## In silico molecular modelling of MTHFR protein across eukaryotic species

ASISH KUMAR SWAIN

A dissertation submitted for the partial fulfilment of the BS-MS dual degree in Biological sciences.



Indian Institute of Science Education and Research, Mohali April 2020

### Certificate of Examination

This is to certify that the dissertation titled **"In silico molecular modelling of MTHFR protein across eukaryotic species"** submitted by **Asish kumar Swain** (Reg. No. MS15174) for the fulfilment of **BS-MS dual degree programme** of the Institute, has been examined by the thesis committee duly appointed by the Institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr.Shashi Bhusan Pandit (Committee member) Dr. Anand Kumar Bachhawat (Committee member) Dr. Monika Sharma (Supervisor)

### **Declaration**

The work presented in this dissertation has been carried out by me under the guidance of Dr. Monika Sharma at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography

> Asish Kumar Swain June, 2020

In my capacity as the supervisor of the candidate's project work, I certify that the above statements by the candidate are true to the best of my knowledge.

Dr. Monika Sharma

(supervisor)

June, 2020

### Acknowledgements

Foremost, I would like to express my deep and sincere gratitude to my supervisor Dr. Monika Sharma for her immense support, kindness, and encouragement during my thesis research. I am thankful to her for giving me the opportunity to work under her guidance.

I would like to thank Muskan Bhatia and Milind kale for their insightful comments during my research work.

Finally, I would like to thank my parents, friends who are always with me in the first place.

## List of Tables

Table 1:Summary of Ramchandran plot of models before and after refinement of models20
Table 2: Physiochemical properties of human, Mus musculus, Gallus gallus, Danio rerio,
Acanthaster Planci, Arabidopsis thaliana, C.elegans MTHFR sequence where EC=Extinction
coefficient, II=Instability Index, GRAVY= Grand Average of hydropathy, AI=Aliphatic
Index
Table 3:Structural validation of all models, details about the tools already explained27
Table 4: Mutational impact of common mutations occur in human MTHFR where these
values represent the $\Delta\Delta G$ value obtained from FoldX
Table 5:Functional consequences of nsSNPs in human MTHFR
Table 6:Percentage of secondary structural elements in models    40

## List of figures

Figure 1:Domain organization of MTHFR orthologs	18
Figure 2:Surface contact to 4 crucial residues(E463, T464, T481, Q485) of SAM in the	
predicted docking site of human MTHFR protein	21
Figure 3.: Phylogenetic tree of MTHFR proteins all across species based on Jones-Taylor-	
Thorton mode, maximum likelihood tree with 1000 bootstrap replication.Bootstrap value	
above the branch suggest confidence/robustness of branch to be form.clades were labelled	1
with boxes ,blue arrow marked sequences are considered to 3D modelling	23
Figure 4Ramchandran plot of – (a) Mus musculus,(b) Gallus gallus,(c) Danio rerio,(d)	
Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans	26
Figure 5Sequence alignment of Mus musculus, Gallus gallus, Danio rerio, Acanthaster	
planci, Arabidopsis thaliana, C.elegans	36
Figure 6: 2D Interaction diagram of residues to FAD – (a) Mus musculus,(b) Gallus	
gallus,(c) Danio rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans	38
Figure7: 2D interaction diagram of residues to SAM –(a) Mus musculus,(b) Gallus gallus,	(c)
Danio rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans	36
Figure.8.Sequence alignment of catalytic domain of all models. Residues inside red box	
interacting to FAD and residues inside blue box are the interface residues2	7
Figure.9.Sequence alignment of regulatory domain on all models. Residues inside red box	are
interacting to SAM and residues inside blue box are the interface residues2	27
Figure.10.3D interaction digram of FAD in (a) Mus musculus,(b) Gallus gallus,(c) Danio	
rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans	37
Figure.11.3D interaction digram of SAM in (a) Mus musculus,(b) Gallus gallus,(c) Danio	
rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans	37

### Abbreviations

- MTHFR Methylene tetra hydro folate reductase
- FAD Flavin Adenine Dinucleotide
- SAM S-Adenosyl Methionine
- SAH S-Adenosyl homocysteine
- **SNP** Single Nucleotide Polymorphisim

## Abstract

Methylenetetrahydrofolate reductase (MTHFR) is a key regulatory enzyme involves in folate and methionine cycle which are important for the biosynthesis of nucleotide, lipid, and amino acids. Deficiency and mutations in MTHFR lead to hyperhomocysteinemia, vascular diseases, neural tube diseases, diabetes, and various cancer diseases in humans. In other eukaryotes like in plants, it has role in photorespiration, germination, root development, and lignification. In mice, MTHFR accelerates aggregation of unmodified keratin in mice hair, in this way MTHFR retains its core function in various eukaryotes. To study the various pathophysiological role of MTHFR in various species, complete 3D structures of different diverged species were modeled using template-based modelling. As loops play a major role in protein, problematic loops were refined and validated using several tools. Impact of experimentally determined mutations analysed on these models, docking of FAD and SAM to get insight into possible binding modes and how they interact with the enzyme. As identification of SNPs in the human genome growing nowadays, damaging SNPs in human MTHFR gene were analysed using SIFT, PROVEAN, PolyPhen2, Mutpred. Total of 14 SNPs were identified which affect the structure and dynamics of human MTHFR protein. As these mutations occur in the course of evolution these deleterious SNPs may having impact on other eukaryotes also.

### Contents

#### Table of Contents

Certificate of Examination
Declaration
Acknowledgements
List of Tables
List of figures
Abbreviations
Introduction17
1.1.Role of MTHFR across eukaryotic phylogeny17
1.2.Structure and regulation of MTHFR17
Methods and computational details
2.1. Sequence and structure retrieval19
2.2.Sequence alignment and phylogenetic analysis19
2.3. 3D model building, refinement and validation19
2.4.Mutational impact on models20
2.5.Cofactor addition and active site identification21
Results
3.1.Sequence Alignment And Phylogenetic Analysis
3.2 Structure Based Sequence Alignment24
3.3.Physiochemical Parameters Of The Sequences24
3.4.Validation Of Models25
3.5.Structural Analysis Of Models27
3.6. Deleterious nsSNPs in human MTHFR gene and their impact on other species31
Conclusion
Biblography
Appendix

## **Chapter 1**

### Introduction

#### 1.1.Role of MTHFR across eukaryotic phylogeny

Folate is one of the B complex vitamin having crucial role in eukaryotes from DNA synthesis to tissue repair. Methylenetetrahydrofolate reductase (MTHFR) is a key regulatory enzyme involves in folate and methionine cycle which are important for biosynthesis of nucleotide, lipid and aminoacids. MTHFR which irreversibly converts 5,10 methylenetetrahydrofolate to 5 methyltetrahydrofolate which requires FAD as cofactor and NADPH as electron donor that remethylate homocysteine to methionine in humans[1]. Deficiency or polymorphism in MTHFR leads to hyperhomocysteinemia, vascular diseases, neural tube diseases, diabetes and various cancer diseases in humans[2][3]. In plants, MTHFR plays a major role in photorespiration, embryogenesis, germination, hypocotyl and root devlopment[4][5][6][7]. Alteration in MTHFR in plants leads to hypomethylation of DNA, in recent studies shows Cosilencing of MTHFR and COMT(Caffeicacid O-methyltransferase) in plants having a significant impact on lignification[8]. Mutation in homocysteine metabolism generating N-Hcy(N-Homocysteine) which accelerates aggregation of unmodified keratins in mouse hair. Same impact also observed in birds feather generation and also in humans[9].

#### **1.2.Structure and regulation of MTHFR**

Mammalian MTHFR is a homodimer of chains, each chain having an N-terminal catalytic domain and C-terminal regulatory domain connected by a linker sequence. Each chain has an N-terminal serine-rich extension region which down regulate the MTHFR activity by protecting thermally unstable SAM[2]. Catalytic domain having binding site for the prosthetic group(FAD), substrates(NADPH & MTHF) and regulatory domain bind to S-AdenosylMethionine (SAM/AdoMet) that allosterically inhibit the enzyme activity. Inhibitory action on SAM can be switch by binding of S-adenosylhomocysteine (SAH) to the regulatory domain[10]. Prokaryotic MTHFR exists as a homotetramer having only catalytic domain which is conserved all across species. E.coli MTHFR have only catalytic domain [Fig.1] consist of 8α/8β TIM barrel which binds to cofactors and NADH as electron donor. Enzyme regulatory mechanisms like SAM binding



Figure 1:Domain organization of MTHFR orthologs

and phosphorylation absent in prokaryotes[11]. Eukaryote MTHFR catalytic domain comprises of 8α/8β Tim barrel as prokaryotes and the C-terminal regulatory domain consist of 2 five stranded βsheets, there is no direct contact between two domains but the linker region having contacts to both catalytic and regulatory domain. Although plant MTHFR sequence have high homology to other eukaryotes but the enzymatic reaction is different. In mammalian cell high NADPH: NADP ratio results MTHFR reaction irreversible which is inhibited by SAM but in plant cytosolic cells, low NADH: NAD which doesn't require inhibition by SAM this is how plant MTHFR is NADH dependent and SAM insensitive[12].

As MTHFR plays major role in humans, deficiency of folate or mutation in MTHFR can cause severe impact as discussed above. C.elegans, zebrafish and mouse are an appropriate model to study these, as in case of C.elegans methionine synthase have 64% identity to humans[13]. To study the various pathophysiological role of MTHFR in various species requires high resolution of complete structure. Experimental methods such as XRAY crystallography, NMR spectroscopy, electron microscopy was expensive and took years to perform. Modeling eukaryotic MTHFR is incomplete using bacterial MTHFR as a template cause it only homolog to the catalytic domain. Recently almost complete 2.5Å resolution human MTHFR crystal structure resolved in Froese, D. Sean et al[2].To obtain a rational 3D structure of MTHFR across different species comparative protein structure modeling was employed. Predicted homology models can be used for modeling substrate specificity, mutation effects, ligand designing, antigenic epitope binding prediction and more[14][15][16]. Resulting models were refined and validated using different tools. Docking of FAD and SAM also provide insight into possible binding modes and how they interact with enzymes.

## Chapter 2

## Methods and computational details

#### 2.1. Sequence and structure retrieval

For phylogenetic analysis MTHFR aminoacid sequence of 23 different diverged species obtained from UniProtKB in FASTA format[17]. Human and yeast MTHFR crystallographic structure obtained from protein data bank(PDB:id-6FCX,6FNU)[18].

#### 2.2. Sequence alignment and phylogenetic analysis

Retrieved sequences were aligned using CLUSTAL Omega saved in phylip format. Clustal omega which generates Multiple sequence alignment using seeded guided trees and HMM profile-profile techniques[19]. Sequence similarity between retrieved sequence and human MTHFR analyzed using EMBOSS Needle[20]. Conserved and consensus positions in the sequence analyzed in CLC sequence viewer 7[21]. To estimating evolutionary relationships between sequences phylogenetic and molecular evolutionary analyses were conducted by MEGA Version7 using maximum likelihood method based on Jones-Taylor-Thorton [JTT] mode with pairwise deletions[22]. Each node of the estimated tree evaluated by the bootstrap method with 1000 replications. Physicochemical properties such as molecular weight, amino acid composition, theoretical pI, extinction coefficient, estimated half-life, aliphatic index, instability index, Grand average of hydropathicity[GRAVY] of the selected sequence computed using protparam tool[23].

#### 2.3. 3D model building, refinement and validation

The three-dimensional model of Mus musculus, Arabidopsis analyzing, Danio rerio(zebrafish), Gallus gallus, Acanthaster planci, C.elegans with sequence similarity higher than 35% indicating strong structural conservation was modeled by homology modeling alignment of these sequences shown in figure.5. Modeller9.22 program performed to build 3D structure of the sequences. Modeller can perform comparative protein structure modelling by satisfaction of special restraints[24]. Crystal structure of human MTHFR(PDB ID:6FCX) taken as a template for modeling, 10 potential structures were generated, conformation with the lowest DOPE(discrete optimized protein energy) score was selected for further refinement.

To increase the accuracy of models, refinement of problematic loops in the protein needed. Loops with high dope score were selected from residue-by-residue energy profile.

Selected loops were refined in analyzing by using loop model class. Modeled structures were further refined in Modrefiner webserver where human MTHFR crystal structure used as a reference structure for C-alpha trace. Modrefiner refines models to their native stare in terms of hydrogen bonds, backbone topology, and side-chain modeling[25]. Rplot summary of models before and after refinement shown in table.1.. Stereochemical properties and accuracy of model analysed by Ramchandran plot in RAMPAGE server[26].Statistics of nonbonded interaction between different atom types determined in ERRAT server.Z score of backbone and sidechain contact estimated in WHAT IF server[27]. Modeled structures are further checked in Verify3D, PROCHECK, QMEAN web interface[28][29].

	R.plot sum refinement	mary before		R.plot summary after refinement			
Organism name	Α	В	С	Α	В	С	
Mus musculus	95.4	2.9	1.7	94.4	5.5	0.2	
Gallus gallus	93.3	6.2	0.5	93.0	7.0	0	
Danio rerio	97.2	2.3	0.5	96.3	2.8	0.9	
Acanthaster planci	91.4	8.5	0.2	92.4	7.1	0.5	
Arabidopsis thaliana	90.3	8.5	1.2	91.3	8.1	0.6	
C.elegans	93.9	4.1	2.0	93.7	6.3	0.0	

Table 1:Summary of Ramchandran plot of models before and after refinement of models

#### 2.4. Mutational impact on models

As MTHFR one of the key regulatory enzyme involves in folate and methionine cycle, mutations in MTHFR leads to abnormal enzymatic functions that affect folate and methionine cycle which can cause severe diseases like Schizophrenia, neural tube defects, and various cancers. Most studies allelic variant C677T which leads to A222V substitution having high frequency throughout the world can cause reduced enzymatic activity, increase in thermolability and vascular diseases. Similarly, missense mutations E429A, G149V, A116T, R157Q, N324S are well studied can cause schizophrenia, homocystinuria in humans[30][31][32][33]. As MTHFR gene highly conserved all across the species, these mutations may affect enzymatic abnormality in other species also. To predict these mutational effects in modeled 3D structures FoldX algorithm was used to calculate  $\Delta\Delta G$  value by substrating energy value from WT to mutant. As some residues in models involve in

Vanderwall clash or bad torsion angles so RepairPDB in FoldX used to optimize the energy of models before mutagenesis[34].

#### 2.5.Cofactor addition and active site identification

Role of FAD, SAM in folate and methionine cycle is essential in MTHFR proteins. Coordinate of FAD extracted from the template PDB file map it into modelled proteins using Pymol. Majority of contacts to FAD are from catalytic domain W95, H127, R157, G158, A175, Y197, K2217 are the major residues interacting to FAD in human MTHFR. Replacing SAM in the position of SAH is more realistic as in the event of enzyme catalyzation, SAM in the regulatory domain who allosterically inhibits the enzyme. Molecular docking of SAM in MTHFR receptor performed using Autodock Vina program[35]. Receptor molecule prepared by removing all water molecules added with polar hydrogen atoms also Kollman charge added to the molecule. 3D conformer of SAM retrieved from Pubchem, Grid box for docking was set such that all residues interacting to the SAH should come in that 3D grid box. Binding affinity between different conformers of SAM and receptor (MTHFR) analyzed via negative Gibbs free energy score(Kcal/mol)which was calculated based on Vina's different scoring function. The most energetically favorable conformer of SAM selected for Further analysis which is shown in figure.2. Residues interacting with SAH(in crystal structure) and SAM(best conformer) analyzed via Ligplot+[36]. Result obtained from Ligplot+ shows, those residues interacting with SAH also interact with SAM. Coordinate of SAM extracted and map in into all modelled structures. After the addition of both cofactors to the models, conserved residues that are interacting with FAD, SAM in all models analyzed in Ligplot+.



Figure 2:Surface contact of 4 crucial residues(E463, T464, T481, Q485[Orange]) to SAM in the predicted docking site of human MTHFR protein

#### 2.6. Prediction of Deleterious nsSNPs in human MTHFR gene

Substitution of a nucleotide at some specific position in the genome called single nucleotide polymorphism(SNP). SNPs that mutate amino acids in the protein called nonsynonymous single nucleotide polymorphism(nsSNP). These nsSNPs in the coding region are related to most human diseases as they influence protein structure, dynamics, and stability of the protein. As these mutations occur in the course of evolution, deleterious SNPs found in the human MTHFR gene can have significant impact on other eukaryotic species also. To find the impact of deleterious nsSNPs in human MTHFR gene via experimental work is costly and took years to perform. Insilico analysis of nsSNPs can be a constructive alternative to experimental study. SNP information of the MTHFR gene retrieved from NCBI dbSNP database[37]. Out of 6390 SNPs 4872 SNPs in intronic region,1492 SNPs in 3 prime UTR region,602 in 5 prime UTR region and 915 SNPs were nsSNPs(missense mutation). Deleterious effect of 915 missense mutation predicted using SIFT, PROVEAN, PolyPhen2, and mutpred webserver. SIFT( Sorting Intolerant from tolerant) predict AA substitution effect based on sequence homology and physical properties of amino acids[38]. SIFT score below 0.05 considered as damaging. PROVEAN(Protein Variation effect analyser) predict impact of AA substitution by using pairwise sequence alignment, scores below -2.5 consider as deleterious substitution[39]. PolyPhen2 predicts the nsSNP effect by analyzing multiple sequence alignment and protein 3D structure. Scores between 0.85-1.00 are more confidently predicted to be damaging[40]. Mutpred predicts pathogenicity of amino acid substitution and molecular mechanisms associated with it[41]. Out of 915 missense mutation after filter out 40 nsSNPs carried out for further analysis by finding out  $\Delta\Delta G$  value via Yasara Foldx tool.

## Chapter 3

## Results



Figure 3.:Phylogenetic tree of MTHFR proteins all across species based on Jones-Taylor-Thorton mode, maximum likelihood tree with 1000 bootstrap replication.Bootstrap value above the branch suggest confidence/robustness of branch to be form.clades were labelled with boxes ,blue arrow marked sequences are considered to 3D modelling .

#### **3.1.Sequence Alignment And Phylogenetic Analysis**

Maximum likelihood tree derived from multiple sequence alignment of 23 diverged species shown in fig.3. Sequences of prokaryotes aligned to the catalytic domain of other species. Evolutionary tree of selected organisms shows delineation of MTHFR into 3 major clades. First clade consists of prokaryotes as they have the only catalytic domain and sequence similarity to human MTHFR is only 49%[to the catalytic domain of human MTHFR]. Second clade consist of fungi species, they have regulatory domain but lack serine-rich phosphorylation region. Clade 3 consists of eukaryotic species ranging from plants to advanced mammal-like humans. Nodes separating major clades having bootstrapping values [60-80] and bootstrapping values of internal nodes [>90] suggest tree is accurate and reliable. In clade 3 there is a subclade comprise of plants and algae as their sequence similarity [60% to human MTHFR] distant from other eukaryotes.

#### 3.2 Structure Based Sequence Alignment

In the sequence alignment, high conservation of Leucine, aspartic acid, glutamic acid, proline residues in all across the species predicts these residues maybe took part in active catalytic activities, as through the process of evolution MTHFR still retain its function from prokaryotes to advanced mammals. Sequence of Mus musculus, Arabidopsis 24nalyzin, Danio rerio(zebrafish), Gallus gallus, Acanthaster planci, C.elegans having sequence similarity greater than 50% with reference to human MTHFR. These sequence were further carried out for modelling where human MTHFR crystal structure (PDB ID-6FCX)choose as template for generating homology models.

organism name	M.weight	sequence length	pl	EC(assuming all pairs of Cys residues form cystines)	EC(assuming all Cys residues are reduced)	Half life(hrs)	Ш	GRAVY	AI
Human	74596.57	656	5.22	119915	119290	1d 6h	49.02	-0.418	80.72
Mus musculus	78924.40	695	5.32	127030	126280	1d 6h	51.17	-0.416	77.17
Gallus gallus	74212.98	651	5.47	119915	119290	1d 6h	47.04	-0.450	80.11
Danio rerio	74655.33	656	5.23	126905	126280	1d 6h	44.36	-0.425	79.21
Acanthaster planci	78138.48	689	5.93	120375	119750	1d 6h	50.89	-0.467	79.97
Arabidopsis thaliana	66803.18	594	5.33	101925	101300	1d 6h	37.31	-0.279	82.88
c.elegans	75486.72	663	5.26	119010	118260	1d 6h	37.97	-0.373	79.25

**3.3.Physiochemical Parameters Of The Sequences** 

Table 2:Physiochemical properties of human, Mus musculus, Gallus gallus, Danio rerio, Acanthaster Planci, Arabidopsis thaliana, C.elegans MTHFR sequence where EC=Extinction coefficient, II=Instability Index, GRAVY= Grand Average of hydropathy, AI=Aliphatic Index

ExPASy protparam tool was used to predict different physical and chemical properties of aminoacid sequences which was summarized in table.2..The result suggested that average molecular weight of MTHFR protein in eukaryotes varies from 66000-79000Da and theoreotical pI predicted to be between 5.22-5.93 which indicates that protein is negatively charged and acidic in nature.ExPASy protparam tool predict pH according to the pKa values of the aminoacid sidechains.Extinction coefficient(EC) indicates rate of transmitted light via scattering and absorption medium .EC of MTHFR protein lies between 119000- $128000m^{-1}cm^{-1}$  at 280nm. However EC of Loa loa was 135635 due to higher number of tyrosine,tryptophan and cysteine compare to other sequences. Aliphatic index(AI) of a protein is defined as the relative volume ocuupied by aliphatic sidechains like alanine, valine, isoleucine and leucine. AI of selected sequences predicted between 77-83 which is stable for wide range of temprature. Hydropathy is the hydrophobic or hydrophilic properties of the aminoacids in the protein. GRAVY(Grand Average of hydropathy)value of protein calculated as sum of hydropathy values of aminoacids divided by number of residues in the sequences. GRAVY value of selected sequences lies between -0.467to -0.279 indicates plausible interaction with water. Higher GRAVY value in C.elegans, Arabidopsis thaliana indicate presence of some hydrophobic residues.Instability index(II)predict stability of protein invitro.Instability index less than 40 predict as stable protein.Sequence of human, Mus musculus, Gallus gallus, Danio rerio and Acanthaster planci(Star fish) rich in PEST[proline(P),glutamic acid(E),serine(S),threonine(T)] may be the reason of their II more than 40. Factos such as some specific dipeptides in protein, disulphide bridge, protease recognition mechanism also influence instability index, In the sequence of C.elegans, Arabidopsis thaliana contain more cysteine(disuphide bridge)than other sequence so this may be reason their II fall below 40.

#### **3.4.Validation Of Models**

Refined models result in RMSD value of 0.972-2.380 Å to the template human MTHFR structure which is summarized in table.3. RAMPAGE server and PROCHECK suite evaluation of models reveals that in Ramchandran plot (normal distribution of points where Phi( $\varphi$ ) angels restricted to a negative value) of all 3D structures 90-95% of residues fall in the most favourable region, 5-8% residues fall in additionally allowed region and 0.7-1.5% of residues in generously allowed region. Only 1-6 residues(0-1.2%) per monomer present in the disallowed region, however, none of these residues fall in any binding site region indicates overall structures to be reasonably good which is summarized in fig.4. Verify 3D



Figure 4Ramchandran plot of – (a) Mus musculus,(b) Gallus gallus,(c) Danio rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans

showed all models have 88-96% of the residues have average 3D-1D score  $\geq 0.2$ [table.3] shows the quality of the model reliable. QMEAN Z score varies from 0 to 1, higher Z score relates to more favourable model and better agreement to predicted feature. QMEAN Z score of the predicted structure varies from 0.68-0.80 indicate greater degree of nativeness. In ProSA-web analysis models lie in the space of protein determined by X-RAY crystallography. Z score value of models ranges from -10 to -13 which is close to Z score of template[PDB ID-6FCX]-11.58 suggests models are so close to experimentally determined structure. WHAT IF fine packing quality control Z score of backbone and sidechain contact differs from -1.93 to -2.94 indicates models are reliable and accurate which all are summarized in table.3.

		Verify 3D	Q MEAN	ProSA	WHAT IF (F	ine packing quality control)
organism name	Template- target RMSD(in Å)	Residues have average 3D-1D score >= 0.2	global score	Z score	All contacts score	Z
Mus musculus	1.481	90.21%	0.78±0.05	-12.39	-0.292	-1.93
Gallus gallus	0.972	92.78%	0.79±0.05	-11.87	-0.339	-2.23
Danio rerio	0.988	95.12%	0.78 ± 0.05	-12.27	-0.450	-2.94
Acanthaster planci	1.517	91.58%	0.70±0.05	-10.83	-0.410	-2.69
Arabidopsis thaliana	2.380	90.54%	0.68±0.05	-12.02	-0.323	-2.13
C.elegans	1.574	88.69%	0.72±0.05	-12.32	-0.548	-2.75

Table 3:Structural validation of all models, details about the tools already explained .

