

Original Article

## Development of humanized steroid and xenobiotic receptor mouse by homologous knock-in of the human steroid and xenobiotic receptor ligand binding domain sequence

Katsuhide Igarashi<sup>1</sup>, Satoshi Kitajima<sup>1</sup>, Ken-ichi Aisaki<sup>1</sup>, Kentaro Tanemura<sup>1</sup>,  
Yuhji Taquahashi<sup>1</sup>, Noriko Moriyama<sup>1</sup>, Eriko Ikeno<sup>1</sup>, Nae Matsuda<sup>1</sup>, Yumiko Saga<sup>2,3</sup>,  
Bruce Blumberg<sup>4</sup> and Jun Kanno<sup>1</sup>

<sup>1</sup>Division of Cellular and Molecular Toxicology, Biological Safety Research Center,

National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo, 158-8501, Japan

<sup>2</sup>Division of Mammalian Development, National Institute of Genetics, Yata 1111, Mishima 411-8540, Japan

<sup>3</sup>The Graduate University for Advanced Studies (Sokendai), Yata 1111, Mishima 411-8540, Japan

<sup>4</sup>Department of Developmental and Cell Biology, 2011 Biological Sciences 3, University of California,  
Irvine, CA 92697-2300, USA

(Received December 7, 2011; Accepted January 12, 2012)

**ABSTRACT** — The human steroid and xenobiotic receptor (SXR), (also known as pregnane X receptor PXR, and NR1I2) is a low affinity sensor that responds to a variety of endobiotic, nutritional and xenobiotic ligands. SXR activates transcription of Cytochrome P450, family 3, subfamily A (CYP3A) and other important metabolic enzymes to up-regulate catabolic pathways mediating xenobiotic elimination. One key feature that demarcates SXR from other nuclear receptors is that the human and rodent orthologues exhibit different ligand preference for a subset of toxicologically important chemicals. This difference leads to a profound problem for rodent studies to predict toxicity in humans. The objective of this study is to generate a new humanized mouse line, which responds systemically to human-specific ligands in order to better predict systemic toxicity in humans. For this purpose, the ligand binding domain (LBD) of the human SXR was homologously knocked-in to the murine gene replacing the endogenous LBD. The LBD-humanized chimeric gene was expressed in all ten organs examined, including liver, small intestine, stomach, kidney and lung in a pattern similar to the endogenous gene expressed in the wild-type (WT) mouse. Quantitative reverse transcription-polymerase chain reaction (RT-PCR) analysis showed that the human-selective ligand, rifampicin induced Cyp3a11 and Carboxylesterase 6 (Ces6) mRNA expression in liver and intestine, whereas the murine-selective ligand, pregnenolone-16-carbonitrile did not. This new humanized mouse line should provide a useful tool for assessing whole body toxicity, whether acute, chronic or developmental, induced by human selective ligands themselves and subsequently generated metabolites that can trigger further toxic responses mediated secondarily by other receptors distributed body-wide.

**Key words:** Steroid and xenobiotic receptor, Pregnane X receptor, Humanized mouse,  
Ligand binding domain, Knock-in mouse

### INTRODUCTION

Most orally administered xenobiotics are metabolized first by the intestine and then by the liver after portal transport. The expression levels of enzymes involved in xenobiotic metabolism are regulated at the transcriptional level by key xenobiotic sensors including the ster-

oid and xenobiotic receptor (SXR), also known as the pregnane X receptor (PXR), pregnane activated receptor (PAR) and NR1I2 (Bertilsson *et al.*, 1998; Lehmann *et al.*, 1998; Blumberg *et al.*, 1998). SXR is important in the field of toxicology for at least two reasons. Firstly, this receptor system induces the expression of CYP3A and CYP2B enzymes, the major metabolizers of pharmaceu-

Correspondence: Jun Kanno (E-mail: kanno@nihs.go.jp)

tics and xenobiotics. Therefore, SXR is a key mediator of drug- and chemical-induced toxicity as well as drug-drug and drug-nutrient interactions (Zhou *et al.*, 2004). Secondly, the orthologous rodent and human receptors exhibit differential sensitivity for a subset of chemical ligands important in the field of toxicology. For example, rifampicin (RIF) is a specific and selective activator of human SXR, whereas pregnenolone 16 $\alpha$ -carbonitrile (PCN) is selective for the rodent orthologue.

Rodent-human differences in CYP3A and CYP2B-mediated responses to xenobiotics can be a profound problem in toxicologic studies where rodents are used to predict the toxicity of a compound in humans (Ma *et al.*, 2007). Therefore, development of a murine model that reconstructs the SXR-mediated systemic response of humans is of a great significance in toxicology.

Human and rodent SXRs share ~95% amino acid sequence identity in the DNA-binding domain (DBD) but only about 77% identity in the LBD. Tirona *et al.* (2004) analyzed the ligand selectivity of a human-rat chimeric protein and showed that the species differences are primarily defined by sequence differences in the LBD. Watkins and colleagues showed that the key residues responsible for the majority of the ligand selectivity were Leu 308 (human) and Phe305 (rat and mouse). Crystallographic analysis located these amino acids within or neighboring the flexible loop that forms a part of the pore to the ligand-binding cavity. Swapping the rodent and human-specific residues was shown to modulate the activation by the human-selective activator RIF *in vitro* (Watkins *et al.*, 2001). According to those findings, a simple replacement of the mouse LBD with the human sequence should be sufficient to "humanize" the ligand binding properties as well as activation of the downstream target genes.

Three kinds of humanized mice have already been generated. One is the SXR-null/Alb-SXR mouse (Alb-SXR mouse) made by crossing the SXR knockout mice with a transgenic mouse line that expresses human SXR in liver under the control of the albumin promoter (Xie *et al.*, 2000). Gonzalez and colleagues generated a transgenic mouse expressing a human BAC containing the entire hSXR gene in a SXR null background, thus controlled under human SXR promoter (SXR BAC mouse) (Ma *et al.*, 2007). Another mouse is the human SXR genome knock-in mice (hSXR genome mouse) (Scheer *et al.*, 2008). The human SXR genomic region from exon 2 to exon 9 was knocked-in to mouse SXR exon 2. This mouse expresses the human full length SXR mRNA under the control of mouse SXR promoter regulation. Although useful for toxicology studies, these mice

have disadvantages in that the human SXR is expressed only in the liver (Alb-SXR mouse), hSXR mRNA is not expressed in all of the tissues where SXR is known to be expressed (SXR BAC mouse), and there might be potential differences in the binding affinities of hSXR DNA-binding domain (DBD) to *cis*-acting elements in mouse SXR target genes (hSXR genome mouse).

As noted above, it is known that the critical differences between human and rodent ligand-selectivity reside in the LBD. Therefore, when our project to generate a humanized SXR mouse was initiated, we reasoned that altering the LBD would be sufficient to generate a humanized ligand selectivity. We decided to retain the mouse DBD to avoid any potential differences between the binding affinities of the chimeric receptor for *cis*-acting elements in the mouse genome. To maintain the tissue-specific expression pattern of the endogenous gene, we inserted the human cDNA encoding the region carboxyl-terminal to the DBD into the mouse gene. This retains all of the 5' and 3' regulatory elements in the mouse gene, as well as introns 1 and 2, which contain important elements for regulating SXR expression (Jung *et al.*, 2006).

Here we report a new line of mouse (hSXRki mouse) in which a cDNA encoding the human LBD is homologously recombined into the mouse gene after exon 3. The tissue distribution of the resulting chimeric mouse DBD-human LBD mRNA is comparable to that of the WT mouse. The hSXRki mouse showed a fully humanized response to the human-selective activator RIF in that the Cyp3a11 mRNA was induced in liver and mucosa of small intestine in response to RIF, but not the rodent-selective compound PCN. This new mouse line should provide a useful tool for assessing the whole body toxicity induced by a human selective SXR ligand itself and its subsequently generated metabolite(s) that can trigger further toxic responses through other pathways body-wide.

## MATERIALS AND METHODS

### Generation of hSXRki knock-in mice

A DNA fragment of mouse SXR intron 2 to exon 3 was PCR amplified using mouse BAC DNA (BAC clone No. RP23-351P21) as a template. Primers used were BAC39486FW and mSXR462RV (for sequences of the primers see Table 1). This fragment was connected to the LBD of human SXR cDNA from amino acid 105 through the carboxyl terminus amplified by the PCR primers: hSXR904FW and hSXR1887RVEcoRI (template; human SXR cDNA). The 3'UTR of bovine growth hormone (BGH) was added to 3' to the terminal codon. This concatenated fragment was introduced to a vector, which

## Humanized SXR Mouse by knock-in of human SXR LBD

**Table 1.** List of primer pairs

Purpose	Primer name	Sequence (5' to 3')
Targeting vector construction	BAC39486FW	CCATGGGTACCACGAATAACAA
	mSXR462RV	CATGCCACTCTCCAGGCA
	hSXR904FW	AAGAAGGAGATGATCATGTCCG
	hSXR1887RVEcoRI	CCGAATTCTCATCATCAGCTACCTGTGATACCGAACA
Genotyping	NeoAL2	GGGGATGCGGTGGGCTCTATGGCTT
	SXR RC RV5	TGAGAGTGCACAAGTTCAAGCT
	WTInt5	AGTGATGGGAACCACTCCTG
	WTEx6RV	TGGTCCTCAATAGGCAGGTC
	mhSXRE4	GTGAACGGACAGGGACTCAG
	mhSXRSARV	CTCTCCTGGCTCATCTCAC
Percellome quantitative RT-PCR	Cyp3a11 FW	CAGCTTGGTGCTCCTCTACC
	Cyp3a11 RV	TCAAACAACCCCATGTTTT
	Ces6 FW	GGAGCCTGAGTTCAGGACAGAC
	Ces6 RV	ACCCTCACTGTTGGGGTTC
	mouse SXR FW	AATCATGAAAGACAGGGTTC
	mouse SXR RV	AAGAGCACAGATCTTTCCG
	human SXR FW	ATCACCCGGAAGACACGAC
	human SXR RV	AAGAGCACAGATCTTTCCG
	mouse-human SXR FW	CCCATCAACGTAGAGGAGGA

has the neomycin resistance gene with loxP sequence at both ends, removable with Cre recombinase (Saga *et al.*, 1999). A 7kb KpnI fragment containing intron 2 was used as a long arm and 1.3kb PstI-EcoRI fragment containing from exon 8 to intron 8 was used as a short arm for homologous recombination (Fig. 1). The resulting targeting vector was linearized with SacII and introduced by electroporation to TT2 ES cell line (Yagi *et al.*, 1993) and neomycin resistant clones were selected, PCR genotyped, and confirmed by the Southern blotting. For generation of chimeric mice, these ES clones were aggregated with ICR 8-cell embryos and transferred to pseudopregnant female recipients. The chimeric mice born were bred with ICR females. Germ line transmission of the targeted allele was confirmed by PCR. A mouse was crossed with a CAG-Cre transgenic mouse (Sakai and Miyazaki, 1997) to evict the neomycin resistance gene, and back crossed to C57BL/6 CrSlc (SLC, Inc., Shizuoka, Japan) at least 6 generations and used for the analysis.

**PCR Genotyping**

(See Table 1 for primer sequences)

Primers for identification of homologously recombined ES clones were NeoAL2 and SXR RC RV5. DNA purified from the tail of each mouse was used for PCR genotyping. Primers for WT detection were WTInt5 and WTEx6RV amplifying a product of 755 bp. Primers for

confirmation of removal of the neomycin resistance gene were mhSXRE4 and mhSXRSARV amplifying a product of 1,223 bp.

**Southern blot analysis**

To confirm homologous recombination, DNA from ES cell cultures was purified and digested with BamHI and XhoI, then electrophoresed and analyzed by Southern hybridization (Saga *et al.*, 1997). Mouse SXR exon 9 region which remains after homologous recombination was used for the probe. The restriction fragments from the WT allele and targeted allele are 2,305 bp and 1,925 bp, respectively.

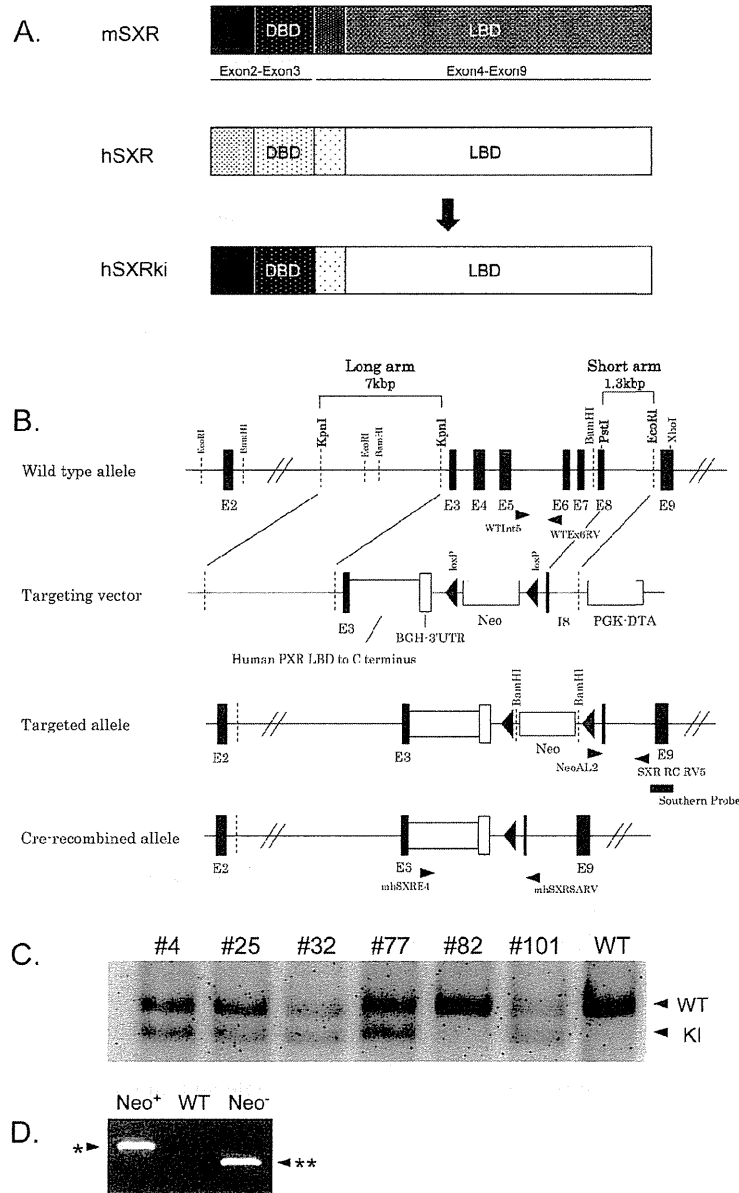
**Chemicals**

RIF (molecular weight 822.95) and PCN (molecular weight 341.49) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Corn oil was purchased from Wako Pure Chemical Industries (Osaka, Japan).

