

Scenario

- We have identified (experimentally or/and computationally) some proteins that interact with our protein of interest
- We want to gather functional information about these proteins: are all their regions/domains annotated?
- What can we say about unannotated regions?
We are going to use sequence searches and homology to investigate their putative functions

Goals

- We will learn more on the advantages of profile-based searches vs single sequence searches
- We will learn how to run an iterative profile search (and associated caveats)
- We will learn more on how to edit a multiple sequence alignment
- We will see an example on how to use homology to make functional hypotheses

Ambitious!

Our protein

C-reactive protein (P02741):

- Displays several functions associated with host defense: it promotes agglutination, bacterial capsular swelling, phagocytosis and complement fixation through its calcium-dependent binding to phosphorylcholine. Can interact with DNA and histones and may scavenge nuclear material released from damaged circulating cells (UniProtKB annotation).
- Over-expressed in many diseases, including malignant tumors and cardiovascular disorders.
- Prognostic Marker in Patients with Hepatocellular Carcinoma
- ...

Identification of the C-Reactive Protein Interaction Network Using a Bioinformatics Approach Provides Insights into the Molecular Pathogenesis of Hepatocellular Carcinoma

Sha She Lingyu Jiang Zhenfang Zhang Min Yang Huaidong Hu
Peng Hu Yong Liao Yixuan Yang Hong Ren

Institute for Viral Hepatitis, Key Laboratory of Molecular Biology for Infectious Diseases (Ministry of Education), Department of Infectious Diseases, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China

Key Words

C-reactive protein (CRP) • Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) • Gene Ontology (GO) • Kyoto Encyclopedia of Genes and Genomes (KEGG)

Abstract

Background/Aims: C reactive protein (CRP) levels are elevated in many diseases, including malignant tumors and cardiovascular disorders. In this study, the protein interaction network for CRP was evaluated to determine the importance of CRP and its interacting proteins in the molecular pathogenesis of hepatocellular carcinoma (HCC). **Methods:** Isobaric tags for relative and absolute quantitation (iTRAQ) and mass spectrometry were used to identify CRP interacting proteins in SMMC7721 cells. Moreover, Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) were used to evaluate enriched genes and pathways for differentially expressed genes using DAVID and WebGestalt. Co-immunoprecipitation and western blot analyses were employed to assess interactions between CRP and KRT8, ANXA2, ENO2, and HSP90B1. **Results:** In total, 52 proteins that interact with CRP were identified. A GO analysis suggested that most of the interacting proteins were involved in CRP complexes and regulated metabolic processes. A KEGG pathway analysis suggested that most CRP-interacting proteins contribute to the TRAIL signaling pathway, Class I PI3K/Akt signaling pathway, plasma membrane estrogen receptor signaling, Nectin adhesion pathway, and S1P1 pathway. Immunoprecipitation and western blot analyses revealed interactions between CRP and KRT8, ANXA2, ENO2, and HSP90B1. **Conclusions:** iTRAQ based proteomic profiling revealed the network of CRP interacting proteins. This network may activate the PI3K/Akt signaling pathway, thereby contributing to the pathogenesis of HCC.

© 2018 The Author(s)
Published by S. Karger AG, Basel

S. She, L. Jiang and Z. Zhang contributed equally to this work.

Prof. Yixuan Yang
and Hong Ren

Department of Infectious Diseases, The Second Affiliated Hospital of Chongqing
Medical University 76# Linjiang Road, Chongqing 400010 (China)
Tel. 86 23 63693215, Fax 86 23 63693215, E-Mail yummy_8813@sina.com

Institute for Viral Hepatitis, Key Laboratory of Molecular Biology for Infectious Diseases (Ministry of Education), Department of Infectious Diseases, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China

Key Words

C-reactive protein (CRP) • Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) • Gene Ontology (GO) • Kyoto Encyclopedia of Genes and Genomes (KEGG)