#### **3.5.Structural Analysis Of Models**

The catalytic domain of all models consist of  $\alpha 8$ - $\beta 8$  Tim barrel with three  $\alpha$ helices( $\alpha 9, \alpha 10, \alpha 11$ ) and regulatory domain comprise of two five stranded  $\beta$ sheets( $\beta 9-\beta 19$ ) and 8 α-helices which confirmed from ENDScript2.0. In the structure of Arabidopsis thaliana where regulatory domain lacks an  $\alpha$ -helix( $\alpha$ 19). In all models, 38-40% of elements comprise of alpha-helix, 12-17% of element comprises of extended strand , 5-9% of elements comprise of beta-turn and 37-43% of element comprises of random coil which all are summarized in table.6. Residue interaction diagram of all models to FAD and SAM showed in fig. 6, fig.7. In the SAM interaction diagram of Arabidopsis thaliana less number of residues interacting to it, as plant MTHFR doesn't require SAM for inhibition. A lot of conserved residues interacting to FAD and SAM in all models, these residues may be the building block of the MTHFR protein in all across the species. T94, H127, T129, M129, L156, R157, Y174, A175, A195, H213, K217 and Y321 (numbering according to human MTHFR) residues in approx all models interacting to FAD. All these residues fall in the catalytic domain. P348, A368, T464, T481, I482, N483, S484, T560 and T575 residues in approx all models interacting to SAM. All these residues fall in the regulatory domain except P348 which lies in the linker region. As these residues conserved and interacting with cofactors, mutation or single nucleotide polymorphism impact on these residues may affect the function of the protein. Mutational impact( $\Delta\Delta G$ ) on all the models having the approx same value maybe the structures are not changing much as MTHFR one of the conserved gene all across the species. However, mutations G149V, N324S, W339G may cause instability( $\Delta\Delta G$ >2.0) in wide range of species.  $\Delta\Delta G$  value obtained from FoldX summarize in table.4.

	_	20		40		60		80	
Homo sapi	MDPERHERIR	FKMRRRIESG	DKWESLEEEP	PRTAFGAVN	LSREDRMAAG	GPLYIDVTWH	PAGDPGS-DK	FTSSMMIAST	79
Mus musculs	LDPERHERLR	EKMRRRMDSG	DKWFSLEFFP	PRTAEGAVNL	ISREDRMAAG	GPLEVDVTWH	PAGDPGS-DK	ETSSMMIAST	79
Gallus gallus	LDAERHERLR	EKMRRRODSG	DRWFSLEFFP	PRTANAAVNL	ISRFDRMAAG	GPLFIDVTWH	PAGDPGS-DK	ETSSMIIANT	79
C. elegans	WSPKHYELLH	ERIERLIDEK	QQFFSLEFFP	PRFVNGVPNF	LERVERLSEG	GSVFVDMTWH	MGSDPANVDK	VTSSSSIAAS	80
Danio rerio	PDADRTDRLR	DKMKRRIESG	DHWFSLEFFP	PRTSSGAVNL	ISRFDRMGSG	GPLFIDITWH	PAGDPGS-DK	ETSSMMIAST	79
Acan planci	SPSPSNTPLI	DRIAQRIACK	DKWFSLEFFP	PRTANGAVNL	LARLERMHVG	GPLFCDITWH	PAGNPGG-ET	ETSSITIATS	79
Ara thaliana	MKV I	DKIQSLADEG	KTAFSFEFFP	PKTEDGVDNL	FERMDRMVAY	GPTFCDITWG	AGGSTADL	T L D I A S R	69
		100		120		140		160	
Homo sani	AVNYCCLETI								156
Mus musculs	AVNYCCLETI			KOLCLKNIMA			SMATDLVKHI	REFGDIFDI	150
Gallus gallus	AVNYCCLETI			KRICIKNIMA		WEEEVHCE		RSEEDDVEDI	156
C elegans	MIDYCOVDTM		ADTIKHLEOA	KAMCLESILA				REFVENCES	156
Danio rerio	AVNYCCLESV			KRICIKNIMA			NEATDIVKHI	RHEEDDVEDI	156
Acan planci	AINYCGLETM	LHMTCVNOTR		KDCGIRNIIA		WOKSDSEDTE		KKMEGDHESI	159
Ara thaliana	MOSVVCVESM	MHITCTNMPV	FKIDHALFTI	RSNGIONVIA		- K E VOVEGGE		RSKYGDYFGI	148
	MQ DI I CI E DIM		ENTERIN	non qui en	Enderindge			Ren de la di	
		180		200		220		240	
		180 I		200		220 		240 I	
Homo sapi	СV <mark>А</mark> GҮРКGНР	EA	– G S F E A D L K <mark>H</mark>	LKEKVSAGAD	FIITQLFFEA	DTFFRFVKAC	TDMGITCPIV	240 I PGIFPIQGYH	227
Homo sapi Mus musculs	CV <mark>A</mark> GYPKGHP CVAGYPRGHP	EA DA	– G S F E A D L K H – E S F E D D L K H	200 I LKEKVSAGAD LKEKVSAGAD	FIITQLFFEA FIITQLFFEA	DTFFRFVKAC STFFSFVKAC	TDMG I T C P I V T E I G I S C P I L	240 I PGIFPIQGYH PGIFPIQGYT	227 227
Homo sapi Mus musculs Gallus gallus	CV <mark>A</mark> GYPKGHP CVAGYPRGHP CVAGYPKGHP	180 I DA EA	– G S F E A D L K H – E S F E D D L K H – E S Y E A D L K H	200 I LKEKVSAGAD LKEKVSAGAD LKEKVLAGAD	FIITQLFFEA FIITQLFFEA FIITQLFFRP	220 I DTFFRFVKAC STFFSFVKAC ETFLKFMKDC	TDMG I T C P I V T E I G I S C P I L Q A I G I T C P I I	240   PG   F P   QG YH PG   F P   QG YT PG   F P   QG YH	227 227 227
Homo sapi Mus musculs Gallus gallus C. elegans	CV <mark>A</mark> GYPKGHP CVAGYPRGHP CVAGYPKGHP GCAGYPLGHP	180 I DA EA QA	– G S F E AD L K H – E S F E DD L K H – E S Y E AD L K H – P S Y K AD L MY	LKEKVSAGAD LKEKVSAGAD LKEKVSAGAD LKEKVLAGAD LKAKCDAGAN	FIITQLFFEA FIITQLFFEA FIITQLFFEA FVITQLFFEA	220 I DTFFRFVKAC STFFSFVKAC ETFLKFMKDC ETFEKFVRDC	TDMGITCPIV TEIGISCPIL QAIGITCPII REIGITQPII	240 I PGIFPIQGYH PGIFPIQGYT PGIFPIQGYH PGIMPIMGYE	227 227 227 227
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio	CVAGYPKGHP CVAGYPRGHP CVAGYPKGHP GCAGYPLGHP CVAGYPTGHP	180 I DA EA QA EA	– G S F E ADLKH – E S F E DDLKH – E S Y E ADLKH – P S Y K ADLMY – E S Y E DDLRH	LK E KV SAGAD LK E KV SAGAD LK E KV SAGAD LK E KV LAGAD LK A KDAGAN LK E KVDAGAH	FIITQLFFEA FIITQLFFEA FIITQLFFEA FVITQLFFEA FIVTQLFFRA	220 1 DTFFRFVKAC STFFSFVKAC ETFLKFMKDC ETFEKFVRDC ETFLKFVKDC	TDMG I T C P I V T E I G I S C P I L QA I G I T C P I I R E I G I T Q P I I R A I G I T C P I L	240 I PGIFPIQGYH PGIFPIQGYT PGIFPIQGYH PGIMPIMGYE PGIFPIQGYH	227 227 227 227 227 227
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci	CVAGYPKGHP CVAGYPRGHP CVAGYPKGHP GCAGYPLGHP CVAGYPTGHP CVAGYPGGHP	180 I DA EA EA EA EA	- GSFEADLKH - ESFEDDLKH - ESYEADLKH - PSYKADLMY - ESYEDDLRH - ESYEEDLKY	200 I LKE (V SAGAD LKE (V SAGAD LKE (V LAGAD LKA (CDAGAN LKE (V DAGAH LKR (V DAGAH	FIITQLFFEA FIITQLFFEA FIITQLFFEA FVITQLFFEA FIVTQLFFEA FIITQLFFEA	220 I DTFFRFVKAC STFFSFVKAC ETFLKFMKDC ETFLKFVKDC KEFLQFYKDC	TDMG I T C P I V TE I G I S C P I L QA I G I T C P I I RE I G I T Q P I I RA I G I T C P I L RA I G I T V P I I	240 I PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGVMPIQGYS	227 227 227 227 227 227 230
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana	CVAGYPKGHP CVAGYPKGHP CVAGYPKGHP GCAGYPLGHP CVAGYPTGHP CVAGYPGGHP TVAGYPEAHP	180 I DA QA EA EA EA DV I GENG LAS	- G S F E AD L K H - E S F E DD L K H - E S Y E AD L K H - P S Y K AD L MY - E S Y E DD L R H - E S Y E E D L K Y N E A Y Q S D L E Y	200 I LKE (V SAGAD LKE (V LAGAD LKA (CDAGAN LKA (VDAGAN LKE (VDAGAD LKK (IDAGAD	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIVTQLFFEA FIVTQLFFKA LIVTQLFYDT	220 I DTFFRFVKAC STFFSFVKAC ETFLKFMKDC ETFLKFVKDC KEFLQFYKDC DIFLKFVNDC	TDMG I T C P I V TE I G I S C P I L QA I G I T C P I I R E I G I T Q P I I R A I G I T C P I L R A I G I T V P I I R Q I G I S C P I V	240 1 PG I F P I QG YH PG I F P I QG YT PG I F P I QG YH PG I F P I QG YH PG I F P I QG YH PG VMP I QG YS PG I MP I NN YR	227 227 227 227 227 227 230 228
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana	CVAGYPKGHP CVAGYPKGHP CVAGYPKGHP GCAGYPLGHP CVAGYPTGHP CVAGYPGGHP TVAGYPEAHP	180 I DA QA EA EA EA DV I G E NG LA S 260 260	- G S F E AD L K H - E S F E DD L K H - E S Y E AD L K H - P S Y K AD L MY - E S Y E DD L R H - E S Y E E D L R H NE A Y Q S D L E Y	200 I LKE KV SAGAD LKE KV SAGAD LKE KV LAGAD LKA CDAGAN LKE KV DAGAD LKK KI DAGAD 280 281	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIVTQLFFEA FIVTQLFFEA LIVTQLFYDT	220 I DTFFRFVKAC STFFSFVKAC ETFLKFMKDC ETFLKFVKDC ETFLKFVKDC DIFLKFVNDC 300	TDMG I T C P I V TE I G I S C P I L QA I G I T C P I I R E I G I T Q P I I R A I G I T C P I L R A I G I T V P I I R Q I G I S C P I V	240 J PGIFPIQGYH PGIFPIQGYT PGIFPIQGYH PGIFPIQGYH PGUFPIQGYS PGIMPINNYR	227 227 227 227 227 227 230 228
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana Homo sapi	CVACYPKGHP CVACYPRGHP CVACYPLGHP CCACYPLGHP CVACYPTGHP CVACYPTGHP TVACYPEAHP SLRQLVKLSK	180 I EA EA EA EA EA DV I GENG LAS 260 I L EVPQE I K DV	- G S F E AD L K H - E S F E DD L K H - E S Y E AD L K H - P S Y K AD L M Y - E S Y E DD L R H - E S Y E E D L K Y N E A Y Q S D L E Y I E P I K D N D A A	2000 I LKE KV SAGAD LKE KV SAGAD LKE KV LAGAD LKA KCDAGAN LKA KVDAGAD LKK KI DAGAD 280 I I RNYG I E LAV	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFEA FIVTQLFFEA FIVTQLFFEA LIVTQLFFKA LIVTQLFYDT	2200 I UTFFRFVKAC STFFSFVKAC ETFLKFVKDC ETFLKFVKDC ETFLKFVKDC DIFLKFVNDC DIFLKFVNDC I UVPGLHF <b>Y</b> T	TDMG ITCPIV TEIGISCPIL QAIGITCPII REIGITQPII RAIGITCPIL RAIGITCPIL RQIGISCPIV NREMATTEVL	2400 I PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGIPIMGYE PGIFPIQGYS PGIMPINNYR KRL 300	227 227 227 227 227 227 230 228
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana Homo sapi Mus musculs	CVACYPKGHP CVACYPRGHP CVACYPKGHP GCACYPLGHP CVACYPGGHP TVACYPEAHP SLRQLVKLSK SLRQLVKLSK	EA DA QA EA EA DV I GENG LAS 2600 LEVPQE I KDV LEVPQE K KDV	-GSFEADLKH -ESFEDDLKH -ESYEADLKH -PSYKADLMY -ESYEDLRH -ESYEEDLKY NEAYQSDLEY IEPIKDNDAA IEPIKDNDAA	2000 I LKE KV SAGAD LKE KV SAGAD LKE KV LAGAD LKA CDAGAN LKK VDAGAD LKK VDAGAD LKK SIDAGAD 2800 IRNYG I ELAV	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFRP FVITQLFFRA FIITQLFFKA LIVTQLFYDT SLCQELLASG SLCRELLDSG	2200 I DTFFRFVKAC ETFLKFVKAC ETFLKFVKDC ETFLKFVKDC ETFLKFVKDC DIFLKFVNDC DIFLKFVNDC UVPGLHFVTL UVPGLHFVTL	TDMGITCPIV TEIGISCPIL QAIGITCPII REIGITQPII RAIGITCPIL RAIGITCPIL RQIGISCPIV NREMATTEVL NREWATMEVL	2400 PGIFPIQGYH PGIFPIQGYT PGIFPIQGYH PGIMPIMGYE PGIFPIQGYH PGVMPIQGYS PGIMPINNYR KRL 300 KQL 300	227 227 227 227 227 227 230 228
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana Homo sapi Mus musculs Gallus gallus	CVACYPKGHP CVACYPKGHP GCACYPLGHP CVACYPKGHP GCACYPLGHP TVACYPGHP TVACYPEAHP SLRQLVKLSK SLRQLVKLSK SLRQLVKLSK	180 E A	- GSFEADLKH - ESFEDDLKH - ESYEADLKH - PSYKADLMY - ESYEDLRH - ESYEEDLKY NEAYQSDLEY IEPIKDNDAA IEPIKDNDAA	2000 I LKE (VSAGAD LKE (VSAGAD LKE (VLAGAD LKA (DAGAD LKK (VDAGAD LKK (IDAGAD 2800 I IRNYGIELAV IRNYGIELAV IRNYGVELAV	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFKA LIVTQLFYDT SLCQELLASG SLCRELLDSG SMCRELLDSG	2200 I UTFFRFVKAC STFFSFVKAC ETFLKFVKDC ETFLKFVKDC UFLKFVKDC IFLKFVNDC I LVPGLHFYTL LVPGLHFYTL UVGLHFYTL	TDMGITCPIV TEIGISCPIL QAIGITCPII RAIGITCPII RAIGITCPIL RAIGITVPII RQIGISCPIV NREWATTEVL NREVATTEVL NREVATTEVL	2400 I PGIFPIQGYH PGIFPIQGYH PGIPIQGYH PGIMPIMGYE PGIMPINGYE PGIMPINNYR KRL 300 KQL 300	227 227 227 227 227 230 228
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana Homo sapi Mus musculs Gallus gallus C. elegans	CVACYPKGHP CVACYPKGHP CVACYPKGHP CVACYPLGHP CVACYPTGHP TVACYPEAHP SLRQLVKLSK SLRQLVKLSK SLRQLVKLSK SIKRIAKLSQ	180 1 EA EA EA EA EA DV I GENG LAS 260 1 LEVPQE I KDV LEVPQE I KDV LEVPQE I KDV LEVPQH I LDD	- G S F E AD L K H - E S F E DD L K H - E S Y E AD L K H - P S Y K AD L M Y - E S Y E DD L R H - E S Y E E D L K Y N E A Y Q S D L E Y I E P I K D N D A A I E P I K D N D A A I E P I K D N D A A I E P I K D N D A A	2000 I LKE KV SAGAD LKE KV SAGAD LKE KV LAGAD LKA KCDAGAN LKK KVDAGAD LKK LIDAGAD 280 I RNYG I E LAV I RNYG I E LAV I RNYG VE LAV VQK YG TERC I	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFEA FITQLFFEA LIVTQLFFEA LIVTQLFYDT SLCQELLASG SLCRELLDSG SMCRELLDSG EMCRRLLDNG	2200 I TFFRFVKAC STFFSFVKAC ETFLKFVKDC ETFLKFVKDC ETFLKFVKDC DIFLKFVNDC UFLKFVNDC I LVPGLHFVTL LVPGLHFVTL LVHGLHFVTL TAPSIHLVTM	TDMG ITCPIV TEIGISCPIL QAIGITCPII RAIGITCPII RAIGITCPIL RAIGITCPIL RQIGISCPIV NREMATTEVL NREVATMEVL NREVATMEVL NREVATTEVL	2400 I PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGVMPIQGYS PGIMPINNYR KRL 300 KRL 300 KRL 300	227 227 227 227 227 227 230 228
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio	CVACYPKGHP CVACYPRGHP CVACYPKGHP GCACYPLGHP CVACYPGGHP TVACYPEAHP SLRQLVKLSK SLRQLVKLSK SIKRIAKLSQ SLRQLVKLSK	180 I EA EA EA EA DV I GENG LAS 260 I LEVPQE I KDV LEVPQE I KDV LEVPQE I KDV LEVPQE I KDV LEVPQE I KDV	- G S F E AD L K H - E S F E DD L K H - E S Y E AD L K H - P S Y K AD L M Y - E S Y E DD L R H - E S Y E E D L K Y NE A Y Q S D L E Y I E P I K D ND A A I E P I K D ND A A L E P I K D ND A A	2000 I LKE KV SAGAD LKE KV SAGAD LKE KV LAGAD LKA KODAGAN LKA KVDAGAD LKK SI DAGAD I RNYG I E LAV I RNYG I E LAV I RNYG V E LAV VQK YG T E RC I I RNYG I QQAY	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFRP FVITQLFFRA FIITQLFFKA LIVTQLFYDT SLCQELLASG SLCRELLDSG SMCRELLDSG EMCRRLLDNG EMCKVLLASG	2200 I T T STFFSFVKAC STFFSFVKAC ETFLKFVKDC ETFLKFVKDC DIFLKFVNDC I LVPGLHFYTL LVPGLHFYTL LVPGLHFYTL LVPGLHFYTL EVSGLHFYTL	TDMG ITCPIV TEIGISCPIL QAIGITCPII RAIGITCPIL RAIGITCPIL RAIGITCPIL RQIGISCPIV NREMATTEVL NREVATTEVL NREGSIREIL NREVATMEVL	2400 I PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGIMPIQGYS PGIMPINNYR KRL 300 KQL 300 KRL 300 KSL 300 RSL 300	227 227 227 227 227 230 228
Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci Ara thaliana Homo sapi Mus musculs Gallus gallus C. elegans Danio rerio Acan planci	CVACYPKGHP CVACYPKGHP GCACYPLGHP CVACYPLGHP CVACYPGGHP TVACYPCGHP TVACYPCAHP SLRQLVKLSK SLRQLVKLSK SLRQLVKLSK SLRQLVKLSK SLRRLVKLSK SLRNLVKLSK	180           EA           DA           CA           QA           EA           QA           UI           EA           DI           EA           EA           DI           EA           UI           EVPQE           IKDV           LEVPQE           LEVPQE	- GSFEADLKH - ESFEDDLKH - ESYEADLKH - PSYKADLMY - ESYEDLRY NEAYQSDLEY IEPIKDNDAA IEPIKDNDAA IEPIKDNDAA IEPIKDNDAA IEPIKDNDAA	2000 I LKE (VSAGAD LKE (VSAGAD LKE (VAGAD LKA (DAGAD LKA (DAGAD LKK (IDAGAD I I RNYG IELAV I RNYG IELAV I RNYG VELAV VQK YGTERCI I RNYG I QAV	FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFEA FIITQLFFEA LIVTQLFFEA LIVTQLFFKA LIVTQLFYDT SLCQELLASG SLCRELLDSG SMCRELLDSG EMCKVLLASG TMCKEMFDSG	2200 I UTFFRFVKAC STFFSFVKAC ETFLKFVKDC ETFLKFVKDC DIFLKFVKDC UIFLKFVNDC I LVPGLHFYTL LVPGLHFYTL LVHGLHFYTL TAPSIHLYTM EVSGLHFYTL DVHGLHIYTL	TDMG ITCPIV TEIGISCPIU QAIGITCPII RAIGITCPII RAIGITCPII RAIGITCPII RAIGISCPIV NREVATTEVL NREVATTEVL NREVATTEVL NREVATTEVL NREVATTEVL NREVATTEVL NREVATIEIL	240 PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGIFPIQGYH PGIMPIMGYE PGIMPINNYR KRL 300 KQL 300 KRL 300 KRL 300 KRL 300 KRL 300	227 227 227 227 230 228

Fig.8.Sequence alignment of catalytic domain of all models. Residues inside red box interacting to FAD and residues inside blue box are the dimeric interface residues [Homo sapi(Homo sapiens), Danio rerio(zebrafish), C.elegans(Caenorhabditis elegans), Acan planci(Acanthaster planci), Ara thaliana(Arabidopsis thaliana)]

		20		40		60		80	
Homo san					S V I V B T O E WD		SCRAECELKD		76
Mus musculs					SVIVETODWD		SPAFGELKD		76
Gallus gallus					SVIVETOEWD		SSPAFGLERD		76
C elegans					SVITRTROWD		SSPAFGELKD		77
Danio reric					SVIVETODWD		SSPAFCELTD	VVIEVIKSKS	71
Acan nlanc					SVILIPTEDWD		SPACOLKO		70
Ara thaliana			ANVERTKEDV	RPIEWANRPK	SVISRTKOWE		RSASVCALSD		75
/ a chanana				120	STISKIKGWL		KJAJIGALJD	1001 J = - KT KA	/ 5
		100		120				<sup>100</sup>	
Homo sap	PKEELLKMWG	EELTSEASVF	EVFVLYLSGE	PNRNGHKVTC	LPWND – – EPL	AAETSLLKEE	LLRVNRQGIL	TINSQPNING	154
Mus musculs	PREELLKMWG	EELTSEESVF	EVFEHYLSGE	PNRHGYRVTC	L P W N D – – E P L	AAETSLMKEE	LLRVNRLGIL	TINSQPNINA	154
Gallus gallus	PREELLKMWG	EELTGEESVF	EVFTCYITGE	PNKNGHRVTC	MPWND – – DPL	ATETSLLKEQ	LEKVNRRGIL	TINSQPNING	154
C. elegans	NADDRLAMFG	ANIESFEDVK	RVFINYITQA	PNADGVKVTV	LPWTEAETGV	QPETSLISEQ	LVWCNENGIL	TVNSQPSVNG	157
Danio rerio	PKEALLQMWG	EELMSEESVY	EVFTNYITGQ	TNRSGHKVMC	LPWND – – DPL	APETNLLKDE	LEKVNRRGVL	TINSQPNING	152
Acan planci	S K E E L Q Q M W G	EELACEEDVF	EVFKCYLDGE	PNRYGHKVKR	MPWND – – DEL	ASETSRLRDR	LISINKRGIL	TINSQPAVNG	157
Ara thaliana	RDKKLQQEWV	VPLKSVEDIQ	EKFKELCLGN	– – – – – – – L K S	SPWSEL-DGL	QPETRIINEQ	LIKVNSKGFL	T I N S Q P S V N A	147
		180		200		220		240	
Homo can	KREEDRIVCW	C R C C C C C C C C C C C C C C C C C C	AVIEFETEDE	TAFALLOVI				VTACLEDCDE	222
Mus muscula	KPSSDPIVGW		ATLEFFISKE	TAEALLQVL-	KKTELKVNTH				232
Callus gallus	KPSSDPVVGW		ATLEFFISKE		KITELKVNTH			VIWGIFPGRE	232
C aloganos	ADSTDRIVGW		ATLEFFISINE						232
Danio rorio	APSIDPLVGW		ATLECTIMIAE						237
Acan plane	ADSTDDVVCW		ATLEFFISSE	NVIALLQVL-					230
Ara thaliana	EREDERTVCW	CCRUCYUYOU	ATLEFFISKE						233
Ald tidialid	EKSUSPIVGW		ATLEFFCSKE	KLDAVVERC-	KALP-STITM	AVN-KGEQWV		VING V F PARE	224
		260 I		280 I		300		320 I	
Homo sap	IICPTVVDPV	SFMFWKDEAF	ALWIEQWGKL	YEEESPSRTI	IQYIHDNYFL	VNLVDNDFPL	DNC – LWQVVE	DTLELLNAEN	311
Mus musculs	IICPTVVDPI	SFMFWKDEAF	ALWIEQWGKL	YEEESPSRMI	IQYIHDNYFL	VNLVDNEFPL	DSC – LWQVVE	DTFELLNRHP	311
Gallus gallus	IICPTVVDPV	SFLSWKDEAF	ALWIEQWAKL	YEEESPSRMI	IQYIHDNYYL	VNLVDNDFPL	ENC-LWQVVD	DTFELLNSPP	311
C. elegans	IACPTVVDPL	SFRAWRDEAY	QMWMAQWG D F	YPKESKSYGV	IKAVHDEFRL	VTLVDNDFQK	PSV-LFDVLQ	KALDELKK – –	314
Danio reric	IICPTVVDPV	SFMYWKDEAF	ALWIEQWAKL	YEDESPSRMI	IQYIHDNYFL	VNLVDNDFPL	DNC – LWQVIE	DMFALLDSSQ	309
Acan planci	IICPTVVDPI	SFMVWKDEAF	GLWKHQWGHI	YSEGSPSRQI	INHIHDTFYL	VNLVNNDYVQ	DDYRLFEVLE	KMLARKEERR	315
Ara thalianə	IIQPTIVDPA	SFNVWKDEAF	ETWSRSWANL	YPEADPSRNL	LEEVKNSYYL	VSLVENDYIN	GDIFAVFA	DL	294
Homo com		215							

Homo sapiLY FQ ------315Mus musculsTERETQAP ---319Gallus gallusQQ --------313C. elegans--------314Danio rerioVPQQEANEL -------314Acan planciR EE LEGPDLD325Ara thaliana--------294

Fig.9.Sequence alignment of regulatory domain on all models. Residues inside red box are interacting to SAM and residues inside blue box are the dimeric interface residues. [Homo sapi(Homo sapiens), Danio rerio(zebrafish), C.elegans(Caenorhabditis elegans), Acan planci(Acanthaster planci), Ara thaliana(Arabidopsis thaliana)]

Sequence alignment of both catalytic and regulatory domain shown in fig.8 and fig.9. Approx all dimeric interface residues present in regulatory domain and they are conserved all across the species.

