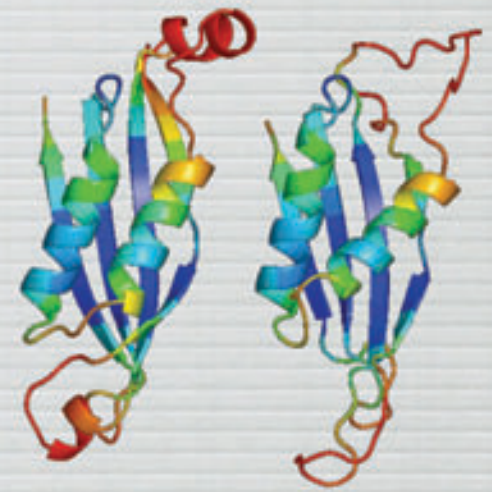
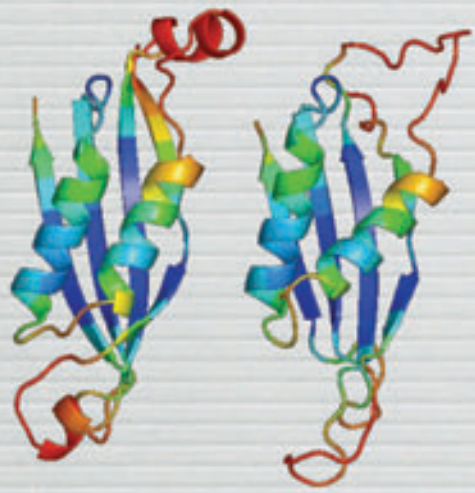
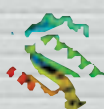
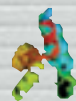




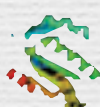
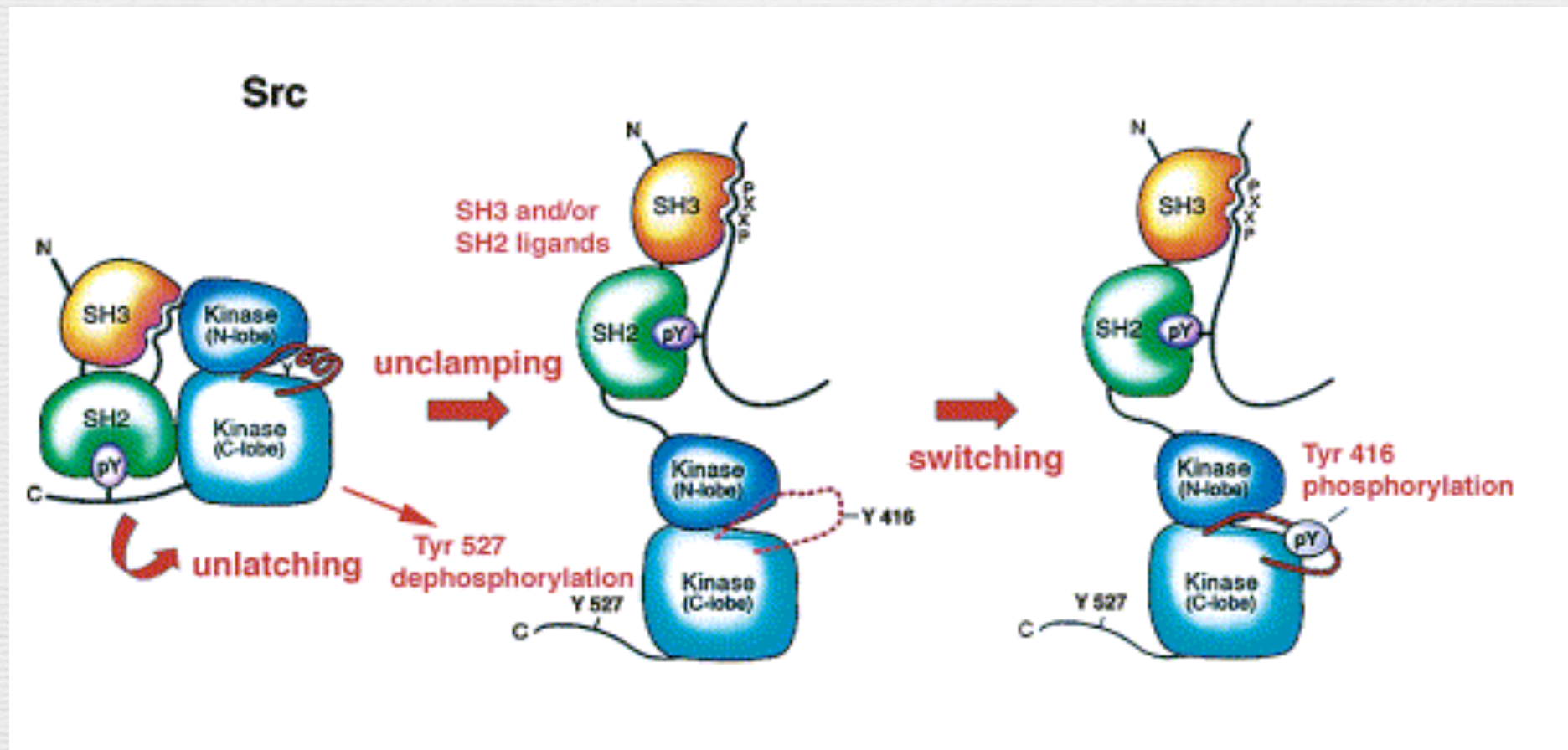
Protein domains, multidomain proteins and protein interactions.

To read:

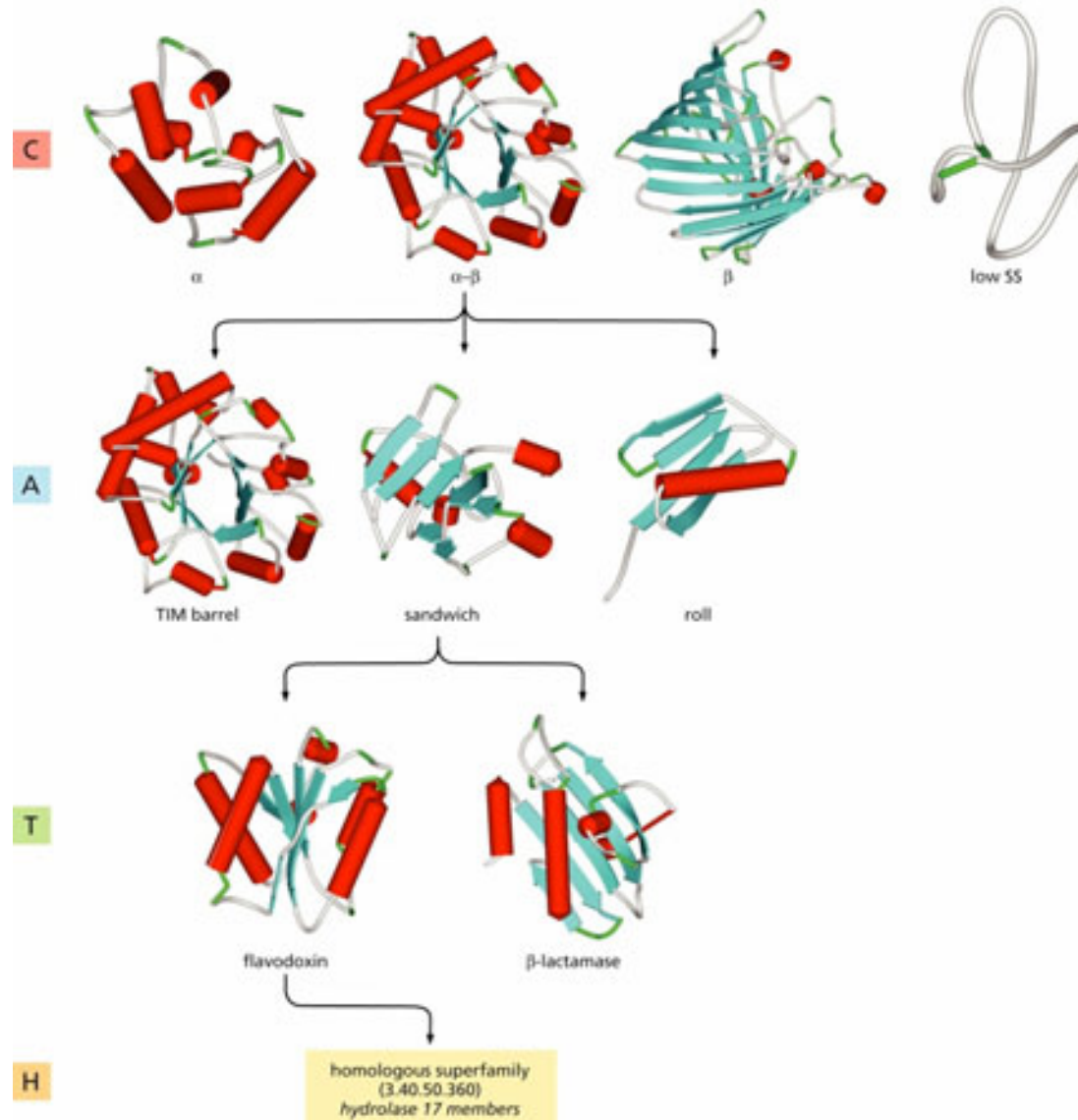
http://nook.cs.ucdavis.edu/~koehl/Teaching/ECS129/Reprints/SCOP_original.pdf
<http://archive.bioinfo.se/papers/16198373.pdf>



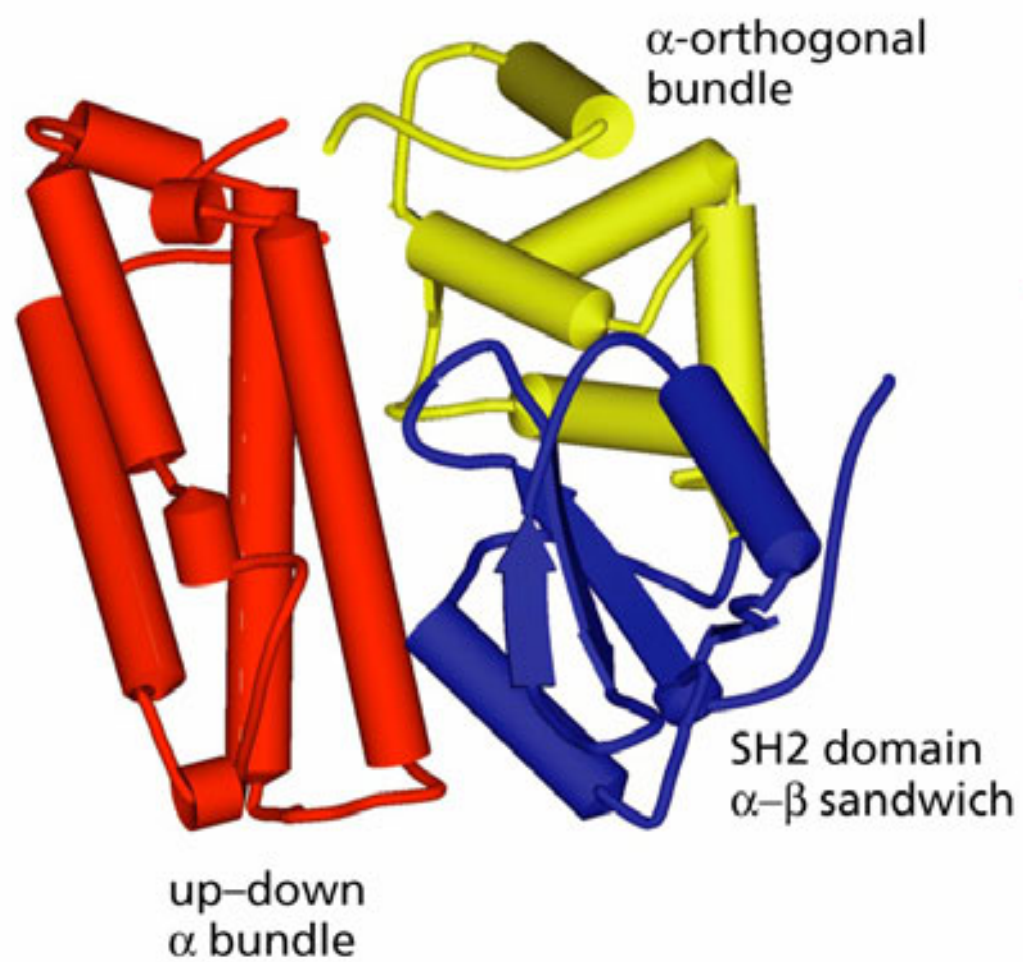
Multi domain proteins



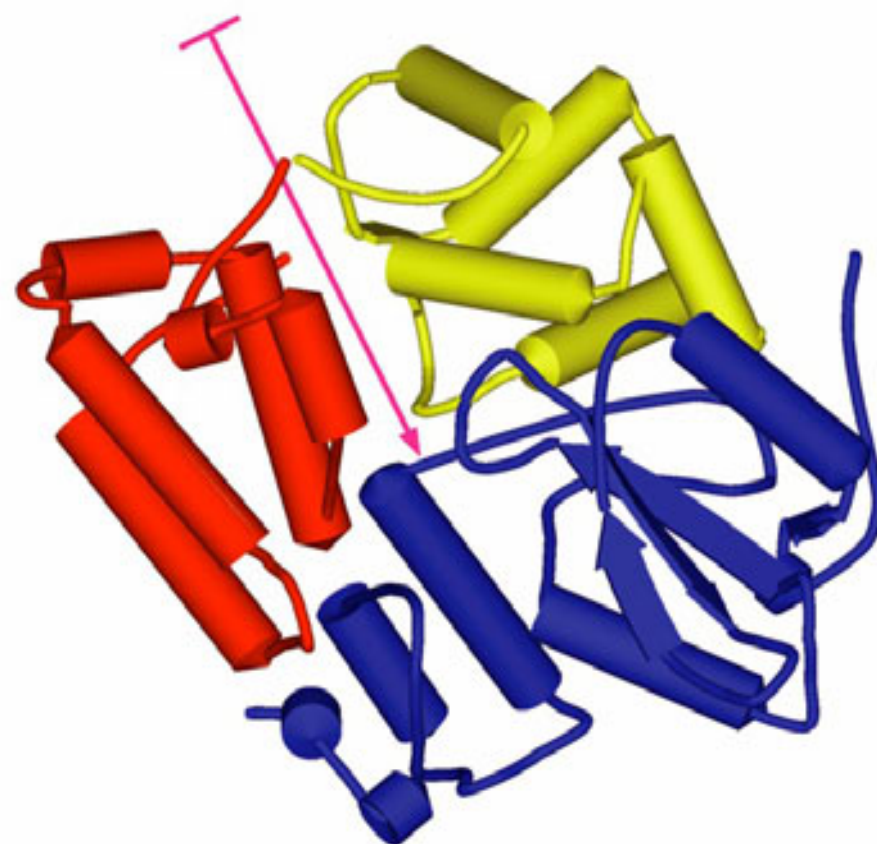
WHAT IS A FOLD SCOP/CATH



(A) X-ray definition



(B) DIAL prediction



Databases: SCOP & CATH

- Hierarchical classification of structures
 - Class
 - - / Architecture
 - Folds / Topology
 - Superfamily / Homology
 - Family
- SCOP manual CATH semi-manual

[Home](#) > [Top](#) > [C \[1\]](#)

Class 1



Mainly Alpha

Displaying 1-5 of 5 architectures

Click on the CATH code below to browse further

	Code	Representative	Architecture description	Image
☐	1.10	1cuk03	<i>Orthogonal Bundle</i>	
☐	1.20	1vdfA0	<i>Up-down Bundle</i>	
☐	1.25	1lrV00	<i>Horshoe</i>	
☐	1.40	1pprM1	<i>Alpha solenoid</i>	
☐	1.50	1cem00	<i>Alpha/alpha barrel</i>	

[View as XML](#)Class representative
1cuk03

Search

Go!

- ☐ PDB code
☐ CATH code
☐ General text

Goto...

[SSAP](#)
[DHS](#)
[Gene3D](#)
[PDBsum](#)

Navigation

[Home](#)
[Top of hierarchy](#)
[Up one level](#)

Help

Select a topic 



Home > Top > C [1] > A [10]

Architecture 1.10



Mainly Alpha



Orthogonal Bundle

Displaying 1-20 of 157 topologies

Click on the CATH code below to browse further

1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | Next

	Code	Representative	Topology description	Image
☐	1.10.8	1cuk03	<i>Helicase, Ruva Protein, domain 3</i>	
☐	1.10.10	1lea00	<i>Arc Repressor Mutant, subunit A</i>	
☐	1.10.12	1dubA2	<i>Lyase 2-enoyl-coa Hydratase, Chain A, domain 2</i>	
☐	1.10.15	1mpgA3	<i>3-methyladenine Dna Glycosylase II; Chain: A, domain 3</i>	
☐	1.10.20	1bfmA0	<i>Histone, subunit A</i>	
☐	1.10.30	1hryA0	<i>DNA Binding (I), subunit A</i>	
☐	1.10.40	1rlr01	<i>Ribonucleotide Reductase Protein R1, domain 1</i>	
☐	1.10.45	1ahuA4	<i>Vanillyl-alcohol Oxidase; Chain: A, domain 4</i>	
☐	1.10.60	2tdx02	<i>Diphtheria Toxin Repressor, domain 2</i>	
☐	1.10.100	6insE0	<i>Insulin-like, subunit E</i>	

View as XML

Architecture representative
1cuk03

Search

Go!

- ☐ PDB code
☐ CATH code
☐ General text

Goto...

SSAP
 DHS
 Gene3D
 PDBsum

Navigation

[Home](#)
[Top of hierarchy](#)
[Up one level](#)
[Next 20 entries](#)

Help


[Home](#) > [Top](#) > [C \[1\]](#) > [A \[10\]](#) > [T \[490\]](#)

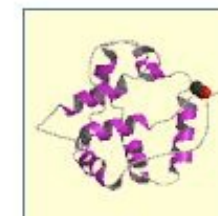
Topology 1.10.490

-  Mainly Alpha
 Orthogonal Bundle
 Globin-like

Displaying 1-4 of 4 homologous superfamilies

Click on the CATH code below to browse further

	Code	Representative	Homologous Superfamily description	Image
☐	1.10.490.10	1hlm00	<i>Globins</i>	
☐	1.10.490.20	1cpcA0	<i>Phycocyanins</i>	
☐	1.10.490.30	1colA0	<i>Colicin</i>	
☐	1.10.490.40	1ddt02	<i>Diphtheria toxin, middle domain</i>	

[View as XML](#)


Topology representative
1hlm00

Search

- ☐ PDB code
☐ CATH code
☐ General text

Goto...

[SSAP](#)
[DHS](#)
[Gene3D](#)
[PDBsum](#)

Navigation

[Home](#)
[Top of hierarchy](#)
[Up one level](#)

Help



Home > Top > C [1] > A [10] > T [490] > H [10]

Homologous Superfamily 1.10.490.10

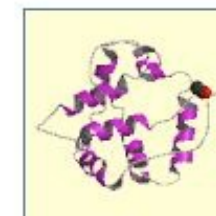
-  Mainly Alpha
-  Orthogonal Bundle
-  Globin-like
-  Globins

Displaying 1-15 of 15 sequence families

Click on the CATH code below to browse further

Code	Representative	Sequence family description	Image
☐ 1.10.490.10.1	1hlm00	<i>OXYGEN TRANSPORT</i>	
☐ 1.10.490.10.2	1hbg00	<i>OXYGEN TRANSPORT</i>	
☐ 1.10.490.10.3	1gdj00	<i>OXYGEN TRANSPORT</i>	
☐ 1.10.490.10.4	1babA0	<i>OXYGEN TRANSPORT</i>	
☐ 1.10.490.10.5	1flp00	<i>OXYGEN TRANSPORT</i>	
☐ 1.10.490.10.6	1lhs00	<i>OXYGEN STORAGE</i>	
☐ 1.10.490.10.7	1eca00	<i>OXYGEN TRANSPORT</i>	
☐ 1.10.490.10.8	1sctA0	<i>OXYGEN TRANSPORT</i>	
☐ 1.10.490.10.9	1mba00	<i>OXYGEN STORAGE</i>	
☐ 1.10.490.10.10	1ash00	<i>OXYGEN STORAGE</i>	

View as XML

Homologous superfamily representative
1hlm00

Search

Go!

- ☐ PDB code
- ☐ CATH code
- ☐ General text

Goto...

[SSAP](#)
[DHS](#)
[Gene3D](#)
[PDBsum](#)

Navigation

[Home](#)
[Top of hierarchy](#)
[Up one level](#)

Help

[Home](#) > [Top](#) > [C \[1\]](#) > [A \[10\]](#) > [T \[490\]](#) > [H \[10\]](#) > [S \[1\]](#)

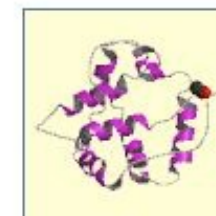
Sequence family 1.10.490.10.1

-  Mainly Alpha
-  Orthogonal Bundle
-  Globin-like
-  Globins
-  OXYGEN TRANSPORT

Displaying 1-2 of 2 non-identical relatives

Click on the CATH code below to browse further

Code	Representative	Non-identical relatives description	Image
☐ 1.10.490.10.1.1	1hlm00	OXYGEN TRANSPORT	
☐ 1.10.490.10.1.2	1h1b00	OXYGEN TRANSPORT	

[View as XML](#)

Sequence family representative
1hlm00

Search

- ☒ PDB code
- ☐ CATH code
- ☐ General text

Goto...

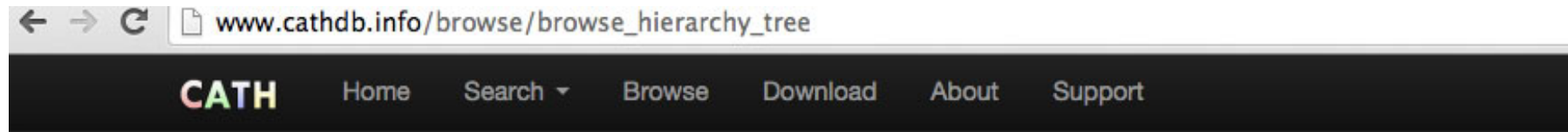
[SSAP](#)
[DHS](#)
[Gene3D](#)
[PDBsum](#)

Navigation

[Home](#)
[Top of heirarchy](#)
[Up one level](#)

Help

CATH

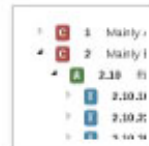


Browse CATH / Gene3D

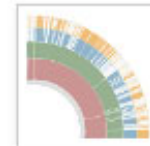
BROWSE LINKS

[Browse CATH Hierarchy](#)

[Compare CATH Superfamilies](#)



Tree



Sunburst

Details on the currently selected CATH node are displayed in the panel below

Top of CATH Hierarchy (4 Classes)

- ▷ **C** 1 Mainly Alpha
- ▷ **C** 2 Mainly Beta
- ▷ **C** 3 Alpha Beta
- ▷ **C** 4 Few Secondary Structures

CATH News

[Support](#)

[Jobs](#)

Get Started

[Documentation](#)

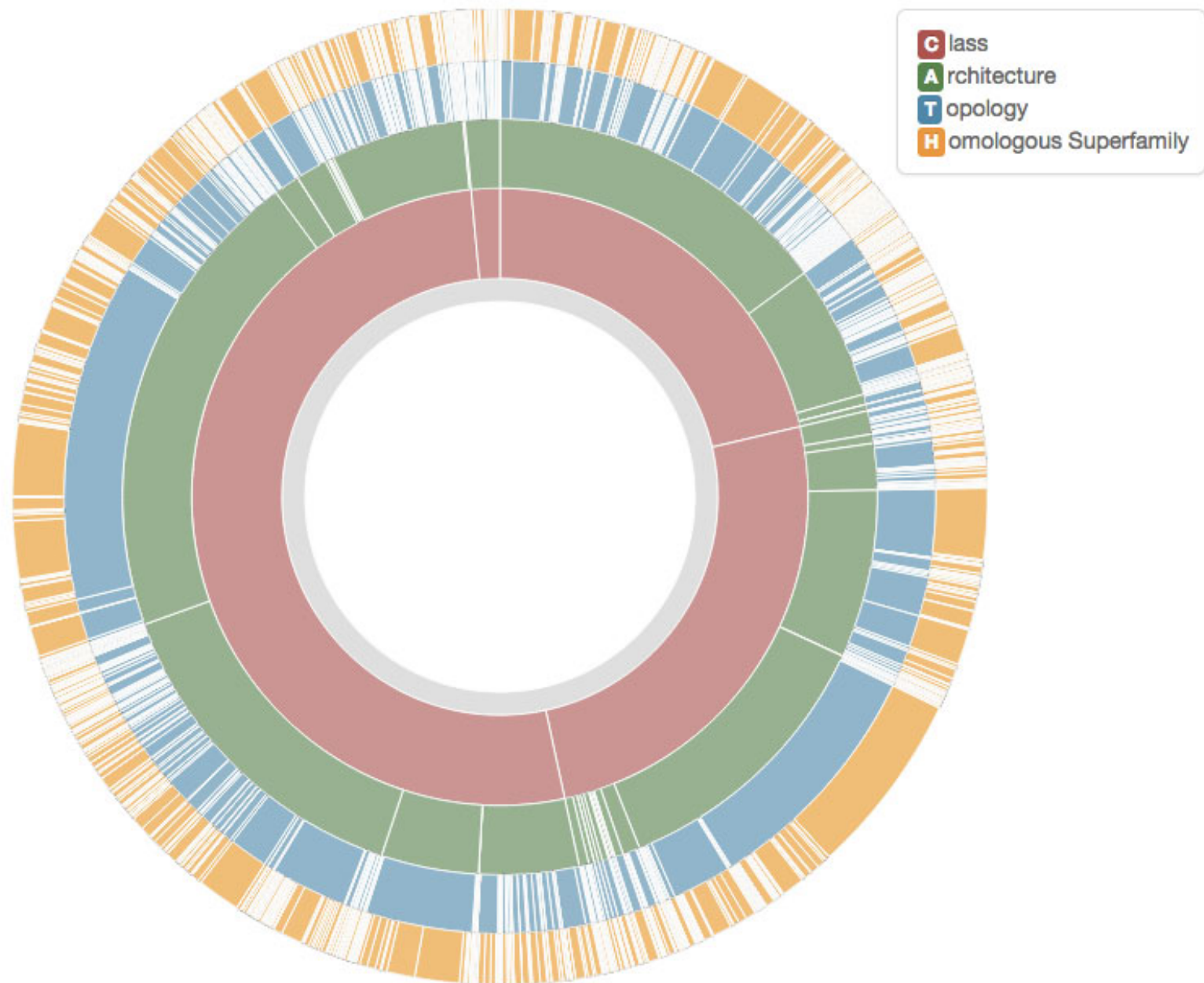
[Tutorials](#)

