

Supplementary Material

Anatomy of the E2 ligase fold: implications for enzymology and evolution of ubiquitin/Ub-like protein conjugation

National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD 20894, USA

Table of contents

- I. E2-like domain family multiple alignments
- II. List of gene identifiers, ordered by E2-like domain family
- III. Phyletic distribution of E2 domains across eukaryotes with completely sequenced genomes
- IV. Comparison of E2 flap architectures with and without Ub/Ubl. (Adapted from pdbids: 1u9a and 1z5s).

Figure Legends

IV. Comparison of E2 flap architectures with and without Ub/Ubl. Adapted from pdbids: 1u9a and 1z5s).

The right depiction is a complex between the E2-like Ubc9 domain, the Rangap1 substrate that is conjugated to the Ubl at a conserved lysine residue, and the Ubl SUMO while the left depicts only the E2-like Ubc9. Strands are depicted as arrows, helices as coils. The E2-like domains are colored green, with the extended regions of the flap colored gray. Conserved residues surrounding the active site are rendered as ball-and-sticks. In the right depiction, the lysine-containing loop of Rangap1 is colored pink, and the Ubl SUMO tail is colored blue.

supplementary_mater

This block contains the full sequence of the 16S rRNA gene for the specified organisms, including the forward and reverse primers used for sequencing.

Pram1000015
90970505
18396376

18593578 *Arabidopsis thaliana* GWDIYKTRKTFPLTIAKCFRD HAWIQ KREKEITRSLATQ MNKNDN DWTNISASNPETI KWIGKRWV VHNELKTEIDQT DIFTIATAP EELFPLDGTQAKTRGKICLIV KFLWANCRP TGAHACLG
 -LAIWLAATIPLFEDSGAIIKRD
 17552592 *Caenorhabditis elegans* -MDDATKSSLKAIPLCKTQKASPRDG -DLWIERLKEEYEAIIAAVQNNKCDRDFWFQLE -SNERGT -LGPWMAVEIPDLIEKGLIOPKA-
 Mbre1000002981 *Monosiga brevicollis* MVDAAHTRKTLAAIPLLRLVNAGPRGS -DLWRARHKEELMSITYVKNNKENDNDWFRLE -SNPEGT -LGPWLAVEIPDLVSKGIVKHET
 Nyec1000012015 *Nematostella vectensis* MVDEATKKTLAAIPLLTKAGPRDG -KDWDVDRILKEEYTSLIKYSNSNKEADNDWFRLE -SNKEGTRKPIITIIVNYYLII -CVE -FSPKIPIYTPTTAPEIALPELDGKTAKMYRYVYWW-----

Nvect00012015	<i>Nematostellavectensis</i>	MVDEATKTLAATPLLKIKAGFRD - KDWDRKEEISLKVSNNEADNDWFRLE - SNKEGIRKPLTIVNIELLCVF - FSPKPIYPTTAPEALPFLDGKTAKMIRYIV -
ci0100146879	<i>Ciona intestinalis</i>	MVDAATKKTLSNIPLLTKCGPRDK - EEWVKRLKEEYLSSLITYVKNNKENDNDWFRLE - SNKDT - KWFGKCWFIKDLKYEFIEFDIPTYPTTAPEIALPQLDGKTAKMIRYGGKICLTDHFKPLWARNPVKGIAHAMALG -
5101123	<i>Danio rerio</i>	MADEATRKAVSEIPLLLKTNSGPRDK - ELWVQRLREYQLAKYVENNKAADNDWFRLE - SNKEGT - RWFGKCWYIHELLKYEFDMEDFIPVTYPATAPEVAIPELDGKTAKMIRGGKICLTDHFKPLWARNPVKGFLAHLMALG -
114560788	<i>Pan troglodytes</i>	MADEATRRVVSEIPVLLKTNAGPRDR - ELWVQRLKEEYQSLIRYVENNKNADNDWFRLE - SNKEGT - RWFGKCWYIHDLLKYEFDIEFDIPTYPTTAPEIAVPELDGKTAKMIRGGKICLTDHFKPLWARNPVKGFLAHLMALG -
55586	<i>Pan troglodytes</i>	MADEATRRVVSEIPVLLKTNAGPRDR - ELWVQRLKEEYQSLIRYVENNKNADNDWFRLE - SNKEGT - RWFGKCWYIHDLLKYEFDIEFDIPTYPTTAPEIAVPELDGKTAKMIRGGKICLTDHFKPLWARNPVKGFLAHLMALG -
118138483	<i>Homo sapiens</i>	MADEATRRVVSEIPVLLKTNAGPRDR - ELWVQRLKEEYQSLIRYVENNKNADNDWFRLE - SNKEGT - RWFGKCWYIHDLLKYEFDIEFDIPTYPTTAPEIAVPELDGKTAKMIRGGKICLTDHFKPLWARNPVKGFLAHLMALG -
114560786	<i>Pan troglodytes</i>	MADEATRRVVSEIPVLLKTNAGPRDR - ELWVQRLKEEYQSLIRYVENNKNADNDWFRLE - SNKEGT - RWFGKCWYIHDLLKYEFDIEFDIPTYPTTAPEIAVPELDGKTAKMYSWVHGWWOKSI -

114560786	<i>Pan troglodytes</i>	MADEATRRVVSEIPVULKTNAGPRDR-ELWVQLKEEYQSLIRYVENNKNADNDWFRLE-SNKEGT-RWFGKCYIHDLLKYEFDFDIPITYPTTAPEIAVPELDGKTA KMYSGKICLTDHFKPLWARNVPKFGLAHLMALG-----LGPWLAVEPDLIQKGVIQHKEK
51172600	<i>Rattus norvegicus</i>	MADEATRRVVSEIPVULKTNAGPRDR-ELWVQLKEEYQSLIRYVENNKNADNDWFRLE-SNKEGT-RWFGKCYIHDFLKYEFDFIEFEIPITYPTTAPEIAVPELDGKTA KMYRGKICLTDHFKPLWARNVPKFGLAHLMALG-----LGPWLAVEPDLIQKGVIQHKEK
13384768	<i>Mus musculus</i>	MADEATRRVVSEIPVULKTNAGPRDR-ELWVQLKEEYQSLIRYVENNKNADNDWFRLE-SNKEGT-RWFGKCYIHDFLKYEFDFIEFEIPITYPTTAPEIAVPELDGKTA KMYRGKICLTDHFKPLWARNVPKFGLAHLMALG-----LGPWLAVEPDLIQKGVIQHKEK
Bflo1000025213	<i>Branchiostoma floridae</i>	MVDAATKKTLSNIPLLTKAGPRDR-DLWVQLKEEYQALIQYVENNKKEDNDWFRLE-SNKEGT-RWFGKCWYQDLMKYEFDFDIPITFPTTAPEIAIPELEGKTA KMYRGKICLTDHFKPLWRGNVPKFGLAHLMALG-----LGPWLAVEPDLISKGLVQHKDD
11597394	<i>Strongylocentrotus purpuratus</i>	MVDAVTKKTLSNIPLLTKAGPRDK-DLWPARLKEEYQSLIKYVGNNEADNDWFRLE-SNKGDT-RWWGKCHIQNLLKYEFDFDIPVTPATSPETAIPELDGKTA KMYRGKICLTDHFKPLWGKVNPKFGIAHAMALG-----LGPWLAVEPDLIEKGIVVHKEG
72013071	<i>Strongylocentrotus purpuratus</i>	-MNKQTAKALLNYVGNNKEADNDWFRLE-SNKDGTRWWGKCHIQNLLKYEFDFDIPVTPATSPETAIPELDGKTA KMYRGKICLTDHFKPLWGKVNPKFGIAHAMALG-----LGPWLAVEPDLIEKGIVVHKEG
Nucel100018661	<i>Naematoctetilla westensis</i>	MVDAATKKTLSNIPLLTKAGPRDR-KDWVPLKEEYQSLIKYVGNNEADNDWFRLE-SNKEGT-RWFGKCYIHDLLKYEFDFDIPITYPTTAPEIAVPELDGKTA KMYRGKICLTDHFKPLWRGNVPKFGLAHLMALG-----LGPWLAVEPDLIEKGIVVHKEK

Nvec100018661	<i>Nematostella vectensis</i>	MVDEATKKTLAAIPLLTKAGPRDG-KDWVDRLKEEYTSLIKVSNNKEADNDWFRLE-SNKEGT-RWFGKCYIHNLKYEFDVEFIDIPITYPTTAPEIALPELDGKTA KMRYRGKICMTDHFKPLWRNVPKFGIAHAMALG-----LGPWLAVEIPDLIEKGLIKHKDK
66561255	<i>Apis mellifera</i>	MVDESTKKTLSNIPLLQTAKAGPRDK-ELWVQLKEEYQALIQYVKNNEKSDNDWFRLE-SNKEGT-RWFGKCYIHNLKYEFDVEFIDIPVTYPTTAPEIALPELDGKTA KMRYRGKICLTDHFKPLWRNVPKFGIAHAMALG-----LGPWLAVEIPDLIEKGAIKHKDK
91084695	<i>Tribolium castaneum</i>	MVDESTRKTLSSIPLLTKAGPRDK-ELWVQLKEEYQSLIKYVQNNKDADNDWFRLE-SNKEGT-RWFGKCFIHDLKYEFDVEFIDIPMYPSTAPEIALPELDGKTA KMRYRGKICLSDHFKPLWRNVPKFGIAHAMALG-----LGPWLAVEIPDLIAKGVVTHKEK
58389907	<i>Anopheles gambiae</i> str PEST	MVDGTRKALSGIPLLTKAGPRDK-ELWVQLKEEYQALIKYVQNNKASDMDWFRLE-SNKEGT-KWFGKCYWMYNLKYEFDVEFIDIPITYPTSTAPEIALPELDGKTA KMRYRGKICLTDHFKPLWRNVPKFGIAHAMALG-----LAPWLAVEVPDLIEKGVISYQEK
19922352	<i>Drosophila melanogaster</i>	MVDDSTRKTLSNIPLLQIRAGPREK-DVWVQLKEEYQALIKYVENNKQSGSDWFRLE-SNKEGT-KWFGKCYMHNLKYEFDVEFIDIPVTYPTTAPEIALPELDGKTA KMRYRGKICLTDHFKPLWRNVPKFGIAHAMALG-----LAPWLAVEIPDLIEKGIVTYKEK
consensus/100%	pb..11.h..s.k..Whpl..ss..sp..s.h.h.p..aP.ss.-h.l.pl.Gbt.KMY..

consensus/95%s.hp.psuPbp...W.RhpEh.uLI.al..NK..spcWhpl..sN..sT..W.GpCW.h.p..pYcFphph-ls.sYP.ss.-l.l.pL-Gbt.KMY+.....
consensus/90%l..IPhp.pSgPR-...W..RL+EEa.uLIpYlp.NK..DscWFplp.uN..GT..pWbGpCW.h.p..+Y-FphpF-IPhsYP.ss.El.lPpL-GKTsKMYRGG+IChs.HF.PLW.+NsP.aGlsH.h.Lu.....LuPWhu.E1P.hl..s.l.....
consensus/85%s1..IP1hp.pSgPRD...W..RLKEEA.uLIpYlp.NK..sDs-WFplp.uN..pGT..pWbGcCW.h.phb+YEfphpF-IPhoYP.ss.El.lPpL-GKTsKMYRGG+IChs.HF.PLW.+NsP.aGluHhhsLG.....LuPWLuhe1P.hl..G.l.....
consensus/80%+ps1p.IP1lpSpAGPRD..c.W.pRLKEYY.uLIpYvpNkpsDs-WFpl-.SNCpGT.+WbG+CW.haphb+YEfchpF-IPhTYP.ssPEIh1PEL-GKTsKMYRGKIChs.HF.PLW.+NsP+FGLuHhhALG.....LuPWLahe1P.LlppGh1.....
consensus/75%-s+ps1p.IP1lpSpAGPRD..-bW.pRLKEYY.uLIpYvpNkpsDsDWF+LE.SNccGT.+WhGKCWahashhKYEF-1pF-IP1TYPsosPEIh1PELDGKTsKMYRGKIClo.HFbPLWu+NsP+FGLAHHhALG.....LuPWLahe1P.LlppGh1.....
consensus/70%D..s+ps1pIP1lpSpAGPRD..-bW.pRLKEYY.uLIpYvcNkpsDNDFW+LE.SNc-GT.+WaGKCWalashhKYEF-1EF-IP1TYPsTsPEIh1PELDGKTuKMYRGKIClo.HFbPLWu+NsP+FGLAHuhALG.....LuPWLahe1PsLlpcGh1.....

Ng100001784	<i>Nucifera gracilis</i>	ACTIVELNSRAS KEELEKRL INSENCRN CWRVLL VDNTVWTVL LINTI TISSTIASG ERSIQE RINGC DIVEVIAN	TPELTIAIR VSTRI QIEQEVQ PGESECVPLTR	MMALNCESMSITSKRKSLSL	NAECL	KTFRP DVAALNTRV	III
Crei100012978	<i>Chlamydomonas reinhardtii</i>	-MRS-LQRCPLATGPKPQISHIEPV--NDNMILLWRLRVLPFAFD-EDVAAGRQLNADLRLGQMGRGSGQDYILMEVSFPQD-	-YPTNPFFLRVSPRC--VMYTGHVTAGGSICIEALVA-	-TGGPGGWQP-DYCEAVILVLVLANMLTAEVAQVRTATGPGGISQPLRVDL-	-SAGLAPYSDFEARAYDRTVANHG		
116058204	<i>Ostreococcus tauri</i>	MAESR--VLYREFKA-LKRREGELFDNLNTMCSEDL--GVWRFELR-SHHFD--DHLIVECSFRLNGSAPFPTCPPLIRVVRPM-	--KWTGHTVLSGGSFCTEMLVN-	-TKGLNGWRS-NYRIEAVIQLVALISAVHCPAIVISTPYNGRQVSGPLRVDL--SCEYTDNVLREYEAVSFTRAEEAHR			
ci0100140471	<i>Ciona intestinalis</i>	LETRPKTVNGR--RLMKEFRF-LRKASEDSE--GAFEVFPYSEDGEGLNDLSEWDIR-YKFID--ESSHLWRS--MQDHGI	--DCVRFRISFPAD--FPFSPPFVRLVSPYI--ENGFMNGGAICLEVLT-	-QGWSS-AYTVEALLVQVAALAH--GAVVSK-HQKRQHQKLTKRAEAEFQRIVKIHS			
Mbre100002317	<i>Monosiga brevicollis</i>	MQK---RLMQELKL-LDKCTS IKD--GVFEVSL--VNNDLFEWDVM-LHKFD-PDTLIAHD--LEMMRRSHGV--	-PDTLIAHD--LEMMRRSHGV--GSIWLRSFPQN--FPFVPFPVRLAPVV--HGGFVLSGGAVCMELLTP-	-DGWSQ-AYRMESVILQTMSTIGK--GQARIVR-QVRRPLTDDEAKRSYDHLVRVHK			
115966376	<i>Strongylocentrotus purpuratus</i>	EGAAEGAEKGAE--AVKKALEDGGFMTNEKLT--GPPYAVEL--VNDSLADWNIK-LFHVN--PKSSLSKD--MKKHNY	--EYILFNMTFPDN--FPFVPFPVRLVSPHV--EYGYVLDGGAI CMELLTP--QGWSS-SYTIDAVIMQLGATLVA-	-GDGRIVQ-SSFRSDVPFNKQEAEVSFRAIVETHE			
91082929	<i>Trifolium castaneum</i>	DGPSHKSVRAR--RLMKEYRD-LORLONSKT--DPVETVEL--VDDNLFEWHVVK-VYKLD-AESELGND--MKELG	-VYKLD-AESELGND--MKELG--NYILLHVVEPEN--FPFAPPFMRVISPRI--EKGFVMEGGAICMELLTP--RGWAS-AYTVEAVIMQFAASVVK-	-GOGRIQR-KTKGOKVESKRATAEESFRSLVKTHD			

