



Human selenophosphate synthetase 1 has five splice variants with unique interactions, subcellular localizations and expression patterns

Jin Young Kim^a, Kwang Hee Lee^{a,b}, Myoung Sup Shim^a, Hyein Shin^a, Xue-Ming Xu^c, Bradley A. Carlson^c, Dolph L. Hatfield^c, Byeong Jae Lee^{a,b,*}

^aLaboratory of Molecular Genetics and Genomics, School of Biological Sciences, Institute of Molecular Biology and Genetics, Seoul National University, Seoul 151-742, Republic of Korea

^bInterdisciplinary Program in Bioinformatics, Seoul National University, Seoul 151-742, Republic of Korea

^cLaboratory of Cancer Prevention, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

ARTICLE INFO

Article history:

Received 5 May 2010

Available online 22 May 2010

Keywords:

Alternative splicing

Cell cycle

Selenium

Selenocysteine

Selenophosphate synthetase 1

ABSTRACT

Selenophosphate synthetase 1 (SPS1) is an essential cellular gene in higher eukaryotes. Five alternative splice variants of human SPS1 (major type, $\Delta E2$, $\Delta E8$, +E9, +E9a) were identified wherein +E9 and +E9a make the same protein. The major type was localized in both the nuclear and plasma membranes, and the others in the cytoplasm. All variants form homodimers, and in addition, the major type forms a heterodimer with $\Delta E2$, and $\Delta E8$ with +E9. The level of expression of each splice variant was different in various cell lines. The expression of each alternative splice variant was regulated during the cell cycle. The levels of the major type and $\Delta E8$ were gradually increased until G2/M phase and then gradually decreased. $\Delta E2$ expression peaked at mid-S phase and then gradually decreased. However, +E9/+E9a expression decreased gradually after cell cycle arrest. The possible involvement of SPS1 splice variants in cell cycle regulation is discussed.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Selenium is an essential trace element and provides many health benefits for animals including humans. For example, this element has been known to play a role in cancer prevention, aging retardation, immune augmentation, prevention of cardiovascular disease, muscle development and immune function [1–4]. Most of these various selenium effects are believed to be mediated by selenoproteins which contain the selenium-containing amino acid, selenocysteine (Sec), at their active site(s). Sec, which is the 21st amino acid in the genetic code, is inserted into a growing peptide during translation in response to the codon, UGA [5–7]. The biosynthesis of Sec is carried out through a unique pathway that is distinct from that of other amino acids. Most amino acids are synthesized prior to being aminoacylated onto their cognate tRNAs; however, Sec is synthesized on its tRNA, designated tRNA^{[Ser]Sec}. In eukaryotes, tRNA^{[Ser]Sec} is initially aminoacylated with serine by seryl-tRNA synthetase, the hydroxyl moiety of serine on seryl-

tRNA^{[Ser]Sec} is then replaced by phosphate group to form O-phosphoseryl-tRNA^{[Ser]Sec} that is catalyzed by a specific kinase, phosphoseryl-tRNA^{[Ser]Sec} kinase [8]. Finally, the phosphate group on O-phosphoseryl-tRNA^{[Ser]Sec} is exchanged with activated selenium to form selenocysteyl-tRNA^{[Ser]Sec} catalyzed by the enzyme, Sec synthase (SecS) [9,10]. The O-phosphoseryl-tRNA^{[Ser]Sec} intermediate is not synthesized in eubacteria wherein SecS in *Escherichia coli* (SelA) removes the hydroxyl group of serine on seryl-tRNA^{[Ser]Sec} to form aminoacrylyl-tRNA^{[Ser]Sec} and the resulting aminoacrylyl intermediate in turn accepts activated selenium to form Sec-tRNA^{[Ser]Sec} [11]. Mono-selenophosphate (SeP) is the activated selenium form in both bacteria and eukaryotes [12], and it is synthesized from selenide and ATP. Selenophosphate synthetase (SPS) catalyzes the synthesis of SeP [13]. Two isoforms of SPS were found in higher eukaryotes, SPS1 and SPS2, whereas only one type of SPS (SelD) exists in lower eukaryotes and eubacteria [14]. The sequence homology between SPS1 and SPS2 is different depending on the species. For example, the amino acid sequence homology between human SPS1 and SPS2 is 72% and that between *Drosophila* SPS1 and SPS2 is approximately 45%. One of the major differences between SPS1 and SPS2 is that SPS1 has an arginine at the position corresponding to Sec in SPS2 [15,16]. The sequence homology between *Drosophila* SPS1 and *Drosophila* SPS2 is 45%, while that between human SPS1 and *Drosophila* SPS1 is 71%, indicating that the SPS1 sequence has been well conserved during evolution (NCBI, Build 37.1 for human sequence and BDGP Release 5.12 for *Drosophila*).

Abbreviations: AA, antibiotic-antimycotic; C_T, cycle-threshold; FACS, fluorescent associated cell sorter; FBS, fetal bovine serum; HA, haemagglutinin; MT, major type; PI, propidium iodide; ROS, reactive oxygen species; SPS, selenophosphate synthetase; Sec, selenocysteine; SecS, Sec synthase.

* Corresponding author. Address: School of Biological Sciences, Seoul National University, Building 504, Room 523, Seoul 151-742, South Korea. Fax: +82 2 872 9019.

E-mail address: imbglmg@snu.ac.kr (B.J. Lee).

Recently, it was shown that SPS1 does not have SeP synthesis activity, but SPS2 catalyzes SeP synthesis. In *in vitro* experiments, SPS2 synthesized SeP from selenide and ATP, but SPS1 did not have this activity [9]. The ablation of SPS2 mRNA in NIH3T3 cells led to the loss of selenoprotein biosynthesis, while the inhibition of SPS1 expression did not affect the biosynthesis of selenoprotein, and the synthesis of selenoproteins in the cells in which SPS2 was knocked down was rescued by expression of SPS2, but not by SPS1 [17]. Although some insects such as the red beetle and silkworm have lost the selenoprotein synthesizing machinery including SPS2, SPS1 is still encoded in the genome of those organisms suggesting SPS1 is not essential for selenoprotein synthesis [18].

Although SPS1 does not catalyze SeP biosynthesis [12], there are several lines of evidence suggesting that it plays important roles in the cell. Targeted removal of the gene encoding SPS1 (also designated *patufet*) in *Drosophila* causes embryonic lethality suggesting that SPS1 participates in pathways important for cell viability [19]. It was also found that knockout of SPS1 caused an increase in reactive oxygen species (ROS) and the haploinsufficiency of SPS1 suppressed the abnormal phenotypes caused by overexpression of genes involved in Ras-regulated signaling pathway [20,21]. These latter findings suggest that SPS1 participates in the regulation of oxidative stress and cell growth. In agreement with this predicted function of SPS1, knockdown of *Drosophila* SPS1 led to growth inhibition and ROS generation [22]. Interestingly, knockdown of *Drosophila* SPS1 caused generation of megamitochondria through the increase of intracellular glutamine levels.

When cDNA of SPS1 was cloned from human lung adenocarcinoma and transformed into a *SelD* deficient mutant of *E. coli*, the *seld* mutation was not complemented in selenite containing medium, but complemented when the cells were cultured in the medium supplemented with L-Sec [23]. From these findings, it was suggested that SPS1 is involved in the recycling of Sec. However, the mechanism of Sec recycling and how SPS1 regulates Sec recycling have not been determined. Interestingly, SPS1 was found to interact with soluble liver antigen (SLA) which was recently identified as eukaryotic Sec synthase (SecS) and the binding reaction was enhanced by a protein designated SECp43 [24].

