eISSN: 09748369, www.biolmedonline.com

Phylogeny reconstruction of ubiquitin conjugating (E2) enzymes

Muhummadh Khan [2], *Kaiser Jamil [1, 2]

[1] School of Life Sciences, Centre for Biotechnology and Bioinformatics, Jawaharlal National Institute of Advanced studies (JNIAS), Budha Bhavan, 6th floor, Secunderabad - 500003, A.P., India.
[2] Genetics Department, Bhagwan Mahavir Medical Research Centre, 10-1-1, Mahavir Marg, Hyderabad-500004, A.P. India.

***Corresponding Author:** kaiser.jamil@gmail.com

Abstract

Phylogeny reconstruction has attracted much attention from biologists and computer scientists from gene-order data over the last few years. Ubiquitin is widespread in eukaryotes and plays a pivotal role in selective ATP dependent protein degradation in cells. It is also highly expressed in several types of cancers. Hence, in order to understand its diversification in eukarya we undertook this study to determine its phylogeny using various computational methods. All sequences of Ubiquitin conjugating enzyme were extracted from protein data bank (PDB), and using Maximum Likelihood method as implemented in PHYML and the unrooted phylogenetic tree was constructed. This tree had four major clusters and one mini cluster. We calculated the divergence for the clusters of the gene from the site-specific change of functional importance in protein sequence evolution. The divergence was then mapped on the 3-D structure of the ube2c protein obtained from the PDB Data bank: (IL7K) using RASMOL. From this study, we conclude that we have been able to point out the exact sites where active evolution of E2s is taking place and it is apparent that these sites are subject to a strong purifying selection. We anticipate that this information would have some useful implications in neoplasia as there are reports that mutations in this protein are likely to promote tumour progression

Keywords: Ubiquitin Conjugating Enzymes; E2; Ube2s; Phylogeny; Divergence; Protein Structure Evolution; PHYML.

Introduction

The phylogenetic analysis of the genes generally determines the evolutionary patterns, conservation, lateral and horizontal distribution of the genes in various species. Universal presence of ubiquitin conjugating enzymes in eukaryotes, its major role in ubiquitination of unwanted proteins in cells, makes it interesting to study its phylogeny. More so because of the importance of ubiquitin conjugating enzymes that are known to be upregulated in cancers, hence this study assumes importance as the knowledge of the diversity and phylogenetic distribution of ube2c, ube2s and their homologs is helpful in understanding their role in ubiquitination in diseases like cancer. Our earlier studies on computation analysis included Homology modelling and Affinity modelling programs on cell membrane receptors like the epidermal growth factor receptor (EGFR) and MEXR efflux pump in humans which were evaluated for drug response studies (Sabitha and Jamil 2006, Sabitha et al. 2007, Suman et al 2006). Further, in another study we proposed the homology model of crystal protein of the microbe *Bacillus thuringiensis var israelensis* (Jamil et al 2007).

Etlinger and Goldberg (1977) discovered the Ubiquitin – Proteosome system, which is an ATP dependent process for protein degradation. Goldstein et al (1975) originally isolated ubiquitin and reported its presence in all tissues. Ubiquitin is a well-conserved, small protein of 76 Amino Acid residues ubiquitously present in all eukaryotic cells. Aaron Clechanover, Avram Hershko and Irwin Rose, who won the Nobel Prize in Chemistry in 2004, described proteins tagged with ubiquitin as “Proteins labelled for destruction”. The ubiquitin-proteosome system is reported to be responsible for the degradation of most short-lived proteins and Ubiquitination is known to occur in various cellular processes other than protein degradation. These include cell division, DNA repair, and quality control of newly produced proteins, and immune defence (Hershko 2005).

Ubiquitin is conjugated to the target protein via a cascade of enzymatic reactions.

Three classes of enzyme are involved in ubiquitination: Ubiquitin-activating enzymes (E1), Ubiquitin-conjugating enzymes (E2) and Ubiquitin-protein ligases (E3). The first step in this cascade is ATP-dependent activation of ubiquitin by an ubiquitin activating enzyme (E1). E1 adenylates the C terminus of Ubiquitin and then forms thiol esters with ubiquitin conjugating enzymes (E2s) that act as mobile carriers of activated Ubiquitin. Ubiquitin ligating enzymes (E3s) are responsible for the substrate specificity and contain either a RING or HECT domain. E3s tag the ubiquitin to the ϵ-amino group of a lysine residue on the target protein through the recruitment of both an E2 thiol ester and a specific substrate (Robinson and Ardley 2004).