91082929	<i>Trifolium castaneum</i>	DGFSHKSVRAR---RLMKERYD-LQRLQNSKI--DPVFIEL---VDDNLFEWHVR-VYKLD--AESELGND---MKELG1-----N YILLHV F PEN---FPFAPPFMVISPRI----EKGFMVEGGAICMELLTP----RGWAS-AYTVEAVIMQFAASVVK----QGQR1QR----K1KGQKVFSK1AEESFRSLVKTHD
48104639	<i>Apis mellifera</i>	MAMPDKRVLRL---RLMKELSE-IQRMQHRL---STFTAEL---VNDNLFEWHVR-LHKID--PESELAAD---MRELNI-----PYILLHVVF PEN---FPFAPPFMRVISPRI----EKGFMVEGGAICMELLTP----RGWAS-AYTIEAVITQFAASIVK----QGGRVAR----KPKTNKEFNRRSAEESFRSLVKTHE
19920874	<i>Drosophila melanogaster</i>	VAAPDHTRTR---RLMKEYRE-MERLQAKND---AVFTVEL---VNDSLFEWHVR-LHVID-PDSPLARD---MAEMGV-----PAILLHLSFPDN---FPFAPPFMRVPHI----EKGYVMEGGAICMELLTP----RGWAS-AYTVEAVIMQFAASVVK----QGGR1AR----PKNSTKEFTRRQAEEESFRSLVKTHE
58386327	<i>Anopheles gambiae str PEST</i>	SKPLDTTIRSR---RLMKELKE-IERLQHSRT--DPCFTVEL---INDNLYEWHAR-LFRID-PDSPLAED---LVELNI-----PFPFAPPFMRVVEPRI----EKGFMVEGGAICMELLTP----RGWAS-AYTVEAILMQFAASLVK----QGGRVS---PKSAKDFSRRSAEAEFRSLVKTHE
Bflo100000133	<i>Branchiostoma floridae</i>	-MRELQE-VFKHKG---VFFVDL---VEDNLFEWNVK-LYKVD-PDSEFYQD---MQEVGT-EFVLLN-TF PEN---FPFAPPFMRLTPKI---ENGYVMDGGAICMELLTP----KGWSS-AYTVEAIIQLQSASLVK----GRGRINR----KTKKKDTYTRSHAEASFRNVKTHE
Nvec1000021943	<i>Nematostella vectensis</i>	-MKEFQD-VSRKTER---IFSAEL---VDDNLFEWNVK-LHTID-GDSLLYRD---MVETGS-KFILLNITF PEN---FPFAPPFMRVLA PRI---EGGFVLDGGAICMELLTP----KGWSS-AYTVEAVVLQFSAAVVK----GKGRIDR----TCKKAFAFSKKEAESAYKRLVKTHD
114598936	<i>Pan troglodytes</i>	ROOHCTOVI NR---RLMKELQD-TAPLTDPEI---SVEI---VDESLFDWNVK-LHQVD-KDSV1WOD---MKEETNT-EFILLNITF PEN---FPFAPPFMRVLSPRI----ENGYVILDGGAICMELLTP----KGKSKKSESRKAETEKS1LVKTHE

114598936	Pan troglodytes	RQQHCTQVLNR---RLMKELQD- IARLIDRFI-----SVEL----VDES LFDWNVK- LHQVD--KDSVLWQD---MKE TNT-----E FILLNLTFPDN---FPFSPPFMRVLSPRL-----ENGYVLDGGAICMELLTP-----RGWSS- AYTVEAVMRQFAASLVK-----GQGRICR-----KAGKSKKFSRKEAEATFKSLVKTHE
94395457	Mus musculus	RQQHCTQVRSR---RLMKELQD- IARLSDRFI-----SVEL----VNENLFDWNVK- LHQVD--KDSVLWQD---MKE TNT-----E FILLNLTFPDN---FPFSPPFMRVLSPRL-----ENGYVLDGGAICMELLTP-----RGWSS- AYTVEAVMRQFAASLVK-----GQGRICR-----KAGKSKKFSRKEAEATFKSLVKTHE
109504540	Rattus norvegicus	RQQHCTQVRSR---RLMKELQD- IARLSDRFI-----SVEL----VNENLFDWNVK- LHQVD--KDSVLWQD---MKE TNT-----E FILLNLTFPDN---FPFSPPFMRVLSPRL-----ENGYVLDGGAICMELLTP-----RGWSS- AYTVEAVMRQFAASLVK-----GQGRICR-----KAGKSKKFSRKEAEATFKSLVKTHE
94396224	Mus musculus	RQQHCTQVRSR---RLMKELQD- IARLSDRFI-----SVEL----VNENLFDWNVK- LHQVD--KDSVLWQD---MKE TNT-----E FILLNLTFPDN---FPFSPPFMRVLSPRL-----ENGYVLDGGAICMELLTP-----RGWSS- AYTVEAVMRQFAASLVK-----GQGRICR-----KAGKSKKFSRKEAEATFKSLVKTHE
88982452	Homo sapiens	RQQHCTQVRSR---RLMKELQD- IARLSDRFI-----SVEL----VDES LFDWNVK- LHQVD--KDSVLWQD---MKE TNT-----E FILLNLTFPDN---FPFSPPFMRVLSPRL-----ENGYVLDGGAICMELLTP-----RGWSS- AYTVEAVMRQFAASLVK-----GQGRICR-----KAGKSKKFSRKEAEATFKSLVKTHE
113416990	Homo sapiens	RQQHCTQVRSR---RLMKELQD- IARLSDRFI-----SVEL----VDES LFDWNVK- LHQVD--KDSVLWQD---MKE TNT-----E FILLNLTFPDN---FPFSPPFMRVLSPRL-----ENGYVLDGGAICMELLTP-----RGWSS- AYTVEAVMRQFAASLVK-----GQGRICR-----KAGKSKKFSRKEAEATFKSLVKTHE
68392292	Danio rerio	RQQHCTQVRSR---RLMKELQD- IARLSDRFI-----SVEL----ADPNLFDWNVK- LHQVD--KDSVLWQD---MKE TNT-----E FILLNLTFPDN---FPFSPPFMRVLSPRL-----ENGYVLDGGAICMELLTP-----RGWSS- AYTVEAVMRQFAASLVK-----GQGRICR-----KAGKSKKFSRKEAEATFKSLVKTHE

68392303	Danio rerio	NRQHCTQVRTR--RLMKELQE-IRRLGDSFI-----TVEL---ADDNLFDWNVK-LHQVD-KDSALWQD--MKETNT-EFILLNVTFPDN---FPFSPPFMRVLTPRL----ENGYVLDGGAICMELLTP-----RGWSS-AYTVEAVMRQFAASLVK-----GQGRICR-KAGKSKKAFSRKEAEATFKSLVKTHE
47229506	Tetraodon nigroviridis	QRPNCTQVRTR--RLMKELQE-IRRLGDNFI-----TVEL---VEDNLFEWNVK-LHQVD-KDSALWQD--MKETNT-EFIVLNVTFPDN---FPFSPPFMRVLSPRL----ENGYVLDGGAICMELLTP-----RGWSS-AYTVEAVMRQFAASLVK-----GQGRICR-KPGKSKKAFNRKEAEATFKSLVKTHE
Bflo1000009658	Branchiostoma floridae	NGEESSESEDD--GMDSDLNEDVDEED-HF-----EMDE-----VTSEK-DKKKE-EDEDIETENLIVLERLKM-NQRRDYLKEN--FPFDPPFVRIIAPII----NGGYVLGGAICMELLTK-----QGWSS-AYTIEAVIMQISATLVK-----GKARINF-NANKTPAVRNPHVCASFIVIV--
71984336	Caenorhabditis elegans	DGKVQGSITATD--RLMKEIRD-IHRSEHFKN--GIYTFELE--KEENLYQWWIK-LHKVD-EDSPLFED--MKKLKDHNQ--DHLLFSFTFNEK--FPCDPFPVRVVAPHI--NQGFVLGGAICMELLTK-----QGWSS-AYSIESCILQIAATLVK-----GRARISF-DAKHTSTYSMARAQQSFKSLQQIHA
47216182	Tetraodon nigroviridis	RRHLLGGTRERQPVRMARQIED-VRNSSCEYF--ILFLPRSF--RGKPAVTCTSL-TIRVD-PDSPLHTD--LQLKEKEGM-DYILLNFSYKDN--FPFDPPFVRVSPVL--CGGYVLGGAICMELLTK-----QGWSS-AYSIESVIMQINATLVK-----GKARVQF-GANKNQYSLARAQQAYKSLVQIHE
Nvec1000009989	Nematostella vectensis	KGTVCGSVQATD--RLMKELRD-VYRSDFKL--GNYSVYL--NNNDNLYDWSIK-IMRVD-PESVLHKD--MVQIEKQEGI--DHILLNMTFTDK--FPFDPPFVRCYVPI--QAGYVLSGGAI CMELLTP-----QGWSS-AYTIEAVIVQISATLVK-----GKARINF-QDTKKVVYSLHRAQQSFKSLVQIHE

ci0100153655	<i>Ciona intestinalis</i>	DGRISGSVQASD--RLLKELKA-IYRSESFKQ--GCYNVEL---VNDSLYEWHVQ-ILKVD--PDSHLHAD--LKEKLKANGQQ--ASIIMGVSFRDN--FPFDPPFVRVVCPVL----TGGFVLGGGAICMELLTK-----QGWSS-AYSIESLIMQIMATLVK-----GKARIQF-----GASSSTYLSRAQQSFQRLVQIHE
115923264	<i>Strongylocentrotus_purpuratus</i>	KGSVSGSGVGATD--RLMKELRD-IYRSDFKK-KKMYSVLD--VNDSLYDWNVK-IYTVD--SDSPLHAD--LCQLKEKEGK--DHILLNMTFKEH--YPDPPFVRVVWPIL----TGGYVLGGGAICMELLTK-----QGWSS-AYTIEAVILQIAATLVK-----GKARIQF-----GAGKSQYSLVRAQQSFRLSVQIHE
48109570	<i>Apis mellifera</i>	KGSVCGSVQATD--RLMKELRD-IYRSDFKK-GMYSIEL--VNDSLYEWNV--LMCVD--PDSPLHSD--LILLKEKEGK--DSILLNMLFKET--YPFEPPFVRVVHPMI----SGGYVLIGGAAICMELLTK-----QGWSS-AYTVEAVIMQISATLVK-----GKARIQFQGPGSASKVCQGQYSLARAQQSFKSLVQIHE
58396115	<i>Anopheles gambiae str PEST</i>	KGSVSGSVQATD--RLMKELRD-IYRSDFKN-NMYSIEL--VNDSIYEWNR-LMSVD--PDSPLHND--LVLLKEREKG--DSILLNIIFKET--YPFEPPFVRVVHPII----SGGYVLVGGAICMELLTK-----QGWSS-AYTVEAVIMQIAATLVK-----GKARIQF-----GPTKSLSQGQYSLARAQQSFKSLVQIHE
24639327	<i>Drosophila melanogaster</i>	KGSVSGSVQATD--RLMKELRD-IYRSDAFKK-NMYSIEL--VNESIYEWNR-LKSVD--PDSPLHSD--LQMLKEKEGK--DSILLNLFKET--YPFEPPFVRVVHPII----SGGYVLIGGAAICMELLTK-----QGWSS-AYTVEAVIMQIAATLVK-----GKARIQF-----GATKALTQGQYSLARAQQSFKSLVQIHE
Bflo100000060	<i>Branchiostoma_floridae</i>	KGAVSGSVQATD--RLMKELKN-VFRSDSLKR-GIYSVEL--VNDNLYDWNIK-LQGVD--PDSALYAD--LMVLKQKEGR--DFILLNMTFKEN--FPFDPPFVRRIAPII----NGGYVLGGGAICMELLTK-----QGWSS-AYTIEAVIMQISATLVK-----GKARINF-----NANKTPANQYSLARAQQSFKSLVQIHE

30425296	Mus musculus	SGATSGSVCASD--RLMKELRE-IYRSQSYKS--GTFSVEL---INDSLYDWHVK-LRKVD--PDSCLYRD---LQLKQKEGI---DYILLNFSFKDN---FPFDPPFVRVLPVL---SDGYVLDGGALCMEELLTN---QGWSS--AYSIESVILQINATLVK-----GKARVRF-----GVDNHYTEQVARVYKSMVLKHE
109465090	Rattus norvegicus	NGAVSGSVQATD--RLMKELRD-IYRSQSFKG--GNYAVEL---VNDSLWNVK-LLKVD--QDSALHND--LQILKEKEGA--DFILLNFSFKDN---FPFDPPFVRVVPVL---SGGYVLLGGGAICMELLTK---QGWSS--AYSIESVIMQISATLVK-----GKARVQF-----GANKSQYSLTRAQQSYKSLVQIHE
31541872	Mus musculus	NGAVSGSVQATD--RLMKELRD-IYRSQSFKG--GNYAVEL---VNDSLWNVK-LLKVD--QDSALHND--LQILKEKEGA--DFILLNFSFKDN---FPFDPPFVRVVPVL---SGGYVLLGGGAICMELLTK---QGWSS--AYSIESVIMQISATLVK-----GKARVQF-----GANKSQYSLTRAQQSYKSLVQIHE
31543906	Homo sapiens	NGAVSGSVQATD--RLMKELRD-IYRSQSFKG--GNYAVEL---VNDSLWNVK-LLKVD--QDSALHND--LQILKEKEGA--DFILLNFSFKDN---FPFDPPFVRVVPVL---SGGYVLLGGGAICMELLTK---QGWSS--AYSIESVIMQISATLVK-----GKARVQF-----GANKSQYSLTRAQQSYKSLVQIHE
114559985	Pan troglodytes	NGAVSGSVQATD--RLMKELRD-IYRSQSFKG--GNYAVEL---VNDSLWNVK-LLKVD--QDSALHND--LQILKEKEGA--DFILLNFSFKDN---FPFDPPFVRVVPVL---SGGYVLLGGGAICMELLTK---QGWSS--AYSIESVIMQISATLVK-----GKARVQF-----GANKSQYSLTRAQQSYKSLVQIHE
47220542	Tetraodon nigroviridis	NGAVSGSVQASD--RLMKELRE-IYRSQSYKT--GIYSVEL---VSDSLWHEVK-LRTVD--PDSPLHSD--LQVLKEKEGM--DYILLNFSYKDND--FPFDPPFVRVVPVL---SGGYVLLGGGAICMELLTKOVRSWFLLPAVVGROVDCLSGAFQGWSS--AYSIESVIMQINATLVK-----GKARVQF-----GANKVKADSFHRRVPROLLCLIVI--