In the present study, we found that hSPS1 produces five alternative splice variants and one of them was generated by mis-splicing. These alternative splice variants showed unique characteristics of intermolecular interaction, subcellular localization, specific tissue and cell cycle expression suggesting that each splice variant has a specific function for cell growth.

2. Materials and methods

2.1. Cell culture and DNA transfection

A549, CRL7407, MCF7, Chang liver, HeLa, and 293 cells were grown in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum (FBS) (Hyclone) and 1% antibiotic-antimycotic (AA), and BJAB cells in RPMI 1640 medium supplemented with 10% FBS and 1% AA. All cultures were maintained in a humidified atmosphere of 5% CO₂ at 37 °C. Transient transfections in HeLa cells were carried out using Lipofectamine Plus (Invitrogen) according to the manufacturer's instructions and calcium phosphate method was used in 293 cell transfection.

2.2. Vector constructions

Vectors that express alternative splice variants were constructed in two steps: 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. Vectors that can express Flag- or HA-tagged proteins were prepared by PCR using

primers that contain the sequences encoding Flag or HA, respectively (see Supplementary Materials and methods for detail).

2.3. Immunostaining and confocal microscopy

HeLa cells transfected with Flag-tagged hSPS1 expression vector were transferred onto a cover glass one day before staining. Immunocytochemistry was carried out as described previously [22] with a slight modification. After the cells were fixed, permeabilized and blocked, anti-Flag antibody (Sigma) diluted in 3% BSA in PBS (1:2500) was treated for 1 h at room temperature. The cells were then incubated with Alexa 488 anti-mouse IgG (Invitrogen) diluted in 3% BSA in PBS (1:2000) for 1 h at room temperature. For counterstaining, cells were stained with 50 µg/ml of propidium iodide (PI) (Sigma) for 20 min at room temperature. Cells were observed with a LSM 510 NLO confocal microscope (Carl Zeiss). Excitation was at 488 nm for Alexa 488 and 543 nm for PI.

2.4. Quantitative real-time RT-PCR

To determine the number of hSPS1 mRNA molecules, total RNA was isolated from cells using TRIzol reagent (Invitrogen), first-strand cDNA was synthesized from 3 µg of total RNA with Mo-MuLV reverse transcriptase (Super-Bio), and used for the PCR reaction. cDNAs were amplified with SYBR Green PCR master mix and specific primers using the ABI 7300 real-time PCR system (Applied Biosystems). Vectors which can express the hSPS1 alternative splice variant were serially diluted from 10¹ to 10⁸ copy numbers and applied to real-time PCR. Standard curve was obtained by plotting the log of the number of DNA templates versus the measured C_T (cycle-threshold). Using the equation calculated from the standard curve, the copy number of hSPS1 mRNA per nanogram of total RNA was determined.

2.5. Immunoprecipitation and immunoblotting

A vector which can express a Flag-tagged splice variant was cotransfected with a HA-tagged splice variant expression vector in 293 cells in a manner that a vector expressing one type of Flag-tagged splice variant was cotransfected with a vector expressing one of all types of HA-tagged splice variants. Cell lysates were coimmunoprecipitated with anti-Flag M2 affinity gel (Sigma) and the precipitates subjected to western blotting using anti-HA antibody or anti-Flag antibody (see Reference [22] and Supplementary Materials and methods for detail).

2.6. Synchronization of cells and flow cytometry

For synchronization, HeLa cells were incubated with 5 µg/ml of aphidicolin (Sigma) for 24 h. Cells were washed with fresh medium and released from the aphidicolin block for indicated times. Cell cycle profiles were analyzed by flow cytometry with standard PI staining methods. Asynchronized and synchronized cells were harvested and washed once with PBS. Cells were fixed overnight in 70% ethanol at 4 °C, washed with PBS twice and then resuspended with PBS containing 200 µg/ml of RNase A. After 10 min incubation at room temperature, cells were further incubated with 100 µg/ml of PI in PBS for 30 min at room temperature. Stained cells were analyzed by using a FACS Calibur flow cytometer and CellQuest Pro software (BD Biosciences).

3. Results

3.1. Identification of alternative splice variants in hSPS1 mRNA

To elucidate the molecular function of hSPS1, its cDNA was amplified using total cDNAs prepared from HeLa cells and cloned

into a vector. The sequences of several clones harboring hSPS1 cDNA were determined. Interestingly, we found that one of the clones was a truncated form in which exon 8 was deleted. We, therefore, assumed that there would be alternative splice variants in hSPS1 mRNA. To find possible alternative splice forms, we searched EST clones from the University of California, Santa Cruz Genome Browser (<http://genome.ucsc.edu>) in scrutiny. Among a total of 283 EST clones of hSPS1, 277 ESTs were the same as known hSPS1 mRNA and designated major type (MT). Except MT, there were two ESTs (CX787008, DB001684) in which exon 2 was deleted ($\Delta E2$), two ESTs (AI741523, AI742171) where exon 8 was deleted ($\Delta E8$) and additional 2 ESTs (BX385308, CR604492) where exon 9 was added to the major type (+E9). These data suggest that there are 4 alternative splice variants of hSPS1 mRNA in human cells.

To confirm that the predicted alternative splice variants are present in human cells, we performed RT-PCR in several steps and surprisingly, detected 5 splice variants instead of the 4 predicted splice variants (see [Supplementary Result and Supplementary Fig. S1](#)). The sequence of each splice variant was determined and found to be consistent with the expected sequence (GenBank Accession No. GU954545, GU954546, GU954547, GU954548, and GU954549; [Supplementary Fig. S2](#)).

As summarized in [Fig. 1](#), hSPS1 pre-mRNA is alternatively spliced to produce 5 different splice variants. The translation initiation codon is located in exon 2 and the termination codon in exon 10. As expected, $\Delta E2$ and $\Delta E8$ do not have exon 2 and exon 8, respectively. Since the normal initiation codon is located in exon 2, $\Delta E2$ cannot use the original initiation codon. Therefore, $\Delta E2$ must use another initiation codon. In exon 3 (at +561), there is an ATG codon and the surrounding Kozak's consensus sequence. Both forms are void of exon 9 as found in MT. The sequence of exon 9 is very short and there is the translation termination codon (TAA) within exon 9 suggesting that the translated protein from +E9 will be truncated from MT, even though the size of mRNA of +E9 is larger than that of MT. +E9a has an additional sequence which is part of 5' region of intron 9. Interestingly, the beginning and end

of the additional sequence contain splice junction sequences (GT and AG), respectively. In the genomic DNA sequence, there is no gap between exon 9 and the additional sequence. This suggests that +E9a is generated by mis-splicing. The protein sequences of +E9 and +E9a should be the same, because there is a termination codon within exon 9.

3.2. Subcellular localizations of splice variants

Recently, it was reported that hSPS1 makes supramolecular complexes by interacting with SecS both *in vivo* and *in vitro* and this supramolecular complex shuttles between the nucleus and the cytoplasm [24]. Therefore, it is important to determine the subcellular localization of each alternative splice variant to elucidate their functions. The Flag-tagged splice variants were expressed and detected by immunocytochemistry using anti-Flag antibody. As shown in [Fig. 2](#), MT was localized on both plasma and the nuclear membrane, but the other three types were localized in the cytoplasm homogeneously. These results suggest that each alternative splice type plays a unique role by interacting with different proteins or molecules.

