

Supplementary Information

Drug-induced cis-regulatory elements in human hepatocytes affect molecular phenotypes associated with adverse reactions

Saki Gotoh-Saito¹, Ryoko Wada¹, Tomoe Nishimura², Hideya Kawaji^{1,3*}

¹ Research Center for Genome & Medical Sciences, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

² Center for Basic Technology Research, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan

³ Graduate School of Medical and Dental Sciences, Institute of Science Tokyo, Tokyo, Japan.

*Correspondence

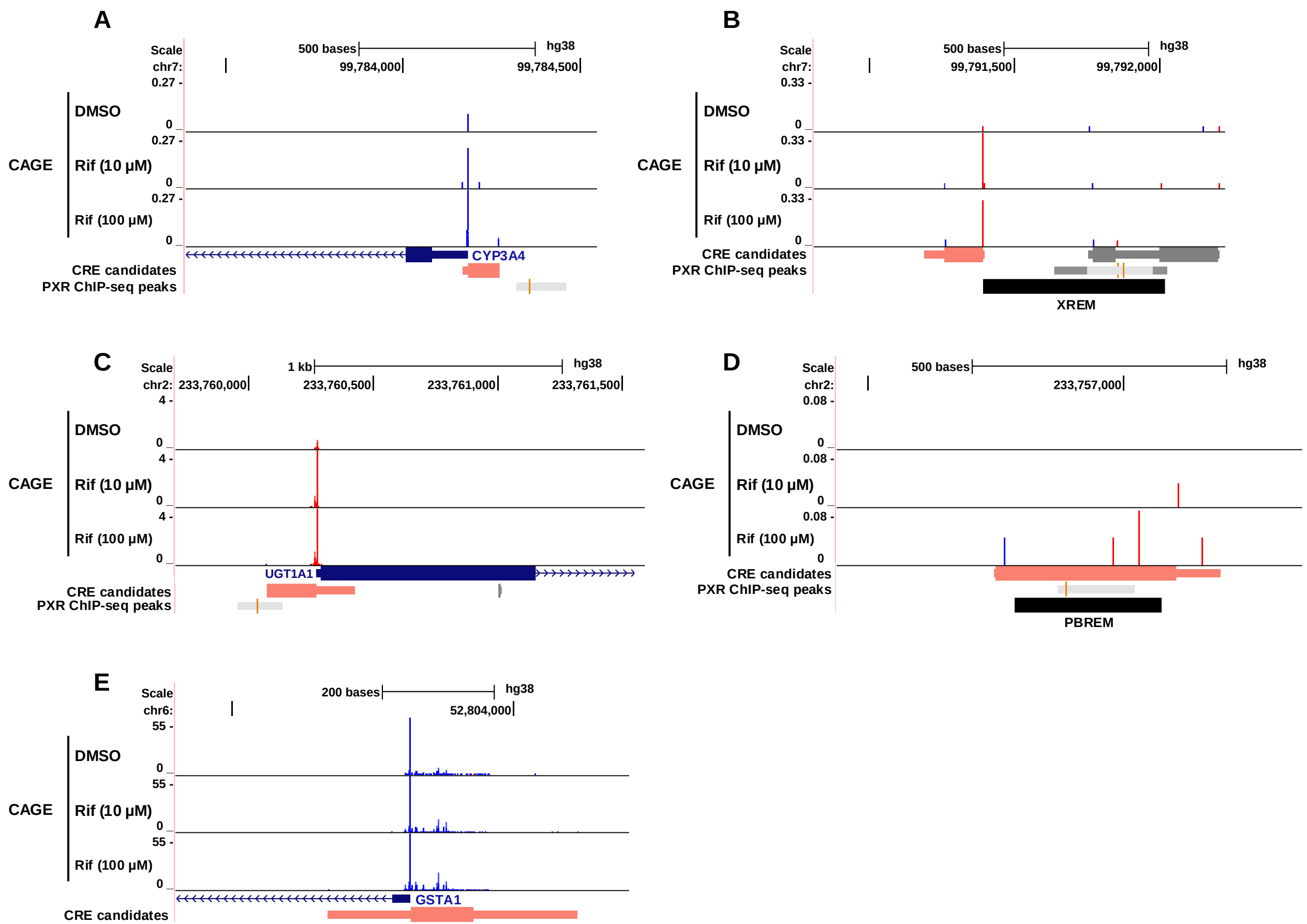
Hideya Kawaji

Research Center for Genome & Medical Sciences, Tokyo Metropolitan Institute of Medical Science

