bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Evolutionary history of Alzheimer Disease causing protein family Presenilins with pathological implications.

Ammad Aslam Khan[1], Bushra Mirza[2] Hashim Ali Raja[3]

Author affiliation

[1] Department of Biochemistry, Barrett Hodgson University, Karachi, Pakistan

[2] Department of Animal Sciences, Quaid-i-Azam University, Islamabad, Pakistan

Corresponding author

Ammad Aslam Khan

Email: ammad.khan@bhu.edu.pk and ammadaslamkhan@gmail.com

Postal address: Barrett Hodgson University, Department of Biochemistry, Korangi Creek, Karachi.

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Abstract

Presenilin proteins are type II transmembrane proteins. They make the catalytic component of Gamma secretase, a multiportion transmembrane protease. Amyloid protein, Notch and beta catenin are among more than 90 substrates of Presenilins. Mutations in Presenilins lead to defects in proteolytic cleavage of its substrate resulting in some of the most devastating pathological conditions including Alzheimer disease (AD), developmental disorders and cancer. In addition to catalytic roles, Presenilin protein is also shown to be involved in many non-catalytic roles i.e. calcium homeostasis, regulation of autophagy and protein trafficking etc. These proteolytic proteins are highly conserved, present in almost all the major eukaryotic groups. Studies on wide variety of organisms ranging from human to unicellular dictyostelium have shown the important catalytic and non-catalytic roles of Presenilins. In the current research project, we aimed to elucidate the phylogenetic history of Presenilins. We showed that Presenilins are the most ancient of the Gamma secretase proteins and might have their origin in last common eukaryotic ancestor (LCEA). We also demonstrated that these proteins have been evolving under strong purifying selection. Through evolutionary trace analysis, we showed that Presenilin protein sites which undergoes mutations in Familial Alzheimer Disease are highly conserved in metazoans. Finally, we discussed the evolutionary, physiological and pathological implication of our findings and proposed that evolutionary profile of Presenilins supports the loss of function hypothesis of AD pathogenesis.

Keywords: Gamma Secretase; Presenilin ; Alzheimer Disease; Amyloid beta; Endoproteolysis; Protein evolution.

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

1. Introduction

Presenilin is a multipass transmembrane protein with nine transmembrane domains (Laudon et al., 2005; Spasic et al., 2006). Presenilin proteins are the major catalytic component of Gamma secretase, a multi-complex protease involved in the catalysis of many membrane proteins (De Strooper et al., 1998; Zhang et al., 2000). Presenilin performs this catalytic activity with the aid of other three Gamma Secretase proteins namely Nicastrin, APH1 and PEN2 (Goutte et al. 2002; Edbauer et al. 2003; Takasugi et al. 2003). In human, there are two Presenilin paralogs, Precesnilin-1 and Presenilin-2 showing 67% sequence similarity (Levy-Lahad et al., 1995). Both these Presenilin proteins are ubiquitously expressed in many human tissues including brain, heart, kidney and muscle (Lee et al., 1996). Inside the cells, Presenilin is abundantly found in Endoplasmic reticulum and Golgi bodies (Area-Gomez et al., 2009; De Strooper et al., 1997; Walter et al., 1996). An important aspect of Presenilin is that under physiological conditions, it undergoes endoproteolysis generating N and C terminal fragments (NTF and CTF respectively) (Brunkan et al., 2005; Thinakaran et al., 1996) involving amino acids 292-299 (Fukumori et al., 2010; Podlisny et al., 1997). This endoproteolysis event is considered important for the Gamma secretase activity. However, there are some Presenilin mutants in which though endoproteolysis is defective but they still exhibit enzymatic activity (Jacobsen et al., 1999; Steiner et al., 1999). There are more than 90 substrates of Gamma Secretase/Presenilin of which APP and the Notch receptor are the predominant substrates (Beel and Sanders, 2008; McCarthy et al., 2009). It is the defect in Presenilin based proteolysis in APP that leads to AD because of the over production of the neurotoxic Aβ42 peptide, which triggers inflammatory responses in brain leading to neuronal dysfunction and ultimately cell death (Duff et al. 1996; Jacobsen et al. 1999; Zhang et al. 2011). However, it is just recently that validity of the amyloid hypothesis as the sole explanation of this AD is questioned and an alternative explanation have been proposed according to which mutations in Presenilin proteins cause loss of many important Presenilin functions (loss of function mutations) in the brain leading to neurodegeneration and dementia, the hallmark features of AD (Kelleher and Shen, 2017; Shen and Kelleher, 2007; Sun et al., 2017).

Quite recently, it has been shown in various studies that in addition to acting as a major protease involved in the regulated intramembrane proteolysis, Presenilin is involved in many non-

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

catalytic processes too. For instance, the role of Presenilin in calcium homoeostasis has been well documented. Presenilin mutations are connected with the defects in calcium signaling in both neuronal and non-neuronal cells and more importantly, also in FAD patients with Presenilin mutations (LaFerla, 2002; Popugaeva and Bezprozvanny, 2013; Supnet and Bezprozvanny, 2010). Presenilin is also involved in the regulation of autophagy in Gammas secretase independent way. The mutations in Presenilin have been shown to be involved in the accumulation of immature autophagic vesicles. The defects in autophagy due to Presenilin mutations contribute towards the neurodegeneration in AD patients because of excess neuronal death (Menzies et al., 2015; Nixon and Yang, 2011; Ohta et al., 2010). Finally, Presenilin have also been found to be involved in the trafficking of proteins in both Gamm secretase dependent (Cai et al., 2003; R. Wang et al., 2006) and independent way (Scheper et al., 2004; Suga et al., 2004).

Though it is known that Presenilins are present in vast variety of eukaryotes, very little work have been done to explore the evolutionary history of this important protein. Even the studies that have been conducted in this direction, are limited either to elucidate the functional aspects of Presenilins in evolutionarily diverse variety of model organisms, or are done before the advent of next generation sequencing technology and hence their coverage is quite limited (Murray et al. 2000; J. Wang et al. 2006; Otto et al. 2016). In this regard, ours is the first such study, elucidating the phylogeny and evolutionary history of Presenilins by exploring their origin and evolution and discussing the physiological implication of the evolutionary profile they exhibits. In our surge, we found that Presenilin is a very ancient protein and might be present in LCA of eukaryotes. We also showed that Presenilin have been evolving under strong purifying selective pressure. Finally we hypothesized that evolutionary history of Presenilin does not support the amyloid hypothesis as an explanation of the etiology of Alzheimer dieses.

2. Materials and Methods

2.1 Identification and Extraction of Gamma secretase proteins

For the extraction of protein sequences of four Gamma secretase proteins, Presenilin, Nicastrin, APHI and PEN-2, BLAST search tool was used (Altschul et al., 1990). Orthologs of each of

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

these four proteins in human genome were used as query sequence to find the corresponding ortholog and paralog in total 47 species, including each major eukaryotic group, from primates to unicellular protists (Supplementary file 1). For all those proteins with not a well-defined annotation, the result was confirmed with reverse BLAST (Wall et al., 2003) and confirming the putative domain structure from CDD database (Marchler-Bauer et al., 2015). For those species whose genome has been sequenced but still not incorporated in NCBI database or is partially incorporated, BLAST search was conducted from the individual genome database websites of each such species.

2.2 Phylogenetic analysis

Phylogenetic analyses were performed using Mega6 (Tamura et al., 2013). Briefly, a total of 153 protein sequences were selected from 47 different species ranging from primates to simplest of protists. Proteins with partial sequences were not included in the analysis. The sequences were aligned by using ClustalW tool employed in MEGA software with affine gap penalties of 5 for gap opening and 1 for gap extension (Chenna et al., 2003). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The analysis gave rise to an optimal tree with the sum of branch length = 39.25074069 (see the figure legend for a detailed account) .

For rate of selection analysis, coding sequences of Vertebrate Presenilin-1and Presenilin-2 gene were extracted from NCBI genome database (Supplementary file 3). Different approaches were adopted to find the rate of evolution of vertebrate Presenilin proteins. Pairwise comparison of the number of synonymous nucleotide substitutions per synonymous (dS) site and non-synonymous nucleotide substitutions per non-synonymous site (dN) was carried out by using Nei-Gojobori method (Nei and Gojobori, 1986). In addition to pairwise methods, the dN/dS ratio in different branches of the maximum-likelihood tree was estimated using the codon-based genetic algorithm implemented in the GA-BRANCH program available at the Datamonkey server (http://www.datamonkey.org/help/ GABranch.php). This approach assigns each branch to an incrementally estimated class of dN/dS ratios without requiring a specification of the branches *a priori* (Kosakovsky Pond and Frost, 2005a). For selection analysis for each Codon, sliding window approach was incorporated in SLAC method employed at the Datamonkey server

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

(Kosakovsky Pond and Frost, 2005b, 2005c) . Evolutionary distance between all possible pairs of vertebrate Presenilin paralogs was estimated by Tajima's relative test employed in Mega6 software (Tajima, 1993; Tamura et al., 2013). Each pair of paralogs was compared with Saccoglossus kowalevskii Presenilin protein sequences which was taken as outgroup.

2.3 Evolutionary Trace Analysis (ETA)

For determining the evolutionary conserved residues, evolutionary trace analysis were employed. ETA is a powerful technique that rank the individual amino acids on the bases of their conservation profile. Lower the rank, more conserve the residue and vice versa, with 1 being the minimum rank value (Lichtarge et al., 1996). The ETA results were incorporated for evolutionary structural analysis by employing pyETV and CHIMERA software tools (Lua and Lichtarge, 2010; Pettersen et al., 2004). The heatmaps for the for the figurative representation of ETA results are generated by using conditional formatting tool embedded in Microsoft excel (Meyer and Avery, 2009).