3.3. Interactions between splice variants

By analyzing the X-ray crystallography structure of hSPS1, it was suggested recently that hSPS1 could form a homodimer [25]. We therefore examined whether all the splice variants could make dimers (homodimers as well as heterodimers). In addition to Flag-tagged constructs, HA-tagged constructs were also prepared (see [Supplementary Materials and methods](#)). After co-transfecting a Flag-tagged construct and a HA-tagged construct into 293 cells, cell lysates were subjected to immunoprecipitation using anti-Flag antibody and then the co-purified products were detected by western blot analysis using anti-Flag or anti-HA antibody. As shown in [Fig. 3](#), MT-Flag interacts with MT-HA and $\Delta E2$ -HA, and $\Delta E2$ -Flag with MT-HA and $\Delta E2$ -HA. Similarly, $\Delta E8$ -Flag interacts with $\Delta E8$ -HA and +E9-HA and +E9-Flag with $\Delta E8$ -HA and +E9-HA.

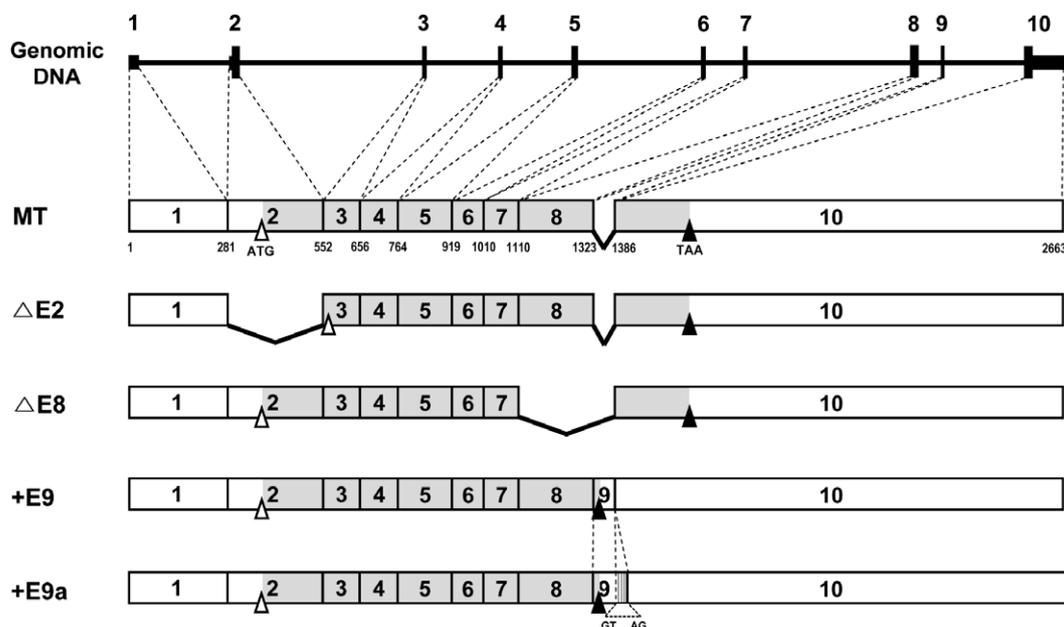


Fig. 1. Five types of alternative splice variants of hSPS1 mRNA. The black line on the top represents genomic DNA, on which exons are shown as a bar. Numbers designate exons of each predicted alternative splice variant. The open and filled area in the exons designate the untranslated region and the open reading frame, while open and closed triangles designate the position of translational initiation and termination codons, respectively. Bent lines within cDNA designate a deleted exon in each splice variant. The numbers underneath MT indicate nucleotide positions.

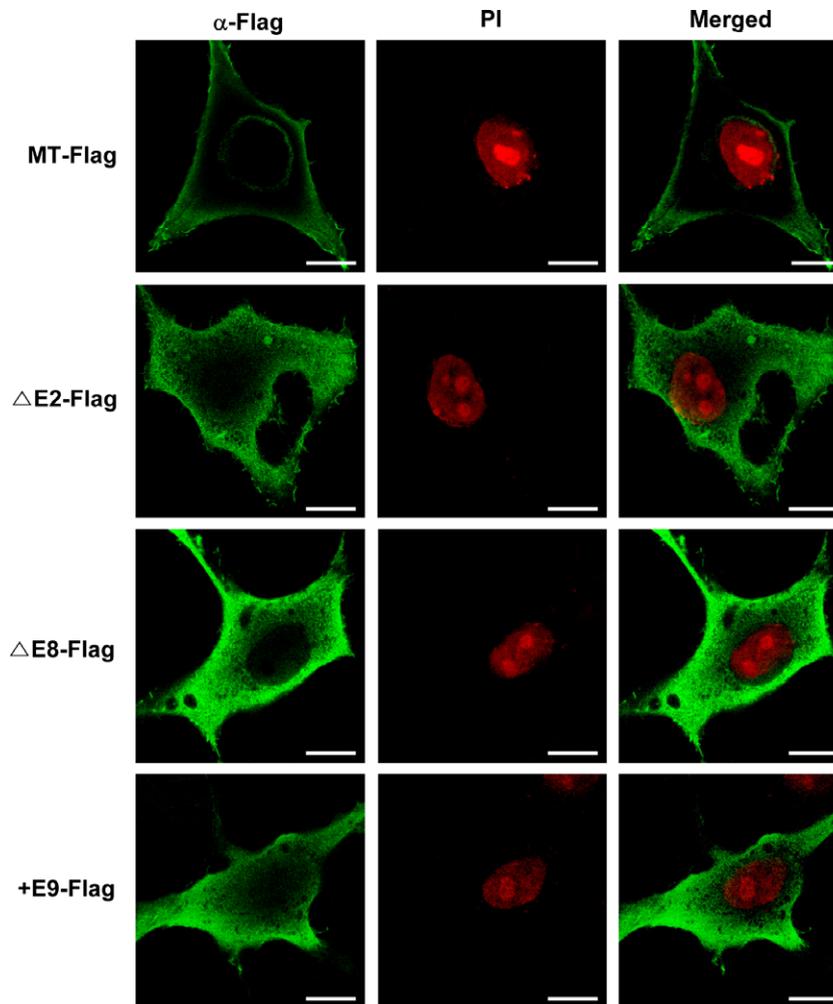


Fig. 2. Subcellular localization of hSPS1 alternative splice variants. Flag-tagged hSPS1 overexpression vectors which can express each alternative splice variant were transfected into HeLa cells. Forty-eight hours after transfection, cells were fixed and permeabilized, then hSPS1 was detected by mouse anti-Flag antibody and Alexa 488 anti-mouse IgG (green). Nucleus was stained with PI (red) and cells were observed under the confocal microscope. Scale bars represent 10 μm . (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

These results suggest that all the splice variants can make homodimers as well as heterodimers between MT and $\Delta\text{E}2$, and between $\Delta\text{E}8$ and +E9. Although it cannot be excluded that they form multimers such as trimers or tetramers, each splice variant makes specific interactions. It should be noted that MT forms a heterodimer only with $\Delta\text{E}2$, and the C-terminal sequences of them are the same suggesting C-terminal sequence is important for heterodimer formation.

3.4. Expression of alternative splice variants during cell cycle progression

It has been reported that *patufet* which is the *Drosophila* homolog of SPS1 is expressed highly in actively dividing cells and its deletion caused embryonic lethality during fly development [19,26]. More recently, it was also reported that knockdown of *Drosophila* SPS1 led to growth inhibition of cells [22]. These reports suggest that the expression of hSPS1 can be controlled during the cell cycle. We examined the expression level of alternative splice variants during the cell cycle. After synchronizing the cells at G1/S phase, the synchronized cells were then released to proceed through the cell cycle by removing aphidicolin. At 6, 12, and 18 h after release, cells were in the mid-S phase, in the G2/M phase, and in the G1/S phase, respectively. However, at 24 h after release, the cells were detected in all phases of the cell cycle (Fig. 4A).

