

## **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.

Alignment of E2 domain families displaying significant structure and/or sequence divergence surrounding the active site. Sequences are aligned based on the structure of the E2 domain from the Schistosoma mansoni protein. Residues from the above secondary structure prediction program are given at the top of the alignment, and residue categories are given at various intervals below the alignment. Residues in grey boxes are conserved in all sequences. Residues in white boxes are conserved in at least one sequence.

Table of Contents

- 1. ACTP2
2. Tsp01 (EUB)
3. Tsp01 (EUB)
4. Aps10
5. Aps1
6. Dc1
7. Dc1
8. D3EKW
9. D3EKW
10. D3EKW

1. ACTP2

FINAL
12443462 Paramonix tetraurelia
5828487 Cryptosporidium parvum var. parvum...\_EJC1
12443462 Paramonix tetraurelia
5828487 Cryptosporidium parvum var. parvum...\_EJC1
12443462 Paramonix tetraurelia
5828487 Cryptosporidium parvum var. parvum...\_EJC1

2. Tsp01 (EUB)

FINAL
5247446 Mus musculus
6350454 Thelazia annulata
7103227 Thelazia parva strain NUG8
6746729 Euzooboa hiolayana NB 1108
6746729 Euzooboa hiolayana NB 1108
6746729 Euzooboa hiolayana NB 1108

3. R9D

FINAL
4934344 Drosophila melanogaster
9104441 Tribolium castaneum
6469979 Tribolium castaneum
9104441 Tribolium castaneum
6469979 Tribolium castaneum
9104441 Tribolium castaneum

4. Aps10

FINAL
1313370 Schistosoma boydii
1313370 Schistosoma boydii
1313370 Schistosoma boydii
1313370 Schistosoma boydii
1313370 Schistosoma boydii

5. Aps1

FINAL
1313370 Schistosoma boydii
1313370 Schistosoma boydii
1313370 Schistosoma boydii
1313370 Schistosoma boydii
1313370 Schistosoma boydii

Main alignment table showing amino acid sequences for various species across multiple columns. Includes species names like Paramonix tetraurelia, Cryptosporidium parvum, Mus musculus, Thelazia annulata, Tribolium castaneum, and Schistosoma boydii. Sequences are aligned based on structural and sequence divergence around the active site.





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: Tpseu  
Nematostella vectens: Nvec  
Branchiostoma floridae: Bflo

