

Supplementary Figure S1. The full sequence of *B. mori* P150 protein. The sequence is divided into 5 regions: N-terminus (orange), Repeat 1 (green), linker (magenta), Repeat 2 (grey) and C-terminus (blue). Sequences of the N-terminal signal peptide and C-terminal conserved motif (CXCXCX) are underlined. Triangles in the schematic figure below indicate the variants of Repeat 1 (R1) and Repeat 2 (R2), which are designated as R1A, R1B, R2A and R2B.

[N-Terminus] -
15 [R1] -1 [R1B] -4 [R1] -1 [R1B] -8 [R1] -1 [R1B] -5 [R1] -1 [R1B] -8 [R1] -1 [R1A] -
[Linker] -
11 [R2] -1 [R2A] -2 [R2] -1 [R2A] -15 [R2] -1 [R2B] -5 [R2] -1 [R2B] -2 [R2] -1 [R2B] -5 [R2] -1 [R2B] -1 [R2] -1 [R2B] -1 [R2] -1 [R2B] -2 [R2] -1 [R2B] -4 [R2] -1 [R2B] -15 [R2] -
[C-Terminus]



>BmP150

[illegible]

SKQTNTPTTEHTVIVAHNQLHKTQTKTTASAHPILETVSGMPSIHLSKNDPPTTDEPIEIQSTKQFAPTTFFISESLVLKNKQTQFKVIPLEIPTTEATQKPVQSTIEPIHPKTKPEQSTIKTVQSTDQSTQATMGSFPSIVKPI
QSRESNVFNKAISSIEPLKSTTEPVHSITQSTVVSKQSSNEPLQVTIELLQSSIEPTQSTVQTQSSAQPGQFATKPVNRIETTENPESKKSNEHRPFIQVTKSTKSPIGSTLKGYYREINPTQGFDDVGENEDKRKGRYDL
SKFNQAPAMHMAVENEITLKTLLIVEDGSKYGRCYCSCDLNSKPVFISLDGSALKPRR

>BmP150_N_Terminus

MKVLCAIVLYIALMQPALCDPPFAKKSNEHHTLDHLLNKGPEGNRYRAPSSFFENSAINKHFFQNYFVNQQGRVQPVQPTRSARHLNAKPFRAA
ENKHNHLHANARI PVRHPSIQPKDTQSETNHAKNSNLPTRVSLNLDVTTESHINENTVVKQPTEGNGQIDSVTNTIDPIIKKPNEVNGTDKKPE
IQWTTNTTHEQTTSAVGKGETSSNNQLKLNKDTVLGKHYIVRPATTNALDKTQNVPELTTDSAYQIPLLPSTQPADSLLSKESNQEVLEKIIEE
VIKDNKYENSDDINPVKFIYKEELVPNAAIENQENVTTTDLKGFLKNIKEIKHSEEKIENFKENKTQEWKSKKNMTIKKAESITQHDVSKKTN
EYTIIEKRNETREESDFESNKQLSKYTEESQFELNTRISLTENSDEFAALNEYLEEUVKKIENQNKAYEKVEQKSLQAEKFDDLEFWRQEAKEEEK
KREKIREMQSKLINGGRHTEMTEKELVSKAEELVENDEDEFWNQEAITYLDNNYENSKTLNKTNSSISSTTNPKQPIITGANIAKTTATEQSTTEA
KVLSTTELKQSTTATSVPSTVASEQSTSEISVPSTTGTEQPTTETSVPSTTATEQSRDIDKLPSTTSTVQS

>BmP150_Repeat1

TTETNVPSTSVTEKSTTETSVSSTTKLKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVPSTTATEQSTTETSVPSTIELKQS
TTETNVPSTTATEQSTTETSVPSTIELKQS
TTETNVPSTSATEKSTTETSVPSTIELKQS
TTETNVPSTSAIEKSTTDTSPSTTELKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVPSTTATEQSTTETSVPSTIELKQS
TTETNVPSTSVTEKSTTETSVSSTTKLKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVPSTTATEQSTTETSVPSTIELKQS
TTETNVPSTSAIEKSTTETSVPSTTELKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVPSTSVTEKSTTETSVSSTTELKQS
TTETNVPSTTATEQSTTETSVPSTIELKQS
TTETSVPSTIELKQS
TTETNVPSTSATEQSTTETSVPSTIELKQS
TTETNVPSTSVTEKSTTETSVSSTTELKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVPSTTATEQSTTETSVPSTIELKQS
TTETSVPSTIELKQS
TTETNVPSTSATEQSTTETSVPSTIELKQS
TTETNVPSTSVTEKSTTETSVSSTTELKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVASTTATEQSTTETSVPSTIELKQS
TTETNVPSTSAIEKSTTDTSPSTTELKQS
TTETNASTPTATEQSTTETSVSSTTELKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETSVPSTIELKQS
TTETNVPSTSATEQSTTETSVSSTTKLKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVASTTATEQSTTETSVPSTIELKQS
TTETNVPSTTATEQSTTETSVPSTTELKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETSVPSTIELKQS
TTETNVPSTSATEQSTTETSVPSTIELKQS
TTETNVPSTSAIEKSTTETSVPSTTELKQS
TTETNASTPTATEQSTTETSVPSTTKLKQS
TTETNASTPTATEQSTTETSVSSTTKLKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVPSTTATEQSTTETSVPSTIELKQS
TTETNVPSTSAIEKSTTETSVPSTTELKQS
TTETNASTPTATEQSTTETSVPSTIELKQS
TTETNVPSTTAEQS