The expression levels of each alternative splice variant were measured by real-time PCR as described in the Materials and methods. As shown in Fig. 4B, the expression of the MT variant was gradually increased until the G2/M phase and then decreased. For example, cells arrested at the G1/S phase by aphidicolin treatment showed that the MT variant expression level was decreased to 56% compared to that in asynchronous cells ($P < 0.05$). The level of MT expression in the cells of the G2/M phase was increased to 160% compared to that of cells arrested at the G1/S phase ($P < 0.05$). The levels of MT were similar at 0 h and 18 h when the cells were in G1/S phase. $\Delta\text{E}2$ expression was peaked at mid-S phase and then gradually decreased. $\Delta\text{E}8$ showed similar expression patterns as MT, but +E9/+E9a levels were decreased gradually as time lapsed suggesting the formation of +E9/+E9a was impaired by synchronization. These results suggest that alternative splicing of hSPS1 mRNA is closely related with cell cycle control.

4. Discussion

Five alternative splice variants of hSPS1 mRNA (MT, $\Delta\text{E}2$, $\Delta\text{E}8$, +E9 and +E9a) were identified in HeLa cells. Through searching the web browser of the University of California at Santa Cruz, four of these variants could be predicted with only the exception of +E9a. Interestingly, +E9a contains an extra 29 nucleotide extension of +E9. In genomic DNA, there is no gap between the 3' end of +E9

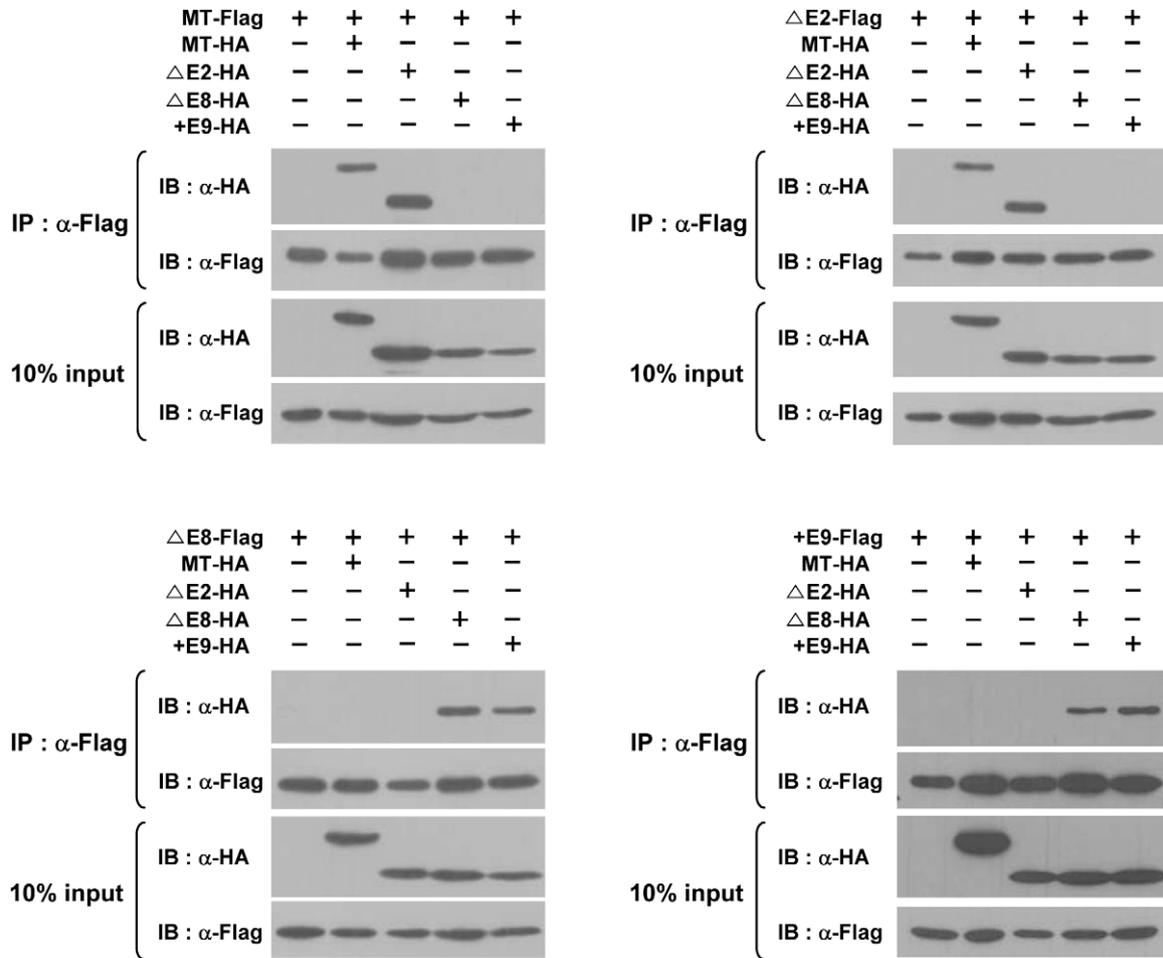


Fig. 3. Interactions between each hSPS1 splice variant. Flag or HA-tagged splice variant expression vectors were transfected in 293 cells in the combinations shown, extracts were prepared, splice variants were immunoprecipitated, electrophoresed and detected by western blotting.

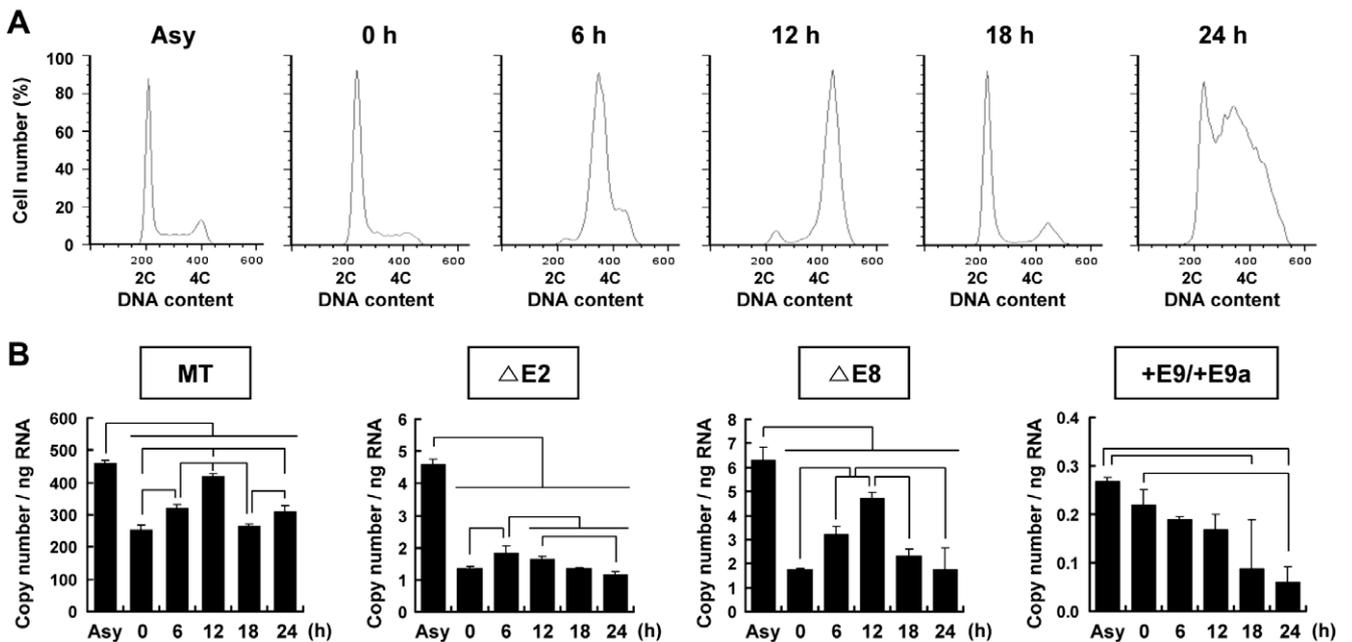


Fig. 4. Expression of hSPS1 alternative splice variants during cell cycle progression. (A) FACS (fluorescent associated cell sorter) analysis of cells after release from synchronization. X axis represents DNA contents and Y axis % of cell number. (B) Quantification of the amount of splice variants during the cell cycle. Real-time RT-PCR was performed using the same primers given in Supplementary Table S2. Experiments were performed in triplicate, and error bars denote the standard deviation from the mean of three experiments. Statistical significance was tested by pairwise Student's *t*-test with Bonferroni correction for multiple testing ($P < 0.05$). Asy represents asynchronous cells.