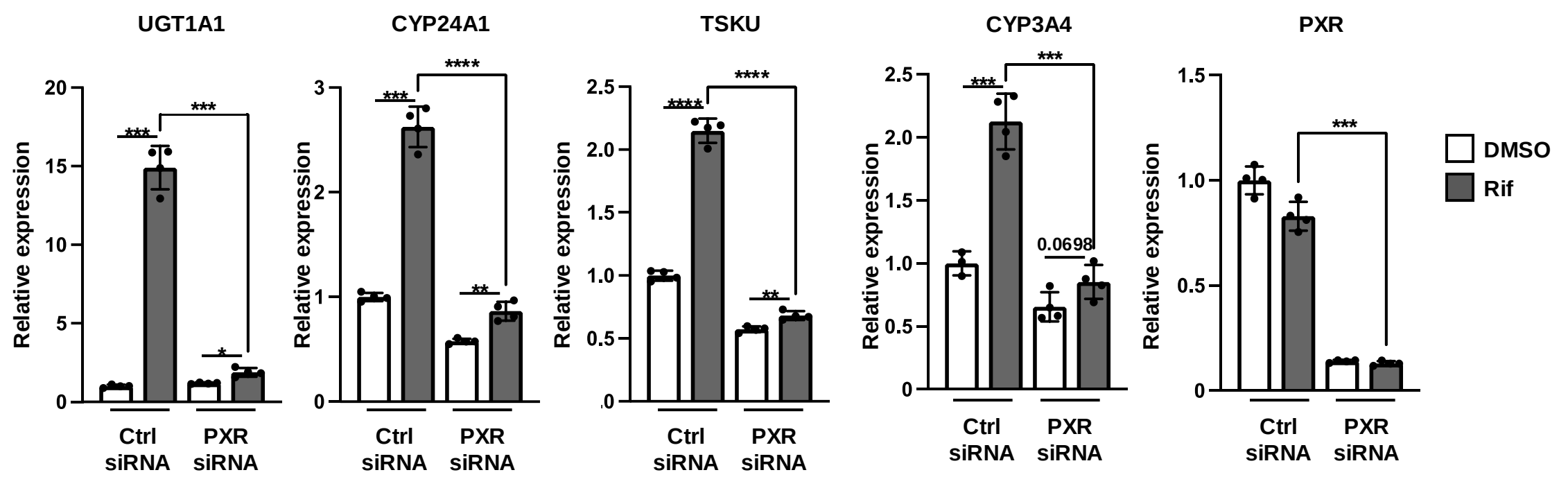
Address: 2-1-6 Kamikitazawa, Setagaya-ku, Tokyo 156-8506, JAPAN

