

Supplementary Data

Supplementary Materials and methods

Vector constructions

A two-step cloning strategy was employed to clone MT, $\Delta E2$ and $\Delta E8$ splice variants: In the first step, a pool of clones containing all types of splice variants was prepared. In the second step, each splice variant was screened using colony PCR. To produce a clone pool, total RNA was isolated from HeLa cells, cDNA was prepared and a pool of splice variants was prepared by PCR using the primers (see Supplementary Table 3) which are located on exon 1 and exon 10. PCR products were then inserted into the KpnI and XhoI sites of pBluescript II KS+. To screen each splice variant, colony PCR was performed on individual clones by using the specific primer sets to identify each splice variant listed (see Supplementary Table 2) and the screened clone was confirmed by DNA sequencing. To construct vectors which can easily express, and to detect and isolate each splice variant, Flag-tagged inserts were prepared from the screened clones by PCR amplification using primer sets containing the Flag sequence (see Supplementary Table 3). To clone Flag-tagged +E9, rapid amplification of cDNA ends method was employed using internal primers located on exon 9 and end primers which were the same as those used for clone pool preparation (see the Supplementary Table 3 for primer sequences). The PCR products were cut with KpnI and XhoI and inserted into pcDNA4/TO.

To construct haemagglutinin (HA) tagged expression vectors, vectors constructed above were used as templates for preparing the inserts using primers that contained the HA sequence. The inserts were cloned into the KpnI and XhoI site of pcDNA4/TO.

Immunoprecipitation and immunoblotting

Twenty four hours after transfection, cells were harvested and lysed in Co-IP lysis buffer [50 mM Tris-Cl (pH 7.4), 150 mM NaCl, 1 mM EDTA, 1% Triton X-100] containing protease inhibitor cocktail. Cell lysates were incubated with anti-Flag M2 affinity gel (Sigma) overnight at 4°C. After removing unbound proteins by centrifugation, the resin was washed with Co-IP lysis buffer five times, then, washed with TBS [50 mM Tris-Cl (pH 7.4), 150 mM NaCl] twice. Finally, bound proteins were eluted with elution buffer [50 mM Tris-Cl (pH 6.8), 2% sodium dodecyl sulfate (SDS), 0.01% bromophenol blue, 10% glycerol] and subjected to western blotting. After 10% SDS-polyacrylamide gel electrophoresis (PAGE), the gel was transferred to Hybond-ECL nitrocellulose membrane and the membrane was blocked with 5% skim milk in TBST for 1 h at room temperature. The membrane was incubated with anti-Flag antibody (1:7,000) (Sigma) or anti-HA antibody (1:20,000) (Sigma) diluted

with 5% skim milk for 1 h at room temperature. After washing with TBST, the membrane was incubated with anti-mouse IgG (1:10,000) for 1 h at room temperature.

RT-PCR and semi-quantification

RT-PCR was performed using Human multiple tissue cDNA (MTC) panels (Clontech) as templates. Primers used for detecting each hSPS1 alternative splice variant are listed in Supplementary Table 2. For quantification, band intensities of PCR products were measured using ImageJ software (Version 1.36b, National Institute of Health, USA) [1]. The value of hSPS1 was normalized to that of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and then the relative band intensities to those from liver were calculated.

Supplementary Results

Detection of alternative splice variants in human SPS1 mRNA

To confirm that the predicted splice variants are present in human cells, we performed RT-PCR in several steps (Supplementary Fig. S1). In the first step, all the splice variant forms of human SPS were amplified using the P1 primer set. The expected size of MT, $\Delta E2$ and $\Delta E8$ and +E9 were 2,315 bp, 2,044 bp, 2,102 bp, and 2,378 bp, respectively. Because the sizes of amplified DNA fragments are similar, it is difficult to identify each splice variant from an agarose gel. In the second step, the amplified DNA fragments were isolated from the gel and used as templates for PCR by using the P2 primer set which can amplify from exon 1 to a part of exon 10 to separate each splice variant. As shown in Fig. S1B, $\Delta E2$ and $\Delta E8$ were separated from MT and +E9 forms. However the later two forms were difficult to separate by using the P2 primer set. To further separate +E9 from MT, the largest band on Fig. S1B was eluted and subjected to a third round of PCR using the P3 primer set. With the P3 primer set, only MT was detected (the P3 lane of Fig. S1C). The most likely reason that +E9 was not detected is that its amount was too low for detection. Therefore, the upper region of the MT band was excised and band stab PCR was performed initially using P4a and P4b primer sets and then using the P3 primer set. As shown in Fig. S1C, an additional band, designated +E9a, was detected [see lane P3 (P4a and P4b)]. Sequencing +E9a showed that it contained part of 5' region of intron 9 (see Supplementary Fig. S2).

Differential expression of splice variants in various human cell lines

Unique characteristics of hSPS1 splice variants such as subcellular localization and heterodimer

formation suggest that their functions are unique and the expression of each variant will be different in various tissues and cells. We examined the expression pattern of each variant in human cell lines which are derived from different tissues. Total RNAs from A549 (lung), CRL7407 (stomach), MCF7 (breast), Chang liver (liver), HeLa (cervix), 293 (kidney), and BJAB (B-lymphocyte) cells were isolated and cDNAs were prepared. To measure copy number of each splice variant, real-time PCR was employed using primer sets which can amplify each variant specifically (see Supplementary Table 2). Copy number was calculated by substituting C_T values obtained from cDNA to the standard curve obtained by real-time PCR performed with serially diluted vectors that can express each splice variant.