3. Results

3.1 Distribution and extent of duplication of Gamma secretase proteins across eukaryotes

In our surge to elucidate the evolutionary history of metazoan Presenilins, we started with investigating the distribution of four Gamma secretase proteins, Presenilin, APH-1, PEN-2 and Nicastrin in 46 species representing all the major groups in eukaryotes (Supplementary file 1 and 2). BLAST analysis revealed that these proteins though are distributed in diverse array of eukaryotic the, the extent of distribution is not uniform (Fig.1). For instance, Presenilin seemed to be present in almost all the eukaryotic lineages with the exception of two alveolates (Vitrella brassicaformis and Tetrahymena thermophile) and one excavate (Naegleria gruberi). On the other hand, other gamma secretase proteins included in this study i.e. Nicastrin, Pen-2 and APH-1 were relatively less widely distributed then Presenilin. Except of APH-1 which seems to be absent in plants, all three the three Gamma secretase proteins are present in the three major divisions of eukaryotes e.g. plants, animals and fungi (Fig.1). Among the major eukaryotic groups Alveolates showed least number of Gamma secretase proteins i.e. in all the five species

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

we included in our sample, both APH-1 and Nicastrin were absent while Pen-2 was present in only one Alveolate species e.g in Hammondia hammondi. Similarly in four Excavates we included in our BLAST analysis, Pen-2 is present in only one species i.e. in Trichomonas vaginalis while both APH-1 and Nicastrin are absent in all four. After opisthokonta, plants and ameobazoa are the eukaryotes in which we found pre-dominant presence of Gamma secretase proteins.

In addition to exploring the distribution of Gamma secretase genes, we also investigated the extent of duplications of Presenilins and their other counterpart Gamma secretase genes in eukaryotes. Interestingly, though the number of paralogs for all four genes were different in different species, it seems that these genes underwent less duplications in general. In vertebrates for instance, neither Pen-2 nor Nicastrin seem to have their counterpart paralog with the only exception of Xenopus in which these proteins had one counterpart paralog each. Similarly, both Presenilin and APH-1 are represented by two paralogs in vertebrates e.g. Presenilin-1 and Presenilin-2 and APH-1A and APH-1B. Only in mice, APH-1 have a third paralog e.g. APH-1C. In invertebrate metazoans, all four proteins are in unduplicated form (singleton) with the exception of C.elegans in which Presenilin are represented by three paralogs (Sel-12, Hop-1 and Spe-4). Similar to the case of multicellular organisms, we did not see significant gene duplication based expansion of three of the Gamma secretase proteins (Nicastrin, APH1 and Pen-2) in major unicellular groups. However Presenilin did expanded to some extent (with maximum of four paralogs) in some of these organisms (Fig 2 and Supplementary file 4).

3.2 Rate of evolution of Presenilins

Exploring the rate of evolution of proteins is an important part towards the elucidation of their evolutionary history. To this direction, we attempted to investigate the rate of evolution of Presenilins, the most widely distributed Gamma secretase protein in eukaryotes. We opted for different approaches including determining the net dn/ds difference in Presenilins, branch site based method (GA-branch) to estimate the lineage specific rate of evolution and sliding window method to estimate the rate of evolution of each site along the whole length of Presenilin proteins (fig. 3 A-D and supplementary file 5). Interestingly, in all these selection analysis, we witnessed both Presenilin predominantly evolving under purifying selection. Finally, we employed Tajima's relative rate test to find out if there is any significant difference between rates of

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

evolution of two Presenilin proteins (fig. 3 E). In this test, we compared the extent of substitutions in both vertebrate Presenilin paralogs by taking the protein sequence of a third species as outgroup in which the respective protein exists as unduplicated singleton e.g. the invertebrate Saccoglossus kowalevskii. The results of Tajima's test showed that there is no significant difference in evolutionary rate between two paralogs.

3.3 Evolutionary Traces in metazoan Presenilins

Evolutionary Trace Analysis (ETA) is a powerful phylogenetic tool to identify evolutionarily conserved residues in protein sequences. The main principle of ETA is that mutations occurring in distant branches in a phylogenetic tree are more important than those occurring at closely related branches. In this way, ETA not only contribute towards our understanding of the evolutionary history of a protein but also provide a mean to identify biologically important residues (Lichtarge et al., 1996). So, in order to identify the highly conserved regions and residues in metazoan Presenilin, we employed ETA. With the help of ETA analysis, we identified 103 residues in Presenilins which were conserved (with rank value of 1) throughout metazoans from primates to choanoflagellates (fig. 4A and supplementary file 6). We also looked for the location of these significantly ranked residues along the whole length of Presenilin protein. Interestingly, majority of the highly ranked residues were present in the transmembrane part of the protein (fig. 4B and C). In order to see the biological importance of these traces, we looked for the site directed mutation analysis by making use of Uniprot (Pundir et al., 2017) (fig. 4D). As was expected, a significant number of mutating sites (60%) were highly conserved among the metazoans as depicted by their very low rank values rank vale of 1 or 2.

3.4 Alzheimer Disease and Presenilin Evolution

Mutations in Presenilin gene are among the hallmark causes of Alzheimer disease. In our surge for elucidation evolutionary history of Presenilin, we were keen to investigate its implication in the pathogenesis of Alzheimer disease. To this direction, we made use of the uniprot database for natural variants (Pundir et al., 2017) in human Presenilin-1 protein causing familial Alzheimer disease (FAD). When we applied the ETA analysis on these AD/FAD causing sites, we were amazed to see a predominant number of sites exhibited very high level of conservation (70%)

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

throughout metazoans from human to choanoflagellates (fig. 5A and B). Of these 70% of rank 1 residues, majority lie in the alpha helix as we observed by the structural analysis in which we mapped the ETA based ranked residues on the existing human Presenilin-1 structure (fig. 5C). Similarly, the surface view of the structure reveals the predominant presence of these conserved sites in the hydrophobic region of the protein which is in consistent with our analysis that a big majority of these residues lie in the transmembrane part of the protein. (fig. 5D).

4. Discussion

Gama secretase is a protein complex that along with β-Secretase is involved in the successive proteolysis of amyloid precursor protein producing amyloid β-protein whose abnormal deposition in the brain leads to AD. Of the four proteins making γ-Secretase, it is the catalytic role of Presenilin which is most prominently involved in the AD pathogenesis. Because of the important role Presenilin play in AD and its other catalytic and non-catalytic roles, it performs in evolutionarily diverse eukaryotes, we considered it important to explore the evolutionary history of this protein and its implication in the roles it performs. A protein BLAST based search for all the γ-Secretase proteins showed that these proteins are widely distributed in all the eukaryotic lineages with Presenilin showing the widest distribution, reflecting its vital role. Among all the species we included in our sample for protein evolution study, only one excavate (Naegleria gruberi) and Ciliates were the groups lacking Presenilin gene in their genome. It is understandable for parasitic species like Naegleria to lack such genes as parasitic organisms tend to have reduced genome because of repeated gene loss during the course of evolution (JACKSON, 2015; Sakharkar and Chow, 2005; Slamovits, 2013). But absence of Presenilin in Tetrahymena (a free living ciliate) may be surprising though. However, Ciliates seem to lack other important substrates of Presenilin too i.e. Amyloid precursor protein and Notch (data not shown). This reflects the possibility of coevolution of Presenilin with its substrates though more study is needed to explore this aspect of Presenilin evolution. Whether the absence Presenilin in Ciliates is result of gene loss or a representative of a feature inherited from a simpler common ancestor cannot be determined at the moment because of the unresolved phylogeny of the basal protist groups (Baldauf, 2003; Keeling et al., 2005; Parfrey et al., 2006). Interestingly, we did not find a single species in which any of the other three γ-Secretase proteins was present when Presenilin was absent. It not only means that Presenilin is the most important protein as for as

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

activity of γ-Secretase is concerned but also mean that perhaps in the eukaryotic LCA, γ-Secretase activity was exhibited solely by Presenilins. It is interesting to notice that Amyloid Beta protein though is highly conserved in metazoans, it is not present in vast number of unicellular organisms. On the other hand, in the same way as is Presenilin, Notch proteins are present throughout the eukaryotes, in metazoans, protists, plants and fungi. Therefore we can say that in LCEA, the main proteolytic function of Presenilins was the catalysis of Notch related proteins while the catalysis of amyloid beta is rather a metazoan innovation. Likewise, incorporation Presenilin to form Gamma secretase complex is also most probably a metazoan innovation, contributing towards the diversity of its proteolytic activity to fulfil the requirements of a complex multicellular organization. The independent evolution of these two catalytic activities of Presenilin is also reflected by the fact that mutations in Presenilin protein that affect the Amyloid beta proteolysis does not primarily affect proteolysis of Notch by Presenilin (Capell et al., 2000; Kulic et al., 2000). The conservation of Asp-257 and Asp-385 throughout the eukaryotic life, from human to hydra to all the protists we included in our analysis, signifies the pivotal role of Presenilins as a protease. This preservation of catalytic residues in species in which Presenilin is the only representative protein of Gamma secretase confirms the fact that assembly of Presenilin to form Gamma secretase is not required for performing its proteolytic activity (Beher et al., 2001). The conservation of Aspartic acid residues in those species in which Presenilin is the only of the Gamma Secretase protein mean that in LCEA, a proto-Presenilin could perform its catalytic activity independent of the Gamma secretase activity. Likewise, as Presenilin perform its non-catalytic functions i.e. protein trafficking, calcium homeostasis and apostasies etc. independent of Gamma secretase activity (Capell et al., 2000; Otto et al., 2016), presence of Presenilin in these simple organisms also hints towards the ancient origin of non-catalytic role of Presenilin which it performs independent of Gamma secretase. So a protein which is present such a diverse group of organisms and performing a variety of important functions of both catalytic as well as non-catalytic nature, must also be evolving under some considerable functional constrains. And this is what we observed in our evolutionary rate analysis too. We employed different strategies to estimate the rate of evolution of vertebrate Presenilin i.e. estimation of net dn/ds difference in Presenilins, Hyphy based estimation of the lineage specific rate of evolution and sliding window method to estimate the rate of evolution of each site along the whole length of Presenilin proteins. As expected, all these strategies proved

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

that Presenilin protein have been evolving under strong purifying selection. This was further confirmed by the Tajima's relative rate test which showed that even duplication of vertebrate Presenilin did not affect this trend and both the Presenilin paralogs went on evolving under negative selection pressure. This is also in accordance with the general understanding that the older genes (and hence present in wide range of organisms) and the genes more important for the survival of the organism evolve slowly than the younger ones (Albà and Castresana, 2005; Domazet-Loso and Tautz, 2003; Hirsh and Fraser, 2001; Wilson et al., 1977; Wolf et al., 2009).

