107.

We developed a novel, low-bias method for genome-wide DNA methylation analysis, and examined the DNA methylation profile of the promoter regions in normal mouse liver by comparing them with those of cerebrum, kidney, and spleen. The results indicate that the resultant methylation profile was implicated in tissue-specific function.

Results

Features of T-DMR profiling by restriction tag-mediated amplification

To illustrate the genome-wide mouse DNA methylation profile, we developed a method involving T-DMR profiling with restriction tag-mediated amplification (D-REAM), which combined microarray technology and modified ligation-mediated polymerase chain reaction (LM-PCR) (Fig. 1A; Supplemental Fig. S1A). D-REAM recognizes various combinations of microarrays and restriction enzymes for LM-PCR. In this study, we used the GeneChip DNA microarray, tiled with ~4.4 million 25-nt oligomers corresponding to regions located from approximately −6 to 2.5 kb of transcription start sites (TSS) of 30,120 Ensembl mouse transcripts (Supplemental Table S1; Supplemental Fig. S2A). To screen unmethylated regions, we used HpyCH4IV, a methylation-sensitive restriction enzyme that recognizes ACGT residues. Such residues are distributed throughout the mouse genome in a less biased manner than HpaII sites, which are localized mainly at CGIs around the TSS on the promoter array (Supplemental Fig. S2B). The microarray probes covered regions comprising 10.1% of all the HpyCH4IV sites (1,751,098) in the mouse genome (Fig. 1C).

When the mouse genome was digested with HpyCH4IV in silico, it generated fragments with median and average sizes of 907 and 1468 bp, respectively (Supplemental Fig. S2B). Because of the hypermethylation status of mouse genome, actual mouse liver HpyCH4IV fragments were deemed too large for efficient PCR amplification. To address this issue, the D-REAM method uses TaqI, a methylation-insensitive restriction enzyme, to reduce the size of the single-digested fragments (Fig. 1B). The modified LM-PCR protocol facilitates the selective amplification of unmethylated HpyCH4IV–TaqI and HpyCH4IV–HpyCH4IV, but not TaqI–TaqI fragments (Supplemental Fig. S1A,B). The selective amplification by this process was confirmed by using fragments digested by rarely occurring restriction enzymes, such as NotI (Supplemental Fig. S1C).

To analyze the D-REAM microarray data, we applied model-based analysis of tiling arrays (MAT) (Johnson et al. 2006). MATscores represent the enrichment of unmethylated fragments and are indicative of the relative methylation status of the HpyCH4IV site. They were visualized using the Integrated Genome Browser (IGB). Intra-genomic comparison can be affected by PCR sequences or fragment length biases; however, these effects were minimized by comparing the same regions in different tissues or cells. In this study, we designated these differentially methylated HpyCH4IV sites as T-DMRtags that represent T-DMRs (Fig. 1A).