Table of Contents

1. Apg3
2. Apg10
3. Tsg101 (UEV)
4. BRUCE-like
5. UBE2W
6. Ufc1
7. Ubc6p
8. UbcI
9. AKTIP

1. Apg3

\*\*COUNTS\*\*

parabasalia	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	autophagocytosis associated protein, putative
66359970	cg8 2630	242	Cryptosporidium parvum Iowa II	alveolata	APG10/ Aut1p like like autophagocytosis protein involved in
71027271	TP03_0261	197	Theileria parva strain Muguga	alveolata	hypothetical protein TP03_0261 [Theileria parva strain Muguga].
23613578	PF0280c	298	Plasmodium falciparum 3D7	alveolata	hypothetical protein [Plasmodium falciparum 3D7].
124396640	GSPATT00032381001	229	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
89292028	THERM_01050620	280	Tetrahymena thermophila SB210	alveolata>ciliophora	Autophagocytosis 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].
8290045	THERM_00012980	225	Tetrahymena thermophila SB210	alveolata>ciliophora	Autophagocytosis associated protein, C-terminal domain containing
56471771	46_t00033	268	Entamoeba histolytica HM-1,IMSS	entamoebidae	autophagocytosis protein, putative [Entamoeba histolytica
71421296	Tc00_1047053510257.90	270	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	autophagocytosis associated protein [Trypanosoma cruzi strain CL
68128829	LmjF33_0295	275	Leishmania major	euglenozoa>kinetoplastida	autophagocytosis protein, putative [Leishmania major].
84043456	Tb927.2.1890	292	Trypanosoma brucei TREU927	euglenozoa>kinetoplastida	autophagocytosis associated protein [Trypanosoma brucei TREU927].
71649849	Tc00_1047053509213.130	270	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	autophagocytosis associated protein [Trypanosoma cruzi strain CL
50420719	DEHA0D10846g	324	Debaryomyces hansenii CBS767	fungi>ascomycota	hypothetical protein DEHA0D10846g [Debaryomyces 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].
85120165	NCU01955.1	352	Neurospora crassa OR74A	fungi>ascomycota	hypothetical protein [Neurospora crassa OR74A].
70998182	AFUA_5G08170	353	Aspergillus fumigatus Af293	fungi>ascomycota	autophagocytosis protein Aut1, putative [Aspergillus fumigatus
50553884	YARL10E24453g	366	Yarrowia lipolytica CLIB122	fungi>ascomycota	hypothetical protein [Yarrowia lipolytica].
19113456	SPBC3B9_06c	275	Schizosaccharomyces pombe 972h-	fungi>ascomycota	hypothetical protein SPBC3B9_06c [Schizosaccharomyces pombe 972h-].
45198343	AGOS AFL178W	282	Ashbya gossypii ATCC 10895	fungi>ascomycota	AFL178Wp [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	fgenesH 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
ci0100131803	ci0100131803	254	Ciona intestinalis	metazoa>Chordata	ci0100131803
Bflo1000027008	Bflo1000027008	1023	Branchiostoma floridae	metazoa>Chordata	estExt fgenesH2 pg.C 400088
Bflo1000029400	Bflo1000029400	318	Branchiostoma floridae	metazoa>Chordata	estExt fgenesH2 pg.C 1590012
ci0100142936	ci0100142936	312	Ciona intestinalis	metazoa>Chordata	ci0100146633
Nvec1000006208	Nvec1000006208	329	Nematostella vectensis	metazoa>Cnidaria	e gw.31.166.1
115956521	LOC594526	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 ENSANGG0000012603	331	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP0000015092 [Anopheles gambiae str. PEST].
66564768	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	Apg3p [Homo sapiens].
114588464	ATG3	318	Pan troglodytes	metazoa>vertebrata	PREDICTED: Apg3p isoform 2 [Pan troglodytes].
19705511	Apg3l	314	Rattus norvegicus	metazoa>vertebrata	APG3 autophagy 3-like [Rattus norvegicus].
114588462	ATG3	314	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein isoform 1 [Pan troglodytes].
13385890	Atg3	314	Mus musculus	metazoa>vertebrata	autophagy Apg3p/Aut1p-like [Mus musculus].
41053345	apg3l	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].
66818030	DBDRAFT_0217934	343	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DBDRAFT_0217934 [Dictyostelium discoideum
28828597	-	338	Dictyostelium discoideum	mycetozoa>dictyosteliida	hypothetical protein [Dictyostelium discoideum].
121911929	TVAG 447140	258	Trichomonas vaginalis G3	parabasalia	autophagocytosis associated protein, putative [Trichomonas
Tpseu1000007918	Tpseu1000007918	315	Thalassiosira pseudonana	stramenopiles	145968
Psoj1000017156	Psoj1000017156	337	Phytophthora sojae	stramenopiles	110627
Ptri1000002031	Ptri1000002031	303	Phaeodactylum tricornutum	stramenopiles	fgenesH1 pm.c_chr_1000016
Pram1000014227	Pram1000014227	323	Phytophthora ramorum	stramenopiles	47287
18424474	ATS661500	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 autophagocytosis during starvation (ISS)