TEL: +81-3-5316-3128

E-MAIL: kawaji-hd@igakuken.or.jp

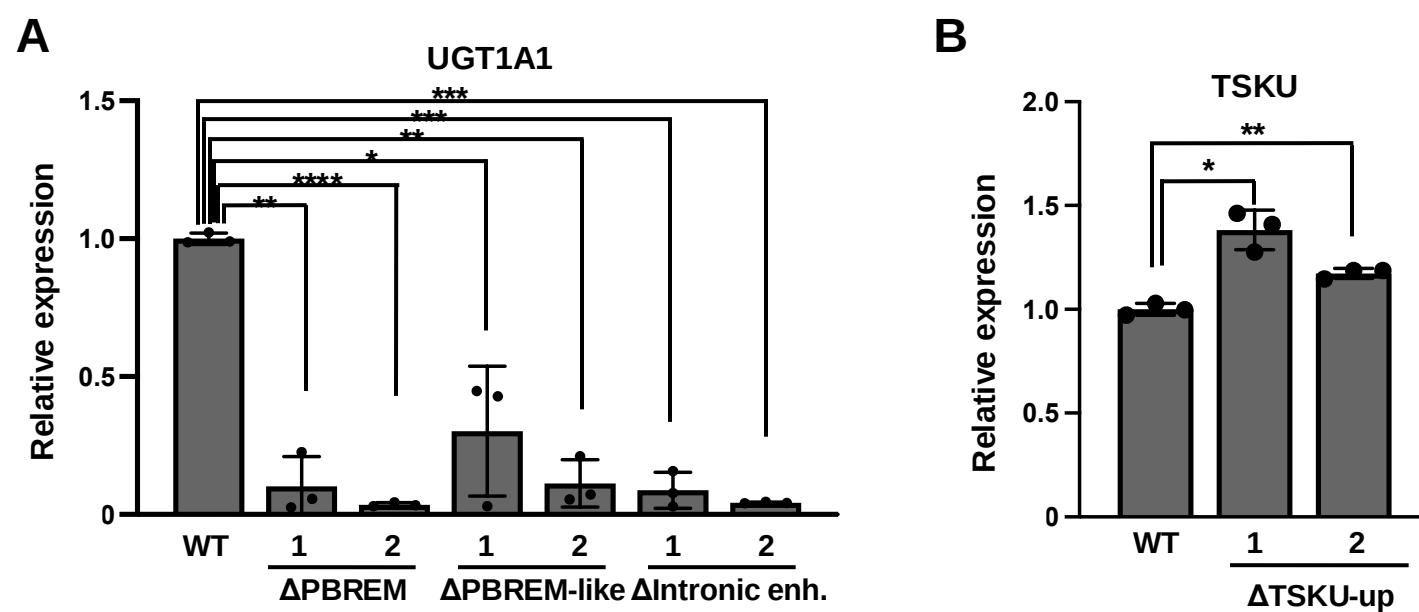


Supplementary Figure 1
 A view of the promoters, *CYP3A4* (A), *UGT1A1* (C), and *GSTA1* (E) and enhancers, XREM (B) and PBREM (D) in UCSC Genome Browser. The tracks for cap analysis of gene expression (CAGE) indicate the frequencies of monitored transcription start sites at single base-pair (bp) resolution in transcripts per million, where red and blue indicate the orientation of transcription (red: forward, blue: reverse).