>BmP150_Link

TNETSVPSTDDNVQPVTKEDVTEPTAAYIKVQS

>BmP150_Repeat2

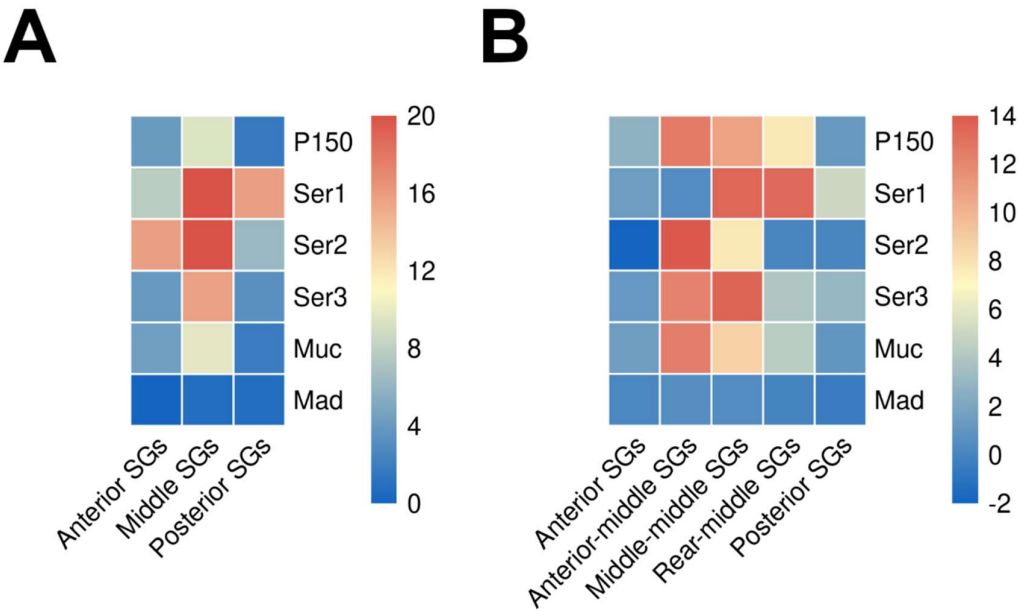
TTVTESNTTGAAVQSTTATENTTTDAEAQS
TTVTESNTTGAAVQSTTATESGTTDAEVRS
TTVTESNTTGAAVQSTTATESATTTDAEAQS
TTVTESNTTGTAVRSTTATESAITDAEVQS
TTVTESNTTGAAVQSSATESATTTDAEAQS
TTVTESNTTGTAVQSTTATESAITDAEVQS
TTVTESNTTGAAVQSTTATESATTTDAEVQS
TTVTESNTTGAAVQSTTATESATTTDAEAQS
TTVTESNTTDAAVQSTTATESASTDAEVQS
TTVTESNTTGTAVQSTTATESAITDAEVQS
TTVTESNTTGAAVQSTTATESATTTDAESQS
TTVTESSTAGAAVQSTTATDAEDQS
TTITESNTTGAAVQSTTATESATTTDAEAQS
TTVTESNTTGAAVQSTTATESATTTDAESQS
TTVTESSTAGAAVQSTTATDAEDQS
TTVTESNTTGAAVQSTTATESATTTDAEAQS
TTVTESNTTGAAVQSTTATESAITDAEVQS

TTVKESNSAGAAVQSSTATESATTDAEQs
TTVTESNTTGTAVQSTTATESAITDAEVQs
TTVTESNTTGAAVQSTTATESATTDAEVQs
TTVTESNTTGAAVQFTTATESATTDAEQs
TTVTESNTTDAAVQSTTATESASTDAEVQs
TTVTESNTTGAAVQSTTATESATTDAEQs
TTVTESNTTDAAVQSTTATESASTDAEVQs
TTVTESNTTGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESAITDAEVQs
TTVKESNSAGAAVQSSTATESATTDAEQs
TTVTESNTTGTAVQSTTATESAITDAEVQs
TTVTESNTTGAAVQSTTTTESATTDAEQs
TTVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESATTDAEQs
TTVTESNTTGTAVQSTTATESAITDAEVQs
TTVTESNTTGAAVQSTTATESATTDAEVQs
TTVTESNTTGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESAITDAEQs
TTVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDSEAQs
TTVTESNTTGAAVQSTTATESAITDAEQs
TTVTTELSTAGAAVQSTTATENATTDAEQs
TTVTESNTTGAAVQSTTTTESATTDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TNVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESAITDAEQs
TTVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATETAITDAEQs
TTVTESNTTGAAVQSTTAKKSTTADAEQF
TTVSESNTGRAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTATESATTDAAATDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESNTTGAAVQSTTAKKSTTADAEQF
TTVSESNTGRAAVQFTTATESATTDAEQs
TTVTESYTTGAAVQSTTATESATTDAEQs
TTVTESNTAGAAVQSTTTTESATTDAEQs
TTVTESSTAGAAAQSTTATDSATTDAEQs
TTVTESSSAGAAVQSTTATESATTNAEQs
TTVTESSSAGAAVQSTTTTESATTDAEQs
TTVTESSTAGAAAQSTTATDSATADAEVQs
TTVTTELSTTGAAVPSTTATESSTTDAEQs
TTVTESSTAGAAVQSTTATESATTDAEQs
TTVTESSSAGAAVQSTTTTESATTDAEQs
TTVTESSTAGAAAQSTTATDSATADAEVQs
TTVTTELSTTGAAVPSTTATKSTTDAEQs
TTVTESSTAGVAVQSTSATERAITDTEVQs