Out of all the different enzymes involved in this pathway, two ubiquitin conjugating enzymes (E2), ube2c and ube2s, are highly expressed in different cancers and are listed in the cancer metasignature (Rhodes et al 2004). The expression of these two genes is low in normal tissues when compared to the expression of other ubiquitin conjugating enzymes (E2s). In cancer cells, their expression shoots up; whereas the level of other E2s drop. Ube2c is highly expressed in various human primary tumors and is reported to have the ability to promote cell growth and malignant transformation (Okamoto et al 2003).

Recent work provides evidence that various E2 proteins play a regulatory role in the cell cycle progression. It has been reported that ube2c is required for APC-dependent ubiquitination of mitotic cyclins (King et al. 1995, Aristarkhov et al. 1996, Yu et al. 1996). Ube2c has been identified as a human homologue of the cyclin-selective E2 (E2-C) that is required for the destruction of mitotic cyclins (Aristarkhov et al. 1996, Townsley et al. 1997, Yu et al. 1996). The enforced overproduction of Ube2c might disrupt the auto regulatory feedback loop and there by lead to deregulated cell growth (Arvand et al. 1998, Fang et al. 1998, Kramer et al 1998, Yamanaka et al. 2000). In addition, the E2s are also required for the destruction of mitotic cyclins and for cell cycle progression (Pagano et al. 1995). Survey of gene expression in a panel of 174 human epithelial tumours shows its presence in both primary and metastatic squamous and adenocarcinomas of prostate, bladder, breast, colon, gastro oesophagus, kidney, liver, ovary, pancreas, and lung (Su et al 2001).

In the present study, the phylogenetic tree of the ubiquitin conjugating enzymes was constructed in order to understand their diversity in eukarya as a part of its evolutionary history and its distribution in various species. The functional divergence was calculated to observe the functional constraints that are site specific. Functionally important sites and regions of biological sequences are under strong purifying selection and therefore evolve slowly according to the rule of functional constraint in molecular evolution. Moreover, our analysis showed that a comprehensive approach including various computational methods and multi-level information (from sequence to experimental data) is beneficial for understanding functional diversity of a large gene family in the post genomics era.

Materials and Methods

In this study, all the protein sequences of Ubiquitin conjugating enzymes were extracted from relevant protein databases for carrying out multiple alignments and subjecting it to various computational processes to build a phylogenetic tree. The various in silico methods used, are listed below sequentially:

Step 1: All the homologs of various organisms of Ubiquitin conjugating enzymes were retrieved from NCBI Protein database using Blast (Altschul et al. 1997). These sequences were downloaded and analyzed, based on their E-values, only those sequences were selected which had the E-value less than 10 or more than 30% amino acid identity discarding the other sequences which did not fall in this range. We found about 104 prominent hits.

Step 2: Using MUSCLE version 3.52 (Edgar 2004) with default settings multiple sequence alignment was constructed with all the selected sequences Keeping in mind the accuracy of the alignment which is critical, as it is the source of the phylogenetic signal and is essential for constructing the phylogenetic tree accurately. Then the resultant alignment was examined using JALVIEW to identify critical motifs or conserved residues (Clamp et al. 2004).

Step 3: After alignment of the dataset of 104 sequences of ubiquitin conjugating enzymes, the phylogenetic tree was constructed using Maximum likelihood method as implemented in PHYML version 2004 (Guindon and Gascuel

2003) under the Jones–Taylor–Thornton (JTT), with four categories of Gamma substitution rates (4G) and invariable sites model (Felsenstein 1996). The maximum likelihood method is a character-based method of tree construction. This method was selected because it is a character-based method and it infers the phylogeny based on all the individual characters like the nucleotides and the amino acids. This is a method for the inference of phylogeny. It evaluates a hypothesis about evolutionary history in terms of the probability that the proposed model and the hypothesized history would give rise to the observed data set. The supposition is that a history with a higher probability of reaching the observed state is preferred to a history with a lower probability. The method searches for the tree with the highest probability or likelihood. Maximum likelihood is probabilistic methods of inference. It implements explicitly models of molecular evolution and allows rigorous statistical inference. However, this approach is very computer intensive. For a given set of sequences, stochastic model of molecular evolution was used to assign a probability (likelihood) to each phylogeny. Maximum likelihood inference then consists of finding the tree, which assigns the highest probability (likelihood) to the data. The statistical support was calculated by generating 100 non-parametric bootstrap replicates using PHYML. Bootstrap is a statistical procedure to evaluate the reliability of the phylogenetic tree. The tree was viewed as unrooted and radial logarithmic and labelled using TREE ILLUSTRATOR version 0.52 (Trooksens et al 2005).

Step 4: Type 1 cluster divergence values (θ_λ) were calculated using DIVERGE 1.04 (Gu and Vander Velden 1999). The sites showing significant divergence were traced on the 3D structure of ube2c obtained from Protein Databank (PDB: IL7K) using RASMOL 2.7.2 (Sayle 1995).