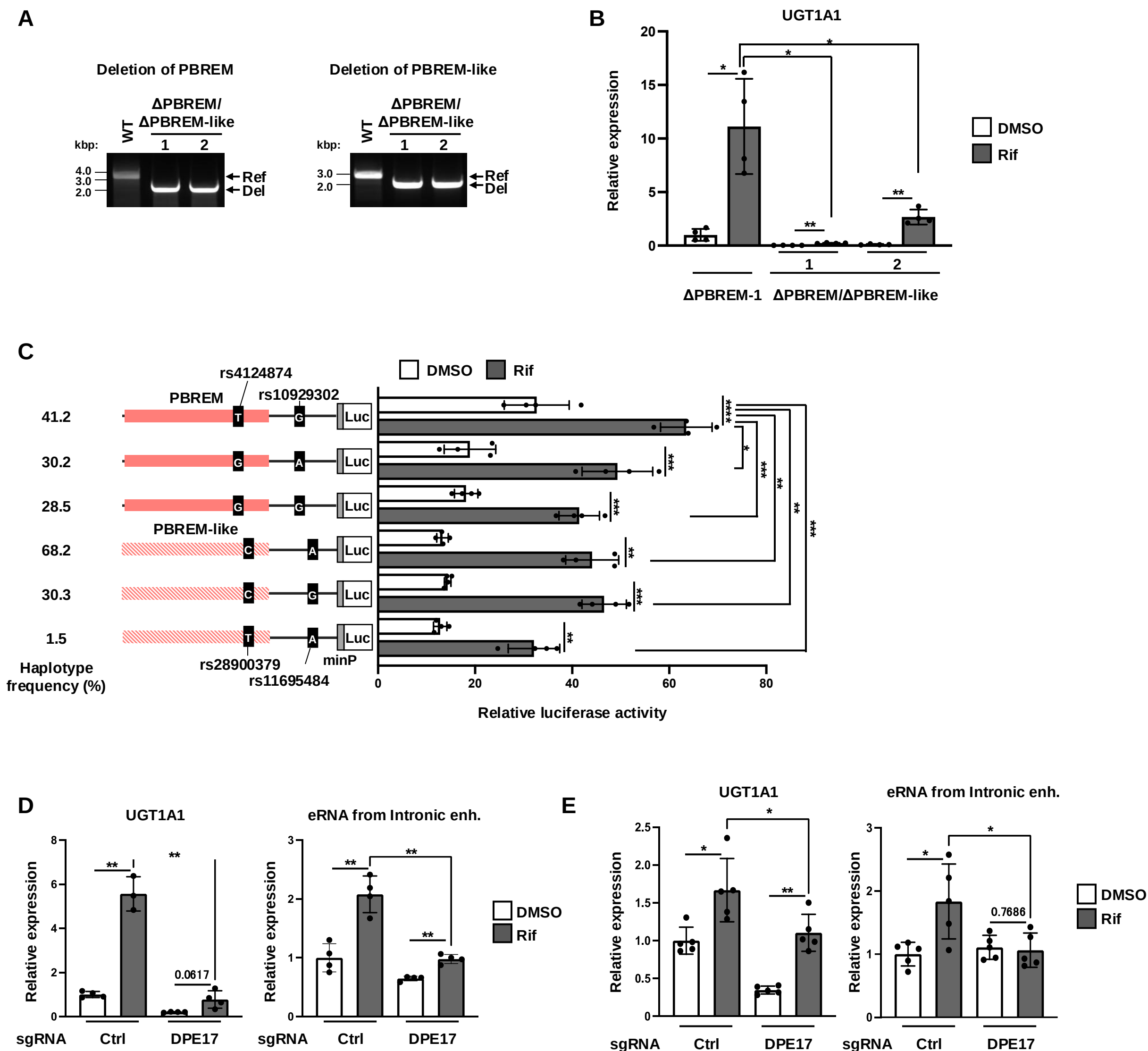
Supplementary Figure 2

qRT-PCR analysis of UGT1A1, CYP24A1, TSKU, CYP3A4, and PXR mRNA in PXR siRNA-transfected ShP51 cells. The experiments were performed using independent transfections of siRNAs (n=4). Data are presented as mean values $\pm$ SD, and a two-sided unpaired Welch's *t*-test was used to calculate the *P* values. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ and **** $P < 0.0001$. Source data are provided as a Source Data file.