47220542	<i>Tetradon nigroviridis</i>	NGAVSGSVQASD--RLMKELRD-IYRSQSYKT--GIYSVEL---INDSLYDWHVK-LRKVD-PDSPLHSD--LQILKEKEGI--EYILLNFSFKDN--FPFDPPFVRLPVL--SGGYVLGGGALCMELLTK-QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKVADSTIRKFRQLECLIVI-
109510440	<i>Rattus norvegicus</i>	NGTVSGSVHASN--RLMKELRD-IYRSQSYKS--GTYSVEL---INDSLYDWHVK-LRKVD-PDSPLYGD--LQILKEKEGI--EYILLNFSFKDN--FPFDPPFVRLPVL--SGGYVLGGGALCMELLTK-QGWSS-AYSIESVIMQINATLVK--GKARVRF-GADKNQYNLEAQQSVDVQMHE
51830637	<i>Mus musculus</i>	NGTVSGSVHASN--RLMKELRD-IYRSQSYKS--GTYSVEL---INDSLYDWHVK-LRKVD-PDSPLYGD--LQILNEKEGI--EYILLNFSFKDN--FPFDPPFVRLPVL--SGGYVLGGGALCMELLTK-QGWSS-AYSIESVIMQINATLVK--GKARVRF-GADKNQYNLEAQQSVDVQMHE
94408620	<i>Mus musculus</i>	NGTVSGSVHASN--RLMKELRD-IYRSQSYKS--GTYSVEL---INDSLYDWHVK-LRKVD-PDSPLYGD--LQILNEKEGI--EYILLNFSFKDN--FPFDPPFVRLPVL--SGGYVLGGGALCMELLTK-QGWSS-AYSIESVIMQINATLVK--GKARVRF-GADKNQYNLEAQQSVDVQMHE
114658251	<i>Pan troglodytes</i>	NGAVSGSVQASD--RLMKELRD-IYRSQSYKT--GIYSVEL---INDSLYDWHVK-LQKVD-PDSPLHSD--LQILKEKEGI--EYILLNFSFKDN--FPFDPPFVRLPVL--SGGYVLGGGALCMELLTK-QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKVLLRILHK
114658255	<i>Pan troglodytes</i>	NGAVSGSVQASD--RLMKELRD-IYRSQSYKT--GIYSVEL---INDSLYDWHVK-LQKVD-PDSPLHSD--LQILKEKEGI--EYILLNFSFKDN--FPFDPPFVRLPVL--SGGYVLGGGALCMELLTK-QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE
29789401	<i>Homo sapiens</i>	NGAVSGSVQASD--RLMKELRD-IYRSQSYKT--GIYSVEL---INDSLYDWHVK-LQKVD-PDSPLHSD--LQILKEKEGI--EYILLNFSFKDN--FPFDPPFVRLPVL--SGGYVLGGGALCMELLTK-QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE

29789401	<i>Homo sapiens</i>	NGAVSGSVQASD--RLMKELRD-IYRSQSYKI--GIYSVEL---INDSLYDWHVK-LQKVD--PDSPLHSD---LQILKEKEGI--EYILLNFSFKDN---FPFDPPFVRLPVL---SGGYVLGGGALCMELLTK--QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE
114658247	<i>Pan troglodytes</i>	NGAVSGSVQASD--RLMKELRD-IYRSQSYKT--GIYSVEL---INDSLYDWHVK-LQKVD--PDSPLHSD---LQILKEKEGI--EYILLNFSFKDN---FPFDPPFVRLPVL---SGGYVLGGGALCMELLTK--QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE
114658249	<i>Pan troglodytes</i>	NGAVSGSVQASD--RLMKELRD-IYRSQSYKT--GIYSVEL---INDSLYDWHVK-LQKVD--PDSPLHSD---LQILKEKEGI--EYILLNFSFKDN---FPFDPPFVRLPVL---SGGYVLGGGALCMELLTK--QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE
30725841	<i>Mus musculus</i>	NGAVSGSVQASD--RLMKELRD-VYRSQSYKA--GIYSVEL---INDSLYDWHVK-QHKVD--SDSPLHSD---LQILKEKEGI--EYILLNFSFKDN---FPFDPPFVRLPVL---SGGYVLGGGALCMELLTK--QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE
109484735	<i>Rattus norvegicus</i>	NGAVSGSVQASD--RLMKELRD-VYRSQSYKA--GIYSVEL---INDSLYDWHVK-LHKVD--SDSPLHSD---LQILKEKEGI--EYILLNFSFKDN---FPFDPPFVRLPVL---SGGYVLGGGALCMELLTK--QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE
94387187	<i>Mus musculus</i>	NGAVSGSVQASD--RLMKELRD-VYRSQSYKA--GIYSVEL---INDSLYDWHVK-LHKVD--SDSPLHSD---LQILKEKEGI--EYILLNFSFKDN---FPFDPPFVRLPVL---SGGYVLGGGALCMELLTK--QGWSS-AYSIESVIMQINATLVK--GKARVQF-GANKNQYNLARAQQSYNSIVQIHE
121908958	<i>Trichomonas vaginalis</i> C ³	EKYKPSNAGEL--RLTQDLKS-TKSMRAKDI--CESAEP--YQCHIISTWETH-LFGEE--PKDSTLYPD--TOKYKOTIGCP--DVYQEMVYERPD--YRIPBPFVWVOPP--KEMTGRVLTIGGSICADILTM--NCWNRP-SYDVSSCGESNLAELES--ONPOIDF--LNMTTRSKFAYNAVLRVAPRHG

121908958 Trichomonas vaginalis G3 EKYKPSNAGEL---RLIQDLKS-IKSMPAKDL---GFSAEPL---YQGHLSTWEIH-LFGFE--PKDSIYPD---IQKYKQLTGR---DYVQFMVFPPD---YPIRPPFVRVVQPRF---KFHTGRVTIGGSICADILTM-----NGWNP-SYDVSSCFNSNIFAEIFS-----QNPQ1DF---LNMTPYSKKEAYNAY1RVAREHG
 121900284 Trichomonas vaginalis G3 NGSSNSSVDE---RIVRDFYT-ISKRHPKEL---GFTVLP---YNNDIHVWEVR-LWSFD-KDTPIYKD---MQIFEAKTGR---NYIELRVSFPPN---YPIHPPFIRVVYPRC---TKEDGWSP-IYDFEDLLSSILVEMRS---TGERVLNGGAFCISALTL-
 121908902 Trichomonas vaginalis_G3 DYQVMNMSATDD-RIVRDYYT-ISRRKPSEL---GFSVKPY---MND-IRTWEVH-LFGFD-KKDPIYAD---IQKYKKQTGK---DYIELRVSFPPD---YPNRPFLRVISPRC-VSHGGRVTLGGAI CVSALTLDNQNGWSP-IYDFESLFMNIIAEMLN---CEPPLRIDF---NNDTPYSTREAFDSFMRLLSDHG
 Ngru1000014952 Naegleria gruberi TGASEGCVN---CIMKEYMK-IQANTEKF---GISANP---VNDDL FHWEIR-FFNF DIKEKGKIAKD---LQEYKKKNGI---DYITLDLTFPLE---FPFKPFIRVLKPRF-AFRTGHTIGGSICMELLTS---SGWSS-VCSLESIFVQIRSEMVA---GEAQLDF---SNTSPYEHEAKEAFFRVAQRYG
 Bflo1000013501 Branchiostoma floridae SYKTSGASSVAIH-RLIMDLKN-MKKTKGKF---GVEGVP---RGDNFLWDVK-LTNID-PKCP LGKD---LQOYAKHQEE-PVIKMEMKFPPD-YPMAPPFVRLKPRF---KFLTGHVTIGGSICMEMLTR---SGWRP-TNDIESILVQIRAEILS---DNNARLDK-NPNWEYTESEAKTAFHRMVNRYG
 Bflo1000003589 Branchiostoma floridae SAPRNESVSFY---VVCPDYLKNMKKTGKF---GVEGVP---RGDNFLWDVK-LTNID-PKCP LGKD---LQOYAKHQEE-PVIKMEMKFPPD-YPMAPPFVRLKPRF---KFLTGHVTIGGSICMEMLTR---SGWRP-TNDIESILVQIRAEILS---DNNARLDK-NPNWEYTESEAKTAFHRMVNRYG
 Bflo1000002552 Phaeodactylum tricornutum DAEEKECCGK---VVCAPLQVIE IMMORTNL---GEGIKPG---KEDGSMERVTKL-LEKED-FDGNLWV---YPMAPPFVRLKPRF---KFLTGHVTIGGSICMEMLTR---SGWRP-TNDIESILVQIRAEILS---DNNARLDK-NPNWEYTESEAKTAFHRMVNRYG

consensus/90%l.p-hb..l.p.p.....hp.....psh..Wpl+.h..hd..cs.l..D....h.....lbhph.F..p....aPh.PPFhRvh.P.h.....G.V..GGulCh-hLT.....pGW.s...hslEulh.ph.s.hh.....s..pl.b.....hs..u...a..h...a.
consensus/85%lhp-hbp.l.c.ps.b.....phc.....slhpWpl+.l.phd..cs.l..D....hbb.....IbhphpF..s....aPh.PPFhRvl.P.h.....psGaV..GGulChEhLT.....pGWss..shslEulh.ph.upllp.....s.uRlpb.....s...as..pA..sa.phs..H.
consensus/80% p...p.p.p....RLh+-lbp.l.+.ps.b.....hsl-....s-slapWpl+.L.phd...-osL.pD...hbbhp.....p.IllphsF.ps....aPhsPPFhRvl.P.h.....psGaVb.GGA1CMELLT.....pGWus..sYolEulh.Qh.Asllp.....GpuRlpb.....sp..as..pAb.sapp1sp.H.
consensus/75% p...ssplp.p...RLh+ELbc.lb+.psbb.....hsl-h....hs-sLacWpl+.L.plD...DSSl.pD...hbbhp.....caIIlphoF.-s....FPFsPPFhRvl.P.l.....psGaVhsGGAlCMELLT.....pGWus..uYolEuVhbQh.As1Vc.....GpuRlpb.....spp.asb.cAbpsapp1lp.Hp
consensus/70% p...ssslpsp...RLMKELb-.lbR.p.psbb....as1EL....lsDsLa-Wp+.L.plD...DSSl.apD...hbbhp.....aILLphoF.-s....FPFsPPF1Rvl.P.l.....psGaVhsGGAlCMELLT.....pGWSS..AYolEuVlhQlsAoLVK.....GpuRlpb.....s.spppasb.cAbpuapplVp.Hp

8. UBE2W

FINAL --HHHHHHH---EE-----EEEEEEE-----EEE-----EEEEE-----HHHHHHHHHH-----HHHH-----EEE-----
Ptri100006839 Phaeodactylum tricornutum NYRIQRELKAFL-SDPPSNLSV--KVG-----KNLRVWIVSIEGAKGT-VYEGERFKLRISEPPQY-PTVPPSVYFL-----PPSIPVHEHVTNGDICLSLLGK-DWRPTMTAQSIASILSILSSAQSKSLPMNDARHA-QN-KPGEYQKDWDVYHDDNC
Tposeu1000010929 Thalassiosira pseudonana NYRIQRELKEFL-KSPPPGLSV--KISG-----KNVRLWVITLSMPENT-VYAGETYKLRVOFPNDY-PTSPPSVYFL-----PPT-PRHEHVTNGDICLSLLGK-DWRPIMTAOSVAQAIIFSILCGAORKSLPMDRNA-GN-KPGOKODDWVYHDDNC

1p6d100001023	<i>Thalassiosira pseudonana</i>	NYRIQKELREFLRSPPFGESV--KISG KIVKREWVITLSMENI-VIAGEITRKEAQFENDI-FPSFPVFL	--PPI-PKHEVITINGDLSSELRK- -DWPKIMTAQSVAQAIISLGCAQKSLFDMSKIA-GN-KFGQKQDWDTHDNC
23619379	<i>Plasmodium falciparum_3D7</i>	NYRIQKELNNFL-KNPPINCTI-DVHP SNIRIWIVQYVGLENT-IYANEVKIKIIFPDNY-PLKKPPIVFL	--QKP-PKHHTVYSNGDICLSQLGD- -DYNPSLSISGLSIISMLSSAKEKQLPIDNYTHA-DA-KPGSSQNDFLYHDDKC
Ia#86.m00396	<i>Toxoplasma gondii</i>	NYRIQKELQAFL-SNPPPNCRVL-YVHP SNIRWLIELMTGMEGS-PYANEMYRKVVIPWPDY-PFSPPTCFFL	--QPA-PVHVHVSNGDVCLNLGS- -DWRPSLSISAIAVAILSMLSAKQKLPTDNAHM-DV-PAGHHGTQFLYHDDKV
50557414	<i>Yarrowia lipolytica</i> CLB122	QRIMKELERLK-TAPPNCEYHYGRSPTLLRLSYCTCTNPAHITLVEANDLETWQI NPLYTGKTFRLQFIIGPNY-PVESPQVQFMPLP-ARPIPIPHIYSNGHICLDILGS	--EWTPVQTAVSCISLQSMGLGSNDRDERPPDDEYIKHA- -PKNPQTRFVYHDDTV
71002042	<i>Aspergillus fumigatus_Af293</i>	GKRLSKELLKMK-EHLPPGISI-VKD DNLEEWQMDIKVLDDNPLYKDQTYRLKFTFGSKY-PIEPPEVQFIELPSTS DTPRIPMHPIYSNGIICLDLLGSA	--GWSPVQTVESVCMSIQSMLTANRNERPQGDAEFVSYN- -KRRPRDIAFMYDDDNV
46124999	<i>Gibberella zeae</i> PH 1	DRRLVKELGKMHKGMPGIFTI-ERND SVAGDWFVIDQVLDENPLYKDQIYRLKFHFPMKY	--PIEPPEVTFNK-QTDRIPMHPIYSNGIICLDLLGQQ -GWSPVQSVSSVCMSIQSMLTSNDKDERPPGDEFVRGN-RQRPRDIEFLYHDDTV
Crei100002793	<i>Chlamydomonas reinhardtii</i>	ORRIQSEI NEWM-RSPPEGCII-ESC EPMTSWVVI MOGAGGVRIYSDEVERVRI DEGEQY-PI.DPPDVIEI	--APA-PIPHIYSNGHICLDILYSGSRNGGWSPAI TMSKVVI SIRSMLASNTDKRKP PPGDAEYCARVG-NRSPKISNWVFEEDDKV

Cre100002793	<i>Chlamydomonas reinhardtii</i>	QRRIQSELNEWM-RSPPEGCCL---ESC-----EPMTSWVIMQAGGVRLYSDEVFRVRIDFGEQY--PLDFPDVIFL-----AP-A1PHPHIYSNGHICLDILYSGRNGWSPALMSKVLSLRSMLASNTDKRKPFGDAEYCARG-NRSPKLSNW-FEDDKV
4609348	<i>Ustilago maydis</i> 521	TKRLQKELSEIKVKGAPEGCEV--IKA-----DQLQEWQFSIQVLGNS-LYQDQKFGLRPFRFSDSY--PMESPEVVFMV---TDGFQAPVHPVYNSNGHICASILGN---EWSPLVTISSLTLQSMLASCKQLQRPPNDAYVKRA-PISPDKSRFVYDDTV
Ngru1000001737	<i>Naegleria gruberi</i>	AKRIQKEIQLG-SGVGIDIFI-NQPI-----NLLETIYCLHGARNT-IYSQEYLLKFTFSNDY-PLQSPVEVFMV---PPHVPPLHEHVYTNGHICLNILYD-GWTPVMNIQSVCMSIQSMSATQKKPIDNDTYVMSC-SHSPKNTRWNFHDDKV
18419831	<i>Arabidopsis thaliana</i>	TNRLQKEFMEWQ-TNPPSGFKH--RVS-----DNLQRWIIEVHGVPGT-LYANETYQLQVEFPEHY-PMEAPQVIFQ-HPA-PLPHPHIYSNGHICLDVLYD-SWSPAMRLSSICLISLMSLSSSVQKPKDNDYLKNCKHGRSPKETRWWFHDDKV
18410856	<i>Arabidopsis thaliana</i>	TNRLQKELEWQ-MNPPTGFKH-KVT-----DNLQRWIIEVIGAPGT-LYANDTYQLQVDPEHY-HPA-PLPHPHIYSNGHICLDILYD-SWSPAMTVSSICISLMSLSSSTEKRPTDNDRVKNCKHGRSPKETRWWFHDDKV
18401461	<i>Arabidopsis thaliana</i>	CNRLQKELSEWQ-LNPPTGFRH--KVT-----DNLQKWTIDVTGAPGT-LYANETYQLQVEFPEHY-PMEAPQVVF-SPA-PSHPHIYSNGHICLDILYD-SWSPAMTVNSVCISILSMSLSSPAKQRPADNDRVKNCKHGRSPKETRWWFHDDKV
18422281	<i>Arabidopsis thaliana</i>	CNRLQKELEWQ-VNPPTGFKH--RVT-----DNLQKWTIEVTCAPGT-LYANETYQLQVEFPEHY-PMEAPQVIFV-HPA-PLPHPHIYSNGHICLDILYD-SWSPAMTVSSVCISILSMSLSSPAKQRPTDNDRVKNCKHGRSPKETRWWFHDDKV