Download

[WebServices](#)

[Software](#)

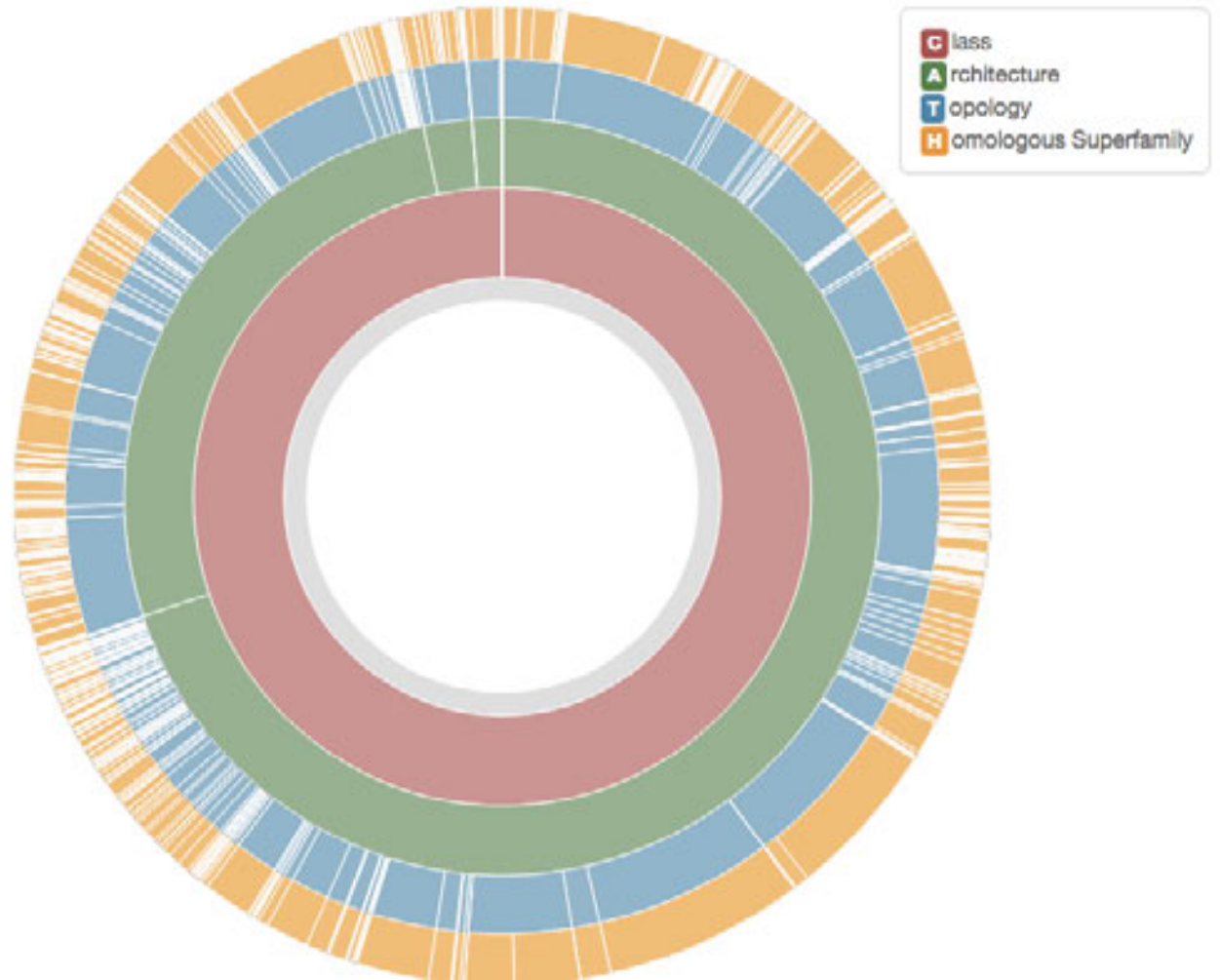
CATH



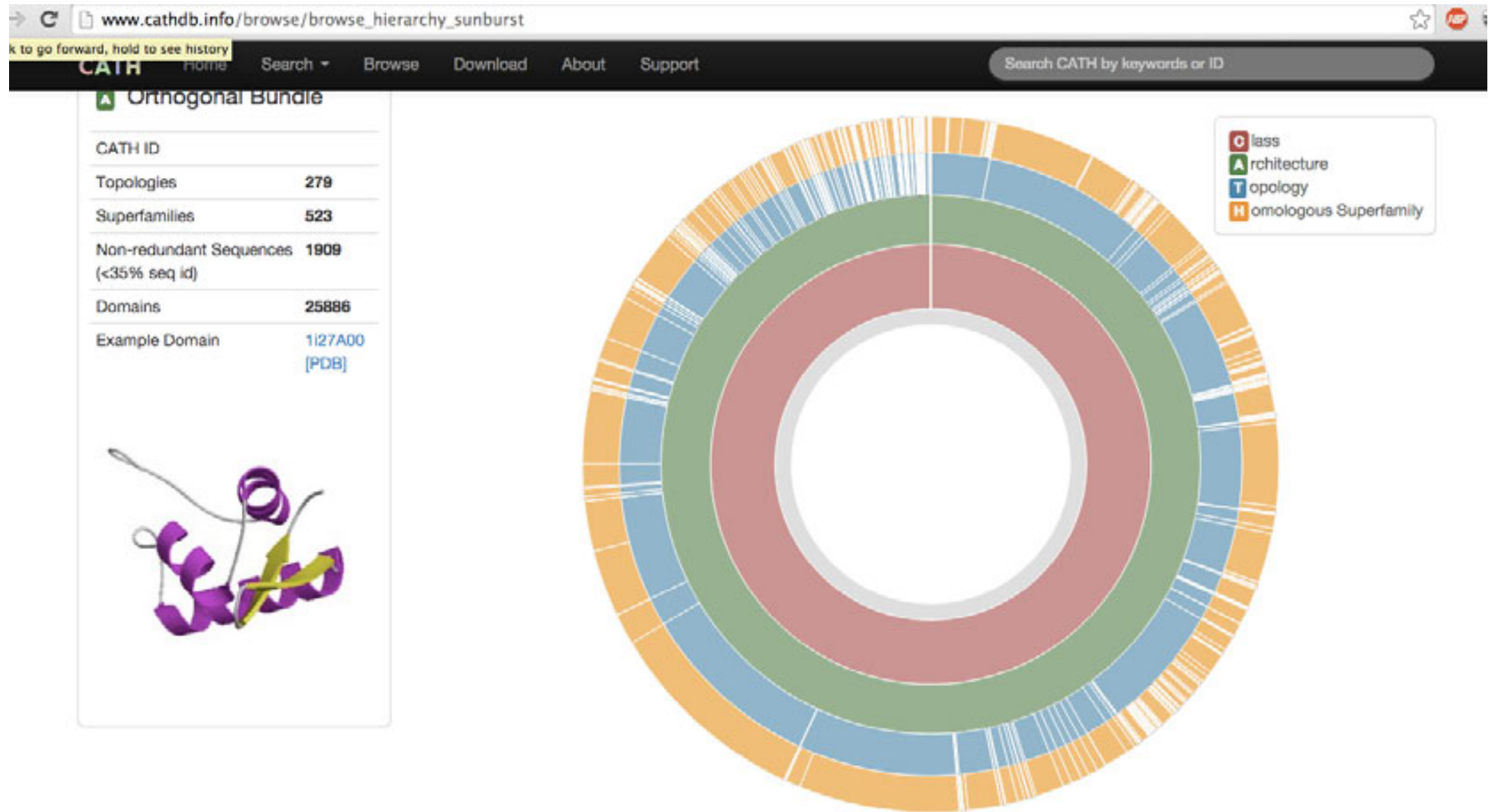
CATH

α Mainly Alpha

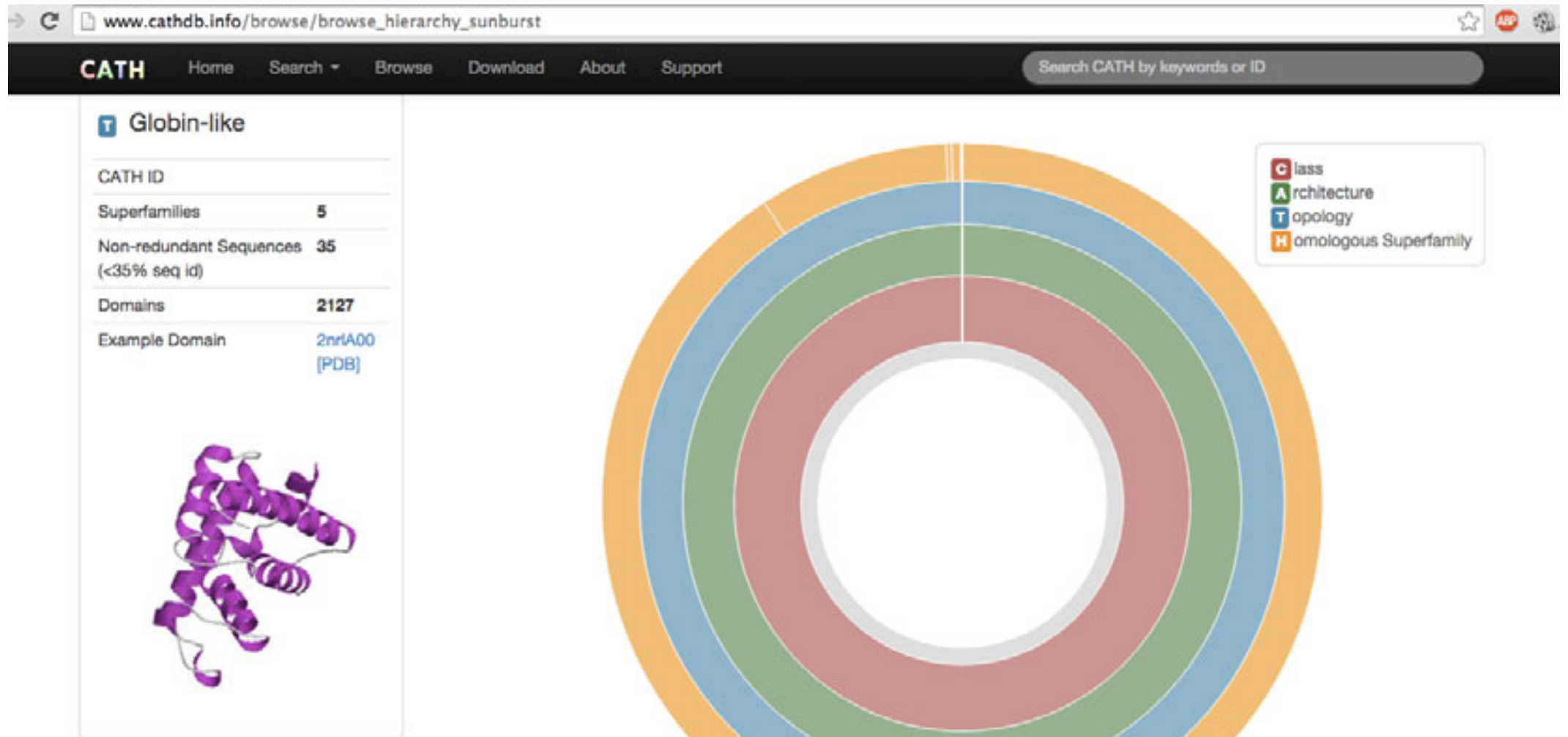
CATH ID	
Architectures	5
Topologies	386
Superfamilies	875
Non-redundant Sequences (<35% seq id)	2917
Domains	37038
Example Domain	1I27A00 [PDB]



CATH



CATH



CATH

← → ↻ www.cathdb.info/version/latest/superfamily/1.10.490.10 ☆   

CATH Home Search ▾ Browse Download About Support

CATH Superfamily 1.10.490.10

Globins

[Home](#) / [Superfamily 1.10.490.10](#)

SUPERFAMILY LINKS

Summary

- [Superfamily Superposition](#)
- [Classification / Domains](#)
- [Alignments](#)
- [Structural Neighbourhood](#)
- [Functional Annotations](#)
- [Taxonomy Browser](#)
- [Multi-Domain Organisation](#)

GO Diversity

Unique GO annotations



105 Unique GO terms >

EC Diversity

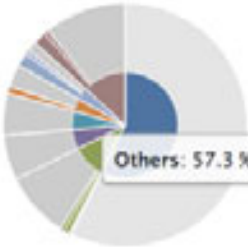
Unique EC annotations



13 Unique EC terms >

Species Diversity

Unique species annotations



4641 Unique species >

Superfamily Summary

A general summary of information for this superfamily.

Structures	
Domains:	2230
Domains (< 95% seq id):	160
Domains (< 35% seq id):	41
Unique PDBs:	955
Alignments	
Structural Clusters:	2

Functional Families

Overview of the Structural Clusters (SC) and Functional Families (FP) within this CATH Superfamily

Structural Diversity

Structural domains within this

Domain Organisation

View multi-domain architectures via

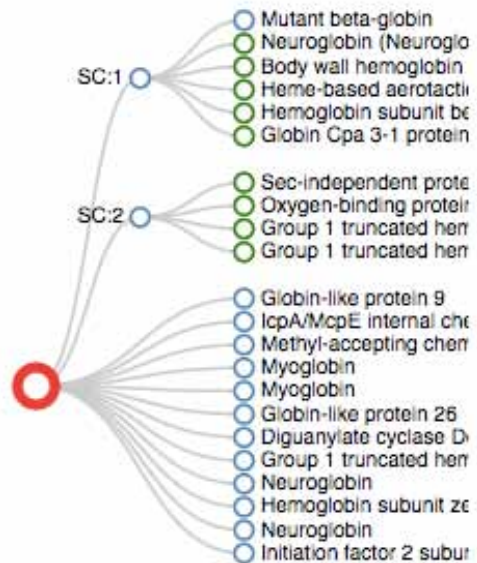
Enzyme Function

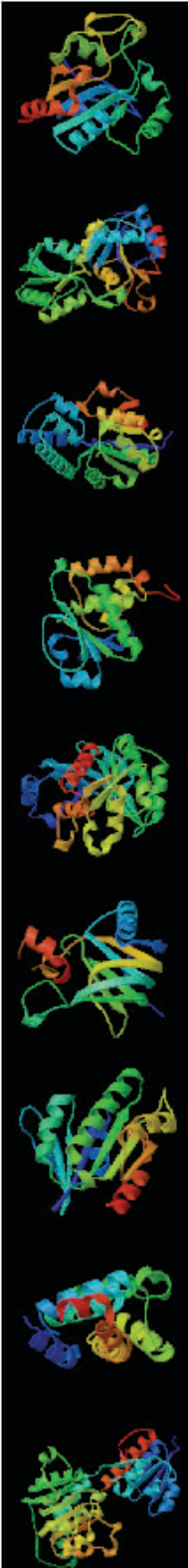
Evolution of Enzyme Function via

CATH - additional information

Functional Families

Overview of the Structural Clusters (SC) and Functional Families (FF) within this CATH Superfamily





SCOP

- Hierarchical classification of structures

- Class

- - / Architecture

- Folds / Topology

- Superfamily / Homology

- Family

- SCOP manual CATH semi-manual

Structural Classification of Proteins



Root: scop

Classes:

1. [All alpha proteins](#) (151)   
2. [All beta proteins](#) (111)   
3. [Alpha and beta proteins \(a/b\)](#) (117)   
Mainly parallel beta sheets (beta-alpha-beta units)
4. [Alpha and beta proteins \(a+b\)](#) (212)   
Mainly antiparallel beta sheets (segregated alpha and beta regions)
5. [Multi-domain proteins \(alpha and beta\)](#) (39)   
Folds consisting of two or more domains belonging to different classes
6. [Membrane and cell surface proteins and peptides](#) (12)   
Does not include proteins in the immune system
7. [Small proteins](#) (59)   
Usually dominated by metal ligand, heme, and/or disulfide bridges
8. [Coiled coil proteins](#) (5)   
Not a true class
9. [Low resolution protein structures](#) (17)  
Not a true class
10. [Peptides](#) (95)   
Peptides and fragments. Not a true class
11. [Designed proteins](#) (36)   
Experimental structures of proteins with essentially non-natural sequences. Not a true class

Enter [search](#) key:

Structural Classification of Proteins



Class: All alpha proteins

Lineage:

1. Root: [scop](#)
2. Class: [All alpha proteins](#)

Folds:

1. [Globin-like](#) (2)   
core: 6 helices; folded leaf, partly opened
2. [Long alpha-hairpin](#) (11)   
2 helices; antiparallel hairpin, left-handed twist
3. [Type I dockerin domain](#) (1)  
tandem repeat of two calcium-binding loop-helix motifs, distinct from the EF-hand
4. [LEM/SAP HeH motif](#) (4)  
helix-extended loop-helix; parallel helices
5. [Cytochrome c](#) (1)   
core: 3 helices; folded leaf, opened
6. [DNA/RNA-binding 3-helical bundle](#) (10)   
core: 3-helices; bundle, closed or partly opened, right-handed twist; up-and down
7. [RuvA C-terminal domain-like](#) (6)  
3 helices; bundle, right-handed twist
8. [Putative DNA-binding domain](#) (1)  
core: 3 helices; architecture is similar to that of the "winged helix" fold but topology is different
9. [Spectrin repeat-like](#) (7)  
3 helices; bundle, closed, left-handed twist; up-and-down
10. [Bacterial immunoglobulin/albumin-binding domains](#) (1)   
3 helices; bundle, closed, left-handed twist; up-and-down
11. [Peripheral subunit-binding domain of 2-oxo acid dehydrogenase complex](#) (1)   

Structural Classification of Proteins



Fold: Globin-like

core: 6 helices; folded leaf, partly opened

Lineage:

1. Root: [scop](#)
2. Class: [All alpha proteins](#)
3. Fold: [Globin-like](#)
core: 6 helices; folded leaf, partly opened

Superfamilies:

1. [Globin-like](#) (4)
2. [alpha-helical ferredoxin](#) (2)
contains two Fe4-S4 clusters

Enter [search](#) key: Search

Generated from scop database 1.61 with scopm 1.096 on Thu Oct 31 18:40:44 2002

[Copyright](#) © 1994-2002 The scop authors / scop@mrc-lmb.cam.ac.uk

Structural Classification of Proteins



Superfamily: Globin-like

Lineage:

1. Root: [scop](#)
2. Class: [All alpha proteins](#)
3. Fold: [Globin-like](#)
core: 6 helices; folded leaf, partly opened
4. Superfamily: [Globin-like](#)

Families:

1. [Truncated hemoglobin](#) (3)
lack the first helix (A)
2. [Nerve tissue mini-hemoglobin \(neural globin\)](#) (1)
lack the first helix but otherwise is more similar to conventional globins than the truncated ones
3. [Globins](#) (61)
Heme-binding protein
4. [Phycocyanin-like](#) (13)
*oligomers of two different types of homologous subunits
each subunit contains 2 additional helices at the N-terminus
binds a chromophore*

Enter [search](#) key:

Structural Classification of Proteins

Family: Globins

Heme-binding protein

Lineage:

1. Root: [scop](#)
2. Class: [All alpha proteins](#)
3. Fold: [Globin-like](#)
core: 6 helices; folded leaf, partly opened
4. Superfamily: [Globin-like](#)
5. Family: [Globins](#)
Heme-binding protein

Protein Domains:

1. Hemoglobin I
 1. [Ark clam \(*Scapharca inaequivalvis*\)](#) (14)  
 2. [Clam \(*Lucina pectinata*\)](#) (4)  
2. Trematode hemoglobin/myoglobin
 1. [Paramphistomum epiclitum](#) (2)  
3. Glycera globin
 1. [Marine bloodworm \(*Glycera dibranchiata*\)](#) (8)  
4. Myoglobin
 1. [Sperm whale \(*Physeter catodon*\)](#) (137)  
 2. [Sea hare \(*Aplysia limacina*\)](#) (7)  
 3. [Common seal \(*Phoca vitulina*\)](#) (1)  
 4. [Pig \(*Sus scrofa*\)](#) (17)  
 5. [Horse \(*Equus caballus*\)](#) (14)  
 6. [Human \(*Homo sapiens*\)](#) (1)  

Structural Classification of Proteins



Protein: Hemoglobin I from Ark clam (*Scapharca inaequivalvis*)

Lineage:

1. Root: [scop](#)
2. Class: [All alpha proteins](#)
3. Fold: [Globin-like](#)
core: 6 helices; folded leaf, partly opened
4. Superfamily: [Globin-like](#)
5. Family: [Globins](#)
Heme-binding protein
6. Protein: Hemoglobin I
7. Species: [Ark clam \(*Scapharca inaequivalvis*\)](#)

PDB Entry Domains:

1. [3sdh](#)   
complexed with hem
 1. [chain a](#)   
 2. [chain b](#)   
2. [3hbi](#)   
complexed with cmo, hem; mutant
 1. [chain a](#)   
 2. [chain b](#)   
3. [4sdh](#)   
complexed with hem
 1. [chain a](#)   
 2. [chain b](#)   
4. [1jzl](#)   
complexed with cmo, hem; mutant
 1. [chain a](#)   

SCOPe 2.0

← → ↻

SCOPe: Structural Classification of Proteins — extended. Release 2.04 (July 2014, new entries added 2014-09-11)

[Browse](#) [Stats & History](#) [ASTRAL Subsets](#) [Downloads](#) [Related Resources](#) [References](#) [Help](#) [About](#)

Welcome to the SCOPe website!

SCOPe is a database developed at the Berkeley Lab and UC Berkeley that extends version 1.75 of the SCOP database, the last SCOP release in the 1.x series, which was conceived at the MRC Laboratory of Molecular Biology. SCOPe classifies many structures released since SCOP 1.75 through a combination of automation and manual curation, and continuing to have the same accuracy as the hand-curated SCOP releases. SCOPe also incorporates and updates the ASTRAL database. For prior releases, click on the [Stats & History](#) tab above. For more info, click on the [About](#) tab above. New PDB entries were last added on **2014-09-11**; for more info click on the [Help](#) tab above.