Results and Discussion

In this study, phylogenetic analyses revealed the divergence of the protein Ube2c which is critical to comparative genomic studies for determining the horizontal or vertical gene transfer or gene function. The computational techniques used here allow precise genome annotation and are invaluable to studies of protein and protein complex evolution. The sequence divergence was determined to measure the evolutionary conservation of the gene and is fundamentally different from gene loss propensity. Although gene loss may be the result of a complete deletion or oblation of a gene, sequence divergence occurs through point mutations, as well as small deletions and insertions, and generally does not lead to elimination of the gene. Hence, these two variables, gene loss propensity and sequence divergence (or its correlate, the evolutionary rate), seem to be complementary measures of the conservation of a gene during evolution.

The phylogenetic analysis of this gene shows the clustering of the sequences into groups based on their similarity to one another. The clustering was done based on models developed by studying molecular evolution. There were a number of such models and software tools, which could be used, but we selected those that allowed diversity and flexibility in phylogenetic tree generation. The ability of the phylogenetic study therefore could be applied to the field of complex diseases in order to understand them better. Our analysis assumes importance in view of the fact that ubiquitin has an important role in cancers. Cancer is one such disease in which large numbers of genes are implicated wherein the cell machinery is pushed into an uncontrolled state of growth and loose its differentiation. The stable or erratic nature of evolution could point out the type of process underlying the gene studied. It is evident that homology modelling is not necessarily a one-to-one relationship, because a single gene in one genome may correspond to a whole family of paralogs in other genomes, which may be functionally diverse (Fig 1 & 2) as illustrated in this study.

The dataset of 104 REFSEQ protein sequences of ubiquitin conjugating enzyme (of E2 homologs) as mentioned above were extracted from public databases of NCBI -like PDB and Blast were aligned using MUSCLE (Edgar 2004). These homologs shared more than 30% identity with the query sequence of Ube2c. The sequences were from 43 different species belonging to different phyla. The alignment was edited using the JALVIEW package (Clamp et al. 2004). Care was taken that the important ubiquitin conjugating domain sequence and active cysteine were aligned correctly. The sequences were submitted in Phylip format to PHYML software for tree generation (Guindon and Gascuel 2003). Maximum Likelihood method was used for the

inference of phylogeny. This technique evaluates a hypothesis about evolutionary history in terms of the probability that the proposed model and the hypothesised history would give rise to the observed data set. The phylogenetic tree thus obtained was unrooted (figure 1). The branches of the resultant tree were grouped into four major clusters and one mini cluster. This grouping was done to calculate the functional divergence amongst the selected cluster groups as seen in figure 2 and 3. The E2s were exclusively eukaryotic with no significant homologs in Bacteria and Archeae. All the protein sequences constituting the tree were orthologs except the sequences of *Trypanosome cruzi* (figure 1: cluster 2 sequences 41 and 66) and *Saccharomyces cerevisiae* (figure 1: cluster 3 sequences 79 and 90). Orthologs are genes that share common ancestral gene and are separated by a speciation event whereas paralogs are those genes, which are a product of a gene duplication event. The species of *Homo sapiens, Mus musculus, Danio rerio* and *Xenopus tropicalis* were well dispersed in the tree in all clusters. The two viral sequences of African swine fever virus and *Acanthamoeba polyphaga mimivirus* were obvious cases of lateral gene transfer. The ube2s and ube2c were found in the cluster 1 and cluster 2 (numbers 101 and 25) respectively.

In order to calculate the functional divergence of the E2, the phylogenetic tree was grouped into 4 major clusters (figure 1). These selected clusters were taken into consideration for calculating the type 1 divergence. Functional divergence of a protein (family) could occur only after major evolutionary events such as gene duplication or speciation. Some of them may result in altered selective constraints (different evolutionary rates) at certain amino acid residues, which are called type I functional divergence, regardless of the underlying evolutionary mechanisms (Gu 2001). In this study, the site specific divergence θ was calculated for the 4 major clusters as shown in the phylogenetic tree (figure 1). For selected sets, if $\theta = 0$ then it implies that there is no site specific divergence. If the $\theta > 0$ then it means that this particular site in the protein has undergone divergence. Therefore, detection of site-specific rate shifts can provide a list of 'predicted' amino acid residues that may be responsible for functional divergence between member genes of a gene family. Some other studies support this 'functional divergence' view (Gaucher et al 2001, Jordan et al. 2001). The type 1 divergence (θ_λ) was calculated as site-specific posterior probability as implemented in DIVERGE 1.04 (Gu and Vander Velden 1999) (Figures 2 & 3). The divergence (θ_λ) between clusters 1, 2 & 4 was calculated in pairs. Divergence was also calculated for the pair of clusters 1, 2 and mini cluster as one set and cluster 4 as the other. Amongst the pairs, clusters 1 /cluster 4 showed the maximum divergence followed by the cluster 1/cluster 2, cluster 1&2/cluster 4 and cluster 2/cluster 4 in the decreasing order of divergence. Given that the θ_λ tests are biased by alignment error care was taken to have accurate multiple alignment. Sequence evolution rate is a traditional measure of the conservation during evolution of a gene. Early molecular evolutionary studies have unequivocally shown that different genes evolve at substantially different rates.