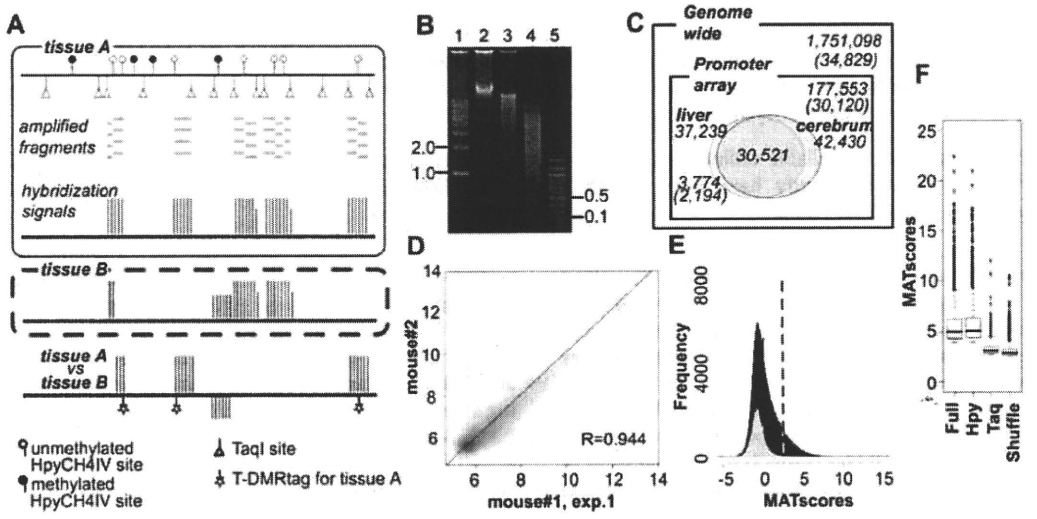


Figure 1. DNA methylation profiles were analyzed by D-REAM. (*A*) Illustration of the D-REAM method. Genomic DNA was digested with methylation-sensitive restriction enzyme HpyCH4IV and amplified by modified LM-PCR (Supplemental Fig. S1). Amplified fragments (gray bars) were hybridized with mouse promoter tiling array (*upper* panel). Array signal intensities (vertical bars) were analyzed to identify regions corresponding to fragments in unmethylated HpyCH4IV loci. Comparison of signals from different samples enabled identification of differentially methylated regions (*lower* panel). HpyCH4IV loci overlapping with regions yielding differential signals were defined as T-DMRtags. (*B*) Agarose gel electrophoresis of undigested (lane *2*), HpyCH4IV-digested (lane *3*), and HpyCH4IV–TaqI-digested (lane *4*) mouse liver DNA. Positions corresponding to 0.1, 0.5, 1.0, and 2.0 kbp (lanes *1*,*5*) are indicated on one side of the gel image. (*C*) Venn diagram of DNA methylation status at HpyCH4IV sites in mouse liver and cerebrum. Numbers without parentheses represent numbers of HpyCH4IV sites, while Ensembl transcripts IDs are in parentheses. *Outer* and *inner* rectangles represent whole mouse genome and regions covered by the promoter tiling array, respectively. Ovals indicate unmethylated HpyCH4IV sites of liver and cerebrum identified by D-REAM. (*D*) Correlation of microarray probe intensities in duplicate mouse liver experiments, plotted on logarithmic axes (base 2). (*E*) MATscore distribution of array regions corresponding to the TaqI–TaqI fragments (gray) and HpyCH4IV-digested fragments (black). The dotted line represents the MATscore cutoff value. (*F*) Reliability of comparative MAT analysis. Bar-plots of MATscores of the hypomethylated regions identified by MAT ($P < 10^{-3}$) using full .bpmap (Full) and subsets of .bpmap corresponding to HpyCH4IV fragments (Hpy) and TaqI–TaqI fragments (Taq). Shuffle column MATscores obtained by using both treatment and control samples containing both liver and cerebrum data from different mice. The boxes, and lines within the boxes, represent the interquartile ranges and medians of the ratios, respectively.

T-DMRs hypomethylated in mouse liver compared to cerebrum

D-REAM assays of liver DNA obtained from two male mice were performed, and reproducible microarray results were obtained (Fig. 1D; Supplemental Fig. S1D). The number of fragments above the MATscore cutoff value of 2.44 was larger at regions corresponding to HpyCH4IV fragments than at those corresponding to TaqI–TaqI fragments (Fig. 1E). This indicated selective amplification of fragments with unmethylated HpyCH4IV sites. At least 20.9% of HpyCH4IV sites were estimated to be unmethylated within the regions covered by the microarray (Fig. 1C). In mouse cerebra (correlation coefficient of 0.955 of the biological duplicate), MAT analysis predicted 42,430 sites of unmethylated HpyCH4IV sites ($P < 10^{-3}$), and 71.9% (30,521 loci) of these sites overlapped those in the liver tissue.

In a comparative analysis of liver and cerebrum, the differences in MATscores were significantly large at regions of HpyCH4IV fragments compared to those of HpyCH4IV fragments with a shuffled combination, and to those of TaqI–TaqI fragments (Fig. 1F). A total of 3774 of differentially hypomethylated T-DMRtags were identified in liver ($P < 10^{-3}$). These tags were located in neighboring 10-kb regions of the TSS for 2194 (7.28% of all IDs) Ensembl transcripts (Fig. 1C).

Positions of T-DMRs relative to the TSS

We classified the genes for these 2194 transcripts into two types, CGI and non-CGI genes, according to TSS position. CGI gene TSS are located within 1 kb from the CGI, and non-CGI genes are located further than 1 kb from the CGI. The positions of T-DMRtags were distinct using this criterion (Fig. 2A). T-DMRtags neighboring non-CGI genes were observed up to −6 kb distal to and 2.5 kb downstream from TSS, and exhibited distribution patterns similar to the probes. No correlations were observed between T-DMR distribution and CpG density (observed/expected), GC percentage, or localization of HpyCH4IV sites (Fig. 2A; Supplemental Fig. S2A). Among the CGI genes, the distribution of T-DMRs seemed to negatively correlate with CpG density and the sequence conservation score among animals, suggesting that CGIs and first exons would be T-DMR-poor regions (Fig. 2A,B).

T-DMRtags in eight liver-specific non-CGI genes were confirmed using combined bisulfite restriction analysis (COBRA) (Supplemental Fig. S3A,B; Xiong and Laird 1997). The degrees of methylation were not proportional to the differences in the MATscores that were observed in three sites of the *Fga* loci, but all 15 HpyCH4IV sites were confirmed to be differentially hypomethylated in liver. The T-DMRs were distributed from the 5′-upstream to the 3′-downstream region of TSS, similar to the distribution of all T-DMRtags in the non-CGI genes (Fig. 2A).