As shown in Supplementary Fig. S3, the average copy number of MT was the largest among all splice variants (772 copies per nanogram of RNA), and those of $\Delta E2$ and $\Delta E8$ were 10 and 14, respectively. The expression level of +E9/+E9a was the lowest (average copy number was 0.4 per nanogram of RNA). MT and +E9/+E9a were expressed most highly in BJAB which was derived from B-cells compared to other cell lines. $\Delta E2$ and $\Delta E8$ were expressed in the highest level in 293 cells derived from kidney cells and in MCF7 cells derived from breast, respectively. These results suggest that cells control their efficiency of splicing to adjust the level of each splice variant.

Tissue specific distributions of hSPS1 splice variants

We also examined the expression pattern of each alternative splice variants in normal human tissues. cDNAs prepared from various human tissues such as lung, liver, pancreas, kidney, spleen, prostate, testis, ovary, small intestine, brain, placenta, thymus, colon, leukocyte, skeletal muscles, and heart were purchased from commercial sources (given in Supplementary Materials and methods) and RT-PCR was performed. Since all commercial cDNAs produced nonspecific bands, measuring copy number of each splice variant by quantitative real-time PCR was not possible. Therefore, the semi-quantitative RT-PCR method was employed by measuring the intensity of the specifically amplified band. Supplementary Fig. S4 shows the relative band intensities generated from various tissues relative to those from liver. MT was expressed abundantly in pancreas, liver, prostate, testis, spleen, small intestine and lung. $\Delta E2$ was expressed in similar levels in most tissues except heart and skeletal muscle which expressed $\Delta E2$ in very low amounts. $\Delta E8$ was expressed in relatively high levels in brain, placenta, lung, and colon. +E9/+E9a were expressed in relatively high levels in brain, placenta, testis, prostate and colon. Skeletal muscle expressed all splice variants in very low levels indicating hSPS1 is not essential in this tissue. It should be noted that MT was expressed in pancreas most abundantly and that the other three splice variants were expressed most abundantly in brain. The significance of these unique expression patterns in tissue specific manner has not been determined.

Supplementary figure legends

Fig. S1. Detection of splice variants of hSPS1 mRNA in HeLa cells. (A) Positions of primer sets (see Supplementary Table 1 for their sequences). Exons are marked as open boxes. The area marked with parenthesis in exon 1 shows an extremely high GC rich region. F and R designate forward and reverse primer, respectively. (B) All splice variants were amplified by PCR using the P1 primer set from total cDNAs and subjected to agarose gel electrophoresis. The band containing all splice variants was eluted, subjected to PCR amplification using the P2 primer set and the products were separated on an agarose gel. (C) The band containing MT, +E9 and +E9a variants in (B) was eluted and subjected to PCR using the P3 primer set (lane P3). The upper region of the band on the P3 lane was eluted and PCRs were performed. Initially, PCRs using P4a and P4b were carried out. The products were then pooled and another PCR using P3 was performed [lane P3 (P4a+P4b)]. SM designates size markers.

Fig. S2. Nucleotide sequence alignment of splice variants of hSPS1 mRNA. Exon number is marked on the top of each exon. Translation initiation codons are in bold and underlined. Kozak's consensus sequences around the initiation codon are boxed. Translation termination codons are shown in bold and italics. The consensus sequence of the splice junction in +E9a are underlined.

Fig. S3. Expression level of hSPS1 splice variants in various human cell lines. The expression of hSPS1 splice variants in various cell lines was studied by real-time PCR analysis. Copy numbers per nanogram of RNA are represented in the graph. Because MT was amplified by using the primer set which can amplify all splice variants, the MT copy number was calculated by subtracting the copy numbers of other splice variants from the total copy number. Experiments were performed in triplicate and error bars denote the standard deviation from the mean of three experiments. Dotted lines indicate mean of all samples.

Fig. S4. Tissue distribution of hSPS1 alternative splice variants. Semi-quantification of the expression levels of hSPS1 alternative splice variants. RT-PCR was performed using human MTC panels as templates, the PCR products run on an agarose gel and the band intensity of each PCR product measured using the ImageJ program (see Supplementary Materials and methods). Experiments were performed in triplicate, and error bars denote the standard deviation from the mean of three experiments. Dotted lines indicate mean of all samples.

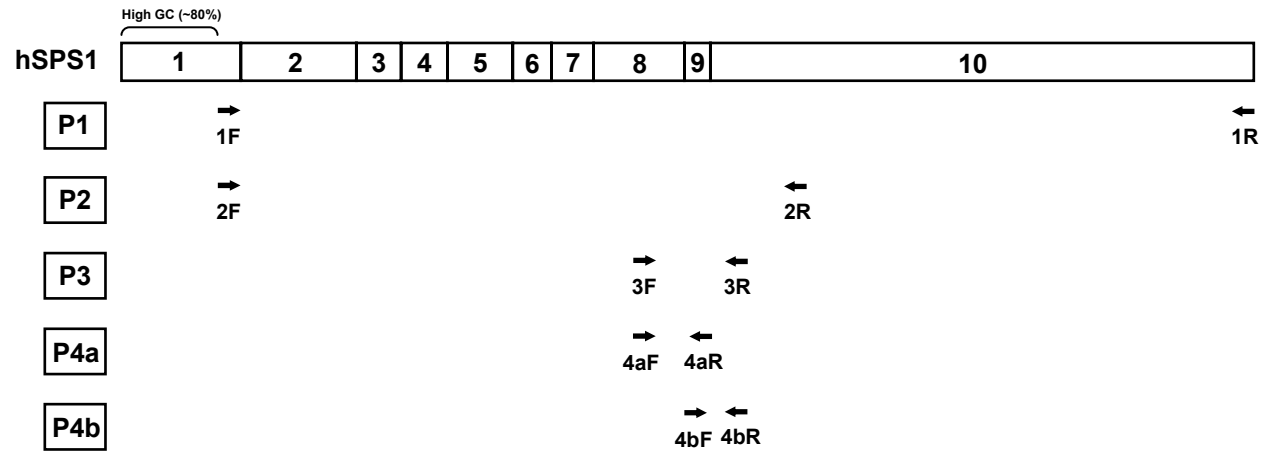
Reference used in Supplementary Data

[1] P. Li, H. Yao, Z. Zhang, M. Li, Y. Luo, P.R. Thompson, D.S. Gilmour, Y. Wang, Regulation of p53 target gene expression by peptidylarginine deiminase 4, *Mol. Cell. Biol.* 28 (2008) 4745-4758.