In the structure of Mus musculus catalytic domain ranges from 1-335 residue, linker region ranges from 336-360 and the regulatory domain ranges from 361 to 654. Main residues interacting with FAD are T93, W94, H95, R156, G157, Y196, D202, Q227. Main residues interacting with SAM are A367, T480, I481, N482, S483. From well-studied mutations in human MTHFR A116T, G149V, R157Q, A222V, N324S(numbers according to human MTHFR) present in catalytic domain W339G present in linker region and E429A present in regulatory domain of mouse . R157Q(in mouse R156) among the main residue interacting with FAD.

In the structure of Gallus gallus(birds) catalytic domain ranges from 1-338 residue, linker region ranges from 339-365 and the regulatory domain ranges from 366 to 651. Main residues interacting with FAD are D94, W97, H129, R159, H176, A177, D179, H203, H215, K219. Main residues interacting with SAM are P380, N458, A463, E465. From well-studied mutations in human MTHFR A116T, G149V, R157Q, A222V, N324S present in catalytic domain W339G present in linker region and E429A present in regulatory domain. R157Q(in birds R159) among the main residue interacting with FAD.

In the structure of Danio rerio(Zebrafish) catalytic domain ranges from 1-336 residue, linker region ranges from 337-363 and the regulatory domain ranges from 364 to 656 which is identical to human MTHFR. Main residues interacting with FAD are W95, H127, R157, G158, A175, W197, H201, A204, D210. Main residues interacting with SAM are S369, N456, A461. From well-studied mutations in human MTHFR G149V, R157Q, A222V, N324S present in catalytic domain, W339G present in linker region, E429A present in regulatory domain and A116T absent in Zebrafish. R157Q(in Zebrafish also R157) among the main residue interacting with FAD.

In the structure of Caenorhabditis elegans catalytic domain ranges from 1-349 residue, linker region ranges from 350-377 and the regulatory domain ranges from 378 to 663. Main residues interacting with FAD are T107, H141, R171, P175, P176, H184, F186, Y210, H214, K230. Main residues interacting with SAM are P361, F383, Q477, E479. From well-studied mutations in human MTHFR G149V, R157Q, A222V, N324S present in the catalytic domain, W339G present in the linker region, E429A present in regulatory domain and A116T absent in C.elegans. R157Q(in C.elegans R171) among the main residue interacting with FAD. In the structure of Acanthaster planci catalytic domain ranges from 1-337 residue, linker region ranges from 338-360 and the regulatory domain ranges from 361 to 689. Main residues interacting with FAD are W93, R155, D211, K218, Y214. Main residues interacting with SAM are T484, S487, Q488, T563, T576. From well-studied mutations in human MTHFR A116T, G149V, R157Q, A222V, N324S present in catalytic domain W339G present in linker region and E429A present in regulatory domain. R157Q(in Starfish R155) among the main residue interacting with FAD.

In the structure of Arabidopsis thaliana catalytic domain ranges from 1-300 residue, linker region ranges from 301-328 and the regulatory domain ranges from 329 to 594. Main residues interacting with FAD are Q118, A131, K182. Main residues interacting with SAM are T438, W411, S441, Q466, T516. As compare to models of other species fewer residues in Arabidopsis thaliana interacting with FAD. From well-studied mutations in human MTHFR G149V, R157Q, A222V, N324S present in the catalytic domain but like other models R157(R111 in A.thaliana )not interact to FAD. A116T, W339G, E429A are absent in plants.

Mutations	Gallus gallus	z.fish	C.elegans	Ara.thaliana	Mus.musculus	A.planci(Star fish)
A116T	0.0505	absent	absent	absent	0.0346	0.0542
G149V	15.0062	12.3015	15.5432	13.0041	11.8753	12.9865
R157Q	0.7466	1.8404	0.8321	0.7635	1.6543	0.7653
A222V	1.1462	-0.1047	0.9843	1.2341	0.6742	1.2341
N324S	2.0623	2.0428	2.4143	2.8890	2.2135	1.9854
W339G	4.7338	3.7752	3.3954	absent	3.6843	4.6753
E429A	0.0594	-0.0521	-0.1654	absent	-0.0342	-0.0321

Table 4: Mutational impact of common mutations occur in human MTHFR where these values represent the  $\Delta\Delta G$  value obtained from FoldX

#### 3.6. Deleterious nsSNPs in human MTHFR gene and their impact

#### on other species

AA variant	ΔΔG	SIFT SCORE	SIFT PREDICTION	PROVEAN SCORE	PROVEAN PREDICTION	PolyPhen2 SCORE	POLY_PHEN2_PRED	MUTPRED SCORE
R157Q	2.1012	0.006	Damaging	-3.93	Deleterious	1	probably damaging	0.912
A220V	3.8608	0.008	Damaging	-3.76	Deleterious	1	probably damaging	0.849
G221R	5.8128	0.001	Damaging	-7.91	Deleterious	1	probably damaging	0.954
F239L	1.9047	0.008	Damaging	-5.64	Deleterious	1	probably damaging	0.848
I256N	3.3092	0	Damaging	-6.61	Deleterious	1	probably damaging	0.951
L323P	4.5113	0.001	Damaging	-6.08	Deleterious	1	probably damaging	0.968
N324S	3.3112	0	Damaging	-4.76	Deleterious	0.983	probably damaging	0.812
R377C	3.3124	0	Damaging	-7.91	Deleterious	1	probably damaging	0.886
G387S	2.3416	0	Damaging	-5.93	Deleterious	1	probably damaging	0.891
W421C	5.8527	0.001	Damaging	-11.35	Deleterious	1	probably damaging	0.941
W421L	3.0653	0.006	Damaging	-11.14	Deleterious	0.968	probably damaging	0.882
G422R	2.6659	0.005	Damaging	-7.17	Deleterious	1	probably damaging	0.871
W500C	6.0318	0	Damaging	-12.39	Deleterious	1	probably damaging	0.956
P572L	4.5507	0	Damaging	-9.81	Deleterious	0.996	probably damaging	0.919

Table 5: Functional consequences of nsSNPs in human MTHFR

Total of 915 missense SNPs were analysed using SIFT. Out of 915 SNPs 89 were predicted deleterious (sift score < 0.05),348 were predicted tolerated and 478 SNPs not found in the database. These 89 SNPs further analysed via various tools, out of 89 SNPs 29 were predicted neutral mutation by PROVEAN, and 20 were predicted no pathogenicity by MutPred. After filtering out, 40 SNPs carried out in FoldX to quantitatively predict change in free energy upon mutagenesis. Further analyzing all the data, 14 SNPs were found damaging and structural impact on protein in all tools and  $\Delta\Delta G$  value greater than 1.8 Kcal/mol shown in table.5. R157Q may have serious impact on various eukaryotic species as this was conserved residue that bind to FAD. Out of 14 damaging SNPs first 7 lies in catalytic domain and another 7 lies in regulatory domain.

# Chapter 4 Conclusion

Role of MTHFR in eukaryote phylogeny is crucial, our modeled 3D MTHFR structure of Mus musculus, Gallus gallus, Danio rerio, Acanthaster planci, Arabidopsis thaliana, C.elegans can be used for mutational effects, ligand designing, antigenic epitope binding prediction and more. Models with cofactors can be accessed through https://doi.org/10.5281/zenodo.3781396 . T94, H127, T129, M129, L156, R157, Y174, A175, A195, H213, K217, Y321 are the conserved residues interacting to FAD and P348, A368, T464, T481, I482, N483, S484, T560, T575 are the conserved residues interacting to SAM. Mutations in these residues can cause abnormality in protein function in all across the species. R157Q is one of the most studied mutation in human MTHFR, as this is the one the conserved residue interacts with FAD this mutation may affect protein function in other species also. Predicted 14 nsSNPs are the most deleterious SNPs among MTHFR nsSNPs as their conservation score is very high, their impact on eukaryote phylogeny can't negligible.

## **Biblography**

[1] "P. Goyette, A. Pai, R. Milos, P. Frosst, P. Tran, Z. Chen, M. Chan, and R. Rozen, 'Gene structure of human and mouse methylenetetrahydrofolate reductase (mthfr),' Mammalian Genome, vol. 9, no. 8, pp. 652–656, 1998."

[2] "D. S. Froese, M. Huemer, T. Suormala, P. Burda, D. Coelho, J.-L. Guéant, M. A. Landolt, V. Kožich, B. Fowler, and M. R. Baumgartner, 'Mutation update and review of severe methylenetetrahydrofolate reductase deficiency,' Human mutation, vol. 37, no. 5, pp. 427–438, 2016."

[3] "S.-C. Liew and E. D. Gupta, 'Methylenetetrahydrofolate reductase (mthfr) c677t polymorphism: epidemiology, metabolism and the associated diseases,' European journal of medical genetics, vol. 58, no. 1, pp. 1–10, 2015."

[4] "B. J. Reyes-Hernández, A. C. Srivastava, Y. Ugartechea-Chirino, S. Shishkova, P. A. Ramos-Parra, V. Lira-Ruan, R. I. D13'053'faz de la Garza, G. Dong, J.-C. Moon, E. B. Blancaflor, et al., 'The root indeterminacy-to-determinacy developmental switch is operated through a folate-dependent pathway in a rabidopsis thaliana,' New phytologist, vol. 202, no. 4, pp. 1223–1236, 2014."

[5] "M. E. Stokes, A. Chattopadhyay, O. Wilkins, E. Nambara, and M. M. Campbell, 'Interplay between sucrose and folate modulates auxin signaling in arabidopsis,' Plant physiology, vol. 162, no. 3, pp. 1552–1565, 2013."

[6] "B. Gambonnet, S. Jabrin, S. Ravanel, M. Karan, R. Douce, and F. Rébeillé, 'Folate distribution during higher plant development,' Journal of the Science of Food and Agriculture, vol. 81, no. 9, pp. 835–841, 2001."

[7] "V. Gorelova, L. Ambach, F. Rébeillé, C. Stove, and D. Van Der Straeten, 'Folates in plants: research advances and progress in crop biofortification,' Frontiers in chemistry, vol. 5, p. 21, 2017."

[8] "Liu, Sijia, et al. 'Simultaneous downregulation of MTHFR and COMT in switchgrass affects plant performance and induces lesion-mimic cell death.' Frontiers in plant science 8 (2017): 982."

[9] "Borowczyk, Kamila, et al. 'Mutations in homocysteine metabolism genes increase keratin N-homocysteinylation and damage in mice.' International journal of genomics 2018 (2018)."

[10] "C. Kutzbach and E. Stokstad, 'Mammalian methylenetetrahydrofolate reductase partial purification, properties, and inhibition by s-adenosylmethionine,' Biochimica et Biophysica Acta (BBA)-Enzymology, vol. 250, no. 3, pp. 459–477, 1971."

[11] "B. D. Guenther, C. A. Sheppard, P. Tran, R. Rozen, R. G. Matthews, and M. L. Ludwig, 'The structure and properties of methylenetetrahydrofolate reductase from escherichia coli suggest how folate ameliorates human hyperhomocysteinemia,' Nature Structural & Molecular Biology, vol. 6, no. 4, p. 359, 1999."

[12] "A. D. Hanson, D. A. Gage, and Y. Shachar-Hill, 'Plant one-carbon metabolism and its engineering,' Trends in plant science, vol. 5, no. 5, pp. 206–213, 2000."

[13] "Ortbauer, Martina, et al. 'Folate deficiency and over-supplementation causes impaired folate metabolism: Regulation and adaptation mechanisms in Caenorhabditis elegans.' Molecular nutrition & food research 60.4 (2016): 949-956."

[14] "M. A. Mart13 053'f-Renom, A. C. Stuart, A. Fiser, R. Sánchez, F. Melo, and A. Šali, 'Comparative protein structure modeling of genes and genomes,' Annual review of biophysics and biomolecular structure, vol. 29, no. 1, pp. 291–325, 2000."

[15] "C. S. Ring, E. Sun, J. H. McKerrow, G. K. Lee, P. J. Rosenthal, I. D. Kuntz, and F. E. Cohen, 'Structure-based inhibitor design by using protein models for the development of antiparasitic agents.,' Proceedings of the national academy of Sciences, vol. 90, no. 8, pp. 3583–3587, 1993."

[16] "J. Westbrook, Z. Feng, L. Chen, H. Yang, and H. M. Berman, 'The protein data bank and structural genomics,' Nucleic acids research, vol. 31, no. 1, pp. 489–491, 2003."

[17] "U. Consortium, 'Uniprot: a hub for protein information,' Nucleic acids research, vol. 43, no. D1, pp. D204–D212, 2014."

[18] "H. M. Berman, P. E. Bourne, J. Westbrook, and C. Zardecki, 'The protein data bank,' in Protein Structure, pp. 394–410, CRC Press, 2003."

[19] "F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, et al., 'Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega,' Molecular systems biology, vol. 7, no. 1, 2011."
[20] "S. B. Needleman and C. D. Wunsch, 'A general method applicable to the search for similarities in the amino acid sequence of two proteins,' Journal of molecular biology, vol. 48, no. 3, pp. 443–453, 1970."

[21] "Knudsen, B., et al. 'Clc sequence viewer.' A/S Cb, version 6.2 (2011)."

[22] "S. Kumar, G. Stecher, and K. Tamura, 'Mega7: molecular evolutionary genetics analysis version 7.0 for bigger datasets,' Molecular biology and evolution, vol. 33, no. 7, pp. 1870–1874, 2016."

[23] "E. Gasteiger, C. Hoogland, A. Gattiker, M. R. Wilkins, R. D. Appel, A. Bairoch, et al., 'Protein identification and analysis tools on the expasy server,' in The proteomics protocols handbook, pp. 571–607, Springer, 2005."

[24] "B. Webb and A. Sali, 'Comparative protein structure modeling using modeller,' Current protocols in bioinformatics, vol. 47, no. 1, pp. 5–6, 2014."

[25] "D. Xu and Y. Zhang, 'Improving the physical realism and structural accuracy of protein models by a two-step atomic-level energy minimization,' Biophysical journal, vol. 101, no. 10, pp. 2525–2534, 2011."

[26] "S. C. Lovell, I. W. Davis, W. B. Arendall III, P. I. De Bakker, J. M. Word, M. G. Prisant, J. S. Richardson, and D. C. Richardson, 'Structure validation by  $c\alpha$  geometry:  $\phi$ ,  $\psi$  and  $c\beta$  deviation,' Proteins: Structure, Function, and Bioinformatics, vol. 50, no. 3, pp. 437–450, 2003."

[27] "C. Colovos and T. O. Yeates, 'Verification of protein structures: patterns of nonbonded atomic interactions,' Protein science, vol. 2, no. 9, pp. 1511–1519, 1993."
[28] "G. Vriend, 'What if: a molecular modeling and drug design program,' Journal of molecular graphics, vol. 8, no. 1, pp. 52–56, 1990."

[29] "D. Eisenberg, R. Lüthy, and J. U. Bowie, '[20] verify3d: assessment of protein models with three-dimensional profiles,' in Methods in enzymology, vol. 277, pp. 396–404, Elsevier, 1997."

[30] "K. Yamada, Z. Chen, R. Rozen, and R. G. Matthews, 'Effects of common polymorphisms on the properties of recombinant human methylenetetrahydrofolate reductase,' Proceedings of the National Academy of Sciences, vol. 98, no. 26, pp. 14853–14858, 2001."

[31] "Lewis, S. J., Zammit, S., Gunnell, D., & Smith, G. D. (2005). A meta-analysis of the MTHFR C677T polymorphism and schizophrenia risk. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 135(1), 2-4."

[32] "Hyland, K., Smith, I., Bottiglieri, T., Perry, J., Wendel, U., Clayton, P. T., & Leonard, J.V. (1988). Demyelination and decreased S-adenosylmethionine in 5.10-

methylenetetrahydrofolate reductase deficiency. Neurology, 38(3), 459-459."

[33] "Burda, P., Schäfer, A., Suormala, T., Rummel, T., Bürer, C., Heuberger, D., ... & Kožich, V. (2015). Insights into severe 5, 10-methylenetetrahydrofolate reductase deficiency: molecular genetic and enzymatic characterization of 76 patients. Human mutation, 36(6), 611-621."

[34] "Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., & Serrano, L. (2005). The FoldX web server: an online force field. Nucleic acids research, 33(suppl\_2), W382-W388."

[35] "AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading - Trott - 2010 - Journal of Computational Chemistry - Wiley Online Library."

https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.21334?casa\_token=KjTRmdZ2FKsAAAA A%3AaPmCUing1-5vp60sA5jmI5WRN2dIRxuYV9L2pTfuCPwxzJfj3XiFTWTGTLQsIqKyLfa\_-14YFdFmdX8 (accessed Nov. 12, 2019).

[36] "LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery | Journal of Chemical Information and Modeling." https://pubs.acs.org/doi/abs/10.1021/ci200227u (accessed Nov. 12, 2019).

[37] " Sherry, S. T., Ward, M. H., Kholodov, M., Baker, J., Phan, L., Smigielski, E. M., & Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic acids research*, *29*(1), 308-311.

[38] "Ng, Pauline C., and Steven Henikoff. "SIFT: Predicting amino acid changes that affect protein function." *Nucleic acids research* 31.13 (2003): 3812-3814."

[39]" Choi, Yongwook, and Agnes P. Chan. "PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels." *Bioinformatics* 31.16 (2015): 2745-2747."

[40]" Adzhubei, Ivan, Daniel M. Jordan, and Shamil R. Sunyaev. "Predicting functional effect of human missense mutations using PolyPhen-2." *Current protocols in human genetics* 76.1 (2013): 7-20."

[41]" Mort, Matthew, et al. "MutPred Splice: machine learning-based prediction of exonic variants that disrupt splicing." *Genome biology* 15.1 (2014): R19."