In liver-specific *Gnmt* gene, T-DMRtags were observed from −5.9 to −4.3 kb 5′ upstream and from 2.2 to 2.4 kb 3′ downstream from TSS. COBRA showed an unmethylated status of HpyCH4IV in CGIs of both liver and cerebrum, and a hypomethylated status of each T-DMRtags in liver (Fig. 2D). Similar meth-

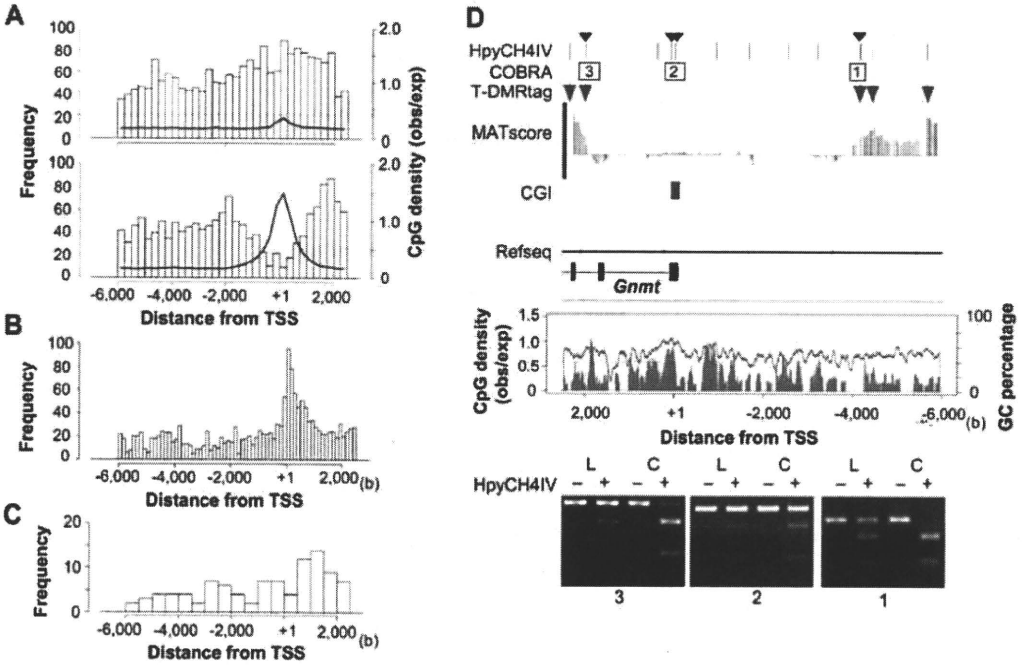


Figure 2. T-DMR positions depend on the genomic context. (*A*) Distribution of positions relative to TSS of hypomethylated T-DMRtags in liver. *Upper* and *lower* panels display distributions in non-CGI and CGI genes, respectively. The width of histogram units is 250 bp. CpG densities are indicated by blue lines. (*B*) Center of phastCons track regions in all CGI genes on chromosomes 5, 12, and 15 (obtained from UCSC genome browser database) plotted with a histogram unit width of 125 bp. (*C*) Positions of T-DMRtags on liver-specific non-CGI genes with HNF1 motifs with expression levels in liver >2-fold those in cerebrum, plotted with a 500-bp histogram category width. (*D*) T-DMRs neighboring the *Gnmt* genes analyzed by COBRA. *Upper* panel displays the position of HpyCH4IV, the regions of restriction mapping with the analyzed HpyCH4IV site, indicated by small arrowheads on the *top*, and positions of T-DMRs plotted over the comparative MATscores on IGB browser from 6000 bp upstream to 2500 bp downstream from the TSS. *Middle* panel shows CpG density (blue) and GC percentage (gray line) in this region. *Bottom* panels show agarose-gel electrophoresis images of COBRA. Hypomethylated fragments converted by bisulfite treatment were resistant to HpyCH4IV digestion (+). L and C indicate liver and cerebrum samples, respectively.

ylation patterns of HpyCH4IV sites in five other CGI genes were confirmed by COBRA (Supplemental Fig. S3C,D).

Biased ontology annotations of genes with T-DMRs with regard to CGIs

To determine features represented by genes with T-DMRs hypomethylated in liver, we applied ontology analysis, using g:GOSt in the g:profiler web database (Table 1; Supplemental Table S2) (Reimand et al. 2007). Significantly enriched ontology terms for biological processes (BPs) ($P < 10^{-5}$) were observed in 1817 g:profiler IDs that had been converted from 2194 Ensembl transcripts with T-DMRs for analysis. These BPs code for genes responsible for metabolism of organic acids and lipids and responses to defense and stress. Among these overrepresented terms, were found liver dominant tissue-specific functions present in the coagulation cascade (described in the Kyoto Encyclopedia of Genes and Genomes [KEGG] pathway database) (Kanehisa et al. 2006), and folate and methyl group metabolism (Fig. 3A,B) (Williams and Schalinske 2007).

