

A. Teraspanin uroplakins (UPK1a and UPK1b)

>UPK1a.human. NP_008931.1
MASAAAAEAEKGSPVVVGLLVGNIIILL**LSGLSLFAETIWVTADQYRVYPLMGVSGKDDVFAGAW**
IAIFCGFSFFMVASFGVGAALCRRRSMVLTYLVLMIVYIFECASCITSYTHRDY**MVSNPSLITK**
QMLTFYSADTDQGQELTRLWDRVMIEQECCGTSGPMDWVNFTSAFRAATPEVFPWPPLCCRRTG
NFIPLNEEGCRLGHMDYLFTK**GCFEHIGHAIDS**YTWGISWFG**FAILMWT**LPVMLIAMFYTML

>UPK1a.mouse NP_081091.1
MASAATEGEKGSPVVVGLLVGNIIILL**LSGLALFAETVWVTADQYRVYPLMGVSGKDDVFAGAWI**
AIFCGFSFFVVASFGVGAALCRRRYMILTYLLLMLIVYIFECASCITSYTHRDY**MVSNPSLITKQ**
MLTYYSADTDQGQELTRLWDRIMIEQECCGTSGPMDWVNFTSAFRAATPEVFPWPPLCCRRTG
FIPINEDGCRVGHMDYLFTK**GCFEHIGHAIDS**YTWGISWFG**FAILMWT**LPVMLIAMFYTTL

>UPK1a.cow NP_788784.1
MASAAAAATTEKGSPVVVGLLVGMNIIILL**LSGLALFAETVWVTADQYRIYPLMGVSGKDDVFAGAW**
IAIFCGFSFFVVASFGVGAALCRRRSMILTYLILMLIIYIFECASCITSYTHRDY**MVSNPSLITK**
QMLTFYSADSNQGRELTRLWDRIMIEQECCGTSGPMDWVNFTSAFRATTPEVFPWPPLCCRRTG
NFIPIVNEEGCRLGHLDYLFTK**GCFEHIGHAIDS**YTWGISWFG**FAILMWT**LPVMLIAMFYTTL

>UPK1a.opossum XP_001370098.1
MAGEKGSPTVVGLLVGNIIILL**MSGLALFAETVWVTADQYRVYPLLGVSGKDDVFAGAWIAIFTG**
FSFFVVASFGIGAVLCHSRGMLLTYLLLMLIVYVFECASSITSYTHRDY**VVSNPALITKQMLTY**
NAPTSQGQELTRFWDRVMIEQQCCGTSGPLDWVNFTSAFREATTPEVFPWPPLCCRRDANFNPLN
EDGCRLGHQDYLFTK**GCFEHIGHAIDS**YTWGISWFG**FAILMWT**LPVMLLAMYYTIL

>UPK1a.aligator profiling AKHW01101299.1| AKHW01101300.1
scaffold-12507_11
MAEKGNPLVVGLLILGNLIILL**LAGLALYAETIWVTADQYKVYPILGVSGKDDVYAGAWIAIFCGF**
AFFCLGVLGIALARGSRPLLMVYLALMVIVYIFESASCITSYTHRDF**VVSNPRLVTKQMLSFYA**
APSRQGRELTRMWDRIEMEQCCGTSGPLDWVNNTSAFRARFPEVAAPWPFLCCRRDPNFVILNQ
EGCRVGHVDYVNTK**GCFEHIENAVNS**YTWGISWFG**FAILMLA**CPVMLLGMYYHTTL

>UPK1a.lizard XP_003225012 gi|294743809|gb |AAWZ02022979.1|
cont2.22978
MAAKGSTLVVTLVGMGNVIIM**LAGLALYAESIWVTADPYKVYPIMGVSGKDDVYAGAWISIFTG**
SFFCVCVFGIISLVNGNRLMVLLYLVLILVVYLFECASCITSYTHRDF**VVSNRLITKQMLSYS**
EDSYQGRELTRMWNRFMEEQCCGTNDPMDWVNFTSVFRRRYPEIVAPWPFCCIRDRNFIIINE
EGCRLGHVDYINTQ**GCFEHIKHAVQSYAWGVS**WFG**FAILMLTV**CPVLLAIYHYTTM

>UPK1a.salamander JK981846.1
MADKGSPGLVLLILGNVVITMSGLALFAETIWATTDPFKVYPILGVTGKDDVFAGGWIAIFCGF
SFFMLGIYGVLATLRGSRTMVMTYLVLMIVYIFECASCITSFTHRDYMINSNVQKQMLQYYTD
SSPQGRELTLWNRMMEKQCCGVKGPADWIQYTSTFRNQYTEFVAPWPFCCIRDRNFIIINE
GCRVGAEGYIYTQGCWDHISNAINSYTWGISWFG**FAILMWTLLV**MLAEMFYTKL

>UPK1a.frog profiling gi|268046463|gb|AAMC01106311.1| |267885090|
gb|AAMC01156946.1|
MAEKGSSTMVACIVFGNIVILL**LSGLALFAETIWATTDPYKVYPILGVTGKDDVFAGGWIAIFCGF**
SFFILGVFGILAVQRGSRTMVLTYLVLMIVYIFECASCITSFTHRDY**MINSNVIKQMLTYSD**
SSTPQGRQITSVWLRMMLEKNCCGVDGPLDWVDYYSYFRQSYNETTAPWPLWCCQRDGNFQILNQ
QGCIVGLSSVYQQ**GCAHISNAINSY**TWGISWFG**FAILMWT**MFVMLATMYHYTKM

>UPK1a.coelacanth Ensemble ENSLACP00000006022;
ENSLACT00000006074
MADGKGGSSVGALLIFGNLIILL**MAGLALFAETIWNTDEFKVYTF**LGVS**GKDDV**

FAGAWIAIFCGFCFFLLGTFGIFAVLKQSRMTVMTYLILMLIVYIFETASCITSFTHRDIYVSNP
NFLKKQMLQLYTSNTSQGIELTEAWNRMVLEECCGVEGPMDWISFSSTYQNSISSPSQTSWPLY
CCKRDKNFIMLSELACLIGHKDFVFSNGCWDFLSFSVNKYTWGVSWFGFAILMWTFFVMCLKMYF
YTII

>UPK1a.zebrafish BE200967; AI396779; E0017774
MGAVTCLMVTVVGLNAIAAAAGLALSVAIWWAVDGYKLYPISGVSGKDDIFAGAWIAIFTGFAF
FLTCIFGIFAALKRSRALMIYLIIMFIIFLFESASAITSATNRDYLVGNSNLVKKQMLQYYADSS
TQGQQITMTWNNVMTQVQCCGADGPTDWIQYNSTYRQLFGAASLWPLGCCKRQSSNFEVVDPIGC
KAGVTSSMFTQGCQFYIESVLSRYTWAVSWYGFVSLMLVFFTLVIAMIYYTQL

>UPK1a.salmo profiling gi|354428581|gbAGKD01030455.1
Contig_030458.
MANGKGFLLIGNIFGAAGLALCALAIWVAVEQYKLYPIAGVSGKDDIFAGAWIAIFTGFAFFCI
CVFGILAATKKSRLMLTYLILMLIIYIFECASCITAVTHRDLVGNSNLVKKQMLKYAAEGDSG
SRITLTWNKVMNEVECCGTDGPVDWIQYNSTFREKFSTDYPWPIHCCKRKNNYQVVNVKACKNGQ
NTTMFTKGCFNHIESVFSLYTWAISWYGFVLMFM

>UPK1A.CatFish. CK412483
MVVVVILNVVAAAAGLALCALAIWVAVDPYKVYPISAVSGKDDIFAAAWIAIFTGFAYFCTAIFG
IYAALKRKRSLVLLYLILMFIIIFIFECASCITAVTNRDYLIGNSNLVKNQMLKYAQDSNQGRQI
TGTWNKVMNDAQCCGTDSPMDWIEYNSTFKQTYGSTYTWPLNCCQRQNSFDPADPVGCMFGQTSA
VFSKGCFFNYIQTVLNRYTWSVSWYGFVQMFVFFLLITIVYFLLLE

>UPK1A.Carp EC392385.1, AU301724
MGAGALTCLMFVVVALNAIAAAAGLALFAVAIWWAVDGYKLYPISGVSGKDDIFAAAWIAIFTGF
AFFLTCTIFGIFAALKRSRALMMVYLIIMFIIFLFECASAITAATNRDYLVGNSNLVKKQMVQYYG
QDSIQGKQITQTNRMVMEQVECCGADSPEDWILYNSTFKQIYRTYNWPLSCCKRLSSFELEDPA
CKVGLTSPVFTKGWFNYIESVLSRYTWAVSR