Search SCOPe ([example](#)):

Classes in SCOPe 2.04:

1.  [a: All alpha proteins](#) [46456] (285 folds)
2.  [b: All beta proteins](#) [48724] (176 folds)
3.  [c: Alpha and beta proteins \(a/b\)](#) [51349] (148 folds)
4.  [d: Alpha and beta proteins \(a+b\)](#) [53931] (380 folds)
5.  [e: Multi-domain proteins \(alpha and beta\)](#) [56572] (68 folds)
6.  [f: Membrane and cell surface proteins and peptides](#) [56835] (57 folds)
7.  [g: Small proteins](#) [56992] (91 folds)
8.  [h: Coiled coil proteins](#) [57942] (7 folds)
9.  [i: Low resolution protein structures](#) [58117] (25 folds)
10.  [j: Peptides](#) [58231] (126 folds)
11.  [k: Designed proteins](#) [58788] (44 folds)

SCOPe 2.0



Search SCOPe ([example](#)):

Lineage for Class a: All alpha proteins

1. Root: [SCOPe 2.04](#)
2.  Class a: All alpha proteins [46456] (285 folds)

Folds:

1.  [a.1: Globin-like](#) [46457] (2 superfamilies)
core: 6 helices; folded leaf, partly opened
2.  [a.2: Long alpha-hairpin](#) [46556] (20 superfamilies)
2 helices; antiparallel hairpin, left-handed twist
3.  [a.3: Cytochrome c](#) [46625] (1 superfamily)
core: 3 helices; folded leaf, opened
4.  [a.4: DNA/RNA-binding 3-helical bundle](#) [46688] (14 superfamilies)
core: 3-helices; bundle, closed or partly opened, right-handed twist; up-and down
5.  [a.5: RuvA C-terminal domain-like](#) [46928] (9 superfamilies)
3 helices; bundle, right-handed twist
6.  [a.6: Putative DNA-binding domain](#) [46954] (1 superfamily)
core: 3 helices; architecture is similar to that of the "winged helix" fold but topology is different
7.  [a.7: Spectrin repeat-like](#) [46965] (16 superfamilies)
3 helices; bundle, closed, left-handed twist; up-and-down
8.  [a.8: immunoglobulin/albumin-binding domain-like](#) [46996] (11 superfamilies)
3 helices; bundle, closed, left-handed twist; up-and-down; mirror topology to the spectrin-like fold
9.  [a.9: Peripheral subunit-binding domain of 2-oxo acid dehydrogenase complex](#) [47004] (1 superfamily)
3 helices; bundle, closed, right-handed twist; up-and-down
10.  [a.10: Protozoan pheromone-like](#) [47013] (2 superfamilies)
3 helices; bundle, closed, left-handed twist, up-and-down
11.  [a.11: Acyl-CoA binding protein-like](#) [47026] (2 superfamilies)
core: 3 helices; bundle, closed, left-handed twist; up-and-down

SCOPe 2.0

SCOPe: Structural Classification of Proteins — extended. Release 2.04 (July 2014, new entries added 2014-09-11)

[Browse](#)[Stats & History](#)[ASTRAL Subsets](#)[Downloads](#)[Related Resources](#)[References](#)[Help](#)[About](#)

Search SCOPe ([example](#)):

Lineage for Fold a.1: Globin-like

1. Root: [SCOPe 2.04](#)
2.  Class [a: All alpha proteins](#) [46456] (285 folds)
3.  Fold a.1: Globin-like [46457] (2 superfamilies)
core: 6 helices; folded leaf, partly opened

Superfamilies:

1.  [a.1.1: Globin-like](#) [46458] (5 families) $\mathcal{S}$
2.  [a.1.2: alpha-helical ferredoxin](#) [46548] (3 families) $\mathcal{S}$
contains two Fe4-S4 clusters

More info for Fold a.1: Globin-like

Timeline for Fold a.1: Globin-like:

- Fold a.1: Globin-like [first appeared \(with stable ids\) in SCOP 1.55](#)
- Fold a.1: Globin-like [appears in SCOPe 2.03](#)

SCOPe

Copyright © 1994-2014 The SCOP and SCOPe authors
scope@compbio.berkeley.edu

SCOPe 2.0

← → ↻ scop.berkeley.edu/sunid=46458

SCOPe: Structural Classification of Proteins — extended. Release 2.04 (July 2014, new entries added 2014-09-1)

[Browse](#) [Stats & History](#) [ASTRAL Subsets](#) [Downloads](#) [Related Resources](#) [References](#) [Help](#) [About](#)

Search SCOPe ([example](#)):

Lineage for Superfamily a.1.1: Globin-like

1. Root: [SCOPe 2.04](#)
2.  Class [a: All alpha proteins](#) [46456] (285 folds)
3.  Fold [a.1: Globin-like](#) [46457] (2 superfamilies)
core: 6 helices; folded leaf, partly opened
4.  Superfamily [a.1.1: Globin-like](#) [46458] (5 families) 

Families:

1.  [a.1.1.1: Truncated hemoglobin](#) [46459] (2 protein domains)
lack the first helix (A)
2.  [a.1.1.2: Globins](#) [46463] (27 protein domains)
Heme-binding protein
3.  [a.1.1.3: Phycocyanin-like phycobilisome proteins](#) [46532] (7 protein domains)
oligomers of two different types of globin-like subunits containing two extra helices at the N-terminus
binds a bilin chromophore
automatically mapped to Pfam [PF00502](#)
4.  [a.1.1.4: Nerve tissue mini-hemoglobin \(neural globin\)](#) [74660] (2 protein domains)
lack the first helix but otherwise is more similar to conventional globins than the truncated ones
automatically mapped to Pfam [PF00042](#)
5.  [a.1.1.0: automated matches](#) [191420] (1 protein)
not a true family

More info for Superfamily a.1.1: Globin-like

SCOPe 2.0

← → X scop.berkeley.edu/sunid=46463

SCOPe: Structural Classification of Proteins — extended. Release 2.04 (July 2014, new entries added 2014-09-11)

[Browse](#)

[Stats & History](#)

[ASTRAL Subsets](#)

[Downloads](#)

[Related Resources](#)

[References](#)

[Help](#)

[About](#)

Search SCOPe ([example](#)):

Lineage for Family a.1.1.2: Globins

1. Root: [SCOPe 2.04](#)
2.  Class [a: All alpha proteins](#) [46456] (285 folds)
3.  Fold [a.1: Globin-like](#) [46457] (2 superfamilies)
core: 6 helices; folded leaf, partly opened
4.  Superfamily [a.1.1: Globin-like](#) [46458] (5 families) *S*
5.  Family [a.1.1.2: Globins](#) [46463] (27 protein domains)
Heme-binding protein

Protein Domains:

1.  [Ascaris hemoglobin, domain 1](#) [46520] (1 species)
 Species [Pig roundworm \(Ascaris suum\) \[TaxId:6253\]](#) [46521] (1 PDB entry)
2.  [Bacterial dimeric hemoglobin](#) [46526] (1 species)
 Species [Vitreoscilla stercoraria \[TaxId:61\]](#) [46527] (7 PDB entries)
3.  [Chimeric hemoglobin beta-alpha](#) [46516] (1 species)
 Species [Synthetic, based on Homo sapiens sequence](#) [46517] (1 PDB entry)
4.  [Cytoglobin](#) [109626] (1 species)
 Species [Human \(Homo sapiens\) \[TaxId:9606\]](#) [109627] (8 PDB entries)
Uniprot Q8WWM9 18-171
5.  [Dehaloperoxidase](#) [46530] (1 species)
 Species [Amphitrite ornata \[TaxId:129555\]](#) [46531] (31 PDB entries)
6.  [Erythrocrutorin](#) [46479] (1 species)

SCOPe 2.0

← → ↻ scop.berkeley.edu/sunid=46520&ver=1.55

SCOP: Structural Classification of Proteins and ASTRAL. Release 1.55 (July 2001)

[Browse](#) [Stats & History](#) [ASTRAL Subsets](#) [Downloads](#) [Related Resources](#) [References](#) [Help](#) [About](#)

Search SCOP ([example](#)):

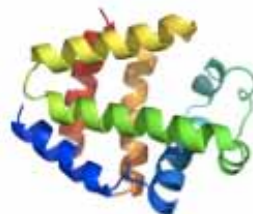
Lineage for Protein: Ascaris hemoglobin, domain 1

1. Root: [SCOP 1.55](#)
2.  Class [a: All alpha proteins](#) [46456] (138 folds)
3.  Fold [a.1: Globin-like](#) [46457] (2 superfamilies)
4.  Superfamily [a.1.1: Globin-like](#) [46458] (3 families) $\mathcal{S}$
5.  Family [a.1.1.2: Globins](#) [46463] (16 protein domains)
6.  Protein Ascaris hemoglobin, domain 1 [46520] (1 species)

Species:

 [Pig roundworm \(Ascaris suum\) \[TaxId:6253\]](#) [46521] (1 PDB entry)

Domain for [lash](#):

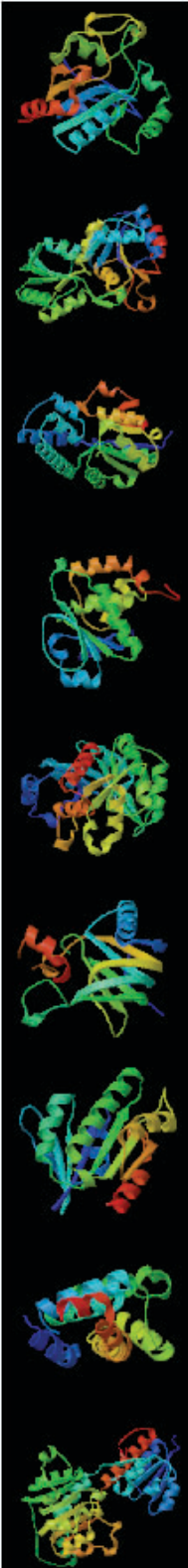


Domain [dlash_](#) : [lash_](#) - [15622]

More info for Protein Ascaris hemoglobin, domain 1 from a.1.1.2: Globins

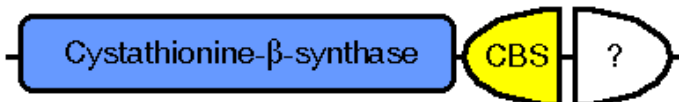
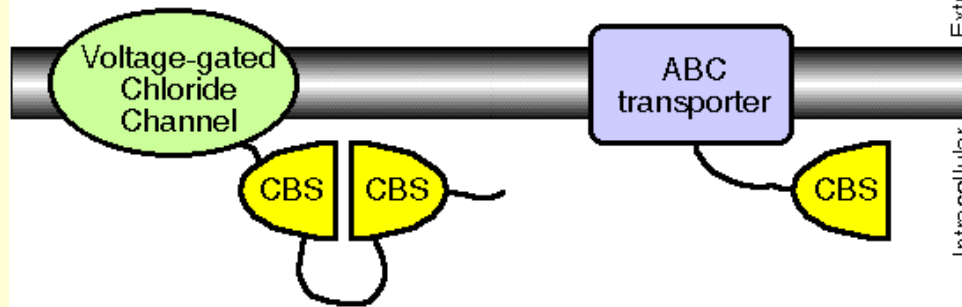
Timeline for Protein Ascaris hemoglobin, domain 1 from a.1.1.2: Globins :

- Protein Ascaris hemoglobin, domain 1 from a.1.1.2: Globins [appears in SCOP 1.57](#)
- Protein Ascaris hemoglobin, domain 1 from a.1.1.2: Globins [appears in the current release, SCOPe 2.04](#)



What is Pfam ?

Proteins containing CBS domains



Domains can be considered as building blocks of proteins.

Some domains can be found in many proteins with different functions, while others are only found in proteins with a certain function.

The presence of a particular domain can be indicative of the function of the protein.

Pfam is a domain database.

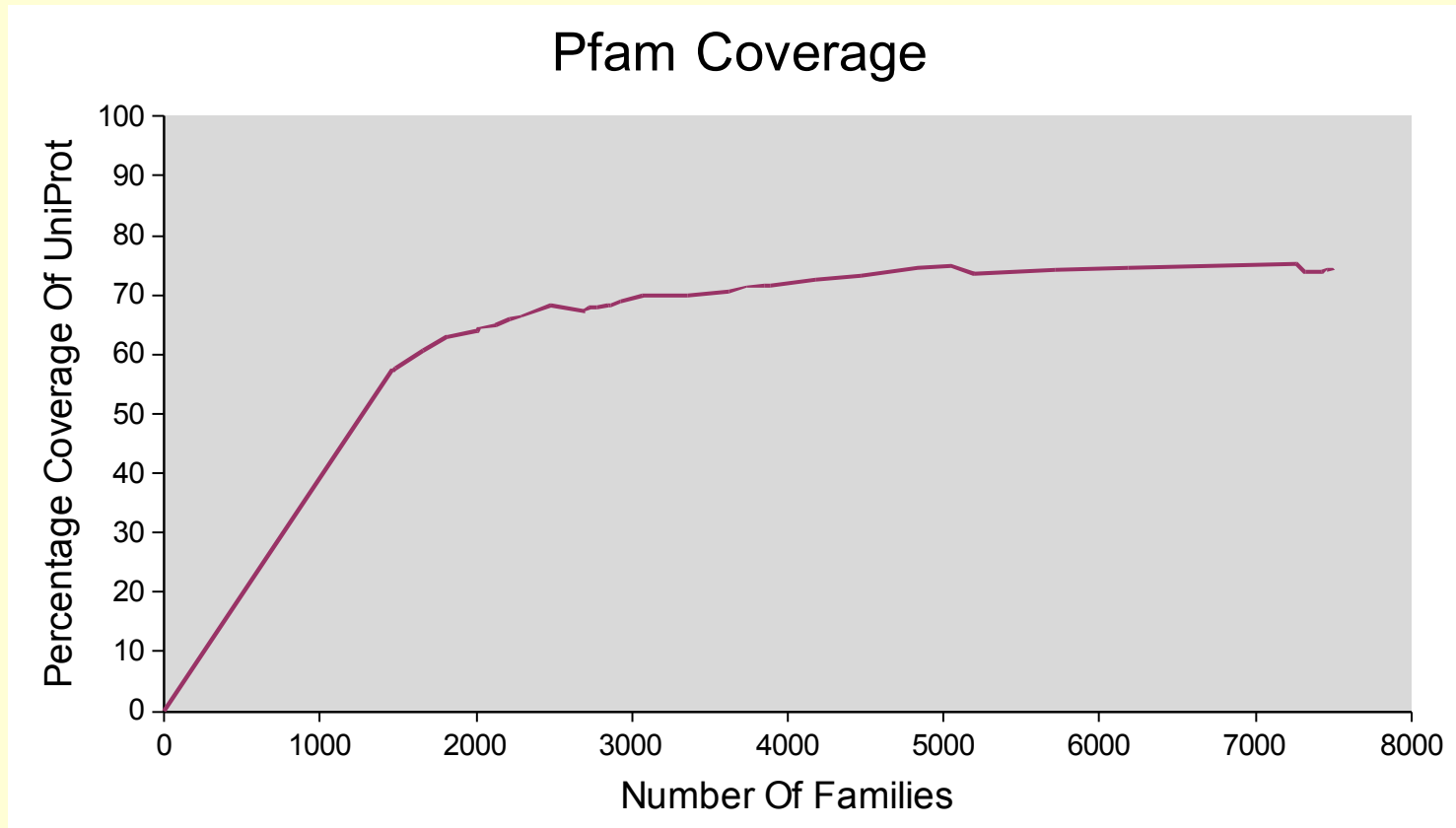
Comprised of two parts – Pfam-A and Pfam-B.

Pfam is use by many different groups in many different ways. Originally set up to aid the annotation the *C. elegans* genomes.



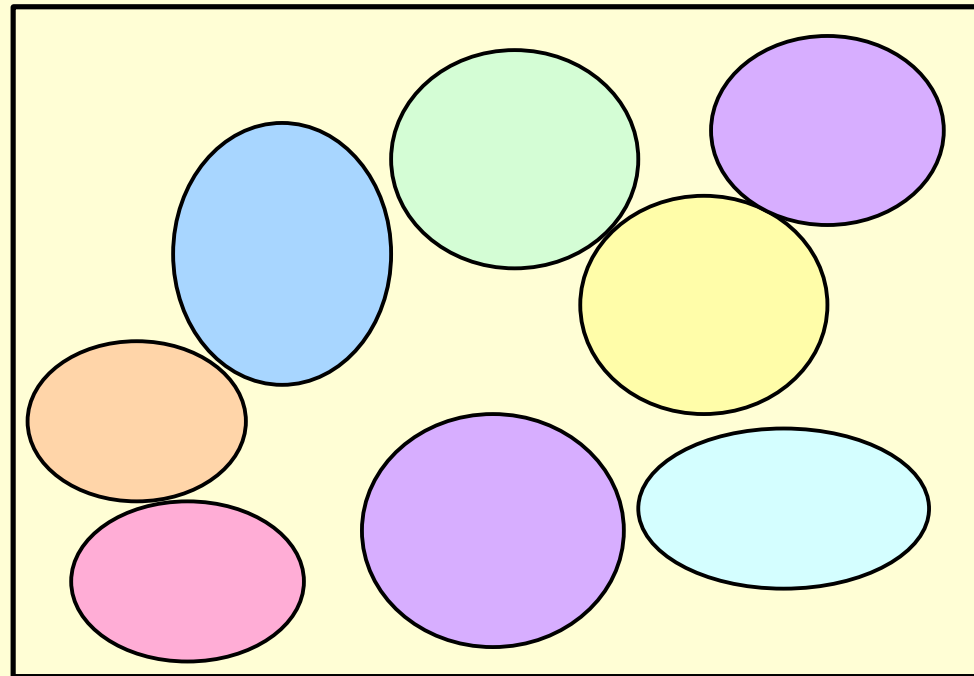
What is a Pfam-A Entry?

- A SEED alignment – contains a set or representative sequences
- HMM – built using the SEED alignment
- A full alignment – contains all (detectable) sequences in the family
- A description of the family, includes thresholds you to create the full alignment
- Rules – No false positives. A family is not allowed to overlap with any other family



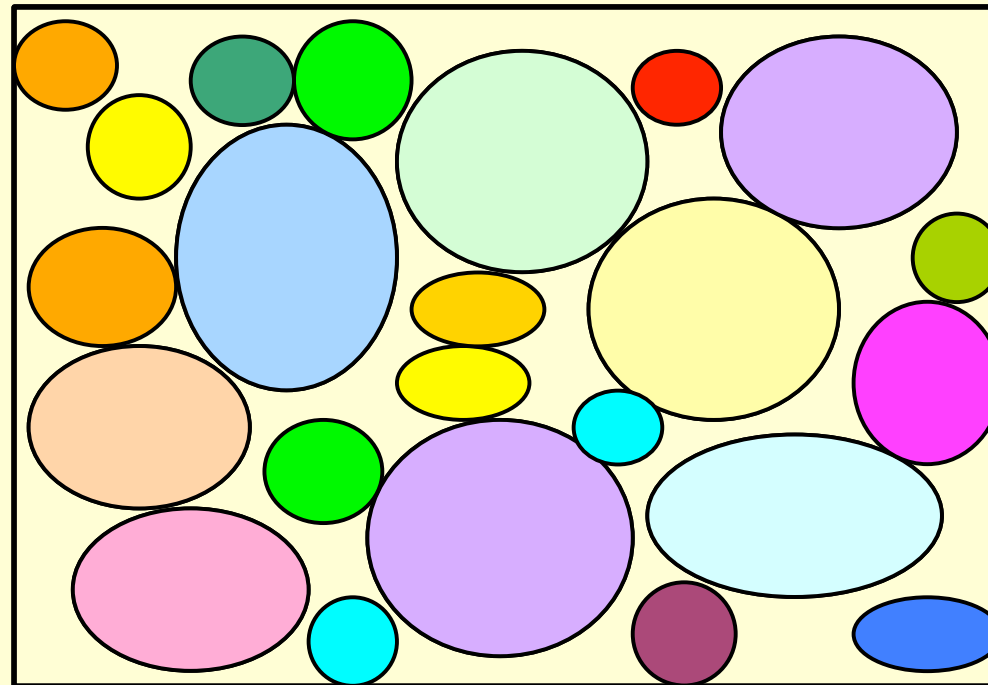
- First 2000 families covered ~ 65% of UniProt
- Currently, 7503 families cover 74% of UniProt

Pfam Sequence Coverage



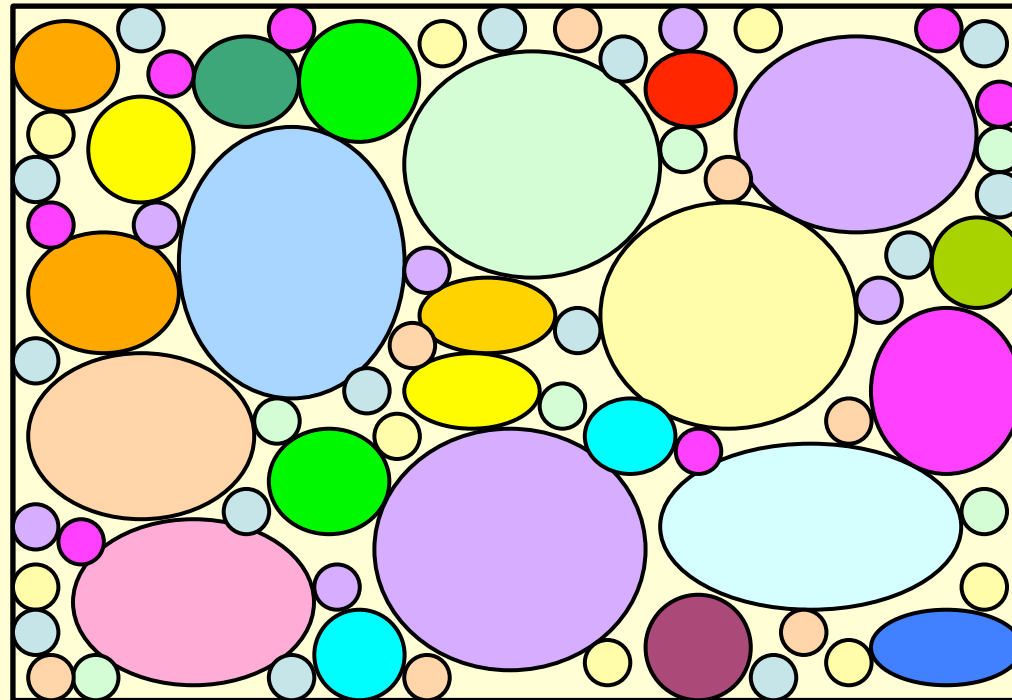
So why does the curve look logarithmic ?

Pfam Sequence Coverage



So why does the curve look logarithmic ?

Pfam Sequence Coverage



So why does the curve look logarithmic ?



Pfam-B

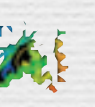
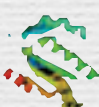
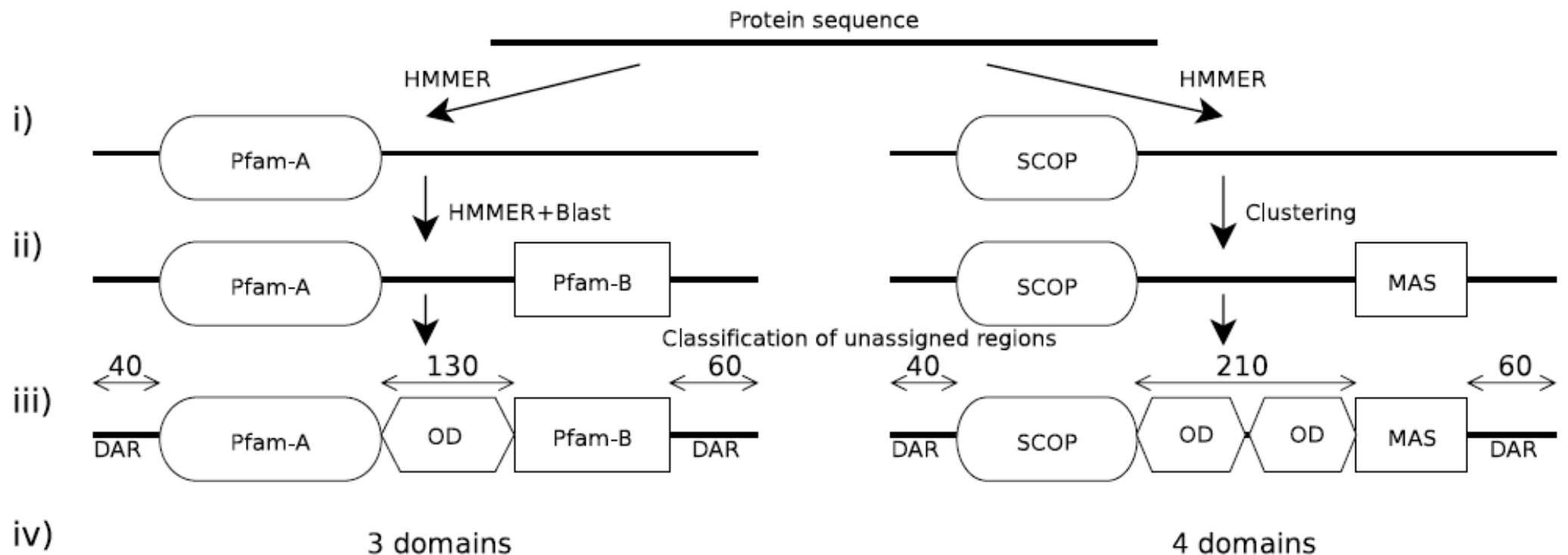
- Pfam-A covers about 74% of sequences
- To be comprehensive we have Pfam-B
- There are over 140,000 Pfam-B
- They cover 24% of UniProt (not covered by Pfam-A)
- Automatically generated clusters that are derived from Prodom



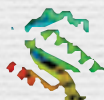
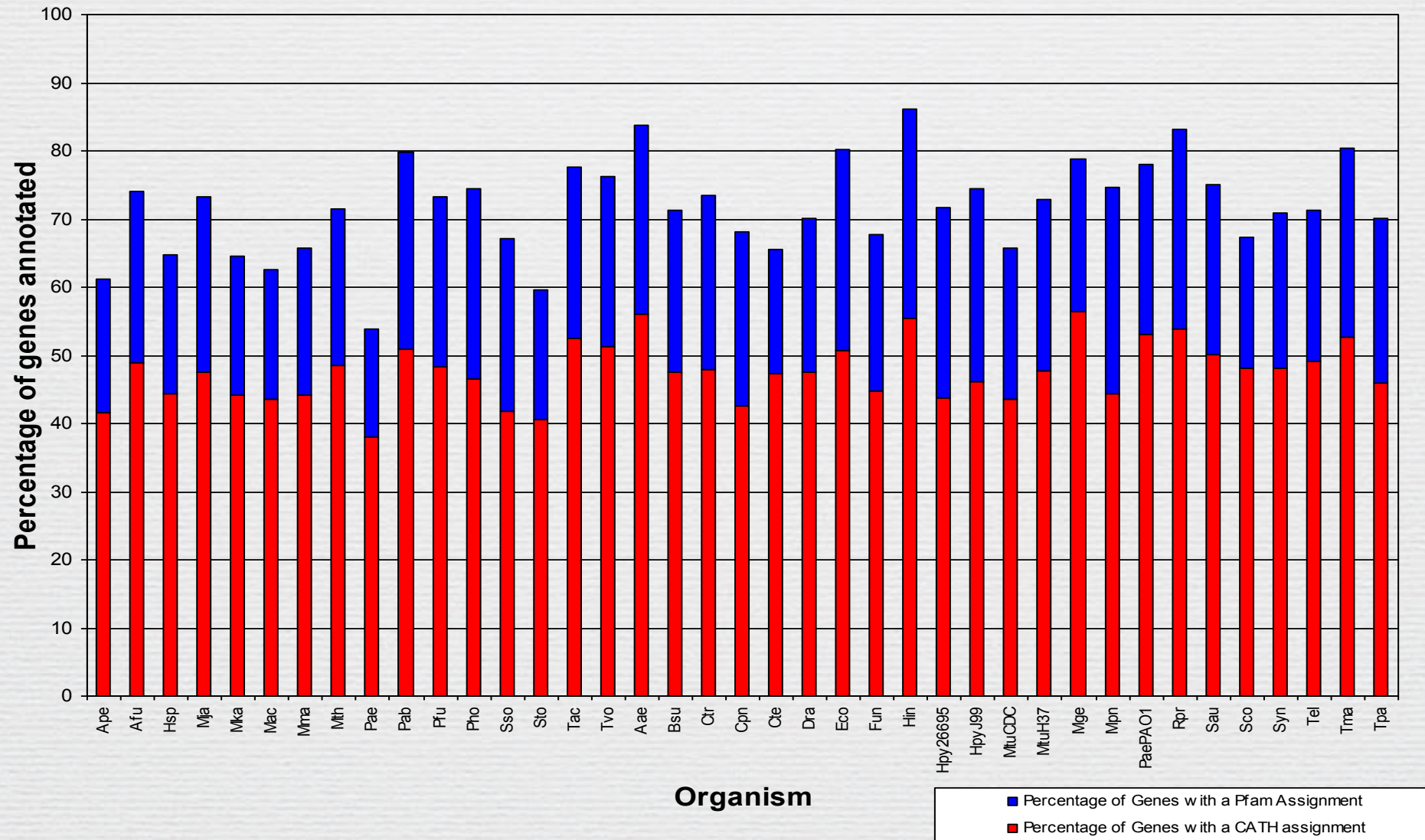
Pfam – Nuts and Bolts

