

Supplement Table 4: Predicted chimaeric protein products of fusion transcripts.

FUSION CATEGORY	pg
I. PRIVATE FUSION PROTEINS	<u>2</u>
A. FUSIONS IN UNTRANSLATED REGIONS	2
1. <i>Fusions in 3'UTR</i>	2
2. <i>Fusions in 5'UTR</i>	4
B. Fusions in coding regions	4
1. <i>Frame shift fusions that give rise to C-terminal truncation events</i>	4
2. <i>In frame fusions that give rise to chimaeric proteins</i>	16
II. REDUNDANT NON-ISOTYPE-SPECIFIC FUSION PROTEINS	<u>30</u>
A. Fusions in untranslated regions	30
1. <i>Fusions in 3'UTR</i>	30
2. <i>Fusions in 5'UTR</i>	31
B. Fusions in coding regions	31
1. <i>Frame shift fusions that give rise to C-terminal truncation events</i>	31
2. <i>In frame fusions that give rise to chimaeric proteins</i>	35
III. REDUNDANT ISOTYPE-SPECIFIC FUSION PROTEINS	<u>37</u>
A. Fusions in untranslated regions	37
1. <i>Fusions in 3'UTR</i>	37
2. <i>Fusions in 5'UTR</i>	37
B. Fusions in coding regions	38
1. <i>Frame shift fusions that give rise to C-terminal truncation events</i>	38
2. <i>In frame fusions that give rise to chimaeric proteins</i>	38

Color code: Amino acids pertaining to 5' fusion partners are highlighted in teal. Amino acids pertaining to 3' fusion partners fused in frame are highlighted in yellow. Amino acids that are inserted at fusions junctions are highlighted in red.

I. PRIVATE FUSION PROTEINS

A. Fusions in untranslated regions

A.1. Fusions in 3'UTR

IGFBP5->AMD1

3'UTR of IGFBP5 fused into AMD1, No AMD1 in the fused protein

APOOL->DCAF8

3'UTR of APOOL fused into DCAF8, No DCAF8 in the fused protein

TMEM119->ARIH2

3'UTR of TMEM119 fused into ARIH2, No ARIH2 in the fused protein

ASAP1->MALAT1

3'UTR of ASAP1 fused into MALAT1, No MALAT1 in the fused protein

C2orf56->SAMD4B

3'UTR of C2orf56 fused into SAMD4B, No SAMD4B in the fused protein

CDK4->UBA1

3'UTR of CDK4 fused into UBA1, No UBA1 in the fused protein

CIRBP->UGP2

3'UTR of CIRBP fused into UGP2, No UGP2 in the fused protein

DNAJA2->COL14A1

3'UTR of DNAJA2 fused into COL14A1, No COL14A1 in the fused protein

HEATR5A->COL1A1

3'UTR of HEATR5A fused into COL1A1, No COL1A1 in the fused protein

POLD3->COL3A1

3'UTR of POLD3 fused into COL3A1, No COL3A1 in the fused protein

COL3A1->ZNF43

3'UTR of COL3A1 fused into ZNF43, No ZNF43 in the fused protein

IFI27->CPNE3

3'UTR of IFI27 fused into CPNE3, No CPNE3 in the fused protein

RHOBTB3->CRNKL1

3'UTR of RHOBTB3 fused into CRNKL1, No CRNKL1 in the fused protein

EPHA2->CTSD

3'UTR of EPHA2 fused into CTSD, No CTSD in the fused protein

LTBP4->CTSD

3'UTR of LTBP4 fused into CTSD, No CTSD in the fused protein

PACSIN3->CTSD

3'UTR of PACSIN3 fused into CTSD, No CTSD in the fused protein

TMEM109->CTSD

3'UTR of TMEM109 fused into CTSD, No CTSD in the fused protein

CTTC->NCRNA00201

3'UTR of CTTN fused into NCRNA00201, No NCRNA00201 in the fused protein

YWHAG->CYB561

3'UTR of YWHAG fused into CYB561, No CYB561 in the fused protein

KRT81->EMP2

3'UTR of KRT81 fused into EMP2, No EMP2 in the fused protein

SBF1->FLNA

3'UTR of SBF1 fused into FLNA, No FLNA in the fused protein

GAPDH->KRT13

3'UTR of GAPDH fused into KRT13, No KRT13 in the fused protein

GAPDH->MRPS18B

3'UTR of GAPDH fused into MRPS18B, No MRPS18B in the fused protein

GNB1->TRH

3'UTR of GNB1 fused into TRH, No TRH in the fused protein

PTMA->GNB4

3'UTR of PTMA fused into GNB4, No GNB4 in the fused protein

NTN1->HDLBP

3'UTR of NTN1 fused into HDLBP, No HDLBP in the fused protein

TES->HNRNPU

3'UTR of TES fused into HNRNPU, No HNRNPU in the fused protein

RAB3IP->IGFBP5

3'UTR of RAB3IP fused into IGFBP5, No IGFBP5 in the fused protein

MAF->IGFBP7

3'UTR of MAF fused into IGFBP7, No IGFBP7 in the fused protein

ITGA3->KHK

3'UTR of ITGA3 fused into KHK, No KHK in the fused protein

SERPINA1->KIAA1217

3'UTR of SERPINA1 fused into KIAA1217, No KIAA1217 in the fused protein

MAPK1IP1L->XPO1

3'UTR of MAPK1IP1L fused into XPO1, No XPO1 in the fused protein

MGP->REPS2

3'UTR of MGP fused into REPS2, No REPS2 in the fused protein

MRPL52->USP22

3'UTR of MRPL52 fused into USP22, No USP22 in the fused protein

NAV2->WDFY1

3'UTR of NAV2 fused into WDFY1, No WDFY1 in the fused protein

YWHAG->PDIA3

3'UTR of YWHAG fused into PDIA3, No PDIA3 in the fused protein

PIKFYVE->TMEM119

3'UTR of PIKFYVE fused into TMEM119, No TMEM119 in the fused protein

POSTN->TM9SF3

3'UTR of POSTN fused into TM9SF3, No TM9SF3 in the fused protein

POSTN->TRIM33

3'UTR of POSTN fused into TRIM33, No TRIM33 in the fused protein

A.2.Fusions in 5'UTR

GEMIN7->SLC39A14

5'UTR of GEMIN7 fused into the 5' UTR of SLC39A14

RBM6->SLC38A3

5'UTR of RBM6 fused into the 5' UTR of SLC38A3

DIDO1->REPS1

5'UTR of DIDO1 fused into the coding region of REPS1, No ORF that corresponds to any known protein.

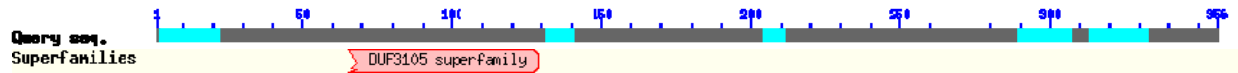
B. Fusions in coding regions

B.1. Frame shift fusions that give rise to C-terminal truncation events

TP53I13->ABCA10

coding to coding fusion—C-terminal truncation product containing aa 1-356 of TP53I13. (full length TP53I13 = 393 aa)

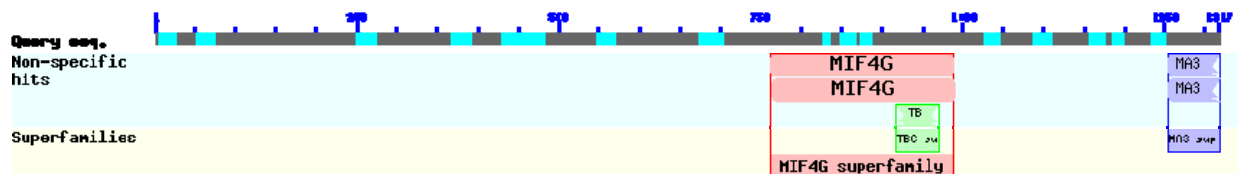
MAPPPSPQLLLLAALARLLGPSEVMAGPAEEAGAHCPESLWPLPPQVSPRVTYTRVSPGQA
 EDVTFLYHPCAH PWLKLQLALLAYACMANPSLTPDFSLTQDRPLVLTAWGLALEMAWVEPAW
 AAHWLMRRRRRKQRKKKAWIYCESLSGPAPSEPTPGRGRLCRRGCVQALALAFALRSWRPP
 GTEVTSQGPRQPSSSGAKRRRLRAALGPQPTRSALRFPSASPGSLKAKQSMAGIPGRESNAP
 SVPTVSLLPGAPGGNASSRTEAQVPNGQGSPGGCVCSQASPAPRAAAPPRAARGPTPRTE
 EAWAAMALTFLLVLLTLATLCTRLHRNFRRGESIWGPTADSQDTVAG*



EIF4G1->ABCC5

Coding region of EIF4G1 fused to 3' UTR of ABCC5—C-terminal truncation containing aa 1-1317 of EIF4G1 .(full length EIF4G1 = 1599 aa)

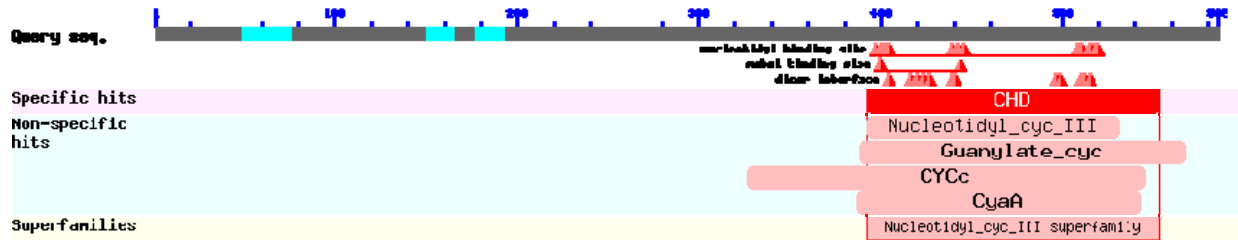
MNKAPQSTGPPAPSPGLPQPAFPPGQTAPVVFSTPQATQMNTPSQPRQHFYPSRAQPPSS
 AASRVQSAAPARPGPAHVYPAGSQVMIPISYPASQGAYYIPGQGRSTYVVTQQYPVQ
 PGAPGFYPGASPTFEGTYAGAYPAQGVQQFPTGVAPT VLMNQPPQIAPKRERKTIRIRDPN
 QGGKDITEEIMSGARTASTPTPPQTGGGLEPQANGETPQVAVIVRPDDRSQGAIADRPLPGP
 EHSPSESQPSSPSPSPVLEPGSEPNLAVLSIPGDTMTTIQMSVEESTPISRETGEPYRLS
 PEPTPLAEPILEVEVTL SKVPPESEFSSSPLQAPTPLASHTVEIHEPNGMVPSLEDLEPEVESSPE
 LAPPACPSESPVPIAPTAQPEELLNGAPSPPAVDLSPVSEPEEQAKEVTASMAPPTIPSATPAT
 APSATSPAQEEEMEEEEEEGEAGEAGEAESEKGGEE LLPESTPIPANLSQNLEAAAATQV
 AVSVPKRRRKIKELNKKEAVGDLLDAFKEANPAVPEVENQPPAGSNPGPESESGVPPRPEEA
 DETWDSKEDIHNAENIQGEQKYEYKSDQWKPLNLEEKRYDRELLGFQFIFASMQKPEGL
 PHISDVVLDKANKTPLRPLDPTRLQGINGCPDFTPSFANLGR TTLSTRGPPRGGPGGELPRGP
 AGLGPRRSQQGPRKEPRKIIATVLMTE DIKLNKA EKAWKPSSKR TAADKDRGEEDADGSKTQD
 LFRRVRSILNKLTPQMFQQLMKQVTQLAIDTEERLKGVIDLIFEKAISEPNFSVAYANMCRLMA
 LKVPTTEKPTVTVNFRKLLLNRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGR LKEELEEA
 RDIARRRSLGNIKFIGELFKL KMLTEAIMHDCVVKLLKNHDEESLECLCRLTTIGKDLDFEKAKP
 RMDQYFNQMEKIIKEKKTSSRIRFMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAEMEEHRE
 HIKVQQLMAKGS DKRRGGPPGPPISRGLPLVDDGGWNTVPI SKGSRPIDTSRLTKITKPGSIDS
 NNQLFAPGGRLSWGKGSSGGSGAKPSDAASEAARPATSTLNRF SALQQAVPTTESTDNRRV
 QRSSLRERGEKAGDRGDR LERSERGGDRGDR LDRARTPATKRSFSKEVEERSRERPSQPE
 GLRKAASLTEDRDRGRDAVKREAA LPPVSPLKAALSEEELEKKS KAIIEEYLHLNDMKEAVQCV
 QELASPSLLFIFVRHGVESTLERSAIAREHMGQLLHQLLCAGHLSTAQYYQG*



CALR->ACACA

Coding region of CALR fused to 3' UTR of ACACA—C-terminal truncation containing aa 1-564 of ADCY9. (full length ADCY9 = 1353 aa)

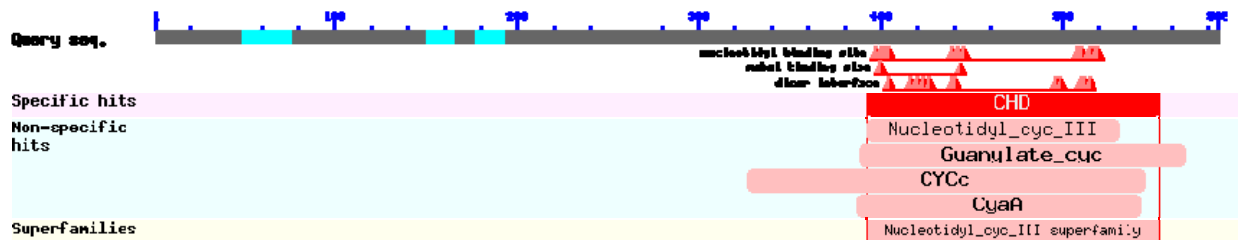
MASPPHQQLLHHHSTEVSCDSSGDSNSVRVKINPKQLSSNSHPKHCKYSSSSCSSSGDSSG
 VPRRVGGGGRLRRQKLPQLFERASSRWWDPKFDSVNLEEACLERCFPQTQRRFRYALFYIG
 FACLLWSIYFAVHMRSRLIVMVAPALCFLLVCVGFFLFTFTKLYARHYAWTSLALTLVVALTLAA
 QFQVLTVPVSGRGDSSNLTATARPTDTCLSQVGSFSMCIEVLFLLYTVMHLPLYLSLCLGVAYSV
 LFETFGYHFRDEACFPSPGAGALHWELLSRGLLHGCIHAIGVHLFVMSQVRSRSTFLKVGQSIM
 HGKDLEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF
 RPFKMQQIEEVSILFADIVGFTKMSANKSAHALVGLLNDLFGRFDRLCEETKCEKISTLGDCYYC
 VAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKEMVMNRVGVHTGTVLCGILGMRRFKFDV
 WSNVNLANLMEQLGVAGKVHISEATAKYLDDRYEMEDGKVIERLGQSVVADQLKALPVPCPS
 RPAAEGASPASLL*



ADCY9->C16orf5

coding region of ADCY9 fused in front of the 5' UTR of C16orf5—C-terminal truncation **aa 1-564 of ADCY9**. (full length ADCY9 = 1353 aa)

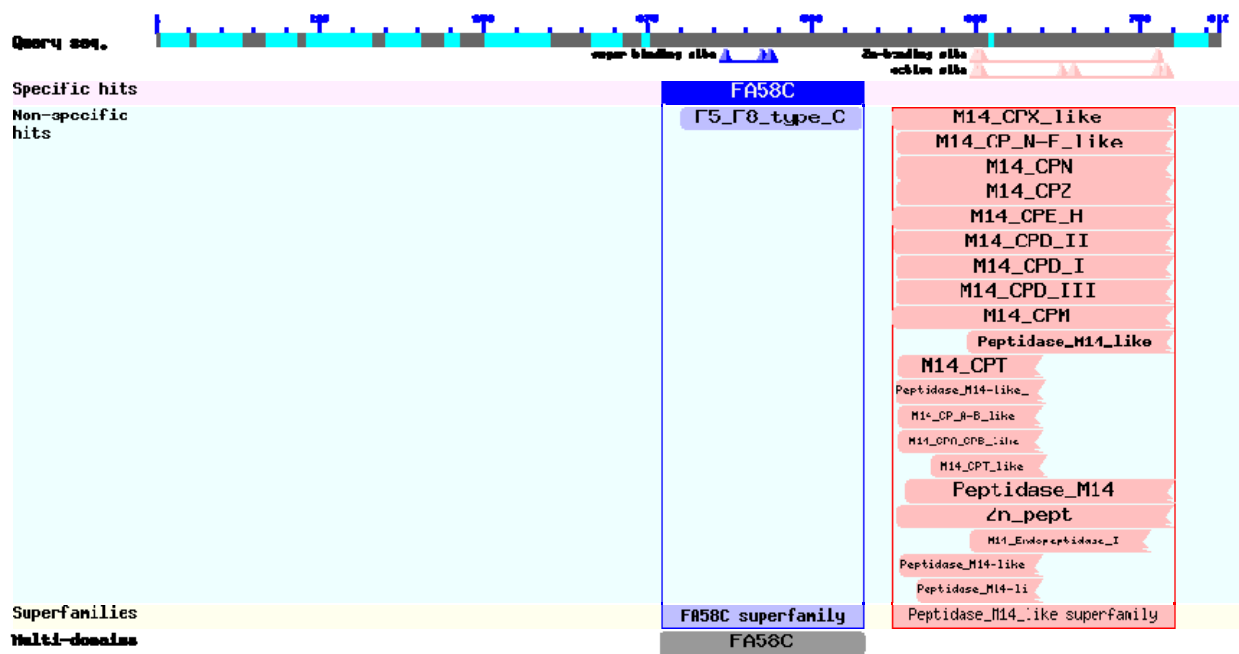
MASPPHQQLLHHHSTEVSCDSSGDSNSVRVKINPKQLSSNSHPKHCKYSSSSCSSSGDSSG
 VPRRVGGGGRLRRQKLPQLFERASSRWWDPKFDSVNLEEACLERCFPQTQRRFRYALFYIG
 FACLLWSIYFAVHMRSRLIVMVAPALCFLLVCVGFFLFTFTKLYARHYAWTSLALTLVVALTLAA
 QFQVLTVPVSGRGDSSNLTATARPTDTCLSQVGSFSMCIEVLFLLYTVMHLPLYLSLCLGVAYSV
 LFETFGYHFRDEACFPSPGAGALHWELLSRGLLHGCIHAIGVHLFVMSQVRSRSTFLKVGQSIM
 HGKDLEVEKALKERMIHSVMPRIIADDLMKQGDEESENSVKRHATSSPKNRKKKSSIQKAPIAF
 RPFKMQQIEEVSILFADIVGFTKMSANKSAHALVGLLNDLFGRFDRLCEETKCEKISTLGDCYYC
 VAGCPEPRADHAYCCIEMGLGMIKAIEQFCQEKEMVMNRVGVHTGTVLCGILGMRRFKFDV
 WSNVNLANLMEQLGVAGKVHISEATAKYLDDRYEMEDGKVIERLGQSVVADQLKALPVPCPS
 RPAAEGASPASLL*



AEBP1->THRA

Coding of AEBP1 fused to 5' of THRA—C-terminal truncation containing **aa 1-776 of AEBP1**. (full length AEBP1 = 1158 aa)

