

SUPPLEMENTAL INFORMATION

Identification and characterization of a novel PPAR α -regulated and 7 α -hydroxyl bile acid-preferring cytosolic sulfotransferase mL-STL (Sult2a8)

Lu Feng,^{1,*} Yee-Lok Yuen,^{1,*} Jian Xu,^{1,*} Xing Liu,^{*} Martin Yan-Chun Chan,^{*} Kai Wang,^{*} Wing-Ping Fong,^{*} Wing-Tai Cheung,^{2,§} and Susanna Sau-Tuen Lee^{2,*}

School of Life Sciences, Faculty of Science,^{*} and School of Biomedical Sciences, Faculty of Medicine,[§]
The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong SAR

¹These authors contributed equally to the work.

²Co-corresponding authors

To whom correspondence should be addressed

Susanna Sau-Tuen Lee, PhD
School of Life Sciences,
Faculty of Science,
The Chinese University of Hong Kong,
Shatin, New Territories,
Hong Kong SAR
Phone: (852) 3943-6333
Fax number: (852) 2603-5646
E-mail: lee2022@cuhk.edu.hk

Wing-Tai Cheung, PhD
School of Biomedical Sciences,
Faculty of Medicine,
The Chinese University of Hong Kong,
Shatin, New Territories,
Hong Kong SAR
Phone: (852) 3943-6104
E-mail: wtcheung@cuhk.edu.hk

Running title: PPAR α regulates bile acid sulfonation during fasting

SUPPLEMENTARY METHODS

Mass spectrometry analysis

The column-purified recombinant His-mL-STL and preparative SDS-PAGE gel semi-purified recombinant mL-STL proteins were confirmed by mass spectrometry analysis. The proteins were first resolved in a 14% SDS-PAGE gel and stained with Coomassie blue R for 1 h. After the gel was destained, the bands corresponding to the His-mL-STL and mL-STL proteins were excised and punched into small discs using a Monoject 202 blunt-end needle (Covidien, Mansfield, MA; Cat. No. 8881-202314). The gel discs were equilibrated with 25 mM NH₄HCO₃ for 1 h with vortexing at room temperature. They were then dehydrated and destained by three consecutive washes each with 50% acetonitrile (ACN)/25 mM ammonium bicarbonate (NH₄HCO₃) for 10 min followed by a final wash with 100% ACN for 10 min. The gel was dried completely by a CentriVap vacuum concentrator (Labconco, Kansas City, MO) and then the proteins in the gel discs were reduced by 10 mM DTT/25 mM NH₄HCO₃ at 60°C for 30 min. After that the proteins were alkylated in 55 mM iodoacetamide/25 mM NH₄HCO₃ for 30 min in dark and the gel discs were then washed, dehydrated, dried, and digested with 400 ng trypsin/20 µl of 25 mM NH₄HCO₃ (Promega, Madison, WI; Cat. No. V5111) at 37°C overnight. The digested peptides were extracted with 25 mM NH₄HCO₃, 5% trifluoroacetic acid (TFA)/50% ACN followed by 100% ACN for 10 min in a Tru-sweep™ ultrasonic cleaner (Crest, Trenton, NJ). The extracted peptides were dried using a speed vacuum and resuspended in 10 µl of 0.1% TFA. The peptide solution was purified with a ZipTip C₁₈ column (Millipore, Billerica) pre-wetted with 50% ACN/0.1% TFA and equilibrated with 0.1% TFA. The peptides were then loaded onto the pre-washed ZipTip C₁₈ column and it was washed with 0.1% TFA. Finally, the purified peptides were eluted in 3 µl of 50% ACN/0.1% TFA and stored frozen at -80°C. Each peptide sample (0.5 µl) was spotted on a MALDI plate and was allowed to air dry. The sample spot was covered with 0.5 µl of 10 mg/ml matrix (α -cyano-4-hydroxy-cinnamic acid) and the sample was analyzed by an Applied Biosystems 4700 proteomics analyzer. The peptide

mass fingerprint spectrum of each peptide spot was generated automatically by plotting the % intensity of positively charged ions against mass-to-charge (m/z) ratio of the ions in the range of 500 - 4000. Peptide analysis was performed using data-dependent acquisition of one MS scan followed by MS/MS scans of the ten most abundant ions in each MS scan (1). Data from peptide mass fingerprinting was processed by the Mascot software using the GPS ExplorerTM Workstation. The database searching was performed using the NCBI non-redundant database with no missed cleavage allowed and peptide tolerance of 100 ppm and MS/MS tolerance of 0.5 Da (2).

SUPPLEMENTARY REFERENCES

1. Tong, W. Y., Y. M. Liang, V. Tam, H. K. Yip, Y. T. Kao, K. M. Cheung, K. W. Yeung, and Y. W. Lam. 2010. Biochemical characterization of the cell-biomaterial interface by quantitative proteomics. *Mol. Cell. Proteomics.* **9**: 2089-2098.
2. Poetsch, A., D. Schlusener, C. Florizone, L. Eltis, C. Menzel, M. Rogner, K. Steinert, and U. Roth. 2008. Improved identification of membrane proteins by MALDI-TOF MS/MS using vacuum sublimated matrix spots on an ultraphobic chip surface. *J. Biomol. Tech.* **19**: 129-138.

Supplementary Table S1. Primers used in this study

Name	Sequences (5' - 3')	Application
AP8	ACGACTCACTATAAGGGTTTTTTTTAA	FDD
ARP1	ACAATTCACACAGGACGACTCCAAG	
5'-RACE primer R	TGGTGCATTGTGGAAGTCTAAGAA	5'-RACE (5'/3' RACE)
5'-RACE primer I	CTGAGGCCAACATCTGATTAGCTCTGT	5'-RACE (GeneRacer)
3'-RACE primer S	AAATGCACCAGATGGCTCCACCCATA	3'-RACE (GeneRacer)
M13 reverse	GTCATAGCTTTCCCTG	Sequencing of 5'- and 3'-RACE clones
M13 forward	CTGGCCGTCGTTTAC	
mL-STL-SEQ110	CAATGAGGCAGAACACT	
mL-STL-SEQ380	ATTGTCTGCTTGATTCTG	
mL-STL-SEQ638	CAGGGAAAGGAACAAAGTG	
mL-STL-SEQ898	AAGGATGATATCTCAGTC	
mL-STL-SEQ6	CTTGGACCACCTCTGTAC	
mL-STL-SEQ7	CAAGCCATGAAGCAGTAG	
mL-STL-FP278-Pst I	<i>aaaactgcagc</i> ATGACAGATGAATTCTGTGGA Pst I	Cloning of His-mL-STL-mpRSETA expression vector
mL-STL-RP1319-Hind III	<i>cccaagctt</i> GTGCCAGGATTGAACTCAGA Hind III	
mL-STL-SEQ-FP379	GCCAATGAGCGCACACCA	Sequencing of His-mL-STL-mpRSETA expression vector
mL-STL-SEQ-RP658	GCCACATGCATGTTCGAA	
mL-STL-SEQ-RP1231	GCTTCCTTCGGGCTTTG	
mL-STL-SEQ-FP3895(-26)	GCGCGTTGGCCGATTCAT	
mL-STL-FP278-Nde I	<i>ggaaatttccat</i> ATGACAGATGAATTCTGTGGA Nde I	Cloning of mL-STL-mmpRSET expression vector
mL-STL-RP1319-Hind III	<i>cccaagctt</i> GTGCCAGGATTGAACTCAGA Hind III	
T7	TAATACGACTCACTATAAGGG	Sequencing of mL-STL-mmpRSET expression vector
mL-STL-SEQ-380	ATTGTCTGCTTGATTCTG	
mL-STL-SEQ-638	CAGGGAAAGGAACAAAGTG	
mL-STL-SEQ-898	AAGGATGATATCTCAGTC	
mL-STL-SEQ5	GTATTGTGCTAGAGAAGT	
mL-STL-FP19	ACAGATTCTCCGACCTTA	Northern blot analysis
mL-STL-RP1884	CTGAGGCCAACATCTGATTAGC	

The flanking regions of the restriction enzyme sites are in lower cases and in italics.