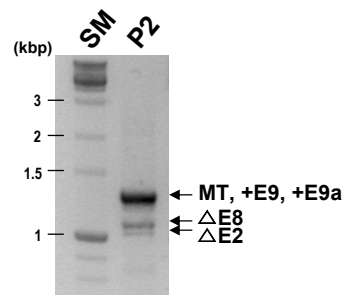
Supplementary Figures

Figure S1

A



B



C

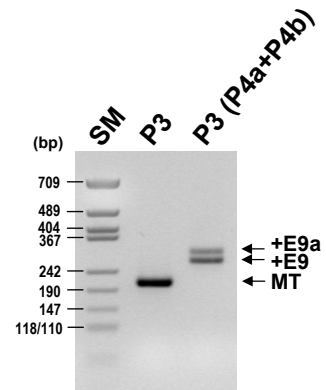


Figure S2

MT	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGCGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
ΔE2	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGCGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
ΔE8	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGCGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
+E9	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGCGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
+E9a	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGCGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
MT	GCCGAGGCAGGCGAGCCCCGCGGCGCCAGGCGCGGGCCCGGGCCCCCGCCGCCGCCGCCCGGGCGCGGCATTTTTATTTCAGGCGACGCTTAAG	: 200
ΔE2	GCCGAGGCAGGCGAGCCCCGCGGCGCCAGGCGCGGGCCCGGGCCCCCGCCGCCGCCGCCCGGGCGCGGCATTTTTATTTCAGGCGACGCTTAAG	: 200
ΔE8	GCCGAGGCAGGCGAGCCCCGCGGCGCCAGGCGCGGGCCCGGGCCCCCGCCGCCGCCGCCCGGGCGCGGCATTTTTATTTCAGGCGACGCTTAAG	: 200
+E9	GCCGAGGCAGGCGAGCCCCGCGGCGCCAGGCGCGGGCCCGGGCCCCCGCCGCCGCCGCCCGGGCGCGGCATTTTTATTTCAGGCGACGCTTAAG	: 200
+E9a	GCCGAGGCAGGCGAGCCCCGCGGCGCCAGGCGCGGGCCCGGGCCCCCGCCGCCGCCGCCCGGGCGCGGCATTTTTATTTCAGGCGACGCTTAAG	: 200
MT	GGAGCCCGGCGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGATCCATAAAGAACC	: 300
ΔE2	GGAGCCCGGCGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGATCCATAAAGAACC	: 282
ΔE8	GGAGCCCGGCGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGATCCATAAAGAACC	: 300
+E9	GGAGCCCGGCGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGATCCATAAAGAACC	: 300
+E9a	GGAGCCCGGCGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGATCCATAAAGAACC	: 300
MT	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGCTACGCGGGAGTCCTTAACCCGAAAGTTACGAATT	: 400
ΔE2	-----	: -
ΔE8	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGCTACGCGGGAGTCCTTAACCCGAAAGTTACGAATT	: 400
+E9	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGCTACGCGGGAGTCCTTAACCCGAAAGTTACGAATT	: 400
+E9a	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGCTACGCGGGAGTCCTTAACCCGAAAGTTACGAATT	: 400
MT	GGACAAAAGCTTCCGGCTAACAGATTCACTGAACTGAAGGGCACAGGCTGCAAAGTGCCCAAGATGTCTGCAAAAATTGCTGGAATCTTTACAGGAG	: 500
ΔE2	-----	: -
ΔE8	GGACAAAAGCTTCCGGCTAACAGATTCACTGAACTGAAGGGCACAGGCTGCAAAGTGCCCAAGATGTCTGCAAAAATTGCTGGAATCTTTACAGGAG	: 500
+E9	GGACAAAAGCTTCCGGCTAACAGATTCACTGAACTGAAGGGCACAGGCTGCAAAGTGCCCAAGATGTCTGCAAAAATTGCTGGAATCTTTACAGGAG	: 500
+E9a	GGACAAAAGCTTCCGGCTAACAGATTCACTGAACTGAAGGGCACAGGCTGCAAAGTGCCCAAGATGTCTGCAAAAATTGCTGGAATCTTTACAGGAG	: 500
MT	AACCACTTCCAAGAAGATGAGCAGTTTCTGGGAGCCGTTATGCCAAGGCTTGGCATTGGAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT	: 600
ΔE2	-----	: 329
ΔE8	AACCACTTCCAAGAAGATGAGCAGTTTCTGGGAGCCGTTATGCCAAGGCTTGGCATTGGAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT	: 600
+E9	AACCACTTCCAAGAAGATGAGCAGTTTCTGGGAGCCGTTATGCCAAGGCTTGGCATTGGAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT	: 600
+E9a	AACCACTTCCAAGAAGATGAGCAGTTTCTGGGAGCCGTTATGCCAAGGCTTGGCATTGGAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT	: 600
MT	CCTTGGTTCAAACCACAGATTACATTTACCCGATCGTAGACGACCCTTACATGATGGCAGGATAGCGTGTGCCAATGCCTCAGTGACCTCTATGCAAT	: 700
ΔE2	CCTTGGTTCAAACCACAGATTACATTTACCCGATCGTAGACGACCCTTACATGATGGCAGGATAGCGTGTGCCAATGCCTCAGTGACCTCTATGCAAT	: 429
ΔE8	CCTTGGTTCAAACCACAGATTACATTTACCCGATCGTAGACGACCCTTACATGATGGCAGGATAGCGTGTGCCAATGCCTCAGTGACCTCTATGCAAT	: 700
+E9	CCTTGGTTCAAACCACAGATTACATTTACCCGATCGTAGACGACCCTTACATGATGGCAGGATAGCGTGTGCCAATGCCTCAGTGACCTCTATGCAAT	: 700
+E9a	CCTTGGTTCAAACCACAGATTACATTTACCCGATCGTAGACGACCCTTACATGATGGCAGGATAGCGTGTGCCAATGCCTCAGTGACCTCTATGCAAT	: 700
MT	GGGGTTCACGGAATGTGACAATATGCTGATGCTCCTTGGAGTCAGTAATAAAATGACCGACAGGGAAAGGGATAAAGTGATGCCTCTGATTATCCAAGGT	: 800
ΔE2	GGGGTTCACGGAATGTGACAATATGCTGATGCTCCTTGGAGTCAGTAATAAAATGACCGACAGGGAAAGGGATAAAGTGATGCCTCTGATTATCCAAGGT	: 529
ΔE8	GGGGTTCACGGAATGTGACAATATGCTGATGCTCCTTGGAGTCAGTAATAAAATGACCGACAGGGAAAGGGATAAAGTGATGCCTCTGATTATCCAAGGT	: 800
+E9	GGGGTTCACGGAATGTGACAATATGCTGATGCTCCTTGGAGTCAGTAATAAAATGACCGACAGGGAAAGGGATAAAGTGATGCCTCTGATTATCCAAGGT	: 800
+E9a	GGGGTTCACGGAATGTGACAATATGCTGATGCTCCTTGGAGTCAGTAATAAAATGACCGACAGGGAAAGGGATAAAGTGATGCCTCTGATTATCCAAGGT	: 800
MT	TTTAAAGACGCAGCTGAGGAAGCAGGAACGCTCTGTAACAGGCGGCCAAACAGTACTAAACCCCTGGATTGTCTGGGAGGAGTGGCTACCACTGTCTGCC	: 900
ΔE2	TTTAAAGACGCAGCTGAGGAAGCAGGAACGCTCTGTAACAGGCGGCCAAACAGTACTAAACCCCTGGATTGTCTGGGAGGAGTGGCTACCACTGTCTGCC	: 629
ΔE8	TTTAAAGACGCAGCTGAGGAAGCAGGAACGCTCTGTAACAGGCGGCCAAACAGTACTAAACCCCTGGATTGTCTGGGAGGAGTGGCTACCACTGTCTGCC	: 900
+E9	TTTAAAGACGCAGCTGAGGAAGCAGGAACGCTCTGTAACAGGCGGCCAAACAGTACTAAACCCCTGGATTGTCTGGGAGGAGTGGCTACCACTGTCTGCC	: 900
+E9a	TTTAAAGACGCAGCTGAGGAAGCAGGAACGCTCTGTAACAGGCGGCCAAACAGTACTAAACCCCTGGATTGTCTGGGAGGAGTGGCTACCACTGTCTGCC	: 900