2. Apg10

\*\*COUNTS\*\*

parabasalia	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

Psoj1000006153	Psoj1000006153	412	Phytophthora sojae	stramenopiles	132761
Ptri1000008763	Ptri1000008763	411	Phaeodactylum tricornutum	stramenopiles	estExt Genewis1.C chr 140041
Ptri1000004948	Ptri1000004948	421	Phaeodactylum tricornutum	stramenopiles	estExt fgeneshl_pg.C_chr_230106
Pram1000004476	Pram1000004476	444	Phytophthora ramorum	stramenopiles	79629
Tpseu1000001553	Tpseu1000001553	400	Thalassiosira pseudonana	stramenopiles	103937
116056391	OT0360280	438	Ostreococcus tauri	viridiplantae	putative human tumor susceptibility gene-like protein (ISS)
18395936	ATELC/ELC	398	Arabidopsis thaliana	viridiplantae	ATELC/ELC; ubiquitin binding [Arabidopsis thaliana].
Crei1000005374	Crei1000005374	525	Chlamydomonas reinhardtii	viridiplantae	Chlre2 kg.scaffold 16000223
15240732	ATELC-LIKE/ELC-LIKE	368	Arabidopsis thaliana	viridiplantae	ATELC-LIKE/ELC-LIKE [Arabidopsis thaliana].
42569747	AT2G38830	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>Cnidaria		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	GSPATT00035572001	2472	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
89285001	THERM_01027520	512	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena
89289979	THERM_00011330	464	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena
89299207	THERM_00481290	2633	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena
89303321	THERM_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	GSPATT00034627001	2123	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
89288400	THERM_00028560	3915	Tetrahymena thermophila SB210	alveolata>ciliophora	ubiquitin-activating enzyme E1 family protein [Tetrahymena
124409420	GSPATT00010919001	1850	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124398217	GSPATT0003569001	2396	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
124412336	GSPATT00012991001	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 zeae PH-1	fungi>ascomycota	hypothetical protein FG01207.1 [Gibberella zeae PH-1].
58261412	CNF05830	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
Bflo1000002074	Bflo1000002074	527	Branchiostoma floridae	metazoa>Chordata	fgenes2 pg.scaffold_8000077
ci0100131458	ci0100131458	1058	Ciona intestinalis	metazoa>Chordata	ci0100131458
Bflo1000037265	Bflo1000037265	301	Branchiostoma floridae	metazoa>Chordata	e gw.8.290.1
Bflo1000027761	Bflo1000027761	1307	Branchiostoma floridae	metazoa>Chordata	estExt fgenes2 pg.C_760117
Bflo1000048584	Bflo1000048584	307	Branchiostoma floridae	metazoa>Chordata	estExt gwp.C_3740030
Bflo1000011222	Bflo1000011222	981	Branchiostoma floridae	metazoa>Chordata	fgenes2 pg.scaffold_138000092
Bflo1000025135	Bflo1000025135	307	Branchiostoma floridae	metazoa>Chordata	estExt fgenes2 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	fgenes1 pg.scaffold_10000074
Nvec1000005680	Nvec1000005680	659	Nematostella vectensis	metazoa>Cnidaria	e gw.18.24.1
115964468	LOC580131	234	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: hypothetical protein [Strongylocentrotus purpuratus].
91089323	LOC56712	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	ENSANGG00000017845	1220	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP00000020334 [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
315455	Ube22	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	Ube20 predicted	5017	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to baculoviral IAP repeat-containing 6 [Rattus
50234896	Ube20	1288	Mus musculus	metazoa>vertebrata	ubiquitin-conjugating enzyme E20 [Mus musculus].
33636750	UBE20	1292	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E20 [Homo sapiens].
12751495	UBE2Z	246	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2Z (putative) [Homo sapiens].
109492276	Ube20 predicted	1312	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E20 [Rattus
109489413	Ube20 predicted	1315	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E20 [Rattus
83320076	Ube2z	246	Rattus norvegicus	metazoa>vertebrata	ubiquitin-conjugating enzyme E2Z (putative) [Rattus norvegicus].
114670605	UBE20	1392	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E20 [Pan troglodytes].
47213169	GSTEN:00010000:G:001	2135	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
68437089	LOC56895	354	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to Baculoviral IAP repeat-containing protein 6
50539720	zgc:92419	248	Danio rerio	metazoa>vertebrata>actinopterygii	PREDICTED: similar to RIKEN cDNA B230113M03 [Danio rerio].
47213168	GSTEN:00009999:G:001	3047	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	hypothetical protein LOC436602 [Danio rerio].
47214958	GSTEN:00032514:G:001	1309	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
47217004	GSTEN:00020460:G:001	403	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
28830048	-	2208	Dictyostelium discoideum	mycetozoa>dictyosteliida	unnamed protein product [Tetraodon nigroviridis].
60472388	DBDRAFT_0217536	910	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	similar to Homo sapiens (Human). Hypothetical protein FLJ13786 (EC
Ptri1000008184	Ptri1000008184	415	Phaeodactylum tricornutum	stramenopiles	hypothetical protein DBDRAFT_0217536 [Dictyostelium discoideum
15219165	UBC26	1163	Arabidopsis thaliana	viridiplantae	estExt fgenes1 pg.C chr 190218
Crei10000012890	Crei10000012890	222	Chlamydomonas reinhardtii	viridiplantae	UBC26 (ubiquitin-conjugating enzyme 25); ubiquitin-protein ligase
30685832	PHO2/UBC24	907	Arabidopsis thaliana	viridiplantae	estExt fgenes2 pg.C_110201
15227320	UBC23	1102	Arabidopsis thaliana	viridiplantae	PHO2/UBC24 (PHOSPHATE 2); ubiquitin-protein ligase [Arabidopsis
15229481	UBC25	609	Arabidopsis thaliana	viridiplantae	UBC23 (ubiquitin-conjugating enzyme 22); ubiquitin-protein ligase
					UBC25 (ubiquitin-conjugating enzyme 23); small conjugating protein