A**ARP1**

```

ACAATTCAC ACAGGACGAC TCCAAGGATA AAGAATCTTG TTAACGACCA TGAGGATCTG AGTTCAATCC TGGCACCAAC 80
ATAGTTGAAG GAAAGAAGTA GTTCTGAAA TTATTCTTAG ACTTCACAA ATGCACCATG ATGGCTCCAC CCATAAAATT 160
TCTGATTA AAATATTCT TGATGTCAA TCCTTAAATA ATGAAATAAT TATACAAATT TGATTCTGA GTGTTTCAGA 240
CATTTAAAG AAAAGTAAA TTCAGAAACA ATCATAACTC TGCTCTTGG CCACCTCTGT ACACCTCTAT GTCTGCCCT 320
GTGGGTCAAA GGTGACTTTTC AAATCTTAGT AAGGATTTC CTAAAATATG ACCAGCTGT GAAAATCCAG GTCATGAGCC 400
TACTCAGTTT TTATGGTCT CTACTTATGG AAGTAGAAAT TTGTACAGTT TGATAAGAAA GAGATTCAAG TGTTTTGAT 480
GTTCTAAC TTCTCCACC CCTTCCAGG GAAGCTTCAT GCCTGGACAA GCCATGAAGC AGTAGAGTGC TCCTCATACC 560
TTTGAAGTAG AAAAGATGCC TGCCGTGG TGTAACAGGC TGAAAAGAT TCCAATACAG AGCTAATCAG ATTGGCCTCA 640
GAAAAAAATAC TAACCTATTG GTTTGTTCC TGTCCTTCAG TGAGAAGAC TTCTGTATTT TTAAATAC AATTTTATT 720
CTTCAACGA ATTTAAAAAA AACACCTTG GAACAACAG ACAACAAAAA GTATAATTAC TTCTCTATT GCTTGATTG 800
AAGAAATGCT TTAAAGTATC ATCATTTTAATTTCTCC ATCATTTCAAC TTATAATCAA TAATGTCGGT AAACAAAAAT 880
TAAAAAAAAA AAAGCCCTAT AGTGAGTCGT 910

```

AP8**B**

ref|NM_175250.5| *Mus musculus* RIKEN cDNA 2810007J24 gene (2810007J24Rik), transcript variant 1, mRNA

Length: 2147

Identities = 874/875 (99%)

mL-STL	18	GACTCCAAGGATAAAGAATCTGTTAACGACCATGAGGATCTGAGTTCAATCCTGGCACC	77
Riken	1261	GACTCCAAGGATAAAGAATCTGTTAACGACCATGAGGATCTGAGTTCAATCCTGGCACC	1320
mL-STL	78	AACATAGTTGAAGGAAAAGAAGTAGTTCTGAAATTATTCTTAGACTTCCACAAATGCACC	137
Riken	1321	AACATAGTTGAAGGAAAAGAAGTAGTTCTGAAATTATTCTTAGACTTCCACAAATGCACC	1380
mL-STL	138	ATGATGGCTCCACCCATAAAATTCTGTAATTAAAATATTCTGATGTCAAATCCTTAA	197
Riken	1381	ATGATGGCTCCACCCATAAAATTCTGTAATTAAAATATTCTGATGTCAAATCCTTAA	1440
mL-STL	198	ATAATGAAATAATTACAAATTGATTTCTGAGTGTGTTCAAGACATTAAAGAAAAGTA	257
Riken	1441	ATAATGAAATAATTACAAATTGATTTCTGAGTGTGTTCAAGACATTAAAGAAAAGTA	1500
mL-STL	258	AAATTCAAGAACAACTCATAACTCTGCTCTGGACCACTCTGTACACCTCTATGCTGGC	317
Riken	1501	AAATTCAAGAACAACTCATAACTCTGCTCTGGACCACTCTGTACACCTCTATGCTGGC	1560
mL-STL	318	CCTGTGGTCACAGGTGACTTTCAAATCTTAGTAAGGATTTCTAAATATGACCAGCT	377
Riken	1561	CCTGTGGTCACAGGTGACTTTCAAATCTTAGTAAGGATTTCTAAATATGACCAGCT	1620
mL-STL	378	GTTGAAAATCCAGGTATGAGCCTACTCAGTTTATGGTTCTCTACTTATGGAAGTAGA	437
Riken	1621	GTTGAAAATCCAGGTATGAGCCTACTCAGTTTATGGTTCTCTACTTATGGAAGTAGA	1680
mL-STL	438	AATTTGTACAGTTGATAAGAAAGAGATTTCAGCTGTTTGTATGTTCTTAACCTCTCCC	497
Riken	1681	AATTTGTACAGTTGATAAGAAAGAGATTTCAGCTGTTTGTATGTTCTTAACCTCTCCC	1740
mL-STL	498	ACCCCTTCCAGGGAGCTTCATGCCTGGACAAGCCATGAAGCAGTAGAGTGCTCCTCAT	557
Riken	1741	ACCCCTTCCAGGGAGCTTCATGCCTGGACAAGCCATGAAGCAGTAGAGTGCTCCTCAT	1800
mL-STL	558	ACCTTGAAAGTAGAAAAGATGCCTGCCTGTTGTAAACAGGCTGGGAAGATTCCAATA	617
Riken	1801	ACCTTGAAAGTAGAAAAGATGCCTGCCTGTTGTAAACAGGCTGGGAAGATTCCAATA	1860
mL-STL	618	CAGAGCTAATCAGATTGGCCTCAGAAAAAAACTAACTAATTCTGTTTGTCTGTCTT	677
Riken	1861	CAGAGCTAATCAGATTGGCCTCAGAAAAAAACTAACTAATTCTGTTTGTCTGTCTT	1920
mL-STL	678	CAGTGTAGAAGACTCTGTTTTAAATAACAATTTTATTCTTCAACGAATTAA	736
Riken	1921	CAGTGTAGAAGACTCTGTTTTAAATAACAATTTTATTCTTCAACGAATTAA	1980
mL-STL	737	AAAAAACACCTTGGAAACAACAAGAACAAACAAAAGTATAATTACTCTTATTGCTTG	796
Riken	1981	AAAAAACACCTTGGAAACAACAAGAACAAACAAAAGTATAATTACTCTTATTGCTTG	2040
mL-STL	797	ATTGAAGAAATGCTTAAAGTATCATCATATTAAATATTCTCCATCATTCACTTATAA	856
Riken	2041	ATTGAAGAAATGCTTAAAGTATCATCATATTAAATATTCTCCATCATTCACTTATAA	2100
mL-STL	857	TCAATAATGTCGGTAAACAAAAATTAAAAAA	891
Riken	2101	TCAATAATGTCGGTAAACAAAAATTAAAAAA	2135

