

# Clinical Epigenetics

## Analysis of DNA methylation landscape reveals the roles of DNA methylation in the regulation of drug metabolizing enzymes

--Manuscript Draft--

<b>Manuscript Number:</b>	CLEP-D-15-00093R1
<b>Full Title:</b>	Analysis of DNA methylation landscape reveals the roles of DNA methylation in the regulation of drug metabolizing enzymes
<b>Article Type:</b>	Research
<b>Abstract:</b>	<p><b>Background:</b> Drug metabolizing enzymes (DMEs) exhibit dramatic inter- and intra-individual variability in expression and activity. However, the mechanisms determining this variability have not been fully elucidated. The aim of this study was to evaluate the biological significance of DNA methylation in the regulation of DME genes by genome-wide integrative analysis.</p> <p><b>Results:</b> DNA methylation and mRNA expression profiles of human tissues and hepatoma cells were examined by microarrays. The data were combined with GEO datasets of liver tissues, and integrative analysis was performed on selected DME genes. Detailed DNA methylation statuses at individual CpG sites were evaluated by DNA methylation mapping. From analysis of 20 liver tissues, highly variable DNA methylation was observed in 37 DME genes, 7 of which showed significant inverse correlations between DNA methylation and mRNA expression. In hepatoma cells, treatment with a demethylating agent resulted in upregulation of 5 DME genes, which could be explained by DNA methylation status. Interestingly, some DMEs were suggested to act as tumor-suppressor or housekeeper based on their unique DNA methylation features. Moreover, tissue-specific and age-dependent expression of UDP-glucuronosyl transferase 1A splicing variants was associated with DNA methylation status of individual first exons.</p> <p><b>Conclusions:</b> Some DME genes were regulated by DNA methylation, potentially resulting in inter- and intra-individual differences in drug metabolism. Analysis of DNA methylation landscape facilitated elucidation of the role of DNA methylation in the regulation of DME genes, such as mediator of inter-individual variability, guide for correct alternative splicing, and potential tumor suppressor or housekeeper.</p>
<b>Response to Reviewers:</b>	<p>Re: CLEP-D-15-00093 Analysis of DNA methylation landscape reveals the roles of DNA methylation in the regulation of drug metabolizing enzymes</p> <p>Dear Dr. Trygve Tollefsbol, Ph.D., D.O. :</p> <p>We are resubmitting this manuscript to Clinical Epigenetics after carefully considering the suggestions made by two reviewers. We appreciate the critical comments and useful suggestions that helped us to improve our manuscript. As indicated in the responses that follow, we have taken all these comments and suggestions into account in the revised version of our manuscript.</p> <p>Reviewer #1</p> <p>Major points</p> <p>1. As pointed by the reviewer, Figure 1 shows the result of cluster analysis on selected CpG sites. We initially performed cluster analysis of DME and control genes for all CpG sites examined by HumanMethylation450 platform. After that, in order to roughly compare the methylation profiles among tissue samples, we selected representative profiles of individual genes by excluding profiles showing the similar pattern each other. We finally carried out methylation mapping and confirmed detailed methylation status for all CpG sites. We have added a supplementary explanation about this in the Figure Legend section of the revised manuscript (page 28, lines 8-9).</p> <p>2. We feel that the reviewer's comment concerns the reliability of <math>\beta</math> value for judging inter-individual differences in DNA methylation. However, many other studies indeed used this value for the statistical analysis of methylation status. In addition, we have</p>

not observed marked differences between  $\beta$  values from HumanMethylation450 and the results of direct sequencing or digestion by restriction enzymes (COBRA assay). In the present study, we would evaluate any detectable inter-individual differences for the purpose of initial screening of candidate DME genes. Although it is difficult to distinguish the maximum and minimum values from outliers, any outliers can reflect significant differences among individuals. For example, CYP2D6 poor metabolizing genotypes are observed in Japanese individuals with minor allele frequency of <1%. Such variants, even if the frequency is quite low, should be considered as clinically significant polymorphisms, but not outliers. Therefore, any data should not be excluded from our study, because such exclusion might miss important signal of the inter-individual differences. As mentioned in the Discussion section of the original manuscript, the criteria using  $\beta$ R value may not be the best for evaluate inter-individual differences at present. We think that a larger scale study would provide more suitable criteria and show the stability of our conclusion.

3. It has recently reported that typical housekeeping genes with higher gene expression exhibited a markedly consistent pattern of low levels of DNA methylation among individuals (Edgar et al. Epingenetics Chromatin 2014). Therefore, in this study, we used several housekeeping genes for negative controls of variation. We have cited this paper in the Results section (page 9, line 2) and the Reference section (Reference 19) in the revised manuscript. In addition, as suggested by the reviewer, we have added the result of 2 housekeeping genes (B2M and TBP) to the revised Figure 3 to strengthen the conclusion. After adding this result (final 8 control genes), the cut-off value (maximum  $\beta$ R >0.296) has not been changed. To clarify the 8 genes examined as negative controls, the name of genes were colored in red in the revised Figure 3.

#### Minor points

1. All control genes examined in this study were listed in the Methods section. However, we should describe these genes in the Results section to make clear to readers. Therefore, we added these gene names to the revised manuscript (page 9, lines 5-6).
2. As the reviewer's comments, the phrase "DNA methylation acted as a driver, but not passenger" seems to be exaggerated and has been omitted from the Discussion section of the revised manuscript.
3. As the reviewer's comments, there were considerable number of DME genes that showed highly variable DNA methylation (maximum  $\beta$ R>0.296) without inverse correlation to gene expression. We think that such DNA methylation is unrelated to gene expression and other mechanisms may be involved in the regulation of these genes. We have not found biological significances of such DNA methylation in our study based on methylation landscape. As described in the original manuscript, we suggest that DME genes are classified into "at least" 3 groups. We think that this is a severe criterion, especially for the identification of housekeeping-type DME genes, and that large numbers of DME genes did not belong to any of the 3 groups. We have added the above explanation in the revised Discussion section (page 15, lines 12-16).
4. As pointed by the reviewer, we recognized the problem of multiple tests. Statistical analysis for fraction of CpG sites can avoid increasing a type I error caused by repeated analyses of individual CpG sites. However, as described in the original manuscript, we wanted to detect significant DNA methylation at the level of single CpG site. For example, CYP1A2 was transcriptionally regulated by DNA methylation at exclusively confined CpG site (Miyajima A et al., cited as Reference 28). Analysis for fraction of CpG sites would miss such significant signs due to averaging effects. In the present study, therefore, we repeatedly evaluated correlations between gene expression and DNA methylation at the level of single CpG sites, but not fraction of CpG sites. The data showing significant correlation were originally listed in the Figure 3 and Additional file 3: Figure S3).
5. The cut-off value of 0.5 was established by us in the present study. This criterion was determined by the methylation status of typical tumor-suppressor genes, whose transcriptional inactivation can be explained by DNA methylation of the promoter CpG sites in tumor cells. As shown in Additional file 6: Figure S6, BMP4 and IGFBP3 genes

were exclusively up-regulated by DAC treatment in JHH1 and HepG2 cells, respectively, in which the levels of methylation were more than 0.5 for the most part of 5'-regulatory regions. Although we examined only 3 hepatoma cells, it is widely accepted that typical tumor-suppressor genes can be functionally suppressed by loss of only one allele (loss of heterozygosity, LOH) (known as "gene dose effect"). In addition, many genetic variants of DMEs, even if they are heterozygous state, can affect drug metabolizing activity. We think that such LOH and heterozygous polymorphism are functionally equal to hypermethylation at the level of  $\beta$  value of 0.5, and that this supports our criterion using cut-off value of 0.5.

6. DME genes with highly variable DNA methylation status among normal livers (maximum  $\beta_R > 0.296$ ) were classified into 2 groups, whether DNA methylation was correlated with gene expression or not (driver or passenger methylation). Moreover, such correlations were not always retained in cancerous livers. Therefore, as pointed by the reviewer, normal liver and hepatoma cells provided distinct list of DME genes. We think that such correlations are determined by kind of DME genes and cell characters. In the present study, we suggested 3 kinds of DME gene types (HVM, TSG and HKG types) based on their DNA methylation landscapes in 2 types of cells (normal liver and hepatoma cells).

7. We feel that these data do not allow us to directly draw conclusions about the classification of DME genes. Therefore, we have excluded the following sentences from the original manuscript.

"large inter-individual differences in the DNA methylation (maximum  $\beta_R > 0.296$ ) were more frequently observed in DME genes without CpG islands (median, 0.244) than in those with CpG islands (median, 0.201;  $p = 0.0267$ ). Therefore," in the Results section.

"In addition, Mann-Whitney U-tests were used to compare variable DNA methylation (maximum  $\beta_R$ ) between 2 different DME gene groups" in the Methods section.

8. We agree with the reviewer that statistical analyses should be required to prove tissue-specific methylation of the UGT1A isoforms. However, we had limited numbers of tissue samples, and could not perform such analyses. In the revised manuscript, in order to solve this problem, we have compared DNA methylation status in individual CpG sites between 18 liver samples (GEO dataset) and one normal small intestine (NSI) or one normal fetal liver (NLF). The results were added as Additional file 10: Table S1. We further focused on the CpG sites most proximal to each transcription start site (excluding more distal CpG sites of TSS1500). In hepatic type UGT1A genes, as a result, we found that  $\beta$  values of the NSI and NLF were statistically higher than medians of 18 individual  $\beta$  values of normal adult livers. This was mentioned on the revised manuscript (page 11, lines 13-20, page 21, lines 8-10, page 33, lines 11-20).

9. As pointed by the reviewer, we cannot rule out the possibility of batch effect. In the study of methylation mapping and correlation analysis, however, we used minimum number of datasets obtained from the same laboratory (Ahrens M et al. Cell Metab 2013, cited as Reference 21). Therefore, we think that such effect was reduced to the lowest level in our study. The discussion has been rewritten to include a consideration of the batch effect (page 13, lines, 2-9) and Reference 22 has been added to the revised manuscript.

10. We expect that further functional analysis with a larger numbers of samples will accumulate evidence of significant roles of DNA methylation in the DME regulation. The results of our study imply that variable drug metabolism and drug response can be caused by not only genetic polymorphisms but also epigenetic mechanisms. Unfortunately, analysis of DNA methylation requires liver tissue specimens from individual patients, and therefore, does not immediately apply to clinical practice. For example, epigenome related to DME gene regulation can be rewritten or modified by "epidrugs" such as DNA methyltransferase inhibitors or histone deacetylase inhibitors. We expect that target-specific epidrug, if developed, will be used in combination with conventional drug therapy to provide optimum treatment for individual patients.

Reviewer #2

We feel that this comment is the same as indicated by the Reviewer#1 (Minor point

10). Our study demonstrates that certain variable drug metabolism can be explained by epigenetic mechanisms. Unfortunately, analysis of DNA methylation requires liver tissue specimens from individual patients, and therefore, does not immediately apply to clinical practice. For example, epigenome related to DME gene regulation can be rewritten or modified by “epidrugs” such as DNA methyltransferase inhibitors or histone deacetylase inhibitors. We expect that target-specific epidrug, if developed, will be used in combination with conventional drug therapy to provide optimum treatment for individual patients.

