Lack of polymorphisms in the coding region of the highly conserved gene encoding transcription elongation factor S-II (*TCEA1*)

Takahiro Ito[1], Kent Doi[2, 3], Naoko Matsumoto[4], Fumiko Kakihara[4], Eisei Noiri[2], Setsuo Hasegawa[4], Katsushi Tokunaga[3], Kazuhisa Sekimizu[1, *]

[1] *Division of Developmental Biochemistry, Graduate School of Pharmaceutical Sciences, University of Tokyo, Tokyo, Japan;*
[2] *Department of Nephrology and Endocrinology, Graduate School of Medicine, University of Tokyo, Tokyo, Japan;*
[3] *Department of Human Genetics, Graduate School of Medicine, University of Tokyo, Tokyo, Japan;*
[4] *Sekino Clinical Pharmacology Clinic, Tokyo, Japan.*

ABSTRACT: Transcription elongation factor S-II stimulates mRNA chain elongation catalyzed by RNA polymerase II. S-II is highly conserved among eukaryotes and is essential for definitive hematopoiesis in mice. In the present study, we report the identification of five novel nucleotide variations in the human S-II gene in the Japanese population. All five variations were located in introns, and no polymorphisms were found in the protein-coding region, suggesting strong negative selection during gene evolution. Together with the SNPs (single nucleotide polymorphisms) reported in the National Center for Biotechnology Information SNP database, our results provide tools for evaluating the role of S-II in complex genetic diseases, such as congenital hematopoietic disorders.

Key Words: Transcription elongation factor, S-II, Japanese population, nucleotide variation

Introduction

Transcription factor S-II, also known as TFIIS, was originally identified as a specific stimulatory protein of RNA polymerase II *in vitro* (*1*). Further biochemical, structural, and genetic studies demonstrated that S-II stimulates transcript elongation by promoting the read-through of transcription arrest sites by RNA polymerase II (*2-5*). The genes encoding S-II have been identified in many eukaryotes including budding yeast, fruit fly, mouse, and human (*6-9*). The human S-II gene, designated *TCEA1*, was initially reported to be a 2.5-kb intronless gene mapped on 3p22- > p21.3 (*10*). We previously reported that the murine S-II gene consists of 10 exons and maps on the proximal region of mouse chromosome 1, which is syntenic to human chromosome 8q (*11*). Consistent with the synteny between the mouse and human chromosomes, recent progress in the human genome project identified another gene composed of 10 exons on 8q11 (UCSC Genome Browser, http://genome.ucsc.edu/). These results suggest that the S-II gene on 8q encodes a functional S-II protein, whereas the other gene on 3q is a pseudogene as proposed by HGNC (HUGO Gene Nomenclature Committee, http://www.gene.ucl.ac.uk/nomenclature/). In the present study, we screened 125 Japanese volunteers for single nucleotide polymorphisms in *TCEA1*, and found 5 nucleotide variations that were not previously reported in the NCBI (National Center for Biotechnology Information) SNP database.

Materials and Methods

Study subjects

Blood samples were collected after obtaining written informed consent from 125 Japanese adult volunteers for this study, which was approved by the ethics committees of the Graduate School of Pharmaceutical Sciences, University of Tokyo, and Sekino Clinical Pharmacology Clinic. Genomic DNAs were purified from whole blood samples using MagExtractor-genome (TOYOBO, Shiga, Japan) and an automated system (SX-6G; Precision System Science, Chiba, Japan).

Amplification of human S-II genomic DNA fragments

**Correspondence to:* Graduate School of Pharmaceutical Sciences, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan;
e-mail: sekimizu@mol.f.u-tokyo.ac.jp