The ratios of CGI genes to non-CGI genes were significantly biased among some ontology terms ($P < 5 \times 10^{-2}$, χ^2 test), with non-CGI genes overrepresented among the genes encoding proteins exported to extracellular regions (Table 1). To avoid preexisting bias in the ontology terms in our criteria, we analyzed them using functional annotation tools in the DAVID Bioinformatics Resources 2007 website (Huang da et al. 2007) using lists of CGI or non-CGI genes as background. Most terms observed in the g:GOSt analysis were significantly overrepresented in the non-CGI genes with T-DMRs compared to CGI genes with T-DMRs. Among the CGI genes, only mitochondrial genes were significantly overrepresented, and one of the identified mitochondrial CGI genes with T-DMRs was *Cpt2* (Supplemental Fig. S3C).

Genes encoding transcription factors responsible for expression of liver genes with T-DMRs

Transcription factor motifs were analyzed by referring to the MAPPER database in the 1-kb 5′-upstream region of TSS (Marinescu et al. 2005). When non-CGI genes with T-DMRs were analyzed, we found significantly enriched motifs, including HNF1A, HNF4, and RXRa (Table 2). In contrast, there were no overrepresented motifs among CGI genes with T-DMRs.

Hnf1a and *Hnf4a* are expressed in a tissue-specific manner and are involved in regulating liver-specific gene expression (Schrem et al. 2002). The MATscores and, hence, the methylation status of these regions were significantly different between liver and cerebrum (Fig. 4A,B). Bisulfite sequencing identified four CpGs hypomethylated in liver, corresponding to the T-DMRtags at −523 bp upstream of *Hnf1a* TSS (Fig. 4A). Among the six T-DMRtags of *Hnf4a*, two located at downstream regions of TSS and two at 5′ distal regions were present in the first exon of *0610008F07Rik*, a gene in which transcription initiates from the 5-kb upstream region of *Hnf4a* TSS in the opposite direction. Bisulfite DNA sequencing revealed that 27 CpGs in these 5-kb regions were hypomethylated in liver but hypermethylated in cerebrum (Fig. 4B).

We further investigated T-DMRs in other liver-enriched transcription factors that support liver-specific gene expression (Giguere 1999; Handschin and Meyer 2005). In the cases of *Nr1h3* (*Lxr*) and *Nr1i2* (Pregnane X receptor, *Pxr*), the T-DMRs in

Table 1. Annotation analysis of genes with liver T-DMRtags

Ontology type[a]	Term	DAVID[b]
GO:BP	Lipid metabolic process	NC
	Cellular lipid metabolic process	NC
	Response to stress	
	Generation of precursor metabolites and energy	NC
	Organic acid metabolic process	NC
	Carboxylic acid metabolic process	NC
	Monocarboxylic acid metabolic process	
GO:MF	Vitamin binding	NC
	Cofactor binding	NC
	FAD binding	NC
Biased to non-CGI genes		
GO:BP	Defense response	
	Response to wounding	NC
	Inflammatory response	
	Acute inflammatory response	
GO:CC	Extracellular region	NC
	Extracellular region part	
	Extracellular space	NC
KEGG	Complement and coagulation cascades	NC
Biased to CGI genes		
GO:CC	Cytoplasm	NC
	Cytoplasmic part	NC
	Mitochondrion	NC, CGI
	Mitochondrial part	
GO:MF	Catalytic activity	NC

Detailed data are shown in Supplemental Table S3. χ^2 tests were applied to examine the difference in the proportions of CGI and non-CGI genes for each criterion among all the genes (1817 genes) containing T-DMRtags. Percentage of non-CGI genes among all 1817 genes is 53.6%.
[a]Ontology types of GO:BP, GO:CC, and GO:MF indicate biological process, cellular component, and molecular function in Gene Ontology criteria, respectively. KEGG represents KEGG pathway database.
[b]The DAVID column indicates the overrepresentation of the terms in DAVID 2007 analysis among all genes classified into the same criterion according to the position of CGIs. NC, non-CGI genes; CGI, CGI genes.

