

Supplementary Data

ValidNESs: a database of validated leucine-rich nuclear export signals

Szu-Chin Fu¹, Hsuan-Cheng Huang², Paul Horton^{3*}, Hsueh-Fen Juan^{4*}

¹Department of Life Science, ²Institute of Molecular and Cellular Biology and Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei 106, Taiwan

³ Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), Tokyo 135-0064, Japan

⁴ Department of Life Science and Institute of Molecular and Cellular Biology, National Taiwan University, Taipei 106, Taiwan

Tables and Figures

Table S1 and S2 give the detailed field descriptions for two tables provided in ValidNESs. Figure S1 shows the submission interface and an example of NES prediction result.

Case Study

We made an investigation among 28 multiple NES-containing proteins whose homologs are available in HomoloGene (<http://www.ncbi.nlm.nih.gov/homologene>), resulting in 13 homolog groups containing at least one NES abrogation. In this supporting online material, we provide multiple sequence alignments of these abrogated NESs and their flanking regions. The NCBI reference sequence (RefSeq) accession number of the validated NES-containing proteins is denoted in black while the homologs with or without NES abrogation are denoted in red or blue respectively. The validated NES is shown in bold and underscored. Mutations into amino acids other than L, I, V, F or M at the essential hydrophobic positions are highlighted in red.

Table S1. Description of fields in the table of NES-containing regions

Field	Description
ACC*	Accession number e.g. S001.
NES_SITE	Hyphen-delimited text consisting of the start position, sequence and end position of the experimentally determined region of NES.
NES_STRC	Hyphen-delimited text consisting of the start position, local structure and end position of the experimentally determined region of NES. The local structure (when applicable) was first downloaded from PDB and then assigned by the DSSP program.
PROTEIN_ACC	Cross-reference to the table of NES-containing proteins.
REF_DB	Cross-reference to UniProtKB and PDB.
TYPE	Classical leucine-rich NES or non-classical NES.
LMB	Whether or not the CRM1-dependance is validated by treatment of leptomycin B.
REFERENCE	The PubMed PMID of the original paper.
SEQUENCE	Sequence of the protein containing specific NES site.

*This denotes the primary key

Table S2. Description of fields in the table of NES-containing proteins

Field	Description
ACC*	Accession number e.g. P001.
PROTEIN_NAME	Name of the protein.
GENE_NAME	Gene name of the protein.
SITE_ACC	Cross-reference to the table of NES-containing regions.
REF_DB	Cross-reference to UniProtKB, DIP, IntAct, MINT, STRING and NLSdb.
ORGANISM	Species name.
SUB-CELLULAR LOCALIZATION	Gene ontology (GO) cellular component annotation derived from UniProtKB.
SITES	Hyphen-delimited text consisting of the start position and end position of the experimentally determined regions of NES. Multiple sites are separated by semicolon.
SEQUENCE	Full length sequence with the NES region denoted in lower case.

*This denotes the primary key

(A)

Input the UniProt ID

e.g. IPKA_HUMAN

IPKA_HUMAN

Or protein sequence in one-letter code

e.g. MTDVETTYADFIASGRTGRRNAIHDLVSSASGNSNELALKLAGLDINKTEGEEDAQRSSTEQSGEAQGEAKSES

MTDVETTYADFIASGRTGRRNAIHDLVSSASGNSNELALKLAGLDINKTEGEEDAQRSSTEQSGEAQGEAKSES

(B)

Your Job Id: 1338934722

Name:

IPKA_HUMAN

Prediction at the protein level:

0.66

Prediction at the site level:

Putative site (Probability)

42-LAGLDI-47 (0.93)

23-IHDILV-28 (0.06)

*NESsential score at the protein level indicates how likely the given protein contains a NES and ranges from 0 to 1

*NESsential score of each putative site indicates the probability to be a functional NES and ranges from 0 to 1

Figure S1. The submission interface and result of NES prediction. (A) The submission interface where users can input a single protein sequence in one-letter code or a UniProt protein name (UniProt ID). (B) An example of a prediction result.