Abstract

Background/Aims: C reactive protein (CRP) levels are elevated in many diseases, including malignant tumors and cardiovascular disorders. In this study, the protein interaction network for CRP was evaluated to determine the importance of CRP and its interacting proteins in the molecular pathogenesis of hepatocellular carcinoma (HCC). **Methods:** Isobaric tags for relative and absolute quantitation (iTRAQ) and mass spectrometry were used to identify CRP interacting proteins in SMMC7721 cells. Moreover, Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) were used to evaluate enriched genes and pathways for differentially expressed genes using DAVID and WebGestalt. Co-immunoprecipitation and western blot analyses were employed to assess interactions between CRP and KRT8, ANXA2, ENO2, and HSP90B1. **Results:** In total, 52 proteins that interact with CRP were identified. A GO analysis suggested that most of the interacting proteins were involved in CRP complexes and regulated metabolic processes. A KEGG pathway analysis suggested that most CRP-interacting proteins contribute to the TRAIL signaling pathway, Class I PI3K/Akt signaling pathway, plasma membrane estrogen receptor signaling, Nectin adhesion pathway, and S1P1 pathway. Immunoprecipitation and western blot analyses revealed interactions between CRP and KRT8, ANXA2, ENO2, and HSP90B1. **Conclusions:** iTRAQ based proteomic profiling revealed the network of CRP interacting proteins. This network may activate the PI3K/Akt signaling pathway, thereby contributing to the pathogenesis of HCC.

Institute for Viral Hepatitis, Key Laboratory of Molecular Biology for Infectious Diseases (Ministry of Education), Department of Infectious Diseases, The Second Affiliated Hospital, Chongqing Medical University, Chongqing, China

Key Words

C-reactive protein (CRP) • Isobaric Tags for Relative and Absolute Quantitation (iTRAQ) • Gene Ontology (GO) • Kyoto Encyclopedia of Genes and Genomes (KEGG)

Abstract

Background/Aims: C reactive protein (CRP) levels are elevated in many diseases, including malignant tumors and cardiovascular disorders. In this study, the protein interaction network for CRP was evaluated to determine the importance of CRP and its interacting proteins in the molecular pathogenesis of hepatocellular carcinoma (HCC). **Methods:** Isobaric tags for relative and absolute quantitation (iTRAQ) and mass spectrometry were used to identify CRP interacting proteins in SMMC7721 cells. Moreover, Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) were used to evaluate enriched genes and pathways for differentially expressed genes using DAVID and WebGestalt. Co-immunoprecipitation and western blot analyses were employed to assess interactions between CRP and KRT8, ANXA2, ENO2, and HSP90B1. **Results:** In total, 52 proteins that interact with CRP were identified. A GO analysis suggested that most of the interacting proteins were involved in CRP complexes and regulated metabolic processes. A KEGG pathway analysis suggested that most CRP-interacting proteins contribute to the TRAIL signaling pathway, Class I PI3K/Akt signaling pathway, plasma membrane estrogen receptor signaling, Nectin adhesion pathway, and S1P1 pathway. Immunoprecipitation and western blot analyses revealed interactions between CRP and KRT8, ANXA2, ENO2, and HSP90B1. **Conclusions:** iTRAQ based proteomic profiling revealed the network of CRP interacting proteins. This network may activate the PI3K/Akt signaling pathway, thereby contributing to the pathogenesis of HCC.

UniProtKB - P05787 (K2C8_HUMAN)

Display

[BLAST](#) [Align](#) [Format](#) [Add to basket](#) [History](#)

Entry

Publications

Feature viewer

Feature table

Protein | Keratin, type II cytoskeletal 8

Gene | KRT8

Organism | Homo sapiens (Human)

Status | Reviewed - Annotation score: - Experimental evidence at protein levelⁱ

None

 Function Names & Taxonomy Subcellular location Pathology & Biotech PTM / Processing Expression Interaction Structure Family & Domains Sequences (2+) Similar proteins Cross-references Entry information Miscellaneous[Top](#)Functionⁱ

Together with KRT19, helps to link the contractile apparatus to dystrophin at the costameres of striated muscle. 1 Publication ▾

Miscellaneous

There are two types of cytoskeletal and microfibrillar keratin: I (acidic; 40-55 kDa) and II (neutral to basic; 56-70 kDa).