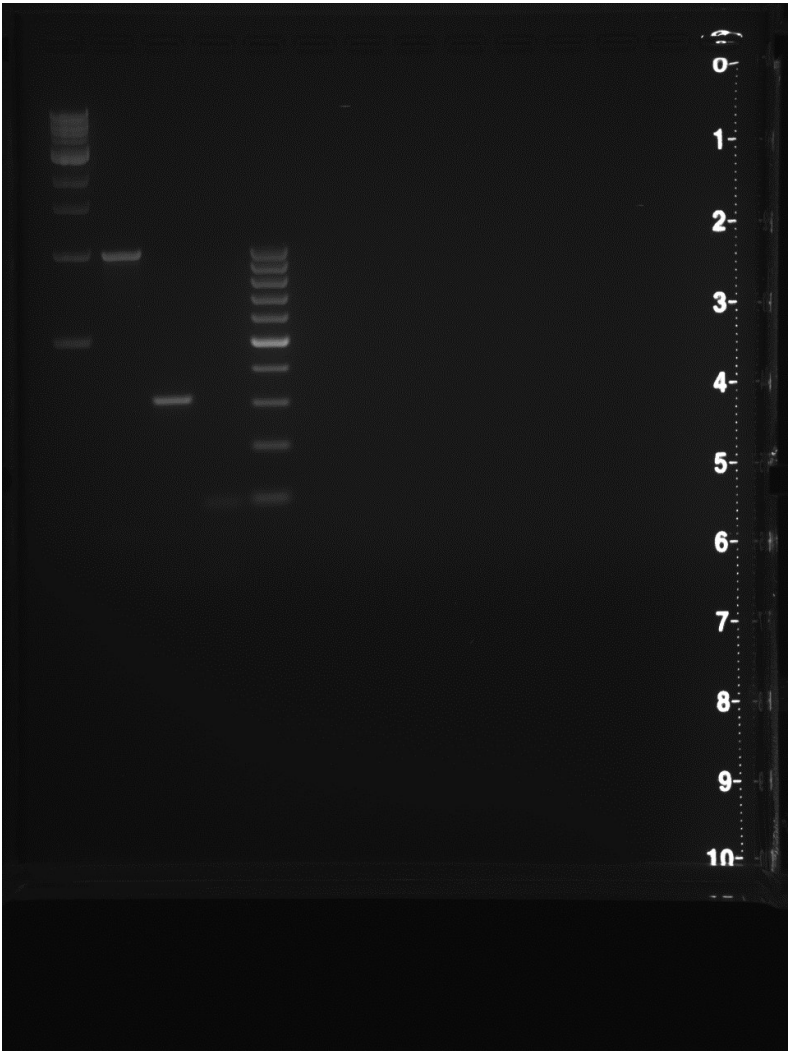
>BmP150_C_Terminus

ATVTESNTAAKVHDQPTRTVFTSPCGNHPTSVATDNTSEKESKQTNTPTTEHTVIVAHNQLHKTQTKTTASAHPILETVSGMPSIHLTSKNDPPT
TDEPIEIQSTKQFAPTTFISESLVLKNKQTQFKVIPLEIPTeATQKPVQSTIEPIHPKTKPEQSTIKTVQSTDQSTQATMGsFPSIVKPIQs
RESNVFNvKAISSIEPLKSTTEPVHSITQSTVVSKQSSNEPLQVTIELLQSSIEPTQSTVQTQQSSAQPGQFATKPVNRIETTTENPESKKSNE
HRPFIQVTKSTKSPIGSTLKGYYREINPTQGGFDDVGENEDKRKGRYDLskFNQAPAMHMAVENEITLKTlIVEDGSKYGRcyCSCDLNSKPVF
ISLDGSALKPRR

Supplementary Figure S2. Expression levels of *B. mori* P150 gene along with major silk genes (Sericin 1-3 and Mucin-12) and the control gene (mothers against dpp; Mad) in different parts of silk glands. Heat maps represent the log2 fold-change values. As shown in both figures, P150 is highly expressed in middle silk glands. Sources of the raw RNA-seq data were listed in **Supplementary Table S1**.



Supplementary Figure S3. The original image of the electrophoresis gel (**Fig. 1C**).



Supplementary Table S1. Sources of the RNA-seq datasets utilized to verify the expression of *B. mori* P150 and other genes.

(A) datasets used in **Supplementary Figure S1A**

Run	BioProject	BioSample	Age	Instrument	LibraryLayout	LibrarySource	Organism	Tissue
SRR10035619	PRJNA559726	SAMN12608940	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior-Silk-Gland
SRR10035620	PRJNA559726	SAMN12608939	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior-Silk-Gland
SRR10035621	PRJNA559726	SAMN12608938	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior-Silk-Gland
SRR10035669	PRJNA559726	SAMN12609122	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Middle-Silk-Gland
SRR10035670	PRJNA559726	SAMN12609121	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Middle-Silk-Gland
SRR10035671	PRJNA559726	SAMN12609120	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Middle-Silk-Gland
SRR10035726	PRJNA559726	SAMN12609071	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Midgut
SRR10035727	PRJNA559726	SAMN12609070	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Midgut
SRR10035728	PRJNA559726	SAMN12609069	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Midgut
SRR10035756	PRJNA559726	SAMN12609044	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior-Silk-Gland
SRR10035757	PRJNA559726	SAMN12609043	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior-Silk-Gland
SRR10035758	PRJNA559726	SAMN12609042	5th instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior-Silk-Gland