and the 5' end of +E9a. Furthermore, the consensus sequence of intron/exon junction (GT/AG) occurs at the 5' and 3' boundary of the additional sequence of +E9a suggesting +E9a is generated by attaching a part of intron 9 to the exon 9 by mis-splicing (see Fig. 1 and Supplementary Fig. S2).

The average amount of each splice variant was different. Approximately 90% of SPS1 population consists of MT. $\Delta E2$ and $\Delta E8$ comprise 2% of the total SPS1 mRNA, while +E9 and +E9a make up less than 1%.

Although there are 5 alternative splice variants in hSPS1 mRNA, only four proteins are produced, because there is a translational termination codon in exon 9. In terms of protein size, the MT protein is the largest, $\Delta E2$ the second, +E9/+E9a the third and $\Delta E8$ the smallest.

The hSPS1 splice variants can interact with specific combinations; MT with $\Delta E2$ and $\Delta E8$ with +E9. It should be noted that the C-terminal sequences of MT and $\Delta E2$ are the same and the C-terminus of $\Delta E8$ and +E9 are truncated. These results suggest that the C-terminus of SPS1 is important for interaction.

As described in the Introduction, SPS1 participates in Ras mediated cell signaling in *Drosophila* [20,21]. MT is located on both the plasma membrane and nuclear membrane and the other splice variants in the cytoplasm. All splice variants interact in a specific manner. Therefore, it can be speculated that the interactions of SPS1 between these molecules mediate cellular signal transductions. However, interactions of splice variants with other proteins have not been found. Because only MT is localized in the membranes, membrane localizing signals may occur in MT. However, we could not find any known common membrane localizing signals suggesting MT may interact with other membrane localizing protein(s).

Each splice variant was expressed differentially depending on the transformed cell lines derived from various tissues. For example, MT and +E9/+E9a are expressed in highest levels in BJAB cells which were derived from B cells; $\Delta E2$ was expressed in the highest level in 293 cells which were derived from kidney; and $\Delta E8$ was expressed in highest level in MCF7 cells which were derived from breast tissue. It is not clear whether these expression patterns are specific to each tissue or they are changed during transformation. However, if we compare Supplementary Fig. S3 with Supplementary Fig. S4, the expression pattern of splice variants at least in liver and kidney was changed during cell transformation. The expression level of splice variants is also regulated in a specific manner during the cell cycle and this result is consistent with the proposal that SPS1 is required for cell growth and that SPS1 is an essential gene; i.e., when *Drosophila* SPS1 was knocked out, development was halted at the embryonic stage. Our findings in the present study suggest that alternative splice variants of SPS1 play important roles in cell growth presumably by controlling cell cycle progression through their specific subcellular localization and interaction.

Acknowledgments

This work was supported by the Korea Research Foundation Grants funded by the Korean Government (MEST) (KRF-2005-070-C00086 and KRF-2008-005-J00201) to B.J.L. and in part by the Intramural Research Program of the National Institutes of Health, National Cancer Institute, Center for Cancer Research. J.Y. Kim, M.S. Shim, and K.H. Lee were supported by Brain Korea 21 Research Fellowship from the Korea Ministry of Education and Human Resources Development.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2010.05.055.

References

- [1] D.L. Hatfield, M.J. Berry, V.N. Gladyshev, Selenium: Its Molecular Biology and Role in Human Health, second ed., Springer, NY, 2006.
- [2] L. Flohé, Selenium in mammalian spermiogenesis, *Biol. Chem.* 388 (2007) 987–995.
- [3] L.V. Papp, J. Lu, A. Holmgren, K.K. Khanna, From selenium to selenoproteins: synthesis identity, and their role in human health, *Antioxid. Redox Signal.* 9 (2007) 775–806.
- [4] M.G. Boosalis, The role of selenium in chronic disease, *Nutr. Clin. Pract.* 23 (2008) 152–160.
- [5] B.J. Lee, P.J. Worland, J.N. Davis, T.C. Stadtman, D.L. Hatfield, Identification of a selenocysteyl-tRNA^{Ser} in mammalian cells that recognizes the nonsense codon, *UGA*, *J. Biol. Chem.* 264 (1989) 9724–9727.
- [6] W. Leinfelder, T.C. Stadtman, A. Böck, Occurrence *in vivo* of selenocysteyl-tRNA^{Ser} in *Escherichia coli*, *J. Biol. Chem.* 264 (1989) 9720–9723.
- [7] R. Longtin, A forgotten debate: Is selenocysteine the 21st amino acid? *J. Natl. Cancer Inst.* 96 (2004) 504–505.
- [8] B.A. Carlson, X.M. Xu, G.V. Kryukov, M. Rao, M.J. Berry, V.N. Gladyshev, D.L. Hatfield, Identification and characterization of phosphoseryl tRNA^{Ser|Sec} kinase, *Proc. Natl. Acad. Sci. USA* 101 (2004) 12848–12853.
- [9] X.M. Xu, B.A. Carlson, H. Mix, Y. Zhang, K. Saira, R.S. Glass, M.J. Berry, V.N. Gladyshev, D.L. Hatfield, Biosynthesis of selenocysteine on its tRNA in eukaryotes, *PLoS Biol.* 5 (2007) e4.
- [10] X.M. Xu, B.A. Carlson, Y. Zhang, H. Mix, G.V. Kryukov, R.S. Glass, M.J. Berry, V.N. Gladyshev, D.N. Hatfield, New developments in selenium biochemistry: selenocysteine biosynthesis in eukaryotes and archaea, *Biol. Trace Elem. Res.* 119 (2007) 234–241.
- [11] K. Forchhammer, A. Böck, Selenocysteine synthase from *Escherichia coli*. Analysis of the reaction sequence, *J. Biol. Chem.* 266 (1991) 6324–6328.
- [12] R.S. Glass, W.P. Singh, W. Jung, Z. Veres, T.D. Scholz, T.C. Stadtman, Monoselenophosphate: synthesis characterization, and identity with the prokaryotic biological selenium donor, compound SePX, *Biochemistry* 32 (1993) 12555–12559.
- [13] A. Ehrenreich, K. Forchhammer, P. Tormay, B. Veprek, A. Böck, Purification and characterisation of the enzyme catalysing selenium activation, *Eur. J. Biochem.* 206 (1992) 767–773.
- [14] W. Leinfelder, K. Forchhammer, B. Veprek, E. Zehelein, A. Böck, *In vitro* synthesis of selenocysteyl-tRNA^{UCA} from seryl-tRNA^{UCA}: Involvement and characterization of the *selD* gene product, *Proc. Natl. Acad. Sci. USA* 87 (1990) 543–547.
- [15] S.C. Low, J.W. Harney, M.J. Berry, Cloning and functional characterization of human selenophosphate synthetase, an essential component of selenoprotein synthesis, *J. Biol. Chem.* 270 (1995) 21659–21664.
- [16] M.J. Guimaraes, D. Peterson, A. Vicari, B.G. Cocks, N.G. Copeland, D.J. Gilbert, N.A. Jenkins, D.A. Ferrick, R.A. Kastelein, J.F. Bazan, A. Zlotnik, Identification of a novel *selD* homolog from Eukaryotes, Bacteria, and Archaea: is there an autoregulatory mechanism in selenocysteine metabolism? *Proc. Natl. Acad. Sci. USA* 93 (1996) 15086–15091.
- [17] X.M. Xu, B.A. Carlson, R. Irons, H. Mix, N. Zhong, V.N. Gladyshev, D.L. Hatfield, Selenophosphate synthetase 2 is essential for selenoprotein biosynthesis, *Biochem. J.* 404 (2007) 115–120.
- [18] A.V. Lobanov, D.L. Hatfield, V.N. Gladyshev, Selenoproteinless animals: selenophosphate synthetase SPS1 functions in a pathway unrelated to selenocysteine biosynthesis, *Protein Sci.* 17 (2008) 176–182.
- [19] B. Alsina, J. Serras, J. Baguña, M. Corominas, *Patufet*, the gene encoding the *Drosophila melanogaster* homologue of selenophosphate synthetase, is involved in imaginal disc morphogenesis, *Mol. Gen. Genet.* 257 (1998) 113–123.
- [20] M. Morey, F. Serras, J. Baguña, E. Hafen, M. Corominas, Modulation of the Ras/MAPK signalling pathway by the redox function of selenoproteins in *Drosophila melanogaster*, *Dev. Biol.* 238 (2001) 145–156.
- [21] M. Morey, M. Corominas, F. Serras, *DIAP1* suppresses ROS-induced apoptosis caused by impairment of the *selD/sp1* homolog in *Drosophila*, *J. Cell Sci.* 116 (2003) 4597–4604.
- [22] M.S. Shim, J.Y. Kim, H.K. Jung, K.H. Lee, X.M. Xu, B.A. Carlson, K.W. Kim, I.Y. Kim, D.L. Hatfield, B.J. Lee, Elevation of glutamine level by selenophosphate synthetase 1 knockdown induces megamitochondrial formation in *Drosophila* cells, *J. Biol. Chem.* 284 (2009) 32881–32894.
- [23] T. Tamura, S. Yamamoto, M. Takahata, H. Sakaguchi, H. Tanaka, T.C. Stadtman, K. Inagaki, Selenophosphate synthetase genes from lung adenocarcinoma cells: *Sps1* for recycling L-selenocysteine and *Sps2* for selenite assimilation, *Proc. Natl. Acad. Sci. USA* 101 (2004) 16162–16167.
- [24] A. Small-Howard, N. Morozova, Z. Stoytcheva, E.P. Forry, J.B. Mansell, J.W. Harney, B.A. Carlson, X.M. Xu, D.L. Hatfield, M.J. Berry, Supramolecular complexes mediate selenocysteine incorporation *in vivo*, *Mol. Cell. Biol.* 26 (2006) 2337–2346.
- [25] K.T. Wang, J. Wang, L.F. Li, X.D. Su, Crystal structures of catalytic intermediates of human selenophosphate synthetase 1, *J. Mol. Biol.* 390 (2009) 747–759.
- [26] B. Alsina, M. Corominas, M.J. Berry, J. Baguña, F. Serras, Disruption of selenoprotein biosynthesis affects cell proliferation in the imaginal discs and brain of *Drosophila melanogaster*, *J. Cell Sci.* 112 (1999) 2875–2884.