Sites

Feature key	Position(s)	Description
Site ⁱ	342	Stutter

GO - Molecular functionⁱ

- protein-containing complex binding
- scaffold protein binding

[Complete GO annotation on QuickGO ...](#)GO - Biological processⁱ

- cell differentiation involved in embryonic placenta development
- cornification
- extrinsic apoptotic signaling pathway
- hepatocyte apoptotic process
- keratinization
- response to hydrostatic pressure
- response to other organism
- sarcomere organization
- tumor necrosis factor-mediated signaling pathway
- viral process

We want to learn what is known
about the domain composition
of KRT8

UniProtKB - P05787 (K2C8_HUMAN)

Display

[BLAST](#) [Align](#) [Format](#) [Add to basket](#) [History](#)

Entry

Publications

Feature viewer

Feature table

Protein | Keratin, type II cytoskeletal 8

Gene | KRT8

Organism | Homo sapiens (Human)

Status | Reviewed - Annotation score: - Experimental evidence at protein level¹Function¹

Together with KRT19, helps to link the contractile apparatus to dystrophin at the costameres of striated muscle.

1 Publication ▾

Miscellaneous

There are two types of cytoskeletal and microfibrillar keratin: I (acidic; 40-55 kDa) and II (neutral to basic; 56-70 kDa).

Sites

Feature key	Position(s)	Description
Site ¹	342	Stutter

GO - Molecular function¹

- protein-containing complex binding
- scaffold protein binding

[Complete GO annotation on QuickGO ...](#)GO - Biological process¹

- cell differentiation involved in embryonic placenta development
- cornification
- extrinsic apoptotic signaling pathway
- hepatocyte apoptotic process
- keratinization
- response to hydrostatic pressure
- response to other organism
- sarcomere organization
- tumor necrosis factor-mediated signaling pathway
- viral process

Family & Domains

Similar proteins

Cross-references

Entry information

Miscellaneous

[Top](#)

Family & Domainsⁱ

Domains and Repeats

Feature key	Position(s)	Description
Domain ⁱ	91 – 402	IF rod  PROSITE-ProRule annotation ▾

Region

Feature key	Position(s)	Description
Region ⁱ	1 – 90	Head
Region ⁱ	91 – 126	Coil 1A
Region ⁱ	127 – 143	Linker 1
Region ⁱ	144 – 235	Coil 1B
Region ⁱ	236 – 259	Linker 12
Region ⁱ	260 – 398	Coil 2
Region ⁱ	261 – 382	Necessary for interaction with PNN  1 Publication ▾
Region ⁱ	399 – 483	Tail

Compositional bias

Feature key	Position(s)	Description
Compositional bias ⁱ	9 – 49	Ser-rich

Let's visit the Pfam website:
<https://pfam.xfam.org>

Pfam 32.0 (September 2018, 17929 entries)

The Pfam database is a large collection of protein families, each represented by **multiple sequence alignments** and **hidden Markov models (HMMs)**. [More...](#)

QUICK LINKS[SEQUENCE SEARCH](#)[VIEW A PFAM ENTRY](#)[VIEW A CLAN](#)[VIEW A SEQUENCE](#)[VIEW A STRUCTURE](#)[KEYWORD SEARCH](#)[JUMP TO](#)**YOU CAN FIND DATA IN PFAM IN VARIOUS WAYS...**

Analyze your protein sequence for Pfam matches

View Pfam annotation and alignments

See groups of related entries

Look at the domain organisation of a protein sequence

Find the domains on a PDB structure

Query Pfam by keywords

Go**Example**

Enter any type of accession or ID to jump to the page for a Pfam entry or clan, UniProt sequence, PDB structure, etc.

Or view the [help](#) pages for more information

Recent Pfam [blog](#) posts Hide this[Pfam 31.0 is released](#) (posted 8 March 2017)

Pfam 31.0 contains a total of 16712 families and 604 clans. Since the last release, we have built 415 new families, killed 9 families and created 11 new clans. We have also been working on expanding our clan

Summary**Sequence****Structures****TreeFam****Jump to...**

enter ID/acc

Go

SummaryThis is the summary of UniProt entry [K2C8_HUMAN](#) (P05787).**Description:** Keratin, type II cytoskeletal 8**Source organism:** [Homo sapiens \(Human\)](#) (NCBI taxonomy ID [9606](#))
[View Pfam proteome data](#).**Length:** 483 amino acids**Reference Proteome:** ✓

Please note: when we start each new Pfam data release, we take a copy of the UniProt sequence database. This snapshot of UniProt forms the basis of the overview that you see here. It is important to note that, although some UniProt entries may be removed after a Pfam release, these entries will not be removed from Pfam until the next Pfam data release.