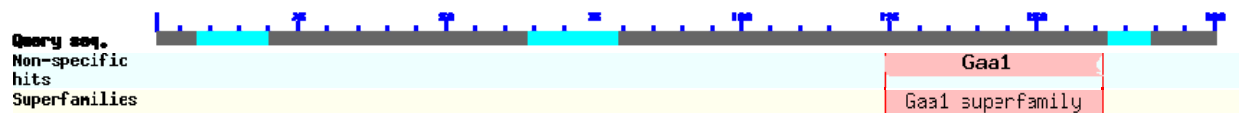
MAAVRGAPLLSCLLALLALCPGGRPQTVLTDDEIEEFLEGFLSELEPEPREDDVEAPPPPEPTP
 RVRKAQAGGKPGKRPGTAAEVPPEKTKDKGKKGKDKGPKVPKESLEGSPPPKKGGKEKPP
 KATKKPKKPKKATKKPKKPKKATKKPKKPKKATKKPPSGKRPPILAPSETLEWPLPPPSP
 GPEELPQEGGAPLSNNWQNPGEETHVEAREHQPEPEEETEQPTLDYNDQIEREDYEDFEYIR
 RQKQPRPPSRRRRPERVWPEPPEEKAPAPAPEERIEPPVKPLLPLPPDYGDGYVIPNYDDM
 DYYFGPPPPQKPAERQTDEEKEELKKPKKEDSSPKEETDKWAVEKGDHKEPRKGEELEEE
 WTPTEKVKCPPIGMESHRIEDNQIRASSMLRHGLGAQRGRLNMQTGATEDDYDGAWCAED
 DARTQWIEVDTRRTTRFTGVITQGRDSSIHDDFVTTFFVGFSDNSQTVWMTNGYEEMTFHG
 NVDKDPVLSPEPVARFIRIYPLTWNGSLCMRLEVLGCSVAPVYSYYAQNEVVATDDLDFR
 HHSYKDMRQLMKVVNEECPTITRTYSLGKSSRGLKIYAMEISDNPGHEHELGEPEFRYTAGIHGN
 EVLGRELLLLLMQYLCREYRDGNPRVRSLVQDTRIHLVPSLNPDPGYEVAAQMGSEFGNWALG
 LWTEEGFDIFEDFPLNSVLWGAEERKWPYRVPNNLPIPERYLSPDATVSTEVRAIIAWMEK
 NPFVLGANLNGGERLVSYPDMAAPRPPRAHSHSHPGAGGGRRPGPHRPPMDAPSTGR*



GPAA1->CD24

coding to coding fusion—C-terminal truncation product containing aa 1-168 of GPAA1. (full length GPAA1 = 621 aa)

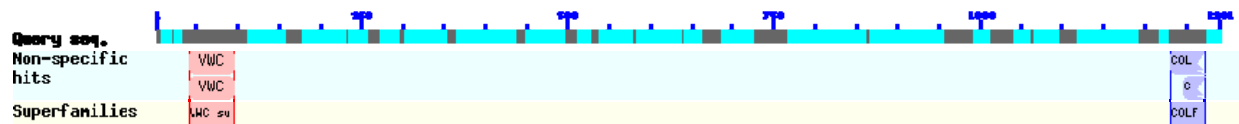
MGLLSDPVRRRALARLVRLNAPLCVLSYVAGIAWFLALVFPPLTQRTYMSENAMGSTMVEEQ
 FAGGDRARAFARDFAAHRKKS GALPVAWLERTMRVSGLEVYTSFSRKLFPDETHERYMVS
 GTVVYGILRAPRAASTESLVTVPCGSDSTNSQAVGLLLALAAAGTAPTHADLFQ*



COL1A1->FGD2

coding to coding fusion—C-terminal truncation containing aa 4-1276 of COL1A1. (full length COL1A1 = 1464 aa)

MFSFVDLRLLLLLAATALLTHGQEEGQVEGQDEDIPPITCVQNGLRYHDRVWKPEPCRICVCD
NGKVLCDVICDETKNCPGAEVPEGECCPVC PDGSESPTDQETTGVVEGPKGDTGPRGPRGP
AGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYDEKSTGGISVPGPMGSPGP
RGLPGPPGAPGPQGFQPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGRPGERG
PPGPQQARGRLPGTAGLPGMKGHRGFSGLDGA KG DAGPAGPKGEPGSPGENGAPGQMGP
GLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAVGAKGEAGPQGPRGSE
GPQGV RGEPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGFPGARGPSGPQGGP
GPPGPKGNSGEPGAPGSKGDTGAKGEPGPVGVQGGPPGAGEEGKRGARGEPTGLPGPP
GERGGPGRGFPGADGVAGPKGPAGERGSPGPAGPKGSPGEAGRPGEAGLP GAKGLT GSP
GSPGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAGERGVPGP
PGAVGPAGKDGEAGAQQPPGPAGPAGERGEQQPAGSPGFQGLPGPAGPPGEAGKPGEQQ
VPGDLGAPGPSGARGERGFPGERGVQGGPPGPAGPRGANGAPGNDGAKGDAGAPGAPGSQ
GAPGLQGM PGERGAAGLP GPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGDK
GESGPSGPAGPTGARGAPGDRGEPGPPGPAGFAGPPGADGQPGAKGEPGDAGAKGDAGP
PGPAGPAGPPGPIGNVGAPGAKGARG SAGPPGATGFPGAAGR VGP GSPGNAGPPGPPGP
AGKEGGKGRGETGPAGRPGEVGP PPGPPG PAGEK GSPGADGPAGAPGTPGPQGIAGQRGV
VGLPGQRGERGFPLPGPSGEPGKQGPSGASGERGPPGPMGPPGLAGPPGESGREGAPGA
EGSPGRDGSPGAKGDRGETGPAGPPGAPGAPGAPGPVGPAGKSGDRGETGPAGPTGPVGP
VGARGPAGPQGRGDKGETGEQDGRGIKHRGFSGLQGGPPGPPGSPGEQQGPSGASGPAGP
RGPPGSAGAPGKDGLNGLPGPIGPPGPRGRTGDAGPVGPPGPPGPPGPPGPPSAGFDF SFL
PQPPQEKAHDGGRYRADDANVVRDRDLEVDTTLSLSQQIENIR SPEGSRKNPARTCRDLK
MCHSDWKSGLHSSRRFSLASRAARLAA*



COL1A1->GORASP2

coding to coding fusion—C-terminal truncation containing aa 1-1338 of COL1A1. (full length COL1A1 = 1464 aa)

MFSFVDLRLLLLLAATALLTHGQEEGQVEGQDEDIPPITCVQNGLRYHDRVWKPEPCRICVCD
NGKVLCDVICDETKNCPGAEVPEGECCPVC PDGSESPTDQETTGVVEGPKGDTGPRGPRGP
AGPPGRDGIPGQPGLPGPPGPPGPPGPPGLGGNFAPQLSYGYDEKSTGGISVPGPMGSPGP
RGLPGPPGAPGPQGFQPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEAGKPGRPGERG
PPGPQQARGRLPGTAGLPGMKGHRGFSGLDGA KG DAGPAGPKGEPGSPGENGAPGQMGP
GLPGERGRPGAPGPAGARGNDGATGAAGPPGPTGPAGPPGFPGAVGAKGEAGPQGPRGSE
GPQGV RGEPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGFPGARGPSGPQGGP
GPPGPKGNSGEPGAPGSKGDTGAKGEPGPVGVQGGPPGAGEEGKRGARGEPTGLPGPP
GERGGPGRGFPGADGVAGPKGPAGERGSPGPAGPKGSPGEAGRPGEAGLP GAKGLT GSP
GSPGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAGERGVPGP
PGAVGPAGKDGEAGAQQPPGPAGPAGERGEQQPAGSPGFQGLPGPAGPPGEAGKPGEQQ
VPGDLGAPGPSGARGERGFPGERGVQGGPPGPAGPRGANGAPGNDGAKGDAGAPGAPGSQ

GAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGSPGKDGVRGLTGPIGPPGPAGAPGDK
 GESGSPGAPPTGARGAPGDRGEPGPPGPAGFAGPPGADGQPGAKGEPGDAGAKGDAGP
 PGPAGPAGPPGPIGNVGAAPGAKGARGSAGPPGATGFPGAAGRVPVGGPSGNAGPPGPPGP
 AGKEGGKGRGETGPAGRPGEVGPVGGPPGAGEKSPGADGPAGAPGTPGPQGIAGQGRGV
 VGLPGQRGERGFPGLPGSPGEPGKQGPSGASGERGPPGPMGPPGLAGPPGESGREGAPGA
 EGSPGRDGPAGKDRGETGPAGPPGAPGAPGAPGVGPAGKSGDRGETGPAGPTGPVGP
 VGARGPAGPQGPRGDKGETGEQDGRGIKGRGFSGLQGPPGPPGSPGEQGPSGASGPAGP
 RGPPGSAGAPGKDGLNGLPGPIGPPGPRGRTGDAGPVGPPGPPGPPGPPSAGFDFLSFL
 PQQPQEKAGDGGRYRADDANVVRDRDLEVDTTLSLSQQIENIRSPEGSRKNPARTCRDLK
 MCHSDWKSGEYWIDPNQGCNLDAIKVFCNMETGETCVYPTQPSVAQKNWYISKNPDKRHHV
 WFGESMTDGFQFEYVIGSEHSFLQL*



GOLPH3L->CTSS

coding to coding fusion—C-terminal truncation containing aa 1-143 of **GOLPH3L**. (full length GOLPH3L = 285 aa)

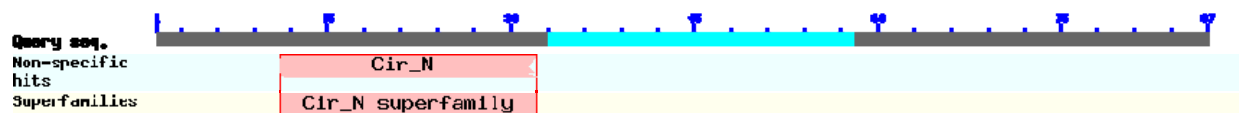
MTTLTHRARRTEISKNSEKKMESEEDSNWEKSPDNEDSGDSKDIRLTLMEEVLLGLKDKEYG
 TSWNDCCISSGLRGGILIELAMRGRIYLEPPTMRKKRLLDRKVLKSDSPTGDVLLDETLKHIKAT
 EPTETVQTWIELLTAGATTLVKKDIFGWQEIKEIIVGLLAFPLTQKSRGSLFITNQEI*



CWC25->ROBO2

coding to coding fusion—C-terminal truncation containing aa 1-63 of **CWC25**. (full length CWC25 = 425 aa)

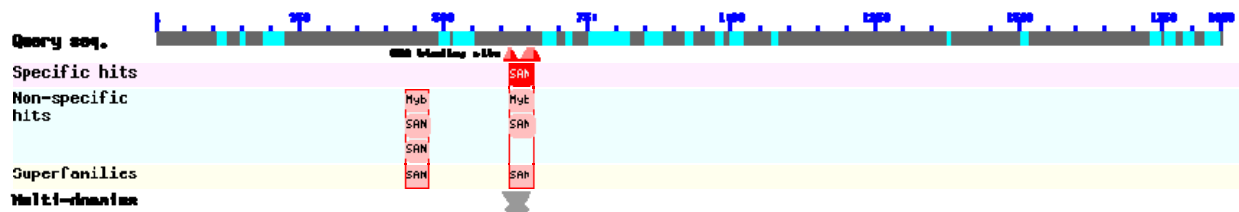
MGGDLNLKKSHPQTLRNVEKVWKAQKHEAERKKIEELQRELREERAREEMQRYAEDVG
 AVNSPTVCGSAKRSDCCSRNSDISL*



NCOR2->ELN

coding to coding fusion—C-terminal truncation containing aa 1-1824 of **NCOR2**. (full length NCOR2 = 2525 aa)

MSGSTQPVAQTRATEPRYPHSLSYVQIARTHTDVGLLEYQHHSRDYASHLSPGSIIQPQR
 RRPSSLSEFQPGNERSQELHLPESHYSYLPPELGKSEMEFIESKRPRLELLPDPLLRPSLLATG
 QPAGEDLTKDRSLTGKLEPVSPSPPHTDPELELVPPRLSKEELIQNMDRVDREITMVEQQIS
 KLKKKQQQLEEEAAKPPPEPEKPVSPPIESKHRSLVQIYDENRKKAEAAHRILEGLGPQVELPL
 YNQPSDTRQYHENIKINQAMRKKLILYFKRRNHARKQWEQKFCQRYDQLMEAWEEKVERIEN
 NPRRRAKESKVREYYEKQFPEIRKQRELQERMQSRVQGRGSGLSMSAARSEHEVSEIIDGLSE
 QENLEKQMRQLAVIPPMLYDADQQRIFINMNGLMADPMKVYKDRQVMNMWSEQEKETFRE
 KFMQHPKNFGLIASFLERKTVAEVLYYYLTKKNENYKSLVRRSYRRRGKSQQQQQQQQQQQ
 QQQQQQPMPRSSQEEKDEKEKEKEAEKEEEEKPEVENDKEDLLKEKTDGSDNDEKEAVA
 SKGRKTANSQGRRKGRITRSMANEANSEEAITPQQSAELASMELESSRWTEEMETAKKGL
 LEHGRNWSAIARMVGSKTVSQCKNFYFNYKKRQNLDEILQQHKLKMEKERNARRKKKKAPAA
 ASEEAAFPPVVEDEEMEASGVSGNEEEMVEEAEATVNNSSDTEIPSPHTEAAKDTGQNGPK
 PPATLGADGPPPGPPTPPPEDIPAPTEPTPASEATGAPTPPPAPPSAPSAPPVVPKEEKEEETA
 AAPVVEEGEEQKPPAAEELAVDTGKAEEPVKSECTEEAEEGPAKGDAAEAEATAEGALKAEK
 KEGGSGRATTAKSSGAPQDSDSSATCSADEVDEAEGGDKNRLSPRPSLLTPTGDPRANASP
 QKPLDLKQLKQRAAAIPIQVTKVHEPPREDAAPTTPAPPAPPQNLQPESDAPQQPGSSPR
 GKSRSAPPADKEAEKPVFFPAFAEAQKLPGDPPCWTSGLPFPVPPREVIKASPHAPDPSAF
 SYAPPGHPLPLGLHDTARVLP RPPTISNPPLISSAKHPSVLERQIGAISQGMVQLHVPYSEH
 AKAPVGPVTMGLPLPMDPKLAPFSGVKQEQLSPRGQAGPPESLGVPTAQEASVLRGTALGS
 VPGGSITKGIPSTRVPSDSAITYRGSITHGTPADVLYKGTITRIIGEDSPSRLDRGREDSLPKGHV
 IYEGKKGHVLSYEGGMSVTQCSKEDGRSSSGPPHETAAPKRTYDMMEGRVGRAISSASIEGL
 MGRAIPPERHSPHHLKEQHHRGSITQGIPRSYEAQEDYLRREAKLLKREGTPPPPPPSRDLT
 EAYKTQALGPLKPKAHEGLVATVKEAGRSIHEIPREELRHTPELPLAPRPLKEGSITQGTPLKY
 DTGASTTGSKKHDVRSLLIGSPGRTFPPVHPLDVMADARALERACYEESLKS RPGTASSSGGSI
 ARGAPVIVPELGKPRQSPLTYEDHGAPFAGHLPRGSPVTTREPTPRLQEGSLSSSKASQDRKL
 TSTPREIAKSPHSTVPEHHPHPISPYEHLLRGVSGVDLYRSHIPLAFDPTSIPRGIPLDAAAAYL
 PRHLAPNPTYPHLYPPYLIRGYPDTAALENRQTIINDYITSQQMHNAATAMAQRADMLRGLSP
 RESSLALNYAAGPRGIIDLSQVPHLPVLVPPTPGTPATAMDRLAYLPTAPQPFSSRHSSSPLSP
 GGPTHLTKPTTSSSERERDRDRERDREREKSILTSTTTVEHAPIWRPGTEQSSGSSGGG
 GPPLPQPKLLPKPPSLA*



HMGN3->PAQR8

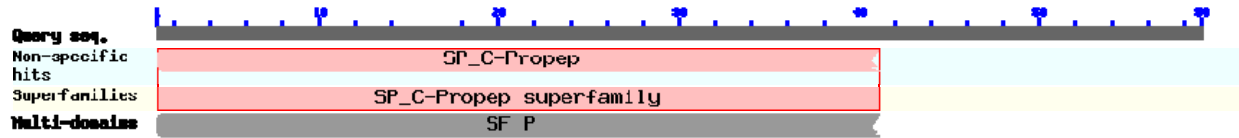
Coding of HMGN3 fused to 5' of PAQR8—121 aa ORF with no significant homology to any known protein.

MPKRKVAYPVLRARHGVHAGRCHDDRHLGAPEHPVGQRAAAAPPAQDPGGWASQDALHCP
 RNGCAPALPGALPHRLPPHGARVALLLLQPLSETQRGGQRLDPFTGSPGRPLAILGLCRG*

SFTPC->IGLL5

Coding of SFTPC fused to 5' of IGLL5—C-terminal truncation containing aa 1-35 of SFTPC. (full length SFTPC = 197 aa)

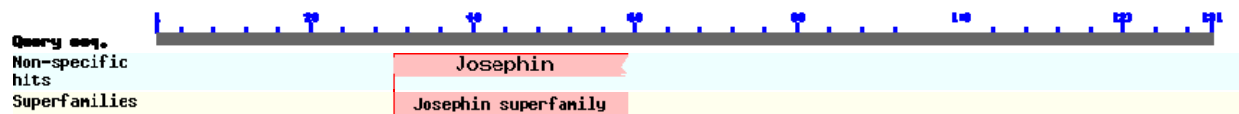
MDVGSKEVLMESPPDYSAAPRGRFGIPCCPVHLKRLIVVVGPPNPPNRATTSTRPAAT*



JOSD1->RPS19BP1

coding to coding fusion—C-terminal truncation containing aa 1-61 of JOSD1. (full length JOSD1 = 202 aa)

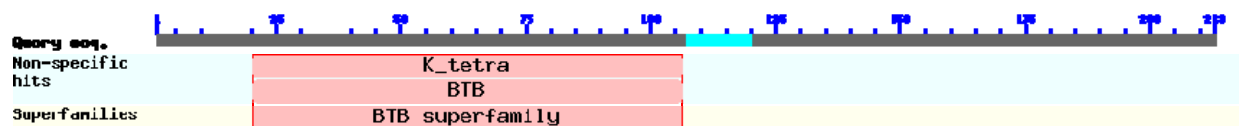
MSCVPWKGDKAKSESELELPQAAPPQIYHEKQRRELCALHALNNVFQDSNAFTRDTLQEIQSP
PGPSRSGQAERGSGETAPEDEGNSGPETAELGQGKGAQVGTGRVPEARVSRPPQSKPEVSD
QDEKHRG*



KCTD3->TXNDC16

coding to coding fusion—C-terminal truncation containing aa 1-208 of KCTD3. (full length KCTD3 = 815 aa)

MAGGHCGSFPAAGSAGEIVQLNVGGTRFSTSRQTLMWIPDSFFSLLSGRISTLRDETGAIFI
DRDPAAFAPILNFLRTKELDLRGVSNVLRHEAEFYGITPLVRRLLLCEELERSSCGSVLFHGYP
PPGIPSRKINNTVRSADSRNGLNSTEGEARGNGTQPVLSGTGEETVRLGFPVDPRKVLIVAGH
HNWIVAAAYAHFAVCYSKRRF*

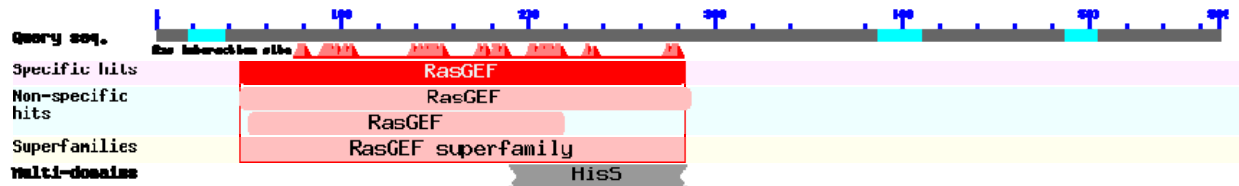


RALGPS2->LAMB3

coding to coding fusion—C-terminal truncation containing aa 1-477 of RALGPS2. (full length RALGPS2 = 583 aa)

MDLMNGQASSVNIAATASEKSSSSESLSDKGSELKKSFDVAVFDVLKVTPEEYAGQITLMDVPV
FKAIQDELSSCGWNNKKEYSSAPNAVAFTRRFNHVSFVWVREILHAQTLKIRAEVLSHYIKTAK
KLYELNNLHALMAVVSGLQSAPIFRLTKTWALLSRKDKTTFEKLEYVMSKEDNYKRLRDYISSLK
MTPCIPYLGIIYLSDLTYIDSAYPSTGSILENEQRSNLMNINLRIISDLQQSCEYDIPMLPHVQKYLN