(B) datasets used in **Supplementary Figure S1B**

Run	BioProject	BioSample	Age	Instrument	LibraryLayout	LibrarySource	Organism	Tissue
DRR186474	PRJDB8614	SAMD00180408	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior silk gland
DRR186475	PRJDB8614	SAMD00180409	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior silk gland
DRR186476	PRJDB8614	SAMD00180410	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior silk gland
DRR186477	PRJDB8614	SAMD00180411	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior part of the middle silk gland
DRR186478	PRJDB8614	SAMD00180412	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior part of the middle silk gland
DRR186479	PRJDB8614	SAMD00180413	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Anterior part of the middle silk gland
DRR186480	PRJDB8614	SAMD00180414	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Middle part of the middle silk gland
DRR186481	PRJDB8614	SAMD00180415	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Middle part of the middle silk gland
DRR186482	PRJDB8614	SAMD00180416	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Middle part of the middle silk gland
DRR186483	PRJDB8614	SAMD00180417	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior part of the middle silk gland
DRR186484	PRJDB8614	SAMD00180418	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior part of the middle silk gland
DRR186485	PRJDB8614	SAMD00180419	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior part of the middle silk gland
DRR186486	PRJDB8614	SAMD00180420	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior silk gland
DRR186487	PRJDB8614	SAMD00180421	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior silk gland
DRR186488	PRJDB8614	SAMD00180422	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Posterior silk gland
DRR186492	PRJDB8614	SAMD00180426	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Midgut
DRR186493	PRJDB8614	SAMD00180427	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Midgut
DRR186494	PRJDB8614	SAMD00180428	fifth instar day 3	Illumina NovaSeq 6000	PAIRED	TRANSCRIPTOMIC	Bombyx mori	Midgut

Supplementary Table S2. Sequences of primers used in (A) exon junction sequencing and (B) qPCR analysis.

(A) Sequencing primers	
P150 #1F	GAATCGAGCACTGCAGGAGT
P150 #1R	AGTAGGCTGATCGTGGACCT
P150 #2F	CCACGAAACAGTTTGCTCCA
P150 #2R	TAGGCTCGATTGTGGACTGC
P150 #3F	CCCAAGGCTTCGGTGATGAT
P150 #3R	TATGCATTGCCGGAGCTTGA

(B) qPCR primers	
P150 F	CCCAAGGCTTCGGTGATGAT
P150 R	TATGCATTGCCGGAGCTTGA
Ser1 F	CACAACCGATAAGACGAG
Ser1 R	GACGAAGTGGAGGAAGC
Ser2 F	CATCGGCTGACTACCA
Ser2 R	AGAGTTGCTGCCCTTAC
Ser3 F	TGTCTCGTCGGTGGA
Ser3 R	TTGTTGTATGACTGGCTCT
Mad F	ACACAAGGCGTCACATAGGG
Mad R	TGGTGATTGCAGTTACGGCT
EF1a F	CAAGTCTGGAGATGCAGCCA
EF1a R	GGGGTGGGAATTCCTGGAAG

Supplementary Table S3. Statistical analysis of quantitative PCR analysis of gene expression of genes in different tissues. Statistical differences were evaluated by Student’s t-test. ASG: anterior silk glands; MSG: middle silk glands; PSG: posterior silk glands; OVA: ovaries. Background colors indicate statistical significance: orange ($p < 0.001$); yellow ($p < 0.01$); blue ($p < 0.05$); blank ($p \geq 0.05$).

	ASG					MSG					PSG					OVA					GUT				
ASG																									
MSG	6.89E-03	9.38E-05	5.64E-02	3.66E-01	8.29E-01																				
PSG	1.16E-01	3.66E-02	8.92E-01	8.28E-02	8.81E-01	1.14E-04	1.98E-02	1.07E-03	9.82E-02	8.89E-01															
OVA	3.89E-01	4.14E-02	1.84E-01	4.36E-01	1.22E-01	2.57E-06	3.05E-04	1.37E-07	2.23E-01	1.99E-03	2.64E-02	2.54E-02	7.48E-03	6.18E-03	9.12E-03										
GUT	2.28E-01	1.06E-02	1.50E-01	1.59E-02	4.47E-01	4.10E-05	1.20E-04	3.08E-08	2.81E-02	1.96E-01	1.13E-01	1.80E-02	4.87E-03	1.10E-01	2.42E-01	1.40E-01	7.94E-02	4.12E-03	4.05E-02	2.71E-03					
	P150	Ser1	Ser2	Ser3	Muc	P150	Ser1	Ser2	Ser3	Muc	P150	Ser1	Ser2	Ser3	Muc	P150	Ser1	Ser2	Ser3	Muc	P150	Ser1	Ser2	Ser3	Muc

Supplementary Table S4. Properties of 2 types of repeat sequences in *B. mori* P150 protein. GRAVY: grand average of hydropathy; M. w.: molecular weight; pI: isoelectric point.

	GRAVY	Residue	M. w. (Da)	pI	Top1 AA	Top2 AA	Top3 AA	Top4 AA	Top5 AA
Repeat 1	-0.8512	1275	133254.12	3.8738	Thr: 417 (32.706%)	Ser: 237 (18.588%)	Glu: 162 (12.706%)	Pro: 75 (5.882%)	Val: 75 (5.882%)
Repeat 2	-0.4541	2220	215927.31	2.6857	Thr: 696 (31.351%)	Ala: 440 (19.820%)	Ser: 320 (14.414%)	Glu: 211 (9.505%)	Val: 161 (7.252%)

Supplementary Table S5. Properties of *B. mori* sericins and mucin-12 proteins. GRAVY: grand average of hydropathy; M. w.: molecular weight; pI: isoelectric point. The conserved cystein residues at the C-terminal region were labeled in red.