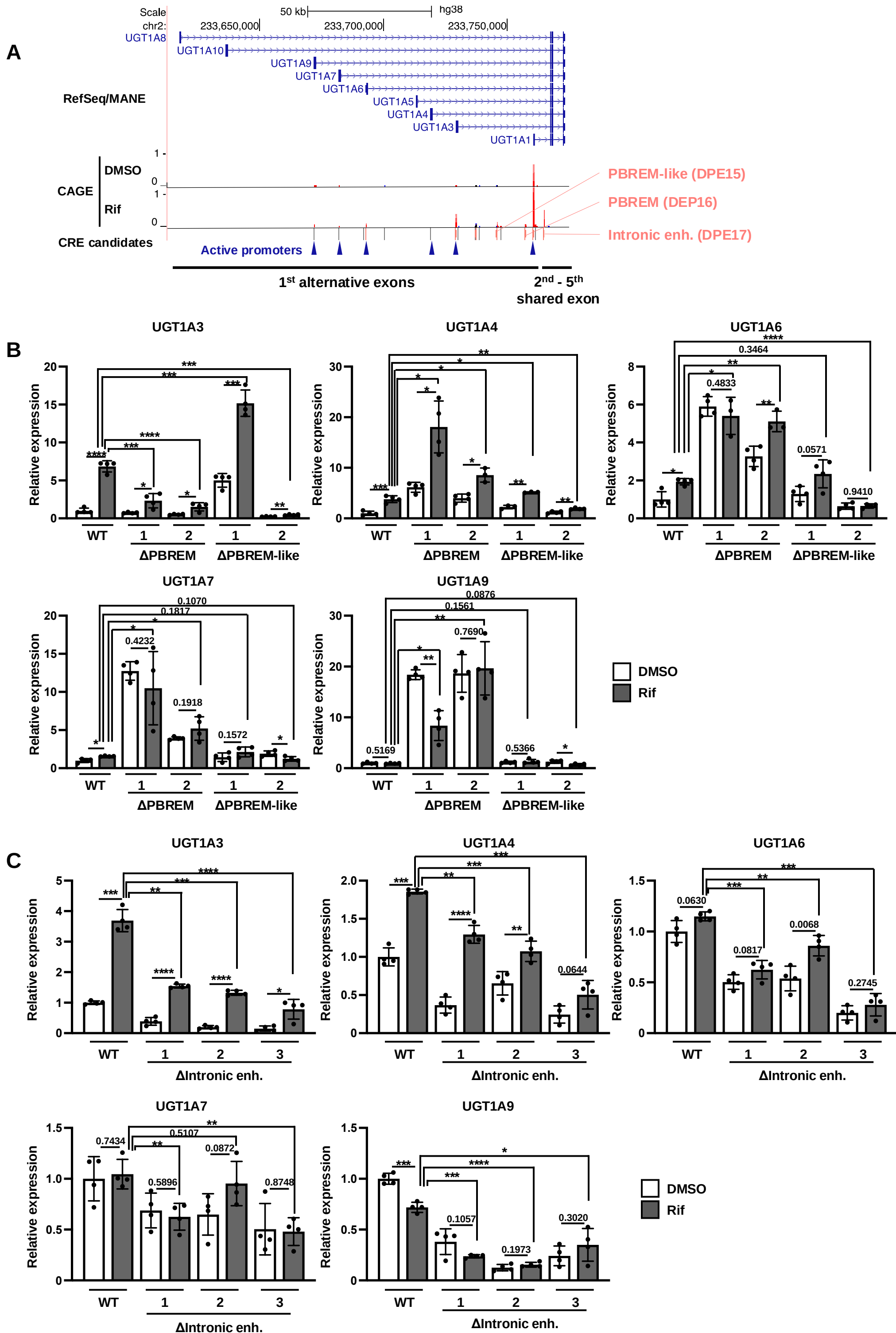
Supplementary Figure 3

Protein expression levels of UGT1A1 and TSKU in each enhancer deletion mutants, determined by LC-MS/MS analyses. (A) UGT1A1 level in rifampicin (Rif)-treated Δ PBREM 1, 2, Δ PBREM-like 1, 2 and Δ Intronic enh. 1-3. (B) TSKU level in rifampicin (Rif)-treated Δ TSKU-up 1, 2. The experiments were performed using independent treatments (n=3). Data are presented as mean values $\pm$ SD, and a two-sided unpaired Welch's *t*-test was used to calculate the *P* values. **P* < 0.05, ***P* < 0.01, ****P* < 0.001 and *****P* < 0.0001. Source data are provided as a Source Data file.