E1

ACCATG

ACCATG

ACCATG

ACCATG

E2

GGAATGG

E3

E4

E5

MT	AACCCAATGAATTTATCAT	BCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGCAGTGCACAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG	: 1000
ΔE2	AACCCAATGAATTTATCAT	BCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGCAGTGCACAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG	: 729
ΔE8	AACCCAATGAATTTATCAT	BCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGCAGTGCACAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG	: 1000
+E9	AACCCAATGAATTTATCAT	BCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGCAGTGCACAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG	: 1000
+E9a	AACCCAATGAATTTATCAT	BCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGCAGTGCACAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG	: 1000
E6			
MT	GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC	: 1100	
ΔE2	GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC	: 829	
ΔE8	GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC	: 1100	
+E9	GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC	: 1100	
+E9a	GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC	: 1100	
E7			
MT	AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC	: 1200	
ΔE2	AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC	: 929	
ΔE8	AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC	: 1110	
+E9	AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC	: 1200	
+E9a	AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC	: 1200	
MT	AGCAGAGGAACGAGGTGTCGTTTGTAAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA	: 1300	
ΔE2	AGCAGAGGAACGAGGTGTCGTTTGTAAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA	: 1029	
ΔE8	-----	: -	
+E9	AGCAGAGGAACGAGGTGTCGTTTGTAAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA	: 1300	
+E9a	AGCAGAGGAACGAGGTGTCGTTTGTAAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA	: 1300	
E8			
MT	CGGGACCTGCCCGGAGACTTCAG-----	: 1323	
ΔE2	CGGGACCTGCCCGGAGACTTCAG-----	: 1052	
ΔE8	-----	: -	
+E9	CGGGACCTGCCCGGAGACTTCAGATGTGCAGTAATTTATTAGAGGAAACATAAATGGAGGATAAACGGGAGATGGAGCAAGCGTAG-----	: 1386	
+E9a	CGGGACCTGCCCGGAGACTTCAGATGTGCAGTAATTTATTAGAGGAAACATAAATGGAGGATAAACGGGAGATGGAGCAAGCGTAGGTAGAGAGAGCCAT	: 1400	
E9			
MT	-----ECGGCCTTCTGATCTGTTTACCAGTGCAGCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA	: 1408	
ΔE2	-----ECGGCCTTCTGATCTGTTTACCAGTGCAGCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA	: 1137	
ΔE8	-----ECGGCCTTCTGATCTGTTTACCAGTGCAGCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA	: 1195	
+E9	-----ECGGCCTTCTGATCTGTTTACCAGTGCAGCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA	: 1471	
+E9a	CAGACAGCAGTGCAGECGGCCTTCTGATCTGTTTACCAGTGCAGCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA	: 1500	
MT	AGCATGGATTATTGGGATTGTAGAGAAGGGCAACCGCACAGCCAGAATCATAGACAAACCCCGGATCATCGAGGTCGCACCACAAGTGGCCACTCAAAAT	: 1508	
ΔE2	AGCATGGATTATTGGGATTGTAGAGAAGGGCAACCGCACAGCCAGAATCATAGACAAACCCCGGATCATCGAGGTCGCACCACAAGTGGCCACTCAAAAT	: 1237	
ΔE8	AGCATGGATTATTGGGATTGTAGAGAAGGGCAACCGCACAGCCAGAATCATAGACAAACCCCGGATCATCGAGGTCGCACCACAAGTGGCCACTCAAAAT	: 1295	
+E9	AGCATGGATTATTGGGATTGTAGAGAAGGGCAACCGCACAGCCAGAATCATAGACAAACCCCGGATCATCGAGGTCGCACCACAAGTGGCCACTCAAAAT	: 1571	
+E9a	AGCATGGATTATTGGGATTGTAGAGAAGGGCAACCGCACAGCCAGAATCATAGACAAACCCCGGATCATCGAGGTCGCACCACAAGTGGCCACTCAAAAT	: 1600	
MT	GTGAATCCCACACCCGGGCCACCTCTTAATCTAGACAGAAATAGCTGTTTGGTTTTGTTTTTAAATAGATCTATTTCCCTTATCACTTCAATTAAGAC	: 1608	
ΔE2	GTGAATCCCACACCCGGGCCACCTCTTAATCTAGACAGAAATAGCTGTTTGGTTTTGTTTTTAAATAGATCTATTTCCCTTATCACTTCAATTAAGAC	: 1337	
ΔE8	GTGAATCCCACACCCGGGCCACCTCTTAATCTAGACAGAAATAGCTGTTTGGTTTTGTTTTTAAATAGATCTATTTCCCTTATCACTTCAATTAAGAC	: 1395	
+E9	GTGAATCCCACACCCGGGCCACCTCTTAATCTAGACAGAAATAGCTGTTTGGTTTTGTTTTTAAATAGATCTATTTCCCTTATCACTTCAATTAAGAC	: 1671	
+E9a	GTGAATCCCACACCCGGGCCACCTCTTAATCTAGACAGAAATAGCTGTTTGGTTTTGTTTTTAAATAGATCTATTTCCCTTATCACTTCAATTAAGAC	: 1700	
E10			
MT	TATAAACAAACAAAAATCTCATTGTGTCTACACATCGGGGTGACCTTAGGTCGGTTTGTAAAGTGGATAACAATTAATAAAATAAAATCCATTGCCTTTTTTT	: 1708	
ΔE2	TATAAACAAACAAAAATCTCATTGTGTCTACACATCGGGGTGACCTTAGGTCGGTTTGTAAAGTGGATAACAATTAATAAAATAAAATCCATTGCCTTTTTTT	: 1437	
ΔE8	TATAAACAAACAAAAATCTCATTGTGTCTACACATCGGGGTGACCTTAGGTCGGTTTGTAAAGTGGATAACAATTAATAAAATAAAATCCATTGCCTTTTTTT	: 1495	
+E9	TATAAACAAACAAAAATCTCATTGTGTCTACACATCGGGGTGACCTTAGGTCGGTTTGTAAAGTGGATAACAATTAATAAAATAAAATCCATTGCCTTTTTTT	: 1771	
+E9a	TATAAACAAACAAAAATCTCATTGTGTCTACACATCGGGGTGACCTTAGGTCGGTTTGTAAAGTGGATAACAATTAATAAAATAAAATCCATTGCCTTTTTTT	: 1800	

MT CCTGTTACATTAAGTGAAGATGCACCTAATCTTGAGGCAGCTTCTGAGTTGAGAATTATATTGTTATCCAATACTGTTGATTCAATTTGAATCTTTAGAC : 1808
ΔE2 CCTGTTACATTAAGTGAAGATGCACCTAATCTTGAGGCAGCTTCTGAGTTGAGAATTATATTGTTATCCAATACTGTTGATTCAATTTGAATCTTTAGAC : 1537
ΔE8 CCTGTTACATTAAGTGAAGATGCACCTAATCTTGAGGCAGCTTCTGAGTTGAGAATTATATTGTTATCCAATACTGTTGATTCAATTTGAATCTTTAGAC : 1595
+E9 CCTGTTACATTAAGTGAAGATGCACCTAATCTTGAGGCAGCTTCTGAGTTGAGAATTATATTGTTATCCAATACTGTTGATTCAATTTGAATCTTTAGAC : 1871
+E9a CCTGTTACATTAAGTGAAGATGCACCTAATCTTGAGGCAGCTTCTGAGTTGAGAATTATATTGTTATCCAATACTGTTGATTCAATTTGAATCTTTAGAC : 1900

MT ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1908
ΔE2 ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1637
ΔE8 ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1695
+E9 ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1971
+E9a ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 2000

MT TATTTATCATATAAGTCTGATTTTTTTTTTAAAGCGTCTTGAATGGTTTTCTGGAGAGACAGCATTGGTAAGTGGCAGTATCCAGTCATAAG : 2008
ΔE2 TATTTATCATATAAGTCTGATTTTTTTTTTAAAGCGTCTTGAATGGTTTTCTGGAGAGACAGCATTGGTAAGTGGCAGTATCCAGTCATAAG : 1737
ΔE8 TATTTATCATATAAGTCTGATTTTTTTTTTAAAGCGTCTTGAATGGTTTTCTGGAGAGACAGCATTGGTAAGTGGCAGTATCCAGTCATAAG : 1795
+E9 TATTTATCATATAAGTCTGATTTTTTTTTTAAAGCGTCTTGAATGGTTTTCTGGAGAGACAGCATTGGTAAGTGGCAGTATCCAGTCATAAG : 2071
+E9a TATTTATCATATAAGTCTGATTTTTTTTTTAAAGCGTCTTGAATGGTTTTCTGGAGAGACAGCATTGGTAAGTGGCAGTATCCAGTCATAAG : 2100

