SUPPLEMENTARY MATERIAL

In the footsteps of sea stars: deciphering the catalogue of proteins involved in underwater temporary adhesion

Morgane Algrain, Elise Hennebert, Philip Bertemes, Ruddy Wattiez, Patrick Flammang, Birgit Lengerer

Published in Open Biology, 2022, Doi:10.1098/rsob.2016xxxx

(1) Whole mount *in situ* hybridization for transcript comp179_c0_seq3 (coding for Sfp11)

RNA probe synthesis and whole-mount in situ hybridization were performed as previously described (Lengerer et al., 2019). Sequence and primers are listed in Table S2. Briefly, transcript-specific primers were designed and a T7 promoter region was added at the 5' end of the reverse primers. The purified polymerase chain reaction (PCR) product was then used to produce single stranded digoxigenin-labelled RNA probes. RNA probes were used at a concentration of 0.1–0.2 ng μl^{-1} and detected with alkaline phosphatase conjugated anti-digoxigenin Fab fragments (Roche) at a dilution of 1: 2000. The signal was developed using the NBT/BCIP system (Roche) at 37°C. Images were taken with a Zeiss Axioscope A1 microscope.

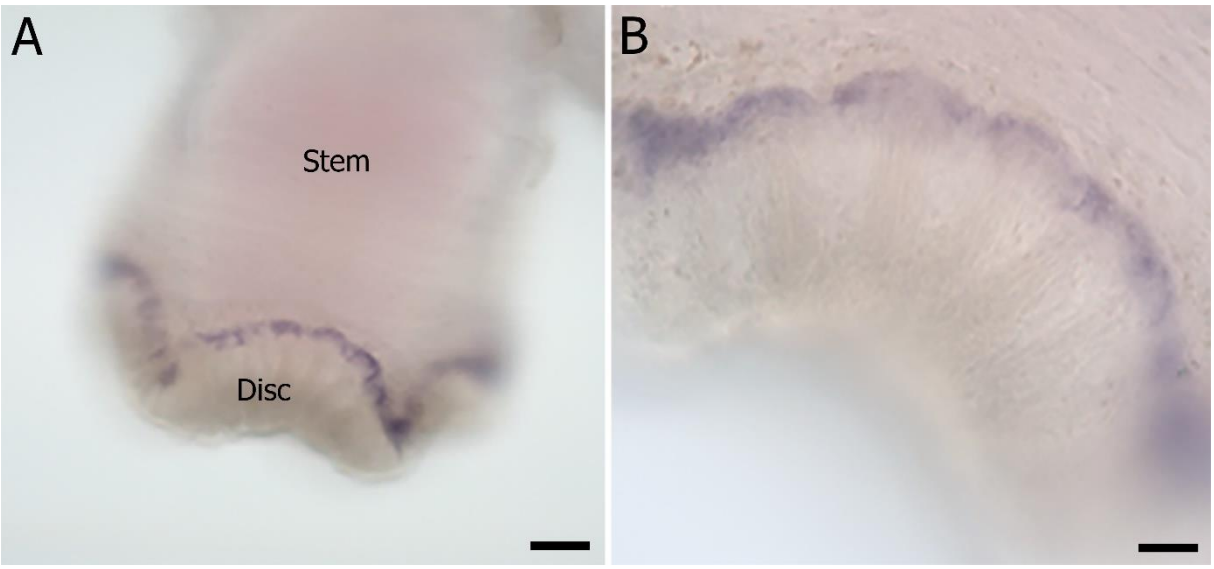
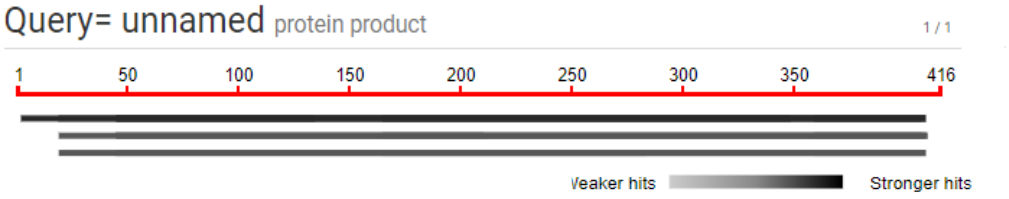


Figure S1. Whole mount *in situ* hybridization for comp179_c0_seq3 (coding for Sfp11) in *A. rubens.* Scale: 100 µm (A) and 25 µm (B).

(2) Sfp sequence elongation based on *Pisaster ochraceus* transcriptomic data

Partial Sfp sequences were elongated using publicly available transcriptomic and genomic resources. For example, the sequence of Arub-12 was originally identified in the tube foot transcriptome of *A. rubens* (Hennebert et al., 2015; Lengerer et al., 2019) but the N-terminal part was missing. To elongate the sequence, a tBLASTn search was performed in the transcriptome of *P. ochraceus* using the translated partial sequence (Fig. S2), and the best hit was the transcript 639619 from *P. ochraceus*, with an e-value of 2.50 10^{-122}. We aligned the translated sequence from *P. ochraceus* with Arub-12 (Fig. S3) and Arub-12 matched with the C-terminal end of the protein encoded by *P. ochraceus* transcript. We then performed a reciprocal tBLASTn search in the tube foot transcriptome of *A. rubens* using the N-terminal end of *P. ochraceus* transcript as query (Fig. S4). Arub-12 (comp15624_c0_seq2), and the other sequences with the best e-values (comp15624_c0_seq6, comp15624_c0_seq4, comp15624_c0_seq5) were aligned. Comp15624_c0_seq4 and seq6 were almost identical, only the 25 first nucleotides did not match. Comp15624_c0_seq5 was almost identical to the transcript coding for Arub-12, with only the 11 last nucleotides different. The comp15624_c0_seq4 sequence matched the *Pisaster* sequence better than seq6 at the N-terminal end, so this sequence was suspected to correspond to the C-terminal end of Arub-12. PCR (list of primers available in Table S1) and Sanger sequencing were performed to validate this elongation. The resulting sequence presented a perfect alignment with both comp15624_c0_seq2 and comp15624_c0_seq4 (elongated protein). Moreover, the reanalysis of the elongated protein sequence with the original mass spectrometry data (Hennebert et al., 2012, 2015) revealed a sequence coverage of 73% (Fig. S5). We concluded that the protein sequence encoded by comp15624_c0_seq4 is the elongation of Arub-12. The corresponding protein sequence is still missing the start, but could be elongated from 416 aa up to 747 aa and the protein was named Sfp10.



View More

Number	Sequences producing significant alignments	Total score	E value
1.	639619-Pisaster-ochraceus	1106.21	2.50×10^{-122}
2.	639618-Pisaster-ochraceus	893.59	8.24×10^{-77}
3.	639620-Pisaster-ochraceus	910.54	2.36×10^{-78}

Figure S2. tBLASTn search for Arub-12 in the transcriptome of *P. ochraceus*

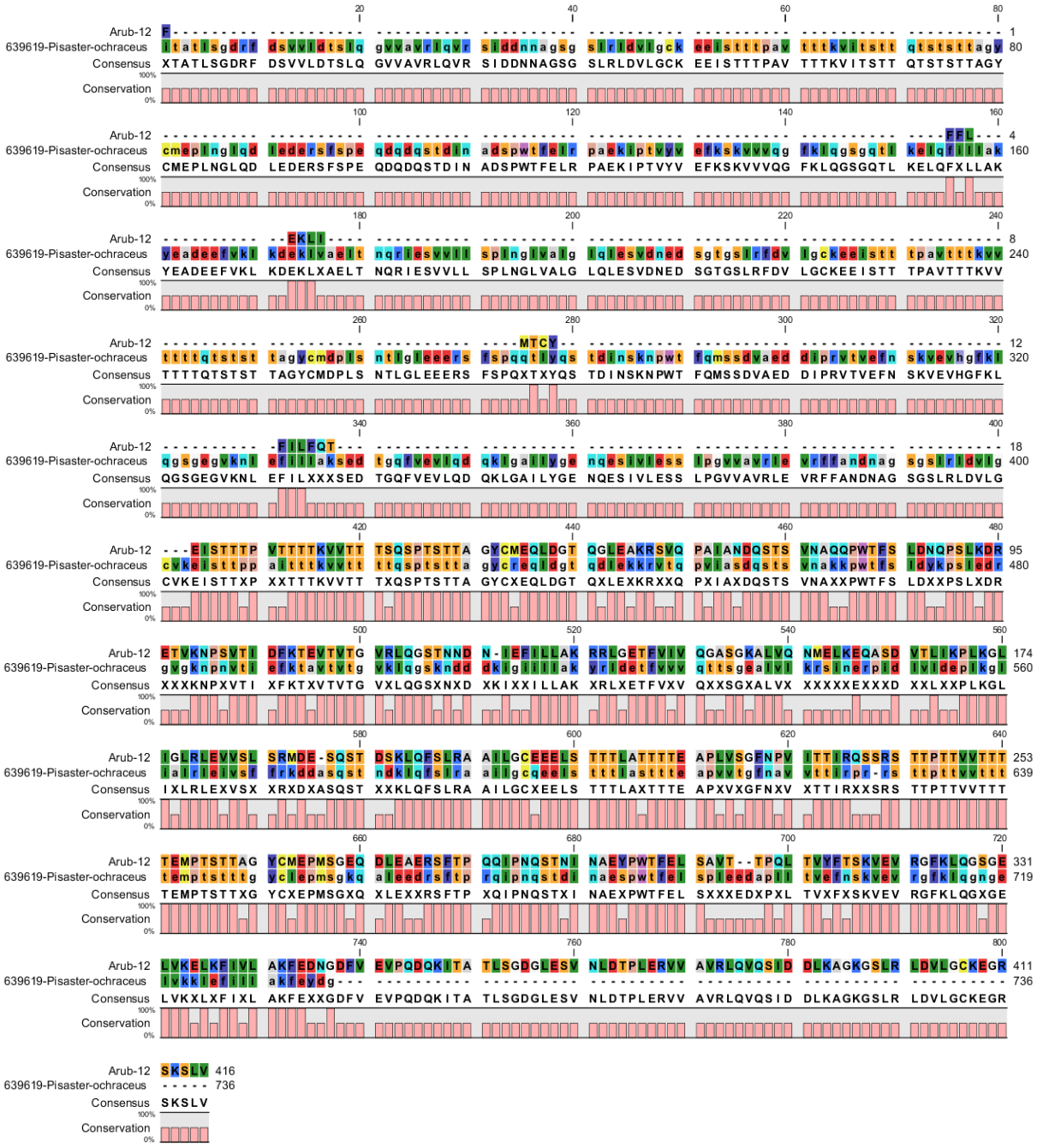
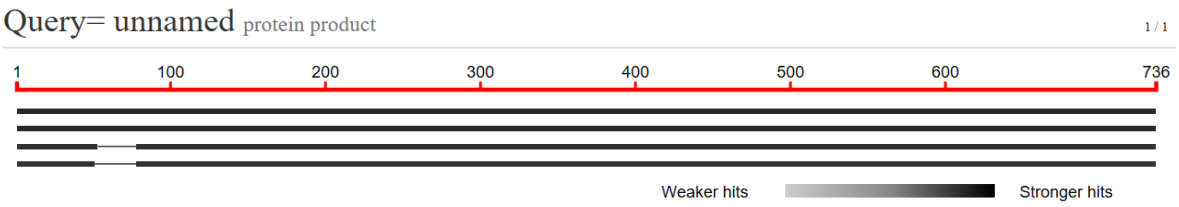


Figure S3. Alignment of Arub-12 and the protein sequence translated from transcript 63619 from *Pisaster ochraceus*.



Number	Sequences producing significant alignments	Total score	E value	Length
1.	comp15624_c0_seq6	997.21	1.19×10^{-120}	1056
2.	comp15624_c0_seq4	999.52	2.69×10^{-120}	1261
3.	comp15624_c0_seq5	982.57	1.25×10^{-108}	1260
4.	comp15624_c0_seq2	974.10	1.76×10^{-108}	1327

Figure S4. tBLASTn search of the N-terminal end of the sequence 63619 from *Pisaster ochraceus* in the tube foot transcriptome of *A. rubens*.

YKNPCLICFAVITTTTQTSTSTTAGECMEPLNGLLDLEERRFISPQQNQSTGINADNPWIFELLPYVD
AENKPTVTVEFKSKVEVHGFKLQGSGDIVKELQFVVSAKYEVDGLFVELEDEKLVADLTNQRVETVAL
PSPLSGLLALSLQLESVDQEDAGFGSLRFDVLGCEEEISTTTPVVTTTKVVTTTTQTSTSTTAGYCMD
PLSELDNLEQRSFAPQQTLENQSTLINSENPWTFQLSSDAGEEYDNPWVIVQYQSKVEVRGFKLQGSG
EDVITLEFILSAKSEDNEQFVDVLEGEKLGATLYQDNLKSITLDSPLPGVVAVKLQVIKKQNGGTGSL
RIEVLGCEKEISTTTPVTTTTKVVTTTSQSPTSTTAGYCMEQLDGTQGLEAKRSVQPAIANDQSTSVN
AQQPWTFSLDNQPSLKDRETVKNPSVTIDFKTEVTVTGVRLQGSTNNDNIEFILLAKRRLGETFVIVQ
GASGKALVQNMELKEQASDVTLIKPLKGLIGLRLEVVSLSRMDESQSTDSKLQFSLRAAILGCEEELS
TTTLATTTTEAPLVSGFNPVITTIRQSSRSTTPTTVVTTTTEMPTSTTAGYCMEPMSGEQDLEAERSF
TPQQIPNQSTNINAEYPWTFELSAVTTPQLTVYFTSKVEVRGFKLQGSGELVKELKFIVLAKFEDNGD
FVEVPQDQKITATLSGDGLESVNLDTPLERVVAVRLQVQSIDDLKAGKGSLRLDVLGCKEGRSKSLV*

Fig. S5. Protein pilot analysis of the elongated protein (Sfp10) with the original mass spectrometry data (Hennebert et al., 2012, Hennebert et al., 2015). The green color in the protein sequence indicates peptides matching the mass spectrometry data with high confidence (≥95%).

(3) Sequences elongation based on *Asterias rubens* genomic data

Using a local genome browser, we mapped the transcripts from the tube foot transcriptome and the corresponding Illumina sequencing raw reads (Hennebert et al, 2015) to the genome of *A. rubens* (eAstRub 1.3). Additionally, we mapped the sequences of the coding regions (CDS) predicted by the Sanger institute. As a typical example, transcript comp7100_c0_seq1 and transcript comp19_c0_seq1 were found to align to the same CDS, now named Sfp8 (Fig. S6). That the transcripts are indeed part of one longer sequence was also confirmed by a PCR and sequencing. Moreover, all elongated proteins were reanalysed with the original mass spectrometry data (Hennebert et al., 2012, 2015).

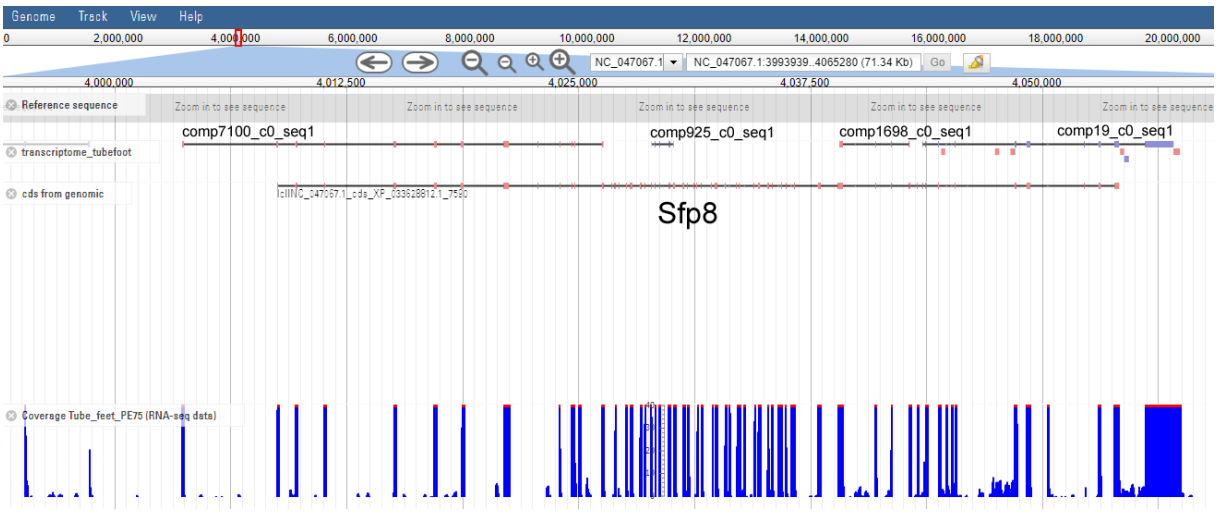


Figure S6. Screenshot from the genome browser in the region of the genome comprising the gene coding for Sfp8. The transcripts sequences from the tube foot transcriptome, the predicted CDS from the *A. rubens* genome, and the Illumina reads were all aligned to the genome sequence. The localization of the transcripts is represented as black line with pink/blue boxes indicating the position of the exons. The coverage of aligned Illumina reads is visualized with blue peaks at the bottom. Note that the predicted CDS from the *A. rubens* genome spans four separate sequences from the tube foot transcriptome.

(4) Protein sequences of Sfps

All Sfp sequences elongated using the *Pisaster ochraceus* transcriptome and the *Asterias rubens* genome were translated *in silico* and reanalyzed with the mass spectrometry data resulting from tube foot and footprint analysis (Hennebert et al., 2012, Hennebert et al., 2015) using the software ProteinPilot 5.0. All peptides identified with high confidence (>95%) in these extracts and present in the sequences are underlined. Amino acid sequences in bold show the parts of the proteins that were confirmed by PCR amplification and sequencing of the corresponding cDNA (to see Table S1 "primers and sequences elongation") and those in italics indicate the signal peptide identified with SignalP 5.0. Peptides used to raise polyclonal antibodies are highlighted in yellow in the corresponding sequences (Sfp7/8, Sfp10 and Astacin-like Sfp).

>Sfp2

*MASAPMIVLLPLLLLALTRTVVTS*DLKMVYVCEGDTLRLNCGEDRFIKMWSALYGRANPTYCPDDTVKASVHHCR
ADKSNFETMRGCDGKNSCEIEASDAVFGDPCPDETSKWYLQVEYYCRSERPS**ECNPKDSPKYQSAHACDDKELVL**
ECPTPGEVVKILDANYGRQGSDYCGYRPGIDDVFCDDDSGESFEYVCLQCRGKDYCMIPVNPLEFKDRCVNTRKY
LSVRYVCVQTEPEIEEEDDSEEEGPEVGDVDLRVTTIAPRRGRTQPATDFPFMMFPVGSAGASGKIKVGEAEHEG
ASFKFEGAPSIIGSANANGKLEVGEIDHEGGSFMFEGSPSRTKTIPGLSSIYTLPAGLVGQSAVLMKITVKYNEH
GEKYFTLTPFDGTSNTKTDGTEFLMKIDLNKILRDDSLQYTFRRLLENYNRGTTYYNNNDGVNTPFQFPVGNGLE
SFFSTLFQMIIETDSSGKMVTTPTNPSSNFQVILNGGNKAIVDYLMTFGSNTMDQNLRTELINHRLFISTSFYNT
LTQKTEEIKIPFYVNFDGLDGSLSTLFKSTSRPLTQFGMKYQVVGPGGAKQIFTMPFTLIEYQGRKRQFLTLPTC
TAFKMYGFVSDKEYLIDPDGINNGEEPFVVNCNMEKGETVVNHDRAESMVASYQAGKNTVDVNINYEGATKKQLL
ALRSVSDMCSQEMTHECVQKSTHFVNGKAHIWWESLDGVKRENWGTKTEEKMCHCGVNNACSDANSLCNCDSRGG
VSDWGTLTDKDTLPVGKVVFQKPVATDSEDNLKVGQFRCGGDQSSSTLQFFIFMNQMLANPKKTCHEFKELGYEL
DGVYLIDPDEKEGERPFTVFCEMEKQNEGITVVHHDLSDQRTQVNGYEEPGSYARKINYRATSRQIDNLKAVSGK
CEQELMYECQGSAFHVGTVLNHWWVSFNGDKMPNWGGAPCDFDGCACKSDSSCVDNLKCNCDKDDNVMRKDGGKL
DDKDKLPVSELRFGDTGSAEEMGYYTLGALKCWDASEKKITRRFLLKSCSEIKRDDPFALSGNYHIRPSDSKDEF
LVYCDMSGTEGVTVVYPDKGETLQIDGFYEEPGTWSRTIQYGASEPQLKALTTVSTTCKQYISHGCYGSLHHSSE
TGKMYTWWESRTGVKMMNWGEVADGCQCSVSNTCKGGPSTKCNCDVNTQSYMTDEGYLTDKTKLPVSGIRSGDVS
DDKGLEYGTAKLGPLMCQDAGSNEQTSRALAEADDVCDEITTEGNVERYLSVGDFKDFHFDLYGLGKATVNLAEA
SSLSASYKIEIGTDLDKQKVILSLCNDDSKCSVAEEKSFTSSTSSDGVERSFTVNVKSENIVLQSGDEALSFTRP
RQDGATHLEPKYFTFSTEEAISWKFDCPSFKYICSVKDAITCADKKPRIKTVRYGVPAANPCNEAVDYSQCSGMS
LLNKKSVSQLCSGKSCFYMKKSLEDGQCTAQGMVLKVEYECDYKEGRTLDLANPCNSNPCQNSGTCDSPDGEDTY
SCECGQYFTGLNCETELVYCSAWGDPHYQNFAGQNYDFMGNCTYVLAQDCIERRFKVLVRNESPKKNPAVSSTVM
VEVILDLVVIEMKQGGEIVVDDVPQSSPTVIEDTFSITDKGDIIELVAPEFQLTVTWNNKNNVEVFLPKYSGDIC
GLCGNTIGEPEVKDPVLFSNRWQVGECKSNTAVGERSNPAESAETLAFRRSIAETFCKAISVSDPFERICPQLYR
DMKSKYLNWCIFDNTQKSIQTYCSYIQEYKRKCDEESSEMFPTNWMKENTECYQGCPSHLNSTTCKSCESTCASL
RLPQTCTQVCNSGCQCPAGHYEEDNECVLEEQCGCAFEGKYYELGSEFLSADCQILHTCVKGNKVQRTTPTDRLC
HSTNPNTLCQVTNGVTECTCMDGFEPSTDKSTCNAISNQGLVRTGRFFNFWGGGFGNDKRRRRDAISEDCNVELG
TTSGKIPLSKITASSSKNQLSTASSARPGTPNSWIPRTNNNNQWIMVDINIPTEITGVTTWGGQREKRGMITDGW
VSAYGVELGMDGQTWTGVHDVQGRQQFSGNIDAISPQTNMLKEMERARFVRIIPISWHGQTRLQFEVIGCKQQCD
SEVGGFPVNRHISSDNWFVVNSAETINCVGVVKSWKVIRARNEPFEAWVLRPEKSDPKKFLLVGSTHIPAGPVDE
EVEFELDENQRIPVKTGDVVGFLYEHNPLYSSEEGEETVDFLWVNIRDITKYEMMRPGSHIKFTGGPARRAFSAV
AVLEPSTGAPVLPVLCNTEMGLNAIKKRIPDTSFTASSHVKGHSASRGRLHIESGQDGGGCGWTPSRRDRNPWLQ
VDLGDQQQLTGIVTQGRHTTAGRKKYWVHKYFVQFSQDAEHWNFIEKPNGRRQHFYGNVDKNTAKYNYLEKAIVA
RYIRIRPKNPRRKKIGLRVEFLGCPELDPCASSPCEQGGTCKQASGLNPDGYVCKCPQGWTGKNCDAEEGVCHSA
GDAHYHTFDGRHFYFPGECSYTLAKDCSVGATPIFDIVTNNAPTGRQIEIATNGIMFEFKQEGVVHVDGEACNLP
IKRYGVSVIHNALTGVLLETTYGLTVEWDGHHRVIVSLPSSFGGTVCGLCGNFNGDKDDDIQDPMTEAQFGESFS
TNAACGECAECILPKPCQQYPDKLTKARALCKPILDTAGPFADCHHVLKPTFIYGDCLEDICGSDLISIVPVCTQ
LAEYADLCAMHGVVIKSWRKTELCSITCPVGMMYNPCGSACPATCGDPHAPNTCTKPCEETCECPDGLLKEGNKM
CVEPSQCGCTDNGCYFESGSTFYRFGCTERGLCGLDGEVVYIPSQCGENAVCGSRNGVYGCYCDSGFSGDGQQCQ
SVEDDASLLVCEGEELSIFCAPGETIVIIEAVYGRSEDFQKCPSATQPRGVCDDGSALSVTRTKCQGKESCILPV
SSEVFGDPCFGVKKYLEVDYQCSALAAVPSESSQEIMKRVCQGDDMEIDCGDLFIEVVSAVYGRLTDGQVCPGGN