Supplementary Figure S1

Supplementary Figure S1. DNA sequence of a 910 bp mL-STL FDD fragment and its alignment to 2810007J24Rik cDNA sequence. (A) DNA sequence of the partial mL-STL cDNA fragment isolated from the fluorescent differential display (FDD) analysis. The sequences of ARP1 and AP8 primers are underlined and bolded. (B) Alignment of the partial mL-STL cDNA sequence with the 3'-end of Riken cDNA 2810007J24. The mL-STL cDNA sequence from nucleotides 18 to 891 showed 99% similarity to the 3'-end of Riken cDNA sequence from nucleotides 1261 to 2135. An extra A nucleotide (boxed) was found in Riken cDNA sequence compared with the mL-STL.

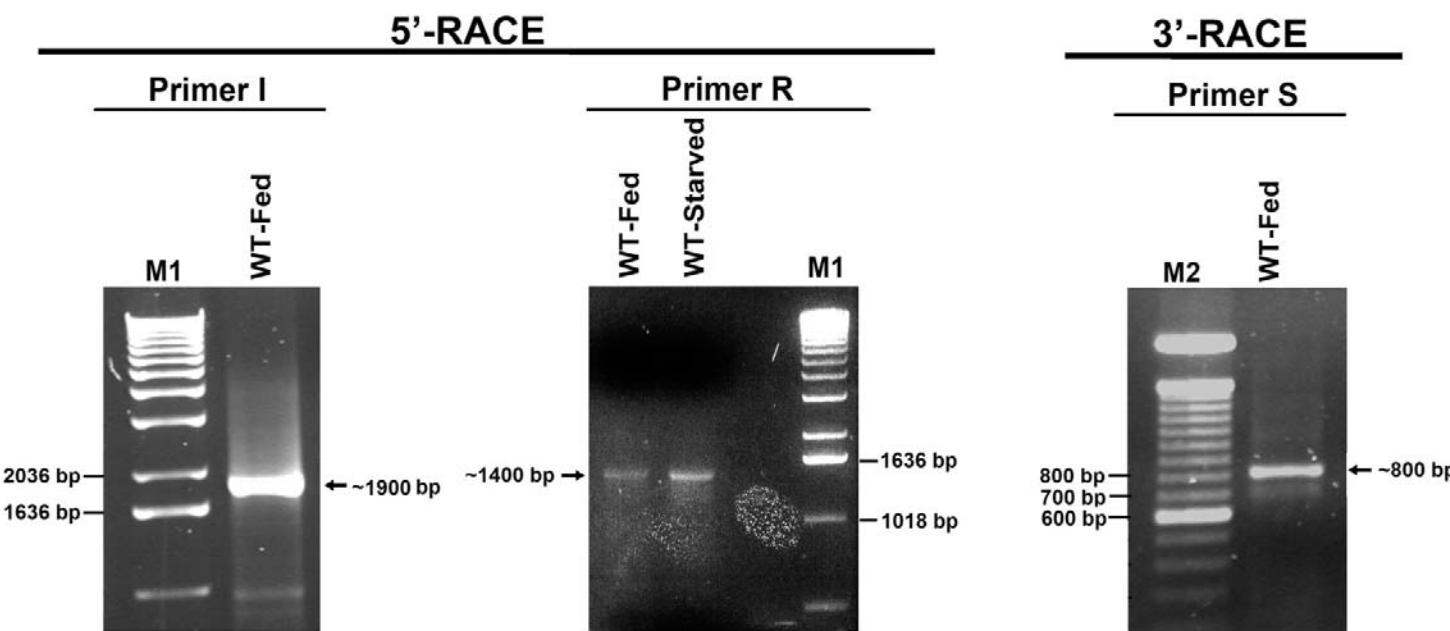
A

ARP1
ACAATTCAC ACAGGACGAC TCCAAGGATA AAGAACCTTG TTAACGACCA TGAGGATCTG AGTTCAATCC TGGCACCAAC ATAGTTGAAG GAAAGAAGTA GTTCTGAAA TT~~T~~**TTCTTAG** 120

3'-RACE primer S→
ACTTCCACAA ATGCCACATG ATGGCTCCAC CCATAAAATT TCTGTAATTA AAATATTTCT TGATGTCAA TCCTTAAATA ATGAAATAAT TATACAAATT TGATTTCTGA GTGTTTCAGA 240
CATTTTAAAG AAAAGTAAAA TTCAGAAACA ATCATAACTC TGCTCTTGG CAACCTCTGT ACACCTCTAT GTCTGGCCCT GTGGGTACAA GGTGACTTTC AAATCTTAGT AAGGATTTTC 360
CTAAATATG ACCAGCTGTT GAAAATCCAG GTCATGAGCC TACTCAGTTT TTATGGTTCT CTACTTATGG AAGTAGAAAT TTGTACAGTT TGATAAGAAA GAGATTGAGC TGTTTTGAT 480
GTTTCTTAAC TTCTCCCACC CCTTTCCAGG GAAGCTTCAT GCCTGGACAA GCCATGAAGC AGTAGAGTGC TCCTCATACC TTTGAAGTAG AAAAGATGCC TGCCTGTTG TGTAACAGGC 600

← 5'-RACE primer I
TGGGGAAAGAT TCCAAATCAG AGCTAATCAG ATTGGCCTCA GAAAAAATAC TAACCTTATTG GTTTTGTCC TGTCTTTAG TGAGAAGAC TTCTGTATTT TTTAAATAC AATTTTATTT 720
CTTTCAACGA ATTTAAAAAA AACACCTTG GAACACAAG AACACACAAA GTATAATTAC TTCTTCTATT GCTTGATTG AAGAAATGCT TTAAAGTATC ATCATATTAA ATATTTCTCC 840

AP8
ATCATTTCAC TTATAATCAA TAATGTCGGT AAACAAAAAT AAAAAAAAAA AAAGCCCTAT AGTGAGTCGT 910

B

Supplementary Figure S2. 5'- and 3'-RACE primer design and amplicon products. (A) Location of the 5'- and 3'-mL-STL gene-specific rapid amplification of cDNA ends (RACE) primers. The sequences of 5'- and 3'-RACE primers are boxed and bolded, and the PCR amplification orientations are indicated. The arbitrary ARP1 and anchored AP8 primer sequences used for amplification of the partial mL-STL cDNA fragment in the FDD analysis are bolded and underlined. (B) 5'- and 3'-RACE mL-STL PCR products. Total RNA (5 µg/reaction) from the livers of a wild-type fed (WT-Fed) and a 72 h-starved (WT-Starved) mice was used in the RACE experiment.