### Appendix

Huma, BCX, moue, Mu, and Apple, LER, KMRRRRD, S. KWRSLEFFP, PRIAECANNI, SER DRAMAG, CPLUTDVIWH PACDPCSD, K. ETSSMMIAST 79 Anddapit, Tablas, S. KWRSLEFF, Multi, Charles, KMRRRRD, S. KWRSLEFFP, PRISCANNI, SER DRAMAG, CPLUTDVIWH PACDPCSD, K. ETSSMMIAST 79 Anddapit, Tablas, S. KWRSLEFF, Multi, Charles, KMRRRRD, S. KWRSLEFFP, PRISCANNI, SKRPRRAG, CELLFOLTWH, PACAPCCSTALL, S. KRRRRD, S. KWRSLEFFF, PRISCANNI, SKRPRRAG, CELLFOLTWH, PACAPCSD, K. ETSSMIAST 79 Multi, Claim, Andhare, S. KRRRRD, S. KWRSLEFFF, PRISCANNI, SKRPRRAG, CELLFOLTWH, PACAPCSD, K. ETSSMIAST 79 Multi, Claim, Andhare, S. KRRRRD, S. KRRRRD, S. KWRSLEFFF, PRISCANNI, SKRPRRAG, CELLFOLTWH, PACAPCSD, K. ETSSMIAST 79 Multi, Claim, Andhare, S. KRRRRD, S. KRRRRD, S. KWRSLEFFF, PRISCANNI, SKRPRRAG, CELLFOLTWH, PACAPCSD, K. ETSSMIAST 79 Multi, Claim, AND MULTI, SKRPRRAG, S. KRRRRD, S. KRRRD, S. KRRD, S. KR			20		40 I		60 I		80 I
<pre>mode Mus LDFEHRERR EXMRRNDSG DEWTSLEFFP RETACAVN. I SEPDEMAGC DEPTFODTWG ACCEPTSD</pre>	Human_6FCX	MDPERHERLR	EKMRRRLESG	DKWFSLEFFP	PRTAEGAVNL	ISRFDRMAAG	GPLYIDVTWH	PAGDPGSD-K	ETSSMMIAST 79
Anadagagi talama Anadagagi ta	mouse_Mus	LDPERHERLR	E KMR R RMD S G	DKWFSLEFFP	PRTAEGAVNL	ISRFDRMAAG	GPLFVDVTWH	PAGDPGSD-K	ETSSMMIAST 79
<pre>um_los ATNGTYPER DETKERARGE [PF5KLARS/VANF FALLORFREG NPF5GDFOH IN ETSSSVAGG I ETSSTLATS p momental particular statements and particular statem</pre>	Arabidopsis_thaliana	MKVI	D K I Q S L A D E G	KTAFSFEFFP	PKTEDGVDNL	FERMDRMVAY	GPTFCDITWG	AGGSTADLT -	LDIASR 69
standa Acathester SPSPANPEL DELAGARACE DEWSLEEPP PRIAMAANL LAALEBMYC EPICES ELSS HILATS 9 Mide, Cake DALEBHERE EKERKBRODS DEWSLEEPP PRIAMAANL LAALEBMYC EPICES ELSS HILATS 9 Mide, Cake DALEBHERE EKERKBRODS DEWSLEEPP PRIAMAANL LAALEBPONAAG ELFEDITIVE PACPCGES ELSS HILATS 9 Mide, Cake DALEBHERE EKERKBRODS DEWSLEEPP PRIAMAANL LAEDPROC HURSLEEPENAAG ELFEDITIVE PACPCGES KETSS HILATS 9 Mide, Cake AVVSCLETT LANDTCCORE ELICALEA KOLCLINI AK KOLCLINI AK KOLCLINI AK KOLCLINI AK KOLCLINI AK Mide, Cake AVVSCLETT LANDTCCORE ELICALEA KOLCLINI AK KOLCLINI AK KOLCLINI AK KOLCLINI AK Mide, Cake AVVSCLETT LANDTCCORE ELICALEA KOLCLINI AK KOLCLINI AK KOLCLINI AK MICHTONOT HURCLINI AK KOLCLINI AK KOLCLINI AK KOLCLINI AK MICHTONOT HURCLINI AK KOLCLINI AK KOLCLINI AK MICHTONOT HURCLINI AK KOLCLINI AK MICHTONOT HURCLINI AK KOLCLINI AK MICHTONOT HURCLINI AK KOLCLINI AK MICHTONOT HURCLINI AK MICHTO	worm_Loa	ATNGTYVPLR	DRIKKRIARG	IPFFSLEFFP	PKTANSVANF	FARLDRFREG	NPMFVDIAWH	FGSDPGNINN	ETSSSSVAAG 80
Jandon Juna (PLANDELE DIENKER152) DIEVELETP *KISSLANK ISSE DIEKS LEPPEIDUNG KAUDESSE VERSTEIN Hand, GEX ANN COLLT. INTEGENER Mana (FX ANN COLLT. INTEGENER) Mana (FX STROLVER) Mana (FX ST	starfish_Acanthaster	SPSPSNTPLI	DRIAQRIACK	DKWFSLEFFP	PRTANGAVNL	LARLERMHVG	GPLFCDITWH	PAGNPGGE - T	ETSSITIATS 79
Human, GEX AVNYCGLETI HATCORAR LEETALIKA KOLGUN MA IRADPILON VEELEETGI NYAVDU VAH RSEFGAYDI 16 Manama, GEX AVNYCGLETI HATCORAR LEETALIKA KOLGUN MA IRADPILON VEELEETGI NYAVDU VAH RSEFGAYDI 16 Anama wom, Loa CLOYCOLDTM HATCORAR LEETALIKA KOLGUN MA IRADPILON VEELEETGI NYAVDU VAH RSEFGAYDI 16 Anama wom, Loa CLOYCOLDTM HATCORAR LEETALIKA KOLGUN MA IRADPILON VEELEETGI NYAVDU VAH RSEFGAYDI 16 Anama wom, Loa CLOYCOLDTM HATCORAR LEETALIKA KOLGUN MA IRADPILON VEELEETGI NYAVDU VAH RSEFGAYDI 16 Hatcora LINYCGLESU HATCORAT SINTHAL KOLGUN MA IRADPILON VEELE VEELEVUK VAADU VAH KRICOHYJ 19 anama GEX CVACYTKEN HATCORAT SINTHAL KRICONAL KRICANANA REGOVAL AN IRADPILON VEELEVUK VAADU VAH KRICOHYJ 19 anama GEX CVACYTKEN FAL HATCORAT SINTHAL KRICANANA KRICANANA REGOVAL AN IRADPILON VEELEVUK VAADU VAH KRICANANA AVNYCGLESU HATCORAT SINTHAL KRICANANA ANANANANANANANANANANANANANANANANA	zebrafish_Danio	PDADRTDRLR	DKMKRRIESG	DHWFSLEFFP	PRTSSGAVNL	ISRFDRMGSG	GPLFIDITWH	PAGDPGSD-K	ETSSMMIAST 79
Human, GRCX         AVNYCGLETI         HMTCCRRE         E E I TGHLHKA         KQLCLKNI MA         IRGDFJG D-Q         WE E - E EGGF         VYAVDUVKHI         R S E F G Y P D 165           Anadopagis, Maina         AVIVYGCLETI         HMTCCRREY         E I TGHLHKA         KQLCLKNI MA         IRGDFJG D-Q         WE E - E EGGF         SYATDU VKHI         R TE F AD YTD 155           Audiopagis, Maina         ALWCCLETI         HMTCCRREY         E I TGHLHKA         KQLCLKNI MA         IRGDFJG D-Q         WE E - E EGGF         SYATDU VKHI         R TE F AD YTD 156           atarina, Acathusa         ALWCCLETI         HMTCCRREY         E I TGHLHKA         KQLCLKNI MA         IRGDFJG D-Q         WE E - E EUGF         NYALDU VKHI         R E F ED YTD 156           atarina, Acathusa         AVIVYCCLETI         HMTCRREY         C ST E AD YTA         INTA YTA YTA         KKHCLKNINA         IRCDFYG D-G         WE E - E EUGF         NYALDU VKHI         R E F ED YTD 156           HMTGCRREY         C ST E AD YTE RY KACT TDMCITCHY         C ST E AD YTE RY KACT TDMCITCHY         F I TG YTE AD YTE RY KACT TDMCITCHY         F I TG YTE AD YTE RY KACT TDMCITCHY         F I TG YTE AD YTE RY KACT TDMCITCHY         F I F T F K KACT TA YTE AD YTE RY KACT TDMCITCHY         F I F T F K KACT TA YTE AD YTE RY KACT TDMCITCHY         F I F T F K KACT TA YTE AD YTE RY KACT TDMCITCHY         F I F T F K KACT TA YTE YTE AD YTE F Y KACT TDMCITCHY	birds_Gallus	LDAERHERLR	EKMERRQDSG	DRWFSLEFFP	PRIANAAVNL	ISRFDRMAAG	GPLFIDVIWH	PAGDPGSD-K	ETSSMITANT /9
<ul> <li>Human, GFCX AVMYCGLETI LIMYTCCRORE EEITGHLMKA KQLGLKNIMA LKGDPYGO WEF. EEGGF NYAVDIVKHI RSFFGYTDI 166</li> <li>Anakdogis, Thalina ANGSVCVESM MKLTCRNPF EKIDALETI SSKGLOVICA. RKGDPHGO KEYOCGT DCALDIVHH ISKYGDYTG 148</li> <li>wom, Lo CLOYCGLENU HUTTCCPNTY EKIDALETI SSKGLOVICA. RKGDPHGO KEYOCGT DCALDIVHH ISKYGDYTG 148</li> <li>Mandonis, Damo AVMYCGLESV HUTTCRNTK EKIDALETI SSKGLOVICA. RKGDPHGO KEYOCGT DCALDIVHI ISKYGDYTG 148</li> <li>Mandonis, Damo AVMYCGLESV HUTTCRNTK EKIDALETI SSKGLOVICA. RKGDPHGO KEYOCGT DCALDIVHI ISKYGDYTG 156</li> <li>Mandonis, Damo AVMYCGLESV HUTCRNTK EKIDDIVGHLKA KKLGLKNIMA. LKGDPYGE WEF. EVGGF NAADUVKHI RFEDDYTDI 56</li> <li>Mandonis, Damo AVMYCGLESV HUTCRNTK EKIDDIVGHLKA KKLGLKNIMA. LKGDPYGE WEF. EVGGF NAADUVKHI SFEDDYTDI 56</li> <li>Mandonis, Markov KVK, PRCHP DV CENCLAX, KKLGLKNIMA. LKGDPYGE WEF. EVGGF NAADUVKHI SFEDDYTDI 56</li> <li>Mandonis, Markov KVK, PRCHP DV CENCLAX, KKLGLKNIMA. LKGDPYGE WEF. EVGGF NAADUVKHI SFEDDYTDI 56</li> <li>Mandonis, Markov KVK, PRCHP DV CENCLAX, KKAVGAGAD HUTGLFFRVKAC TEGGISCPIY PCHPVGVG 722</li> <li>Mandonis, Markov KVK, KKVP RCHP DV CENCLAX, KKAVGAGAD HUTGLFFRVKAC TEGGISCPIY PCHPVGVG 722</li> <li>Mandonis, Markov KVK, KKVP RCHP DV CENCLAX, KKAVGAGAD HUTGLFFRVKAC TEGGISCPIY PCHPVGVG 722</li> <li>Mandonis, Markov KVK, KKVP RCHP DV CENCLAX, KKVVGAGAD HUTGLFFRVKVKG CALGCITVPI PCLIPIQYKI 222</li> <li>Mandonis, Markov KVKV, KKVP RCHP DV CENCLAX, KKVVGAGAD HUTGLFFRVKVKG CALGCITVPI PCLIPIQYKI 223</li> <li>Mandonis, Danis VXAVPEAH PA</li> <li>Markov KVKVKVKVKI KKVVKVKI KKVVKVKVKVKVKVKVKVK</li></ul>			100		120 I		140 I		160 I
<ul> <li>mouse, Mas, ALMY CGLET I, LIMITCOORP, EE, ITCHLIRAR, KGLCLKN, MA, LEGOPY, CD-H, WAA, EEGOP, SYATOLYKH, RITEFAVFO, IKA Anabogis, thalian, MGS VC VCAR, MILLITCONYK, EGO IRRIEGS, GALCLKN, IFA, REGDERGD, K-FVQVEGOP, CALDLIANY, KEEVHOFT, IKA BWORLDA, ANNYCCLET I, LIMITCONYK, EGO IRRIEGS, GALCLKN, IFA, REGDERGD, K-FVQVEGOP, KATOLYKH, KEEVHOFT, IKA BWORLDA, ANNYCCLET I, LIMITCONYK, EKITONIKA, KRICLKN, MA, REGDEVCJ, DWCE, EQUID, FATOLYKH, KEEVHOFT, IKA BWORLDA, ANNYCCLET I, LIMITCONYK, KRICLKN, KRICLKN, MA, REGDEVCJ, DWCE, EQUID, KATOLYKH, KEEVFOPTDI KG BWORLDA, ANNYCCLET I, LIMITCONYK, KRICLKN, KRICLKN, MA, REGDEVCJ, DWCE, EQUID, KATOLYKH, KEEVFOPTDI KG BWORLDA, ANNYCCLET I, LIMITCONYK, KRICLKN, KRICLKN, MA, REGDEVCJ, DWCE, LOGY, FATOLYKH, KRIFPDYPDI GWC, ANAYCCLET I, LIMITCONYK, KRIFT, KRIFT, KRIK, KRICLKN, MA, REGDEVCJ, DWCE, LOGY, FATOLYKH, KRIFFDYCH, MILLIGAR, ANNYCCLET I, LIMITCONYK, KRIFT, KRIF</li></ul>	Human_6FCX	AVNYCGLETI	LHMTCCRQRL	EEITGHLHKA	KQLGLKNIMA	LRGDPIGD-Q	WEEEEGGF	NYAVDLVKHI	RSEFGDYFDI 156
Anablogis thaliama MGS VVCVESM MILLICITEMPY E KIDALETI RSKOLGNYLA LEGOPPIGD - K-VVESGP DCALDUNHI RSYKOPYG I B safah Akashastar ALMYCGLETM LHMTCGWQTR IDITSYLMRA (DCGURNILA LEGOPPIGDE NPLLYK YRADDL RWY KELYKOPYT I B basis Galus AVNYCGLETM LHMTCGWQTR IDITSYLMRA (DCGURNILA LEGOPPIGLE WCSSDD)F NYADDURYT KEVKOPYT I B basis Galus AVNYCGLETM LHMTCGWQTR IDITSYLMRA (DCGURNILA LEGOPPIGLE WCSSDD)F NYADDURYT KEVKOPYT I B basis Galus AVNYCGLETM LHMTCGWQTR IDITSYLMRA (DCGURNILA LEGOPPIGLE WCSSDD)F NYADDURYT KEVKOPYT I B basis Galus AVNYCGLETI LHMTCSWQTR (DDITCHLQKA KELCKNIMA LEGOPPIGLE WCSSDD)F NYADDURYT (FGIPPIGLY moxek Mas CVACYFGHP DA	mouse_Mus	AVNYCGLETI	L H M T C C Q Q R P	EEITGHLHRA	KQLGLKNIMA	LRGDPVGD-H	WEA EEGGF	SYATDLVKHI	RTEFADYFDI 156
worn_Loa CLDYCGLDTM LHHITCCPYTK EQS IRHLEGS KALGLKN IFA LEGGLPAQDE NPLLYK YRALDL RWI KEFYDYFT IS aehafh, Admixat AINYCGLESY LHHITCCNGYK EKITDHLNKA KRCGLWN IAL REGOPYGE WCKSDS DIF NYALDL RWI KEFYDYFT IS bing, Guis AWNCGLESY LHHITCCNGYK EKITDHLNKA KRCGLWN IAL REGOPYGE WCKSDS DIF NYALDL RWI KHFGDHYD IS worn_Loa AXGYYACHP EA	Arabidopsis_thaliana	MQSVVCVESM	MHLTCTNMPV	EKIDHALETI	RSNGIQNVLA	LRGDPPHGQD	- K F <mark>V Q</mark> V E G <mark>G</mark> F	DCALDLVNHI	RSKYGDYFGI 148
sandh, Aanthater ALNYCGLETM LIMIT(CYNGTR I DI TKYLNRA (CBCCIRN LLA LEGGVPCDLE WKELEVHOF NYALDLYRY) KKMFCOHFS I B budg Callus ANNYCGLET LIMIT(CSNGTK DDI TCHLGKA (KRCLKN) MA LRGOPYCE D WEL-ENCH Human, GFCK CYACYFKCHP EA mouse, Mas CYACYFKCHP EA 	worm_Loa	CLDYCGLDTM	LHITCCPYTK	EQSIRHLEQS	KALGLKNIFA	LRGDLPRQDE	N P L L Y K	YRALDLIRWI	KEEYNDYFTI 156
aebath. Danio ANYCGLESY EHLUCKOVIK EKITDHINKA KRUCIKNIMA LRGPYCE - WEE-EUCGI MIADUKHI KHEPDYDI D bung. Galas ANYCGLESY EHLUCKOVIK EKITDHINKA KRUCIKNIMA LRGPYCE - WEE-EUCGI MIADUKHI RSEPDYDI D mouse. Mas CYACYPECHF EA mouse. Mas CYACYPECHF EA antionautor of the second for	starfish_Acanthaster	AINYCGIETM	LHMTCVNQTR	IDITKYLNRA	KDCGIRNILA	LRGDVPGDLE	WQKSDSEDTF	NYALDLVRYI	KKMFGDHFSI 159
Human GEX CVACYTRCHE FA	zebrafish_Danio	AVNYCGLESV	LHLTCCNQTK	EKITDHLNKA	KREGEKNIMA	LRGDPVGE - D	WVE EDGGF	NFATDLVKHI	RHEFDDYFDI 156
Human GFCX CVACYPEGHP DA mouse Mas CVACYPEGHP DA Muddogist staliana TVACYPEAHP DV GENERALS NEAVOS DELY LIKK KVSAGAD FITIQLEFEA DTFFFFVKAC TEGISEPIL PGIFFIGQYI 22 mouse Mas CVACYPEGHP DV GENERALS NEAVOS DELY LIKK KVSAGAD FITIQLEFEA DTFFFFVKAC TEGISEPIL PGIFFIGQYI 22 mouse Mas CVACYPEGHP EA mouse Mas SLEQVECHP EA Muman GFCX SLEQUVKLSK LEVPELIKUK LEVPELIKUK KKVSAGAD Human GFCX SLEQUVKLSK LEVPELIKUV LEVPELIKUK KKVSAGAD Human GFCX SLEQUVKLSK LEVPELIKUV LEVPELIKUK KKVSAGAD Human GFCX SLEQUVKLSK LEVPELIKUV LEVPELIKUV LEVPELIKUK KKVSAGAD Human GFCX SLEQUVKLSK LEVPELIKUV Human GFCX SLEQUVKLSK STEDPIGAN MUMAN GFCX SLEQUVKLSK STEDPIGAN MUMAN GFCX SLEQUVKLSK STEDPIGAN HUMAN GFCX S	birds_Gallus	AVNYCGLETT	LHMICSNQIK	DDTTGHLQKA	KREGEKNIMA	LRGDPVGE - E	WEEEVHGF	NHAVDLVKHI	RSEFUDYFUT 156
<ul> <li>Human, GKX, CYAKYPKGKP, PA,</li></ul>			180 I		200 		220 I		240 
<ul> <li>mouse_Mus_CVAXCYPRCHP_DA</li> <li>LSEEDDIKH_LKEKVSACAD</li> <li>FITOLFFEA</li> <li>STFFSVKAC TEIGISCFIL PGIFFIGUT 22 wom_Loa</li> <li>ACACYPRCHP_DA</li> <li>LSEEDDIKH_LKEKVDACAD</li> <li>FITOLFFEA</li> <li>STFFSVKAC TEIGISCFIL PGIFFIGUT 22 wom_Loa</li> <li>ACACYPRCHP_EA</li> <li>STFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama cVAXFTCHP_EA</li> <li>STFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama cFILKTKICK TEIFVERVAK</li> <li>FFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama cFILKTKICK TEIFVERVAA</li> <li>FFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama cFILKTKICK TEIFVERVAA</li> <li>FFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama sFILKTKICK TEIFVERVERVENT</li> <li>FFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama sFILKTKICK TEIFVERVENT</li> <li>FFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama sFILT</li> <li>FFFSVKAC TEIGISCFIL PGIFFIGUT 22 adafish dama sFILT</li></ul>	Human_6FCX	CVAGYPKGHP	E A	- G S F E A D L K H	LKEKVSAGAD	FIITQLFFEA	D T F F R F V K A C	TDMGITCPIV	PGIFPIQGYH 227
Anabdogis, thalian TVACYPEAHP DVIGENCLAS NEAYOSDLEY LKKKIDACAD LIVIGLEYDR DFLKEVNDC RQIGISCETV PIGUMEINNE 228 wwm.a AGSCPACHP EA ESYEBDLEY LKKKVDACAD FITTOLEFKK EFLQEVKDC RAIGITVFI PCUPICYE 23 stafish.Acanthaster CVACYPECHP EA ESYEBDLEY LKKKVDACAD FITTOLEFKK EFLQEVKDC RAIGITVFI PCUPICYE 23 wmman GCX SEPERDENE LKEVKVDACAN FITTOLEFKK EFLQEVKDC RAIGITVFI PCUPICYE 23 wmman GCX SEPERDENE LKEVKVDACAN FITTOLEFKK EFLQEVKDC RAIGITVFI PCUPICYE 23 wmman GCX SEPERDENE LKEVKVDACAN FITTOLEFKK EFLQEVKDC RAIGITVFI PCUPICYE 23 wmman GCX SEPERDENE LKEVKVDACAN FITTOLEFKK EFLQEVKDC RAIGITVFI PCUPICYE 23 wmman GCX SEPERDENE LKEVPOEIKDV IEPIKDADAA IRVGIELAV SLCQELLASC LVPCLHFYTL NREVATMEVL KQLGWTED 307 mouse Mus SLRQUVKLSK LEVPOEIKDV IEPIKDADAA IRVGIELAV SLCQELLASC VYSLHFYTL NREVATMEVL KQLGWTED 307 stafish.Acanthaster SLRNUVKLSK LUPIPEILAS LEPIKDADAA IRVGIELAV SLCQELLASC VYSLHHYTL NREVATMEVL KQLGWTED 307 stafish.Acanthaster SLRNUVKLSK LEVPOEIKDV IEPIKDADAA IRVGIELAV VLCATLETS G SATSVHLFTL NRETS REIL QQLGWTES 307 woman SLRQUVKLSK LEVPOEIKDV IEPIKDADAA IRVGIQAV MCKENLDSC VVGLHHYTL NREVATIELI KRLGUWEED 307 stafish.Acanthaster SLRNUVKLSK LEVPOEIKDV IEPIKDADAA IRVGIQAV MCKENLDSC VVGLHHYTL NREVATIELI KRLGUWEED 307 stafish.Acanthaster SLRNUVKLSK LEVPOEIKDV IEPIKDADAA IRVGIQAV MCKENLDSC VVGLHHYTL NREVATIELI KRLGUWEED 307 wmman KGX R.RPIFWA - ISANFRRE EDVEPIFYKS RYSLYRTQ DVDEPPNCRV CNSSSPAFCE LKDYVLFYLK SKSPKELLK 383 mouse MSR - PIFWA - ISANFRRE EDVEPIFYKS RYSLYRTQ EVDEPPNCRV CDSSSPAFCE LKDYVLFYLK SKSPKELLK 383 mouse MSR - PIFWA - ISANFRRE EDVEPIFYKS RYSLYRTQ DVDEPPNCRV CDSSSPAFCE LKDYVLFYLK SKSPKELLK 383 mouse MSR - PIFWA - ISANFRRE EDVEPIFYKS RYSLYRTQ DVDEPPNCRV CDSSSPAFCE LKDYVLFYLK SKSPKELLK 383 mouse MSR - PIFWA - ISANFRRE EDVEPIFYKS RYSLYRTQ DVDEPPNCRV CDSSSPAFCE LKDYVLFYLK SKSPKELLG 383 stafish.Acanthaster IRPRLPWK - ISANFRRE EDVEPIFYKS RYSLYRTQ DVDEPPNCRV CDSSSPAFCE LKDYVLFYLK SKSPKELLG 383 stafish.Acanthaster IRPRLPWK - ISANFRRE EDVEPIFYKS RYSLYRTQ DVDEPPNCRV CDSSSPAFCE LKDYVLFYLK SKSPKELG 333 mouse STAFFE SYSEVFLYF	mouse_Mus	CVAGYPRGHP	DA	- ESFEDDLKH	LKEKVSAGAD	FIITQLFFEA	S T F F S F V K A C	TEIGISCPIL	PGIFPIQGYT 227
wom, Lak ASS, VYAGHP EA	Arabidopsis_thaliana	ΤΥΑGΥΡΕΑΗΡ	DVIGENGLAS	NEAYQSDLEY	LKKKIDAGAD	LIVTQLFYDT	DIFLKFVNDC	RQIGISCPIV	PGIMPINNYR 228
stardih Aanthater CVACYPCGHP EA	worm_Loa	ACSGYPAGHP	E A	- PSYRADLLY	LKAKVDAGAD	FIISQIIFDG	KIFQSFLRDC	REIGITVPII	PGILPIQSYE 227
abadan Danio CVARVPIGHP EA	starfish_Acanthaster	CVAGYPGGHP	EA	- ESYEEDLKY	LKRKVDAGAD	FIITQLFFKA	KEFLQFYKDC	RAIGITVPII	PGVMPIQGYS 230
<pre>DBd_Gallas_CVALIPPCHP EAL</pre>	zebrafish_Danio	CVAGYPIGHP	EA	- ESYEDDLRH	LKEKVDAGAH	FIVIQLEERA	ETFLKFVKDC	RAIGITCPIL	PGTEPTQGYH 22/
Human GFCX SLRQLVKLSK LEVPQEIKDV IFPIKDNDAA IRNYGIELAV SLCQELLASG LVPCHFYTL NREMATTEVL KRLCMVEDP 307 mouse Mus SLRQLVKLSK LEVPQEIKDV IFPIKDNDAA IRNYGIELAV SLCQELLASG LVPCHFYTL NREMATTEVL KRLCMVEDP 307 wom.Loa SIRRIAELSQ LIIPDSILAS LEPIKDDDAA IRNYGIELAV SLCQELLASG LVPCHFYTL NREVATHEVL KQLCMWTEDP 307 wom.Loa SIRRIAELSQ LIIPDSILAS LEPIKDDDAA VRNFGIRHAV DLCQTLFSG SATSVHLFTL NREVSTELL KRLCLMVEDP 310 astafsh Acanthast SLRRIVKLSK LEPPEIKQV IFPIKDNDAA IRNYGIQQAV EMKKVLASG EVSCHFYTL NREVATHEVL KQLCLWAEDP 310 bids_Galus SLRQLVKLSK LEVPEIKQV IFPIKDNDAA IRNYGIQQAV EMKKVLASG EVSCHFYTL NREVATHEVL KQLCLWAEDP 307 Human GFCX RRPLFWALSAHPKRE EDVRPIFKNS RREX SVRTER UDDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSPKEELLK 383 Anatdopsis_thalana IS-RSLPWRLSAHPKRE EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSPKEELLK 383 Anatdopsis_thalana IS-RSLPWRLSAHPKRE EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSPKEELLK 383 stafsh Acanthaster IRPR-LPWKLSAHPKRE EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSPKEELLK 383 bids_Galus RRPLFWALSAHPKRE EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSPKEELLK 383 stafsh Acanthaster IRPR-LPWKSAHPKRE EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSPKEELLK 383 bids_Galus RRPLFWAVSAHPKREV EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSFKELQQ 388 bids_Galus RRPLFWAVSAHPKREV EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSFKELQQ 385 bids_Galus RRPLFWAVSAHPKREV EDVRPIFKNS RREXSIYIRTQ EWDEFPNGRW GNSS SPAFGE LKDYYLFYLK SKSFKELQQ 385 bids_Galus RRPLFWAVSAHPKREV EDVRPIFKSK UNDEDFPNGRW GNSS SPAFGE LKDYYLFYLK SKSFKELQQ 385 bids_Galus MWCELLTSE SVFEVFLYLFI SGEPNRHCKY VTCLFWNDD EPLAAFTSLL KELLRNNRG GILTDNSPN NNGKPSSDPI 461 mouse MMS WUCELTSE SVFEVFLYLFI SGEPNRHCKY VTCLFWNDD EPLAAFTSLL KELLRNNRG GILTDNSPN NNGKPSSDPI 461 Arabdopsithalana RRPLFWAVSAHPKRKY EDVRFIFKSKVELVELVENDY HVUNKGC-E NITNAPELQP NAVWGEFG REILQPIVD S40 wom.Loa MVCEELTAE DQKKKKLC LGNKNKVEENT VTCLFWNDD EPLAAFTSLL KELLRNNRG	birds_Gallus	CVAGYPKGHP	EA	- ESYEADLKH	LKEKVLAGAD	FILIQLEEKP	ETFLKFMKDC	QAIGITCPIT	PGTEPTQGYH 22/
Human_GFCX SLRQLYKLSK LEVPQEIKDV IEPIKDNDAA IRNYGIELAV SLCQELLASG LVPGLHFYTL NREMA TEVL KRLGMTEPP 307 mouse Mus SLRQLYKLSK LEVPQEIKDV IEPIKDNDAA IRNYGIELAV SLCRELDSG LVPGLHFYTL NREMA TEVL KRLGLWREPP 307 stafish_Acanthaster SLRNLYKLSK LDIPEDILAK LEPIKDNDAA IRNYGIHQAV TMCKRMAHG -VKSHLYTL NRESALAIL MILGUNERPS 307 stafish_Acanthaster SLRNLYKLSK LDIPEDILAK LEPIKDNDAA IRNYGIHQAV TMCKRMADG EVSGLHFYTL NREVA TEIL KRLGLWREPP 310 birds_Gallus SLRQLYKLSK LEVPQEIKDV IEPIKDNDAA IRNYGIHQAV TMCKRMEDSG DVHGLHIYTL NREVA TEIL KRLGLWREPP 310 birds_Gallus SLRQLYKLSK LEVPQEIKDV IEPIKDNDAA IRNYGVELAV SMCKELLDSG LVHGLHFYTL NREVA TEIL KRLGLWREPP 307 mouse.Mus RR -PI PWA - LSAHPKRRE EDVRPIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383 Anaddogsisthalma IS -RSJPWR - RPANYERTK EDVRPIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 384 radiafda.Kanthaster IRRPRIPWR - FANYERTK EDVRPIFWAS RPKSYIYRTQ DWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383 radiafda.Kanthaster IRRPRIPWK - HTANHRTK EDVRPIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383 radiafda.Kanthaster IRRPRIPWK - TANHARTK EDVRPIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383 radiafda.Kanthaster IRRPRIPWK - TANHARTK EDVRPIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383 radiafda.Kanthaster IRRPRIPWK - TANHARTK EDVRPIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 radiafda.Gallus RR - PIPWA - VSAHPKRK EDVRPIFWAS RPKSYIYRTQ EWDEFNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 radiafda.Gallus RR - PIPWA - VSAHPKRK EDVRPIFWAS RPKSYIYRTQ EWDEFNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 radiafda.Gallus RR - PIPWA - VSAHPKRK EDVRPIFWAS RPKSYIYRTQ EWDEFNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 radiafda.Gallus RK - PIPWA - VSAHPKRK EDVRPIFWAS RPKSYIYRTQ EWDEFNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 radiafda.Gallus RK - PIPWA - VSAHPKRK EDVRPIFWAS RPKSYIYRTQ DWDEFNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 radiafda.Gallus RK - PIPWA - VSAHPKRK EDVRPIFWAS RPKSYIYRTQ DWDEFNGRW GNSSSPAFGE LKDYYLFYLK			260 I	·····	- 280 I		300		320 I
mouse Mus SLRQLUKLSK LEVPQKIKDV IEPIKDNDAA IRNYGIELAV SLCRELLDSG LVPCHFITL NREVAIMEVL KQLGMVTEDP 307 worm.Loa SIRRIAELSQ LIIPDSILAS LEPIKDNDAA IRNYGIGUAV TMCKEMPDSG OVHGLHITL NREVAIMEVL QQLGUPRRSP 307 stafish.Aanthaster SIRRIVKLSK LEVPEDILKV IEPIKDNDAA IRNYGIGUAV TMCKEMPDSG OVHGLHITL NREVAIMEVL QQLGUPRRSP 307 bids_Galos SIRQLVKLSK LEVPETIKQV IEPIKDNDAA IRNYGIGUAV TMCKEMPDSG OVHGLHITL NREVAIMEVL RQLGUWEDP 307 bids_Galos SIRQLVKLSK LEVPETIKQV IEPIKDNDAA IRNYGIGUAV TMCKEMPDSG VCHALTUL NREVAIMEVL RQLGUWEDP 307 human_6FCX RR - PIPWA - LSAHPKRRE EDVRPIFKAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 Arabidopsithalana IS-RSIPWA - LSAHPKRRE EDVRPIFKAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 stafish.Aanthaster IRRPIPWA - HEANVRRTK EDVRPIFKAS RPKSYIYRTQ DWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 arabidopsithalana IS-RSIPWA - VSAHPKRRE EDVRPIFKAS RPKSYIYRTQ DWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 stafish.Aanthaster IRRPIPWA - TTANHKRK EDVRPIFKAS RPKSYIYRTQ DWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 bids_Galos RR - PIPWA - VSAHPKRKE EDVRPIFKAS RPKSYIYRTQ DWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 bids_Galos RR - PIPWA - VSAHPKRKE EDVRPIFKAS RPKSYIYRTQ DWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 bids_Galos RR - PIPWA - VSAHPKRKE VDVRPIFKAS RPKSYIYRTQ DWDEFNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLK 303 commouse Mus RCELTSEE SVFEVFLYL GURDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLX 303 bids_Galos RR - PIPWA - VSAHPKRKE VDVRPIFKAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLX 303 commouse Mus WCELLSE SVFEVFLYL GURDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLX 303 commouse Mus MCGELTSEE SVFEVFLYL GURDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLX 303 commouse Mus WCELTSEE SVFEVFLYL GURDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELLX 303 commouse Mus WCELSE SVFEVFLYL GURDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELZ GURDEFPNGRW GNSSSPAEGE LKDYLFYLK SKSPKELZ 303 mouse Mus WCELTSEE SVFEVFLYL SCEPNRCYK VTCLWNND - EPILAAETSLK KEELLRVNRQ GILTINSQPN INKKSSPTPFE 400005 MUSGENSYSFYTYTYTT GURDEFPNGRW GNSSSFFFFFYL SKSTALWY VTCLW	Human_6FCX	SLRQLVKLSK	LEVPQEIKDV	IEPIKDNDAA	IRNYGIELAV	S L C Q E L L A S G	LVPGLHFYTL	NREMATTEVL	KRLGMWTEDP 307
Arabdopsis thalana GFLRMTGFCK TKI FVEYMAA LEPIKDNEA VKAYGIHLGT EMCKMLAHG -VKSLHLYTL NKESSA BLAIL MILGNUDESS 307 starfah Acanthaster SLRNLYKLSK LDIPEDILAS LEPIKDNDAA IRNYGIGHQAV TMCKRMFDSG DVHGLHIYTL NKESSA BEIL QGLGUPRSP 300 birds Gallus SLRQLYKLSK LDIPEDILAS LEPIKDNDAA IRNYGIGHQAV TMCKRMFDSG DVHGLHIYTL NKEVATIEIL KRLGLWAEDP 300 birds Gallus SLRQLYKLSK LEVPGEIKQV IEPIKDNDAA IRNYGIGHQAV TMCKRMFDSG DVHGLHIYTL NKEVATIEIL KRLGLWAEDP 307 400 400 400 400 400 400 400 4	mouse_Mus	S L R Q L V K L S K	LEVPQKIKDV	IEPIKDNDAA	IRNYGIELAV	S L C R E L L D S G	LVPGLHFYTL	NREVATMEVL	KQLGMWTEDP 307
wom Lea SIRRIAELSQ LIIPDSILAS LEPIKNDDA VRNFGIRHAV DLCQTLFTSG SATSVHLFTL NRETSSREIL QQLGLVPRSP 307 zebafish_Danio SLRQLVKLSK LEVPEEIKQV IEPIKDNDAA IRNYGIQQAV EMCKVLLASG EVSGLHFYTL NREVATTEVL KRLGLWNEDP 300 bids_Galus SLRQLVKLSK LEVPEEIKQV IEPIKDNDAA IRNYGIQQAV EMCKVLLASG EVSGLHFYTL NREVATTEVL KRLGLWNEDP 307 30 30 40 40 40 40 40 40 40 40 40 4	Arabidopsis_thaliana	GFLRMTGFCK	TKIPVEVMAA	LEPIKDNEEA	VKAYGIHLGT	EMCKKMLAHG	- V K S L H L Y T L	NMEKSALAIL	MNLGMIDESK 307