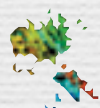
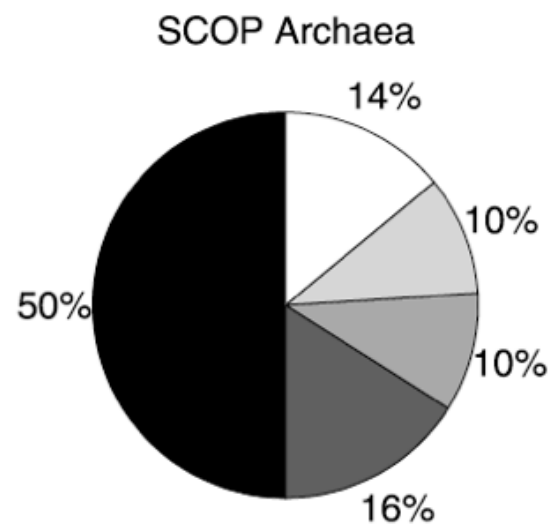
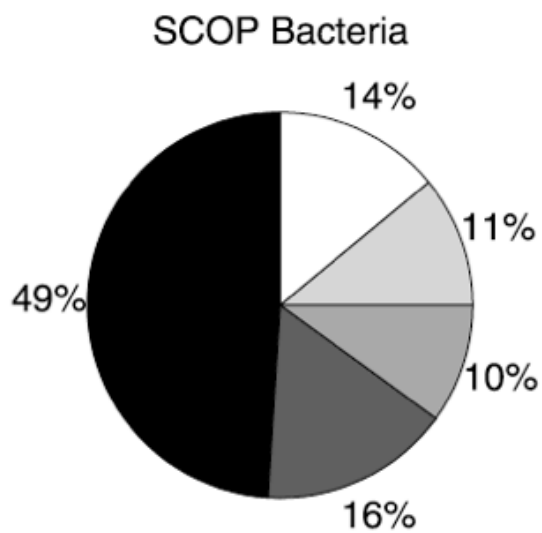
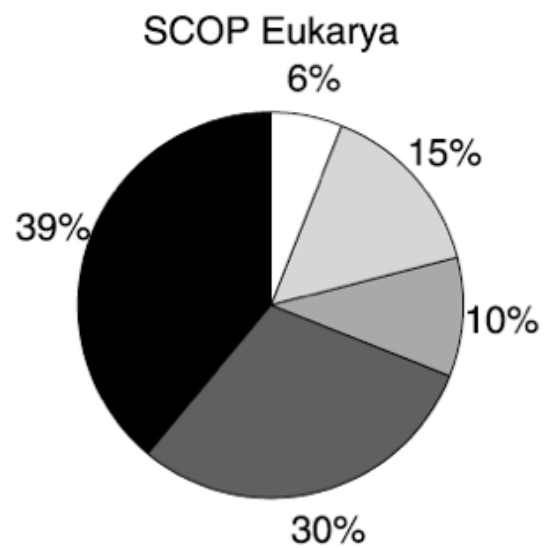
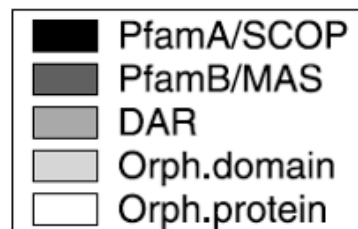
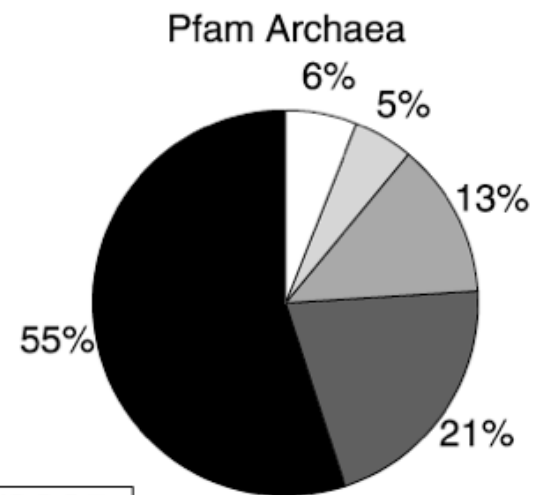
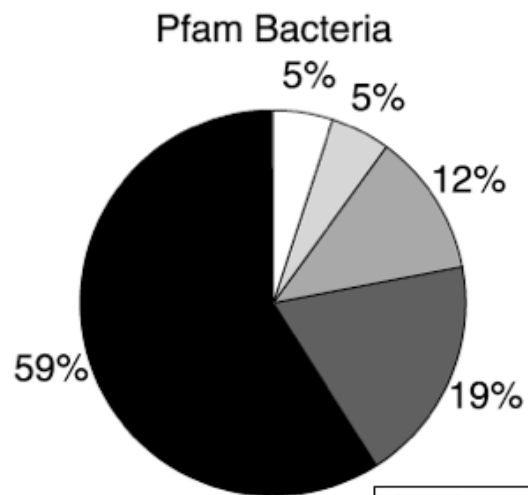
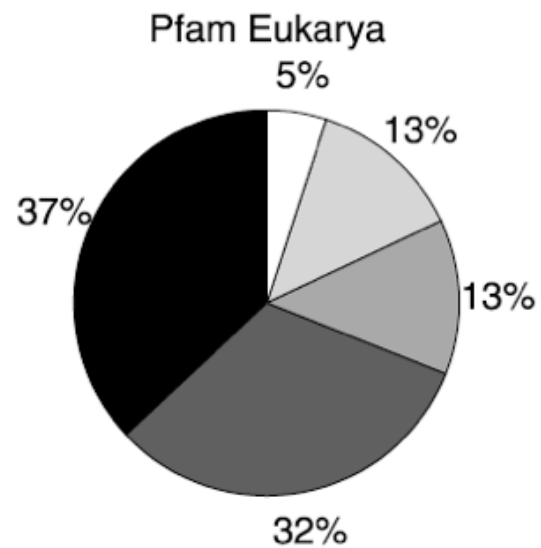
- **Collection of sequence alignments and profile hidden Markov models (HMMs)**
- **Over 7,500 families**
- **mySQL database**
- **Bi-Monthly Releases - flatfiles and relational tables**
- **Current Release – 15.0**
- **Mirrored around the World**

Method for Domain Assignment

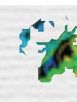
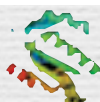
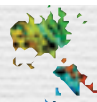
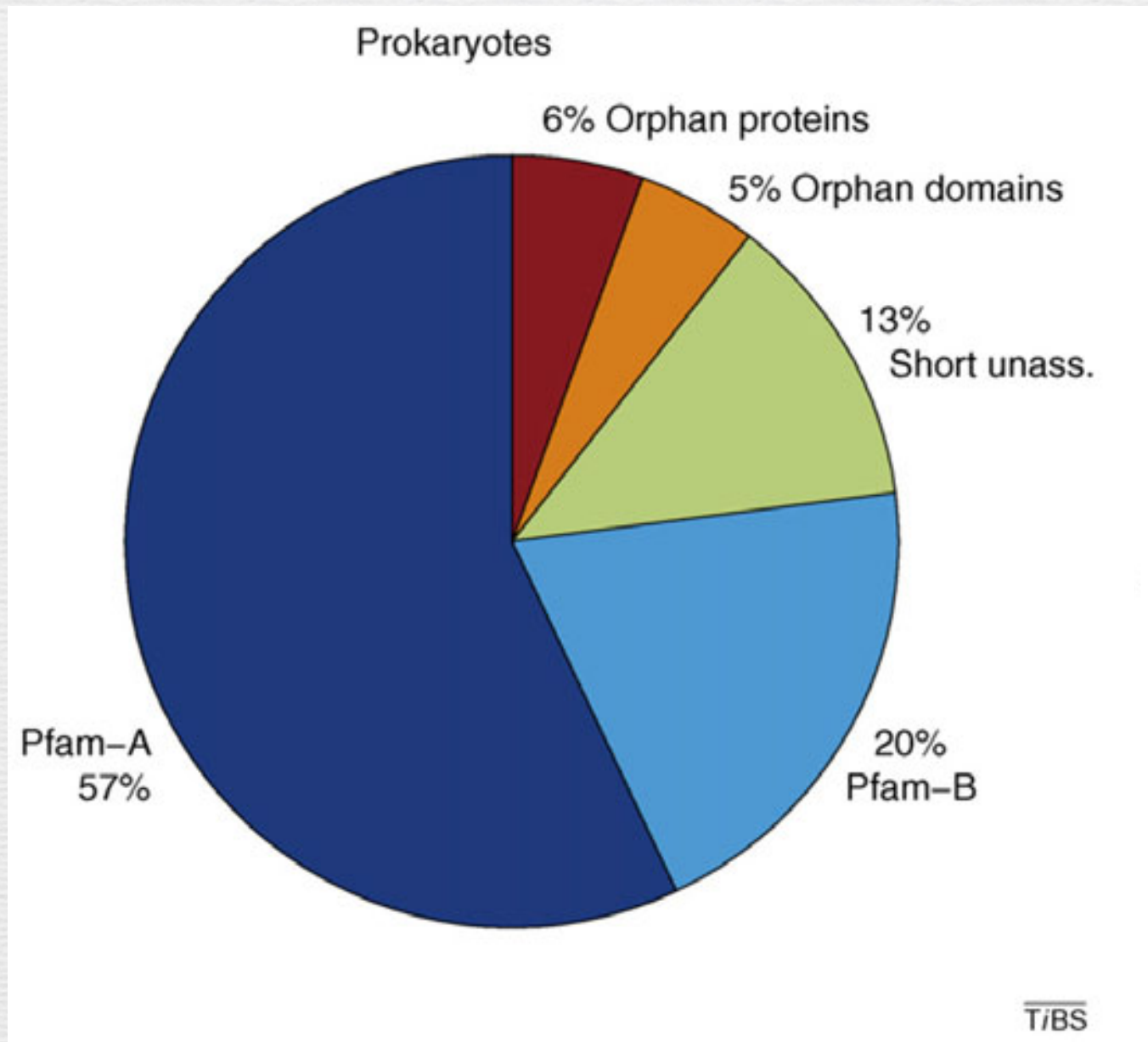


Coverage of CATH+Pfam



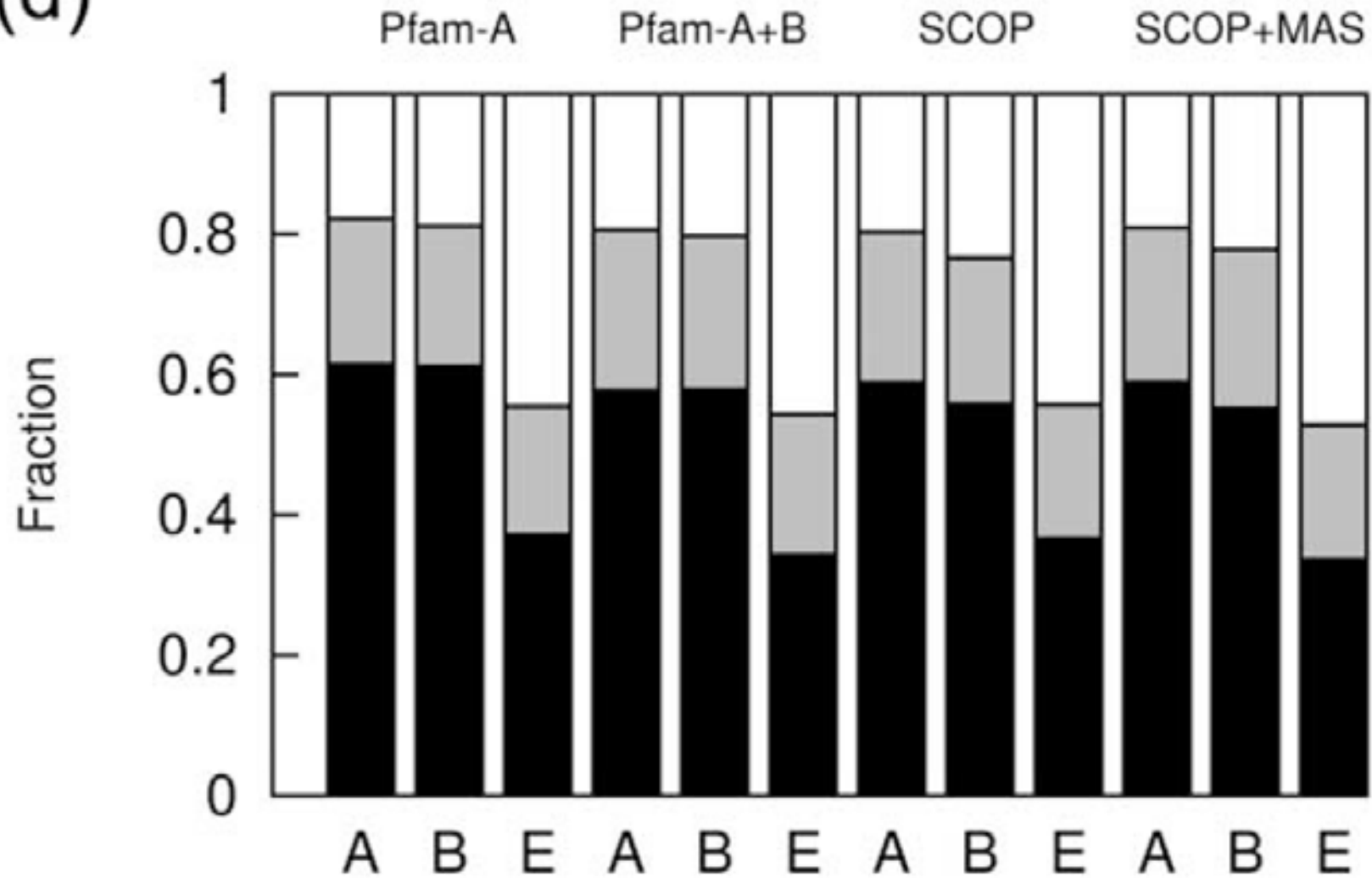


Domain Assignments

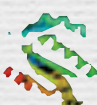
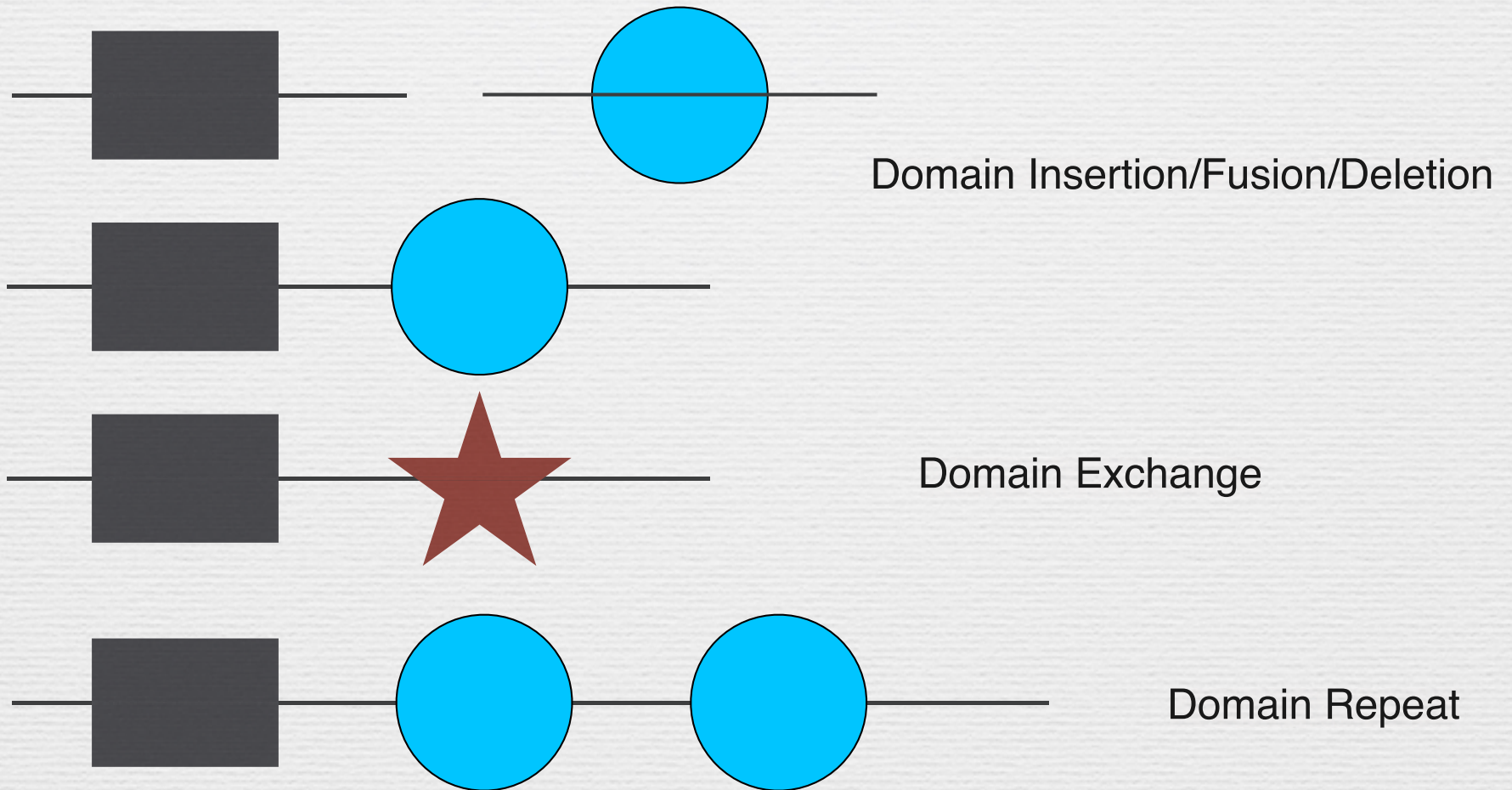


Fraction multidomain proteins

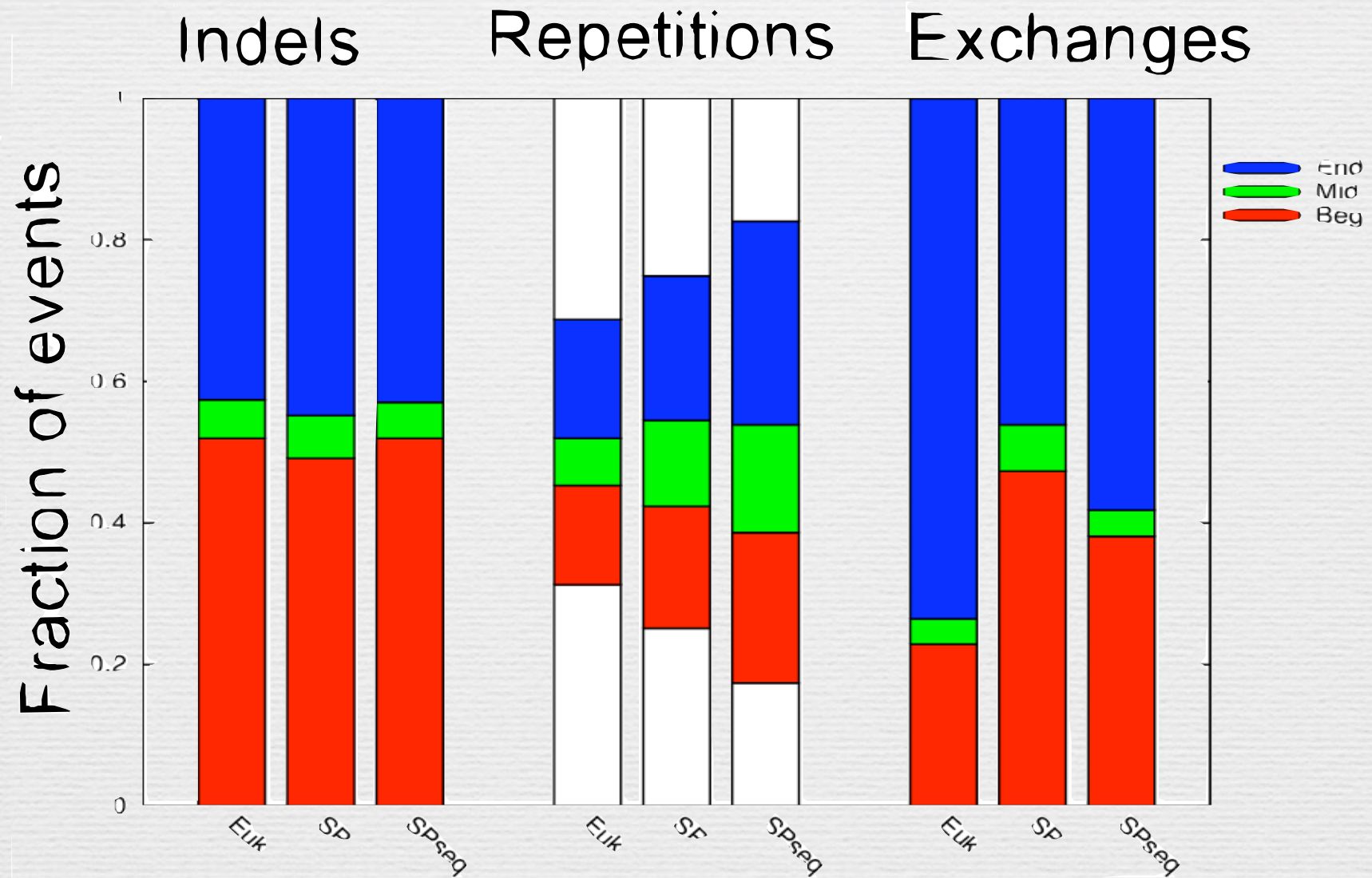
(d)



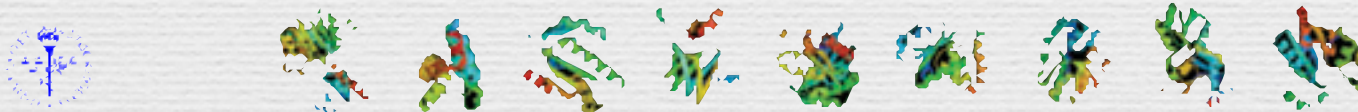
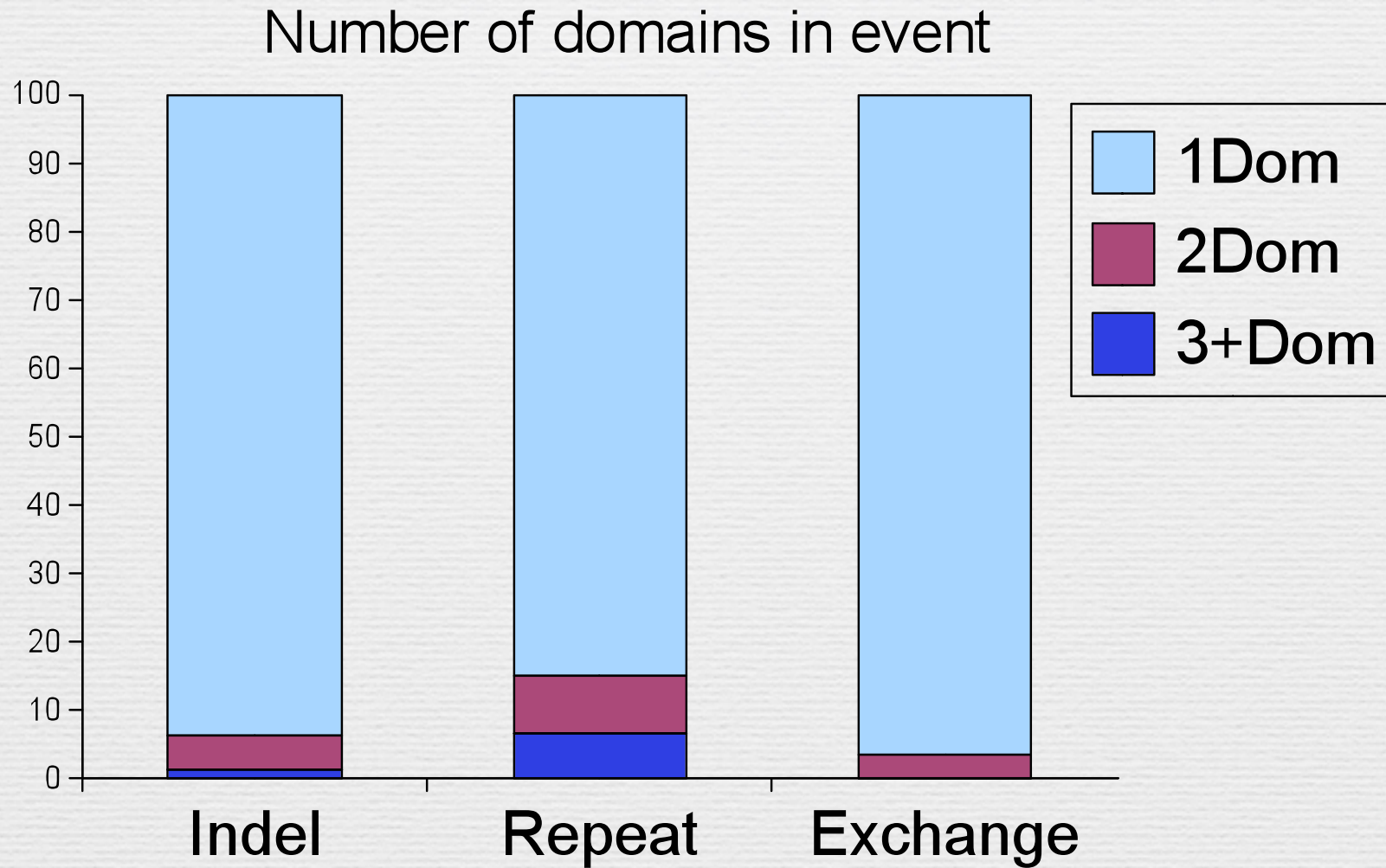
Domain evolution