Exon 1

	1	15 16	30 31	45 46	60 61	75 76	90	
5' #20-Starved(R)	-	-	-	-	-	-	-	0
5' #29-Starved(R)	-	-	-	-	-	-	-	0
5' #23-Starved(R)	-	-	-	-	-	-	-	0
5' #35-Starved(R)	-	-	-	-	-	-	-	0
5' #37-Starved(R)	-	-	-	-	-	-	-	0
5' #34-Starved(R)	-	-	-	-	-	-	-	0
5' #39-Starved(R)	-	-	-	-	-	-	-	0
5' #4-Fed(R)	-	-	-	-	-	-	-	0
5' #33-Starved(R)	-	-	-	-	-	-	-	0
5' #22-Fed(I)	-	-	-	-	-	-	-	0
5' #17-Fed(I)	-	-	-	-	-	-	-	0
5' #37-Fed(I)	-	-	-	-	-	-	-	0
5' #17-Starved(R)	-	-	-	-	-	-	-	0
5' #100-Fed(I)	-	-	-	-	-	-	-	0
5' #11-Starved(R)	-	-	-	-	-	-	-	0
5' #25-Starved(R)	-	-	-	-	-	-	-	0
5' #1-Fed(R)	-	-	-	-	-	-	-	0
5' #30-Starved(R)	-	-	-	-	-	-	-	0
5' #28-Starved(R)	-	-	-	-	-	-	-	0
5' #6-Fed(I)	-	-	-	-	-	-	-	0
5' #69-Fed(I)	-	-	-	-	-	-	-	0
5' #9-Starved(R)	-	-	-	-	-	-	-	0
5' #31-Starved(R)	-	-	-	-	-	-	-	16
5' #35-Fed(I)	-	-	-	-	-	-	-	68
5' #23-Fed(I)	-	-	-	-	-	-	-	72
5' #9-Fed(I)	-	-	-	-	-	-	-	83
5' #12-Starved(R)	-	-	-	-	-	-	-	87
5' #2-Starved(R)	Riken_v1	TATCTTGTTAGTTAC	ACCAACGATTCTCCC	GACCTTACTCCATGT	AAAAGCCTATACCCCT	CAGTGTTCATTCA	ACATTCTCGGTCA	90
Riken_v2	TATCTTGTTAGTTAC	ACCAACGATTCTCCC	GACCTTACTCCATGT	AAAAGCCTATACCCCT	CAGTGTTCATTCA	ACATTCTCGGTCA	90	
				-AGTAGTAGT	TCAGTTCACCCCTGA	GTTACACAGTAGTT		40

Exon 2

	91	105 106	120 121	135 136	150 151	165 166	L10	
5' #20-Starved(R)	-	-	-	-	-	-	-	12
5' #29-Starved(R)	-	-	-	-	-	-	-	34
5' #23-Starved(R)	-	-	-	-	-	-	-	34
5' #35-Starved(R)	-	-	-	-	-	-	-	34
5' #37-Starved(R)	-	-	-	-	-	-	-	34
5' #34-Starved(R)	-	-	-	-	-	-	-	34
5' #39-Starved(R)	-	-	-	-	-	-	-	34
5' #4-Fed(R)	-	-	-	-	-	-	-	34
5' #33-Starved(R)	-	-	-	-	-	-	-	34
5' #22-Fed(I)	-	-	-	-	-	-	-	34
5' #17-Fed(I)	-	-	-	-	-	-	-	34
5' #37-Fed(I)	-	-	-	-	-	-	-	34
5' #17-Starved(R)	-	-	-	-	-	-	-	36
5' #100-Fed(I)	-	-	-	-	-	-	-	36
5' #11-Starved(R)	-	-	-	-	-	-	-	36
5' #25-Starved(R)	-	-	-	-	-	-	-	36
5' #1-Fed(R)	-	-	-	-	-	-	-	36
5' #30-Starved(R)	-	-	-	-	-	-	-	36
5' #28-Starved(R)	Riken_v1	TATGACTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	88
Riken_v2	TATGACTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	106	
5' #31-Starved(R)	GTTAGTCTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	106	
5' #35-Fed(I)	GTTAGTCTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	158	
5' #23-Fed(I)	GTTAGTCTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	162	
5' #9-Fed(I)	GTTAGTCTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	173	
5' #12-Starved(R)	GTTAGTCTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	177	
5' #2-Starved(R)	GTTAGTCTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	180	
Riken_v1	GTTAGTCTAAAGAAG	ACTTCAGAACAGAT	TTGCTTCTCAATT	CCCTGCTTCACAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	180	
Riken_v2	TGTGTTCAAGGATGG	TCACATAAAACCTG	CAAAGCAAATCTA	TATAGGCAT	CAAC	TTGATTCTTAACCA	ATGAGGCAGAACAC	130

	181	195 196	210 211	225 226	240 241	255 256	270	
5' #20-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCC		102
5' #29-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #23-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #35-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #37-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #34-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #39-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #4-Fed(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #33-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #22-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #17-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #37-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #17-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		124
5' #100-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #11-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #25-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #1-Fed(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #30-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #28-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #6-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #69-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		126
5' #9-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		178
5' #31-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		196
5' #35-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		248
5' #23-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		252
5' #9-Fed(I)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		263
5' #12-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		267
5' #2-Starved(R)	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		270
Riken_v1	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		270
Riken_v2	TGATCATTTGTTCTT	CACTATAAAATTGAG	ACTCAAAGCACCCTG	GTGATCTCCACAG	AAGAGGTCAAACAGA	TAATTCCTCAAAGCT		220

Supplementary Figure S3 (p. 1)

Supplementary Figure S3 (p. 2)

Supplementary Figure S3 (p. 3)

Supplementary Figure S3 (p. 4)

Supplementary Figure S3 (p. 5)