stardin Acanthaster SLRNLVKLSK LDIPEDILKV IEPIKDNDAA IRNYGIHQAV TMCKEMEDSG DVHGLHIVTL NREVATTEIL KRLGLWNEDP 307 birds_Gallus SLRQLVKLSK LEVPGEIROV IEPIKDNDAA IRNYGUQAV EMCKVLLASG EVSGLHFYTL NREVATTEIL KRLGLWNEDP 307 mouse_Mus RRPLIPWA - LSAHPKRRE EDVRPIIFWAS RPKSVIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYVLFYLK SKSPKEELLK 383 mouse_Mus RRPLIPWA - LSAHPKRRE EDVRPIIFWAS RPKSVIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYVLFYLK SKSPKEELLK 383 mouse_Mus RRPLIPWA - LSAHPKRRE EDVRPIIFWAS RPKSVIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYVLFYLK SKSPKEELLK 383 mouse_Mus RRPLIPWA - ISAHPKRRE EDVRPIIFWAS RPKSVIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYVLFYLK SKSPKEELLK 383 mouse_Mus RRPLIPWA - TANHNRTK LDVRPIIFWAR RPKSVIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYVLFYLK SCPKEELLK 383 mouse_Mus RRPLIPWA - TANHNRTK LDVRPIIFWAS RPKSVIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYVLFYLK GQPTKDEQLR 385 birds_Gallus RRPLIPWA - TANHNRTK LDVRPIIFWAS RPKSVIYRTQ EWDDFPNGRW GNSSSPAFGE LTDYVLFYLK SKSPKEELLK 383 birds_Gallus RRPLIPWA - TSAHPKRKV EDVRPIIFWAS RPKSVIYRTQ EWDDFPNGRW GNSSSPAFGE LTDYVLFYLK SKSPKELLQ 388 birds_Gallus RRPLIPWA - VSAHPKRV EDVRPIIFWAS RPKSVIYRTQ EWDDFPNGRW GNSSSPAFGE LTDYVLFYLK SKSPKELLQ 388 starfish.Acanthster MWGEELTSEE SVFEVFLYL SGEPNRKOK VTCLIPWND EPLAAETSLK KELLRVNRQ GILTINSOPN INCKPSSDPI 461 mouse Mus WGEELTSEE SVFEVFLYL SGEPNRKOK VTCLIPWND EPLAAETSLK KELLRVNRQ GILTINSOPN INCKPSSDPI 461 sof wom Lam QGERLTSEE SVFEVFLYL SGEPNRKOK VTCLIPWND EPLAAETSLK KELLRVNRQ GILTINSOPN INCKPSSDPI 461 sof wom Lam QGERLTSEE SVFEVFLYL TGQDNRENGVV VSQLPWNEH DGIRAETSLK KELLRVNRQ GILTINSOPN INCKPSSDPI 461 sof wom Lam QGERLTSEE SVFEVFLYL TGQDNRENGVV VSQLPWNEH DGIRAETSLK KELLRVNRQ GILTINSOPN INCKPSSDPI 461 sof wom Lam QGENCPSGGVV FQXAYLEFFT SRETAEALUQ VLKKYENRN VHLVNKC-E NITNAPELOP NAVTWGIFPG REIIQPTVVD 540 sof human_GFCX VGWCPSGGVV FQXAYLEFFT SRETAEALUQ VLKKYENRN VHLVNKC-E NITNAPELOP NAVTWGIFPG REIIQPTVVD 540 sof human_GFCX PVSFMFWKDE AFALWIEGYGK KLYEESPSR TIIQYHDNY FLVNLVDNF PLDSC-LWQV 599 mouse Mus QGCPSGGVV FQXAYLEFFT SRETAEALLQ VLKKYENRN YLVNNKGC R NITNAPDLOP NAV	worm_Loa	SIRRIAELSQ	LIIPDSILAS	LEPIKNDDDA	VRNFGIRHAV	DLCQTLFTSG	SATSVHLFTL	NRETSSREIL	QQLGLWPRSP 307
2007 2017 2017 2017 2017 2017 2017 2017	starfish_Acanthaster	SLRNLVKLSK	LDIPEDILKV	IEPIKDNDAA	IRNYGIHQAV	TMCKEMFDSG	DVHGLHIYTL	NREVATIEIL	KRLGLWNEDP 310
Human 6FCX RR PLEPWA - LSAHPKRRE EDVRPIEWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383     mouseJMus RR PLEPWA - LSAHPKRRE EDVRPIEWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383     mouseJMus RR PLEPWA - LSAHPKRRE EDVRPIEWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383     analdopsis.thaliana IS RSLPWA RPANDYRTK EDVRPIEWAM RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383     analdopsis.thaliana IS RSLPWA RPANDYRTK EDVRPIEWAM RPKSYIFRTR EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK GQPTKDEQLR 385     abrids_Gallus RR PLEPWA ISAHPKRKV EDVRPIEWAM RPKSYIFRTR EWDEFPNGRW GNSSSPAFGE LTDYYLFYLK SKSSPKEQLQ 388     abrids_Gallus RR PLEPWA ISAHPKRKV EDVRPIEWAS RPKSYIYRTQ EWDEPPNGRW GNSSSPAFGE LTDYYLFYLK SKSSPKEQLQ 388     abrids_Gallus RR PLEPWA ISAHPKRKV EDVRPIEWAS RPKSYIYRTQ EWDDPPNGRW GNSSSPAFGE LTDYYLFYLK SKSPKEALLQ 388     abrids_Gallus RR PLEPWA VSAHPKRKV EDVRPIEWAS RPKSYIYRTQ EWDDPPNGRW GNSSSPAFGE LTDYYLFYLK SKSPKEALLQ 388     abrids_Gallus RR PLEPWA VSAHPKRKV EDVRPIEWAS RPKSYIYRTQ     EWDFPNGRW GNSSSPAFGE LTDYYLFYLK SKSPKEALLQ 389     adva VSAHPKRKV EDVRPIEWAS RPKSYIYRTQ     EWDFPNGRW GNSSSPAFGE LTDYYLFYLK SKSPKEALLQ 388     abrids_Gallus RR PLEPWA VSAHPKRKV EDVRPIEWAS RPKSYIYRTQ     EWDFPNGRW GNSSSPAFGE LTDYYLFYLK SKSPKEALLQ 389     adva VSAHPKRKV EDVRPIEWAS RPKSYIYRTQ     EWDFPNGRW GNSSSPAFGE LTDYYLFYLK SKSPKEALLQ 389     adva VSAHPKRKV EDVRPIEWAS RPKSYIYRTQ     EWDFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 389     adva VSAHPKRVK EDVRPIEWAS RPKSYIYRTQ     EWDFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 389     adva VSAHPKRVK EDVRPIEWAS RPKSYIYRTQ     EWDFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEALLQ 389     adva VSAHPKRVK VSAFFFFF SKSYVRVFHYLVND EPLAAETSLL KELLRVNRG GLLTINSQPN INGKPSSPH 461     avaldabasis MWGEELTSEE SVFFFYL     TQONNENGVK VSQLWVND DPLAFETSLK KELLRVNRG GLTINSQPN INGKPSSDH 461     avaldabasis MWGEELTSEE SVFFFYL     TGONNENGVK VSQLWVND DPLAFETSLL KELLRVNRG GLTINSQPN INGKPSSDH 464	zebrafish_Danio	SERQEVKESK	LEVPEEIKQV	TEPTKONDAA	TRNYGTQQAV	EMCKVLLASG	EVSGLHFYIL	NREVAIMEVL	RQLGLWAEDP 30/
Human 6FCX RR - PLEWA - LSAHPKRKE EDVRPIFWAS mouse Mus RR - PLEWA - LSAHPKRKE EDVRPIFWAS Arabidopsis thaliana IS - RS LPWR - RPANY RTK EDVRPIFWAS starfsh Acanthaster IS - RS LPWR - RPANY RTK EDVRPIFWAS starfsh Acanthaster IS - RS LPWR - RPANY RTK EDVRPIFWAS starfsh Acanthaster IS - RS LPWR - RPANY RTK EDVRPIFWAS MWC ELLTSEE SVFEVFEN - LOVRPIFWAR RKSYIFRTR EWDEPPNGRW GNSSSPAFGE LKDYYLFYLK SKSP REELLK 303 mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAR RKSYIFRTR EWDEPPNGRW GNSSSPAFGD LKDYYLFYLK SKSP KEALQ 382 mouse Mus MWC ELLTSEA SVFEVFEN - TTANHMRTK EDVRPIFWAR RKSYIFRTR EWDEPPNGRW GNSSSPAFGD LKDYYLFYLK SKSP KEALQ 388 mouse Mus MWC ELLTSEA SVFEVFEN - VSAHPKRKV EDVRPIFWAS RKSYIFRTR EWDEPPNGRW GNSSSPAFGD LKDYYLFYLK SKSP KEALQ 388 mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAS RKSYIFRTR EWDEPPNGRW GNSSSPAFGE LKDYYLFYLK SKSP KEALQ 388 mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAS RKSYIFRTR EWDEPPNGRW GNSSSPAFGE LKDYYLFYLK SKSP KEALQ 388 mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAS RKSYIFRTR EWDEPPNGRW GNSSSPAFGE LKDYYLFYLK SKSP KEALQ 388 mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAS RKSYIFRTR EWDEPPNGRW GNSSSPAFGE LKDYYLFYLK SKSP KEALQ 388 mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAS RKSYIFRTR EWDEPNGRW GNSSSPAFGE LKDYYLFYLK SKSP KEALQ 388 mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAS RKSYIFRTR EWDEPNGRW GNSSSPAFGE LKDYYLFYLK SKSP KEALG mouse Mus MWC ELLTSEA SVFEVFEN - LOVRPIFWAS RKSYIFRTR EWDEPNGRW GNSSSPAFGE LKDYYLFYLK SKSP KEALG mouse Mus MWC ELLTSEE SVFEVFEN - LOVRPIFWAS RKSYIFRTR EVELLEWNR GILTINSQPN INAKPSSDPF 465 mouse Mus MWC ELLTSEE SVFEVFEN - LCN								NDEVATENCE	
Human 6FCX RR - PLEWA - LSAHPKRE EDVRPIIFWAS RPKSYIYTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383 Arabidopsis thaliana IS - RSLEWR - RPANYFRTK EDVRPIIFWAS RPKSYIYTR DWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLK 383 Arabidopsis thaliana IS - RSLEWR - RPANYFRTK EDVRPIIFWAS RPKSYIYTR GWDEFPNGRW GNSSSPAFGE LKDYHLFYLK SKSPKEELLK 383 starfish_Acanthaster IRPRPLPWK - TTANHMRTK EDVRPIIFWAR RPKSYIFTR EWDEFPNGRW GNSSSPAFGD LKDYHLFYLK SKSPKEELQ 388 zebrafish_Dania RR - PLWA - ISAHPKRKV EDVRPIIFWAR RPKSYIFTR EWDDFFPNGRW GNSSSPAFGE LKDYHLFYLK SKSSKELQ 383 birds_Galus RR - PLWA - VSAHPKRRV EDVRPIIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYHLFYLK SKSPKEELLQ 383 birds_Galus RR - PLWA - VSAHPKRRV EDVRPIIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYHLFYLK SKSPKEELLQ 383 birds_Galus RR - PLWA - VSAHPKRRV EDVRPIIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLQ 383 birds_Galus RR - PLWA - VSAHPKRRV EDVRPIIFWAS RPKSYIYRTQ EWDEFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPKEELLZ 383 dWGEELTSEE SVFEVFFYLYL SGEPNRHGKV VTCLFWND - EPLAAETSLL KEELLRVNRQ GILTINSQPN INGKPSSDPI 461 mouse_Mus MWGEELTSEE SVFEVFFYLYL SGEPNRHGKV VTCLFWND - EPLAAETSLL KEELLRVNRQ GILTINSQPN INGKPSSDPI 461 mouse_Mus MWGEELTSEE SVFEVFFYL I GEPNRHGKV VTCLFWND - EPLAAETSLL KEELLRVNRQ GILTINSQPN INGKPSSDPI 461 mouse_Mus MWGEELTSEE SVFEVFFYL I GEPNRHGKV VTCLFWND - EPLAAETSLL KEQLLKVNRG GLITINSQPN INGKPSSDPI 461 mouse_Mus WGEELTSEE SVFEVFFYL I TOQINKSKYK VKRMPWND - DELASETSRL RDRLISINKR GULTINSQPN INGKASSDAI 461 birds_Galus WGEELTGEE SVFEVFTCYI TCQINKSCHK VKRMPWND - DELASETSRL RDRLISINKR GULTINSQPN INGKASSDAI 461 birds_Galus WGEELTGEE SVFEVFFYKCYL DGEPNRGKW VTCMPWND - DELASETSRL RDRLISINKR GULTINSQPN INGKASSDAI 461 birds_Galus WGEELTGEE SVFEVFTCYI TCQINKSCHK VKRMPWND - DELASETSRL RDRLISINKR GULTINSQPN INGKASSDAI 461 birds_Galus WGEELTGEE SVFEVFFYKCYL DGEPNRGKW VTCMPWND - DELASETSRL RDRLISINKR GULTINSQPN INGKASSDAI 461 birds_Galus WGEELTGEE SVFEVFFYKCYL VI TCQINKSCHK VKRMPWND - DELASETSRL RDRLISINKR GULTINSQPN INGKASSDAI 461 000000000000000000000000000000000000	birds_Gallus	SERQEVELSE	LEVPQEIKDV	IEPIKDNDAA	IRNYGVELAV	SMCRELLDSG	LVHGLHFYTL	NREVATTEVL	KRLGIWNEDP 307
mouse_Mus_RR PLEWA LSAHPKRE EDVRPIIFWAS Rabidopsis_thalian 1S-REV_VR LSAHPKRE EDVRPIIFWAS Rabidopsis_thalian 1S-REV_VR LSAHPKRE EDVRPIIFWAS Rabidopsis_thalian 2K-REV_VR LSAHPKRKV EDVRPIIFWAS REKSYIFRT EWDEFPORW GDSSSPAFD LKDYVLFYLK GQPTKDEQLR 385 staffsh.Aanthaster IRPRPIIPWA ISAHPKRKV EDVRPIIFWAS REKSYIFRT EWDEFPORW GDSSSPAFGE LTDYVLFYLK GQPTKDEQLR 385 birds_Gallus RR PLEWA ISAHPKRKV EDVRPIIFWAS REKSYIFRT Q EWDEFPORW GDSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 birds_Gallus RR PLEWA ISAHPKRKV EDVRPIIFWAS REKSYIFTQ EWDDFPNGRW GDSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 birds_Gallus RR PLEWA VSAHPKRVV EDVRPIIFWAS REKSYIFTQ EWDDFPNGRW GDSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 birds_Gallus RR PLEWA VSAHPKRVV EDVRPIIFWAS REKSYIFTQ EWDDFPNGRW GDSSSPAFGE LKDYYLFYLK SKSPKEALLQ 383 birds_Gallus RR PLEWA VSAHPKRVV EDVRPIIFWAS REKSYIFTQ EWDDFPNGRW GDSSSPAFGE LKDYYLFYLK SKSPKEALLQ 384 mouse_Mus MWCEELTSEE SVFEVFEHYL SGEPNRNGHK VTCLPWND EPLAAETSLK KEELLRVNRL GLIFINSQPS NIGKPSSDPY 461 mouse_Mus MWCELLSEE SVFEVFEHYL SGEPNRNGHK VTCLPWND EPLAAETSLK KEELLRVNRL GLIFINSQPS VNGAPSSDPY 454 worm_Loa MYCERLETLA DIQKVFYNFI TQDNNENCVK VSQLPWNERH DCIRAETSLK KEELLRVNRL GLIFINSQPS VNGAPSSDPY 464 zebafish.Danio MWCEELMSEE SVFEVFEHYL DCEENRYCHK VKRWPND DPLAFETSLK KEQLEKVNR GULTINSQPS NIGAPSSDPY 466 staffsh.Aanthaster MWCEELAGEE DVFEVFKYL DCEENRYCHK VKWLPWND DPLAFETSLK KEQLEKVNR GULTINSQPS NIGAPSSDA1461 50 MOUSE_MUS VGWCPSGGYV FQKAVLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REILQPTVVD 540 mouse_Mus VGWCPSGGYV FQKAVLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REILQPTVVD 540 mouse_Mus VGWCPSGGYV FQKAVLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REILQPTVVD 540 birds_Gallus VGWCPACGYV FQKAVLEFFT SRETAAALLD VLKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REILQPTVVD 540 safafsh.Zanhastr VGWCAPKCYV FQKAVLEFFT SRETAEALLQ VLKKYELRVN YHVNNKG-E NITNAPELQP NAVTWGIFPG REILQPTVVD 540 safafsh.Zanhastr VGWCAPKCYV FQKAVLEFFT SRETAEALLQ VLKKYELRVN YHVNNKG QNSTAPALD NAVTWGIFPG REI	Dirds_Gallus	SERQEVRESK	LEVPQEIKDV 340 I		IRNYGVELAV 360 I	SMCRELLDSG		NREVATTEVL	K R L G I WN E D P 307 400 I
Arabidopsis_thaliana IS - RS LPWR RPANVFRTK EDVRPI FWAM RPKSYI FRTK EWDEFPQGRW GDSSSPAFOD LKDYHLFYLK GQPTKLEQQ 382 starfish_Acanthaster IRPRPLPWKTTANHMRTK EDVRPI FWAM RPKSYI FRTK EWDEFPNGRW GDSSSPAFOD LKDYHLFYLK SKSSKEELQQ 388 zebrafish_Acanthaster ISPRPLPWK TSAHPKRKV EDVRPI FWAA RPKSYI FRTK EWDEFPNGRW GDSSSPAFOE LKDYYLFYLK SKSSKEELQQ 388 birds_Gallus RR - PLPWK VSAHPKRRV EDVRPI FWAS RPKSYI YRTQ EWDDFPNGRW GDSSSPAFOE LKDYYLFYLK SKSPRELLK 383 #40 #40 #40 #40 #40 #40 #40 #40	Human_6FCX	RRPLPWA-	1 - L S A H P K R R E	E D V R PI FWA S	I RNYGVELAV 360 I RPKSYIYRTQ	SMCRELLDSG EWDEFPNGRW	L V H G L H F Y T L 380 I G N S S S P A F G E	NREVATTEVL LKDYYLFYLK	K R L G I WN E D P 307 400 I S K S P K E E L L K 383
wom_Loa VR ALDWRS FGENNPIRCK EDVRPIEWAM RPKSYIHETK EWDEFPNCRW CNSSSPAFGD LKDYHLFYLK GGPTKDEQLR 385 staffsh_Danio RR PLPWA - TTANHMRT EDVRPIEWAA RPKSYIHETK EWDEFPNCRW CDSSSPAFGD LKDYHLFYLK SKSELQQ 383 birds_Galus RR PLPWA - VSAHPKRV EDVRPIEWAS RPKSYIHYTQ DWDDFPNCRW GNSSSPAFGE LKDYYLFYLK SKSELQQ 383 birds_Galus RR PLPWA - VSAHPKRV EDVRPIEWAS RPKSYIYRTQ EWDDFPNGRW GNSSSPAFGE LKDYYLFYLK SKSELQQ 383 birds_Galus RR PLPWA - VSAHPKRV EDVRPIEWAS RPKSYIYRTQ EWDDFPNGRW GNSSSPAFGE LKDYYLFYLK SKSELQQ 383 birds_Galus RR PLPWA - VSAHPKRV EDVRPIEWAS RPKSYIYRTQ EWDDFPNGRW GNSSSPAFGE LKDYYLFYLK SKSPREELLK 383 40 40 40 40 40 40 40 40 40 40 40 40 40	Human_6FCX mouse_Mus	RR PLPWA - RR PLPWA -	400 400 400 1 - L S AHP K R R E - L S AHP K R R E	EDVR PI FWAS	I RNYGVELAV 360 I RPKSYIYRTQ RPKSYIYRTQ	EWDEFPNGRW DWDEFPNGRW	L V HG L H F Y T L I GN S S S P A F G E GN S S S P A F G E	LKDYYLFYLK LKDYYLFYLK	K R L G I WN E D P 307 400 I S K S P K E E L L K 383 S K S P R E E L L K 383
Startin Acanchaster I REPKP LEWK FLAMHMKIK EDVKPILEWAA REPKSTIYRTQ EWDDFPNCRW GDSSSPAFGE LKDTYLFTK SKSFKEELQ 388 zebrafsh Danio RR - PLEWA - VSAHPKRV EDVRPILEWAS REPKSTIYRTQ EWDDFPNCRW GNSSSPAFGE LKDTYLFTK SKSFKEELQ 388 birds_Gallus RR - PLEWA - VSAHPKRV EDVRPILEWAS REPKSTIYRTQ EWDDFPNGRW GNSSSPAFGE LKDTYLFTK SKSFKEELQ 388 dirds_Gallus RR - PLEWA - VSAHPKRV EDVRPILEWAS REPKSTIYRTQ EWDDFPNGRW GNSSSPAFGE LKDTYLFTK SKSFKEELQ 388 dirds_Gallus RR - PLEWA - VSAHPKRV EDVRPILEWAS REPKSTIYRTQ EWDDFPNGRW GNSSSPAFGE LKDTYLFTK SKSFKEELQ 388 dirds_Gallus MWGEELTSEE SVFEVFHYLL Arabidopsis thaliana EWVVPLKSVE DIQEKFKELC LGN	Human_6FCX mouse_Mus Arabidopsis_thaliana	RR PLPWA - RR PLPWA - IS - RSLPWR -	400 340 - L S AH P K R R E - L S AH P K R R E - R P AN V F R T K	E D V R P I FWA S E D V R P I FWA S E D V R P I FWA S E D V R P I FWA N	I RNYGVELAV 360 I RPKSYIYRTQ RPKSYIYRTQ RPKSYISRTK	EWD E F PN G RW DWD E F PN G RW GWE D F PQ G RW	L VHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSASYGA	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LSDHQFS R	K R L G I WN E D P 307 400 I S K S P K E E L L K 383 S K S P R E E L L K 383 P R A R D K K L Q Q 382
<pre>bids_callus RRPLIPWAVSAHPRRV EDVRP IFWAS RPKSTIYTK Q DWDPPNGRW GNSSSPACE LD7LTFLK SKSPRELLQ 383 bids_callus RRPLIPWAVSAHPRRV EDVRP IFWAS RPKSTIYTK Q DWDPPNGRW GNSSSPACE LD7LTFLK SKSPRELLQ 383 fullos RRPLIPWAVSAHPRRV EDVRP IFWAS RPKSTIYTK Q DWDPPNGRW GNSSSPACE LD7LTFLK SKSPRELLQ 383 fullos RRPLIPWAVSAHPRRV EDVRP IFWAS RPKSTIYTK Q DWDPPNGRW GNSSSPACE LD7LTFLK SKSPRELLQ 383 fullos RRPLIPWAVSAHPRRV EDVRP IFWAS RPKSTIYTK Q DWDPPNGRW GNSSSPACE LD7LTFLK SKSPRELLQ 383 fullos RRPLIPWAVSAHPRRV EDVRP IFWAS RPKSTIYTK Q EWDPPNGRW GNSSSPACE LD7LTFLK SKSPRELLQ 383 fullos RRPLIPWAVSAHPRRV EDVRP IFWAS RPKSTIYTK Q EWDPPNGRW GNSSSPACE LD7LTFLK SKSPRELLQ 383 fullos RMS MWGEELTSE SVFEVFKLYL SCEPNROGK VTCLPWND- EPLAFETSLK KELLRVNRL GLLTINSQPS VNAERSDSPT 454 wom_Loa MYGERLETLA D IQKVFVNFI TQDNNENGVK VSQLPWNEH DGIRAETSLI KDQLWCNAN GFLTINSQPS VNAERSDSPT 454 stafish Aanthaster MYGELACEE DVFEVFKCYL L CONNENGVK VSQLPWNEH DGIRAETSLI KDQLWCNAN GLLTINSQPS VNAERSDSPT 454 ibids_Gallus MWGEELMSEE SVFEVFTCYI TQDNNESGKK VMCLPWND- DELAFETSLL KDQLWCNAN GLLTINSQPN INGKASSDAI 461 bids_Gallus MWGEELMSEE SVFEVFTCYI fullos CFU FUNCTUN for mouse Mus VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 mouse Mus VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 stafish Aanthaster VGWGAPGGYC YQKAYLEFFT SRETVEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 stafish Aanthaster VGWGAPGGYV FQKAYLEFFT SRENAALLD VLKKYELRVN YHLVNKG-Q NITNAPDLQP NAVTWGIFPG REIIQPTVVD 540 stafish Aanthaster VGWGAPGGYC YQKAYLEFFT SRENAALLD VLKKYELRVN YHLVNKG-Q NITNAPDLQP NAVTWGIFPG REIIQPTVVD 540 stafish Aanthaster VGWGAPGGYV FQKAYLEFFT SRENAALLD VLKKYELRVN YHLVNVKG-Q NITNAPDLQP NAVTWGIFPG REIIQPTVVD 540 fullos Gallus FGWGPSGGYV FQKAYLEFFT SRENAALLD VLKKYELRVN YHLVNVKG-Q NITNAPDLQP NAVTWGIFPG REIIQPTVVD 540 stafish Aanthaster VSFKMKDE AFALWIEQWG KLYEEESPSR MIIQYHDNY FLVNLVDNF PLDSC-LWQV S99 mouse Mus PISFRWKDE AFALWIEQWG KLYEEESPSR</pre>	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa	R R P L PWA - R R P L PWA - I S - R S L PWR - V R A L PWR S	440 450 471 472 473 474 475 475 475 475 475 475 475	E DV R PI FWA S E DV R PI FWA S E DV R PI FWA S E DV R PI FWA N E DV R PI FWS M	I RNYGVELAV 3600 I RPKSYIYRTQ RPKSYIYRTQ RPKSYISRTK RPKSYIFRTR	EWD E F PN G RW DWD E F PN G RW GWE D F PN G RW EWD E F PN G RW	L V HG L H F Y T L 3000 I GN S S S P A F G E GD S R S A S Y G A GN S S S P A F N D	LKDYYLFYLK LKDYYLFYLK LSDHQFS R LKDYYLFYLK	KRLGIWNEDP 307 400 1 SKSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 GQPTKDEQLR 385
Bub_datus       RKK+TELWK       POWAIT KKW       EDVITING 2       Mitter KKW       GRUPTING 40       40 <td>Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster</td> <td>RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R PR P L PWK -</td> <td>LEVPQETKDV - LSAHPKRRE - LSAHPKRRE - RPANVFRTK FGENHPIRCK - TTANHMRTK</td> <td>E DV R P I FWA S E DV R P I FWA S E DV R P I FWA S E DV R P I FWA M E DV R P I FWA A E DV R P I FWA A</td> <td>I RNYGVELAV 360 I RPKSYIYRTQ RPKSYIYRTQ RPKSYISRTK RPKSYIFRTR RPKSYIFRTK</td> <td>EWDEFPNGRW DWDEFPNGRW GWEDFPQGRW EWDEFPNGRW DWDDFPNGRW</td> <td>L VHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSASYGA GNSSSPAFGD GDSSSPAFGD</td> <td>NREVATTEVL LKDYYLFYLK LKDYYLFYLK LSDHQFSR LKDYYLFYLK LKDYHLFYLK</td> <td>KRLGIWNEDP 307 400 I SKSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSSKEELQQ 388</td>	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R PR P L PWK -	LEVPQETKDV - LSAHPKRRE - LSAHPKRRE - RPANVFRTK FGENHPIRCK - TTANHMRTK	E DV R P I FWA S E DV R P I FWA S E DV R P I FWA S E DV R P I FWA M E DV R P I FWA A E DV R P I FWA A	I RNYGVELAV 360 I RPKSYIYRTQ RPKSYIYRTQ RPKSYISRTK RPKSYIFRTR RPKSYIFRTK	EWDEFPNGRW DWDEFPNGRW GWEDFPQGRW EWDEFPNGRW DWDDFPNGRW	L VHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSASYGA GNSSSPAFGD GDSSSPAFGD	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LSDHQFSR LKDYYLFYLK LKDYHLFYLK	KRLGIWNEDP 307 400 I SKSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSSKEELQQ 388
Human 6FCX       MWGEELTSEA       SVFEVFYLVL       SGEPNRNGHK       VTCLPWND       EPLAAETSLK       KEELLRVNRQ       GILTINSOPN       INAKPSSDP1 461         Arabidopsis thaliana       EWVVPLKSVE       DIQEKFKELC       LGA       LGSPNRHGYR       VTCLPWND       EPLAAETSLK       KEELLRVNRQ       GILTINSOPN       INAKPSSDP1 461         Arabidopsis thaliana       EWVVPLKSVE       DIQEKFKELC       LGA       LGSPNRHGYR       VTCLPWND       EPLAAETSLK       KEELLRVNRQ       GILTINSOPN       INAKPSSDP1 461         worm_Loa       MYGERLETLA       DIQKVFVNFI       TODNNENGVK       VSQLPWNERH       DGLOPETRII       NEQLIXNAK       GELTINSOPN       VNAERSDSPT 454         starfish_Acanthaster       MWGEELACEE       DVFEVFKCYL       DGEPNRYGHK       VKRMPWND       DELASETSRL       RDRLISINKR       GILTINSOPN       INAKPSSDP1 466         zebrafish_Danio       MWGEELTGEE       SVFEVFTCYI       TGGPNRNGHK       VTCLPWND       DPLAFETSLL       KEQLEKVNRR       GILTINSOPN       INGKASSDA1 461         birds_Gallus       MWGEELTGEE       SVFEVFTYI       TGEPNKNGHR       VTCLPWND       DPLAFETSLL       KEQLEKVNRR       GILTINSOPN       INGKASSDA1 461         worm_Loa       GWGCPPSGCYV       FQKAYLEFFT       SRETAEALLQ       VLKKYELRVN <t< td=""><td>Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birde_colluc</td><td>R R P L PWA - R R - P L PWA - I S - R S L PWR - V R - A L PWR S I R PR P L PWK - R R - P L PWA - P - P L PWA</td><td>LEVPQETKDV 340 - LSAHPKRRE - LSAHPKRRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKPKV</td><td>E DV R P I FWA S E DV R P I FWA A E DV R P I FWA A E DV R P I FWA S</td><td>IRNYGVELAV 360 RPKSYIYRTQ RPKSYIYRTQ RPKSYISRTK RPKSYIFRTR RPKSYIHRTK RPKSYIYRTQ PPKSYIYRTQ</td><td>EWDEFPNGRW DWDEFPNGRW GWEDFPQGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDEPNCRW</td><td>LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSSPAFGE GDSRSSPAFGD GDSSSPAFGD GNSSSPAFGE CNSSSPAFGE</td><td>NREVATTEVL LKDYYLFYLK LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYHLFYLK LTDYYLFYLK</td><td>KRLGIWNEDP 307 400 5KSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 CVCPDELLV 383</td></t<>	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birde_colluc	R R P L PWA - R R - P L PWA - I S - R S L PWR - V R - A L PWR S I R PR P L PWK - R R - P L PWA - P - P L PWA	LEVPQETKDV 340 - LSAHPKRRE - LSAHPKRRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKPKV	E DV R P I FWA S E DV R P I FWA A E DV R P I FWA A E DV R P I FWA S	IRNYGVELAV 360 RPKSYIYRTQ RPKSYIYRTQ RPKSYISRTK RPKSYIFRTR RPKSYIHRTK RPKSYIYRTQ PPKSYIYRTQ	EWDEFPNGRW DWDEFPNGRW GWEDFPQGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDEPNCRW	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSSPAFGE GDSRSSPAFGD GDSSSPAFGD GNSSSPAFGE CNSSSPAFGE	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYHLFYLK LTDYYLFYLK	KRLGIWNEDP 307 400 5KSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 CVCPDELLV 383
Human_6FCX       YWGEELTSEA       SYFEVFYLYL       SGEPNRNGHK       YTCLPWND       EPLAAETSLL       KEELLRVNRQ       GILTINSQPN       INGKPSSDP1       461         mouse_Mus       MWGEELTSEE       SVFEVFEHYL       SGEPNRHGYR       YTCLPWND       EPLAAETSLM       KEELLRVNRQ       GILTINSQPN       INAKPSSDP1       461         worm_Loa       MYGERLETLA       DIQKYFKELC       LGN       LKSSPWSEL       DGLQPETRII       NEQLWCNAN       GFLTINSQPS       VNAERSDSP1       454         starfish_Canthaster       MWGEELACEE       DYFVFKCYL       DGEPNRYGHK       VKRMPWND       DELASETSL       KDQLWCNAN       GFLTINSQPS       VNAERSDSP1       466         zebrafish_Danio       MWGEELTGEE       SVFEVFTCYI       DGEPNRYGHK       VKRMPWND       DPLAPETNL       KDELKVNRR       GILTINSQPN       INGKPSSDP1       466         birds_Gallus       MWGEELTGEE       SVFEVFTCYI       TGEPNKNGHR       VTCMPWND       DPLAPETNL       KEELLKVNRR       GILTINSQPN       INGKPSSDP1       461         mouse_Mus       VGWGPSGGYV       FQKAYLEFFT       SRETAEALLQ       VLKKYELRVN       YHLVNVKG-E       NITNAPELQP       NAVTWGIFPG       REIIQPTVVD 540         mouse_Mus       VGWGPSGGYV       YQKAYLEFFT       SRETAEALLQ       VLKKYELRVN	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus	R R P L PWA - R R P L PWA - I S - R S L PWR - V R A L PWR S I R PR P L PWK - R R P L PWA - R R P L PWA -	Allow - LSAHPKRRE - LSAHPKRRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKRRV	E DV R P I FWA S E DV R P I FWA S E DV R P I FWA S E DV R P I FWA M E DV R P I FWA M E DV R P I FWA A E DV R P I FWA S E DV R P I FWA S	IRNYGVELAV Jaco RPKSYIYRTQ RPKSYIYRTQ RPKSYISRTK RPKSYIFRTR RPKSYIHRTK RPKSYIYRTQ RPKSYIYRTQ	EWDEFPNGRW DWDEFPNGRW GWEDFPQGRW EWDEFPNGRW DWDDFPNGRW EWDDFPNGRW	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSASYGA GNSSSPAFGD GNSSSPAFGD GNSSSPAFGE	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYHLFYLK LTDYYLFYLK LKDYYLFYLK	KRLG I WN EDP 307 400 5 KS P KEELL K 383 5 KS P REELL K 383 PRARDKKLQQ 382 GQ P T KDEQLR 385 5 KS S K EELQQ 388 5 KS P KEALLQ 383 5 KS P REELL K 383