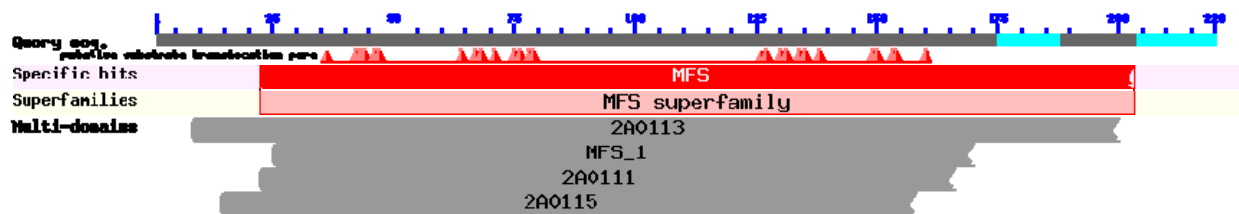
SVQYIEELQKFVEDDNYKLSLKIIEPGTSTPRSAASREDLVGPEVGPASQSGRKSVAEAGALLPQ
 TPPSPRNLIPIHGHRKCHSLGYNFIHKMNTAEFKSATFPNAGPRHLLDDSVMEPHAPSRGQAES
 STLSSGISIGSSDGSELSEETSWPAFERNRLYHSLGPVTRVARNGYRSHMKASSSAESEDLAV
 HLYPGA VTIQGVLRRTLLKEGKKPTPVTVISGEQRARAATRHQAAASAALA*PGPAVTSASEA
 TVIATRC AWPATLASRPM MRTSGSRPCALVDSAMP PPACGQGLGWRTVAWPPGS*



SLC16A3->MRPL4

Coding of SLC16A3 fused to the 3' UTR of MRPL4—C-terminal truncation containing aa 1-213 of SLC16A3. (full length SLC16A3 = 465 aa)

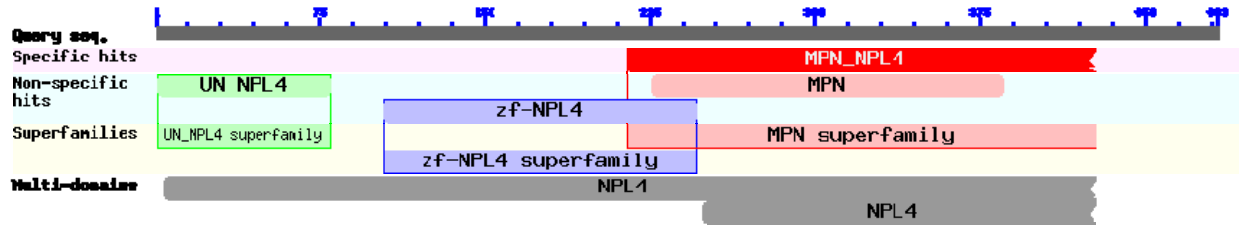
MGGAVVDEGPTGVKAPDGGWGWAVLFGCFVITGFSYAFPKAVSVFFKELIQEFGIGYSDTAWI
 SSILLAMLYGTGPLCSVCVNRFGCRPVMLVGGFLFASLGMVAASFCSRSIQVYLTTGVITGLGLAL
 NFQPSLIMLNRYFSKRRPMANGLAAAGSPVFLCAL SPLGQLLQDRYGWRGGFLILGGLLLNCC
 VCAALMRPLVVTAQPGSRPPAPPPPRPQA*



NPLOC4->PDE6G

coding to coding fusion—C-terminal truncation containing aa 1-427 of NPLOC4. (full length NPLOC4 = 608 aa)

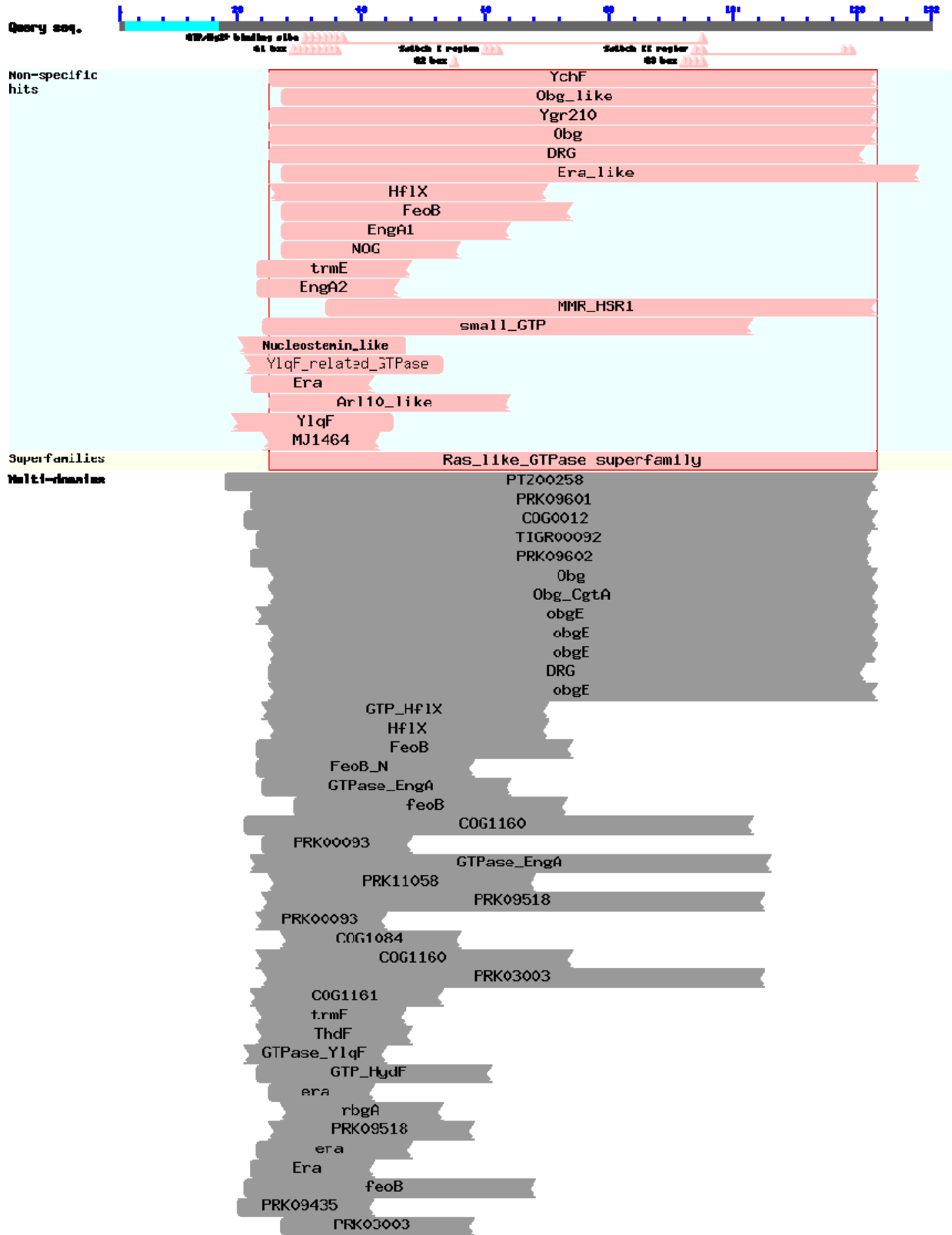
MAESIIRVQSPDGVK RITAKRETAATFLKKVAKEFGFQNNGFSVYINRNKTGEITASSNKS LNL
 LKIKHGDLLFLFPSSLAGPSSEMETSVP PGFKVFGAPNVVEDEIDQYLSKQDGKIYRSRDPQLC
 RHGPLGKCVHCVPLEPDEDYLNHLEPPVKHMSFHAYIRKLTGGADK GK FVALENISCKIKSGC
 EGHLWPNGICTKCQPSAITLNRQKYRHVDNIMFENHTVADRFLDFWRKTGNQHFGYLYGRY
 TEHKDIPLGIRAEVAAYEPPQIGTQNSLELLEDPKAEVVDEIAAKLGLRKGWIFTDLVSEDTRK
 GTVRYSRNKD TYFLSSEECITAGDFQNKHPNMCRLSPDGHFGSKFVTAVATGGPDNQVHFEG
 YQVSNQCMALVRDECLLPCKDAPELGYAKESSEQYVPDV FYK VVGRHPWN GRPGRNRHHS
 HLPLGGLQPPGAARAGPIWHLARGPAEVQTLPLLP TVL*



OLA1->ORMDL3

coding region of OLA1 fused in front of the 5' UTR of ORMDL3—C-terminal truncation containing aa 1-124 of OLA1. (full length OLA1 = 396 aa)

MPPKKGDDGIKPPPIIGRFGTSLKIGIVGLPNVGKSTFFNVLNLSQASAENFPFCTIDPNESRVP
 VPPERFDLFCQYHKPASKIPAFNLNVVDIAGLVKGAHNGQGLGNAFLSHISACDGIHFLTQQQRG
 AAG*



SFI1->YPEL1

coding region of SF11 fused in front of the 5' UTR of YPEL1—C-terminal truncation containing aa 1-150 of SF11. (full length SF11 = 1242 aa)

MKNLLTEKCISSHNHFHQKVIKQRMEKKVDSRYFKDGAVKKPYSAKTLNKKSSASFGIRRELPS
TSHLVQYRGTHCTRQGRLRELRIKRVARKFLYLWIRMTFGRVFPKARFYEQRLLRKVFEE
WKEEWWVFQHEWKLCVRADCHYRLRFSLVPAVWTVPHALLDNDGSCVHFPVLLLSERPAEP
RFFQNQP*

No putative conserved domains detected.

PCNX->MKKS

coding region of PCNX fused in front of the 5' UTR of MKKS—C-terminal truncation containing aa 1-814 of PCNX. (full length PCNX = 2341 aa)

MGSQTLQILRQGVWAALSGGWYYDPHQATFVNALHLYLWLFLLGLPFTLYMALPSTMIIVAVYC
PVIAAVFIVLKMVNYRLHRLDAGEVVDRTANEFTDQRTKAEQGNCSSTRKDSNGPSDPGGGI
EMSEFIREATPPVGCSSRNSYAGLDPSNQIGSGSSRLGTAATIKGDTDTAKTSDDISLSLGQSS
SLCKEGSEEQDLAADRKLFRVSNDSFISIQPSLSSCGQDLPRDFSDKVNLP SHNHHHHVDQS
LSSACDTEVASLVPLHSHSYRKDHRPRGVPRTSSSAVAFPDTSLNDFPLYQRRGLDPVSELE
SSKPLSGSKESLVENSGLSGEFQLAGDLKINTSQPPTKSGKSKPLKAEKSMDSLRSLSTRSSG
STESYCSGTDRDTNSTVSSYKSEQTSSTHIESILSEHEESPAGTKSGRKKKECCAGPEEKNSC
ASDKRTSSEKIAMEASTNSGVHEAKDPTPSDEMHNQRGLSTSASEEANKNPHANEFTSQGDR
PPGNTAENKEEKSDKSAVSVDSKVRKDVGGKQKEGDVVRPKSSSVIHRTASAHKSGRRRTGKK
RASSFDSSRRDYVCFRGVSGTKPHSAIFCHDEDSSDQSDLSRASSVQSAHQFSSDSSSSTT
SHSCQSPEGRYSAKTKHCHKERTDSEHTHKAHLVPEGTSKKRATRRTSSTNSAKTRARVLS
LDSGTVAACLNDNRLMAPESIKPLTTSKSDLEAKEGEVLDELSLLGRASQLETVTRSRNSLPNQ
VAFPEGEEQDAVSGAAQASEEAVSFRRERSTFRRQAVRRRHAGSNPTPPTLLIGSPLR*

No putative conserved domains detected.

SLC39A6->LRIG1

coding to coding fusion—C-terminal truncation containing aa 1-234 of SLC39A6. (full length SLC39A6 = 755 aa)

MARKLSVILITFALSVTNPLHELKAAAFPQTTEKISPWESGINVDLAISTRQYHLQQLFYRYGE
NNSLSVEGFRKLLQNIQIDKIKRIHHHDHSDHEHSDHERHSDHEHHSEHEHSDHDH
SHHNHAASGKNKRKALCPDHDSDSSGKDPRNSQGKGAHRPEHASGRRNVKDSVSASEVTST
VYNTVSEGTHFLETIETPRPGKLFKDVSSSTPPSVTSKSRVSRRCPGTCPPGRGA*

No putative conserved domains detected.

GLI3->FAM3B

coding to coding fusion—C-terminal truncation containing aa 1-41 of GLI3. (full length GLI3 = 1580 aa)

MEAQSHSSTTEKKNVENSIVKCSTRTDVSEKAVASSTTSND*

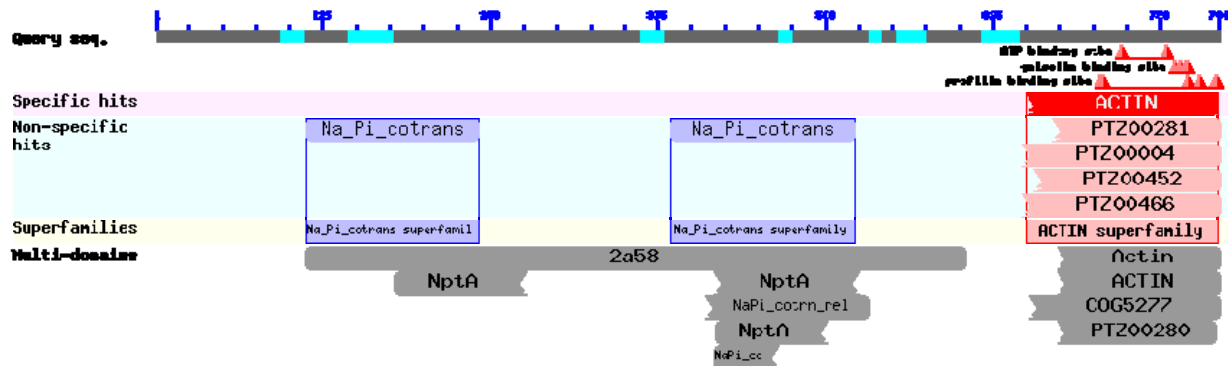
No putative conserved domains detected

B.2. In frame fusions that give rise to chimaeric proteins

SLC34A2->ACTB

coding to coding fusion—aa 1-671 of SLC34A2 fused in frame to aa 33-159 of ACTB with a single aa insertion at the fusion junction.

MAPWPELGDAQPNPDKYLEGAAGQQPTAPDKSKETNKTDNTEAPVTKIELLPSYSTATLIDEPT
 EVDDPWNLPTLQDSGIKWSERDTKGKILCFFQGIGRLILLGLFLYFFVCSLDLSSAFQLVGGKM
 AGQFFSNSSIMSNPLLGLVIGVLVTVLVQSSSTSTSIVVSMVSSLLTVRAAIIIMGANIGTSITNT
 IVALMQVGRSEFRRAFAGATVHDFFNWLSVLVLLPVEVATHYLEIITQLIVESFHFKNGEDAPD
 LLKVITKPFTKLIVQLDKKVISQIAMNDEKAKNKSLVKIWCKTFTNKTQINVTVPSTANCTSPSLC
 WTDGIQNWTKMKNVYKENIAKQCQHIFVNFHLPDLAVGTILLILLSLLVLCGCLIMIVKILGSVLKGQV
 ATVIKKTINTDFPPFAWLTGYLAILVGAGMTFIVQSSSVFTSALTPLIGIGVITIERAYPLTLGNSIG
 TTTTAILAALASPGNALRSSLQIALCHFFFNISGILLWYPIPFTRLPIRMAKGLGNISAKYRWFVAVF
 YLIIFFFLIPLTVFGLSLAGWRVLVGVGVPVVFIIILVLCRLLLQSRCPRVLPKQLQWNFLPLWMR
 SLKPWDAVVSKFTGCFQMRCCCCRVCRCACLLCDCPKCCRCCKCEDLEEAQEGQDVPV
 KAPETFDNITISRFRCPEALFQPSFLGMESCGIHETTFFNSIMKCDVDIRKDLYANTVLSGGTTM
 YPGIADRMQKEITALAPSTMKIKIIPPERKYSVWIGGSILASLSTFQQMWISKQEYDESGPSIVH
 RKCF*

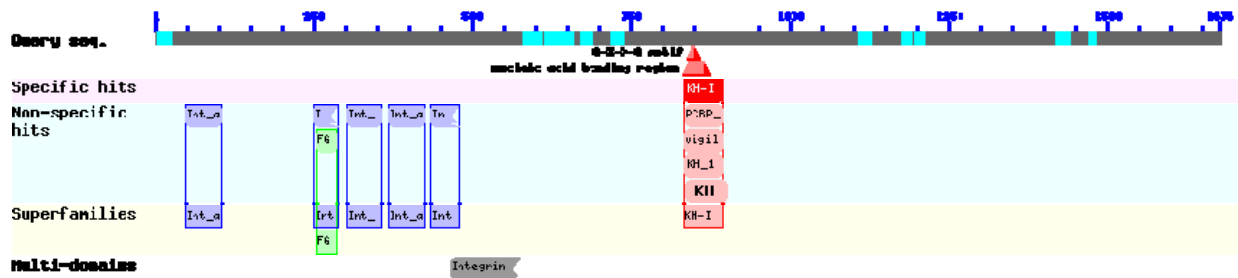


ITGAV->ANKHD1

coding to coding fusion—aa 1-478 of ITGAV fused in frame to aa 1439-2542 of ANKHD1.

MAFPRRRLRLGPRGLPLLLSGLLLPLCRAFNLVDVSPAKEYSGPEGSYFGFAVDFFVPSASSR
 MFLLVGAPKANTTQPGIVEGGQVLKCDWSSTRRCQPIEFDATGNRDYAKDDPLEFKSHQWFG
 ASVRSKQDKILACAPLYHWRTMKQEREPVGTCLFQDGTKTVEYAPCRSQDIDADGGQFCQG
 GFSIDFTKADRVLLGGPGSFYWQQLISDQVAEIVSKYDPNVYSIKYNNQLATRTAQAIFFDSSYL
 GYSVAVGDFNGDGIDDFVSGVPRAARTLGMVYIYDGKNMSSLYNFTGEQMAAYFGFSVAATI
 NGDDYADVFIGAPLFMDRGSQGLQEVGQVSVSLQRASGDFQTTKLNQFEVFAFGSAIAPLG
 DLDQDGFNDIAIAAPYGGEDKKGIVYIFNGRSTGLNAVPSQILEGQWAARSMPPSFGYSMKGA
 TDIDKNGYPDLIVGAFGVDRAILYRARPVITVNAGLEVYPSILNQDNKTCSLPGTALKVSCFNVR
 FCLKADGKGVLPKLNQVVELLLDKLKQKGAIRRALFLYSRSPSHSKNMTISRGGMLMQCEELIA
 YLRSREESRKQALAAKREKRKEKRKKKKEEQKRKQEEDEENKPKENSELPEDEDEEENDEVDV
 EQEVPPEPPSATTTTTIGISATSATFTNVFGKKRANVVTPSTNRKNKKNKTKETPPTAHLILPEQ
 HMSLAQQKADKNKINGEPRGGGAGGNSDSDNLDSTDCNSESSSGGKSQELNFVMDVNSSKY

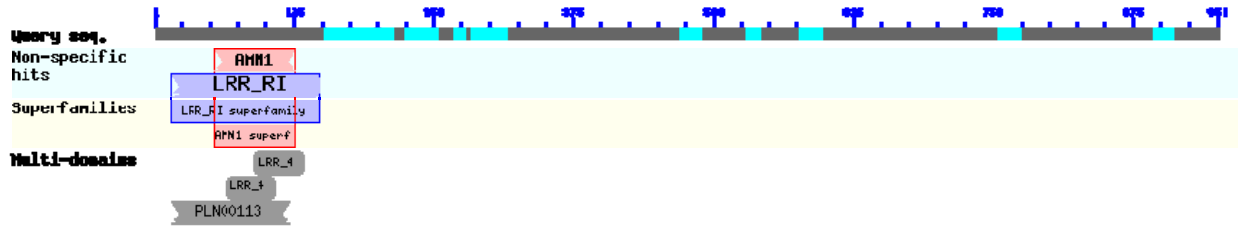
PSLLLHSQEEKTSTATSKTQTRLEGEVTPNSLSTSYKTVSLPLSSPNIKLNLTSPKRGQKREEG
WKEVRRSKKLSVPASVVSRLMGRGGCNITAIQDVTGAHIDVDKQKDKNGERMITIRGGTESTR
YAVQLINALIQDPAKELEDLIPKNHIRTASTKSIHANFSSGVGTTAASSKNAFPLGAPTLVTSQA
TTLSTFQPANKLNKNVPTNVRSSFPVSLPLAYPHPHFALLAAQTMQQIRHPRLPMAQFGGTF
PSPNTWGPFPVRPVNPGNTNSSPKHNNTSRLPNQNGTVLPSESAGLATASCPITVSSVVAASQ
QLCVTNRTRTPSSVRKQLFACVPKTSPPATVISSVTSTCSSLPSVSSAPITSGQAPTTFLPASTSQ
AQLSSQKMESFSAVPPTKEKVSTQDQPMANLCTPSSTANSCSSSSASNTPGAPETHPSSSPTP
TSSNTQEEAQSSVSDLSPMSMPFASNSEPAPLTTSRPMVAADNQDTSNLPQLAVPAPRVS
HRMQPRGSFYSMVPNATIHQDPQSIFVTNPVTLTPPQGPPAAVQLSSAVNIMNGSQMHINPAN
KSLPPTFGPATLNFHSSLFDSSQVPANQGWGDGPLSSRVATDASFTVQSAFLGNSVLGHLEN
MHPDNSKAPGFRPPSQRVSTSPVGLPSIDPSGSSPSSSSAPLASFSGIPGTRVFLQGPAPVGT
PSFNRQHFSHPWTSASNSSTSAPPTLGQPKGVSASQDRKIPPIGTERLARIRQGGSSVAQAP
AGTSFVAPVGHSGIWSFGVNAVSEGLSGWSQSVMGNHMPHQQLSDPSTFSQHQPIMERDDS
GMVAPSNIFHQPMASGFVDFSKGLPISMYGGTIIPSHPQLADVPGGPLFNGLHNPDPAWNPMI
KVIQNSTECTDAQIWPGTWAPHIGNMHLKYVN*



ANP32E->MYST4

coding to coding fusion—aa 1-246 of ANP32E fused in frame to aa 1377-2073 of MYST4.