1842281	<i>Arabidopsis thaliana</i>	CNRIQKELSEWQ-VNPPTGFKH--RVT-----	DNLQKWVIEVTGAPGT-LYANETYNIQVEFPQH--PMEAPQVIFV-----PPA-PLPHPIYSNGHICLDILYD---SWSPAMTVSSVCISIQLSMLSSPEQRPTDNDRYVKNCNGRSPKETRWWFDDKV
124399351	<i>Paramecium tetraurelia</i>	SRRLSKELEQMQ-KSFANEFNI--KLPN-----	NEISHWIVGFEGAKGT-LYEGEKFELQFKFPNSYVEPIESPEVVFL-----GKP-PEHEHIYSNGFICLISILYD---EWSAAHNSSLCLSIQSMMSSATIKMPPNDADFVKQAT-GRGPKSYKWTFHDTKC
89300310	<i>Tetrahymena thermophila</i> SB210	TKRLQKDLREAMQ-KNYKDQFHV--TLPN-----	NDLKLWHVEFTAAQGT-VFQGEKFQKFSPEY----VIFI----GKI-PDHEHIYSNGFICLISILYD---EWSAALTVCSSVCLSLMSLSSATKKGRPFNDAEFCKRQSQ-GRGPKAFLCRQNIKKQ
124394210	<i>Paramecium tetraurelia</i>	AKRLQKDLQEQMQ-KSYVDQFNV--RMPN-----	NDIKHWIVAFEGAKGT-LYQGEKFELQFKFSNEY--PIESPEVIFI----GKP-PEHEHIYSNGFICLISILFD---EWSAALTVCSSVCLSIQSMLSSATKKMPPNDAEFKVRAA-GRGPKSFLWSYHDEKC
124397544	<i>Paramecium tetraurelia</i>	AKRLQKDLQEQMQ-KSYTDQFNV--RMPN-----	NDIKHWIVAFEGAKGT-LYQGEKFELQFKFSNEY--PIESPEVIFI----GKP-PEHEHIYSNGFICLISILFD---EWSAALTVCSSVCLSIQSMLSSATKKMPPNDAEFKVRAA-GRGPKSFLWSYHDEKC
124416285	<i>Paramecium tetraurelia</i>	AKRLQKDLQEQMQ-KSYTDQFNV--RMPN-----	NDIKHWIVAFEGAKGT-LYQGEKFELQFKFSNEYVEPIESPEVIFI----GKP-PEHEHIYSNGFICLISILFD---EWSAALTVCSSVCLSIQSMLSSATKKMPPNDAEFKVRAA-GRGPKSFLWSYHDEKC

VIIb#55.m04812	Toxoplasma gondii	SSARQRVLAVLP-PHLTDYSMM--DTD-----HDGFVWYIALEGAAGT-LYEKEVFLVRFRSPKY--PIEAPEVTFV-----PPFLPVPHVYSNGHICLISILYD---SWSPALGVSSCGMSLLSMVSSCRQKQPKADDAYCKVWG-SKSPKNVKWFHDDRI
23509349	Plasmodium falciparum_3D7	KRRLKERLELL-SQKENTIKL---LQ-----EHADKWIQITGAENT-LYSNETFQMQFKFTEKY--PIESPEVIFL-----GQP-PIHPHIYSNGHICLISILYD---HWSPVLSVNSICLSIISMSSCKKRKPLDDILYCSTGP-RISPKNMKWMFHDDKV
66827055	Dictyostelium discoideum_AX4	AKRLQKELLDLK-TNPPPCISI---TEG-----DNLDKWVIAVDGTEGS-IYQGEHFKLQFKFSSGY--PLDSPEVIFI-----GTP-PIHPHIYSNGHICLISILYD---NWSPALTVSSVCLSLMSLSGCTEKIRPTDDSKYVSRL-NKSPKEVRWMFHDDTV
17510293	Caenorhabditis elegans	TRRLMKELAQLK-SEAPEGLLVDNTSTS-----NDLKQWKIGVVGAEQT-LYAGEVFMLQFTFGPQY--PFNSPEVMFV-----GETIPAHPHIYSNGHICLISILSD---DWTPALSQSVCLSLMSLSSKEKHPIDDAIYVRTC-SKNPSKTRWWFHDHSV
Nvec100014187	Nematostella vectensis	KKRLQKELCELQ-KRPPSGMKINKDSVS-----S SLAVWVI LDGATGT-LYENEKFLLQFKFGARY--PFESPEVTFI-----GGHVPVPHVYSNGHICLISILTD---DWSPALSVEAVCISIVSMLSSCHEKRPDNNFYVSTC-HKNPKKTRWWYHDDSV
ci010014013	Ciona intestinalis	-KRIQKEILTMR-KSPPPGLRCEDSL-----RSAPEFMDVLTGATGT-LYSEQIFKLLFKFGERY--PFESPQVTFV-----GNCIPVPHVYSNGHICLISILTE---DWSPALSTEAVCLSVISMLSSCTEKLPDPNAFYIRTC-SKNPKDTKWWFHGKT-

58387116	<i>Anopheles gambiae</i> str PEST	QRLQKELMSLI-KEPPGVSVDEESV-----	-QNLQWIIINIDGVEGT-LYEGEHFQLFKNNKY--PFDSPEVTI-----	-GSNIPIPHVVYNSNGHICLSILTD---DWSPALSVQSVCLSISSMLSSCREKRPPDNGIYVKTC--KNPKKTWWYHDDSV
28573347	<i>Drosophila melanogaster</i>	ERRLHKELMSLI-KEPPGVTIDTESVQ-----	-QNLSEWKINIKGFEQT-LYEGEDFQLFKNNKY--PFDSPEVTI-----	-GTNIPVPHVVYNSNGHICLSILTE---DWSPALSVQSVCLSIASMLSSCREKRPPDNTIYVKTC--KNPKKTWWYHDDSV
115663109	<i>Strongylocentrotus purpuratus</i>	QKRLHKELMQIQ-KDPPPGIKINEEKA-----	-TVLNTWHVLDGAPNS-IYAGEKFQLQFKFSNKY--PFDSPEVVF-----	-GSNIPVPHVVYNSNGHICLSILTE---DWSPALSVQSVCLSIISMLSSCTEKRPPDNSFYVKTC--NKSPKKTWWYHDNN-
91079008	<i>Tribolium castaneum</i>	ERRLQKELMSLM-KEPPPGEVDFDLAE-----	-QNLHWWIINMEAGAT-LYEGERFQLQFKFSNKY--PFDSPEVTF-----	-GNNIPVPHIYSNGHICLSILTD---DWSPALSVQSVCLSIIVSMLSSCKEQRPPDNAFYVKTC--KNPKKTWWYHDDSV
66549055	<i>Apis mellifera</i>	KRRLQKELTSLI-REPPPGVHDEDLTS-----	-QNLQWIVHMEGAKGT-LYEQEQQFQLQFRFSSKY--PFDSPEVTI-----	-GGNIPIPHVVYNSNGHICLSILTE---DWSPALSVQSVCLSIIVSMLSSCKEKRPPDNSFYVKTC--SKNPKKTWWYHDDNV
Bflo1000025168	<i>Branchiostoma floridae</i>	QKRLQKELLALQ-KDPPPGVRVDEASVT-----	-KSLSTWQVMDGAPGT-LFEGEKFQLFKFGPRY--FDSPQVFT-----	-GPNIIPVPHVVYNSNGHICLSILTE---DWSPALSVQSVCLSIISMLSSCKEKRPPDNNFYVRTC--SKNPKKTWWYHDDSV

4793383	<i>Homo sapiens</i>	QKRLQKELLALQ-NDPPPGMTLNEKSVQ-----NSITQWIVDMEGAPGT-LYEGERFQLEIKFSSRY-PFDSPQVMFT-----GENIPVPHVYSNGHICSLTE-DWSPALSVQSVCLSIISMLSSCKEKRPPDNSFYVRTC-NKNPKKTWWYHDDTC
7576089	<i>Homo sapiens</i>	QKRLQKELLALQ-NDPPPGMTLNEKSVQ-----NSITQWIVDMEGAPGT-LYEGERFQLLFKFSSRY-PFDSPQVMFT-----GENIPVPHVYSNGHICSLTE-DWSPALSVQSVCLSIISMLSSCKEKRPPDNSFYVRTC-NKNPKKTWWYHDDTC
8279563	<i>Mus musculus</i>	QKRLQKELLALQ-NDPPPGMTLNEKSVQ-----NSITQWIVDMEGAPGT-LYEGERFQLLFKFSSRY-PFDSPQVMFT-----GENIPVPHVYSNGHICSLTE-DWSPALSVQSVCLSIISMLSSCKEKRPPDNSFYVRTC-NKNPKKTWWYHGRNT
114620514	<i>Pan troglodytes</i>	QKRLQKELLALQ-NDPPPGMTLNEKSVQ-----NSITQWIVDMEGAPGT-LYEGERFQLLFKFSSRY-PFDSPQVMFT-----GENIPVPHVYSNGHICSLTE-DWSPALSVQSVCLSIISMLSSCKEKRPPDNSFYVRTC-NKNPKKTWWYHDDTC
94364832	<i>Mus musculus</i>	QKRLQKELLALQ-NDPPPGMTLNEKSVQ-----NSITQWIVDMEGAPGT-LYEGERFQLLFKFSSRY-PFDSPQVMFT-----GENIPVPHVYSNGHICSLTE-DWSPALSVQSVCLSIISMLSSCKEKRPPDNSFYVRTC-NKNPKKTWWYHDDTC
109475992	<i>Rattus norvegicus</i>	DKRLQKELLALQ-NDPPPGMTLNEKSVQ-----NSITQWIVDMEGAPGT-LYEGERFQLLFKFSSRY-PFDSPQVMFT-----GENIPVPHVYSNGHICSLTE-DWSPALSVQSVCLSIISMLSSCKEKRPPDNSFYVRTC-NKNPKKTWWYHDDTC
4793381	<i>Homo sapiens</i>	PKRLQKELLALQ-NDPPPGMTLNEKSVQ-----NSITQWIVDMEGAPGT-LYEGERFQOLIEKESRY-PFDSPQVMFT-----GENIPVPHVYSNGHICSLTE-DWSPALSVQSVCLSIISMLSSCKEKRPPDNSFYVRTC-NKNPKKTWWYHDDTC

supplementary_material_E2_alignments.txt

```

124399658 Paramecium tetraurelia
124400960 Paramecium tetraurelia
15219165 Arabidopsis_thaliana
68437089 Danio rerio
47214958 Tetraodon nigroviridis
50234896 Mus_musculus
109492276 Rattus norvegicus
109489413 Rattus norvegicus
114670605 Pan_troglodytes
33636750 Homo_sapiens
124409420 Paramecium tetraurelia
89289979 Tetrahymena_thermophila SB210
58261412 Cryptococcus_neoformans_var_neofo
15229481 Arabidopsis thaliana
124412336 Paramecium tetraurelia
124406526 Paramecium tetraurelia
89299207 Tetrahymena_thermophila_SB210
124401200 Paramecium tetraurelia
89288400 Tetrahymena_thermophila SB210
89303321 Tetrahymena_thermophila SB210
89285001 Tetrahymena_thermophila SB210
115964468 Strongylocentrotus_purpuratus
31542455 Mus_musculus
114666526 Pan_troglodytes
12751495 Homo_sapiens
83320076 Rattus_norvegicus
47217004 Tetraodon_nigroviridis
50539720 Danio_rerio
60472388 Dictyostelium_discoideum_AX4
28830048 Dictyostelium_discoideum
91086823 Tribolium_castaneum
45550729 Drosophila_melanogaster
110756922 Apis_mellifera
58382934 Anopheles_gambiae_str_PEST
61744456 Homo_sapiens
109478992 Rattus_norvegicus
10048468 Mus_musculus
47213168 Tetraodon_nigroviridis
47213169 Tetraodon_nigroviridis
68387366 Danio_rerio
consensus/100%
consensus/95%
consensus/90%
consensus/85%
consensus/80%
consensus/75%
consensus/70%

```

10.Ubc6p

FINAL

```

45201468 Ashbya_gossypii_ATCC_10895
50309721 Kluyveromyces_lactis_NRRL_Y_1
6320947 Saccharomyces_cerevisiae
50290111 Candida_glabrata CBS_138
32563946 Caenorhabditis_elegans
21358599 Drosophila_melanogaster
58388043 Anopheles_gambiae_str_PEST
ci0100138685 Ciona_intestinalis
48132665 Apis_mellifera
91082969 Tribolium_castaneum
115969448 Strongylocentrotus_purpuratus
Bflo1000025151 Branchiostoma_floridae
37577126 Homo_sapiens
37577124 Homo_sapiens
114550572 Pan_troglodytes
85662413 Mus_musculus
56090425 Rattus_norvegicus
68442601 Danio_rerio
68442599 Danio_rerio
47212316 Tetraodon_nigroviridis
Tpseu1000010010 Thalassiosira_pseudonana
Ptri1000001394 Phaeodactylum_tricornutum
Pram1000009035 Phytophthora_ramorum
Psoj1000000332 Phytophthora_sojae
consensus/100%
consensus/95%
consensus/90%
consensus/85%
consensus/80%
consensus/75%
consensus/70%

```

List of gis, ordered by family

Each cluster of gi numbers represents a non-classical, eukaryotic E2 ligase domain variant and is preceded by a count of total members from selected lineages. Please refer to previous work for lists of prokaryotic E2 proteins (<ftp://ftp.ncbi.nih.gov/pub/aravind/UB/>). Each sequence is identified by gi, gene name, protein length, species name, lineage name, and genbank protein description.