Pfam domains

This image shows the arrangement of the Pfam domains that we found on this sequence. Clicking on a domain will take you to the page describing that Pfam entry. The table below gives the domain boundaries for each of the domains. [More...](#)



[Download the data used to generate the domain graphic in JSON format.](#)

Source	Domain	Start	End
Pfam	Keratin_2_head	6	87
disorder	n/a	18	26
low_complexity	n/a	23	48
disorder	n/a	31	34
disorder	n/a	36	38
low_complexity	n/a	41	61
disorder	n/a	89	91
Pfam	Filament	90	401
coiled_coil	n/a	95	115
coiled_coil	n/a	144	192
coiled_coil	n/a	200	234
disorder	n/a	297	298
disorder	n/a	300	309
coiled_coil	n/a	300	380
disorder	n/a	311	312
low_complexity	n/a	393	404
disorder	n/a	402	411
disorder	n/a	413	415
low_complexity	n/a	415	431
low_complexity	n/a	426	459
disorder	n/a	477	479

Summary**Sequence****Structures****TreeFam****Jump to...**

enter ID/acc

Go

SummaryThis is the summary of UniProt entry [K2C8_HUMAN](#) (P05787).**Description:** Keratin, type II cytoskeletal 8**Source**Homo sapiens (Human) (NCBI taxonomy ID [9606](#))

with Pfam proteome data.

3 amino acids

Reference Proteomes

Please note: when we start each new Pfam data release, we take a copy of the UniProt sequence database. This snapshot of UniProt forms the basis of the overview that you see here. It is important to note that, although some UniProt entries may be removed after a Pfam release, these entries will not be removed from Pfam until the next Pfam data release.

Pfam domains

This image shows the arrangement of the Pfam domains that we found on this sequence. Clicking on a domain will take you to the page describing that Pfam entry. The table below gives the domain boundaries for each of the domains. [More...](#)



[Download](#) the data used to generate the domain graphic in JSON format.

Source	Domain	Start	End
Pfam	Keratin_2_head	6	87
disorder	n/a	18	26
low_complexity	n/a	23	48
disorder	n/a	31	34
disorder	n/a	36	38
low_complexity	n/a	41	61
disorder	n/a	89	91
Pfam	Filament	90	401
coiled_coil	n/a	95	115
coiled_coil	n/a	144	192
coiled_coil	n/a	200	234
disorder	n/a	297	298
disorder	n/a	300	309
coiled_coil	n/a	300	380
disorder	n/a	311	312
low_complexity	n/a	393	404
disorder	n/a	402	411
disorder	n/a	413	415
low_complexity	n/a	415	431
low_complexity	n/a	426	459
disorder	n/a	477	479

Family: Filament (PF00038)

119 architectures

9043 sequences

3 interactions

327 species

43 structures

Summary

Domain organisation

Clan

Alignments

HMM logo

Trees

Curation & model

Species

Interactions

Structures

Jump to... ⓘ

enter ID/acc

Go

Summary: Intermediate filament protein

Pfam includes annotations and additional family information from a range of different sources. These sources can be accessed via the tabs below.

[Wikipedia: Intermediate filament](#)

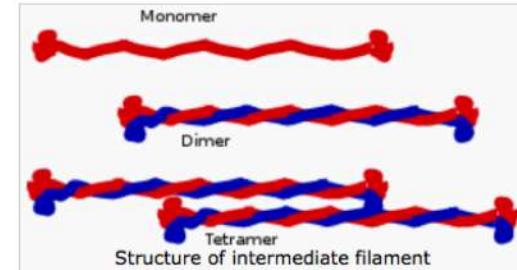
[Pfam](#)

[InterPro](#)

This is the Wikipedia entry entitled "[Intermediate filament](#)". [More...](#)

Intermediate filament [Edit Wikipedia article](#)

Intermediate filaments (IFs) are cytoskeletal components found in the cells of vertebrate animal species,^{[1][2]} and perhaps also in other animals, fungi, plants, and unicellular organisms.^[3] They are composed of a family of related proteins sharing common structural and sequence features. Initially designated 'intermediate' because their average diameter (10 nm) is between those of narrower microfilaments (actin) and wider myosin filaments found in muscle cells, the diameter of intermediate filaments is now commonly compared to actin microfilaments (7 nm) and microtubules (25 nm).^{[1][4]} Most types of intermediate filaments are cytoplasmic, but one type, the lamins, are nuclear. Unlike microtubules, IFs distribution in cells show no good correlation with the distribution of either mitochondria or endoplasmic reticulum.^[5]