Gene	Accession	GRAVY	Residue	M. w. (Da)	pI	Top1 AA	Top2 AA	Top3 AA	Top4 AA	Top5 AA	C-terminus
Sericin 1	XP_037869538.1	-1.1181	3385	331903.34	4.7726	Ser: 1300 (38.405%)	Gly: 428 (12.644%)	Thr: 317 (9.365%)	Asn: 270 (7.976%)	Asp: 221 (6.529%)	LLHKPGQGKICLCFENIFDIPYHLRKNIGV
Sericin 2	NP_001166287.1	-2.1516	1758	198657.39	9.1575	Lys: 302 (17.179%)	Ser: 266 (15.131%)	Asp: 208 (11.832%)	Glu: 195 (11.092%)	Asn: 123 (6.997%)	SSSSSSSSSSSRSSSSSTYTGS HDDSSEE
Sericin 3	NP_001108116.1	-1.4651	1271	123298.08	5.7076	Ser: 554 (43.588%)	Gly: 152 (11.959%)	Gln: 94 (7.396%)	Asn: 85 (6.688%)	Lys: 79 (6.216%)	QSKASSFSASSASESSSLSSDVNFEEKTD
Sericin 4	BGIBMGA011896	-0.9876	2270	251440.57	6.7156	Lys: 223 (9.824%)	Thr: 218 (9.604%)	Ser: 215 (9.471%)	Glu: 205 (9.031%)	Gly: 169 (7.445%)	IHGGENTEAGGHSPGLVGGLFKYLLGGKTQ
Sericin 5	KWMTBOMO06271	-2.0155	1976	218868.3	9.6116	Lys: 399 (20.192%)	Asp: 254 (12.854%)	Glu: 211 (10.678%)	Ser: 203 (10.273%)	Gly: 202 (10.223%)	SESASSSRSSSYHSSSMQRNQKTSFDDQNE
Sericin 6 / P150	Sericin 6 / P150	-0.6718	4552	467241.21	3.8151	Thr: 1243 (27.307%)	Ser: 658 (14.455%)	Ala: 547 (12.017%)	Glu: 479 (10.523%)	Val: 297 (6.525%)	SKYGRCYCSDLNSKPVFISLDGSALKPRR
Mucin-12	XP_021205679.2	-1.4977	3173	368381.36	4.137	Thr: 583 (18.374%)	Glu: 522 (16.451%)	Gln: 478 (15.065%)	Pro: 218 (6.870%)	Leu: 202 (6.366%)	QGQNCFCDCDNTSAPKLIPLKDLPA GITGS

Supplementary Table S6. Properties of P150 proteins in *B. mori*, *E. kuehniella* and *G. mellonella*. GRAVY: grand average of hydropathy; M. w.: molecular weight; pI: isoelectric point.

Accession	Species	Family	Superfamily	GRAVY	Residue	M. w. (Da)	pI	Top1 AA	Top2 AA	Top3 AA	Top4 AA	Top5 AA
P150	<i>Bombyx mori</i>	Bombycidae	Bombycoidea	-0.6718	4552	467241.21	3.8151	Thr: 1243 (27.307%)	Ser: 658 (14.455%)	Ala: 547 (12.017%)	Glu: 479 (10.523%)	Val: 297 (6.525%)
WDD44665.1	<i>Ephestia kuehniella</i>	Pyralidae	Pyraloidea	-0.8492	1741	173261.02	4.212	Ser: 603 (34.635%)	Gln: 193 (11.086%)	Gly: 160 (9.190%)	Pro: 154 (8.845%)	Ile: 94 (5.399%)
XP_026763958.2	<i>Galleria mellonella</i>	Pyralidae	Pyraloidea	-1.2812	1465	160622.85	4.7732	Ser: 334 (22.799%)	Gln: 251 (17.133%)	Asn: 116 (7.918%)	Glu: 101 (6.894%)	Thr: 101 (6.894%)

Supplementary Table S7. List of proteins detected in *B. mori* cocoon silk by MaxQuant. Intensity is the sum of intensity values of all replicates.