Note: eukaryotic proteomes with unassigned gi numbers have been assigned temporary ids, prefixed by species abbreviations which follow:

Naegleria gruberi: Ngru
Phytophthora sojae: Psoj
Phytophthora ramorum: Pram
Phaeodactylum tricornutum: Ptri
Chlamydomonas reinhardtii: Crei
Thalassiosira pseudonana: Tpsen
Nematostella vectensis: Nvec
Branchiostoma floridae: Bflo

Table of Contents

1. Apg3
2. Apg10
3. Tsg101 (UEV)
4. BRUCE-like
5. UBE2W
6. Ufc1
7. Ub60p
8. Ubcl
9. AKTIP

1. Apg3

****COUNTS****

parabasalidae	1
heterolobosea	1
euglenozoa>kinetoplastida	4
alveolata	5
alveolata>ciliophora	6
stramenopiles	4
viridiplantae	3
entamoebidae	1
mycetozoa>dictyosteliida	2
fungi>ascomycota	10
fungi>basidiomycota	2
metazoa	1
metazoa>Choanoflagellida	1
metazoa>Chordata	4
metazoa>Cnidaria	1
metazoa>echinodermata	1
metazoa>hexapoda	4
metazoa>vertebrata	5
metazoa>vertebrata>actinopterygii	4

65303096	TA03605	203	Theileria annulata	alveolata	hypothetical protein, conserved [Theileria annulata].
X#46_m01688	X#46_m01688	389	Toxoplasma gondii	alveolata	autophagy associated protein, putative
66359970	cgd8_2630	242	Cryptosporidium parvum Iowa II	alveolata	APG10/ Autip like like autophagy protein involved in
71027271	TP03_0261	197	Theileria parva strain Muguga	alveolata	hypothetical protein TP03_0261 [Theileria parva strain Muguga].
23613578	PF10280c	298	Plasmodium falciparum 3D7	alveolata	hypothetical protein [Plasmodium falciparum 3D7].
124396640	GSPATT00032381001	229	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
89292028	TTHERM_01050620	280	Tetrahymena thermophila SB210	alveolata>ciliophora	Autophagy associated protein, C-terminal domain containing
124405511	GSPATT00000653001	240	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124430297	GSPATT00027413001	108	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124419471	GSPATT00018613001	241	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
892900405	TTHERM_00012980	225	Tetrahymena thermophila SB210	alveolata>ciliophora	Autophagy associated protein, C-terminal domain containing
56471771	46.t00033	268	Entamoeba histolytica HM-1:IMSS	entamoebidae	autophagy protein, putative [Entamoeba histolytica]
71421296	Tc00_1047053510257.90	270	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	autophagy associated protein [Trypanosoma cruzi strain CL
68128829	LmjF33.0295	275	Leishmania major	euglenozoa>kinetoplastida	autophagy protein, putative [Leishmania major].
84043456	Tb927.2.1890	292	Trypanosoma brucei TREU927	euglenozoa>kinetoplastida	autophagy associated protein [Trypanosoma brucei TREU927].
71649849	Tc00_1047053509213.130	270	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	autophagy associated protein [Trypanosoma cruzi strain CL
50420719	DEHA0D10846g	324	Debaromyces hansenii CBS767	fungi>ascomycota	hypothetical protein DEHA0D10846g [Debaromyces hansenii CBS767].
50310697	KLLA0F06402g	302	Kluyveromyces lactis NRRL Y-1140	fungi>ascomycota	unnamed protein product [Kluyveromyces lactis].
50290141	CAGL0I05808g	309	Candida glabrata CBS 138	fungi>ascomycota	hypothetical protein CAGL0I05808g [Candida glabrata CBS138].
6324334	ATG3	310	Saccharomyces cerevisiae	fungi>ascomycota	Atg3p [Saccharomyces cerevisiae].
42548126	FG08900.1	349	Gibberella zeae PH-1	fungi>ascomycota	hypothetical protein FG08900.1 [Gibberella zeae PH-1].
85112065	NCU01955.1	352	Neurospora crassa OR74A	fungi>ascomycota	hypothetical protein [Neurospora crassa OR74A].
70998182	AFU4_5G08170	353	Aspergillus fumigatus Af293	fungi>ascomycota	autophagy protein Atu1, putative [Aspergillus fumigatus
50553884	YALI0E24453g	366	Yarrowia lipolytica CLIB22	fungi>ascomycota	hypothetical protein [Yarrowia lipolytical].
19113456	SPBC3B9.06c	275	Schizosaccharomyces pombe 972h-	fungi>ascomycota	hypothetical protein SPBC3B9.06c [Schizosaccharomyces pombe 972h-].
45198343	AGOS_APL178W	282	Ashbya gossypii ATCC 10895	fungi>ascomycota	APL178Wp [Ashbya gossypii ATCC 10895].
58260390	CNK00360	385	Cryptococcus neoformans var. neoformans JEC21	fungi>basidiomycota	autophagy-related protein [Cryptococcus neoformans var. neoformans
71003271	UM00169.1	439	Ustilago maydis 521	fungi>basidiomycota	hypothetical protein UM00169.1 [Ustilago maydis 521].
Ngru1000009601	Ngru1000009601	304	Naegleria gruberi	heterolobosea	fgeneshNG pg scaffold 40000109
17543646	Y55F3AM_4	305	Caenorhabditis elegans	metazoa	Y55F3AM_4 [Caenorhabditis elegans].
Mbre1000005126	Mbre1000005126	342	Monosiga brevicollis	metazoa>Choanoflagellida	fgenesh2_pg_scaffold_25000124
cio100131803	cio100131803	254	Ciona intestinalis	metazoa>Chordata	cio100131803
Bfl01000027008	Bfl01000027008	1023	Branchiostoma floridae	metazoa>Chordata	estExt fgenesh2_pg.C_4000088
Bfl01000029400	Bfl01000029400	318	Branchiostoma floridae	metazoa>Chordata	estExt fgenesh2_pg.C_1590012
cio100142936	cio100142936	312	Ciona intestinalis	metazoa>Chordata	cio100142936
Nvec1000006208	Nvec1000006208	329	Nematostella vectensis	metazoa>Cnidaria	e_gw_31_166.1
115956521	LOC594522	268	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: similar to MGC80121 protein [Strongylocentrotus
91092526	LOC655634	316	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to CG6877-PA [Tribolium castaneum].
58379448	Agap_ENSANGG00000012603	331	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP00000015092 [Anopheles gambiae str. PEST].
66567468	LOC552315	327	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to Aut1 CG6877-PA [Apis mellifera].
21357935	Aut1	330	Drosophila melanogaster	metazoa>hexapoda	Aut1 CG6877-PA [Drosophila melanogaster].
19526773	ATG3	314	Homo sapiens	metazoa>vertebrata	Atg3p [Homo sapiens].
114588464	ATG3	318	Pan troglodytes	metazoa>vertebrata	PREDICTED: Atg3p isoform 2 [Pan troglodytes].
19705511	Apg31	314	Rattus norvegicus	metazoa>vertebrata	PREDICTED: Atg3p isoform 1 [Pan troglodytes].
114588462	ATG3	314	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein isoform 1 [Pan troglodytes].
13385890	Atg3	314	Mus musculus	metazoa>vertebrata	autophagy Atg3p>Autip-like [Mus musculus].
41053345	apg31	317	Danio rerio	metazoa>vertebrata	APG3 autophagy 3-like [Danio rerio].
47226783	GSTEN:00027065:G:001	133	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
47226782	GSTEN:00027064:G:001	135	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
68355530	LOC563944	66	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to APG3 autophagy 3-like [Danio rerio].
66810830	DBBDRAFT_0217934	343	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DBBDRAFT_0217934 [Dictyostelium discoideum
28828597	-	338	Dictyostelium discoideum	mycetozoa>dictyosteliida	hypothetical protein [Dictyostelium discoideum].
121911929	TVAG_447140	258	Trichomonas vaginalis G3	parabasalidae	autophagy associated protein, putative [Trichomonas
Tpsen1000007918	Tpsen1000007918	315	Thalassiosira pseudonana	stramenopiles	145968
Psj1000017156	Psj1000017156	337	Phytophthora sojae	stramenopiles	110627
Ptri1000002031	Ptri1000002031	303	Phaeodactylum tricornutum	stramenopiles	fgenesh1_pg.C_chr_1000016
Pram1000014227	Pram1000014227	323	Phytophthora ramorum	stramenopiles	47287
18424474	AT5G61500	313	Arabidopsis thaliana	viridiplantae	autophagy 3 (APG3) [Arabidopsis thaliana].
Crei1000012355	Crei1000012355	307	Chlamydomonas reinhardtii	viridiplantae	estExt fgenesh2_pg.C_50319
116058478	Ot06g04060	81	Ostreococcus tauri	viridiplantae	Protein involved in autophagy during starvation (ISS)

2. Apg10

****COUNTS****

parabasalidae	1
stramenopiles	2
viridiplantae	2
entamoebidae	1
mycetozoa>dictyosteliida	1
fungi>ascomycota	3
fungi>basidiomycota	1
metazoa	1
metazoa>Chordata	3
metazoa>Cnidaria	1

supplementary_material_E2_gis.txt

metazoa>hexapoda	3	
metazoa>vertebrata	4	
metazoa>vertebrata>actinopterygii	1	
56467139	215.t00010	198 Entamoeba histolytica HM-1:IMSS
19113870	SPAC227.04	179 Schizosaccharomyces pombe 972h-
50425969	DEHA0G01309g	221 Debaryomyces hansenii CBS767
42544793	FG00611.1	313 Gibberella zae PH-1
58267542	CNE03590	782 Cryptococcus neoformans var. neoformans JEC21
71984851	D2085.2	157 Caenorhabditis elegans
cio100138026	ci0100138026	101 Ciona intestinalis
Bflc1000011200	Bflc1000011200	218 Branchiostoma floridae
Bflc1000028874	Bflc1000028874	241 Branchiostoma floridae
Nvec1000017690	Nvec1000017690	218 Nematostella vectensis
110761124	LOC726781	197 Apis mellifera
58378916	ENSANGG000000018063	135 Anopheles gambiae str. PEST
28573298	CG12821	130 Drosophila melanogaster
18594496	ATG10	220 Homo sapiens
55624036	LOC461768	220 Pan troglodytes
109464245	LOC688555	214 Rattus norvegicus
27777646	Atg10	215 Mus musculus
42707141	GSTEN:00005155:G:001	141 Tetraodon nigroviridis
60475073	DBBDRAFT_0190076	230 Dictyostelium discoideum AX4
121893860	TVAG_457890	145 Trichomonas vaginalis G3
Psoj1000014309	Psoj1000014309	239 Phytophthora sojae
Pram1000006756	Pram1000006756	186 Phytophthora ramorum
30680336	AT3G07525	226 Arabidopsis thaliana
30680332	AT3G07525	225 Arabidopsis thaliana
		entamoebidae
		fungi>ascomycota
		metazoa>Chordata
		metazoa>Cnidaria
		metazoa>hexapoda
		metazoa>vertebrata
		mycetozoa>dictyosteliida
		hypothetical protein 215.t00010 [Entamoeba histolytica HM-1:IMSS].
		hypothetical protein SPAC227.04 [Schizosaccharomyces pombe 972h-].
		hypothetical protein DEHA0G01309g [Debaromyces hansenii CBS767].
		hypothetical protein FG00611.1 [Gibberella zae PH-1].
		hemolysin [Cryptococcus neoformans var. neoformans JEC21].
		D2085.2 [Caenorhabditis elegans].
		ci0100138026
		fgenesh2 pg.scaffold_138000045
		estExt fgenesh2 pg.C 13800027
		fgenesh1 pg.scaffold_108000021
		PREDICTED: similar to autophagy-related 10-like [Apis mellifera].
		ENSANGP00000020552 [Anopheles gambiae str. PEST].
		CG12821-PA [Drosophila melanogaster].
		APG10 autophagy 10-like [Homo sapiens].
		PREDICTED: hypothetical protein [Pan troglodytes].
		PREDICTED: similar to autophagy-related 10-like [Rattus
		autophagy-related 10-like [Mus musculus].
		unnamed protein product [Tetraodon nigroviridis].
		hypothetical protein DBBDRAFT_0190076 [Dictyostelium discoideum
		hypothetical protein TVAG_457890 [Trichomonas vaginalis G3].
		142106
		76873
		autophagocytosis-associated family protein [Arabidopsis thaliana].
		autophagocytosis-associated family protein [Arabidopsis thaliana].

3. Tsg101 (UEV)

COUNTS	
parabasalidae	2
euglenoza>kinetoplastida	13
alveolata	2
alveolata>ciliophora	3
stramenopiles	5
viridiplantae	5
entamoebidae	6
fungi>ascomycota	8
fungi>basidiomycota	2
metazoa	1
metazoa>Choanoflagellida	1
metazoa>Chordata	10
metazoa>Cnidaria	1
metazoa>echinodermata	1
metazoa>hexapoda	5
metazoa>vertebrata	14
metazoa>vertebrata>actinopterygii	12
mycetozoa>dictyosteliida	1

65305040	TA06765	294 Theileria annulata
71033227	TP01_0734	309 Theileria parva strain Muguga
124407857	GSPATT00009545001	444 Paramecium tetraurelia
89285001	TTHERM_01027520	512 Tetrahymena thermophila SB210
124396952	GSPATT00005706001	410 Paramecium tetraurelia
67478665	46_t00043	434 Entamoeba histolytica HM-1:IMSS
67466467	310_t00002	322 Entamoeba histolytica HM-1:IMSS
67481843	15_t00074	250 Entamoeba histolytica HM-1:IMSS
56468853	134_t00016	212 Entamoeba histolytica HM-1:IMSS
56469948	96_t00001	111 Entamoeba histolytica HM-1:IMSS
67466719	Tc00_10470535005	494 Entamoeba histolytica HM-1:IMSS
71651794	Tc00_1047053505025	309 Trypanosoma cruzi strain CL Brener
71408683	Tc00_1047053509031	295 Trypanosoma cruzi strain CL Brener
71413390	Tc00_1047053509718	293 Trypanosoma cruzi strain CL Brener
70834758	Tb11_01_4703	241 Trypanosoma brucei
71666894	Tc00_1047053506459	235 Trypanosoma cruzi strain CL Brener
71666892	Tc00_1047053506459_195	247 Trypanosoma cruzi strain CL Brener
71419864	Tc00_1047053508869	293 Trypanosoma cruzi strain CL Brener
71656365	Tc00_1047053508871	247 Trypanosoma cruzi strain CL Brener
71413135	Tc00_1047053509097	247 Trypanosoma cruzi strain CL Brener
71659407	Tc00_1047053504081	293 Trypanosoma cruzi strain CL Brener
71659181	Tc00_1047053510377	293 Trypanosoma cruzi strain CL Brener
71655155	Tc00_1047053506967	293 Trypanosoma cruzi strain CL Brener
71662378	Tc00_1047053510021	247 Trypanosoma cruzi strain CL Brener
45184818	AGOS_AAL006C	445 Ashbya gossypii ATCC 10895
50556412	VPS23	378 Yarrowia lipolytica CLIB122
50426875	DEHA0G12397g	496 Debaryomyces hansenii CBS767
42547092	FG02656.1	519 Gibberella zae PH-1
85093619	NCU05753.1	580 Neurospora crassa OR74A
70998933	AFU4_3G14380	579 Aspergillus fumigatus Af293
50289071	CAGLOH03927g	468 Candida glabrata CBS 138
42759852	STP22	385 Saccharomyces cerevisiae
46098756	UM02831.1	1157 Ustilago maydis 521
58264686	CNC0240	550 Cryptococcus neoformans var. neoformans JEC21
71982582	C09G12.9	425 Caenorhabditis elegans
Mbre1000007539	Mbre1000007539	1347 Monosiga brevicollis
cio100154542	ci0100154542	317 Ciona intestinalis
Bflc1000012054	Bflc1000012054	484 Branchiostoma floridae
cio100154515	ci0100154515	419 Ciona intestinalis
Bflc1000010581	Bflc1000010581	401 Branchiostoma floridae
Bflc1000016007	Bflc1000016007	306 Branchiostoma floridae
Bflc1000047714	Bflc1000047714	372 Branchiostoma floridae
Bflc1000042242	Bflc1000042242	69 Branchiostoma floridae
Bflc1000036388	Bflc1000036388	116 Branchiostoma floridae
Bflc1000042438	Bflc1000042438	119 Branchiostoma floridae
Bflc1000048372	Bflc1000048372	144 Branchiostoma floridae
Nvec1000024634	Nvec1000024634	432 Nematostella vectensis
115681892	LOC590965	108 Strongylocentrotus purpuratus
91086823	LOC662712	4316 Tribolium castaneum
17648053	TSG101	408 Drosophila melanogaster
66531777	LOC409438	416 Apis mellifera
91080041	LOC661457	392 Tribolium castaneum
58388004	ENSANGG000000011072	376 Anopheles gambiae str. PEST
109458775	LOC691172	683 Rattus norvegicus
114636495	LOC748881	215 Pan troglodytes
23943814	UEV3	379 Homo sapiens
114636489	LOC748881	379 Pan troglodytes
5454140	TSG101	390 Homo sapiens
109461971	LOC688000	470 Rattus norvegicus
114636476	TSG101	390 Pan troglodytes
48374087	Tsg101	391 Rattus norvegicus
11230780	Tsg101	391 Mus musculus
114636491	LOC748881	357 Pan troglodytes
114636478	TSG101	403 Pan troglodytes
70780379	Attp	250 Mus musculus
114636483	LOC748881	449 Pan troglodytes
114636481	LOC748881	471 Pan troglodytes
68368272	LOC562443	816 Danio rerio
68356050	LOC567642	984 Danio rerio
68364140	LOC558584	396 Danio rerio
68405599	zgc:100959	471 Danio rerio
68364142	LOC558584	393 Danio rerio
47203297	GSTEN:00012595:G:001	192 Tetraodon nigroviridis
50344832	GSTEN:00019782:G:001	390 Danio rerio
47215244	GSTEN:00019782:G:001	470 Tetraodon nigroviridis
68405637	LOC561016	185 Danio rerio
47212486	GSTEN:00004508:G:001	463 Tetraodon nigroviridis
47217083	GSTEN:00021501:G:001	371 Tetraodon nigroviridis
68405668	LOC564878	185 Danio rerio
66807691	DBBDRAFT_0218848	478 Dictyostelium discoideum AX4
121907565	TVAG_128690	264 Trichomonas vaginalis G3
121915555	TVAG_193190	293 Trichomonas vaginalis G3
		parabasalidae
		fungi>ascomycota
		metazoa>Chordata
		metazoa>Cnidaria
		metazoa>hexapoda
		metazoa>vertebrata
		mycetozoa>dictyosteliida
		hypothetical protein, conserved [Theileria annulata].
		hypothetical protein, conserved [Theileria parva strain Muguga].