**Quantitative RT-PCR (Percellome PCR)**

(See Table 1 for primer sequences)

The method for Percellome quantitative RT-PCR was described previously (Kanno *et al.*, 2006). Briefly, tissue pieces stored in RNAlater (Ambion, Austin, TX, USA) were homogenized and lysed in RLT buffer (Qiagen GmbH, Germany) and 10 µl aliquots were used



**Fig. 1.** Targeting strategy used to generate the hSXRki mouse. A) Diagram of hSXRki chimeric protein. Hinge region and ligand binding domain (LBD) of human SXR are knocked-in to mouse SXR, resulting in chimeric protein having murine N-terminal domain and DNA binding domain (DBD). B) Targeting strategy used to generate the hSXRki mouse. The chimeric mouse DBD and human LBD fragment, followed by the BGH 3' UTR were knocked-in to the mouse SXR gene. The genomic region spanning from exon 3 to exon 8 was substituted by the inserted fragment with the remainder of the gene remaining intact. C) Confirmation of homologous recombination by southern blot analysis. Six ES clones positive for recombination by PCR genotyping were further analyzed by southern blot (clones #4 ~ #101). Lower bands (1925 bp) indicate successful homologous recombination; upper bands (2305 bp) correspond to WT allele. Clones #4, #25, #32, #77 and #101 were confirmed as homologous recombinants; clones #4 and #25 were used for the generation of chimeric mice. D) Confirmation of Cre-mediated removal of the neomycin resistance gene. Mouse tail genome DNA was PCR amplified with the primer set, mhSXR E4 and mhSXR SARV. \*: 2,858 bp (for the mice having the neomycin resistance gene), \*\*: 1,223 bp (for the mice without the neomycin resistance gene).

## Humanized SXR Mouse by knock-in of human SXR LBD

for genomic DNA quantification with PicoGreen fluorescent dye (Invitrogen, Carlsbad, CA, USA). A prepared spike mRNA cocktail solution containing known quantity of five mRNAs of bacillus subtilis was added to the tissue lysate in proportion to the DNA quantity. Total RNA was purified from the lysate using the RNeasy kit (Qiagen). One microgram of total RNA was reverse-transcribed with SuperScript II (Invitrogen). Quantitative real time PCR was performed with an ABI PRISM 7900 HT sequence detection system (Applied Biosystems) using SYBR Green PCR Master Mix (Applied Biosystems), with initial denaturation at 95°C for 10 min followed by 40 cycles of 30 sec at 95°C and 30 sec at 60°C and 30 sec at 72°C, and Ct values were obtained. Primers for Cyp3a11 were Cyp3a11 FW and Cyp3a11 RV. Primers for Ces6 were Ces6 FW and Ces6 RV. Primers for mouse SXR selective quantification were mouse SXR FW and mouse SXR RV. Primers for hSXRki selective quantification were human SXR FW and human SXR RV. Primers for both mouse SXR and hSXRki quantification were mouse-human SXR FW and mouse-human SXR RV that amplify the DBD region of the chimera.

**In Situ Hybridization analysis**

Digoxigenin-labeled cRNA probe for Cyp3a11 was synthesized according to Suzuki *et al.* (2005) by RT-PCR using mouse liver cDNA as a template. The primers used were as follows: forward 5'-GATTGGTTTTGATGCCTGGT-3' and reverse 5'-CAAGAGCTCACATTTTTCATCA-3'. The amplified product was sequence confirmed

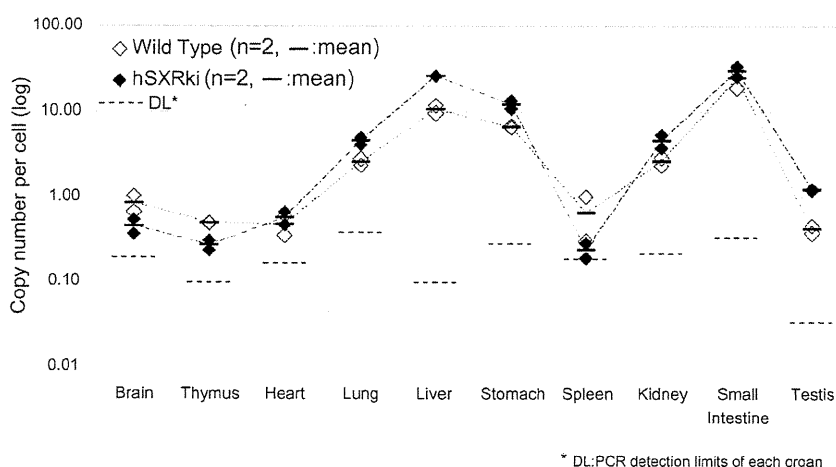
and ligated with Block-iT T7-TOPO (Invitrogen) Linker, which contains the T7 promoter site. A secondary PCR was performed to generate the sense and antisense DNA templates. For antisense template, Block-iT T7 Primer and Cyp3a11 forward primer (or reverse primer for generation of sense DNA template), the same primer as for the first PCR amplification, were used. With these DNA templates, both sense and antisense digoxigenin-labeled riboprobes were synthesized using a DIG RNA labeling kit (Roche Diagnostics, Germany) according to the manufacturer's protocol.

ISH on paraffin sections was carried out according to Suzuki *et al.* with a modification; permeabilization condition 98°C for 15 min in HistoVT One (Nacalai tesque, Japan).

**Animals experiments**

Male hSXRki and WT mice were maintained under a 12 hr light/12 hr dark cycle with water and chow (CRF-1, Oriental Yeast Co. Ltd., Tokyo, Japan) provided *ad libitum*. The animal studies were conducted in accordance with the Guidance for Animal Studies of the National Institute of Health Sciences under Institutional approval. The expression level of the hSXRki and WT SXR mRNA of ten organs (brain, thymus, heart, lung, liver, stomach, spleen, kidney, small intestine and testis) were analyzed on 15 weeks old male mice (n = 2) by the Percellome quantitative RT-PCR.

For the demonstration of selective gene induction by RIF and PCN in hSXRki and WT male mice on 13 weeks



**Fig. 2.** Conservation of tissue expression patterns of hSXRki mRNA in the knock-in mouse. Percellome quantitative RT-PCR analysis was performed to measure the absolute expression levels of WT SXR mRNA and hSXRki mRNA in ten organs of WT and hSXRki mice. The expression levels of hSXRki mRNA among organs were comparable to WT.

old, three mice per group were singly dosed orally with vehicle (corn oil+0.1% DMSO), 10, 30, or 100 mg/kg of RIF, or 20, 70, or 200 mg/kg PCN (approximately equivalent in molar dose). Eight hours later, mice were sacrificed by exsanguination under ether anesthesia and the liver and the small intestine mucosa were sampled. Liver samples in small pieces were stored in RNA later (Applied Biosystems, Foster City, CA, USA) for further analysis. The small intestine under ice-cooled condition was longitudinally opened, gently rinsed with RNase-free saline and the epithelium was scraped with a glass slide and immersed in RNAlater. For *in situ* hybridization (ISH) of Cyp3a11 in the liver, 15 weeks old male hSXRki and WT mice were dosed orally with vehicle (corn oil), RIF (10 mg/kg), or PCN (40 mg/kg) daily for 3 days and liver sampled 24 hr later. All mice were sacrificed by exsanguination under ether anesthesia.

### Statistical analysis

All values are expressed as the means  $\pm$  S.D. and group differences analyzed by unpaired Student's *t* test or one-way ANOVA followed by Dunnett's post hoc comparison. Level of significance was set at  $p < 0.05$ .

## RESULTS

### Generation of hSXRki knock-in mice

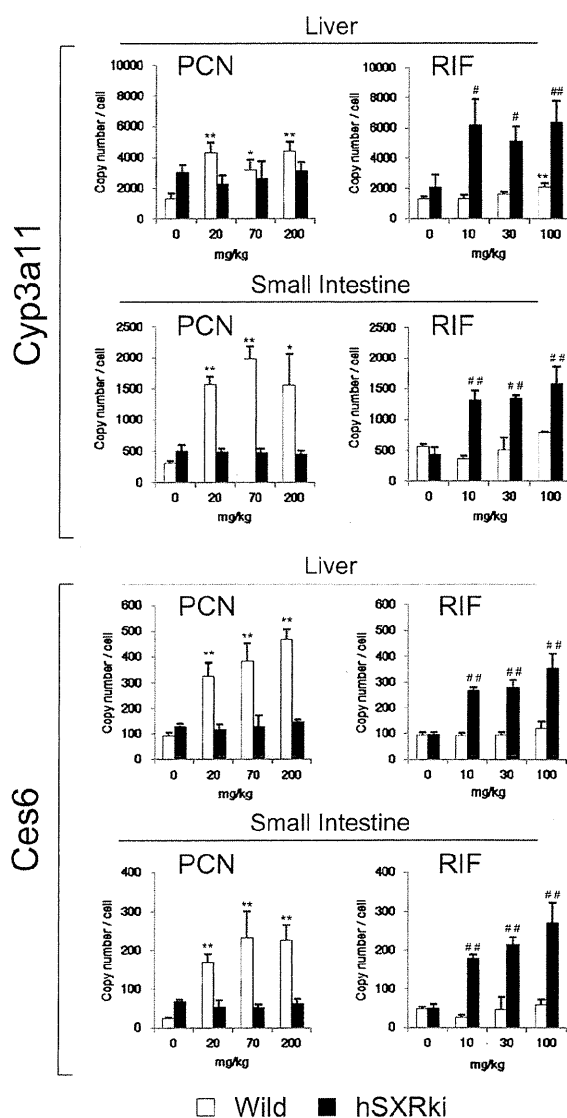
Among 144 neomycin resistant TT2 ES clones, six PCR positive clones were further submitted to Southern blotting for the confirmation of homologous recombination. As shown in Fig. 1C, five clones were confirmed, and two (#4 and #25) were used to generate chimeric mice. The resulting mice were backcrossed to ICR strain to confirm germline transmission. One clone (#4) was crossed to a mouse constitutively expressing Cre recombinase to remove the neomycin resistance gene (Fig. 1D) and backcrossed to C57BL/6 CrSlc for at least 6 generations before further analysis.

### Tissue distribution of hSXRki mRNA

Ten tissues, i.e., brain, thymus, heart, lung, liver, stomach, spleen, kidney, small intestine and testis from both hSXRki and WT mice were measured for hSXRki or WT SXR mRNA expression by the Percellome quantitative RT-PCR. As shown in Fig. 2, the levels of hSXRki mRNA are comparable to that of SXR in WT mouse and expressed in all tissues analyzed.

### Humanized responses in hSXRki mouse

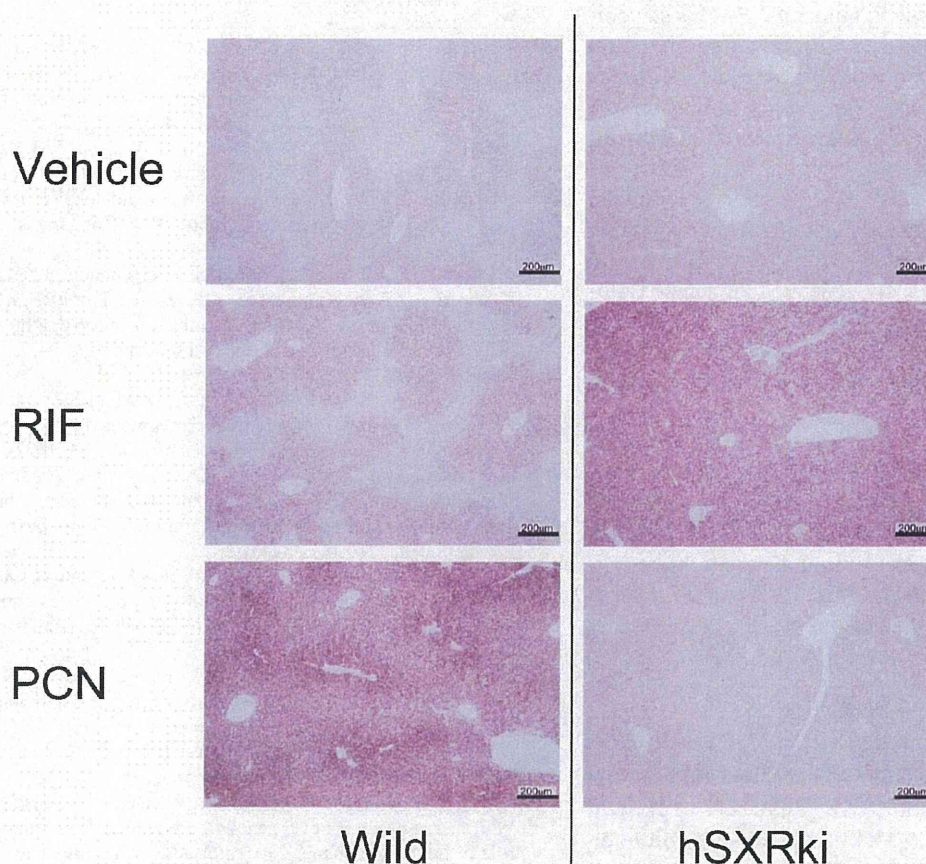
Humanized response of hSXRki was demonstrated by administration of the mouse-specific ligand PCN and the



**Fig. 3.** Humanized response of hSXRki mice to RIF and PCN; Percellome quantitative RT-PCR. WT mice and hSXRki mice ( $n = 3$  each) were singly dosed orally with vehicle (corn oil+0.1% DMSO), 20, 70, or 200 mg/kg PCN, or 10, 30, or 100 mg/kg of RIF (approximately equivalent in molar dose each other). Percellome quantitative RT-PCR data of Cyp3a11 and Ces6, both known as SXR target genes, in liver and small intestinal mucosa showed humanized responses in hSXRki. Bars = S.D., \*,  $p < 0.05$ , \*\*,  $p < 0.01$  compared with vehicle group of WT, #,  $p < 0.05$ , ##,  $p < 0.01$  compared with vehicle group of hSXRki. Analyzed by one-way ANOVA followed by Dunnett's post hoc comparison. Level of significance was set at  $p < 0.05$ .

## Humanized SXR Mouse by knock-in of human SXR LBD

## ISH of Cyp3a11



**Fig. 4.** Humanized response of hSXRki mice to RIF and PCN; *In situ* hybridization for Cyp3a11 mRNA in liver. A DIG-labeled cRNA probe for Cyp3a11 was hybridized and developed for purplish blue chromogenic reaction. Histologically, Cyp3a11 induction was localized around the central veins in both mice with species-specific ligands, respectively.

human-specific ligand RIF to the mice. Induction of the well-known SXR-regulated genes, Cyp3a11 and Ces6 was monitored by Percellome quantitative RT-PCR. As shown in Fig. 3, in the liver and small intestinal mucosa, RIF, but not PCN, induced Cyp3a11 and Ces6 in hSXRki mice (closed column), whereas PCN exclusively induced these genes in WT mice (open column). ISH of Cyp3a11 of the liver also showed humanized responses in hSXRki mice (Fig. 4).

## DISCUSSION

We generated a new humanized mouse model in which the ligand binding domain (LBD) of human SXR was homologously knocked-into the murine SXR gene so that systemic response induced by human-selective SXR ligands can be monitored in mice. Firstly, we showed that mRNA from this chimeric gene was expressed at appropriate levels in the same tissues as the endogenous mouse SXR gene in WT mice. Then the humanized response of the mouse was confirmed by monitoring its response to the human-selective activator RIF, and the lack of response to the rodent-selective activator PCN.