These changes have addressed all of the critiques of the reviewers. We hope you now find that this manuscript is suitable for publication.

Sincerely yours,

Wataru Habano, Ph.D.

**Title:**

Analysis of DNA methylation landscape reveals the roles of DNA methylation in the regulation of drug metabolizing enzymes

**Authors:**

Wataru Habano<sup>1§</sup>, Kohei Kawamura<sup>1</sup>, Natsuki Iizuka<sup>1</sup>, Jun Terashima<sup>1</sup>, Tamotsu Sugai<sup>2</sup>, and Shogo Ozawa<sup>1</sup>

<sup>1</sup>Department of Pharmacodynamics and Molecular Genetics, School of Pharmacy, Iwate Medical University

<sup>2</sup>Department of Diagnostic Pathology, School of Medicine, Iwate Medical University, Iwate Medical University

§Corresponding author:

Wataru Habano, PhD

Department of Pharmacodynamics and Molecular Genetics, School of Pharmacy, Iwate Medical University

2-1-1 Nishitokuta, Yahaba-Cho, Shiwa-Gun 028-3694, Japan

E-mail: whabano@iwate-med.ac.jp

Telephone: +81-19-651-5111 ext. 5241

Fax: +81-19-698-1879

Other authors:

Kohei Kawamura: aquosphone.12151@gmail.com

Natsuki Iizuka: p008\_natsuki@yahoo.co.jp

Jun Terashima: jterashi@iwate-med.ac.jp

Tamotsu Sugai: tsugai@iwate-med.ac.jp

Shogo Ozawa: sozawa@iwate-med.ac.jp

## Abstract

**Background:** Drug metabolizing enzymes (DMEs) exhibit dramatic inter- and intra-individual variability in expression and activity. However, the mechanisms determining this variability have not been fully elucidated. The aim of this study was to evaluate the biological significance of DNA methylation in the regulation of DME genes by genome-wide integrative analysis.

**Results:** DNA methylation and mRNA expression profiles of human tissues and hepatoma cells were examined by microarrays. The data were combined with GEO datasets of liver tissues, and integrative analysis was performed on selected DME genes. Detailed DNA methylation statuses at individual CpG sites were evaluated by DNA methylation mapping. From analysis of 20 liver tissues, highly variable DNA methylation was observed in 37 DME genes, 7 of which showed significant inverse correlations between DNA methylation and mRNA expression. In hepatoma cells, treatment with a demethylating agent resulted in upregulation of 5 DME genes, which could be explained by DNA methylation status. Interestingly, some DMEs were suggested to act as tumor-suppressor or housekeeper based on their unique DNA methylation features. Moreover, tissue-specific and age-dependent expression of UDP-glucuronosyl transferase 1A splicing variants was associated with DNA methylation status of individual first exons.

**Conclusions:** Some DME genes were regulated by DNA methylation, potentially resulting in inter- and intra-individual differences in drug metabolism. Analysis of DNA methylation landscape facilitated elucidation of the role of DNA methylation in the regulation of DME genes, such as mediator of inter-individual variability, guide for

correct alternative splicing, and potential tumor suppressor or housekeeper.

**Keywords:** DNA methylation mapping, differences in drug metabolism, integrative analysis, alternative splicing



## Background

Inter-individual differences in responses to drug therapy vary widely. Such differences are due in part to variable pharmacokinetics, which can be explained by diversity in genes affecting drug absorption, distribution, metabolism, and excretion (ADME) [1]. For example, genetic polymorphisms in cytochrome P450 (CYP) genes, such as *CYP2C9*, *CYP2C19*, and *CYP2D6*, affect drug metabolizing activity and lead to different drug responses [2-4]. Inter-individual differences are also observed in *CYP3A4* expression and activity [5,6], but are rarely associated with any detectable polymorphisms in the *CYP3A4*. Additionally, UDP-glucuronosyl transferase 1A1 (*UGT1A1*) mutant genotypes \*28 and \*6 are related to poor glucuronidation activity and have been shown to act as useful indicators of adverse reactions to cancer chemotherapy with irinotecan (CPT-11) [7,8]. However, some individuals who do not have these mutant genotypes also suffer from adverse reactions to irinotecan [9].

Intra-individual differences are also involved in the regulation of ADME-related genes. For example, the expression and metabolizing activity of UGT1A isoforms vary among tissues and during normal development [10,11]. The *UGT1A* locus encodes 9 functional isoforms through an exon sharing mechanism in which the transcripts of individual first exon cassettes are spliced to exons 2–5, leading to the expression of individual UGT1A isoforms [12]. Although the tissue-specific and age-dependent expression of the UGT1A isoforms are important determinants of drug efficacy and adverse reactions, the regulatory mechanisms involved in UGT1A expression cannot be explained by genetic polymorphisms. Thus, the mechanisms underlying inter- and intra-individual differences in responses to drug therapy have not been fully elucidated.

In order to examine these mechanisms, we investigated the involvement of epigenetics,

1  
2  
3 the mechanism of heritable changes in gene regulation without DNA sequence  
4 alteration. We previously found that epigenetic mechanisms, such as DNA methylation,  
5 are involved in the regulation of drug metabolizing enzymes (DMEs) in colon cancer  
6 cells [13,14]. The *CYP1B1* and *CYP3A4* genes were shown to be transcriptionally  
7 upregulated by treatment with a demethylating agent, and such upregulation could be  
8 explained by hypermethylation at CpG sites located in the 5'-promoters of the *CYP1B1*  
9 and pregnane X receptor (*PXR*) genes. To date, more than 50 ADME-related genes have  
10 been reported as targets for epigenetic regulation [15,16]. Although these  
11 ADME-related genes were found to be aberrantly regulated by epigenetic mechanisms  
12 in tumor cells, most ADME-related genes have been discovered unexpectedly by  
13 multiple independent studies during the search for tumor-suppressor genes. Therefore,  
14 we still have a limited understanding of the biological significance of epigenetic  
15 mechanisms in the regulation of ADME-related genes.

16  
17 In the present study, we examined global DNA methylation and mRNA expression  
18 profiles of human tissues and hepatoma cell lines using microarray platforms. These  
19 two omics datasets were combined with similar corresponding datasets derived from the  
20 Gene Expression Omnibus (GEO) database at the National Center for Biotechnology  
21 Information (NCBI). In order to evaluate the significance of DNA methylation in the  
22 regulation of DME genes, we examined which DME genes were regulated by DNA  
23 methylation in normal liver tissues and hepatoma cells, whether the tissue-specific and  
24 age-dependent expression of UGT1A isoforms could be regulated by DNA methylation,  
25 and whether DNA methylation profiles could be used to elucidate the specific roles of  
26 DME genes. To this end, we performed DNA methylation mapping to determine the  
27 methylation levels and variations at individual CpG sites in relation to the structure of

individual DME genes. Recent studies demonstrated that global and detailed analysis of dynamic DNA methylation profiles facilitated the finding of the informative fraction of CpG sites [17,18]. Therefore, we would be able to identify the role of DNA methylation of DME genes by analysis of DNA methylation landscapes.

## Results

### DNA methylation profiles of DME genes in human tissues

We examined DNA methylation profiles of 2 adult liver tissues (NLA and NL2), fetal liver tissue (NLF), adult small intestinal tissue (NSI), and 3 hepatoma cell lines (HepG2, HuH7, and JHH1) by HumanMethylation450 Bead Chip. These data were combined with GEO-registered datasets of 18 healthy adult liver tissues. The levels of DNA methylation for more than 480,000 CpG sites were determined as the  $\beta$  values ( $0 < \beta < 1$ ). Cluster analysis demonstrated that methylation levels of DME genes, including 55 *CYP* genes and 62 phase II DME genes, differed markedly among distinct tissue groups. The representative profiles of 6 CYPs (*CYP1A2*, *CYP1B1*, *CYP2C9*, *CYP2C19*, *CYP2D6*, and *CYP3A4*), and 2 controls (*ACTB* and *BMP4*) were shown in Figure 1. Although the 20 adult liver tissues were derived from 2 different race populations (2 Chinese and 18 German), they exhibited similar DNA methylation profiles. However, some DME genes such as *CYP1A2*, *CYP2C9*, *CYP2C19*, *CYP2D6*, and *CYP3A4* showed relatively variable methylation statuses, compared to *CYP1B1* and 2 control genes. To examine the landscape of DNA methylation in more detail, we performed the DNA methylation mapping of individual DME genes (Figure 2 and Additional file 1: Figure S1). For each CpG site, the range of variation was defined as a  $\beta_R$  value, which was calculated from the difference between the highest and lowest  $\beta$  values among the 20 livers. Next, for all CpG sites located in the 5' regulatory region (defined as TSS1500, TSS200, or 5'UTR in the array platform), the maximum  $\beta_R$  value was used for estimation of the degree of inter-individual differences in DNA methylation. The distribution of maximum  $\beta_R$  values (degrees of variation) of the genes is shown as a histogram in Figure 3. We expected that the control genes would show the least variation because the expression of

these genes should be tightly regulated by DNA methylation on their promoter CpG islands, as reported by Edgar et al. [19]. Therefore, these genes were used as negative controls of inter-individual variation. We finally identified 37 (32%) DME genes with significantly variable DNA methylation statuses, for which the maximum  $\beta_R$  values were more than those of all 8 control genes (*ACTB*, *B2M*, *GAPDH*, *TBP*, *BMP4*, *IGFBP3*, *MLH1* and *MGMT*) (maximum  $\beta_R > 0.296$ ).

### DME genes regulated by DNA methylation in adult livers

Transcript datasets were also registered for 10 of the 18 liver tissues examined for DNA methylation analysis. We analyzed the mRNA expression profiles of DME genes in these 10 adult livers and detected the highest expression of *CYP2E1* and *CYP3A4* genes, consistent with a previous report [20] (Additional file 2: Figure S2A). In addition, the majority of the DME genes exhibited variable levels of mRNA expression compared to the housekeeping genes (with coefficients of variation (CVs) of more than 17.1%; Additional file 2: Figure 2B). We found inverse correlations between mRNA expression and DNA methylation levels in 7 DME genes (*CYP1A2*, *CYP2C19*, *CYP2D6*, *GSTA4*, *GSTM5*, *GSTT1*, and *SULT1A1*). In particular, the *CYP2C19*, *GSTA4*, and *GSTM5* genes had CpG sites that simultaneously showed inverse correlations and highly variable methylation statuses ( $\beta_R > 0.296$ ; Figure 4 and Additional file 3: Figure S3).

### DME genes regulated by DNA methylation in 3 hepatoma cell lines

Cluster analysis revealed that the 3 hepatoma cell lines showed clearly distinct methylation profiles compared to normal liver tissues (Figure 1). Aberrant DME gene methylation detected in hepatoma cells is summarized in Additional file 4: Figure S4.