## 5. UBE2W

**COUNTS**					
heterolobosea		1			
alveolata		2			
alveolata>ciliophora		5			
viridiplantae		5			
fungi>ascomycota		3			
fungi>basidiomycota		1			
metazoa		1			
metazoa>Chordata		3			
metazoa>Cnidaria		1			
metazoa>echinodermata		1			
metazoa>hexapoda		4			
metazoa>vertebrata		7			
metazoa>vertebrata>actinopterygii		3			
mycetozoa>dictyosteliida		1			
23509349	PF14_0128	299	Plasmodium falciparum 3D7	alveolata	ubiquitin conjugating enzyme, putative [Plasmodium falciparum 3D7].
VIib#55.m04812	VIib#55.m04812	287	Toxoplasma gondii	alveolata	ubiquitin-conjugating enzyme domain-containing protein
89300310	THERM_00285290	256	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena
94364832	Ube2w	180	Mus musculus	metazoa>vertebrata	PREDICTED: similar to CG7220-PA, isoform A isoform 5 [Mus
50557414	YALIOF1955g	178	Yarrowia lipolytica CLIB122	fungi>ascomycota	hypothetical protein [Yarrowia lipolytica].
82795563	Ube2w	178	Mus musculus	metazoa>vertebrata	PREDICTED: similar to CG7220-PA, isoform A isoform 11 [Mus
114620512	UBE2W	162	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2W (putative) isoform 1
47933381	UBE2W	162	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2W (putative) isoform 1 [Homo
18401461	ATUBC2-1	161	Arabidopsis thaliana	viridiplantae	ATUBC2-1 (Arabidopsis thaliana ubiquitin-conjugating enzyme);
18410856	UBC16	161	Arabidopsis thaliana	viridiplantae	UBC16 (ubiquitin-conjugating enzyme 15); ubiquitin-protein ligase
18419831	UBC17	161	Arabidopsis thaliana	viridiplantae	UBC17 (UBIQUITIN-CONJUGATING ENZYME 17); ubiquitin-protein ligase
18422281	UBC18	161	Arabidopsis thaliana	viridiplantae	UBC18 (ubiquitin-conjugating enzyme 16); ubiquitin-protein ligase
Crei1000002793	Crei1000002793	161	Chlamydomonas reinhardtii	viridiplantae	e gwW.24.158.1
46124999	FG06877.1	157	Gibberella zeae PH-1	fungi>ascomycota	hypothetical protein FG06877.1 [Gibberella zeae PH-1].
71002042	AFUA_2G13690	158	Aspergillus fumigatus Af293	fungi>ascomycota	ubiquitin conjugating enzyme, putative [Aspergillus fumigatus
28573347	CG7220	154	Drosophila melanogaster	metazoa>hexapoda	CG7220-PA, isoform A [Drosophila melanogaster].
58387116	ENSANGG00000018586	154	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP00000024002 [Anopheles gambiae str. PEST].
66549055	LOC409100	154	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to CG7220-PA, isoform A [Apis mellifera].
91079008	LOC663661	154	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to CG7220-PA, isoform A [Tribolium castaneum].
124416285	GSPATT00016175001	152	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].



supplementary\_material\_E2\_gis.txt

17510293	ubc-16	152	Caenorhabditis elegans	metazoa	Ubiquitin Conjugating enzyme family member (ubc-16) [Caenorhabditis]
46099348	UM03443.1	152	Ustilago maydis 521	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	flj110111	151	Danio rerio	metazoa>vertebrata>actinopterygii	hypothetical protein LOC402907 [Danio rerio].
4793385	UBE2W	151	Homo sapiens	metazoa>vertebrata	ubiquitin-conjugating enzyme E2W (putative) isoform 2 [Homo]
Bf1o1000025168	Bf1o1000025168	152	Branchiostoma floridae	metazoa>Chordata	estExt_fgfnsh2.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	DDBDRAFT_0202295	149	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DDBDRAFT_0202295 [Dictyostelium discoideum]
124399351	GSPATT00034321001	148	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
Ngru1000001737	Ngru1000001737	144	Naegleria gruberi	heterolobosea	fgfnshHS.pm.scaffold_72000004
47223533	GSTEN:00015564:G:001	148	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
ci0100140113	ci0100140113	145	Ciona intestinalis	metazoa>Chordata	ci0100140113
Bf1o1000032579	Bf1o1000032579	144	Branchiostoma floridae	metazoa>Chordata	estExt_fgfnsh2.pg.C_5140030
68356096	LOC586071	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	THERM_01108500	180	Tetrahymena thermophila SB210	alveolata>ciliophora	hypothetical protein THERM_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
17525292	C40H1_6	162	Caenorhabditis elegans	metazoa	C40H1.6 [Caenorhabditis elegans].
Mbre1000002981	Mbre1000002981	167	Monosiga brevicollis	metazoa>Choanoflagellida	fgfnsh2.pm.scaffold_3000061
ci0100146879	ci0100146879	166	Ciona intestinalis	metazoa>Chordata	ci0100146879
Bf1o1000011222	Bf1o1000011222	981	Branchiostoma floridae	metazoa>Chordata	fgfnsh2.pg.scaffold_138000092
Bf1o1000025213	Bf1o1000025213	170	Branchiostoma floridae	metazoa>Chordata	estExt_fgfnsh2.pm.C_3610002
Nvec1000018661	Nvec1000018661	170	Nematostella vectensis	metazoa>Cnidaria	fgfnsh1.pg.scaffold_199000025
Nvec1000012015	Nvec1000012015	116	Nematostella vectensis	metazoa>Cnidaria	e.gw.1650.2.1
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	ENSANGP00000008034	164	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGP00000010523 [Anopheles gambiae str. PEST].
66512255	LOC581106	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].
7705481	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	ufc1	164	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	ubiquitin-like (UBL) post-translational modifier [Dictyostelium
Pram1000015146	Pram1000015146	158	Phytophthora ramorum	stramenopiles	53495
Crei1000013700	Crei1000013700	167	Chlamydomonas reinhardtii	viridiplantae	estExt_fgfnsh2.pg.C_240032
18396376	ATIG27530	174	Arabidopsis thaliana	viridiplantae	unknown protein [Arabidopsis thaliana].