IQSTDCVAASSLLVVQEMCQGEHKCTGLQAASHVFGANPCEGTFKYLEVRYMCNAEVLPKESIIVNPCALEPCEN
GGQCVTVHTVEGFMCKCPHDWTGPTCENEEVSCRISGLSHFIDFNQQSFTFLGDSKYILAEGCLGSEHTFKVMVE
MSRSPINNEVTSVKDVHVHYDDRIVQLKHEHRVLVDGEKASSSVNLNFGKIKICRVGNYLLLRTDFGMIVAWDAL
NRVQVRIPDIVARKSCGLCGTYQGVSDEPVGFVNSDGNKVDDLHAFGSSWRVAADPPAFVHTNLIDTPACIKWPE
NLQYAKYACSSLDNDDQSSPFKSCFGKVNPSRYYENCLYDLCSSLPSRELFCQALAEYADACLDDGIIIENWRST
VLGGQCAMTCPKGQEYSACIQSCPKSCGEPHAGKEPVDRCYEGCTCKAGEDLKNFQDGDQCLPESKCGCARRGIY
YPPGQRAVNEVCTEQFVCLPGGDWISNPLTCDVNAHCGLVDYKHGCVCNAQHVGNGKRCYHHGDIKHEEASQGGK
LSINCGSEYVDIISGTYSVFDRSSAGSPTDRDCPTDNTLDVLANHIIGGSKLKVISSAMQFGDSCQPENNLLNVT
YLCSEELPVHEVEHGRVVHMRACENRMVQLKCKKGQKMEILSAKYGRNAGSEVCPSKQIRTTSCLAPQDRTLKKL
EKHCRGRDKCIIHVNGRRFFRNDPCSGTHKYLDVTFKCNDY*

>Sfp3

*MARLWLAAALCALVFIA***LANDAEGHGHHYGHRHGYHHHKSKGACKSKGRNIGHHHARKSYPRGNIGVLGSLYAKK**
GCSKRRHKSHSGHKHGHGLFKAFGDIKKCGSRRRKHGLSVSGNKKSHSHSHSGGRSHKLNINNKRRNRSRKRQKH
GGSGSKDSSGSSSHHVNKNNKRRNGTRKRQKHGGSGSKDSSGSSSHHVNKNNKRRNRTRRQKKRGSSSSSSSSSG
GNSDHQQTNTVDKKRPVPSKKEHPCRSNPCQNGAECYEKPNGDYTCSCKAGFRGTRCEQKIVVKVRPCDSNPCLN
HGQCIPSGSGYRCRCERGFAGENCQIVPKSFDPCAGVVCHNGGTCYAENGEAKCKCRTGFKGTLCRNKVDRSGAC
ANNPCKHNGDCHEIDNGNTFYCNCMGDWEGVTCKIAIDHSGPCDNNPCAHGHCVPSRNIFRCLCSGGYTGDLCDV
PPKKDPCLQNPCLNNGQCISLTTGDFNCKCRSGFEGAFCETGKMRPGLCPISNTGNCVRGIDGCNSDSECPGEDE
KCCSTGCSRVCRKPCRGVCYNFCPNGFVYDNDGCPTCQCKELPLPPKEEVVPPKPKVADQVGKENIAFNTKIDLS
KTTRVRLSVIWDDDLRISFKHPNSPNAFYVIYIGWETARIERCSSSGTCKMVRDVNTKGILKKSPVESINFYIII
SPEGDFIIGLDNPDYELMSYNDPNPVTGFANMYIFAGGVYPVDWTVATPTIEDSPMVQIDTDGRDTVVDLQRIRA
IDVRQTVKIRFSVIHGEQLQIGLKSTINLKKYIIIHIGKESSHVEECNEDEECTVSENFDTAGNIASDHFVNYFI
LFRRDGSVSVVYEDTQTEFITYSIIEPREGSFNVFISSGHGGQASWKVHQSTTIQEDTITLKPHPIEQNDKTWTF
ITEPLRSFEKEIQVGLAADAQIHIVLTGDKEKADSAYEIIIGAESNTKVIVKKGSQSVSANHVLSSGSYDDYSDY
WIGWSDGNLCVGDIDNYKKNDALACIAVPVKDQVFTWAGFKGDMMDYIIYYSFEYAVSYGIFSTYDSNMAYPLRF
KRARMTNNVLEFGVMTDSDAHMAVSNCQGADTPTAPMYELVMGTQGSEGCLFRRCKGCEAVAEGACPGLLDETEP
KDYTFEFNNGHLKLHKGRGVSKKLLLEWDDPNPIQNVGHLATATGFQSSGTSRWILYHP*

>Sfp4a

TQNPTKSASKENRPSSPGSSSHTDPDHVNSHASGSKSGSGSGRGSGSNERRGYYRSPCEKNPCKRGSCKEFGKTF
KCKCPSGFKGRRCQYTKRPRTTHTHGSKSSGSSSSAEIEAPQPCDYNPCWNGGTCLPEGDPYTCNCAPGYEGRRC
EIAPTPHPCEQDPCSGHGKCYSTRDGGYRCECDPNYGGERCNEVIKVNPCDRNPCKNDGRCLTRGDDYQCHCMFP
YEGTNCEQVQPGPCEDSPCTNGGTCFPDGQDFFCQCPKQFEGRRCEIALEVGPCQTEPCFNGGTCTPTATDFECA
CKPSYTGKRCEDSPDVNDHVILESTNGEMEDVDIPLLAFYTHSSYLLRFSVMCYDSVRIGFKSAVYSHVYYYVYL
GKTSARIEHCENTCQSVSEVDISGQLSATAFADYYIRFTSTGSLIVGILDASFTELMSFNGEPPTCPLLVFVSSG
DGYVATWKIYEAITEQVNTITWRTVNVNLRSDQYTYLGESFQVYDRQLRLGLAIDSEVTVLLSKGYDDANNYEIV
FRAAGNTKLQIKKNGEVVKEGDCVVNSSSLDDFEDFWISFAKSTISVGRVGLDMQPAFVSWSDTSSLEVSLAAFK
GASSSYLVYYKFEYQTARGIFWTFDCDMAYPLPFEKIKMPQTALEIGLTADCDAHFSIAATDEDSPDAPMYEVVI
GAEDNTKCLIRKCKECECMAEVDCPHILDKNKQVDFTFEFEFGHLKLFKGKGASKSAQPIMSCEDPEPLIGLEYL
ALSTGFESDGESHWRTYNV*

>Sfp4b

*MGRFCFDIAFCSLVCLFLAIYSVDGY*SISRGGSGSQESTKGSGGSSTGGSHEQHAPSPQPCDSHPCQHGGICHNH
GTTYECVCPSGCSGARCELENKLLDTIKGATGSGPSSSSSQTYPDHVNSKSNSGSNEVPALGDTQNPTKSASKES
RPSSSGSSSHTDPDHVNSKPSSESNEVPAPNPQPCDSHPCQHGGVCQNKGQLFECVCPSGCSGRRCENGPVQDPC
ANNPCEHGGACSPYGSSYRCSCPSAYMGSRCQLGKTQNPTKSASKESGHSSHGSSSHKDPDHVNPKSNSGSNEVP
ALGDTQNPTKSASKENRPSSPGSSSHTDPDHVNS*

>Sfp5

*MAKITLLVVLLVSVLPAI*ARKDQSYKCNFLESNGDGAYVFDDKQLNGGKFSVSFNVIANRDVRIALSPIKEEHNS
MIVISIGGPNNDDSTIGKSDGTILSEYLMTTGSRESELQCLTTCMNNYWICYEDNMLNVGRGGEIEPFLKYSVDA
ASSFDPGYVGFSTGSGVGMFGNIQAYESCADRQENPTMTSTLQPRSCATIDCGDVTCEWGFKKNAEGCMTCECKT
SPCSGLTCSETCANGYEVDDHGCFTCDCKSGPCDSQPCQNDGYCVSEDADEYRCICVNGFTGTNCERDEVPPTIE
CPANMEFNTDQGSAFATVTYELPEASDNDGEPSIECIPDFLTLGIGANTIKCKAVDAAGNEASCEYTITVKDEEP
PQMTCPVDMDVPTDEGLPTSSFDFENPTATDNDGGEPSVVCTADNSPMVLGENTITCTATDSGGLSVECSYTVTV
KDEEPPKMTCPVDMDVPTDEGLPTSSFDLENPTATDNDGGEPSVVCTADNSPMVLGENTITCTATDSAGLSVECS
YTITAKKKK

>Sfp6

*MLRTLILVTLLGYVVPAAVM*QRCEECAPPPPPPKRCHWAKSTPSECCPSKLVCEEPLTACRSGAKVAKYCPDSLC
DDTTCKGHPHAKCLISRCGGCSVKFYKHDELVKCAGHNKEIKPICPYLPVATCENSPCMYGKCKAEPSAPCRVDF
CQPCTAVFHDENNKEMQCSPS**VDTEPLPDFELDGETEPLQKDSKPHGETEPAQEEFGQQHSKPTQGEAISHEESE**
PAQEEPDQEQNKPTLGETVHQQENELVQRETESVEEIEPSTGAVESQEESEPVGKVEPSTGAVESQEESEPVGEV
EPSTGAVETQEESEPVGEIKPSTGESEPTQEEPDQEQNKPTLGETVHQQENELVQQETESVEVIEPSTGAVESHE
ESEPVGEIEPSTGESEPAQEEPDQEQNKPTLGETVHQQENKLVQRETESVEEIEPSTGAVESQEESEPTQEEPDQ
EQNKPTLGETVHQQEHELVQQETESVEEIEPSTGAVESQEESEPTQEEPDQEQNKPTLRETVHQQENELVQHETE
SVEEIEPSIGAVESQEESEPVGEIGPSTGESEPTQEEPDQEQNKPTLGETVHQQENELVQHETESVEEIEPSIGA
VESQEESEPVGEIGPSTGESEPTQEEPDQEQNKPTLGETVHQQENELVQQETESFEEIEPSTGESEPAQEEPDQE
QNKPTLGETVHQQENELVQRETESVEEIEPSTGAVESQEESEPTQEEPDQEQNKPTLRETVHQQENELVQHETES
VEEIEPSTGAVESQEESEPTQEEPDQEQNKPTLGETVHQQENELVQRETESVEEIEPSTGAVESQEESEPTQEEP
DQEQNKPTLGETVHQQENELVQRETESVEEIEPSIGVVETQEESEPTQEEPDQEQNKPTLGETVHQQENELVQQE
TESVEEIEPSIGVVETQEESEPTQEEPDQEQNKPTLGETAHQQENELVQRETESVEEIEPSTGAVESQEESEPVG
EIEPQTAKSEEVGHYCVDGETWKDGCADCQCVNGFPTCFWVGCGTFQVPEGIICNVRPGTENDCCPKYDCEEVES
HCVDKEGKEYAEGEQWKEGCEDCECSNGIPSCKLIPCALFMPPEGTQCKVRPGTENDCCPQFDCEEVENHCVDKE
GKKYADGEKWKEGKCGMDCQCSNGVPSCYSKGCSLFTVGLGVRCKVRPGTEEDCCPQHDCEQVESYCVDKEGKEY
ADGEKWKEGCVDCECSNGIPSCKLIPCALFMPPEGTQCKVRPGTENDCCPQFDCEDVGSHCVDKEGNKYAEGEKW
KDGCAECECLTGFSHICKSTSCPIYEVLEGSVCKVRPGTEEDCCPQYDCEEVESHCVDKEGKKYADGEKWKEGCV
DCECSNGIPSCKCALFMPPEGCKVRPGTENGCCPQFDCEDVGSHCVDKEGNKYAEGEKWKDGCAQCECLTGFSHI
CKSTSCPIYEVLEGSVCKVRPGTEEDCCPQYDCEEVESHCVDKEGKKYADGEKWKEGCVDCECSNGIPSCKLIPC
ALFMPPEGTQCKVRPGTENDCCPQFDCEDVGSHCVDKEGNKYAEGEKWKDGCAQCECLTGFSHICKSTSCPIYEV
LEGSVCKVRPGTEEDCCPQYDCEEVENHCVDKEGKKYADGEKWKEGKCGPDCQCSNGVLSCYSSGCSIFLVGEGT
QCKIRPGTEEDCCPQYDCEEAPSNKPCIDANGNTRKHREEYFLGKCDPCVCIDGIPVCAVVDCVHHPPVSPFEIC
HVRPRTEDQCCPEYDCTVKRERRKSVGANQHEVIDQPEEIDQPTDINQAGYTEQPDIA**GIEQPIETEQLVDSEHE**
DSEQPLVIEHTVGEQNPSSKLDCSTVRCKLARTKLPPLCRSIVLPGECCADGMVCEEPLSDCPIAVSVPESCPDD
LCLKATCESHPEAKCIISRCGGCFTKFFDDSGTVIEECSKDFANKTEMCPPPPLSFWCISPPCAYGQHCKAYPDA
ICHSTYCSIKCGLPEIFHDKYNNEIKDCSIEGFKNKCEEMGPCEDPPEGVLPDNSLCRWIVQPEECCPRDFGPNG
FVCEEPISACGDGVNTLQLDCPDDLCQRATCENHPGAKCKISHCGKCHTVFYDNDGNKVDCGNDPVESDTYCPRN
EWGRCSEPACPLPILPLPGLILICPRCFLDWCDSCKAYNFDAEFKLMDCDKEPCQCTMLIEGTEQPVDIEQPVVS
EQEDSEPPIVMEHSVGQQNMGGIEHQVDTEQLVDSEQEDSEQQLVIEHTVGEQNPSSKLDCSTVRCKLARTKLPP
LCRSIVLPGECCADGMVCEESLSKCPIAVSVPESCPDDLCLKATCESHPEAKCIISRCGGCFTKFFDDSGTAIEE
CNKDFSDLISCPPLPDTISCVSPPCAYGQHCKAYPDAICHSTYCSIKCGLPEIFHDKYNNEIKDCSIEGFKNKCE
EMGPCEDPPKGVLPMFCRWIVQPEECCPRDFGPNGFVCEEPISACGDEVNTLQLDCPDDLCQRATCENHPGAKCK
ISHCGKCHTVFYNCNGNKVDCGNNPVESDTFCPTNFWGRCAEPACPIPPPPQPGLICPRCFLDWCDSCKAYNYNA
EFKLMDCNNESCHCIMSQVIEQPVDNEQLVDEQEDSKQPIVMEHSVGDKMIH*

>Sfp7

*MFQLLLLIMVTVPASSLVQS*DVADMTCGVIRE**TGHSEEFMWDLPQLHNDPVVVEFSVYAKSEVYLALTDKPGPNL**
PAKATVYYIVIGGDGNTASYISRCRGCQPLTPASTEVLQCDRLNNYWLHFEDGFIKLGLRGKLEPITTYEVPEKD
FQPRYLGYSSNMLSQGMFTFSFDKFGYHTYSEHVYYTGEAFCLDWFLKEIPVGVKQFCFQAQVEQGPIYVVLSPQ
RFRNPGEIAYEIVIGEEENSKTFVFDKVRSGEQAIKEVFTPSVLNPNSFVEYCFTFNEKVITFGPASRRDLVVWV
RDESIQIKHVGFAGSKKPAQFLFPVDFIPSYYQGVIDLPPESPAGSDKCLEKPCAENGKCYVDHSFSHGYRCECS
NGYTGKNCDILPPETVPAVEESAVKSTKAPEKRPTPKPDTTPPTIDCNVELGDNYVTSADSKTMQVPFSFPRAED

NADKNP**YVVCVPSSNSEFPLGTSTVKCTATDESGNEATCDLVVEVKNSCENNPCGLHGTCVVDVFQPSEFNCVCI**
DDYEGELCTKAPKVDEEPLVEESDVKSSTKAPGRSSTTPKPPKPDVDEGETFTTSGDSSQEIKRTLDTKVISAFE
ISPFDDKVSKEFQIGCEIGEFKLDVNSETFSLKNGKGEVVDSETTPETSGSLKFSIEITESAFIISYILDQDFTT
LIEYSPPGREDGKELNCKAFTVKNVDKRRQRGCQIKYTKITQHKWRSYHRKPRRTRTRNIEKGETYVRKTSDEKT
RQRRPRTDRTKTRTIEKGVSERKKSSSGSNSGSSSEENKAISVGDTQVLPTNGQTSGQRRRPRTDRSKTRTIEKG
VSERKKSSSGSNSGSSSEENKAISVGDTQVLPTNGQTSGKRRRPRTDRTKTRTIEKGVSGRQKSGSRSSSGSSSE
EKNAISVGATVNKNTDVTTRRRTPAAKVKSSSATLVSGDTE

>Sfp8

*MARTFVCAIVVAIFVTIPAVA*SQLCQLYQTERNIGVQWLPETLPAERFEITFRVLALGDVYIYLGDGPHARLAIG
IGVDQNSATIIKGNSGKTYAKTPGIIKCDTFSEFWLLYSPSFVKLGRVGQTSPIVTLFHSESYVPTRLGFSTSSV
SVGLFSFDPFGKTAPTFFEQEFNTCTGFCYDWRLNPMPLGERICFQARTEGVAAIALSPMREPNAGQSLYEIVFG
YDENSMIVLRRHMHGEEIEVVEAKDTLNCDEFVSYCIEASGPHIQITRLGDPKILMDFTDSVGIPINYRYIGFAG
GFDHCVEWRWPVDIIPKLILETGTQLTEETGEAPAVDPVLPDAGETSIVVPDLPAAETDVQPPLDESNNPSGSQG
CAPVRCKKFCPFGR**ATDKNGCEKCKCNRKGKSSEEDPEISVGETDVQPPSDKPDTPRPRKDRPNNNDKPKTETDV**
QPPSDETDTPSGNQAVCPEVPEGVFGICSEECGADKKCDDGQLCCSNGCGHVCKKGCSPVRCRRYCPFGWATDED
GCDMCKCKRKGGSKGKKENRRRGESSEEDPEISVGETDVQPPSDKPDTPRPRKDRPNNNNKPKKETDVQPPSDET
DTPSGNQAVCPEVPEGVFGICSEECGADKKCDDGQLCCSNGCGHVCKKGCSPVRCRRNCPFGWATDEDGCDMCKC
KRKGGSKGKNKNPRRGESSEEDPEISVGETDVQPPSDKPDTPRPRKDRPNNNNKPKKETDVQPPSDETDTPSGNQ
AVCPEVPEGVFGICSEECGADKKCDDGQLCCSNGCGHVCKKGCSPVRCRRYCPFGWATDEDGCDMCKCKRKGGSK
GKKENRRRGESSEEDPEISVGETDVQPPSDKPDTPRPRKDRPNNNNKPKKETDVQPPSDETDTPSGNQVVCPEVP
EGVFGICSEECGADKKCDDGQLCCSNGCGHVCKKGCSPVRCRRYCPFGWATDEDGCDMCKCKRKGGSKGKNKNPR
RGESSEEDPEISVGETDVQPPSDKPDTPRPRKDRPNNNNKPKKGESSEEDPEISVGETDVQPSSDDPSVPRGNQD
GDEGETFTTSGDSSQEIKRTLDKKVISVFEISPFDDKVSKEFQIGCEVGDFKLDVNSETFSLKNGKGEEVDSQTT
PDTSGRLKFSIEITESAFIISYILDQDFTTLIEYSPPGREDGKELNCKAFTVKNVDKRRQRGCQIKYTKISQTRW
RSYNRRNRPSNTRKVQTGATYVPKKTSGSKPRRVREHSRGSGSGSGNSKGSKSGGGRRGRGSGSGSGKKDKILSA
GETFVPKKTSGSRPRPTVRERSSGSGSGSGNSKGSKSGGGRRGRGSGSGSGKKDKILSAGETFVPKKTSGSRPRP
TVRERSSGSESRRKKKS*

>Sfp9

*MGLSKSGTRAWIVVVCVIYVLPYATG*QNLKIQPVVDEPLIMPAPMPPAPVPPVDIDNEEAEEKIALEKQAEIRNR
AFGAGYLAIAPKVMCAGGHEKVFLTFTNFTQPVDVQFAVLDKNDESVGTSKLETFSTPCGCLDVHLNPSDSPRKT
VTLVMYAKRTLMQCEEFEIVKKIKVYLDACYTETFIETDKPMYKAGQNVHFRVLTLHPDLRPDISEVDKIWIEAP
GGVHMAQWLGVETNQGLIDETMKLSTNPTQGEWTIKVLHHNRDFSQKFTVGEYVLPKYEVMVQGPDYVVVDVEQI
SVVVCGRYTHGQPVVGSIILKVGVLNYPDKSVETSAEAKAGCFTFEVDMTKLDLMSSSYNVWNAKLHLEATFEEA
STNHIMLKTSDSTKIVNKPFHLEFSAPSTFKPGLSYSGALLFTDPSDKPLADTKSEINIEADGVSIFKETVISDA
KGMAFFSVTTIPSGSKSVTLKAVSKGYFRQSTWHPKNGGTHGIQDPTAYDVAQPQISPSNSFLHIDPVLETAGVN
SFQSLTIHLTSDKPEYNGVYLHTVMMSRGNILLSVESIIPPEDIDGFRRPGTQGVYPVDTDEYGRKEYPTRTNME
VTDIPVNILNKGQSKTTRRKIRILYEMAPAVSVMVYYIREDGEVVADTITIPVEEVFQNEVDVSFDEKTAAPGDS
AVLSVRAKASSLCAFGVVDKSVHLLGGDNRLTKKKVFEALEGLQLSSEGGSAEPSRCSSGGGMFRGYMPPQNLND
ASQAFKDLGVVYLTNLKVDTAPCPEVNPYMYFRGEVQNDGGDMMMAMPMMMMAVMEDSAIEGAVMKGAESDAAGP
TIRSYFPETWLYQLVEIDDSGHKDINVHVPHTITKWIGHGFCTSTKYGAGVSDLTSITAFQAFFIDMRLPYSVIR
YELVPVTVVVFNYITECLMVELTLEASSSYEIDDQNRVEKLCVCAGESQSVSFYVKAIELGDIPIQVAAVSVEAS
GACIGREMDDSLLGVSDTVLKKLLVKPEGIETEQVFSTFFCPVDYENGVYIEDIILKDLPENFVSGSERGIITLT
GDMMGPSISNLDQLLRMPTGCGEQNMIGFVPNIFVLHYLTGINKLTPGVEAKAKQHMEIGYQRELTYQHTDGSYS
AFGMSDPSGSTWLTAFVVRSFAQASKYIFIDPEELAKSKVWLLSLQQKDGCFKSVGKVIHKDMMGNVNDQATLTA
YVMISLIEAGDASEAEHIQKGVQCLNKEVDIQTADMYTLALSSYAYILAGSPHTDTLLERLNELAVESEGLMHWE
GAEVLNTDKRSPYLHRQASSQNIEITSYVLLTYVRSYDKSQALVKGNPVARWIVGQRNAEGGFSSTQDTVMALQA
LADYSMMAYSGGNLDISLGIKFSCDDDIRKLHVTQENSLLLQIEQIPEVPVEIKMTATGSGCALTQVIVTYNIIP
EVSAEPPFTIDISSVDNRNPRQFPCISYSMTICAHYNGEDEFSNMAVIQVRMVSGFEPESNSLQALEQEEGGPKM
VEHEGKEVNFYYDQLEKEDTCVSFMVTREEFVKEPRKGIVTVFDYYEKSLSVSELYDLECKR*

>Sfp10

YKNPCLICFAVITTTTQTSTSTTAGECMEPLNGLLDLEERRFISPQQNQSTGINADNPWIFELLPYVDAENKPTV
TVEFKSKVEVHGFK**LQGSGDIVKELQFVVSAKYEVDGLFVELEDEKLVADLTNQRVETVALPSPLSGLLALSLQL**
ESVDQEDAGFGSLRFDVLGCEEEISTTTPVVTTTKVVTTTTQTSTSTTAGYCMDPLSELDNLEQRSFAPQQTLEN
QSTLINSENPWTFQLSSDAGEEYDNPWVIVQYQSKVEVRGFKLQGSGEDVITLEFILSAKSEDNEQFVDVLEGEK
LGATLYQDNLKSITLDSPLPGVVAVKLQVIKKQNGGTGSLRIEVLGCEKEISTTTPVTTTTKVVTTTSQSPTSTT
AGYCMEQLDGTQGLEAKRSVQPAIANDQSTSVNAQQPWTFSLDNQPSLKDRETVKNPSVTIDFKTEVTVTGVRLQ
GSTNNDNIEFILLAKRRLGETFVIVQGASGKALVQNMELKEQASDVTLIKPLKGLIGLRLEVVSLSRMDESQSTD
SKLQFSLRAAILGCEEELSTTTLATTTTEAPLVSGFNPVITTIRQSSRSTTPTTVVTTTTEMPTSTTAGYCMEPM
SGEQDLEAERSFTPQQIPNQSTNINAEYPWTFELSAVTTPQLTVYFTSKVEVRGFKLQGSGELVKELKFIVLAKF
EDNGDFVEVPQDQKITATLSGDGLESVNLDTPLERVVAVRLQVQSIDDLKAGKGSLRLDVLGCKEGRSKSLV*