mouse_Mus_MWGEELTSEE_SVFEVFEHYLL Arabidopsis_thaliana EWVPUKSYD DIQEKFKELC worm_Loa MYGERLETLA DIQKVFVNFI Staffsh_Acanthaster WWGEELACEE DVFEVFKCYL zebrafish_Danio MYGEELACEE DVFEVFKCYL birds_Gallus MWGEELTGEE SVFEVFTCYI birds_Gallus MWGEELTGEE SVFEVFTCYI soo Human_6FCX VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 Arabidopsis_thaliana VGWGGPVGYV YQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGVFPA KEIIQPTIVD 540 arbifsh_Canthaster VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGVFPA KEIIQPTIVD 540 arbifsh_Canthaster VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNKG-E NITNAPELQP NAVTWGVFPA KEIIQPTIVD 540 arbifsh_Canthaster VGWGPAGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGVFPA KEIIQPTIVD 540 arbifsh_Canthaster VGWGPAGGYV FQKAYLEFFT SRETAEALLD VLKKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGVFPA KEIIQPTIVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SRENAALLD VLRVPS VHVVNHOKCF NITNAPELQP NAVTWGVFPA KEIIQPTIVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SRENAALLD VLRVPS VHVNHOKCF Q NITNAPDLQP NAVTWGVFPA KEIIQPTIVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SRENAALLD VLRVPS VHVNHOKCF Q NITNAPDLQP NAVTWGIFPG REIIQPTIVD 540 soo soo soo soo birds_Gallus PISFMFWKDE AFALWIEQWG KLYEESPSR TIIQYIHDNY FLVNLVDNFF PLDSC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus	RR P L PWA - RR P L PWA - I S - R S L PWR - VR - A L PWR S I R P R P L PWK - RR P L PWA - RR P L PWA -	AND	E D V R P I FWA S E D V R P I FWA S E D V R P I FWA S E D V R P I FWA M E D V R P I FWA M E D V R P I FWA S E D V R P I FWA S	I RNYGVELAV 3600 1 RPKSYIYRTQ RPKSYISRTK RPKSYIFRTR RPKSYIHRTK RPKSYIYRTQ 4000 1 4000 1	EWD EF PNGRW DWD EF PNGRW GWE DF PQGRW EWD EF PNGRW DWD DF PNGRW EWD DF PNGRW	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSASYGA GNSSSPAFGD GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE	NREVATTEVL LKDYYLFYLK LSDHQFSR LKDYYLFYLK LKDYHLFYLK LKDYYLFYLK LKDYYLFYLK	KRLGIWNEDP 307 400 I SKSPKEELLK 383 SKSPKEELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPREELLK 383 480 I
Arabidopsis_thaliana EWVVPLKSVE DIQEKFKELC LGK LKSSPWSEL DGLQPETRIT NEQLIKVNSK GFLTINSQPS VNAERSDSPT 454 wom_Loa MVGERLETLA DIQKVFVNFI TQDNNEGVKV VSQLPWNERH DGIRAFTSLL KDQLLWNAN GFLTINSQPS VNAERSDSPT 454 starfish_Acanthaster MWGEELACEE DVFEVFKCYL DGEPNRYGHK VKRMPWND - DELASETSRL RDRLISINKR GILTINSQPA VNGAPSTDPV 466 zebrafish_Danio MWGEELMSEE SVYEVFTNYI TGQTNRSGHK VKRMPWND - DPLAPETNLL KDQLEKVNRR GVLTINSQPA VNGAPSTDPV 466 mouse_Mus VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 mouse_Mus VGWGPSGGYV FQKAYLEFFT SRETVEALLQ VLKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 arabidopsis_thaliana VGWGFVGYY YQKAYLEFFT SRETVEALLQ VLKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 wom_Loa VGWGKPGGYC YQKAYLEFFT SRETVEALLQ VLKYELRVN YHVVNHNGKE DITNTDRQP IAVTWGVFPA KEIIQPTVVD 544 starfish_Acanthaster VGWGAPNGYY YQKAYLEFFT SRETVALLQ VLKYEPRVN YQIVNVG-R NITNAPELQP NAVTWGFPG REIIQPTVVD 544 starfish_Acanthaster VGWGAPGGYC KQKAYLEFFT SRETVALLQ VLKYEPRVN YQIVNVGC-R NITNAPGVPA KEIIQPTVVD 544 starfish_Acanthaster VGWGAPGGYV FQKAYLEFFT SRETVALLQ VLKYVEPRVN YQIVNVG-R NITNAPGVPPA KEIIQPTVVD 545 starfish_Acanthaster VGWGAPGGYV FQKAYLEFFT SRETVALLQ VLKYVEPRVN YQIVNVG-R NITNAPGVPPA KEIIQPTVVD 545 starfish_Acanthaster VGWGAPGGYV FQKAYLEFFT SRETVALLQ VLKYVEPRVN YQIVNVG-R NITNAPDQP NAVTWGFPG REIIQPTVVD 545 starfish_Acanthaster VGWGAPGGYV FQKAYLEFFT SRETVALLQ VLKKYELRVN YHVVNHNGKE DITNTDRQPPI NAVTWGFPG REIIQPTVVD 545 starfish_Acanthaster PISFMFWKDE AFALWIEQWG KLYEESPSR TIIQYIHDNY FLVNLVDNFF PLDSC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Callus Human_6FCX	RR P L PWA -           RR - P L PWA -           I S - RS L PWR -           I S - RS L PWR -           I R P R P L PWK -           RR P L PWA -           RR P L PWA -           MWGEELTSEA	LEVPQETRDV and LSAHPKRE - LSAHPKRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKRKV 420 SVFEVFVL	E D V R P I FWAS E D V R P I FWAS S G E P N R N G H K	I RNYGVELAV ASO RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ ASO VTCL	EWDEFPNGRW DWDEFPNGRW GWEDFPQGRW GWEDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW	UNDEL HEYTL GNSSSPAFGE GNSSSPAFGE GDSRSASYGA GNSSSPAFGD GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE HEELLRVNRQ	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 DKSPKEALLQ 383 SKSPKEALLQ 480 INGKPSSDPI 461
worm_Loa MYGERLETLA DIQKVFYNFT TQUNNENGYK VSQLWNERH DGTRAFISLT KDQLWNAAN GFLITNSQPS VNGAPSTDPF 465 starfish_Danio MWGEELATGEE DVFEYFKCYL zebrafish_Danio MWGEELATGEE SVYEVFTNYT birds_Gallus MWGEELATGEE SVYEVFTNYT TGTNRSGHK VMCLPWND - DPLAPETNLL KDELEKVNRR GVLTINSQPN INGKASSDAT 461 birds_Gallus MWGEELTGEE SVYEVFTNYT TGTNRSGHK VMCLPWND - DPLAPETNLL KDELEKVNRR GUTTINSQPN INGKASSDAT 461 so 40 40 40 40 40 40 40 40 40 40	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I RP RP L PWK - RR P L PWA - RR P L PWA - RR P L PWA - MWG E E L T S E A	LEVPQETKDV and LSAHPKRRE - LSAHPKRRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - TSAHPKRRV - VSAHPKRRV 420 SVFEVFVLYL SVFEVFEHYL	E D V R P I FWA S E D V R P I FWA S E D V R P I FWA S E D V R P I FWS M E D V R P I FWS M E D V R P I FWA A E D V R P I FWA S E D V R P I FWA S S G E P N R N G H K S G E P N R N G H K	I RNYGVELAV 360 777 787 797 797 797 797 797 797	EWDEFPNGRW DWDEFPNGRW EWDEFPNGRW EWDEFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLL	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGD GNSSSPAFGD GNSSSPAFGE GNSSSPAFGE KEELLRVNRL	NREVATTEVL LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPN	KRLGIWNEDP 307 400 SKSPKEELLK 383 SKSPKEELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPREELLK 383 INGKPSDPI 461 INAKPSSDPV 461
stanish Adaitulate mwGEELACEE DYPEVFKLL DGEPKKIGHK VMKUMD- DELASETSKE KOLLSINK GILTINSGPA VIOAPSTDPV 466 zebrafish_Danio MWGEELACEE SVYEVFTCTI birds_Gallus MWGEELACEE SVYEVFTCTI GGTNRSGHK VMCUPWND- DPLATETSLL KDELEKVNRR GILTINSGPA VIOAPSTDP1 461 500 Human_6FCX VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 mouse_Mus VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 Arabidopsis_thaliana VGWGGPVGYV YQKAYLEFFT SRETAEALLQ VLKTYELRVN YHLVNVKG-E NITNAPELQP NAVTWGVFPA KEIIQPTIVD 542 worm_Loa VGWGPSGGYV YQKAYLEFFT SRETAEALLD VLKTYELRVN YHLVNVKG-E NITNAPELQP NAVTWGVFPA KEIIQPTIVD 542 birds_Gallus FGWGPSGGYV FQKAYLEFFT SRETAEALLD VLKTYELRVN YHVNHKGE DITNTDRQP IAVTWGVFPA KEIIQPTIVD 544 starfish_Acanthaster VGWGAPAGYV YQKAYLEFFT SRENAALLD VLRDYP-SVT YHVVNHNGKE DITNTAPDQP NAVTWGVFPG REIIQPTIVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SNEIVTALLQ VLKKYERVN YQIVNVQG-R NITNAPDQP NAVTWGIFPG REIIQPTIVD 540 son mouse_Mus PISFMFWKDE AFALWIEQWG KLYEESPSR TIIQYIHDNY FLVNLVDNDF PLDNC-LWQV 599 mouse_Mus PISFMFWKDE AFALWIEQWG KLYEESPSR MIIQYIHDNY FLVNLVDNFF PLDSC-LWQV 599 Arabidopsis_thalian ASFNVWKDE AFGLWKHQWG HIYSEGSPSR QIINHIHDTF YLVNLVNDY QDDYRLFEV 605 zebrafish_Canthaster PISFMWKDE AFALWIEQWG KLYEESPSR MIIQYIHDNY FLVNLVDNDF PLDNC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR P L PWA - MWG E E L T S EA MWG E E L T S EA	LEVPQETRDV ano LSAHPKRE LSAHPKRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKRKV SVFEVFVLYL SVFEVFEHYL DIGEKFKELC	E D V R P I FWA S E D V R P I FWA S E D V R P I FWA S E D V R P I FWA M E D V R P I FWA S E D V R P I FWA S E D V R P I FWA S S G E PN R NG HK S G E PN R NG HK	I RNYGVELAV 360 360 367 177 178 178 178 178 178 178 17	EWDEFPNGRW DWDEFPNGRW GWEDFPNGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLL EPLAAETSLL	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSRSASYGA GNSSSPAFGD GNSSSPAFGE GNSSSPAFGE KEELLRVNRL KEELLRVNRL NEQLIKVNSL	NREVATTEVL LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LSDHQFS - L KDYHLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GFLTINSQPN GFLTINSQPS	KRLGIWNEDP 307 400 SKSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPREELLK 383 480 INGKPSDPV 461 INAKPSDPV 461 VNAERSDSPT 454
birds_Gallus MWGEELINGLE SVFEVTRUT I GQINKSUIR WIGHT BELANDE BELANDE KURK GILTINSON INGKSSUN 1401 birds_Gallus MWGEELINGLE SVFEVTRUT I GQINKSUIR WIGHT BELANDE BELANDE KURK GILTINSON INGKSSUN 1401 birds_Gallus MWGEELINGLE SVFEVTRUT I GGPNKNGHR VTCMPWD 540 Human_6FCX VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 Arabidopsis_thaliana VGWGPVGYV YQKAYLEFFT SRETAEALLQ VLKKYELRVN YHIVDVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 worm_Loa VGWGPPGGYC YQKAYLEFFT SRETAEALLQ VLKYELRVN YHIVDVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 starfish_Canthaster VGWGAPNGYV YQKAYLEFFT SQANAEVLWS LLPSYS-RLN CHIVNHDATI DWINNETIMP IAVTWGVFPA KEIIQPTVVD 544 starfish_Canthaster VGWGAPNGYV YQKAYLEFFT SRENAALLD VLRDYP-SVT HYVNHNKKE DITNIDQOP IAVTWGIFPG REIIQPTVVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SSENVTALLQ VLKKYEPRVN YQIVNVGG-R NITNAHDMQP NAVTWGIFPG REIIQPTVVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SSENVTALLQ VLKKYEPRVN YQIVNVGG-R NITNAHDMQP NAVTWGIFPG REIIQPTVVD 540 Arabidopsis_thaliana PASFNVKDE AFALWIEQWG KLYEESPSR TIIQYIHDNY FLVNLVDNFF PLDSC-LWQV 599 Arabidopsis_thaliana PASFNVKDE AFALWIEQWG KLYEESPSR MIIQYIHDNY FLVNLVDNFF PLDSC-LWQV 599 Starfish_Canthaster PISFMWKDE AFGLWKHQWG HIYSEGSPSR QIINHIHDTF YLVNLVNDY VQDPYLFEV 605 zebrafish_Danio PVSFMTWKDE AFGLWKHQWG HIYSEGSPSR MIIQYIHDNY FLVNLVDNFF PLDSC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR - P L PWA - MWG E E L TS EA MWG E E L TS EA C R - C R C R C R C R C R C R C R C R C	LEVPQETKDV 340 - LSAHPKRE - LSAHPKRE - RPANVFRTK - TTANHMRTK - TTANHMRTK - TSAHPKRKV - VSAHPKRKV - VSAHPKRV - SVFEVFVLYL DIQEKFKELC DIQEKFKELC DIQEKFKELC	E D V R P I FWAS E D V R P I FWAS S C E P N R N G Y R S C E P N R N G Y R C O N N E N G V C W	I RNYGVELAV 360 RPKSYIYRTQ RPKSYIYRTQ RPKSYIFRTR RPKSYIYRTQ RPKSYIYRTQ 440 10 VTCL PWND VTCL PWND	EWDEFPNGRW DWDEFPNGRW GWEDFPNGRW DWDDFPNGRW DWDDFPNGRW BWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLL DGLQPETRII DGLQPETRII	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLIKVNSK	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 5KSSKEELQQ 388 SKSPKEALLQ 383 5KSPKEALLQ 383 5KSPKEALLQ 383 1NGKPSSDPI 461 INGKPSSDPI 461 INGKPSSDPI 461 VNAERSDSPT 455 VNAERSDSPT 465
Side_datas       Side_datas <td>Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Callus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster</td> <td>RR P L PWA - RR - P L PWA - IS - RS L PWR - IR - A L PWRS IR PR P L PWK - RR - P L PWA - RR - P L PWA - MWGEELTSEA MWGEELTSEE EWV V PLKSVE MYGEELETLA MWGEELACEE</td> <td>LEVPQETKDV and LSAHPKRE - LSAHPKRE - RPANYFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKRKV 200 SVFEVFVLYL DIQEKFKELC DIQEVFVNFI DVFEVFKCYL</td> <td>E D V R P I FWAS E D V R P I FWAS S C E P N R HG Y R L G N</td> <td>I RNYGVELAV 360 1 RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIFRTR RPKSYIYRTQ 440 1 VTCL PWND LKSS PWSE LKSS PWSE VSQL PWND VKRM PWND</td> <td>SMCRELLDSG EWDEFPNGRW DWDEFPNGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLI DGLAAETSLI DGLAETSLI DGLASETSLI</td> <td>LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRU KEELLRVNRU KEQLIKVNSK KDQLLWCNAN RDRLISINKR</td> <td>NREVATTEVL LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GILTINSQPS GFLTINSQPA</td> <td>KRLGIWNEDP 307 400 5KSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPREELLK 383 480 INGKPSSDP 461 VNAERSDSPT 454 VNCAPSTDPF 465 VNCAPSTDPF 466</td>	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Callus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster	RR P L PWA - RR - P L PWA - IS - RS L PWR - IR - A L PWRS IR PR P L PWK - RR - P L PWA - RR - P L PWA - MWGEELTSEA MWGEELTSEE EWV V PLKSVE MYGEELETLA MWGEELACEE	LEVPQETKDV and LSAHPKRE - LSAHPKRE - RPANYFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKRKV 200 SVFEVFVLYL DIQEKFKELC DIQEVFVNFI DVFEVFKCYL	E D V R P I FWAS E D V R P I FWAS S C E P N R HG Y R L G N	I RNYGVELAV 360 1 RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIFRTR RPKSYIYRTQ 440 1 VTCL PWND LKSS PWSE LKSS PWSE VSQL PWND VKRM PWND	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLI DGLAAETSLI DGLAETSLI DGLASETSLI	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRU KEELLRVNRU KEQLIKVNSK KDQLLWCNAN RDRLISINKR	NREVATTEVL LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GILTINSQPS GFLTINSQPA	KRLGIWNEDP 307 400 5KSPKEELLK 383 SKSPREELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPREELLK 383 480 INGKPSSDP 461 VNAERSDSPT 454 VNCAPSTDPF 465 VNCAPSTDPF 466
Human_6FCX_VGWGPSGGYV       FOKAYLEFFT       SRETAEALLQ_VLKKYELRVN_YHLVNVKG-E       NITNAPELQP       NAVTWGIFPG       REIIQPTVVD 540         Arabidopsis_thaliana       VGWGPSGGYV       FOKAYLEFFT       SRETVEALLQ_VLKTYELRVN_YHLVDVKG-E       NITNAPELQP       NAVTWGIFPG       REIIQPTVVD 540         Arabidopsis_thaliana       VGWGPSGGYV       YQKAYLEFFT       SRETVEALLQ_VLKTYELRVN_YHLVDVKG-E       NITNAPELQP       NAVTWGIFPG       REIIQPTVVD 540         Arabidopsis_thaliana       VGWGRPGGYC       YQKAYLEFFT       SRENAALLD       VLRVP-SVT       YMVNHNCKE       NAVTWGVFPG       KEIIQPTVVD 542         staffsh_Acanthaster       VGWGRPGGYC       YQKAYLEFFT       SRENAAALD       VLRVP-SVT       YHVNHNCKE       DITNTDRQQP       NAVTWGIFPG       REIIQPTVVD 544         staffsh_Canthaster       VGWGPAGGYV       YQKAYLEFFT       SSENVALLQ       VLKKYERVN       YHVNHNCKE       NITNAHDMOP       NAVTWGIFPG       REIIQPTVVD 540         birds_Gallus       FGWGPSGGYV       FQKAYLEFFT       SNEIVTALLQ       VLKKYERVN       YHVNHVNKG-Q       NITNAPDLQP       NAVTWGIFPG       REIIQPTVVD 540         schaftsh_Canthaster       FGWGPSGGYV       FQKAYLEFFT       SNEIVTALLQ       VLKKYERVN       YHVNLVDNDF       PLDNC-LWQV       599         Muman_6FCX       PVSFMFWKDE       AFALWIEQWG       K	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR - P L PWA - MWG E E L T S EA MWG E E L T S EA MWG E E L T S EA MWG E E L A C EE MWG E E L A C EE MWG E E L A C EE MWG E E L A C EE	LEVPQETKDV and LSAHPKRE - LSAHPKRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRRV - VSAHPKRRV A200 SVFEVFVLYL DIQEKFKELC DIQEKFKELC DIQEKFKCYL SVFEVFVFIN SVFEVFCYL	E D V R P I FWA S E D V R P I FWA S E D V R P I FWA S E D V R P I FWA M E D V R P I FWA M E D V R P I FWA M E D V R P I FWA S E D V R P I FWA S S G E P N R NG H K S G E P N R NG H K G G E P N R G H K G G E P N R G H K	IRNYGVELAV 360 TRPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIFRTR RPKSYIFRTR RPKSYIYRTQ TCL PWND UTCL PWND UKSQ PWND VKRMPWND VTCM PWND VTCM PWND VTCM PWND VTCM PWND	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLI DGIRAETSLI DGIRAETSLI DFLAAETSLI DPLAFTSLI	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGD GNSSSPAFGD GNSSSPAFGC GNSSSPAFGC KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLIKVNAK RDRLISINKR KDQLLWCNAN	NREVATTEVL LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYHLFYLK GILTINSQPN GILTINSQPS GFLTINSQPA GYLTINSQPA GYLTINSQPA	KRLGIWNEDP 307 400 SKSPKEELLK 383 SRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 INGKPSDPI 461 INAKPSDPI 461 VNGAPSTDPF 465 VNGAPSTDPF 466 INGKASSDAI 461 INGKPSDPI 461
Human_6FCX VGWGPSGGYV FQKAYLEFFT SRETAEALLQ VLKYELRVN YHLVNVKG-E NITNAPELQP NAVTWGIFPG REIIQPTVVD 540 Arabidopsis_thaliana VGWGFVGYV FQKAYLEFFT SRETVEALLQ VLKTYELRVN YHLVNVKG-EQ WVSNTVQADV NAVTWGVFPA KEIIQPTIVD 532 worm_Loa VGWGFVGCYC YQKAYLEFFT SQANAEVLWS LLPSYS-RLN CHIVNHDATI DWTNNETTMP IAVTWGVFPG KEIIQPTIVD 544 starfish_Canthaster VGWGAPAGGYV YQKAYLEFFT SRENAALLD VLRDYF-SVT YHVVNHNGKE DITNTDRQP IAVTWGGVFPG KEIIQPTIVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SSENVTALLQ VLKKYEPRVN YQIVNVQG-R NITNAHDMQP NAVTWGIFPG REIIQPTIVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SSENVTALLQ VLKKYEPRVN YQIVNVQG-R NITNAHDMQP NAVTWGIFPG REIIQPTIVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SNEIVTALLK VLKKYELRVN YHIVNVKG-Q NITNAPDLQP NAVTWGIFPG REIIQPTIVD 540 son for	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Callus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R P P L PWK - RR P L PWA - RR P L PWA - MWGEELTSEE EWV VP LKSVE MYGERLETLA MWGEELASEE MWGEELMSEE MWGEELTGEE	1         340           -         LSAHPKRE           -         LSAHPKRE           -         RPANYFRTK           FGENHPIRCK         TSAHPKRKV           -         VSAHPKRKV           -         VSAHPKRKV           SVFEVFVLYL         TSVFEVFHYL           DIQEKFKELC         DIQKYFVNFI           DVFEVFKLYL         SVFEVFVLYL           SVFEVFVLYL         SVFEVFVLYL           SVFEVFVLYL         SVFEVFVLYL	E D V R P I FWAS E D V R P I FWAS C E P N R NGHK C C P N R NGHK C C P N R NGHK	I RNYGVELAV 360 RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ RPKSYIYRTQ PKSYIYRTQ 10 10 10 10 10 10 10 10 10 10	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDFPNGRW DGLQPETRII DGLAETSLI DFLAFETSLL DFLAFETSLL	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEELLRVNSK KDQLLWCNAN RDRLISINKR KDQLEKVNRR KEQLEKVNRR	NREVATTEVL LKDYYLFYLK LSDHQFS R LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GFLTINSQPS GFLTINSQPS GILTINSQPN GILTINSQPN	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKDQLR 385 5KSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKELLK 383 1NGKPSSDP1 461 INAKPSSDP1 461 INAKPSSDP1 461 INGKPSTDP1 461 INGKPSTDP1 461
mouse_Mus_VGWGPSGGYV_FQKAYLEFFT_SRETVEALLQ_VLKTYELRVN_YHIVDVKG-E_NITNAPELQP_NAVTWGIFPG_REIIQPTVVD 540 Arabidopsis_thaliana_VGWGGPVGYV_YQKAYLEFFT_SRENDAVUE_KCKALP-SIT_YMAVNK-GEQ_WVSNTVQADV_NAVTWGVFPA_KEIIQPTIVD_532 worm_Loa_VGWGPRGYV_YQKAYLEFFT_SQANAEVLWS_LLPSYS-RLN_CHIVNHDATI_DWTNNETTMP_IAVTWGVFPG_SEIAQPTIVD_543 starfish_Aranthaster_VGWGPNGYV_YQKAYLEFFT_SRENAAALLD_VLRDYP-SVT_YHVVNHNGKE_DITNTDRQOP_IAVTWGVFPG_SEIAQPTVVD_545 zebrafish_Danio_VGWGPAGGYV_FQKAYLEFFT_SRENAAALLD_VLRDYP-SVT_YHVVNHNGKE_DITNTDRQOP_IAVTWGIFPG_REIIQPTVVD_540 birds_Gallus_FGWGPSGGYV_FQKAYLEFFT_SRENVALLQ_VLKKYEPRVN_YQIVNVQG-R_NITNAHDMQP_NAVTWGIFPG_REIIQPTVVD_540 	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus	RR P L PWA - RR - P L PWA - I S - RS L PWR - VR A L PWRS I R PR P L PWK - RR P L PWA - RR - P L PWA - MWGEELTSEE MWGEELTSEE MWGEELACEE MWGEELASEE MWGEELTGEE	LEVPQETRD 300 -LSAHPKRE -LSAHPKRE -RPANVFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRKV -VSAHPKRKV -VSAHPKRKV SVFEVFVLYL DIQEKFKELC DIQKVFVNFI DVFEVFKCYL SVFEVFTYI SVFEVFTCYI 	E D V R P I FWAS E D V R P I FWAS S G E P N R HG Y R S G E P N R HG Y R C Q T N R S G H K T G Q T N R S G H K	I RNYGVELAV 360 777 787 787 787 787 787 787 78	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW GWEDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLU DGLQPETRII DGLASETSLU DPLAFTSLL DPLAFTSL	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLIKVNSK KDQLLWCNAN RDRLISINKR KEQLEKVNRR KEQLEKVNRR	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GILTINSQPN GILTINSQPN	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 VNGAPSDP1 461 INGKPSDP1 461 INGKPSDP1 461 INGKPSTDP1 461 500 1
Arabidopsis_thaliana VGWG6PVGYV YQKAYLEFFT SKEKLDAVVE KCKAP-STT YMAVNK-GEQ WVSNTVQADV NAVIWGVFPA KETIQPTTVD 542 wom_Loa VGWG7PGCYV YQKAYLEFFT SQANAEVLWS LLPSYS-RIN CHIVNHDATI DWTNKTTMP IAVTWGVFPA KETIQPTVVD 543 starfish_Danio VGWG7PGCYV FQKAYLEFFT SSENVALLQ VLKKYEPRVN YQTVNHQG-R NITNAHDMQP NAVTWG1FPG RETIQPTVVD 540 birds_Gallus FGWG7SGYV FQKAYLEFFT SSENVALLQ VLKKYEPRVN YQTVNVG-C R NITNAHDMQP NAVTWG1FPG RETIQPTVVD 540 starfish_Danio VGWG7PGCYV FQKAYLEFFT SSENVALLQ VLKKYEPRVN YQTVNVG-C R NITNAHDMQP NAVTWG1FPG RETIQPTVVD 540 second Starfish_Comparison of the start of t	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Callus Human_6FCX Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Callus Human_6FCX	RR P L PWA - RR P L PWA - IS - RS L PWR - IR - A P WRS IR PR P L PWK - RR P L PWA - RR P L PWA - MWGEELTSEE EWV VP LKSVE MWGEELTSEE MWGEELTA WGEELTGEE VGWGPSGGYV	LEVPQETRB -LSAHPKRRE -LSAHPKRRE -RPANYFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRRV -VSAHPKRRV -VSAHPKRV -VSA	E D V R P I FWAS E D V R P I FWAA E D V R P I FWAA E D V R P I FWAS E D V R P I FWAS SCE P N R NGHK SCE P N R NGYK D C P N R SGHK T G P N K NGHR S R E T A E A L L Q	I RNYGVELAV 360 777 787 787 787 787 787 787 78	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDFFFNGRW EWDFFFNG EWDFFFFNG EWDFFFNG EWDFFFFNG EWDFFFFNG EWDFFFFNG EWDFFFFNG EWDFFFFNG EWDFFFFFNG EWDFFFFNG EWDFFFFNG EWDFFFFNG EWDFFFFNG EWDFFFFFNG EWDFFFFFFFFFNG EWDFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNR KEELLRVNRK KDQL LWCNAN RDRL I SINKR KDQL LKVNSK KDQL LWCNAN RDRL I SINKR KDQL LKVNRR KDQL LKVNRR	NREVATTEVL LKDYYLFYLK LSDHQFS - R LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GILTINSQPS GILTINSQPA GYLTINSQPN GILTINSQPN NAVTWGIFPG	K R L G I WN E D P 307 400 5K S P K E E L L K 383 S K S P R E E L L K 383 PR A R D K K L QQ 382 G Q P T K D E Q L R 385 S K S S K E E L L Q 383 S K S P K E A L L Q 383 S K S P R E E L L Q 383 S K S P R E E L Q 384 100 K P S S D P I 461 I N A K P S S D P V 461 I N A K P S T D P V 466 I N G K P S T D P V 466 I N G K P S T D P V 466 I N G K P S T D P I 461 I N G K P S T D P I 460 I N G K P S T D P I 460 I N G K P S T D P I 460 I N G K P S T D P I 460 I N G K P S T D P I 460 I N G K P S T D P I 400 S 600 R E I L Q P I V V D 540
Worm_Lob VOWGRPNGUTU YQKAYLEFFT SQANAEVLWS LLYRDYP-SYT YHVVNHNGKE DITNTDRQP TAVTWGPFG SETAQFTVVD 544 zebrafish_Danio VGWGPNGUY YQKAYLEFFT SRNAALLD VLRDYP-SYT YHVVNHNGKE DITNTDRQP TAVTWGPFG AETIQPTVVD 540 birds_Gallus FGWGPSGYV FQKAYLEFFT SSENVTALLQ VLKKYEPRVN YQTVNQG-R NITNAPDLQP NAVTWGTFPG RETIQPTVVD 540 	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R PR P L PWK - RR P L PWA - RR - P L PWA - MWG E E L T S EA MWG E E L T S EA MWG E E L A C EE MWG E E L A C EE MWG E E L A C EE WWG E E L T G EE WWG E C C C G C Y VGWG P S G G Y V VGWG P S G G Y V	LEVPQETKB0 300 -LSAHPKRE -LSAHPKRE -RPANYFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRKV -VSAHPKRV -VSAHPKRV SVFEVFVLYL DIQEKFKELC DIQKVFVNFI DIQEKFKELC DIQKVFVNFI SVFEVFCYI SVFEVFTCYI - SVFEVFTCYI - SVFEVFECFT - - - - - - - - - - - - -	E D V R P I FWAS E D V R P I FWAS E D V	I RNYGVELAV 360 RPKSYIYRTQ RPKSYIYRTQ RPKSYIFRTR RPKSYIFRTR RPKSYIYRTQ RPKSYIYRTQ 400 41 VTCL PWND VTCL PWND-	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW GWEDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDFPNGRW EWDFPNGRW WDWFPNGRW EWDFPNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDWFFNGRW WDFFNGRW WDWFFNGRW WDFFNGRW WDFFNGRW WDFFNGRW WDFFNGRW WDFFNGRW WDDFPNGRW WDFFNGRW W	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GNSSSPAFGD GDSSSPAFGD GDSSSPAFGE MELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLLWVNAK KDQLLWVNAK KDQLLWVNAK KDQLLWVNAK KDQLLWVNAK KDQLLWVNAK KDQLLWVNAK STAN KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR	NREVATTEVL LKDYYLFYLK LSDHQFS R LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GILTINSQPS GILTINSQPN GILTINSQPN NAVTWGIFPG NAVTWGIFPG	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKKDQLR 385 SKSPKEELLK 383 SKSPKELQ 383 SKSPKELQ 383 SKSPKELLQ 383 SKSPKELLK 383 400 INGKPSSDP1 461 VNAERSDSPT 465 VNAERSDSPT 465 VNACPSTDP4 466 INGKASSDAI 461 INGKPSTDP1 461 S60 REIIQPITVVD 540 REIIQPITVVD 540