Contents [hide]

- 1 Structure
- 2 Biomechanical properties
- 3 Types
 - 3.1 Types I and II – acidic and basic keratins
 - 3.2 Type III
 - 3.3 Type IV
 - 3.4 Type V - nuclear lamins
 - 3.5 Type VI
 - 3.6 Unclassified

Intermediate filament tail domain



Click

Family: Filament (PF00038)

119 architectures

9043 sequences

3 interactions

327 species

43 structures

Summary

Domain organisation

Clan

Alignments

HMM logo

Trees

Curation & model

Species

Interactions

Structures

Jump to... ⌂

enter ID/acc

Go

Summary: Intermediate filament protein

Pfam includes annotations and additional family information from a range of different sources. These sources can be accessed via the tabs below.

Wikipedia: Intermediate filament

Pfam

InterPro

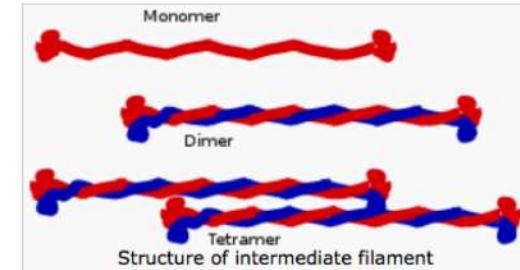
This is the Wikipedia entry entitled "[Intermediate filament](#)". [More...](#)

Intermediate filament Edit Wikipedia article

Intermediate filaments (IFs) are cytoskeletal components found in the cells of vertebrate animal species,^{[1][2]} and perhaps also in other animals, fungi, plants, and unicellular organisms.^[3] They are composed of a family of related proteins sharing common structural and sequence features. Initially designated 'intermediate' because their average diameter (10 nm) is between those of narrower microfilaments (actin) and wider myosin filaments found in muscle cells, the diameter of intermediate filaments is now commonly compared to actin microfilaments (7 nm) and microtubules (25 nm).^{[1][4]} Most types of intermediate filaments are cytoplasmic, but one type, the lamins, are nuclear. Unlike microtubules, IFs distribution in cells show no good correlation with the distribution of either mitochondria or endoplasmic reticulum.^[5]

Contents [hide]

- 1 Structure
- 2 Biomechanical properties
- 3 Types
 - 3.1 Types I and II – acidic and basic keratins
 - 3.2 Type III
 - 3.3 Type IV
 - 3.4 Type V - nuclear lamins
 - 3.5 Type VI
 - 3.6 Unclassified



Intermediate filament tail domain



Family: Filament (PF00038)

119 architectures

9043 sequences

3 interactions

327 species

43 structures

[Summary](#)[Domain organisation](#)[Clan](#)[Alignments](#)[HMM logo](#)[Trees](#)[Curation & model](#)[Species](#)[Interactions](#)[Structures](#)[Jump to... ↗](#)

enter ID/acc

Go

Summary: Intermediate filament protein

Pfam includes annotations and additional family information from a range of different sources. These sources can be accessed via the tabs below.

[Wikipedia: Intermediate filament](#) [Pfam](#) [InterPro](#)

This tab holds annotation information from the [InterPro](#) database.