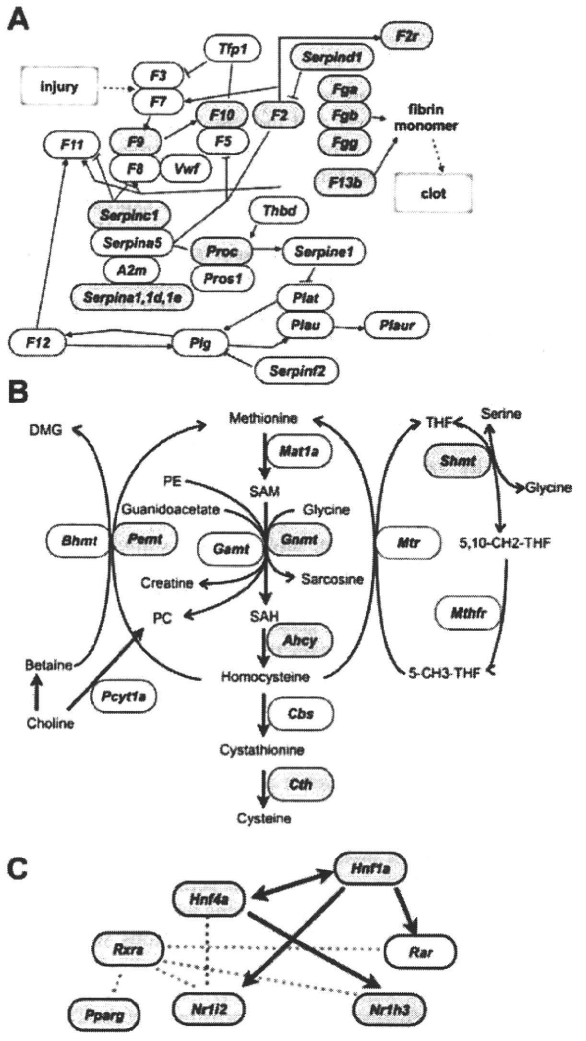


Figure 3. Genes with T-DMRtags (gray) in the complement and coagulation cascade in the modified KEGG pathway map (ID 04610) (*A*), and in the folate and methyl group metabolism pathway (Williams and Schalinske 2007) (*B*). (C) Transcription factor network for liver-specific gene expression. Arrows indicate that the gene expression is controlled by transcription factors. Dotted lines represent molecular interaction between factors.

liver were localized downstream from their TSS (Fig. 4C,D). T-DMRs in the *Rxra* gene were localized between exons 1 and 2 (Fig. 4E), where an alternative TSS for the testis-specific transcript, which is detected in liver, is located (Brocard et al. 1996). These data suggested that T-DMRs could be involved in gene regulation of transcription factors, aiding liver-specific gene expression (Fig. 3C).

Expression of genes with HNF1 motifs and T-DMRs

Previously, 222 HNF1-binding promoters were identified in human hepatocytes by chromatin immunoprecipitation combined with DNA microarray analysis (ChIP-chip) experiments (Odom et al. 2004). Assuming that mouse orthologs of these genes would be bound to HNF1 in mouse liver, we searched for these orthologs using g:Orth in the g:Profiler web database, and observed that 43 genes of 174 orthologs (unique Entrez gene IDs) were accompanied with T-DMRs (Supplemental Fig. S4A). Tissue specificity of expression was investigated in these 43 orthologs and in 180 Entrez gene IDs with HNF-1A motifs (TRANSFAC ID M00790) and T-DMRs (Supplemental Fig. S4B), which were selected by referring to the MAPPER database. The gene expression levels in liver were more than twofold of those in cerebrum in 32 of 43 orthologs and in 68 of 180 genes with HNF-1A motifs (Fig. 5A). In cerebrum, 75% of genes exhibited expression levels <229.6 (gcRMA preprocessed data), which indicated that gene hypermethylation might repress these genes in cerebrum (Fig. 5B). Among CGI genes with T-DMRs, liver-specific expressed genes were significantly overrepresented in genes with HNF-1A motifs, but less significantly in those with HNF4 motifs (Fig. 5C). These data supported the conclusion that HNF1 and T-DMRs are involved in the regulation of these genes in mice.

Correlations between expression levels of genes and T-DMRs in somatic tissues

We examined correlations between T-DMRs and gene expression in liver (Fig. 5C). The expression levels of genes with T-DMRs in liver were significantly higher than those in cerebrum, especially with regard to non-CGI genes. In 77 genes (Ensembl gene IDs) with HNF1-binding motifs that were expressed in liver (Fig. 2C), the T-DMR distribution profile was similar to that of all non-CGI genes (Fig. 2A).

In liver-specific CGI genes, the positions of T-DMRs were biased toward the regions 0.5–2.5 kb, which are 3′ downstream from the TSS (Fig. 5C,D). These regions corresponded to the first introns, judged by the distribution of the gene conservation index and by CpG density (Fig. 2B). These data suggested that T-DMRs in noncore promoter regions could be involved in the regulation of gene plurality.

We analyzed the DNA methylation status at T-DMRs in mouse kidney and spleen in genes with expression levels higher in liver than in cerebrum. The regions corresponding to these T-DMRs were clustered into four groups according to their

Table 2. Transcription factor motifs overrepresented in genes with T-DMRs among non-CGI genes