One of the important functional aspects of Presenilin is the structure-function relationship of Presenilin (Bai et al., 2015). The study of the topology of Presenilin has revealed important aspects of this structure function relationship. For instance, studies conducted towards this direction have shown the importance of transmembrane portions of Presenilin involved in the catalytic activity of Gamma secretase. For instance, studies based upon substituted cysteine accessibility method (SCAM) and Nuclear Magnetic Resonance (NMR) have shown the importance of TMD1, TMD5, TMD6, TMD7 and TMD9 in the catalysis of transmembrane proteins (Sato et al., 2008, 2006, Tolia et al., 2008, 2006; Watanabe et al., 2010). In this direction, we also looked to see if we can see any selective constrain also on transmembrane domains of Presenilin as compared to the cytosolic or luminal portions. As was expected, the ETA analysis also showed that a significant majority (65%) of the highly conserved amino acids (with ETA rank value = 1) were present in the transmembrane region of the protein. This conservation of the transmembrane portion of the Presenilin reflects the functional constrain TMDs exhibit by providing a hydrophilic catalytic pore in the hydrophobic milieu for an efficient catalysis of the membrane proteins. Intrigued by the evolutionary history of Presenilin, we were curious to look at the evolutionary profile of amino acid sites in Presenilin whose mutation give rise to Familial Alzheimer disease (FAD). This evolutionary analysis of FAD causing mutations gave some interesting results. For instance, as was expected from structure-function evolutionary analysis, majority of these conserved FAD causing sites were located in the alpha helical transmembrane part of the Presenilin. In addition, many of these mutations seem to be exposed to the hydrophilic aqueous environment as depicted by the surface view of structure analysis. Intriguingly, of the total amino acid sites susceptible for FAD, a significant majority were highly conserved among metazoans from human to choanoflagellates (70%). The genes/loci that are involved in old age diseases usually tend to evolve under lesser functional

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

constrain. It reflects that the sites whose mutation lead to FAD might be involved in performing some highly important functions (Drenos and Kirkwood, 2010; Kirkwood and Austad, 2000). These functions may involve eliciting the signaling cascades and networks controlling physiological manifestations as complex as learning and memory. In other words, we can say that the mutation in these sites might have disturbed important physiological functions in addition to disruption of normal amyloid catalysis. In addition it has been shown that genes involved in a disease due to dominant negative mutations evolve under purifying selection (Blekhman et al., 2008). Interestingly, it has been proposed recently that only amyloid hypothesis cannot account for the occurrence and progression of AD. According this alternative hypothesis, mutations in Presenilin leading to AD/FAD should rather be considered as loss of function dominant negative mutations rather than gain of function mutations as significant number of these mutations severely impair the functioning of Gamma secretase instead of only increasing Aβ42 production (Kelleher and Shen, 2017; Sun et al., 2017; Watanabe and Shen, 2017). This later hypothesis is in conformity with our findings. Therefore, the outcome of our evolutionary analysis showing the ancient origin of Presenilins, their evolution under strong negative selection and strong conservation of the amino acid sites involved in AD pathogenesis favors idea that through loss of function mutations, Presenilins play a dominant negative role leading to AD pathogenesis. Our findings also favor to explore candidates other than amyloid beta also to elucidate the role of Presenilin in AD etiology not only to gain more mechanistic details of AD pathogenicity but also to formulate better therapeutic strategies for this devastating disease.

5. Conclusion

Presenilins are type II membrane proteins involved in the proteolysis of quite an important number of proteins. Of the four Gamma secretase proteins, Presenilins are the most primitive one. Their primitive origin means that they may be present in the LCEA and may be the only representative of Gamma secretase activity in LCEA, performing both catalytic as well as non-catalytic roles. Presenilin have been evolving under strong purifying selection representing the physiological vitality of this protein. Remarkably, the Presenilin residues susceptible for AD show significantly high conservation in metazoans manifesting the involvement of these residues in important physiological processes. The mutation in such functionally constrained residues

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

would eventually lead to AD like complex and multifaceted disorder. Based upon their evolutionary profile we can predict that involvement of Presenilin in AD pathogenicity cannot be ascribed to a single cause like amyloid beta production but other factors should also be explored for more effective therapeutic solutions.

Competing interest

We have no competing interest for this article.

Author's contribution

AAK conceived the idea. AAK and BM extracted and analyzed the data, AAK wrote the article.

.

Funding

This research was done through the support from by Higher Education Commission of Pakistan (HEC) and Barret Hodgson University, Karachi, Pakistan.

References

Albà, M.M., Castresana, J., 2005. Inverse relationship between evolutionary rate and age of mammalian genes. Mol. Biol. Evol. 22, 598–606. https://doi.org/10.1093/molbev/msi045

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2

Area-Gomez, E., de Groof, A.J.C., Boldogh, I., Bird, T.D., Gibson, G.E., Koehler, C.M., Yu, W.H., Duff, K.E., Yaffe, M.P., Pon, L.A., Schon, E.A., 2009. Presenilins Are Enriched in Endoplasmic Reticulum Membranes Associated with Mitochondria. Am. J. Pathol. 175, 1810–1816. https://doi.org/10.2353/ajpath.2009.090219

Bai, X.C., Yan, C., Yang, G., Lu, P., Ma, D., Sun, L., Zhou, R., Scheres, S.H., Shi, Y., 2015. An atomic structure of human gamma-secretase. Nature 525, 212–217. https://doi.org/10.1038/nature14892

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Baldauf, S.L., 2003. The Deep Roots of Eukaryotes. Science (80-.). 300, 1703–1706. https://doi.org/10.1126/science.1085544

Beel, A.J., Sanders, C.R., 2008. Substrate specificity of gamma-secretase and other intramembrane proteases. Cell. Mol. Life Sci. 65, 1311–34. https://doi.org/10.1007/s00018-008-7462-2

Beher, D., Wrigley, J.D.J., Nadin, A., Evin, G., Masters, C.L., Harrison, T., Castro, J.L., Shearman, M.S., 2001. Pharmacological knock-down of the presenilin 1 heterodimer by a novel γ-secretase inhibitor: Implications for presenilin biology. J. Biol. Chem. 276, 45394–45402. https://doi.org/10.1074/jbc.M103075200

Blekhman, R., Man, O., Herrmann, L., Boyko, A.R., Indap, A., Kosiol, C., Bustamante, C.D., Teshima, K.M., Przeworski, M., 2008. Natural Selection on Genes that Underlie Human Disease Susceptibility. Curr. Biol. 18, 883–889. https://doi.org/10.1016/j.cub.2008.04.074

Brunkan, A.L., Martinez, M., Walker, E.S., Goate, A.M., 2005. Presenilin endoproteolysis is an intramolecular cleavage. Mol. Cell. Neurosci. 29, 65–73. https://doi.org/10.1016/j.mcn.2004.12.012

Cai, D., Leem, J.Y., Greenfield, J.P., Wang, P., Kim, B.S., Wang, R., Lopes, K.O., Kim, S.-H., Zheng, H., Greengard, P., Sisodia, S.S., Thinakaran, G., Xu, H., 2003. Presenilin-1 regulates intracellular trafficking and cell surface delivery of beta-amyloid precursor protein. J. Biol. Chem. 278, 3446–54. https://doi.org/10.1074/jbc.M209065200

Capell, a, Steiner, H., Romig, H., Keck, S., Baader, M., Grim, M.G., Baumeister, R., Haass, C., 2000. Presenilin-1 differentially facilitates endoproteolysis of the beta-amyloid precursor protein and Notch. Nat. Cell Biol. 2, 205–11. https://doi.org/10.1038/35008626

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T.J., Higgins, D.G., Thompson, J.D., 2003. Multiple sequence alignment with the Clustal series of programs. Nucleic Acids Res. 31, 3497–3500. https://doi.org/10.1093/nar/gkg500

De Strooper, B., Beullens, M., Contreras, B., Levesque, L., Craessaerts, K., Cordell, B., Moechars, D., Bollen, M., Fraser, P., George-Hyslop, P.S., Van Leuven, F., 1997. Phosphorylation, subcellular localization, and membrane orientation of the Alzheimer's disease-associated presenilins. J. Biol. Chem. 272, 3590–3598. https://doi.org/10.1074/jbc.272.6.3590

De Strooper, B., Saftig, P., Craessaerts, K., Vanderstichele, H., Guhde, G., Annaert, W., Von Figura, K., Van Leuven, F., 1998. Deficiency of presenilin-1 inhibits the normal cleavage of amyloid precursor protein. Nature 391, 387–90. https://doi.org/10.1038/34910

Domazet-Loso, T., Tautz, D., 2003. An evolutionary analysis of orphan genes in Drosophila. Genome Res. 13, 2213–2219. https://doi.org/10.1101/gr.1311003