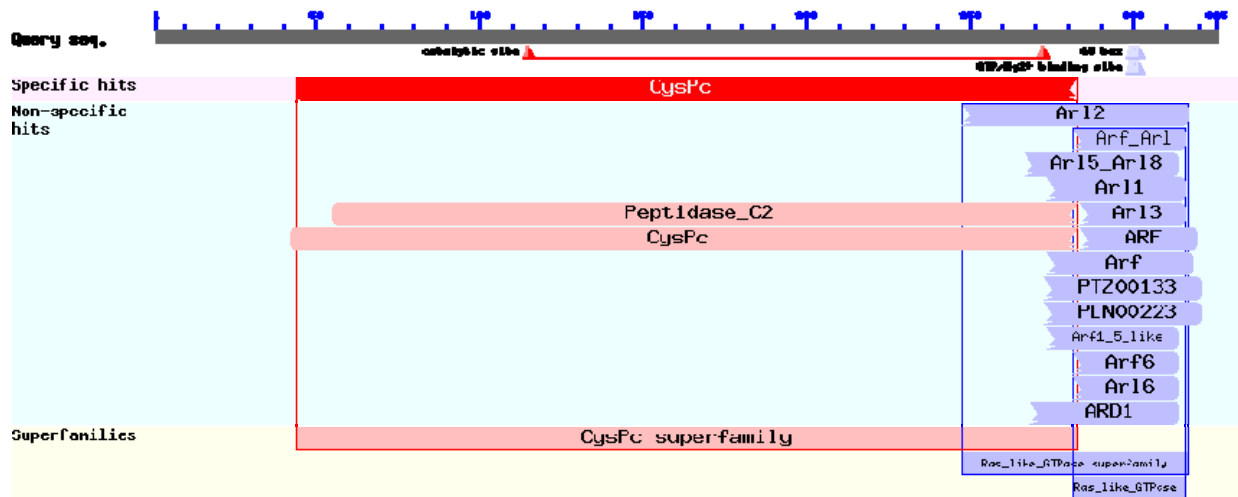
MEMKKKINLELRNRSPEEVELVDNCLCVNGEIEGLNDTFKELEFLSMANVELSSLARLPSLNK
LRKLELSDNIISGGLEVLAEKCPNLTYLNLSGNKIKDLSTVEALQNLKNLKSDFLNFCEITNLEDY
RESIFELLQQITYLDGFDQEDNEAPDSEEEDEDEDGDEDDEEEEEENEAGPPEGYE
DEDEDDEDEAGSELGEGEEVGLSYLMKEEIQDEEDDDDYVEEGEEEEEEEGEEEEGGGNV
EKDPDGAKSQEKEEPEISTEKEDSARLDDHEEEEEEEDEEPSHNEDHDADDEDSSHME
SAEVEKEELPRESFKEVLENQETFLDLNVQPGHSNPEVLMDCGVDLTASCNSEPKELAGDPEAVPESD
EPPPPGEQAQKQDQKNSKEVDTEFKEGNPATMEIDSETVQAVQSLTQESSEQDDTFQDCAET
QEACRSLQNYTRADQSPQIATTLDDCQQSDHSSPVSSVHSHPGQSVRSVNSPSVPALENSYA
QISPDQSAISVPSLQNMETSPMMDVPSVSDHSQQVVDSGFSDLGSIESTTENYENPSSYDSTM
GGICGNGSSQNSCSYSNLTSSSLTQSSCAVTQQMSNISGSCSMLQQTSSISSPPTCSVKSPQG
CVVERPPSSSQQLAQCMAANFTPPMQLAEIPETSNANIGLYERMGGQDFGAGHYPQPSATF
SLAKLQQLTNTLIDHSLPYSHSAAVTSYANSASLSTPLSNTGLVQLSQSPHSVPGGPQAQATMT
PPPNTLPPPMNLPPPLLQRNMAASNIGISHSQRLLTQIASKGHISMRTKSASLSPAAATHQSQIY
GRSQTVMQGPARTLTMQRGMNMSVNLMPAPAYNVNSVNMNMNTLNAMNGYSMSQPMMN
SGYHSNHGYMNQTPQYPMQMGMGMMGTQPYAQQPMQTPPHGNMMYTAPGHHGYMNTGM
SKQSLNGSYMRR*



CAPN1->ARL2

coding to coding fusion—aa 1-282 of CAPN1 fused in frame to aa 132-184 of ARL2.

MSEIITPVYCTGVSAQVQKQRARELGLGRHENAIKYLGGDYEQLRVRCVLSGTLFRDEAFPP
 VPQSLGYKDLGNSSKTYGIKWKRPTELLSNPQFIVDGATRTDICQGALGDCWLLAAIASLTLN
 DTLHRVVPHGQSFQNGYAGIFHFQLWQFGEWVDVVDDLLPIKDGKLVFVHSAEGNEFWSA
 LLEKAYAKVNGSYEALSGGSTSEGFEDFTGGVTEWYELRKAPSDLYQIILKALERGSLLGCSIDI
 SSVLDMEAITFKKLVKGHAYSVTGAKQVLELDSIRSHHWCIQGCSAVTGENLLPGIDWLLDDISS
 RIFTAD*



GPATCH8->C8orf46

coding to coding fusion—aa 16-64 of GPATCH8 fused in frame to aa 22-206 of C8orf46.

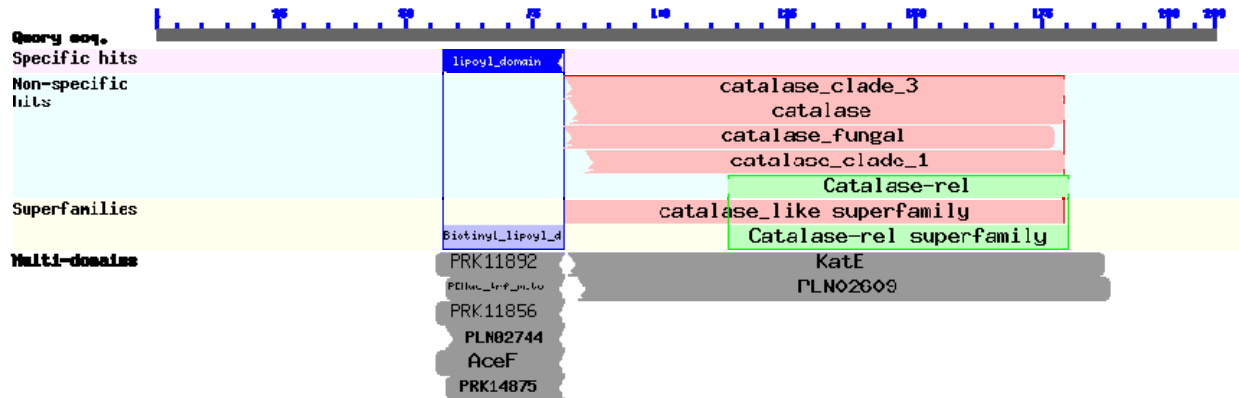
GNHFDQYEEGHLEIEQASLDKPIESDNIGHRLQKHGWKLGQGLGKSLQVSSPARRRAKSSQH
 LLTKNVVIESDLYTHQPLELLPHRGDRRDPGDRRRFGRLLQTARPPTAHPAKASARPVGISEPKT
 SNLCGNRAYGKSLIPPVPRISVKTSASASLEATAMGTEKGAVLMRGSRHLKMKMTEEYPALPQG
 AEASLPLTGSASCGVPGILRKMWTRHKKKSEYVGATNSAFEAD*



PDHX->CAT

coding to coding fusion—aa 1-83 of PDHX fused in frame to aa 99-228 of CAT

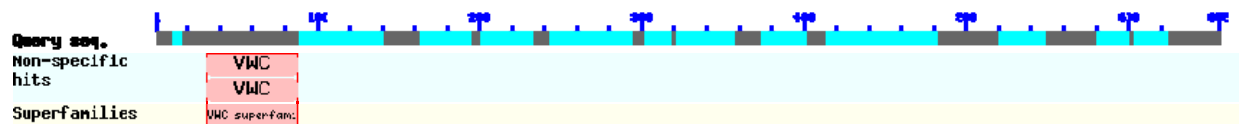
MAASWRLGCDPRLRLRYLVGFPGRRSVGLVKGALGWSVSRGANWRWFHSTQWLRGDPIKILM
PSLSPTMEEGNIVKWLKKEGGAPNYYPNSFGAPEQQPSALEHSIQYSGEVRRFNTANDDNVT
QVRAFVYVNVLNEEQRKRLCENIAGHLKDAQIFIQKKAVKNFTEVHPDYGSHIQALLDKYNAEKP
KNAIHTFVQSGSHLAAREKANL*



COL3A1->COL16A1

coding to coding fusion—aa 1-528 of COL3A1 fused in frame to aa 721-1186 of COL16A1.

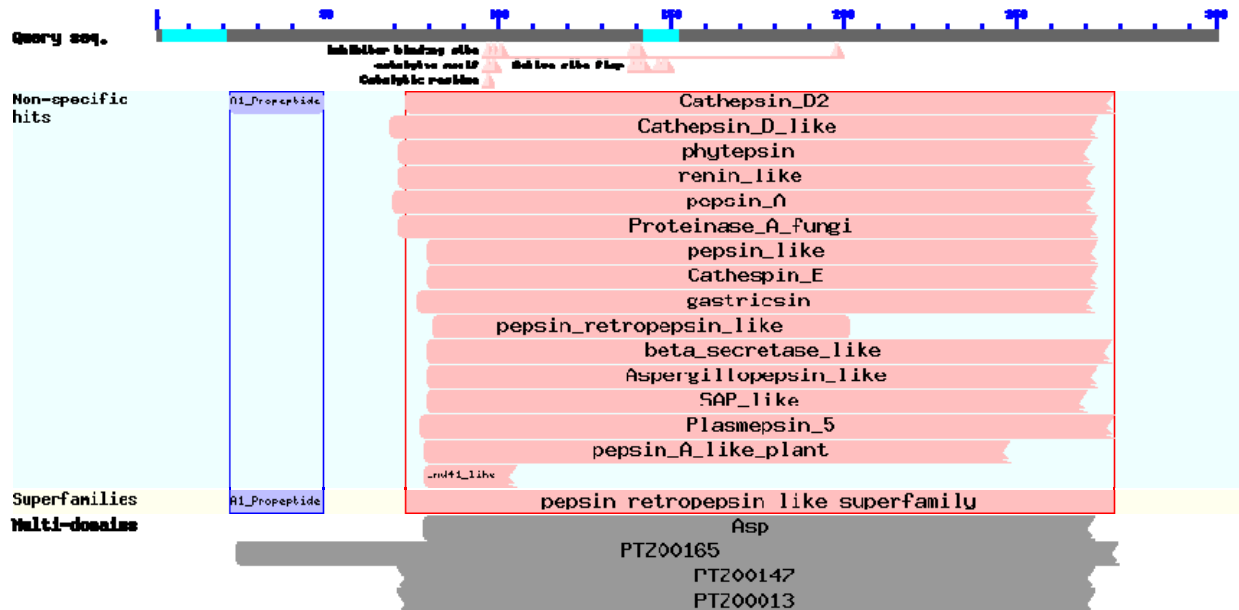
MMSFVQKGSWLLLALLHPTIILAQQEAVEGGCSHLGQSYADRDVWKPEPCQICVCDSGSVLCD
DIICDDQELDCPNPEIPFGECCAACPPTAPTRPPNGQGPQGPKGDPGPPGIPGRNGDPGIP
GQPGSPGSPGPPGICESCPTGPQNYSPQYDSYDVKSGVAVGGLAGYPGPA GLPGPPGPAGP
RGERGPQGNSEKGDQGFQGGPFPGPPGPPGFPKVGSPGPPGPQAEKGESEGIRGPSGL
PGSPGPPGPPGIQGPAGLDGLDGKDGKPLRGDPGPAGPPGLMGPPGFKGKTGHPGLPGPK
GDCGKPGPPGSTGRPGAEGEPGAMGPQGRPGPPGHVGGPPGPPGQPGPAGISAVGLKGDGRG
ATGERGLAGLPQQGPPGHPGPPGEPGTDGAAGKEGPPGKQGFYGGPPGPKGDPGAAGQKG
QAGEKGRAGMPGGPGKSGSMGPVGGPPGAGERGHPGAPGPGSGSPGLPGVPGSMGDMVNY
DEIKRFIRQEIIMFDERMAYYTSRMQFPMEMAAAPGRPGPPGKDGAPGRPGAPGSPGLPGQI
GREGRQGLPGVRGLPGTKGEKGDIGIGIAGENLPGPPGPQGGPPGYGKMGATGPMGQQGIP
GIPGPPGPMGQPGKAGHCNPSCFCGAMPMEQQYPPMKTMTKGPF*



CTSD->PRKAR1B

coding to coding fusion—aa 1-271 of CTSD fused in frame to aa 345-381 of PRKAR1B with a single aa insert.

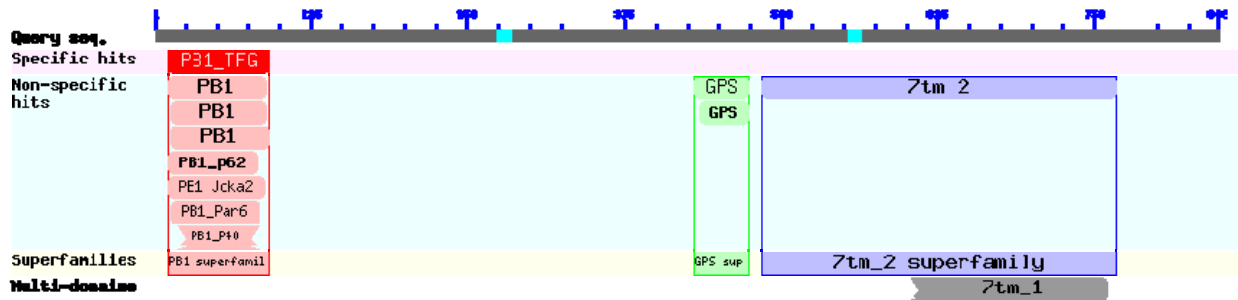
MQPSSLLPLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVPAVTEG
 PIPEVLKNYMDAQYYGEIGIGTPPQCFTVVFDTGSSNLWVPSIHCLLDIACWIHHKYNSDKSST
 YVKNGTSFDIHYGSGSLSGYLSQDTSVPCQSASSASALGGVKVERQVFGEATKQPGITFIAAK
 FDGILGMAYPRISVNNVLPVFDNLMQQLVDQNFYSYLSRDPDAQPGGELMLGGTDSKYKYG
 SLSYLNVTRKAYWQV **KCVKLD RPRFERVLGPCSEILKRNIQRYNSFISLTV***



TFG->GPR128

coding to coding fusion—aa 1-89 of TFG fused in frame in aa 39-797 of GPR128.

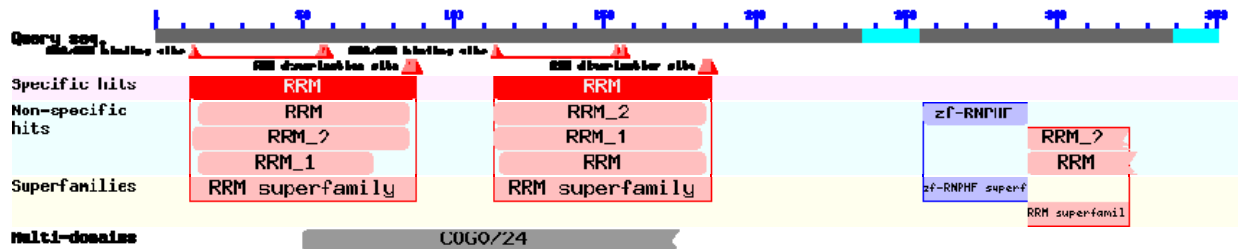
MNGQLDLGKLIKAQLGEDIRRIPIHNEDITYDELVLMQVRVFRGKLLSNDEVTIKYKDEDGDLI
 TIFDSSDLSFAIQCSRILKLTGKSTSSSSTPTEFCRNGGTWENGRICITEEWKGLRCTIANFC
 ENSTYMGFTFARIPVGRYGPSLQTCGKDTNAGNPMVRLCSLSLYGEIELQKVTIGNCNENLE
 TLEKQVKDVTAPLNNISSEVQILTSDANKLTAENITSATR VVGQIFNTRNASPEAKKVAIVTVSQ
 LLDASEDAFQRVAATANDDALTTLIEQMETSLSLGNQSVVEPNIAIQSANFSSENA VGPSNVR
 FSVQKGASSSLVSSSTFIHTNVDGLNPDAQTELQVLLNMTKNYTKTCGFVVYQNDKLFQSKTFT
 AKSDFSQKIISKTDENEQDQSASVDMVFPKYNQKEFQLYSYACVYWNLSAKDWDTYGCQK
 DKGTDGFLRCRCNHTTNFAVLMTFKKDYQYPKSLDILSNVGCALSVTGLALTVIFQIVTRKVRKT
 SVTWLVNLCISMLIFNLLFVFGIENSKNLQTS DGDINNIDFDNNDIPRTDTINIPNPMCTAIAALL
 HYFLLVTFTWNALSAAQLYLLIRTMKPLPRHFILFISLIGWGVPAIVVAITVGVVYSQNGNPNQW
 ELDYRQEKICWLAIPENGVIKSPLLWSFIVPVTIILSNVVMFITISIKVLWKNQNLSTKKVSSM
 KKIVSTLSVAVVFGITWILAYLMLVNDSDIRIVFSYIFCLFNTTQGLQIFILYTVRTRKVFQSEASKVL
 MLLSSIGRRKSLPSVTRPRLRVKMYNFLRSLPTLHERFRLLTSPSTEEITLSESDNAKESI*



HNRNPH1->VAPA

coding to coding fusion—aa 1-327 of HNRNPH1 fused in frame to aa 218-242 of VAPA with single aa insertion at the fusion junction.