There are relatively few reports about the regulation of SXR expression to date. Aouabdi *et al.* (2006) reported the presence of a PPAR alpha binding site 2.2 kb upstream of the transcription start site in human SXR. This site corresponded to the induction site with clofibrate in the rat and they further confirmed its importance using human liver cancer cell line (Huh7). Jung *et al.* (2006) reported the presence of four FXR binding sites in intron 2 of the mouse SXR gene that were required for FXR regulation of SXR expression. This intron 2 region is completely intact in our hSXRki mouse. Therefore, the regulation by FXR should be preserved in our mice.

Compared to the previously generated humanized Alb-SXR, SXR BAC, and hSXR genome mice, we contend that our hSXRki mouse has an advantage because the human-mouse chimeric gene is expressed in the same tissues and at similar levels to endogenous SXR in WT mice under control of the mouse promoter. This feature would make this model suitable not only for systemic toxicity but also toxicity at various stages of development of the embryo and fetus, maturation of infant, and of senescence, where the *cis* and *trans* regulations might be critical in its regulation (Sarsero *et al.*, 2004) (Konopka *et al.*, 2009). Thus, we believe that our system has a broader application range for toxicological studies.

### ACKNOWLEDGMENTS

The authors thank Ms. Yuko Matsushima, Mr. Masaki Tsuji, Ms. Maki Otsuka, Mr. Yusuke Furukawa, Mr. Kouichi Morita, Ms. Maki Abe, and Ms. Shinobu Watanabe for technical support. This study was supported in part by the Health Sciences Research Grants H19-Toxico-Shitei-001 from the Ministry of Health, Labour and Welfare, Japan.

### REFERENCES

- Aouabdi, S., Gibson, G. and Plant, N. (2006): Transcriptional regulation of the PXR gene: identification and characterization of a functional peroxisome proliferator-activated receptor alpha binding site within the proximal promoter of PXR. *Drug Metab. Dispos.*, **34**, 138-144.
- Bertilsson, G., Heidrich, J., Svensson, K., Asman, M., Jendeborg, L., Sydow-Bäckman, M., Ohlsson, R., Postlind, H., Blomquist, P. and Berkenstam, A. (1998): Identification of a human nuclear receptor defines a new signaling pathway for CYP3A induction. *Proc. Natl. Acad. Sci., USA*, **95**, 12208-12213.
- Blumberg, B., Sabbagh, W.Jr., Juguilon, H., Bolado, J.Jr., van Meter, C.M., Ong, E.S. and Evans, R.M. (1998): SXR, a novel steroid and xenobiotic-sensing nuclear receptor. *Genes. Dev.*, **12**, 3195-3205.
- Jung, D., Mangelsdorf, D.J. and Meyer, U.A. (2006): Pregnane X receptor is a target of farnesoid X receptor. *J. Biol. Chem.*, **281**, 19081-19091.
- Kamo, J., Aisaki, K., Igarashi, K., Nakatsu, N., Ono, A., Kodama, Y. and Nagao, T. (2006): "Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays. *BMC genomics*, **7**, 64.
- Konopka, G., Bomar, J.M., Winden, K., Coppola, G., Jonsson, Z.O., Gao, F., Peng, S., Preuss, T.M., Wohlschlegel, J.A. and Geschwind, D.H. (2009): Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature*, **462**, 213-217.
- Lehmann, J.M., McKee, D.D., Watson, M.A., Willson, T.M., Moore, J.T. and Kliewer, S.A. (1998): The human orphan nuclear receptor PXR is activated by compounds that regulate CYP3A4 gene expression and cause drug interactions. *J. Clin. Invest.*, **102**, 1016-1023.
- Ma, X., Shah, Y., Cheung, C., Guo, G.L., Feigenbaum, L., Krausz, K.W., Idle, J.R. and Gonzalez, F.J. (2007): The PREgnane X receptor gene-humanized mouse: a model for investigating drug-drug interactions mediated by cytochromes P450 3A. *Drug Metab. Dispos.*, **35**, 194-200.
- Saga, Y., Hata, N., Koseki, H. and Taketo, M.M. (1997): Mesp2: a novel mouse gene expressed in the presegmented mesoderm and essential for segmentation initiation. *Genes. Dev.*, **11**, 1827-1839.
- Saga, Y., Miyagawa-Tomita, S., Takagi, A., Kitajima, S., Miyazaki, J. and Inoue, T. (1999): MesP1 is expressed in the heart precursor cells and required for the formation of a single heart tube. *Development*, **126**, 3437-3447.
- Sakai, K. and Miyazaki, J. (1997): A transgenic mouse line that retains Cre recombinase activity in mature oocytes irrespective of the cre transgene transmission. *Biochem. Biophys. Res. Commun.*, **237**, 318-324.
- Sarsero, J.P., Li, L., Holloway, T.P., Voullaire, L., Gazeas, S., Fowler, K.J., Kirby, D.M., Thorburn, D.R., Galle, A., Cheema, S., Koenig, M., Williamson, R. and Ioannou, P.A. (2004): Human BAC-mediated rescue of the Friedreich ataxia knockout mutation in transgenic mice. *Mamm. Genome*, **15**, 370-382.
- Scheer, N., Ross, J., Rode, A., Zevnik, B., Niehaves, S., Faust, N. and Wolf, C.R. (2008): A novel panel of mouse models to evaluate the role of human pregnane X receptor and constitutive androstane receptor in drug response. *J. Clin. Invest.*, **118**, 3228-3239.
- Suzuki, T., Akimoto, M., Mandai, M., Takahashi, M. and Yoshimura, N. (2005): A new PCR-based approach for the preparation of RNA probe. *J. Biochem. Methods*, **62**, 251-258.
- Tirona, R.G., Leake, B.F., Podust, L.M. and Kim, R.B. (2004): Identification of amino acids in rat pregnane X receptor that determine species-specific activation. *Mol. Pharmacol.*, **65**, 36-44.
- Watkins, R.E., Wisely, G.B., Moore, L.B., Collins, J.L., Lambert, M.H., Williams, S.P., Willson, T.M., Kliewer, S.A. and Redinbo, M.R. (2001): The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science*, **292**, 2329-2333.
- Xie, W., Barwick, J.L., Downes, M., Blumberg, B., Simon, C.M., Nelson, M.C., Neuschwander-Tetri, B.A., Brunt, E.M., Guzelian, P.S. and Evans, R.M. (2000): Humanized xenobiotic response in mice expressing nuclear receptor SXR. *Nature*, **406**, 435-439.
- Yagi, T., Tokunaga, T., Furuta, Y., Nada, S., Yoshida, M., Tsukada, T., Saga, Y., Takeda, N., Ikawa, Y. and Aizawa, S. (1993): A novel ES cell line, TT2, with high germline-differentiating potency. *Anal. Biochem.*, **214**, 70-76.
- Zhou, C., Tabb, M.M., Sadatrafiei, A., Grün, F. and Blumberg, B. (2004) Tocotrienols activate the steroid and xenobiotic receptor, SXR, and selectively regulate expression of its target genes. *Drug Metab. Dispos.*, **32**, 1075-1082.



# GlcNAcylation of histone H2B facilitates its monoubiquitination

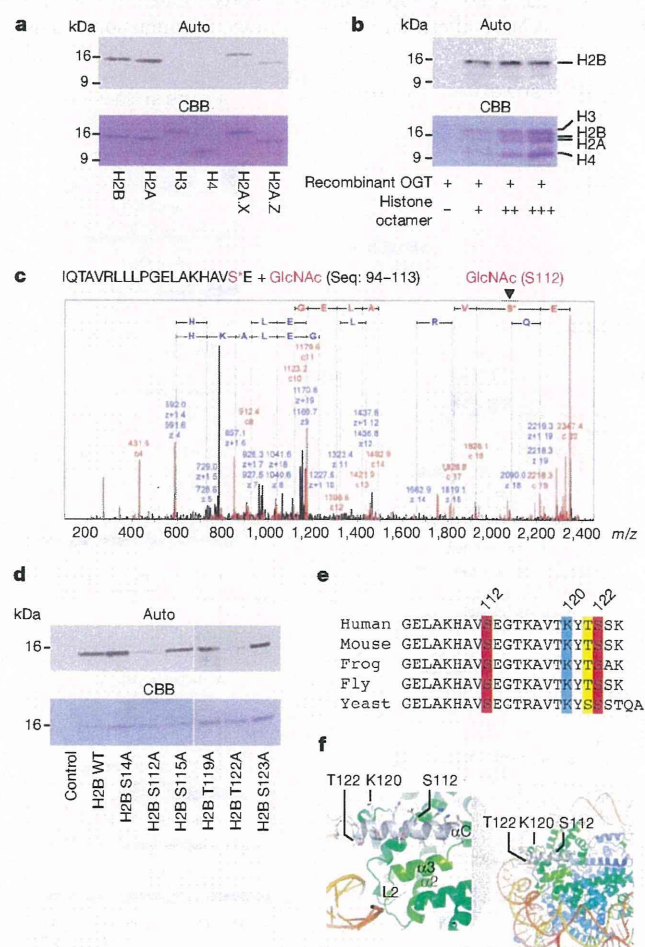
Ryoji Fujiki<sup>1</sup>, Waka Hashiba<sup>1</sup>, Hiroki Sekine<sup>1</sup>, Atsushi Yokoyama<sup>1</sup>, Toshihiro Chikanishi<sup>1</sup>, Saya Ito<sup>1</sup>, Yuuki Imai<sup>1</sup>, Jaehoon Kim<sup>2</sup>, Housheng Hansen He<sup>3</sup>, Katsuhide Igarashi<sup>4</sup>, Jun Kanno<sup>4</sup>, Fumiaki Ohtake<sup>1</sup>, Hirochika Kitagawa<sup>1</sup>, Robert G. Roeder<sup>2</sup>, Myles Brown<sup>3</sup> & Shigeaki Kato<sup>1,5</sup>

Chromatin reorganization is governed by multiple post-translational modifications of chromosomal proteins and DNA<sup>1,2</sup>. These histone modifications are reversible, dynamic events that can regulate DNA-driven cellular processes<sup>3,4</sup>. However, the molecular mechanisms that coordinate histone modification patterns remain largely unknown. In metazoans, reversible protein modification by *O*-linked *N*-acetylglucosamine (GlcNAc) is catalysed by two enzymes, *O*-GlcNAc transferase (OGT) and *O*-GlcNAcase (OGA)<sup>5,6</sup>. However, the significance of GlcNAcylation in chromatin reorganization remains elusive. Here we report that histone H2B is GlcNAcylated at residue S112 by OGT *in vitro* and in living cells. Histone GlcNAcylation fluctuated in response to extracellular glucose through the hexosamine biosynthesis pathway (HBP)<sup>5,6</sup>. H2B S112 GlcNAcylation promotes K120 monoubiquitination, in which the GlcNAc moiety can serve as an anchor for a histone H2B ubiquitin ligase. H2B S112 GlcNAc was localized to euchromatic areas on fly polytene chromosomes. In a genome-wide analysis, H2B S112 GlcNAcylation sites were observed widely distributed over chromosomes including transcribed gene loci, with some sites co-localizing with H2B K120 monoubiquitination. These findings suggest that H2B S112 GlcNAcylation is a histone modification that facilitates H2BK120 monoubiquitination, presumably for transcriptional activation.

Some nuclear proteins have been shown to be GlcNAcylated by OGT, for example the enzymatic activity of histone H3K4 methyltransferase 5 (MLL5) is modulated by GlcNAcylation<sup>7–9</sup>. To identify chromatin substrates for OGT further, we screened for unknown GlcNAcylated glycoproteins in HeLa cell chromatin. GlcNAcylated proteins were purified by WGA lectin column chromatography and anti-GlcNAc antibody (clone RL2). Liquid chromatography–mass spectrometry (LC–MS)/MS analysis of the fraction revealed 284 factors, including previously reported GlcNAcylated glycoproteins<sup>6,10</sup> (Supplementary Table 1). Among the candidates, the enrichment of nucleosomes was confirmed by silver staining and western blotting (Supplementary Fig. 2), suggesting one or more histone(s) might have been GlcNAcylated. As OGT is the only known nuclear enzyme for protein GlcNAcylation<sup>5</sup>, we asked whether histones served as substrates for OGT *in vitro* (Supplementary Fig. 3). H2A and H2B, as well as H2A variants (H2A.X and H2A.Z), but not H3 and H4, appeared to be GlcNAcylated (Fig. 1a). With histone octamers, H2B, but not H2A, appeared to serve as a substrate (Fig. 1b). Likewise, H2B in *Drosophila* histone was also GlcNAcylated (Supplementary Fig. 4), implying that H2B GlcNAcylation is conserved in metazoans.

A quadrupole (Q)-time of flight (TOF) MS assessment of the *in vitro* GlcNAcylated H2B showed that OGT could transfer three GlcNAc moieties to H2B (Supplementary Fig. 5). Electro-transfer-dissociation (ETD)–MS/MS mapped the sites to S91, S112 and S123 (Fig. 1c and Supplementary Fig. 6). Unlike a recent report<sup>11</sup>, we were unable to

detect the reported sites in H2B S36 and H4 S47. However, H2A T101 was detected as a GlcNAc site when H2A protein alone was used (data not shown). This discrepancy in identified GlcNAc sites might be due to differences in experimental approaches.



**Figure 1 | H2B is GlcNAcylated at the C-terminal S112.** **a, b**, *In vitro* OGT assay with recombinant histones (**a**) or the octamers reconstituted *in vitro* (**b**). Histones were GlcNAcylated by uridine diphosphate (UDP)-[<sup>3</sup>H]GlcNAc and OGT, and the radiolabelled histones were subjected to autoradiography (top) and CBB staining (bottom). **c**, ETD–MS/MS scanned the GlcNAcylated peptides (2349.43 *m/z*) in Supplementary Fig. 5b. **d**, A series of H2B mutants at the indicated S/T was assessed by *in vitro* OGT assays. **e**, Sequence alignment of  $\alpha$ C. **f**, The locations of the GlcNAc sites and the ubiquitination site of H2B in a nucleosome. The  $\alpha$ C helix is illustrated as a white ribbon.

<sup>1</sup>Institute of Molecular and Cellular Biosciences, University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan. <sup>2</sup>Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, New York 10065, USA. <sup>3</sup>Department of Medical Oncology, Dana-Farber Cancer Institute and Harvard Medical School, Boston, Massachusetts 02115, USA. <sup>4</sup>Division of Cellular and Molecular Toxicology, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. <sup>5</sup>ERATO, Japan Science and Technology Agency, Kawaguchi, Saitama 332-0012, Japan.

Next, *in vitro* OGT assays using peptide arrays covering full-length H2B revealed peaks at 101–115 peptides in the carboxy (C)-terminal  $\alpha$ -helix ( $\alpha$ C)<sup>12</sup> (Supplementary Fig. 7). This peptide was found to bear only one moiety by matrix-assisted laser desorption/ionization–time of flight (MALDI-TOF)/MS (Supplementary Fig. 8). Indeed, substitutions of S112 and T122 to A significantly reduced *in vitro* GlcNAcylation by OGT (Fig. 1d), but not mutations in the amino (N)-terminal tail (Supplementary Fig. 9). On the basis of these data, we concluded that the conserved S112 was a GlcNAc site in H2B, whereas T122 might be needed for recognition by OGT (Fig. 1e, f).