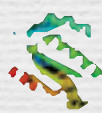
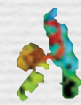
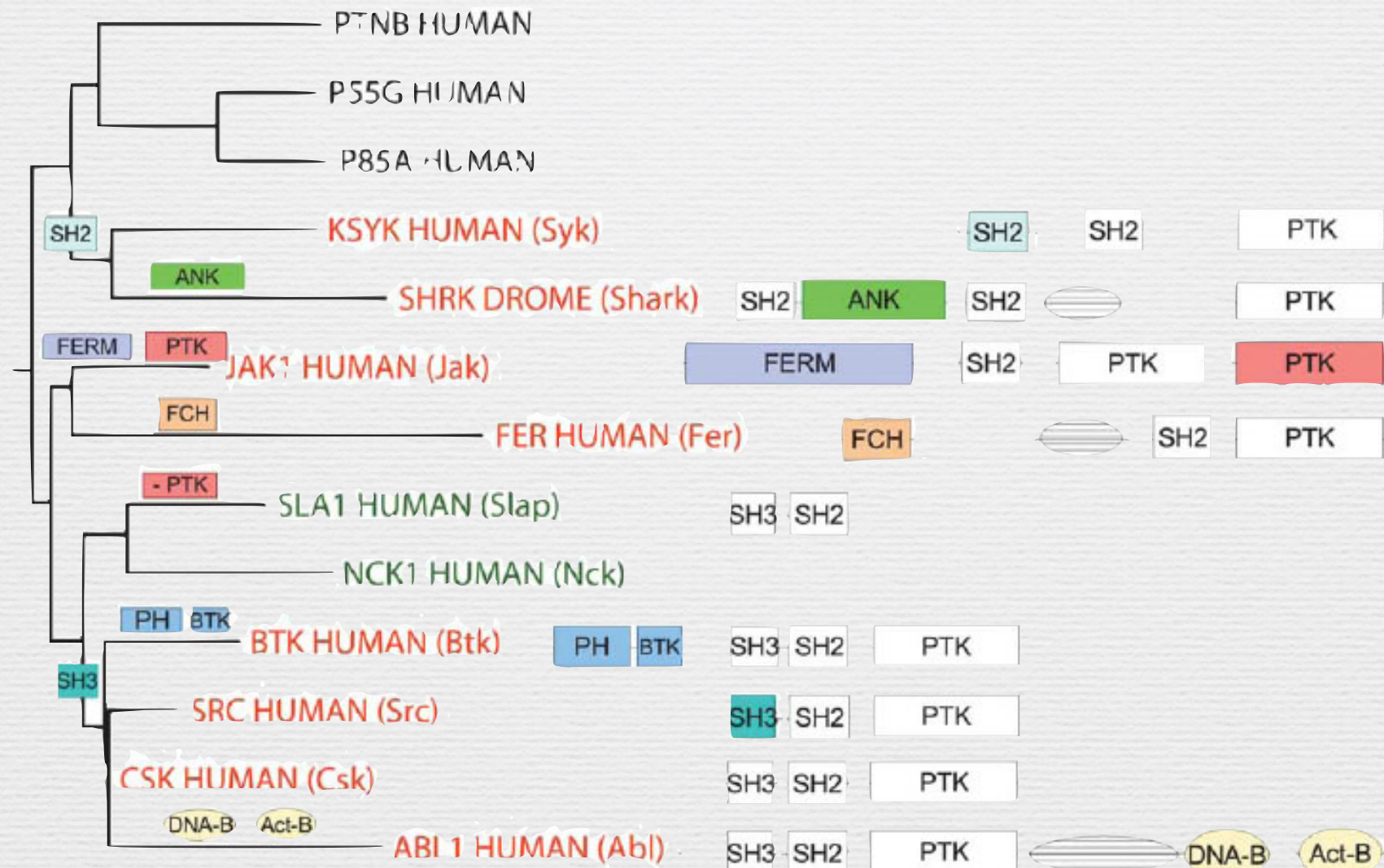


Insertions C- or N-terminal

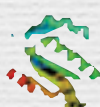
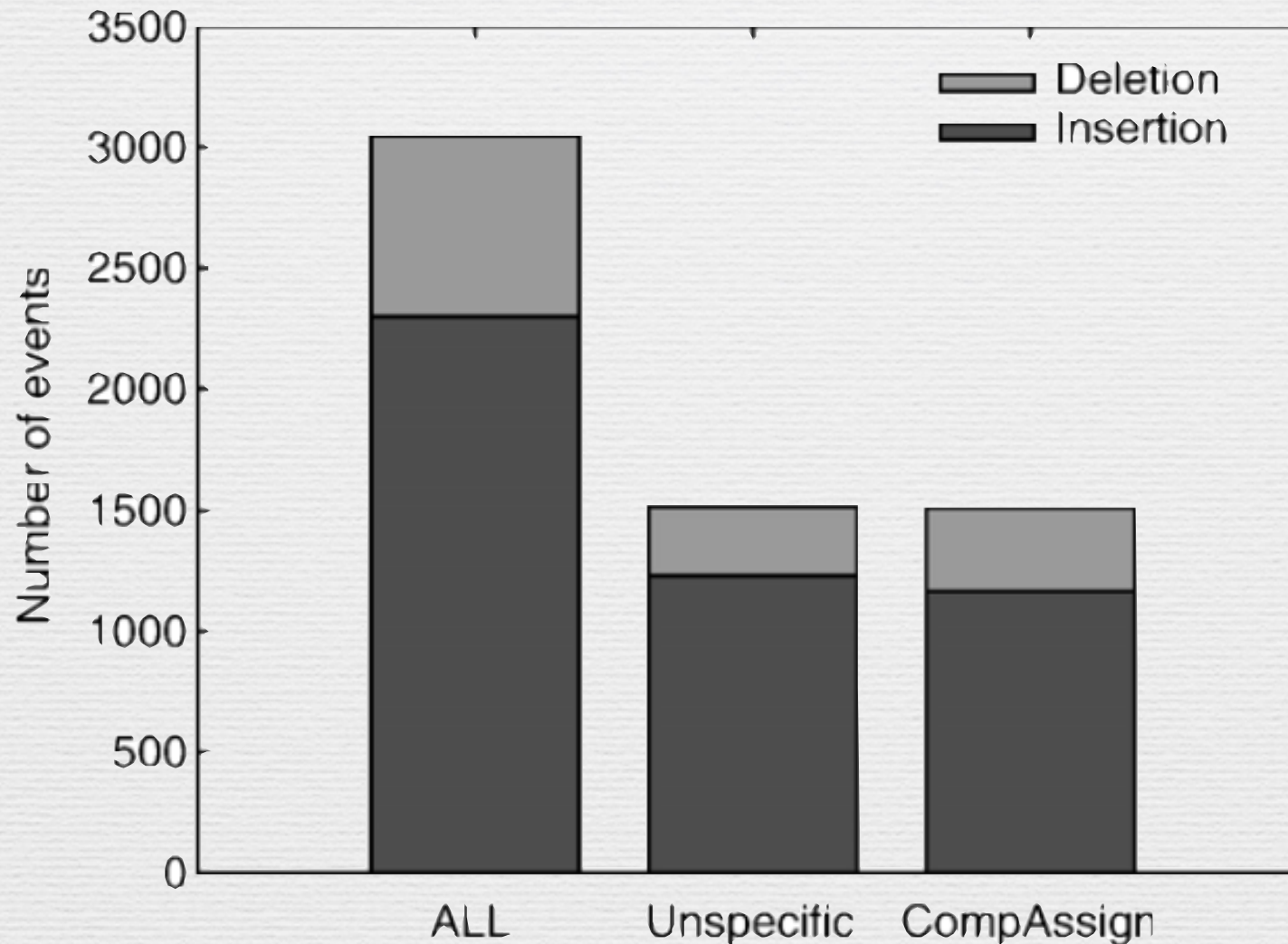


Single domain insertions

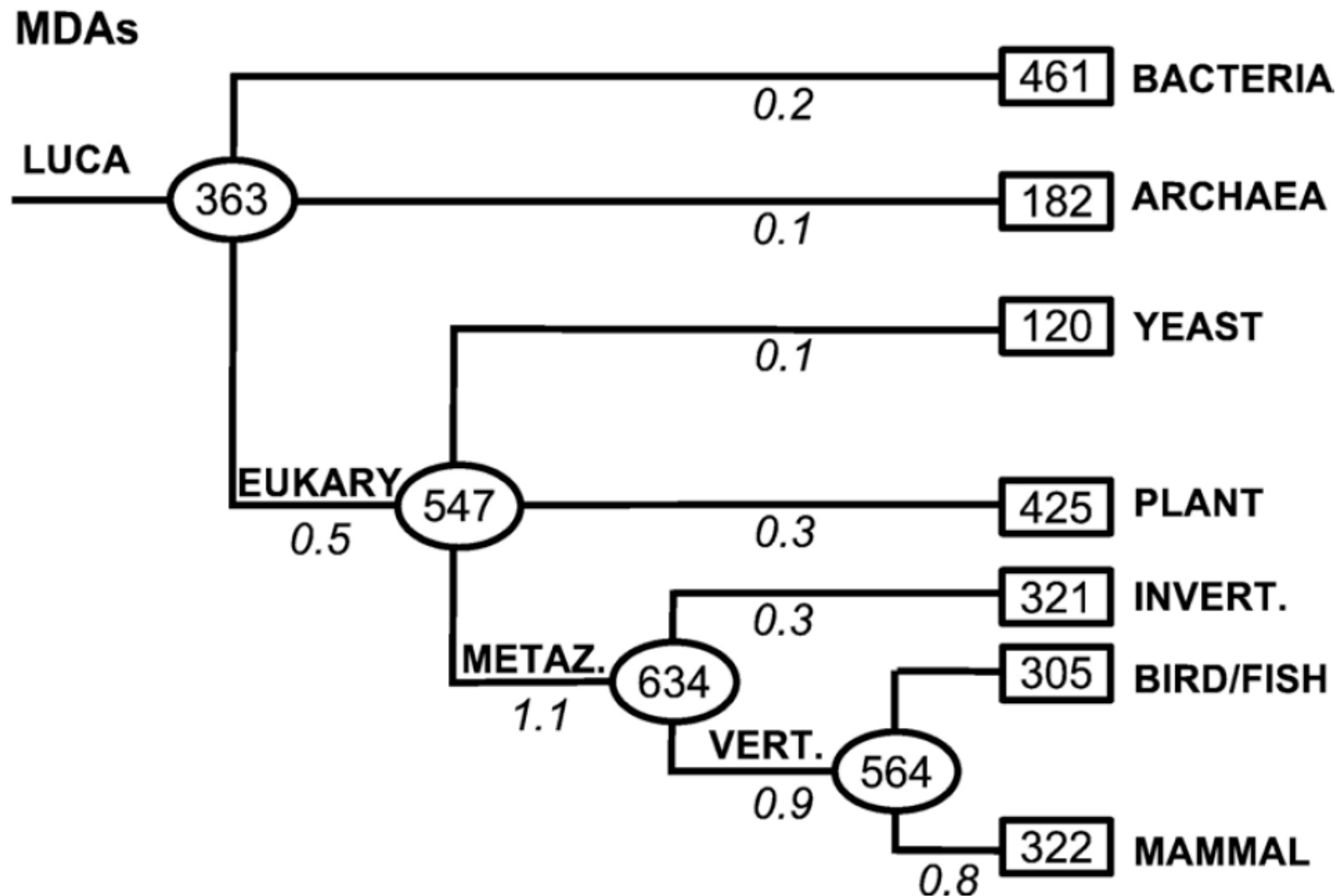




Insertion more frequent than deletions



Evolution of multidomain proteins



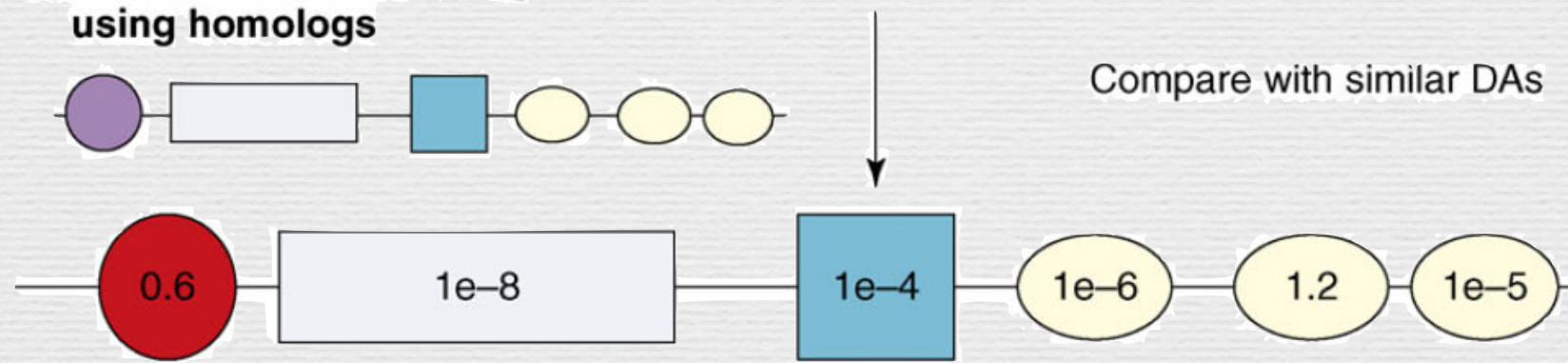
(a) Initial domain assignment



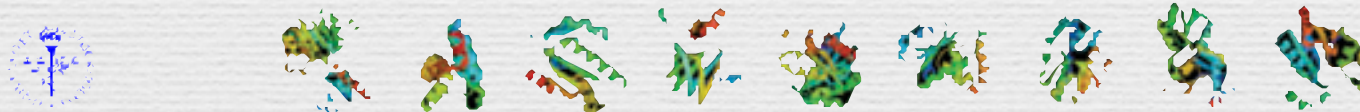
(b) Additional repeat assignment

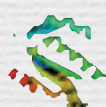


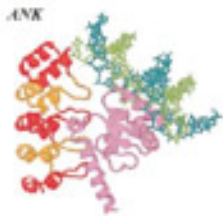
(c) Additional domain assignment using homologs



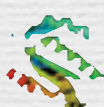
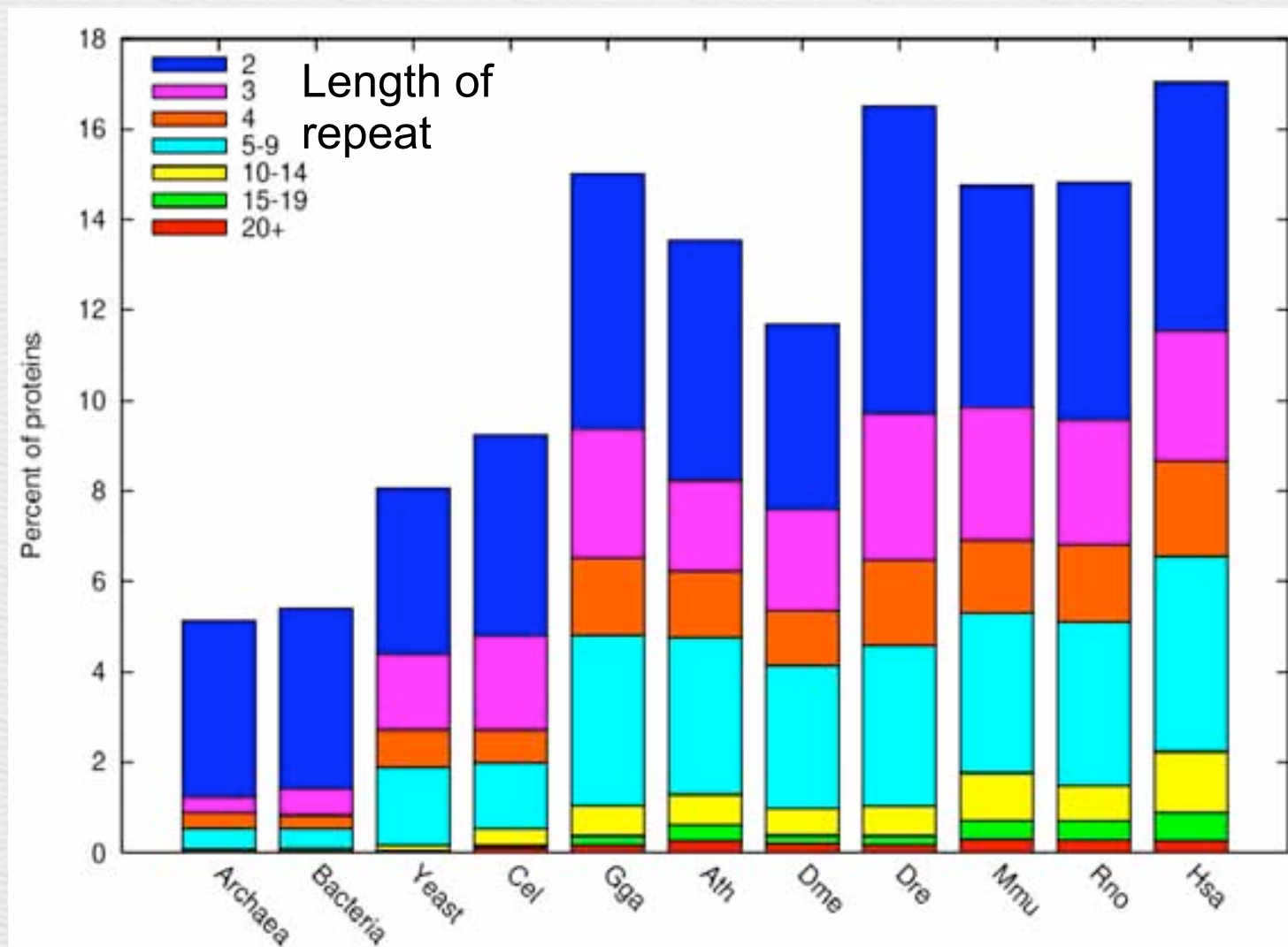
TiBS



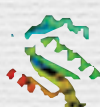
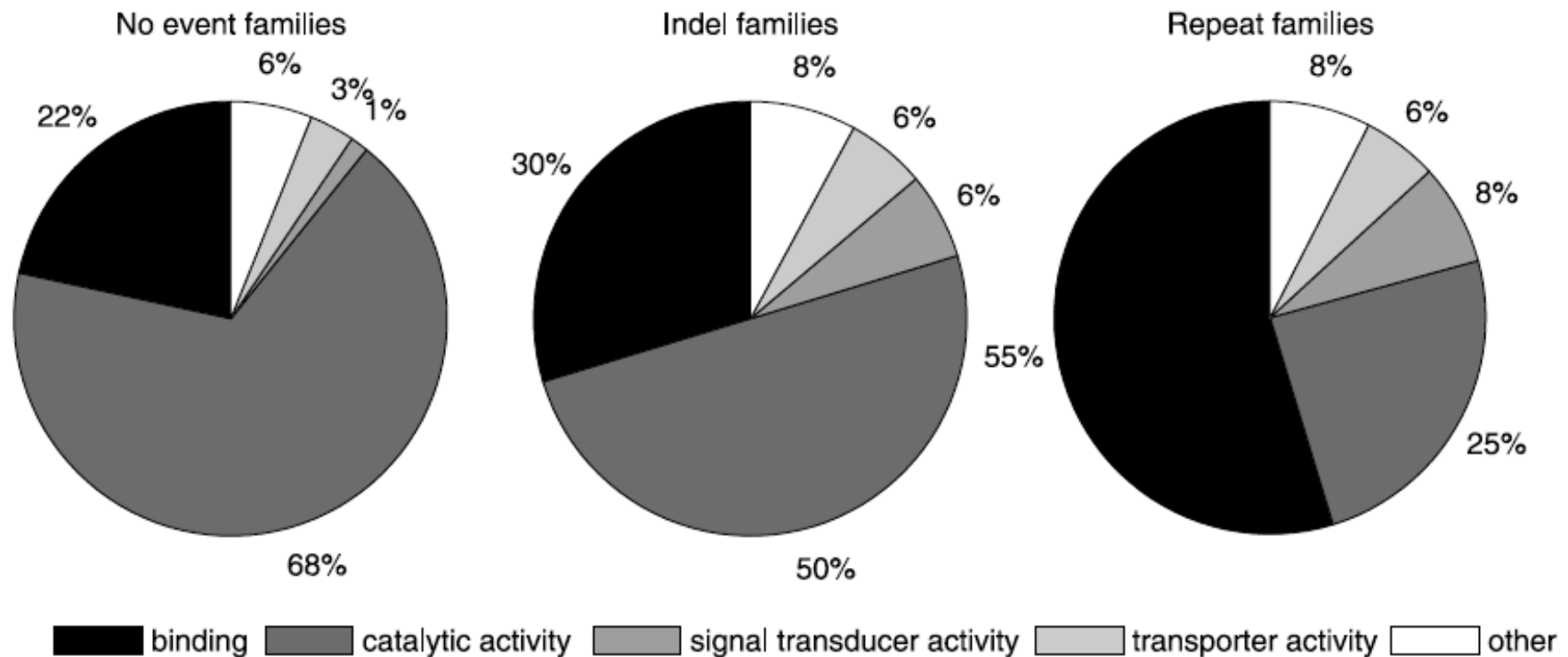


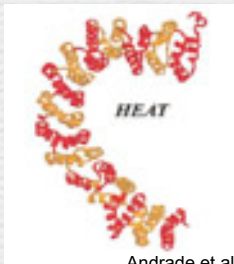


Many repeats in vertebrates



Repeat proteins are involved in binding

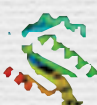
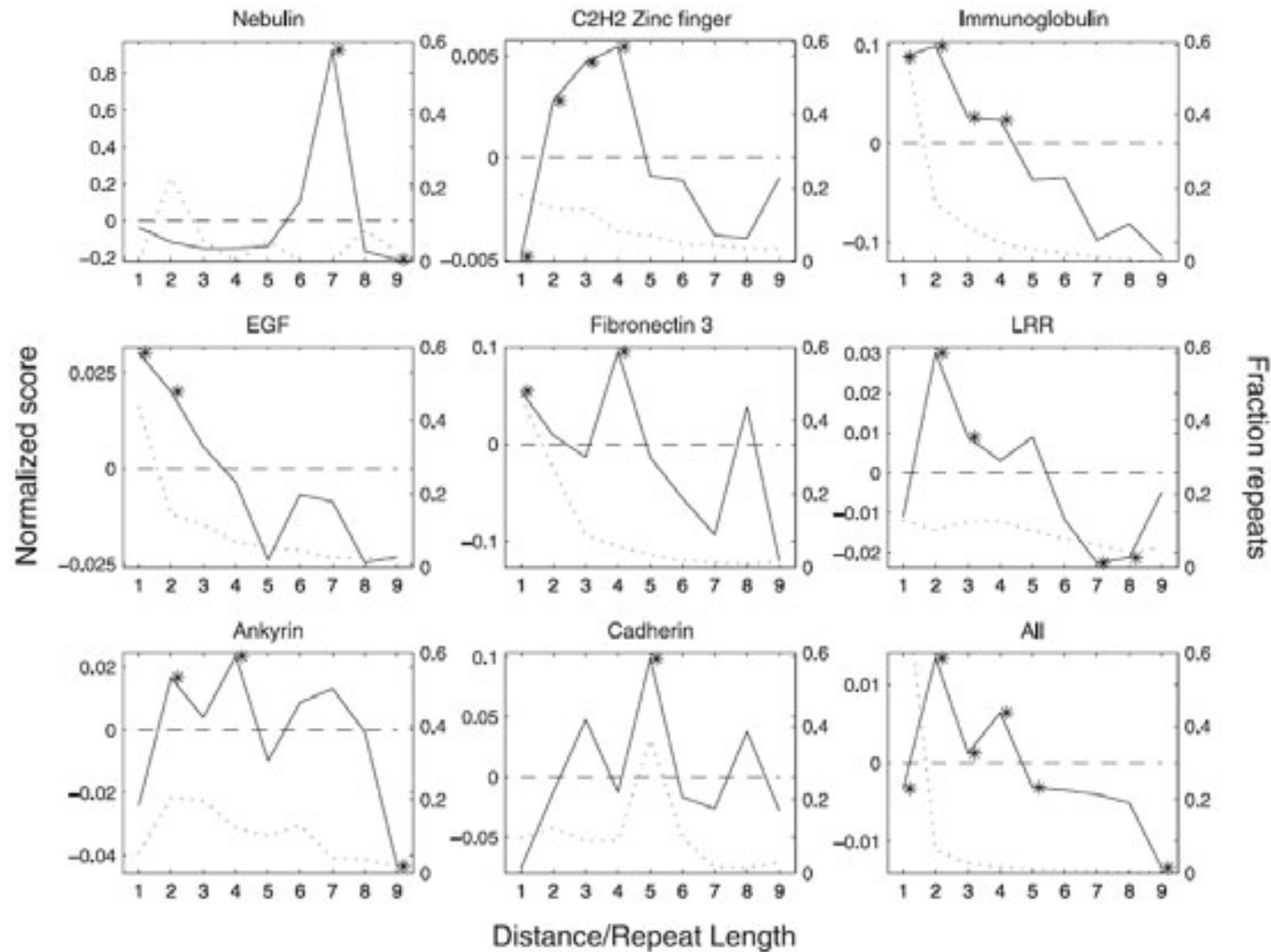
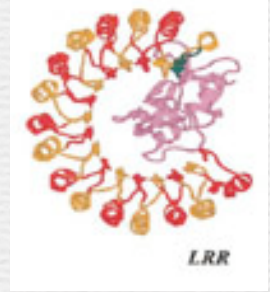