>Sfp11

*MEFFILVLLFSLLTTKGAVA*QDDLDWDGFGEADGSLYELKPLSIIGFHTEQSSDYDSNLYPSSKGVDGNVDHGFF
HTQNDEHSFWTVQLKEEHCIGKVTLFNRKDCCGARLANAVVRAGVSPFYSLNQRCGSPVESHLGNGGVIDVYCDP
PLLASEVNVVLPNREYLHLHEVQVEEYPIQRCHQIERRISILNSPTEQSSGYDDTKYPAERAADGFIGKFVGFSH
TGHDFDPWWRVD**LLDTHCISKVVLFNRYDCCSERLSNAVVRAGTSKDIHTNTRCGSAITAEQAKIKGGVLAVYCG**
PPLSGRYVSVNIPGRNEYLQLREVEIYELDVNECESDERHLSIIGKPTEQSSYYDNRYTPDKAADGNLDTKLSPE
FICTHTKEEDDPWWIVDLEDEHCISKVVLLNRGDCCSERLTNAVVRAGTSHNIQSNPQCGDVVSRAMAAPRGGTI
ELVCDPPLVARYISVDIPSDEHSFLQLCEVTVKEFIGEDACREPEPIVIKAPEPQQPVEEPKEAVEEDNQKKVLG
GGVKIQQDVPDDMGP*

>Sfp12a

CAEEFSTTEYVTTKSPTT**KGLPTGDTFLSSTTINSRHPPTTTATTTKTSSTTTHQLSTTQGYCLENMDGEEGLQE**
GFTPTSDLITLQVENGTTYIKFESLPLEVGDNNLPSVDGILTRSASIRAISIRLVPSSSSSTPSPTGNEELVSVT
INFFTKKERNDDFKPLLLEDNQEPKVITFTPVSDQPLQKFLKIGNEVIEGVVAFRVTVLEVTGVSSATLELHVLG
CAEEFTTTQVTTTAVSVTTPTTTTEESTSTVAVTTPTQTTTERTTTSTTSLVSTTEGYCLQNMDSESGVETKDQT
RGNLLTDESGFLTFSELPIEQSVGNSAVLERKFNSNINVRGVLITLKKITSSVTSGQAEGMTTPETGSTSPQSVK
FTFTLLAKRVDETDFKQVVIKNTITEFTLTIEGIQSEVKVFTEDEPSLQGITELQIKFIDVSGVSGA

>Sfp12b

Q**GITELQIKFIDVSGVSGASFELTVLGCAEEVSTTTQTTTSSTVSTTTPQLSTTPGYCLENMDGEEGLQEGFTPT**
SDLITLQVENGTTYIKFESLPLEVGDNNLPSVDGILTRSASIRAISIRLVPSSSSSTPSPTGNEELVSVTINFFT
KKERNDDFKPLLLEDNQEPKVITFTPVSDQPLQKFLKIGNEVIEGVVAFRVTVLEVTGVSSATLELHVLGCAEEF
TTTQVTTTAVSVTTPTTTTEESTSTVAVTTPTQTTTERTTTSTTSLVSTTEGYCLQNMDSESGVETKDQTRGNLL
TDESGFLTFSELPIEQSVGNSAVLERKFNSNINVRGVLITLKKITSSVTSGQAEGMTTPETGSTSPQSVKFTFTL
LAKRVDETDFKQVVIKNTITEFTLTIEGIQSEVKVFTEDEPSLQGITELQIKFIDVSGVSGA

>Sfp12c

PTTTTEESTSTVAVTTPTQTTTERTTTSTTSLVSTTEGYCLQNMDSESGVETKDQTRGNLLTDESGFLTFSELPI
EQSV**GNSAVLERKFNSNINVRGVLITLKKITSSVTSGQAEGMTTPETGSTSPQSVKFTFTLLAKRVDETDFKQVV**
IKNTITEFTLTIEGIQSEVKVFTEDEPSLQGITELQIKFIDVSGVSGASFELTVLGCAEEVSTTTQTTTSSTVST
TTPQLSTTPGYCLENMDGEEGLQEGFTPTSDLITLQVENGTTYIKFESLPLEVGDNNLPSVDGILTRSASIRAIS
IRLVPSSSSSTPSPTGNEELVSVTINFFTKKERNDDFKPLLLEDNQEPKVITFTPVSDQPLQKFLKIGNEVIEGV
VAFRVTVLEVTGVSSATLELHVLGCAEEFTTTQVTTTAVSVTTPTTTTEE

>Sfp13

MEEQVVCMLEFKTSAEVSCAPVLVTKLNAQGSIKVSTGDISQSQK**LPFEMKLKDGSTLIARQVSSIFSRMTIADL**
GIDIQLGLDKRIYVTLNPSLQGKVSGLCGNYNNKQIDDFTTSSGVNAKANINKFGNSWKVSLHCGDVEEKNLRPC
DLKSQQRPLAEEICRKLREEPFASCNQEIPVDKYVEQCEQDVCGCPPSDKECQCLAFATYARECSVAGFDFLWRG

LHNCPIKCPVGQVYKMCGSNCHVNCKDFLTQQPCHEECVEGCQCPEDKVRDDESGQCVKVDYCPCIVDGIVYQSH
QSWKKGCNDCFCSSGTHTCTELDCSTYETCADGMEWIPCLQCERTCQSANKACSIRSGCTPGCGCPKGLVLHKGS
CIDVSDCTCEYNDKSYKPFDTTAMDCHTCMCHTDYRWICDEDQCPSTCRSYGESHFQTLDAKWYSFQGNCEYTLV
ENFCGGSNVEGFFRVTIENVPCGRNGVTCTKAVKFTLHDTTIHLVRGAEYTVAKDPRVTTKARFKMEDAGLFLVI
KTAEGILLKWDYGTSVLVTLDPSHAGQVCGLCGNFNGDSSDDFYTRSGAMESSPQLFGRSWRTSENCPDPKELTL
PCKSEPHRRAWAESVCSIIKKSIFASCHNLVDSEPFFQACVSDTCFCDRGADCECMCTAVTAYATACNENNVAIA
WRAEGYCPLQCESGRVYQACGKICPDKCYPSHAEQNYGCEERCVEGCHCPEGTVLFRDECIERKLCPCFINNGQE
EIRHGALVVEDCMQCMCNGGHLECTGRNCTESVVAETTTPCAHKTKQCSCAPNFSCCADKSVCIHDSKFCDGEYN
CKDFSDEVHCNCIYNDIVYLTGDNWPGVGACEECKCEAGKTTCYKKCQLTCEFPEILNLYEDDDRKCCSCEGGED
TVRRTTTPPRTRPPRSPPTTTQSTTQECVDPFVVGEPTCCQYLDVCAINTKFSGSENCNACREPLVFNGTDCVRP
NECLCKDKVGYHLPGDTWENTEKACSVCTCFNNNIECSDVMCDDTMLDCSKYGEVFAYGACEECTCLDEIVCVKS
ECKLTLEDCEEGEVVFPIGDSCCECRAIETTTIVPTTTLPACVAPFVYRESCERHVMCETLGDDTNEVNCEEGCF
CPEGKLMSEKECIPSHQCPCTDIFGIKREFEEMWAQSDCQQCVCYENMTIHCTKECDLKCDESEEVLVNPDGSMC
CYCEPIPETTTMFIPETTTATTTPSGECFGCYDDKSDTCHAVDTSWKDDVCTSCECKYKNEFIMNCEHMQCGKDL
ECKEGHVLVKDDQECCYHCVPDICVGDFNCNSTFCIPNKWVCDGNMDCANEYDEENCPLPPTLPPPSPACAKTEE
RAFCKDCRERVCHDMECFEMPQELCCECKQGTVYNGTQCVHPKECQCLDENGNIRLPFDKWETADCQECKCHYNR
ITCYETGLCHSTTPSTSTLPPRPYYCTETCTCNLDCEGDFIDEKCLFAVPQECNDTCLCLPPMKKDDVAGRSCIY
QPSQEECIKATTDATTTRELHPISFIEKQIEIGKEKQGKRLQPGSSTPYHL*

>Sfp14

LITGGVQIKL**EGEQRVFLEYVTQGRKSKCTRVRETYTRTKSTVQQGSRRRVYSRQRGGEGGSGSRGGSNSGSSSE**
ENKAISVGATVNKNTDVSTRRRTPAATVQSSSATLVPGEMQQVDWTQKKYEVNFVDKIEDILDLTFQDGSVLTID
VTRTRMTIKDQSGKELKSESLRGKVRDNRLVFTIYLITGGVKIKLEGGQQTFLEYVKAGRTSKCNRVRESYKSSY
RRYL*

>Sfp15

*MWVLRG**FVVLLVIAECLC**RSPPFMKMCPVVPEDVGGICVEECGPDKKCDDGQLCCSNGCGHVCKAACPPLICGLT*
GCLYGRVTDENGCETCSCRGPPAVCPKAPEGVFGICVEECGPNKKCDDGQLCCPNGCGHVCKAGCAPVLCRMGCR
NGWATDEDGCEICQCKRPAQ*

>Astacin-like Sfp

*MKTLLITLLLVGFTLA*FVKNRGMRDDFNSKAVEFLSGRDGEAARAEMTPSPGCVEGDMELTSEQKDVYFAMREAT
AKNRARVRRDAQAAFSIWPNKTLVYRFDASLDNGIKQVFREAIEIYREHTGLSFEEAFDDYSGAIHPLVVYGLND
GCSSSVGYRERSGGRMHLETEKCGALGTILHELGHALGRFHEQTRSDRDDYITVIESNILPGFEHNFRKADRPIL
ALYDLQSIMHYSRDAFSFDRRSSTIMPHNAKDSEKMGTQRRLSRLDVFTVNAAYGVDALCPADLWDKCQRGGVPN
KHCSCDCLPEFEGPYCEIVVIQTECSGHYQLEESGLISSPGYPNTYSSNAKCSYVIECAKDEVVQLDFLEFSLEG
WGCFFDKMNMMTGDLVTEEYPDEYCDDNLKGESVFSKSNVALMTFESDDMFNYEGYQIFYHCIHSDLVSK*

Coding nucleotide sequences of Sfps

All Sfp cDNA coding sequences are listed here. Sequences in bold show the parts of the sequences that were confirmed by PCR amplification and sequencing of the corresponding cDNA (see Table S1 “primers and sequences elongation”). Primers used for ISH probe production are listed in Table S2. Note that some primers were designed in the non-coding UTR and therefore cannot be found in the coding sequences.

>Sfp2

ATGGCTTCAGCCCCTATGATTGTCCTGCTTCCTCTGCTGCTGCTGGCACTAACTCGAACGGTCGTTACTAGTGAC
CTCAAGATGGTCTACGTCTGCGAAGGAGATACTCTTCGACTCAACTGCGGGGAGGACAGATTTATCAAGATGTGG
AGCGCTCTCTACGGGCGTGCCAACCCGACTTACTGCCCAGACGACACCGTCAAGGCCTCTGTGCATCACTGCAGA
GCAGATAAGTCCAATTTCGAAACCATGCGTGGATGTGACGGAAAGAATAGCTGCGAGATCGAGGCAAGTGATGCA
GTCTTTGGTGACCCTTGTCCGGACGAGACGAGCAAGTGGTATCTCCAGGTGGAATACTACTGTCGCTCCGAGAGA
CC**ATCTGAATGCAACCCAAAGGATTCACCAAAATACCAGAGCGCGCATGCATGCGACGACAAAGAGTTGGTTCTG**
GAATGTCCGACCCCAGGAGAGGTTGTCAAGATCCTTGATGCCAATTACGGTCGCCAGGGATCCGACTACTGCGGG
TACAGACCGGGGATCGATGACGTCTTTTGCGACGACGATTCTGGGGAATCGTTTGAGTATGTTTGTCTTCAATGT
CGAGGGAAGGACTACTGCATGATTCCCGTGAACCCACTGGAGTTCAAAGACCGCTGTGTCAACACAAGGAAGTAC
CTCTCTGTTCGCTATGTGTGTGTACAAACAGAGCCGGAAATTGAAGAGGAAGATGATTCAGAGGAAGAGGGGCCA
GAGGTTGGTGATGTCGATCTTCGTGTTACAACCATAGCTCCAAGACGAGGAAGAACTCAACCTGCTACCGATTTT
CCTTTCATGATGTTTCCAGTTGGTTCAGCTGGTGCCAGTGGAAAGATCAAAGTGGGAGAAGCTGAACACGAAGGG
GCTTCATTTAAGTTTGAAGGTGCACCAAGCATAATTGGTTCAGCTAACGCCAATGGAAAGCTTGAGGTAGGAGAA
ATTGACCACGAAGGAGGCTCATTTATGTTCGAGGGTTCACCATCAAGAACTAAAACTATTCCTGGATTGTCAAGT
ATTTACACATTGCCTGCAGGTTTGGTGGGCCAGTCAGCGGTGCTGATGAAGATCACAGTGAAGTATAATGAACAC
GGAGAGAAGTATTTTACCTTGACACCCTTTGACGGAACTTCAAATACCAAAACAGATGGGACTGAATTTTTAATG
AAGATTGACCTGAATAAGATACTCCGAGATGACTCCTTACAGTACACCTTCAGGAGATTGTTGGAAAACTACAAT
AGAGGAACAACTTACTACAATAACAACGACGGGGTGAACACGCCGTTCCAGTTTCCCGTCGGTAATGGACTGGAG
AGTTTCTTCTCAACACTGTTTCAGATGATCATTGAAACGGACTCCAGCGGGAAGATGGTTACAACTCCAACCAAC
CCATCGTCCAACTTCCAAGTTATTCTGAACGGAGGCAACAAGGCCATCGTGGACTACCTCATGACCTTTGGCTCC
AACACAATGGATCAAAATCTACGAACTGAACTGATCAACCATAGGCTTTTTATATCAACAAGTTTCTACAATACG
TTGACCCAAAAAACTGAAGAAATTAAAATTCCCTTCTATGTCAACTTTGATGGATTGGACGGCTCGCTCTCGACT
TTATTCAAATCAACTTCAAGACCTCTGACCCAATTTGGAATGAAGTATCAAGTTGTGGGACCGGGCGGAGCCAAG
CAAATCTTCACCATGCCTTTTACTCTGATTGAATATCAGGGAAGAAAGAGGCAGTTTCTGACCCTGCCAACATGT
ACTGCTTTCAAGATGTATGGATTTGTGTCTGATAAAGAGTACCTCATTGATCCAGATGGGATCAACAACGGAGAG
GAACCATTTGTGGTAAACTGCAACATGGAAAAAGGTGAAACTGTTGTCAATCACGACCGCGCTGAATCCATGGTC
GCCAGTTACCAAGCCGGCAAGAATACAGTGGACGTTAATATCAACTACGAGGGCGCTACAAAGAAACAGCTCCTC
GCTCTGCGGAGCGTCTCGGATATGTGCAGCCAGGAGATGACGCACGAATGCGTGCAGAAGTCGACTCATTTTGTC
AACGGGAAAGCGCATATCTGGTGGGAGAGTCTGGACGGGGTCAAAAGAGAGAATTGGGGCACGAAAACTGAAGAA
AAGATGTGTCACTGTGGTGTTAACAATGCCTGCAGTGATGCCAACAGCCTCTGTAACTGTGACAGTAGAGGAGGT
GTCAGCGATTGGGGAACCCTCACAGACAAAGACACCCTTCCTGTCGGAAAGGTCGTCTTTCAGAAACCTGTGGCT
ACGGACAGTGAGGACAATCTCAAAGTCGGGCAATTCAGATGCGGCGGAGATCAGTCCTCAAGCACTCTTCAGTTT
TTCATCTTCATGAATCAGATGCTAGCCAATCCCAAGAAGACTTGCCATGAGTTTAAGGAGCTCGGTTACGAGTTG
GACGGTGTTTACTTAATCGACCCGGATGAGAAGGAAGGAGAACGACCTTTCACTGTCTTCTGTGAGATGGAAAAG
CAGAATGAAGGAATCACCGTTGTCCACCATGACCTCAGCGACCAACGAACCCAAGTCAATGGCTACGAAGAACCA
GGTAGCTACGCCAGGAAAATTAACTATCGAGCTACCTCGAGACAAATCGATAACTTGAAGGCTGTTTCCGGTAAA
TGTGAACAGGAGCTCATGTACGAGTGTCAAGGGTCTGCGTTTCACGTCGGCACGGTTCTCAACCATTGGTGGGTG
TCATTCAATGGGGATAAAATGCCAAACTGGGGTGGGGCCCCGTGCGATTTCGATGGGTGTGCTTGCAAGTCTGAT
TCGAGTTGCGTCGACAACCTCAAATGCAACTGTGACAAAGATGACAATGTCATGCGTAAGGACGGTGGCAAGCTG
GATGACAAGGACAAGCTACCCGTCTCCGAGCTTCGCTTCGGTGACACTGGCAGCGCTGAGGAGATGGGGTACTAC
ACTCTAGGAGCATTGAAATGCTGGGATGCCTCTGAAAAGAAAATAACAAGGCGCTTCTTGCTCAAATCTTGCAGT
GAAATCAAAAGGGATGATCCTTTTGCGCTGAGTGGTAACTACCATATAAGACCTTCAGACTCCAAGGATGAATTC
CTCGTCTATTGTGACATGTCCGGCACTGAAGGAGTGACGGTCGTTTACCCCGACAAAGGAGAGACTCTTCAAATC
GATGGTTTCTACGAGGAGCCCGGAACCTGGTCGCGAACGATCCAATATGGCGCATCGGAGCCGCAGCTAAAAGCC
CTGACAACGGTTTCTACTACCTGTAAACAATACATCAGCCACGGTTGTTACGGCTCACTCCATCACAGCAGCGAG
ACTGGCAAGATGTACACATGGTGGGAGTCTAGAACGGGAGTGAAGATGATGAATTGGGGGGAGGTGGCTGATGGC
TGCCAGTGCTCTGTCAGTAACACTTGTAAGGGAGGTCCAAGCACAAAGTGTAACTGTGACGTCAACACTCAGAGC
TACATGACGGACGAAGGTTACCTGACCGATAAAACTAAACTGCCTGTGAGTGGGATCCGCTCTGGTGATGTCAGC

GATGATAAGGGACTGGAATATGGAACGGCCAAGCTGGGGCCGCTGATGTGTCAAGACGCCGGTTCAAATGAACAG
ACCAGTAGAGCTCTCGCCGAAGCAGATGACGTCTGTGACGAAATCACAACGGAAGGCAACGTAGAACGTTATCTC
TCGGTAGGAGACTTTAAAGATTTTCACTTTGACCTGTACGGTTTGGGAAAAGCTACTGTGAATCTAGCCGAAGCT
AGTTCTTTGTCTGCGTCTTACAAAATTGAAATTGGAACGGACTTGGACAAACAAAAAGTGATCCTGAGCCTGTGC
AATGACGACAGCAAGTGTTCTGTCGCAGAAGAGAAATCCTTTACTTCCTCCACCAGTAGTGATGGTGTGGAACGT
AGCTTCACTGTCAACGTTAAGAGCGAGAACATCGTGCTGCAATCTGGAGACGAGGCCTTGAGTTTTACCAGGCCT
CGTCAGGATGGAGCGACACACTTGGAGCCCAAGTATTTCACCTTCTCAACAGAAGAAGCGATCAGCTGGAAGTTT
GACTGTCCATCTTTCAAATATATCTGCAGTGTAAAAGACGCAATCACATGCGCGGATAAAAAACCACGCATCAAG
ACGGTCCGCTACGGAGTCCCGGCCGCGAACCCGTGCAACGAAGCAGTAGATTATTCACAGTGCAGCGGCATGTCA
CTGCTGAACAAAAAATCTGTGAGTCAATTGTGTTCGGGGAAGTCCTGTTTCTACATGAAGAAATCTTTGGAGGAT
GGCCAGTGCACTGCTCAAGGGATGGTTCTTAAGGTGGAATATGAATGCGACTACAAGGAGGGTAGGACGTTAGAC
TTGGCCAACCCATGCAATAGTAACCCATGCCAGAACTCTGGGACGTGTGACTCCCCAGACGGTGAAGACACCTAC
AGTTGCGAGTGTGGACAATACTTTACGGGACTCAACTGTGAAACAGAACTGGTGTATTGCAGTGCCTGGGGCGAC
CCTCACTACCAAAACTTCGCAGGGCAGAATTATGACTTCATGGGTAATTGCACATACGTGTTAGCCCAAGACTGC
ATAGAACGTCGCTTTAAAGTTCTAGTGAGGAATGAGTCGCCTAAGAAGAACCCTGCCGTTTCCAGCACTGTTATG
GTTGAAGTTATTTTAGATCTTGTGGTAATTGAAATGAAGCAAGGTGGAGAAATCGTGGTCGATGATGTGCCTCAG
TCTTCACCAACAGTTATTGAGGATACTTTCTCTATAACCGACAAGGGCGATATCATTGAGCTTGTGGCACCTGAA
TTTCAGCTGACGGTGACGTGGAACAACAAAAATAACGTCGAAGTGTTTCTCCCAAAGTATTCAGGGGATATCTGT
GGATTATGCGGTAACACCATTGGAGAACCCGAGGTGAAGGACCCAGTTCTATTCAGCAATAGGTGGCAGGTAGGG
GAATGCAAATCAAATACCGCTGTTGGAGAACGGTCCAATCCAGCAGAATCGGCTGAGACTCTAGCTTTTAGAAGA
AGCATTGCTGAGACATTTTGCAAAGCCATCTCAGTCAGTGATCCTTTTGAGAGAATATGTCCACAATTATACCGC
GACATGAAATCGAAATACTTGAATTGGTGCATTTTTGACAACACACAAAAAAGCATCCAGACTTATTGCTCTTAC
ATCCAAGAGTACAAGAGGAAATGTGACGAGGAAAGCAGTGAAATGTTCCCAACCAACTGGATGAAAGAGAACACT
GAATGCTATCAAGGATGTCCATCCCATCTGAACTCAACCACATGCAAGTCATGCGAGTCTACTTGTGCGTCTTTA
CGTCTCCCACAAACGTGCACACAAGTGTGTAACTCGGGATGTCAATGCCCGGCCGGTCATTACGAAGAGGATAAC
GAGTGTGTACTTGAGGAGCAATGTGGATGCGCATTCGAGGGAAAGTACTATGAGTTGGGCTCGGAGTTCCTGTCC
GCGGATTGTCAAATTCTCCACACCTGTGTTAAAGGTAACAAAGTCCAAAGAACAACACCCACTGATCGGTTATGC
CACTCGACCAATCCAAACACTCTGTGCCAGGTGACTAACGGGGTCACAGAGTGTACCTGCATGGACGGATTTGAA
CCTAGTACCGACAAGTCAACATGTAACGCTATCTCAAACCAAGGCCTGGTCAGAACAGGGCGATTCTTTAACTTC
TGGGGTGGAGGATTTGGCAATGACAAACGTCGAAGGAGAGATGCAATTAGTGAAGACTGCAATGTGGAGCTTGGC
ACAACCTCGGGTAAAATACCTCTTTCTAAAATCACGGCGTCATCCAGCAAGAATCAGCTTTCAACAGCGAGTTCG
GCACGCCCTGGTACCCCCAACTCCTGGATACCCAGAACGAACAACAACAACCAATGGATCATGGTGGACATCAAT
ATTCCAACAGAGATTACTGGTGTGACTACTTGGGGTGGACAACGAGAGAAGAGAGGTATGATTACTGACGGCTGG
GTGTCGGCCTACGGAGTGGAACTTGGCATGGATGGTCAAACGTGGACAGGTGTTCATGATGTTCAAGGCCGTCAA
CAGTTTTCGGGAAACATCGACGCGATCTCTCCACAGACCAATATGTTGAAAGAGATGGAACGTGCACGCTTTGTA
CGTATCATTCCTATATCGTGGCATGGACAAACACGCTTGCAGTTTGAGGTCATTGGATGTAAACAGCAATGTGAT
TCAGAGGTCGGTGGTTTCCCGGTCAACCGCCACATCTCATCCGACAACTGGTTCGTCGTCAACAGCGCAGAAACC
ATCAACTGCGTGGGCGTGGTCAAGTCCTGGAAGGTAATCCGCGCTCGAAATGAGCCCTTCGAGGCATGGGTGCTC
CGCCCCGAGAAGTCGGACCCGAAAAAATTCCTTCTCGTAGGGAGCACTCACATCCCCGCCGGTCCGGTCGACGAG
GAAGTGGAATTTGAACTGGATGAAAACCAGCGCATACCCGTTAAGACAGGAGATGTTGTTGGGTTTTTGTACGAG
CACAATCCTCTGTACTCCAGCGAAGAAGGAGAAGAGACAGTGGATTTCTTGTGGGTGAACATCCGAGACATTACC
AAGTATGAGATGATGCGGCCCGGGTCGCATATCAAGTTTACTGGGGGGCCGGCCCGTCGTGCTTTTTCAGCTGTT
GCCGTATTGGAGCCTTCAACAGGTGCGCCAGTTTTGCCAGTTCTTTGCAACACAGAAATGGGATTAAATGCCATT
AAGAAACGGATCCCAGACACGTCCTTTACGGCTAGCAGTCATGTGAAGGGACACTCGGCATCGAGAGGACGTCTT
CATATTGAATCGGGGCAGGATGGCGGAGGATGTGGCTGGACACCTAGTCGTCGTGACAGAAATCCATGGCTGCAG
GTTGACCTCGGTGATCAGCAGCAGTTGACTGGAATCGTAACACAGGGTCGCCATACTACAGCTGGCCGAAAGAAA
TATTGGGTTCACAAATATTTCGTCCAGTTCAGCCAAGACGCAGAGCATTGGAACTTCATTGAAAAGCCAAATGGA
AGACGCCAGCATTTCTACGGCAACGTGGATAAGAACACAGCTAAGTACAACTACTTGGAGAAGGCCATCGTCGCA
AGGTACATTCGCATCAGACCTAAGAATCCACGCAGGAAAAAGATTGGTCTTCGTGTGGAGTTCCTTGGTTGCCCT
GAATTGGATCCATGTGCTAGCAGCCCATGCGAACAGGGAGGTACGTGCAAGCAAGCGTCAGGCCTAAACCCAGAC
GGTTACGTCTGCAAATGTCCGCAGGGATGGACCGGGAAGAATTGTGATGCAGAGGAGGGAGTTTGTCACAGTGCA
GGTGATGCACACTACCATACATTTGACGGACGCCACTTCTACTTCCCCGGCGAGTGCTCTTACACTCTAGCCAAA
GACTGTAGCGTCGGCGCAACTCCTATCTTTGATATCGTAACGAACAACGCCCCAACGGGACGGCAGATTGAAATC
GCCACTAACGGCATAATGTTTGAGTTTAAACAAGAGGGAGTCGTTCATGTTGACGGTGAGGCTTGCAATCTTCCC
ATCAAGCGTTACGGTGTATCGGTCATTCATAACGCACTGACTGGCGTACTGCTTGAAACAACTTATGGTTTGACT
GTAGAATGGGACGGACATCACAGAGTTATCGTAAGCCTGCCATCTTCATTTGGAGGCACCGTTTGTGGTCTTTGT
GGCAACTTCAACGGTGACAAAGACGATGATATTCAGGACCCTATGACCGAAGCACAGTTTGGAGAATCCTTTTCT
ACAAATGCTGCGTGTGGAGAATGCGCGGAATGCATATTGCCCAAACCGTGTCAACAATACCCTGATAAGCTCACC
AAGGCCAGGGCATTGTGCAAACCTATTCTTGACACTGCTGGTCCTTTTGCAGACTGCCATCATGTTCTGAAACCG
ACGTTCATTTACGGCGACTGTCTTGAAGACATTTGCGGCTCTGACTTGATCAGCATAGTACCCGTTTGCACCCAG
CTGGCTGAGTATGCTGATTTGTGCGCCATGCATGGAGTAGTTATCAAGTCATGGAGAAAGACGGAACTGTGCAGT