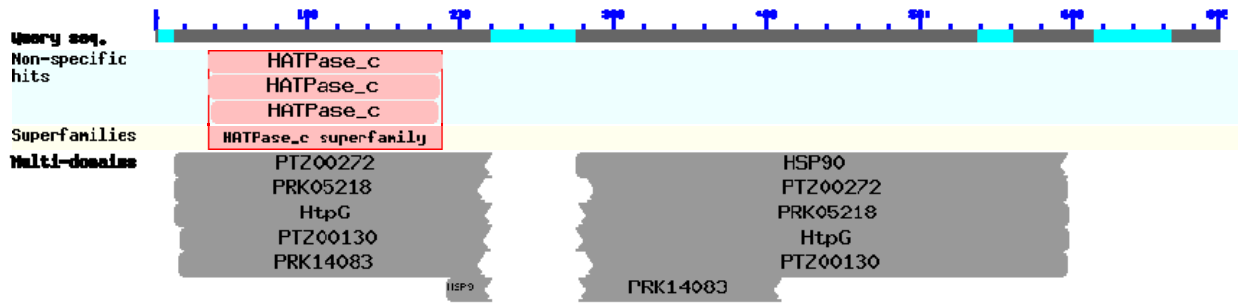
MMLGTEGGEGFVVKVRGLPWSCSADEVQRFFSDCKIQNGAQQGIRFIYTREGRPSGEAFVELE
 SEDEVKLALKKDRETMGHRYVEVFKSNNVEMDWLKHGTGPNSPDTANDGFVRLRGLPFGCSK
 EEIVQFFSGLIIVPNGITLPVDFQGRSTGEAFVQFASQEI AEKALKKHKERIGHRYIEIFKSSRAE
 VRTHYDPPRKLMMAMQRPQPYDRPGAGRGYNSIGRGAGFERMRRGAYGGGYGGYDDYNGYN
 DYGFGSDRFGRDLNYCFSGMSDHRYGDDGSTFQSTTGHCVHMRGLPYRATENDIYNFFSP
 LNPVRVHIEIGPDGRVIT**TSPLPSLLVVIAAIFIGFFLGKFIL***



HSP90AB1->PCGF2

coding to coding fusion—aa 1-599 of HSP90AB1 fused in frame to aa 241-344 of PDGF2

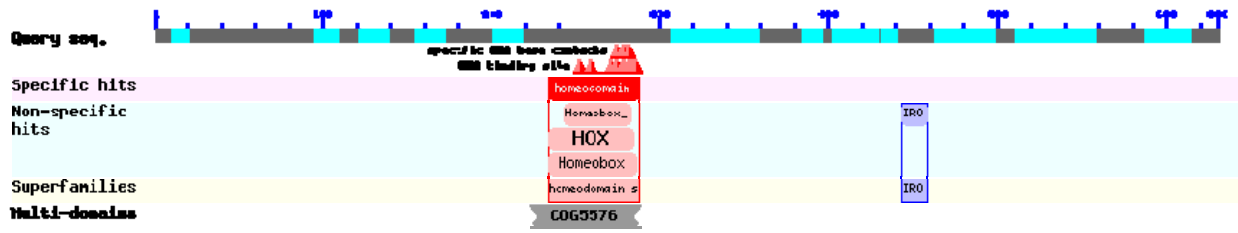
MPEEVHHGEEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNASDALDKIRYESLTDPSKLD
 GKELKIDIIPNPQERTLTLVDTGIGMTKADLINNLGTIAKSGTKAFMEALQAGADISMIGQFGVGF
 YSAYLVAEKVVVITKHNDDEQYAWESSAGGSFTVRADHGEPGRGTVKILHLKEDQTEYLEERR
 VKEVKKHSQFIGYPITLYLEKEREKEISDDEAEKEEEDKDDEEKPKIEDVGSDEEDDS
 GKDKKKKTKKIKEKYIDQEELNKTPIWTRNPDDITQEEYGEFYKSLTNDWEDHLAVKHFSVEG
 QLEFRALLFIPRRAPFDLFENKKKNNIKLYVRRVFIMDSCDELIPEYLNFRGVDSEDLPLNISR
 EMLQQSKILKVRKNIVKKCLELFSELAEDKENYKFFYEAFSKNLKLGIHEDSTNRRRSELLRYH
 TSQSGDEMTSLSEYVSRMKETQKSIYYITGESKEQVANSFAVERVRKRGFEVVMTEPIDEYC
 VQQLKEFDGKSLVSVTKEGLELPEDEEEKKMEESKAKFENLCKLMKEILDKKVEKVTISNRLV
 SSPCCIVTSTYEGTNTSGASECESVSDKAPSPATLPATSSSLPSPATPSHGSPSSHGPPATHPT
 SPTPPSTASGATTAANGGSLNCLQTPSSTSRRGRKMTVNGAPVPPLT*



USF2->IRX3

coding to coding fusion—aa 1-124 of USF2 fused in frame to aa 28-501 of IRX3

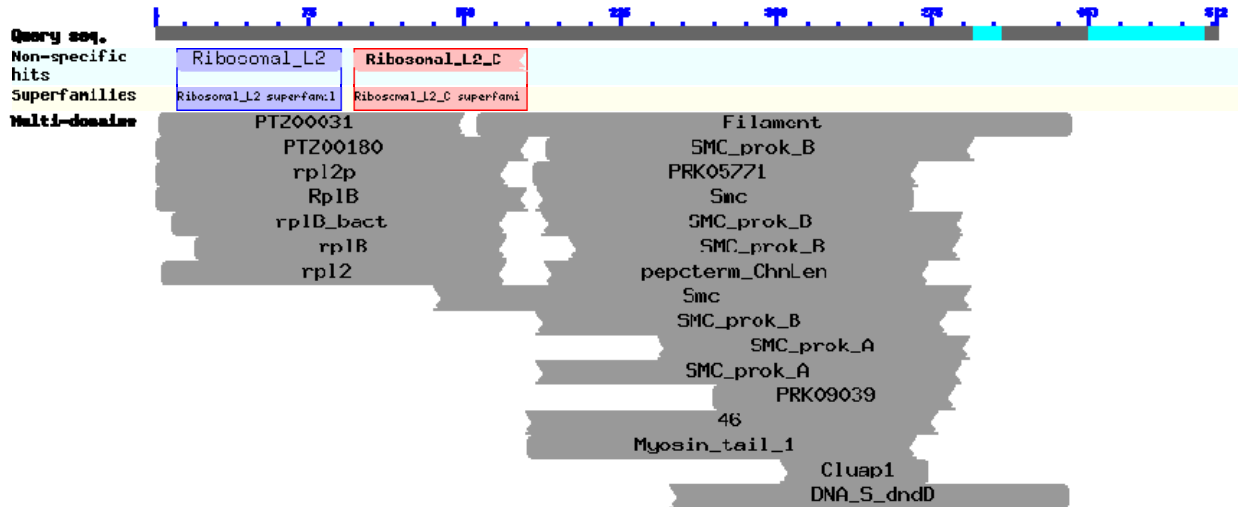
MDMLDPGLDPAASATAAAAASHDKGPEAEEGVELQEGGDGPGAEEQTAVAITSVQQAFFGDH
 NIQQFRFTETNGGQVYTRVVQVTDGQLDGQDGTAGAVSVVSTAAFAGGQQAVTQVGVGDGAA
 QRGSAGARGGLGAGASELNASGSLSNVLSVYGAPYAAAAAAAAAQQGYGAFLPYAAELPIFPQ
 LGAQYELKDSPGVQHPAAAAAFPHPHPAFYQYQFGDPSRPKNATRESTSTLKAWLNEHR
 KNPYPKGEKIMLAITKMTLTQVSTWFANARRRLKKNKMTWAPRSRTDEEGNAYGSEREE
 DEEDEEDGKRELELEEEELGGEEDTGGGLADDEDEEIDLENLDGAATEPELSLAGAARR
 DGDGLGPISDSKNSDSEDSSEGLEDRPLPVLAPAPPVAVASPSLPSPPVSLDPCAPAPAP
 ASALQKPKIWSLAETATSPDNPRRSPPGAGGSPPGA AVAPSALQLSPAAAAAAHRLVSAPLG
 KPAWTNRPFPGPPPGRHLPLSLLGSAPPHLLGLPGAAGHPAAAAAFARPAEPEGGTDRCS
 ALEVEKLLKTAQVPRRPQNHLDALVLSALSS*FFKKNKTKKTTFFNRCNNCIKKIALYSY
 NL*



RPL8->KRT4

coding to coding fusion—aa 1-167 of RPL8 fused in frame to aa 67-408 of KRT4 with a single aa insertion at the fusion junction.

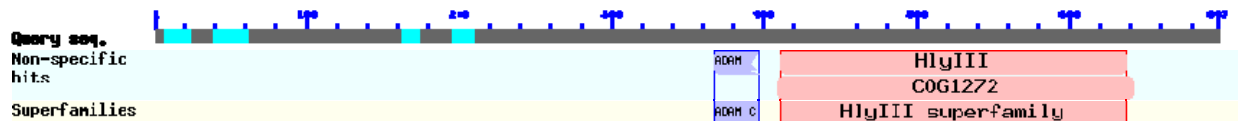
MGRVIRGQRKGAGSVFRAHVKHRKGAARLRAVDFAERHGYIKGIVKDIHDPGRGAPLAKVVF
 RDPYRFFKRTIELFIAAEGIHGTGQFVYCGKKAQLNIGNVLPVGTMPGEGTIVCCLEEKPGDRGKLA
 RASGNATVISHNPETKKTRVKLPSGSKKVISSANRAVVGWATSSKNLEPLFETYLSVLRKQL
 DTLGNDKGRQLQSELKTMQDSVEDFKTKYEEEINKRTAAENDFVVLKKDQVDAAYLNKVELEAKV
 DSLNDEINFLKVLYDAELSQQMTHVSDTSVVL SMDNNRNLDLDSIIAEVRAQYEEIAQRSKAEAE
 ALYQTKVQQLQISVDQHGDNLKNTKSEIAELNRMIQRLRAEIEIENIKKQCQTLQVSVADAEQRGE
 NALKDAHASKRVELEAALQQAKEELARMLREYQELMSVKLALDIEIATYRKLLEGEYRMSGECQ
 SAVSISVVGSTSTGGISGGLGSGSGFGLSSGFGSGSGSGFGFGGSVSGSSSSKIISTTTLNKR
 R*



THSD4->PAQR5

coding to coding fusion—aa 1-384 of THSD4 fused in frame to aa 18-324 of PAQR5 (with numerous mismatches?)

MVSHFMGSLSVLCFLLLLGFQFVCPQPSTQHRKVPQRMAAEGAPEDDGGGGGAPGVWGAWG
 PWSACSRSCSGGVMEQTRPCLPRSYRLRGGQRPGAPARAFADHVVS AVR TSVPLHRSRDET
 PALAGTDASRQGPTVLRGSRHPQPQGLEVTGDRRSRTRGTIGPGKYGYGKAPYILPLQDTAH
 TPQRLRRQKLSSRHSRSQGASSARHGYSSPAHQVPQHGPLYQSDSGPRSGLQAAEAPIYQLP
 LTHDQGYPAASSLFHSPETSNNHGVGTHGATQSFSQPARSTAISCIGAYRQYKLCNTNVCPE
 SRSIREVQCASYNKPFMGRFYEWEPFAEVKGNRKCELNCQAMGYRFYVRQAEKVIDGTPCD
 QNGTAICVSGQCKVFHEQGILFGYRHPQSSATACILSLFQMTNETLNIWTHLLPFWFFAWRFVT
 ALYMTDIKNDSSWPMLVYMCTSCVYPLVSSCAHTFSSMSKNARHICYFLDYGAVNLFSLGSAI
 AYSAYTFPDALMCTTFHDYYVALAVLNTILSTGLSCYSRFLEIQKPRCLKVIRVLAFAYPYTWDS
 LPIFYRLFLFPGESAQNEATS YHQKHMIMTLLASFLYSAHLPERLAPGRFDYIGHSHQLFHVCVI
 LATHMQMEAILLDKTLRKEWLLATSKPFSFSQIAGAAILLCIIFSLSNIIYFSAALYRIPKPELHKKET*

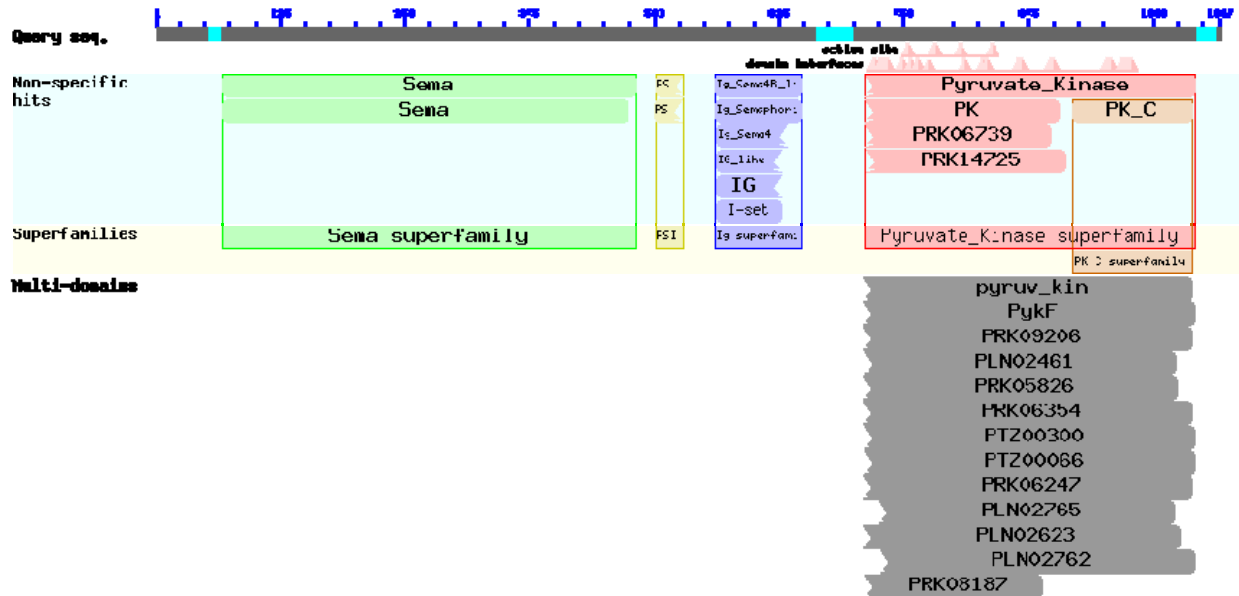


SEMA4C->PKM2

coding to coding fusion—aa 1-710 of SEMA4C fused in frame to aa 13-343 of PKM2

MAPHWAVWLLAARLWGLGIGAEVWWNLVPRKTVSSGELATVVRFRFSQTGIQDFLTLTLTEPT
 GLLYVGAREALFAFSMEALELQGAISWEAPVEKKTECIQKGNQTECFNFIRFLQPYNASHLY
 VCGTYAFQPKCTYVNMLTFTLEHGEFEDGKKGKCPYDPAKGHAGLLVDGELYSATLNNFLGTEP
 IILRNMGPHHSMKTEYLAFWLNEPHFVGSAYVPESVGSFTGDDDKVYFFFRERAVESDCYAEQ
 VVARVARVCKGDMGGARTLQRKWTTFLKARLACSAPNWQLYFNQLQAMHTLQDTSWHNTTF
 FGVFQAQWGDMYLSAICEYQLEEIQRVFEGPYKEYHEEAQKWDRYTDPVPSRPGSCINNWH
 RRHGYTSSLELPDNLNFVKKHPLMEEQVGRWSRPLL VKKGTNFTHLVADRVTGLDGATYTV

LFIGTGDGWLLKAVSLGPVWHLIEELQLFDQEPMRSLVLSQSKLLFAGSRSQLVQLPVADCM
 KYRSCADCVLARDPYCAWSVNTSRCVAVGGHSGSLLIQHVMTSDTSGICNLRGSKKVRPTPK
 NITVVAGTDLVLPCHLSSNLAHARWTFGGRDLPAEQPGSFLYDARLQALVMAAQPRHAGAY
 HCFSEEQGARLAAEGYLVAVVAGPSVTLERAPLENLGLVWLAVVALGAVCLVLLLLVLSLRRR
 LREELEKGAKATERTLVYPLEGSLGSKKGVNLPAAVDLPAVSEKDIQDLKFGVEQDVMVFA
 SFIRKASDVHEVRKVLGEKGNIKIISKIENHEGVRRFDEILEASDGIMVARGDLGIEIPAQVFLA
 QKMMIGRCNRAKGPVICATQMLESMIKKPRPTRAEGSDVANAVLDGADCIMLSGETAKGDYPL
 EAVRMQHIAIEAEAAIYHLQLFEELRRLAPITSDPTEATAVGAVEASFKCCSGAIIVLTKSGRSA
 HQVARYRPRAPIAVTRNPQTARQAHLRGIFPVLCKDPVQEAWAEDVDLRVNFAMNVGKARG
 FFKKGDVVIVLTGWRP GSGFTNTMRVVPV*

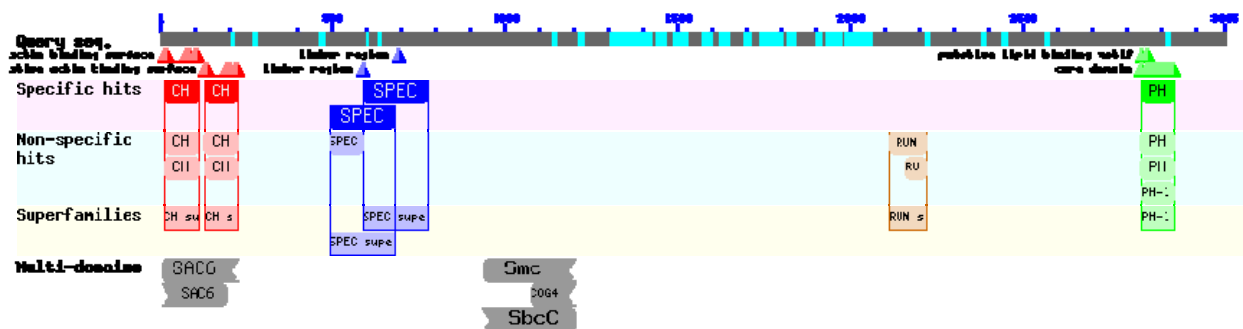


PLEC->PLEKHM2

Coding of PLEC fused to 5' of PLEKHM2—aa 1-2049 of PLEC fused in frame to aa 61-1095 of PLEKHM2

MKIVPDERDRVQKKTFTKWVNKHLIKAQRHISDLYEDLRDGHNLISLLEVLSGDSLPREKGRMR
 FHKLQNVQIALDYLRHRQVKLVNIRNDDIADGNPKLTGLIWTIILHFQISDIQVSGQSEDMTAKE
 KLLLWSQRMVEGYQGLRCDNFTSSWRDGRFLFNAIIHRHKPLLIDMNKVYRQTNLENLDQAFSV
 AERDLGVTRLLDPEDVDVPQPDEKSIITYVSSLYDAMPVPDVQDGV RANELQLRWQEYRELV
 LLLLQWMRHHTAAFEERRFPSSFEEIEILWSQFLKFKEMELPAKEADKNRSKGIYQSLEGAVQA
 GQLKVPPGYHPLDVEKEWGLHVAILEREKQLRSEFERLECLQRIVTKLQMEAGLCEEQLNQA
 DALLQSDVRLAAGKVPQRAGEVERDLKADSMIRLLFNDVQTLKDGRHPQGEQMYRRVYRL
 HERLVAIRTEYNLRLKAGVAAPATQVAQVTLQSVQRRPELEDSTLRYLQDLLAWVEENQHRVD
 GAEWGVDLPSVEAQLGSHRGLHQSIEEFRAKIERARSDEGQLSPATRGAYRDCLGRDLQYA
 KLLNSSKARLRSLESLSFVAAATKELMWLNEKEEEVGFSDRNTNMTAKKESYSALMREL
 ELKEKKIKELQNAGDRLLREDHPARPTVESFQAALQTQWSWMLQLCCIEAHLKENAAYFQFF
 SDVREAEGQLQKLQEALRRKYSCDRSATVTRLEDLLQDAQDEKEQLNEYKGHLSGLAKRAKA
 VVQLKPRHPAHPMRGRPLLA VCDYKQVEVTVHKGDECQLVGPAQPSHWKVLSSSGSEAAV
 PSVCFLVPPPNQEAQEA VTRLEAQHQALVTLWHQLHVDKMSLLAWQSLRRDVLIRSWSLAT
 FRTLKPEEQRQALHSLELHYQAFLRDSQDAGGFGPEDRLMAEREYGSCSHHYQQLLQSLEQG
 AQEESRCQRCISELKDIRELQLEACETRTVHRLRLPLDKEPARECAQRIAEQQAQAEVEGLGK