	1351	1365	1366	1380	1381	1395	1396	1410	1411	1425	1426	1440	
◀ 5'-RACE primer R													
5' #20-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1213
5' #29-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #23-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #35-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #37-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #34-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #39-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #4-Fed(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #33-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #22-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1191
5' #17-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1286
5' #37-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1286
5' #17-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1235
5' #100-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1288
5' #11-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1237
5' #25-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1237
5' #1-Fed(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1237
5' #30-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1237
5' #28-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1237
5' #6-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1288
5' #69-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1288
5' #9-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1289
5' #35-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1410
5' #23-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1414
5' #9-Fed(I)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1425
5' #12-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1291
5' #2-Starved(R)	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCA										1381
Riken_v1	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1432
Riken_v2	AGTCCCTGAAATTAT	TCTTAGACTTCACA	AATGCACCATGATGG	CTCCACCCATAAAA	TTCTGAATTAAAAT	ATTTCTTGATGTC	AA						1382
1441	1455	1456	1470	1471	1485	1486	1500	1501	1515	1516	1530		
5' #22-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAAGTAAATTCA	GAAACAATCATTA	ACT						1281
5' #17-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1376
5' #37-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1376
5' #100-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1378
5' #6-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1378
5' #69-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1378
5' #35-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1500
5' #23-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1504
5' #9-Fed(I)	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1515
Riken_v1	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1522
Riken_v2	ATCCTTAAATATGAA	AATAATTATACAAAT	TTGATTTCTGAGTGT	TTGATTTCTGAGTGT	TTCAAGACATTAAA	GAAACAATCATTA	ACT						1472
1531	1545	1546	1560	1561	1575	1576	1590	1591	1605	1606	1620		
5' #22-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1371
5' #17-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1466
5' #37-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1466
5' #100-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1468
5' #6-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1468
5' #69-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1468
5' #35-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1590
5' #23-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1594
5' #9-Fed(I)	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1605
Riken_v1	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1612
Riken_v2	CTGCTCTGGACCAC	TTCTGTACACCTCA	TGTCTGGCCCTGTGG	GTCACAGGTGACTTT	CAAATCTTAGTAAGG	ATTTCTTAAATAT							1562
1621	1635	1636	1650	1651	1665	1666	1680	1681	1695	1696	1710		
5' #22-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1461
5' #17-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1556
5' #37-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1556
5' #100-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1558
5' #6-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1558
5' #69-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1558
5' #35-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1680
5' #23-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1684
5' #9-Fed(I)	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1695
Riken_v1	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1702
Riken_v2	GACCAGCTGTTGAAA	ATCCAGGTATGAGC	CTACTCAGTTTTAT	GTTGATCTACTTTAT	GAAGTAGAAAATTGT	ACAGTTTGATAAGAA							1652
1711	1725	1726	1740	1741	1755	1756	1770	1771	1785	1786	1800		
5' #22-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1551
5' #17-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1646
5' #37-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1646
5' #100-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1648
5' #6-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1648
5' #69-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1648
5' #35-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1770
5' #23-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1774
5' #9-Fed(I)	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1785
Riken_v1	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1792
Riken_v2	AGAGATTCACTGTTG	TTTGATGTTCTTAA	CTCTTCCCACCCCT	TCCAGGGAGCTTC	TGCCCTGGACAAAGCA	TGAAGCAGTAGAGTG							1742
1801	1815	1816	1830	1831	1845	1846	1860	1861	1875	1876	1890		
◀ 5'-RACE primer I													
5' #22-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1641
5' #17-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1736
5' #37-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1736
5' #100-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1738
5' #6-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1738
5' #69-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1738
5' #35-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1738
5' #23-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1774
5' #9-Fed(I)	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1785
Riken_v1	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1792
Riken_v2	CTCCTCATACCTTG	AAGTAGAAAAGATGC	CTGCGTGTGTTGTA	ACAGGCTGGGGAGA	TTCCAATACAGAGCT	AATCAGATTGGCCTC							1742

Supplementary Figure S3 (p. 6)

	1891	1905	1906	1920	1921	1935	1936	1950	1951	1965	1966	1980	
5' #22-Fed(I)	AG												1643
5' #17-Fed(I)	AG												1738
5' #37-Fed(I)	AG												1738
5' #100-Fed(I)	AG												1740
5' #6-Fed(I)	AG												1740
5' #69-Fed(I)	AG												1740
5' #35-Fed(I)	AG												1862
5' #23-Fed(I)	AG												1866
5' #9-Fed(I)	AG												1877
Riken_v1	AGAAAAAAATCAAC	TTATTCGTTTGTTC	CTGTCTTCAGTGT	GAAGACTTCTGTATT	TTTTAAAATACAATT	TTATTTCTTCACG							1972
Riken_v2	AGAAAAAAATCAAC	TTATTCGTTTGTTC	CTGTCTTCAGTGT	GAAGACTTCTGTATT	TTTTAAAATACAATT	TTATTTCTTCACG							1922
	1981	1995	1996	2010	2011	2025	2026	2040	2041	2055	2056	2070	
Riken_v1	AATTAAAAAAAC	ACCTTGAAACAACA	AGAACAAACAAAAGTA	TAATTACTCTCTCA	TTGCTTGATTTGAAG	AAATGCTTTAAGTA							2062
Riken_v2	AATTAAAAAAAC	ACCTTGAAACAACA	AGAACAAACAAAAGTA	TAATTACTCTCTCA	TTGCTTGATTTGAAG	AAATGCTTTAAGTA							2012
	2071	2085	2086	2100	2101	2115	2116	2130	2131	2145	2146	2160	
Riken_v1	TCATCATATTAATA	TTCTCCATCATTTC	ACTTATAATCAATAA	TGTCGGTAAACAAAA	ATTAAAAAAAAAACCA	ATCATGACTA----							2147
Riken_v2	TCATCATATTAATA	TTCTCCATCATTTC	ACTTATAATCAATAA	TGTCGGTAAACAAAA	ATTAAAAAAAAAACCA	ATCATGACTA----							2097

Supplementary Figure S3 (p. 7)

Supplementary Figure S3. Alignment of 28 mL-STL 5'-RACE clone nucleotide sequences. All 5'-RACE clones showed identical nucleotide sequence except that a 95 bp fragment (boxed and shaded) was absent in two clones 5' #22-Fed(I) and 5' #12-Starved(R), suggesting the presence of alternative splicing variants. Ten (L1 – L10) different 5'-termini are indicated by downward arrows from nucleotides 1 to 169 relative to clone 5' #2-Starved(R). An extra short fragment (TTTTTCAG) was found at the end of exon 6 in clone 5' #12-Starved(R). The mismatched nucleotides are indicated by boxed nucleotides throughout the sequences.

Supplementary Figure S4 (p. 1)