InterPro entry IPR039008 ↗

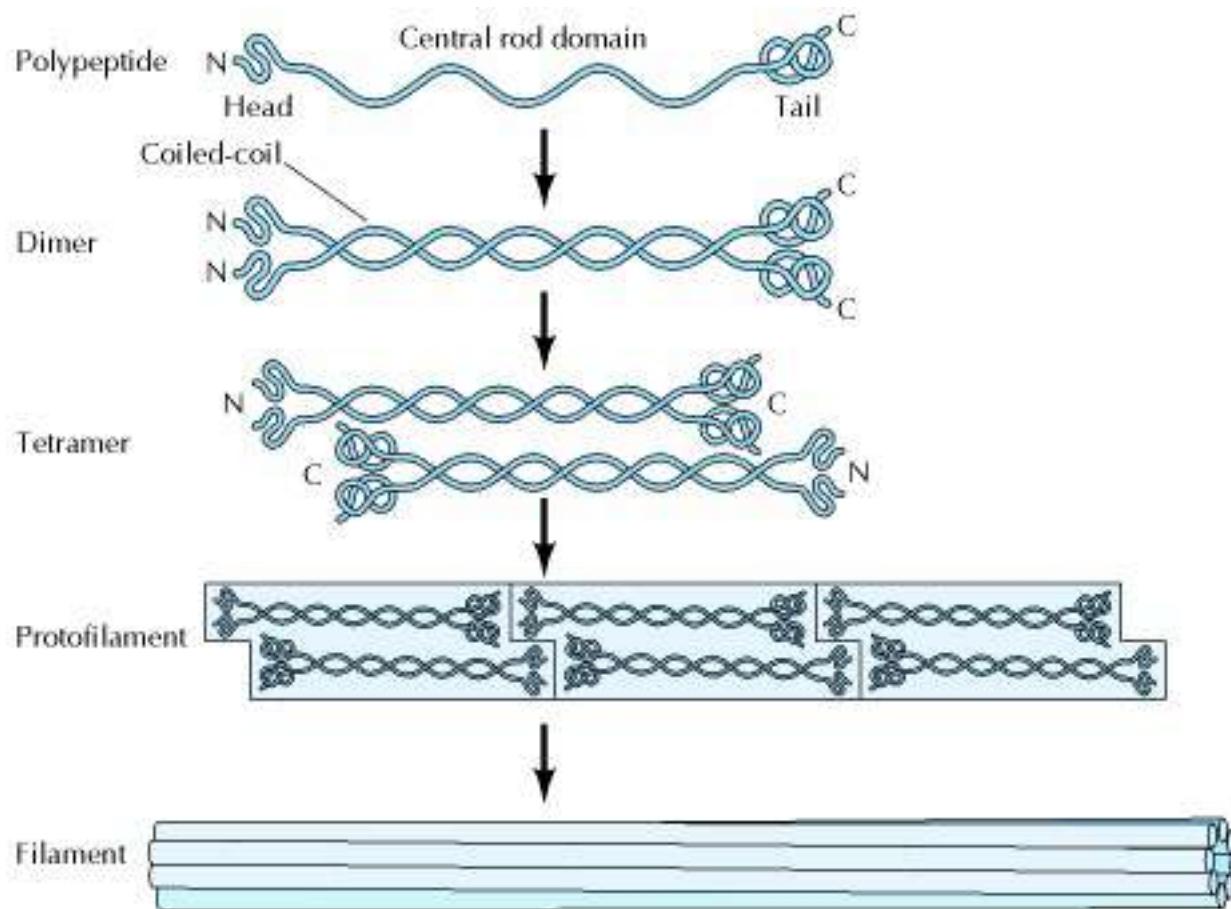
Intermediate filaments (IF) [PUBMED:2183847, PUBMED:28101862] are proteins which are primordial components of the cytoskeleton and the nuclear envelope. They generally form filamentous structures 8 to 14 nm wide. IF proteins are members of a very large multigene family of proteins which has been subdivided in six types:

- Type I: Acidic cytokeratins.
- Type II: Basic cytokeratins.
- Type III: Vimentin, desmin, glial fibrillary acidic protein (GFAP), peripherin, and plastin.
- Type IV: Neurofilaments L, H and M, alpha-internexin and nestin.
- Type V: Nuclear lamins A, B1, B2 and C.
- Type VI: 'Orphan' IF proteins, which are more distant in terms of their amino acid sequences.

All IF proteins are structurally similar in that they consist of: a central rod domain comprising some 300 to 350 residues which is arranged in coiled-coiled alpha-helices, with at least two short characteristic interruptions; a N-terminal non-helical domain (head) of variable length; and a C-terminal domain (tail) which is also non-helical, and which shows extreme length variation between different IF proteins.