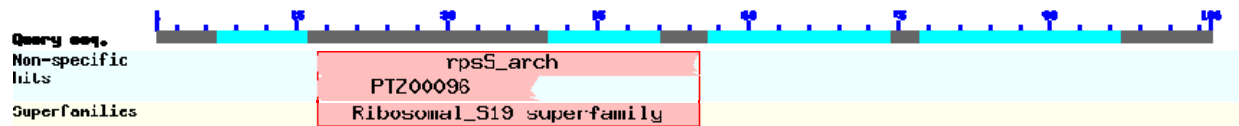
GVARLSAEAEKVLALPEPSPAAPT LRSELELT LGKLEQVRSLSAIYLEKLTISLVIRGTQGAEV
 LRAHEEQLKEAQAVPATLPELEATKASLKKLRAQAEAAQPTFDALRDELGAQEVGERLQQRH
 GERDVEVERWRERVAQLLERWQAVLAQTDVRRQRELEQLGRQLRYYRESADPLGAWLQDAR
 RRQEQIQAMPLADSQAVREQLRQE QALLEEIERHGEKVEECQRFAKQYINAIDYELQLVTYKA
 QLEPVASPAKKPKVQSGSESVIQEYVDLRTHYSEL TLT SQYIKFIS ETLRRMEEEEERLAEQQRA
 EERERLAEVEAALEKQRQLAEAHAQAQAEREAKELQQRMQEEVVRREEAAVDAQQQKRSI
 QEELQQLRQSSEAEIQAKARQAEAAERSRLRIEEEEIRVVRLQLEATERQRGGAEGELQALRAR
 AEEAEAQKRQAQEEAERLRRQVQDESQRKRQAEVELASRVKAEAEAAAREKQRALQALEELRL
 QAEEAERRLRQAEVERARQVVALETAQRSAEAE LQSKRASFAEKTAQLERSLQEEHVAVAQ
 LREEAERRAQQQAEAEERAREEAERELERWQLKANEALRLRLQAE EVAQQKSLAQAEAEKQKE
 EAEREARRRGKAEQAVRQRELAEQELEKQRQLAEGTAQQRLAAEQELIRLRAETE QGEQQR
 QLLEELARLQREAAAATQKRQELEAE LAKVRAEMEVLLASKARAEESRSTSEKSKQRLEAE
 AGRFRELAE EAARLRALAE EAKRQRQLAEEDAAQRQAEAEVLAEKLA AIGEATRLKTEAEIAL
 KEKEAENERLRLLAEDEAFQRRRLEEQA AQHKADIEERLAQLRKASDSELERQKGLVEDTLRQ
 RRQVEEELALKASFEKAAAGKAELELELGRIRSNAEDTLRSKEQAELEAARQRQLAAEEERRR
 REAEERVQKSLAE EEAARQRKAAL EEVERLKAKVEEARRLRERAEQESARQLQLAQEAAQK
 RLQAE EKAHAFVQQKEQELQQT LQQEQSVLDQLRGE GGGGGGGGGGGG DGGSSAMEPGEV
 KDRILENISL SVKKLQSYFAACEDEIPAIRNHDKVLQRLCEHLDHALLYGLQDLSSGYWVLVWHF
 TRREAIKQIEVLQHVATNLGRSRAWLYLALNENSLESYLRLFQENLGLLHKYYVKNALVCSHDH
 LTLFLT VSGLEFIRFELDL DAPYLDLAPYMPDYYPQYLLDFEDRLPSSVHGSDSLSLNSFNSV
 TSTNLEWDDSAIAPSS EYDFGDVFP AVPSVPSTDWEDGDLDTVSGPRSTASDLTSSKASTR
 SPTQRQNPFN EEP AETVSSSDTTPVHTTSQEKEEAQALDPPDACTELEVIRVT KKKKIGK KKKKS
 RSDEEASPLHPAC SQKKCAKQGDGDSRNGSPSLGRDSPDTMLASPQEEGEGPSSTTESSER
 SEPGLLIPEMKDTSMERL GQPLSKVIDQLNGQLDPSTWCSRAEPPDQSFRTGSPGDAPERPPL
 CDFSEGLSAPMDFYRFTVESPSTVTSGGGHDPAGLGQPLHVPSSPEAAGQEEEGGGGEGQ
 TPRPLEDTTREAQELEAQLSLVREGPVSEPEPGTQEVLCQLKRDQPSPCLSSAEDSGVDEGQ
 GSPSEMVHSSEFRVDNNHLLLLMIHVFRENEEQ LFKMIRMSTGHMEGNLQLLYVLLTDCYVYLL
 RKGATEKPYLVEEAVSYNELDYVSVGLDQQT VKL VCTNRRKQFLD TADVALAEFFLASLKSA
 MIKGCREPPYPSILT DATMEKLALAKFVAQESKCEASAVTVRFYGLVHWEDPTDES LGPTPCH
 CSPPEGTITKEGMLHYKAGTSYL GKEHWKTCFVVL SNGILYQYPDR TDVIPLL SVNMGGEQCG
 GCRRANTTDRPHAFQVILSDRPCLELSAESEAEMA EWMLCQAVSKGVIPQGVAPSPCIPCC
 LVLTD DRLFTCH EDCQTSFFRSLGTAKLGDISAVSTEPGKEYCVLEFSQDSQQLLPPWVIYLS
 TSELDRLLSALNSGWKTIYQVDLPHTAIQEASNKKKFEDALSLIHS AWQRSDSLRCRGRASRDP
 WC*



RPS15->PLEC

coding to coding fusion—aa 1-73 of RPS15 fused in frame to aa 4554-4574 of PLEC with a single aa insertion at the fusion junction.

MAEVEQKKKRTFRKFYRQVLDQLLDMSYEQLMQLYSARQRRRLNRGLRRKQHSLKRLRK
 AKKEAPPMEKPE^ETFSSSSYSSSSGYGRRYASGSSASLGGPESAVA*



PROM1->TAPT1

coding to coding fusion—aa 1-497 of PROM1 fused in frame to aa 439-567 of TAPT1 with a **single** aa insertion at the fusion junction.

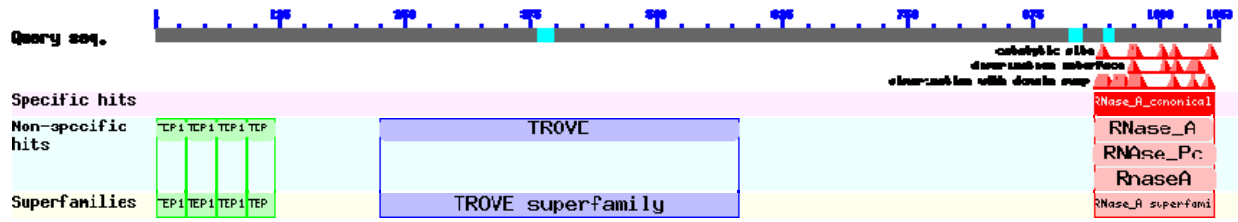
MALVLGSLLLLGLCGNSFSGGQPSSTDAPKAWNYELPATNYETQDSHKAGPIGILFELVHIFLYV
 VQPRDFPEDTLRKFLQKAYESKIDYDKPETVILGLKIVYYEAGIILCCVLGLLFIILMPLVGYFFCM
 CRCCNCKGGEMHQRQKENGPFRLKCFAILLVICIIISIGIFYGFVANHQVTRIKRSRKLADSNF
 KDLRTLLNETPEQIKYILAQYNTTKDKAFTDLNSINSVLLGGGILDRLRPNIIPVLDEIKSMATAIKET
 KEALENMNSTLKSLLHQQSTQLSSSLTSVKTSRSSLNDPLCLVHPSSETCNSIRLSLSQLNSNP
 ELRQLPPVDAELDNVNNVLRDLDGLVQQGYQSLNDIPDRVQRQTTTAVAGIKRVLNSIGSDID
 NVTQRLPIQDILSAFSVYVNNTESYIHRNLPTELEYDSYWWLGGGLVICSLLLIVIFYLLGLLCGV
 CGYDRHATPTTRGCVSNTGGVFLM^VLISLKVLSIVLLGKSCQYVKEAKMEEKLSNPPATCTPG
 KPSSKSQNKCKPSQGLSTEENLSASITKQPIHQKENIIPLLVTSNSDQFLTTPDGDEKIDITQDNS
 ELKHRSSKKDLLEIDRFTICGNRID*



TEP1->RNASE1

coding region of TEP1 fused in front of the 5' UTR of RNASE1—aa 1-895 of TEP1 fused in frame to aa 1-156 of RNASE1 with **9 aa insertion** at fusion junction.

MEKLGHVSAHPDILSLENRCLAMLPDLQPLEKLHQHVSTHSDILSLKNQCLATLPDLKTMEKP
 HGYVSAHPDILSLENQCLATLSDLKTMEKPHGHVSAHPDILSLENRCLATLSSLKSTVSASPLFQ
 SLQISHMTQADLYRVNNSNCLLSEPPSWRAQHFSKGLDLSTCPIALKSISATETAQEATLGRWF
 DSEKKGAEQMPYSLSLGEVEEDLAVKLTSGDSESHPEPTDHVLQEKKALLSLLCSTL
 VSEVMNNTSDPTLAAIFEICRELALLEPEFILKASLYARQQLNVRNVANNILAAFLPACRPHL
 RRYFCAIVQLPSDWIQVAELYQSLAEGDKNKLVPLPACLRMTAMDKFAQFDEYQLAKYNPRKH
 RAKRHPRRPPRSPGMEPPFSHRCFPRYIGFLREEQRKFEKAGDTVSEKKNPPRFTLKKLVQRL
 HIHKPAQHVQALLGYRYPNSLQLFSSRSLPGPWDSSRAGKRMKLSRPETWERELSLRGNKAS
 VWEELIENGKLPFAMLRNLCNLLRVGISSRHHELILQRLQHAHSVHSRQFPFRFLNAHDAIDA
 LEAQLRNQALPFPSNITLMRRILTRNEKNRPRRFLCHLSRQQLRMAMRIPVLYEQLKREKLRV
 HKARQWKYDGEMLNRYRQALETAVNLSVKHSLPLLPGRVTVLVYLTANADRLCPKSNPQGGP
 LNYALLIGMMITRAEQVDVVLGGDTLKTAVLKAEEGILKTAIKLQAQVQEFDENDGWSLNTFG
 KYLLSLAGQRPVDRVILLGQSMDDGMINVAKQLYWQRVNSKCLFVGILLRRVQYLSTDLNPN
 DVTLSGCTDAILKFAIEHGASHLLEHVGQMDKIFKIPPPGKTGVQSLRPLEEDTPSPLAPVSQQ
GLFWESEATMALEKSLVRLLLLVLILLVLGWVQPSLGKESRAKKFQRQHMDSDSSPSSSSTYC
 NQMMRRRNMTQGRCKPVNTFVHEPLVDVQNVCFQEKVTCKNGQGNCYKSNSSMHITDCRLT
 NGSRYPNCAVRTSPKERHIIVACEGSPYVPVHFDASVEDST*

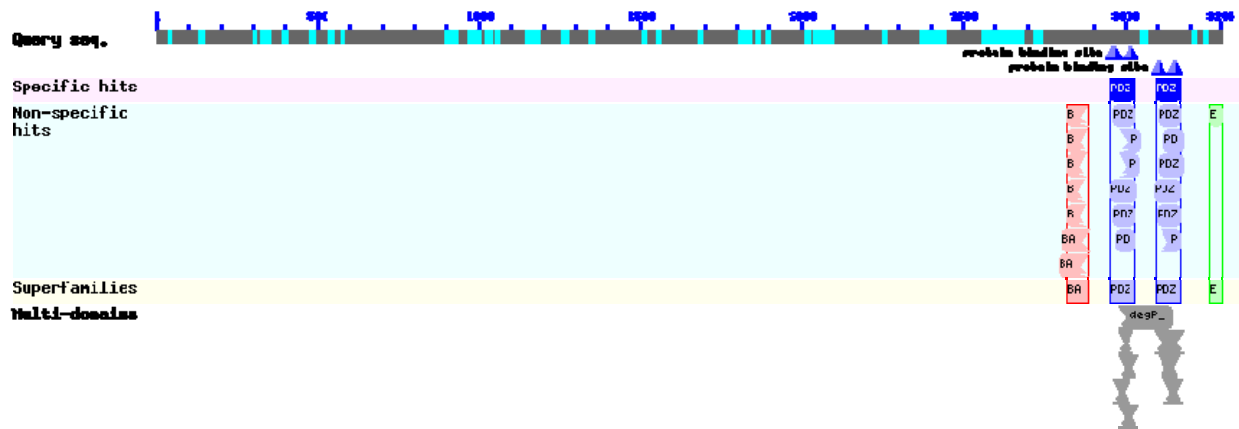


TNRC18->SLC9A3R1

Coding of TNRC18 fused to 5' of SLC9A3R1—aa 1-2884 of TNRC18 fused in frame to aa 20-428 of SLC9A3R1.

MDGRDFGPQRSVHGPPPLLSGLAMDSHRVGAATAGRLPASGLPGPLPPGKYMAGLNLHPH
PGEAFLGSFVASGMGPSASSHGSPVPLPSDLFRSPTPSNLPMVQLWAAHAHEGFSHLPSGL
YPSYLHLNHLEPPSSGSPLLSQLGQPSIFDTQKGQGGPGGDFYLPTAGAPGSLHSHAPSARTP
GGGHSSGAPAKGSSSRDGPAKERAGRGGEPPPLFGKKDPRARGEESGPRGVVDLTQEARA
EGRQDRGPPRLAERLSPFLAESKTKNAALQPSVLTMCNGGAGDVGLPALVAEAGRGGAKEAA
RQDEGARLLRRTETLLPGPRPCPSPLPPPAPPKGPAPPAAATPAGVYTVFREQGREHRVVA
PTFVPSVEAFDERPGPIQIASQARDARAREREAGRPGVLQAPPGSPRPLDRPEGLREKNSVIR
SLKRPPPADAPTVRATRASPDPRAYVPAKELLKPEADPRPCERAPRGPAGPAAQQAAKLFGLE
PGRPPPTGPEHKWKPFELGNFAATQMAVLAAQHHSRAEEEEAAVVAASSSKKAYLDPGAVLP
RSAATCGRPVADMHSAAHGSGEASAMQSLIKYSGSFARDAVAVRPGGCGKKSPPFGGLGTMK
PEPAPTSAGASRAQARLPHSGGPAAGGGRQLKRDPERPESAKAFGREGSGAQGEAEVRHPP
VGIAVAVARQKDSGGSGRLGPGLVDQERSLSLSNVKKGHGRADEDCVDDRARHREERLLGARL
DRDQEKLLRESKELADLARLHPTSCAPNGLNPNLMVTGGPALAGSGRWSADPAAHLATHPWL
PRSGNASMWLAGHPYGLGPPSLHQMAPAFPPGLGGSLSAYQFVRDPQSGQLVVIPSDHL
PHFAELMERATVPPLWPALYPPGRSPLHHAQQLQLFSQQHFLRQQEFLYLQQQAAQALELQR
SAQLVQERLKAQEHRAMEEKGSKRGLEAAGKAGLATAGPGLLPRKPPGLAAGPAGTYGKAV
SPPSPRASPVAAKAKVIQKLEDVSKPPAYAYPATPSSHPTSPPPASPPPTPGITRKEEAPEN
VVEKKDLELEKEAPSPFQALFSDIPPRYPFQALPPHYGRYPFLLQPTAADADGLAPDVPLPA
DGPERLALSPEDKPIRLSPSKITEPLREGPEEEPLAEREVKAIVEDMDEGPTLPPLESPLPLPA
AEAMATPSPAGGCGGLLEAQUALSATGQSCAEPSECPDFVEGPEPRVDSPGRTEPCTAALDL
GVQLTPETLVEAKEEPEVVPVAVPVVEAVPEEGLAQVAPSESQPTLEMSDCDVPAGEGQCPS
LEPQEAVPVLGSTCFLEEASSDQFLPSLEDPLAGMNALAAAELPQARPLPSPGAAGAQALEK
LEAAESLVLEQSFLHGITLLSEIAELELERRSQEMGGAERALVARPSLESLLAAGSHMLREVLDG
PVVDPLKNLRLPRELKPNKKYSWMRKKEERMYAMKSSLEDMDALELDFRMRLAEVQRQYKEK
QRELVKLQRRRDSEDRREEPHRSARRGPRPRKRTHAPSALSPPRKRGKSGHSSGKLSSK
SLLTSSDYELGAGIRKRHKGSEEEHDALIGMGKARGRNQWDEHEASSDFISQLKIKKKKMAS
DQEQLASKLDKALSLTKQDKLKSPFKFSDSAGGKSKTSGGCGRYLTPYDSLLGKNRKALAKGL
GLSLKSSREGKHKRAAKTRKMEVGFKARGQPKSAHSPFASEVSSYSYNTDSEEDDEEFLKDEW
PAQGPSSSKLTPSLLCSMVAKNSKAAGGPKLTRGLAAPRTLKPKPATSRKQPFCLLLREAEA
RSSFSDSSEESFDQDESSEEEDEEEDEEASGGGYRLGARERLSPGLEESGLGLLARFA
ASALPSPVGPVSVVQLEAKQKARKKEERQSLLGTEFEYTDSESEVKVRKRSPAGLLRPKKG
LGEPGPSLAAPTPGARGPDPSSPKAKLAVEKGRKARKLRGPKEPGFEAGPEASDDDLWTRR
RSERIFLHDASAAAPAPVSTAPATKTSRCAKGGPLSPRKDAGRAKDRKDPRKGGGKEAGPG
AGLPPPRAPALPSEARAPHASLTAARKSKAKAKGKEVKKENRGKGGAVSKLMESMAAEEDF
EPNQDSSFSEDEHLPRGGAVRPLTPAPRSCIIDKDELKDGLRVLIPMDDKLLYAGHVQTVHSP
DIYRVVVEGERGNRPHIYCLEQLLQEAIIDVRPASTRFLPQGTRIAAYWSQQYRCLYPGTVVRG
LLDLEDDGDLITVEFDDGDTGRIPLSHIRLLPPDYKIQCAEPPALLVPSAKRRSRKTSKDTGEG
KDGGTAGSEEPGAKARGRGRKPSAKAKGDRAATLEEGNPTDEVPSTPLALEPSSTPGSKKSP

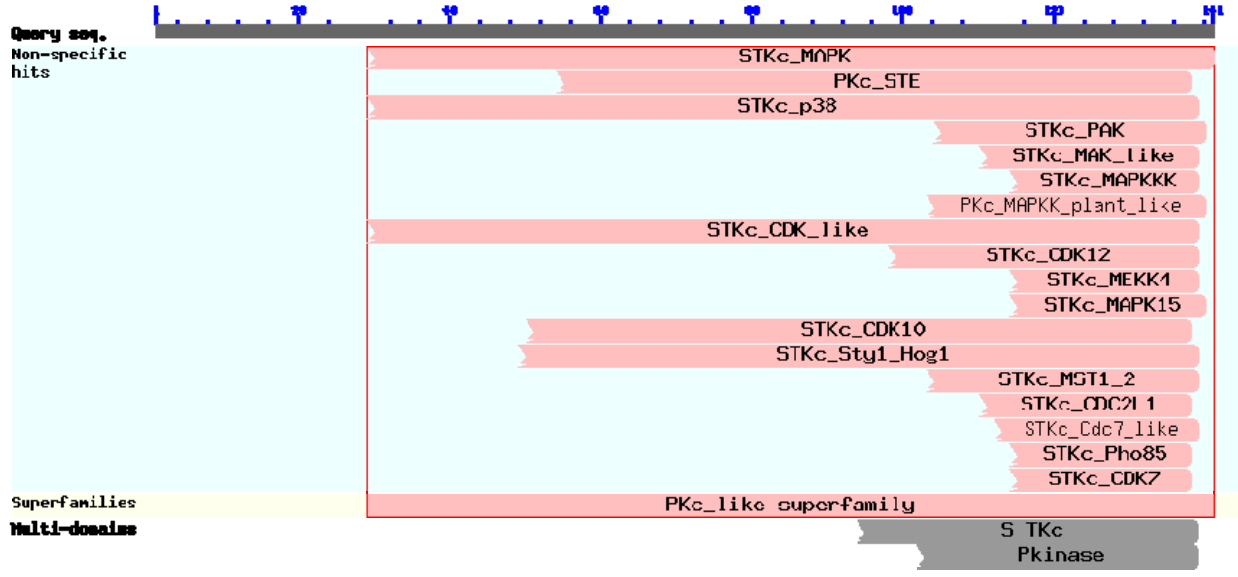
PEPVDKRAKAPKARPAPPQPSPAPPAFTSCPAPEPFAELPAPATSLAPAPLITMPATRPKPKKA
 RAAEESGAKGPRRPGEAAELLVKLDHEGVTSPKSKKAKEALLLREDPGAGGWQEPKSLLSLG
 SYPPAAGSSEPKAPWPKATDGDLAQEPGPGLTFEDSGNPKSPDKAQAEQDGAEESESSSSS
 SSGSSSSSSSSSSSSGSETEGEEEGDKNGDGGCGTGGRNCSAASSRAASPASSSSSSSSSSS
 SSTTDEDSSCSDDEAAP
 APTAGPSAQAALPTKATKQAGKARPSAHSPGKKTAPQPQAPPPQPTQPLQPKAQAGAKSRP
 KKREGVHLPTTKELAKRQRLPSVENRPKIAAFLPARQLWKWFGKPTQRRGMKGKARKLFYKAI
 VRGKEMIRIGDCAVFLSAGRPNLPYIGRIQSMWESWGNMNVVRVKWIFYHPEETSPGKQFHQG
 QHWDQKSSRSPTVPGTPACLARPAAQGFSALPVRWTGRRAGPSRPVPIGTPSRAADPSQG
 EMSADAAAGAPLPRCLCCLEKGPNGYGFHLHGEKGLGQYIRLVEPGSPAEEKAGLLAGDRLVEV
 NGENVEKETHQQVVSIRAAALNAVRLLVDPETDEQLQKLGQVQVREELLRAQEAPGQAEPPAA
 AEVQGAGNENEPREADKSHPEQRELRPRLCTMKKGPSGYGFNLHSDKSKPGQFIRSVDPDSP
 AEASGLRAQDRIVEVNGVCMEGKQHGDVVS AIRAGGDETLLVVDRETDEFFKKCRVIPSQEHL
 LNGPLVPFTNGEIQKENSREALAEAALESPPALVRSASSDTSEELNSQDSPPKQDSTAPSS
 SSSDPILDFNISLAMAKERAHQKRSSKRAPQMDWSKKNELFSNL*