## 7. Ubc6p

\*\*COUNTS\*\*

diplomonadida group>diplomonadida group	3
parabasaliida	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	YALI0E30173g	132	Yarrowia lipolytica CLIB122	fungi>ascomycota	hypothetical protein [Yarrowia lipolytica].
50550481	YALI0D11748g	252	Yarrowia lipolytica CLIB122	fungi>ascomycota	hypothetical protein [Yarrowia lipolytica].
46101486	UM05955.1	223	Ustilago maydis 521	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
71655980	Tc00_1047053508387.100	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 E2 [Trypanosoma cruzi strain CL
71665910	Tc00_1047053506789.200	410	Trypanosoma cruzi strain CL Brener	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme [Trypanosoma cruzi strain CL Brener].
62359387	Tb927_4_3190	220	Trypanosoma brucei	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme E2, putative [Trypanosoma brucei].
70802852	-	310	Trypanosoma brucei	euglenozoa>kinetoplastida	ubiquitin-conjugating enzyme, putative [Trypanosoma brucei].
121894423	TVAG_091080	352	Trichomonas vaginalis G3	parabasaliida	Ubiquitin-conjugating enzyme family protein [Trichomonas vaginalis
121911847	TVAG_066800	344	Trichomonas vaginalis G3	parabasaliida	Ubiquitin-conjugating enzyme family protein [Trichomonas vaginalis
91082969	LOC662641	225	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to ubiquitin conjugating enzyme E2, J2 isoform 2
91079666	LOC655113	268	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to ubiquitin-conjugating enzyme E2, J1
Ib#25.m01819	Ib#25.m01819	676	Toxoplasma gondii	alveolata	hypothetical protein
XII#65.m01972	XII#65.m01972	202	Toxoplasma gondii	alveolata	ubiquitin-conjugating enzyme E2, putative
XII#50.m03405	XII#50.m03405	207	Toxoplasma gondii	alveolata	ubiquitin-conjugating enzyme E2, putative
Tpseu1000010010	Tpseu1000010010	175	Thalassiosira pseudonana	stramenopiles	160728
Tpseu1000002627	Tpseu1000002627	318	Thalassiosira pseudonana	stramenopiles	109475
Tpseu1000004117	Tpseu1000004117	183	Thalassiosira pseudonana	stramenopiles	119221
47212316	GSTEN:00003577:G:001	185	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
47190747	GSTEN:00037077:G:001	81	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
89287327	THERM_00470910	443	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena
89303616	THERM_00931940	221	Tetrahymena thermophila SB210	alveolata>ciliophora	Ubiquitin-conjugating enzyme family protein [Tetrahymena
115969448	LOC9010	249	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: similar to LOC495424 protein [Strongylocentrotus
72159165	LOC587822	370	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: similar to ubiquitin-conjugating enzyme E2, J1 (UBCG
19114168	SPAC10F6.05c	227	Schizosaccharomyces pombe 972h-	fungi>ascomycota	hypothetical protein SPAC10F6.05c [Schizosaccharomyces pombe
6320947	UBC6	250	Saccharomyces cerevisiae	fungi>ascomycota	Ubc6p [Saccharomyces cerevisiae].
56090425	MGC94113	259	Rattus norvegicus	metazoa>vertebrata	similar to Ubc6p homolog [Rattus norvegicus].
109474783	Ube2j1 predicted	386	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E2, J1 [Rattus
109476293	Ube2j1 predicted	426	Rattus norvegicus	metazoa>vertebrata	PREDICTED: similar to ubiquitin-conjugating enzyme E2, J1 [Rattus
Psoj1000015568	Psoj1000015568	249	Phytophthora sojae	stramenopiles	143578
Psoj1000000332	Psoj1000000332	214	Phytophthora sojae	stramenopiles	108600
Pram1000004878	Pram1000004878	332	Phytophthora ramorum	stramenopiles	79160
Pram1000009034	Pram1000009034	185	Phytophthora ramorum	stramenopiles	74102
Pram1000009035	Pram1000009035	225	Phytophthora ramorum	stramenopiles	74101



supplementary\_material\_E2\_gis.txt

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
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].
114559985	UBE2Q1	422	Pan troglodytes	metazoa>vertebrata	PREDICTED: ubiquitin-conjugating enzyme E2Q [Pan troglodytes].
71984336	ubc-25	387	Caenorhabditis elegans	metazoa	Ubiquitin Conjugating enzyme family member (ubc-25) [Caenorhabditis
Ngru1000014952	Ngru1000014952	330	Naegleria gruberi	heterolobosea	estExt fgenshNG_pg.C 610113
Ngru1000013764	Ngru1000013764	772	Naegleria gruberi	heterolobosea	estExt_fgenshNG_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