NES1 of NF2L2_HUMAN

HomoloGene:2412

NP_878309.1	181	SLPELQQCLNMPMQETLDMNAFMKPSTEAPTQ--NYSQY--LPGMDHLGS	226
NP_990448.1	175	SLPEL-QCLNIENDNLAEVSTITSPETKPAEMHNSYDYYNSLPIMRK---	220
NP_035032.1	185	SIPEL-QCLNTENKQLADTTAVPSPEATLTEMDSNYHFYSSISSLEK---	230
NP_113977.1	185	SIPEL-QCLNTENKQQAETTVPSPEATLTEMDSNYHFYSSIPSLEK---	230
XP_535975.1	188	SIPEL-QCLNIQNDKLVETSMVPSPETKMTEIDNNYHFYPSMPSLEK---	233
NP_001011678.2	192	SIPEL-QCLNIQNDKLAETSTVPSPETKLTEID-NYHFYSSMPSLDK---	236
XP_001096024.2	193	SIPEL-QCLNIENDKLVETTMVPSPEAKLTEVD-NYHFYSSISMEK---	237
NP_006155.2	193	SIPEL-QCLNIENDKLVETTMVPSPEAKLTEVD-NYHFYSSIPSMEK---	237
XP_001145876.2	193	SIPEL-QCLNIENDKLVETTMVPSPEAKLTEVD-NYHFYSSIPSMEK---	237

NES2 of NF2L2_HUMAN

HomoloGene:2412

NP_878309.1	513	LKEEKERLMKEKSERSNLKEMKQQLSTLY Q EVFGMLRDENGKAFSPNEF	562
NP_990448.1	510	LKDEREKLLKEKGENDKSLRQMKKQLTTLYIEVFSMLRDEDGKSYSPSEY	559
NP_035032.1	524	LKDEREKLLREKGENDRNLHLLKRLSTLYLEVFSMLRDEDGKPYSPEY	573
NP_113977.1	524	LKDEREKLLREKGENDRNLHLLKRKLSTLYLEVFSMLRDEDGKPYSPEY	573
XP_535975.1	528	LKDEKEKLLKEKGENDKSLHLLKKQLSTLYLEVFSMLRDEDGKPYSPEY	577
NP_001011678.2	534	LKDEKEKLLKRGENDKSLHLLKKQLSTLYLEVFSMLRDEDGKPYSPEY	583
XP_001096024.2	533	LKDEKEKLLKEKGENDKSLHLLKKQLSTLYLEVFSMLRDEDGKPYSPEY	582
NP_006155.2	532	LKDEKEKLLKEKGENDKSLHLLKKQLSTLY LKKQLSTLYL EVFSMLRDEDGKPYSPEY	581
XP_001145876.2	532	LKDEKEKLLKEKGENDKSLHLLKKQLSTLYLEVFSMLRDEDGKPYSPEY	581

NES2 of ZO2_CANFA

HomoloGene:3541

NP_001188499.1	784	FGPISDAANEKLAN-----ELPNEFTIAKTEPKDAGSEKSSGVVRLNT	826
NP_990249.1	705	FGPIADVAMEKLST-----DLPHLYQTAKTEPRDAGSEKATGVVRLNT	747
NP_035727.2	718	FGPIADIAMERLAT-----ELPDLFQTAKTEPKDAGSEKSSGVVRLNT	760
NP_446225.1	716	FGPIADIAMERLTT-----ELPDLFQTAKTEPKDAGSEKSSGVVRLNT	758
NP_001003204.1	720	FGPIADIAL EKLAN-----ELPDLFQTAKTEPKDAGSEKSSGVVRLNT	762
XP_002808149.1	734	CAVIADARWEYLIVTVFXTIIXLPA-FAFAATEPKDAGSEKSTGVVRLNT	782
NP_001163887.1	767	FGPIADIAMEKLAN-----ELPDWFQTAKTEPKDAGSEKSTGVVRLNT	809
XP_003312175.1	767	FGPIADIAMEKLAN-----ELPDWFQTAKTEPKDAGSEKSTGVVRLNT	809