Drenos, F., Kirkwood, T.B.L., 2010. Selection on alleles affecting human longevity and late-life disease: The example of apolipoprotein E. PLoS One 5. https://doi.org/10.1371/journal.pone.0010022

Duff, K., Eckman, C., Zehr, C., Yu, X., Prada, C.-M., Perez-tur, J., Hutton, M., Buee, L., Harigaya, Y., Yager, D., Morgan, D., Gordon, M.N., Holcomb, L., Refolo, L., Zenk, B.,

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Hardy, J., Younkin, S., 1996. Increased amyloid-β42(43) in brains of mice expressing mutant presenilin 1. Nature 383, 710–713. https://doi.org/10.1038/383710a0

Edbauer, D., Winkler, E., Regula, J.T., Pesold, B., Steiner, H., Haass, C., 2003. Reconstitution of gamma-secretase activity. Nat. Cell Biol. 5, 486–8. https://doi.org/10.1038/ncb960

Fukumori, A., Fluhrer, R., Steiner, H., Haass, C., 2010. Three-Amino Acid Spacing of Presenilin Endoproteolysis Suggests a General Stepwise Cleavage of -Secretase-Mediated Intramembrane Proteolysis. J. Neurosci. 30, 7853–7862. https://doi.org/10.1523/JNEUROSCI.1443-10.2010

Goutte, C., Tsunozaki, M., Hale, V. a, Priess, J.R., 2002. APH-1 is a multipass membrane protein essential for the Notch signaling pathway in Caenorhabditis elegans embryos. Proc. Natl. Acad. Sci. U. S. A. 99, 775–9. https://doi.org/10.1073/pnas.022523499

Hirsh, A.E., Fraser, H.B., 2001. Protein dispensability and rate of evolution. Nature 411, 1046–1049. https://doi.org/10.1038/35082561

JACKSON, A.P., 2015. Genome evolution in trypanosomatid parasites. Parasitology 142, S40–S56. https://doi.org/10.1017/S0031182014000894

Jacobsen, H., Reinhardt, D., Brockhaus, M., Bur, D., Kocyba, C., Kurt, H., Grim, M.G., Baumeister, R., Loetscher, H., 1999. The influence of endoproteolytic processing of familial Alzheimer's disease presenilin 2 on Aβ42 amyloid peptide formation. J. Biol. Chem. 274, 35233–35239. https://doi.org/10.1074/jbc.274.49.35233

Keeling, P.J., Burger, G., Durnford, D.G., Lang, B.F., Lee, R.W., Pearlman, R.E., Roger, A.J., Gray, M.W., 2005. The tree of eukaryotes. Trends Ecol. Evol. https://doi.org/10.1016/j.tree.2005.09.005

Kelleher, R.J., Shen, J., 2017. Presenilin-1 mutations and Alzheimer's disease. Proc. Natl. Acad. Sci. 114, 629–631. https://doi.org/10.1073/pnas.1619574114

Kirkwood, T.B.L., Austad, S.N., 2000. Why do we age? Nature 408, 233–238. https://doi.org/10.1038/35041682

Kosakovsky Pond, S.L., Frost, S.D.W., 2005a. A genetic algorithm approach to detecting lineage-specific variation in selection pressure. Mol. Biol. Evol. 22, 478–485. https://doi.org/10.1093/molbev/msi031

Kosakovsky Pond, S.L., Frost, S.D.W., 2005b. Datamonkey: Rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21, 2531–2533. https://doi.org/10.1093/bioinformatics/bti320

Kosakovsky Pond, S.L., Frost, S.D.W., 2005c. Not so different after all: A comparison of methods for detecting amino acid sites under selection. Mol. Biol. Evol. 22, 1208–1222. https://doi.org/10.1093/molbev/msi105

Kulic, L., Walter, J., Multhaup, G., Teplow, D.B., Baumeister, R., Romig, H., Capell, a, Steiner, H., Haass, C., 2000. Separation of presenilin function in amyloid beta-peptide generation and endoproteolysis of Notch. Proc. Natl. Acad. Sci. U. S. A. 97, 5913–8. https://doi.org/10.1073/pnas.100049897

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

LaFerla, F.M., 2002. Calcium dyshomeostasis and intracellular signalling in alzheimer's disease. Nat. Rev. Neurosci. 3, 862–872. https://doi.org/10.1038/nrn960

Laudon, H., Hansson, E.M., Melén, K., Bergman, A., Farmery, M.R., Winblad, B., Lendahl, U., Von Heijne, G., Näslund, J., 2005. A nine-transmembrane domain topology for presenilin 1. J. Biol. Chem. 280, 35352–35360. https://doi.org/10.1074/jbc.M507217200

Lee, M.K., Slunt, H.H., Martin, L.J., Thinakaran, G., Kim, G., Gandy, S.E., Seeger, M., Koo, E., Price, D.L., Sisodia, S.S., 1996. Expression of presenilin 1 and 2 (PS1 and PS2) in human and murine tissues. J. Neurosci. 16, 7513–7525.

Levy-Lahad, E., Wasco, W., Poorkaj, P., Romano, D.M., Oshima, J., Pettingell, W.H., Yu, C.E., Jondro, P.D., Schmidt, S.D., Wang, K., et al., 1995. Candidate gene for the chromosome 1 familial Alzheimer's disease locus. Science (80-.). https://doi.org/10.1126/science.7638622

Lichtarge, O., Bourne, H.R., Cohen, F.E., 1996. An Evolutionary Trace Method Defines Binding Surfaces Common to Protein Families. J. Mol. Biol. 257, 342–358. https://doi.org/10.1006/jmbi.1996.0167

Lua, R.C., Lichtarge, O., 2010. PyETV: A PyMOL evolutionary trace viewer to analyze functional site predictions in protein complexes. Bioinformatics 26, 2981–2982. https://doi.org/10.1093/bioinformatics/btq566

Marchler-Bauer, A., Derbyshire, M.K., Gonzales, N.R., Lu, S., Chitsaz, F., Geer, L.Y., Geer, R.C., He, J., Gwadz, M., Hurwitz, D.I., Lanczycki, C.J., Lu, F., Marchler, G.H., Song, J.S., Thanki, N., Wang, Z., Yamashita, R.A., Zhang, D., Zheng, C., Bryant, S.H., 2015. CDD: NCBI's conserved domain database. Nucleic Acids Res. 43, D222–D226. https://doi.org/10.1093/nar/gku1221

McCarthy, J. V., Twomey, C., Wujek, P., 2009. Presenilin-dependent regulated intramembrane proteolysis and γ-secretase activity. Cell. Mol. Life Sci. https://doi.org/10.1007/s00018-009-8435-9

Menzies, F.M., Fleming, A., Rubinsztein, D.C., 2015. Compromised autophagy and neurodegenerative diseases. Nat. Rev. Neurosci. 16, 345–357. https://doi.org/10.1038/nrn3961

Meyer, D.Z., Avery, L.M., 2009. Excel as a Qualitative Data Analysis Tool. Field methods 21, 91–112. https://doi.org/10.1177/1525822X08323985

Murray, M.M., Jones, C.J., Clark, R.F., 2000. The presenilin gene is very highly conserved throughout evolution. Neurobiol. Aging 21, 48. https://doi.org/10.1016/S0197-4580(00)82893-6

Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol. Biol. Evol. 3, 418–26.

Nei, M., Kumar, S., 2000. Molecular Evolutionand Phylogenetics, Archives of virology.

Nixon, R.A., Yang, D.S., 2011. Autophagy failure in Alzheimer's disease-locating the primary defect. Neurobiol. Dis. https://doi.org/10.1016/j.nbd.2011.01.021

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Ohta, K., Mizuno, A., Ueda, M., Li, S., Suzuki, Y., Hida, Y., Hayakawa-Yano, Y., Itoh, M., Ohta, E., Kobori, M., Nakagawa, T., 2010. Autophagy impairment stimulates PS1 expression and γ-secretase activity. Autophagy. https://doi.org/10.4161/auto.6.3.11228

Otto, G.P., Sharma, D., Williams, R.S.B., 2016. Non-Catalytic Roles of Presenilin Throughout Evolution. J. Alzheimer's Dis. https://doi.org/10.3233/JAD-150940

Parfrey, L.W., Barbero, E., Lasser, E., Dunthorn, M., Bhattacharya, D., Patterson, D.J., Katz, L.A., 2006. Evaluating support for the current classification of eukaryotic diversity. PLoS Genet. 2, 2062–2073. https://doi.org/10.1371/journal.pgen.0020220

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera—A Visualization System for Exploratory Research and Analysis. J Comput Chem 25, 1605–1612. https://doi.org/10.1002/jcc.20084

Podlisny, M.B., Citron, M., Amarante, P., Sherrington, R., Xia, W., Zhang, J., Diehl, T., Levesque, G., Fraser, P., Haass, C., Koo, E.H.M., Seubert, P., St. George-Hyslop, P., Teplow, D.B., Selkoe, D.J., 1997. Presenilin Proteins Undergo Heterogeneous Endoproteolysis between Thr291and Ala299and Occur as Stable N- and C-Terminal Fragments in Normal and Alzheimer Brain Tissue. Neurobiol. Dis. 3, 325–337. https://doi.org/10.1006/nbdi.1997.0129

Popugaeva, E., Bezprozvanny, I., 2013. Role of endoplasmic reticulum Ca2+ signaling in the pathogenesis of Alzheimer disease. Front. Mol. Neurosci. 6. https://doi.org/10.3389/fnmol.2013.00029

Pundir, S., Martin, M.J., O'Donovan, C., 2017. UniProt protein knowledgebase, in: Methods in Molecular Biology. pp. 41–55. https://doi.org/10.1007/978-1-4939-6783-4_2

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425. https://doi.org/citeulike-article-id:93683

Sakharkar, K.R., Chow, V.T.K., 2005. Strategies for genome reduction in microbial genomes. Genome Inform. 16, 69–75.