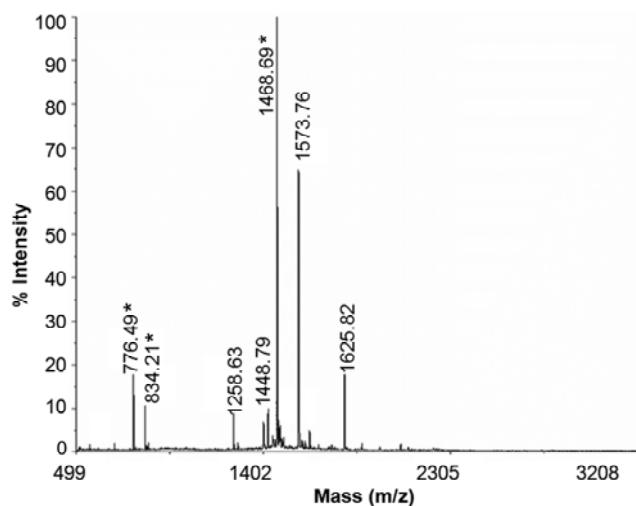
3' #3-Fed(S)	451	465	466	480	481	495	496	510	511	525	526	540
3' #13-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #24-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #20-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						519
3' #1-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #14-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #21-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #2-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #9-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #16-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #23-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
3' #25-Fed(S)	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						518
Riken_v1												1889
Riken_v2	ACCTTGAAGTAGAA	AAAGATGCCCTGCCTGT	TTGTGTAACAGCGCTG	GGGAA=GATTCCAAAT	ACAGAGCTAATCAGA	TTGGCCTCAGAAAAAA						1839
3' #3-Fed(S)	541	555	556	570	571	585	586	600	601	615	616	630
3' #13-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #24-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #20-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						609
3' #1-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #14-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #21-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #2-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #9-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #16-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #23-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
3' #25-Fed(S)	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						608
Riken_v1												1979
Riken_v2	ATACTAACATTTCG	TTTGTCTCTGTCTT	TCAGTGTAGAAGACTT	CTGTATTTTTAAA	ATACAATTTTATTC	TTTCAACGAATTAA						1929
3' #3-Fed(S)	631	645	646	660	661	675	676	690	691	705	706	720
3' #13-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #24-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #20-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #1-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #14-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #21-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #2-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #9-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #16-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #23-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
3' #25-Fed(S)	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						697
Riken_v1												2069
Riken_v2	AAAAAA-CACCTTTG	GAACACAAAGAACAA	CAAAGATATAATTAC	TTCTCTATTGCTTG	CATTGAAGAAATGCT	TAAAGTATCATCAT						2019
	721	735	736	750	751	765	766	780	781	795	796	810
	PAS											
3' #3-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAACAAATCATGA	--AAAAAAAAAAAAAA						785
3' #13-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAA--CAATCATGA	--AAAAAAAAAAAAAA						784
3' #24-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						785
3' #20-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						787
3' #1-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						782
3' #14-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						784
3' #21-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						785
3' #2-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						787
3' #9-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	CTAAAAAAAAAAAAAA						787
3' #16-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						785
3' #23-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						785
3' #25-Fed(S)	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	--AAAAAAAAAAAAAA						785
Riken_v1												2147
Riken_v2	ATTTAATTTCTCC	ATCATTTCACTTATA	ATCAATAATGTCGGT	AAACAAAAATTAAA	AAAAAA-CAATCATGA	CTA-----						2097

Supplementary Figure S4 (p. 2)

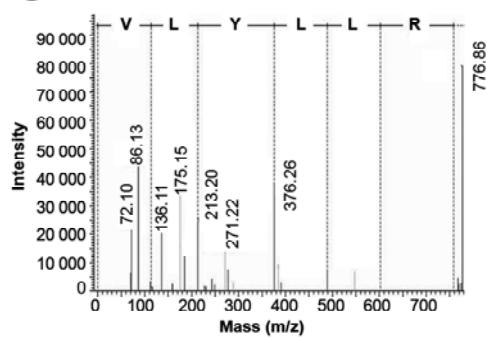
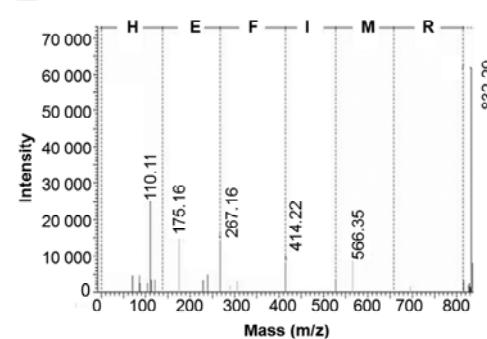
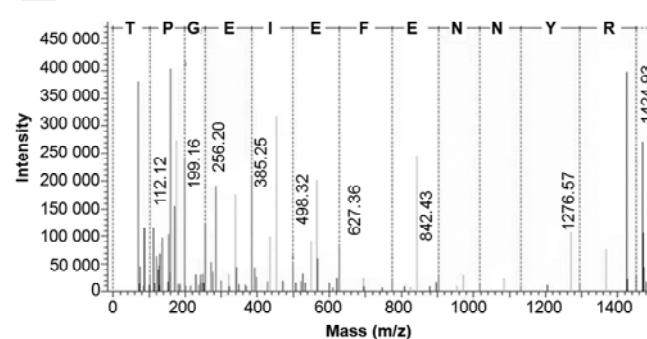
Supplementary Figure S4. Alignment of 12 mL-STL 3'-RACE clone nucleotide sequences. All 3'-RACE clones showed identical nucleotide sequence except the presence of some mismatched nucleotides as indicated by boxed nucleotides throughout the sequences. The longest 3'-RACE cDNA was 797 bp and found in clones 3'#9-Fed(S) and 3'#16-Fed(S), while the shortest 786 bp in clones 3'#13-Fed(S) and 3'#14-Fed(S). The solid triangles show the four different polyadenylation cleavage site (PACS) variants. The polyadenylation signal (PAS) ATTAAA is double-underlined.

mL-STL2	MTDEFILWIEGIPFPTVYYSQEIIIREVRDRF	30
mL-STL1	MTDEFILWIEGIPFPTVYYSQEIIIREVRDRF	30
mL-STL2	VVRDEDTIIVT YPKSGTHW LNEIVCLILTK	60
mL-STL1	VVRDEDTIIVT YPKSGTHW LNEIVCLILTK	60
	N-terminal/5'-PSB motif	
mL-STL2	GDPTWVQSTIANERTPWIEFENNYRILNSK	90
mL-STL1	GDPTWVQSTIANERTPWIEFENNYRILNSK	90
mL-STL2	EGPRLMASLLPIQLFPKSF FSSKA KVIYLI	120
mL-STL1	EGPRLMASLLPIQLFPKSF FSSKA KVIYLI	120
	FSSKA motif	
mL-STL2	RNPRDVLVSGY HYFNALKQGKEQVPWKIYF	150
mL-STL1	RNPRDVLVSGY HYFNALKQGKEQVPWKIYF	150
	3'-PB motif	
mL-STL2	ENFLQGK-----	157
mL-STL1	ENFLQGKSYFGSWFEHACGWISLRKRENIL	180
mL-STL2	-----	157
mL-STL1	VLSYEQLKKDTRNTIKKICEFLGENLESGE	210
mL-STL2	-----	157
mL-STL1	LELVILKNISFQIMKERMISQSCLSNIKHE	240
mL-STL2	-----	157
mL-STL1	FIMRKGITGDWKNHFT VAQAEAFDKAFQEK	270
	C-terminal/3'-PSB motif	
mL-STL2	-----	157
mL-STL1	AADFPQELFSWE	282