Intensity	Accession	Gene	Residue	M. w. (kDa)	GRAVY	pI	Top3 AA	Signal peptide
190419400000	XP_037895938.1	sericin 1 isoform X1	3385	331.903	-1.118	4.773	Sr (38.405%); Gly (12.644%); Thr (9.365%)	MFVPLCCTGLAALAAVLVSKAFG
41472000000	XP_004926112.1	uncharacterized protein LOC101740726	239	25.909	-0.016	8.824	Leu (15.900%); Ala (10.460%); Gly (8.368%)	MKLLVLAVCAVFAA
27399600000	XP_004922636.1	fibrin(ogen)-like	241	27.201	-0.058	4.555	Leu (9.950%); Asn (7.884%); Ser (7.884%)	MKMLLVLAVCSGLLTLTA
22181230000	XP_037874919.1	LOW QUALITY PROTEIN: zonadhesin	3973	432.78	-0.563	6.444	Cys (14.372%); Gly (7.759%); Asn (3.568%)	MKFLTFVAGIANVLLVLDSTEA
17189600000	NP_001108116.1	sericin 3 precursor	1271	123.298	-1.465	5.708	Sr (45.588%); Gly (11.895%); Gln (7.589%)	MKNCFLVLAIVAAVGA
11367900000	NP_001106733.1	fibrin heavy chain precursor	5915	591.593	0.216	4.128	Gly (30.268%); Asn (12.000%); Thr (12.000%)	MKRFKTVLCGLVAYVNTA
7049000000	NP_001037488.1	fibrin light chain precursor	262	27.754	0.024	5.188	Ala (13.740%); Ser (9.542%); Gly (8.397%)	MKPIFLVLLVATVTA
6734500000	XP_004921814.1	uncharacterized protein LOC101741789	474	50.925	-0.448	5.343	Lys (9.705%); Ser (9.494%); Val (8.228%)	MKVAALFTVLQVNA
4989900000	NP_001037045.1	sericin 1 precursor	108	11.882	-0.388	4.059	Sr (12.037%); Asp (10.185%); Gly (9.259%)	MATFKFLVLTITLISA
33431100000	NP_001078333.1	beta-N-acetylglucosaminidase 1 precursor	111	69.655	-0.287	6.124	Leu (9.493%); Val (7.529%); Arg (6.710%)	MMLHRLMLMTLMLHGLVLC
2228430000	XP_004921708.1	autophagoraptor adhesion BpAc-1	261	24.119	-0.133	3.879	Ala (29.502%); Ser (19.923%); Thr (17.241%)	MSLSRVAFVASILVYVSVISA
17895600000	XP_004926242.1	zapoliprotein isoform X1	3317	369.106	-0.315	7.973	Gly (9.828%); Lys (9.014%); Ser (7.266%)	MTGTSFLVLLVLLVLPVLA
17895600000	NP_001106745.1	carboxypeptidase inhibitor precursor	113	12.561	-0.296	4.244	Gly (13.212%); Gly (10.619%); Gly (8.888%)	MLFLFCAMCLVLA
1700100000	XP_004925684.1	hN5-interacting protein isoform X1	394	43.157	-0.34	4.366	Cys (13.959%); Pro (10.650%); Asn (8.376%)	MLFLFCSTVTA
1198871000	XP_012546338.1	glycine-rich cell wall structural protein 1.0-like	208	18.431	-0.801	3.207	Gly (36.538%); Asp (18.750%); Ala (12.500%)	MKIKLWLLSHVVAVTVVEA
1102090000	XP_037874647.1	zonadhesin isoform X1	683	75.835	-0.606	4.663	Cys (14.614%); Gly (9.956%); Pro (7.174%)	MGRGCFVLTAAWLVAVKADA
1100470000	XP_037867528.1	zonadhesin isoform X1	2721	296.36	-0.549	5.952	Cys (14.672%); Pro (13.120%); Asn (8.012%)	MLVFLVGLMLLSSTTA
1081719000	XP_004928913.2	vanin-like protein 2	515	57.687	-0.124	5.014	Val (8.932%); Leu (7.573%); Ser (7.379%)	MYKFLVGLFLSTKATKS
1060700000	XP_004923812.1	alcohol dehydrogenase 2	139	35.343	-0.202	6.788	Val (10.460%); Ile (8.458%); Asn (6.448%)	MKFLVFLVFLVAVFTST
1034700000	XP_021294310.1	uncharacterized protein LOC101744556 isoform X1	1979	222.181	-0.727	4.753	Asn (9.807%); Asp (8.599%); Gly (8.599%)	MKLYLVFLVFLVSTVSPIS
9618000000	XP_037871220.1	fibrin(ogen)-like	240	27.327	-0.179	4.591	Leu (9.583%); Asn (8.333%); Gly (7.500%)	MLRVVYSLVLAFLGUTA
8809600000	XP_037873722.1	glucose xeroderma [FAD, quinone]	577	64.896	-0.155	7.01	Leu (10.052%); Val (8.146%); Lys (7.452%)	-
7499530000	XP_037869877.1	uncharacterized protein LOC101740721 isoform X1	2270	251.441	-0.988	6.716	Lys (9.824%); Thr (9.604%); Ser (9.471%)	-
7171880000	XP_037874060.1	myosinase 1	567	65.135	-0.044	6.173	Leu (8.466%); Ser (8.113%); Ala (7.760%)	MDWTYKLVFLVLMVVVVS
7007560000	NP_001109433.1	fibrin(ogen) precursor	220	25.167	-0.102	7.162	Leu (10.000%); Ala (7.273%); Phe (6.818%)	MGFLLSMMLVFLDFPFRASP
6779150000	NP_001138339.1	integrin epsilon subunit 1 precursor	561	63.237	-0.425	6.835	Lys (8.021%); Gly (7.843%); Pro (7.843%)	MLRCLAFCAVAVVADA
6779150000	NP_012551240.1	uncharacterized protein LOC105842476	209	26.439	-0.007	4.163	Asn (10.402%); Asp (8.478%); Thr (7.892%)	MKFLICFACVYLIA
4741860000	XP_004924748.1	proton-coupled folate transporter	508	57.334	-0.229	5.884	Leu (10.872%); Ile (9.646%); Ser (8.661%)	-
456522000	XP_012546430.1	lysosomal acid glucosylceramidase isoform X1	536	61.151	-0.243	7.051	Leu (8.676%); Lys (7.836%); Ile (7.090%)	-
4032735000	XP_004928916.1	glucose dehydrogenase [FAD, quinone]	657	73.389	-0.294	7.026	Leu (8.767%); Val (7.610%); Gly (7.154%)	MHTRKRIKFMFLFVNANCCDA