ATTACCTGTCCGGTGGGCATGATGTACAATCCTTGTGGCAGTGCTTGTCCAGCCACCTGCGGAGATCCTCATGCT
CCTAATACTTGCACCAAACCATGTGAAGAGACCTGTGAATGCCCAGATGGACTCCTGAAGGAAGGTAACAAGATG
TGCGTGGAACCCTCCCAGTGTGGATGCACCGACAACGGCTGTTATTTTGAGAGTGGCTCTACTTTCTACCGTTTT
GGTTGTACTGAGAGGGGTCTGTGCGGCTTAGATGGAGAGGTGGTATACATACCTTCACAGTGTGGAGAGAATGCT
GTCTGCGGATCTAGAAATGGAGTGTATGGTTGCTACTGCGATAGCGGATTCTCTGGAGATGGCCAACAGTGTCAA
AGTGTTGAAGACGACGCCTCTCTGCTCGTCTGTGAAGGTGAGGAGTTAAGCATCTTCTGCGCTCCTGGTGAGACT
ATCGTCATCATTGAAGCCGTATACGGAAGGAGCGAAGATTTCCAGAAGTGTCCGAGCGCGACACAACCACGCGGA
GTTTGCGATGACGGGTCCGCACTATCCGTTACGAGAACCAAGTGTCAAGGGAAGGAATCATGTATATTGCCGGTG
TCTTCTGAGGTTTTTGGAGATCCTTGTTTTGGTGTCAAGAAGTACCTTGAGGTTGATTACCAATGCAGTGCCTTG
GCTGCCGTACCTTCTGAATCCAGCCAAGAGATTATGAAGAGAGTATGCCAAGGAGATGATATGGAGATAGATTGT
GGAGATCTCTTTATTGAGGTTGTGAGTGCTGTCTACGGACGGTTGACCGATGGTCAAGTTTGTCCAGGAGGTAAC
ATTCAGAGCACGGACTGTGTAGCGGCATCGTCCCTATTGGTCGTTCAAGAGATGTGTCAGGGAGAACACAAGTGT
ACTGGTCTTCAAGCGGCCAGTCATGTGTTTGGTGCCAATCCATGTGAAGGAACATTCAAATACTTGGAGGTCCGG
TACATGTGTAACGCAGAGGTGTTGCCGAAAGAGAGCATTATCGTAAACCCTTGCGCCCTTGAACCGTGTGAGAAT
GGTGGTCAGTGTGTTACTGTTCATACTGTGGAAGGGTTCATGTGCAAATGCCCTCATGATTGGACTGGACCAACC
TGTGAGAATGAGGAAGTCAGTTGTCGCATATCGGGACTATCTCATTTTATCGACTTCAACCAGCAGAGTTTCACA
TTCTTGGGCGATTCTAAGTACATTTTAGCGGAAGGATGCCTGGGCAGCGAGCACACATTCAAAGTGATGGTCGAG
ATGAGCAGGTCACCCATCAATAATGAAGTGACCTCAGTTAAAGATGTCCATGTGCATTATGATGACAGGATTGTT
CAGCTCAAACATGAGCACAGGGTTCTAGTTGATGGAGAGAAGGCATCTAGCTCAGTCAACCTCAACTTTGGTAAA
ATAAAGATCTGTCGAGTCGGTAACTACTTGTTGCTGAGAACCGATTTTGGAATGATAGTTGCATGGGATGCTCTT
AATCGTGTTCAAGTAAGAATACCTGACATCGTAGCCAGGAAATCTTGCGGTCTGTGCGGCACATATCAAGGAGTC
TCAGATGAACCCGTTGGGTTTGTCAATAGCGATGGGAACAAGGTTGATGACTTGCACGCATTTGGATCGAGCTGG
CGTGTTGCAGCTGACCCTCCAGCCTTTGTTCACACCAATCTCATAGACACACCTGCCTGTATTAAGTGGCCAGAG
AATCTGCAGTACGCAAAATACGCTTGCTCCTCTCTGGACAATGATGACCAAAGCAGTCCGTTCAAGAGCTGTTTT
GGTAAGGTCAATCCATCAAGGTACTATGAGAACTGCTTGTATGACCTCTGTTCCTCGCTGCCTAGTCGGGAGTTA
TTTTGCCAGGCTTTGGCTGAGTATGCTGATGCTTGTCTGGACGATGGTATTATTATTGAAAACTGGAGGAGCACC
GTCTTAGGAGGACAGTGTGCAATGACATGTCCTAAAGGTCAAGAGTACAGCGCTTGCATCCAATCCTGTCCCAAA
TCATGCGGTGAGCCCCATGCTGGCAAAGAACCAGTAGATAGGTGTTATGAGGGATGTACTTGCAAAGCAGGAGAA
GATCTTAAAAACTTCCAAGATGGAGATCAATGTTTGCCAGAGTCAAAGTGTGGCTGCGCCCGTCGGGGTATCTAC
TACCCACCCGGACAGCGAGCTGTGAATGAGGTCTGCACAGAACAGTTTGTGTGTCTACCCGGGGGTGACTGGATC
TCCAACCCGTTGACATGTGATGTCAACGCTCACTGTGGATTGGTAGATTACAAGCACGGCTGCGTATGCAATGCT
CAACACGTTGGAAACGGCAAGAGATGCTACCACCATGGAGACATCAAACACGAAGAAGCGTCTCAAGGAGGGAAG
CTTAGCATCAACTGCGGATCAGAGTATGTAGACATCATCTCCGGCACCTACAGTGTCTTCGATAGAAGCAGTGCA
GGGTCACCCACCGACCGAGACTGCCCTACCGACAACACCCTGGACGTACTCGCCAATCACATCATTGGTGGTAGT
AAACTCAAGGTCATCTCCTCAGCCATGCAGTTCGGCGACAGCTGCCAGCCGGAGAACAACCTCTTGAACGTCACT
TATTTATGCTCGGAGGAGCTTCCAGTGCACGAGGTGGAACACGGTAGGGTGGTTCACATGAGAGCTTGTGAGAAT
AGGATGGTTCAGCTGAAGTGCAAGAAGGGACAGAAGATGGAGATTCTGAGCGCCAAGTACGGGAGGAACGCCGGG
AGTGAGGTATGCCCGAGCAAACAGATTCGTACGACAAGTTGCTTAGCACCACAGGACAGGACCCTTAAAAAGCTT
GAGAAGCATTGCCGTGGACGTGACAAGTGTATAATTCACGTGAATGGACGACGTTTCTTCAGAAATGATCCATGC
AGCGGAACTCACAAGTACCTTGATGTAACCTTCAAGTGCAACGACTATTAG

>Sfp3

ATGGCGCGTCTCTGGTTGGCTGCTGCTTTGTGTGCTTTAGTATTCATTG**CTCTAGCTAACGATGCCGAAGGACAT**
GGACATCATTATGGGCATCGCCATGGATATCACCATCATAAAAGCAAAGGCGCATGCAAAAGCAAAGGACGAAAT
ATAGGACACCATCATGCTCGCAAGAGCTACCCACGTGGAAACATAGGAGTTTTGGGAAGTCTATATGCCAAAAAA
GGATGCTCAAAGAGAAGACATAAGAGTCATAGTGGTCATAAACACGGACATGGTCTTTTCAAGGCTTTTGGTGAC
ATCAAGAAATGTGGATCAAGGAGAAGAAAACATGGCCTGTCCGTCTCAGGAAACAAGAAAAGTCACAGTCACAGT
CACAGTGGGGGCAGAAGTCATAAATTAAACATAAACAATAAGAGAAGAAACAGGAGCAGGAAACGACAGAAACAT
GGTGGGTCTGGCAGCAAGGACAGCAGCGGCAGCAGCAGTCATCACGTGAACAAAAACAATAAGAGAAGGAACGGG
ACAAGGAAACGACAGAAACATGGTGGGTCTGGCAGCAAGGACAGCAGCGGCAGCAGCAGTCATCACGTGAACAAA
AACAATAAGAGAAGGAACAGGACAAGGAGGCAAAAGAAACGTGGTTCATCATCCAGTTCATCATCCAGTTCTGGT
GGTAACTCCGACCATCAACAAACAAATACAGTTGATAAGAAGAGGCCAGTTCCCAGTAAAAAGGAGCATCCTTGT
CGAAGTAACCCGTGTCAGAATGGAGCTGAATGTTACGAGAAACCTAACGGGGATTACACTTGTTCTTGCAAAGCT
GGATTTAGAGGAACGCGCTGTGAGCAAAAAATAGTCGTCAAGGTGCGTCCTTGCGATAGCAACCCGTGCTTGAAT
CATGGACAATGCATTCCCAGCGGCTCGGGGTACAGGTGTAGATGTGAAAGAGGCTTCGCTGGGGAAAATTGTCAA
ATTGTGCCGAAATCATTTGACCCCTGTGCAGGCGTGGTTTGCCACAATGGAGGTACGTGTTACGCAGAAAATGGC
GAAGCAAAGTGCAAATGCCGGACTGGATTCAAAGGGACACTCTGCAGAAACAAGGTCGATAGGAGTGGAGCATGT
GCAAACAACCCATGTAAACACAATGGTGATTGTCACGAGATCGACAACGGCAACACCTTCTATTGCAACTGCATG
GGCGACTGGGAAGGTGTGACTTGTAAAATCGCCATCGATCATAGCGGACCTTGCGACAACAACCCGTGTGCACAT
GGCCATTGCGTGCCGTCTAGAAACATATTCCGTTGCTTGTGCTCGGGCGGTTACACTGGAGATCTATGTGACGTT
CCTCCCAAAAAAGATCCATGTCTACAAAATCCATGTCTAAATAATGGACAGTGTATTTCACTGACAACGGGTGAC

TTCAACTGCAAGTGCCGCAGTGGCTTTGAAGGCGCGTTCTGTGAGACGGGTAAGATGAGACCCGGACTCTGCCCT
ATCTCCAACACAGGTAACTGTGTGAGAGGTATAGATGGGTGCAATTCAGATAGCGAATGTCCAGGAGAAGATGAG
AAATGCTGCTCAACTGGATGCAGTCGCGTTTGCCGAAAACCATGTAGAGGTGTATGCTACAACTTCTGCCCGAAC
GGTTTCGTGTATGATAACGATGGCTGTCCAACATGCCAATGCAAAGAGTTACCGCTGCCACCTAAAGAAGAAGTT
GTACCACCAAAGCCTAAAGTAGCAGATCAAGTGGGAAAGGAAAACATTGCATTCAACACCAAAATCGATCTGTCT
AAGACAACTAGAGTTCGTCTCAGCGTCATTTGGGACGATGATCTGCGCATCTCCTTCAAACACCCTAATTCACCT
AATGCATTCTACGTCATATATATTGGTTGGGAGACTGCCCGCATCGAAAGGTGTTCTTCAAGTGGAACTTGTAAA
ATGGTTAGAGACGTCAATACCAAGGGCATCCTTAAGAAGTCTCCGGTTGAATCCATCAATTTCTACATCATCATT
TCACCAGAAGGCGATTTCATTATAGGCCTTGACAATCCGGACTACGAGTTGATGAGTTACAACGATCCCAACCCT
GTGACAGGTTTCGCTAACATGTACATCTTCGCTGGTGGCGTGTACCCAGTGGACTGGACGGTTGCAACACCAACA
ATTGAAGATTCTCCAATGGTACAGATAGACACAGATGGTCGCGACACAGTGGTGGATCTCCAGCGAATCAGGGCT
ATCGATGTCCGTCAAACTGTCAAAATAAGGTTCAGCGTCATCCACGGAGAGCAGTTGCAGATTGGACTTAAATCG
ACGATCAATTTGAAGAAGTACATCATCATACACATCGGCAAGGAGTCTTCTCATGTTGAAGAGTGTAATGAGGAC
GAGGAATGCACGGTGAGCGAAAACTTCGACACGGCAGGGAACATTGCTAGTGATCACTTCGTCAACTACTTTATC
TTATTCCGCCGTGACGGTTCTGTAAGCGTCGTGTATGAGGATACACAGACAGAGTTTATCACTTACAGTATCATT
GAACCACGAGAGGGCTCTTTTAACGTTTTCATCAGTTCTGGACATGGTGGCCAGGCCAGCTGGAAAGTACACCAG
TCAACCACTATACAAGAGGACACGATTACGTTGAAGCCACACCCTATCGAACAGAATGACAAAACATGGACCTTC
ATAACCGAACCTCTTCGTTCATTCGAAAAGGAGATCCAAGTGGGCTTGGCTGCTGATGCACAGATTCACATTGTG
CTTACGGGAGACAAAGAAAAGGCAGATAGTGCATACGAAATCATTATCGGAGCTGAGTCCAACACCAAGGTCATT
GTCAAGAAGGGTTCCCAGAGCGTTTCCGCGAACCATGTCCTCTCCAGTGGCAGTTATGATGACTACTCAGACTAT
TGGATTGGTTGGTCTGATGGTAACCTGTGCGTGGGTGACATTGACAACTACAAGAAGAACGACGCATTAGCCTGC
ATTGCTGTGCCAGTCAAAGATCAAGTCTTCACATGGGCAGGTTTCAAGGGCGACATGATGGACTACATCATATAC
TACTCGTTTGAGTACGCAGTCTCTTATGGCATTTTCTCTACCTACGACAGCAACATGGCATACCCATTGCGATTC
AAACGAGCTCGCATGACGAACAACGTATTGGAGTTTGGAGTCATGACAGACAGTGATGCACATATGGCCGTATCT
AACTGCCAAGGTGCGGACACACCGACTGCACCGATGTACGAGCTTGTCATGGGAACTCAAGGTTCCGAAGGGTGT
CTCTTCCGTCGTTGCAAGGGATGTGAGGCGGTGGCAGAGGGGGCTTGTCCAGGCTTACTTGACGAAACTGAACCC
AAAGATTACACCTTCGAGTTCAATAATGGTCATCTGAAACTTCACAAAGGAAGAGGAGTTTCCAAGAAGCTTCTT
CTGGAATGGGACGACCCTAACCCAATACAAAACGTGGGTCACCTGGCAACGGCTACTGGATTCCAAAGTTCTGGA
ACATCTCGCTGGATACTGTACCATCCATAA

>Sfp4a

TACACAAAACCCAACCAAGTCTGCGAGTAAAGAGAACCGCCCCTCTTCCCCTGGCAGCAGCAGTCATACAGACCC
TGATCACGTCAACTCACATGCATCAGGAAGTAAATCAGGGAGTGGTTCAGGGCGTGGCTCAGGCAGTAACGAAAG
AAGAGGCTACTACAGAAGTCCGTGTGAGAAAAACCCTTGCAAACGGGGATCTTGCAAAGAGTTCGGCAAAACATT
CAAGTGCAAATGTCCAAGTGGCTTCAAAGGAAGACGCTGCCAATATACAAAAAGACCACGCACAACTCATACCCA
TGGGAGCAAATCGAGCGGCTCGTCGTCTTCTGCAGAAATCGAAGCACCCCAACCATGTGATTACAACCCATGTTG
GAACGGAGGAACCTGTTTGCCCGAGGGTGACCCGTATACCTGTAATTGCGCACCAGGATATGAAGGAAGACGTTG
TGAAATTGCTCCCACACCTCATCCATGTGAACAAGACCCTTGCTCTGGCCACGGTAAATGCTACTCGACCCGGGA
CGGAGGATACAGGTGTGAATGTGATCCTAATTACGGGGGTGAACGTTGCAATGAAGTTATCAAGGTAAATCCATG
TGACAGAAACCCATGTAAGAATGACGGCAGATGTTTAACAAGAGGTGACGACTACCAATGCCACTGCATGTTTCC
ATACGAAGGCACTAACTGTGAACAAGTCCAACCTGGTCCCTGTGAGGACAGCCCATGCACCAACGGTGGAACCTG
CTTTCCGGACGGTCAAGATTTCTTCTGTCAATGCCCAAAACAATTCGAGGGTAGACGATGCGAAATAGCGCTTGA
AGTTGGACCATGTCAAACTGAGCCCTGTTTCAACGGAGGAACTTGTACACCAACAGCAACTGATTTTGAATGTGC
ATGTAAGCCAAGCTACACTGGGAAAAGATGTGAAGACAGTCCAGATGTAAATGATCACGTGATACTAGAATCAAC
CAATGGCGAGATGGAAGATGTTGATATCCCACTCCTGGCTTTCTACACACATTCGTCCTATTTGCTGCGGTTCAG
TGTCATGTGTTACGACAGTGTTCGCATCGGATTCAAATCTGCAGTTTATTCGCATGTCTACTACTACGTTTATTT
GGGTAAAACCTCAGCTCGGATTGAACACTGTGAGAACACATGTCAAAGTGTATCTGAGGTTGACATTAGCGGTCA
ACTGAGTGCGACAGCCTTTGCTGATTACTACATCCGATTTACATCAACTGGGTCTCTCATAGTTGGTATCCTTGA
TGCATCATTTACTGAGCTGATGTCATTCAACGGTGAACCACCAACATGTCCTTTGCTCGTGTTTGTCAGCTCTGG
GGACGGTTATGTGGCAACCTGGAAGATATATGAAGCCATCACCGAGCAAGTTAATACAATTACCTGGCGAACAGT
CAACGTGAACCTGAGATCCGATCAGTACACTTACCTGGGTGAAAGCTTCCAGGTATACGACCGACAATTACGACT
TGGTCTAGCGATAGATTCAGAGGTGACGGTATTGCTATCCAAGGGATATGATGACGCCAATAATTATGAGATTGT
CTTCCGAGCCGCAGGAAACACAAAACTGCAAATCAAGAAAAACGGTGAAGTTGTGAAAGAAGGGGACTGCGTGGT
TAATAGCTCCAGCCTCGATGACTTTGAAGACTTCTGGATAAGCTTCGCTAAGAGTACCATCTCAGTTGGTCGAGT
TGGTCTGGACATGCAACCTGCATTTGTCTCGTGGTCGGACACCTCGTCGCTTGAGGTCTCACTTGCAGCCTTCAA
GGGAGCGTCATCTTCATATCTTGTGTACTACAAATTCGAATATCAAACTGCAAGGGGTATATTCTGGACCTTCGA
TTGCGACATGGCGTACCCACTACCATTTGAGAAAATTAAAATGCCACAAACAGCGCTGGAGATTGGTTTGACCGC
AGACTGTGATGCTCATTTCAGTATTGCTGCTACAGACGAAGACAGCCCAGATGCTCCGATGTACGAAGTTGTGAT
TGGAGCTGAGGATAACACCAAATGCCTCATTCGCAAATGCAAAGAATGTGAGTGTATGGCCGAAGTAGACTGCCC
ACATATCCTAGATAAGAACAAGCAGGTCGACTTCACCTTCGAGTTTGAGTTTGGTCACTTGAAACTCTTCAAGGG

TAAGGGGGCTTCCAAGTCGGCCCAACCCATCATGTCATGCGAAGATCCAGAACCATTGATTGGTCTGGAGTACTT
GGCGTTAAGTACTGGATTCGAAAGCGACGGAGAATCACACTGGAGGACATATAATGTGTAA

>Sfp4b

ATGGGGCGATTCTGCTTTGATATCGCCTTCTGCAGTCTGGTATGTTTGTTCCTCGCTATCTACAGCGTCGACGGT
TACTCTATTTCTAGAGGAGGTTCTGGGTCTCAGGAAAGTACAAAAGGAAGCGGGGGTTCTTCAACAGGAGGCTCT
CATGAACAACACGCACCAAGTCCTCAGCCTTGTGATAGTCATCCATGCCAACATGGCGGTATATGCCATAATCAT
GGAACAACATATGAGTGTGTTTGCCCCAGTGGCTGCTCAGGAGCACGCTGTGAACTAGAAAATAAACTGTTGGAC
ACCATCAAAGGAGCTACAGGAAGCGGCCCCTCCAGCAGCAGCAGTCAAACTTACCCTGACCATGTCAACTCAAAA
TCAAATTCTGGAAGCAATGAAGTGCCAGCTTTAGGTGATACACAAAACCCAACCAAGTCCGCAAGTAAAGAGAGC
CGCCCCTCTTCCTCTGGCAGCAGCAGTCATACAGACCCTGACCACGTCAACTCAAAACCAAGTTCTGAAAGCAAT
GAAGTGCCAGCACCTAATCCTCAGCCTTGCGATAGTCATCCATGCCAACATGGCGGTGTATGCCAAAATAAAGGA
CAATTATTTGAGTGTGTTTGCCCCAGTGGCTGCTCAGGAAGACGCTGTGAAAATGGACCTGTTCAAGACCCATGT
GCGAATAACCCTTGCGAGCATGGGGGAGCATGTTCACCTTATGGATCCAGCTACCGATGCAGTTGCCCGTCAGCG
TACATGGGTTCTAGATGCCAACTAGGCAAAACACAAAACCCAACCAAATCTGCAAGTAAAGAAAGCGGACATTCT
TCCCATGGCAGCAGCAGTCATAAAGACCCTGACCATGTCAACCCAAAATCAAATTCTGGAAGCAATGAAGTGCCA
GCCTTAGGAGATACACAAAACCCAACCAAGTCTGCGAGTAAAGAGAACCGCCCCTCTTCCCCTGGCAGCAGCAGT
CATACAGACCCTGATCACGTCAACTCGTAA