We defined a hypermethylated or hypomethylated CpG site as having a  $\beta$  value of more than 0.5 for comparisons between hepatoma cell lines and normal liver (NL2). This cut-off value was determined by validation analysis, as described below. Using this criterion, 36 DME genes had hypermethylated CpG sites within the 5' regulatory region in at least one hepatoma cell line. To determine whether the hypermethylation observed in hepatoma cells was associated with downregulation of the DME gene, DNA methylation was reversed by treatment with DAC, and mRNA expressions were examined using a SurePrint G3 Human Gene Expression 8 × 60K v2 microarray. As a result, 44 (38%) DME genes showed upregulation by more than 2-fold following DAC treatment (Additional file 5: Figure S5). Therefore, these DME genes were downregulated by specific DNA methylation events. Among these DME genes, the downregulation of *CYP1B1*, *CYP8B1*, *GSTM2*, *GSTP1*, *UGT2B15*, and *UGT3A2* in hepatoma cells could be explained by DNA methylation status. The results of array-based DNA methylation and mRNA expression analyses were examined for validity using combined bisulfite restriction analysis (COBRA) and quantitative real-time PCR assays, respectively. The criterion for judging hypermethylation was determined by the results of the tumor-suppressor genes used as positive controls (Additional file 6: Figure S6). By using a cut-off value of  $\beta > 0.5$ , we identified 5 DME genes (*CYP1B1*, *CYP8B1*, *GSTM2*, *GSTP1*, and *UGT3A2*) that were regulated by DNA methylation in hepatoma cells. Representative results for the *CYP1B1* gene are shown in Figure 5.

### **DNA methylation and alternative splicing of UGT1A isoforms**

Next, the transcript levels of *UGT1A* isoforms were examined in different tissues using

quantitative real-time PCR. The tissue-specific expression profiles shown in Additional file 7: Figure S7 were similar to the results of a previous study [10-12]. The *UGT1A* genes were then classified into 2 major groups according to the tissues in which they were dominantly expressed: hepatic type (i.e., *UGT1A1*, *UGT1A3*, *UGT1A4*, *UGT1A6*, and *UGT1A9*) and intestinal type (i.e., *UGT1A5*, *UGT1A7*, *UGT1A8*, and *UGT1A10*). Although the level of *UGT1A8* expression in adult livers was somewhat higher than that in small intestines, *UGT1A8*, unlike the other hepatic-type *UGT1A*s, showed relatively high expression in the small intestine as well. Therefore, we classified *UGT1A8* as an intestinal-type gene. DNA methylation mapping on the *UGT1A* locus revealed that DNA methylation status was largely variable among different tissues (Additional file 8: Figure S8). We focused on the first exon of each isoform and found that the hepatic-type genes *UGT1A1*, *UGT1A4*, *UGT1A6*, and *UGT1A9* tended to be methylated at higher levels in the small intestine than in the liver (Figure 6). In order to perform statistical analysis, moreover, DNA methylation status in individual CpG sites were also compared between 18 adult livers and one normal small intestine (NSI) or one normal fetal liver (NLF). We further focused on the CpG sites most proximal to each transcription start site (excluding more distal CpG sites of TSS1500). In hepatic type *UGT1A* genes, as a result, we found that  $\beta$  values of the NSI and NLF were statistically higher than medians of 18 individual  $\beta$  values of normal adult livers (Additional file 9, Table S1). In contrast, the methylation levels of the intestinal-type genes *UGT1A10* and *UGT1A8* were relatively higher in the liver than in the small intestine. This suggested that DNA methylation status around the first exon determined the splicing isoforms of the *UGT1A* gene and led to tissue-specific expression. In addition, higher levels of DNA methylation were observed in fetal livers

1  
2  
3 than in adult livers. This tendency was found in hepatic-type genes rather than  
4  
5 intestinal-type genes, implying that the low levels of expression observed in the fetal  
6  
7 liver may result from downregulation by DNA methylation.  
8  
9

## 10 11 12 **Classification of DME genes based on their DNA methylation landscape** 13

14  
15 Interestingly, DME genes had unique features of the DNA methylation landscape  
16  
17 (Figure 1), which could be classified into at least 3 groups, as summarized in Figure 7.  
18  
19 The first group showed highly variable methylation among normal livers (maximum  $\beta_R$   
20  
21  $> 0.296$ ) and inverse correlations with mRNA expression. DME genes in this first group  
22  
23 may be candidates for explaining inter-individual variations in drug metabolizing  
24  
25 activity; we classified genes in this group as the highly variable methylation (HVM)  
26  
27 type. Genes in the second group, similar to the *CYP1B1* gene, showed stable  
28  
29 methylation statuses among normal livers but were hypermethylated in tumor cells.  
30  
31 These DNA methylation features were similar to those of tumor-suppressor genes (i.e.,  
32  
33 the TSG type). The last group retained low levels of methylation in both normal and  
34  
35 tumor livers, suggesting these genes may act as housekeeping genes. Typical  
36  
37 housekeeping genes, such as *ACTB* and *GAPDH*, had similar features of methylation  
38  
39 mapping and expression profiles in our study. For example, the level of DNA  
40  
41 methylation was highly stable and low in the 5' regulatory region (hypomethylated  
42  
43 5'UTR), but considerably high within the gene body (hypermethylated exons). In  
44  
45 addition, the mRNA expression of the housekeeping genes was stable among normal  
46  
47 livers, with CVs of less than 17.1%. Although it was difficult to fulfill all these criteria,  
48  
49 we found 2 DME genes (*CYP2R1* and *CYP46A1*) that had characteristics similar to  
50  
51 those of housekeeping genes (i.e., HKG type).  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65



## Discussion

We used GEO datasets previously registered elsewhere [21] and combine these datasets with our original data. We demonstrated that genome-wide integrative analysis using GEO datasets was a powerful tool for identification of novel roles of DNA methylation. We cannot rule out the possibility of batch effects, which are the technical artifacts such as laboratory conditions, experiment time, reagent lots and/or laboratory personnel differences [22]. However, the datasets obtained from the same laboratory were used for the study of methylation mapping and correlation analysis. Therefore, we think that such effect was reduced to the lowest level in our study.

The DNA methylation statuses of DME genes varied among normal liver tissues. We found 7 DME genes (i.e., *CYP1A2*, *CYP2C19*, *CYP2D6*, *GSTA4*, *GSTM5*, *GSTT1*, and *SULT1A1*) that showed significant inverse correlations between DNA methylation and mRNA expression levels. Some other DME genes also tended to show such correlations, but statistical significance was not reached. These results suggested that a small but significant fraction of DME genes were transcriptionally regulated by DNA methylation, resulting in different mRNA expression levels among individuals. Further investigations with larger sample sizes may allow for identification of additional targets for DNA methylation. Recent studies demonstrated that considerable numbers of hepatic genes are also regulated by 5-hydroxymethylcytosine (5hmC) [23]. Therefore, the levels of 5hmC should be evaluated in our next studies.

Interestingly, *CYP1A2*, *CYP2C19*, *CYP2D6*, and *SULT1A1* tend to show inter-individual variations in activity, and this has become a major problem in clinical practice. Our data supported that these variations may result from the variable DNA methylation statuses of these DME genes. For example, DNA methylation status could

1  
2  
3 explain the discrepancy in the relatively variable expression of *CYP2C19* among  
4  
5 Caucasian populations in which there is a low frequency of mutant genotypes [4].  
6  
7 Although considerable inter-individual differences were observed in *CYP3A4* expression  
8  
9 (CV: 22.9%; Figure 2B), we did not detect a significant relationship between DNA  
10  
11 methylation and mRNA expression. As reported by Kacevska et al. [24], DNA  
12  
13 methylation at more 5' distal regions (position -1,547 or -10,762) may be associated with  
14  
15 *CYP3A4* regulation. Alternatively, *PXR* gene methylation, but not *CYP3A4* gene  
16  
17 methylation, may contribute to *CYP3A4* downregulation in colon cancer cells, as  
18  
19 previously described by our laboratory [14].  
20  
21  
22  
23

24 We also found 5 DME genes (*CYP1B1*, *CYP8B1*, *GSTM2*, *GSTP1*, and *UGT3A2*) that  
25  
26 were likely to be downregulated by DNA methylation. Further analyses, such as  
27  
28 chromatin immunoprecipitation, will be required to establish functional evidence for  
29  
30 this relationship. Interestingly, the methylation statuses of these 5 DME genes were  
31  
32 highly stable in normal livers, suggesting that DNA methylation may play different  
33  
34 roles in the regulation of DME genes between normal and cancerous livers. As expected,  
35  
36 DNA methylation in the 5'-regulatory region of housekeeping and tumor-suppressor  
37  
38 genes was stable and restricted to low levels. In contrast, the methylation levels of these  
39  
40 genes were relatively high within the exons. These results were supported by a previous  
41  
42 study showing that DNA methylation plays critical roles as markers of exon-intron  
43  
44 boundaries and transcriptional silencing [25,26].  
45  
46  
47  
48  
49

50 In this study, we demonstrated that DNA methylation could function as a guide for  
51  
52 correct alternative splicing of the *UGT1A* gene in a tissue-specific manner. Despite the  
53  
54 limited cases in our study, our findings were supported by recent genome-wide analyses  
55  
56 showing that many genes are regulated by alternative splicing via DNA methylation  
57  
58  
59  
60  
61  
62  
63  
64  
65

[18,25,26]. Oda et al. also reported that tissue-specific *UGT1A10* expression is regulated by DNA methylation [27]. We found that the relatively low expression levels of DME genes in the fetal liver could be associated with the higher methylation levels observed in these genes. Thus, DNA methylation mapping may provide novel insights into the tissue-specific and age-dependent splicing switch of DME genes regulated by DNA methylation.

From these results, we propose that the roles of DME genes in the liver depend on the DNA methylation landscape (Figure 7). Variable methylation of the DME gene observed in normal livers may be one of the mechanisms mediating inter-individual variation in drug metabolizing activity and drug efficacy if the methylation status is associated with mRNA expression. Therefore, the DNA methylation statuses of the 7 HVM-type genes may be indicators for prediction of drug efficacy and safety. Although there were considerable number of DME genes that showed highly variable DNA methylation (maximum  $\beta_R > 0.296$ ) without inverse correlation to gene expression, such DNA methylation was probably unrelated to gene expression and other mechanisms might be involved in the regulation of these DME genes. The cut-off value as maximum  $\beta_R > 0.296$  may not be the best for categorizing HVM-type, because maximum  $\beta_R$  can depend on the number of CpG sites examined for each gene. However, some DME genes such as *CYP1A2* are regulated by DNA methylation at exclusively confined CpG sites [28]. In order to detect any informative signals without averaging effect, therefore, we evaluated  $\beta_R$  value with single (maximum) CpG site rather than with the fraction of CpG sites. We also found 5 TSG-type genes that were thought to act as tumor-suppressor genes because they exhibited DNA methylation features similar to those observed in the *BMP4* and *IGFBP3* genes. Therefore, these TSG-type genes may

1  
2  
3 have novel functions related to gatekeeping or genome stability in tumor cells. On the  
4  
5 other hand, some HKG-type genes were suggested to function as housekeeping genes.  
6  
7 Based on this extremely stable DNA methylation status, we predicted that the  
8  
9 expression of these HKG-type genes should be tightly regulated, which may explain  
10  
11 why DME genes with stable DNA methylation had conserved CpG islands. Interestingly,  
12  
13 DMEs with a small range of variation in DNA methylation were known to metabolize  
14  
15 endogenous substrates rather than xenobiotics. In contrast, DMEs with highly variable  
16  
17 DNA methylation (including *CYP1A2*, *CYP2C19*, and *CYP2D6*) catalyze the metabolism  
18  
19 of many xenobiotics, including drugs used in the clinical setting. These types of DMEs  
20  
21 may flexibly modify their DNA methylation and expression profiles to metabolize and  
22  
23 excrete various xenobiotics. We hypothesize that DNA methylation may play different  
24  
25 roles in the regulation of DME genes depending on context. For example, HKG-type  
26  
27 genes are regulated by “rigid” DNA modifications to strictly retain the epigenome. On  
28  
29 the other hand, HVM-type genes are regulated by “plastic” DNA modifications to  
30  
31 flexibly rewrite the epigenome for adjustment to new environments. Thus, HVM-type  
32  
33 genes tend to show higher inter-individual variations in DNA methylation status.  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65

## Conclusions

We described the global DNA methylation landscape of DME genes in human tissues and demonstrated that a small but specific fraction of DME genes was regulated by DNA methylation. Variations in DNA methylation may result in inter-individual differences in the efficacies and toxicities of many drugs. DNA methylation of DME genes may represent landmarks of tissue-specific and age-dependent splicing switches and may be useful indicators for predicting the unknown functions, such as tumor-suppressor or housekeeper.