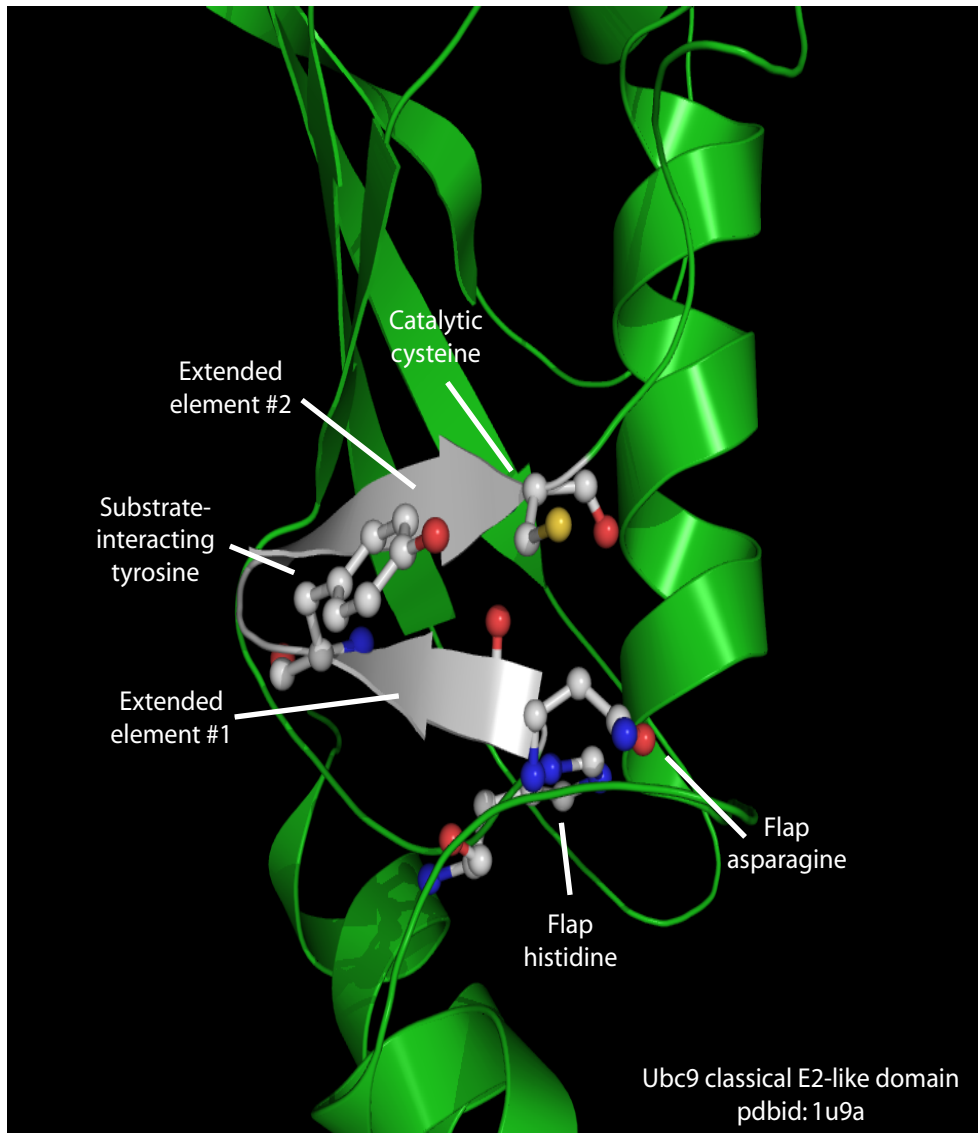
121917901	TVAG 476010	182	Trichomonas vaginalis G3	parabasalidea	Ubiquitin-conjugating enzyme family protein [Trichomonas vaginalis
23509112	PFL2100w	465	Plasmodium falciparum 3D7	alveolata	hypothetical protein PFL2100w [Plasmodium falciparum 3D7].
Mbre1000000752	Mbre1000000752	386	Monosiga brevicollis	metazoa>Choanoflagellida	fgenshl pg.scaffold 7000053
114662500	FTS	332	Pan troglodytes	metazoa>vertebrata	PREDICTED: fused toes homolog isoform 5 [Pan troglodytes].
VIb#55.m05100	VIb#55.m05100	303	Toxoplasma gondii	alveolata	ubiquitin-conjugating enzyme domain-containing protein
Nvec1000022094	Nvec1000022094	298	Nematostella vectensis	metazoa>Cnidaria	fgsh est.C_scaffold 25000015
114662502	FTS	293	Pan troglodytes	metazoa>vertebrata	PREDICTED: fused toes homolog isoform 1 [Pan troglodytes].
17566194	ubc-19	292	Caenorhabditis elegans	metazoa	Ubiquitin Conjugating enzyme family member (ubc-19) [Caenorhabditis
58865430	Fts	292	Rattus norvegicus	metazoa>vertebrata	AKT interacting protein [Rattus norvegicus].
61743933	AKTIP	292	Homo sapiens	metazoa>vertebrata	AKT interacting protein [Homo sapiens].
6753918	Aktip	292	Mus musculus	metazoa>vertebrata	AKT interacting protein [Mus musculus].
47216620	GSTEN:00025428:G:001	290	Tetraodon nigroviridis	metazoa>vertebrata>actinopterygii	unnamed protein product [Tetraodon nigroviridis].
114600887	LOC736031	281	Pan troglodytes	metazoa>vertebrata	PREDICTED: similar to fused toes homolog [Pan troglodytes].
Pram10000008760	Pram10000008760	281	Phytophthora ramorum	stramenopiles	74424
115929982	LOC578956	277	Strongylocentrotus purpuratus	metazoa>echinodermata	PREDICTED: hypothetical protein [Strongylocentrotus purpuratus].