>Sfp5

ATGGCTAAAATCACACTACTTGTGGTTCTTCTAGTGAGTGTTCTCCCTGCGATTGCAAGGAAAGATCAAAGTTAC
AAGTGCAACTTTCTGGAGTCCAATGGCGACGGGGCATACGTCTTTGATGATAAGCAACTGAACGGGGGAAAGTTT
TCAGTCTCCTTCAATGTAATTGCGAACAGAGATGTTCGTATTGCGCTATCGCCTATTAAAGAGGAACACAACAGC
ATGATTGTCATATCTATTGGTGGACCCAATAACGACGACTCCACCATAGGAAAGTCGGATGGTACGATATTATCC
GAGTACCTAATGACAACTGGCTCCCGGGAGAGTGAACTCCAATGCTTGACAACTTGCATGAACAACTACTGGATC
TGCTATGAAGATAACATGCTCAATGTTGGACGGGGCGGTGAGATCGAACCGTTCCTAAAGTACTCTGTCGATGCG
GCATCGTCATTCGATCCAGGGTATGTAGGCTTTTCGACAGGAAGCGGCGTTGGAATGTTCGGGAACATCCAAGCC
TATGAATCATGCGCAGATCGACAAGAGAACCCAACAATGACATCAACACTTCAACCAAGAAGCTGTGCAACTATT
GACTGCGGGGACGTGACGTGCGAATGGGGTTTCAAGAAAAATGCCGAAGGATGCATGACTTGTGAATGCAAAACA
TCTCCATGCAGTGGGCTGACGTGCTCTGAAACGTGTGCAAACGGCTACGAAGTCGACGACCATGGTTGCTTTACG
TGTGACTGCAAAAGTGGACCATGTGACAGTCAGCCCTGTCAGAATGATGGATACTGCGTTTCAGAAGACGCTGAT
GAGTATCGATGCATCTGCGTAAACGGTTTCACTGGAACAAATTGTGAACGAGATGAGGTTCCGCCAACAATCGAG
TGTCCCGCGAATATGGAATTCAACACGGACCAAGGTTCCGCATTTGCTACAGTCACTTACGAACTTCCCGAGGCT
AGTGATAATGACGGAGAGCCGAGTATCGAATGTATCCCAGACTTCCTAACATTGGGTATAGGCGCAAATACAATC
AAATGCAAAGCCGTTGATGCAGCCGGAAATGAAGCTAGCTGTGAATATACGATCACCGTGAAAGATGAAGAACCG
CCACAGATGACATGCCCAGTAGACATGGACGTACCGACAGACGAAGGATTGCCTACTTCTTCCTTTGATTTTGAA
AATCCAACAGCTACAGATAATGACGGCGGCGAACCCTCAGTTGTATGCACCGCTGACAACAGCCCTATGGTACTT
GGTGAAAACACAATTACATGCACAGCCACTGATTCTGGTGGACTGTCAGTGGAATGCTCGTACACAGTTACCGTA
AAAGATGAAGAACCGCCAAAGATGACATGCCCAGTAGACATGGACGTACCGACAGACGAAGGATTGCCTACTTCT
TCCTTTGATTTGGAAAATCCAACAGCTACAGATAATGACGGCGGCGAACCCTCAGTTGTATGCACCGCTGACAAC
AGCCCTATGGTACTTGGTGAAAACACAATTACATGCACAGCCACTGATTCTGCTGGACTGTCAGTGGAATGCTCG
TACACAATAACCGCAAAAAAAAAAAAAA

>Sfp6

TCAATGTATCATTTTGTCACCTACTGAATGCTCCATAACTATTGGTTGTTTGCTATCTTCTTGTTCATCCACTAG
TTGTTCATTGTCAACTGGTTGTTCAATAACTTGTGACATGATACAGTGGCAAGATTCATTGTTACAATCCATTAG
TTTAAACTCTGCATTATAGTTGTATGCCTTGCAAGAATCACACCAGTCCAAGAAACACCTGGGACATATCAACCC
AGGCTGTGGTGGTGGGGGAATTGGACACGCAGGTTCCGCACACCTCCCCCAAAAGTTTGTTGGGCAGAATGTATC
ACTTTCCACAGGGTTGTTTCCACAGTCCACCTTGTTACCATTGCAGTTGTAGAAGACAGTGTGGCATTTTCCACA
GTGACTGATCTTACATTTAGCACCAGGGTGGTTCTCACATGTTGCTCTCTGGCACAAGTCATCTGGACAGTCTAG
TTGCAATGTATTGACTTCATCACCGCAAGCTGATATTGGCTCTTCACACACAAAACCATTAGGACCAAAATCCCG
AGGGCAGCATTCTTCTGGTTGGACAATCCATCGGCAGAACATAGGAAGAACTCCTTTTGGAGGATCTTCGCAGGG
ACCCATTTCTTCACATTTGTTCTTGAAGCCCTCAATAGAGCAATCCTTGATCTCGTTGTTGTATTTATCATGGAA
TATCTCAGGAAGACCACACTTTATGCTGCAGTAAGTGGAATGACAAATAGCGTCTGGATATGCTTTACAATGTTG
CCCATATGCACAAGGAGGTGAAACACAAGATATGGTATCTGGTAGTGGTGGACAGCTGATCAAATCCGAAAAATC
CTTGTTGCATTCCTCAATGGCTGTGCCACTATCATCAAAGAACTTAGTAAAGCATCCTCCACAACGACTGATGAT
GCATTTAGCTTCAGGGTGACTCTCGCAGGTTGCCTTCAAACACAGATCATCAGGACAAGATTCTGGGACACTCAC

AGCGATTGGGCATTTAGAAAGAGATTCTTCACACACCATACCATCAGCACAACATTCCCCTGGTAAAACAATCGA
TCGGCACAGAGGGGGGAGCTTGGTTCTTGCCAATTTGCAACGCACAGTTGAACAGTCTAGCTTGGAAGATGGATT
TTGTTCACCTACGGTATGTTCAATAACTAACTGTTGTTCACTATCTTCTTGTTCACTATCCACTAGTTGTTCAGT
ATCAACTTGATGTTCAATACCTCCCATATTTTGTTGACCTACTGAATGTTCCATAACTATTGGTGGTTCACTATC
TTCTTGTTCACTAACCACTGGTTGTTCAATATCAACTGGTTGTTCAGTACCTTCAATTAACATTGTACATTGGCA
AGGTTCTTTGTCACAATCCATTAGTTTAAACTCTGCATCAAAGTTATATGCCTTGCAAGAATCACACCAGTCCAA
GAAGCACCTGGGACATATCAATATCAACCCAGGCAGTGGTAGTATGGGGAGGGGACACGCAGGTTCTGAACAACG
CCCCCATTCATTTCGTGGGCAGTATGTATCACTTTCCACAGGGTCGTTTCCGCAGTCCACCTTGTTACCATCGTT
GTCGTAGAAGACAGTGTGGCATTTTCCACAGTGACTGATCTTACATTTAGCACCAGGGTGGTTCTCACATGTTGC
TCTCTGGCACAAGTCATCTGGACAGTCTAGTTGCAATGTATTGACTCCATCACCGCAAGCTGATATTGGCTCTTC
ACACACAAAACCATTAGGACCAAAATCCCGAGGGCAGCATTCTTCTGGTTGGACAATCCATCGGCAGAGAGAGTT
ATCAGGAAGAACGCCTTCTGGGGGATCTTCGCAGGGACCCATTTCTTCACATTTGTTCTTGAAGCCCTCAATAGA
GCAATCCTTGATCTCGTTGTTGTATTTATCATGGAATATCTCAGGAAGACCACACTTTATGCTGCAGTAAGTGGA
ATGACAAATAGCGTCTGGATATGCTTTACAATGTTGCCCATATGCACAAGGAGGTGAAATACACCAAAAGCTTAA
TGGGGGTGGTGGACACATCTCGGTTTTATTGGCAAAATCCTTGCTGCATTCCTCAATAACTGTGCCACTATCATC
AAAGAACTTGGTAAAGCATCCTCCACAGCGACTGATGATGCATTTAGCTTCAGGGTGACTCTCGCAGGTTGCCTT
CAAACACAGATCATCAGGACAGGATTCTGGGACACTCACAGCGATTGGGCAATCAGAAAGAGGTTCTTCACACAC
CATACCATCAGCACAACATTCCCCTGGTAAAACAATCGATCGGCACAGAGGCGGGAGCTTGGTTCTTGCCAGTTT
GCAACGCACAGTTGAACAGTCTAGCTTGGAAGAGGGATTTTGTTCACCTACTGTATGTTCTATAACCAAAGGTTG
TTCACTATCTTCATGTTCACTATCCACTAGTTGCTCAGTTTCAATTGGTTGCTCAATACCTGCAATATCTGGTTG
TTCAGTATATCCTGCTTGATTGATGTCTGTTGGTTGATCAATCTCTTCTGGTTGATCAATGACCTCATGTTGATT
GGCTCCTACTGATTTTCTACGTTCTCGCTTTACTGTGCAATCATACTCTGGACAGCATTGGTCCTCTGTTCGAGG
CCTCACATGGCAGATTTCAAAGGGCGAGACTGGAGGGTGATGGACGCAATCTACAACTGCACACACAGGGATACC
ATCAATGCAGACGCAAGGATCACATTTTCCCAAAAAATATTCCTCTCTATGCTTACGCGTGTTTCCATTGGCATC
AATACATGGTTTGTTAGATGGAGCCTCTTCGCAATCATATTGTGGGCAACAATCTTCTTCAGTTCCTGGTCGTAT
TTTACATTGTGTACCCTCACCTACTAAAAAGATAGAACACCCACTGGAATAGCAAGAAAGAACACCATTGCTACA
CTGACAGTCCGGCCCGCATTTACCCTCTTTCCATTTCTCTCCGTCAGCATACTTCTTTCCCTCCTTGTCGACGCA
GTGATTTTCCACCTCTTCACAATCATATTGTGGGCAACAATCTTCTTCAGTTCCTGGTCGTACTTTACATACTGA
ACCCTCAAGTACTTCATAGATTGGACACGATGTGGATTTGCAAATATGACTGAAACCAGTCAAACACTCACACTG
TGCGCATCCATCTTTCCATTTCTCTCCCTCAGCATACTTATTTCCCTCCTTGTCAACGCAGTGACTTCCCACATC
TTCACAATCAAATTGTGGGCAACAGTCATTTTCTGTTCCTGGTCGTACTTTACATTGTGTACCCTCAGGTGGCAT
AAAGAGTGCACATGGTATTAATTTGCAAGAAGGAATACCATTGCTACACTCGCAATCCACGCATCCCTCTTTCCA
TTTCTCTCCATCAGCATACTTCTTTCCCTCCTTGTCAACGCAGTGACTTTCCACCTCTTCACAATCATATTGTGG
GCAACAATCTTCTTCAGTTCCTGGTCGTACTTTACATACTGAACCCTCAAGTACTTCATAGATTGGACACGATGT
GGATTTGCAAATATGACTGAAACCAGTCAAACACTCACACTGTGCGCATCCATCTTTCCATTTCTCTCCCTCAGC
ATACTTATTTCCCTCCTTGTCAACGCAGTGACTTCCCACATCTTCACAATCAAATTGTGGGCAACAGCCATTTTC
TGTTCCTGGTCGTACTTTACAACCCTCAGGTGGCATAAAGAGTGCACATTTGCAAGAAGGAATACCATTGCTACA
CTCGCAATCCACGCATCCCTCTTTCCATTTCTCTCCATCAGCATACTTCTTTCCCTCCTTGTCAACGCAGTGACT
TTCCACCTCTTCACAATCATATTGTGGGCAACAATCTTCTTCAGTTCCTGGTCGTACTTTACATACTGAACCCTC
AAGTACTTCATAGATTGGACACGATGTGGATTTGCAAATATGACTGAAACCAGTCAAACACTCACACTCTGCGCA
TCCATCTTTCCATTTCTCTCCCTCAGCATACTTATTTCCCTCCTTGTCAACGCAGTGACTTCCCACATCTTCACA
ATCAAATTGTGGGCAACAGTCATTTTCTGTTCCTGGTCGTACTTTACATTGTGTACCCTCAGGTGGCATAAAGAG
TGCACATGGTATTAATTTGCAAGAAGGAATACCATTGCTACACTCGCAATCCACGCATCCCTCTTTCCATTTCTC
TCCATCAGCATACTCCTTTCCCTCCTTGTCAACGCAGTAACTTTCCACCTGTTCGCAATCATGTTGTGGGCAACA
GTCTTCTTCAGTTCCTGGTCGTACTTTACATCGTACACCCAACCCTACTGTGAAGAGAGAACACCCCTTGGAATA
GCAAGAAGGAACACCATTGCTACACTGACAGTCCATCCCGCATTTTCCCTCTTTCCATTTCTCTCCGTCAGCATA
CTTCTTTCCCTCCTTATCGACGCAGTGATTTTCCACTTCTTCACAATCAAATTGTGGGCAACAGTCATTTTCTGT
TCCTGGTCGCACTTTACATTGTGTACCCTCAGGTGGCATAAAGAGTGCACATGGTATTAATTTGCAAGAAGGAAT
ACCATTGCTACACTCGCAATCCTCGCATCCCTCTTTCCATTGCTCTCCCTCAGCATACTCCTTTCCCTCCTTGTC
AACGCAGTGACTTTCCACCTCTTCACAATCATATTTTGGGCAACAATCATTTTCTGTTCCTGGTCGTACATTACA
TATTATACCTTCAGGTACTTGAAAGGTTCCACACCCTACCCAAAAGCAAGTAGGAAAACCA**TTAACACACTGACA**
ATCCGCGCATCCATCTTTCCATGTCTCTCCGTCAACGCAGTAATGTCCCACTTCTTCTGATTTAGCTGTTTGAGG
TTCGATCTCTCCCACTGGTTCACTTTCCTCCTGTGATTCAACTGCTCCGGTTGATGGTTCAATCTCTTCCACTGA
TTCGGTTTCCCGTTGTACTAGTTCATTCTCTTGTTGATGTGCAGTTTCTCCCAGTGTTGGTTTGTTCTGTTCTTG
GTCAGGTTCTTCTTGTGTTGGTTCACTTTCCTCTTGTGTTTCAACTACTCCAATTGATGGTTCAATCTCTTCCAC
TGATTCGGTTTCCTGTTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGGTTTGTTCTGTTC
TTGGTCAGGTTCTTCTTGTGTTGGTTCACTTTCCTCTTGTGTTTCAACTACTCCAATTGATGGTTCAATCTCTTC

CACTGATTCGGTTTCCCGTTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGGTTTGTTCTG
TTCTTGGTCAGGTTCTTCTTGTGTTGGTTCACTTTCCTCTTGTGATTCAACTGCTCCAGTTGATGGTTCAATCTC
TTCCACTGATTCGGTTTCCCGTTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGGTTTGTT
CTGTTCTTGGTCAGGTTCTTCTTGTGTTGGTTCACTTTCCTCTTGTGATTCAACTGCTCCAGTTGATGGTTCAAT
CTCTTCCACTGATTCGGTTTCATGTTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCTCAGTGTTGGTTT
GTTCTGTTCTTGGTCAGGTTCTTCTTGTGTTGGTTCACTTTCCTCTTGTGATTCAACTGCTCCAGTTGATGGTTC
AATCTCTTCCACTGATTCGGTTTCCCGTTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGG
TTTGTTCTGTTCTTGGTCAGGTTCTTCTTGTGCTGGTTCACTTTCCCCAGTTGATGGTTCAATCTCTTCAAATGA
TTCGGTTTCCTGTTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGGTTTATTCTGTTCTTG
GTCAGGTTCTTCTTGTGTTGGTTCACTTTCCCCAGTTGATGGTCCGATCTCTCCCACTGGTTCACTTTCCTCTTG
TGATTCAACTGCTCCAATTGATGGTTCAATCTCTTCCACTGATTCGGTTTCATGTTGTACTAGTTCATTCTCTTG
TTGATGTACAGTTTCTCCCAGTGTTGGTTTGTTCTGTTCTTGGTCAGGTTCTTCTTGTGTTGGTTCACTTTCCCC
AGTTGATGGTCCGATCTCTCCCACTGGTTCACTTTCCTCTTGTGATTCAACTGCTCCAATTGATGGTTCAATCTC
TTCCACTGATTCGGTTTCATGTTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCTCAGTGTTGGTTTGTT
CTGTTCTTGGTCAGGTTCTTCTTGTGTTGGTTCACTTTCCTCTTGCGATTCAACTGCTCCGGTTGATGGTTCAAT
CTCTTCCACTGATTCGGTTTCCTGTTGTACTAGTTCATGCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGGTTT
GTTCTGTTCTTGGTCAGGTTCTTCTTGTGTTGGTTCACTTTCCTCTTGCGATTCAACTGCTCCGGTTGATGGTTC
AATCTCTTCCACTGATTCGGTTTCCCGTTGTACTAGTTTATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTAGG
TTTATTCTGTTCTTGGTCAGGTTCTTCTTGTGCTGGTTCACTTTCCCCAGTTGATGGTTCGATCTCTCCCACTGG
TTCACTTTCCTCGTGTGATTCAACTGCTCCAGTTGATGGTTCAATCACTTCCACTGATTCGGTTTCCTGTTGTAC
TAGTTCATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGGTTTATTCTGTTCTTGGTCAGGTTCTTCTTGTGT
TGGTTCACTTTCCCCAGTTGATGGTTTGATCTCTCCCACTGGTTCACTTTCCTCTTGTGTTTCAACTGCTCCAGT
TGATGGTTCGACTTCTCCCACTGGTTCACTTTCCTCTTGTGATTCAACTGCTCCAGTTGATGGTTCGACTTTTCC
CACTGGTTCACTTTCCTCTTGTGATTCAACTGCTCCAGTTGATGGTTCAATCTCTTCCACTGATTCGGTTTCCCG
TTGTACTAGTTCATTCTCTTGTTGATGTACAGTTTCTCCCAGTGTTGGTTTATTCTGTTCTTGGTCAGGTTCTTC
TTGTGCTGGTTCACTTTCCTCATGTGATATTGCTTCCCCTTGTGTCGGTTTACTATGTTGTTGTCCAAATTCTTC
TTGAGCTGGTTCTGTCTCCCCGTGTGGTTTAGAATCTTTTTGAAGAGGTTCTGTTTCTCCGTCTAACTCAAAGTC
TGGCAGCGGTTCAGTATCAACAGATGGGGAACATTGCATCTCTTTATTGTTTTCATCATGAAACACTGCTGTACA
TGGCTGACAGAAGTCAACTCGACAAGGAGCACTTGGTTCAGCCTTACATTTACCATACATACATGGTGAATTCTC
GCAGGTGGCAACAGGAAGATAAGGACAGATTGGTTTGATTTCTTTATTATGTCCTGCACATTTTACAAGTTCATC
GTGCTTATAGAACTTGACAGAACAGCCTCCACATCTACTGATCAGACACTTCGCATGAGGGTGACCCTTACAGGT
TGTGTCATCACAGAGTGAATCTGGACAATATTTAGCCACTTTTGCTCCTGAACGACAGGCAGTCAGAGGTTCTTC
ACACACCAGTTTTGAAGGGCAGCATTCGCTCGGCGTACTTTTTGCCCAGTGACAACGTTTGGGAGGCGGCGGAGG
AGGTGCACATTCCTCACATCTCTGCATAACAGCCGCCGGAACGACATATCCCAGCAACGTCACCAATATCAGCGT
TCTTAACAT

>Sfp7

ATGTTTCAACTTCTCCTACTTATCATGGTCACTGTACCCGCGTCGTCATTGGTGCAAAGTGACGTTGCTGATATG
ACTTGCGGCGTTATAAGGGA**GACGGGACACAGTGAGGAATTTATGTGGGATCTTCCCCAACTTCATAACGACCCA**
GTTGTAGTTGAATTTTCTGTCTACGCAAAGAGTGAAGTTTACCTTGCCTTAACAGACAAACCGGGCCCTAATTTG
CCAGCTAAAGCAACTGTATATTACATTGTGATTGGTGGGGATGGGAACACTGCTTCTTACATCTCACGGTGTAGG
GGCTGTCAACCACTGACTCCAGCATCGACTGAAGTACTACAGTGCGACCGTCTTAACAACTACTGGCTCCACTTC
GAAGACGGCTTCATCAAGCTTGGCTTACGAGGAAAGCTTGAACCAATCACAACGTACGAGGTTCCTGAAAAAGAC
TTCCAGCCCCGGTATTTGGGATACTCCTCCAACATGCTGAGCCAGGGGATGTTCACTTTCTCCTTCGATAAATTT
GGATATCATACCTACAGTGAGCACGTGTACTATACAGGCGAAGCGTTCTGCCTCGACTGGTTCTTGAAAGAAATC
CCAGTAGGGGTTAAACAGTTTTGTTTCCAAGCCCAGGTAGAGCAAGGACCGATTTACGTTGTTTTGAGCCCTCAG
CGATTCAGAAATCCTGGTGAAATAGCTTACGAAATAGTGATTGGAGAAGAAGAAAATAGCAAGACTTTCGTCTTC
GACAAAGTTCGTTCCGGGGAGCAAGCAATCAAGGAAGTTTTCACCCCTTCAGTTTTAAATCCAAACAGTTTCGTA
GAGTACTGTTTTACTTTCAATGAGAAGGTGATTACTTTCGGCCCCGCGAGTAGAAGAGATTTGGTGGTTTGGGTG
AGAGACGAGTCTATTCAAATCAAACATGTCGGATTCGCAGGAAGTAAGAAACCGGCCCAGTTCTTGTTTCCCGTG
GACTTCATTCCTTCATATTACCAAGGCGTTATCGACTTACCGCCAGAATCACCCGCCGGGTCAGATAAATGCTTG
GAGAAACCGTGTGCAGAAAACGGTAAATGTTATGTTGACCATTCGTTCAGCCACGGTTACAGATGCGAATGTTCT
AATGGGTACACCGGGAAAAACTGCGATATACTTCCACCGGAAACTGTACCGGCGGTCGAGGAGTCAGCTGTCAAA
TCGACTAAGGCTCCTGAAAAAAGGCCTACACCTAAGCCCGATACAACACCGCCGACTATCGATTGCAATGTTGAG
CTCGGGGACAACTATGTGACATCAGCAGATTCAAAGACGATGCAAGTTCCGTTTAGTTTTCCACGGGCCGAAGAT

AATGCGGACAAAAACCCATACGTCGTGTGTGTTCCAT**CTTCTAACAGTGAATTCCCCCTTGGTACTTCCACGGTG**
AAGTGTACAGCTACTGACGAATCCGGAAACGAAGCGACATGCGACTTGGTAGTCGAGGTGAAAAATTCCTGTGAA
AACAATCCTTGTGGACTACACGGAACTTGCGTGGTTGATGTGTTTCAACCTTCAGAATTCAACTGTGTGTGCATC
GATGATTACGAAGGGGAATTATGCACAAAGGCACCAAAGGTTGATGAGGAACCGTTGGTCGAGGAATCAGATGTC
AAATCATCCACCAAAGCTCCTGGACGGAGCTCAACTACTCCTAAGCCCCCTAAGCCAGATGTTGATGAAGGGGAG
ACTTTTACAACTTCAGGAGATTCTTCACAGGAGATCAAGAGGACTTTGGATACGAAAGTAATCAGTGCCTTTGAA
ATATCTCCATTCGACGATAAGGTATCAAAGGAGTTCCAAATCGGTTGTGAAATCGGCGAGTTCAAACTTGATGTC
AATAGTGAAACCTTTTCACTCAAAAACGGCAAAGGCGAAGTGGTGGACTCTGAGACCACGCCTGAAACTTCGGGT
AGTCTCAAGTTCTCAATTGAAATTACCGAGTCAGCTTTCATAATTAGCTACATCTTGGACCAAGACTTCACTACC
CTTATCGAATACTCGCCTCCTGGAAGGGAGGATGGCAAGGAATTGAACTGCAAGGCTTTCACCGTTAAGAACGTG
GATAAGCGTCGTCAGCGAGGTTGTCAGATTAAATACACCAAGATCACTCAACACAAATGGAGGAGCTACCATAGG
AAACCGAGACGAACAAGGACTCGAAACATCGAGAAGGGAGAAACGTACGTTCGTAAAACCAGCGATGAAAAGACC
AGGCAGAGACGACCGAGAACGGACCGAACAAAGACGCGGACAATCGAGAAGGGAGTGAGTGAACGCAAGAAAAGT
AGCAGTGGATCGAATTCAGGGTCAAGTTCTGAAGAGAATAAAGCCATATCTGTTGGAGATACACAGGTTCTTCCA
ACCAACGGTCAAACGAGCGGCCAGAGACGACGACCGAGAACGGACCGATCAAAGACGCGGACAATCGAGAAGGGA
GTGAGTGAACGCAAGAAAAGTAGCAGTGGATCGAATTCAGGGTCAAGTTCTGAAGAGAATAAAGCCATATCTGTT
GGAGATACACAGGTTCTTCCAACCAACGGTCAAACGAGCGGCAAGAGACGACGACCGAGAACGGACCGAACAAAG
ACGCGGACAATCGAGAAGGGAGTGAGTGGACGCCAGAAAAGTGGCAGTAGATCGAGTTCAGGGTCAAGTTCTGAA
GAAAAAAATGCCATATCTGTTGGAGCAACGGTAAACAAGAACACAGATGTAACTACTAGGAGACGAACACCAGCT
GCAAAGGTCAAAAGCTCATCGGCTACATTGGTCTCAGGAGATACAGAA