## Methods

### Cell culture and treatment

Human hepatoma HepG2, HuH7, and JHH1 cells were obtained from American Type Culture Collection (Manassas, VA, USA) and cultured in Dulbecco's minimal essential medium (Invitrogen, Carlsbad, CA, USA) containing 10% fetal bovine serum at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>. In order to reverse DNA methylation, the cells were treated with 0.5 or 5 µM 5-aza-2'-deoxycytidine (DAC; Sigma-Aldrich, St. Louis, MO, USA) for 72 h. After DAC treatment, the cells were treated with trichostatin A (TSA; Sigma-Aldrich) at 200 nM (HepG2 and HuH7 cells) or 20 nM (JHH1 cells) for 24 h.

### DNA and RNA samples

Genomic DNA was extracted from hepatoma cells using a standard proteinase K/sodium dodecyl sulfate and phenol/chloroform method. Total cellular RNA was isolated using an RNeasy Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Genomic DNA and total RNA of human adult liver tissues (NLA and NL2), fetal liver tissue (NLF), adult small intestinal tissue (NSI), and adult colonic tissue (NC) were obtained from commercially available products (Capital Biosciences, Gaithersburg, MD, USA and BioChain Institute, Newark, CA, USA). Because we had no information other than age, gender, and race (Chinese) for these samples, we did not submit our research proposal to an ethical committee.

### DNA methylation analysis

Genomic DNA from tissues and hepatoma cells was subjected to sodium bisulfite

modification using an EpiTect Plus DNA Bisulfite Kit (Qiagen) and examined for DNA methylation using an Infinium HumanMethylation450 Bead Chip (Illumina, San Diego, CA, USA), which interrogates over 480,000 CpG sites in the genome [29]. The level of DNA methylation for each CpG site was reported as the  $\beta$  value, which ranged from 0 (fully unmethylated) to 1 (fully methylated). Data have been deposited in the GEO repository with accession numbers GSE67477 and GSE67484. DNA methylation data of 55 *CYP* genes and 62 phase II DME genes, including 17 glutathione S-transferase (*GST*), 10 *N*-acetyltransferase (*NAT*), 13 sulfotransferase (*SULT*), and 22 *UGT* genes, were selected and used for subsequent analyses. For the control, 4 housekeeping genes (*ACTB*, *B2M*, *GAPDH*, *TBP*), 2 tumor-suppressor genes (*BMP4* and *IGFBP3*), and 2 DNA repair genes (*MLH1* and *MGMT*) were also examined. The results were visualized in heat maps using Cluster 3.0 software (<http://bonsai.hgc.jp/~mdehoon/software/cluster/software.htm#ctv>). The DNA methylation status of candidate DME genes was validated by COBRA method [30]. Briefly, after polymerase chain reaction (PCR) using a TaKaRa EpiTaq HS kit (Takara Bio Inc., Shiga, Japan), the resulting products were digested with an appropriate restriction enzyme, such as HpyCH4IV, TaqI, or BstUI (New England Biolabs, Ipswich, MA, USA). The digested products were electrophoresed on 2% agarose gels followed and visualized by ethidium bromide staining. Primers used are listed in Additional file 10: Table S2.

## mRNA expression analysis

First-strand cDNA was synthesized from total RNA samples using a Transcriptor First Strand cDNA Synthesis Kit (Roche Diagnostics, Mannheim, Germany) according to the

1  
2  
3 manufacturer's protocol. The mRNA expression profiles were examined using a  
4  
5 SurePrint G3 Human Gene Expression 8 × 60K v2 microarray (Agilent Technologies,  
6  
7 Palo Alto, CA, USA). The signal intensity for each probe was normalized by the 75th  
8  
9 percentile. The data have been deposited in the GEO repository with the accession  
10  
11 number GSE67318. The mRNA expression data for the *CYP* gene, phase II DME genes,  
12  
13 and control genes were selected and used for subsequent analyses. The level of each  
14  
15 transcript was validated by quantitative real-time PCR analysis using a FirstStart  
16  
17 Universal SYBR Green Master (ROX) kit (Roche Diagnostics). Transcript levels of  
18  
19 individual *UGT1A* isoforms and *ACTB* were evaluated by TaqMan Gene Expression  
20  
21 Assays (Life Technologies, Gaithersburg, MD, USA). All primer sets used were  
22  
23 described in Additional file 10: Table S2. Each real-time PCR analysis was performed in  
24  
25 triplicate.  
26  
27  
28  
29  
30  
31  
32  
33

## 34 **GEO datasets used for integrative analysis**

35  
36 DNA methylation datasets of liver tissues derived from 18 healthy German individuals  
37  
38 were found in GEO records (GSE48325) [21]. These data were obtained in another study  
39  
40 using the same HumanMethylation450 platform (GPL13534) and could be combined  
41  
42 with our data for 2 adult liver tissues (NLA and NL2). Moreover, transcript datasets  
43  
44 were also found for 10 of the 18 liver tissues in GEO records (GSE48452) and used for  
45  
46 correlation analysis between DNA methylation and mRNA expression.  
47  
48  
49  
50  
51  
52

## 53 **DNA methylation mapping**

54  
55 DNA methylation mapping was carried out for two purposes. First, in order to estimate  
56  
57 the variable DNA methylation statuses of the 20 adult liver tissues, the  $\beta$  value of each  
58  
59  
60  
61  
62  
63  
64  
65



CpG site was expressed as a box-and-whisker plot. Second, we aimed to identify informative CpG sites in which DNA methylation levels were different between tissues or cell lines. We constructed line graphs connecting each CpG site arranged in the 5' to 3' direction according to the relative location and distance of each CpG site.

## Statistical analysis

Correlations between DNA methylation levels of each CpG site were examined for all CpG sites in the 5' regulatory region by Spearman's rank correlation coefficient in 10 liver tissues. In the DNA methylation mapping of the *UGT1A* isoforms, Wilcoxon signed rank test was used to compare median of 18 individual  $\beta_R$  values of normal adult liver and one normal small intestine (NSI) or one normal fetal liver (NLF). These analyses were performed using GraphPad Prism 5 software (GraphPad Software, La Jolla, CA, USA), and differences or correlations with  $p$ -values of less than 0.05 were considered significant.

## Availability of supporting data

(1) GSE67477 (DNA methylation)

(2) GSE67484 (DNA methylation)

(3) GSE67318 (mRNA expression)

This series is linked to GSE67485.

Additional file 1: Figure S1

Additional file 2: Figure S2

Additional file 3: Figure S3

Additional file 4: Figure S4

Additional file 5: Figure S5

Additional file 6: Figure S6

Additional file 7: Figure S7

Additional file 8: Figure S8

Additional file 9: Table S1

Additional file 10: Table S2

## Abbreviations

ACTB, actin, beta; B2M,  $\beta$ 2 microglobulin; BMP4, bone morphogenetic protein 4;

COBRA, combined bisulfite restriction analysis; CYP, cytochrome P450; DAC,

5-aza-2'-deoxycytidine; DME, drug metabolizing enzyme; GAPDH,

glyceraldehyde-3-phosphate dehydrogenase; GEO, Gene Expression Omnibus; GST,

glutathione S-transferase; IGFBP3, insulin-like growth factor binding protein 3; MGMT,

O-6-methylguanine-DNA methyltransferase; MLH1, mutL homolog 1; NAT,

N-acetyltransferase; SULT, sulfotransferase; TBP, TATA box-binding protein; TSA,

trichostatin A; TSS, transcription start site; UGT, UDP-glucuronosyltransferase; UTR,  
untranslated region

### **Competing interests**

The authors declare that they have no competing of interest.

### **Authors' contributions**

WH designed the study, analyzed all data and drafted the manuscript. KK and NI performed all experiments and collected the data. JT, TS and SO contributed to manuscript preparation. All authors read and approved the final manuscript.

### **Acknowledgements**

This work was supported by a Grant-in-Aid for Scientific Research (C) (24590451) from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (2012–2014) and by Medical Innovation by Advanced Science and Technology (MIAST): a Grant-in-Aid for Strategic Medical Science Research Center for the Ministry of Education, Culture, Sports, Science, and Technology of Japan (2010–2014).

## References

1. Evans WE, Relling MV. Pharmacogenomics: translating functional genomics into rational therapeutics. *Science* 1999;**286**:487-91.
2. Van Booven D, Marsh S, McLeod H, Carrillo MW, Sangkuhl K, Klein TE, et al. Cytochrome P450 2C9-CYP2C9. *Pharmacogenet Genomics* 2010;**20**:277-81.
3. Scott SA, Sangkuhl K, Shuldiner AR, Hulotn J-S, Thorn CF, Altman RB, et al. PharmGKB summary: very important pharmacogene information for cytochrome P450, family 2, subfamily C, polypeptide 19. *Pharmacogenet Genomics* 2012;**22**:159-65.
4. Owen RP, Sangkuhl K, Klein TE, Altman RB. Cytochrome P450 2D6. *Pharmacogenet Genomics* 2009;**19**:559-62.
5. Shimada T, Yamazaki H, Mimura M, Inui Y, Guengerich P. Inter-individual variations in human liver cytochrome P-450 enzymes involved in the oxidation of drugs, carcinogens and toxic chemicals: studies with liver microsomes of 30 Japanese and 30 Caucasians. *J Pharmacol Exp Ther* 1994;**270**:414-23.
6. Lamba JK, Lin YS, Schuetz EG, Thummel KE. Genetic contribution to variable human CYP3A-mediated metabolism. *Adv Drug Deliv Rev* 2002;**54**:1271-94.
7. Minami H, Sai K, Saeki M, Saito Y, Ozawa S, Suzuki K, et al. Irinotecan pharmacokinetics/pharmacodynamics and UGT1A genetic polymorphisms in Japanese: roles of UGT1A1\*6 and \*28. *Pharmacogenet Genomics* 2007;**17**:497-504.
8. Sai K, Saeki M, Saito Y, Ozawa S, Katori N, Jinno H, et al. UGT1A1 haplotypes associated with reduced glucuronidation and increased serum bilirubin in irinotecan-administered Japanese patients with cancer. *Clin Pharmacol Ther* 2004;**75**:501-15.