Sato, C., Morohashi, Y., Tomita, T., Iwatsubo, T., 2006. Structure of the catalytic pore of gamma-secretase probed by the accessibility of substituted cysteines. J. Neurosci. 26, 12081–8. https://doi.org/10.1523/JNEUROSCI.3614-06.2006

Sato, C., Takagi, S., Tomita, T., Iwatsubo, T., 2008. The C-terminal PAL motif and transmembrane domain 9 of presenilin 1 are involved in the formation of the catalytic pore of the gamma-secretase. J. Neurosci. 28, 6264–71. https://doi.org/10.1523/JNEUROSCI.1163-08.2008

Scheper, W., Zwart, R., Baas, F., 2004. Rab6 membrane association is dependent of Presenilin 1 and cellular phosphorylation events. Mol. Brain Res. 122, 17–23. https://doi.org/10.1016/j.molbrainres.2003.11.013

Shen, J., Kelleher, R.J., 2007. The presenilin hypothesis of Alzheimer's disease: evidence for a loss-of-function pathogenic mechanism. Proc. Natl. Acad. Sci. U. S. A. 104, 403–409. https://doi.org/10.1073/pnas.0608332104

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Slamovits, C.H., 2013. Extreme genome reduction in microbial parasites, Comparative Genomics in Neglected Human Parasites.

Spasic, D., Tolia, A., Dillen, K., Baert, V., De Strooper, B., Vrijens, S., Annaert, W., 2006. Presenilin-1 maintains a nine-transmembrane topology throughout the secretory pathway. J. Biol. Chem. 281, 26569–26577. https://doi.org/10.1074/jbc.M600592200

Steiner, H., Romig, H., Grim, M.G., Philipp, U., Pesold, B., Citron, M., Baumeister, R., Haass, C., 1999. The biological and pathological function of the presenilin-1 Δexon 9 mutation is independent of its defect to undergo proteolytic processing. J. Biol. Chem. 274, 7615–7618. https://doi.org/10.1074/jbc.274.12.7615

Suga, K., Tomiyama, T., Mori, H., Akagawa, K., 2004. Syntaxin 5 interacts with presenilin holoproteins, but not with their N- or C-terminal fragments, and affects beta-amyloid peptide production. Biochem. J. 381, 619–628. https://doi.org/10.1042/BJ20040618

Sun, L., Zhou, R., Yang, G., Shi, Y., 2017. Analysis of 138 pathogenic mutations in presenilin-1 on the in vitro production of Aβ42 and Aβ40 peptides by γ-secretase. Proc. Natl. Acad. Sci. 114, E476–E485. https://doi.org/10.1073/pnas.1618657114

Supnet, C., Bezprozvanny, I., 2010. The dysregulation of intracellular calcium in Alzheimer disease. Cell Calcium. https://doi.org/10.1016/j.ceca.2009.12.014

Tajima, F., 1993. Simple methods for testing the molecular evolutionary clock hypothesis. Genetics 135, 599–607. https://doi.org/citeulike-article-id:785176

Takasugi, N., Tomita, T., Hayashi, I., Tsuruoka, M., Niimura, M., Takahashi, Y., Thinakaran, G., Iwatsubo, T., 2003. The role of presenilin cofactors in the γ-secretase complex. Nature 422, 438–441. https://doi.org/10.1038/nature01506

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729. https://doi.org/10.1093/molbev/mst197

Thinakaran, G., Borchelt, D.R., Lee, M.K., Slunt, H.H., Spitzer, L., Kim, G., Ratovitsky, T., Davenport, F., Nordstedt, C., Seeger, M., Hardy, J., Levey, A.I., Gandy, S.E., Jenkins, N.A., Copeland, N.G., Price, D.L., Sisodia, S.S., 1996. Endoproteolysis of presenilin 1 and accumulation of processed derivatives in vivo. Neuron 17, 181–190. https://doi.org/10.1016/S0896-6273(00)80291-3

Tolia, A., Chávez-Gutiérrez, L., De Strooper, B., 2006. Contribution of presenilin transmembrane domains 6 and 7 to a water-containing cavity in the □??-secretase complex. J. Biol. Chem. 281, 27633–27642. https://doi.org/10.1074/jbc.M604997200

Tolia, A., Horré, K., De Strooper, B., 2008. Transmembrane domain 9 of presenilin determines the dynamic conformation of the catalytic site of γ-secretase. J. Biol. Chem. 283, 19793–19803. https://doi.org/10.1074/jbc.M802461200

Wall, D.P., Fraser, H.B., Hirsh, A.E., 2003. Detecting putative orthologs. Bioinformatics 19, 1710–1711. https://doi.org/10.1093/bioinformatics/btg213

Walter, J., Capell, A., Grünberg, J., Pesold, B., Schindzielorz, A., Prior, R., Podlisny, M.B.,

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Fraser, P., Hyslop, P.S., Selkoe, D.J., Haass, C., 1996. The Alzheimer's disease-associated presenilins are differentially phosphorylated proteins located predominantly within the endoplasmic reticulum. Mol. Med. 2, 673–91.

Wang, J., Beher, D., Nyborg, A.C., Shearman, M.S., Golde, T.E., Goate, A., 2006. C-terminal PAL motif of presenilin and presenilin homologues required for normal active site conformation. J. Neurochem. 96, 218–227. https://doi.org/10.1111/j.1471-4159.2005.03548.x

Wang, R., Tang, P., Wang, P., Boissy, R.E., Zheng, H., 2006. Regulation of tyrosinase trafficking and processing by presenilins: partial loss of function by familial Alzheimer's disease mutation. Proc. Natl. Acad. Sci. U. S. A. 103, 353–8. https://doi.org/10.1073/pnas.0509822102

Watanabe, H., Shen, J., 2017. Dominant negative mechanism of Presenilin-1 mutations in FAD. Proc. Natl. Acad. Sci. U. S. A. 201717180. https://doi.org/10.1073/pnas.1717180114

Watanabe, N., Takagi, S., Tominaga, A., Tomita, T., Iwatsubo, T., 2010. Functional analysis of the transmembrane domains of presenilin 1: Participation of transmembrane domains 2 and 6 in the formation of initial substrate-binding site of γ-secretase. J. Biol. Chem. 285, 19738–19746. https://doi.org/10.1074/jbc.M110.101287

Wilson, A.C., Carlson, S.S., White, T.J., 1977. Biochemical Evolution. Annu. Rev. Biochem. 46, 573–639. https://doi.org/10.1146/annurev.bi.46.070177.003041

Wolf, Y., Novichkov, P., Karev, G., Koonin, E., Lipman, D., 2009. The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. Proc. Natl. Acad. Sci. 106, 7273–7280. https://doi.org/10.1073/pnas.0901808106

Zhang, S., Zhang, M., Cai, F., Song, W., 2013. Biological function of Presenilin and its role in AD pathogenesis. Transl. Neurodegener. 2, 15. https://doi.org/10.1186/2047-9158-2-15

Zhang, Y., Thompson, R., Zhang, H., Xu, H., 2011. APP processing in Alzheimer's disease. Mol. Brain 4, 3. https://doi.org/10.1186/1756-6606-4-3

Zhang, Z., Nadeau, P., Song, W., Donoviel, D., Yuan, M., Bernstein, a, Yankner, B. a, 2000. Presenilins are required for gamma-secretase cleavage of beta-APP and transmembrane cleavage of Notch-1. Nat. Cell Biol. 2, 463–465. https://doi.org/10.1038/35017108

Figure Legends

FIG.1. Distribution of Gamma secretase proteins in 47 species belonging to major eukaryotic groups from human to protists. Solid and hollow circles represent the presence and absence of the respective proteins respectively.

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIG.2. Evolutionary Tree of Gamma Secretase proteins. Phylogenetic tree depicting the phylogenetic history of four Gamma secretase proteins. The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum of branch length = 39.25074069 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the p-distance method (Nei and Kumar, 2000) and are in the units of the number of amino acid differences per site. The analysis involved 153 amino acid sequences. All ambiguous positions were removed for each sequence pair. There were a total of 2897 positions in the final dataset. Evolutionary analyses were conducted in MEGA 6 (Tamura et al., 2013).

FIG.3. Selection analysis of Presenilin proteins in Vertebrates. (A and B) For both Presenilin-1 and Presenelin-2, the difference between the nonsynonymous and synonymous differences per site between different species are shown. Analyses were conducted using the modified Nei-Gojobori (assumed transition/transversion bias = 2) model. A negative value means the sequence is under purifying selection. (C) GA- Branch based analysis of lineage specific selective pressure in Presenilin genes. A cladogram is shown with maximum-likelihood estimates of lineage specific *d*N/*d*S during vertebrate Presenilin evolution. Percentages for branch classes in the legend reflect the proportion of total tree length (measured in expected substitutions per site per unit time) evolving under the corresponding value of *dN*/*dS* with letters A, B, C, D and E represents each branch class. (D) Sliding window display of nonsynonymous and synonymous differences for each site (codon) in vertebrate Presenilins. SLAC method from datamonkey server is employed for this analysis. (E) Tajima's relative rate test to depict the difference in rate of evolution of two Presenilin paralogs in vertebrates with Presenilin protein sequence of Saccoglossus kowalevskii as outgroup.