>UPK1a.spottedgar profiling gi|363782398| gb|AHAT01019516.1|
contig019516,
MSEGRGSTLMMALLVAGNLFLALCGLALYAVAIWWATDGYRLYPLSAVSGKDDIFAGSWIAIFTG
FAFFCAAVYGVFAALRESRAMMLLYLVLMVIYIFEAAASITAATHRDYLVGNSNLIKKQMLTY
ADSDPGRQVTTTWNRVNCCGTDGPLDWISYNSTFRSKFPTQEYPWPLNCCKRKDNFEVLNLDAC
RIGDWNMYMNYKGCFDHIEFVFNQYAWAICWYGFVLMFVLPLMLLAMVYYLKL

>UPK1a.shark EB688237.1; ES415544.1; CV889144.1|
MGETSASPVLSVLIFGKLVLMAGIALFAETIWWVTDQFQVYPVLGASGKDDVFAGAWIAIFVG
FAFFCLAVFGILAVLSESRMTVITYLALMLVYIFECASCITSITHRDYLTSPNPKFIKKQMLQFY
GDTSSNGGRDLTTMWNHVMPEKQCCGSTGPADWIQYTSFFRTMYNETFAPWPFQCKRNANSQNH
QAACAVGHKDFL

>UPK1a.lamprey EE740596; UPK1aTRACE name:PMAH-aab87f10.g1;
Ensemble scaffold: Pmarinus_7.0:GL480986:-1269:28785:
MAKESNAGCVQLVIIGNVILLCGITLTAETIWWVTDGYKVYPILGMAHNDVDFAGAWIAIFTG
MAFFLLGFVGIVAALRMTRKLLLYIVCMLIVFAFESASCITSFTQRDYLVDQNFLLKQMLTEY
PSDTYPAYTDTWNMFMRQKCCGSNGPTDWLVYTSKFSEAHNDDVNFPMHCCVLSKDGPPLN
ITYCRLGMDGYVNTAGCFDYFSAAVNRYTWGVAVSDSHLCFTFFVLSGAMYLYTVA

>UPK1a2.lamprey EG022567, profiling Ensemble
ENSPMAT00000009881;scaffold:Pmarinus_7.0:GL476495:371924:395954:-
1
LCGMALMATAIWWVTDPYKIYPFLAADNNTDIFAAAWIAIFCGFAFFLLGIFGMYAVWKMQRSAL
LAYFILMLIVFIFEAAASCIVIFTHRDYLVGTKNLMLKHLRDYGEIPSFTDNWKDLMQKCCGVNG
PEDWISYTSSFSIYHKEDDADSPWPTKCCVLNENDDYLYGRMGCILGKPDIAIFQKGCYGFSSAV
DGYTFPLAWFGFSILVLFVEIGTLYLYTVS

>UPK1b.human. NP_008883.2
MAKDNSTVRCFQGLLIFGNVLIIGCCGIALTAECIFFVSDQHSLYPLLEATDNDDIYGAAWIGIFV
GICLFCLSVLGIVGIMKSSRKILLAYFILMFIVYAFEVASCITAATQQDFFTPNLFKQMLERYQ

NNSPPNDDQWKNNGVTKTWDRLMLQDNCCGVNGPSDWQKYTSAFRTENNDADYPWPRQCCVMNN
LKEPLNLEACKLGVPGFYHNQGCYELISGPMNRHAWGVAWFGFAILCWTFWVLLGTMFYWSRIEY

>UPK1b.mouse. NM_178924

MAKDDSTVRCFQGLLIFGNVIIGMCSIALMAECIFFVSDQNSLYPLLEATNNDIYAAAWIGMSV
GICLFCLSVLGIVGIMKSNRKILLVYFILMFIVYAFEVASCITAATQRDFFTPNLFLKQMLERYQ
NNSPPNDDQWKNNGVTKTWDRLMLQDNCCGVNGPSDWQKYTSAFRTENSADADYPWPRQCCVMNS
LKEPLNLDACKLGVPGGYHSHGCEYELISGPMNRHAWGVAWFGFAILCWTFWVLLGTMFYWSRIDY

>UPK1b.cow NP_776907.2

MAKDDSTVRCFQGLLIFGNVIIGMCSIALMAECIFFVSDQNSLYPLLEATNNDIYAAAWIGMFV
GICLFCLSVLGIVGIMKSNRKILLVYFILMFIVYAFEVASCITAATQRDFFTPNLFLKQMLERYQ
NNSPPNDDQWKNNGVTKTWDRLMLQDNCCGVNGPSDWQKYTSAFRTENSADADYPWPRQCCVMNS
LKEPLNLDACKLGVPGGYHSQGCYELISGPMNRHAWGVAWFGFAILCWTFWVLLGTMFYWSRIDY

>UPK1b.opossum XP_001363056.1

MAKGDSTVRCFQGLLIFGNVIIGMCGIALTAECIFFVSDQHSLYPLLEATNNDIYGAAWIGLHV
GICLFCLSVLGIVGIMKSSRKILLAYFILMFIVYAFEVASSITAQVQDFFFTNLFLRQMLERYQ
NKSQSSNDDIWKNNMAVTETWDRLMLQEKCCGVNGPSDWQKYKSAFRDENNDADYPWPHQCCVMDN
LYKPLNLDACKLGVSQGYHSHGCEYELISGPMNRHAWGVAWFGFAILCWTFCVLLGTMFYWSRIEY

>UPK1b.gallus XP_416567.3; DR426684.1; BU224014.1

MAKTDNGIRICQGLLILGNVVIIGMCGIALTAECIFFVSDPHGLYPLLEATENNDIYAAAWIGIFV
GFALLALSILGIVGVMKSSKTLVYIILMLITYAFEMASCITAATHRDFTPNLFLKQMLERYM
KSDTDNNDKQMTGVTKTWDNLMLQNCCGVHGPLDWQEYTSAFRMTNDADYPWPHNCCVMNT
RSEPINLDGCKLGVPGFYNSNGCYDAISGPLNRQAWGVAWFGFAILCWTFCVLLGTMFYWSRIAY

>UPK1b.lizard FG709884, FG750007; gb|AAWZ02001347.1|este exons
3-7; gb|AAWZ02001348.1| FIRST EXON chromosome 3 cont2.1347,
MVKGDTGVRFCFQGLLIMGVLGCCGLALMAECIYFVSDQHSLYPLLEATNDDIFGAAWIGIFTG
CFFCLSVLGIVGVKSNRTMLLVYILLMMVFAFEVASSITAVVHRDFTPNLFLKQMLEKYQNP
NPINNDNRWKSOGITRTWNRLMVLDDCCGVTGPRDWQNYTSVFRTVSNDAFPWPRQCCVIDVLG
KPTNLDGCKLGVSQGYHEMGCYDTISGPMNRHAWGVAWYGFVLCWATWVLLCTMFYWSRIEY

>UPK1b.turtle Ensembl. ENSPSIP00000002711; gb| AGCU01013879.1|
scaffold639_322,

MAKGDNGVRFFQALLILGNVIIGMCGIALTAECIFFVSDQQTLYPLLEATNNDIYGAAWIGIFV
GFALFNLSVLGIVGVKSNRTMLLAYIILMLITYGFEVASCITAATHRDFTPNLFLRQMLEKYQ
NPEPASNDDKWMSEGVTRTWDRMLQNCCGVNGPSDWQNYMSVFRKGNDSSEFPWPRQCCVMDV
QGIPINLDGCKLGVSQGYNNKGCYDIISGPMNRHAWGVAWFGFAILCWTFWVLLICTMFYWSRIEY

>UPK1b.aligator profiling 397239340|gb|AKHW01070659.1| scaffold-
8202_6,

MAKGDNGVRFCFQGLLVGNVIIGLCGIALTAECIFFVSDQQTLYPLLEATNNDIYGAAWIGIFV
GFALFNLAFLGIVGVKSNRAMLLAYIILMLITYAFETASCITAATHRDFTNRNLFKQMLEKYQ
NPEPLNDDKWMSEGVTRTWDRMLQHECCGVNGPSDWQNYMSKFRNINSSEFPWPRQCCVMNV
QGIPVNLDGCKLGVSQGYNDKGCYELIAGPMNRHAWGVAWFGFAILCWTFWVLLCTMFYWSRIEY