9. Li M, Wang Z, Guo J, Liu J, Li C, Liu L, et al. Clinical significance of UGT1A1 gene polymorphisms on irinotecan-based regimens as the treatment in metastatic colorectal cancer. *Onco Targets Ther* 2014;**7**:1653-61.
10. Nakamura A, Nakajima M, Yamanaka H, Fujiwara R, Yokoi T. Expression of UGT1A and UGT2B mRNA in human normal tissues and various cell lines. *Drug Metab Dispos* 2008;**36**:1461-4.
11. Strassburg CP, Strassburg A, Kneip S, Barut A, Tukey RH, Rodeck B, et al. Developmental aspects of human hepatic drug glucuronidation in young children and adults. *Gut* 2002;**50**:259-65.
12. Tukey RH, Strassburg CP. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annu Rev Pharmacol Toxicol* 2000;**40**:581-616.
13. Habano W, Gamo T, Sugai T, Otsuka K, Wakabayashi G, Ozawa S. CYP1B1, but not CYP1A1, is downregulated by promoter methylation in colorectal cancers. *Int J Oncol* 2009;**34**:1085-91.
14. Habano W, Gamo T, Terashima J, Sugai T, Otsuka K, Wakabayashi G, et al. Involvement of promoter methylation in the regulation of Pregnane X receptor in colon cancer cells. *BMC Cancer* 2011;**11**:81.
15. Kim IW, Han N, Burckart GJ, Oh JM. Epigenetic changes in gene expression for drug-metabolizing enzymes and transporters. *Pharmacotherapy* 2014;**34**:140-50.
16. Ivanov M, Kacevska M, Ingelman-Sundberg M. Epigenomics and inter-individual differences in drug response. *Clin Pharmacol Ther* 2012;**92**:727-36.
17. Ziller MJ, Gu H, Müller F, Donaghey J, Tsai LT, Kohlbacher O, et al. Charting a dynamic DNA methylation landscape of the human genome. *Nature*. 2013;**500**:477-81.

- 1  
2  
3 18. Anastasiadou C, Malousi A, Maglaveras N, Kouidou S. Human epigenome data  
4 reveal increased CpG methylation in alternatively spliced sites and putative exonic  
5  
6 splicing enhancers. *DNA Cell Biol* 2011;**30**:267-75.  
7  
8  
9  
10 19. Edgar R, Tan PP, Portales-Casamar E, Pavlidis P. Meta-analysis of human  
11 methylomes reveals stably methylated sequences surrounding CpG islands  
12 associated with high gene expression. *EpiGenetics Chromatin* 2014;**7**:28.  
13  
14  
15  
16 20. Bièche I, Narjoz C, Asselah T, Vacher S, Marcellin P, Lidereau R, et al. Reverse  
17 transcriptase-PCR quantification of mRNA levels from cytochrome (CYP)1, CYP2  
18 and CYP3 families in 22 different human tissues. *Pharmacogenet Genomics*  
19 2007;**17**:731-42.  
20  
21  
22  
23  
24  
25  
26 21. Ahrens M, Ammerpohl O, von Schönfels W, Kolarova J, Bens S, Itzel T, et al.  
27 DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct  
28 disease-specific and remodeling signatures after bariatric surgery. *Cell Metab*  
29 2013;**18**:296-302.  
30  
31  
32  
33  
34  
35  
36 22. Sun Z, Chai HS, Wu Y, White WM, Donkena KV, Klein CJ, et al. Batch effect  
37 correction for genome-wide methylation data with Illumina Infinium platform.  
38 *BMC Med Genomics* 2011;**4**:84.  
39  
40  
41  
42  
43 23. Ivanov M, Kals M, Kacevska M, Barragan I, Kasuga K, Rane A, et al. Ontogeny,  
44 distribution and potential roles of 5-hydroxymethylcytosine in human liver function.  
45 *Genome Biol.* 2013;**14**:R83.  
46  
47  
48  
49  
50 24. Kacevska M, Ivanov M, Wyss A, Kasela S, Milani L, Rane A, et al. DNA  
51 methylation dynamics in the hepatic CYP3A4 gene promoter. *Biochimie*  
52 2012;**94**:2338-44.  
53  
54  
55  
56  
57 25. Choi JK. Contrasting chromatin organization of CpG islands and exons in the  
58  
59  
60  
61  
62  
63  
64  
65

human genome. *Genome Biol* 2010;**11**:R70.

26. Gelfman S, Cohen N, Yearim A, Ast G. DNA-methylation effect on cotranscriptional splicing is dependent on GC architecture of the exon-intron structure. *Genome Res* 2013;**23**:789-99.

27. Oda S, Fukami T, Yokoi T, Nakajima M. Epigenetic regulation of the tissue-specific expression of human UDP-glucuronosyltransferase (UGT) 1A10. *Biochem Pharmacol* 2014;**87**:660-7.

28. Miyajima A, Furihata T, Chiba K. Functional analysis of GC Box and its CpG methylation in the regulation of CYP1A2 gene expression. *Drug Metab Pharmacokinet.* 2009;**24**:269-76.

29. Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. High density DNA methylation array with single CpG site resolution. *Genomics* 2011;**98**:288-95.

30. Xiong Z, Laird PW. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Res* 1997;**25**:2532-4.

## Figure Legends

**Figure 1.** Heat maps of the DNA methylation profiles of the representative 6 *CYP* and 2 control genes by cluster analysis

The top of the map indicates samples from 20 normal adult livers (L1 to L81, NLA and NL2), 1 fetal liver (NLF), 1 adult small intestine (NSI), and 3 hepatoma cell lines. The right of the map indicates the target ID and gene name (*CYP1A2*, *CYP1B1*, *CYP2C9*, *CYP2C19*, *CYP2D6*, *CYP3A4*, *ACTB*, and *BMP4*) examined for individual CpG sites. The representative methylation profiles of individual genes were selected by excluding profiles showing the similar patterns each other. Higher DNA methylation levels are shown in red, while lower DNA methylation levels are shown in black.

**Figure 2.** Representative results of DNA methylation mapping for 20 liver tissues

In each panel, the horizontal axis indicates the positions of CpG sites arranged in the 5' to 3' direction. The 5' regulatory region (TSS1500, TSS200, or 5'UTR) is underlined. The vertical axis indicates the  $\beta$  value, with the variations expressed as box-and-whisker plots. The maximum  $\beta_R$  value within the 5' regulatory region is indicated by a circle for each gene. (A) *CYP1B1*, (B) *CYP3A4*, (C)  $\beta$ -actin (*ACTB*), (D) bone morphogenetic protein 4 (*BMP4*).

**Figure 3.** Distribution of maximum  $\beta_R$  values of the DME and control genes

The maximum  $\beta_R$  values of the 8 control genes colored in red were relatively low ( $\leq 0.296$ ). We defined highly variable DNA methylation status as a  $\beta_R$  value of more than 0.296. The  $\beta_R$  values of representative DME genes are also shown in the histogram. *ACTB*, actin, beta; *B2M*,  $\beta 2$  microglobulin; *BMP4*, bone morphogenetic protein 4;



*GAPDH*, glyceraldehyde-3-phosphate dehydrogenase; *IGFBP3*, insulin-like growth factor binding protein 3; *MGMT*, O-6-methylguanine-DNA methyltransferase; *MLH1*, mutL homolog 1; *TBP*, TATA box-binding protein.

**Figure 4.** Representative results of correlation analysis

A significant inverse correlation between DNA methylation level ( $\beta$  value) and mRNA expression level was detected for the *CYP2C19* (A) and *GSTA4* (C) genes ( $p < 0.05$ , Spearman's rank correlation test). The level of mRNA expression, shown as the vertical axis, was normalized to the expression level of *ACTB* (set as 100%). CpG sites with significant correlations are indicated by asterisks in methylation mapping for the *CYP2C19* (B) and *GSTA4* (D) genes. Additionally, CpG sites with  $\beta_R$  values of more than 0.296 are indicated by circles.

**Figure 5.** Validation of mRNA expression and DNA methylation of the *CYP1B1* gene in hepatoma cells

(A) The level of DNA methylation was mapped on the *CYP1B1* gene locus in the 3 hepatoma cell lines. The gene structure and location of the CpG sites are shown in the lower panel. (B) The levels of *CYP1B1* mRNA were detected by quantitative real-time PCR in hepatoma cells treated with TSA and/or DAC. The vertical axis indicates the mRNA level of treated cells relative to paired control cells without treatment. Each column represents the mean  $\pm$  SD ( $n = 3$ ). (C) The *CYP1B1* gene methylation status of the 3 control cell lines and adult liver tissues was examined by COBRA assay. DNA fragments cleaved by TaqI digestion represent methylated DNA (M), while noncleaved fragments represent unmethylated DNA (U).

**Figure 6.** Representative results of DNA methylation mapping of *UGT1A* isoforms

The levels of the DNA methylation of hepatic type (*UGT1A1*, *UGT1A4*, *UGT1A6*, and *UGT1A9*) and intestinal type (*UGT1A10* and *UGT1A8*) genes were examined in adult liver tissues (NLA and NL2), fetal liver tissue (NLF), and adult small intestinal tissue (NSI). In each panel, the open or closed circles located on the horizontal axis indicate the positions of CpG sites arranged in the 5' to 3' direction according to the relative distance of each CpG site. The closed circles indicate CpG sites in which tissue-specific expression could be explained by the DNA methylation status.

**Figure 7.** Classification of DME genes based on the corresponding DNA methylation landscape

## Additional files

### Additional file 1: Figure S1. Results of methylation mapping of 20 liver tissues

In each panel, the horizontal axis indicates the positions of CpG sites arranged in the 5' to 3' direction. The 5' regulatory region (TSS1500, TSS200, or 5'UTR) is underlined. The vertical axis indicates the  $\beta$  values of individual CpG sites with the variations expressed as box-and-whisker plots. The maximum  $\beta_R$  value within the 5' regulatory region is indicated by a circle for each gene.

### Additional file 2: Figure S2. Distribution of the mRNA expression levels and CV values of DME and control genes

(A) The horizontal axis of the histogram represents the logarithm of the mean value of the transcript (normalized values by the 75th percentile) for 10 liver specimens. (B) The horizontal axis of the histogram represents the CV (%) of the transcript.

### Additional file 3: Figure S3. Representative results of correlation analysis

Significant inverse correlations between DNA methylation levels and mRNA expression levels were detected for the other 5 DME genes (*CYP1A2*, *CYP2D6*, *GSTM5*, *GSTT1*, and *SULT1A1*;  $p < 0.05$ , Spearman's rank correlation test). The mRNA expression level, shown as the vertical axis, was normalized to the level of *ACTB* mRNA (set as 100%). CpG sites with significant correlation are indicated by asterisks in the methylation map for each gene. CpG sites showing  $\beta_R$  values of more than 0.296 are indicated by circles.