statish_Acattoriste VGWGPAGGYV FQKAYLEFFT SKENATALLO VLKKYEPRVN YQIVNYQG-R NITNAHDMQP NAVTWGIFPG REIIQPTVVD 540 birds_Gallus FGWGPSGGYV FQKAYLEFFT SKENYTALLK VLKKYELRVN YHIVNYGG-Q NITNAPDLQP NAVTWGIFPG REIIQPTVVD 540 600 600 Human_6FCX PVSFMFWKDE AFALWIEQWG KLYEEESPSR TIIQYIHDNY FLVNLVDNDF PLDNC-LWQV 599 Arabidopsis thaliana PASFNVWKDE AFALWIEQWG KLYEEESPSR MIIQYIHDNY FLVNLVDNEF PLDSC-LWQV 599 Arabidopsis thaliana PASFNVWKDE AFETWSRSWA NLYPEADPSR NLLEEVKNSY YLVSLVENDY INGDIFAY 590 wom_Loa PLSFRVWKDE AFGLWKHQWG HIYSEGSPSR QIINHIHDTF YLVNLVDNDF PLDNC-LWQV 599 starfish_Acanthaster PISFMVWKDE AFGLWKHQWG HIYSEGSPSR QIINHIHDTF YLVNLVDNDF PLDNC-LWQV 599 birds Gallus PVSFLSWKDE AFALWIEQWA KLYEEESPSR MIIQYIHDNY FLVNLVDNDF PLDNC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR - P L PWA - MWGEELTSEA MWGEELTSEA MWGEELASEE MWGEELASEE MWGEELASEE MWGEELASEE VGWGPSGGYV VGWGCPSGGYV	LEVPQETRD 300 -LSAHPKRE -LSAHPKRE -RPANVFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRKV -VSAHPKRKV -VSAHPKRKV -VSAHPKRV - SVFEVFVLYL DIQEKFKELC DIQKVFVNFI DVFEVFKCYL - SVFEVFTC - - - - - - - - - - - - -	E D V R P I FWAS E D V R P I FWAS S GE P N R NG V K G E P N R NG V K G C E P N R NG V K G C E P N R Y G H K T G C T N R S G H K T S G H K	I RNYGVELAV 360 777 787 787 787 787 787 787 78	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EPLAAETSLL EPLAAETSLL DPLAFETSLL DPLAFETSLL DPLAFETSLL DPLAFETSLL PHIVDVKG-E YMAVNK-GEQ	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLIKVNSK KDQLLWCNAN RDRLISINKR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR NITNAPELQP WVSNTVQADV	NREVAITTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GFLTINSQPS GILTINSQPN GILTINSQPN NAVTWGIFPG NAVTWGFPG	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKVDQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 480 INGKPSDPI 461 INGKPSDPI 461 INGKPSTDPI 461 INGKPSTDPI 461 INGKPSTDPI 461 INGKPSTDPI 461 Store 100 KEIIQPI VVD 540 KEIIQPI VVD 540 KEIIQPI VVD 540 KEIIQPI VVD 540
birds_Gallus FGWGPSGGYV FQKAYLEFFT SNEIVTALL& VLKKYELRVN YHIVNVKG-Q NITNAPDLQP NAVTWGIFPG REIIQPTVVD 540 500 600 600 Human_6FCX PVSFMFWKDE AFALWIEQWG KLYEEESPSR TIIQYIHDNY FLVNLVDNDF PLDNC-LWQV 599 mouse_Mus PISFMFWKDE AFALWIEQWG KLYEEESPSR TIIQYIHDNY FLVNLVDNEF PLDSC-LWQV 599 Arabidopsis_thaliana PASFNVWKDE AFALWIEQWG KLYEEGSPSR NLLEEVKNSY YLVSLVENDY INGDIFAV 590 worm_Loa PLSFRWWKDE AFALWIEQWG HIYSEGSPSR CLLEKIYNEY CLVTLVDNDY PKPTI-IFDQ 602 starfish_Acanthaster PISFMWKDE AFALWIEQWG HIYSEGSPSR MIIQYIHDNY FLVNLVNNDY VQDDYRLFEV 605 zebrafish_Danio PVSFLSWKDE AFALWIEQWA KLYEDESPSR MIIQYIHDNY FLVNLVNDDF PLDNC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa	RR P L PWA - RR P L PWA - IS - RS L PWR - IR - R P L PWA - IR - P L PWA - RR P L PWA - RR P L PWA - MWGEELTSEE EWV VP LKSVE MWGEELTSEE MWGEELTSEE MWGEELTGEE VGWGPSGGYV VGWGCPSGGYV VGWGCPVGYV	LEVPQETRID A A A A A A A A A A A A A	E D V R P I FWAS E D V R P I FWAS S C E P N R HG Y R L G N	I RNYGVELAV 360 370 371 372 374 375 375 375 375 375 375 375 375	EMDEFPNGRW DWDEFPNGRW GWEDFPQGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDFPNGRW EWDFFNGRW EWDFPNGRW EWDFFNGRW EWDFFNGRW EWDFFNGRW EWDFFNGRW EWDFFNGRW EWDFFNGRW EVDFFNGRW EWDFFNGRW EWDFFNGRW EVDFFNGRW EVDFFNGRW EVDFFNGRW EVDFFNGRW EVDFFNGRW EVDFFNGRW EVDFFNGRW EVDFFNGRW EVDFFNGRW COMFON EVGFFNGRW COMF EVGFNGRW COMF EVGFNGRW	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEQLIKVNSK KDQLUWCNAN RDRLISINKR KDELEKVNRR SITNAPELQP NITNAPELQP NITNAPELQP UNITNAPTMP	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPA AVTWGIFPG NAVTWGFPG NAVTWGFPA IAVTWGFPG	KRLGIWNEDP 307 400 5KSPKEELLK 383 SKSPKEELLK 383 SRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKELLK 383 450 INGKPSDPI 461 INAKPSDPV 461 VNAERSDSPT 454 VNCAPSTDPF 465 INGKASSDAI 461 INGKPSTDPI 461 INGKPSTDPI 461 S00 REIIQPTVVD 540 REIIQPTVVD 540 KEIIQPTVVD 544 ACLIOPTVVD 544
SND_decides       SND_transfer       600       620         Human_6FCX       PVSFMFWKDE       AFALWIEQWG       KLYEEESPSR       TIIQYIHDNY       FLVNLVDNDF       PLDNC-LWQV       599         mouse_Mus       PISFMFWKDE       AFALWIEQWG       KLYEEESPSR       TIIQYIHDNY       FLVNLVDNEF       PLDSC-LWQV       599         Arabidopsis_thaliana       PASFNVKDE       AFALWIEQWG       KLYEEESPSR       NILQYIHDNY       FLVNLVDNEF       PLDSC-LWQV       599         Arabidopsis_thaliana       PASFNVKDE       AFALWIEQWG       KLYEEGSPSR       NILQEVKNSY       YLVSLVENDY       INGDIFAV       590         worm_Loa       PLSFRVWKDE       AFGUWKHQWG       HIYSEGSPSR       CLLEKIYNEY       CLVTLVDNDY       PKPTI-IFDC       602         starfish_Acanthaster       PLSFRVWKDE       AFGUWHQWG       HIYSEGSPSR       MIIQYIHDNY       YLVNLVNNDY       VQDDYRLFEY       605         zebrafish_Danio       PVSFLSWKDE       AFALWIEQWA       KLYEDESPSR       MIIQYIHDNY       FLVNLVDNDF       PLDNC-LWQV       599         birds Gallus       PVSFLSWKDE       AFALWIEQWA       KLYEDESPSR       MIIQYIHDNY       FLVNLVNDDF       PLENC-LWQV       599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR - P L PWA - RR - P L PWA - MWG E E L TS EA MWG E E L TS EA MWG E E L A C EE WWG E L A C EE WWG E C S C S C V VGWG C S G G Y V VGWG C P G G Y V VGWG C P G G Y V VGWG C C Y V	LEVPQEIND 300 -LSAHPKRE -LSAHPKRE -LSAHPKRE -RPANYFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRV -VSAHPKRV -VSAHPKRV -VSAHPKRV -VSAHPKRV -VSAHPKRV - SVFEVFVLYL DIQEKFKELC DIQKVFVNFI SVFEVFCYI SVFEVFTNYI SVFEVFTCYI - FQKAYLEFFT - - - - - - - - - - - - -	E D V R P I FWAS E D V R P I FWAS E D V	I RNYGVELAV 360 RPKSYIYRTQ RPKSYIYRTQ RPKSYIFRTR RPKSYIFRTR RPKSYIFRTR RPKSYIYRTQ PKSYIYRTQ 400 100 100 100 100 100 100 100	EWDEFPNGRW DWDEFPNGRW GWEDFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW UNDGLPETRII DGLAAETSLI DFLAAETSLI DFLAAETSLI DFLAAETSLI DFLAFETSLI	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GNSSSPAFGD GDSSSPAFGD GNSSSPAFGE GNSSSPAFGE KELLRVNRQ KELLRVNRQ KELLRVNRL KDQLLWCNAN KDQLLWCNAN KDQLLWCNAN KDQLLWCNAN KDQLLWCNAN KDQLLWCNAN KDQLLWCNAN KTNAPELQP NITNAPELQP WVSNTVQADV WTNNTTMP	NREVAITTEVL LKDYYLFYLK LSDHQFS R LKDYYLFYLK LSDHQFS R LKDYYLFYLK LTDYYLFYLK LKDYHLFYLK GILTINSQPN GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPN AVTWGIFPG NAVTWGYFPA IAVTWGYFPG IAVTWGYFPG	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKVEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 384 100 100 100 100 100 100 100 10
Human_6FCX       PVSFMFWKDE       AFALWIEQWG       KLYEEESPSR       TIIQYIHDNY       FLVNLVDNDF       PLDNC-LWQV       599         mouse_Mus       PISFMFWKDE       AFALWIEQWG       KLYEEESPSR       MIIQYIHDNY       FLVNLVDNF       PLDNC-LWQV       599         Arabidopsis       thaliana       PASFNVWKDE       AFETWSRSWA       NLYPEADPSR       NLIEEVKNSY       YLVSLVENDY       INGDIFAY       590         Worm Loa       PLSFRVWKDE       AFSGLWKHQWG       HIYSEGSPSR       QLIEKIVNEY       CLUEKIVNEY       CLUEKIVNEY       CLUEKIVNEY       602         starfish_Acanthaster       PISFMVWKDE       AFGLWKHQWG       HIYSEGSPSR       QIINHIHDTF       YLVNLVNNDY       VQDDYRLFEV       605         zebrafish_Danio       PVSFLSWKDE       AFALWIEQWA       KLYEEDSPSR       MIIQYIHDNY       FLVNLVNDDF       PLDNC-LWQV       599         birds <galus< td="">       PVSFLSWKDE       AFALWIEQWA       KLYEESSPSR       MIIQYIHDNY       FLVNLVNDNDF       PLDNC-LWQV       599</galus<>	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Canthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio	RR P L PWA - RR P L PWA - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR P L PWA - MWG E L T S EA WWG E L T S EA WWG E L T S EA WWG E L T S EA C C C C C C C C C C C C C C C C C C C	LEVPQEIND - LSAHPKRE - LSAHPKRE - LSAHPKRE - RPANYFRTK - TTANHMRTK - TTANHMRTK - TSAHPKRKV - VSAHPKRKV - VSAHPKRKV - VSAHPKRV - VSAHV - VSAHPKRV - VSAHVV - VSAHPKRV - VSAHPKRV - VSAHPKRV - VSAHPKRV - VSAH	E D V R P I FWAS E C D V R P I FWAS E D V R A I FWAS E D V R P I FWAS E D	I RNYGVELAV 360 777 787 787 787 787 787 787 78	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDFPNGRW EPLAAETSLL DGLQPETRII DGLQPETRII DGLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSLL DPLAAETSL DP	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLLWCNAN RDQLLWCNAN RDQLLSINKR KDQLLWCNAN RDQLLWCNAN RDQLLWCNAN RDQLLWCNAN RDQLLWCNAN DNITNAPELQP WVSNTVQADV DWTNNETTMP DITNTDRQQP NITNAHDMQP PITNAHDMQP	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GFLTINSQPS GFLTINSQPS GILTINSQPN GILTINSQPN NAVTWGIFPG NAVTWGFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFFF	KRLGIWNEDP 307 400 5KSPKEELLK 383 PRARDKKLQQ 382 GQPTKDEQLR 385 SKSSKEELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKEALLG 383 SKSPKEALLG 383 SKSPKEALLG 383 SKSPKEALLG 383 SKSPKEALG 383 SKS
Human_6FCX       PVSFMFWKDE       AFALWIEQWG       KLYEEESPSR       TIQYIHDNY       FLVNLVDNDF       PLDNC-LWQV       599         mouse_Mus       PISFMFWKDE       AFALWIEQWG       KLYEEESPSR       MIIQYIHDNY       FLVNLVDNDF       PLDSC-LWQV       599         Arabidopsig_thaliana       PASFNVWKDE       AFETWSRSWA       NLYPEADPSR       NLLEEVKNSY       YLVSLVENDY       INGDI - FAV       590         worm_Loa       PLSFRVWKDE       AFSAWL-NWS       SIYAEGTSSR       CLLEKIYNEY       CLVTLVDNDY       PKPTI - IFDC       602         starfish_Acanthaster       PISFMVWKDE       AFGLWKHQWG       HIYSEGSPSR       CLIEKIYNEY       CLVTLVDNDY       VQDYRLFEV       605         zebrafish_Danio       PVSFMYWKDE       AFALWIEQWA       KLYEDESPSR       MIIQYIHDNY       FLVNLVNDNF       PLDNC-LWQV       599         birds Gallus       PVSFLSWKDE       AFALWIEQWA       KLYEESPSR       MIIQYIHDNY       FLVNLVDNDF       PLENC-LWQV       599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus	RR P L PWA - RR - P L PWA - I S - RS L PWR - VR A L PWR S I R PR P L PWK - RR P L PWA - RR - P L PWA - MWGEELTSEE MWGEELTSEE MWGEELTSEE MWGEELACEE WGEPSGGYV VGWGPSGGYV VGWGCPSGGYV VGWGCPSGGYV VGWGCPGGYC VGWGCPGGYV VGWGCPGGYV VGWGCPGGYV FGWGPSGGYV	LEVPQETRD 340 -LSAHPKRRE -LSAHPKRRE -RPANYFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRKV -VSAHPKRKV 420 SVFEVFVLYL DIQEKFKELC DIQEVFKCYL DIQEKFKELC DIQEVFKCYL FQKAYLEFFT YQKAYLEFFT FQKAYLEFT FQKAYLEFT	E DVR PI FWAS E DVR PI FWAS S CE PNR HGYR C N TQDNN E NGVK D CE PNR YGHK T G C PNR	I RNYGVELAV 360 370 371 372 372 374 375 375 375 375 375 375 375 375	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDFPNGRW EWDFPNGRW EWDFPNGRW EWDFPNGRW EWDFPNGRW EWDFFNGRW EVGGG EFRI DGLASETSLL DFL	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNR KEELLRVNRL KEQL KVNSK KDQL LWCNAN RDRL ISINKR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR SITNAPELQP WVSNTVQADV DWTNNETTMP DITNTDRQP NITNAPDLQP	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPN NAVTWGIFPG NAVTWGFPG NAVTWGFPG NAVTWGFPG NAVTWGFPG	K R L G I WN E D P 307 400 400 5 K S P K E L L K 383 P R A R D K K L Q 382 G Q P T K D E Q L R 385 S K S P K E L L Q 388 S K S P K E L L Q 388 S K S P K E L L Q 388 S K S P K E L L K 383 480 I N G K P S D P 1 461 I N A K P S D P 1 461 I N A K P S D P T 465 V N G A P S T D P 1 461 1 N G K P S T D P 1 461 1 N G K P S T D P 1 461 1 N G K P S T D P 1 461 1 N G K P S T D P 1 461 1 R E I I Q P T V V D 540 R E I I Q P T V V D 540 R E I I Q P T V V D 540 R E I I Q P T V V D 540 R E I I Q P T V V D 540
MOUSE_MUS_FISEMEWKDE_AFALWIEQWG_KLYEEESPSR_MILQYIHDNY_FLVNLVDNEF_PLDSC-LWQV 599 Arabidopsis_thaliana_PAS_FNVWKDE_AFSWL-NWS_SIYAEGTSSR_CLLEKIYNEY_CLVTLVDNDY_NKPTI-IFDC_602 worm_Loa_PLSFRWWKDE_AFSAWL-NWS_SIYAEGTSSR_CLLEKIYNEY_CLVTLVDNDY_VKPTI-IFDC_602 starfish_Acanthaster_PISFMWWKDE_AFGLWKHQWG_HIYSEGSPSR_QIINHIHDTF_YLVNLVNNDY_VQDDYRLFEV_605 zebrafish_Danio_PVSFMYWKDE_AFALWIEQWA_KLYEDESPSR_MIIQYIHDNY_FLVNLVDNDF_PLENC-LWQV 599 birds Gallus_PVSFLSWKDE_AFALWIEQWA_KLYEEESPSR_MIIQYIHDNY_FLVNLVDNDF_PLENC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Acanthaster zebrafish_Acanthaster	RR P L PWA - RR P L PWA - IS - RS L PWR - VR A L PWR S IR P R P L PWK - RR P L PWA - RR P L PWA - RR P L PWA - RR - P L PWA - RR - P L PWA - RG C L TS EA MWG E E L TS EA MWG E E L TS EA C C C P S G C V V G C C P S G C V V G C C P S G C V V G C C C P S G C V F G C C P S G C V	LEVPQEIND 300 -LSAHPKRE -LSAHPKRE -LSAHPKRE FGENHPIRCK -TTANHMRTK FGENHPIRCK -TTANHMRTK SVFEVFVLYL SVFEVFHYL DIQEKFKELC DIQEKFKELC DIQEKFKELC DIQEKFKELC DIQEKFKELC JVFEVFKCYL SVFEVFTNYI FQKAYLEFFT FQKAYLEFT FQK	E D V R P I FWAS E D V R P I FWAS S C E P N R HG Y R G C P N R HG Y R G C P N R Y G H T G Q T N R S G H T G Q T N R S G H S R E T A E A L L Q S R E T V E A L L Q S R E T V E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S R E T A E A L L Q S S E N T A L L K	I RNYGVELAV 360 370 370 371 372 374 375 375 375 375 375 375 375 375	SMCRELLDSG EWDEFPNGRW WDEFPNGRW EWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EWDCFPNGRW EVDCFPNGRW EVDCFPNGRW EVDCFPNGRW EVDCFPNGRW EVDCFPNGRW EVDCFPNGRW EVDCFPNGRW EVCCFF EPLAAETSLL DPLAFTSL PLAAETSL PLAAETSL PLAAETSL PLAAETSL PLAAETSL PLAAETSL PLAAETSL PLAAETSL PLAAETSL PLAFTSL PLAFFNGRW YHLVNVKG-E YHLVNVKG-E YHIVNVKG-Q YHIVNVGG-R YHIVNVKG-Q	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEQLIKVNSK KDQLLWCNAN RDRLISINKR KDQLLWCNAN RDRLISINKR KDQLLKVNSK COLEKVNRR SU NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPS GILTINSQPA GYLTINSQPN NAVTWGIFPG NAVTWGYFPA IAVTWGYFPA IAVTWGYFPA	K R L G I WN E D P 307 400 5 K S P K E E L L K 383 S K S P K E E L L K 383 PR A R D K K L QQ 382 G Q P T K D E Q L R 385 S K S S K E E L QQ 388 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 F K S D K E A L Q 7 450 I N G K P S D P I 461 I N G K P S D P I 461 I N G K P S T D P I 461 I N G K P S T D P I 461 I N G K P S T D P I 461 I N G K P S T D P I 461 I N G K P S T D P I 461 I N G K P S T D P I 461 I N G K P S T D P I 461 I N G K P S T D P I 461 S 00 R E I I Q P T V V D 540 R E I I Q P T V V D 544 A E I I Q P T V V D 540 R E I Q P T V V D 540
Worm Loa PLSFRVWKDE AFELWSKSWA NLTFEADF3K NLLTERDF3K NLTFEADF3K NL	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Acanthaster zebrafish_Acanthaster zebrafish_Acanthaster zebrafish_Acanthaster zebrafish_Gallus	RR P L PWA - RR P L PWA - I S - RS L PWR - I S - RS L PWR - VR A L PWRS I R PR P L PWK - RR P L PWA - MWG E E L TS EA EWV V PLKS VE MYG E E L A C EE MWG E E L A C EE MWG E E L A C EE WWG E E L A C EE WWG E S G G V V GWG P S G G V V GWG C P S G G V V GWG C P S G G V V GWG C P S G G V C WGG C C C C C C C C C C C C C C C C C	LEVPQEIND A A A A A A A A A A A A A	E DVR PIFWAS EDVR PIFWAS EDVR PIFWAS EDVR PIFWAS EDVR PIFWAS EDVR PIFWAS EDVR PIFWAS EDVR PIFWAS EDVR PIFWAS EDVR PIFWAS CEPNR NGHK SGE PNR NGHK GE PNR YGHK TGQ TNR SGHK TGQ TNR SGHK TGE PNK NGHR SRETAEALLQ SRETVEALLQ SKEK LDAVVE SRENAALLD SSEN VTALLQ SNE VTALLQ	I RNYGVELAV 360 777 787 787 797 797 797 797 797	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EVDFFNGRW EVDF	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLLKVNSK KDQLLWCNAN RDRLISINKR KDQLLWCNAN RTNAPELQP WVSNTVQADV WTNNETTMP DITNTDRQQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GFLTINSQPS GILTINSQPN GILTINSQPN NAVTWGIFPG NAVTWGIFPG NAVTWGFPG NAVTWGIFPG NAVTWGIFPG NAVTWGIFPG S99	KRLGIWNEDP 307 400 SKSPKELLK 383 PRARDKKLQQ 382 GQPTKLDQLR 385 SKSPKELLK 383 PRARDKLQQ 388 SKSPKELLQ 388 SKSPKELLQ 388 SKSPKELLK 383 INGKPSDPI 461 INGKPSDPI 461 INGKPSDPI 461 INGKPSDPI 461 INGKPSTDPI 461 INGKPSTDPI 461 S60 REIIQPTVVD 540 REIIQPTVVD 540 REIIQPTVVD 540 REIIQPTVVD 540
starfish_Acanthaster PISFMVWKDE AFSAWLFWWS SITAEOTSSK CULENTINET CUVIL/UNDT FNFTFTFDD 602 starfish_Acanthaster PISFMVWKDE AFSAWLFWG HIYSEGSPSR QLINHIHDTF YLVNLVNNDY VQDDYRLFEV 605 zebrafish_Danio PVSFMVWKDE AFSAWLFQWA KLYEDESPSR MILQYIHDNY FLVNLVDNDF PLDNC-LWQV 599 birds Gallus PVSFLSWKDE AFSAWLFQWA KLYEEESPSR MILQYIHDNY YLVNLVDNDF PLENC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsi_thali	RR P L PWA - RR P L PWA - I S - RS L PWR - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR - P L PWA - MWGEELTSEA MWGEELTSEA MWGEELASEE MWGEELASEE MWGEELASEE WGGPSGGYV VGWGCPSGGYV VGWGCPSGGYV VGWGCPSGGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV VGWGCPSGYV PVSFMFWKDE P I SFMFWKDE P I SFMFWKDE	LEVPQETRAD and LSAHPKRE - LSAHPKRE - RPANVFRTK FGENHPIRCK - TTANHMRTK - ISAHPKRKV - VSAHPKRKV - VSAHPKRV - SAHPKRV - SAHPKRV	E D V R P I FWAS E D V R P I FWAS E D V	I RNYGVELAV 360 777 787 787 787 787 787 787 78	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EWDTFPNGRW EVNVKG-E YHIVDVKG-E YHIVNVKG-Q YHIVNVKG-Q FLVNLVDNDF FLVNLVDNF	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLLWCNAN RDRLISINKR KDQLLWCNAN RDRLISINKR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR MITNAPELQP WYSNTVQADV WTNNETTMP DITNTAPQQP NITNAPLQP NITNAPLQP SAUGON	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GFLTINSQPS GFLTINSQPS GILTINSQPN GILTINSQPN GILTINSQPN AVTWGIFPG NAVTWGIFPG NAVTWGIFPG NAVTWGIFPG S99 599	KRLGIWNEDP 307 400 5KSPKELLK 333 PRARDKKLQQ 382 GQPTKDQLR 385 SKSSKELQQ 388 SKSPKEALLQ 383 SKSPKEALLQ 383 SKSPKELLK 383 480 INGKPSSDP 461 INGKPSSDP 461 VNAERSDSPT 4461 INGKPSSDP 461 VNGAPSTDPF 465 VNGAPSTDPF 465 VNGAPSTDPF 465 VNGAPSTDPF 465 SKSSD 41 461 INGKPSTDPI 461 Store 100 100 KEIIQPI VVD 540 KEIIQPI VVD 540 REIIQPI VVD 540 REIIQPI VVD 540
zebrafish_Danio PVSFMYWKDE AFALWIEQWA KLYEDESPSR MIIQYIHDNY FLVNLVDNDF PLDNC-LWQV 599 birds Gallus PVSFLSWKDE AFALWIEQWA KLYEEESPSR MIIQYIHDNY YLVNLVDNDF PLENC-LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa	RR P L PWA - RR P L PWA - IS - RS L PWR - VR A L PWR S IR P R P L PWK - RR P L PWA - RR P L PWA - RR P L PWA - RR - P L PWA - RR - P L PWA - RG - P L PWA -	LEVPQETRD 340 340 -LSAHPKRE -LSAHPKRE -RPANYFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRKV -VSAHPKRKV 420 SVFEVFVLYL DIQKVFVNFI DIQKVFVNFI DIQKVFVNFI DIQKVFVNFI DIQKVFVNFI FQKAYLEFFT FQKAYLEFFT FQKAYLEFFT FQKAYLEFFT FQKAYLEFFT FQKAYLEFFT FQKAYLEFFT FQKAYLEFFT SN AFALWIEQWG AFALWIEQWG AFALWIEQWG AFELWNAW	E D V R P I FWAS E D V R P I FWAS S CE P N R HG Y R L G N T Q D N N E N G V K G C F P N K N G H R S R E T A E A L L Q S R E T V E A L L Q S R E T V E A L L Q S R E T V E A L L Q S R E T V E A L L Q S R E T V E A L L Q S N E I V T A L L K K L Y E E S P S R N L Y P E A D P S R	I RNYGVELAV 360 370 371 372 374 375 375 375 375 375 375 375 375	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW GWEDFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EYLAAETSLM DGLQPETRII DFLAAETSLM DFLAAETSLM DFLAFETSLL DFLAFETSLL PLAFETSL QLAFETSL QLAFETSL QLAFETSL CHIVNVKG-E QHIVNVKG-E QHIVNVKG-Q FLVNLVDNF FLVNLVDNF FLVNLVDNF VLVSLVENDY	LVHGLHFYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRU KEELLRVNRU KEQLLWUNAN RDRLISINKR KDQLLWUNAN RDRLISINKR KDQLEKVNRR SITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPELQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP NITNAPDUQP	NREVATTEVL LKDYYLFYLK LSDHQFS R LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GFLTINSQPS GFLTINSQPS GILTINSQPS GILTINSQPA GYLTINSQPA GYLTINSQPA GYLTINSQPS NAVTWGIFPG NAVTWGYFPA IAVTWGYFPG NAVTWGIFPG NAVTWGIFPG S99 599 599	K R L G I WN E D P 307 400 5K S P K E E L L K 383 S K S P K E E L L K 383 P R A R D K K L QQ 382 G Q P T K D E Q L R 385 S K S S K E E L QQ 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L L Q 383 S K S P K E A L Q 7 N A E S D P I 461 I N A K P S D P I 461 I N A K P S T D P I V V D 540 R E I Q P I V V D
birds Gallus PVSFLSWKDE AFALWIEQWA KLYEEESPSR MILQVIHDNY YLVNLVDNDF PLENC - LWQV 599	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus	RR P L PWA - RR P L PWA - IS - RS L PWR - IS - RS L PWR - VR A L PWRS IR PR P L PWK - RR P L PWA - RR - P L PWA - RR - P L PWA - MWG E L TS EA MWG E L TS EA MWG E L AC EE MWG E L AC EE MWG E L AC EE WGG P SGG YV VGWG P SGG YV VGWG C P SGG YV VGWG C P SGG YV VGWG C P SGG YV FGWG P SGG YV P V SFM FWK DE P AS FN VWK DE P L S FR VWK DE P L S FR VWK DE	LEVPQEIND 300 -LSAHPKRE -LSAHPKRE -LSAHPKRE -RPANYFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRV -VSAHPKRV VSAHPKRV SVFEVFVLYL DIQEKFKELC DIQKVFVNFI DIQEKFKELC DIQKVFVNFI SVFEVFCYI SVFEVFCYI SVFEVFCYI FQKAYLEFFT YQKAYLEFFT YQKAYLEFFT YQKAYLEFFT AFALWIEQWG AFALWIEQWG AFETWSRSWA AFSAWL-NWS	E DVR PIFWAS EDVR	I RNYGVELAV 360 777 787 787 797 797 797 797 797	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW GWEDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW HUDYKG-E YHLVNVKG-E YLVNV	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GDSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE KELLRVNRQ KEELLRVNRQ KEELLRVNRL KDQLLWCNAN RDRLISINKR KDQLLWCNAN RDRLISINKR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR KEQLEKVNRR SOULS NITNAPELQP SOUNT	NREVATTEVL LKDYYLFYLK LSDHQFS R LKDYYLFYLK LKDYHLFYLK LKDYYLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GFLTINSQPS GILTINSQPN GILTINSQPN NAVTWGIFPG NAVTWGIFPG NAVTWGIFPG NAVTWGIFPG NAVTWGIFPG NAVTWGIFPG NAVTWGIFPG S99 599 590 602	K R L G I WN E D P 307 400 5K S P K E E L L K 383 S K S P R E E L L K 383 P R A R D K K L QQ 382 G Q P T K H D Q 182 S K S P K E A L QQ 388 S K S P K E A L Q S S D P 461 I N G K P S T D P 461 I N
	Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_Danio birds_Gallus Human_6FCX mouse_Mus Arabidopsis_thaliana worm_Loa starfish_Acanthaster zebrafish_banio birds_Gallus	RR P L PWA - RR P L PWA - I S - RS L PWR - I S - RS L PWR - VR A L PWR S I R P R P L PWK - RR P L PWA - RR - P L PWA - MWG E L TS EA WWG E L TS EA WWG E L TS EA WWG E L TS EA WWG E L TG EE WWG E C L TG EE WWG E L TG EE WWG E C L TG EE P L S FM WKD E P L S	LEVPQEIND 300 -LSAHPKRE -LSAHPKRE -LSAHPKRE -RPANVFRTK FGENHPIRCK -TTANHMRTK -ISAHPKRKV -VSAHPKRKV -VSAHPKRV -	E D V R P I FWAS E D V R P I FWAS E D V	I RNYGVELAV 360 777 787 787 787 787 787 787 78	SMCRELLDSG EWDEFPNGRW DWDEFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW DWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDDFPNGRW EWDJFPNGRW EWDJFPNGRW EWDJFPNGRW EPLAAETSLL DFLAAETSLL	LVHGLH FYTL GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE GNSSSPAFGE KEELLRVNRQ KEELLRVNRQ KEELLRVNRQ KEQLLWCNAN RDRLISINKR KDQLLWCNAN RDRLISINKR KEQLEKVNRR KEQLEKVNRQ NITNAPELQP WVSNTVQADV DWTNNETTMP DITNTAPELQP NITNAPELQP	NREVATTEVL LKDYYLFYLK LKDYYLFYLK LKDYYLFYLK LKDYHLFYLK LKDYHLFYLK LKDYYLFYLK GILTINSQPN GILTINSQPS GFLTINSQPS GFLTINSQPS GFLTINSQPN GILTINSQPN GILTINSQPN GILTINSQPN NAVTWGIFPG NAVTWGFPG NAVTWGFPG NAVTWGFPG NAVTWGFPG S99 599 599 599	K R L G I WN E D P 307 400 5K S P K E E L L K 383 PR A R D K K L Q 382 GQ P T K D C L R 385 S K S P K E E L L K 383 PR A R D K K L Q 382 S K S P K E A L Q 388 S K S P K E A L Q 383 S K S P K E A L Q 383 S K S P K E A L Q 383 S K S P K E A L Q 383 S K S P K E A L Q 383 K S P K E A L Q 383 K S P K E A L Q 383 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 183 S K S P K E A L Q 19