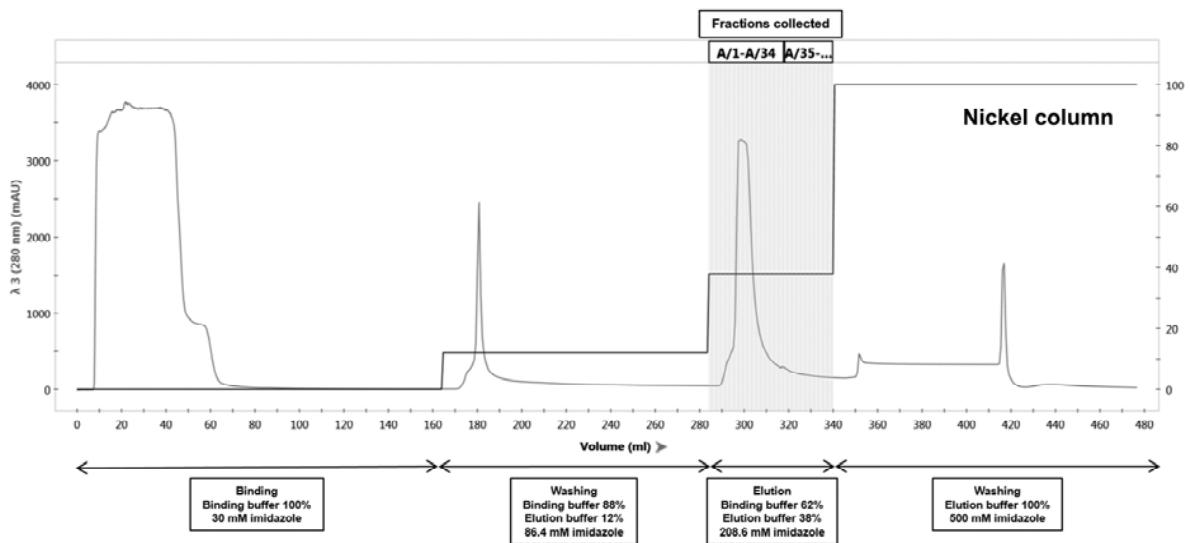
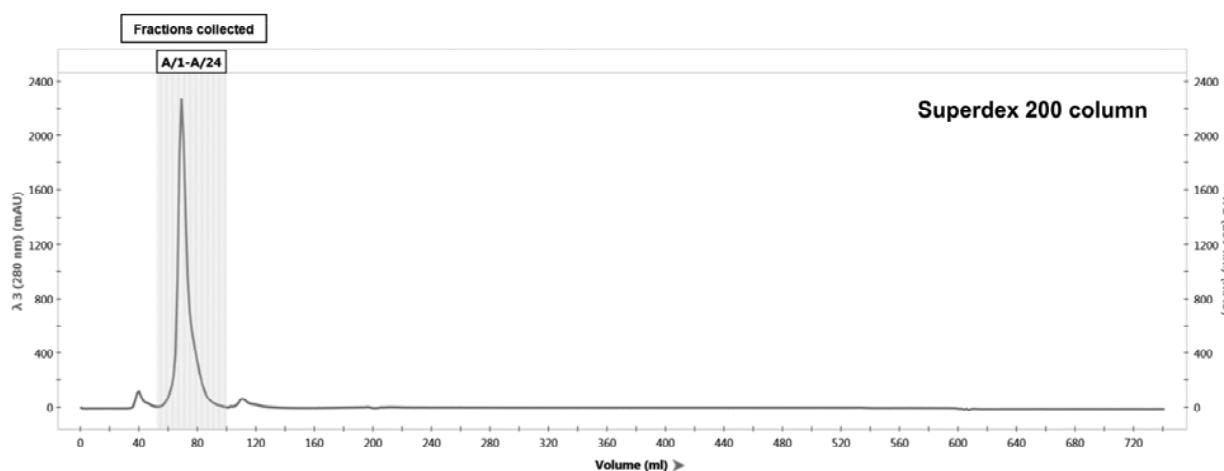
Supplementary Figure S5. Comparison of mL-STL1 and mL-STL2 amino acid sequences. The amino acid residues of mL-STL1 and mL-STL2 are 282 and 157, respectively. The N-terminal conserved sequence [N-terminal/5'-phosphosulfate binding (PSB) motif], SULT2 FSSKA signature motif sequence (FSSKA motif), 3'-phosphate binding (PB) motif, and C-terminal conserved sequence (C-terminal/3'-PSB motif) are boxed. The C-terminal/3'-PSB motif is missing in the mL-STL2 isoform.

A**B**

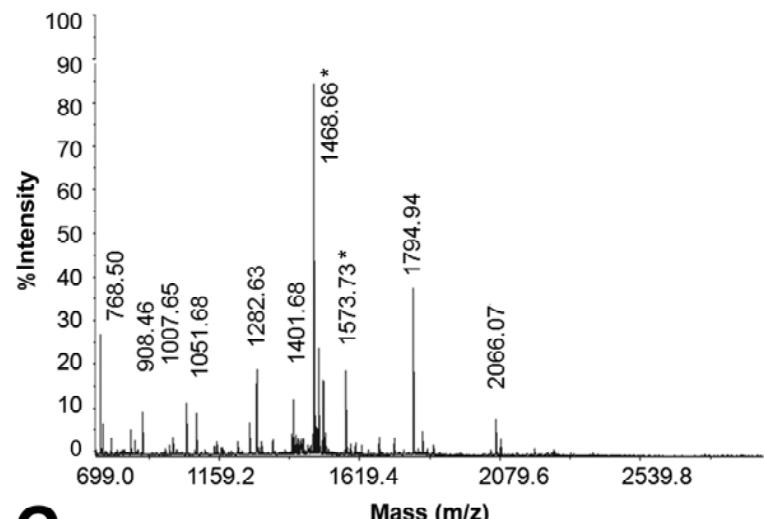
Protein name	GenBank accession no.	Position in amino acid sequences	Peptide sequences	Observed mass (m/z)	Experimental mass (m/z)	Calculated mass (m/z)	Delta mass (Da)	Mascot score	Matched peptides	Sequence coverage (%)
Sulfo-transferase-like protein 1	NP_780459	30 - 33	FVVR	520.32	519.31	519.32	-0.01			
		116 - 121	<u>VIYLIR</u>	776.49	775.48	775.50	-0.01			
		142 - 147	EQVPWK	786.42	785.41	785.41	0.01			
		239 - 244	<u>HEFIMR</u>	832.41	831.40	831.41	-0.01			
		148 - 157	IYFENFLQGK	1258.63	1257.63	1257.64	-0.01	593	10/23	34
		177 - 188	ENILVLSYEQLK	1448.80	1447.79	1447.79	-0.01			
		75 - 85	<u>TPWIEFENNYR</u>	1468.69	1467.68	1467.68	0.00			
		253 - 265	NHFTVAQAEAFDK	1477.69	1476.68	1476.70	-0.02			
		61 - 74	GDPTWVQSTIANER	1573.76	1572.76	1572.75	0.00			
		125 - 138	DVLVSGYHYFNALK	1625.82	1624.82	1624.82	-0.01			

C**D****E**

Supplementary Figure S6. Mass spectrum of trypsin-digested recombinant mL-STL protein. The protein band corresponding to the expected recombinant mL-STL protein was excised from the 14% SDS-PAGE, digested with trypsin, and analyzed by MALDI-TOF peptide sequencing. The tryptic peptides of the excised protein bands were submitted for searching in NCBI database (http://www.matrixscience.com/search_form_.select.html). The peptides that matched with the theoretical predicted peptides of the mouse liver sulfotransferase-like protein 1 (NP_780459) are shown in the MS spectrum in (A) and summarized in a table in (B). The peptides that are indicated with an asterisk (*) in (A) or underlined in (B) were analyzed by MS/MS sequencing. The MS/MS spectra of doubly charged peptides at mass-to-charge (m/z) ratio of 776.49 (C), 834.21 (D), and 1468.69 (E) were sequenced as VLYLLR, HEFIMR, and TPGEIEFENNYR, respectively.