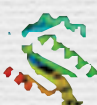
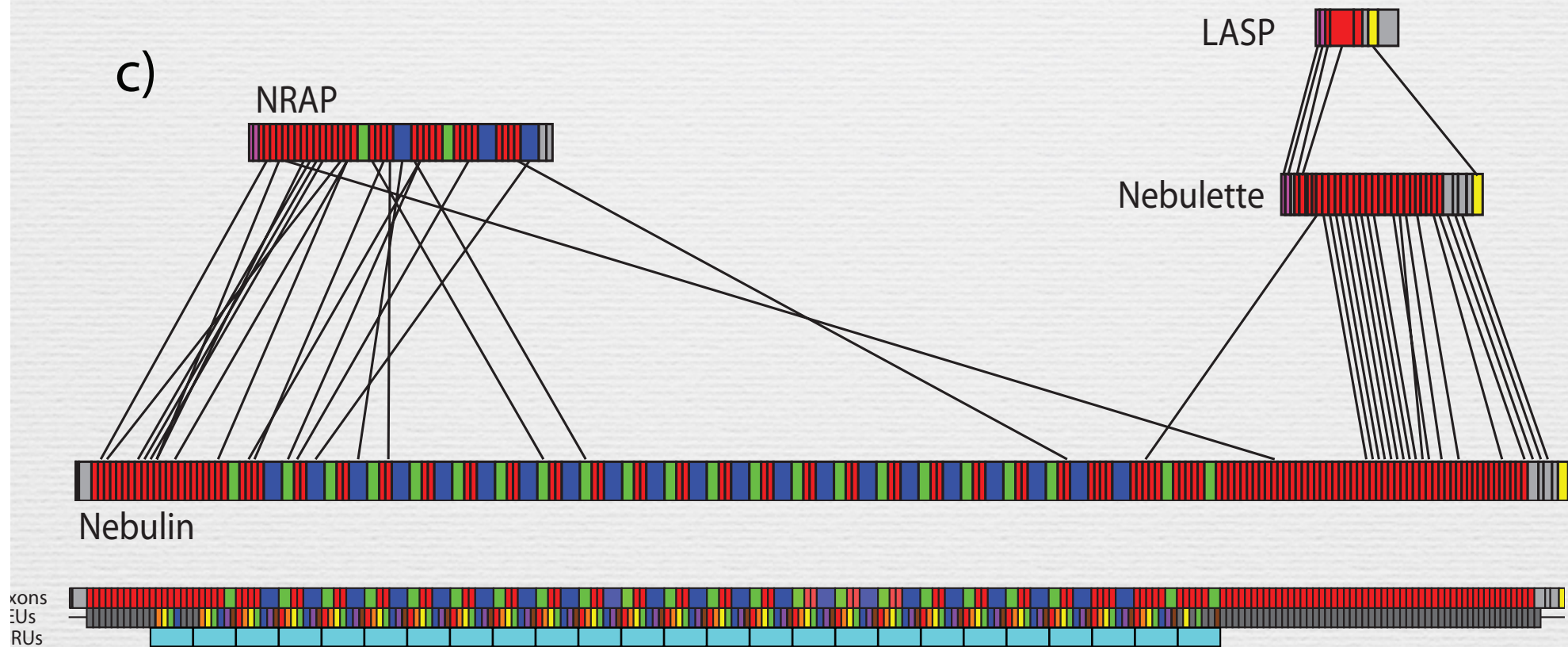
Andrade et al. 2001

Analysis of Repeats

(Björklund 2006 JMB)

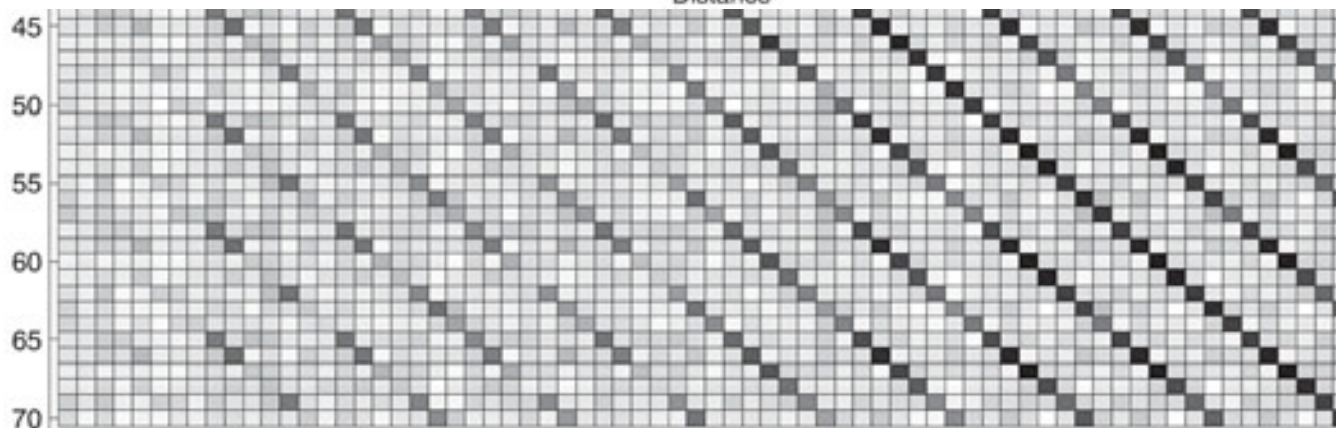
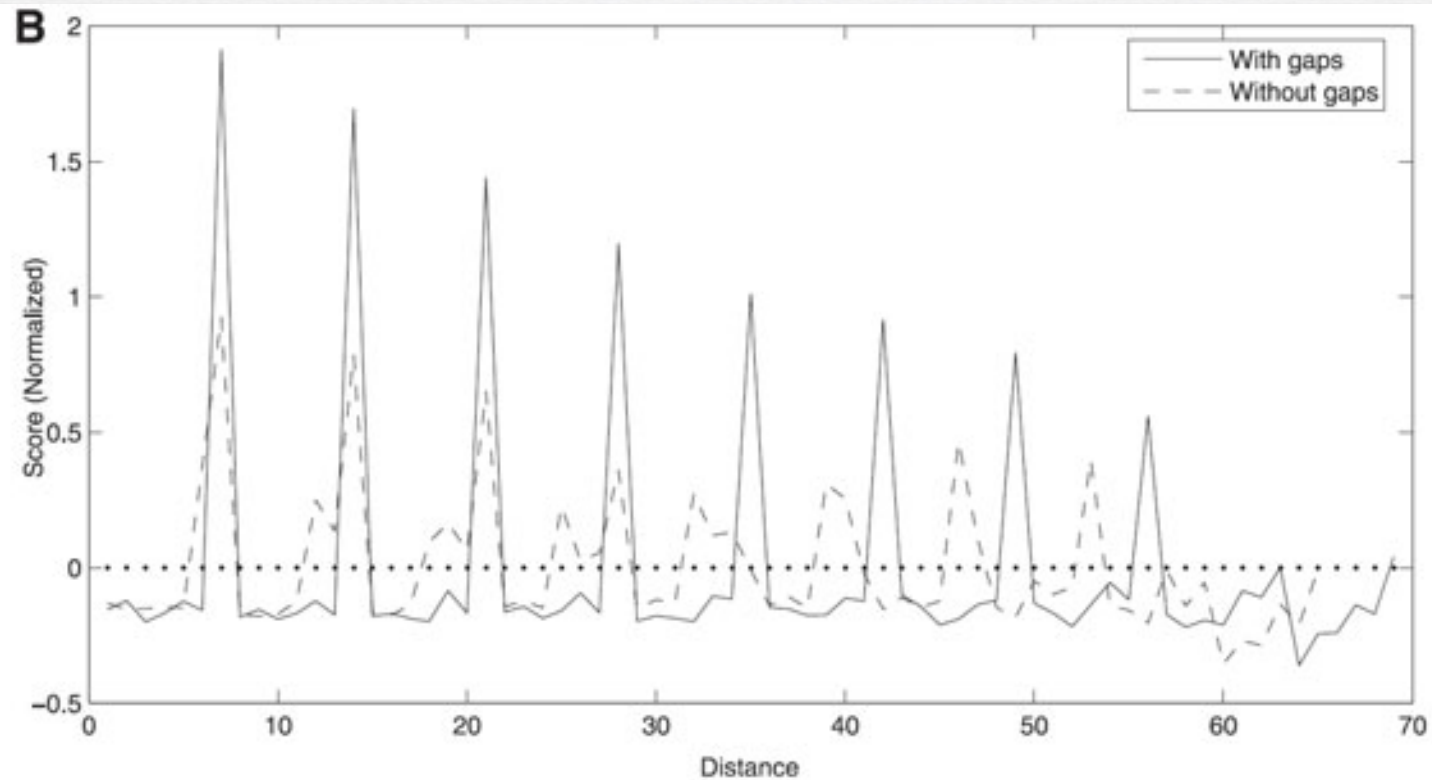


Overview of Nebulin containing proteins

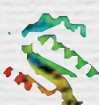
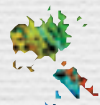
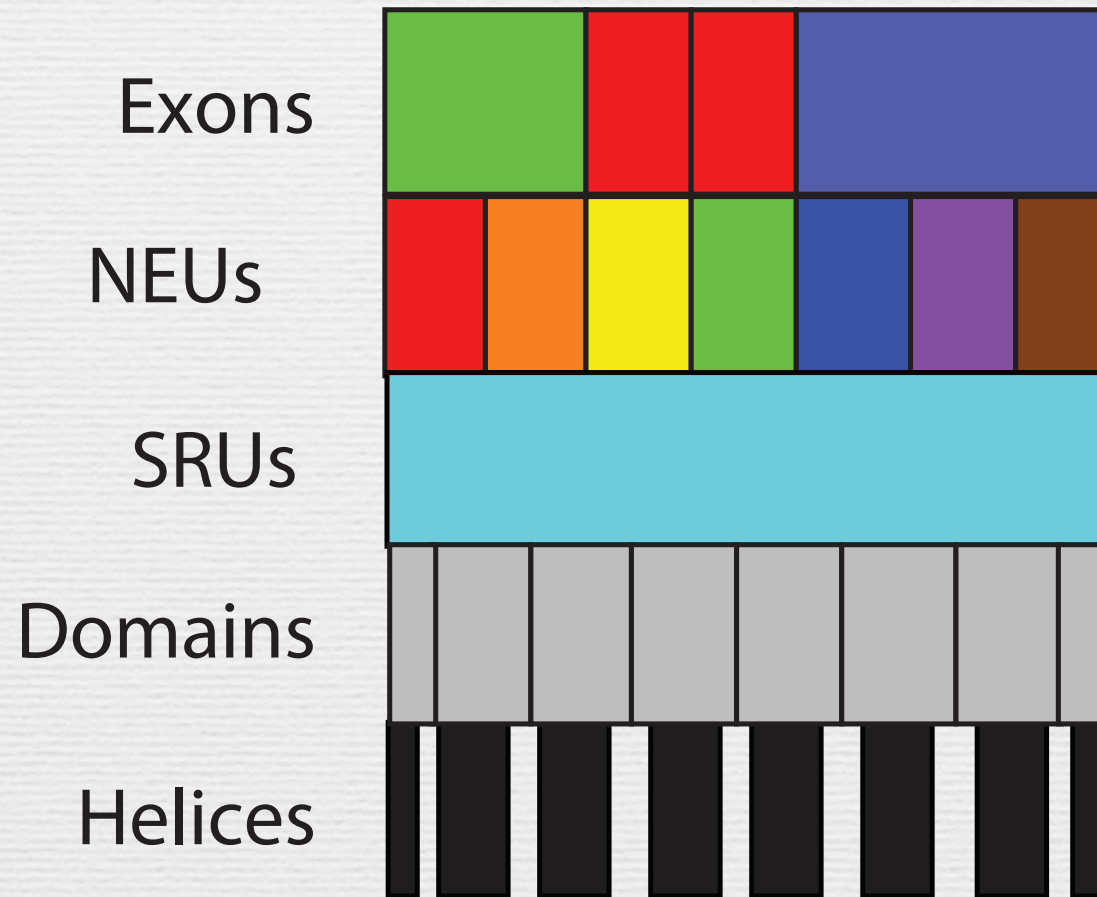


Nebulin

(Björklund 2010 JMB)



Domains and evolutionary units



Different exon patterns in Branchiostoma

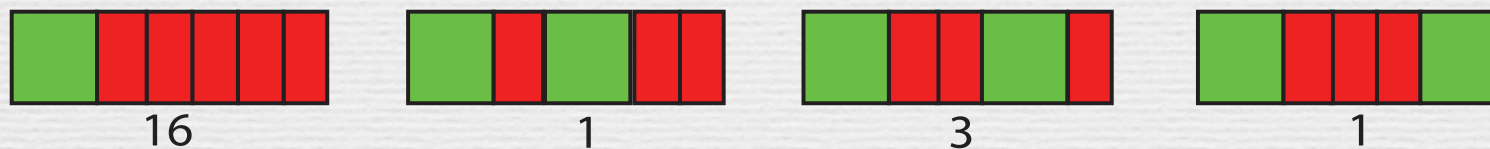
Human Neb SR (25 SR units)



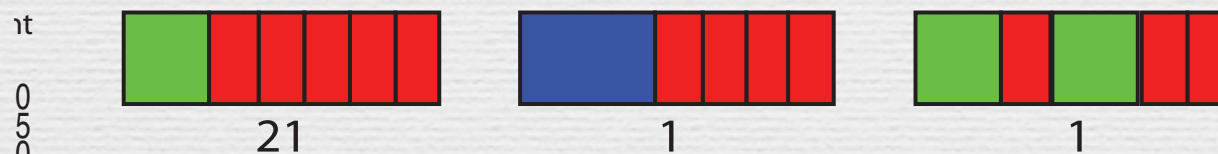
Human NRAP SR (5 SR units)



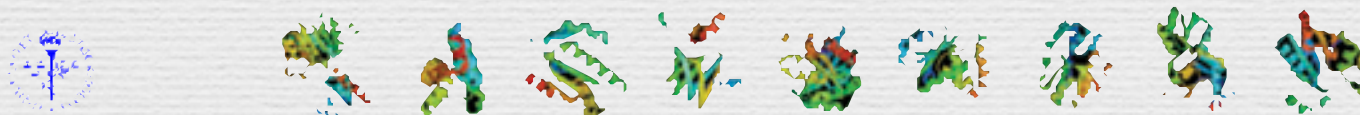
Branchiostoma SR1 (21 SR units)

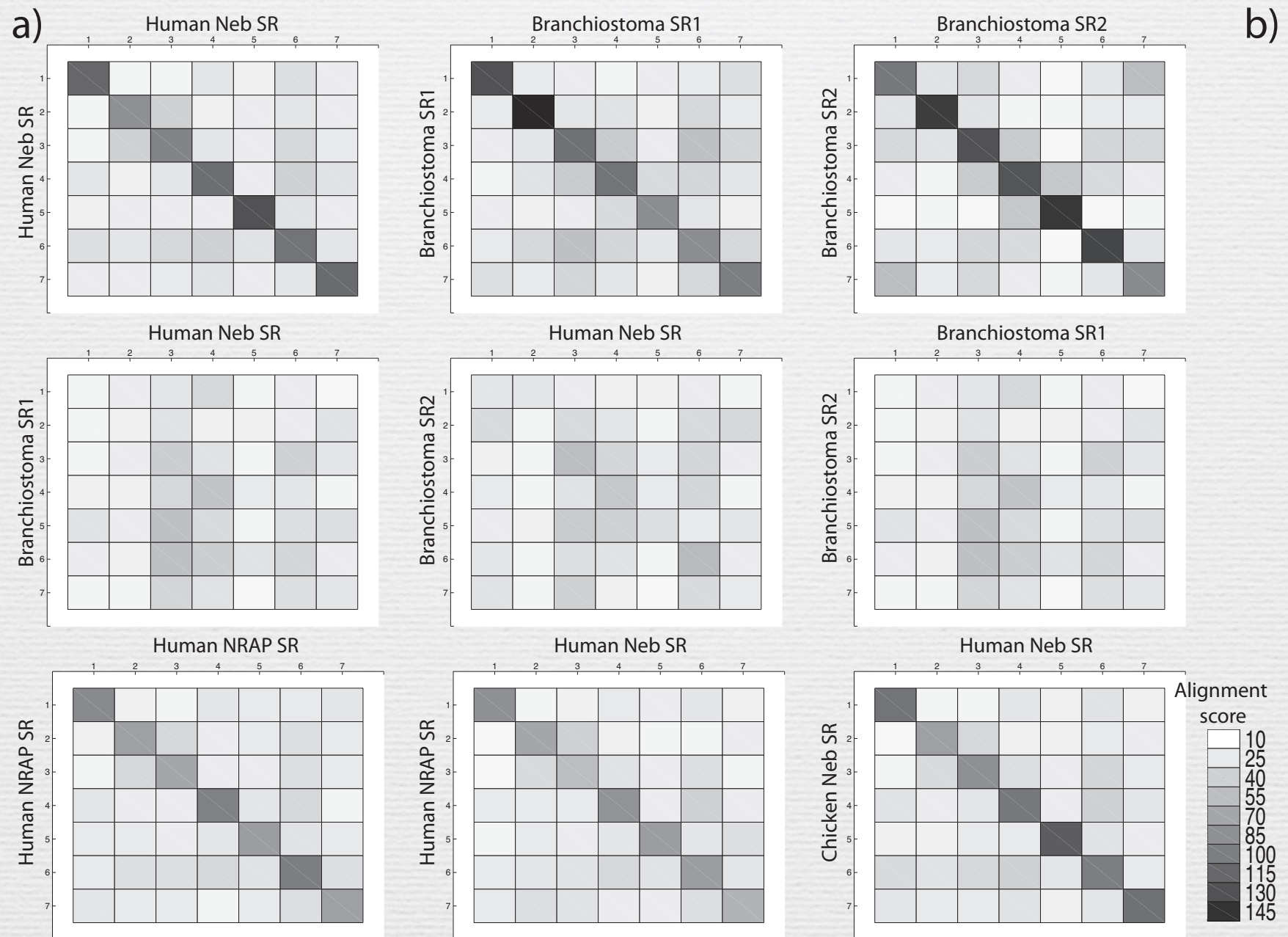


Branchiostoma SR2 (23 SR units)



it
0550





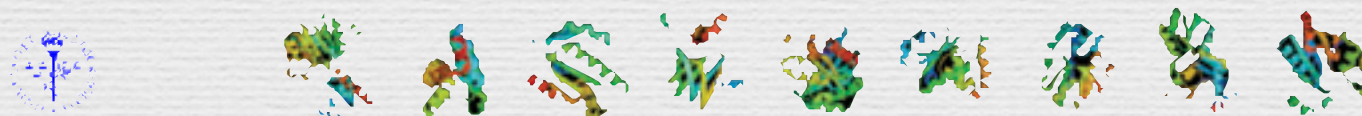
b)

I

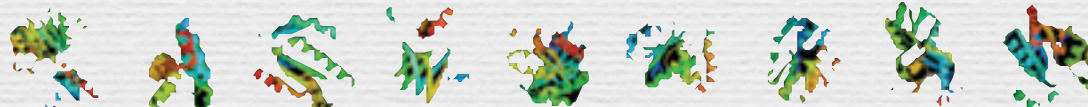
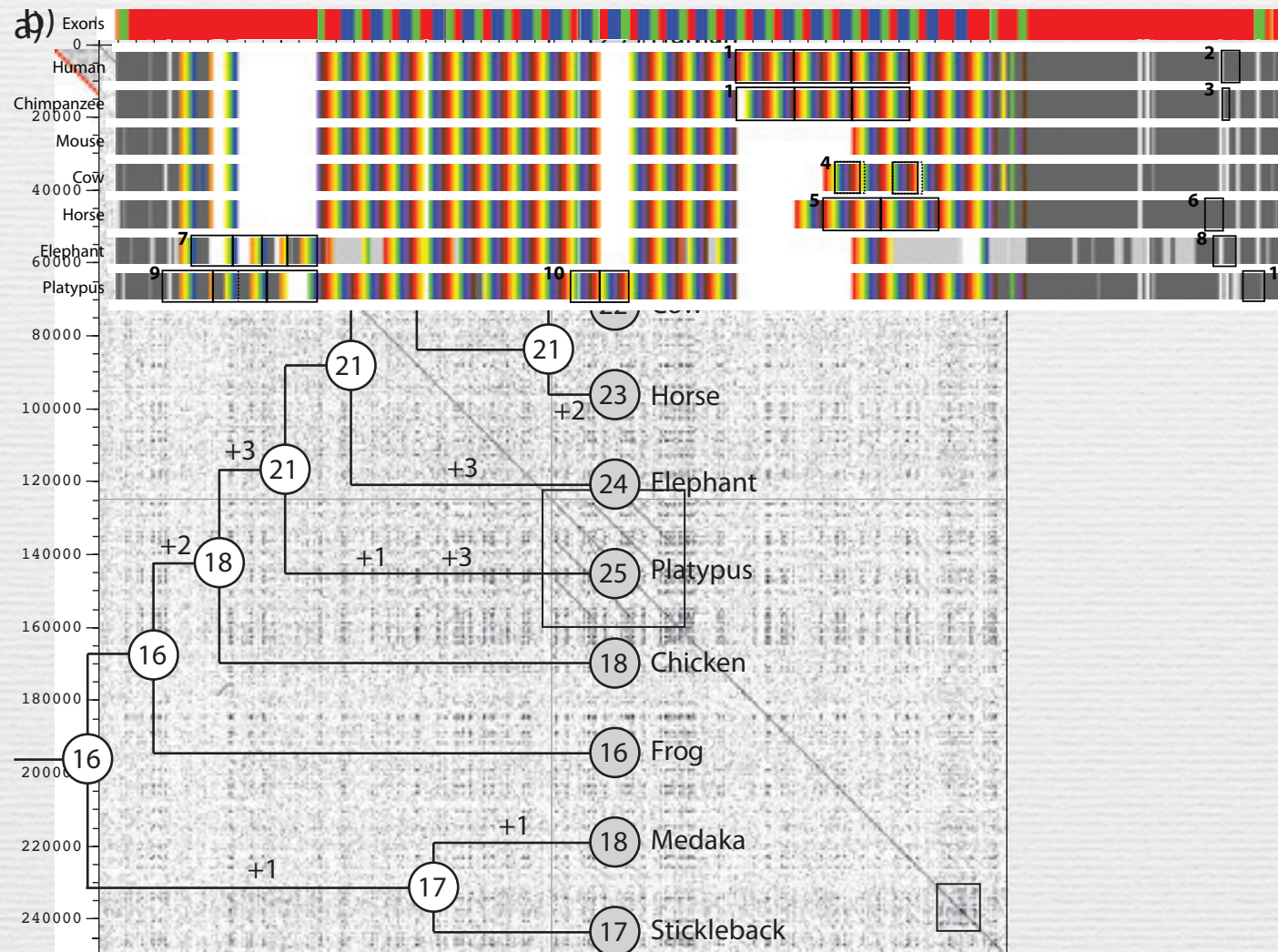
I