>Sfp8

ATGGCCAGAACATTTGTGTGTGCCATCGTGGTTGCCATCTTTGTGACCATACCCGCCGTTGCGTCTCAACTGTGT
CAATTGTATCAAACGGAGAGGAACATCGGAGTTCAGTGGCTCCCCGAGACTCTTCCAGCAGAGCGATTCGAGATT
ACTTTCAGGGTGTTGGCGCTAGGGGATGTTTACATCTACCTTGGTGATGGTCCTCACGCACGTTTGGCTATTGGA
ATCGGTGTGGATCAAAACTCCGCGACTATCATCAAAGGAAATTCCGGCAAAACTTACGCAAAGACTCCTGGGATA
ATCAAATGTGACACATTCAGCGAGTTTTGGCTGCTATATAGTCCGAGTTTTGTGAAACTGGGTAGAGTTGGTCAG
ACTTCACCAATCGTGACACTCTTTCACTCTGAGAGTTATGTACCGACTAGATTGGGTTTCAGTACTAGTTCGGTC
AGTGTTGGGCTGTTTTCATTTGATCCATTCGGAAAGACGGCACCAACGTTTTTCGAGCAAGAGTTTAACACTTGC
ACAGGATTCTGCTACGACTGGAGATTGAATCCAATGCCTCTCGGCGAGAGGATATGCTTCCAGGCACGCACAGAA
GGAGTGGCTGCTATTGCCTTATCTCCAATGAGAGAACCAAACGCTGGGCAGTCTCTCTATGAAATAGTATTTGGC
TATGACGAAAATAGCATGATCGTGCTCCGCCGCCATATGCACGGAGAAGAGATCGAAGTTGTCGAGGCAAAGGAC
ACCCTAAACTGTGATGAATTCGTGTCTTACTGTATCGAGGCTTCAGGACCTCACATCCAAATCACCCGCCTGGGA
GACCCCAAGATACTGATGGACTTCACGGACTCTGTAGGCATTCCCATCAACTATCGATACATCGGATTCGCCGGA
GGGTTTGATCACTGTGTCGAATGGAGATGGCCCGTTGATATTATTCCAAAACTCATTCTAGAGACAGGTACTCAA
TTGACGGAGGAAACAGGGGAAGCTCCTGCAGTGGATCCAGTGTTACCCGATGCAGGAGAAACTTCCATAGTGGTT
CCAGATTTACCCGCCGCAGAAACTGATGTTCAGCCGCCGTTAGATGAATCCAATAACCCAAGTGGAAGTCAGGGC
TGTGCTCCAGTGCGTTGTAAAAAATTTTGTCCTTTCGGTCGG**GCCACTGATAAGAACGGATGCGAGAAATGTAAA**
TGCAACCGAAAAGGGAAATCTTCTGAAGAGGATCCAGAGATATCAGTAGGAGAAACTGATGTTCAACCGCCGTCA
GATAAACCCGACACTCCAAGGCCAAGAAAAGACCGTCCTAACAACAACGACAAGCCGAAGACAGAAACTGATGTT
CAGCCGCCGTCAGATGAAACCGATACTCCAAGTGGAAATCAGGCCGTGTGTCCCGAAGTCCCTGAAGGTGTTTTC
GGAATCTGTTCGGAGGAGTGTGGAGCCGATAAGAAGTGTGACGATGGTCAACTGTGTTGCTCTAATGGGTGTGGT
CACGTGTGTAAAAAGGGCTGTTCTCCAGTGCGTTGTAGAAGGTATTGTCCTTTTGGTTGGGCCACTGATGAGGAC
GGATGTGACATGTGTAAATGCAAGCGAAAAGGTGGTTCCAAGGGCAAGAAAGAGAACCGGAGGAGAGGGGAATCT
TCTGAAGAGGATCCAGAGATATCAGTAGGAGAAACTGATGTTCAACCGCCGTCAGATAAACCCGACACTCCAAGG
CCAAGAAAAGACCGTCCTAACAACAACAACAAGCCGAAGAAAGAAACTGATGTTCAGCCGCCGTCAGATGAAACC
GATACTCCAAGTGGAAATCAGGCCGTGTGTCCCGAAGTCCCTGAAGGTGTTTTCGGAATCTGTTCGGAGGAGTGT
GGAGCCGATAAGAAGTGTGACGATGGTCAACTGTGTTGCTCTAATGGGTGTGGTCACGTGTGTAAAAAGGGCTGT
TCTCCAGTGCGTTGTAGAAGAAATTGTCCTTTTGGTTGGGCCACTGATGAGGACGGATGTGACATGTGTAAATGC
AAGCGAAAAGGTGGTTCCAAGGGCAAGAATAAGAACCCGAGGAGAGGGGAATCTTCTGAAGAGGATCCAGAGATA
TCAGTAGGAGAAACTGATGTTCAACCGCCGTCAGATAAACCCGACACTCCAAGGCCAAGAAAAGACCGTCCTAAC
AACAACAACAAGCCGAAGAAAGAAACTGATGTTCAGCCGCCGTCAGATGAAACCGATACTCCAAGTGGAAATCAG
GCCGTGTGTCCCGAAGTCCCTGAAGGTGTTTTCGGAATCTGTTCGGAGGAGTGTGGAGCCGATAAGAAGTGTGAC
GATGGTCAACTGTGTTGCTCTAATGGGTGTGGTCACGTGTGTAAAAAGGGCTGTTCTCCAGTGCGTTGTAGAAGG
TATTGTCCTTTTGGTTGGGCCACTGATGAGGACGGATGTGACATGTGTAAATGCAAGCGAAAAGGTGGTTCCAAG

GGCAAGAAAGAGAACCGGAGGAGAGGGGAATCTTCTGAAGAGGATCCAGAGATATCAGTAGGAGAAACTGATGTT
CAACCGCCGTCAGATAAACCCGACACTCCAAGGCCAAGAAAAGACCGTCCTAACAACAACAACAAGCCGAAGAAA
GAAACTGATGTTCAGCCGCCGTCAGATGAAACCGATACTCCAAGTGGAAATCAGGTCGTGTGTCCCGAAGTCCCT
GAAGGTGTTTTCGGAATCTGTTCGGAGGAGTGTGGAGCCGATAAGAAGTGTGACGATGGTCAACTGTGTTGCTCT
AATGGGTGTGGTCACGTGTGTAAAAAGGGCTGTTCTCCAGTGCGTTGTAGAAGGTATTGTCCTTTTGGTTGGGCC
ACTGATGAGGACGGATGTGACATGTGTAAATGCAAGCGAAAAGGTGGTTCCAAGGGCAAGAATAAGAACCCGAGG
AGAGGGGAATCTTCTGAAGAGGATCCAGAGATATCAGTAGGAGAAACTGATGTTCAACCGCCGTCAGATAAACCC
GACACTCCAAGGCCAAGAAAAGACCGTCCTAACAACAACAACAAGCCGAAGAAAGGGGAATCTTCTGAAGAGGAT
CCAGAGATATCAGTAGGAGAAACTGATGTTCAACCTTCGTCAGATGATCCGTCTGTTCCTAGAGGAAACCAAGAT
GGAGATGAAGGGGAGACTTTTACAACTTCAGGAGATTCTTCACAAGAGATCAAGAGGACTTTGGATAAGAAAGTA
ATCAGTGTATTCGAAATATCTCCATTCGACGATAAAGTATCAAAGGAGTTCCAAATCGGTTGTGAAGTCGGCGAT
TTTAAACTGGATGTCAATAGTGAAACCTTTTCACTTAAAAATGGCAAAGGCGAAGAGGTGGATTCGCAGACCACA
CCTGATACTTCAGGTCGTCTCAAGTTCTCAATTGAAATTACCGAGTCAGCTTTCATAATTAGCTACATCTTGGAC
CAAGACTTCACTACCCTTATCGAATACTCACCTCCTGGAAGGGAGGATGGCAAGGAATTGAACTGCAAGGCTTTC
ACCGTTAAAAACGTGGATAAGCGTCGTCAGCGAGGTTGTCAGATTAAATACACCAAGATCAGTCAAACCAGATGG
AGGAGCTACAATAGGAGAAACAGACCTTCAAATACGAGGAAAGTCCAGACGGGAGCAACGTATGTGCCAAAGAAG
ACATCGGGATCTAAACCACGAAGAGTGAGGGAACACAGCCGTGGAAGTGGTAGTGGAAGTGGCAATAGCAAGGGA
TCGAAATCAGGTGGTGGACGCCGTGGACGTGGCAGTGGATCAGGCTCAGGGAAAAAGGATAAAATCTTATCTGCT
GGAGAAACGTTCGTGCCAAAGAAGACTTCGGGATCTCGACCACGACCGACAGTGAGGGAACGTAGCAGTGGAAGT
GGCAGTGGAAGTGGCAATAGCAAGGGATCGAAATCAGGTGGTGGACGCCGTGGACGTGGCAGTGGATCAGGCTCA
GGGAAAAAGGATAAAATCTTATCTGCTGGAGAAACGTTCGTGCCAAAGAAGACTTCGGGATCTCGACCACGACCG
ACAGTGAGGGAACGTAGCAGTGGAAGTGAGTCAAGACGAAAGAAGAAATCATAA

>Sfp9

CTATCGTTTACACTCGAGGTCGTACAGTTCAGACACGGAAAGACTCTTTTCATAGTAGTCGAAGACAGTTACGAT
ACCCTTCCTTGGCTCCTTGACAAATTCTTCTCGAGTCACCATGAAACTGACACAGGTGTCCTCTTTCTCAAGCTG
ATCATAGTAGAAGTTAACTTCCTTTCCTTCGTGTTCGACCATCTTCGGACCTCCTTCCTCTTGTTCCAGTGCTTG
CAGACTGTTTGATTCCGGTTCAAACCCAGACACCATACGCACTTGTATGACTGCCATATTGGAAAATTCATCTTC
ACCATTGTAGTGTGCACAGATGGTCATGCTGTATGAAATACATGGAAACTGTCTAGGATTACGATTATCGACTGA
AGAGATATCAATGGTAAATGGCGGTTCAGCTGACACCTCAGGTATGATGTTATACGTCACTATCACCTGAGTCAA
GGCACAGCCGGAACCTGTAGCAGTCATCTTGATTTCTACTGGAACCTCCGGTATCTGCTCAATCTGTAGAAGAAG
ACTGTTTTCTTGGGTGACGTGAAGTTTTCTTATGTCATCATCGCAAGAGAACTTAATCCCTAATGATATGTCTAG
GTTACCGCCTGAGTACGCCATCATGCTGTAGTCAGCCAAAGCTTGTAAGGCCATCACTGTGTCCTGAGTGGAAGA
GAACCCTCCCTCAGCATTACGTTGCCCGACAATCCAGCGAGCCACGGGGTTGCCCTTGACTAGAGCTTGCGATTT
GTCATAAGACCTGACGTAGGTGAGAAGCACGTAAGATGTAATCTCAATGTTCTGGGACGACGCTTGACGATGGAG
ATATGGGGATCGTTTGTCTGTGTTGAGAACCTCCGCTCCTTCCCAATGCATCAAGCCTTCAGATTCCACTGCAAG
CTCATTCAGTCTTTCCAATAGAGTATCTGTGTGAGGACTCCCAGCCAAGATGTAAGCATACGATGATAAAGCCAA
TGTGTACATGTCTGCCGTTTGGATGTCCACTTCCTTGTTCAGACACTGCACACCTTTTTGGATATGTTCAGCCTC
ACTTGCGTCTCCTGCTTCAATGAGTGATATCATAACATAAGCAGTGAGTGTTGCTTGGTCGTTAACATTCCCCAT
CATATCCTTATGGATGACCTTGCCCACAGATTTAAAACATCCATCCTTCTGTTGAAGACTCAACAACCAAACTTT
ACTCTTTGCCAACTCCTCGGGATCAATAAAGATATACTTGGACGCCTGCGCAAATGACCGTACGACAAACGCCGT
CAGCCACGTACTTCCGGACGGGTCACTCATTCCAAAAGCTGAGTAGGAACCATCGGTGTGTTGGTACGTTAGCTC
TCGCTGGTAGCCAATCTCCATGTGTTGCTTTGCTTTAGCTTCTACGCCTGGTGTGAGCTTGTTGATCCCTGTAAG
ATAGTGCAAGACAAAGATGTTGGGAACGAAACCGATCATATTTTGTTCACCGCATCCAGTTGGCATACGCAGAAG
CTGATCCAGGTTGCTGATTGAAGGACCCATCATATCACCTGTGAGAGTGATGATTCCTCTTTCAGAGCCAGACAC
AAAGTTCTCAGGTAGATCTTTCAGAATAATATCCTCAATATAGACGCCGTTCTCGTAGTCGACAGGACAGAAGAA
TGTGCTGAAAACTTGTTCAGTTTCTATACCCTCAGGCTTGACTAGTAACTTCTTTAAGACTGTATCAGAAACTCC
AAGCAGAGAGTCATCCATTTCTCGTCCTATACAAGCACCAGATGCCTCTACGGAAACAGCTGCCACTTGAATGGG
AATATCTCCAAGTTCGATAGCCTTGACGTAGAATGATACCGATTGAGACTCGCCAGCACATACACAAAGCTTCTC
TACTCGATTCTGATCGTCAATCTCGTAGCTACTTGACGCTTCTAAAGTTAGCTCAACCATTAGACATTCGGTTAT
GTAGTTGAATACTACAACAGTGACTGGCACAAGCTCATAGCGAATAACGGAGTAAGGTAATCTCATATCAATGAA
GAACGCTTGGAAGGCTGTGATGCTGGTCAAATCAGACACTCCAGCCCCGTACTTGGTGGAGGTACAGAATCCATG
GCCAATCCACTTCGTGATGGTGTGTGGAACATGTACATTGATGTCCTTGTGGCCGCTATCGTCAATTTCAACAAG
TTGATACAACCAAGTCTCTGGAAAGTAGGATCTGATAGTCGGCCCAGCAGCATCACTTTCAGCTCCTTTCATAAC
AGCCCCTTCTATTGCAGAGTCCTCCATCACCGCCATCATCATCATTGGCATCGCCATCATCATGTCACCACCATC

ATTTTGAACTTCCCCTCTAAAATACATGTACGGGTTAACCTCAGGACAGGGTGCAGTGTCTACTTTGAGGTTGGT
CAGATACACAACACCTAAATCTTTAAAAGCCTGAGAGGCGTCGTTAAGATTTTGGGGTGGCATATAACCTCTAAA
CATTCCCCCTCCGCTGCTACATCTTGAAGGTTCCGCAGAGCCACCCTCCGATGAGAGTTGTAGCCCCTCAAGTGC
CTCAAACACCTTTTTTTTTGTAAGTCGGTTGTCACCGCCCAAAAGATGGACACTCTTATCAACGACACCGAACGC
ACATAAAGAGGATGCTTTGGCTCGGACCGACAAGACAGCACTATCTCCAGGTGCTGCTGTCTTTTCATCGAACGA
AACATCCACCTCATTCTGAAACACTTCCTCCACGGGAATGGTAATGGTATCCGCTACGACCTCTCCATCTTCCCT
TATATAATAAACCATGACAGAAACAGCGGGTGCCATCTCATATAAGATTCGGATCTTGCGTCTCGTCGTCTTGGA
CTGCCCCTTGTTCAGAATGTTGACTGGGATGTCTGTTACTTCCATGTTGGTCCGAGTAGGATACTCTTTCCTGCC
ATATTCATCAGTATCTACGGGGTAAACGCCTTGGGTTCCTGGGCGTCTGAAGCCATCAATGTCCTCTGGTGGAAT
AATGGACTCAACAGAAAGCAGAATATTTCCACGAGACATCATCACAGTGTGCAAGTAGACTCCATTATATTCTGG
TTTATCACTAGTCAGATGAATGGTGAGAGACTGGAAACTGTTGACTCCAGCCGTCTCCAGAACGGGATCAATGTG
AAGGAAGCTGTTGGACGGTGAGATTTGCGGCTGTGCAACATCGTAGGCAGTGGGGTCTTGAATGCCGTGAGTTCC
ACCATTCTTTGGATGCCATGTTGACTGGCGGAAATAGCCTTTTGAAACAGCCTTAAGAGTGACGCTCTTTGAACC
GCTAGGAATTGTCGTGACTGAGAAGAAAGCCATTCCTTTGGCATCACTGATGACCGTCTCCTTAAAGATACTCAC
ACCGTCTGCTTCTATATTGATTTCTGATTTAGTATCAGCGAGAGGTTTATCACTAGGGTCGGTGAACAAAAGCGC
TCCAGAGTATGAAAGTCCTGGCTTGAACGTTGAAGGTGCAGAGAACTCCAGATGAAAAGGCTTATTAACAATCTT
AGTTGAGTCACTCGTTTTCAACATTATATGATTTGTTGAAGCCTCCTCAAACGTTGCTTCTAGATGTAGTTTAGC
GTTCCAGACATTATAACTACTTGACATCAAGTCCAGTTTAGTCATGTCTACCTCAAACGTGAAACATCCGGCTTT
GGCCTCAGCTGATGTTTCCACACTCTTGTCTGGGTAGTTTAGTACGCCAACTTTCAAAATAATGGAACCAACCAC
CGGCTGGCCATGGGTATACCTTCCGCAGACTACTACACTAATCTGCTCAACATCCACCACAACGTAGTCAGGTCC
TTGTACCATCACTTCATATTTGGGTAAAACATACTCCCCGACCGTAAATTTCTGACTGAAATCTCTGTTATGATG
AAGAACTTTAATCGTCCACTCTCCTTGAGTGGGATTGGTTGATAGTTTCATCGTCTCATCGATCAGTCCCTGGTT
TGTTTCGACACCTAACCACTGGGCCATATGAACGCCTCCTGGTGCCTCAATCCAAATCTTATCCACCTCACTTAT
GTCGGGTCGAAGGTCAGGATGCAAGGTCAGAACTCGGAAATGTACATTTTGGCCTGCCTTGTACATCGGCTTGTC
TGTCTCGATAAAGGTCTCAGTGTAGCAAGCATCTAGGTATACCTTGATCTTCTTGACGATTTCAAACTCCTCGCA
TTGCATGAGGGTTCTCTTGGCATACATGACCAGGGTGACAGTCTTCCGGGGTGAATCACTGGGATTAAGATGAAC
ATCCAAGCAGCCACATGGAGTACTGAAAGTCTCAAGTTTAGAAGTGCCAACAGATTCATCATTCTTGTCCAAAAC
AGCAAATTGGACGTCAACTGGCTGTGTGAAGTTGGTAAAGGTGAGGAAGACTTTCTCGTGACCGCCGGCACACAT
GACTTTGGGAGCAATGGCCAAATAACCTGCTCCAAAAGCTCGGTTTCTTATTTCTGCCTGTTTCTCCAACGCAAT
TTTCTCCTCAGCTTCCTCGTTGTCGATGTCAACAGGTGGAACGGGAGCAGGAGGCATAGGCGCAGGCATAATGAG
TGGTTCATCAACAACGGGCTGAATCTTCAAGTTCTGCCCGGTCGCATACGGCAAAACATAGATGACACAGACAAC
AACAATCCAAGCGCGGGTTCCACTTTTGCTGAGCCCCAT

>Sfp10

TATAAGAATCCATGTTTAATTTGTTTTGCAGTTATTACTACTACGACCCAAACGTCAACCTCTACTACAGCAGGA
GAATGTATGGAACCATTAAATGGCCTGCTGGATCTAGAAGAAAGACGTTTCATATCACCACAGCAAAATCAATCA
ACCGGTATTAATGCTGACAATCCCTGGATATTCGAACTACTGCCATATGTTGATGCAGAGAACAAACCTACGGTG
ACTGTTGAGTTCAAGTCCAAGGTGGAAGTGCACGGATTTA**AGCTGCAAGGAAGTGGTGATATTGTCAAGGAATTG**
CAATTTGTTGTGTCGGCTAAGTACGAAGTTGACGGACTGTTTGTAGAGCTGGAAGATGAAAAACTAGTTGCAGAT
CTGACAAACCAGAGAGTCGAAACCGTCGCACTACCTTCTCCTTTATCTGGTTTGCTGGCTCTTAGCTTGCAACTT
GAGTCTGTAGATCAAGAAGATGCCGGATTTGGTAGTTTGCGCTTTGATGTTCTTGGCTGCGAAGAGGAAATCTCC
ACTACTACACCGGTTGTCACAACTACTAAAGTGGTAACAACTACAACCCAAACATCAACCTCGACTACTGCAGGT
TACTGCATGGATCCCTTGAGTGAGCTGGATAATCTTGAACAGCGTTCATTCGCACCACAGCAAACACTAGAAAAT
CAATCAACACTTATCAACTCTGAAAATCCGTGGACCTTCCAGTTGTCGTCAGATGCTGGTGAGGAATATGACAAT
CCATGGGTGATTGTTCAGTACCAATCTAAAGTTGAAGTGCGTGGTTTTAAACTACAAGGGAGTGGTGAGGATGTC
ATTACCCTGGAATTCATTCTCTCAGCCAAATCTGAAGATAATGAACAGTTTGTTGATGTGCTGGAAGGTGAGAAA
CTCGGAGCAACATTGTACCAAGACAACTTAAAGTCCATCACTCTTGACTCTCCCTTGCCTGGAGTAGTTGCAGTT
AAATTACAAGTTATCAAAAAGCAGAATGGCGGAACTGGTAGTTTGCGAATTGAGGTCCTCGGCTGCGAGAAGGAG
ATTTCAACTACCACTCCAGTGACCACAACGACTAAAGTTGTAACCACCACGTCTCAGTCTCCAACTTCTACAACT
GCAGGTTACTGTATGGAGCAGCTTGACGGAACGCAGGGACTGGAAGCCAAGCGATCAGTACAACCGGCTATTGCC
AACGATCAGTCAACTAGTGTCAACGCACAGCAACCTTGGACGTTCTCCCTTGATAACCAGCCATCTTTGAAAGAT
CGAGAAACTGTCAAAAACCCCAGTGTAACCATAGATTTCAAGACGGAAGTCACAGTTACTGGAGTAAGGCTTCAG
GGCAGTACAAACAACGACAACATTGAATTCATCTTGTTAGCTAAACGCAGACTGGGTGAAACGTTCGTTATAGTA
CAGGGAGCCTCTGGTAAAGCATTGGTGCAGAATATGGAGCTTAAAGAGCAAGCTTCAGATGTAACATTGATAAAG
CCGCTGAAAGGGTTAATAGGATTGCGCTTAGAAGTAGTTTCCCTCTCTCGAATGGATGAAAGTCAATCCACTGAC

AGCAAACTTCAGTTCAGTTTACGTGCTGCTATCCTTGGTTGTGAAGAAGAACTGTCAACTACTACTCTTGCGACA
ACGACAACAGAAGCGCCACTTGTTTCAGGGTTCAATCCTGTCATCACAACAATTCGACAATCAAGTAGATCGACA
ACACCCACAACCGTGGTTACTACTACAACTGAAATGCCAACCTCCACCACAGCAGGTTACTGTATGGAACCCATG
AGTGGCGAGCAAGATCTTGAAGCGGAGCGTTCATTCACACCGCAGCAGATTCCAAATCAATCAACTAATATCAAC
GCTGAATATCCGTGGACCTTCGAGCTGTCAGCAGTTACCACCCCTCAGTTGACTGTTTATTTCACTTCTAAAGTT
GAAGTGCGCGGATTTAAGCTACAAGGCAGTGGTGAGCTAGTCAAGGAATTGAAATTCATCGTCCTGGCCAAATTT
GAAGACAATGGAGACTTTGTTGAGGTGCCACAAGATCAGAAAATCACAGCAACCTTGTCCGGAGATGGACTAGAG
TCCGTCAATCTGGACACTCCATTAGAACGAGTAGTTGCAGTTCGATTACAAGTTCAATCAATTGATGACTTAAAA
GCTGGAAAAGGCAGTTTGCGACTCGATGTGCTAGGATGCAAAGAAGGTAGGTCGAAATCCTTAGTGTAA