Received May 1, 2007
Accepted May 15, 2007

Table 1. Primer sequences for amplification and sequencing used in this study

Amplicon ID	Feature	PCR Primer sequence (5' to 3')	Amplified Region[1]	Sequencing Primer Sequence (5' to 3')	Sequenced region[1]
E1	Covers exon 1	FW - AGCGATCTGCAGTCAGTTGGTAGC RV - CAGGGACTGGAAATACAAGAGCGA	-256 - 284	FW - TTCGTAAGGAAGGGGGCCTA RV - same as PCR	-143 - 284
E2	Covers exon 2	FW - AGGTGGTGCTGTTGCTCCTTATC RV - GAGATTTCACTGCTACTGCCAAC	11500 - 11940	FW, RV - same as PCR	11500 - 11940
E3	Covers exon 3	FW - GCAGCTGGTGCTCTATGAAGTAATCCATG RV - TCGCCTTTATTACGAGGCACTGCTTTCAG	21663 - 22326	FW, RV - same as PCR	21663 - 22326
E4	Covers exon 4	FW - GGGAGTGTTGACTGAACTTGCATTG RV - AGACAGGGGAATTGATGCAGGAAGT	28189 - 28529	FW, RV - same as PCR	28189 - 28529
E56	Covers exons 5& 6	FW - AGCTGTCATTCTCTGGTCCCATC RV - ACCGTGCCTGGCATTTGCATATG	33501 - 35236	FW - same as PCR RV - CCCGTATCAGCAACAACTTC	33501 - 34162 and
				FW - GCAGTGAAGTGATCAGATTC RV - same as PCR	34944 - 35236
E7	Covers exon 7	FW - GGTGACAGAGTGAGACTCCTTCTACCCA RV - CCTAATCAGTCAGCAGCCGTCAACATCC	37057 - 39065	FW - CCAGGTCAAGTGGTTGTCCTGCCT RV - AGATCGTGCCATTGCACTCCAGCCT	37318 - 38076
E8	Covers exon 8	FW - GACAGCCAGCCTGTGACACTAGAAG RV - TCCGTGGTTTCAGGCATCCACTCAG	42629 - 43256	FW, RV - same as PCR	42629 - 43256
E9	Covers exon 9	FW - GGAAGTGCGTGCTCTGTTGACAT RV - GTCAGGGAGTACAAGAGTACTGT	51489 - 51870	FW, RV - same as PCR	51489 - 51870
E10	Covers exon 10	FW - GGAAAGTCAGGCACTTAAGATAGG RV - GGAAAACTAGTTGCTTGGCCTAGTT	53847 - 54107	FW, RV - same as PCR	53847 - 54107

[1] Nucleotide numbering is according to the mutation nomenclature *(den Dunnen and Antonarakis 2000).*

for sequencing

To amplify the coding regions and exon-intron boundaries from genomic DNA, we developed primer sets using the human chromosome 8 sequence assembled by the International Human Genome Sequencing Consortium (Nucleotide number 55041469-55097366, May 2004 freeze, hg17, http://genome.ucsc.edu/). Using the primers listed in Table I, we amplified nine genomic DNA fragments covering all the coding exons of human S-II by polymerase chain reactions (PCR) using individual genomic DNAs extracted from blood samples as templates. The reactions were performed with LA Taq DNA polymerase and its reaction buffer (Takara Bio Inc. Shiga, Japan) supplemented with 2.5 mM magnesium chloride. Primer concentrations in the reaction were 0.4 μM each, except for the E56 and E8 amplicons in which 0.2 μM was used. For E1 amplification, we added dimethyl sulfoxide at a final concentration of 4% to the reaction. Amplification conditions were as follows; for the E1 amplification, 2 min at 94°C followed by 35 cycles of 30 sec at 94°C, 30 sec at 54°C, 1 min at 72°C, followed by a final extension of 1 min at 72°C; E2 amplification, 2 min at 94°C followed by 35 cycles of 30 sec at 94°C, 30 sec at 54°C, 30 sec at 72°C, followed by a final extension of 1 min at 72°C; E3, E4, E7, E9, and E10 amplifications, 2 min at 94°C followed by 35 cycles of 30 sec at 94°C, 30 sec at 56°C, 1 min at 72°C, followed by a final extension of 1 min at 72°C; E56 and E8 amplifications, 2 min at 94°C followed by 35 cycles of 30 sec at 94°C, 30 sec at 58°C, 30 sec at 72°C, followed by a final extension of 1 min at 72°C. Specific amplifications of the PCR products were evaluated by agarose gel electrophoresis before performing the sequencing reactions.

Nucleotide sequence determination and screening for polymorphisms

Nucleotide sequences of the PCR products were determined using BigDye terminator cycle sequence reaction mix Ver 3.1 (Applied Biosystems) and the sequencing primers listed on Table 1. Sequencing reactions with E1 and E7 amplicons contained dimethyl sulfoxide at final concentrations of 10% and 5%, respectively. After purification, the samples were sequenced on an ABI PRISM 3100 Genetic Analyzer (Applied Biosystems). Nucleotide variations were first identified by SeqScape software (Applied Biosystems), and verified visually. We used the human chromosome 8 genomic sequence from the International Human Genome Sequencing Consortium (Nucleotide number 55041469-55097366, May 2004 freeze, hg17) as a reference sequence.