352649000	XP_037868215.1	beta-glucuronidase isoform X1	688	79.06	-0.304	8.715	Val (7.849%); Leu (9.777%); Thr (6.831%)	-
342386000	XP_004927001.2	ispase 1	399	46.491	-0.293	5.885	Lys (9.023%); Glu (8.020%); Ile (7.519%)	MKVGIRFTFMFLSIGHLASA
342524000	XP_004927001.2	ispase D2	209	23.055	-0.266	6.052	Val (11.005%); Leu (10.948%); Ala (6.162%)	MGALIRACILTVATVDFVRLTA
319436000	XP_037872749.1	uncharacterized protein LOC105842476	585	65.476	-0.206	6.052	Lys (9.573%); Leu (9.090%); Thr (8.547%)	-
314065000	NP_001037046.2	sericin 2 precursor	112	12.273	-0.287	9.883	Sr (12.500%); Lys (10.714%); Thr (10.714%)	MATFKFLMLSLITISA
298483000	XP_037869036.1	beta-galactosidase isoform X1	2942	328.542	-0.211	6.857	Leu (9.007%); Val (7.988%); Gly (7.104%)	MTLTIATFALLCGLQGTYS
283373000	XP_012545566.1	alpha-1 isoform X1	475	53.869	-0.208	6.511	Val (8.211%); Leu (7.158%); Ser (7.158%)	-
252201950	NP_001038650.1	cathespin B precursor	337	37.556	-0.441	6.382	Gly (9.792%); Ser (7.122%); Asp (6.825%)	MFISRAAYVTLVCVLA
245267000	XP_004924612.2	esterase FE4	550	61.919	-0.231	5.06	Leu (8.000%); Gly (7.636%); Ile (7.273%)	MKLTFTFVIAVSVYVLG
152476000	XP_037876444.1	lateral-transferrin binding factor beta-binding protein 4	1359	150.321	-0.342	5.769	Cys (10.500%); Leu (9.500%); Thr (7.000%)	MKIVFTVFLFTSTVNA
165783000	XP_004929234.1	enzyme dipeptidyl peptidase 4	740	83.673	-0.199	5.527	Leu (8.243%); Val (7.973%); Ile (7.626%)	MAMEQAVLLMLGLUICHTQSSA
149592000	NP_001093080.1	erythrocyte-regulated 16 kDa protein isoform X1	145	15.824	-0.179	6.336	Leu (11.724%); Val (8.276%); Ala (7.588%)	MLFFITAAVLASAEA
148759000	XP_037874096.1	myosinase 1 isoform X1	495	57.471	-0.473	5.019	Leu (8.081%); Gly (7.879%); Ser (7.071%)	MNL.SQWFAVSLAMACAWG
145575000	XP_037868236.1	juvenile hormone esterase-like	562	63.516	-0.189	9.306	Leu (9.253%); Asn (8.541%); Ile (7.295%)	MSNTFYTLVTLVLLNTINA
143543000	NP_001288822.1	neutral alpha-glucosidase AB-like precursor	925	104.635	-0.371	5.706	Leu (8.342%); Val (8.324%); Ala (8.000%)	MKVVALLVVAIVGISA
137424000	XP_004926880.2	lysosomal acid glucosylceramidase isoform X1	530	59.677	-0.19	4.426	Leu (8.115%); Asp (7.358%); Ile (7.170%)	MDNKICISASLVAAVVLFGSADA
132450000	XP_012547740.1	heat shock 70 kD protein cognate isoform X1	960	106.18	-0.4	5.488	Lys (10.542%); Thr (8.125%); Thr (7.928%)	-
129430000	XP_037868883.1	arylsulcatase	537	60.465	-0.397	7.796	Leu (8.755%); Gly (8.194%); Lys (7.635%)	MFVFRVFLVLLMLVAES
116001000	XP_012544831.3	uncharacterized protein LOC101741510	1999	228.101	-0.553	5.916	Val (8.054%); Leu (8.004%); Lys (7.404%)	MSKISAYLLFOIVIVNCCS
114229100	NP_001037047.1	silk proteinase inhibitor precursor	65	6.962	-0.584	8.134	Cys (13.846%); Gly (10.769%); Thr (10.769%)	MTSLVFLVFLCCTALGAEIS
113206000	XP_021066228.1	poly(U)-specific endoribonuclease-D isoform X1	536	60.005	-0.79	9.886	Sr (10.634%); Thr (9.888%); Asn (8.769%)	-
102791000	NP_001106274.1	antennal binding protein precursor	140	15.49	-0.11	7.206	Lys (11.429%); Ala (10.000%); Leu (7.857%)	MMYLSFVLVLLCAFAVFNCGA
97291000	XP_004927018.1	C3 and P2P-like alpha-2-macroglobulin domain-containing protein 8	1448	155.934	-0.031	5.938	Ala (11.119%); Leu (10.152%); Ser (8.840%)	MEKIKTLFLFILSVPAVTCG
936170000	XP_004927018.1	lysosomal metalloproteinase inhibitor protein-like isoform X1	209	23.055	-0.274	7.206	Cys (10.500%); Leu (9.500%); Thr (7.000%)	-
93201000	XP_02130913.1	uncharacterized protein LOC101747119 isoform X1	722	80.01	-0.235	6.747	Val (9.695%); Leu (7.895%); Ser (7.895%)	MSNSEIIVFLVTLCCSSVAG
77323000	XP_012545193.1	uncharacterized protein LOC101735738 isoform X1	671	75.143	-0.309	6.349	Leu (10.581%); Gly (8.197%); Val (7.899%)	MSVRVFVTLVLAIVHNSRA
75014000	NP_001040174.1	alpha-esterase 13 precursor	540	61.416	-0.241	5.012	Leu (9.615%); Phe (7.593%); Lys (7.593%)	MLFAICVQLVLSVFG
74413000	XP_037876352.1	4-hydroxy-tetrahydrodipicolinate synthase-like	328	35.694	-0.094	4.9	Leu (11.585%); Ala (8.537%); Ile (7.927%)	MYSIGSLTFLCVLINVDFTKC
73489010	XP_037867171.1	15-hydroxyprostaglandin dehydrogenase [NAD(+)]=like	289	31.605	-0.025	9.45	Lys (10.727%); Ile (10.035%); Thr (9.957%)	MTLSLKLTFVFSVFINVDSIA
71931700	NP_001037075.1	cathepsin L precursor	398	45.802	-0.292	4.245	Asp (13.819%); Lys (12.060%); Glu (11.307%)	MYKLVVYVLLSLASLNS
6869500	XP_02106073.1	uncharacterized protein LOC105842440	118	127.974	-0.974	5.852	Thr (10.698%); Thr (9.891%); Thr (9.101%)	-GARLLAAVLLAAILAVSC