### Additional file 4: Figure S4. DNA methylation profiles of DME genes in 3 hepatoma cell lines

DME genes with hypermethylated or hypomethylated CpG sites are listed and shown by closed and shaded boxes, respectively. The number of CpG sites showing hypermethylation or hypomethylation is also indicated in each box.

**Additional file 5: Figure S5.** Changes in the mRNA expression levels of DME genes in 3 hepatoma cell lines after DAC treatment

DME genes showing upregulation in at least one hepatoma cell line ( $> 2$  fold change) following DAC treatment are listed. The fold change is indicated in each box, and genes with fold changes of more than 2 are indicated by closed boxes. The 6 DME genes exhibiting correlations between mRNA expression and DNA methylation status are shown in highlighted boxes.

**Additional file 6: Figure S6.** DNA methylation mapping of *BMP4* and *IGFBP3* tumor-suppressor genes in the 3 hepatoma cell lines

In each panel, open circles located on the horizontal axis indicate the positions of CpG sites arranged in the 5' to 3' direction according to the relative distance of each CpG site. Exons are indicated by green boxes, and translated regions are shown in dark green. The fold change in mRNA expression is shown on the right side of the panel. Following DAC treatment, the *BMP4* and *IGFBP3* genes were exclusively upregulated in JHH1 (7.0 fold) and HepG2 (8.8 fold) cells, respectively, and hypermethylation was also detected in methylation mapping for these two cell lines, with  $\beta$  values of more than 0.5.

**Additional file 7: Figure S7.** Expression levels of *UGT1A* transcripts in different tissues

The vertical axis indicates the mRNA expression levels of *UGT1A* isoforms normalized to the level of *UGT1A1* expression in adult liver (NL2; set as 100%). Each column represents the mean  $\pm$  SD (n = 3).

**Additional file 8: Figure S8.** DNA methylation mapping on the *UGT1A* locus

The vertical axis of the upper panel shows the level of DNA methylation on the *UGT1A* locus for different tissues. Open circles on the horizontal axis indicate the positions of CpG sites corresponding to the *UGT1A* gene structure shown in the lower panel. The *UGT1A* isoforms are transcribed by an exon sharing mechanism in which the transcripts of the individual first exon cassette are spliced to exons 2–5. The CpG sites boxed in red correspond to the first exon of each isoform.

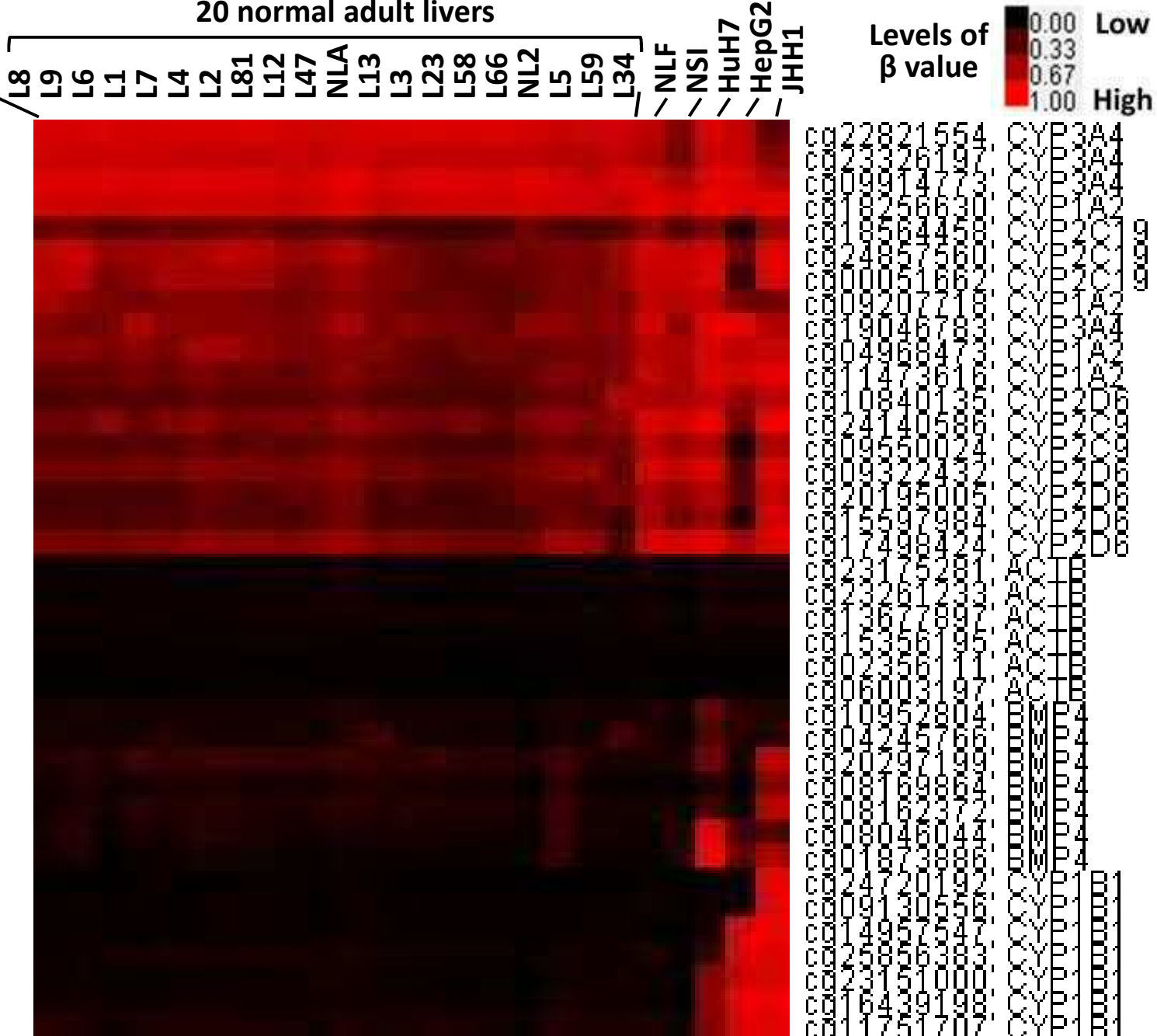
**Additional file 9: Table S1.** Levels of DNA methylation around the first exons of the *UGT1A* isoforms

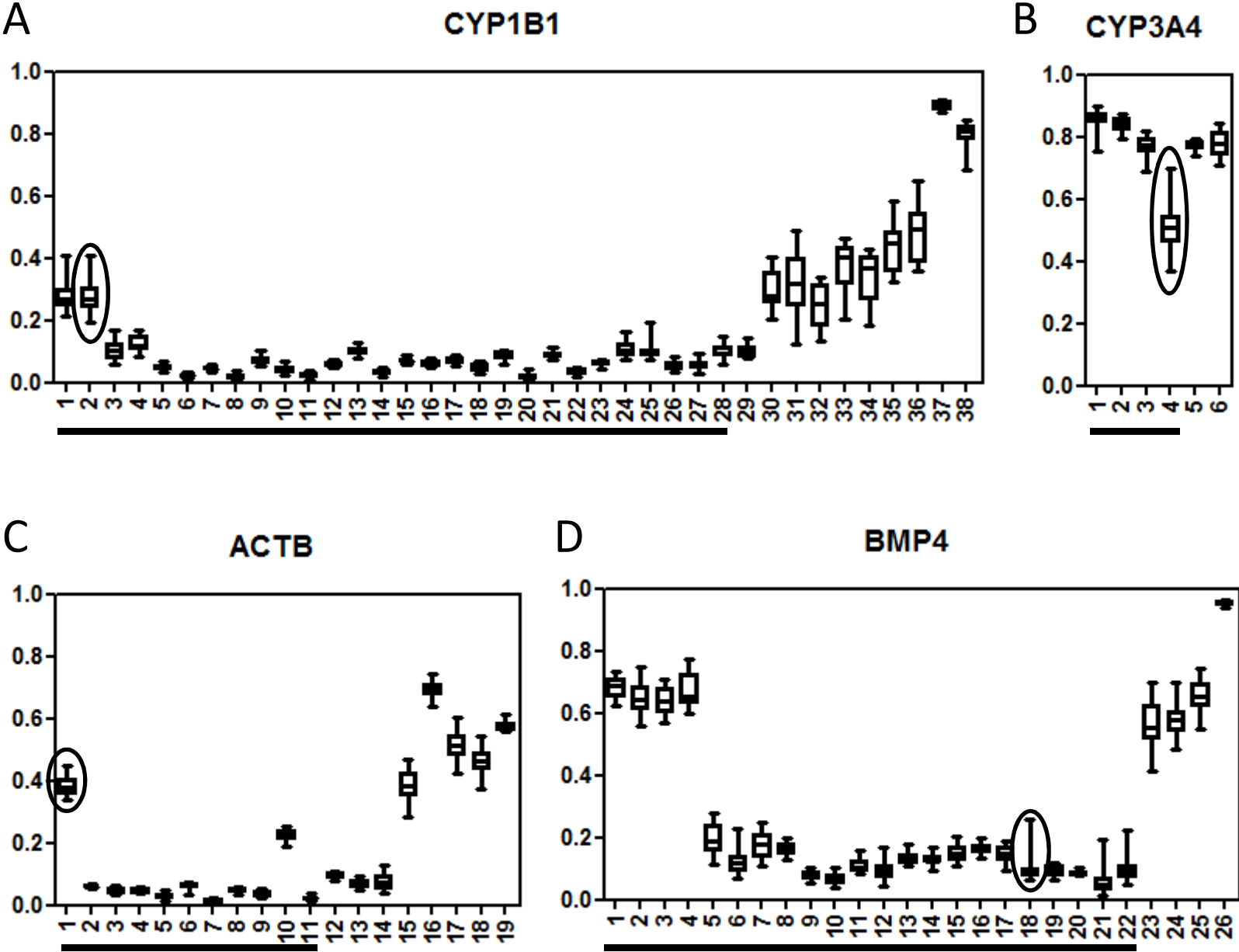
The 9 *UGT1A* isoforms were classified into 2 groups (hepatic and intestinal types). The levels of DNA methylation at individual CpG sites of the first exons were expressed as  $\beta$  values, and separately compared between 18 adult livers and one normal small intestine (NSI) or one normal fetal liver (NLF). The  $\beta$  values colored in red were statistically higher than the median of  $\beta$  values of normal adult livers. The  $\beta$  values colored in blue were statistically lower than the median of  $\beta$  values of normal adult livers. The CpG sites proximal to TSS200 region (except TSS1500 region) were indicated by bold frame.