>UPK1b.frog. NP_001037968.1|

MKEDSGVRCYQSIIIFGNVVLGLCGVALTAECIFFVSDQSGIYPLLEATNNADIFAAAWIGIFTG
FCFFILSIVGIIGIMKSNRRMLMVYILMFVYAFEVASAITAATQENFFIPNLFLKQMLDFYQN
PNPTNNDNLWKINGVTNTWNRFMLLNGCCGVNGPQDWQTYTSVFRQSNDSAYPWPQCCAMNSL
GQPVNLDACKLGVSQGYVNLNGCYDQMGAPMTRHAWGVAWFGFSILCWTFWVLLGSMLYWTRIEY

>UPK1b.salamander JK975385 CN045036

MAKDAGVQFFQGLLIFGNVVIAMCGIALTAECIFFVSDQSSLYPLLEATNDDIFGAAWIGIFTG
FCLFVLSIVGIVGVMKSNRRMLLVYILMFIVYCFEVASSITAATQRDFFIPNLFLRQMLELYQN
PNPVNNDQAWKINGVTATWNRLMLLNCCGVNGPTDWQAYASSFRVTNSDSDFPWPQCCVMNDL
GQPENIDGCKLGVSQGHYYNTGCYDLIAGPLMRHAWGVAWFGFAILCWTFWVLLGSMFY

>UPK1b.coelacanth Ensemble ENSLACP00000016648;
ENSLACT00000016762MALKGSSPVRCFQGLLIFGNVVIALTGLTLFAACIFFLSDDNRYRL
LDASENDDIFAAAWISLFGVGSFFLLSILGIIGVMKSNRTMLLVYIILMLVVYCFEVASCITAIT
HRDFFIPNMFLKTLLEKYQKPKARDLDEQKIDGVTTIWNRIIMPQHKCCGVNGPMDWQEYNSSFR
MQNSDADFPWPRQCCVRAINQEPQNLDAKLGIVGSYYAQGCFEFIAGPLNRHAWGVAWYGFAIL
CWVFWVLAVTMFYFSRIDSTE

>UPK1b.trout FP322998, CU068699
MTPDLDTKDRCFRALLIQGNLVIACCGIALMAMCIIFFISDRAGLYVLVYATGNDSIWRGAWIGLF
TGFALFCTSIIFGMHAIIVVSKRNILLAYILLMVIIYAFEVASAITAATHKDWFPFLKQMLQNY
NKPLPDNL PSTQDQIYVINGITEAWNRFMTESKCCGVYGPEDWVKYESHFRQNTDADNPWPRQC
CQQDAMGAISHLEACKIGVSPFLHAQGCYDYIAGPLITHGFGVSWFGFAILCWTFFVILGVIFHY
TQLDI

>UPK1b.salmo NP_001134735.1; EG792520; EG874850; gi|354302482
|gb|AGKD01156046.1| Contig_156071; gi|354418425
|gb|AGKD01040611.1
Contig_040615, MTPDLDTKDRCFRALIIWGNLVIACCGIALMAMCIIFFISDRGGLYVLVYAT
GNDSIWRGAWIGLFTGFALFCTSIIFGMHAIIVVSKRNILLAYILLMVIIYAFEVASAITAATHKDW
FVPDLFLKQMLQNYNKPLPDNL PSTQDQIYVINRVTETWNRFMTESKCCGVYGPEDWVKYESHFR
QQNTDADYPWPRQCCQDAMGAISHLDACKIGVSPFLHAQGCYDYIAGPLIRHGFGVSWFGFAIL
CWTFFVILGVIFHYTQLDF

>UPK1b.spottedgar profiling gi|363798646| gb|AHAT01003268.1|
contig003268,
MVAKSDIGVRCFQAFLOWGNFCCGLALMAMCIIFFIFDRDHLVVLVYATGNDSIWRGAWIGLFTGF
ALFCTAVFGMYAVLKS KRGLLLVYILLMIFIFAFEVASAITAATHKDW FVPNLFLKQMLEYNKPL
LPQNL PSTQDEIYKNSGITSTWNMIMTEYQCCGVYGPQDWLQYNSQFRQQNSDSEYPWPRQCCLQ
DGTGGIVNVNACKIGVEPYLFTQGCYDYIAGPLIRQGLGVSFWFGFAILCWTFFVIGVMFYTTQL
DSZ

>UPK1B.skate CV547311;.C0049535; C0050503; AESE010043837 exons 1-
5 gi|363355253|gb|.1| LER_WGS_1_CONTIG_4389;
gi|363324303|gb|AESE010074792.1| CONTING EXON 6
LER_WGS_1_CONTIG_74854.
MSKSTGVRVQGLLIFGNVIMLSGLALTAECIFHVSDQHNLWPFLAAADNSDIFAAAWIGL FAG
FCLFCLSIYGIFGIMKS AKKVILTYLILMLIVYIFEVASCITAATQRDFFVPNFFLKQMLQLYGN
PNPLTEAEIRNTFGVTNAWNRVMPDNDCCGVNGPRDWVQHNSAFRVVNSDSRFPWPRQCCTTDQF
GRMLNVTGCKLGIPGYVE

>UPK1b.Lamprey FD710561; EB084275 1220567368gi; TRACE gn1|ti
|1470569568 reverse exon4 name:PMAC-bnw68a06.g1 mate: 1470753301;
gn1|ti|1423504665 name:PMAC-axy03h10.g1
7MRDVSSGVRCVQILLIVGNVLLGCGLALFAVSVFVADPARLWPLLSAFNNTDIFAGAWISIF
TGFTFFLLAMLGIYAVMRLARGLLLAYILLTLFVYIFEVASCITAITHRDFLTKNAYLQQLQLY
QNPNWVDENSRFNIEGVTAVWDRTMFNYQCCGVNGPQDWQAYGSAFREAHPGDFPWLQCCSLD
PASGLIRDLDACRLGLAGYLNTPGCYDYVSGPLHRAAWGVAWFGFAILCWEFWLLGTMVLYNIS

>UPK1a.platypus exons 2-6 but exon 5 coo intron Boundary CT instead of GT;
missing exons 1 and 7; profiling AAPN01309122.1 plus AAPN01141210.1
Cont44002.2,
LSGLALFAESVWASDPYRVYPTLGVS GKDDVFAGAWISIFTGFAFFLVGSLGLVALLRRSRPMV
LTYLVLMIVYVFESASSITAHTHRDYMVSDPSLITKEMLTHYGAPSRQGRELTRLWDSIMVEKE
CCGTGLGPLDWVHYTSTFRTSIPEAVFPWPPPCRRDPNFIPLSEEGRVGRHDYIFTKGCFEHI
HAIDSYAWGVSWFGFAILMWT

>UPK1b.platypus Exon 3 and 6 missing 24 and 28 aa; profiling
AAPN01272338.1 reverse exons 1,2 | AAPN01106002.1|exon 3
AAPN01297767.1 exon4 AAPN01302484.1 exon 5 AAPN01104274.1 exon
Cont11876.1.

MAKDDGSRCLQGLLVFGNVVIA**MSGLALMAECIFFV**TD**WPRLYGLLEATD**ND**DDVFAAAWIGIFT**
GFSLFCLAAALGVVAVLRAGRKLLLAXXXXXXXXXXXXXXXXXXXXXXXXXXXX**FTPNLF**L**RQMLERYQ**
NQSGPTNDDQYKNKKVTD**TDWDRMLQ**EKCCGVNGPSDWQDYHSAFRRTHLDADFPWPHQCCARDG
QYRPLSLDGCKLGVADYYHEEXXXXXXXXXXXXXXXXXXXXXXXXXXXX**FCVLLGVMFYWSRIEP**

>UPK1a.dog XP_541693.1
MASAATEAEKGSPVVVGLLVVGNIIIL**LSGLALFAETVWVTADQYRVYPLMGVSGKDDVFAGAWI**
AIFCGFSFFVVASLGVGAALCRRRSMIVTYLVLM**LIVYIFECASCITSYTHRDMVSNPSLITKQ**
MLTFYSADTDQGQELTRLWDRIMIEQECCGTSGPMDWVNFTSAFRATPEVFPWPPLCCRRNGN
FIPLNEEGCRLGHTDYLFT**EGCFEHIGHAIDSYTWGISWFGFAILMWT**LPVMLIAMFYTTL

>UPK1b.dog XP_850140.2
MAKDDSTVRFCQGILLIFGNVIVG**MCGIALMAECIFFVSDQHSLYPLLEATD**ND**DDIYGAAWIGMFV**
GICLFCLSVLGIVGIMKSSRKILLAYFILMFIVYGFEVASCITAATQRDF**FTPNLF**L**FKQMLERYQ**
NSSPPNDDQWKNNGVTKTWDRMLQDHCCGVNGPSDWQKYTSAFRTENNDADYPWPRQCCVMNN
LQEPLNLEACKLGVPGYHKQ**GCYELISGPMNRHAWGVAWFGFAILCWT**FWVLLGTMFYWSRIEY