FIG.4. Evolutionary traces analysis of Presenilin-1 protein**. (A)** A heatmap of evolutionary traces found in metazoan Presenilin. The color shades from dark brown to white correspond to most conserved to least conserved residues. The TA is carried out in by using 15 species, covering major metazoan groups from primates to choanoflagellates. (B) A Schematic heatmap based representation of Presenilin-1 protein across the cell membrane. The color shades corresponds to same conservation profile as mentioned in A and B. The scheme of arrangement of Presenilin across the membrane is taken from the work of Zhang et al. (Zhang et al., 2013). (C) Pi graph showing the distribution of highly conserved residues –as depicted by the ETA rank profile- in the transmembrane and non-transmembrane part of the Presenilin protein. (D) List of site directed mutations which have been carried out in Preselin-1 gene. Heatmap shown in column 2 depicts the conservation profile of mutated sites with color shades from dark brown to white correspond to most conserved to least conserved residues. Phenotypic effects due to these mutations are given in columns 5. Φ= Loss of interaction with GFAP, Ω= Abolished protease activity, Ψ= increased protease activity, □= alters gamma secretase specificity, □= Reduced Notch processing, Ю= Abolished PKA signaling. Δ= Abolished caspase cleavage.

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

FIG.5. Evolution of Presenilin and Alzheimer disease. (A) List of Presenilin-1 mutations leading to AD along with conservation profile of each mutated site in the form of heatmap. The color shades from dark brown to white correspond to most conserved to least conserved residues. (B) Pi graph showing the extent of conservation of evolutionary traces mutation in whom leads to AD. (C and D) ETA based structural analysis of Presenilin protein by using Cryo-EM structure of human Presenilin-1 (5NF2) as template. Both ribbon (C) and surface (D) cartoons are shown with the red spots depicting the location of highly conserved residues (traces) susceptible for AD. ETA based ranks are generated by comparing protein sequences of 15 species, covering major metazoan groups from primates to choanoflagellates.

Figure 1

	Group	Species	Presenilin	APH-1	Nicastrin	PEN-2
1		Homo sapiens	●	●	●	●
2		Pan troglodytes	●	●	●	●
3		Bos taurus	●	●	●	●
4		Mus musculus	●	●	●	●
5		Xenopus tropicalis	●	●	●	●
6		Anolis carolinensis	●	●	●	●
7		Daniorerio	●	●	●	●
8		Takifugu rubripes	●	●	●	●
9		Branchiostoma floridae	●	●	●	●
10		Saccoglossus kowalevskii	●	●	●	●
11		Strongylocentrotus purpuratus	●	●	●	●
12		Caenorhabditis elegans	●	●	●	●
13		Drosophila melanogaster	●	●	●	●
14		Hydra vulgaris	●	●	●	●
15		Amphimedon queenslandica	●	●	●	●
16	**Opisthokonts**	Monosiga brevicollis	●	●	●	○
17		Chlamydomonas reinhardtii	●	●	●	●
18	**Archeaplastida**	Populus trichocarpa	●	○	●	●
19	**Red Algea**	Chondrus crispus	●	○	●	○
20		Trypanosoma brucei	●	○	○	○
21		Naegleria gruberi	○	○	○	○
22		Trichomonas vaginalis	●	○	○	●
23	**Excavata**	Rozella allomycis	●	○	○	○
24		Fonticula alba	●	●	●	●
25		Batrachochytrium salamandrivorans	●	●	●	○
26	**Fungi**	Capsaspora owczarzaki	●	●	●	●
27		Neospora caninum Liverpool	●	○	○	●
28		Toxoplasma gondii	●	○	○	○
29		Hammondia hammondi	●	○	○	●
30		Vitrella brassicaformis	○	○	○	○
31	**Alveolates**	Tetrahymena	○	○	○	○
32		Blastocystis hominis	●	○	○	○
33		Aphanomyces invadans	●	●	●	●
34		Saprolegnia diclina	●	○	●	●
35		Thalassiosira pseudonana	●	○	●	○
36		Phaeodactylum tricornutum	●	●	○	○
37		Ectocarpus siliculosus	●	●	●	○
38		Aureococcus anophagefferens	●	○	○	●
39	**Stramenopiles**	Nannochloropsis gaditana	●	○	●	○
40		Bigelowiella natans	●	○	○	○
41	**Rhizaria**	Reticulomyxa filose	●	○	●	○
42		Entamoeba histolytica	●	●	○	●
43		Polysphondylium pallidum	●	●	●	●
44	**Amoebozoa**	Dictyostelium discoideum	●	●	●	●
45		Guillardia theta	●	●	●	○
46		Emiliania huxleyi	●	●	○	○
47	**Haptophytes and Cryptophytes**	Chrysochromulina	●	○	●	●

bioRxiv preprint doi: https://doi.org/10.1101/394270; this version posted August 20, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