>Sfp11

ATGGAGTTCTTCATCCTTGTGTTACTGTTCAGCCTCTTGACAACCAAGGGAGCAGTGGCTCAAGATGATTTGGAC
TGGGACGGTTTTGGAGAGGCTGATGGCAGTCTGTATGAACTCAAACCATTGTCCATCATCGGATTCCATACGGAA
CAAAGTTCTGACTATGATAGCAACCTATATCCATCCTCTAAAGGTGTAGACGGAAACGTAGATCATGGCTTCTTC
CATACACAGAATGATGAGCATTCGTTCTGGACAGTTCAACTCAAGGAGGAACACTGTATTGGAAAGGTCACCCTT
TTCAATCGTAAAGACTGTTGCGGTGCACGTTTAGCGAACGCCGTTGTCCGAGCTGGGGTCAGTCCATTCTATAGT
CTGAATCAACGATGCGGGTCACCAGTGGAATCCCACTTGGGCAATGGGGGTGTCATTGATGTTTACTGTGACCCG
CCGTTGTTGGCGAGTGAGGTCAACGTTGTTCTACCAAACAGGGAATACCTACATCTTCATGAAGTTCAGGTTGAG
GAGTATCCTATTCAACGATGCCACCAAATAGAAAGGAGAATCTCCATTCT**GAATTCCCCGACCGAACAAAGCTCC**
GGATACGATGATACCAAATACCCAGCGGAGAGGGCCGCCGATGGTTTTATTGGCAAGTTTGTAGGTTTCAGTCAC
ACCGGTCATGACTTTGATCCATGGTGGAGAGTAGATCTTCTAGACACGCACTGCATCAGTAAAGTGGTTTTGTTC
AACCGTTATGATTGCTGCAGCGAACGTTTATCGAATGCTGTTGTGAGAGCAGGAACAAGTAAAGACATACACACA
AATACCCGATGCGGATCAGCAATCACCGCTGAACAAGCCAAGATTAAAGGCGGAGTTCTGGCCGTCTACTGTGGA
CCACCGTTATCGGGGCGTTACGTCAGTGTTAACATACCAGGCAGAAACGAATATTTGCAACTACGTGAAGTAGAG
ATCTACGAACTTGACGTCAACGAGTGTGAAAGTGACGAACGCCACTTGTCTATAATCGGTAAGCCGACGGAGCAA
AGTTCTTACTATGACAACAGGTACACGCCAGATAAAGCTGCAGACGGAAACCTGGACACTAAGCTGTCACCAGAG
TTTATCTGTACACACACAAAGGAAGAGGATGATCCATGGTGGATAGTAGACCTTGAAGATGAACACTGCATCAGT
AAAGTTGTCCTTCTCAATCGTGGGGATTGTTGTAGTGAGCGTTTGACGAATGCTGTTGTTCGCGCGGGTACCAGT
CACAACATTCAAAGCAATCCCCAGTGCGGAGATGTAGTCAGTCGTGCCATGGCAGCTCCACGCGGAGGCACAATT
GAGTTGGTATGTGACCCACCGTTGGTGGCGCGTTACATTAGCGTCGACATACCCTCTGATGAGCATTCGTTTCTG
CAGCTGTGTGAGGTGACAGTAAAGGAGTTCATCGGAGAAGACGCGTGTCGAGAACCAGAACCGATAGTCATAAAA
GCACCAGAACCGCAACAACCTGTCGAAGAACCAAAAGAAGCTGTCGAAGAGGATAACCAAAAGAAGGTTCTTGGT
GGTGGGGTCAAGATTCAACAGGATGTTCCAGATGACATGGGTCCCTAA

>Sfp12a

TGTGCTGAAGAGTTTTCAACAACAGAATATGTGACAACCAAATCCCCTACAACT**AAAGGTTTGCCAACTGGAGAC**
ACTTTTCTATCGTCTACTACCATCAACTCCCGTCACCCTCCAACAACGACCGCTACAACAACCAAGACCAGCAGC
ACTACAACTCATCAGCTATCAACAACTCAAGGTTACTGTTTGGAGAATATGGATGGTGAAGAAGGCCTTCAGGAA
GGTTTTACACCTACCTCAGATCTAATTACTTTACAAGTAGAAAATGGAACAACTTACATCAAGTTCGAAAGCTTG
CCTTTGGAGGTGGGAGACAACAACCTGCCCAGTGTTGATGGCATTTTAACAAGATCTGCCTCTATCAGAGCAATC
AGTATTAGACTTGTTCCATCATCATCTTCCTCGACACCATCTCCGACTGGAAATGAGGAACTTGTGTCTGTGACT
ATAAATTTCTTCACCAAGAAGGAGAGAAACGATGACTTTAAGCCCCTGTTGCTGGAAGACAACCAAGAACCCAAG
GTGATCACATTCACACCAGTCAGTGATCAACCCTTGCAGAAGTTCCTTAAAATTGGAAATGAAGTCATTGAAGGG
GTTGTGGCATTCCGTGTGACTGTCCTAGAGGTCACTGGTGTTTCGTCTGCAACTTTAGAACTACATGTATTGGGC
TGTGCAGAAGAATTCACCACAACCCAGGTAACAACTACAGCAGTCTCTGTGACAACACCAACAACGACAACCGAA
GAATCAACAAGCACTGTGGCAGTTACAACACCAACTCAGACAACCACAGAACGAACAACAACCTCAACAACATCT
CTGGTTTCAACCACAGAGGGGTACTGTCTTCAAAATATGGATAGCGAGAGCGGTGTAGAGACAAAGGACCAGACA
AGAGGAAATCTCTTGACAGATGAGTCAGGCTTTCTCACATTTTCAGAGTTACCAATTGAACAGTCTGTTGGCAAC
AGTGCTGTTCTGGAGAGAAAGTTCAATAGCAATATCAATGTTCGTGGAGTTCTAATCACTTTGAAGAAAATCACA
TCAAGTGTTACTTCTGGACAAGCAGAAGGAATGACCACACCAGAGACAGGATCAACGTCACCGCAGTCAGTCAAA
TTCACTTTCACATTACTTGCAAAGAGAGTAGACGAGACAGATTTCAAGCAAGTGGTGATCAAAAATACAATCACG
GAGTTTACCTTGACTATTGAGGGTATTCAATCGGAAGTGAAGGTATTTACAGAGGACGAACCATCACTACAAGGT
ATAACGGAACTTCAGATCAAGTTCATTGATGTCTCAGGTGTCAGCGGAGCAA

>Sfp12b

C**AAGGTATAACGGAACTTCAGATCAAGTTCATTGATGTCTCAGGTGTCAGCGGAGCAAGTTTTGAGTTAACTGTC**
TTGGGATGTGCTGAAGAAGTCTCAACAACTACGCAAACTACAACTAGCAGTACTGTTAGCACCACAACTCCTCAG
CTATCAACAACTCCAGGTTACTGTTTGGAGAATATGGATGGTGAAGAAGGCCTTCAGGAAGGTTTTACACCTACC
TCAGATCTAATTACTTTACAAGTAGAAAATGGAACAACTTACATCAAGTTCGAAAGCTTGCCTTTGGAGGTGGGA
GACAACAACCTGCCCAGTGTTGATGGCATTTTAACAAGATCTGCCTCTATCAGAGCAATCAGTATTAGACTTGTT
CCATCATCATCTTCCTCGACACCATCTCCGACTGGAAATGAGGAACTTGTGTCTGTGACTATAAATTTCTTCACC
AAGAAGGAGAGAAACGATGACTTTAAGCCCCTGTTGCTGGAAGACAACCAAGAACCCAAGGTGATCACATTCACA
CCAGTCAGTGATCAACCCTTGCAGAAGTTCCTTAAAATTGGAAATGAAGTCATTGAAGGGGTTGTGGCATTCCGT
GTGACTGTCCTAGAGGTCACTGGTGTTTCGTCTGCAACTTTAGAACTACATGTATTGGGCTGTGCAGAAGAATTC
ACCACAACCCAGGTAACAACTACAGCAGTCTCTGTGACAACACCAACAACGACAACCGAAGAATCAACAAGCACT
GTGGCAGTTACAACACCAACTCAGACAACCACAGAACGAACAACAACCTCAACAACATCTCTGGTTTCAACCACA
GAGGGGTACTGTCTTCAAAATATGGATAGCGAGAGCGGTGTAGAGACAAAGGACCAGACAAGAGGAAATCTCTTG
ACAGATGAGTCAGGCTTTCTCACATTTTCAGAGTTACCAATTGAACAGTCTGTTGGCAACAGTGCTGTTCTGGAG
AGAAAGTTCAATAGCAATATCAATGTTCGTGGAGTTCTAATCACTTTGAAGAAAATCACATCAAGTGTTACTTCT
GGACAAGCAGAAGGAATGACCACACCAGAGACAGGATCAACGTCACCGCAGTCAGTCAAATTCACTTTCACATTA
CTTGCAAAGAGAGTAGACGAGACAGATTTCAAGCAAGTGGTGATCAAAAATACAATCACGGAGTTTACCTTGACT
ATTGAGGGTATTCAATCGGAAGTGAAGGTATTTACAGAGGACGAACCATCACTACAAGGTATAACGGAACTTCAG
ATCAAGTTCATTGATGTCTCAGGTGTCAGCGGAGCAA

>Sfp12c

CCAACAACGACAACCGAAGAATCAACAAGCACTGTGGCAGTTACAACACCAACTCAGACAACCACAGAACGAACA
ACAACCTCAACAACATCTCTGGTTTCAACCACAGAGGGGTACTGTCTTCAAAATATGGATAGCGAGAGCGGTGTA
GAGACAAAGGACCAGACAAGAGGAAATCTCTTGACAGATGAGTCAGGCTTTCTCACATTTTCAGAGTTACCAATT
GAACAGTCTGT**TGGCAACAGTGCTGTTCTGGAGAGAAAGTTCAATAGCAATATCAATGTTCGTGGAGTTCTAATC**
ACTTTGAAGAAAATCACATCAAGTGTTACTTCTGGACAAGCAGAAGGAATGACCACACCAGAGACAGGATCAACG
TCACCGCAGTCAGTCAAATTCACTTTCACATTACTTGCAAAGAGAGTAGACGAGACAGATTTCAAGCAAGTGGTG
ATCAAAAATACAATCACGGAGTTTACCTTGACTATTGAGGGTATTCAATCGGAAGTGAAGGTATTTACAGAGGAC
GAACCATCACTACAAGGTATAACGGAACTTCAGATCAAGTTCATTGATGTCTCAGGTGTCAGCGGAGCAAGTTTT
GAGTTAACTGTCTTGGGATGTGCTGAAGAAGTCTCAACAACTACGCAAACTACAACTAGCAGTACTGTTAGCACC
ACAACTCCTCAGCTATCAACAACTCCAGGTTACTGTTTGGAGAATATGGATGGTGAAGAAGGCCTTCAGGAAGGT
TTTACACCTACCTCAGATCTAATTACTTTACAAGTAGAAAATGGAACAACTTACATCAAGTTCGAAAGCTTGCCT
TTGGAGGTGGGAGACAACAACCTGCCCAGTGTTGATGGCATTTTAACAAGATCTGCCTCTATCAGAGCAATCAGT
ATTAGACTTGTTCCATCATCATCTTCCTCGACACCATCTCCGACTGGAAATGAGGAACTTGTGTCTGTGACTATA
AATTTCTTCACCAAGAAGGAGAGAAACGATGACTTTAAGCCCCTGTTGCTGGAAGACAACCAAGAACCCAAGGTG
ATCACATTCACACCAGTCAGTGATCAACCCTTGCAGAAGTTCCTTAAAATTGGAAATGAAGTCATTGAAGGGGTT
GTGGCATTCCGTGTGACTGTCCTAGAGGTCACTGGTGTTTCGTCTGCAACTTTAGAACTACATGTATTGGGCTGT
GCAGAAGAATTCACCACAACCCAGGTAACAACTACAGCAGTCTCTGTGACAACACCAACAACGACAACCGAAGAA
TC

>Sfp13

ATGGAAGAACAAGTTGTGTGCATGCTTGAATTCAAGACCTCAGCCGAGGTCTCATGCGCTCCAGTTCTCGTTACC
AAACTAAATGCACAGGGATCAATAAAGGTGTCAACAGGTGATATTTCACAGTCTCAAAAGTTACCGTTTGAGATG
AAGCTGAAGGATGGATCAACCCTCATTGCTCGTCAGGTGTCCTCGATATTCTCTCGTATGACAATAGCCGATCTT
GGGATTGACATTCAATTGGGACTTGACAAGCGTATCTATGTCACGCTGAATCCAAGCTTGCAAGGAAAGGTAAGT
GGGCTGTGCGGCAATTACAACAACAAACAAATAGACGACTTCACCACTTCCAGTGGTGTGAACGCTAAAGCCAAC
ATCAACAAGTTCGGTAACTCCTGGAAAGTCAGCCTTCACTGTGGTGACGTAGAGGAGAAGAACCTACGACCTTGC
GACCTGAAATCCCAGCAACGACCCCTGGCCGAGGAGATTTGCCGTAAACTCCGAGAAGAGCCTTTTGCTAGCTGT
AACCAAGAGATCCCAGTTGAT**AAGTACGTTGAGCAATGCGAGCAGGATGTATGTGGCTGCCCTCCTAGCGACAAG**
GAATGCCAGTGCCTAGCCTTTGCTACCTATGCTCGGGAGTGTTCTGTGGCTGGGTTTGACTTCTTGTGGAGAGGC

CTCCATAACTGTCCCATAAAGTGTCCAGTTGGCCAGGTTTACAAGATGTGTGGTTCAAATTGTCACGTGAATTGC
AAAGATTTCTTGACACAACAACCTTGCCACGAGGAGTGTGTAGAAGGCTGCCAATGTCCAGAAGATAAAGTTCGA
GATGACGAATCAGGACAGTGTGTAAAAGTTGATTACTGCCCATGCATAGTTGACGGTATCGTGTACCAATCTCAT
CAGTCTTGGAAGAAAGGATGCAATGACTGTTTTTGTTCTTCGGGGACACACACCTGCACAGAGCTCGACTGTTCA
ACTTATGAGACTTGTGCCGATGGTATGGAATGGATTCCCTGTCTGCAATGTGAGAGAACCTGCCAAAGTGCCAAC
AAGGCATGTAGTATCCGAAGCGGCTGTACCCCTGGGTGTGGATGTCCGAAAGGTCTAGTCTTACACAAAGGATCG
TGCATAGATGTCTCGGATTGTACTTGCGAGTACAATGATAAAAGCTACAAGCCTTTCGACACCACTGCAATGGAT
TGCCATACGTGCATGTGTCACACTGATTACCGATGGATCTGTGACGAAGACCAATGTCCATCTACTTGCCGGTCA
TATGGAGAGTCGCATTTCCAAACACTGGACGCAAAATGGTATTCATTTCAAGGCAACTGCGAATACACCTTGGTG
GAAAATTTCTGCGGTGGGTCCAACGTTGAGGGGTTCTTCCGAGTCACGATCGAAAACGTTCCTTGTGGTAGAAAT
GGGGTAACTTGCACCAAGGCAGTGAAGTTCACACTTCATGACACCACAATCCATCTGGTTAGAGGAGCAGAGTAT
ACTGTAGCCAAGGATCCCAGAGTAACCACAAAAGCCCGATTTAAGATGGAAGATGCTGGGCTCTTTCTTGTCATC
AAGACAGCTGAAGGAATCCTGTTGAAATGGGACTATGGCACAAGTGTGTTAGTGACCCTGGATCCATCGCACGCA
GGACAAGTGTGTGGTCTTTGCGGTAACTTCAATGGGGACTCATCTGATGATTTCTACACAAGATCCGGTGCAATG
GAATCTTCGCCACAACTCTTTGGTAGAAGTTGGAGAACCTCAGAAAACTGCCCAGATCCTAAAGAACTGACCTTG
CCGTGCAAATCGGAGCCTCATCGAAGAGCTTGGGCTGAGTCTGTTTGCAGCATAATCAAGAAGTCCATCTTTGCT
TCATGTCACAACCTTGTTGACTCTGAACCTTTCTTCCAAGCGTGTGTCTCAGACACCTGTTTCTGCGACAGAGGG
GCTGACTGTGAGTGCATGTGCACTGCCGTCACAGCATACGCAACCGCTTGCAATGAAAACAACGTTGCAATCGCG
TGGAGGGCAGAGGGATACTGTCCGTTGCAATGCGAGAGCGGAAGGGTATACCAGGCATGTGGGAAGATTTGCCCC
GACAAATGTTACCCTTCACATGCTGAACAGAACTACGGATGCGAAGAAAGGTGTGTGGAGGGATGCCACTGCCCA
GAAGGAACAGTTCTGTTCCGGGACGAGTGTATCGAGAGAAAGCTTTGTCCGTGTTTCATCAACAATGGGCAAGAG
GAGATCAGACACGGCGCACTTGTTGTTGAAGACTGCATGCAGTGTATGTGTAATGGTGGTCATTTAGAATGCACA
GGAAGAAACTGCACAGAGTCAGTTGTTGCTGAAACGACCACACCTTGTGCACATAAGACTAAACAATGCTCGTGT
GCTCCTAACTTCTCTTGCTGTGCGGATAAGAGTGTGTGCATTCATGACAGTAAGTTCTGCGATGGAGAATATAAC
TGCAAAGACTTCTCGGACGAAGTTCACTGCAACTGCATTTACAATGACATCGTTTATCTTACTGGTGATAATTGG
CCAGGTGTCGGAGCTTGTGAGGAATGCAAGTGTGAAGCCGGAAAAACAACTTGTTACAAGAAATGTCAGCTGACA
TGCGAGTTCCCTGAAATATTGAACTTATACGAGGATGACGACAGGAAGTGTTGTAGTTGTGAAGGCGGAGAAGAT
ACTGTACGCAGAACAACAACTCCACCACGAACACGACCACCACGATCACCACCAACCACAACACAATCAACAACC
CAAGAATGCGTTGATCCATTTGTTGTTGGAGAACCAACTTGCTGTCAGTACCTTGATGTCTGCGCCATAAACACC
AAGTTCTCTGGATCAGAGAACTGCAATGCTTGTAGAGAGCCACTTGTCTTTAATGGAACGGACTGCGTTCGCCCG
AATGAATGTTTGTGCAAAGACAAAGTCGGCTACCACCTGCCCGGAGATACTTGGGAAAATACAGAGAAAGCCTGC
AGCGTATGCACATGTTTCAATAATAACATTGAATGCAGTGATGTAATGTGTGATGACACCATGTTAGACTGCTCC
AAGTATGGAGAAGTATTTGCCTATGGTGCTTGTGAAGAGTGCACTTGTCTGGACGAGATAGTTTGTGTTAAGAGT
GAATGCAAACTAACACTAGAGGACTGTGAAGAGGGTGAAGTCGTATTTCCAATCGGAGACTCGTGTTGTGAATGC
AGAGCAATTGAAACAACGACGATTGTTCCAACCACAACCCTTCCAGCCTGCGTTGCTCCATTTGTCTACCGAGAG
AGTTGTGAACGCCACGTAATGTGCGAGACATTGGGTGATGATACCAATGAAGTAAATTGTGAGGAAGGGTGCTTC
TGTCCCGAAGGAAAACTGATGAGCGAGAAAGAATGCATTCCTTCACACCAGTGTCCATGCACAGATATATTTGGA
ATCAAAAGAGAGTTTGAGGAGATGTGGGCTCAAAGCGATTGCCAACAGTGTGTTTGTTACGAAAACATGACCATC
CACTGCACGAAGGAATGTGATCTTAAGTGCGATGAGAGCGAAGAAGTGCTTGTTAATCCGGATGGAAGTATGTGC
TGTTACTGTGAACCAATACCTGAGACAACAACGATGTTTATCCCTGAAACAACAACAGCAACAACAACTCCATCT
GGAGAGTGCTTCGGATGTTACGACGACAAATCAGACACCTGCCATGCGGTCGACACATCATGGAAGGATGATGTG
TGTACATCATGTGAATGCAAGTATAAAAACGAATTCATCATGAATTGTGAACACATGCAATGCGGAAAGGACCTG
GAATGCAAAGAAGGACATGTTTTAGTGAAGGACGATCAAGAATGTTGTTACCACTGTGTACCAGACATTTGCGTA
GGAGACTTCAACTGCAACTCAACATTCTGTATCCCCAATAAATGGGTGTGTGACGGAAATATGGACTGTGCAAAC
GAGTATGATGAAGAAAATTGCCCCTTGCCTCCCACACTGCCACCACCCTCTCCAGCGTGTGCCAAAACAGAGGAA
AGAGCTTTCTGCAAAGACTGTCGTGAAAGAGTTTGCCACGACATGGAGTGCTTCGAAATGCCACAGGAGTTGTGT
TGTGAGTGCAAGCAAGGCACCGTTTACAATGGCACACAATGTGTTCATCCAAAGGAATGCCAGTGTCTGGATGAA
AATGGCAACATTCGTTTGCCATTTGACAAATGGGAAACTGCTGACTGTCAGGAGTGTAAATGTCACTACAACAGA
ATAACTTGTTATGAAACGGGTCTTTGTCACTCAACGACGCCCTCAACATCTACATTGCCTCCTAGACCCTATTAT
TGCACTGAAACATGCACATGCAACCTTGACTGTGAGGGTGACTTTATTGACGAAAAGTGCCTATTCGCTGTTCCC
CAGGAGTGTAATGACACATGTTTATGCCTACCACCAATGAAAAAAGATGATGTAGCTGGTAGATCATGCATTTAC
CAGCCATCGCAAGAGGAATGCATCAAGGCAACTACAGATGCGACTACAACCAGGGAGCTCCACCCCATATCATTT
ATAGAAAAACAAATAGAAATAGGCAAGGAGAAACAAGGCAAGCGACTACAACCAGGGAGCTCCACCCCATATCAT
TTATAG

>Sfp14

TCTTGATAACTGGAGGTGTACAAATCAAAC**TGGAAGGTGAACAACGAGTGTTCCTGGAGTATGTTACGCAGGGTA**
GGAAGTCCAAATGCACAAGAGTTAGAGAGACATATACAAGAACCAAATCAACGGTACAACAGGGATCAAGGCGCC
GTGTTTACAGTCGACAAAGAGGCGGTGAAGGTGGCAGTGGAAGTAGGGGTGGATCAAATTCAGGGTCAAGTTCTG
AAGAGAATAAAGCCATATCTGTTGGAGCAACGGTAAACAAGAACACAGATGTATCTACTCGAAGACGAACACCAG
CTGCAACGGTTCAAAGTTCATCAGCTACATTGGTCCCAGGAGAAATGCAACAGGTGGATTGGACTCAGAAAAAGT
ATGAAGTCAACTTTGTTGATAAAATCGAGGATATCCTTGACTTGACGTTTCAAGATGGTAGCGTTTTGACAATTG
ATGTCACCAGAACAAGAATGACCATCAAAGATCAGAGTGGTAAAGAACTGAAATCGGAATCGCTACGAGGAAAAG
TCCGGGACAACCGTTTGGTATTCACAATCTACTTGATAACTGGAGGTGTAAAAATCAAACTGGAAGGTGGACAAC
AAACTTTCCTGGAGTATGTTAAGGCGGGAAGAACGTCCAAGTGCAACAGAGTGAGAGAGTCGTACAAATCATCAT
ACAGAAGATACCTTTAA

>Sfp15

TCACTGAGCTGGTCGCTTGCATTGACATATCTCA**CATCCGTCCTCATCAGTGGCCCAACCATTCCGACATCCCAT**
GCGGCAAAGCACTGGAGCACAGCCTGCTTTACACACGTGCCCACATCCATTAGGGCAGCACAGCTGACCATCATC
ACACTTCTTATTGGGTCCACACTCCTCAACACAGATTCCGAAAACACCTTCAGGGGCTTTGGGACACACGGCAGG
AGGGCCTCTGCAGCTACATGTTTCACATCCATTCTCATCAGTAACACGGCCATAGAGACATCCGGTAAGCCCACA
GATAAGAGGTGGACAGGCTGCTTTACATACATGACCACACCCATTAGAGCAACACAGCTGACCATCATCACACTT
CTTATCGGGTCCACACTCCTCAACACAAATACCACCAACATCTTCAGGGACTACGGGACACATTTTCATGAAGGG
AGGCGACCTACAAAGACATTCTGCGATTACCAGGAGCACAACGAAGCCTCTTAGAACCCACAT

>Astacin-like Sfp

ATGAAGACCTTGCTGATCACCCTGTTGCTGGTTGGGTTCACCCTCGCATTTGTCAAAAATCGTGGGATGCGGGAT
GACTTCAACAGCAAAGCGGTGGAATTCCTGAGCGGGCGTGACGGAGAGGCTGCCCGGGCAGAGATGACGCCCTCC
CCAGGGTGTGTAGAGGGTGACATGGAGCTCACATCGGAACAGAAAGATGTCTACTTCGCCATGAGGGAAGCCACA
GCTAAGAACAGAGCTCGGGTACGGCGTGACGCCCAAGCCGCATTCTCTATCTGGCCGAACAAGACCCTCGTGTAT
CGGTTCGACGCATCACTTGACAACGGCATCAAGCAAGTATTTAGGGAGGCGATTGAAATCTACAGAGAGCATACG
GGTCTATCATTTGAGGAGGCGTTTGATGACTACAGTGGCGCTATACATCCTCTGGTTGTATATGGACTCAACGAC
GGATGCTCCTCTTCTGTGGGTTACCGCGAACGCTCTGGCGGACGGATGCATCTTGAAACAGAAAAATGTGGAGCG
CTTGGTACTATACTCCATGAGCTCGGCCACGCCCTCGGGCGTTTCCACGAGCAGACCAGAAGCGACCGTGATGAC
TACATCACCGTAATTGAATCCAACATCCTACCGGGATTTGAGCACAACTTCAGGAAGGCCGATCGACCAATACTT
GCCCTGTACGATCTCCAGTCCATCATGCACTACAGTCGTGATGCATTCAGTTTCGACCGCAGGTCATCAACGATC
ATGCCACACAATGCGAAGGATTCAGAGAAGATGGGCACACAACGGAGACTGTCTCGTCTTGACGTATTCACTGTT
AACGCTGCCTACGGTGTGGATGCATTGTGTCCCGCCGACCTTTGGGATAAGTGTCAGCGTGGTGGAGTTCCTAAC
AAGCATTGCTCTTGCGACTGCCTGCCGGAATTTGAAGGCCCGTATTGCGAAATCGTCGTCATTCAAACAGAATGT
AGCGGCCACTACCAGCTGGAAGAATCGGGGCTTATTTCATCCCCAGGTTACCCTAATACCTACTCCAGCAACGCC
AAATGCAGCTATGTCATCGAGTGTGCGAAGGACGAGGTCGTCCAACTGGACTTTTTAGAATTCTCCTTAGAGGGC
TGGGGCTGCTTCTTTGACAAGATGAATATGATGACGGGCGATTTGGTTACAGAGGAGTACCCAGATGAATATTGC
GACGATAACCTGAAAGGAGAATCGGTGTTTTCCAAGAGCAATGTTGCATTGATGACGTTCGAGAGTGACGACATG
TTTAACTACGAAGGATATCAGATATTTTACCACTGTATTCATTCGGACTTGGTATCAAAATAA

(5) Expression of Sfps in secretory cells of *Asterias rubens*

Primers used for ISH probe production are listed in Table S2. Note that the ISH primers were designed based on the transcriptomic data (Lengerer et al 2019).