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.

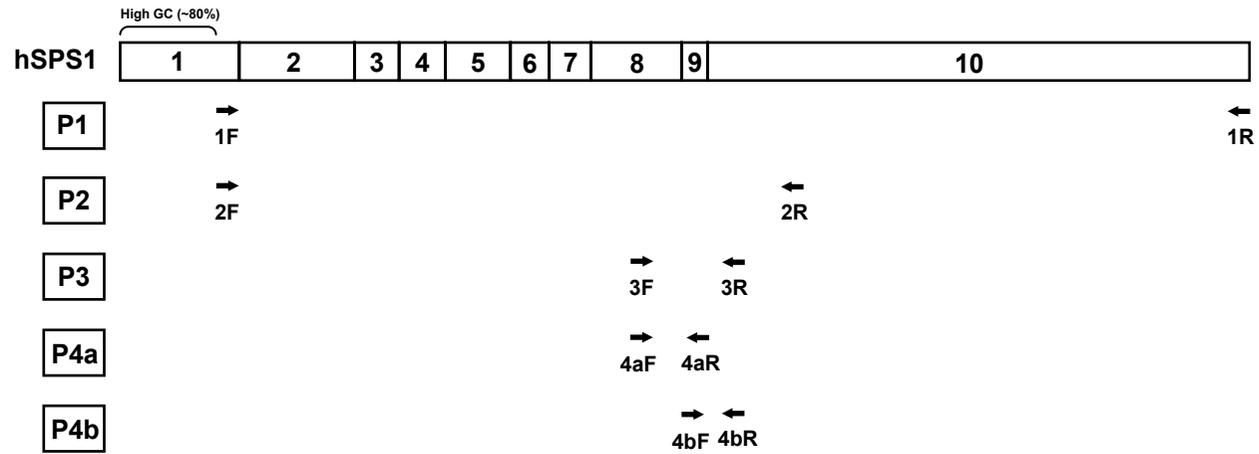
Reference used in Supplementary Data

[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.