NES1 of RIPK3_HUMAN

HomoloGene:31410

XP_001343827.1	261	ELLKLMM Q CWDQEPRRPSFLDCVHAITAIYTMHARRLNDVDHVLEQLQ	310
XP_001114079.1	257	GLKELMQLCWSTEPKDRPSFQECLPKTDEAFQMVENNINAAVSTVKNFLS	306
NP_006862.2	257	GLKELMQL CWSSEPKDRPSFQECLPKTDEVFQMVENNMNAAVSTVKNFLS	306
XP_001169864.1	257	GLKELMQLCWS SKPKDRPSFQECLPKTDEAFQMVENNMNAAVSTVKNFLS	306
XP_850611.2	339	RLKALMQ H CWSHEPKDRPSFQECPITEEALNLVNKEMDAAVSTVKKFLC	388
NP_001095354.1	222	-----SKTDSGI-----	228
NP_064339.2	262	KLKELM I HCWGGSQSENRPSPFQDCEPKTNEVYNLVKDKVDAAVSEVKHYLS	311
NP_647558.1	259	GLKELM T HCWSSEPKDRPSFQDCESKTNNVYILVQDKVDAAVSKVKHYLS	308

NES2 of RIPK3_HUMAN

HomoloGene:31410

XP_001343827.1	311	DDEICSSLRSVQISRKPHEVD-----PQDLSC	338
XP_001114079.1	307	ELRSSNPRFSTPESGQGGTEMDGFRRRTGNQHSRNDVMVSELLNKLNLLEG	356
NP_006862.2	307	QLRSSNRRFSIPESGQGGTEMDGFRRTIENQHSRNDVM MVSEWLNLNL EE	356
XP_001169864.1	307	QLRGSNRRFSIPESGQGGTEMDGFRRTIENQHSRNDVMSEWLNLNLLEE	356
XP_850611.2	389	ENRGS-RILSAPQPQPGGGIETEGPRGTTGSH---DITASEMLSKLSLKE	433
NP_001095354.1	229	-----LHLSA	233
NP_064339.2	312	QHRSSGRNLSAREPSQRGTEMDCPRET-----MVKMLDRLHLEE	351
NP_647558.1	309	QYRSSDTKLSARESSQKGTEVDCPRET-----IVYEMLDRLHLEE	348

NES1 of ELF3_HUMAN

HomoloGene:3265

XP_002666146.1	60	NVLEWIGFHVEESRFDAGLLNLYCTMDGLTLCATS KEALMSMFGPELGN	109
XP_419257.3	71	QVLEWISYHVEKNKYDASSIDFSCCNMDGHALCHCT RDQMRLIFGP-LGD	119
NP_001092379.1	71	QVLDWVSYQAEKHRCDPGTIDLCSRCELDGAALCRCA PEELRLLVFGP-LGD	119
XP_547356.2	177	QVLDWISYQVEKNKYDASAIDFSRCDMDGATLCNC PEELRLLVFGP-LGD	225
NP_001156603.1	90	QVLEWISYQVEKNKYDASSIDFSRCDMDGATLCNSCALEELRLVFGP-LGD	138
NP_001019939.1	94	QVLDWISYQVEKNKYDASSIDFSRCDMDGATLCNC CALEELRLVFGP-LGD	142
XP_001096374.2	153	QVLDWISYQVEKNKYDASAIDFSRCDMDGATLCNC CALEELRLVFGP-LGD	201
NP_001107781.1	71	QVLDWISYQVEKNKYDASAIDFSRCDMDGAT LCNCALEELRLVFGP-LGD	119
XP_001150545.1	71	QVLDWISYQVEKNKYDASAIDFSRCDMDGATLCNC CALEELRLVFGP-LGD	119