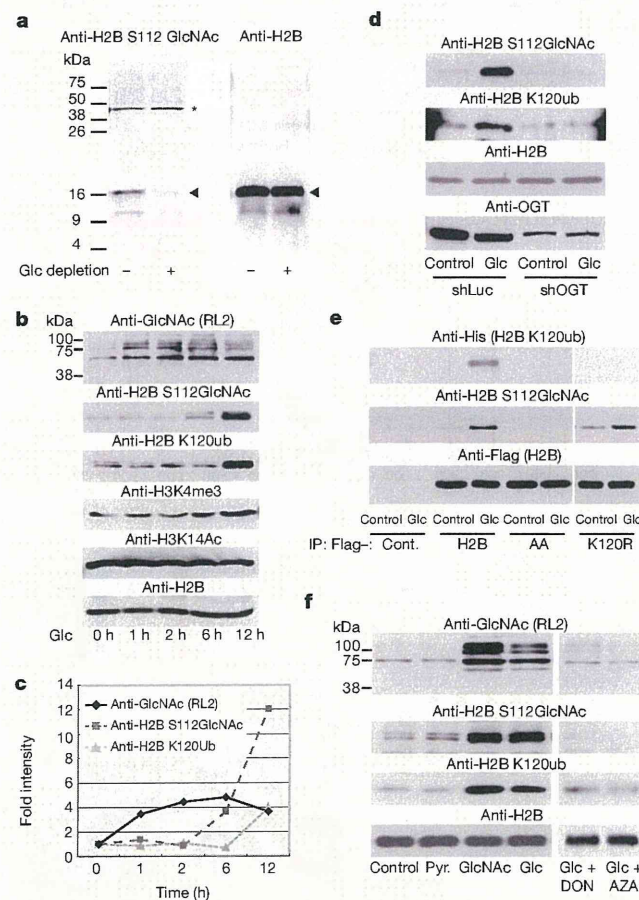
With our newly developed antibody (Supplementary Fig. 10), H2B S112 GlcNAc was detected in histones of HeLa cells. Depletion of glucose from the media for 24 h induced deglycosylation with neither overt cell death (Fig. 2a and Supplementary Fig. 11) nor alteration in histone acetylation marks of cell state indicators (H3 K14, H3 K56, H4 K16)<sup>13,14</sup> (Supplementary Fig. 12). H2B S112 GlcNAc could be restored by re-treatment with glucose at physiological concentrations (Supplementary Fig. 13).

Because many histone modifications are orchestrated, we tested if H2B S112 GlcNAc influenced H2B K120 monoubiquitination because

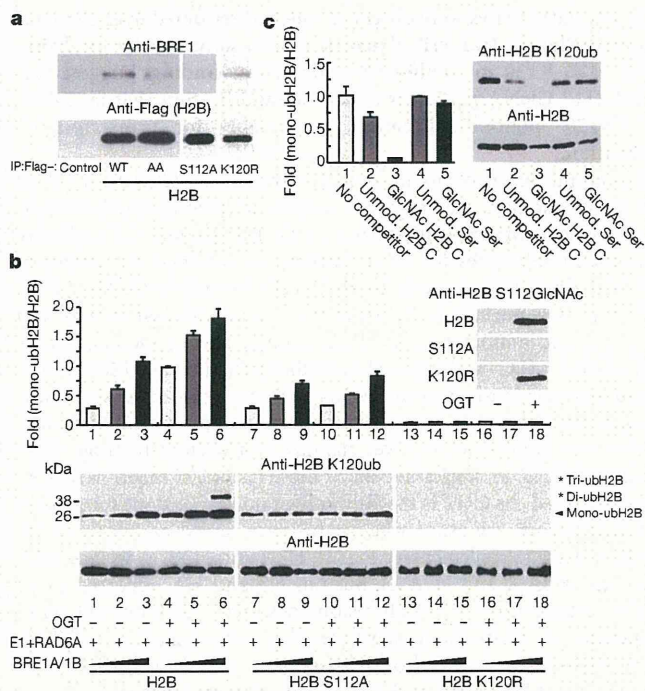
of their proximity. After glucose depletion, replenishment of glucose gradually increased global GlcNAcylation of proteins, followed by H2B S112 GlcNAc and H2B monoubiquitination (Fig. 2b, c). Their reciprocal modifications disappeared when OGT was knocked down (Fig. 2d and Supplementary Fig. 14). In addition, in the immunoprecipitates of H2B containing the S112A and T122A double mutations (H2B AA), no response of K120 monoubiquitination to extracellular glucose was detected (Fig. 2e and Supplementary Fig. 15). Conversely, GlcNAcylation of H2B S112 was observed, even when K120 was mutated to R (Fig. 2e). From these findings, we conclude that H2B K120 monoubiquitination is mediated, at least in part, through S112 GlcNAcylation.

As glucosamine, but not pyruvate, potentiated H2B S112 GlcNAc (Fig. 2f), it appeared that this GlcNAcylation step was dependent on the HBP. To clarify this point, two HBP inhibitors (DON and AZA) were tested (Supplementary Information). After glucose depletion from media, these inhibitors attenuated the effect of glucose in H2B S112 GlcNAcylation along with K120 monoubiquitination (Fig. 2f).

In yeast, it was previously shown that H2B K120 monoubiquitination was induced by carbohydrates by glycolysis<sup>15</sup>. To address this issue, inhibitors of both glycolysis and deGlcNAcylation were applied to assess the crosstalk between the two modifications. When the cells were treated with iodoacetate, which blocks glycolysis but not HBP<sup>15</sup>, the glucose effects on histone modifications were impaired, whereas the additional treatment of an OGA inhibitor (PUGNAc) restored both H2B S112 and K120 monoubiquitination (Supplementary Fig. 16). These data support the notion that H2B S112 GlcNAc senses decreases in glucose levels below normal levels and acts to promote H2B monoubiquitination, a modification that is associated with active transcription. Together with the fact that OGT is absent in yeast<sup>6</sup>, the present H2B S112 GlcNAc-dependent pathway appears to constitute a system capable of sensing nutritional states in metazoans.



**Figure 2 | H2B S112 GlcNAc is a glucose-responsive modification linked to K120 monoubiquitination (ub).** **a**, Chromatin was prepared from HeLa cells cultured in media with or without  $1 \text{ g l}^{-1}$  glucose (Glc) for 24 h, and subjected to western blotting. Arrowheads show the indicated proteins. Asterisks indicate non-specific band. **b, c**, After 24 h Glc depletion, chromatin samples were prepared from HeLa cells treated with  $4.5 \text{ g l}^{-1}$  Glc for the indicated time. The intensities of the western blotting bands (**b**) were quantified (**c**). **d, e**, The effects of OGT knockdown (**d**) or H2B mutations (**e**) on H2B modifications after Glc replenishment. **f**, Western blotting analysis of the H2B modifications in HeLa cells that were cultured in DMEM without Glc (Cont.), or supplemented with 1 mM pyruvate (Pyr.), 10 mM GlcNAc or  $4.5 \text{ g l}^{-1}$  Glc with or without HBP inhibitors, 6-diazo-5-oxo-L-norleucine (100  $\mu\text{M}$ , DON) or azaserine (100  $\mu\text{M}$ , AZA).



**Figure 3 | GlcNAcylation at S112 facilitates ubiquitination at K120 in H2B.** **a**, Western blotting analysis of the interaction of H2B mutants with BRE1A. **b, c**, *In vitro* monoubiquitination assay with GlcNAcylated H2B (**b**), or in the presence of competitor peptides (**c**). H2B was GlcNAcylated *in vitro* (**b**, top right), and the reactants were subsequently ubiquitinated by H2B monoubiquitination ligase. The reaction was performed with the indicated competitor peptides (0.25  $\mu\text{g ml}^{-1}$ ) (**c**). H2B K120 monoubiquitination was detected by western blotting (**b**, bottom; **c**, right) and quantified (**b**, top; **c**, left). Error bars, means and s.d. ( $n = 3$ ).

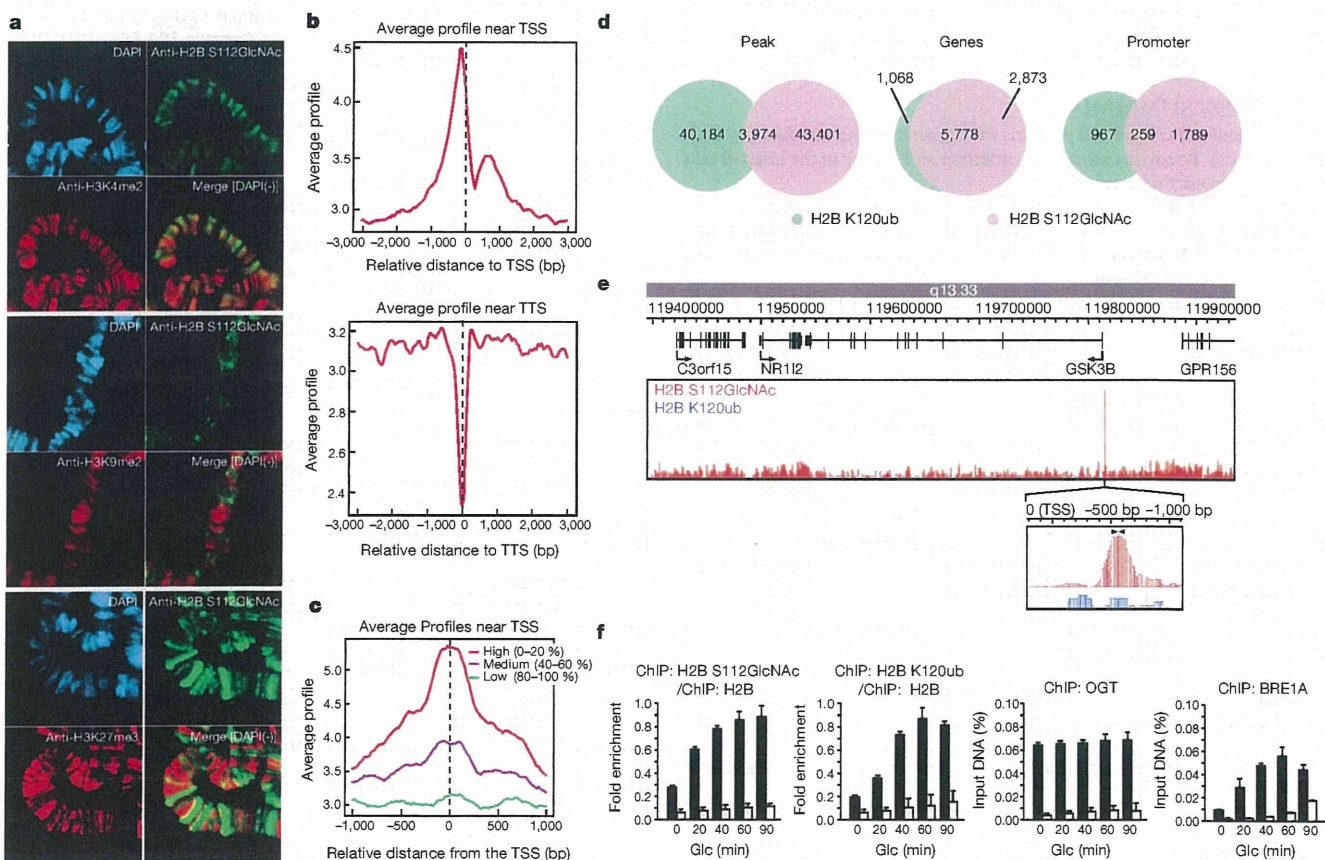
The terminal GlcNAc of polysaccharides reportedly serves as a recognition moiety for E3 monoubiquitination ligase<sup>16</sup>. Therefore, we proposed that H2B S112 GlcNAc affected K120 monoubiquitination by the BRE1A/1B complex<sup>17</sup>. Flag-tagged H2B, but not AA or S112A, was co-immunoprecipitated with BRE1A (Fig. 3a). This association was observed in the presence of physiological levels of glucose in the media, and BRE1A was bound to H2B S112 GlcNAc (Supplementary Fig. 17). We then assessed how the GlcNAcylation of H2B influenced its *in vitro* ubiquitination by E1, RAD6A (E2) and the BRE1A/1B complex (E3). Although H2B K120 could be substantially ubiquitinated only by the ligases (Supplementary Fig. 18), GlcNAcylation of H2B promoted subsequent H2B ubiquitination, but not its S112A mutant (Fig. 3b). Likewise, ubiquitination was significantly attenuated by the presence of an H2B-S112-GlcNAcylated peptide, but not by either the unmodified control peptide or by GlcNAcylated serine (Fig. 3c). On the basis of these results, we conclude that the GlcNAc moiety at H2B S112 may anchor H2B monoubiquitination ligase.

To illustrate the role of H2B S112 GlcNAc in chromatin regulation, its location was visualized on fly polytene chromosomes. H2B S112 GlcNAc was detected widely in euchromatin, and, as anticipated, its signal disappeared in an OGT-disrupted fly, *sxc<sup>1</sup>/sxc<sup>78</sup>* (Supplementary Fig. 19). H2B S112 GlcNAc overlapped with H3K4 me2 more than with H3K9 me2 or H3K27 me3 (Fig. 4a). Similarly, in immunostained HeLa cells, H2B S112 GlcNAc sites appeared exclusively in 4',6-diamidino-2-phenylindole (DAPI)-poor areas (Supplementary Fig. 20).

Thus, H2B S112 GlcNAc probably accumulates in active chromatin rather than inactive chromatin.

To determine the precise loci of H2B S112 GlcNAc in HeLa cells, we performed chromatin immunoprecipitation (ChIP) and high-throughput sequencing (ChIP-seq). We confirmed ChIP quality by enrichments of H2B GlcNAc as well as H3K4 me2 and H2B K120 monoubiquitination, but neither H3K9 me2 nor H3K27 me3 (Supplementary Fig. 21). A total of 47,375 peaks were found widely distributed over the genome (Supplementary Fig. 22). However, H2B S112 GlcNAc peaked near transcription start sites (TSS), whereas the distribution decreased at transcription termination sites (TTS) (Fig. 4b), suggesting that it correlated with transcriptional regulation. To test this assumption, the activities of genes harbouring H2B S112 GlcNAc near TSS were estimated by microarray analysis (Supplementary Table 2). The average profiles near TSS significantly correlated with gene activity (Fig. 4c). Moreover, the expression levels of the 1,299 genes were reliably measured, and 1,021 genes showed high expression (Supplementary Fig. 23a and Supplementary Table 3b). Moreover, gene ontology analysis revealed that there was an association of the genes harbouring H2B S112 GlcNAc to cellular metabolic processes (Supplementary Fig. 23b and Supplementary Table 3c).