MT AGGGTTGCATGATTCCCTTTGAGTGTTGATTTGAAAAGCCTAGTCTTGCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 2108
ΔE2 AGGGTTGCATGATTCCCTTTGAGTGTTGATTTGAAAAGCCTAGTCTTGCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 1837
ΔE8 AGGGTTGCATGATTCCCTTTGAGTGTTGATTTGAAAAGCCTAGTCTTGCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 1895
+E9 AGGGTTGCATGATTCCCTTTGAGTGTTGATTTGAAAAGCCTAGTCTTGCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 2171
+E9a AGGGTTGCATGATTCCCTTTGAGTGTTGATTTGAAAAGCCTAGTCTTGCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 2200

MT AACACAGCAAGTGCTTCATTGCATGGAAAACACTTTGAAGACAAAAAGAAATCTTATTTCTTTTTTGTAGCCTTCTGATATTTACAGTAATACCATT : 2208
ΔE2 AACACAGCAAGTGCTTCATTGCATGGAAAACACTTTGAAGACAAAAAGAAATCTTATTTCTTTTTTGTAGCCTTCTGATATTTACAGTAATACCATT : 1937
ΔE8 AACACAGCAAGTGCTTCATTGCATGGAAAACACTTTGAAGACAAAAAGAAATCTTATTTCTTTTTTGTAGCCTTCTGATATTTACAGTAATACCATT : 1995
+E9 AACACAGCAAGTGCTTCATTGCATGGAAAACACTTTGAAGACAAAAAGAAATCTTATTTCTTTTTTGTAGCCTTCTGATATTTACAGTAATACCATT : 2271
+E9a AACACAGCAAGTGCTTCATTGCATGGAAAACACTTTGAAGACAAAAAGAAATCTTATTTCTTTTTTGTAGCCTTCTGATATTTACAGTAATACCATT : 2300

MT AACTGTTTTATCGATAGCAAAAAAGGATACTTTTTGCAATGTTATTAGATGTTCTATAGTGCTACAAGGAATTGCCTTCCGAATGGAGGTTTCATGTATAA : 2308
ΔE2 AACTGTTTTATCGATAGCAAAAAAGGATACTTTTTGCAATGTTATTAGATGTTCTATAGTGCTACAAGGAATTGCCTTCCGAATGGAGGTTTCATGTATAA : 2037
ΔE8 AACTGTTTTATCGATAGCAAAAAAGGATACTTTTTGCAATGTTATTAGATGTTCTATAGTGCTACAAGGAATTGCCTTCCGAATGGAGGTTTCATGTATAA : 2095
+E9 AACTGTTTTATCGATAGCAAAAAAGGATACTTTTTGCAATGTTATTAGATGTTCTATAGTGCTACAAGGAATTGCCTTCCGAATGGAGGTTTCATGTATAA : 2371
+E9a AACTGTTTTATCGATAGCAAAAAAGGATACTTTTTGCAATGTTATTAGATGTTCTATAGTGCTACAAGGAATTGCCTTCCGAATGGAGGTTTCATGTATAA : 2400

MT TACTCATTTACAATTC AATATATAATTACACAAATAATTTTTAAATATAATCAATAGTAAAGACTGTTCTGTGGATGGTAGTGTTAATACATTTTCTAT : 2408
ΔE2 TACTCATTTACAATTC AATATATAATTACACAAATAATTTTTAAATATAATCAATAGTAAAGACTGTTCTGTGGATGGTAGTGTTAATACATTTTCTAT : 2137
ΔE8 TACTCATTTACAATTC AATATATAATTACACAAATAATTTTTAAATATAATCAATAGTAAAGACTGTTCTGTGGATGGTAGTGTTAATACATTTTCTAT : 2195
+E9 TACTCATTTACAATTC AATATATAATTACACAAATAATTTTTAAATATAATCAATAGTAAAGACTGTTCTGTGGATGGTAGTGTTAATACATTTTCTAT : 2471
+E9a TACTCATTTACAATTC AATATATAATTACACAAATAATTTTTAAATATAATCAATAGTAAAGACTGTTCTGTGGATGGTAGTGTTAATACATTTTCTAT : 2500

MT TTTGTACAGTGATTTTCAGGCCCTTTGTTTTCTTAAAATCAGCAGCTGTTTGGCCTAATTCCTAGCATTATTTTGTCCCTTTGCGCCAGTACTTTTTTGTGC : 2508
ΔE2 TTTGTACAGTGATTTTCAGGCCCTTTGTTTTCTTAAAATCAGCAGCTGTTTGGCCTAATTCCTAGCATTATTTTGTCCCTTTGCGCCAGTACTTTTTTGTGC : 2237
ΔE8 TTTGTACAGTGATTTTCAGGCCCTTTGTTTTCTTAAAATCAGCAGCTGTTTGGCCTAATTCCTAGCATTATTTTGTCCCTTTGCGCCAGTACTTTTTTGTGC : 2295
+E9 TTTGTACAGTGATTTTCAGGCCCTTTGTTTTCTTAAAATCAGCAGCTGTTTGGCCTAATTCCTAGCATTATTTTGTCCCTTTGCGCCAGTACTTTTTTGTGC : 2571
+E9a TTTGTACAGTGATTTTCAGGCCCTTTGTTTTCTTAAAATCAGCAGCTGTTTGGCCTAATTCCTAGCATTATTTTGTCCCTTTGCGCCAGTACTTTTTTGTGC : 2600