Psoj1000006153	Psoj1000006153	412	Phytophthora sojae	stramenopiles	132761
Ptri1000008763	Ptri1000008763	411	Phaeodactylum tricornutum	stramenopiles	estExt Genewise1.C chr 140041
Ptri1000004948	Ptri1000004948	421	Phaeodactylum tricornutum	stramenopiles	estExt_fgenesh1_pg.C_chr_230106
Pram1000004476	Pram1000004476	444	Phytophthora ramorum	stramenopiles	79629
Tpseu1000001553	Tpseu1000001553	400	Thalassiosira pseudonana	stramenopiles	103937
116056391	Ot03g0280	438	Ostreococcus tauri	viridiplantae	putative human tumor susceptibility gene-like protein (ISS)
18399596	ATELC/ELC	398	Arabidopsis thaliana	viridiplantae	ATELC/ELC; ubiquitin binding [Arabidopsis thaliana].
Crei1000005374	Crei1000005374	525	Chlamydomonas reinhardtii	viridiplantae	Chlre2 kg.scaffold 1600223
15240732	ATELC-LIKE/ELC-LIKE	368	Arabidopsis thaliana	viridiplantae	ATELC-LIKE/ELC-LIKE [Arabidopsis thaliana].
42569747	AT2G3883	331	Arabidopsis thaliana	viridiplantae	tumor susceptibility protein-related [Arabidopsis thaliana].

4. BRUCE-like

COUNTS					
alveolata>ciliophora	12				
fungi>ascomycota	3				
fungi>basidiomycota	2				
metazoa	1				
metazoa>Chordata	7				
metazoa>Chidaria	5				
metazoa>echinodermata	1				
metazoa>hexapoda	8				
metazoa>vertebrata	12				
metazoa>vertebrata>actinopterygii	7				
mycetozoa>dictyosteliida	2				
stramenopiles	1				
viridiplantae	6				

124406526	GSPATT00008424001	4620	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124400960	GSPATT0035572001	2472	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
89285001	TTHERM_01027520	512	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena]
89289979	TTHERM_00011330	464	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena]
89299207	TTHERM_00481290	2633	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena]
89303321	TTHERM_00149190	6304	Tetrahymena thermophila SB210	alveolata>ciliophora	ubiquitin-activating enzyme E1 family protein [Tetrahymena]
124401200	GSPATT00035810001	5133	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124399658	GSPATT0034627001	2123	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
89288400	TTHERM_00028560	3915	Tetrahymena thermophila SB210	alveolata>ciliophora	ubiquitin-activating enzyme E1 family protein [Tetrahymena]
124409420	GSPATT0010919001	1850	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124398217	GSPATT0033569001	2396	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124412336	GSPATT0012991001	2601	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
85118516	NCU02688_1	440	Neurospora crassa OR74A	fungi>ascomycota	hypothetical protein [Neurospora crassa OR74A].
70999189	AFUA_3G13060	1062	Aspergillus fumigatus Af293	fungi>ascomycota	ubiquitin conjugating enzyme, putative [Aspergillus fumigatus]
42544747	FG01207.1	431	Gibberella zaeae PH-1	fungi>ascomycota	hypothetical protein FG01207.1 [Gibberella zaeae PH-1].
58261412	CNL05830	927	Cryptococcus neoformans var. neoformans JEC21	fungi>basidiomycota	hypothetical protein [Cryptococcus neoformans var. neoformans]
58268232	CNF01270	920	Cryptococcus neoformans var. neoformans JEC21	fungi>basidiomycota	hypothetical protein [Cryptococcus neoformans var. neoformans]
71981623	ubc-17	679	Caenorhabditis elegans	metazoa	UBiquitin Conjugating enzyme family member (ubc-17) [Caenorhabditis
Bf1o1000002074	Bf1o1000002074	527	Branchiostoma floridae	Branchiostoma>Chordata	fgenesh2_pg.scaffold_800077
ci0100131458	ci0100131458	1058	Ciona intestinalis	metazoa>Chordata	ci0100131458
Bf1o1000037265	Bf1o1000037265	301	Branchiostoma floridae	metazoa>Chordata	e gw.8.290.1
Bf1o1000027761	Bf1o1000027761	1307	Branchiostoma floridae	metazoa>Chordata	estExt_fgenesh2_pg.C_760117
Bf1o1000048584	Bf1o1000048584	307	Branchiostoma floridae	metazoa>Chordata	estExt_gwp_3740030
Bf1o1000011222	Bf1o1000011222	981	Branchiostoma floridae	metazoa>Chordata	fgenesh2_pg.scaffold_13800092
Bf1o1000025135	Bf1o1000025135	307	Branchiostoma floridae	metazoa>Chordata	estExt_fgenesh2_pm.C_2360005
Nvec1000007435	Nvec1000007435	158	Nematostella vectensis	metazoa>Cnidaria	e gw.69.171.1
Nvec1000012274	Nvec1000012274	222	Nematostella vectensis	metazoa>Cnidaria	e gw.2654.4.1
Nvec1000013121	Nvec1000013121	275	Nematostella vectensis	metazoa>Cnidaria	estExt_gwp.C_30788
Nvec1000015586	Nvec1000015586	1039	Nematostella vectensis	metazoa>Cnidaria	fgenesh1_pg.scaffold_10000074
Nvec1000005680	Nvec1000005680	659	Nematostella vectensis	metazoa>Cnidaria	e gw.18.24.1
159644648	LOC580131	234	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: hypothetical protein [Strongylocentrotus purpuratus].
91086823	LOC662712	4316	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to CG6303-PA [Tribolium castaneum].
91084477	LOC659924	1041	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to CG10254-PA, isoform A [Tribolium castaneum].
24649371	CG10254	1398	Drosophila melanogaster	metazoa>hexapoda	CG10254_PB, isoform B [Drosophila melanogaster].
110760609	LOC724799	1287	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to CG10254-PA, isoform A [Apis mellifera].
58383761	ENSANGP0000002034	1220	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP0000002034 [Anopheles gambiae str. PEST].
45550729	Bruce	4876	Drosophila melanogaster	metazoa>hexapoda	Bruce CG6303-PA [Drosophila melanogaster].
58382934	IAP6	4760	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP0000002826 [Anopheles gambiae str. PEST].
110756922	LOC411115	4673	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to Bruce CG6303-PA [Apis mellifera].
61744456	BIRC6	4829	Homo sapiens	metazoa>vertebrata	baculoviral IAP repeat-containing 6 [Homo sapiens].
114666526	LOC455196	371	Pan troglodytes	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E2Z [Pan
31542455	Ube2z	350	Mus musculus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2Z (putative) [Mus musculus].
10048468	Birc6	4845	Mus musculus	metazoa>vertebrata	baculoviral IAP repeat-containing 6 [Mus musculus].
109478992	Birc6_predicted	5017	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to baculoviral IAP repeat-containing 6 [Rattus
50234896	Ube2o	1288	Mus musculus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2O [Mus musculus].
33636750	UBE2O	1292	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2O [Homo sapiens].
12751495	UBE2Z	246	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2Z (putative) [Homo sapiens].
109492276	Ube2o_predicted	1312	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E2O [Rattus
109489413	Ube2o_predicted	1315	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E2O [Rattus
83320076	Ube2z	246	Rattus norvegicus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2Z (putative) [Rattus norvegicus].
114670605	UBE2O	1392	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2O [Pan troglodytes].
47213169	GSTEN:00010000:G:001	2135	Tetradon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
68387366	LOC566895	354	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to Baculoviral IAP repeat-containing protein 6
68437089	LOC564215	837	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to RIKEN cDNA B230113M03 [Danio rerio].
50539720	Zgc:92419	248	Danio rerio	metazoa>vertebrata>actinopterygii	hypothetical protein LOC436602 [Danio rerio].
47213168	GSTEN:00009999:G:001	3047	Tetradon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
47214958	GSTEN:00032514:G:001	1309	Tetradon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
47217004	GSTEN:00020460:G:001	403	Tetradon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
28830048	-	2208	Dictyostelium discoideum	mycetozoa>dictyosteliida	similar to Homo sapiens (Human). Hypothetical protein FLJ13786 (EC
60472388	DBDBRAFT 0217536	910	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DBDBRAFT 0217536 [Dictyostelium discoideum
Ptri1000008184	Ptri1000008184	415	Phaeodactylum tricornutum	stramenopiles	estExt_fgenesh_pg.C.chr.190218
15219165	UBC26	1163	Arabidopsis thaliana	viridiplantae	UBC26 (ubiquitin-conjugating enzyme 25);

17510293	ubc-16	152	Caenorhabditis elegans	metazoa	Ubiquitin Conjugating enzyme family member (ubc-16) [Caenorhabditis elegans].
46099348	UM03443.1	152	Ustilago maydis	fungi>basidiomycota	hypothetical protein UM03443.1 [Ustilago maydis 521].
Nvec1000014187	Nvec1000014187	153	Nematostella vectensis	metazoa>Cnidaria	estExt_gwp.C 3490026
114620514	UBE2W	151	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2W (putative) isoform 2
45387655	f1j110111	151	Danio rerio	metazoa>vertebrata>actinopterygii	hypothetical protein LOC402907 [Danio rerio].
47933385	UBE2W	151	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2W (putative) isoform 2 [Homo sapiens].
Bflc1000025168	Bflc1000025168	152	Branchiostoma floridae	metazoa>Chordata	estExt_fgenesh2_pm.C 2780002
124394210	GSPATT00030369001	150	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124397544	GSPATT00032900001	150	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
115663109	LOC582071	149	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: similar to Ubiquitin-conjugating enzyme E2W (putative)
66827055	DDBBRAFT 0202295	149	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DDBBRAFT 0202295 [Dictyostelium discoideum].
124399351	GSPATT00034321001	148	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
Ngru1000001737	Ngru1000001737	148	Naegleria gruberi	heterolobosea	fgeneshHS pm.scaffold 7200004
47223533	GSTEN:00015564:G:001	144	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
cio100140113	cio100140113	145	Ciona intestinalis	metazoa>Chordata	cio100140113
Bflc1000032579	Bflc1000032579	144	Branchiostoma floridae	metazoa>vertebrata	estExt_fgenesh2_pg.C 5140030
68356096	LOC556071	126	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to CG7220-PB, isoform B [Danio rerio].
109475992	LOC682704	396	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to CG7220-PA, isoform A [Rattus norvegicus].

6. Ufc1

COUNTS	
heterolobosea	1
alveolata	3
alveolata>ciliophora	3
euglenozoa>kinetoplastida	3
metazoa	1
metazoa>Choanoflagellida	1
metazoa>Chordata	3
metazoa>Cnidaria	2
metazoa>echinodermata	2
metazoa>hexapoda	4
metazoa>vertebrata	6
metazoa>vertebrata>actinopterygii	1
mycetozoa>dictyosteliida	1
stramenopiles	1
viridiplantae	2

65305321	TA20810	162	Theileria annulata	alveolata	protein CGI-126, putative [Theileria annulata].
71032621	TP01_0425	172	Theileria parva strain Muguga	alveolata	hypothetical protein TP01_0425 [Theileria parva strain Muguga].
V#31.m00087	V#31.m00087	158	Toxoplasma gondii	alveolata	hypothetical protein.
124411671	GSPATT00012387001	156	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124416518	GSPATT00016381001	156	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
89284506	TTHERM_01108500	180	Tetrahymena thermophila SB210	alveolata>ciliophora	hypothetical protein TTHERM_01108500 [Tetrahymena thermophila].
68125492	LmjF15.1250	171	Leishmania major	euglenozoa>kinetoplastida	hypothetical protein, conserved [Leishmania major].
71650234	Tc00.1047053506445.100	164	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	hypothetical protein [Trypanosoma cruzi strain CL Brener].
71744660	Tb09.160.4150	165	Trypanosoma brucei TREU927	euglenozoa>kinetoplastida	hypothetical protein Tb09.160.4150 [Trypanosoma brucei TREU927].
Ngru1000001149	Ngru1000001149	176	Naegleria gruberi	heterolobosea	e gw1.20.112.1
17552592	C40H1.6	162	Caenorhabditis elegans	metazoa	C40H1.6 [Caenorhabditis elegans].
Mbre1000002981	Mbre1000002981	167	Monosiga brevicollis	metazoa>Choanoflagellida	fgenesh2_pm.scaffold_3000061
cio100146879	cio100146879	166	Ciona intestinalis	metazoa>Chordata	cio100146879
Bflc1000011222	Bflc1000011222	981	Branchiostoma floridae	metazoa>Chordata	fgenesh2_pg.scaffold_138000092
Bflc10000025213	Bflc10000025213	170	Branchiostoma floridae	metazoa>Chordata	estExt_fgenesh2_pm.C 3610002
Nvec1000018661	Nvec1000018661	170	Nematostella vectensis	metazoa>Cnidaria	fgenesh1_pg.scaffold_199000025
Nvec1000012015	Nvec1000012015	116	Nematostella vectensis	metazoa>Cnidaria	e gw1.650.2.1
72013071	LOC582152	139	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: similar to conserved hypothetical protein
115937394	LOC582152	170	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: similar to conserved hypothetical protein
58389907	ENSANG000000008034	164	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP0000010523 [Anopheles gambiae str. PEST].
66561255	LOC551106	170	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to CG8386-PA [Apis mellifera].
19922352	CG8386	164	Drosophila melanogaster	metazoa>hexapoda	CG8386-PA [Drosophila melanogaster].
91084695	LOC657656	166	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to Ufm1-conjugating enzyme 1 (Ubiquitin-fold).
13384768	Ufc1	167	Mus musculus	metazoa>vertebrata	ubiquitin-fold modifier conjugating enzyme 1 [Mus musculus].
114560786	LOC457449	122	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein isoform 1 [Pan troglodytes].
705481	Ufc1	167	Homo sapiens	metazoa>vertebrata	ubiquitin-fold modifier conjugating enzyme 1 [Homo sapiens].
55588606	LOC457449	167	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein isoform 3 [Pan troglodytes].
114560788	LOC457449	148	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein isoform 2 [Pan troglodytes].
51172600	Ufc1	167	Rattus norvegicus	metazoa>vertebrata	Ufm1-conjugating enzyme 1 [Rattus norvegicus].
51011123	zgc:100800	166	Danio rerio	metazoa>vertebrata>actinopterygii	hypothetical protein LOC445256 [Danio rerio].
90970505	ufci	164	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	ubiquitin-like (UBL) post-translational modifier [Dictyostelium discoideum].
Pram1000015146	Pram1000015146	158	Phytophthora ramorum	stramenopiles	53495
Crei1000013700	Crei1000013700	167	Chlamydomonas reinhardtii	viridiplantae	estExt_fgenesh2_pg.C 240032
18396376	AT1G27530	174	Arabidopsis thaliana	viridiplantae	unknown protein [Arabidopsis thaliana].