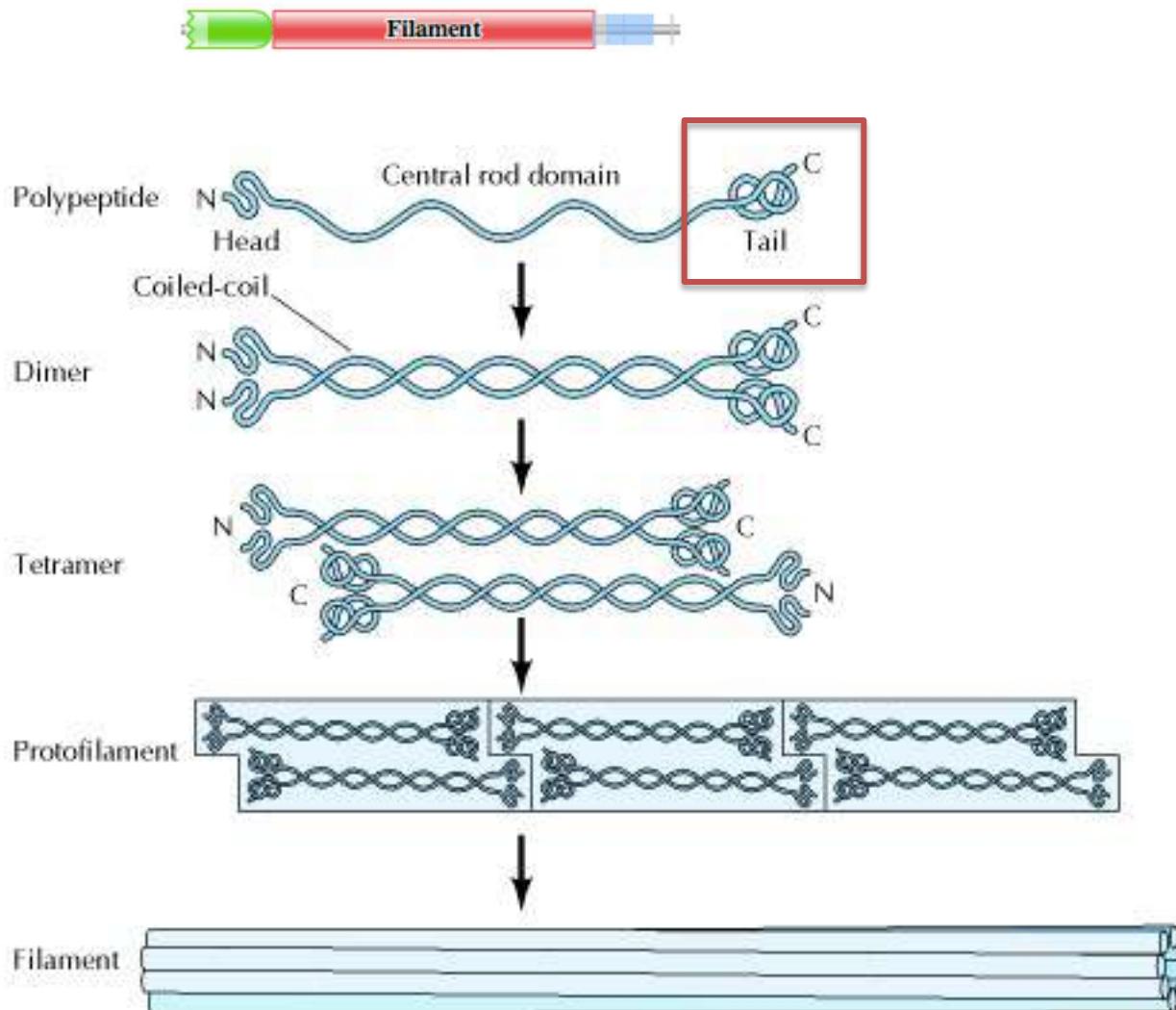
While IF proteins are evolutionary and structurally related, they have limited sequence homologies except in several regions of the rod domain. The IF rod domain is approximately 310 residues long in all cytoplasmic IF proteins and close to 350 residues in the nuclear ones. The IF rod domain exhibits an interrupted alpha-helical conformation and reveals a pronounced seven-residue periodicity in the distribution of apolar residues. The heptad periodicity within the rod domain is interrupted in several places, which generates four consecutive alpha-helical segments: 1A and 1B, which together form the so-called coil 1, and 2A and 2B, which form coil 2. The four alpha-helical segments are interconnected by relatively short, variable linkers L1, L12 and L2 [PUBMED:12596228, PUBMED:22869704].