Next, we analysed the genome-wide overlap of H2B S112 GlcNAc with K120 monoubiquitination. A total of 44,158 peaks of H2B K120 monoubiquitination were detected, and their average profiles near TSS were similar to those profiles of H2B S112 GlcNAc (Supplementary



**Figure 4** | GlcNAcylated H2B is associated with transcribed genes.

**a**, Polytene staining with  $\alpha$ -H2B S112 GlcNAc (green) and DAPI (blue) along with  $\alpha$ -H3K4me2 (red, top),  $\alpha$ -H3K9me2 (red, middle) or  $\alpha$ -H3K27me3 (red, bottom). **b**–**e**, ChIP-seq analysis of the H2B S112 GlcNAc and K120 monoubiquitination. The distributions of H2B S112 GlcNAc were averaged near TSS (top) and TTS (bottom) (b). The average profiles of H2B S112 GlcNAc near TSS were calculated based on the associated gene activities (c). Venn diagrams

showing overlap of the peaks (d, left), and the genes (d, middle) and the promoter (d, right) harbouring the modifications. The ChIP-seq profile surrounding the *GSK3B* gene (e). Arrowhead, position of qPCR primer. **f**, ChIP-qPCR validation in the *GSK3B* promoter. After Glc depletion, the control HeLa cells (black bar) and the OGT-knockdown cells (white bar) were replenished with Glc for 24 h. Then, the cells were subjected to ChIP with the indicated antibody and qPCR analysis. Error bars, means and s.d. ( $n = 3$ ).

Fig. 24). Among the H2B K120 monoubiquitination peaks, nearly 10% (3,974 peaks) overlapped with H2B S112 GlcNAc peaks (Fig. 4d, left), and this evaluation was confirmed by a sequential ChIP–reChIP assay (Supplementary Fig. 25). Although 5,778 genes (66.8% of H2B S112 GlcNAc and 84.4% of K120 monoubiquitination) were found at the same loci (Fig. 4d, middle, and Supplementary Table 3d), 259 genes were identified when the two peaks were compared only within the promoters (Fig. 4d, right). The results of the ChIP-seq analysis were validated by ChIP–quantitative PCR (qPCR) assessment for the glycogen synthase kinase 3 $\beta$  (*GSK3B*) gene (Fig. 4e, f). These findings suggest that at several H2B S112 GlcNAc sites, it aids H2B monoubiquitination ligase recruitment whereas at others additional or different factors may be operational.

Here we provide evidence that histone GlcNAcylation is a post-translational modification correlated with active transcriptional events, and is responsive to serum glucose levels and/or cellular energy states in certain cell types (Supplementary Fig. 1). Using an antibody that specifically recognizes the S112 GlcNAc moiety of endogenous H2B, H2B was shown to serve as an OGT substrate. We have focused on the role of H2B S112 GlcNAcylation in gene regulation (Supplementary Fig. 1). Genome-wide analysis revealed that H2B S112 GlcNAc was frequently located near transcribed genes, suggesting that histone GlcNAcylation facilitates transcription of the genes. This idea is supported by previous reports that transcriptional output driven by several transcription factors is co-activated by OGT<sup>9,18–20</sup>. However, recent papers reported that *Drosophila* OGT is itself a polycomb group protein<sup>8,21</sup>, and that many O-GlcNAcylated factors are involved in transcriptional repression and gene silencing<sup>7,8</sup>. In this respect, it will be interesting to identify other histone glycosylation sites and investigate their roles in transcriptional repression as well as activation.

## METHODS SUMMARY

**Plasmids and cell culture.** All plasmids were generated with standard protocols (see Methods). Retrovirus production, infection and sorting of the infected cells followed previously reported protocols<sup>8</sup>.

**Purification of GlcNAc proteins from chromatin.** Chromatin pellets were prepared from HeLa cells as previously described<sup>22</sup>. GlcNAc proteins were enriched with  $\alpha$ -O-GlcNAc (RL2) antibody (Abcam) immobilized on Dynabeads (Invitrogen), and released with GlcNAc-O-serine.

**Generation of monoclonal antibody.** The synthetic H2B S112 GlcNAc peptide (CKHAV S(GlcNAc) EGTK) was used to immunize mice. The hybridomas were selected by enzyme-linked immunosorbent assay (ELISA) and western blotting analysis.

**In vitro OGT and monoubiquitination assays.** Flag-OGT, Flag-E1, and Flag-BRE1A/BRE1B were purified by baculoviral systems, whereas histones and 6  $\times$  His-RAD6A were prepared from bacteria as previously reported<sup>17,23</sup>. H2B was incubated with OGT or H2B monoubiquitination ligases *in vitro*, and its modification was detected by western blotting as previously reported<sup>9,23</sup>.

**ChIP-seq and ChIP-qPCR.** ChIP and ChIP-seq library construction was performed as previously described<sup>24,25</sup>, and the libraries were sequenced to 50 base pairs (bp) with HiSeq2000 (Illumina). The fragments of interest in the libraries were quantified with specific promoter sets (Methods) by qPCR.

**Full Methods** and any associated references are available in the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

Received 16 July 2010; accepted 20 October 2011.

Published online 27 November 2011.

1. Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).

2. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693–705 (2007).
3. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
4. Berger, S. L. The complex language of chromatin regulation during transcription. *Nature* **447**, 407–412 (2007).
5. Hart, G. W., Housley, M. P. & Slawson, C. Cycling of O-linked  $\beta$ -N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446**, 1017–1022 (2007).
6. Love, D. C. & Hanover, J. A. The hexosamine signaling pathway: deciphering the 'O-GlcNAc code'. *Sci. STKE* **2005**, re13 (2005).
7. Yang, X., Zhang, F. & Kudlow, J. E. Recruitment of O-GlcNAc transferase to promoters by corepressor mSin3A: coupling protein O-GlcNAcylation to transcriptional repression. *Cell* **110**, 69–80 (2002).
8. Gambetta, M. C., Oktaba, K. & Muller, J. Essential role of the glycosyltransferase *sxc/Ogt* in polycomb repression. *Science* **325**, 93–96 (2009).
9. Fujiki, R. et al. GlcNAcylation of a histone methyltransferase in retinoic-acid-induced granulopoiesis. *Nature* **459**, 455–459 (2009).
10. Wang, Z. et al. Extensive crosstalk between O-GlcNAcylation and phosphorylation regulates cytokinesis. *Sci. Signal.* **3**, ra2 (2010).
11. Sakabe, K., Wang, Z. & Hart, G. W.  $\beta$ -N-acetylglucosamine (O-GlcNAc) is part of the histone code. *Proc. Natl Acad. Sci. USA* **107**, 19915–19920 (2010).
12. Luger, K. et al. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**, 251–260 (1997).
13. Das, C., Lucia, M. S., Hansen, K. C. & Tyler, J. K. CBP/p300-mediated acetylation of histone H3 on lysine 56. *Nature* **459**, 113–117 (2009).
14. Dang, W. et al. Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* **459**, 802–807 (2009).
15. Dong, L. & Xu, C. W. Carbohydrates induce mono-ubiquitination of H2B in yeast. *J. Biol. Chem.* **279**, 1577–1580 (2004).
16. Yoshida, Y. et al. E3 ubiquitin ligase that recognizes sugar chains. *Nature* **418**, 438–442 (2002).
17. Kim, J. et al. RAD6-Mediated transcription-coupled H2B ubiquitylation directly stimulates H3K4 methylation in human cells. *Cell* **137**, 459–471 (2009).
18. Dentin, R. et al. Hepatic glucose sensing via the CREB coactivator CRIC2. *Science* **319**, 1402–1405 (2008).
19. Chikanishi, T. et al. Glucose-induced expression of MIP-1 genes requires O-GlcNAc transferase in monocytes. *Biochem. Biophys. Res. Commun.* **394**, 865–870 (2010).
20. Jackson, S. P. & Tjian, R. O-glycosylation of eukaryotic transcription factors: implications for mechanisms of transcriptional regulation. *Cell* **55**, 125–133 (1988).
21. Sinclair, D. A. et al. *Drosophila* O-GlcNAc transferase (OGT) is encoded by the Polycomb group (PcG) gene, super sex combs (*sxc*). *Proc. Natl Acad. Sci. USA* **106**, 13427–13432 (2009).
22. Sawatsubashi, S. et al. A histone chaperone, DEK, transcriptionally coactivates a nuclear receptor. *Genes Dev.* **24**, 159–170 (2009).
23. Fujiki, R. et al. Ligand-induced transrepression by VDR through association of WSTF with acetylated histones. *EMBO J.* **24**, 3881–3894 (2005).
24. He, H. H. et al. Nucleosome dynamics define transcriptional enhancers. *Nature Genet.* **42**, 343–347 (2010).
25. Minsky, N. et al. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nature Cell Biol.* **10**, 483–488 (2008).

**Supplementary Information** is linked to the online version of the paper at [www.nature.com/nature](http://www.nature.com/nature).

**Acknowledgements** We thank A. Miyajima, S. Saito and N. Moriyama for experimental support, and M. Yamaki for manuscript preparation. We also thank Y. Maekawa, J. Seta and N. Iwasaki for support with MS. This work was supported in part by The Naito Foundation, the Astellas foundation (to R.F.), the Ministry of Education, Culture, Sports, Science and Technology (MEXT) and the Japan Society for the Promotion of Science (to R.F. and S.K.).

**Author Contributions** S.K. planned the study with H.K.; R.G.R. and M.B. provided support and general guidance; R.F. designed the study and performed the experiments with H.S. ( $\alpha$ -O-GlcNAc purification), A.Y. (LC-MS/MS), W.H. (O-GlcNAc site mapping), T.C. (*in vitro* OGT assay), S.I. (*Drosophila* analysis), Y.I., H.H.H. (ChIP-seq), F.O., J.K. (*in vitro* monoubiquitination assay), K.I. and J.K. (microarray).

**Author Information** Reprints and permissions information is available at [www.nature.com/reprints](http://www.nature.com/reprints). The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at [www.nature.com/nature](http://www.nature.com/nature). Correspondence and requests for materials should be addressed to S.K. ([uskato@mail.ecc.u-tokyo.ac.jp](mailto:uskato@mail.ecc.u-tokyo.ac.jp)).

## METHODS

**Plasmids and retroviruses.** Complementary DNAs (cDNAs) of N-terminally Flag-tagged H2B and its mutant were subcloned into pcDNA3 (Invitrogen). A series of H2B point mutants were subcloned into the pET3 vector (Novagen). shRNA sequences targeting hOGT (5'-GCACATAGCAATCTGGCTCC-3') and *Renilla* luciferase (5'-TGCGTTGCTAGTACCAAC-3', as a control) were inserted into the pSIREN-RetroQ-ZsGreen vector (Clontech). For retroviral production, the constructed shRNA vectors were transfected into PLAT-A cells. The virus contained in the medium was used for infection.

**Generation of stable cell lines.** To generate OGT-KD cells by retroviral infection,  $10^5$  cells were plated in 60 mm culture dishes, treated with 3 ml of retroviral cocktail (1 ml of the prepared retroviral solution plus 2 ml of DMEM with 10% FBS and  $8 \mu\text{g ml}^{-1}$  polybrene), then cultured for another 48 h. A FACS Vantage (BD) sorter was used to isolate the retrovirally transduced, enhanced green fluorescent protein (eGFP)-positive cells, as previously described<sup>9</sup>. To generate the cells stably expressing Flag-tagged constructs, HeLa cells were transfected with the pcDNA vectors encoding the Flag-tagged H2B or the AA mutant. The cells containing the integrated vectors were selected by exposure to  $0.5 \text{ mg ml}^{-1}$  G418.

**Generation of monoclonal antibody.** H2B S112 GlcNAc peptide (CKHAV S(GlcNAc) EGTK) was synthesized (MBL Institute) and used as an antigen (Operon Biotechnologies). The hybridomas were briefly screened using ELISA with the GlcNAc peptide, and finally selected by immunoblot analysis with the *in vitro* GlcNAcylated H2B.

**Antibodies.** Antibodies were obtained as follows:  $\alpha$ -Flag M2 agarose (Sigma),  $\alpha$ -H2A,  $\alpha$ -H2B,  $\alpha$ -H3,  $\alpha$ -H4 (Abcam),  $\alpha$ -H2B K120 monoubiquitination (Upstate),  $\alpha$ -GlcNAc (RL2 or CTD110.6) (Abcam),  $\alpha$ -OGT (Sigma),  $\alpha$ -Flag (Sigma) and  $\alpha$ -RNF20/BRE1A (Bethyl).

**Purification and identification of GlcNAc proteins.** The  $\alpha$ -O-GlcNAc-immobilized beads were prepared with  $15 \mu\text{g}$   $\alpha$ -O-GlcNAc (RL2) antibody and 0.5 ml of Dynabeads M-280 sheep  $\alpha$ -mouse IgG (Invitrogen) according to the manufacturer's instructions. Chromatin extracts from HeLa cells (0.5 g protein) were prepared essentially as previously described<sup>22</sup>. In brief, the chromatin pellet, which consisted of residual material from the nuclear extract preparation with buffers supplemented with 1 mM streptozotocin (STZ), was re-suspended with micrococcal nuclease (MNase) buffer (20 mM Tris-HCl, 1 mM  $\text{CaCl}_2$ , 2 mM  $\text{MgCl}_2$ , 0.1 M KCl, 0.1% (v/v) Triton-X, 0.3 M sucrose, 1 mM DTT, 1 mM benzamidine, 0.2 mM PMSF, 1 mM STZ, pH 7.9). After addition of  $3 \text{ U ml}^{-1}$  MNase, the samples were incubated for 30 min at room temperature with continuous homogenization and the reaction was stopped by adding 5 mM EGTA and 5 mM EDTA. After centrifugation at 2,000g for 30 min at 4 °C, the supernatant (chromatin extract) was used for the following purification steps. The chromatin extracts were passed through a WGA agarose column (Vector). The flow-through fraction was further mixed with  $\alpha$ -O-GlcNAc-immobilized beads and rotated for 8 h at 4 °C. After three washes with buffer D (20 mM Tris-HCl, 0.2 mM EDTA, 5 mM  $\text{MgCl}_2$ , 0.1 M KCl, 0.05% (v/v) NP-40, 10% (v/v) glycerol, 1 mM DTT, 1 mM benzamidine, 0.2 mM PMSF, 1 mM STZ, pH 7.9), glycoproteins were eluted twice with buffer D plus  $0.4 \text{ mg ml}^{-1}$  GlcNAc-O-serine (MBL) (elutions 1 and 2)

and finally with 0.1 M glycine-HCl (pH2.0) (elution 3). Eluted proteins were desalted by methanol-chloroform precipitation, digested with trypsin (Promega) then loaded on the automated LC-MS/MS system, which was assembled with Zaplous nano-LC (AMR) plumbed with a reverse-phase C18 electrospray ionization (ESI) column (LC assist) and a Finnigan LTQ ion-trap mass spectrometer (Thermo). The LC-MS/MS data were processed using Thermo BioWorks (Thermo) and SEQUEST (Thermo) for protein identification. The list of the identified proteins was further analysed by using the 'gene functional classification tool' in DAVID bioinformatics resources 6.7 (<http://david.abcc.ncifcrf.gov/>).