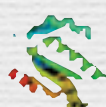
I

I



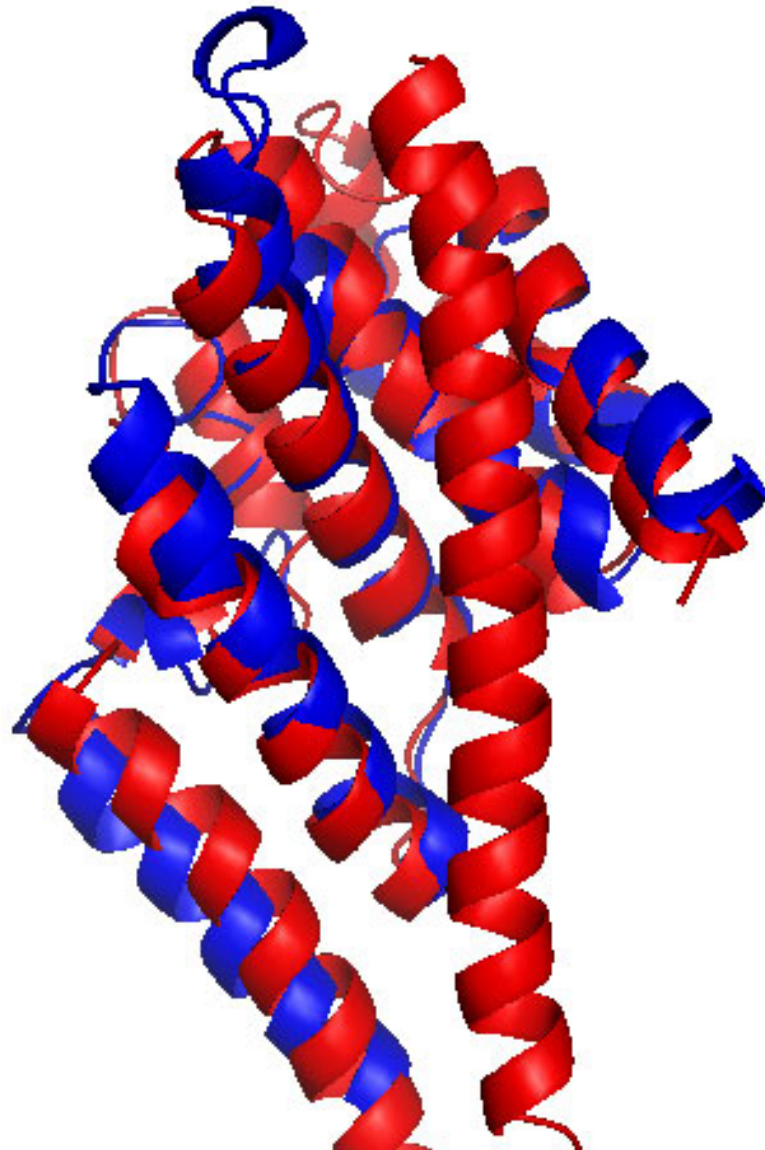
Detection of recent duplications



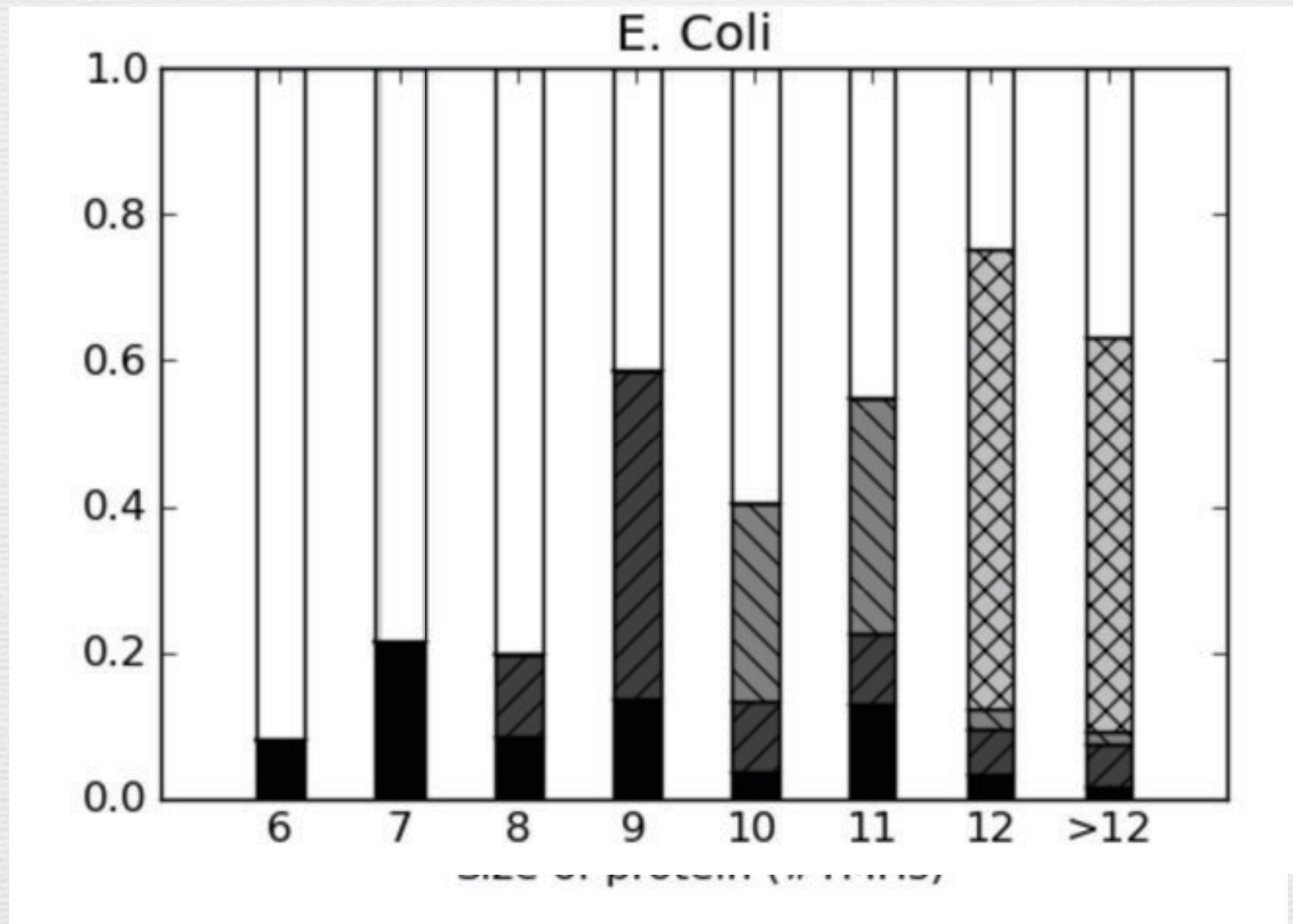


Repeats in TM proteins

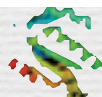
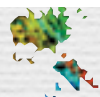
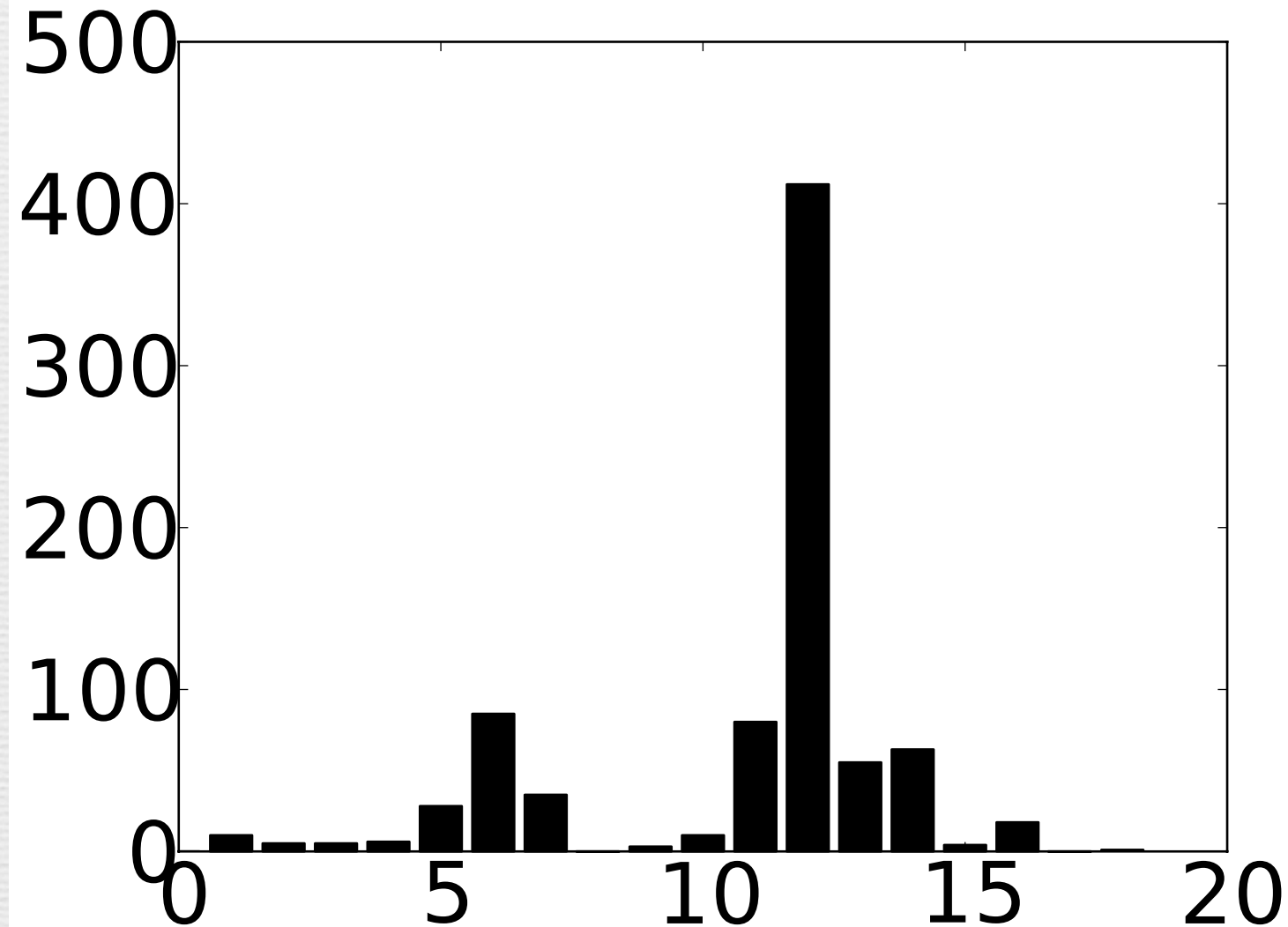
(Hennerdal, Protein Science 2010)



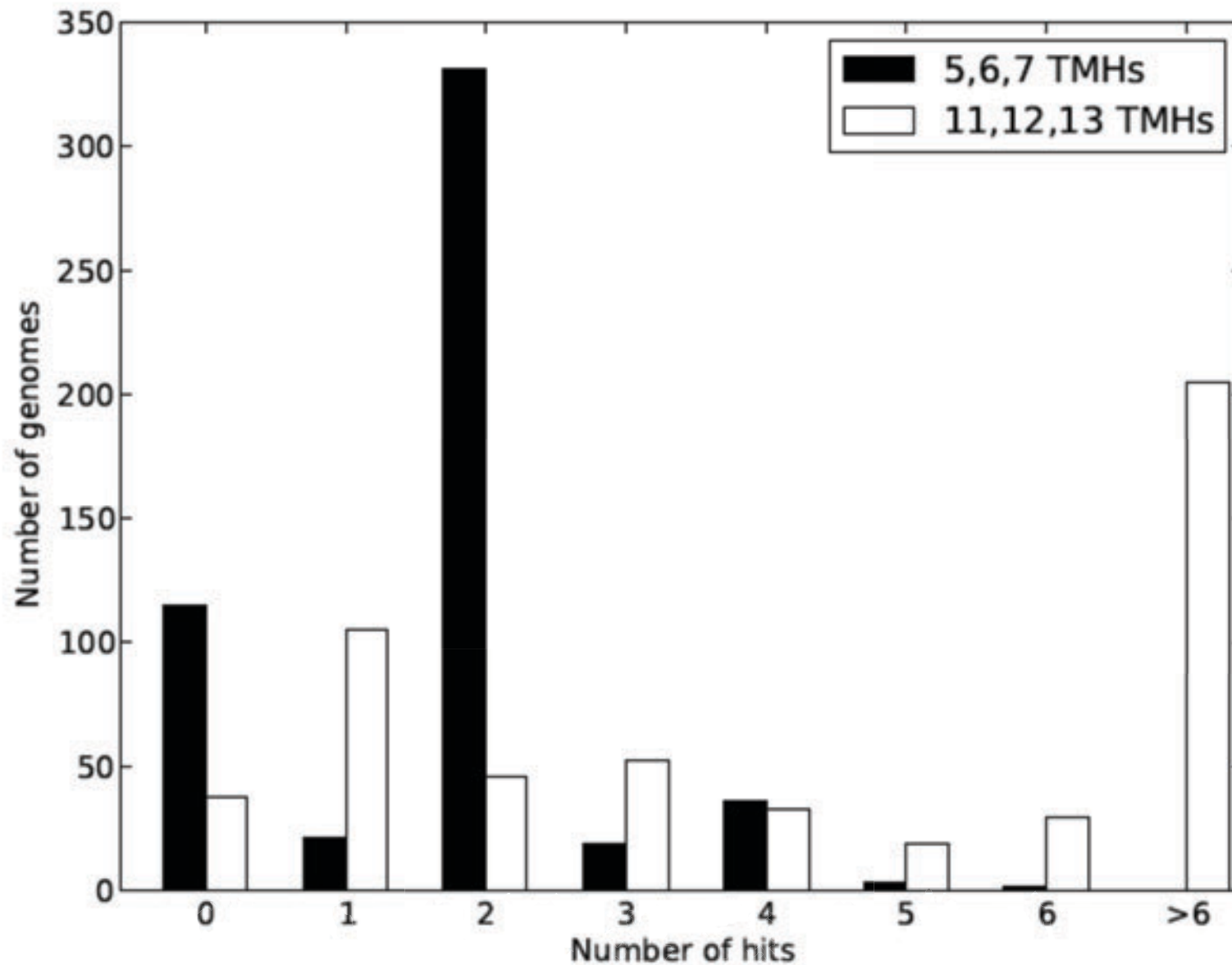
How frequent are duplications



Genome scans



Short version in zero or two copies

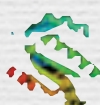
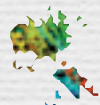
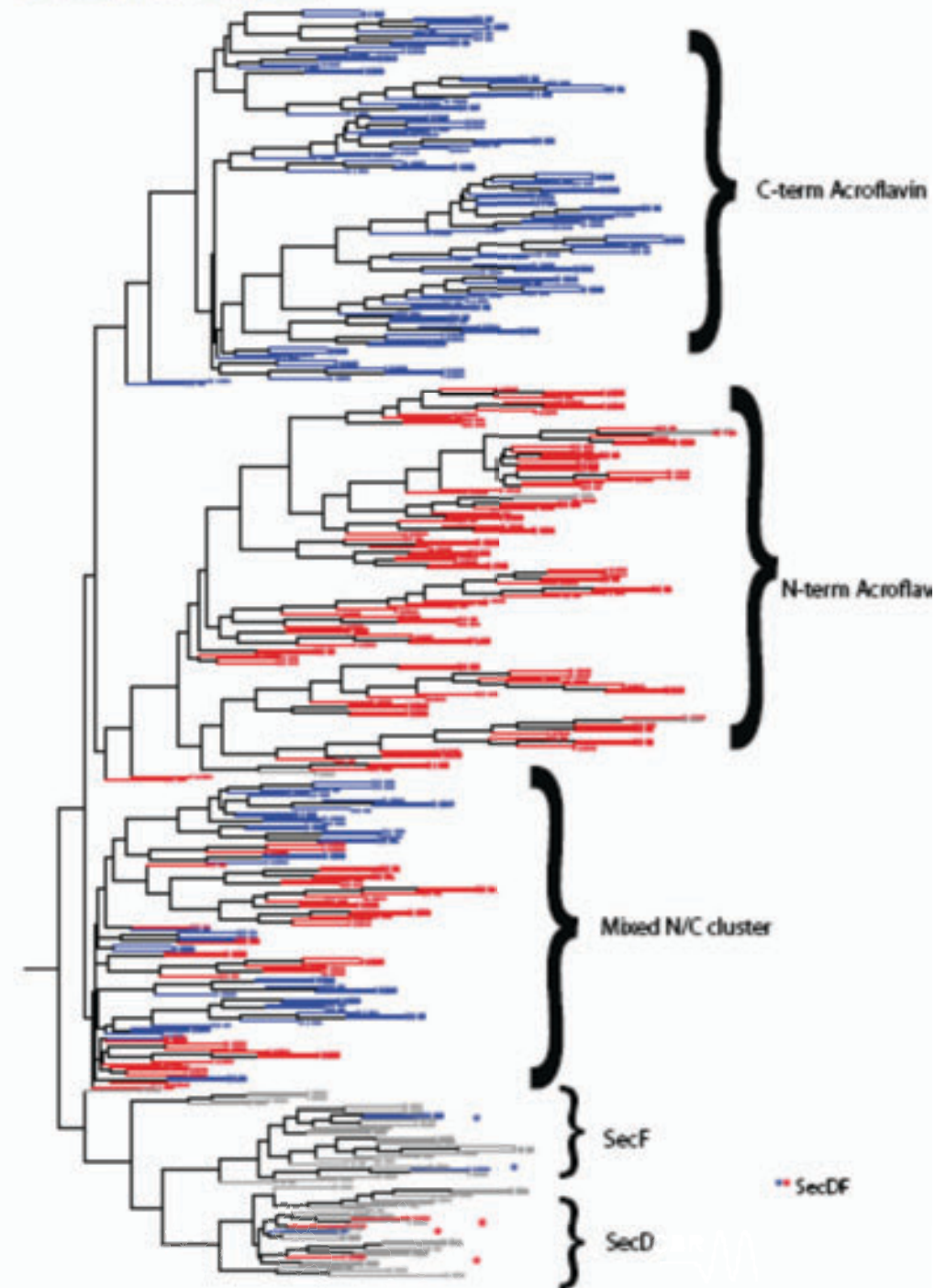