>UPK1a.elephant XP_003420803.1
MASAAAEVKKGSPVVVGLLVVGNIIIL**LSGLALFAETVWVTADQYHVYPLMGVSGKDDVFAGAWI**
AIFCGFSFFVVASFGVGAALCRRRSMILTYLVLMLTVYVFECASCITSYTHRDF**MVSTPSLITKQ**
MLTFYGADTDQGQELTRLWDRIMIEQECCGNSGPMDWVNFTSTFRKATPEVFPWPPLCCRRNGN
FIPLSEEGCRVGHVDYLF**TKGCFEHIGHAINS**YTWGISWFG**FAILMWT**LPVMLIAMFYTTL

>UPK1b.elephant XP_003413030.1
MAKDNSTVRFCQGVLIIFGNVIIG**MCSIALMAECIFFVSDQHSLYPLLEATN**ND**DDIYGAAWIGMFV**
GICLFCLSVLGIVGIMKSNRKILLAYFILMFIAYGFEVASCITAATQRDF**FTTNLF**L**FKQMLERYQ**
NNSPPNDDKWKNNGVTKTWDRMLQDNCCGVNGPSDWQKYTSAFRIENNDADYPWPRQCCVTNN
LKEPLNVEACKLGVPGYHHSQ**GCYELISGPMNRHAWGVAWFGFAILCWT**FWVLLGAMFYWSRIEY

B. UPK2/3 uroplakins (UPK2a, UPK2b, UPK3a, UPK3b, UPK3c AND UPK3d). Exons 2-5

>UPK3a.human NP_008884.1
AVNLQPQLASVTFATNNPT**LTTVALEKPLCMFDSKEAL**TG**THEVYLYVLVDS**AISRNASVQDSTNTPL
GSTFLQTEGGRTGPYKAVAFDLIPCSDLPSLDAIGDVSKASQILNAYLVRVGANGTCLWDPNFQGLCN
APLSAATEY**RFKYVLVNMSTGLVEDQTLWSDPIRTNQ**LTPYSTIDTWPGRSSGGMIVITSILGSLPFF
LLVGFAGAIALS**LV**

>UPK3a.mouse NP_075967.2
TVNLQPQLASVTFATNNPT**LTTVALEKPLCMFDSSEPLSGSYEVYLYAMVDS**AMSRNVSVQDSAGVPL
STTFRQTQGGRTGPYKAAAFDLTPCGDLPSLDAVGDTVQASEILNAYLVRVGNGTCTFWDPNFQGLCN
PPLTAATEY**RFKYVLVNMSTGLVQDQTLWSDPIWTNR**PPIPYSAIDTWPGRSSGGMIVITSILGSLPFF
LLVGFAGAIILS**FV**

>UPK3a.cow NP_777134.1
GVNLQPQLASVTFATNNPT**LTTVALEKPLCMFDSSAALHGTYEYLYVLVDS**ASFRNASVQDSTKTPL
SSTFQQTQGGRTGPYKAAAFDLTPCSDSPSLDAVRDVSRASEILNAYLIRVGTNGTCLLDPNFQGLCN
PPLSAATEY**RFKYVLVNMSSGLVQDQTLWSDPIRTDR**LTLYSAIDTWPGRSSGGMIVITSILGSLPFF
LLIGFAGAIVLS**LV**

>UPK3a.elephant XP_003420946.1

AVNLQPQLASVTFTTNPTLTVALEKPLCMFESNTALVGTYEVYLYVLAESA SLRNASI QDHTSAPL
SSTFQQTEGGRTGPYKAAAFDLQPCSDLPSLDAVEDVSRASEILSTYLVRVGANGTCLSDPNFQGLCN
PPLSVATEYRFKYVLVNMSTGLVQDQTLWSDPIRTNRLTPYSAIDTWPGRRS GGMIVITAILGSLPFF
LLLGFAGAVLLSFV

>UPK3a.dog XP_851414.1

GVNLQPQLASVTFATNNPTLTVALEKPLCMFDSAATLNGTYEYLYVLVNLASSRNASVQD GARAPL
SSTVQQTEGGRTGPYKAVAFGLIPCSLDLPSLDAVG DVARASEILNAYLVRVGANGTCLSDPNFGGLCN
APLSAATEYRFKYVLVNMSTGLVQDQTLWSDPIRTNRP TPYAAIDTWPGRRS GGMIVITSILGSLPFF
LLVAFAGAVVLSLL

>UPK3a.opossum XP_001378435.2

AVDLEPQMASITFATNNPTLTITLEKPF CMFNASNLGNVSYEVNLYVMENSGSVRYAVIKNNRSFPI
NSTFQETAGGQRAPYKAASFILPQCGDLPNLDEAGDVSKVAEILSAYLVRVG DNGYCSLNP NFKGTCN
PPLTRATEYRFKYVLINPSSGFVEDQTLWSQPIRTNQISPYLEIDTWPGRRS GAMIVITSILSTLVFF
LLVGFAAAVIFS FV

>UPK3a.platypus profilin AAPN01155619.1 PLUS

AAPN01135149.1Cont1295.23,
AQIKPQIAGLAFATNNPTLTITIALEKPF CMFDGAGPAQLPSEIYLYVMMDSA DAHPVLDNGSRPLRTT
FQQAAGRRGPYWAATFSVPRCEDLPQLWDAGD PARAPQILDAYLFRVGGDTACMWDPDFSGACNPPL
AGETGYRFKYVLVNTTSGSVVDQSLWSDPIRTKR T LSWSRVDTWPGRRS GSMIVITSILSSLMFVLLV
GLAAAVTCRV

>UPK3a.chicken BU410839, BU464912 and profiling NW_001471513.1

Gga1_WGA14_2:27108888-27154218
DQSMKPQLAAPELATNNPTLTVALEKPF CMFDSSLHPNKS YAIYLYVMKSSANTISSVVT DSSSKPL
DSTFQQTHGGHLGPYKAASFDPNCVSPPR LADAGDINKVSDVLKQYLFRVGDDGTCLYDPNFLDVCN
PPLAPDTTYRFKYVLVDNTEGIVKDQTLWSDPIKTRK AKLPMKIDIWPGRRS GSMIVITSILSVSVFL
LLAGLLASVFSALV

>UPK3a.aligator profiling gb|AKHW01039965.1| scaffold-4401_2,

AQSLRPQIASPQLATNNPTFTTIALEKPF CVFDGSLSPGKSYEVLYAMMDSM SMISSAVTDNSSKPL
DSTFQEVNGGQLGPYKAAVLNVPDCASPPKLADIRNVKKASDVLKQYLFRVGDDV SCLYDPNFLGVCN
PPLASDTTYRFKYILVDETLGIMKDETLWSDPIKTNS VKKSSTIDTWPGRRS GGMIVITSILSVLIFL
LLAGLFASVFFAVM

>UPK3a.turtle profiling gi|350386162|gb|AGCU01117018.1|

scaffold678_371,
ALNLRPQIANPKLATSNPTLTITIALEKPF CMFDDSLSPGSSYEVLYAMADSESTVSSAVTDNSSKPL
NTTFQDTNGGQLGPYRAALFNVPNCASPPMLADVNVKKVSDVLKQYLFRVGDDVTCLYDPNFP GACN
PPLAQDTTYRFKYLLVDVNAGVVKDQTLWSDPMKTRR VKQSSSTIDTWPGRRS GGMIVITSILSTLMFI
LVAGFLASLYFIVM

>UPK3a.Xtropicalis BX737880 CX441249 ENSEMBL

scaffold:JGI_4.2:GL173822.1:32034:48169:1
AGADMPLLANSDFFS LNPTQTITIALEQPICMF KSAVNVYLIGIVAGAPNTPLYDGNKKV NASTYSGTQ
GGKTGPYIVAKLPNQQCINIQALS NMADPTQVQSILSKYVVRVGADVTCLTNPNFVGYC NAPLQGNTQ
YSFKYLFTDSGDIVQSETSWSLGITTVNGKASSTIDTWPGRRS GGMIVLTSILSTLMFFVFIAYVIGF
AYSIA

>UPK3a.Xlaevis DC114372, DT404162

ARSAVPLLANSDFFS LNPTQTITILERPF CMYKDAINVYLFAIVKGATNIQVADA AKKVIASNYTGTQ
GGLLGPYQVAKLDNPKCENIQASNIMADPNKYIVRVG DNVNCLTDPNFKGICNPPLQNNLQYRFTYVF
TIGDVVQYQTDWSPPISTVNVKSSGTIDTWPGRRS GGMIVLTSILSTLMFFVFFAYIVGFAYSIL

>UPK3a.salamander JK978030, JK980634, DT404162.