The type 1 divergence was calculated to observe the functional constraints on proteins which limit their evolutionary rates at specific sites. These constraints allow for the interpretation of sites of conserved residues and sites with a rate change as those most likely underlying the functional similarities and differences among protein subfamilies. In figure 2, each bar in the graph represents site-specific divergence at the site number provided at the bottom of the bar. These sites were also marked in the multiple alignments (figure 3). Each bar consists of different colour coded segments which represent a value as given in the table below the graph. The y-axis lists these site-specific probabilities and the x-axis consists of the specific sites of the protein sequence. The significance of these values is that these values show the extent of conservation and divergence of the protein's amino acid residues. The sites with significant rate change are marked in figure 2 of the multiple alignments and in colour red in figure 4 of the protein; in other words, the values show the amount of divergence between the selected clusters of the tree at the specific sites on the sequence of the protein. The ubiquitin conjugating enzymes contains a highly conserved core. The sites showing significant divergent rates were found in the structurally less important regions like the loops as seen in figure 4. This clearly implies that these proteins were under evolutionary pressure to maintain their structural and functional characteristics throughout the eukaryotic evolution.

The 3D structure of the ube2c was obtained from the PDB and the sites with

significant divergence (cut off value 0.05) were traced on the structure (figure 4). The E2 protein is remarkable because despite its relatively small size (typically 20kDa), it interacts with 3 or 4 different proteins, namely ubiquitin, E1, E3, and perhaps the target protein. Therefore, the E2 protein has to maintain its structural features in order to allow interactions with common elements of the system, such as ubiquitin and E1, and yet specify interactions with its cognate E3 and target protein (Passmore and Barford 2004). The Ube2c protein is classified as alpha-beta protein with Ubc-like fold belonging to the ubiquitin conjugating family. The ube2c protein is 179 amino acids in length. This protein is relatively inflexible. It contains the ubiquitin binding domain which is about 150 residues (from 34 to 175 residues) with highly conserved catalytic cysteine at the active site at 114th position in a shallow groove on the protein surface (Lin et al 2002). The conserved core is a central beta sheet with two flanking helices. The more variable regions flank the conserved region. The protein ube2s is of 222 residues in length with same conserved core and the cysteine at 95th position. The function of ube2s is characterized by its similarity to the other ubiquitin conjugating proteins whose function has been proven experimentally. The ubiquitin binding domain is shared by all the conjugating enzymes. The secondary structural elements are highly conserved in all known Ubiquitin conjugating protein structures. The central beta sheet had 3 significant divergence sites with flanking helix showing two sites. Other sites showing divergence are dispersed in the loop regions (fig 4). Throughout the phylogenetic tree, we observed varying sequence divergence amongst the clusters, but the functional and the structural characteristics of these proteins have been preserved indicating functional conservation.

Conclusion

The phylogeny of ubiquitin conjugating enzymes has given us evidence of varied and diverse history of these proteins. Ubiquitin conjugating enzymes are widespread in eukarya with multiple homologs in representative species. It is evident from these results that despite gene duplications and mutations, E2s are under strong evolutionary pressure to maintain structure and relatively low pressure to maintain sequence suggesting that the regions showing divergence are more important for intrinsic structural reasons than for specific protein-protein interactions. It is concluded that the E2s are undergoing active evolution and have apparently been subject to strong purifying selection. It is clear that cellular processes underlying the E2 genes are highly active resulting in wide variety of the ubiquitin conjugating enzymes.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 17: 3389 –3402.

Aristarkhov A, Eytan E, Moghe A, Hershko A, Ruderman JV, 1996. E2-C, a cyclin-selective ubiquitin carrier protein required for the destruction of mitotic cyclins. Proceedings of National Academy of Sciences USA, 93: 4294–4299.

Arvand A, Bastians H, Welford SM, Thompson AD, Ruderman JV, Denny CT, 1998. EWS/FLI1 upregulates mE2-C, a cyclin-selective ubiquitin conjugating enzyme involved in cyclin B destruction. Oncogene, 17: 2039–2045.

Clamp M, Cuff J, Searle SM, Barton GJ, 2004. The Jalview Java alignment editor. Bioinformatics, 12: 426-7.