NES2 of MEX67_SCHPO

HomoloGene: 38176

NP_595996.1	467	-----PQ <u>NKRLLRSFDRTLLILPGGSTGILIINDLLVIRSFAGSLGW</u> -	508
XP_452635.1	431	NHGFTSTPNIKLSQKSFDRTMVLVPSPNGSVIIASDSLTIIRPFCDSD-AW-	478
NP_015156.1	429	NHGYNSTSNNKLSKKSFDRTWIVPM-NNSVIIASDLTVRAYSTG-AWK	476
NP_982438.1	426	NHQQQNSAQNKLSKKSFDRWTIVPT-QGSVIIASDLTVRPYAAAG-AW-	472
NP_506567.1	494	LTQTPS-----PSF--FSRSFLVSPRENDSVAVIDSQLFITVASLD----	532
NP_506568.2	345	SAIKPENTEEHDNY--FTRTFMVAPRGEKGKVAIVSDQLFISSMSKR----	388
XP_001923961.1	527	GKSR-----ETYRAFSRVFI AVPAGSTGLCIVNDELFVRNATTE ----	565
NP_006353.2	504	GKSR-----DSLRAFTRTFIAVPASNGLCIVNDELFVRNASSE----	542
XP_001157986.1	504	GKSR-----DSLRAFTRTFIAVPASNGLCIVNDELFVRNASSE----	542
XP_001115982.1	506	GKSR-----DSLRAFTRTFIAVPASNGLCIVNDELFVRNASSE----	544
XP_540901.3	444	GKSR-----DSLRAFTRTFIAVPASNGLCIVNDELFVRNASAD----	482
NP_001069023.1	505	GKSR-----DSLRAFTRTFIAVPASNGLCIVNDELFVRNASAD----	543
NP_058093.2	503	GKSR-----DSLRAFTRTFIAVPASNGLCIVNDELFVRNASPE----	541
NP_067590.1	503	GKSR-----DSLRAFTRTFIAVPASNGLCIVNDELFVRNASPE----	541
NP_524660.1	491	DETNNPASMELYDVRHFART YVVVP -QNNGFCIRNETIFITNATHE----	535
XP_307993.4	512	GSGN-----MEQIRSfqRTLIVIVP-SNGGFCIRNEMMHVNTVTRA----	550

NES2 of DAB1_MOUSE

HomoloGene: 32084

NP_001035775.1	462	FSSYFSRVGMAQDTDDCDDFDISQMNN-----SHRHFTTPPQLTPPRPSAE	506
NP_989569.1	443	FSSYFNKVGMAQEADDCCDFDISQLNLTPVTSTTPSTNSPPTPAPRQSSP	492
NP_796233.2	443	FSSYFNKVGVAQDTDDCDD <u>FDISQLNL</u> TPVTSTTPSTNSPPTPAPRQSSP	492
NP_705885.1	443	FSSYFNKVGVAQDTDDCDDFDISQLNLTPVTSTTPSTNSPPTPAPRQSSP	492
XP_002686414.1	443	FSSYFNKVGVAQDTDDCDDFDISQLNLTPVTSTTPSTNSPPTPAPRQSSP	492
XP_003434697.1	443	FSSYFNKVGVAQDTDDCDDFDISQLNLTPVTSTTPSTNSPPTPAPRQSSP	492
XP_001114475.1	441	FSSYFNKVGVAQDTDDCDDFDISQLNLTPVTSTTPSTNSPPTPAPRQSSP	490
NP_066566.3	443	FSSYFNKVGVAQDTDDCDDFDISQLNLTPVTSTTPSTNSPPTPAPRQSSP	492
XP_001155092.2	443	FSSYFNKVGVAQDTDDCDDFDISQLNLTPVTSTTPSTNSPPTPAPRQSSP	492