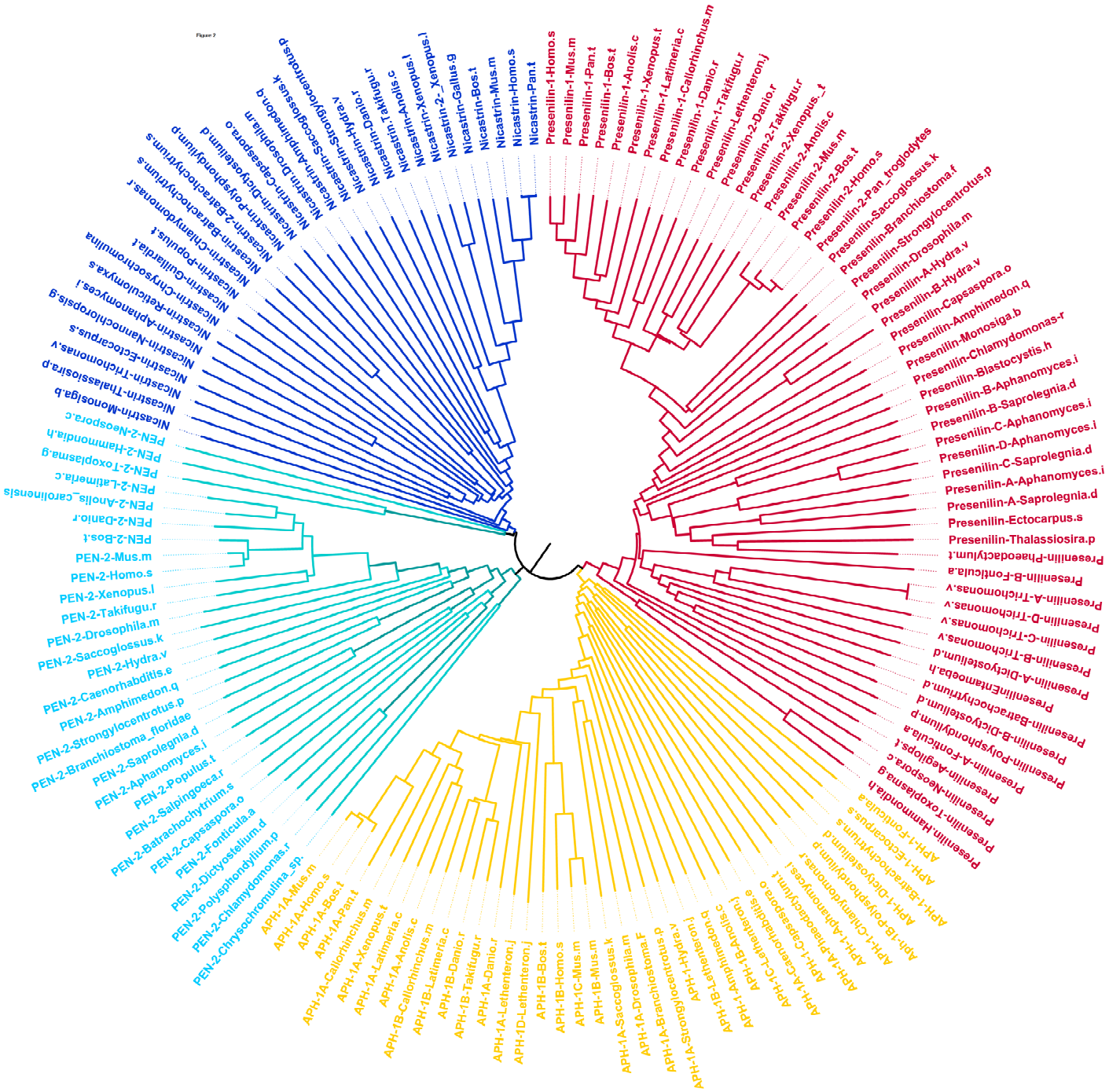


Figure 3

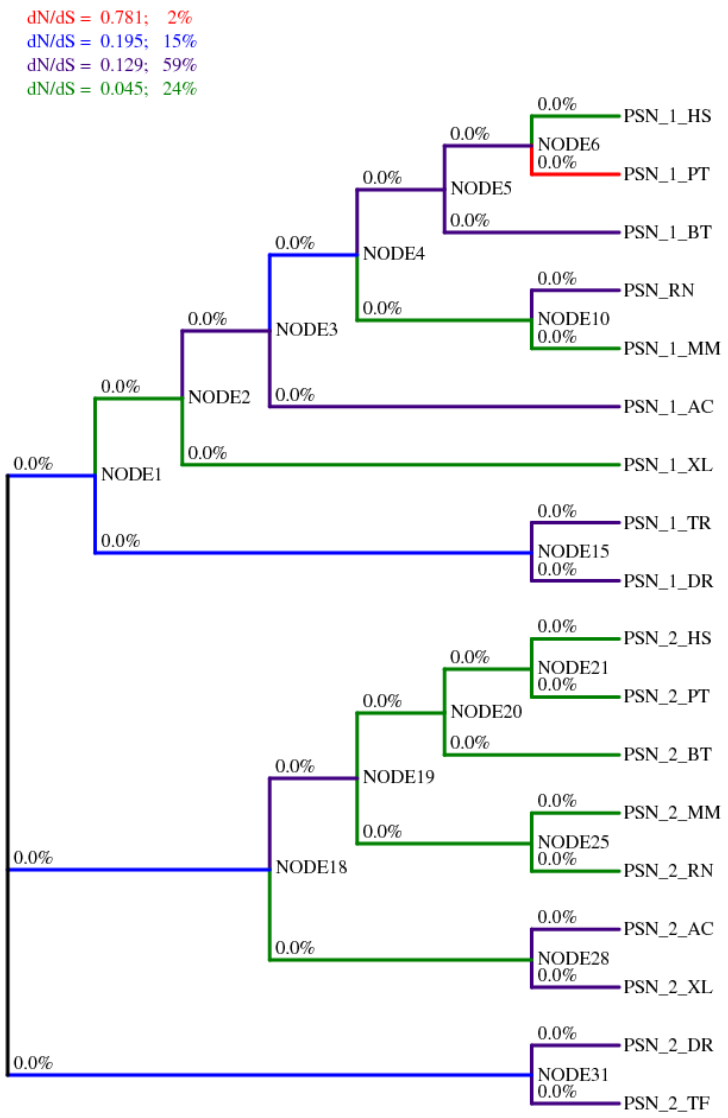
A

PSN-1	Hs	Bt	Pt	Rn	Mm	Ac	Xl	Tr	Dr
Hs									
Bt	-0.2								
Pt	-0	-0.2							
Rn	-0.3	-0.3	-0.3						
Mm	-0.3	-0.3	-0.3	-0.1					
Ac	-0.4	-0.4	-0.4	-0.5	-0.5				
Xl	-0.5	-0.5	-0.5	-0.5	-0.5	-0.5			
Tr	-0.4	-0.4	-0.4	-0.4	-0.4	-0.5	-0.4		
Dr	-0.5	-0.4	-0.5	-0.4	-0.4	-0.5	-0.5	-0.4	

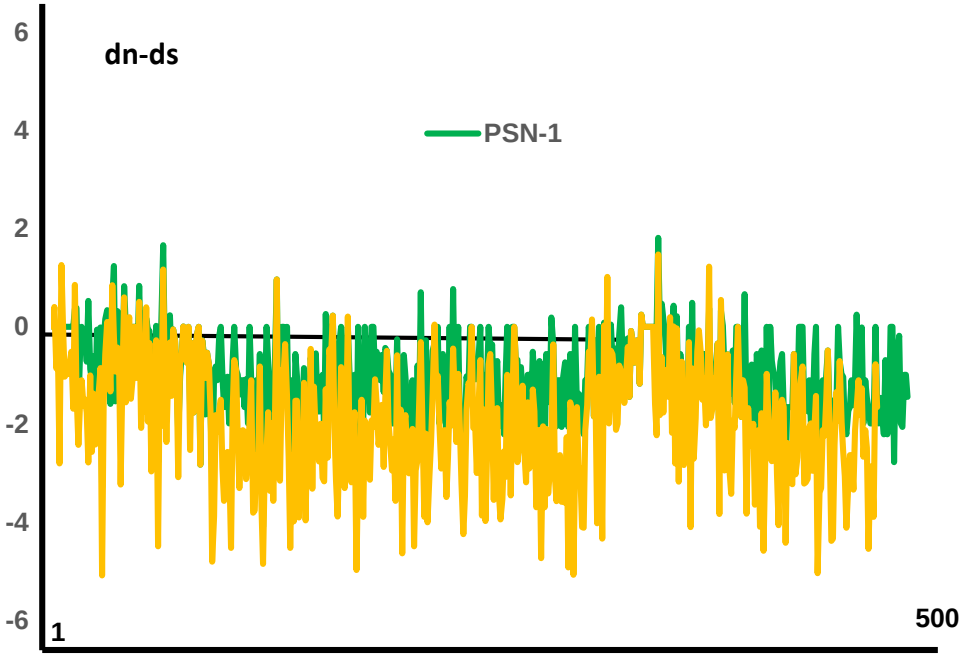
B

PSN-2	Hs	Bt	Pt	Rn	Mm	Ac	Xl	Tr	Dr
Hs									
Bt	-0.02								
Pt	-0.28	-0.27							
Rn	-0.29	-0.29	-0.2						
Mm	-0.22	-0.22	-0.3	-0.3					
Ac	-0.5	-0.5	-0.5	-0.5	-0.5				
Xl	-0.52	-0.51	-0.5	-0.5	-0.6	-0.4			
Tr	-0.37	-0.37	-0.4	-0.4	-0.3	-0.4	-0.5		
Dr	-0.43	-0.43	-0.4	-0.4	-0.4	-0.4	-0.5	-0.4	

C



D



E

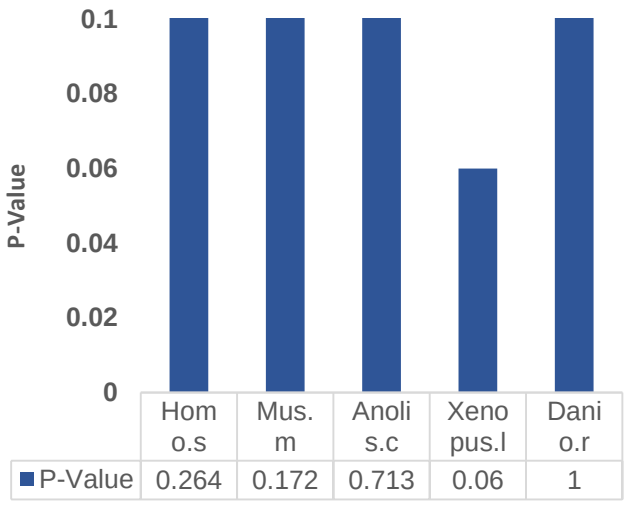
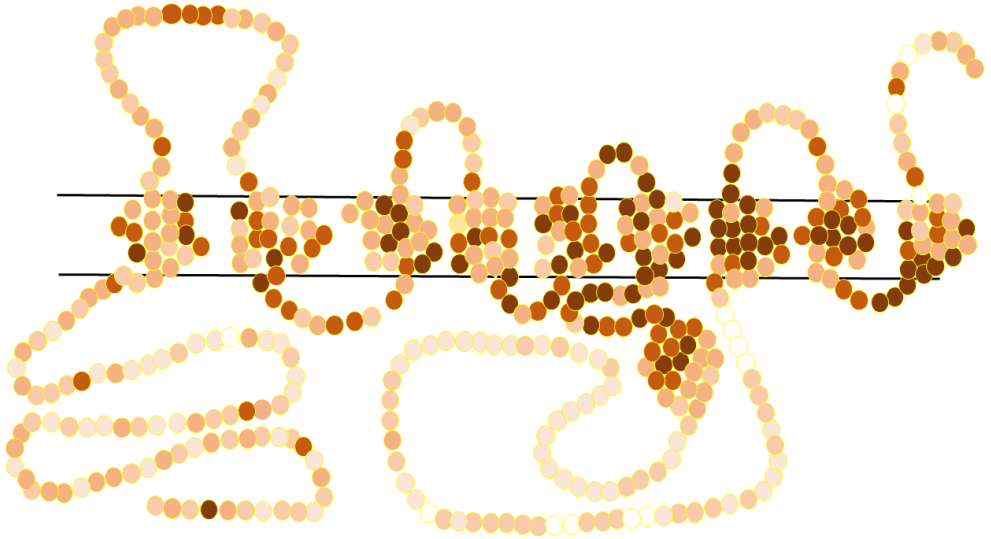