Edgar RC, 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research, 32: 1792-97.

Etlinger JD, Goldberg AL, 1977. A soluble ATP-dependent proteolytic system responsible for the degradation of abnormal proteins in reticulocytes. Proceedings of National Academy of Sciences USA, 74: 54–58.

Fang G, Yu H, Kirschner MW, 1998. Direct binding of CDC20 protein family members activates the anaphase-promoting complex in mitosis and G1. Molecular Cell, 2: 163–171.

Felsenstein J, 1996. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. Methods in Enzymology, 266: 418–427.

Gaucher EA, Miyamoto MM, Benner SA, 2001. Function-structure analysis of proteins using covarion-based evolutionary approaches: Elongation factors. Proceedings of National Academy of Sciences USA, 98: 548–552.

Goldstein G, Scheid M, Hammerling U, Boyse EA, Schlesinger DH, Niall HD, 1975. Isolation of a

polypeptide that has lymphocyte-differentiating properties and is probably represented universally in living cells. Proceedings of National Academy of Sciences USA, 72: 11-15.

Gu X, Vander Velden K, 1999. Diverge: a phylogeny based analysis for functional-structural divergence of a protein family. Molecular Biology and Evolution, 16: 1664–1674.

Gu X, 2001. Maximum-likelihood approach for gene family evolution under functional divergence. Molecular Biology and Evolution, 18: 453–464.

Guindon S, Gascuel O, 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Systems Biology, 52: 696–704.

Hershko A, 2005. The ubiquitin system for protein degradation and some of its roles in the control of the cell-division cycle (Nobel lecture). Angewandte Chemie International Edition England, 44: 5932–43.

Jordan K, Bishop GR, Gonzalez DS, 2001. Sequence and structural aspects of functional diversification in class I - mannosidase evolution. Bioinformatics, 17: 965–976.

Jamil K, Devi S, Khan M, 2006. In silico and in vitro investigations on cry4a and cry11a toxins of *Bacillus thuringiensis var israelensis*, Iranian Journal of Biotechnology, 4: 17-25.

King RW, Peters JM, Tugendreich S, Rolfe M, Hieter P, Kirschner MW, 1995. A 20 S complex containing CDC 27 and CDC16 catalyzes the mitosis-specific conjugation of ubiquitin to cyclin B. Cell, 81: 279–288.

Kramer ER, Gieffers C, Holzl G, Hengstschlager M, Peters JM, 1998. Activation of the human anaphase-promoting complex by proteins of the CDC20/Fizzy family. Current Biology, 18: 1207–1210.

Lin Y, Hwang WC, Basavappa R, 2002. Structural and functional analysis of the human mitotic-specific ubiquitin-conjugating enzyme, UbcH10. Journal of Biological Chemistry, 277: 21913-21921.

Okamoto Y, Ozaki T, Miyazaki K, Aoyama M, Miyazaki M, Nakagawara A, 2003. UbcH10 Is the Cancer-related E2 Ubiquitin-conjugating Enzyme. Cancer Research, 63: 4167–4173.

Pagano M, Tam SW, Theodoras AM, Beer-Romero P, DelSal G, Chau V, Yew PR, Draetta GF, Rolfe M, 1995. Role of the ubiquitin-proteasome pathway in regulating abundance of the cyclin-dependent kinase inhibitor p27. Science, 269: 682–685.

Passmore LA, Barford D, 2004. Getting into position: the catalytic mechanisms of protein ubiquitylation. Biochemical Journal, 379: 513–525.

Pruitt KD, Maglott DR, 2001. RefSeq and LocusLink: NCBI gene-centered resources. Nucleic Acids Research, 29: 137-140.

Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pandey A, Chinnaiyan AM, 2004. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. Proceedings of National Academy of Sciences USA, 22: 9309-9314.

Robinson PA, Ardley HC, 2004. Ubiquitin-protein ligases. Journal of Cell Science, 117: 5191–5194.

Sabita K, Jamil K, 2006. Analysis of EGFR mutations which have a response to quinazolin inhibitors. In: Proceedings Of The Fifth International Conference On Bioinformatics Of Genome Regulation And Structure (July 18-22, 2006, Novosibirsk, Russia), 3: 229-232.

Sabita K, Krishna Kishore M, Jamil K, 2008. Homology Models of the Mutated EGFR and their Response towards Quinazolin Analogues. Journal of Molecular Graphics and Modeling, 27: 244-254.

Suman G, Khan M, Sabitha K, Jamil K, 2006. Mutation in mexR-gene leading to drug resistance in comeal keratitis in human. Indian Journal of Experimental Biology, 44: 929-936.