**Recombinant proteins.** Preparation of recombinant proteins was performed as previously reported<sup>23,24</sup>. Recombinant Flag-OGT, Flag-E1, Flag-BRE1A/B complexes were isolated by baculovirus expression and immunoprecipitation-based purification with  $\alpha$ -Flag M2 agarose (Sigma). Recombinant  $6 \times \text{His-RAD6A}$  was expressed in bacteria and partly isolated with a HIS-Select Nickel Affinity Gel (Sigma). The eluate was diluted 1:20 with BC0 (20 mM HEPES, 0.2 mM EDTA, 10% (v/v) glycerol, pH 7.9), and fractionated with a Resource Q column (GE Healthcare) using a linear gradient (0–0.5 M KCl) method. Preparation of recombinant *Xenopus* histone H2B and its mutants was performed as previously described<sup>23,24</sup>.

***In vitro* GlcNAcylation assay (autoradiographic analysis).** Recombinant Flag-OGT protein (0.5  $\mu\text{g}$ ) was incubated with 0.5  $\mu\text{g}$  of recombinant histone and 0.2 mM (0.2  $\mu\text{Ci}$ ) UDP- $^3\text{H}$ GlcNAc (PerkinElmer) in a 25  $\mu\text{l}$  reaction (50 mM Tris-HCl, 12.5 mM  $\text{MgCl}_2$ , 1 mM DTT, pH 7.5) for 24 h at 37 °C. The reaction was resolved with SDS-PAGE, blotted onto a polyvinylidene difluoride (PVDF) membrane, then subjected to autoradiography after spraying EN<sup>3</sup>HANCE (NEN Lifescience).

***In vitro* GlcNAcylation assay (MS analysis).** Recombinant histones (1  $\mu\text{g}$ ) or recombinant histone octamers assembled *in vitro* (1  $\mu\text{g}$ ) were GlcNAcylated by recombinant Flag-OGT in 25  $\mu\text{l}$  reactions (50 mM Tris-HCl, 2 mM UDP-GlcNAc, 12.5 mM  $\text{MgCl}_2$ , 1 mM DTT, pH 7.5) for 24 h at 37 °C. The reactions were directly subjected to a nano-LC ESI-TOF mass spectrometer system, which was assembled with a 1100 nanoLC (Agilent) plumbed with a ZORBAX 300SB-C18 column (Agilent) and micrOTOF (Bruker). Or, the reactions were digested with trypsin (Promega) and subjected to purification of glycopeptides with an MB-LAC WGA kit (Bruker). The enriched glycopeptides were loaded on the nano-LC ESI-ETD ion-trap mass-spectrometer system, which was assembled with the Agilent HP1200 Nano (Agilent) plumbed with ZORBAX 300SB-C18 (Agilent) and amaZon ETD (Bruker).

***In vitro* monoubiquitination assay.** GlcNAcylated histones (1  $\mu\text{g}$ ) were ubiquitinated with the E1 (0.1  $\mu\text{g}$ ), RAD6 (0.2  $\mu\text{g}$ ), BRE1 complex (0.5  $\mu\text{g}$ ), ubiquitin (3  $\mu\text{g}$ ) in 50 mM Tris (pH7.9), 5 mM  $\text{MgCl}_2$ , 4 mM ATP at 37 °C for 24 h.

**ChIP-seq and ChIP-qPCR.** ChIP and ChIP-seq libraries were constructed as previously described<sup>24,25</sup>. For ChIP-seq analysis, the libraries were sequenced to 50 bp with Hiseq2000 (Illumina). For ChIP-qPCR analysis, the fragments of interest in the libraries were quantified with Thermal Cycler TP800 (TAKARA) and SYBR Premix Ex Taq II (Takara). The qPCR primer sets for the GSK3B gene were 5'-TGCAAGCTCTCAGACGCTAA-3' and 5'-CTCATTCTCATGGCG TTT-3'.



Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: [www.elsevier.com/locate/ybbrc](http://www.elsevier.com/locate/ybbrc)

## Genistein promotes DNA demethylation of the steroidogenic factor 1 (SF-1) promoter in endometrial stromal cells

Hiroshi Matsukura<sup>a</sup>, Ken-ichi Aisaki<sup>b</sup>, Katsuhide Igarashi<sup>b</sup>, Yuko Matsushima<sup>b</sup>, Jun Kanno<sup>b</sup>, Masaaki Muramatsu<sup>a</sup>, Katsuko Sudo<sup>a,c</sup>, Noriko Sato<sup>a,\*</sup>

<sup>a</sup> Department of Molecular Epidemiology, Medical Research Institute, Tokyo Medical and Dental University, 2-3-10 Kanda-surugadai, Chiyoda-ku, Tokyo 101-0062, Japan

<sup>b</sup> Division of Cellular and Molecular Toxicology, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

<sup>c</sup> Animal Research Center, Tokyo Medical University, 6-1-1 Shinjuku, Shinjuku-ku, Tokyo 160-8402, Japan

### ARTICLE INFO

#### Article history:

Received 21 July 2011

Available online 29 July 2011

#### Keywords:

Genistein

DNA methylation

Ovariectomized mice

Primary culture

Steroidogenic factor 1

High-resolution melting analysis

### ABSTRACT

It has recently been demonstrated that genistein (GEN), a phytoestrogen in soy products, is an epigenetic modulator in various types of cells; but its effect on endometrium has not yet been determined. We investigated the effects of GEN on mouse uterine cells, *in vivo* and *in vitro*. Oral administration of GEN for 1 week induced mild proliferation of the endometrium in ovariectomized (OVX) mice, which was accompanied by the induction of steroidogenic factor 1 (SF-1) gene expression. GEN administration induced demethylation of multiple CpG sites in the SF-1 promoter; these sites are extensively methylated and thus silenced in normal endometrium. The GEN-mediated promoter demethylation occurred predominantly on the luminal side, as opposed to myometrium side, indicating that the epigenetic change was mainly shown in regenerated cells. Primary cultures of endometrial stromal cell colonies were screened for GEN-mediated alterations of DNA methylation by a high-resolution melting (HRM) method. One out of 20 colony-forming cell clones showed GEN-induced demethylation of SF-1. This clone exhibited a high proliferation capacity with continuous colony formation activity through multiple serial clonings. We propose that only a portion of endometrial cells are capable of receiving epigenetic modulation by GEN.

© 2011 Elsevier Inc. All rights reserved.

### 1. Introduction

Genistein (GEN), a major phytoestrogen in dietary soy, is a substantial component of the typical Asian and Western vegetarian diets, as well as recently developed infant soy milk formulas. There are several well known potential health benefits of GEN intake [1,2], one of which is an apparent decreased risk of breast and prostate cancers, based on human observational studies [1,3]. But GEN also paradoxically stimulates growth of breast cancer cells in culture [2] and uterine enlargement in rodents [4]. These effects may be mediated through estrogen receptor interactions and/or modulation of endogenous estrogen metabolism [5,6]. Since GEN can bind to estrogen receptors (ERs)  $\alpha$  and  $\beta$ , with a stronger affinity to ER $\beta$  [5], it is categorized as a phyto-selective estrogen receptor modulator (SERM) [6,7]. The variations in GEN's agonistic or antagonistic effects may be affected by variations in endogenous estrogen levels. Previous studies have not determined whether the pleiotropic effects of GEN involve distinct epigenetic alteration.

Recently, GEN was shown to alter DNA methylation in various types of cells, including ES cells [8], but most studies have been performed using cancer cell lines [9–11]. There have been few reports of the effects of GEN on DNA methylation in intact cells or *in vivo* [12]. In the present study, we utilized a uterotrophic assay in ovariectomized (OVX) mice, as a model system to analyze epigenetic regulation by GEN.

In a previous study, high-dose GEN administration to OVX rats resulted in increased uterine weight and changed endometrial cell gene expression [6]. However, no epigenetic alterations were demonstrated under this condition. We selected the steroidogenic factor 1 (SF-1; official symbol: Nr5a1) gene as a target for the methylation analysis. SF-1 is an orphan nuclear receptor and transcription factor for key enzymes involved in steroidogenesis, such as StAR, Cyp11a1 (p450scc), Cyp17a1 (p450c17), and Cyp19a1 (aromatase) [13]. The SF-1 gene is not expressed in normal endometrium; however, SF-1 expression is reactivated in the disease state of human ectopic endometriosis, in which the SF-1 promoter is abnormally demethylated by an unknown mechanism [14]. The subsequent enhancement of steroidogenic genes and resultant local steroidogenesis are proposed to be important etiologies [15]. Therefore, we hypothesized that in mouse endometrial cells,

\* Corresponding author. Fax: +81 3 5280 8058.

E-mail addresses: [hmatsukura.epi@mri.tmd.ac.jp](mailto:hmatsukura.epi@mri.tmd.ac.jp) (Hiroshi Matsukura), [nsato.epi@tmd.ac.jp](mailto:nsato.epi@tmd.ac.jp) (N. Sato).

SF-1 might be subjected to epigenetic modulation by some external stimuli. Here we show that the SF-1 promoter was demethylated *in vivo* and *in vitro* by GEN treatment. This is the first demonstration of a phytoestrogen altering the epigenetic state of adult endometrium.

## 2. Materials and methods

### 2.1. Ethics statement

All procedures described here were performed according to protocols approved by the Animal Care Committee of the National Institute of Health Sciences, and Tokyo Medical and Dental University (No. 0110306A).

### 2.2. Oral administration of genistein to ovariectomized mice

C57BL/6JmsSlc female mice (SLC) were used in this study. All mice were fed a phytoestrogen-free diet (Oriental Yeast) and were ovariectomized (OVX) 2 weeks prior to the genistein (GEN) treatment. OVX mice were divided into three different treatment groups, each consisting of 3–5 independent replicates, which orally received low-dose GEN (60 mg/kg/day), high-dose GEN (200 mg/kg/day), or vehicle (0.5% CMC-Na (Maruishi Pharmaceutical); 5 ml/kg/day) for 1 week. At the end of treatment (9 weeks of age), all mice were euthanized by exsanguination under ether anesthesia.

### 2.3. Uterotrophic assay and gene expression study after oral administration of genistein

Whole uteri were harvested, blotted, and weighted. Each uterus was divided into two horns, immediately placed into 2 ml plastic tubes of RNeasy lysis solution (Qiagen), and stored at 4 °C. From each sample, one horn was processed for mRNA expression analyses; RNeasy lysis solution was replaced with 1.0 ml of RLT buffer (Qiagen), and the horn was homogenized by addition of a 5 mm diameter Zirconium bead (Funakoshi) and shaking with a MixerMill 300 (Qiagen) at 20 Hz for 5 min (only the outermost row of the shaker box was used). Further sample preparation and analysis were performed as previously described [16]. mRNA expressions were analyzed using Affymetrix Murine Genome 430 2.0 GeneChips, and calculated as copy number per cell by the PerCellome method [16]. The second uterine horn of each sample was subjected to genomic DNA isolation.

### 2.4. Isolation of colony-forming cells derived from intact uteri

Five 8- to 9-week-old C57BL/6JmsSlc female mice (SLC) were euthanized by cervical dislocation and whole uteri were harvested. Uterine horns were collected in Dulbecco's modified Eagle's medium/Hams F-12 (DMEM/F-12; Nacalai Tesque) containing 0.05 mg/ml gentamicin (Sigma-Aldrich). Each horn was dissected longitudinally and the endometrial tissue was divided into two portions: the luminal side and the myometrium side. A single cell suspension of endometrial cells was obtained using enzymatic digestion and mechanical means adapted from Chan et al. [17]. The tissue samples were minced and dissociated in 500  $\mu$ l DMEM/F-12 containing 0.12 mg/ml (0.56 Wünsch U/ml) Blendzyme 2 and 40  $\mu$ g/ml deoxyribonuclease type I (both from Roche Applied Science) in a shaking incubator (~90 rpm) at 37 °C. At 15 min intervals, the digests were pipetted to promote separation and cell dissociation was monitored microscopically. After 45 min, debris was filtered out using a 40- $\mu$ m sieve (BD Biosciences). The single-cell suspensions were collected in DMEM/F-12 containing

10% FBS, 0.05 mg/ml gentamicin and stored on ice. Then the sieves were backwashed, and myometrial and glandular debris were further digested to single cells for 45 min as described above. All cell suspensions were filtered as described above, and combined. To remove erythrocytes, the cells were resuspended in 500  $\mu$ l of HLB solution (Immuno Biological Laboratories) and incubated for 3 min. After washing twice with PBS, viable cell numbers were counted with trypan blue (Sigma-Aldrich). Cells were seeded on gelatin (0.1%, Sigma-Aldrich)-coated dishes at various densities of  $0.1\text{--}3 \times 10^5$  cell/60-mm dish. After 14 days, non-overlapping clones were distinguished. Primary cell clones were expanded in DMEM/F-12 containing 5% FBS (SAFC Biosciences) and 0.05–0.1 mg/ml gentamicin, on gelatin-coated dishes.

### 2.5. Serial cloning of colony-forming cell clones

Self renewal was assessed by serial cloning of individual clones as described by Gargett et al. [18]. Cells were seeded on gelatin-coated 100-mm dishes at 10 cells/cm<sup>2</sup> (600 cells/100-mm dish). Culture medium was changed every 4 days and secondary clones formed distinct colonies by 14 days after plating. Secondary clones were similarly recloned to generate tertiary clones, and were also expanded in the same manner as the primary culture.

### 2.6. *In vitro* genistein exposure to colony-forming cells

From 70 isolated cell clones, we selected 20 clones from colonies that were composed of fibroblastic-shaped, homogenous cells with an average doubling time of less than 100 h. The selected clonal cells, whose passage number was less than 10, were subjected to *in vitro* GEN exposure. Cells were seeded on gelatin-coated 60-mm dishes, treated with or without 10  $\mu$ M of GEN (dissolved in dimethylsulfoxide (DMSO)) in DMEM/F-12 containing 5% FBS and 0.05 mg/ml gentamicin for 7 days. The final DMSO concentration was 0.02%. The culture medium was changed every 2 days.

### 2.7. Genomic DNA preparation and bisulfite sequencing

Genomic DNA was isolated using a QIAamp DNA Mini Kit (QIAGEN) and 180 ng–1  $\mu$ g was subjected to sodium bisulfite modification with an EpiTect Bisulfite Kit (QIAGEN) according to manufacturer's protocols. Bisulfite sequencing primers are shown in Supplementary Table 1. PCR products were cloned into the pT7 blue T vector (Novagen) and transformed into *Escherichia coli*. Plasmid DNA from positive colonies was purified and sequenced at the Tokyo Medical and Dental University Genome Laboratory (Tokyo, Japan). Sequence and statistical analyses were performed with the QUantification tool for Methylation Analysis; [http://quma.cdb.riken.jp/top/quma\\_main\\_j.html](http://quma.cdb.riken.jp/top/quma_main_j.html) [19]. The statistical significance of the difference between two bisulfite sequence groups at each CpG site was evaluated with Fisher's exact test.

### 2.8. Screening of DNA methylation status by high-resolution melting assay

All assays were performed on the LightCycler 480 using the LightCycler 480 High Resolution Melting Master kit, according to the manufacturer's instructions. Primers, designed using LightCycler Probe Design Software 2.0 (All, Roche Applied Science) are shown in Supplementary Table 1. All data were analyzed using LightCycler Gene Scanning Software.