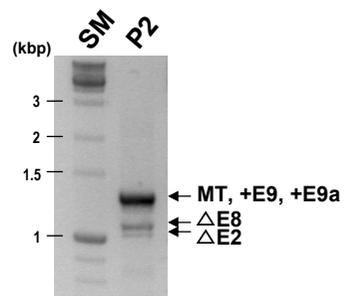
Supplementary Figures

Figure S1

A



B



C

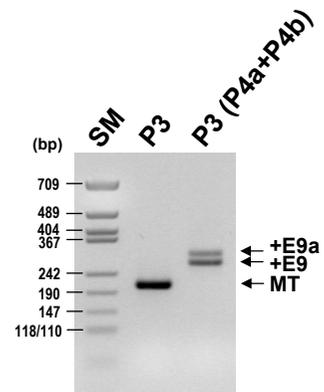


Figure S2

MT	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
ΔE2	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
ΔE8	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
+E9	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
+E9a	CAGCAGCCAGGGCCGCTCTTAAAGGGAGGTGGCCGGCCTTAAAGGACCCCCGCGGCCAGCCGGCGGGAGCGCGGGCCCGGCTCCCGCAGGGCCGC	: 100
MT	GCCGAGGCAGGCGAGCCCCGCGGGCCAGGCGGGCCCGGGCCCCCGCCGCCGCCGCCGCCGGGCGGGCATTTTTATTCAGGCGACGCTTAAG	: 200
ΔE2	GCCGAGGCAGGCGAGCCCCGCGGGCCAGGCGGGCCCGGGCCCCCGCCGCCGCCGCCGCCGGGCGGGCATTTTTATTCAGGCGACGCTTAAG	: 200
ΔE8	GCCGAGGCAGGCGAGCCCCGCGGGCCAGGCGGGCCCGGGCCCCCGCCGCCGCCGCCGCCGGGCGGGCATTTTTATTCAGGCGACGCTTAAG	: 200
+E9	GCCGAGGCAGGCGAGCCCCGCGGGCCAGGCGGGCCCGGGCCCCCGCCGCCGCCGCCGCCGGGCGGGCATTTTTATTCAGGCGACGCTTAAG	: 200
+E9a	GCCGAGGCAGGCGAGCCCCGCGGGCCAGGCGGGCCCGGGCCCCCGCCGCCGCCGCCGCCGGGCGGGCATTTTTATTCAGGCGACGCTTAAG	: 200
MT	GGAGCCCGGGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGGATCCATAAAGAACCC	: 300
ΔE2	GGAGCCCGGGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGGATCCATAAAGAACCC	: 282
ΔE8	GGAGCCCGGGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGGATCCATAAAGAACCC	: 300
+E9	GGAGCCCGGGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGGATCCATAAAGAACCC	: 300
+E9a	GGAGCCCGGGCGCCCGGTGCATTGTGGGAGCGGCCCGGGCCCGTTTCGGGAGGAGGCGGAGGGCGCAAAGCGAGCCCGTGGATCCATAAAGAACCC	: 300
MT	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGTTCTACGCGGGAGTCCTTAACC CGGAAAGTTACGAATT	: 400
ΔE2	-----	: -
ΔE8	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGTTCTACGCGGGAGTCCTTAACC CGGAAAGTTACGAATT	: 400
+E9	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGTTCTACGCGGGAGTCCTTAACC CGGAAAGTTACGAATT	: 400
+E9a	AGCCAACCCGAGAGGGAGGGGAGGGGCTGAGCTGTGAGGAGAGCGGGGCCAAGAACCATGTTCTACGCGGGAGTCCTTAACC CGGAAAGTTACGAATT	: 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

ACCATGTT

ACCATGTT

ACCATGTT

ACCATGTT

E2

GCAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT

GCAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT

GCAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT

GCAATGGATACTTGTGTCAATTCCTTTGAGGCACGGTGGGCTTT

E3

E4

E5

MT AACCCAATGAATTTATCATGCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGTGCACAAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG : 1000
ΔE2 AACCCAATGAATTTATCATGCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGTGCACAAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG : 729
ΔE8 AACCCAATGAATTTATCATGCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGTGCACAAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG : 1000
+E9 AACCCAATGAATTTATCATGCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGTGCACAAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG : 1000
+E9a AACCCAATGAATTTATCATGCCAGACAATGCAGTGCCAGGGGACGTGCTGGTGTGCACAAAACCCCTGGGGACACAGGTGGCAGTGGCTGTGCACCAGTG : 1000

E6

MT GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC : 1100
ΔE2 GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC : 829
ΔE8 GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC : 1100
+E9 GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC : 1100
+E9a GCTGGATATCCCTGAGAAATGGAATAAGATTAAGATTAACTAGTGGTCACCCAAGAAGATGTAGAGCTGGCCTACCAGGAGGCGATGATGAACATGGCGAGGCTC : 1100

E7

MT AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC : 1200
ΔE2 AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC : 929
ΔE8 AACAGGACAG----- : 1110
+E9 AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC : 1200
+E9a AACAGGACAGCTGCAGGACTCATGCACACGTTCAATGCCACGCGGCCACTGACATCACGGGCTTCGGGATTTTGGGCCATGCGCAGAACCTGGCCAAGC : 1200

MT AGCAGAGGAACGAGGTGTCTGTTGTAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA : 1300
ΔE2 AGCAGAGGAACGAGGTGTCTGTTGTAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA : 1029
ΔE8 ----- : -
+E9 AGCAGAGGAACGAGGTGTCTGTTGTAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA : 1300
+E9a AGCAGAGGAACGAGGTGTCTGTTGTAATTCACAACCTCCCGGTGCTGGCCAAGATGGCTGCGGTGAGCAAGGCTGCGGAAACATGTTGCGCCTCATGCA : 1300

E8

MT CGGGACCTGCCCGGAGACTTCAG----- : 1323
ΔE2 CGGGACCTGCCCGGAGACTTCAG----- : 1052
ΔE8 ----- : -
+E9 CGGGACCTGCCCGGAGACTTCAGATGTGCAGTAATTTATTAGAGGAAACATAAATGGAGGATAAACGGGAGATGGAGCAAGCGTAG----- : 1386
+E9a CGGGACCTGCCCGGAGACTTCAGATGTGCAGTAATTTATTAGAGGAAACATAAATGGAGGATAAACGGGAGATGGAGCAAGCGTAGGTAGAGAGAGCCAT : 1400

E9

MT -----ECGGCCTTCTGATCTGTTTACCAGTGTGACCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA : 1408
ΔE2 -----ECGGCCTTCTGATCTGTTTACCAGTGTGACCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA : 1137
ΔE8 -----ECGGCCTTCTGATCTGTTTACCAGTGTGACCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA : 1195
+E9 -----ECGGCCTTCTGATCTGTTTACCAGTGTGACCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA : 1471
+E9a CAGACAGCAGTGCAGECGGCCTTCTGATCTGTTTACCAGTGTGACCAAGCAGCTCGGTTCTGTGCAGAGATAAAGTCCCCAAATATGGTGAAGGCCACCA : 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 ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTCACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1908
ΔE2 ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTCACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1637
ΔE8 ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTCACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1695
+E9 ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTCACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 1971
+E9a ACTTATCTCTTGCCGCATAGGCTTTTTAAAGGTGCTTTTCACATAGCACAGGCATTACCCGTAGTCGTGTCAAATAGCAGTTGGTGTCTTCATTTTATGTA : 2000

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

MT AGGGTTGCATGATTCCCTTTGAGTGTTTGATTTGAAAAGCCTAGTCTTGTCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 2108
ΔE2 AGGGTTGCATGATTCCCTTTGAGTGTTTGATTTGAAAAGCCTAGTCTTGTCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 1837
ΔE8 AGGGTTGCATGATTCCCTTTGAGTGTTTGATTTGAAAAGCCTAGTCTTGTCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 1895
+E9 AGGGTTGCATGATTCCCTTTGAGTGTTTGATTTGAAAAGCCTAGTCTTGTCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 2171
+E9a AGGGTTGCATGATTCCCTTTGAGTGTTTGATTTGAAAAGCCTAGTCTTGTCTCTCAAGAGCATCTCGGACCCAGAACATTCTCCAGTAGTGCATTGAGTTC : 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 ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATATTTAATTTTTTAAA : 2600
ΔE2 ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATATTTAATTTTTTAAA : 2329
ΔE8 ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATATTTAATTTTTTAAA : 2387
+E9 ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATATTTAATTTTTTAAA : 2663
+E9a ACGCTTTTTGTGATCTGTGTTAAAAACCTGCATTGCCAACATTGCAGCTCGAACTTAAACTTGTTATTCAAATAAATATTTAATTTTTTAAA : 2692