UBR2->SRPK1

Coding region of UBR2 fused to 3' UTR of SRPK1—aa 1-26 of UBR2 fused in frame to aa 434-548 of SRPK1

MASELEPEVQAIDRSLLLECSAEEIAGAFELATGDYLFEPHSGEEYTRDEDHIALIIELLGKVPRKLI
 VAGKYSKEFFTKKGDCLKHITKLPWGLFEVLVEKYEWSQEEAAGFTDFLLPMLLELIPEKRATAA
 ECLRHPWLNS*



II. REDUNDANT NON-ISOTYPE-SPECIFIC FUSION PROTEINS

A. Fusions in untranslated regions

A.1. Fusions in 3'UTR

RHOB->GATA3

3'UTR of RHOB fused into GATA3, No GATA3 in the fused protein

LOC728606->KCTD1

3'UTR of LOC728606 fused into KCTD1, No KCTD1 in the fused protein

OGT->ACTB

3'UTR of OGT fused into ACTB, No ACTB in the fused protein

H1F0->ACTB

3'UTR of H1F0 fused into ACTB, No ACTB in the fused protein

PTP4A2->MALAT1

Not found

COL1A2->LAMP2

3'UTR of COL1A2 fused into LAMP2, No LAMP2 in the fused protein

SPATS2L->COL3A1

3'UTR of SPATS2L fused into COL3A1, No COL3A1 in the fused protein

PLXNA1->CTSD

3'UTR of PLXNA1 fused into CTSD, No CTSD in the fused protein

VPS35->DCN

3'UTR of VPS35 fused into DCN, No DCN in the fused protein

FTL->ADD3

3'UTR of FTL fused into ADD3, No ADD3 in the fused protein

COL1A1->BASP1

3'UTR of COL1A1 fused into BASP1, No BASP1 in the fused protein

BAT2L2->COL3A1

3'UTR of BAT2L2 fused into COL3A1, No COL3A1 in the fused protein

CD68->NEAT1

3'UTR of CD68 fused into NEAT1, No NEAT1 in the fused protein

CD68->PSAP

3'UTR of CD68 fused into PSAP, No PSAP in the fused protein

RAB8A->EIF4G2

3'UTR of RAB8A fused into EIF4G2, No EIF4G2 in the fused protein

A.2. Fusions in 5'UTR

MALAT1->IGF2

5'UTR of MALAT1 fused into the 3' UTR of IGF2

LGMN->NAP1L1

5'UTR of LGMN fused into the coding region of NAP1L1

LOC96610->IGLL5

5'UTR of LOC96610 fused into the coding region of IGLL5

B. Fusions in coding regions

B.1 Frame shift fusions that give rise to C-terminal truncation events

FLNA->ABCA2

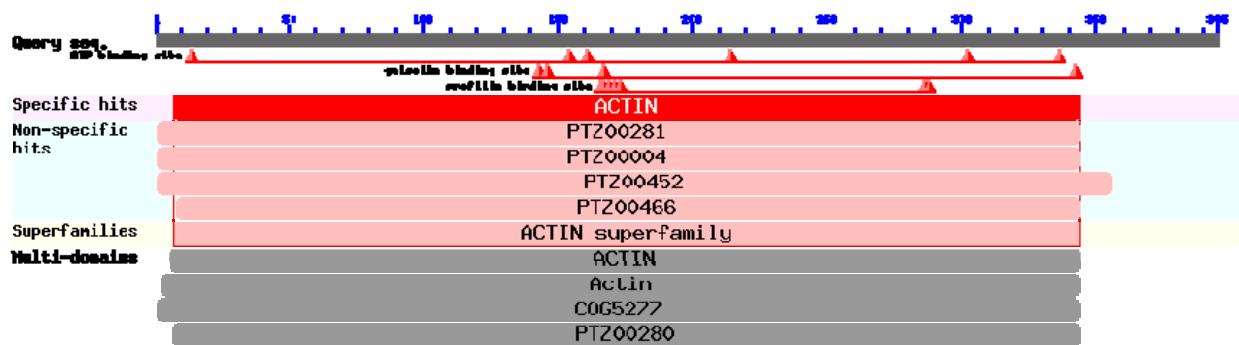
Coding region of FLNA fused to 3' UTR of ABCA2—C-terminal truncation containing **aa 1-2492 of FLNA**. (full length FLNA = 2647aa)

MSSSHSRAGQSAAGAAPGGGVDTDRDAEMPATEKDLAEDAPWKKIQNTFTRWCNEHLKCVS
KRIANLQTDLSDGLRLIALLEVLSSQKKMHRKHNRPTFRQMQLENVSVALEFLDRESIKLVSIDS
KAIVDGNLKLILGLIWTLILHYSISMPMWDEEEDEEAKKQTPKQRLLGWIQNKLPQLPITNFSRD
WQSGRALGALVDSCAPGLCPDWDSWDASKPVTNAREAMQQADDWLGIPQVITPEEIVDPNV
DEHSVMTYLSQFPKAKLKPGAPLRPKLNPKKARAYGPGIEPTGNMVKKRAEFTVETRSAGQGE
VLVYVEDPAGHQEEAKVTANNDKNRTFSVWYVPEVTGTHKVTVLFAGQHIAKSPFEVYVDKSQ
GDASKVTAQQPGLEPSGNIANKTTYFEIFTAGAGTGEVEVVIQDPMGQKGTVEPQLEARGDST
YRCSYQPTMEGVHTVHVTFAGVPIPRSPYTVTGGACNPSACRAVGRGLQPKGVRVKETADF
KVYTKGAGSGELKVTVKGPKEERVKQKDLGDGVYGFYYPMPVPGTYIVTITWGGQNIGRSPF
EVKVGTECGNQKVRWGPGLGGVVGKSADVFVEAIGDDVGTGLGFSVEGPSQAKIECDDKGD
GSCDVRYPQEAGEYAVHVLNSEDIRLSPFMADIRDAPQDFHPDRVKARGPGLKGTGVAVN
KPAEFTVDAKHGGKAPLRVQVQDNEGCPVEALVKDNGNGTYSCSYVPRKPKVHTAMVSWG
VSIPNSPFRVNVGAGSHPNKVKVYGPVAKTGLKAHEPTYFTVDCAEAGQGDVSIKICAPGV
VGPAEADIDFDIIRNDNDFTVKYTPRGAGSYTIMVLFADQATPTSPIRVKVEPSHDASKVKAEG
PGLSRTGVELGKPTHFTVNAKAAGKGLDVQFSGLTGDAVRDVIIDHHDNTYTVKYTPVQQ
GPVGVNVTYGGDPIPKSPFSVAVSPSLDLSKIKVSGLGEKVDVGKDQEFTVKSAGAGGQKVA

ACTG1->PPP1R12C

Coding region of ACTG1 fused to 3' UTR of PPP1R12C—C-terminal truncation containing aa 1-343 of ACTG1. (full length ACTG1 = 375aa)

MEEIEAALVIDNGSGMCKAGFAGDDAPRAVFPVGRPRHQGVMVGMGQKDSYVGVDEAQS
RGILTLKYPIEHGIVTNWDDMEKIWHHTFYNELRVAPPEHPVLLTEAPLNPKANREKMTQIMFET
FNTPAMYVAIQAVLSLYASGRTTGIVMDSGDGVTHTVPIYEGYALPHAILRLDLAGRDLTDYLMK
ILTERGYSFTTTAEREIVRDIKEKLCYVALDFEQEMATAASSSSLEKSYELPDGQVITIGNERFRC
PEALFQPSFLGMESCGIHETTFNSIMKCDVDIRKDLYANTVLSGGTTMYPGIADRMQKEITALAP
STMKIKIIPPERKYSVWIGGEARMEVGGRAKQGDVARGDQGQRRDQEGTEDQEVQGSVART
RVAAWRSPSHV*



MTF2->ARL3

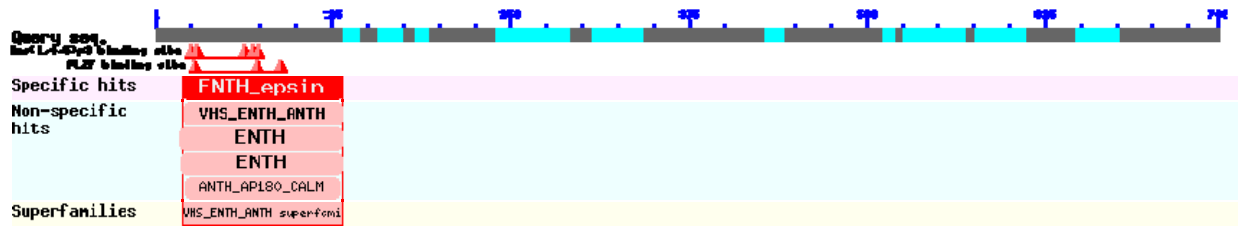
coding to coding fusion—no significant ORF

MRACSQFCAS*

EPN1->COL1A1

coding to coding fusion—C-terminal truncation containing aa 1-538 of EPN1. (full length EPN1 = 576 aa)

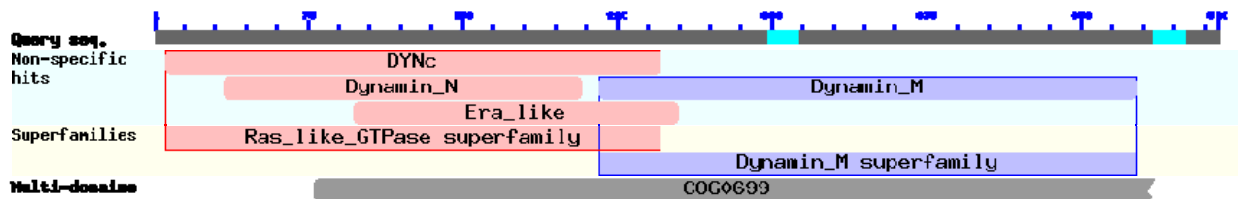
MSTSSLRRQMKNIVHNYSEAEIKVREATSNDPWGPSSSLMSEIADLTYNVAVFSEIMSMIWKRL
NDHGKNWRHVYKAMTLMELYIKTGSERVSQCKENMYAVQTLKDFQYVDRDGKDGQGVNVRE
KAKQLVALLRDEDRLREERAHALKTKEKLAQTATASSAAVGSPPPEAEQAWPQSSGEEELQL
QLALAMSKEEADQEERIRRGDDLRLQMAIEESKRETGGKEESSLMDLADVFTAPAPAPTTDPW
GGPAPMAAAVPTAAPTSDPWGGPPVPPAADPWGGPAPTASGDPWRPAAPAGPSVDPWGG
TPAPAAGEGPTDPWGSSDGGVPVSGPSASDPWTPAPAFSDPWGGSPAKPSTNGTTAGGF
DTEPDEFSDFDRLRTALPTSGSSAGELELLAGEVPARSPGAFDMSGVIRGSLAEAVGSPPPAAT
PTPTPPTRKTPESFLGPNAALVDLDSLVS RGPPTPPGAKASNPFLPGGGPATGPSVTNPFQPA
PPATLTLNQLRLSPVPPVPGAPPTYISPLGGGGPLPPMMPPLLALLVNKVPLEPLVLLVPEVPLA
LLVLLAKMDSTVSLAPLGPLVLAVALVMLVLLVPPALLDLLVPLVLPALVSTSASCPSHLKRRLTM
VAATTGLMMPMWFVTVTSRWTPPSRA*ASRSRTSGAQRAAARTPPAPAVT SRCATLTGRVES
TGLTPTKAATWMPKSSATWRLVRPACTPLSPVWPRRTGTSARTPRTRGMSGSARA*



DNM2->PIN1

coding to coding fusion—C-terminal truncation containing aa 1-474 of DNM2. (full length DNM2 = 870 aa)

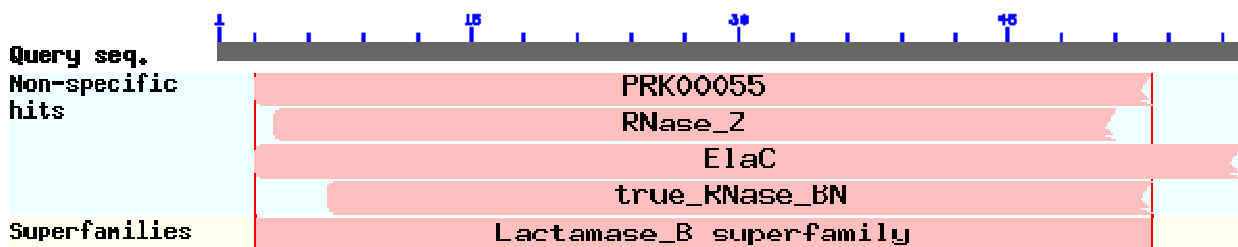
MGNRGMEELIPLVNLQDAFSSIGQSCHLDLPQIAVVGQSQAGKSSVLENFVGRDFLPRGSGI
 VTRRPLILQLIFSKTEHAFLHCKSKKFTDFDEVQRQIEAETDRVTGTNKGISPVPINLRVYSPHV
 LNLTLIDLPGITKVPVGDQPPDIEYQIKDMILQFISRESSLILAVTPANMDLANSALKLAKEVDPQ
 GLRTIGVITKLDLMDEGTDARDVLENKLLPLRRGYGVVNRSQKDIEGKKDIRAALAAERKFFLS
 HPAYRHMADRMGTPHLQKTLNQQLTNHIRESLPALRSKLQSQLLSLEKEVEEYKNFRPDDPTR
 KTKALLQMVGQFGVDFEKRIEGSGDQVDTLELSSGARINRIFHERFPFELVKMEFDEKDLRREI
 SYAIKNIHGVRTGLFTPDLAFAIVKKQVVKLKEPCLKCVDLVIQELINTVRQCTSKLSSYPRLRE
 ETERIVTTYIRERERGRTKDQAECTTSTTSLTPASGSGPAATAAVVAKTGRGSLPGSAARTCW*



ELAC1->SMAD4

Coding of ELAC1 fused to 5' of SMAD4—C-terminal truncation containing aa 1-53 of ELAC1. (full length ELAC1 = 363 aa)

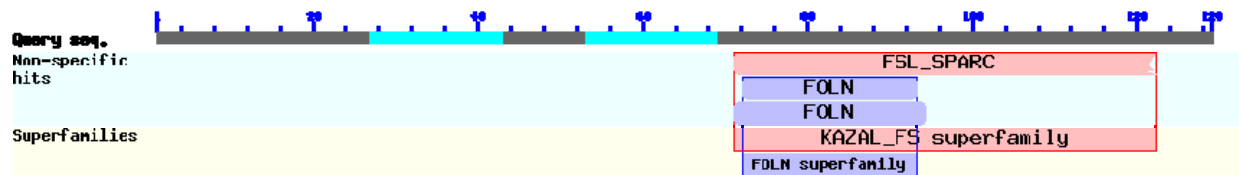
MSMDVTFLLGTGAAYPSPTRGASAVVLRCEGECWLFDCGEGTQTQLMKSQKAGYPEYMSNN
 FPCNVSCCFSLFPKDQNCFRNWRHI*



SPARC->TRPS1

coding region of SPARC fused in front of the 5' UTR of TRPS1—C-terminal truncation containing aa 1-120 of SPARC. (full length SPARC = 303 aa)

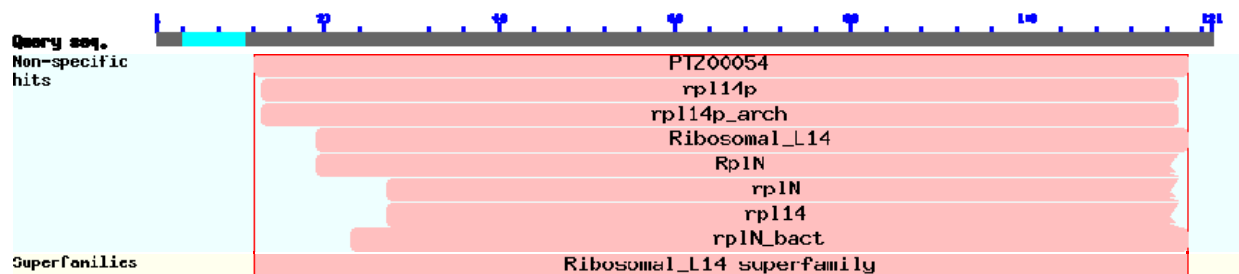
MRAWIFFLLCLAGRALAAPQQEALPDETEVVEETVAEVTEVSVGANPVQVEVGEFDDGAEETE
EEVVAENPCQNHHCXHGKVCVCELDENNTPMCVCQDPTSCPAPIGEFEKVCNSNDNKTFD¹²⁰NSVGV
LKLT*



RPL23->MUCL1

coding to coding fusion—C-terminal truncation containing aa 1-118 of RPL23.(full length RPL23 = 140 aa)

MSKRGRGGSSGAKFRISLGLPVGAVINCADNTGAKNLYIISVKGIGRLNRLPAAGVGDMMVMA
VKKGKPELRKKVHPAVVIRQRKSYRRKDGVFLYFEDNAGVIVNNKGEMKGSATW¹¹⁸SC*



TTC7A->SOCS5

Coding of TTC7A fused to 5' of SOCS5—C-terminal truncation containing aa 1-62 of TTC7A.(full length TTC7A = 858 aa)

MAAKGAHGSYLKVESELERCRAEGHWDRMP⁶²ELVRQLQTLSPGGGGNRRGSPSAAFTFPD
DFIING*

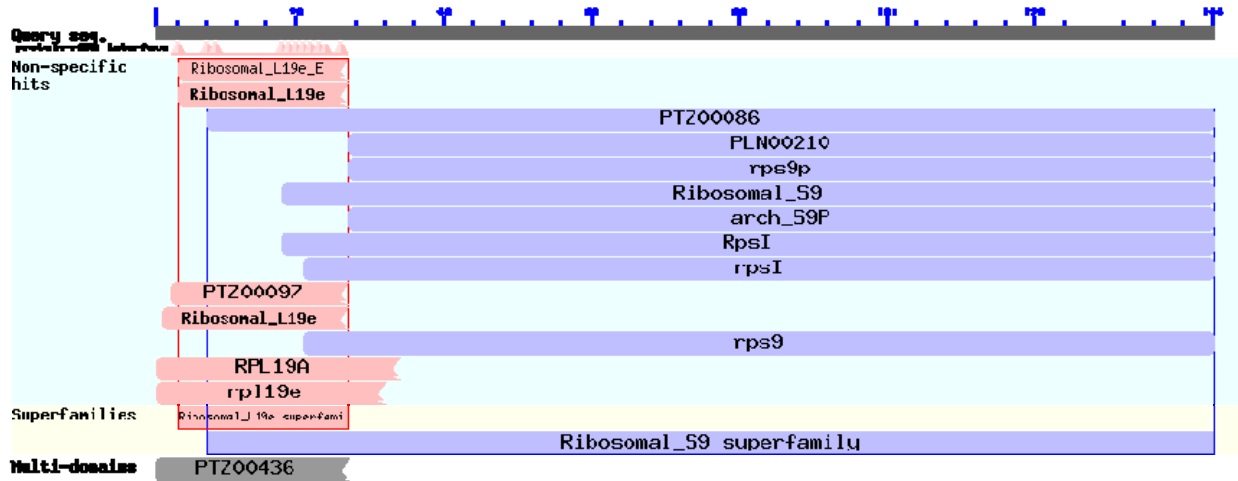
No putative conserved domains detected.