MAPPER factor name[a]	All	T-DMR	χ^2 test (*P*-value)[b]
ARP-1 (NR2F2, HNF4A)	1580	160	8.76×10^{-4}
COUP direct repeat 1 (NR2F2)	1198	131	1.27×10^{-4}
COUP-TF:HNF-4 (NR2F2, HNF4A)	1292	136	5.17×10^{-4}
ER-alpha (ESR1)	2250	232	6.03×10^{-6}
FOXD3 (FOXD3)	4111	400	2.73×10^{-8}
GCR1[c]	1268	133	7.11×10^{-4}
HNF-1A (HNF1A)	1661	196	2.24×10^{-9}
HNF-4 direct repeat 1 (HNF4A)	1051	116	2.41×10^{-4}
HNF-4alpha (HNF4A)	1232	131	4.45×10^{-4}
HNF-4alpha1 (HNF4A)	3113	348	6.45×10^{-15}
NF-kappaB (NFKB)	4484	430	2.08×10^{-8}
RAP1[c]	4201	379	5.03×10^{-4}
RXR-alpha (RXRA)	1230	142	3.11×10^{-6}
Zic3 (ZIC3)	386	54	4.42×10^{-5}

[a]Detailed information is available from MAPPER factors table (http://bio.chip.org/mapper/factors-table, free registration required). Symbols in parentheses correspond to the factors described in the models.
[b]χ-Square tests were applied to examine the difference in the proportions of each transcription factor motif between all non-CGI and non-CGI carrying T-DMRs.
[c]Models are based on the yeast transcription factors.

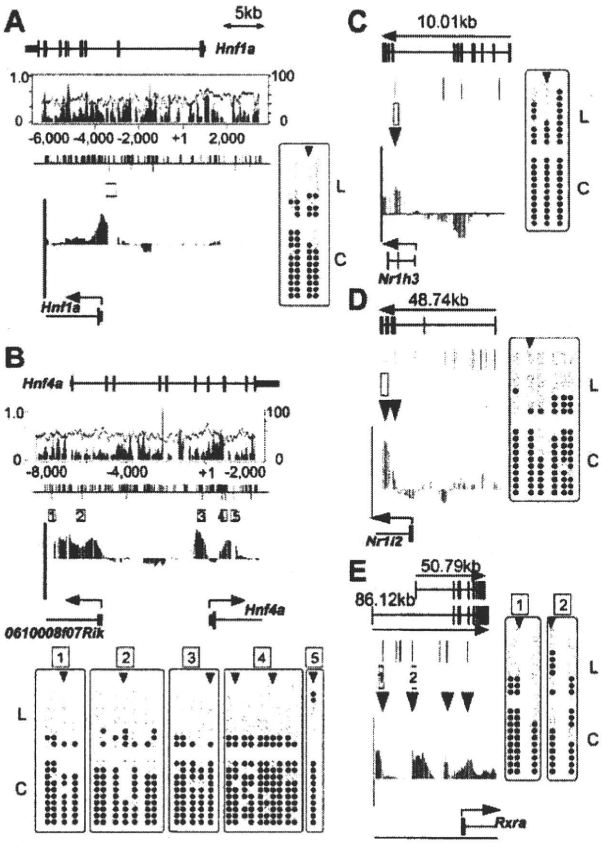


Figure 4. Bisulfite sequencing of T-DMRs of liver-specific transcription factors, as *Hnf1a* (*A*), *Hnf4a* (*B*), *Nr1h3* (*C*), *Nr1i2* (*D*), and *Rxra* (*E*). Genomic structures are presented at the *top* of each figure section. The graphs in boxes toward the center in *A* and *B* represent CpG density (blue) and GC percentages (gray). The bars visible along the *top* of the center lines in *A* and *B* represent CpG dinucleotide positions; bars *below* represent HpyCH4IV sites. Boxes and arrowheads represent T-DMRs and T-DMRtags, respectively. IGB plots of comparative microarray signals corresponding to the regions in the abovementioned figures are displayed toward the *bottom* of the middle sections. Bisulfite sequencing data obtained for 10 isolates from liver (L) and cerebrum (C) are summarized at the *bottom* or *side* of the figure section. Open and closed circles represent unmethylated and methylated CpG, respectively.

MATscores of comparative D-REAM analysis, representing relative methylation status (Fig. 5E). Distinct DNA methylation profiles were observed among tissues, with spleen hypomethylated T-DMRs classified into clusters 1 and 2. Cluster 3 contained genes that hypomethylated in both liver and kidney, while cluster 4 contained genes concerned with liver-specific hypomethylation. Although the expression of these genes was repressed in both spleen and cerebrum (Supplemental Fig. S5A), the expression levels in kidney were similar to those in liver in clusters 1 and 2, exhibited a greater distribution toward kidney in cluster 3, and toward liver in cluster 4 (Fig. 5F; Supplemental Fig. S5). In the genes with HNF-1A motifs, *Hnf4a* was classified into cluster 3, while *Nr1h3*, *Nr1i2*, and *Serpina1e* were placed in cluster 4 (Supplemental Figs. S4, S5B). These data indicated that tissue specificity of gene expression in these regions would reflect patterns of T-DMRs of the genes containing tissue-specific transcription factors and their targets.