NES1 of MLXPL_RAT

HomoloGene:32507

NP_001104311.1	1	--MAEAQLGLPGPFAQPPVGPCTPDSDFSDPEDAAAGGAGPGAAVQNPSQ	48
NP_067430.2	1	--MARAL LADLSVNLQV PRVVPSPDSDSDTDLEDPSPRRSAGGLH---RSQ	45
NP_116569.1	1	--MAGALAGLAAGLQVPRVAPSPDSDSDTDSEDPSLRRSAGGLL---RSQ	45
XP_003318716.1	1	--MAGALAGLAAGLQVPRVAPSPDSDSDTDSEDPSLRRSAGGLL---RSQ	45
XP_001115131.2	1	--MAGALAGLAAGLQVPRVGPSPDSDSDTDSEDPSLRRSAGGLL---RSQ	45
XP_546925.2	1	--MAGALAGLAAGLH GPRGVPSQDS SDTDSEDPSARRSAGGLL---RSQ	45
NP_001192337.1	1	MTMAGALAGLVAGLQ GPRVVPS PDSDSDTDSEDPNTRRSAGGLL---RSQ	47

NES2 of PCNT_HUMAN

HomoloGene:86942

XP_002662350.2	2080	ATQQRQDLQTA RDKLQ RIVKEMLRITITTEEHISHKLGSCVSGGQ---SE	2126
XP_421895.3	1419	SKDERKKMQDS YQCVLKW LVELVKATIAVEDLICSKIGLCLDNSMASADS	1468
XP_548735.4	1330	SDRERRALRDALRRLLSIFGEMLKVAIALKSRIERVGCLCLEDPPGAW	1379
XP_002685207.2	1061	SHGERRALQDTLRRLLGLFGETLQTAIALKSRIERVGCLCLEDSPPTTQ	1110
NP_006022.3	1151	SHSERGA LQDALRRLLGL FGETLRAAVTLRSRIERVGCLDDAG---A-	1196
XP_003319184.1	1086	SHSERGA LQDALRRLLGL FGETLRAAVTLRSRIERVGCLDDAG---A-	1131
NP_032813.3	925	AQGEGKALRDALRRLLDLFGDTLKAATLKSRIERAGLLDHED---AA	971
XP_002728982.1	927	AQSEGKALRDALRRLLDLFGETLKAATLKSRIERAGLLDHED---AV	973

NES4 of PCNT_HUMAN

HomoloGene:86942

XP_002662350.2	2469	QVESLQALIKDKMEDY SVLLLAKEQCQRDVEERNDEIEKLATIRELEQA	2518
XP_421895.3	1783	GVESLQAEVEEKVDDY NKLLKEQKHKQEIAARDKEIEKL-----	1822
XP_548735.4	1715	EVELLQEKLREKSDGFNELVIKKELADRQVMIQEEEIRRLEE-----	1756
XP_002685207.2	1446	EVELLQEKLREKSDGLNELILKKELVDRQVLIQGEEIRRLEE-----	1487
NP_006022.3	1528	EVELLQQKLREK LDEFNELAI QKESADRQVLMQEEEIKRLEE-----	1569
XP_003319184.1	1463	EVELLRQKLREKLDEFNELAIQKESADRQVLMQEEEIKRLEE-----	1504
NP_032813.3	1305	EVELLQEKLREKLDFNELVIKKDFADQQLIQQEEEIKRLEE-----	1346
XP_002728982.1	1307	EVELLQEKLREKLDFNELVIKKDLADQQLIQQEEEIKHLEE-----	1348