Results and Discussion

By designing sets of PCR primers within the introns (Table 1), we could specifically amplify genomic DNA fragments encoding human S-II gene located on 8q11. Sequencing of these PCR products and comparison of these sequences with a reference sequence revealed a total of five single nucleotide variations in introns 2, 4, and 7 (Table 2; nucleotides are numbered according to the mutation nomenclature (*12*)). Comparison of our data with the SNPs deposited in the NCBI SNP database (dbSNP build 127) indicated that these five nucleotide variations were novel. Chi-square analysis revealed that the observed frequencies of the genotypes did not

Table 2. Nucleotide variations from *TCEA1* locus

ID[1]	Position[1]	Genotype	Number of genotype	Allele frequencies of		*P* value for Hardy-Weinberg equilibrium test
				Major allele	Minor allele	
		TT	101	T	C	
IVS2 (-274) T > C	21802	TC	23	0.9	0.1	0.80
		CC	1			
		AA	124	A	T	
IVS4 (-220) A > T	33647	AT	1	0.996	0.004	0.96
		TT	0			
		CC	124	C	T	
IVS4 (-169) C > T	33698	CT	1	0.996	0.004	0.96
		TT	0			
		GG	120	G	A	
IVS7 + 46 G > A	37809	GA	5	0.98	0.02	0.82
		AA	0			
		TT	124	T	C	
IVS7 (-155) T > C	42800	TC	1	0.996	0.004	0.96
		CC	0			

[1] Nucleotide numbering is according to the mutation nomenclature *(den Dunnen and Antonarakis 2000).*

deviate from the expectations according to the Hardy-Weinberg equilibrium (Table 2). No variations were detected in the protein-coding regions, suggesting the operation of negative selection due to a strong functional constraint of the transcription elongation factor S-II protein. Consistent with this notion, the genes encoding S-II are highly conserved among eukaryotes (*2*), and gene disruption in mice leads to embryonic lethality due to defects in definitive hematopoiesis (*13*). It remains to be determined whether nucleotide polymorphisms in the S-II gene are associated with congenital disorders of hematopoiesis.

Acknowledgements

This work was supported by grants from the Japan Society for the Promotion of Science, and the Ministry of Education, Science, Sports and Culture. We thank Taku Fujinawa for his excellent technical assistance.

References

1. Sekimizu K, Nakanishi Y, Mizuno D, Natori S. Purification and preparation of antibody to RNA polymerase II stimulatory factors from Ehrlich ascites tumor cells. Biochemistry 1979;18:1582-1588.
2. Wind M, Reines D. Transcription elongation factor SII. Bioessays 2000;22:327-336.
3. SivaRaman L, Reines D, Kane CM. Purified elongation factor SII is sufficient to promote read-through by purified RNA polymerase II at specific termination sites in the human histone H3.3 gene. J Biol Chem 1990;265:14554-14560.
4. Kettenberger H, Armache KJ, Cramer P. Architecture of the RNA polymerase II-TFIIS complex and implications for mRNA cleavage. Cell 2003;114:347-357.
5. Shimoaraiso M, Nakanishi T, Kubo T, Natori S. Transcription elongation factor S-II confers yeast resistance to 6-azauracil by enhancing expression of the SSM1 gene. J Biol Chem 2000;275:29623-29627.
6. Chen HC, England L, Kane CM. Characterization of a HeLa cDNA clone encoding the human SII protein, an elongation factor for RNA polymerase II. Gene 1992;116:253-258.
7. Hirashima S, Hirai H, Nakanishi Y, Natori S. Molecular cloning and characterization of cDNA for eukaryotic transcription factor S-II. J Biol Chem 1988;263:3858-3863.
8. Nakanishi T, Nakano A, Nomura K, Sekimizu K, Natori S. Purification, gene cloning, and gene disruption of the transcription elongation factor S-II in Saccharomyces cerevisiae. J Biol Chem 1992;267:13200-13204.
9. Marshall TK, Guo H, Price DH. Drosophila RNA polymerase II elongation factor DmS-II has homology to mouse S-II and sequence similarity to yeast PPR2. Nucleic Acids Res 1990;18:6293-6298.
10. DiMarco SP, Glover TW, Miller DE, Reines D, Warren ST. Transcription elongation factor SII (TCEA) maps to human chromosome 3p22 - > p21.3. Genomics 1996;36:185-188.
11. Ito T, Seldin MF, Taketo MM, Kubo T, Natori S. Gene structure and chromosome mapping of mouse transcription elongation factor S-II (Tcea1). Gene 2000;244:55-63.
12. den Dunnen JT, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. Hum mutat 2000;15:7-12.
13. Ito T, Arimitsu N, Takeuchi M, Kawamura N, Nagata M, Saso K, Akimitsu N, Hamamoto H, Natori S, Miyajima A, Sekimizu K. Transcription elongation factor S-II is required for definitive hematopoiesis. Mol Cell Biol 2006;26:3194-3203.