Sayle RA, Milner-White EJ, 1995. RASMOL: biomolecular graphics for all. Trends in Biochemical Sciences, 20: 374-376.

Su AL, Welsh JB, Sapinoso LM, Kern SC, Dimitroy P, Lapp H, Schultz PG, Powel SM, Moskaluk CA, Frierson HF Jr, Hampton GM, 2001. Molecular classification of human carcinomas by use of gene expression signatures. Cancer Research, 15: 7388-93.

Townsley FM, Aristarkhov A, Beck S, Hershko A, Ruderman JV, 1997. Dominant-negative cyclin-selective ubiquitin carrier protein E2-C/UbcH10 blocks cells in metaphase. Proceedings of National Academy of Sciences USA, 94: 2362–2367.

Trooskens G, De Beule D, Decouttere F, Van Criekinge W, 2005. Phylogenetic trees: visualizing, customizing and detecting incongruence. Bioinformatics, 21: 3801-3802.

Wang Y, Gu X, 2001. Functional divergence in the caspase gene family and altered functional constraints: statistical analysis and prediction. Genetics, 158: 1311–1320.

Yamanaka A, Hatakeyama S, Kominami K, Kitagawa M, Matsumoto M, Nakayama K, 2000. Cell cycle-dependent expression of mammalian E2-C regulated

by the anaphase-promoting complex / cyclosome. Molecular Biology of the Cell, 11: 2821–2831.

Yu H, King RW, Peters JM, Kirschner MW, 1996. Identification of a novel ubiquitin-conjugating enzyme involved in mitotic cyclin degradation. Current Biology, 6: 455–466.

Figure 1 - Phylogenetic tree of ubiquitin conjugating enzymes.