NES5 of PCNT_HUMAN

HomoloGene:86942

XP_002662350.2	2718	IEELRSLVERLQCDQERLQQAKEEEMEQLHEVINKLQEEIS Q LDPNHHEI	2767
XP_421895.3	1958	IEELRSVIEYLRGDQERLCKDKDEEVQLHEVIEKLQKELA Q IGPVCHEV	2007
XP_548735.4	1901	IEELRSIIENLRENQQRLQKEKAEEEMEQLHEVIERLQEELSLGAPAVPAA	1950
XP_002685207.2	1632	VEELKSIIETLRENQARLQKDNAEEIEQLHEVIEKLQRELPLGGPEAPKA	1681
NP_006022.3	1714	IEELKATIENLQENQKRLQKEKAEEIEQLHE VIEKLQHEL SLMGPVVHEV	1763
XP_003319184.1	1649	IEELKATIENLQENQKRLQKEKAEEIEQLHEVIEKLQHEL SLMG PVVHEV	1698
NP_032813.3	1439	IDELKSLIENLQENQRQLQKDKAEEIEQLHEVIEKLQSEL SLMGPK VHEV	1488
XP_002728982.1	1384	IEELKAIIEHLQENQEQLQKAKAEEIEQLHEVIEKLQSEL SLMGPT VHEM	1433

NES2 of MALT1_HUMAN

HomoloGene:4938

NP_495424.2	382	AISAFVEYSEQFKKF <u>HRAHRNMVGY</u> STSGVGAYEVK--GEVNGVFMKY	429
XP_690409.3	436	DAMPNV-----MLKVTANIVFGYATCQDAEAFELSSSGFTNGVFIKF	477
XP_003642952.1	456	DTILILD-----ALKVTANIVFGYATCQGAEAFEIQQLGLANGIFMKF	498
XP_533392.2	632	DTIPILD-----ALKVTANIVFGYATCQGAEAFEIQHSGLANGIFMKF	674
NP_766421.1	469	DTIPILD-----ALKVTANIVFGYATCQGAEAFEIQHSGLANGIFMKF	511
XP_225927.5	469	DTIPILD-----ALKVTANIVFGYATCQGAEAFEIQHSGLANGIFMKF	511
XP_002697882.1	707	DTIPILD-----ALRVVTANIVFGYATCQGAEAFEIQHSGLANGIFMKF	749
XP_001086151.1	472	DTIPILD-----ALKVTANIVFGYATCQGAEAFEIQHSGLANGIFMKF	514
NP_006776.1	472	DT <u>IPILD</u> -----ALKVTANIVFGYATCQGAEAFEIQHSGLANGIFMKF	514
XP_523938.2	472	DTIPILD-----ALKVTANIVFGYATCQGAEAFEIQHSGLANGIFMKF	514

NES1 of ATF2_HUMAN

HomoloGene:31061

XP_001345237.4	1	-----MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	32
NP_990235.1	1	-----MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	32
NP_112280.1	1	-----MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	32
NP_001075053.2	1	MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	50
NP_001020264.1	1	-----MSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	32
NP_001871.2	1	MKFKLHV NSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	50
XP_001151932.1	1	MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	50
XP_535970.2	1	MKFKLHVNSARQYKDLWNMSDDKPFLCTAPGCGQRFTNEDHLAVHKHKHE	50