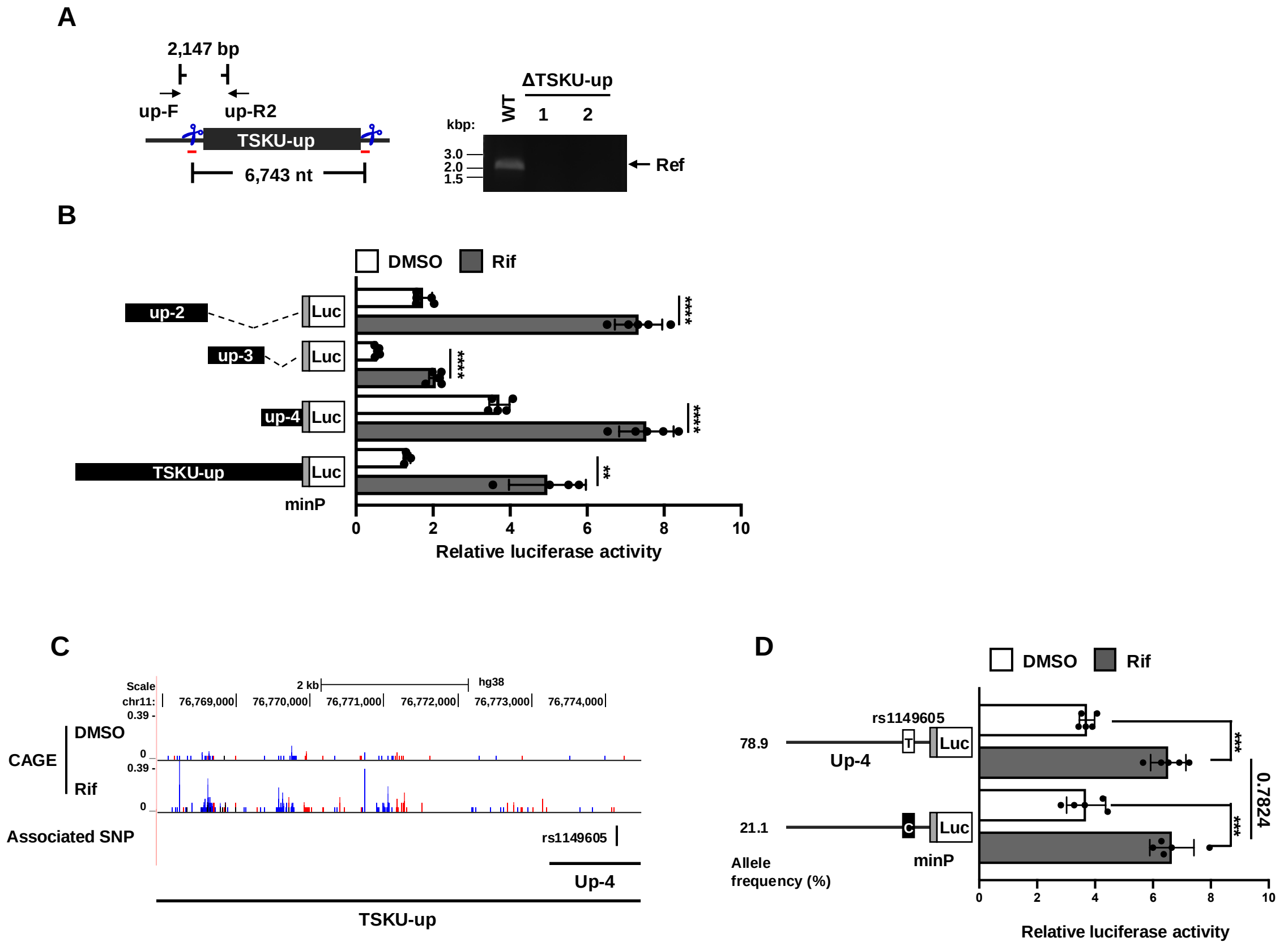
Supplementary Figure 4

A Gel images show the PCR products from genomic DNA extracted from the deletion mutants Δ PBREM/ Δ PBREM-like 1 and 2. Deletion of PBREM (left), Deletion of PBREM-like (right) **B** qRT-PCR analysis of UGT1A1 mRNA normalized to GAPDH mRNA in rifampicin (Rif)-treated Δ PBREM-1 and Δ PBREM/ Δ PBREM-like 1 and 2 cells with independent treatments (n=4). **C** Luciferase activity of -3,490/-3,060 (PBREM, DPE16) region or -14,848/-14,415 (PBREM-like, DPE15) region from *UGT1A1* transcription start site with minimal promoter (minP) in Rif-treated ShP51 cells with independent transfections (n=4). There are three major haplotypes with the frequencies in the population respectively. **D** qRT-PCR analysis of UGT1A1 mRNA and eRNA transcribed from *UGT1A1* Intronic enh. (DPE17) in ShP51 cells expressing dCas9-KRAB and sgRNAs targeting the DPE17 with independent treatments of Rif and DMSO (n=4). **E** The same qRT-PCR analysis in differentiated HepaRG cells with independent treatments (n=5). Data are presented as mean values $\pm$ SD, and a two-sided unpaired Welch's *t*-test was used to calculate the *P* value. **P* < 0.05, ***P* < 0.01, ****P* < 0.001 and *****P* < 0.0001. Source data are provided as a Source Data file.