TFAVTPQVAGPALVPNTPTLNTVALQKPVCLFDSQTGGDPANYQVELFAMAASAPSSTPLASGNTFRN
TSGGTTGPYVAGKFGVPNCTLVGSLTPSQDDFLRQYIFRVGDNPTCLTDPNFSGICNPPLANSTAYRF
LYSLVHNNGSRVAN THWSDRISTKNVKTPD TLDTWPGRKRS GGMIVITSILSTLLFFLLSGFVAAANVL

>UPK3a.coelacanth Ensemble
scaffold:LatCha1:JH126738.1:234530:268960:-1; gbAFYH01050375
gi|346164461|gb|.1| contig050375.
AQFPKPEIANPQFVAGNPTLTTISLEKPFVFDRTISASSGDFTVAVFAVKSTVAENINDFTQTYQ
SSREGTTAPYKAASFAVPNCDSPLVLSNPVNVNIRSLNQLIRIGDDTMCSNMPGICNGPLTPNTA
YRFKFVLLNGNRPAAPQTPWSSMISTRKSKPFEDIDTWP GKRTGGQVVVTTILVILLFLLLCGFVTTLV
ASI

>UPK3a.spottedgar genomic profiling AHAT01033516.1 contig033516
EALKVKPEAVSPRLLRFNPTQSTVSLAKPLCVFDSVKPTNEMMVDVYVHSLSATLTFETGKTYKETN
GGTETPYKATSFGIPNCTSPPNPADLSVPQRIKDLDEYLVRIGSNPTCVGEPEAEAFCNAPLSDGTS
YRFKYLLVNGTTTIAETEWSESILTRKALSPDEIDTWIGKRSGGGIVVTIVLSLLLFLLLGAAIFMGV
LDVI

>UPK3a.skate EE989195; GH546922; AESE010876980.1WGS_1_CONTIG_884036,
AESE010997029.1;WGS_1_CONTIG_1005178;AESE012684067.1WGS_1_CONTIG_284
8843
PLSFKPEVSSIALTLGLRTSTTVTLAKPICVFSSGDDVVEVFGVQTTADSIPIEIGNRTILTYQQTGG
ARGPYRAARFMTPICTSLPFVPSRDPVIRVQIEQYLFRVGDDQCLNQAPFVSGNCNAPLKENVAYR
FKYAVLNSSTNIILNETSWSDPILLRVADFALIDTWP GARTGGMVVITLLVILLSLLLCGYGALLV
YACC

>UPK3a.Shark JK928319; AAVX01399128.1|PLUS AAVX01255679.1
ILQFKPEMPSVSTIDGLRTARTVTL SKPLCVLPTGHMIELLVVQGNVPIATNLDGTYQATKGATGP
YRAARFANPECTSSSFTPSTDP TKIQTLDIRYFIRVGSDTQCLSGGTAQSVPCNAPLNENVNYRFKYV
ARDPITFILRDETVWSKPITLLQVQDPARIITWP GARTGGMVVLTTLLVIILFLLLCFAVAF

>lamprey3a.3 profiling Ensemble scaffold:Pmarinus_7.0:GL480097.
GERYGNLFPSTSSSTRCANQFQSFVSSAHSSSLLRLYDTTTERSAPGYSFSRVLGAVGSRVATGAGDRG
PYLAGYFPVPPCAQPPPSVVGTGGSSVALALLGMVIFRLGGDPACATGIADDPAAALVCNGPLWNGT
AYRVKVFLLARSSLLSPPAMETPWSDPISKPVVSVEQIHVWPGKRTGGMVVVTSVLC SILFVLLFAL
TAALLAPS

>lamprey3a.1 FD718834.1 FD706071 FD704820; profiling Ensemble
scaffold: Pmarinus_7.0:GL477389.
GSGTKPQLVSSSTAVPYNPTETTIVWSKPFVQKPVPTTQYVVDVYASITNNSYAFDNSIGAVLSSYW
TNAVSPSPYLAATFKVPDCASQPSIYDAMAVKTNATFRLGGDTACVNSIGPSTSVCNGPLVPGMKYRV
KYTLSEESPQFPRTIVDQTPWSDPVSTKSPAASTINTWP GKRTGGMVVVTAVLSTLLFLLLAALLLV
VIFKACS

>lamprey3a.2 FD709144 profiling Ensemble scaffold:_7.0:GL483426.
VSQVVPTVVNPNLLGAVTQTVALQAPFCSALDAEVVALSSVSADLRLFVMATAQRNVSNEMITSSST
IGLDKGYAGSGAGTSSWY LALGGRPLQNCTISTPSLSTPTSYYRVGADSKCSVAVTCNGPLNAGTIYW
FKYIIGTVATHGAIDRSYLESSWSKPIRLNKAGELNAIGVTPGPXSGGMVVVTVILVLLFIAVDRAR
PALGQ

>UPK3b.human NP_872625.1
ELVPYTPQITAWDLEGKVTATTFSLAQPRCVFDGLASASDTVWL VVAFSNASRGFQNPETLADIPASP
QLLTDGHYMTLPLSPDQLPCGDPMAGSGGAPVLRVGHGCHQQPFCNAPLPGPGPYRVKFLLM DTRG
SPRAETKWSDPITLHQGKTPGSIDTWPGRRS GSMIVITSILSSLAGLLLLAFLAASTMR

>UPK3b.cow NP_001012690.1
DLIPYTPRITSWDLEGKVTATTFSLAQPRCVLDRHSSAADTVWL VVAFSNASRVFQNPQTAEIPASP
RLLTDGHYMTLPLTMDQLPCEDPADGSGRAPVLRVGNDA GCLADLHQPRYCNAPLPGPGPYRVKFLLT
NSRGSPPAQETRWSDLIALRQGKSPGSIDTWPGRRS GDMIIITSILSSLAGLLLLAFLAASSVR

>UPK3b.mouse NP_780518.1|
DLIAYVPQITAWDLEGKITATTFSLAQPRCVFDEHVSTKDTIWL VVAFSNASRDFQNPQTA AKIPTFP
QLLTDGHYMTLPLSLDQLPCEDLTGGSGGVPVLRVGNDFGCYQRPYCNAPLPSQGPYSVKFLVMDAAG
PPKAETKWSNPIYLHQGNPN SIDTWPGRRS GCMIVITSILSALAGLLLLAFLAASTTR

>UPK3b.elephant XP_003416614.1

DPLHATDNGLGKKVTATTFSLQPRCVLDGHARATDTVWL VVAFSNAASKDFQKPKTQAEIPTFAQLLT
DGHCMTLPPSPVQLPCTDSVGGSGSALLLRVGNDAARCLADLQQSPYCNAPLPSPGPY SVKFLLMDTEG
SPKAETRWSDSIALHQGRAPGSIDTWPGRRS GGMIVITAVLSSLAGLLLLLAFLAASTVHF

>UPK3b.dog XP_849277.2

ELIPYTPQITAWDLEGKVTATTFSLQPRCVLDGPASVASTVWL VVTFSNAASKDFHNPQTLAEIPAF
RLLTGYYMTLPLSLDQLPCEDPEGGRSIPLLRVGNDPGCLADFYEPYCINNLPSPGPY RVKFLLM
DARGSPQAETRWSDPITLHQGRAPGSIDTWPGRRS GDMIIITSILSSLAGLLLLLAFLAASTVHF

>UPK3b.platypus profiling gi|125702340|gb|DS191120.1| Scfld14595

LMSYKPVQSAHPLEGKVTASTFTLDQPRCVFDGQVVSTDTIWL VVAFSNAASRDFQNPPTAASIPAYPR
LLTDYYMTLKVSPDLPCADAGLSVLRVGTDPGCLRDPGREYCNAPLPAPGPY RTKFLLVMD EASHP
KAETQWSDPITLKQGRDPSSVDTPGRRGSGMVVIA SILSVLAGLLLLLALLTAAAGACTV

>UPK3b.opossum XP_001378885.2

DQIPYTPQISALALEGKVTAAATFSLEQPRCIFSELAAPADAVWL VVAFSNAATEDFQNPKTAAEIPSYT
ELSSSFYYMTLKLSPDLPC EEEEDIAVLRVGSNTNCLRNLSQEYCNAPLLAPGPY RVKFLLVMDNNGQP
KAETWSDPITLNQGRDPRSIDTWPGRRS GCMIVITSILSTFAGLLVIAFLIASTVQF

>UPK3b.Xtropicalis EL866947; Ensemble scaffold:JGI_4.2:

GL172708.1:135056:150415:-1

DITSYVPQLTLSPIVGTVTSTTFVLDKPPQCVFGNTGNQVWLL VARSNV SANVLTTPPSMYSSFATKGY
YHVPFGTESLYHCSNTAEYIRVGDTAQCNNTNCGPLDPGPY RVKYLVMNNALVSQSLWSQQITL
LTGKSSSQLDTPGRRS GGMIVLTSILSVLMGILTLCLFAAFFVGC

>UPK3b.Xlaevis BP703707 BJ034026

DITTYVPQLTLMPIQGSVTSTFTLDKPPQCFGSRTNQVWLL VARSNVSVSITNAMLKPPSMYSSFPT
QGYHYVPLGT EASYPCNTADYIRVGDTVYCTDNTYCNAPLPDSGPYRVKFVVMNNALVSSSLWSGL
ITLRTGKNPSTIDTPGRRS GGMIVLTSILSLLMGILTLCLIAAFFVGC

>UPK3b1.Xtropicalis profiling Ensemble scaffold:JGI_4.2:

GL172728.1:2871266:2881345:1

IPYYVPQITTKPILGKLTSSSFVLEQPQCIFQYKTSLVWL VVALNRVIPQLSYTQLSNPANISSFET
NGFYHTLPVFGGDYPCADTSGQLSAMIYVGS DVNCSNPLFCNGPVPSRGTY RVRFVVLNGTVMETGTR
WSEVITMHIAINSSTIKTEPKRPSSGMIVIITILASLLFILLVCLIAAVSLGS

>UPK3b2.Xtropicalis XP_002934406.1; EL802701; Ensemble scaffold:

JGI_4.2:GL172728.1:2884693: 2898749:1

DVGSYVPKITTSILGNLTFSTFVLEQPQCIFSSNYPTQDVWL VVALDTEPFLTDNLSTPVTYSSFT
TNKFYHTLRVRGADYPCFNESAMSLALLQVGAD EKCNEFCNGPLTSPGPY RVRFVVLNNTGMVAKTN
RSDLIRLPIGINYTTIDTWPLSRSGSMIVITILSILLAVLLACLLAALCSE

>UPK3b.chicken one extra cysteine; profiling chromosome:

WASHUC2:19:4127260:4157406:-1

ALLPYVPRVAPGAMPKVTATTFVLERPRCIFDPFANASDAVWLAVAFADA
SAAFKNPSTSSAEVPPYEGLPTARAYMTLQMAAAAYGCSAPGA AVLRVGGDTACHGRAPCNGPLPSPGP
YRVKFLLMGCGGPKAETKWSDPILLRRARSLSTIDTPARRSSTAVVIAAILASLGAALAMAVLGAVG

>UPK3b.aligator profiling gi|397229275|gb|AKHW01080724.1 scaffold-
9559_4

ELVPYTPQITGNSLEGKLTASTFSLDQPICIFDQYVNATDDIWL VVAFANATSSLKNPSTRTDIPPYQ
RLSTALHYXTLRTTIXFFPCPRSTNTSVLRVGS DAFCRNDNSQQHCNGPLPNPGPY RVKFLLIDSNGA
KAETRWSERITLKQGRHSSTIDTPGRRS GTMVVITVILSSLVGILTIAFLCT SAYD

>UPK3d.coelacanth profiling AFYH01004438 gi|346210398|gb|.1|

contig004438.

DVPYLPEVAFNNVAGSVTATFTLQQPRCIFKDV FSSCPLCQLWVAVATEKGANNLITTVGRPVT FAG
YQKLSQRGYYFTMRSSKDLYKCVENDDKVRVLR LGADTSCPQSIGESDCNGPLPDGSGSYRARFLIVDS
QSGNMLKAQSQWSKSIQLKRPKDPASIDTPGKR TGGMVVTTILVILLAILLLFIVALITAC

>UPK3d.1.spottedGar genomic profiling AHAT01005644.1 contig005644
VTSVEYTPETITAYNMAGRVTGTTVMKQPRCYFDNQLLLPCTPDKCEIWLVAAVGSAAGIQNFDADK GK
STILSESPYPTAFTGSPPKNYLT KVGRQQVFPCQQSHGIVYFRVGDEGNCTSANCNGILPAGSTVRV
KYVLVDPASRNVTTETYSQNTLYSSSTDPVVIFDGIQRSGMIVITSILSVLLFLLLLLLIAGLIY
AK

>UPK3d.salmo BG936628; AGKD01191862.1| Contig_191888
IVSLDHIPEILPYEASGRVTSTTIVLGQPLCYFNTLTQLKCSQSTCQVWAAIASGPGINNFDIDKLVA
VQIVSASPYPPIAFSSQTNRMYFVTKLGRPKDFPCGQLPGIKYFRVGAEGNCTNTNCNGILPPGSTVRV
KYILIDPVSRRGVVSESKWSYPISLTSTRSWSSIDEWLGKRSGGMVVITVISSCLLAVLLLLLLGAVLLL
DG

>UPK3d.trout CX140558, CA369219, BX073041, CX140559.
IVSLDHIPEILPYEASGRVTSTTIVLGQPLCYFNTLTQLKCSQSTCQVWAAIASGPGINNFDIDKLVA
VQIVSASPYPPIAFSSQTNRMYFVTKLGRPKDFPCGQLPGIKYFRVGAEGNCTNTNCNGILPPGSTVRV
KYILIDPVSRRGVVSESKWSYPISLTSTRSWSSIDEWIGKHS GGMVVITVISSCLLAVLLLLLLGAVLLL
DG

>UPK3d.zebrafish A2CE76; DT080652.
GQIFQPQLAPANFLGRITSNTVILQQPYCVFTQTCPGCEIWLVAALSTGTGNFNALVNISSPISLSVS
PYPTAFLPSSAQFFLTRVGPLANFPCNTAPAFPYFTVGADGICTGINCNGVLPVGSIVSFRYLLIDPS
NYTVVNMTNWGGPFNLTL LLSYQTINDGLSARSGAMVVITTLCCVAVALLLVFFIMLCVSC

>UPK3b.skate EE991990; AESE010072473.1| _WGS_1_CONTIG_72535;
AESE012558708 WGS_1_CONTIG_2721782; |AESE011082469.1| Leucoraja
erinacea LER_WGS_1_CONTIG_1092019; AESE012622692.1
_WGS_1_CONTIG_2785766; AESE010900346.1 _WGS_1_CONTIG_907428,
AESE012509864.1|WGS_1_CONTIG_2672938
ADEADFVPEIFKDNVIGGVSQTTFALQQPQCVDFRFPACTLCEIWLVDNPA NVSTFDTNMNTLTPTS
AATYKDFLLNGFYLT VTKTARTDYACPQTMGQVYTLRVGDEDPCTTPNCNAPLNAGSLV RVRYVMINPL
ATTNNVIAVTKWSNPIQLENA VDPNIIDTSTRSGAMVVITILSILLFLLLVLFI VMLAA

>UPK3b.Shark AAVX01026120.1| PLUS TRACE gnl|ti|1573944991 PLUS
AAVX01044278.1|Callorhinchus milii scaf_1099306886782, whole genome
shotgun sequence
VKNYTPEVTNANLAGRITATTITFQQPLCQDFSEINATPGLIFWL VVANSQ RGQNTFDAITFVQPSLM
THQTASTADYYLTLRLAERYLCPTSSSNGEMYFVRVGQETTCRTETCNVPLNNSKSY SVKFILINP
SLQNSPNVIAQTNWSSFISLKT VVEPRNIDPSRGRSAGMIVITAILSVLLFLLLAFFVAMLLMVC

>UPK3d1.salmo profiling AGKD01066640.1| Contig_066647.
IPVNNTPEITPSSLAAKLTSNSVILTSPSCYFNGLGNLSCNSTTTCEIWLVS AKDTGVSNYDADNKMP
YIDTRSPYPTAFSSNTSKKYFLTKLGLQKAYPCPIVAGTGYFRVGS DGNCTPNCNGILPVGSTARFK
YVLINPANKTVVAESLWSNNITLYSLKDPEKIDNGFAGRSAAMIVITAILCSFLALLLLLLLLIMLIYV
L

>UPK3d1.trout CX145617, CX153372.
PVNNTPAINPSSLAAKLTINSVILTSPSCYFNSLANLPCNSTTTCELWLVS AIDTGVSNYDADKNRPF
IDTLSPYPTAFSSNTSKKYFLTKLGFQKDYPPIVAGTDYFRVGS DGCSTPNCNGILPVGSTARFRY
VLINPENKTVAESLWSNNITLYPLKDLESIDHGFAGRSAMIVITAILCSFLALLLLLLLLIMLIYV
CC