Figure S3

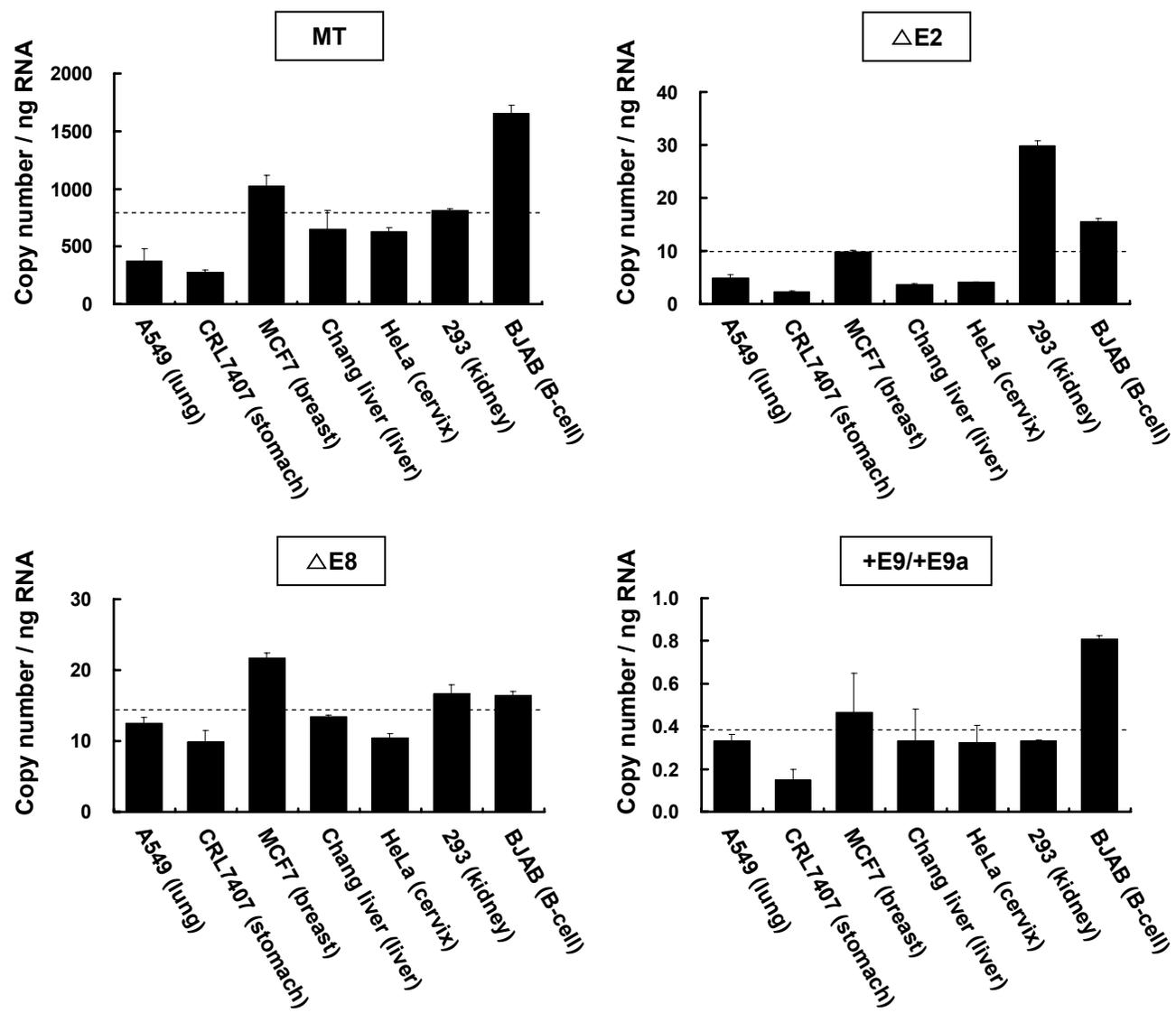
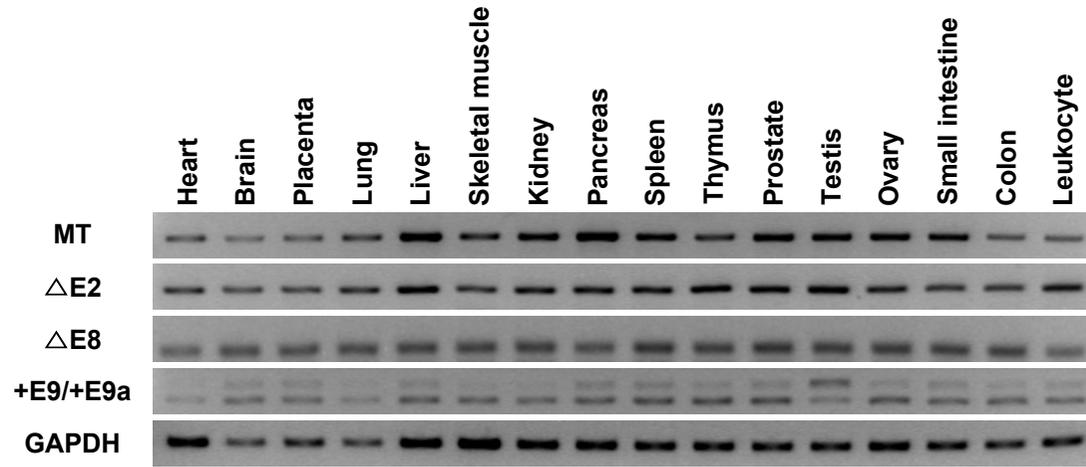
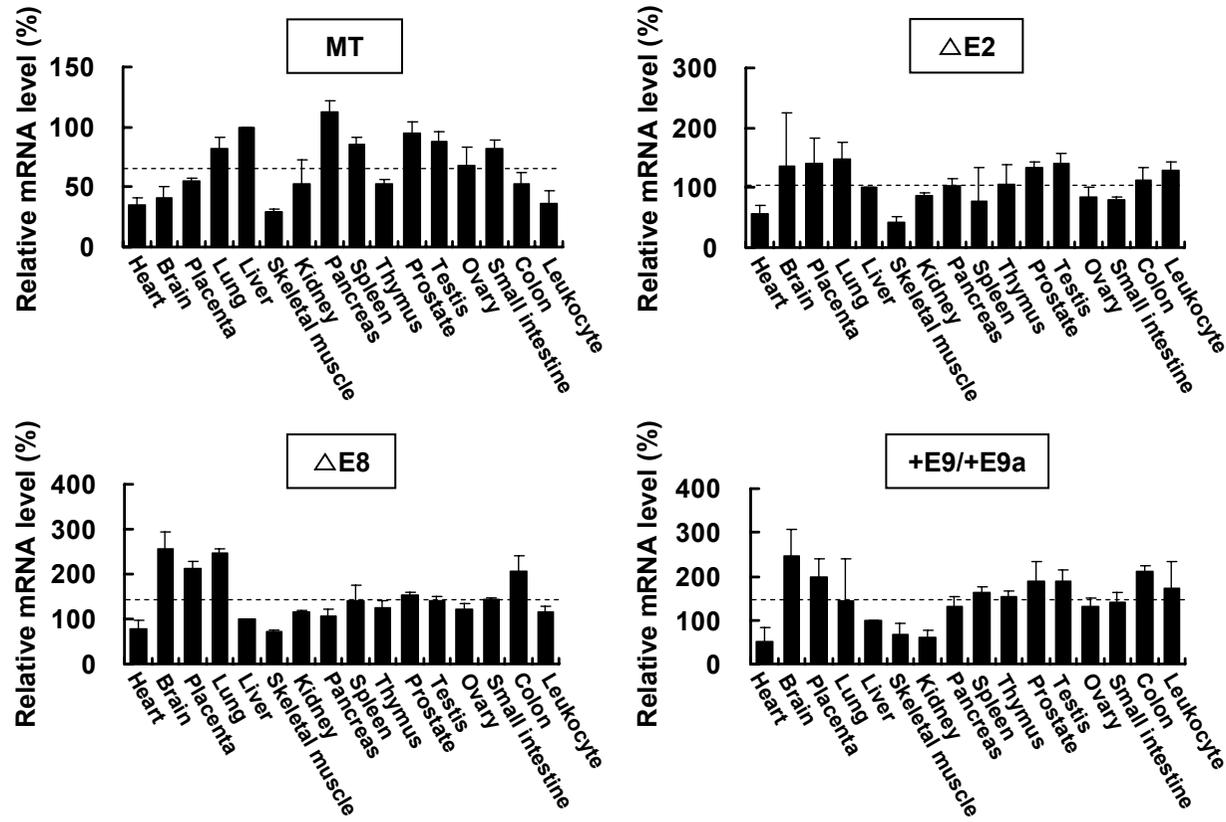


Figure S4

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