Nvec1000020327	Nvec1000020327	273	Nematostella vectensis	metazoa>Cnidaria	fgenshl pg.scaffold 156500002
Nvec1000015413	Nvec1000015413	272	Nematostella vectensis	metazoa>Cnidaria	fgenshl pg.scaffold 6000043
91084089	LOC656942	267	Tribolium castaneum	metazoa>hexapoda	PREDICTED: similar to fused toes (predicted) [Tribolium
24656026	CG16894	266	Drosophila melanogaster	metazoa>hexapoda	CG16894-PA [Drosophila melanogaster].
66510005	LOC552831	266	Apis mellifera	metazoa>hexapoda	PREDICTED: similar to fused toes [Apis mellifera].
Bflo1000043060	Bflo1000043060	264	Branchiostoma floridae	metazoa>Chordata	e gw.431.37.1
41056191	aktip	261	Danio rerio	metazoa>vertebrata>actinopterygii	akt interacting protein [Danio rerio].
Ptri1000004528	Ptri1000004528	250	Phaeodactylum tricorntum	stramenopiles	fgenshl pg.C chr 3000244
17933724	cbx	244	Drosophila melanogaster	metazoa>hexapoda	crossbronx CG10536-PB, isoform B [Drosophila melanogaster].
58258487	CNA02640	242	Cryptococcus neoformans var. neoformans JEC21	fungi>basidiomycota	hypothetical protein [Cryptococcus neoformans var. neoformans
71003864	UM00451.1	240	Ustilago maydis 521	fungi>basidiomycota	hypothetical protein UM00451.1 [Ustilago maydis 521].
124429462	GSPATT00026577001	239	Paramecium tetraurelia	alveolata>ciliophora	unnamed protein product [Paramecium tetraurelia].
114662514	FTS	237	Pan troglodytes	metazoa>vertebrata	PREDICTED: hypothetical protein isoform 6 [Pan troglodytes].
68816191	DDBDRAFT_0204404	232	Dictyostelium discoideum AX4	mycetozoa>dictyosteliida	hypothetical protein DDBDRAFT_0204404 [Dictyostelium discoideum
58381786	ENSANGS00000011255	203	Anopheles gambiae str. PEST	metazoa>hexapoda	ENSANGS00000011255 [Anopheles gambiae str. PEST].
Tpseu1000010538	Tpseu1000010538	161	Thalassiosira pseudonana	stramenopiles	165255