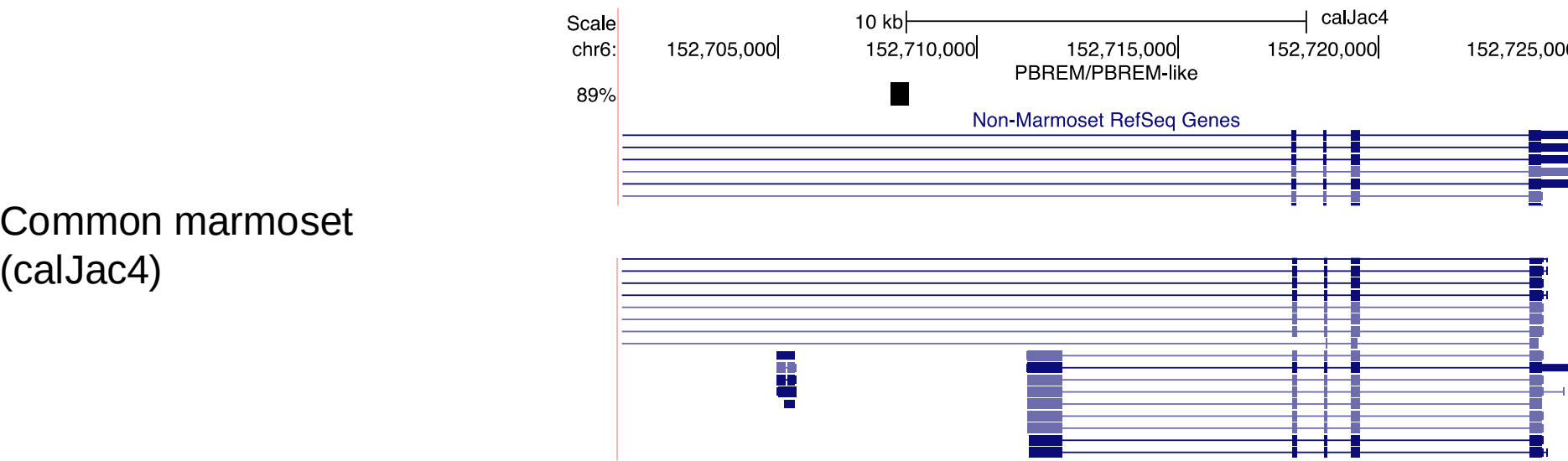
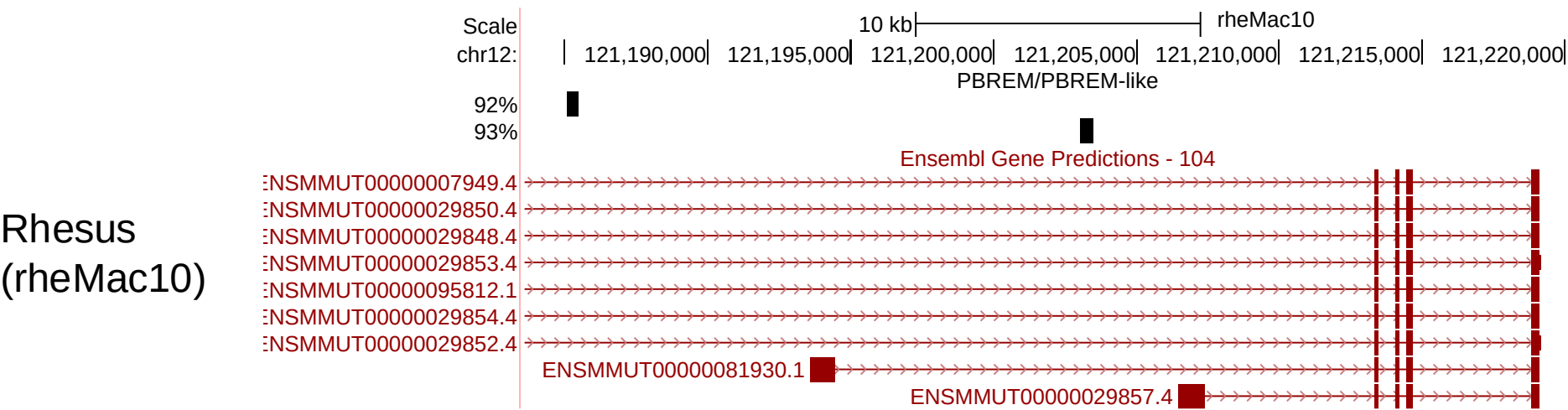
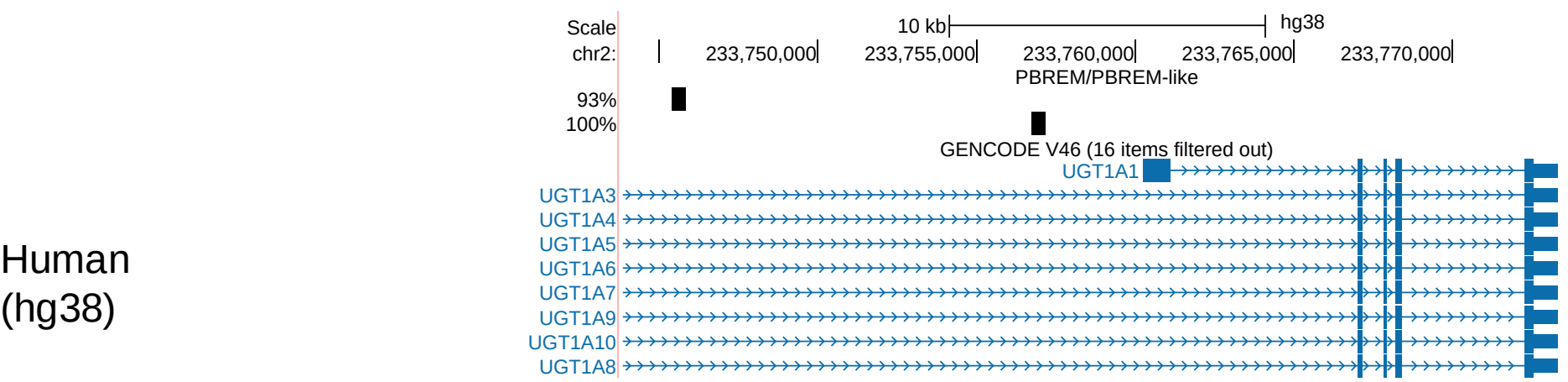
Supplementary Figure 5