**Additional file 10: Table S2.** Primers used for mRNA expression and DNA

methylation analyses

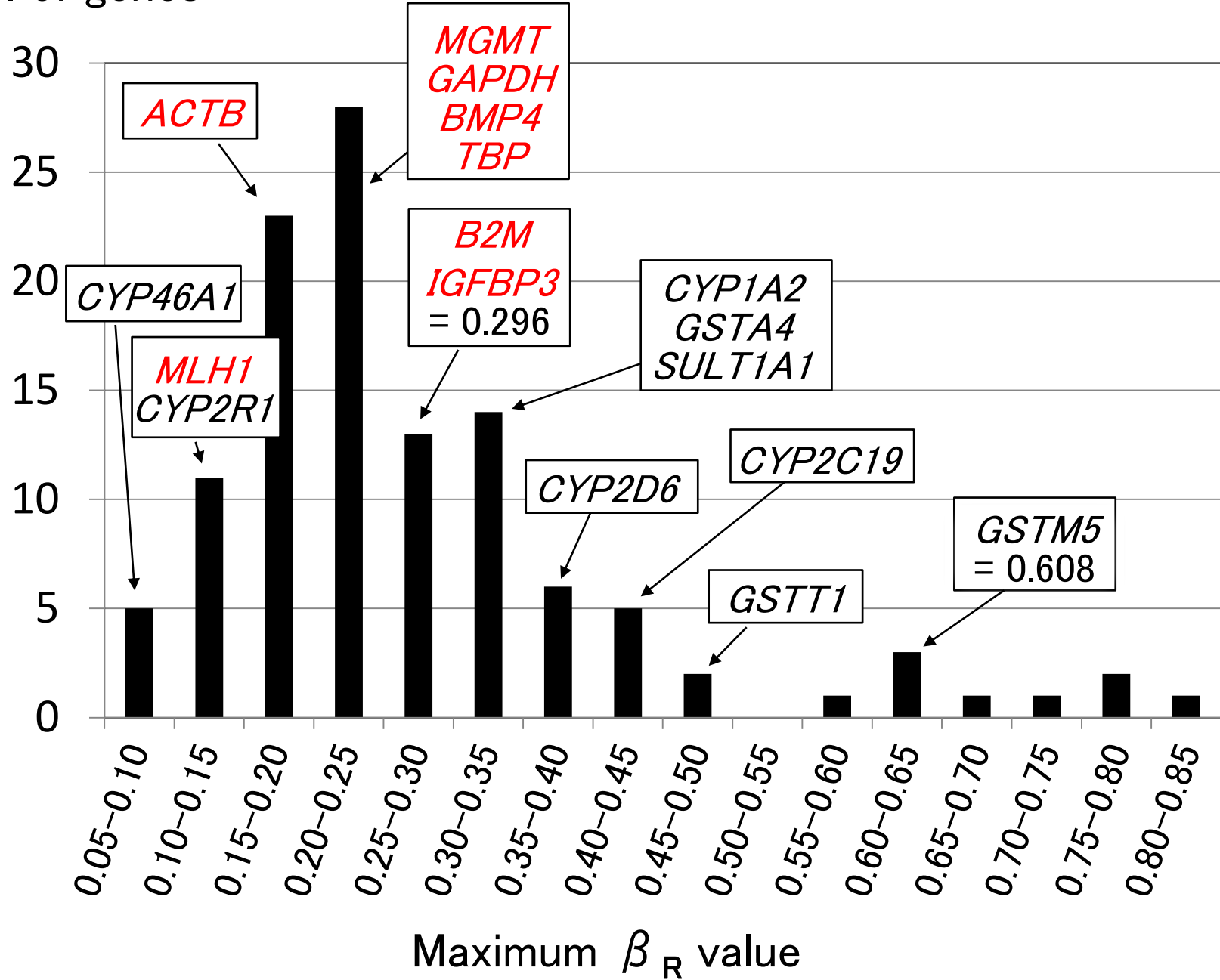
Figure  
[Click here to download Figure: Figure 1\\_Habano W.pdf](#)