NES1 of TOP2A_HUMAN

HomoloGene:830

XP_369302.1	1102	QKNVAITNLVAFDTNGQIRKYDKVEDILEEFYHYRLNVYAARKAHWLRVF	1151
XP_962372.2	1086	TKTVATTNLVAFDTRGKIHKYENPQEIMEEYYHYRLNLYGERKKHWLKVY	1135
NP_985265.1	975	ISTVSLSNMVLFDAHNKIKKYNDVKEILTDFFYVRLEFYQKRKDYMIE RL	1024
NP_014311.1	967	ISPISLMNMVAFDPHGKIKKYNSVNEILSEFYVRLEYYQKRKDHM SERL	1016
XP_452818.1	971	VSTVSLSNMVAFDPQGKIKKYDDVRDILSDFYYTRLTYYQKRKDYG ERL	1020
NP_496536.1	1028	QAVINTTCMVLFDAAGCLRTYTSPEAITQE FYDSRQE KYVQRKEYLLGV L	1077
NP_001003834.1	995	QNLLTCNSMVLFDHVGSLLKKYESVQDILKEFFELRMKYYVLRKD WLVGML	1044
NP_990122.1	993	QTNLTCNSMVLFDHVGFLLKKYESPQDILKEFFELRLRYYGLRKEW LGML	1042
NP_035753.2	991	QSSLTCNSMVLFDHVGC LLKKYDTVLDILRDFFELRLKYYGLRKEWLLGML	1040
NP_071519.2	990	QTSLTCTNSMVLFDHVGC LLKKYDTVLDILRDFFELRLKYYGLRKEWLLGML	1039
XP_002696055.2	992	QTSLTCTNSMVLFDHVGC LLKKYDTVLDILRDFFELRLKYYGLRKEWLLGML	1041
XP_537646.2	992	QTSLTCTNSMVLFDHVGC LLKKYDTVLDILRDFFELRLKYYGLRKEWLLGML	1041
XP_002808175.1	959	QTSLTCTNSMVLFDHVGC LLKKYDTVLDILRDFFELRLKYYGLRKEWLLGML	1008
NP_001058.2	992	QTSLTCTNSMVLFDHVGC LLKKYDTVLDILRDFFELRLKYYGLRKEWLLGML	1041
XP_003315524.1	961	QTSLTCTNSMVLFDHVGC LLKKYDTVLDILRDFFELRLKYYGLRKEWLLGML	1010
NP_476760.1	974	TTTLSTNQMHAFDQNNCLRRFPTAIDILKEYYKLRREYYARRDFLV QL	1023
XP_319665.4	973	TSSISTASMNAFDDKNYLKRYAHANEIFMEFYTVRLDFYGKRRAYLLGML	1022
XP_001348490.1	1013	KSTLTNTNMTLFDPNLKLQRYSTEILDILKEFCYQRLKAYENRKSYLISK L	1062
NP_595805.1	1021	SRTQATSNMIAFDASGRIKKYDSVEDILTEFYEVRLRTYQRRKEHMVN EL	1070
NP_189031.1	982	TTTIATSNMHLFDKKGVIKKYVTPEQILEEFFDLRF EYYEKRKETVVKN M	1031
NP_001047833.1		-----	

NES1 of NPM_HUMAN
HomoloGene: 81697

NP_955460.1	1	-----MDLEQM-----PQTFLYGCELKAGKDITFNPEDDDYDHQLSVRM A C	42
NP_990598.1	1	MEDSAMDME SMGPLRPQTFLFGCELKAEKEYQFKVDEENEHQLSLRTVT	50
NP_032748.1	1	MEDS-MDMD-MSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRTVS	48
NP_037124.1	1	MEDS-MDMD-MSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRTVS	48
NP_001030518.1	1	MEDS-MDMD-MSPLRPQNYLFGCELKADRDYHFKVDNDENEHQLSLRTVS	48
XP_001092135.1	1	MEDS-MDMD-MSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRTVS	48
NP_001239100.1	1	MEDS-MDMD-MSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRTVS	48
NP_002511.1	1	MEDS-MDMD-MSPLRPQNYLFGCELKADKDYHFKVDNDENEH QLSLRTVS	48
XP_001151979.1	1	MEDS-MDMD-MSPLRPQNYLFGCELKADKDYHFKVDNDENEHQLSLRTVS	48
NP_955460.1	43	VDPSTKDELNVVEIEGQDSEGQKVKAVLATLKPSLPSVCLGGFEITPPV	92
NP_990598.1	51	LGAGAKDELHVEAEALDYEGNPTKVVLASLKMSVQPTVSLGGFEITPPF	100
NP_032748.1	49	LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPV	98
NP_037124.1	49	LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPV	98
NP_001030518.1	49	LGAGAKDELHVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPV	98
XP_001092135.1	49	LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPV	98
NP_001239100.1	49	LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPV	98
NP_002511.1	49	L LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPV	98
XP_001151979.1	49	LGAGAKDELHIVEAEAMNYEGSPIKVTLATLKMSVQPTVSLGGFEITPPV	98