B.2. In frame fusions that give rise to chimaeric proteins

RPL19->RPS16

coding to coding fusion--aa 1-27 of RPL19 fused in frame to aa 33-129 of RPS16.

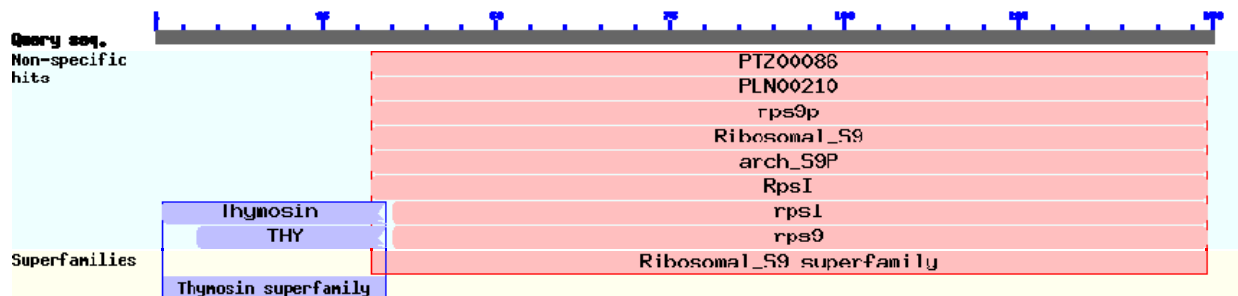
MSMLRLQKRLASSVLRGKKKVWLDPNGLIKVNGRPLEMIEPRTLQYKLLPEVLLL GKERFAGV
DIRVRVKGGGHVAQIYAIRQSISKALVAYYQKYVDEASKKEIKDILIQYDRTELLVADPRRCESKKF
GGPGARARYQKSYR*



TMSB10->RPS16

coding to coding fusion—aa 1-34 of TMSB10 fused in frame to aa 29-145 of RPS16 with single aa insertion at fusion junction.

MADKPDMGEIASFDKAKLKKTTETQEKNTLPTKEINGLIKVNGRPLEMIEPRTLQYKLLPEVLLL G
KERFAGVDIRVRVKGGGHVAQIYAIRQSISKALVAYYQKYVDEASKKEIKDILIQYDRTELLVADPR
RCESKKFGGPGARARYQKSYR*



III. REDUNDANT ISOTYPE-SPECIFIC FUSION PROTEINS

A. Fusions in untranslated regions

A.1. Fusions in 3'UTR

COL1A1->FMNL3

3'UTR of COL1A1 fused into FMNL3, No FMNL3 in the fused protein

CYB5R3->TXNIP

3'UTR of CYB5R3 fused into TXNIP, No TXNIP in the fused protein

ELF3->SLC39A6

3'UTR of ELF3 fused into SLC39A6, No SLC39A6 in the fused protein

TAX1BP1->MALAT1

3'UTR of TAX1BP1 fused into MALAT1, No MALAT1 in the fused protein

MGP->NCRNA00188

3'UTR of MGP fused into NCRNA00188, No NCRNA00188 in the fused protein

APOL1->ACTB

3'UTR of APOL1 fused into ACTB, No ACTB in the fused protein

ACTB->C20orf112

3'UTR of ACTB fused into C20orf112, No C20orf112 in the fused protein

CD74->MBD6

3'UTR of CD74 fused into MBD6, No MBD6 in the fused protein

GNB2->CTSD

3'UTR of GNB2 fused into CTSD, No CTSD in the fused protein

TSPAN14->HLA-E

3'UTR of TSPAN14 fused into HLA-E, No HLA-E in the fused protein

YWHAZ->ZBTB33

3'UTR of YWHAZ fused into ZBTB33, No ZBTB33 in the fused protein

DCLK1->COL3A1

Not found

A.2. Fusions in 5'UTRs

Not detected

B. Fusions in coding regions

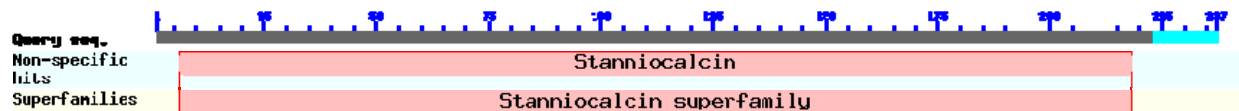
B.1. Frame shift fusions that give rise to C-terminal truncation events

STC2->RNF11

Coding of STC2 fused to 5' of RNF11—C-terminal truncation containing **aa 1-212 of STC2**. (full length STC2 = 302 aa)

MCAERLGQFMTLALVLATFDPARGTDATNPPEGPQDRSSQQKGRSLQNTAEIQHCLVNAGD
VGCGVFECFENNSCEIRGLHGICMTFLHNAGKFDAQGKSFIKDALKCKAHALRHRFGCISRKCP
AIREMVSQLRQRECYLKHDLCAAAQENTRVIVEMIHFKDLLHEPYVDLVNLLTCGEEVKEAITH
SVQVQCEQNWGSLCSILSFCFTRPPPPPIAASQEANAATPAEPAAQP*

B.2. In frame fusions that give rise to chimaeric proteins

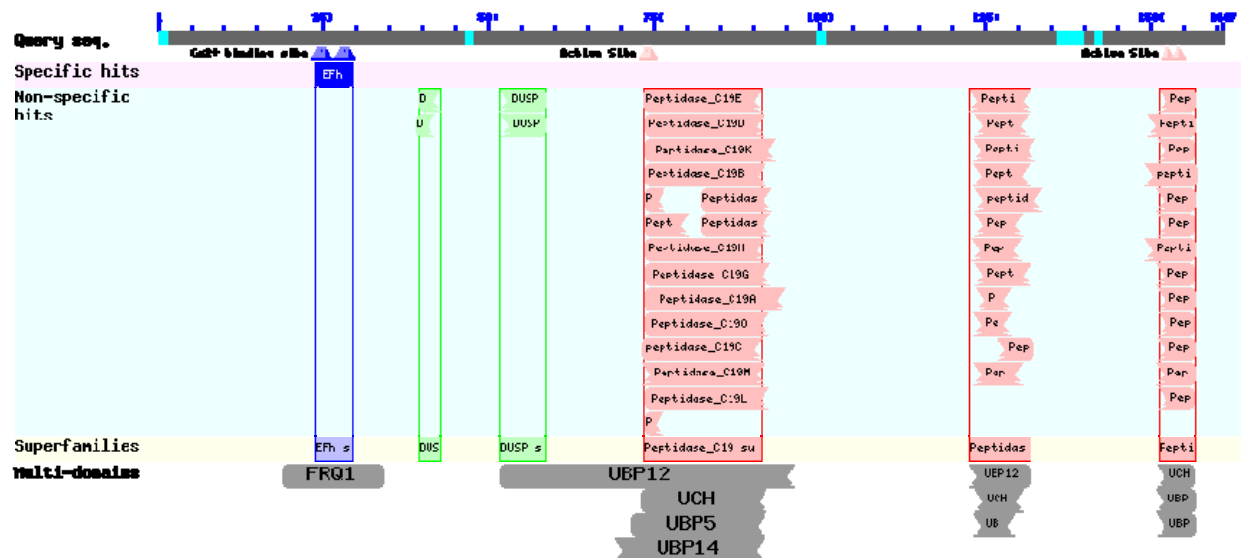


AATK->**USP32**

coding to coding fusion—**aa 1-18 of AAKT** fused in frame to **aa 20-1604 of USP32**

MSSSFFNPSFAFSSHFDPVTDVELKRLKDAFKRTCGLSYYMGQHCFIREVLGDGVPPKVAEVIY
CSFGGTSKGLHFNNLIVGLVLLTRGKDEEKAKYIFSLFSSESGNYVIREEMERMLHVVDGKVPD
TLRKCSEGEKVNIEKFRNWLFLNKDAFTFSRWLLSGGVYVTLTDDSDTPTFYQTLAGVTHLE
ESDIIDLEKRYWLLKAQSRTGRFDLETFGPLVSPPIRPSLSEGLFNAFDENRDNDHDFKEISCGLS
ACCRGPLAERQKFCFKVFDVDRDGVLSRVELRDMVVALLEVWKNRRTDDIPELHMDLSDIVEG
ILNAHDTTKMGHLTLEDYQIWSVKNVLANEFLNLLFQVCHIVLGLRPATPEEEGQIIRGWLERES
RYGLQAGHNWFIISMQWWQQWKEYVKYDANPVVIEPSSVLNNGGKYSFGTAAHPMEQVEDRI
GSSLSYVNTTEEFSDNISTASEASETAGSGFLYSATPGADVCFARQHNTSDNNNQCLLGANG
NILLHLNPQKPGAIDNQPLVTQEPVKATSLTLEGGRLKRTPQLIHGRDYEMVPEPVWRALYHW
YGANLALPRPVIKNSKTDIPELELFPYLLFLRQQPATRTQQSNIWVNMGNVPSNPAPLKRVL
YTGCFSRMQTIKEIHEYLSQRLRIKEEDMRLWLYNSENYLTLDDDEDHKLEYLKIQDEQHLVIEV
RNKDMSWPEEMSFIANSSKIDRHKVPTTEKATGLSNLGNTECFMNSSIQCVSNTQPLTQYFISG
RHLYELNRTNPIGMKGHMAKCYGDLVQELWSGTQKNVAPLKLRTIAKYAPRFNGFQQQDSQ
ELLAFLLDGLHEDLNRVHEKPYVELKDSGDPDWEVAAEAWDNHLRRNRSIVVDLFDHGQLRS
QVKCKTCGHISVRFDPFNFLSLPLPMDSYMHEITVIKLDGTPVRYGLRLNMDEKYTGLKKQL
SDLCGLNSEQILLAEVHGSNIKNFPQDNQKVRLSVSGFLCAFEIPVPVSPISASSPTQDFSSSP
STNEMFTLTTNGDLPRPIFIPNGMPNTVVP CGTEKNFTNGMVNGHMPSLPDPSPTGYIIAVHRK
MMRTELYFLSSQKNRPSLFGMPLIVPCTVHTRKKDLYDAVWIQVSRLASPLPPQEASNHAQDC
DDSMGYQYPFTLRVVQKDGNSCAWCPWYRFCRGCKIDCGEDRAFIGNAYIAVDWDPTALHLR
YQTSQERVVDEHESVEQSRRAQAEPINLDSCLRAFTSEEELGENEMYCCKKTHCLATKKLD
LWRLPPIIIHLKRFQFVNGRWIKSQKIVKFPRESFDPSAFLVPRDPALCQHKPLTPQGDELSEP
RILAREVKKVDAQSSAGEEDVLLSKSPSSLSANISSPKGSPSSSRKSGTSCPSSKNSSPNSSP
RTLGRSKGRLRLPQIGSKNKLSSSKENLDASKENGAGQICELADALSRGHVVGGSQPELVTPQ
DHEVALANGFLYEHEACGNGYSNGQLGNHSEEDSTDDQREDTRIKPIYNLYAISCHSGILGGG

HYVTYAKNPNCWKYCYNDSSCKELHPDEIDTDSAYILFYEQQGIDYAQFLPKTDGKKMADTSS
 MDEDFESDYKKYCVLQ*



KRT18->PLEC

Coding to coding fusion—aa 1-325 of KRT18 fused in frame to aa 1396-4574 of PLEC

MSFTTRSTFSTNYRSLGVSQAPSYGARPVSSAASVYAGAGGSGSRISVSRSTSFRGGMGSGG
 LATGIAGGLAGMGGIQNEKTMQSLNDRLASYLDRVRSLETENRRLESKIREHLEKKGQVRD
 WSHYFKIIEDLRAQIFANTVDNARIVLQIDNARLAADDFRVKYETELAMRQSVENDIHGLRKVIDD
 TNITRLQLETEIEALKEELLFMKKNHEEEVKGLQAQIASSGLTVEVDAPKSQDLAKIMADIRAQYD
 ELARKNREELDKYWSQQIEESTTVVTTQSAEVGAAETTLTELRRTVQSLEIDLDSMRNLKASLE
 NSLRVEKQRQLAEAHAQAQAQAEREAKELQQRMQEEVVRREEAAVDAQQQKRSIQEELQQL
 RQSSEAEIQAKARQAEAAERSRLRIEEEEIRVVRLQLEATERQRGGAEGELQALRARAEAEAEQ
 KRQAQEEAERLRRQVQDESQRKRQAEVELASRVKAEAEAAAREKQRALQALEELRLQAEAAER
 RLRQAEVERARQVQVALETAQRSAAEELQSKRASFAEKTAQLERSLQEEHVAVAQLREEAER
 RAQQQAEAEERAREEAERELERWQLKANEALRLRLQAEVVAQQKSLAQAEAEKQKEEAEREAR
 RRGKAEQAVRQRELAEQELEKQRQLAEGTAQQRLAAEQELIRLRAETEQQGEQQRQLLEEL
 ARLQREAAAATQKRQLEAEALAKVRAEMEVLLASKARAEESRSTSEKSKQRLEAEAGRFREL
 AEEAARLRLALAEAKRQRQLAEEDAARQRAEAERVLAEKLAIGEATRDKTEAEIALKEKEAEN
 ERLRRLAEDEAFQRRRLEEQAQHKADIEERLAQLRKASDSELERQKGLVEDTLRQRRQVEEE
 ILALKASFEEKAAAGKAELELELGRIRSNAEDTLRSKEQAELEAARQRQLAAEEERRRREAEERV
 QKSLAAEEEAARQRKAAL EEVERLKAKVEEARLRERAEQESARQLQLAQEAAQKRLQAEK
 AHAFVQQKEQELQQTLLQEQSVLDQLRGEAEARRRAEEEAARVQAEREAQAQSRQVEE
 AERLKQSAEEQAQARAQAQAAAEKLRKEAEQEAARRAQAEQAALRQKQAADAEMEKHKKFA
 EQTLRQKAQVEQELTTLRLQLEETDHQKNLLDEELQRLKAEATEAARQRSQVEEELFSVRVQM
 EELSKLKARIEAENRALILRDKDNTQRFLQEEAEKMKQVAEEAARLSVAAQEAARLRQLAEEDL
 AQRALAEKMLKEKMQAVQEATRLKAEALLQQQKELAQEQARRLQEDKEQMAQQLAEETQ
 GFQRTLEAERQRQLEMSAEERLKLVAEMSRAQARAEEDAQFRKQAEIEGKLRHRELAT
 QEKVTLVQTLQIRQQSDHDAERLREIAELEEREKEKLQEQAKLLQLKSEEMQTVQQEQLLQE
 TQALQQSFLSEKDSLLQRERFIEQEAKLEQLFQDEVAKAQQRLREEQQRQQQQMEQERQLRV
 ASMEEARRRQHEAEEGVRRKQEELQQLEQQRRQQEELLAEENQRLREQLQLLEEQHRAALA
 HSEVTSQVAATKTLPNGRDALDGPAAEAPEHSFDGLRRKVSQAQRLQEAGILSAEELQLRA

QGHTTVDELARREDVRHYLQGRSSIAGLLLKATNEKLSVYAALQRQLLSPGTALILLEAQAASG
 FLLDPVRNRRLTVNEAVKEGVVGPPELHHKLLSAERAVTGYKDPYTGQQISLFQAMQKGLIVRE
 HGIRLLEAQIATGGVIDPVHSHRVPVDVAYRRGYFDEEMNRVLADPSDDTKGFFDPNTHENLT
 YLQLLERCVEDPETGLCLLPLTDKAAKGELVYTDSEARDVFEKATVSAPFGKFQGKTVTIWEII
 NSEYFTAQRDRLRQFRTGRITVEKIIIIITVVEEQEQKGRLCFEGRLSLVPAELLESRVIDRE
 LYQQLQRGERSVRDVAEVDTVRRALRGANVIAGVWLEEAGQKLSIYNALKDLLPSDMAVALL
 EAQAGTGHIIDPAT SARLTVDEAVRAGLVGPEFHEKLLSAEKAVTGYRDPYTGQSVSLFQALKK
 GLIPREQGLRLLDAQLSTGGIVDPSKSHRVPLDVACARGCLDEETSRLSAPRADAKAYS DPST
 GEPATY GELQQRCPDQLTGLSLLPLSEKAARARQEELYSELQARETFEKT PVEVPVGGFKGR
 TVTVWELISSEYFTAQRQELLRQFRTGKVTVEKVIKILITIVEEVETLRQERLSFSGLRAPVPAS
 ELLASGVLSRAQFEQLKDGKTTVKDLSELGSVRTLLQGSGCLAGIYLEDTKEKVS IYEAMRRGL
 LRATTAALLLEAQAATGFLVDPVRNQRLYVHEAVKAGVVGPELHEQLLSAEKAVTGYRDPYSG
 STISLFQAMQKGLVLRQH GIRLLEAQIATGGIIDPVHSHRVPVDVAYQRGYFSEEMNRVLADPS
 DDTKGFFDPNTHENLTYRQLLERCVEDPETGLRLLPLKGAEKA EVVETTQVYTEEETRRAFEET
 QIDIPGGGSHGGSTMSLWEVMQSDLIPEEQRAQLMADFQAGRVTKERMIIIIIEIEKTEIIRQQGL
 ASYDYVRRRLTAEDLFEARIISLETYNLLREGTRSLREALEAESAWCYLYGTG SVAGVYLPGSR
 QTLSIYQALKKGLLSAEVARLLLEAQAATGFLDPVKGERLTVDEAVRKGLVGP ELHDRLLSAE
 RAVTGYRDPYTEQTISLFQAMKKE LIPTTEALRLLDAQLATGGIVD PRLGFHLPLEVAYQRGYLN
 KDTHDQLSEPSEVRSYVDPSTDERLSYTQLLRRCRRDDGTGQLLLPLSDARKLTFRGLRKQIT
 MEELVRSQVMDEATALQLREGLTSIEEVTKNLQKFLEGTSCIAGVFVDATKERLSVYQAMKKGII
 RPGTAFELLEAQAATGYVIDPIKGLKLTVEEAVRMGIVGPEFKDKLLSAERAVTGYKDPYSGKLI
 SLFQAMKKG LILKDHGIRLLEAQIATGGIIDPEESHRLPVEVAYKRGLFDEEMNEILDPSSDDTKG
 FFDPNTEENLTYLQLMERCITDPQTGLCLLPLKEKKRERKTS SKSSVRKRRVVIVDPETGKEMS
 VYAYRKGLIDHQTYLELSEQECEWEEITISSSDGVVKSMIIDRRSGRQYDIDDAIAKNLIDRSAL
 DQYRAGTLSITEFADMLSGNAGGFRSRSSSVGSSSSYPISPAVSRTQLASWSDPTEETGPVAG
 ILDTETLEKVSITEAMHRNLVDNITGQRLLLEAQA CTGGIIDPSTGERFPVTD AVNKGLVDKIMVDR
 INLAQKAFCGFEDPRTTKMSAAQALKKGWLYYEAGQRFLEVQYLTGGLIEPDT PGRVPLDEA
 LQRGTVDARTAQKLRDVGAYSKYL TCPPKTKLKISYKDALDRSMVEEGTGLRLL EAAAQSTKGY
 YSPYSVSGSGSTAGSRTGSRTGSRAGSRRG SFDTATGSGFSMTFSSSSYSSSGYGRRYASGS
 SASLGGPESAVA*