>UPK3d.Spotted Gar genomic profiling contig005644, Sequence ID:
gb|AHAT01005644.1|
TVPSVDYVPEIIPMKLAGRLTGTTFLLRQPLCYFSNQQLNCSLSTCEIWLAVAREAG VNNFDTDKVQ
PSFDIVSASPYPEAFQKNKYVTRLGVQNNFLCAELPGIRYFRVGAEGNCTPTCNGILPAGSTASCA
VLGLAVVSQLQDQLVKGALLWAQQGLVVVDV MPSSIDDWTWKRS GGMVVITVITSCLLAILLLLVA
LLLG

>UPK3d.2 Spotted Gar genomic profiling |AHAT01005643.1 contig005643

VPYAPEVTPHNLLGRVTSTTLTQQPVCFNNLSGLPCAPDNCEIWLVIARGVPKF EAIKGNTSVLMS
SPYPDAFRNNSPPNYFLTKLGVQNAFPCAESTGFRFFRVGDEGLCSTSNCGILPAGSTVRAKYVLLD
PGSKQVVSESQWSSPISLIAALSAGIDEWIGKRSGGMIVVTSILSSLLAILLLFLLTAFVLR S

>UPK3c.human NP_001107875

AAPEHISYVPQLSNDTLAGRLTLSTFTLEQPLGQFSSHNISDLDTIWL VVALSNATQSFTAPRTNQDI
PAPANFSQRGYYLT L RANRVLYQTRGQLHVL RVGNDTHCQPTKIGCNHPLPGPGPYRVKFLVMNDEGP
VAETKWSSDTRLQQAQALRAVPGPQSPGT VVI IAILSILLAVLLTVLLAVLIYT C

>UPK3c.mouse XP_485688.3

ESINYAPQLLGATLEGRLTQSTFTLEQPLGQFKNVNLSDPDP IWL VVAHSNAQNF TAPRKVEDRHAP
ANFDRNGYYLT L RANRVHYKGGQPDSQLRVLRVGN DNNCSLESQGCNSPLPGAGPYRVKFLAMSAEGP
VAETLWSEEIYLQQAQTFREAPGSQKGTVV I IAFLSILLAILLVFLVLVISAC

>UPK3c.dog XP_850037.1

EHISYVPQLSNRSLAGTLTQSTFTLEQPRGQFSHRSISDSDAIWL VVAHSNATQNF SAPQRVEDIPVP
EDFTRRGYYLT L MANRLLYPGNQPGNQLRVLRVGN D TSCSPTKRGCNHPLPGPGPYRVKFLVMSDKGP
VAETEWSNETHLQRAERLQAAPGPQSTGT VVI IAILS VLLAVLLTALLALLIYT C

>UPK3c.cow gb|AAI49699.1

ERISYVPQLSSATLAGRLTQSTFTLEQPRGQFSHPSISDSDAIWL VVAHSNATQKFTAPQKVEDTPVP
ADFPQRGYYLT L RASRALYPGGPPSNQLRVLRVGN DTRCSPRTRGCNRPLPGPGPYRVKFLVMSDRGP
MAETEWSSETRLQQA EVLQAAPGPQTAGT VVI IAILS VLLAVLLAALLALLIFT W

>UPK3c.elephant profiling gi|253616662|gb |AAGU03064090.1
cont3.64089.

ELISYVPRLSNATLAGKLTQSTFTLQQPRGRFSHKISDFDAIWL VVAHSNATQSF IAPQRVKDSPAPA
DLPLKGYLT L RASRALYPGDQAGNQLQVLRVGN DTRCSLTTKGCNHPLPGPGPYRVKFLVLSDDRGL
VAETEWSGETHLQRAEVLQASPGPQTAGT VVI ITFLSVLLVLLTALLALLIYT C

>UPK3c.opossum XP_001378846.2

EPINYTPAITREPLEGSITSSFTLDQPNDDQFNGSGISDLDDIWL VVAFSNASQSFEPPQSAQDIPYA
ATFLDKKYYLT I RASRDLYSSKRGSQGISVLRVGN ETNCTRSDCNKPLPGPGPYRVKFLVMNTNGPVA
GTNWSEDITLRKPVEFSES RPPSKSAGTIV I IAILSILLSLLFLALVALLVYT C

>UPK3c.platypus profiling gi|125692247|gb|DS201210.1| Scf1d26655

KIAYTPVITKLPMEGKITLSTFTLDQPLGQFNSSAVDDLDDIWL VVAYSNATDNFSNPVKPDEEYNLS
ELSKKQYYMTMRATRDLYPGGNSSFLHVL RVGQEANCTSKSCNGYLPGP GPYRVKFLVMNERGPVAE
TDWSDNIALREAKDPGDPNPTRRSTFMIVITILSVLFALLLGALTGVLIQAC

>UPK3c.Lyzard profiling gb|AAWZ01050405.1| cont1.50404; whole genome
shotgun sequence REVERSE

VNYTPRLASENLGGKVTAFTFTLDQPRCVFNDVVNATDGIWLLVARSDAARNFTRPGSPSELPFQDLE
KNGLYLT LNTAPASYPCPEPGAAGGPLTVLRVGNEVQCASNRARPDCNGPLPRPGPYRVKFLAINPDG
VTAESEWSEEIALVQA
QSPETIDVSPGRRSASAIASLLSILCAVLLAALIAALVYKYT

>UPK3c.chicken profiling ref|NW_001471508.1|Gga19_WGA256_2

DKLSYKPTLVGGNVEGRMTGSTFVLEQPRCVFDSYSTANIWL VVATRAGMNAFNDSAQGMPEWSFQR
FPTNTSAYLT LGAMQYHYGCPKPDRELTVLRVGSETGCADNISVPNCNGPLPGPGPYWKFLALNGSE
PTATTEWSGPITLKTAREPQSIPGMGGARSGAMIAITAILSVLLAILLAALLATLCS

>UPK3c.turtle profiling Ensemble scaffold:PeISin_1.0: JH207163.1:

SIDYTPVITARELEGKITSSFTFVLEQPRCVFNDSVSNTDEIWL VVALSNGTSTFTNPTSLQSLPAF
QKFPGPSHYMTMGTSLLNYPCEKSSGQITVLRVGN ETGCVSDTTRPDCNGPLPGLGPYRVKFLAMSPV
TGPTAETRWSDPILLKAGKDPATIDTWPGKRSAGMIVITILSILLAILLACFIAALTYRC

>UPK2a.human. NP_006751.1

DFNISSLSGLLSPALTESLLVALPPCHLTGGNATLMVRRANDSKVVTSSFVVPPCRGRRELVSVD SG
AGFTVTRLSAYQVTNLVPGTKFYISYLVKKGTTATESSREIPMSTLPRNMESIGLMARTGGMVVITV
LLSVAMFLLVLGFIIALALGSRK

>UPK2a.mouse. P38575.1

DFNISSLSGLLSPALTESLLIALPPCHLTGGNATLMVRRANDSKVKSDFFVVPPCRGRRELVSVD SG
SGYTVTRLSAYQVTNLTPGTKYISYRVQKGTSTESSPETPMSTLPRKNMESIGLMARTGGMVVITV
LLSVAMFLLVVGLIVALHWDARK

>UPK2a.dog XP_853307.1

DFNISSVSGLLSPALTESLLVALPPCHLTGGNATLMVRRANDSKVKSSFFVVPSCRGRRELVSVD SG
AGFTVTRLSAYQVTNLVPGTKYISYLVKKGSTESSREIPMSTLPRQAESIGLMARTGGMVVITV
LLSVAMFLLVLGLIIALALGARK

>UPK2a.cow NP_776639.1

ADFISSLSGLLSPVMTESLLVALPPCHLTGGNATLTVRRANDSKVVRSSFFVVPPCRGRRELVSVD S
GSGFTVTRLSAYQVTNLAPGTKYISYLVTKGASTESSREIPMSTFPRKAESIGLAMARTGGMVVIT
VLLSVAMFLLVLGLIIALALGARK

>UPK2a.elephant XP_003418225.1

DFHISSLSGVLPALTESLLVALAPCRLTGGNATLMVRRANDSKVKSSFMVPPCRGHRELVSVD SG
AGFTVTRLSAYQVTNLVPGTKYISYLVKKGTTATESSKESAMSTLPRKMESTGLGMARTGGMVVITV
LLSVAMFLLVVGFITALALGAQK

>UPK2a.opossum XP_001380648.2

EFNISSLSGLLSPALAESLLVALPPCHLTGGKASLTVRRVNESA GMTHNFTVPPCRARRDLVS VYNS
GSFSITRLSAYQVTNLIPGTKYFVYYSVEKGTAVESSNKVQMATLPRKRVETLGLGMARTGGMIVITV
LLSVAMFLLVVGLIVALALGVHK

>UPK2a.Platypus profiling AAPN01121652.1 exon2 AAPN01340066.1 exon3
| AAPN01066803.1 exon4; AAPN01269979.1, exon5.