### 2.9. Statistical analysis

Data are shown as means  $\pm$  SD. Unpaired t-tests were used to compare the significance between two groups. Statistical analysis

was performed using Dr. SPSS 2 for Windows. Results were considered statistically significant at a  $P$  value of  $<0.05$ .

### 3. Results

#### 3.1. Effects of genistein in uteri of ovariectomized (OVX) mice

OVX mice were fed with either vehicle (control) or low (60 mg/kg) or high (200 mg/kg) doses of GEN for 7 days and blotted uterus weights were determined (Fig. 1A). Compared to the control,

low-dose GEN treatment did not significantly increase the uterus weight; high-dose treatment induced a slight but significant uterus enlargement (1.4-fold of control;  $P < 0.005$ ). We then determined the mRNA expression levels of SF-1 (Fig. 1B) and steroidogenic genes (Fig. 1C–F) by the Percellome method. The mRNA levels of these genes were very low in the endometria from control and low-dose GEN-treatment groups, but were significantly increased (still less than one copy per cell on average) in the high-dose treatment group ( $P < 0.05$ ), indicating that high-doses of GEN induced expression of these genes. Next, we determined the methylation status of SF-1 in

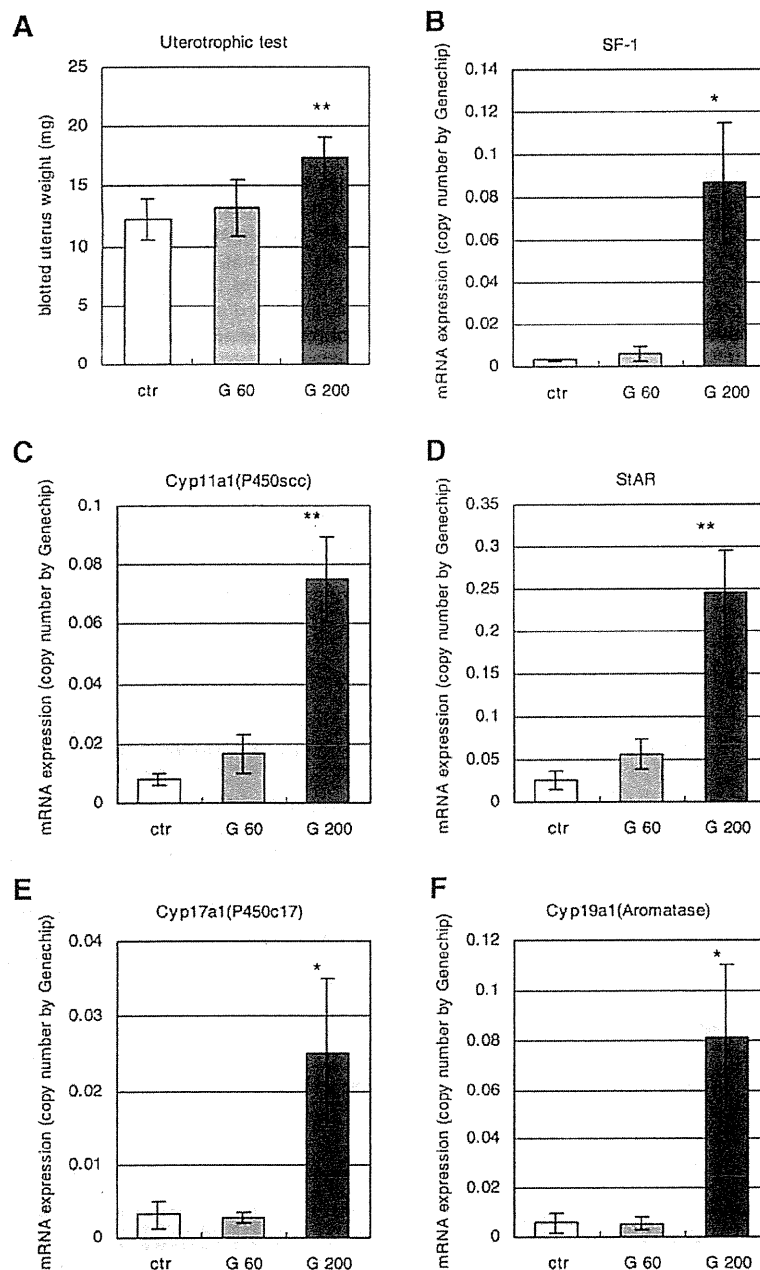


Fig. 1. Genistein induced endometrial regeneration and SF-1 mRNA expression in uterine tissue of OVX mice. (A) Blotted uterine weights were recorded. Control and GEN-treated groups comprised five and four mice, respectively. ctr; control, G 60; GEN 60 mg/kg/d, G200; GEN 200 mg/kg/d. (B–F) mRNA expressions of steroidogenic genes were determined using GeneChip analysis and were calculated by the Percellome method. Y axis indicates mRNA expression as copy number per cell. (B) SF-1 was determined by 1418315\_at, (C) Cyp11a1 by 1439947\_at, (D) StAR by 1418729\_at, (E) Cyp17a1 by 1417017\_at, and (F) Cyp19a1 by 1449920\_at. \*Statistically significant at  $P < 0.05$ . \*\*Statistically significant at  $P < 0.005$ .



the total uterus tissues of the three groups of mice by bisulfite sequencing (Fig. 2A–C). There were 15 CpG sites spanning –272 to +199 of the promoter and the 5'-UTR (exon 1) region of SF-1. The percentages of total methylated CpG sites in this region, in control, and low and high-dose GEN, were 78.3%, 73.9%, and 54.4%, respectively, indicating that GEN dose-dependently induced demethylation in this region. Most CpG sites in the 5'-UTR (exon1) were demethylated by high-dose GEN. In particular, the methylation levels of 5 CpG sites between +45 and +89 were significantly lower in high-dose GEN than in control (Fisher's exact test,  $P < 0.01$ ). We further split the endometrium to separate the luminal side (LU) from the basilar myometrial side (MY), and both specimens were separately subjected to bisulfite sequencing. This procedure was applied to the samples from the GEN-treated groups but not to control samples due to uterus atrophy. In low-dose GEN treated mice, the mean methylation levels of LU and MY were 84.8% and 65.0%, respectively (not shown). In high-dose GEN treated mice, the mean methylation levels of LU and MY were 42.1% and 66.7%, respectively (Fig. 2D). Thus, the demethylation induced by high-dose GEN occurred predominantly in the LU, rather than in the MY.

3.2. Effect of genistein on primary endometrial cell culture

In order to study the SF-1 promoter methylation at the cellular level, we employed an endometrial cell primary culture. Intact

murine endometrium were divided into LU and MY portions, and cells were separately isolated. Primary cell clones were established by colony-formation, following the plating of a serially-titrated cell suspension (see Section 2). Efficient isolated colony formation was achieved with cells seeded at a density of 4,500–15,300 cells/cm<sup>2</sup>. For LU and MY, the average frequencies of colony appearance were 7.5 per 10<sup>5</sup> cells and 15 per 10<sup>5</sup> cells, respectively. The growth curves of representative clones derived from LU and MY are shown in Fig. 3. Cell clones with highest and lowest proliferative activities were obtained from LU and MY, separately. More highly proliferative cells were obtained from MY than from LU (Fig. 3A and C). We selected 20 highly proliferative clones for further study (see Section 2). Two rapid growing clones obtained from MY showed self-renewal activity when secondarily seeded at a very low cell density (10 cells/cm<sup>2</sup>) (not shown). To screen for primary cultured cells that responded to GEN, we set up a high-resolution melting (HRM) assay that identified region-specific methylation levels. The region analyzed by HRM assay exclusively contained the 7 CpG sites between +19 and +89 bp that were most differentially demethylated following oral administration of GEN (Fig. 2). Each clone was treated with or without GEN for 1 week; cells treated with a similar concentration of DMSO served as control. Among the 20 clones that we screened, only one GEN-treated clone (No. 16) exhibited a significant shift in the melting curve compared to control cells (Fig. 4A). This clone had the highest proliferation

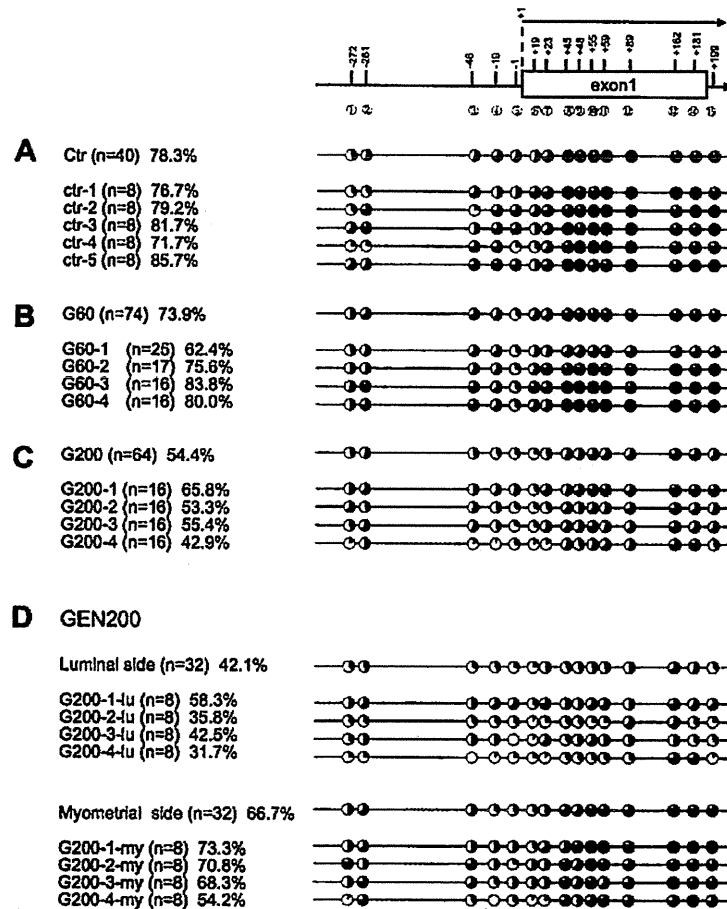
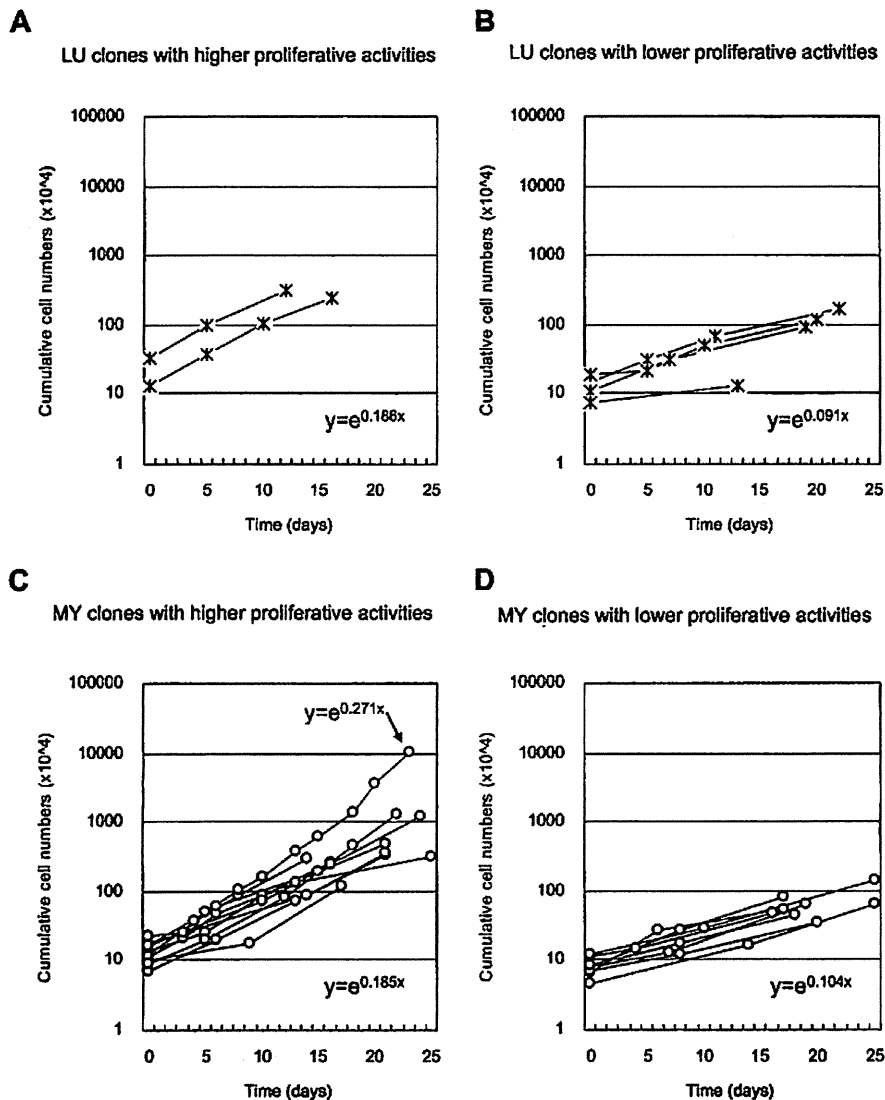


Fig. 2. Genistein induced demethylation of the SF-1 gene promoter in endometrial tissues of OVX mice. The schematic diagram indicates CpG locations on the SF-1 promoter region, spanning 5'-flanking to exon 1. The CpG position relative to the first base of exon 1 (+1) is shown. The bisulfite sequencing fragment contains 15 CpG sites. The black inlay represents the mean methylation levels of each CpG, and the left panel contains the number of sequenced clones and the mean methylation level of all CpGs. Summarized methylation results are shown at the top. (A–C) Difference in methylation levels between (A) vehicle (control), (B) low-dose GEN (60 mg/kg/d), and (C) high-dose GEN (200 mg/kg/d) exposed uteri. (D) Difference in methylation levels between luminal and myometrial sides of a uterus exposed to high-dose GEN.