MT ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATAATTTAATTTTTTAAA : 2600
ΔE2 ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATAATTTAATTTTTTAAA : 2329
ΔE8 ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATAATTTAATTTTTTAAA : 2387
+E9 ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATAATTTAATTTTTTAAA : 2663
+E9a ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATAATTTAATTTTTTAAA : 2692

Figure S3

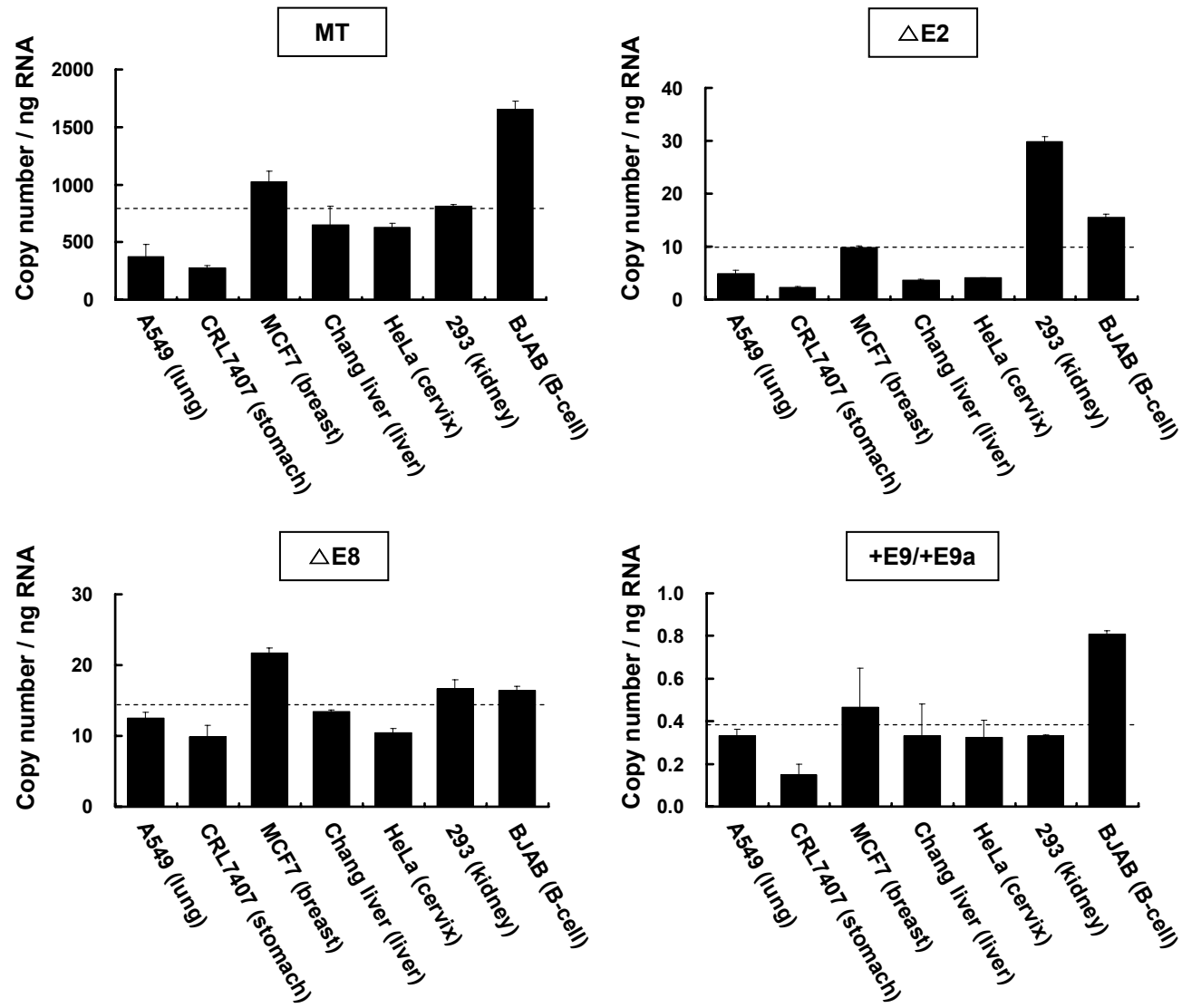
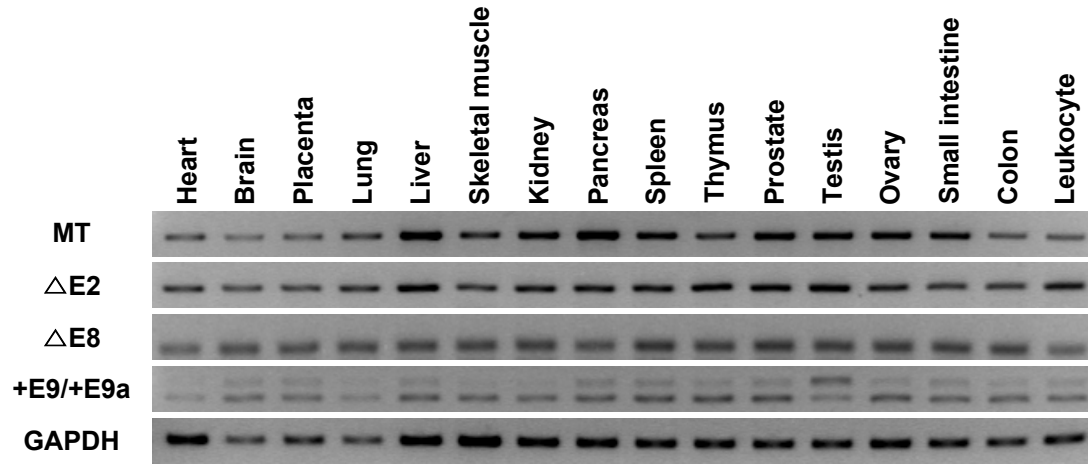
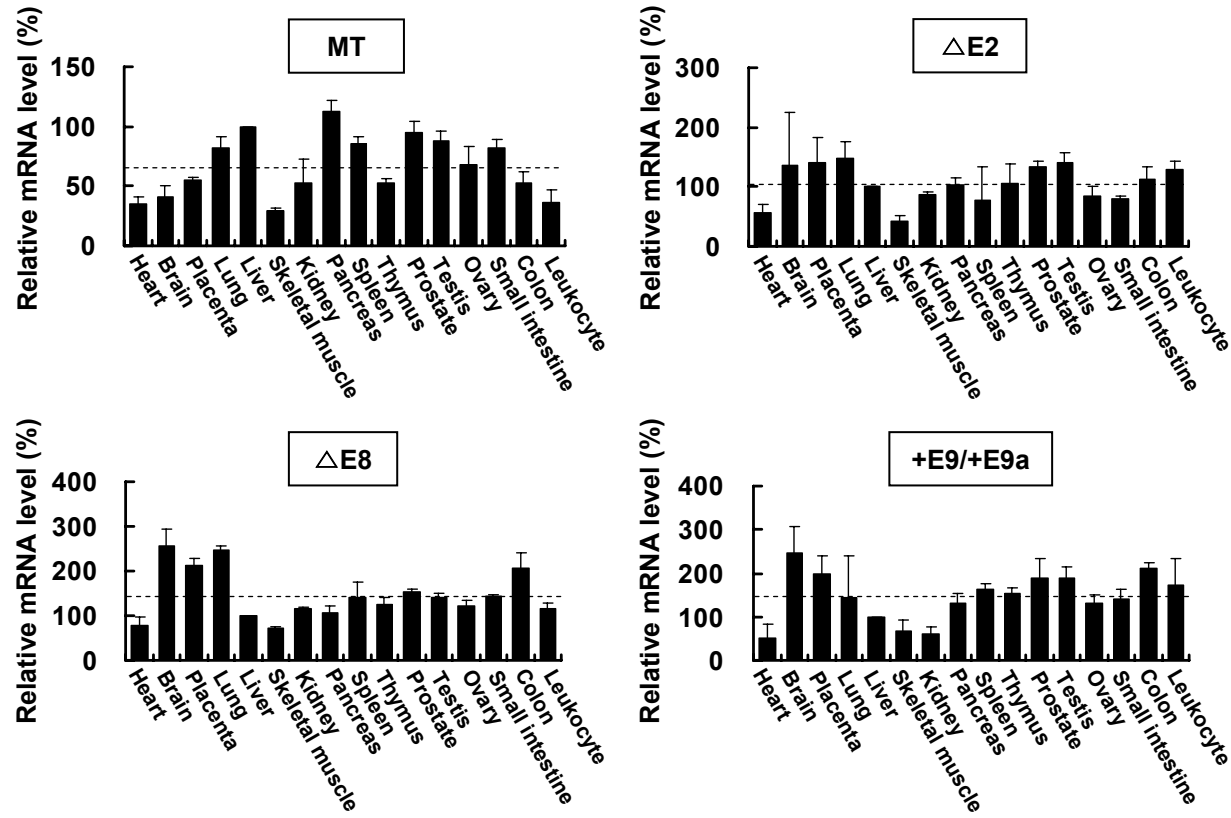