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

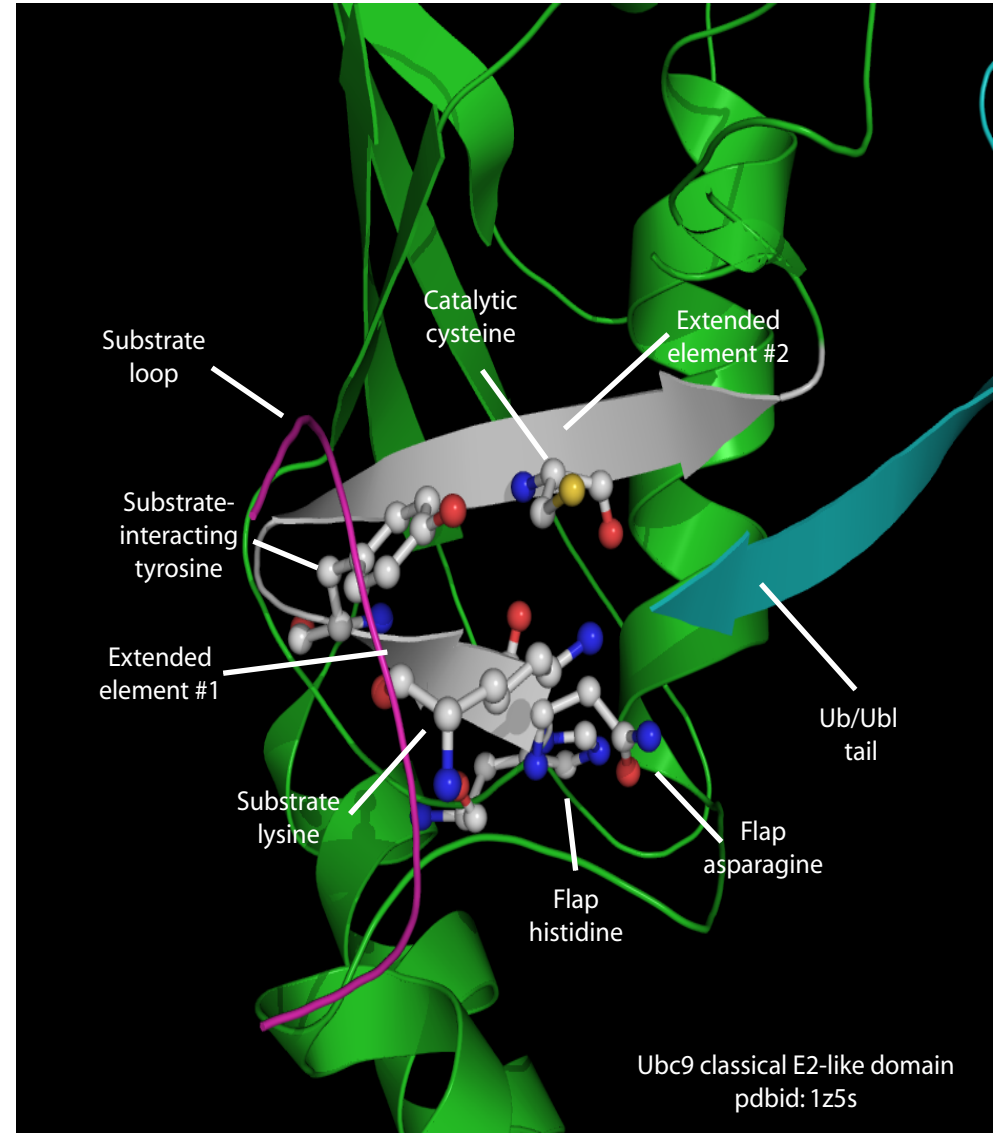
Lineage	**Asco	Afum	Cgla	Dhan	Gzea	Klac	Ncra	Scer	Spom	Ylip	**Basid	Cneo	Umay	**Micro	Ecun	**Met	Agam	Amel	Bflo	Cele	Cint	Drer	Dmel	Hsap	Mmus	Ptro	Rnor	Spur	Tnig	Tcas	Mbre	Nvec	**E&D	Ehis	Ddis	**Plants	Crei	Otau	Atha	**stram	Ptri	Psoj	Pram	TPse	**Alv	Tthe	Ptet	Tgon	Tpar	Tann	Cpar	Pfal	**Kinet	Tcru	Tbru	Lmaj	Ngru	Glam	Tvag	**Misc	Gthe
all E2 domains		20	12	13	17	12	16	14	14	14		19	16		8		27	25	52	26	32	46	33	54	70	69	55	32	47	27	26	39		17	29		22	13	53		18	19	19	19		40	71	17	11	12	8	14		23	13	15	33	11	42		3

- Abbreviations:
- \*\*Asco: Ascomycetes
  - Afum: Aspergillus fumigatus
  - Cgla: Candida glabrata
  - Dhan: Debaryomyces hansenii
  - Gzea: Gibberella zeae
  - Klac: Kluyveromyces lactis
  - Ncra: Neurospora crassa
  - Scer: Saccharomyces cerevisiae
  - Spom: Schizosaccharomyces pombe
  - Ylip: Yarrowia lipolytica
  - \*\*Basid: Basidiomycetes
  - Cneo: Cryptococcus neoformans
  - Umay: Ustilago maydis
  - \*\*Micro: Microsporidia
  - Ecun: Encephalitozoon cuniculi
  - \*\*Met: Metazoa
  - Agam: Anopheles gambiae
  - Amel: Apis mellifera
  - Bflo: 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: Tetraodon nigroviridis
  - Tcas: Tribolium castaneum
  - Mbre: 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
  - Psoj: 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
  - Tvag: 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