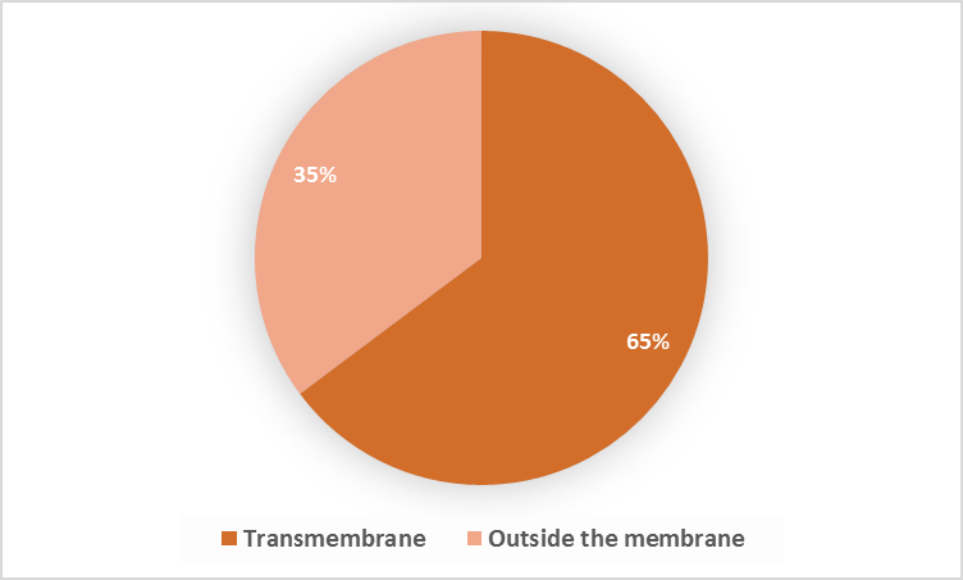
Figure 4

A

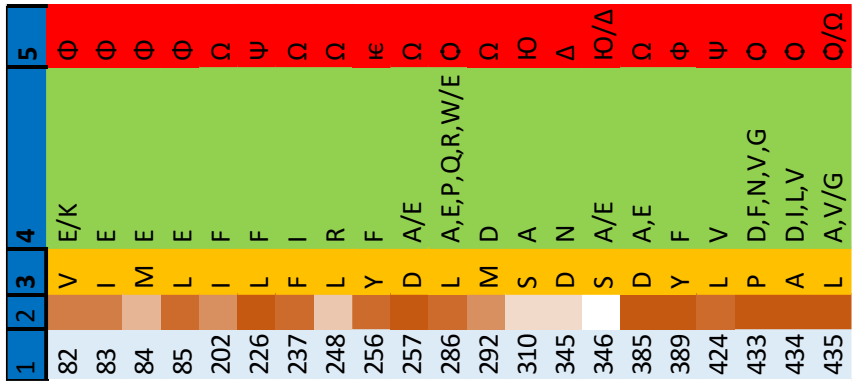
B



C



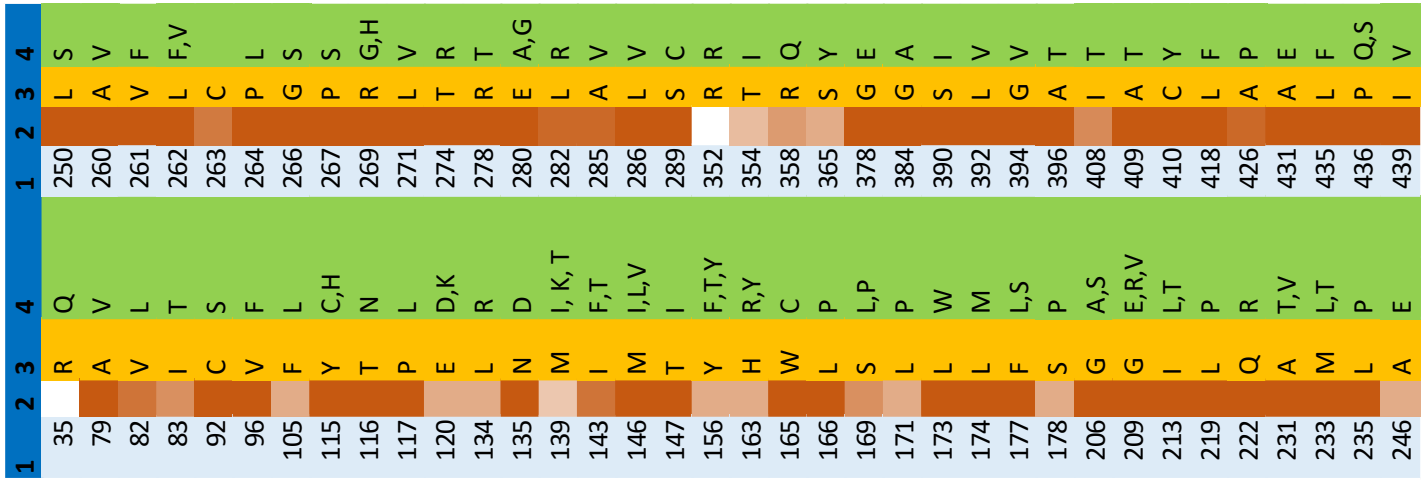
D



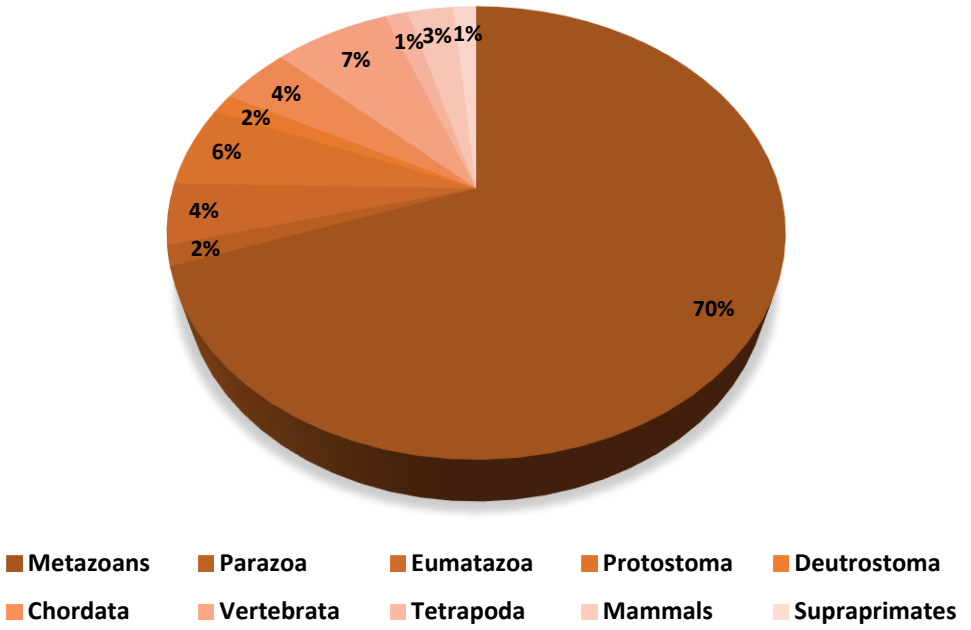
1	2	3	4	5
82		V	E/K	Φ
83		I	E	Φ
84		M	E	Φ
85		L	E	Φ
202		I	F	Ω
226		L	F	Ψ
237		F	I	Ω
248		L	R	Ω
256		Y	F	Ж
257		D	A/E	Ω
286		L	A,E,P,Q,R,W/E	O
292		M	D	Ω
310		S	A	Ю
345		D	N	Δ
346		S	A/E	Ю/Δ
385		D	A,E	Ω
389		Y	F	Φ
424		L	V	Ψ
433		P	D,F,N,V,G	O
434		A	D,I,L,V	O
435		L	A,V/G	O/Ω

Figure 5

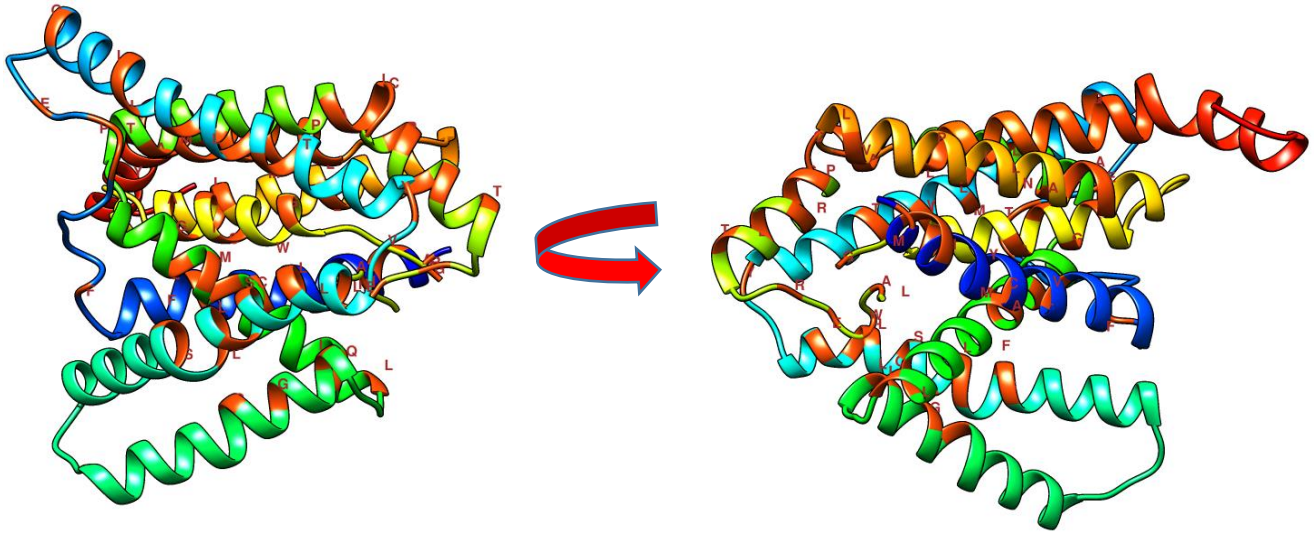
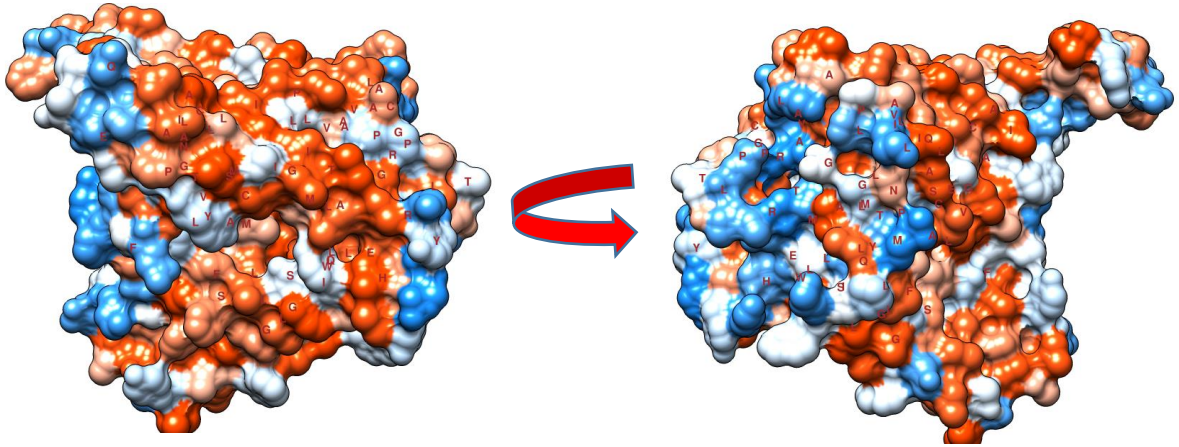
A



B



C



D