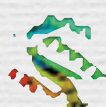
Evolutionary history

6-TM proteins

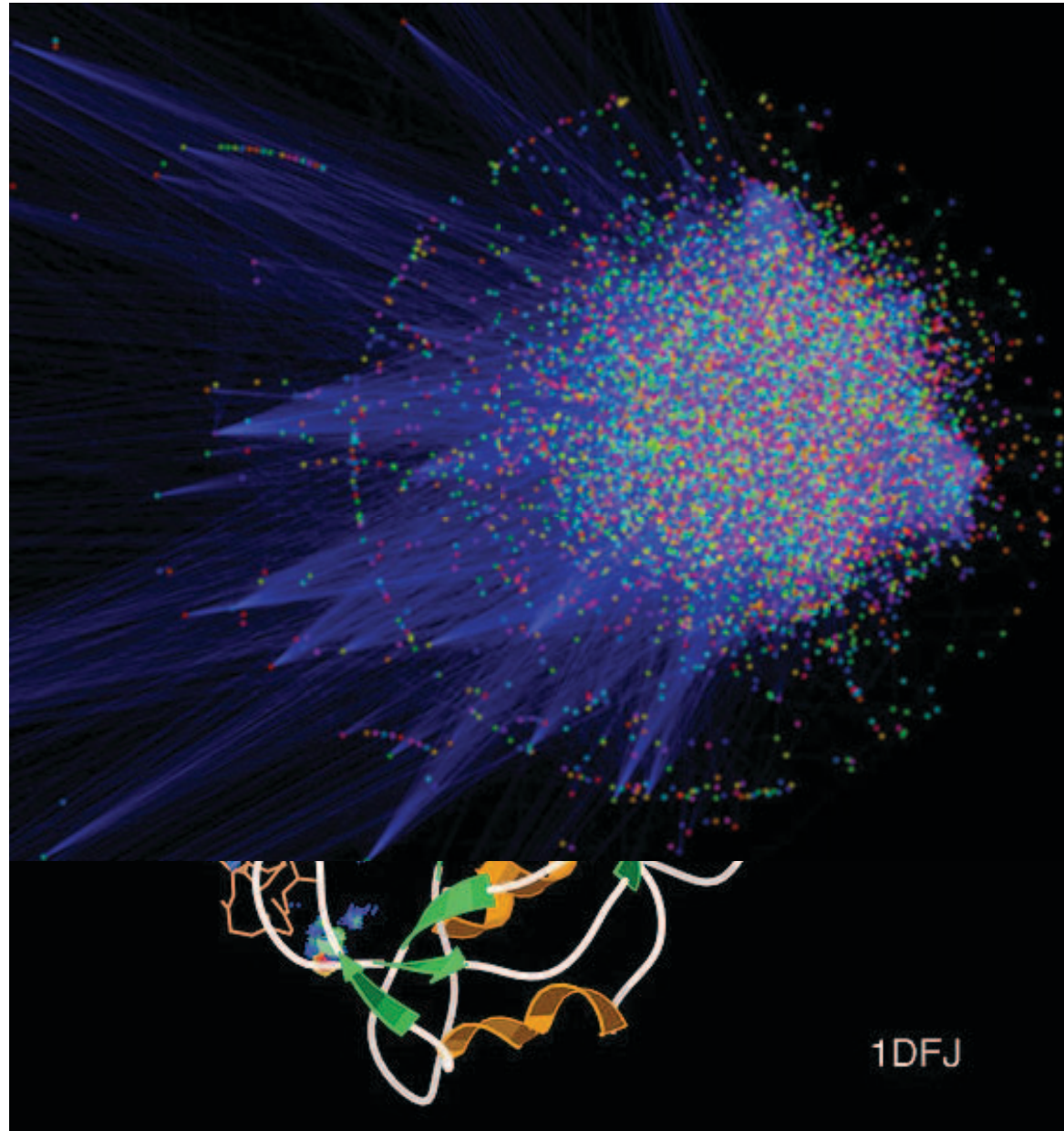
N-terminal part of 12 TM protein

C-terminal part of 12 TM protein

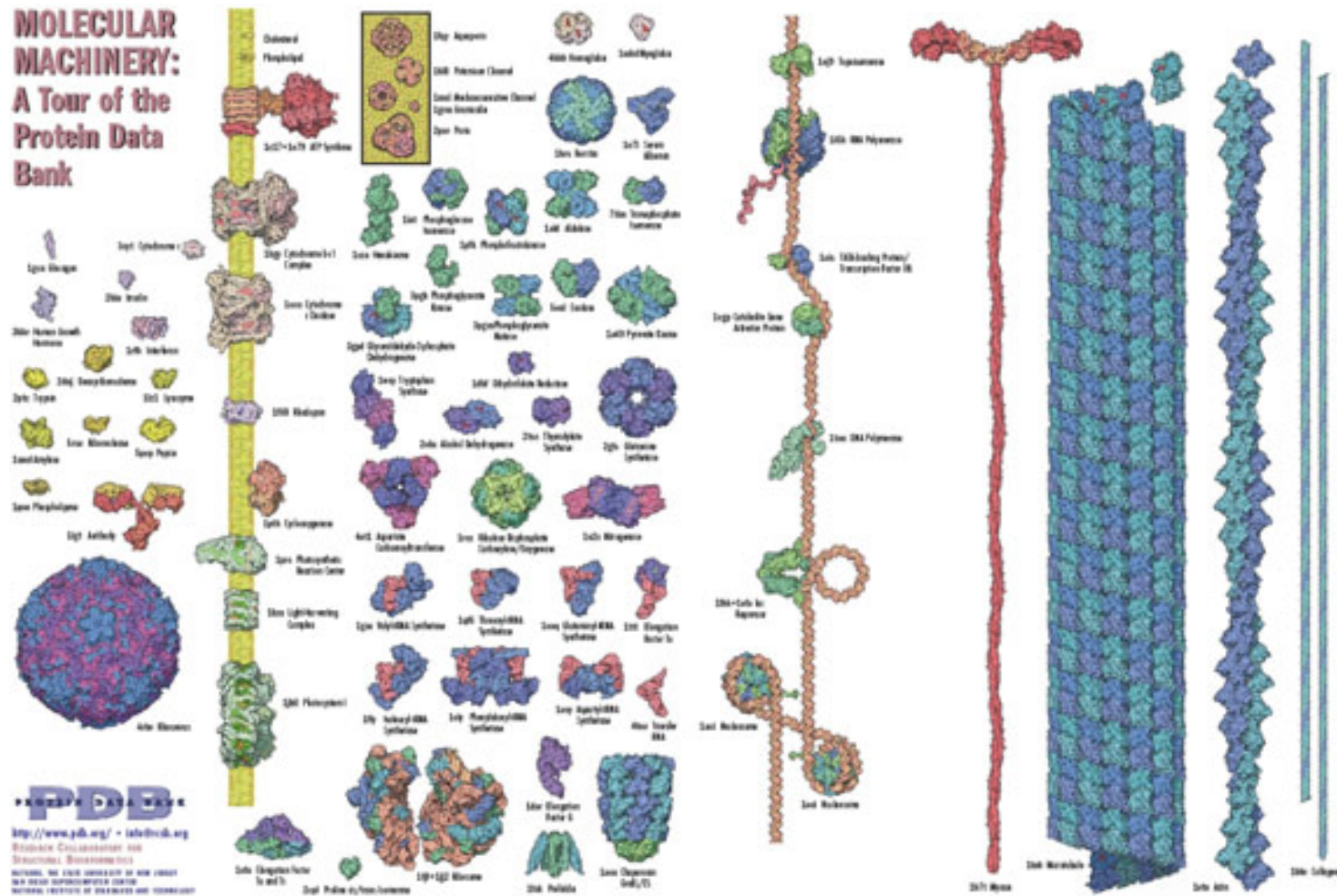




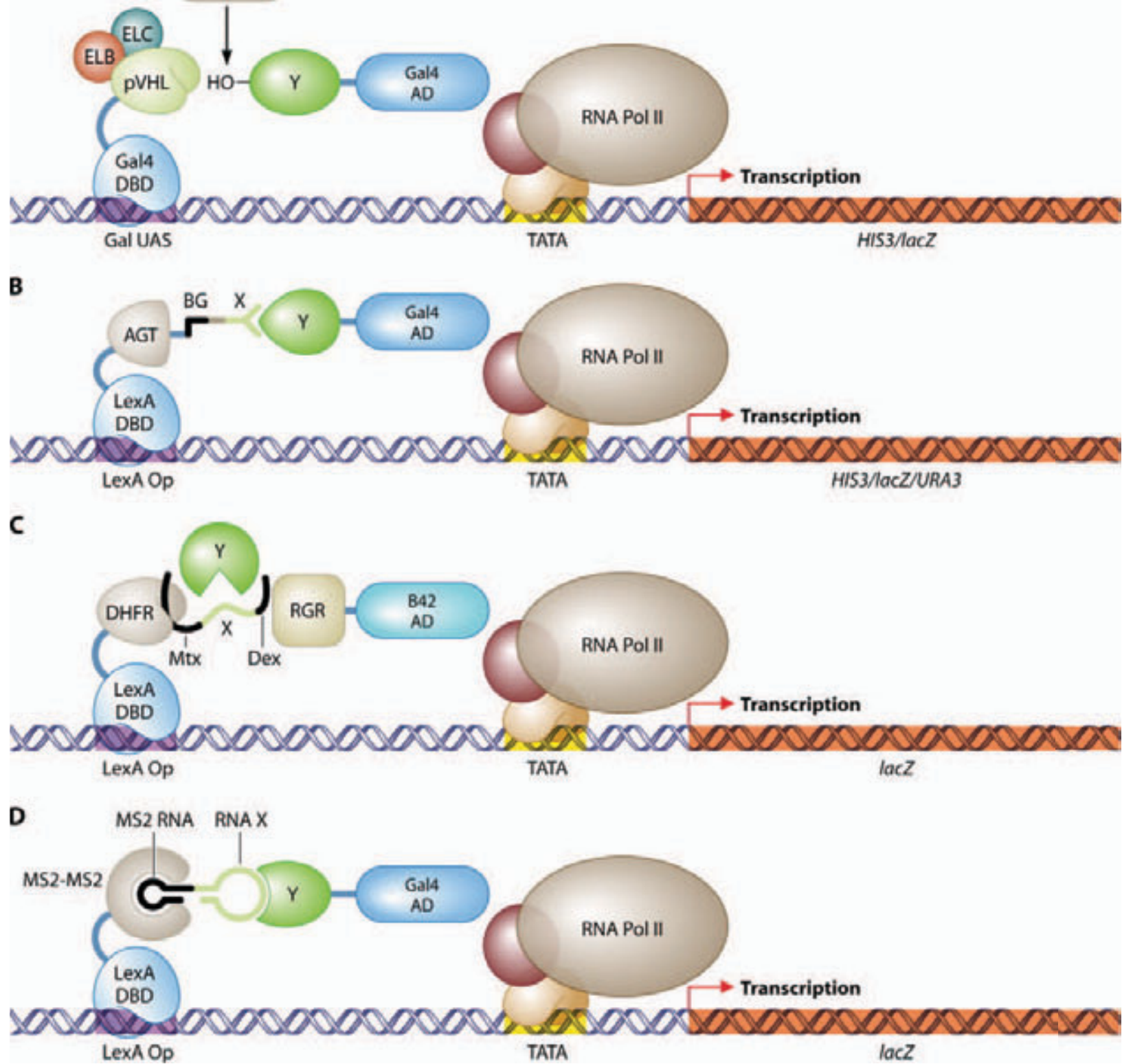
Protein Interactions



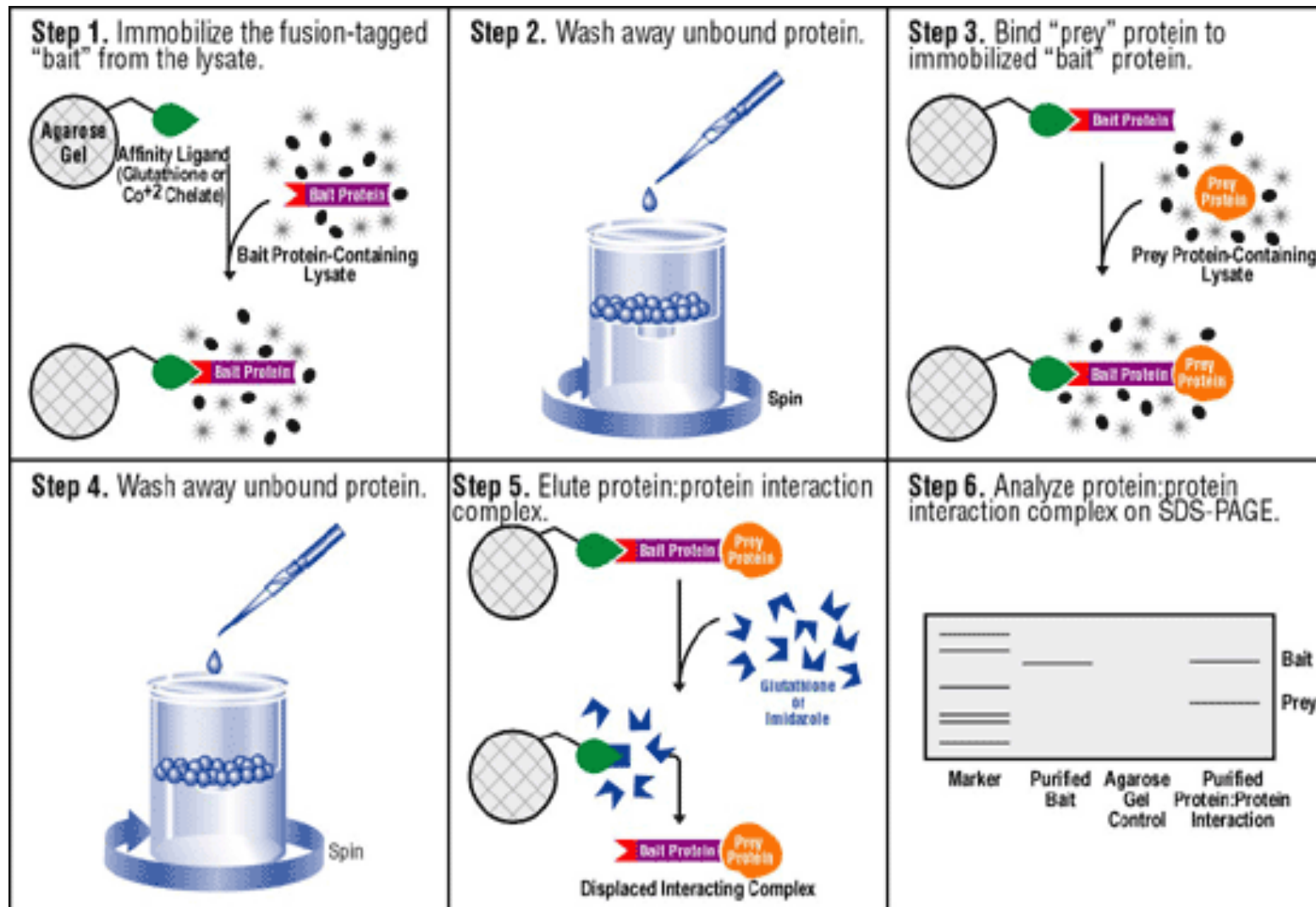
Studies of interactions and networks



Y2H

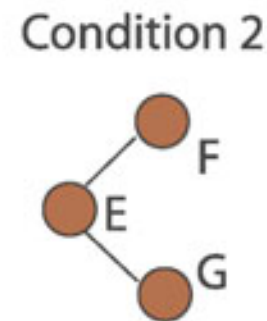
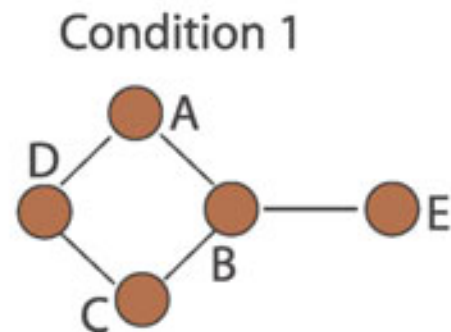
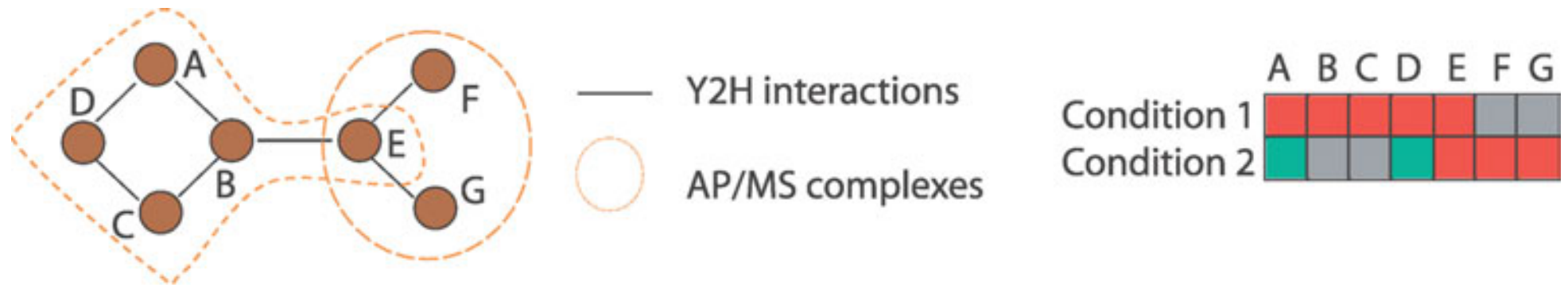


Tap-Tag

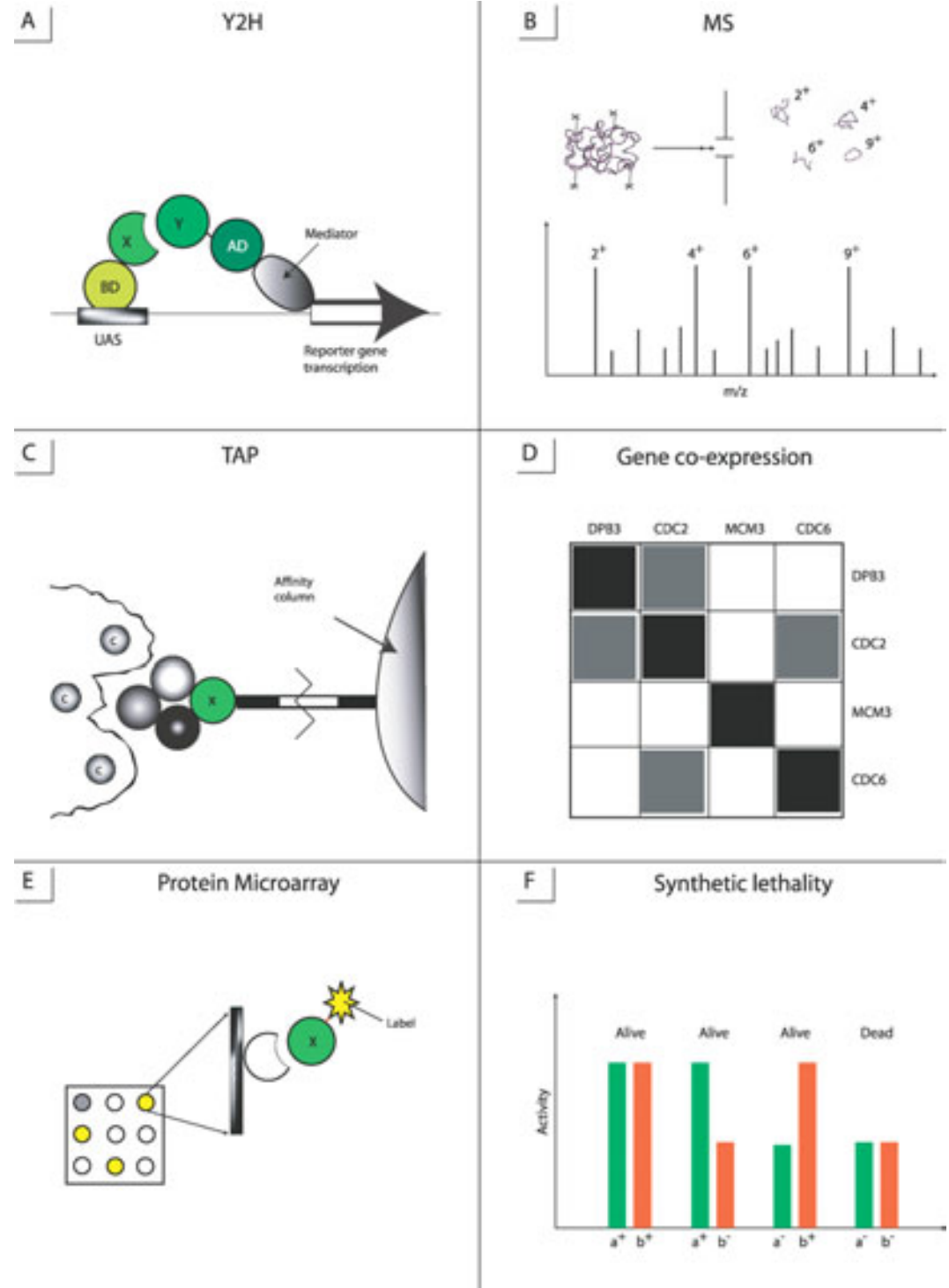


▶ = Fusion Tag (GST or polyHis)

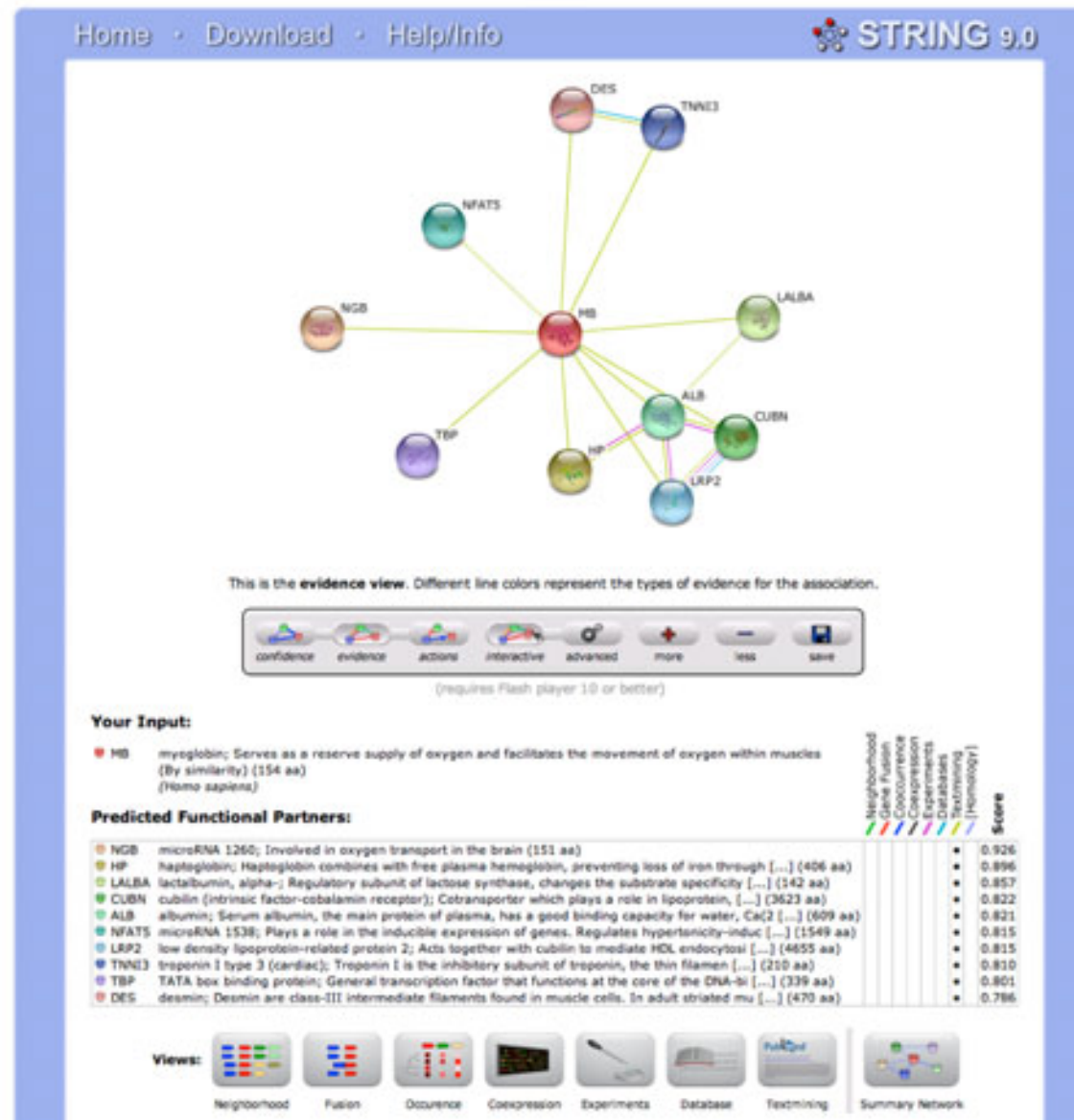
Results depend on conditions



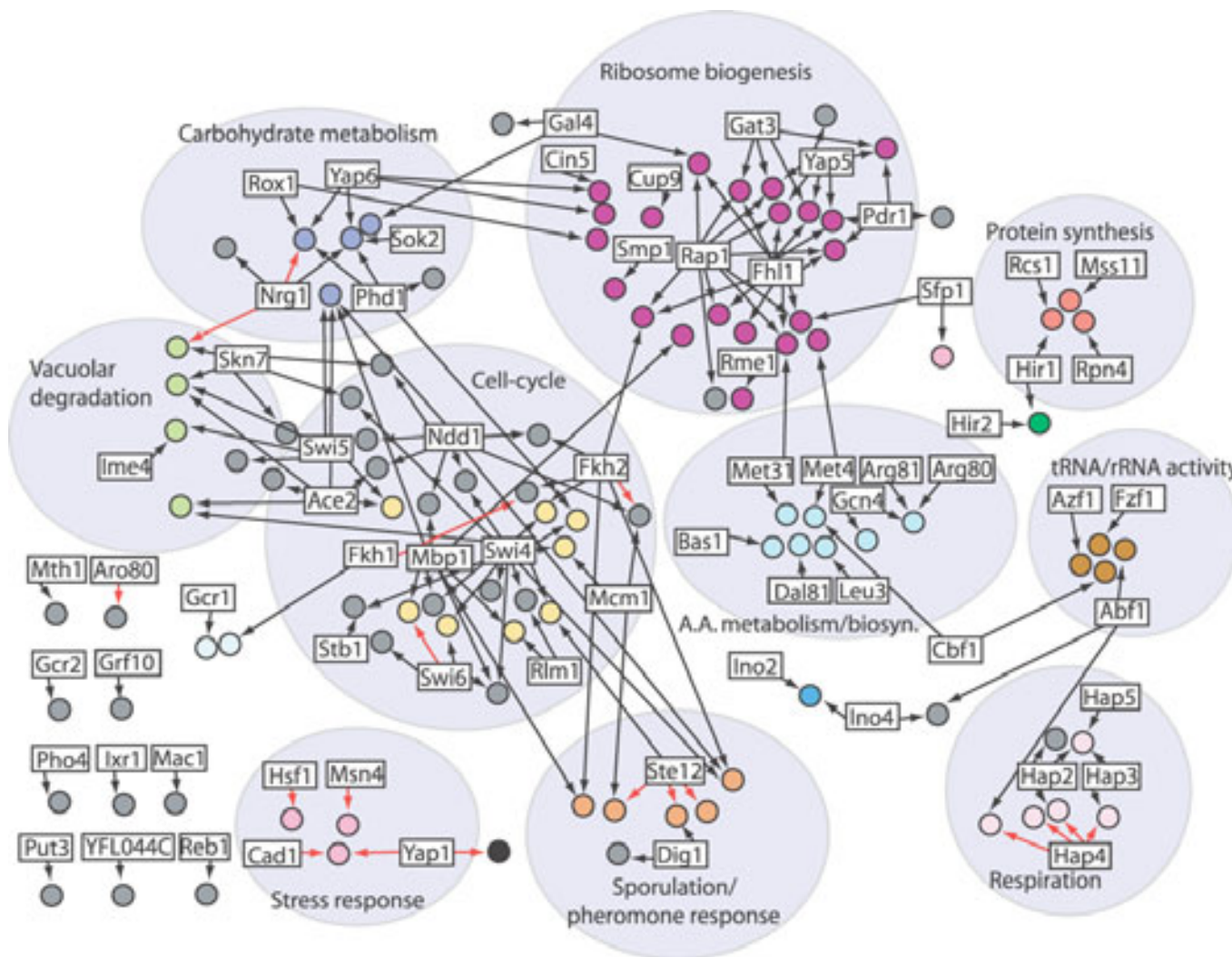
Other methods to detect interactions



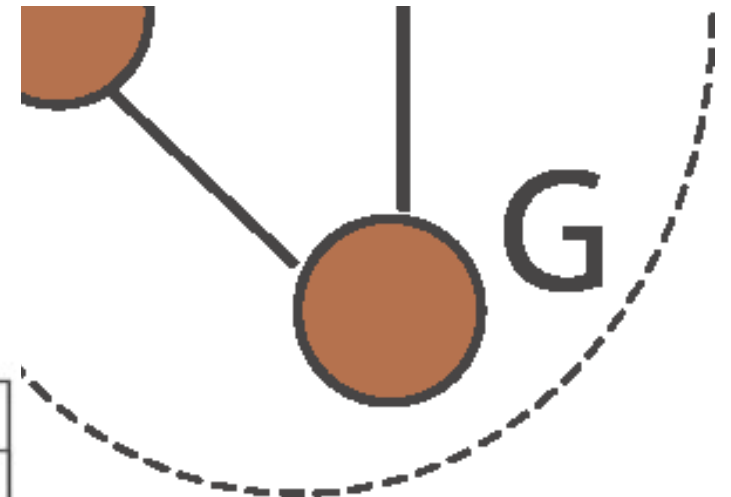
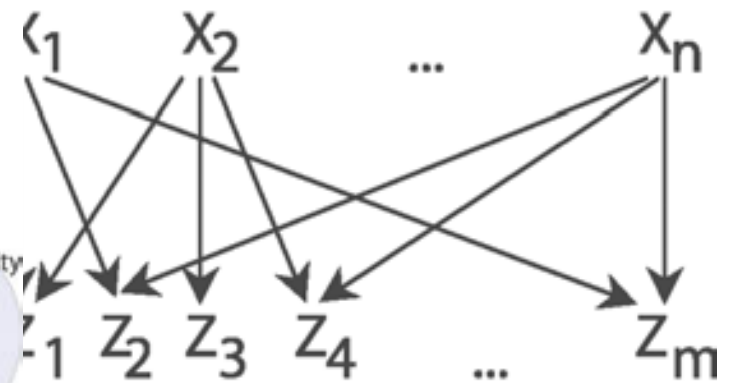
string.embl.de



Interaction networks are modular

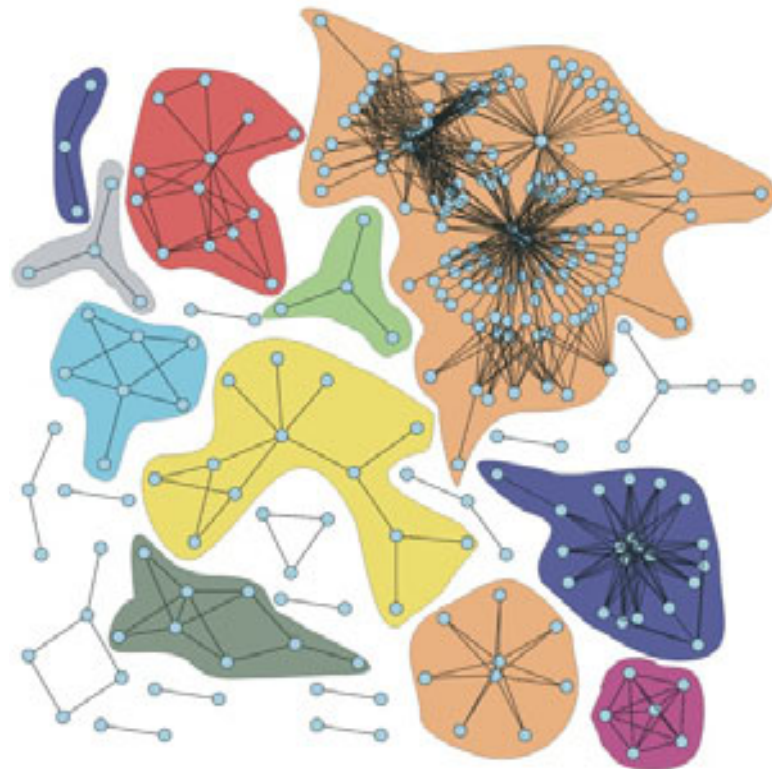


Multiple input motif



Transcription factor	Activator	Function not determined
Module categories:		
tRNA/rRNA activity	Respiration	Carbohydrate metabolism
Ribosome biogenesis	Stress response	Sporulation/pheromone resp.
Amino acid met./biosynth.	Protein synthesis	Chromosome/histone
Glycolysis/metabolism	Fermentation	Lipid/fatty acid biosynth.
		Unknown

A



Signalling

ATP synthase

Secretory pathway

Ubiquitin-proteases

Folding

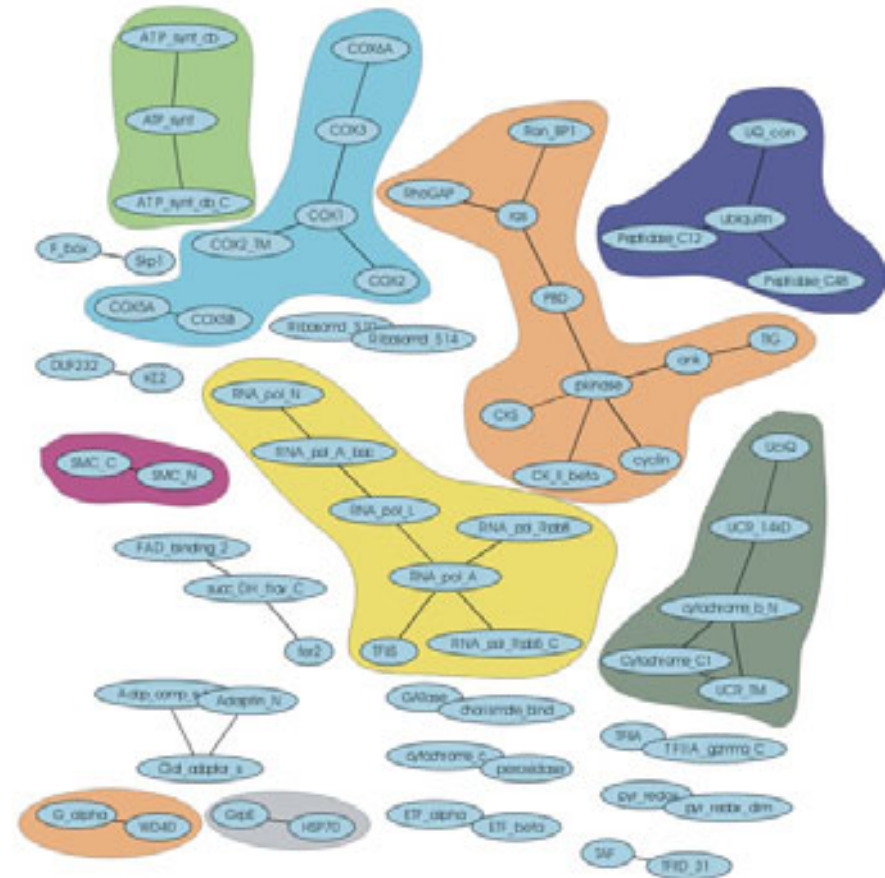
Cytochrome oxidase

RNA polymerase

Cytochrome C1

Chromosome structure

B



Signalling

ATP synthase

Secretory pathway

Ubiquitin-proteases

Folding

Cytochrome oxidase

RNA polymerase

Cytochrome C1

Chromosome structure

Proteomics



MS-based proteomics