Figure S4

A



B



Supplementary Tables

Table S1. Oligonucleotides used as primers for detecting hSPS1 alternative splice variants

Name	Sequence (5' → 3')
P1 - forward	GAGGCGGAGGGCGCAAAGCG
P1 - reverse	CAAGTTTAAGTTCGAGCTGCAA
P2 - forward	GAGGCGGAGGGCGCAAAGCG
P2 - reverse	TCTGTCTAGATTAAGAGGTGGC
P3 - forward	CAGAGGAACGAGGTGTCGTT
P3 - reverse	TTGGTGGCCTTCACCATATT
P4a - forward	CAGAGGAACGAGGTGTCGTT
P4a - reverse	CTACGCTTGCTCCATCTCCC
P4b - forward	TGGAGGATAAACGGGAGATG
P4b - reverse	TTGGTGGCCTTCACCATATT

Table S2. Oligonucleotides used as primers for measuring expression of splice variants

Name	Sequence (5' → 3')
MT-forward	CTTGGTTCAAACCACAGATT
MT-reverse	CCAGCTCTACATCTTCTTGG
ΔE2-forward	GGCGCAAAGCGAGCCGGCAT
ΔE2-reverse	GGTCGTCTACGATCGGGTAA
ΔE8-forward	TAGAGCTGGCCTACCAGGAG
ΔE8-reverse	AGATCAGAAGGCCGCCTGTC
+E9/+E9a-forward	TGGAGGATAAACGGGAGATG
+E9/+E9a-reverse	TTGGTGGCCTTCACCATATT

Table S3. Oligonucleotides used as primers for constructing hSPS1 expression vectors

Name	Sequence (5' → 3')
E1-forward	AAGCTTGGTACCGAGGCGGAGGGCGCAAAGCG
E10-reverse	AAGCTTCTCGAGTCTGTCTAGATTAAGAGGTGGC
Over-forward	AAGCTTGGTACCGAGGCGGAGGGCGCAAAGCG
Over-Flag-reverse	AAGCTTCTCGAGTTACTTGTGTCATCGTCTTTGTAGTCAGAGGTGGCCCCGGGTGTGG
Internal E9-forward	GGAGACTTCAGATGTGCAGT
Internal E9-reverse	CTACGCTTGCTCCATCTCCC
Over-HA-reverse	TCTAGACTCGAGTTAAGCGTAATCTGGAACATCGTATGGGTAAGAGGTGGCCCCGGGTGT