NES1 of CHP1_RAT

HomoloGene: 5235

<u>NP_566610.1</u>	107	CNGKVSFKDIMEVLRDLSGS F MSDEQREQVLSQLKESGYTSDSFLTLED	156
<u>NP_001061916.2</u>	107	CNGKVA <u>FDDILSILRDLSGSF</u> MTEQQRQKVLIHVLEEAGYTKDSNFTLPD	156
<u>NP_505623.1</u>	126	KNDYITREEFKVILNSMVG A NITS <u>DQLDKIADRTIEEADADRGKISFDE</u>	175
<u>NP_503830.1</u>	144	KNNYITREEFKVILNSMVG A NITS <u>DQLDKIADKTLEEADQDRDGKISFED</u>	193
<u>NP_956009.1</u>	125	RDDKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADTNGDMCISFNE	174
<u>NP_001007931.1</u>	126	KDDKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGCAISFAE	175
<u>NP_062743.1</u>	126	KDDKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGDSAISFTE	175
<u>NP_077053.1</u>	126	KDDKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGDSAISFTE	175
<u>NP_009167.1</u>	126	KDEKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGDSAISFTE	175
<u>XP_523056.2</u>	126	KDEKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGDSAISFTE	175
<u>XP_001099956.1</u>	126	KDEKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGDSAISFTE	175
<u>XP_535438.1</u>	126	KDDKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGDSAISFTE	175
<u>NP_001069044.1</u>	126	KDDKISRDELLQVLRMMVGVNISDEQLGSIADRTI QEADQDGDSAISFTE	175
<u>NP_649568.1</u>	120	DDGVISRDELLSILHMMVG A NIS <u>SQDQLVSIAERTILEADLCCQGKISFED</u>	169
<u>XP_321323.3</u>	120	DDETISRDELLNILQMMVG A NIS <u>SQDQLNSIAERTIVEADTVGVGKISFDD</u>	169

NES2 of CHP1_RAT

HomoloGene: 5235

<u>NP_566610.1</u>	157	FIKIFGS <u>SRPEMDVEIP</u> -VD	175
<u>NP_001061916.2</u>	157	FMKILDNS <u>SDVKMEEV</u> P-ID	175
<u>NP_505623.1</u>	176	FCRAMEK <u>T</u> DIEEKMSIRFLN	195
<u>NP_503830.1</u>	194	FCRAMEK <u>T</u> DIEEKMSMRFLN	213
<u>NP_956009.1</u>	175	FTKVLEKVDVEQKMSIRFLH	194
<u>NP_001007931.1</u>	176	FVKVLEKVDVEQKMSIRFLH	195
<u>NP_062743.1</u>	176	FVKVLEKVDVEQKMSIRFLH	195
<u>NP_077053.1</u>	176	FVKVLEKVDV EQKMSIRFLH	195
<u>NP_009167.1</u>	176	FVKVLEKVDVEQKMSIRFLH	195
<u>XP_523056.2</u>	176	FVKVLEKVDVEQKMSIRFLH	195
<u>XP_001099956.1</u>	176	FVKVLEKVDVEQKMSIRFLH	195
<u>XP_535438.1</u>	176	FVKVLEKVDVEQKMSIRFLH	195
<u>NP_001069044.1</u>	176	FVKVLEKVDVEQKMSIRFLH	195
<u>NP_649568.1</u>	170	FCKALDR <u>TDVDQKMSIRFLN</u>	189
<u>XP_321323.3</u>	170	FCRALERT <u>TEVEQKMSIRFLN</u>	189