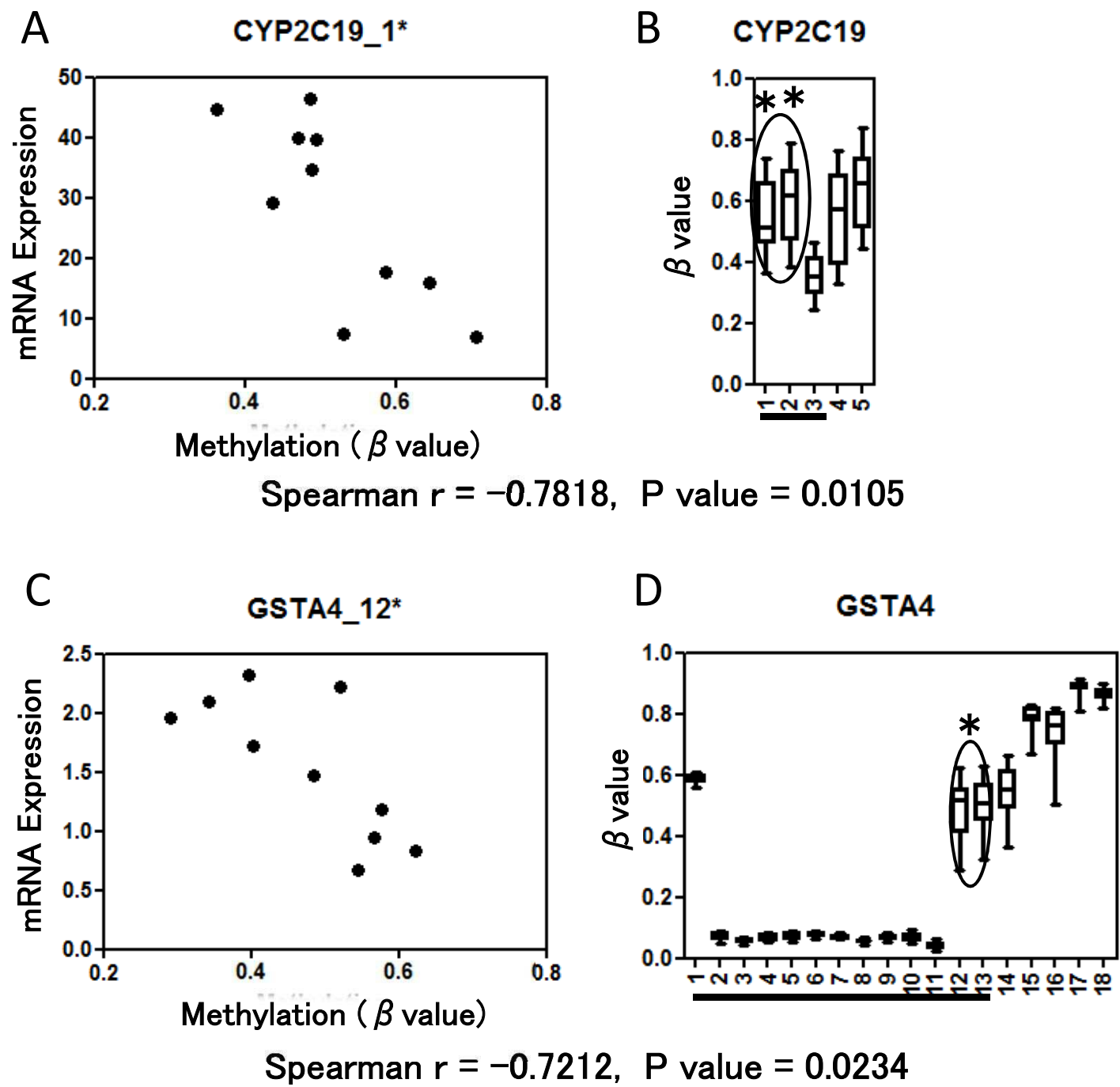






No. of genes





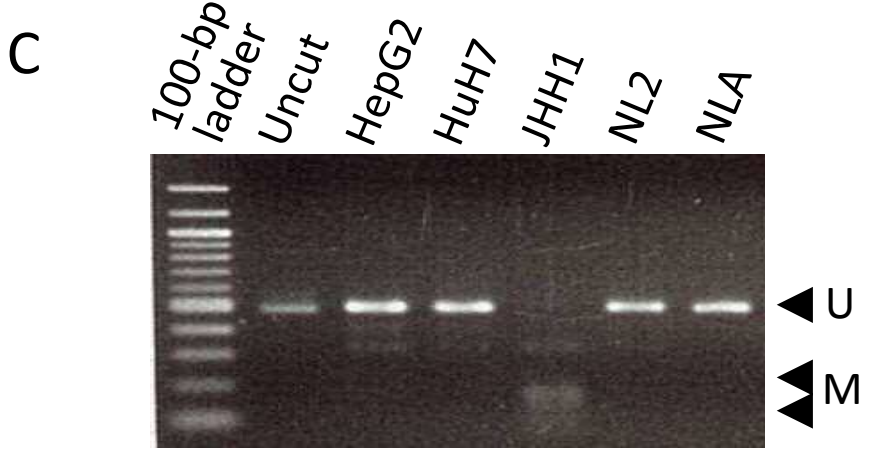
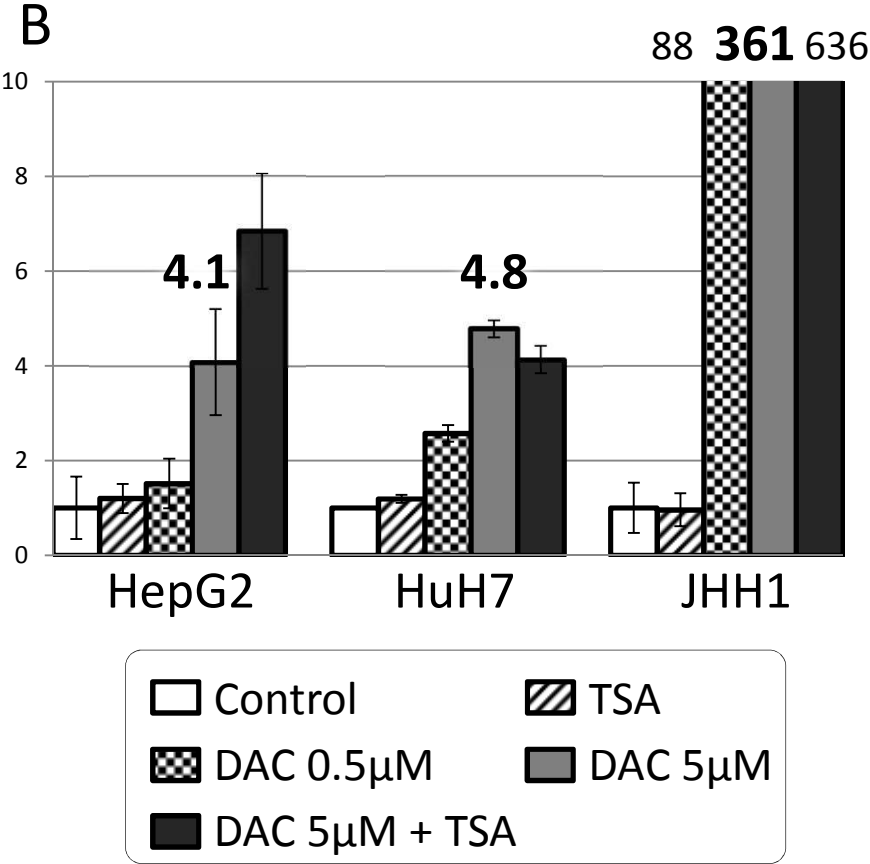
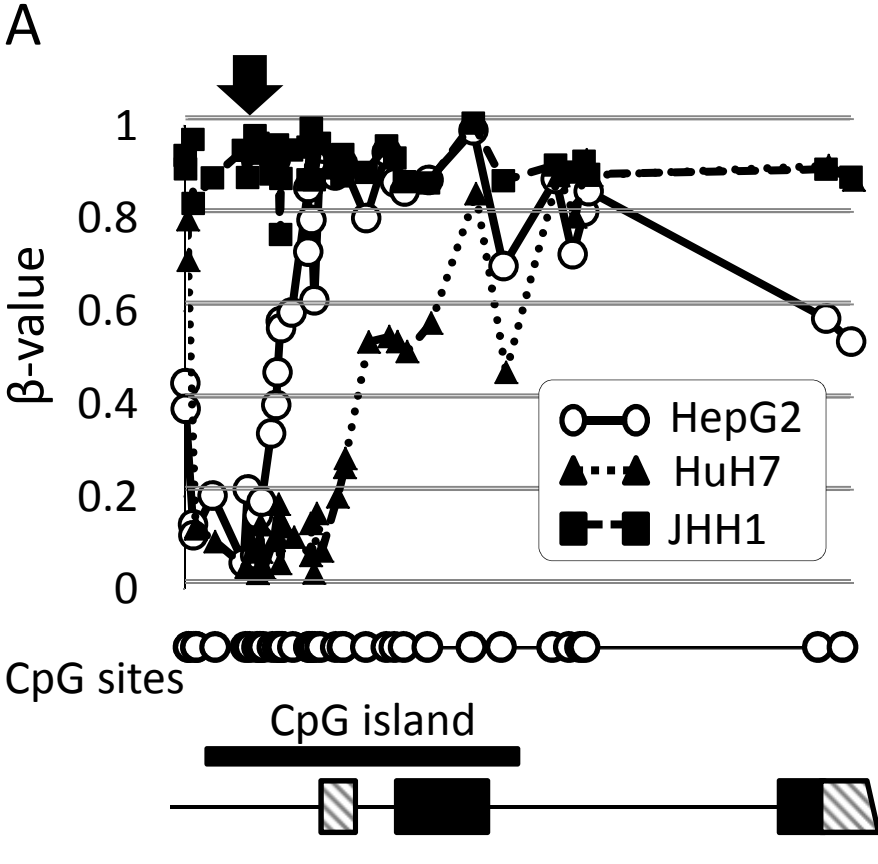
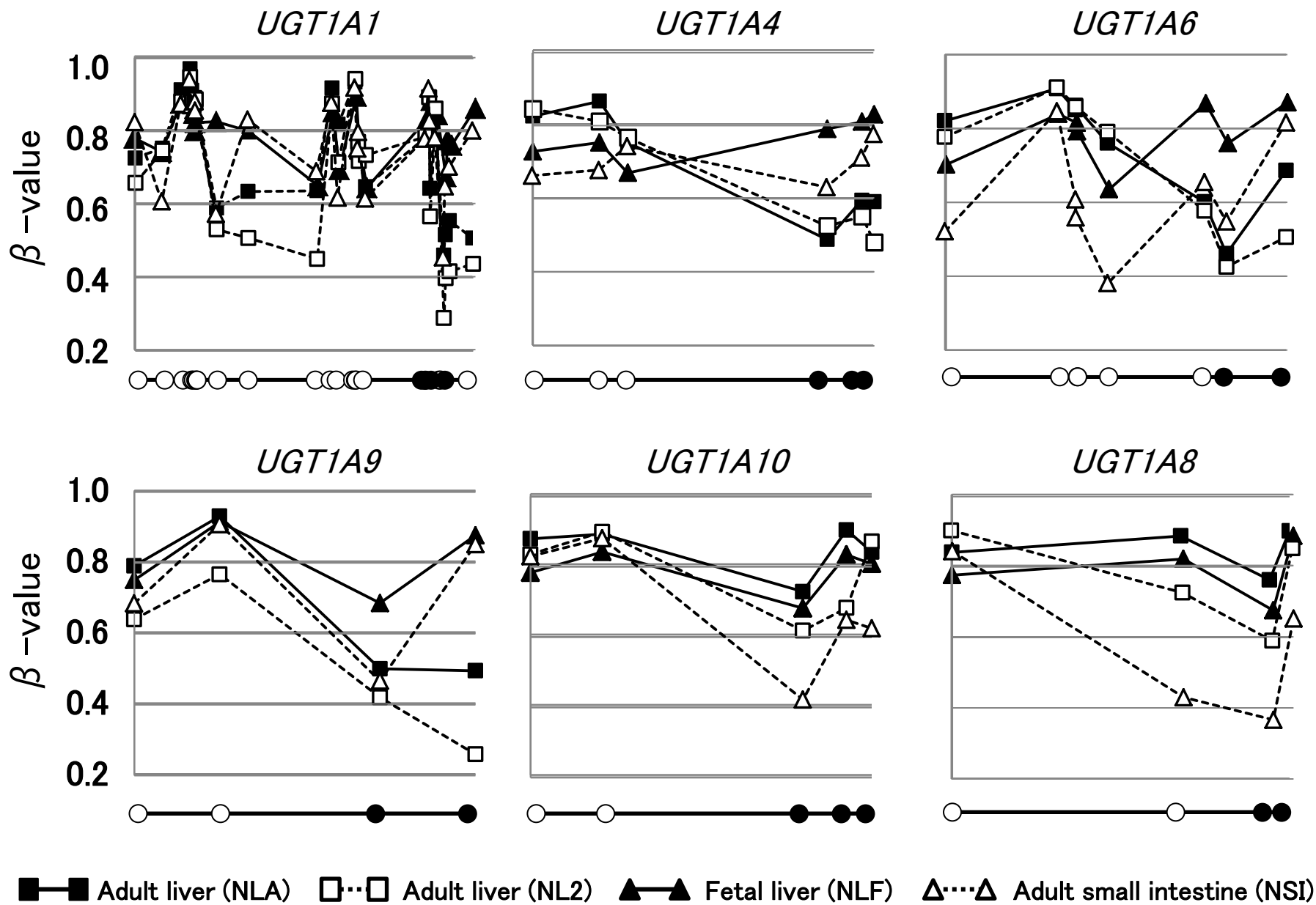
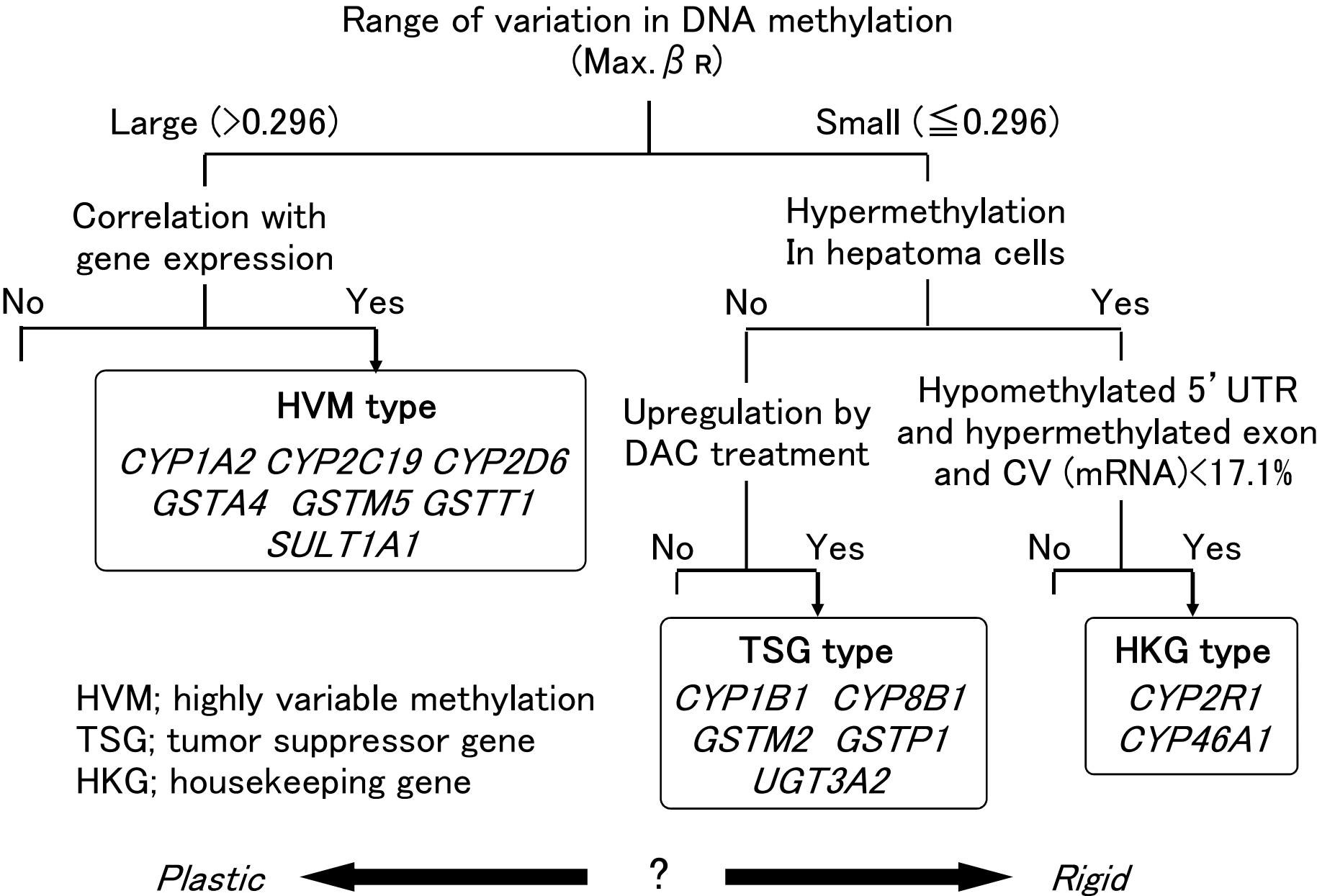


Figure  
[Click here to download Figure: Figure 6 \(revised\)\\_Habano W.pdf](#)





## Supplementary Material

[Click here to download Supplementary Material: Additional file 1\\_Figure S1\\_Habano W.pdf](#)

## Supplementary Material

[Click here to download Supplementary Material: Additional file 2\\_Figure S2\\_Habano W.pdf](#)

## Supplementary Material

[Click here to download Supplementary Material: Additional file 3\\_Figure S3\\_Habano W.pdf](#)



## Supplementary Material

[Click here to download Supplementary Material: Additional file 4\\_Figure S4\\_Habano W.pdf](#)

## Supplementary Material

[Click here to download Supplementary Material: Additional file 5\\_Figure S5\\_Habano W.pdf](#)

## Supplementary Material

[Click here to download Supplementary Material: Additional file 6\\_Figure S6\\_Habano W.pdf](#)

## Supplementary Material

[Click here to download Supplementary Material: Additional file 7\\_Figure S7\\_Habano W.pdf](#)

## Supplementary Material

[Click here to download Supplementary Material: Additional file 8\\_Figure S8\\_Habano W.pdf](#)

Supplementary Material

[Click here to download Supplementary Material: Additional file 9\\_Table S1\\_Habano W.xlsx](#)

Supplementary Material

[Click here to download Supplementary Material: Additional file 10\\_Table S2\\_Habano W.pdf](#)