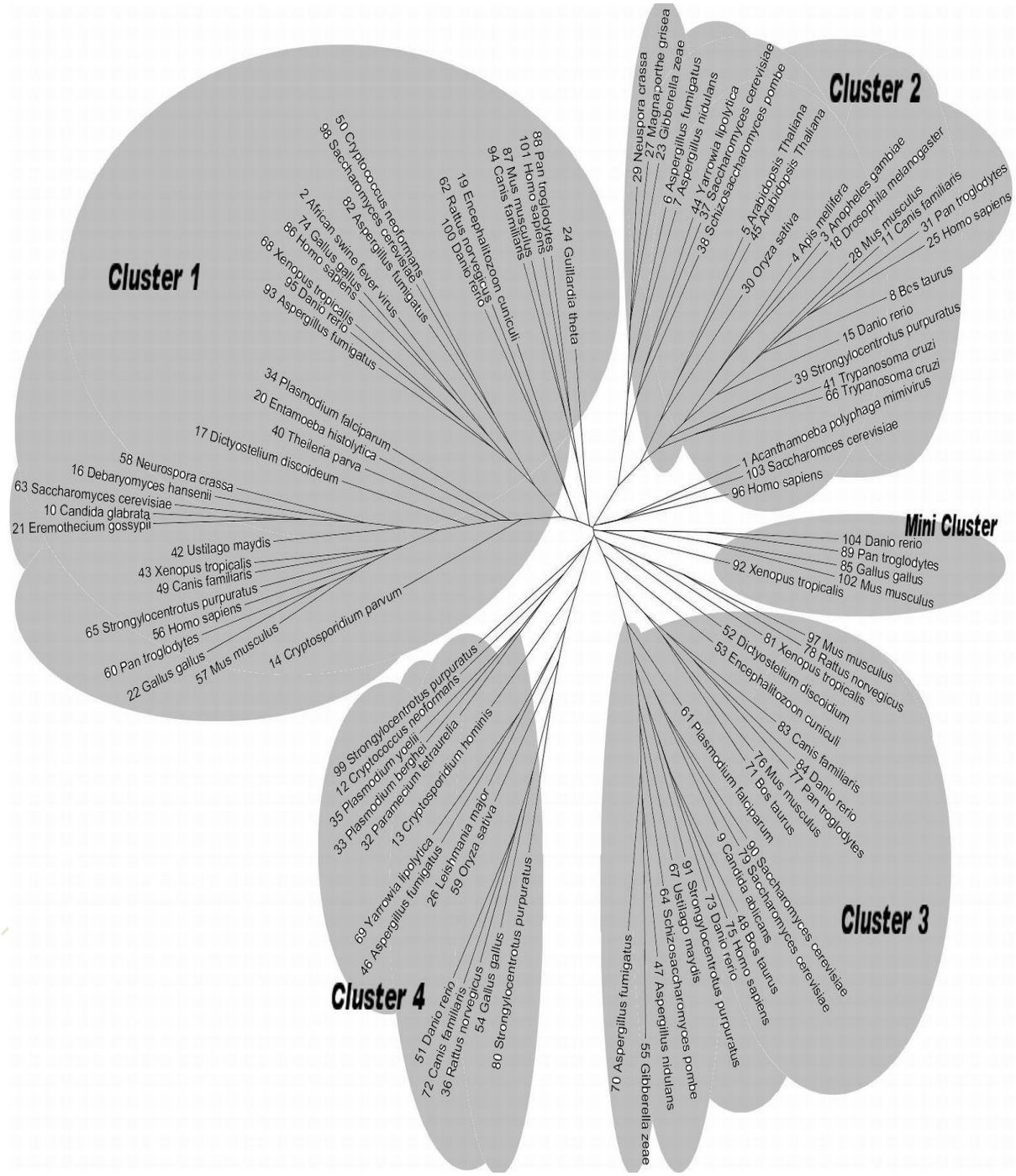


Figure 2 - Significantly divergent sites marked on the multiple alignments.

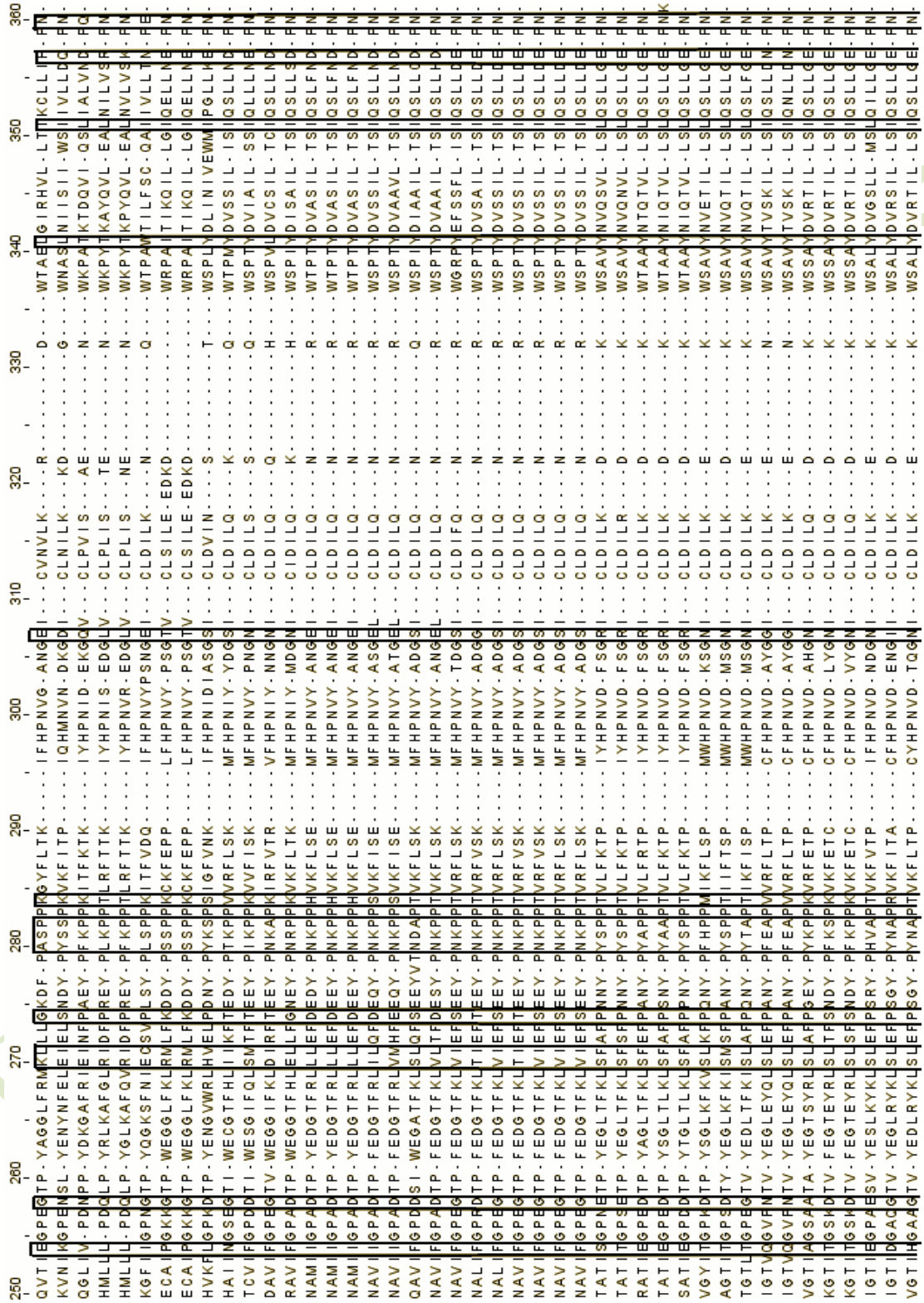
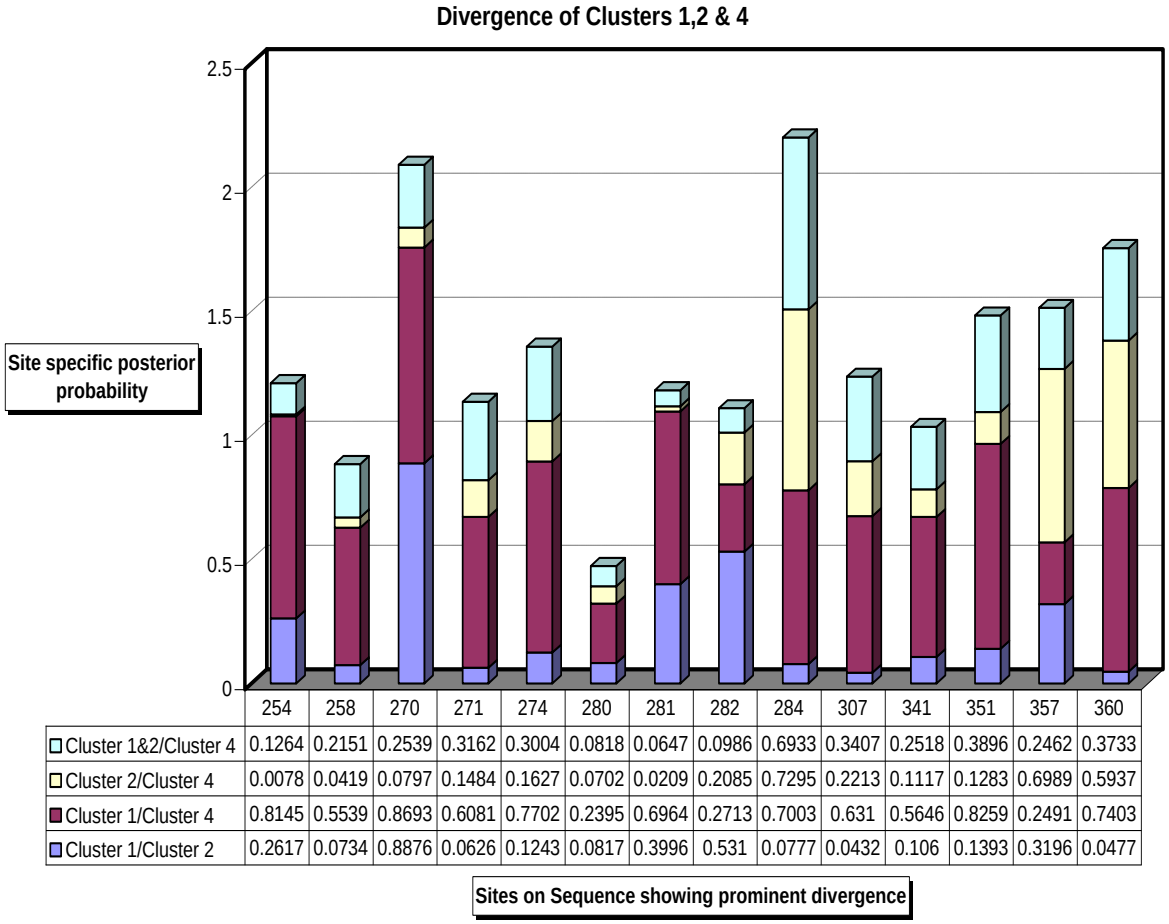


Figure 3 - Graph and calculated values of divergence for cluster pairs of the phylogenetic tree.



	254	258	270	271	274	280	281	282	284	307	341	351	357	360
Cluster 1&2/Cluster 4	0.1264	0.2151	0.2539	0.3162	0.3004	0.0818	0.0647	0.0986	0.6933	0.3407	0.2518	0.3896	0.2462	0.3733
Cluster 2/Cluster 4	0.0078	0.0419	0.0797	0.1484	0.1627	0.0702	0.0209	0.2085	0.7295	0.2213	0.1117	0.1283	0.6989	0.5937
Cluster 1/Cluster 4	0.8145	0.5539	0.8693	0.6081	0.7702	0.2395	0.6964	0.2713	0.7003	0.631	0.5646	0.8259	0.2491	0.7403
Cluster 1/Cluster 2	0.2617	0.0734	0.8876	0.0626	0.1243	0.0817	0.3996	0.531	0.0777	0.0432	0.106	0.1393	0.3196	0.0477

Figure 4 - Sites plotted on the 3D structure of ube2c showing significant divergence (in red colour).