A A view of *UGT1A* locus in UCSC Genome Browser. The locus includes unique alternate first exons followed by four common exons. DPE15 (PBREM-like), DPE16 (PBREM), and DPE17 (Intronic Enh) are the *cis*-regulatory element (CRE) candidates. **B**, **C**. qRT-PCR analysis of mRNA levels of *UGT1A* isoforms (UGT1A3, 1A4, 1A6, 1A7 and 1A9) in rifampicin (Rif)-treated Δ PBREM 1, 2 and Δ PBREM-like 1, 2 (**B**) or Rif-treated Δ Intronic enh. 1-3 (**C**). The expression levels are normalized to GAPDH mRNA. The experiments were performed using independent treatments (n=4). Data are presented as mean values $\pm$ SD, and a two-sided unpaired Welch's *t*-test was used to calculate the *P* value. **P* < 0.05, ***P* < 0.01, ****P* < 0.001 and *****P* < 0.0001. Source data are provided as a Source Data file.



Supplementary Figure 6

A Schematic view of CRISPR/Cas9-based deletion of TSKU-up in ShP51 cells with PCR primers used for amplification of genomic DNA (Left). Gel images show the PCR products from genomic DNA extracted from the deletion mutants (Δ TSKU-up 1, 2) (Right). **B** Luciferase activity of each of Up-2, -3, -4 and TSKU-up with the minimal promoter (minP) in rifampicin (Rif)-treated ShP51 cells. **C** A view of the TSKU-up region (chr11:76,767,944-76,774,353) in UCSC Genome Browser. up-4 contains GWAS SNP, rs1149605 associated with vitamin D levels. **D** Effect of rs1149605 allele on luciferase activity of up-4 with minP in Rif-treated ShP51 cells. In luciferase reporter assays, Firefly luciferase activity was normalized to *Renilla* luciferase activity. Results are expressed as fold change compared with empty vector control. The experiments were performed using independent transfections (n=5). Data are presented as mean values $\pm$ SD, and a two-sided unpaired Welch's *t*-test was used to calculate the *P* value. ***P* < 0.01 and ****P* < 0.001. Source data are provided as a Source Data file.



Supplementary Figure 7
Homologous sequences of PBREM in primates. The DNA sequence corresponding to the PBREM (or DPE16, located at chr2:233756748-233757190 on hg38) was aligned with three primate species with BLAT. Two regions were identified in human and rhesus macaque (old world monkey), while a single region found in common marmoset (new world monkey). No corresponding sequence were found in mouse.