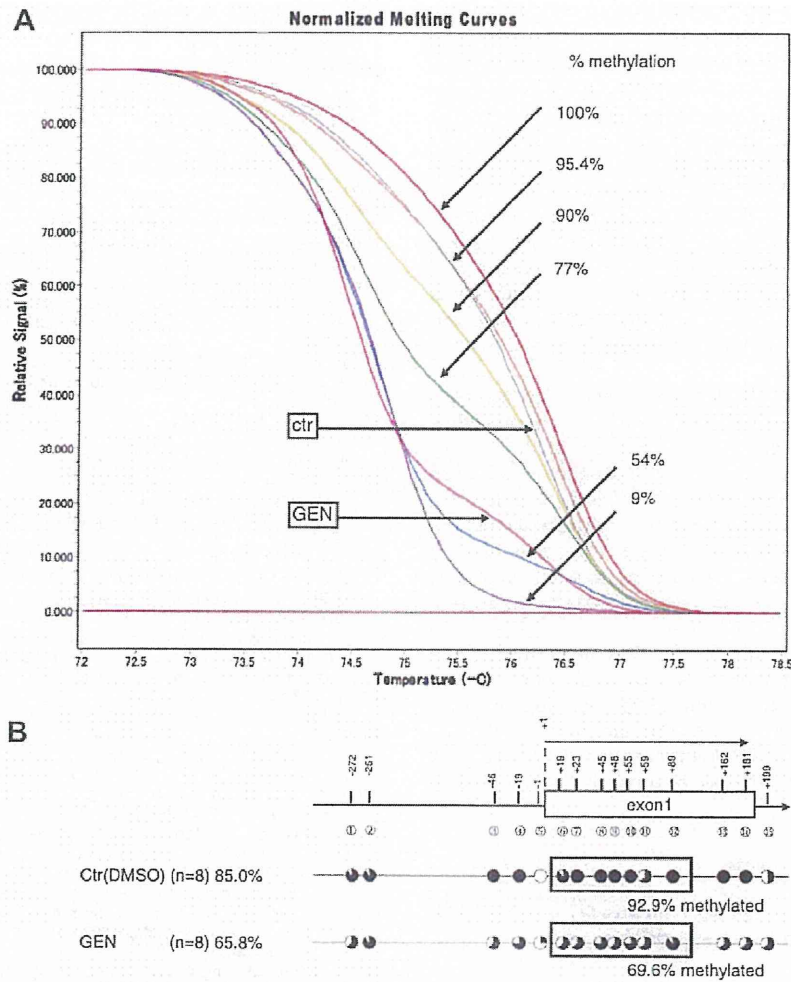
**Fig. 3.** Proliferation properties of isolated clones: one clone showed highly proliferative activity. Representative growth curves of clones harvested from intact murine endometrium are shown. Formula indicates the slope of fitted growth curves. Clones were divided into two groups: high and low proliferative (see Section 2). A total of 20 highly proliferative clones were analyzed by HRM after *in vitro* GEN treatment. (A, B) Asterisks indicate representative growth curves of cell clones isolated from luminal side. (A) Clones with higher proliferative activities, derived from luminal side. Six highly proliferative clones were obtained in total; 2 representatives are shown. (B) Clones with lower proliferative activities, derived from luminal side. (C, D) Each circle indicates representative growth curves of cell clones isolated from myometrial side. (C) Clones with higher proliferative activities, derived from the myometrial side. In total, 14 highly proliferative clones were obtained; 10 representatives are shown. The arrow indicates the clone with the most rapid growth. (D) Clones with lower proliferative activities, derived from myometrial side.

activity (Fig. 3C, arrow). GEN treatment of the other clones did not result in significant changes to the melting curve patterns (not shown). We further confirmed the methylation status of clone No. 16 by bisulfite sequencing. The percentages of CpG methylation in the SF-1-272 to +199 promoter regions for untreated and GEN-treated cells were 85.0% and 65.8%, respectively (Fig. 4B).

#### 4. Discussion

Growing evidence suggests that the manner in which nutrients can either help maintain health, or conversely, promote disease development may be mediated by epigenetic regulation [12,20]. However, relatively little is known about tissue-specific sensitivity or how much plasticity exists in regards to the effect that a given environmental factor can exert on a certain epigenetic target

[20,21]. GEN, a non-nutrient dietary component of soy products, exhibits mixed estrogen agonist and antagonist properties, and multiple functions both *in vivo* and *in vitro* [7,22]. Several animal studies have demonstrated that GEN acts as an epigenetic modulator [20]. We focused on the effects of GEN on endometrium, because endometrium is not only hormone responsive, but also a highly proliferative organ. Epigenetic alterations of proliferative tissue or cells may then be expanded through tissue proliferation. We used OVX rodents, which are a widely used model for studying estrogen withdrawal and replacement [23], as well as for the assessment of endocrine-disrupting chemicals in the environment [4]. In our experiment, GEN induced proliferation of the endometrium and increased uterine weight (Fig. 1A) to extents similar to those previously reported in OVX rats [4]. Our findings also suggested that GEN treatment induced marked demethylation of



**Fig. 4.** Genistein induced demethylation of the SF-1 gene in the cell clone that showed the highest proliferative activity. (A) The GEN-mediated demethylation of CpGs (at the positions of +19, 23, 45, 48, 55, 59, and 89) in the SF-1 gene *in vitro* in one out of 20 isolated clones (Fig. 3A and C). HRM analysis enabled to clear separation of the PCR product mixtures with different CpG compositions. The methylation standards were prepared as 100% (red), 95.4% (orange), 90% (yellow), 77% (green), 54% (blue), and 9% (violet) methylated template in demethylated background. Grey and pink indicate the control and GEN-treated clone sample, respectively. (B) SF-1 DNA methylation status, obtained from bisulfate sequencing, of the clone that showed demethylation in HRM assay. Percentage under box (HRM region) indicates percent methylation of HRM region. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

CpG sites in the SF-1 promoter (Fig. 2), and substantial increase in SF-1 mRNA level (Fig. 1B). Expressions of genes downstream of transcription factor SF-1 were also significantly enhanced (Fig. 1C–F). However, it should be noted that the induced mRNA levels were still very low (Fig. 1B–F), less than one copy per cell, as determined by the Percollome method [16]. Our results are consistent with those of a previous study reporting that a physiological concentration of GEN increased Cyp19a1 enzymatic activity in endometrial cells derived from a normal uterus, whereas GEN did not affect Cyp19a1 activity in a cell-free assay [24]. It is unknown whether GEN stimulated Cyp19a1 activity by epigenetic modulation.

In the present study, we also identified primary cultured endometrial cells that were competent for epigenetic regulation by GEN, which were present at a very low frequency. In our *in vitro* study, GEN treatment did not enhance, but rather inhibited the proliferation of colony-derived cells. Taken together, these findings indicate that a minor population of endometrium cells can respond to GEN and that, in these cells, GEN induces demethylation and

activation of SF-1, followed by the induction of the SF-1 steroidogenic cascade. This might lead to local steroidogenesis and enhanced endometrium proliferation *in vivo* in GEN-treated OVX mice. Since demethylation of the SF-1 promoter was observed at the whole tissue level (Fig. 2), and the induced expression of steroidogenic genes occurred only to a subcellular amount (Fig. 1B–F), the demethylation event in each cell may not be sufficient for the SF-1 induction.

The demethylation in the endometrium that occurred after 1 week of treatment with high-dose GEN was more prominent in the LU than in the MY. Following GEN treatment, the methylation level in the MY was similar to that in the untreated endometrium as a whole. After 7 days of GEN exposure, the endometrial cells in the LU of OVX mice were composed of regenerated cells moving from MY to LU; in light of this, our results indicate that the initial epigenetic alteration might be expanded through proliferation of regenerating cells. This is also consistent with our observation that more endometrial cells derived from MY showed higher colony-formation activity and rapid proliferation than did those derived

from LU (Fig. 3). We thus speculate that there are GEN-sensitive cells in the MY, which might contain endometrial stem cells [25].

Recently, human endometrial stem cells were identified; they reside in endometrial stromal tissue and possess fibroblastic-shape and self-renewal ability, thus forming large, densely packed, homogenous colonies [17,18]. Some candidates for murine endometrial progenitor cells have been suggested to reside in the luminal epithelial or area of adjacent to the myometrium [25]. The rapidly growing endometrial cell that we obtained had a fibroblastic-shape, formed large, relatively homogenous, densely packed colonies, and showed self-renewal activity when seeded at a very low cell density (data not shown). Although there are differences between the species, we speculate that our rapid growing cell clones may correspond to the human endometrial stromal progenitor cells [17].

Aberrant SF-1 expression with lack of promoter CpG methylation has been reported in ectopic endometriosis [14], and thus endometriosis is now considered an epigenetic disease [26]. Endometriosis is classically defined as the growth of endometrial tissue at extrauterine sites; it has been suggested that each endometriotic lesion originates from a single epigenetically deregulated endometrial progenitor cell [27]. Further studies are required to identify cells which are competent for epigenetic changes following GEN exposure, and to elucidate the relationships between these cells, GEN, and endometriosis.

In conclusion, we demonstrated that GEN demethylates the promoter region of the SF-1 gene. This is the first demonstration of phytoestrogen participation in epigenetic alterations in adult endometrial tissue. These findings are important from standpoints of nutrition, public health, and disease prevention. Further study is warranted to characterize the nature of the cells that respond to GEN in the endometrium.

#### Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2011.07.104.

#### References

- [1] H. Adlercreutz, Phyto-oestrogens and cancer, *Lancet Oncol.* 3 (2002) 364–373.
- [2] M. Messina, W. McCaskill-Stevens, J.W. Lampe, Addressing the soy and breast cancer relationship: review, commentary, and workshop proceedings, *J. Natl. Cancer Inst.* 98 (2006) 1275–1284.
- [3] A.H. Wu, R.G. Ziegler, A.M. Nomura, D.W. West, L.N. Kolonel, P.L. Horn-Ross, R.N. Hoover, M.C. Pike, Soy intake and risk of breast cancer in Asians and Asian Americans, *Am. J. Clin. Nutr.* 68 (1998) 1437S–1443S.
- [4] J. Kanno, L. Onyon, S. Peddada, J. Ashby, E. Jacob, W. Owens, The OECD program to validate the rat uterotrophic bioassay. Phase 2: dose-response studies, *Environ. Health Perspect.* 111 (2003) 1530–1549.
- [5] K. Pettersson, J.A. Gustafsson, Role of estrogen receptor beta in estrogen action, *Annu. Rev. Physiol.* 63 (2001) 165–192.
- [6] P. Diel, T. Hertrampf, J. Seibel, U. Laudendach-Leschowsky, S. Kolba, G. Vollmer, Combinatorial effects of the phytoestrogen genistein and of estradiol in uterus and liver of female Wistar rats, *J. Steroid Biochem. Mol. Biol.* 102 (2006) 60–70.
- [7] V. Beck, U. Rohr, A. Jungbauer, Phytoestrogens derived from red clover: an alternative to estrogen replacement therapy?, *J. Steroid Biochem. Mol. Biol.* 94 (2005) 499–518.
- [8] N. Sato, N. Yamakawa, M. Masuda, K. Sudo, I. Hatada, M. Muramatsu, Genome-wide DNA methylation analysis reveals phytoestrogen modification of promoter methylation patterns during embryonic stem cell differentiation, *PLoS One* 6 (2011) e19278.
- [9] S.M. Meeran, A. Ahmed, T.O. Tollefsbol, Epigenetic targets of bioactive dietary components for cancer prevention and therapy, *Clin. Epigenetics* 1 (2010) 101–116.
- [10] S. Majid, A.A. Dar, V. Shahryari, H. Hirata, A. Ahmad, S. Saini, Y. Tanaka, A.V. Dahiya, R. Dahiya, Genistein reverses hypermethylation and induces active histone modifications in tumor suppressor gene B-Cell translocation gene 3 in prostate cancer, *Cancer* 116 (2010) 66–76.
- [11] A.K. Jha, M. Nikbakht, G. Parashar, A. Shrivastava, N. Capalash, J. Kaur, Reversal of hypermethylation and reactivation of the RARbeta2 gene by natural compounds in cervical cancer cell lines, *Folia. Biol. (Praha)* 56 (2010) 195–200.
- [12] Y. Li, T.O. Tollefsbol, Impact on DNA methylation in cancer prevention and therapy by bioactive dietary components, *Curr. Med. Chem.* 17 (2010) 2141–2151.
- [13] K. Morohashi, S. Honda, Y. Inomata, H. Handa, T. Omura, A common trans-acting factor, Ad4-binding protein, to the promoters of steroidogenic P-450s, *J. Biol. Chem.* 267 (1992) 17913–17919.
- [14] Q. Xue, Z. Lin, P. Yin, M.P. Milad, Y.H. Cheng, E. Confino, S. Reierstad, S.E. Bulun, Transcriptional activation of steroidogenic factor-1 by hypomethylation of the 5' CpG island in endometriosis, *J. Clin. Endocrinol. Metab.* 92 (2007) 3261–3267.
- [15] S.E. Bulun, H. Utsunomiya, Z. Lin, P. Yin, Y.H. Cheng, M.E. Pavone, H. Tokunaga, E. Trukhacheva, E. Attar, B. Gurates, M.P. Milad, E. Confino, E. Su, S. Reierstad, Q. Xue, Steroidogenic factor-1 and endometriosis, *Mol. Cell. Endocrinol.* 300 (2009) 104–108.
- [16] J. Kanno, K. Aisaki, K. Igarashi, N. Nakatsu, A. Ono, Y. Kodama, T. Nagao, "Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays, *BMC Genomics* 7 (2006) 64.
- [17] R.W. Chan, K.E. Schwab, C.E. Gargett, Clonogenicity of human endometrial epithelial and stromal cells, *Biol. Reprod.* 70 (2004) 1738–1750.
- [18] C.E. Gargett, K.E. Schwab, R.M. Zillwood, H.P. Nguyen, D. Wu, Isolation and culture of epithelial progenitors and mesenchymal stem cells from human endometrium, *Biol. Reprod.* 80 (2009) 1136–1145.
- [19] Y. Kumaki, M. Oda, M. Okano, QUMA: quantification tool for methylation analysis, *Nucleic Acids Res.* 36 (2008) W170–W175.
- [20] J.A. McKay, J.C. Mathers, Diet induced epigenetic changes and their implications for health, *Acta Physiol. (Oxf)* (2011) 103–118.
- [21] B.C. Christensen, E.A. Houseman, C.J. Marsit, S. Zheng, M.R. Wrensch, J.L. Wiemels, H.H. Nelson, M.R. Karagas, J.F. Padbury, R. Bueno, D.J. Sugarbaker, R.F. Yeh, J.K. Wiencke, K.T. Kelsey, Aging and environmental exposures alter tissue-specific DNA methylation dependent upon CpG island context, *PLoS Genet.* 5 (2009) e1000602.
- [22] C.B. Klein, A.A. King, Genistein genotoxicity: critical considerations of in vitro exposure dose, *Toxicol. Appl. Pharmacol.* 224 (2007) 1–11.
- [23] G. Rimoldi, J. Christoffel, D. Seidlova-Wuttke, H. Jarry, W. Wuttke, Effects of chronic genistein treatment in mammary gland, uterus, and vagina, *Environ. Health Perspect.* 115 (Suppl. 1) (2007) 62–68.
- [24] K.M. Edmunds, A.C. Holloway, D.J. Crankshaw, S.K. Agarwal, W.G. Foster, The effects of dietary phytoestrogens on aromatase activity in human endometrial stromal cells, *Reprod. Nutr. Dev.* 45 (2005) 709–720.
- [25] C.E. Gargett, H. Masuda, Adult stem cells in the endometrium, *Mol. Hum. Reprod.* 16 (2010) 818–834.
- [26] S.W. Guo, Epigenetics of endometriosis, *Mol. Hum. Reprod.* 15 (2009) 587–607.
- [27] Y. Wu, Z. Basir, A. Kajdacsy-Balla, E. Strawn, V. Macias, K. Montgomery, S.W. Guo, Resolution of clonal origins for endometriotic lesions using laser capture microdissection and the human androgen receptor (HUMARA) assay, *Fertil. Steril.* 79 (Suppl. 1) (2003) 710–717.