Discussion

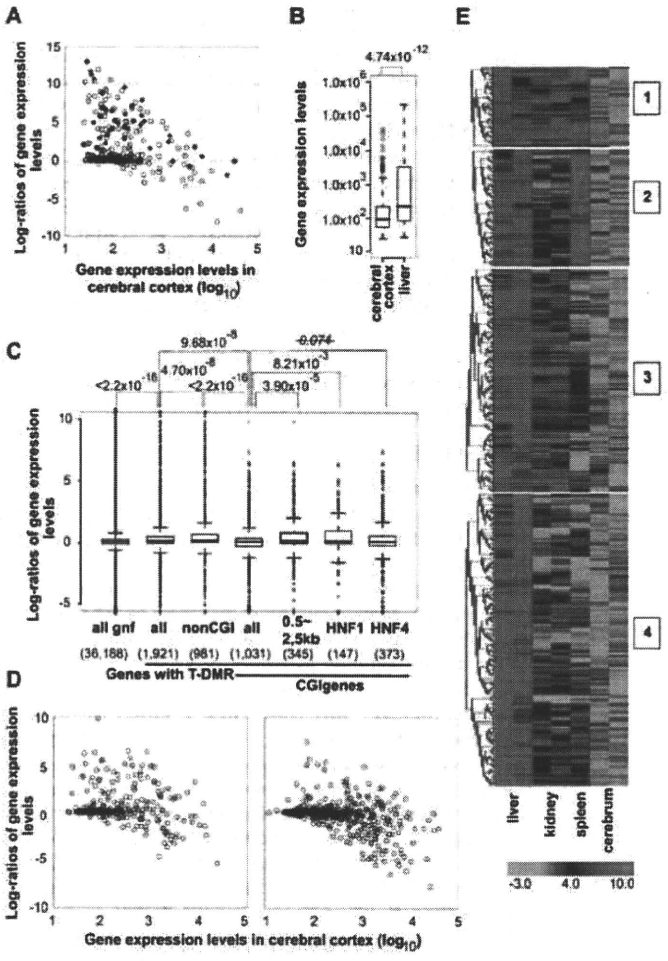
D-REAM, performed using a high-density genome tiling array for the mouse promoter regions, revealed thousands of T-DMRs that are hypomethylated in liver and indicated that these T-DMRs represent the profiles of genes specifically expressed in liver, including those responsible for the liver phenotype and for transcription factors that regulate liver-specific gene expression.

A combination of transcription factors (e.g., HNF1 and HNF4a) is reported to confer liver-specific gene expression (Schrem et al. 2002). HNF1 affects the gene expression of HNF4a, and vice versa (Ktistaki and Talianidis 1997), and these factors, along with other liver-enriched transcription factors such as NR1H3 (LXR), NR1I2 (PXR), RXR, and PPARγ, affect each other's expressions and functions (Geier et al. 2007). D-REAM indicated that T-DMR is involved in the expression of these transcription factors, and that many genes with HNF-1A motifs and T-DMRs are specifically expressed in liver and kidney (Figs. 3, 5F). D-REAM revealed coordinated DNA methylation in transcription factors and the target genes in somatic tissues including liver, kidney, spleen, and cerebrum, and that the combinations of them were distinct among tissues. These results indicated that the DNA methylation profile, comprising T-DMRs of transcription factors and their target genes, is responsible for tissue-specific gene expression in somatic tissues and, consequently, for their function.

Correlations between T-DMRs and transcriptional regulation have been revealed in several CGI and non-CGI genes by the previous investigations on T-DMRs focusing primarily on short core promoter regions (Imamura et al. 2001; Hattori et al. 2004a; Nishino et al. 2004). Genome-wide DNA methylation analysis of the human promoter regions (−700 to 200 bp from the TSS) suggested that the DNA methylation status of a limited number of genes correlated with transcriptional activity (Weber et al. 2005, 2007). A HELP assay, another genome-wide study, revealed plurality of T-DMRs within 1 kb in promoter regions, suggesting their involvement in tissue-specific expression (Khulan et al. 2006). In our study, D-REAM revealed that T-DMRs are localized at a few kilobases both upstream and downstream of the TSS and that the DNA methylation status of these T-DMRs correlates with the transcriptional activities of the neighboring genes. Preferential tissue-specific expression of non-CGI genes has been reported previously (Yamashita et al. 2005), but in the present study, hypomethylation and gene expression were also observed in CGI genes. Thus, the DNA methylation profile is expected to reflect the tissue-specific gene expression profile.

It has become feasible to identify the DNA methylation status of every CpG dinucleotide in plant (Cokus et al. 2008; Lister et al. 2008) and mammalian (Meissner et al. 2008) genomes; however, high redundancy in mammalian genome sequences narrows the window of analysis. All known analysis methods have bias windows, and different windows illustrate different results. In the mouse genome, the HpaII/MspI sites are concentrated in the region between −1 kb and +1 kb from the TSS. In the case of meDIP, the precipitation efficiency of methylated DNA depended on the density of CpG, whose distribution is biased (Keshet et al. 2006; Weber et al. 2007). We suggest that the different T-DMR profiles observed in this study were due to the relatively low bias of HpyCH4IV.