6716000	XP_012547984.1	uncharacterized protein LOC101744060	934	101.918	-0.275	4.433	Leu (8.950%); Ser (8.655%); Glu (8.458%)	MNANKKAVLLTLLNNIS
66627000	XP_037874369.1	fibrillin-1 isoform X1	3047	325.331	-0.468	4.496	Gly (12.504%); Cys (12.143%); Asp (7.844%)	MGDGKASVSWYRLVLAALLASVAG
60053000	XP_012550868.1	uncharacterized protein LOC778506 isoform X1	292	30.66	-0.599	4.054	Leu (16.781%); Leu (12.329%); Lys (11.644%)	-
59817000	XP_037868167.1	uncharacterized protein LOC101736658 isoform X1	456	51.569	-0.228	6.318	Leu (9.649%); Ser (8.333%); Val (8.114%)	MGSIETLVLLQLVYVTS
59417900	XP_004929039.1	phospholipase A2	177	20.411	-0.24	4.652	Glu (7.910%); Gly (7.345%); Gly (7.345%)	MFRFLFCVLIVAHNTKCA
58991600	XP_037867041.1	alpha-crystallin B chain	241	26.543	-0.263	4.361	Val (11.618%); Ala (9.129%); Thr (9.129%)	MSRPLRFVAFVAFGATVTA
58991600	XP_004929039.1	uncharacterized protein LOC101743953 isoform X1	472	52.872	-0.063	7.143	Leu (11.444%); Thr (8.263%); Ala (8.051%)	MNVYVFLVLCVAGCGG
55294500	XP_037868227.1	mucin-SAC	1768	198.075	-0.847	6.971	Thr (14.988%); Ser (8.484%); Glu (8.258%)	MGIDFYKFKGFKMESTMTLLLVAVIATPVVPS
50392000	NP_001037090.1	serine protease inhibitor 4 precursor	410	46.3	-0.095	7.322	Leu (9.024%); Ser (7.805%); Ile (7.561%)	MCLKFLVLAIVLPSLA
44188000	NP_001037171.1	protein disulfide isomerase precursor	494	55.589	-0.278	4.325	Glu (11.134%); Ala (8.704%); Lys (8.502%)	MRVLTFIAILLGLFA
41498000	NP_001119727.1	actin, cytoplasmic A4	376	41.822	-0.193	5.156	Ala (7.979%); Glu (7.447%); Gly (7.447%)	-
40515000	XP_037874307.1	basement membrane-specific heparan sulfate proteoglycan core protein isoform X1	4228	463.254	-0.538	4.432	Gly (8.609%); Ser (8.538%); Asp (7.427%)	MARGGLAAALLLSTFTQVLKA
39652000	XP_00266666.2	diserpin	104	45.955	-0.194	5.018	Leu (10.44%); Ile (8.024%); Glu (8.263%)	-
38959300	XP_012551293.2	hemo peptidase 2	1404	149.897	-0.304	6.082	Leu (10.000%); Glu (8.929%); Thr (7.065%)	-
37942900	NP_001066591.1	silk gland derived serine protease 1 precursor	392	42.098	-0.134	7.383	Gly (9.694%); Ser (8.049%); Asp (7.693%)	MGAFKASIRSLKTLFLTAL
37469400	NP_001040479.1	peptidylprolyl isomerase B precursor	205	22.397	-0.267	8.842	Gly (12.195%); Lys (10.244%); Thr (7.870%)	MQLVLLVLLVNLGSLIS
37622200	NP_001037430.1	peptidylprolyl isomerase A precursor	457	50.803	-0.156	5.119	Leu (9.409%); Ala (8.972%); Asn (7.440%)	MGYTLTSLSMIAIVLAAVAAAAA
36225300	NP_001037013.1	glucosidase precursor	491	55.596	-0.373	4.575	Asp (7.943%); Leu (7.943%); Gly (7.536%)	MAWLTLSLIVLAVCHTGLA
33477200	NP_001288827.1	low molecular mass 30 kDa lipoprotein 21G1-like precursor	256	29.858	-0.714	4.942	Asp (10.938%); Lys (8.984%); Ser (8.594%)	MKIKLITLTVIAVTFMMSVEA
32917000	XP_001037111.1	probable pyruvate oxidase 5 isoform X1	485	53.878	-0.216	6.602	Leu (9.978%); Lys (7.835%); Thr (7.735%)	MKLVFLVLCVAGCGG
29882500	NP_001036944.1	juvenile hormone binding protein an-095 precursor	243	29.047	-0.044	5.868	Leu (10.000%); Val (8.053%); Thr (7.892%)	MYTGLFLVGLVYSGS
27055300	XP_02105588.1	probable cyclin-dependent serine/threonine/protein kinase DDB_G0292550 isoform X2	544	59.257	-0.054	6.833	Sr (18.199%); Asn (11.944%); Asp (8.272%)	MNARSIFSVTLVGLVSPSTG
25743100	XP_004926879.1	apolipoprotein D	269	29.851	-0.199	6.508	Val (8.922%); Ser (8.550%); Leu (8.178%)	MWVKTLVGLVFLTVVYVS
25471600	XP_004926990.1	ATP-dependent (S)-NAD(P)-H dehydratase	327	35.913	-0.099	7.438	Sr (10.092%); Ile (9.786%); Gly (8.257%)	MTPSFNILVFLVTLFTQITNG
25363500	XP_037867475.1	calsenitin-1 isoform X1	953	105.096	-0.324	5.94	Ala (9.654%); Val (8.919%); Leu (8.185%)	MVRFVSLVLCGVFLTAVVG
23648900	XP_004923234.1	putative alpha-L-fucosidase	477	55.387	-0.509	7.75	Leu (7.757%); Thr (7.128%); Gly (6.918%)	MKSFVFLVFGVNG
20068000	XP_004933849.1	15-hydroxyprostaglandin dehydrogenase [NAD(+)]=like	283	31.156	-0.188	6.252	Ile (8.834%); Ala (8.481%); Lys (7.774%)	MALVTLTFLVCLVSLFA
19455700	NP_001091826.1	ubiquitin and ribosomal protein S27a	1747	188.87	-0.707	4.483	Lys (15.129%); Leu (7.742%); Thr (7.097%)	MALVTLTFLVCLVSLFA
19077000	NP_001166287.1	sericin 2 isoform 1 precursor	1758	198.657	-1.052	2.158	Lys	

Supplementary Table S8. Coding sequences of P150/ser6 and Mucin-12 homologs for alignment and tree construction.

[illegible]