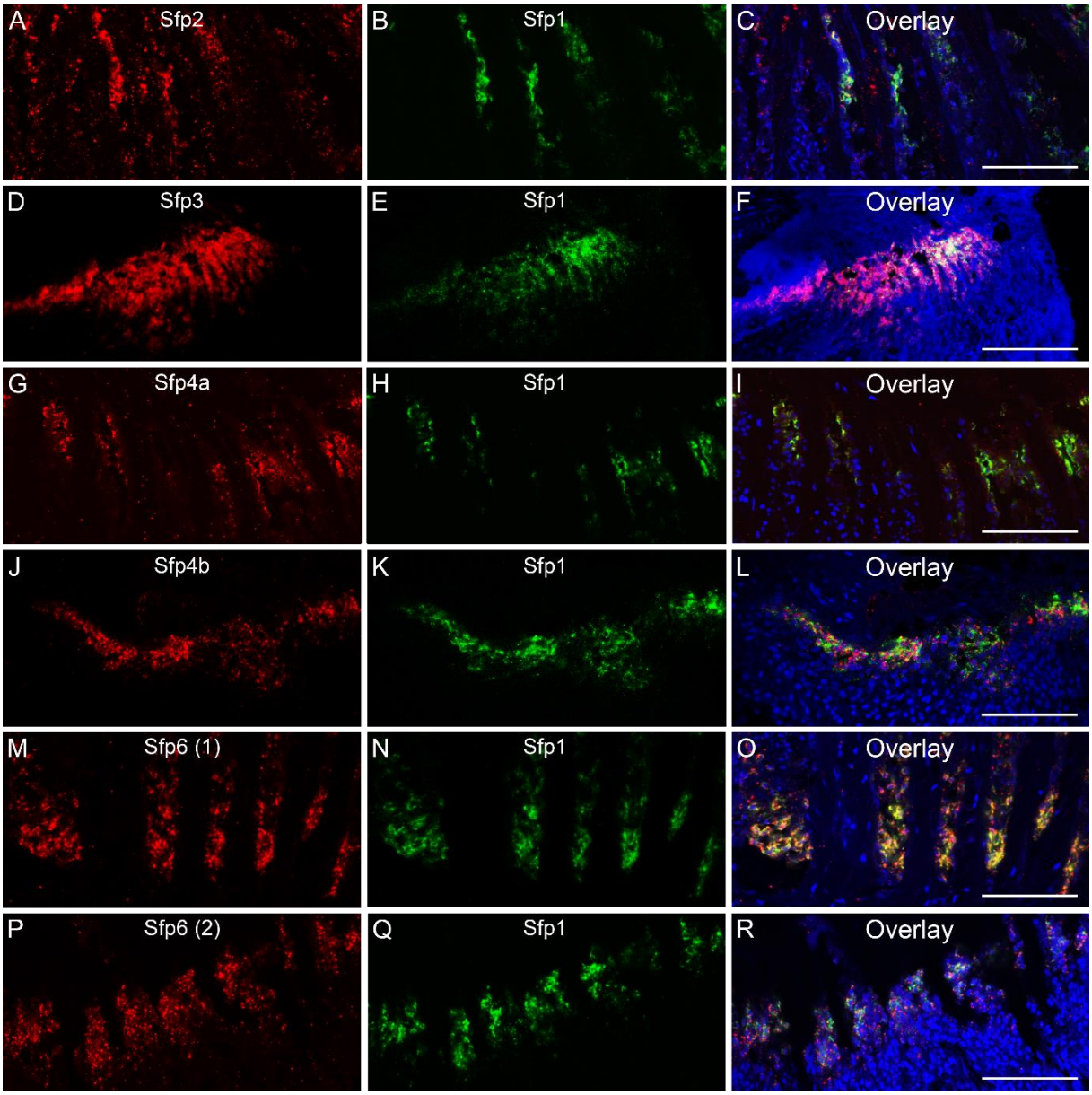


Figure S7. Localization of the expression sites of mRNA coding for Sfp1 (in green) and five Sfps candidates (Sfp2, 3, 4a, 4b, and 6) (in red) on tube foot sections of *Asterias rubens* by double ISH. Co-localization of mRNAs coding for the proteins Sfp1 and Sfp2 (A-C), Sfp1 and Sfp3 (D-F), Sfp1 and Sfp4a (G-I), Sfp1 and Sfp4b (J-L), Sfp1 and Sfp6 (M-R). Two different ISH probes were used for Sfp6, e.g. Sfp6 (1) (M-O) and Sfp6 (2) (P-R) (see Table S2). Scale bar: 50µm.

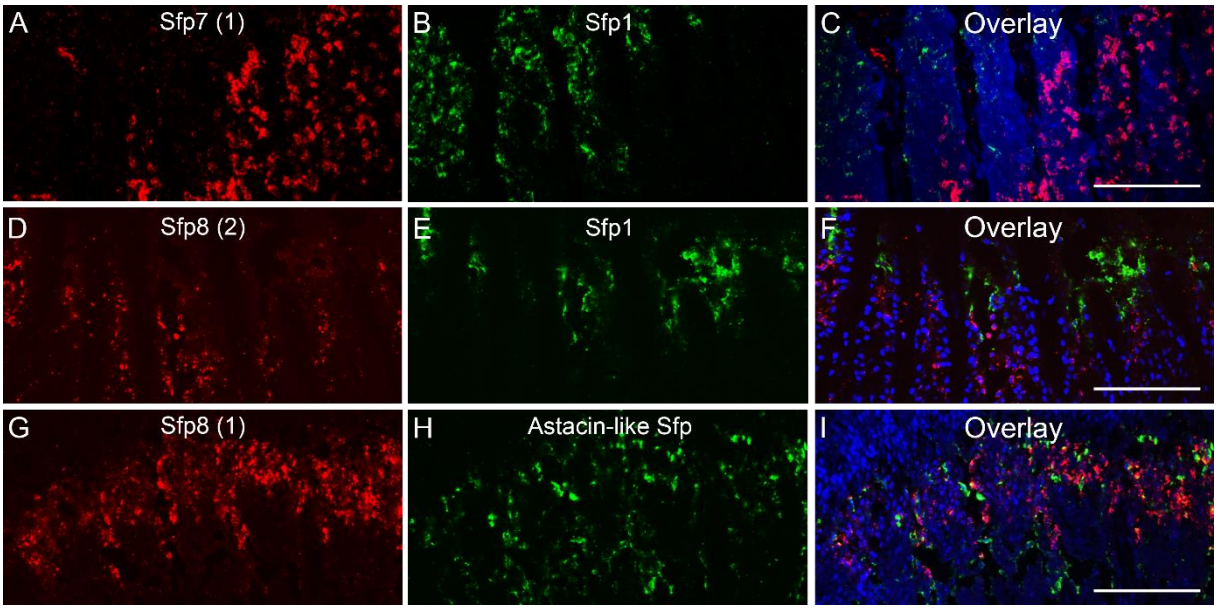


Figure S8. Localization of the expression sites of mRNA coding for Sfp1 or Astacin-like Sfp (in green) and two Sfps candidates (Sfp7 and 8) (in red) on tube foot sections of *Asterias rubens* by double ISH. Co-localization of mRNAs coding for the proteins Sfp1 and Sfp7 (A-C), Sfp1 and Sfp8 (D-F), Astacin-like Sfp and Sfp8 (G-I). Two different ISH probes were used for Sfp8, e.g., Sfp8 (1) (D-F) and Sfp8 (2) (G-I) (see Table S2). Scale bar: 50µm.

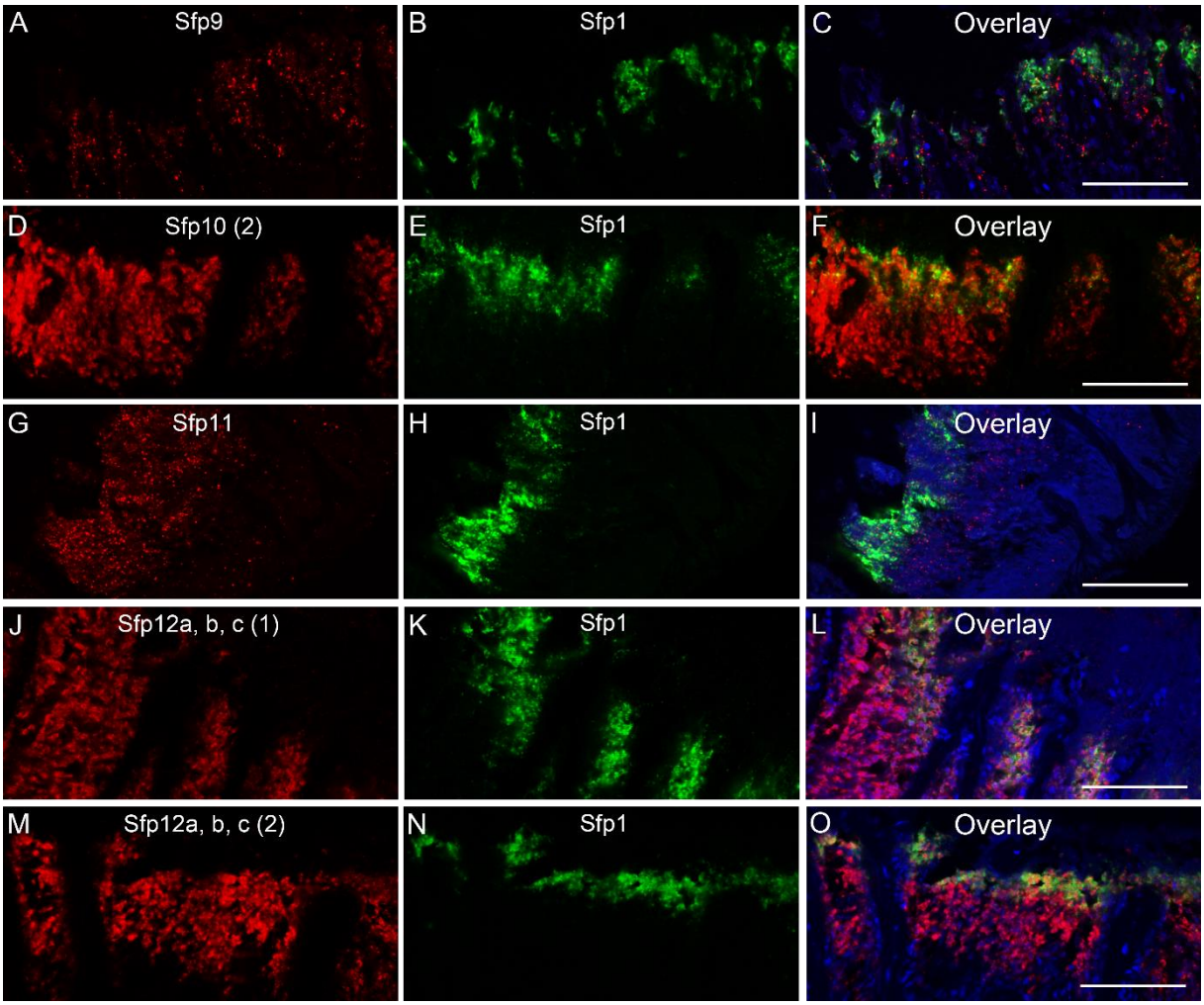


Figure S9. Localization of the expression sites of mRNA coding for Sfp1 (in green) and four Sfps candidates (Sfp9, 10, 11 and 12a, b, c) (in red) on tube foot sections of *Asterias rubens* by double ISH. Co-localization of mRNAs coding for the proteins Sfp1 and Sfp9 (A-C), Sfp1 and Sfp10 (D-F), Sfp1 and Sfp11 (G-I), Sfp1 and Sfp12a, b, c (J-O). Two different ISH probes were used for Sfp12a, b, c, e.g. Sfp12a, b, c (1) (J-L) and Sfp12a, b, c (2) (M-O) (see Table S2). Scale bar: 50µm.

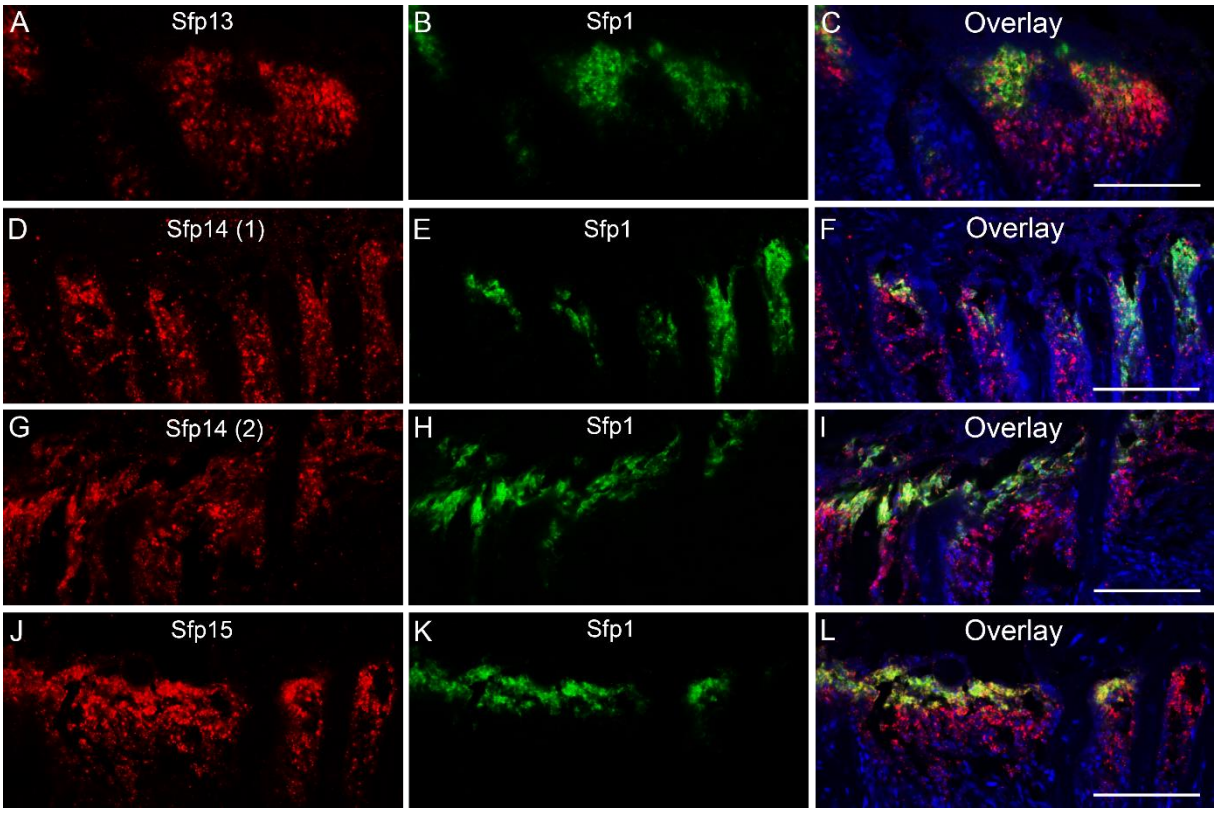


Figure S10. Localization of the expression sites of mRNA coding for Sfp1 (in green) and three Sfps candidates (Sfp13, 14 and 15) (in red) on tube foot sections of *Asterias rubens* by double ISH. Co-localization of mRNAs coding for the proteins Sfp1 and Sfp13 (A-C), Sfp1 and Sfp114 (D-I), Sfp1 and Sfp15 (J-L). Two different ISH probes were used for Sfp14, e.g. Sfp14 (1) (D-F) and Sfp14 (2) (G-I) (see Table S2). Scale bar: 50µm.

(6) Western blot analyses of selected Sfps

Protein extraction and Western blotting

Adhesive footprints were collected as previously described (Hennebert et al., 2012). Tube feet and footprint proteins were extracted as previously described (Hennebert et al., 2014). Briefly, tube feet were rapidly dissected from one sea star and immediately frozen in liquid nitrogen. Freeze-dried footprint material (approximately 2 mg) and tube feet (approximately 400 mg) samples were homogenized in a 1.5 M Tris-HCl buffer (pH 7.8) containing 0.5 M DTT, 5 M urea, 2% (w/v) SDS, and protease inhibitors (cOmplete™, EDTA-free Protease Inhibitor Cocktail, Roche) using a Heidolph Silent Crusher-M homogeniser. The extracts were stored at 4 °C for 15 min to 1 h until use. Footprint and tube foot extracts were centrifuged for 10 min at 19,000 × g and loaded on 10% (w/v) SDS/PAGE gels. After electrophoresis, proteins were transferred onto PVDF membranes (GE Healthcare) using 25 mM Tris, 192 mM glycine, 0.05% (w/v) SDS, 20% methanol (v/v) as transfer buffer. The blots were then immunodetected as previously described (Hennebert et al., 2012) using the anti-Sfp7/8, anti-Sfp10 and anti-

Astacin-like Sfp antibodies diluted 1:1000, followed by HRP-conjugated anti-rabbit antibodies (Invitrogen) and chemiluminescence detection (ECL Western Blotting Substrate or SuperSignal™ West Femto Maximum Sensitivity Substrate, Pierce).

Based on the protein sequences, the predicted sizes are >104 kDa for Sfp7, 149 kDa for Sfp8, >81 kDa for Sfp10 and 48 kDa for Astacin-like Sfp (with propeptide). With the Sfp7/8-antibody, we consistently observed multiple bands in tube foot and footprint extracts (Fig. S11A), presumably deriving from protein fragmentation during protein extraction. With the Sfp10-specific antibody, a large protein band (>200 kDa) was observed in both tube foot and footprint extracts (Fig. S11B). In tube foot extract additionally two bands at around 190 kDa and 110 kDa were present (Fig. S11B). The Astacin-like Sfp-specific antibody detected two protein bands close to the predicted size (55kDa) in the tube foot extract and a faint band at double the size at around 100 kDa in the footprint extract (Fig. S11C).

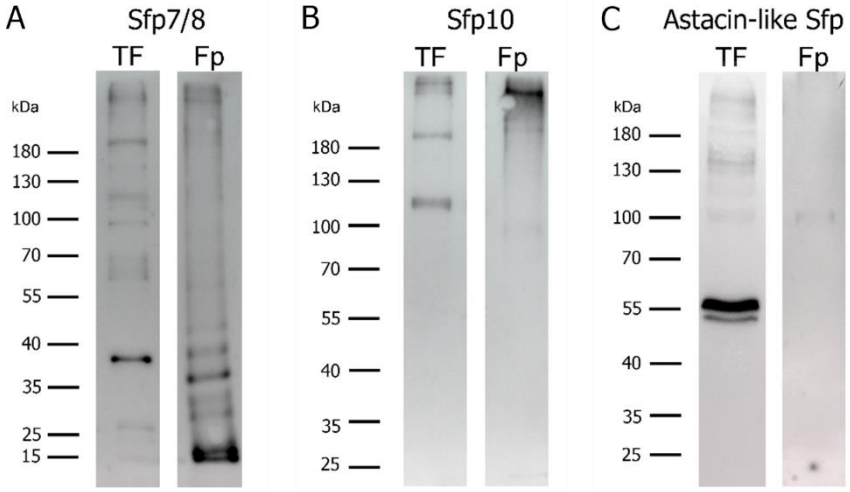


Figure S11. Western blot analysis of selected Sfps in *Asterias rubens* tube feet and footprints. Western blots of the protein extracts immunolabelled with polyclonal antibodies raised against the peptide GDSSQEIKRTLDTK from Sfp7/8 (A), the peptide SLSRMDESQSTDSKL from Sfp10 (B), and the peptide KMNMMTGDLVTEEY from Astacin-like Sfp (C). TF: Tube feet and Fp: Footprint.

(7) Footprint immune-histochemistry

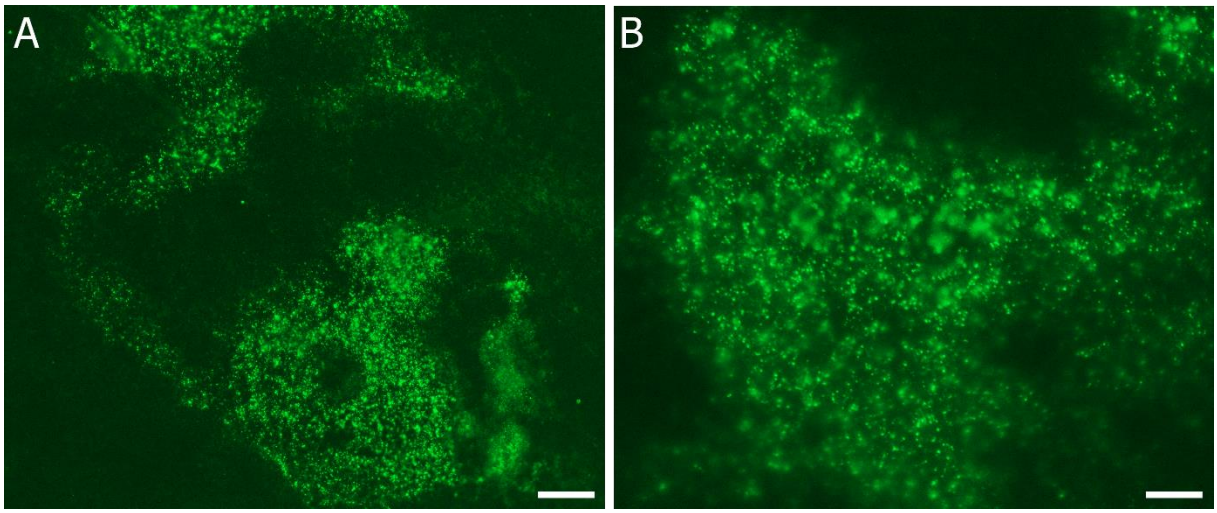


Figure S12. Immunofluorescence localization of Sfp10 in the footprints of *Asterias rubens*. Footprints were labelled with antibodies directed against Sfp10 (A) and high magnification (B). Scale bars: 100 µm (A) and 30 µm (B).

(8) Footprint immune- and lectin histochemistry

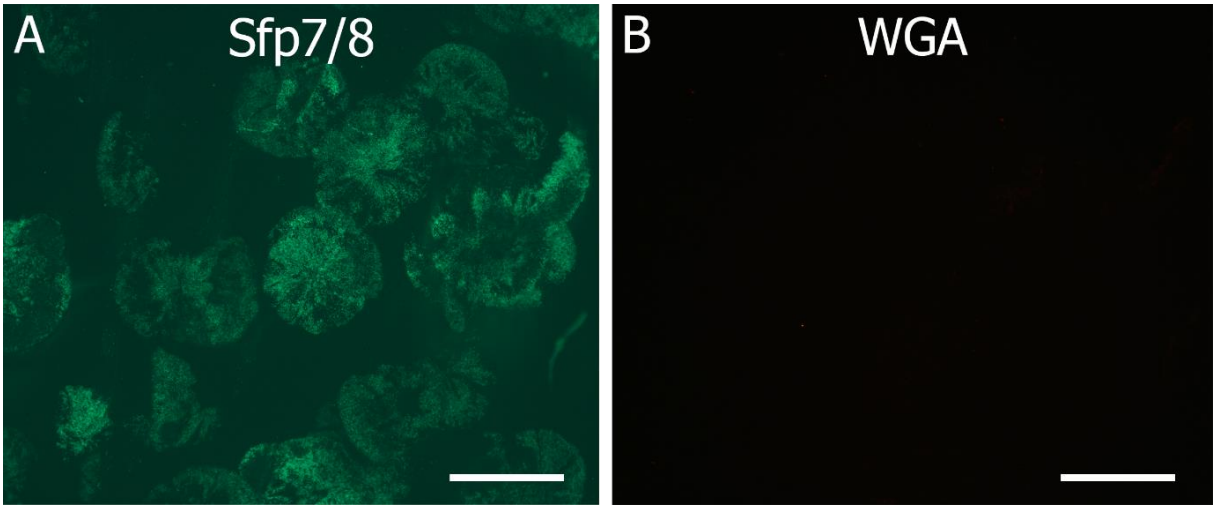


Figure S13. Double immuno- and lectin-histochemistry on footprints of *Asterias rubens*. Footprints obtained from sea stars walking on glass slides were labelled with antibodies directed against Sfp7/8 (A) and with the lectin WGA (B). Scale bars: 500 µm.