*Figure 5*Sequence alignment of Mus musculus, Gallus gallus, Danio rerio, Acanthaster planci, Arabidopsis thaliana, C.elegans. Important conserve resdidues are in boxes.



Figure.10.3D interaction digram of FAD in (a) Mus musculus,(b) Gallus gallus,(c) Danio rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans



Figure.11.3D interaction digram of SAM in (a) Mus musculus,(b) Gallus gallus,(c) Danio rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans



Figure 6: 2D Interaction diagram of residues to FAD – (a) Mus musculus,(b) Gallus gallus,(c) Danio rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans



Figure 7: 2D interaction digram of residues to SAM –(a) Mus musculus,(b) Gallus gallus,(c) Danio rerio,(d) Acanthaster planci, (e)Arabidopsis thaliana, (f) C.elegans

	alpha helix (%)	extend strand	beta turn	random coil	number of residues
human	38.72	12.35	5.79	43.1	656
Mus musculus	39.91	14.68	8.26	37.16	654
z.fish	38.57	13.41	5.64	42.38	656
c.elegans	39.82	15.23	7.09	37.84	663
Arabidopsis thaliana	38.72	17.34	9.34	34.51	594
Acanthaster planci	38.17	14.51	6.10	41.22	689
Gallus gallus	39.38	12.14	5.53	42.86	651

Table 6:Percentage of secondary structural elements in models