7. Ubc6p

COUNTS					
dipomonadida group>dipomonadida group	3				
parabasalidae	2				
euglenozoa>kinetoplastida	8				
heterolobosea	2				
alveolata	4				
alveolata>ciliophora	6				
stramenopiles	11				
viridiplantae	8				
entamoebidae	3				
mycetozoa>dictyosteliida	3				
fungi>ascomycota	14				
fungi>basidiomycota	3				
metazoa	3				
metazoa>Choanoflagellida	3				
metazoa>Chordata	4				
metazoa>Cnidaria	3				
metazoa>echinodermata	2				
metazoa>hexapoda	6				
metazoa>vertebrata	10				
metazoa>vertebrata>actinopterygii	5				
50554345	YAL10E30173g	132	Yarrowia lipolytica CLIB122	fungi>ascomycota	hypothetical protein [Yarrowia lipolytica].
50550481	YAL10D1748g	252	Yarrowia lipolytica CLIB122	fungi>ascomycota	hypothetical protein [Yarrowia lipolytica].
46101486	UM05955.1	223	Ustilago maydis	fungi>basidiomycota	hypothetical protein UM05955.1 [Ustilago maydis 521].
71423664	Tc00.1047053504055.81	225	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme E2 [Trypanosoma cruzi strain CL Brener].
71655980	Tc00.104705350837.100	437	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme [Trypanosoma cruzi strain CL Brener].
71668166	Tc00.1047053506435.300	225	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme [Trypanosoma cruzi strain CL Brener].
71665910	Tc00.1047053506789.200	410	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme E2, putative [Trypanosoma brucei].
62359387	Tb927.4.3190	220	Trypanosoma brucei	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme, putative [Trypanosoma brucei].
70802852	-	310	Trypanosoma brucei	parabasalidea	Ubiquitin-conjugating enzyme family protein [Trichomonas vaginalis].
121894423	TVAG 091080	352	Trichomonas vaginalis G3	parabasalidea	Ubiquitin-conjugating enzyme family protein [Trichomonas vaginalis].
121911847	TVAG 066800	344	Trichomonas vaginalis G3	metazoa>hexapoda	PREDICTED: similar to ubiquitin-conjugating enzyme E2, J2 isoform 2
91082969	LOC662641	225	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to ubiquitin-conjugating enzyme E2, J1
91079666	LOC655113	268	Tribolium castaneum	alveolata	hypothetical protein
Ib#25					

Ptri1000008341	Ptri1000008341	119	Phaeodactylum tricornutum	stramenopiles	gw1.25.109.1
Ptri1000001394	Ptri1000001394	223	Phaeodactylum tricornutum	stramenopiles	fgenesh1_pg.C chr_12000047
Ptri1000004266	Ptri1000004266	143	Phaeodactylum tricornutum	stramenopiles	e_gw1.10.176.1
124397229	GSPATT0032588001	322	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124408394	GSPATT0010030001	214	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124394906	GSPATT0005174001	232	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124398877	GSPATT0003384901	388	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
114550572	LOC457188	267	Pan troglodytes	metazoa>vertebrata	PREDICTED: similar to ubiquitin conjugating enzyme 6 [Pan
114608464	UBE2J1	318	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2, J1 [Pan troglodytes].
116058293	Ot06g02230	437	Ostreococcus tauri	viridiplantae	Non-Canonical Ubiquitin Conjugating Enzyme 1 (ISS) [Ostreococcus
116054737	Ot14g00680	209	Ostreococcus tauri	fungi>ascomycota	Ubiquitin-protein ligase (ISS) [Ostreococcus tauri].
85099285	NCU01268.1	313	Neurospora crassa OR74A	fungi>ascomycota	hypothetical protein [Neurospora crassa OR74A].
85098002	NCU05592.1	232	Neurospora crassa OR74A	fungi>ascomycota	e_gw1.1577.2.1
Nvec1000011974	Nvec1000011974	81	Nematostella vectensis	metazoa>Cnidaria	e gw1.11.71.1
Nvec1000008459	Nvec1000008459	129	Nematostella vectensis	metazoa>Cnidaria	estExt_gwp.C_1190068
Nvec1000013828	Nvec1000013828	218	Nematostella vectensis	metazoa>Cnidaria	gw1.2.320.1
Ngru1000000443	Ngru1000000443	140	Naegleria gruberi	heterolobosea	estExt_fgeneshNG_kg.C_40012
Ngru1000015437	Ngru1000015437	209	Naegleria gruberi	heterolobosea	ubiquitin-conjugating enzyme E2, J2 homolog isoform a [Mus
85662413	Ube2j2	271	Mus musculus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2, J2 homolog isoform a [Mus
31980960	Ube2j1	318	Mus musculus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2, putative [Leishmania major].
Mbre1000001520	Mbre1000001520	3469	Monosiga brevicollis	metazoa>Choanoflagellida	ubiquitin-conjugating enzyme E2, putative [Leishmania major].
Mbre1000003059	Mbre1000003059	227	Monosiga brevicollis	metazoa>Choanoflagellida	unnamed protein product [Kluyveromyces lactis].
Mbre1000002602	Mbre1000002602	111	Monosiga brevicollis	metazoa>Choanoflagellida	ubiquitin-conjugating enzyme E2, J1 [Homo sapiens].
68223781	LmjF05_0930	464	Leishmania major	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme E2, J2 isoform 1 [Homo sapiens].
68129257	LmjF34_0900	289	Leishmania major	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme E2, J2 isoform 2 [Homo sapiens].
50309721	KLJA0E20493g	251	Kluyveromyces lactis NRRL Y-1140	fungi>ascomycota	hypothetical protein [Leishmania major].
37577122	UBE2J1	318	Homo sapiens	metazoa>vertebrata	hypothetical protein [Kluyveromyces lactis].
37577126	UBE2J2	275	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2, J1 [Homo sapiens].
37577124	UBE2J2	259	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2, J2 isoform 1 [Homo sapiens].
46127873	FG08314.1	264	Gibberella zae PH-1	fungi>ascomycota	ubiquitin-conjugating enzyme E2, J2 isoform 2 [Homo sapiens].
42565438	FG07464.1	328	Gibberella zae PH-1	fungi>ascomycota	hypothetical protein FG08314.1 [Gibberella zae PH-1].
29251108	-	177	Giardia lamblia ATCC 50803	dipomonadida group>dipomonadida group	hypothetical protein FG07464.1 [Gibberella zae PH-1].
29250521	-	324	Giardia lamblia ATCC 50803	dipomonadida group>dipomonadida group	GLP 487_28716_29249 [Giardia lamblia ATCC 50803].
7078496	GLP 546_71955_72440	161	Giardia lamblia ATCC 50803	dipomonadida group>dipomonadida group	GLP 68_18546_19520 [Giardia lamblia ATCC 50803].
67480959	22.t00035	145	Entamoeba histolytica HM-1:IMSS	entamoebidae	ubiquitin conjugating enzyme [Giardia lamblia ATCC 50803].
56472977	22.t00037	159	Entamoeba histolytica HM-1:IMSS	entamoebidae	ubiquitin conjugating enzyme [Entamoeba histolytica HM-1:IMSS].
67483836	4.t00050	344	Entamoeba histolytica HM-1:IMSS	entamoebidae	ubiquitin conjugating enzyme, putative [Entamoeba histolytica HM-1:IMSS].
21358599	CG5823	283	Drosophila melanogaster	metazoa>hexapoda	CG5823-PA [Drosophila melanogaster].
60462031	DDBDRAFT_0229815	170	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DDBDRIFT_0229815 [Dictyostelium discoideum].
60466117	DDBDRAFT_0218828	241	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DDBDRIFT_0218828 [Dictyostelium discoideum].
66813106	DDBDRAFT_0218298	401	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DDBDRIFT_0218298 [Dictyostelium discoideum].
50408085	DEHA0A09999g	242	Debaromyces hansenii CBS767	fungi>ascomycota	DEHA0A09999g [Debaromyces hansenii CBS767].
50421749	DEHA0E02871g	296	Debaromyces hansenii CBS767	fungi>ascomycota	DEHA0E02871g [Debaromyces hansenii CBS767].
68442601	LOC563187	259	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to LOC495424 protein isoform 1 [Danio rerio].
68442599	LOC563187	185	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: hypothetical protein XP_697630 [Danio rerio].
68393616	LOC554884	314	Danio rerio	metazoa>vertebrata>actinopterygii	Ubc6p like ubiquitin conjugating enzyme E2, possible transmembrane
46228870	cgd_3850	368	Cryptosporidium parvum Iowa II	alveolata	ubiquitin-conjugating enzyme E2-28.KD [Cryptosporidium parvum].
58268762	CNF02130	162	Cryptococcus neoformans var. neoformans JEC21	fungi>basidiomycota	ubiquitin-conjugating enzyme [Cryptococcus neoformans var. neoformans].
58267460	CNE03270	297	Cryptococcus neoformans var. neoformans JEC21	fungi>basidiomycota	cio100148642
cio100148642	cio100148642	293	Ciona intestinalis	metazoa>Chordata	cio100138685
cio100138685	cio100138685	233	Ciona intestinalis	metazoa>Chordata	estExt_gwp.IH.C_640021
Crei1000003831	Crei1000003831	167	Chlamydomonas reinhardtii	viridiplantae	estExt_fgenesh2_kg.C_470012
Crei1000011977	Crei1000011977	179	Chlamydomonas reinhardtii	viridiplantae	hypothetical protein CAGLO105478g [Candida glabrata CBS138].
50290111	CAGLO105478g	246	Candida glabrata CBS 138	fungi>ascomycota	UBiquitin Conjugating enzyme family member (ubc-6) [Caenorhabditis elegans].
17532649	ubc-6	314	Caenorhabditis elegans	metazoa	UBiquitin Conjugating enzyme family member (ubc-15) [Caenorhabditis elegans].
17537739	ubc-15	218	Caenorhabditis elegans	metazoa	UBiquitin Conjugating enzyme family member (ubc-26) [Caenorhabditis elegans].
32563946	ubc-26	203	Caenorhabditis elegans	metazoa	estExt_fgenesh2_pm.C_559003
Bf1o1000025151	Bf1o1000025151	240	Branchiostoma floridae	metazoa>Chordata	estExt_fgenesh2_pg.C_1470011
Bf1o1000029101	Bf1o1000029101	275	Branchiostoma floridae	metazoa>Chordata	ubiquitin conjugating enzyme (UbcF), putative [Aspergillus fumigatus Af293].
71002606	AFUA_2G16470	312	Aspergillus fumigatus Af293	fungi>ascomycota	AGR372Mp [Ashbya gossypii ATCC 10895].
45201468	AGOS_AGR372W	242	Ashbya gossypii ATCC 10895	fungi>ascomycota	UBC33; ubiquitin-protein ligase [Arabidopsis thaliana].
15240671	UHC33	243	Arabidopsis thaliana	viridiplantae	UBC34; ubiquitin-conjugating enzyme 32; ubiquitin-protein ligase [Arabidopsis thaliana].
15220062	UHC32	237	Arabidopsis thaliana	viridiplantae	UBC32; ubiquitin-conjugating enzyme 31; ubiquitin-protein ligase [Arabidopsis thaliana].
1841338	UHC32	309	Arabidopsis thaliana	viridiplantae	UBC33; ubiquitin-protein ligase [Arabidopsis thaliana].
79330462	UHC33	230	Arabidopsis thaliana	viridiplantae	PREDICTED: similar to Ubiquitin-conjugating enzyme E2 J2
48132665	LOC413253	230	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to Ubiquitin-conjugating enzyme E2, J2 [Apis mellifera].
66507431	LOC409794	265	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to ubiquitin-conjugating enzyme E2, J1 [Apis mellifera].
58388043	ENSANGP000000008285	127	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP00000010774 [Anopheles gambiae str. PEST].

8. UbcI

94495457	3110006E14Rik	296	Mus musculus	metazoa>vertebrata	PREDICTED: similar to CG4502-PA, isoform A [Mus musculus].
94408620	LOC245350	372	Mus musculus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E2Q (putative) 2
94387187	Ube2q2	364	Mus musculus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E2Q 2 [Mus
30725841	Ube2q2	378	Mus musculus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2Q 2 [Mus musculus].
51830637	LOC245350	387	Mus musculus	metazoa>vertebrata	PREDICTED: similar to Ubiquitin-conjugating enzyme E2Q (putative) 2
94396224	3110006E14Rik	304	Mus musculus	metazoa>vertebrata	PREDICTED: similar to CG4502-PA, isoform A [Mus musculus].
30425296	E330021D16Rik	371	Mus musculus	metazoa>vertebrata	hypothetical protein LOC243676 [Mus musculus].
31541872	Ube2q1	422	Mus musculus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2Q [Mus musculus].
115923264	LOC581772	341	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: similar to GA15522-PA [Strongylocentrotus purpuratus].
115966376	LOC591210	448	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: hypothetical protein [Strongylocentrotus purpuratus].
116058204	Ot06g01340	200	Ostreococcus tauri	viridiplantae	Predicted ubiquitin-conjugating enzyme (ISS) [Ostreococcus tauri].
70993182	AFUA_4G13940	1158	Aspergillus fum		

68392303	LOC557472	279	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to CG4502-PA, isoform A [Danio rerio].
114598936	LOC750981	291	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein [Pan troglodytes].
114658247	UBE2Q2	375	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2Q (putative) 2 isoform 4
114658255	UBE2Q2	357	Pan troglodytes	metazoa>vertebrata	PREDICTED: similar to Ubiquitin-conjugating enzyme E2Q (putative) 2 isoform 1
114658251	UBE2Q2	353	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2Q (putative) 2 isoform 1
114658249	UBE2Q2	340	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein isoform 5 [Pan troglodytes].
114559885	UBE2Q1	422	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2Q [Pan troglodytes].
71984336	ubc-25	387	Caenorhabditis elegans	metazoa>vertebrata	Ubiquitin Conjugating enzyme family member (ubc-25) [Caenorhabditis elegans].
Ngru1000014952	Ngru1000014952	330	Naegleria gruberi	heterolobosea	estExt fgeneshNG pg.C 610113
Ngru1000013764	Ngru1000013764	772	Naegleria gruberi	heterolobosea	estExt_fgeneshNG_pg.C_160052

9. AKTIP

```
**COUNTS**
parabasalidea      1
alveolata          2
alveolata>ciliophora 1
stramenopiles     3
fungi>basidiomycota 2
mycetozoa>dictyosteliida 1
metazoa           1
metazoa>choanoflagellida 1
metazoa>Chordata   1
metazoa>Cnidaria    3
metazoa>echinodermata 1
metazoa>hexapoda    5
metazoa>vertebrata  7
metazoa>vertebrata>actinopterygii 2
```