Microarray technology for the analysis of DNA methylation has advantages and disadvantages (Khulan et al. 2006). When compared with systems using isoschizomers, D-REAM is flexible



F

Gene expression in kidney	>1000		≤1000		number of genes
	number	p-value*	Liv > 2x Kid	p-value*	
cluster 1	12	7.4×10^{-1}	8	2.2×10^{-1}	59
cluster 2	18	8.3×10^{-1}	13	9.1×10^{-2}	93
cluster 3	54	1.2×10^{-6}	7	8.0×10^{-2}	154
cluster 4	54	6.9×10^{-2}	37	1.1×10^{-4}	228
all T-DMR	363		171		1982

*: χ square test

Figure 5. The tissue-specific DNA methylation profiles and gene expression in mouse tissues. (*A*) Expression levels of human gene orthologs identified by the ChIP-Chip experiment using anti-HNF1 (closed rectangles) and of genes with HNF-1A motifs (model IDs T01211 and T00368) classified by the MAPPER database (open circles). (*B*) The distribution of expression levels of the genes listed in *A*, in cerebrum and liver represented by box plots. Expression levels were box-plotted with logarithmic scale (base = 10). (*C*) Box plots of log ratios (base = 2) of liver and cerebrum gene expression indicate factors affecting liver-specific expression of genes with T-DMRs. *P*-values obtained from Wilcoxon's matched-pair signed rank test are indicated on the *top* of the plot. (*D*) Correlation of expression between two tissues. CGI genes containing T-DMRs are divided into two groups by the position of T-DMRs: T-DMRs within 0.5 to 2.5 kb downstream from the TSS (*left* panel) and those outside of this region (*right* panel). Gene expression levels and ratios of gene expression levels are expressed with logarithmic scale of base 10 and base 2, respectively. Numbers of liver-specific genes expressed, with expression levels <1000 in cerebral cortex and at a liver:cerebral cortex level ratio of >2, in the *left* and *right* panels are 60 out of 346 and 45 out of 685, respectively. (*E*) K-means clustering of regions corresponding to T-DMRs by Pearson's correlations of their MATscores. The ranges of the MATscores represented in the plot are shown at the *bottom* of the panels. The MATscores were obtained by MAT analysis of D-REAM data from liver, kidney, and spleen using cerebrum data as the control. (*F*) χ^2 test for distributions of genes, classified by the expression levels, first in kidney (>1000 or not) and those in the later set divided by their expression in liver (>2-fold of those in kidney or not), in each cluster. Statistically significant distributions are shadowed in pink.

because it enables selective amplification of DNA fragments digested with any restriction enzyme, e.g., HpyCH4IV and NotI, a methylation-sensitive enzyme with recognition sites without isoschizomers. One of the disadvantages in D-REAM is that intra-genomic comparison of MATscores between different genomic loci does not always represent differences in DNA methylation. However, a panel of profiles comprising several samples allowed us to conduct a broad intragenomic comparison due in part to the presence of numerous T-DMRs that could provide standard sample data.

A recent study revealed that the T-DMR distribution in randomly selected NotI sites was disproportionate in non-CGI loci, which were located both upstream of the TSS and in the intronic regions (Sakamoto et al. 2007). Present study demonstrates that the T-DMR distribution pattern appears to be independent of CpG density and GC content, except in regions around the TSS in CGI genes. In addition, numerous T-DMRs are localized at 3′ regions downstream from the TSS. T-DMRs were also identified in distal regions (up to 6 kb from the TSS). Furthermore, the distribution patterns of hypomethylated T-DMRs in cerebral and liver genes are similar (S. Yagi, K. Hirabayachi, T. Hirakawa, S. Sato, C. Maeda, J. Ohgane, S. Tanaka, and K. Shiota, unpubl.). These findings suggest that the function of T-DMRs might differ with respect to their positions in the genes.

In our analyses, especially those involving CGI genes, hypomethylation of T-DMRs in the 3′-regions downstream from the TSS was correlated with high gene expression levels in liver. The first exon and the first intron are hot spots of antisense RNA transcription, and antisense RNA TSS is preferentially observed in relatively long CGIs covering exon 1 and extending into intron 1 (Finocchiaro et al. 2007). Both sense and antisense RNA play a role in gene expression by affecting DNA methylation status (Sleutels et al. 2002; Pickford and Cogoni 2003; Imamura et al. 2004). ChIP-chip experiments on regions adjacent to the TSS have indicated that H3K4 methylation occurs in a 2-kb region downstream from the TSS and that the peaks in the region are located 1 kb downstream (Barski et al. 2007). In addition, histone modification has been reported to affect DNA methylation in a locus-specific manner (Ikegami et al. 2007). These data suggest that T-DMRs located