FNISSLSGLLTPTLAESLLVALPPCHLTGGNATLKVQGVNGSSVLQQRFFVPPPCRGRRELVSVDSSA
GFATTRLDAYQITGLRPATTYVXXXXXXXXXXXXXXXXXXXXXXXXXQKAELLVGMARTGGMVVITVL
LSVAMFLLLLGFIVALVLGSHD

>UPK2b.snake profiling gb|AEQU010390653.1| contig26931230;

gb|AEQU010390653.1| Python molurus contig26931230;

gb|AEQU010042415.1| Python molurus contig26234771,

MDFFRNNSMLLTFTTGTTFIVNIPKCVSSSKFSPATVRIAIAQLPDASSLPGITD TDQIENLR RTPQA
LVYFADEFKDISCRVARDLLVLDLDDSQYELITVVGYQVGGEFCRQTKGPFCNQALKPSTFYRVNFFF
LDDKSVIRAHTDWSTAIQTRNV TNYESADVMFEGRAGGMIVITILVSVGGAVLLVALIVAVALSSKK

>UPK2b.turtle profiling |AGCU01117019.1| Pelodiscus sinensis

scaffold678_372, LDFFPNASLILAARMSTSFIVNIPKCISSSQFTPTTIRPAVAIVGDKVAMP
AVTDTNQIQSLHSSKFKSVScriARDLVSMIDDDGNYKLTTVVGYQVGVECEKTKGPFCNQALQPSS
VYRVNFFILDEKAVIRAHTGWS DVIQTNNVTTFMAYDGSFGRAGGMIVITVLLSVAMFVLVGLIVA
AALGGKKS

>UPK2b.lyzard XP_003221456.1; FG735771.1 gb|AAWZ02009508.1|

cont2.9507,

ELDFFPNNSMLLTFSFMSTYFIVNVPKCVSPKDFSPVKIRLAVAQVPDTSTLPGVIDTDDIANLR RTPQ
ATIYFAGEFDSVSCRVT RDLLVMDMDDSQFELITVLGYQVGAEFCRQTKGPYCNQALKPSTYYRVNFF
FLDNKSVIRAHTDWSTALQTRNVSDHESADVMFGGRAGGMIVITILVSVGGAVLAIALIVAVALSNKK

>UPK2b.aligator profiling gi|397270035|gb|AKHW01039964.1| scaffold-
4401_1,

TLNFFPNNSNILAARLSSSFIVNIPKCISSSHYTPTTIRPAIAVLGETVGLPAVTD TNQIRSLRNFTD
APIYYAGEFASVSCRMARALVKMDIDDENYKLTTIVGYQVGAEVCENTKGPFCNRVLKPSSFYRVNFF
VL DENAVVRAHTDWSDPIQTNNVTFSFAYDGSFGRAGGMIVITVLLSVGMFLVVVGLIVAAALGGRK
S

>UPK2b.salamander JK979875, JK978999, JK977433, C0783569
PAFSFFGNQDNIIGNRMGYSFITNIPSCISGAGYTPASILLAVSTSAAVPGVSNTDAIKSLLNDSRAQ
VYYAGQFNTVPCRVS RDVVQVSSKAADNSFTLTTLV LGYQVGSEVCTSVKGLYCNQVLEPGTPYRVNFF
ILDASN VIRAYTDWSDIVTTLNVTNNAQLDSGLSRRSGGMVVITVLLSIALFMLVPAL IATLVVGREK
SPLP

>UPK2b.Xlaevis DC123476, DY574479, BJ622989.
SNFQFFTNTDDVLGAVLSQSFMVNVPSCINAVGYVPSTLKVAVANRNPTCMVDTSIKSLKTDPTAPV
YYTGQMKVPQCRLRRDLEPVKMNSMRDLGYQVGTENCTEVS GPFCNQFLQPGTSYWNFIILDETDP
RAYTGWSEPRTTTRQVRRRLDAVDLGLSGHSGGMV IITVLLSVSVFLLLLGFMV VVVS RAPNLFTSESK
DCLDMQTQP

>UPK2b.Shark JK941564, JK956747; AAVX01097325.1| scaf_1099306956075,
DYDIALFNGEVSRLATLVWLSPAYCLFEKWVEQRGGELTVRSSASVQVELLREDNATFVLPQRYSVPF
CND FSPKPAAVSTPLAYQMGP NVACIDGSCIEHVLPGQRFVRVYIIYSVSEEALVTTKWSAPIATRDD
PPSYFSIDADTT PRSGAMVVITTVLVVALFLLLLGFV SMLAARSKFAR

>UPK2A.salamander C097794
QTAQNFSTSLADLPINPLQTLAIVAFPPCWVGVSQSNANLILVSKNSSGSSITNNQTPVPVPPCRLRRD
AVFSSDSSSGGTVITNIGFRVTNL TANTTYTASYQSNQVTIGLPTNFTTVQPTNYTAMPEVFARSGGM
VVITVLLSIAMAILVIALILTFVMGRKK

>UPK2a.Xlaevis BP683995
QNTSLADGVLTP LSTSVIIAFPGCKDSGKT VNLIVANGTTTVQNISLQVPQCRLKRDVVVINNSQSGN
VQTVNVGYQIQNLQPGA IYTTYAVDGSNIPSITFSTRSVSQTPDIMARSGGMVVITVLLSIAMFVL
LVGLIAVLVIGRK

>UPK2a.Xtropicalis genomic profiling;gb|AAMC01017077.1
scaffold_39_Cont17077,
QNVSQATGVLTP LATS AIFAFPDCTYSGQTVSLVITNSTSTATIIQNATFQVPQCRLKRDIVVINNGQ
SGNVQTVNVGYQIQNLQPGTNYMATYSSGGISGPSFQFSTRTVYPAVANIMARSGGMVVITVLLSIAM
FVLLAGLIAVLILGRK

>UPK2a.coelacanth profiling gi|346202989|gb|AFYH01011847.1|
contig011847
NFNTSLSDNAALISNVYSSFVILSLPPCTYAGKNASVTYSKNSTSEDSKTESFVVPPCRFRREVVEVA
RQMEGFTVTDLLGFRVGNL KAGTQYDFRYTIDNTTNVLKSNIIQITTSLVTSNVLIDEGFKLHSGGMI
VITILLSFAMLF LIIGVIVVLVLGNKS

>UPK2a.skate CV067582; AESE010437608.1
WGS_1_CONTIG_439473;AESE012541543.1| WGS_1_CONTIG_2704617;
AESE011527864.1|WGS_1_CONTIG_1648366,
GFTISLANDNMGEVVASRRSMSAIITMDPNSCNLAGETVIVTVNNTSNGQVIAQPNFVRPVCNRNRDL
ISLVSNADGTPQTLNLGYMLEMLQPSTTYNVYLRAGTIRSNMLGVTTISPVDYRTIDLGFGRSGAMVV
ITVILSIAMLALIIAFIVVLVLSK

>UPK2a.zebrafish EH444580, EH464141; gi|312125210:188150-200000
scaffold, Zv9_scaffold619
DIPISLLNPNTDGV LASTFPNSFLLQMPDCSIYGNQSVLLLYTEAPTNLNTVNFTVQPCPV SQSWYL
LGNLKN GTTYMSYKIGNDTSSVLTNTTTNNDYQQIDTGLRARGAMVVITVILSLAMVFLLVGIILV
FFFFSG

>UPK2a.salmo EG847319 EST_ssal_eve_49206; AGKD01043111.1|
Contig_043115
EFQVSLLKESDGVVTGRFADSLLLSLPPCALATQSVTLEYNNTDTNESKTLVNI FKVLP CRFR RDIIS
TIENNAQFTTSRNLGYQVTNLTTGSTYRFQYVVGAEKSNILEVSTRQVKDHNQIDSGLPACSGAMMVI
TVILSVSMFILLVALIFTVAHSLGGD

>UPK2a.lamprey

DYQVKLLNTSVVRAAQT PQSVAFVPSCEIAM IAGTMLNIFVTPNTTGVEIVGQVAVPGCRVARAVS
GQVVSGNDGTGPGIPVPSGVAYRVTGLTPSTSYSVLRHSTLGLQSVPGILSTANARTPRSAVDAEAF
SRSGGMVVITVILSVLIFLLLAILVAALLLGKN