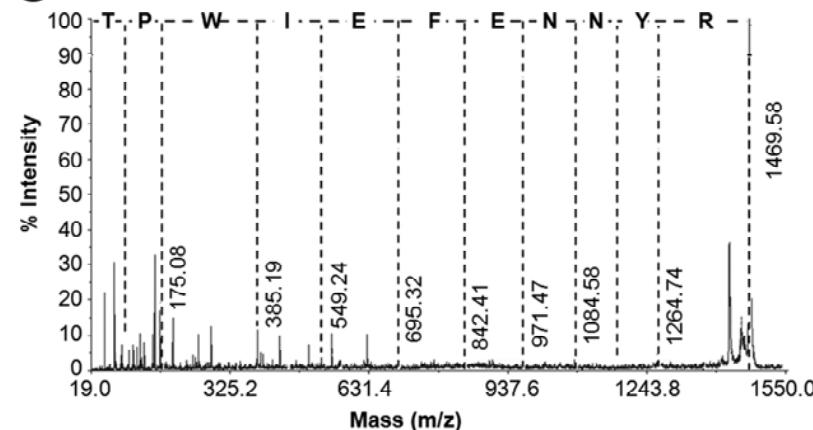
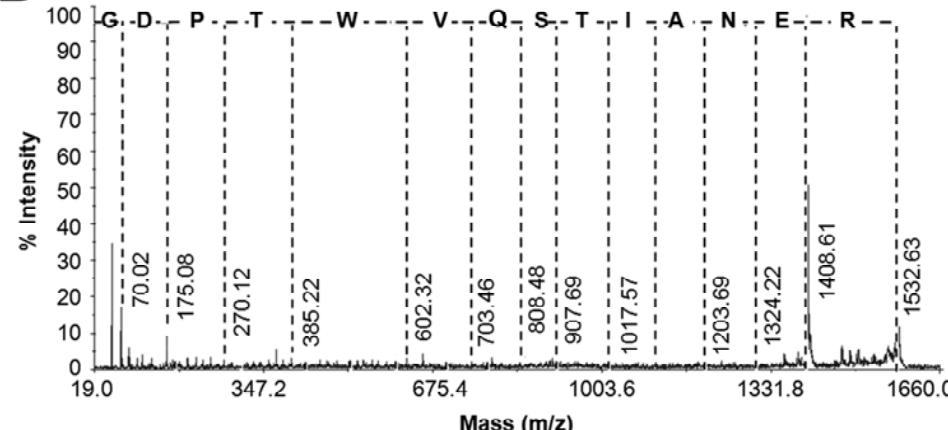
A**B**

Supplementary Figure S7. Purification of recombinant His-mL-STL protein. A two-step column chromatographic method using a Bio-Rad NGC™ Discover 10 chromatography system was used. (A) The supernatant collected from crude cell lysate was subject to two 5-ml nickel affinity HiTrap columns coupled in series. The columns were charged with nickel ion and pre-equilibrated with a binding buffer (50 mM Tris-HCl, pH 8, 500 mM NaCl, and 30 mM imidazole). The samples were then loaded onto the column and 170 ml of binding buffer was passed through the column at a flow rate of 4 ml/min to remove unbound proteins. Then the column was washed with 120 ml of 86.4 mM imidazole in a mixed binding and elution buffer (50 mM Tris-HCl, pH 8, 500 mM NaCl, and 500 mM imidazole) to remove non-specific protein binding. The bound His-mL-STL proteins were eluted from the column with 28 ml of 208.6 mM imidazole in elution buffer. (B) The eluents containing the His-mL-STL proteins were pooled, concentrated to 3 ml and then loaded, at a flow rate of 1 ml/min, onto a HiLoad Superdex 200 prep grade size exclusion column pre-equilibrated with a size-exclusion buffer (50 mM Tris-HCl, pH 8, and 150 mM NaCl). The His-mL-STL proteins were eluted with 100 ml of the same buffer and the eluted fractions containing the His-mL-STL proteins were pooled and collected for sulfotransferase activity assays.

A**B**

Protein name	GenBank accession no.	Position in amino acid sequences	Peptide sequences	Observed mass (m/z)	Experimental mass (m/z)	Calculated mass (m/z)	Delta mass (Da)	Mascot score	Matched peptides	Sequence coverage (%)
Sulfo-transferase-like protein 1	NP_780459	30 - 44	FVVRDEDTIVTYPK	1794.94	1793.93	1793.96	-0.03			
		61 - 74	<u>GDPTWVQSTIANER</u>	1573.73	1572.72	1572.75	-0.03			
		75 - 85	TPWIEFENNYR	1468.66	1467.65	1467.68	-0.03			
		95 - 107	LMASLLPIQLFPK	1486.84	1485.83	1485.86	-0.03			
		116 - 121	VIVLIR	776.48	775.47	775.50	-0.03			
		116 - 124	VIVLIRNPR	1143.67	1142.66	1142.69	-0.03			
		125 - 138	DVLVSGYHYFNALK	1625.81	1624.80	1624.82	-0.03			
		148 - 157	IYFENFLQGK	1258.61	1257.61	1257.64	-0.04			
		177 - 188	ENILVLSYEQLK	1448.79	1447.80	1447.79	-0.01	283	17	57
		198 - 216	ICEFLGENLESGEELVLK	2192.09	2191.10	2191.11	-0.02			
		217 - 224	NISFQIMK	996.49	995.48	995.51	-0.03			
		217 - 226	NISFQIMKER	1281.62	1280.63	1280.65	-0.03			
		227 - 238	MISQSCLSNIEK	1409.61	1408.60	1408.67	-0.07			
		227 - 244	MISQSCLSNIEKHEFIMR	2255.06	2254.05	2254.05	-0.00			
		239 - 244	HEFIMR	848.37	847.37	847.40	-0.03			
		253 - 265	NHFTVAQAEAFDK	1477.68	1476.67	1476.70	-0.03			
		253 - 270	NHFTVAQAEAFDKAFQEK	2080.98	2079.98	2080.00	-0.02			

Peptides sequenced by tandem mass spectrometry are underlined.

C**D**

Supplementary Figure S8. Mass spectrum of trypsin-digested recombinant His-mL-STL protein. The protein band corresponding to the expected recombinant His-mL-STL protein was excised from the 14% SDS-PAGE, digested with trypsin, and analyzed by MALDI-TOF peptide sequencing. The tryptic peptides of the excised protein bands were submitted for searching in NCBI database (http://www.matrixscience.com/search_form_select.html). The peptides that matched with the theoretical predicted peptides of the mouse liver sulfotransferase-like protein 1 (NP_780459) are shown in the MS spectrum in (A) and summarized in a table in (B). The peptides that are indicated with an asterisk (*) in (A) or underlined in (B) were analyzed by MS/MS sequencing. The MS/MS spectra of doubly charged peptides at mass-to-charge (m/z) ratio of 1468.66 (C) and 1573.73 (D) were sequenced as TPWIEFENNYR and GDPTWVQSTIANER, respectively.