121917901	TVAG 476010	182	Trichomonas vaginalis G3	parabasalidea	Ubiquitin-conjugating enzyme family protein [Trichomonas vaginalis].
23509112	PFL2100W	465	Plasmodium falciparum 3D7	alveolata	hypothetical protein PFL2100W [Plasmodium falciparum 3D7].
Mbre1000000752	Mbre1000000752	386	Monosiga brevicollis	metazoa>Choanoflagellida	fgenesh1 pg.scaffold 7000053
114662500	FTS	332	Pan troglodytes	PREDICTED: fused toes homolog isoform 5 [Pan troglodytes].	
VIIb#5..m05100	VIIb#5..m05100	303	Toxoplasma gondii	alveolata	ubiquitin-conjugating enzyme domain-containing protein fgh est.C_scaffold 2500015
Nvec1000022094	Nvec1000022094	298	Nematostella vectensis	metazoa>Cnidaria	PREDICTED: fused toes homolog isoform 1 [Pan troglodytes].
114662502	FTS	293	Pan troglodytes	metazoa>vertebrata	UBiquitin Conjugating enzyme family member (ubc-19) [Caenorhabditis elegans].
17566194	ubc-19	292	Caenorhabditis elegans	metazoa	AKT interacting protein [Rattus norvegicus].
58865430	Fts	292	Rattus norvegicus	metazoa>vertebrata	AKT interacting protein [Homo sapiens].
61743933	AKTIP	292	Homo sapiens	metazoa>vertebrata	AKT interacting protein [Mus musculus].
6753918	Aktip	292	Mus musculus	metazoa>vertebrata	unnamed protein product [Tetraodon nigroviridis].
47216620	GSTEN:00025428:G:001	290	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	PREDICTED: similar to fused toes homolog [Pan troglodytes].
114600887	LOC736031	281	Pan troglodytes	metazoa>vertebrata	74424
Pram1000008760	Pram1000008760	281	Phytophthora ramorum	stramenopiles	PREDICTED: hypothetical protein [Strongylocentrotus purpuratus].
11592982	LOC578956	277	Strongylocentrotus purpuratus	metazoa>echinodermata	fgenesh1 pg.scaffold 1565000002
Nvec1000020327	Nvec1000020327	273	Nematostella vectensis	metazoa>Cnidaria	fgenesh1 pg.scaffold 6000043
Nvec1000015413	Nvec1000015413	272	Nematostella vectensis	metazoa>hexapoda	PREDICTED: similar to fused toes (predicted) [Tribolium CG16894-PA [Drosophila melanogaster].
91084089	LOC656942	267	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to fused toes [Apis mellifera].
24656026	CG16894	266	Drosophila melanogaster	metazoa>hexapoda	akt interacting protein [Danio rerio].
66510005	LOC552831	266	Apis mellifera	metazoa>hexapoda	fgenesh1 pg.C chr 3000244
Bf1o1000043060	Bf1o1000043060	264	Branchiostoma floridae	metazoa>Chordata	crossbrnx CG10536-PB, isoform B [Drosophila melanogaster].
41056191	aktip	261	Danio rerio	metazoa>vertebrata>actinopterygii	fgenesh1 pg.scaffold 1565000002
Ptri1000004528	Ptri1000004528	250	Phaeodactylum tricornutum	stramenopiles	hypothetical protein [Cryptococcus neoformans var. neoformans JEC21].
17933724	cbx	244	Drosophila melanogaster	metazoa>hexapoda	fungi>basidiomycota
58258487	CNA02640	242	Cryptococcus neoformans var. neoformans	metazoa>hexapoda	hypothetical protein [Ustilago maydis 521].
71003864	UM00451..1	240	Ustilago maydis 521	fungi>basidiomycota	unnamed protein product [Paramecium tetraurelia].
124429462	GSPATT00026577001	239	Paramecium tetraurelia	alveolata>ciliophora	PREDICTED: hypothetical protein isoform 6 [Pan troglodytes].
114662514	FTS	237	Pan troglodytes	metazoa>vertebrata	hypothetical protein DDBDRIFT_0204404 [Dictyostelium discoideum ENSANGP0000013744 [Anopheles gambiae str. PEST].
66816191	DDDBRAFT 0204404	232	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	ENSANGP0000013744 [Anopheles gambiae str. PEST].
58381786	ENSANGG00000011255	203	Anopheles gambiae str. PEST	metazoa>hexapoda	
Tpseu1000010538	Tpseu1000010538	161	Thalassiosira pseudonana	stramenopiles	165255

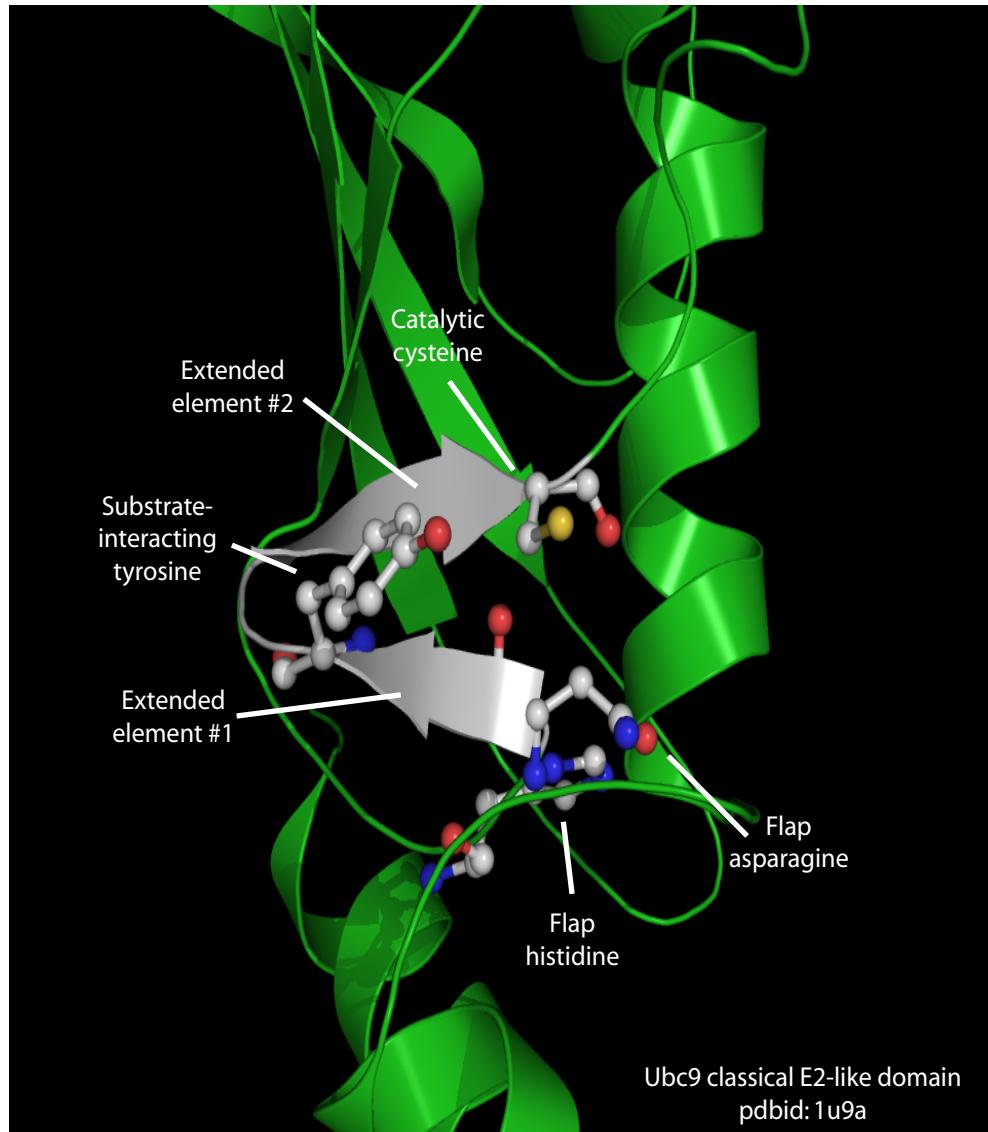
III. Phyletic distribution of E2 domains across eukaryotes with completely sequenced genomes

Lineage	**Asco	Afum	Cgla	Dhan	Gzea	Klac	Ncra	Scer	Spom	Ylip	**Basid	Cneo	Umay	**Micro	Ecun	**Met	Agam	Amel	Bfla	Cele	Cint	Drer	Dmel	Hsap	Mmus	Ptro	Rnor	Spur	Tnig	Tcas	Mbre	Nvec	**E&D	Ehis	Ddis	**Plants	Crei	Otau	Atha	**stram	Ptri	Psjo	Pram	Tpse	**Alv	Tthe	Plet	Tgon	Tpar	Tann	Cpar	Pfai	**Kinet	Tcru	Tbru	Lmaj	Ngru	Glam	Ttag	**Misc	Gthe	
all E2 domains	20	12	13	17	12	16	14	14	14	14	**Basid	19	16	**Micro	8	**Met	Agam	27	25	52	26	32	46	33	54	70	69	55	32	47	27	26	39	**E&D	Ehis	Ddis	**Plants	Crei	Otau	Atha	**stram	Ptri	Psjo	Pram	Tpse	**Alv	Tthe	Plet	Tgon	Tpar	Tann	Cpar	Pfai	**Kinet	Tcru	Tbru	Lmaj	Ngru	Glam	Ttag	**Misc	Gthe

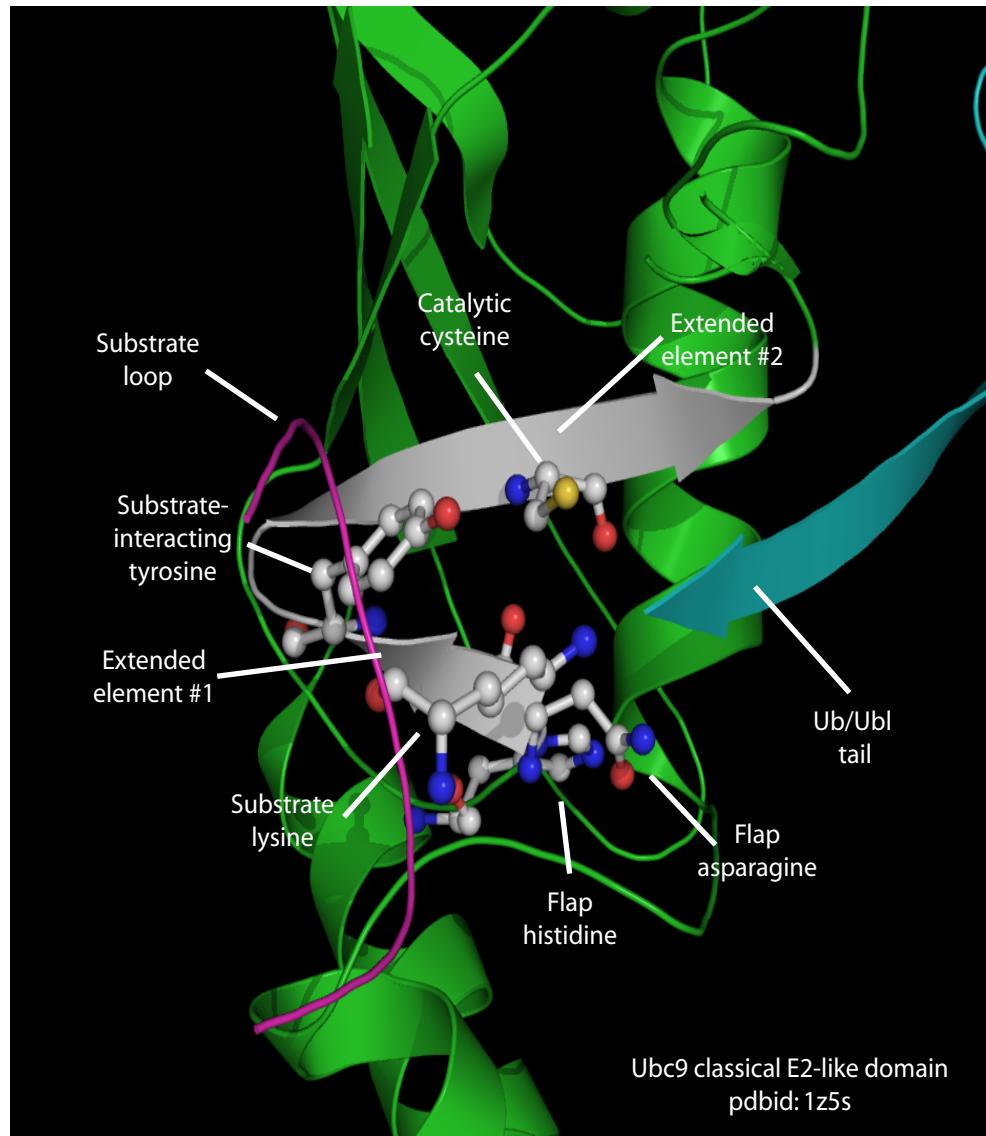
Abbreviations:
 **Asco: Ascomycetes
 Afum: Aspergillus fumigatus
 Cgla: Candida glabrata
 Dhan: Debaryomyces hansenii
 Gzea: Gibberella zaeae
 Klac: Kluyveromyces lactis
 Ncra: Neurospora crassa
 Scer: Saccharomyces cerevisiae
 Spom: Schizosaccharomyces pombe
 Ylip: Yarrowia lipolytica
 **Basid: Basidiomycetes
 Cneo: Cryptococcus neoformans
 Umay: Utilago maydis
 **Micro: Microsporidia
 Ecun: Encephalitozoon cuniculi
 **Met: Metazoa
 Agam: Anopheles gambiae
 Amel: Apis mellifera
 Bfla: Branchiostoma floridae
 Cele: Caenorhabditis elegans
 Cint: Ciona intestinalis
 Drer: Danio rerio
 Dmel: Drosophila melanogaster
 Hsap: Homo sapiens
 Mmus: Mus musculus
 Ptro: Pan troglodytes
 Rnor: Rattus norvegicus
 Spur: Strongylocentrotus purpuratus
 Tnig: Tetrahymena thermophila
 Tcas: Tribolium castaneum
 Mbire: Monosiga brevicollis
 Nvec: Nematostella vectensis
 **E&D: Amoebozoans
 Ehis: Entamoeba histolytica
 Ddis: Dictyostelium discoideum

**Plants:
 Crei: Chlamydomonas reinhardtii
 Otau: Ostreococcus tauri
 Atha: Arabidopsis thaliana
 **stram: Stramenopiles
 Ptri: Phaeodactylum tricornutum
 Psjo: Phytophthora sojae
 Pram: Phytophthora ramorum
 Tpse: Thalassiosira pseudonana
 **Alv: Alveolates
 Tthe: Tetrahymena thermophila
 Ptet: Paramecium tetraurelia
 Tgon: Toxoplasma gondii
 Tpar: Theileria parva
 Tann: Theileria annulata
 Cpar: Cryptosporidium parvum
 Pfal: Plasmodium falciparum
 **Kinet: Kinetoplastids and Heterolobosea
 Tcru: Trypanosoma cruzi
 Tbru: Trypanosoma brucei
 Lmaj: Leishmania major
 Ngru: Naegleria gruberi
 Glam: Giardia lamblia
 Ttag: Trichomonas vaginalis
 **Misc:
 Gthe: Guillardia theta

IV. Comparison of E2 flap architectures with and without Ub/Ubl. (Adapted from pdbid: 1u9a and 1z5s)



Ub/Ubl tail not present



Ub/Ubl tail present