IF proteins have a very strong tendency to dimerize via the formation of an alpha-helical coiled coil (CC) by their rod domains [PUBMED:22869704].



Pfam domains

This image shows the arrangement of the Pfam d entry. The table below gives the domain boundaries.



Summary**Sequence****Structures****TreeFam****Jump to...**

enter ID/acc

Go

SummaryThis is the summary of UniProt entry [K2C8_HUMAN](#) (P05787).**Description:** Keratin, type II cytoskeletal 8**Source organism:** [Homo sapiens \(Human\)](#) (NCBI taxonomy ID [9606](#))
[View Pfam proteome data](#).**Length:** 483 amino acids**Reference Proteome:** ✓

Please note: when we start each new Pfam data release, we take a copy of the UniProt sequence database. This snapshot of UniProt forms the basis of the overview that you see here. It is important to note that, although some UniProt entries may be removed after a Pfam release, these entries will not be removed from Pfam until the next Pfam data release.

Pfam domains

This image shows the arrangement of the Pfam domains that we found on this sequence. Clicking on a domain will take you to the page describing that Pfam entry. The table below gives the domain boundaries for each of the domains. [More...](#)



[Download](#) the data used to generate the domain graphic in JSON format.

Source	Domain	Start	End
Pfam	Keratin_2_head	6	87
disorder	n/a	18	26
low_complexity	n/a	23	48
disorder	n/a	31	34
disorder	n/a	36	38
low_complexity	n/a	41	61
Pfam	Filament	90	401
coiled_c	n/a	93	113
coiled_c	n/a	144	192
coiled_c	n/a	200	234
disorder	n/a	297	298
disorder	n/a	300	309
coiled_c	n/a	300	380
disorder	n/a	311	312
low_complexity	n/a	393	404
disorder	n/a	402	411
disorder	n/a	413	415
low_complexity	n/a	415	431
low_complexity	n/a	426	459
disorder	n/a	477	479

Central Rod domain

Summary**Sequence****Structures****TreeFam****Jump to...**

enter ID/acc

Go

SummaryThis is the summary of UniProt entry [K2C8_HUMAN](#) (P05787).**Description:** Keratin, type II cytoskeletal 8**Source organism:** [Homo sapiens \(Human\)](#) (NCBI taxonomy ID [9606](#))
[View Pfam proteome data](#).**Length:** 483 amino acids**Reference Proteome:** ✓

Please note: when we start each new Pfam data release, we take a copy of the UniProt sequence database. This snapshot of UniProt forms the basis of the overview that you see here. It is important to note that, although some UniProt entries may be removed after a Pfam release, these entries will not be removed from Pfam until the next Pfam data release.

Pfam domains

This image shows the arrangement of the Pfam domains that we found on this sequence. Clicking on a domain will take you to the page describing that Pfam entry. The table below gives the domain boundaries for each of the domains. [More...](#)



[Download](#) the data used to generate the domain graphic in JSON format.

Source	Domain	Start	End
Pfam	Keratin_2_head	6	87
disorder	n/a	18	26
low_complexity	n/a	23	48
disorder	n/a	31	34
disorder	n/a	36	38
low_complexity	n/a	41	61
disorder	n/a	89	91
Pfam	Filament	90	401
coiled_coil	n/a	95	115
coiled_coil	n/a	144	192
coiled_coil	n/a	200	234
disorder	n/a	297	298
disorder	n/a	300	309
coiled_coil	n/a	300	380
disorder	n/a	311	312
low_complexity	n/a	393	404
disorder	n/a	402	411
disorder	n/a	413	415
low_complexity	n/a	415	431
low_complexity	n/a	426	459
disorder	n/a	477	479

Head domain**Tail domain region**

Family: Keratin_2_head (PF16208)

32 architectures

2224 sequences

0 interactions

119 species

0 structures

Summary**Domain organisation****Clan****Alignments****HMM logo****Trees****Curation & model****Species****Interactions****Structures****Jump to...** ↻

Summary: Keratin type II head

Pfam includes annotations and additional family information from a range of different sources. These sources can be accessed via the tabs below.

This tab holds annotation information from the [InterPro](#) database.

InterPro entry [IPR032444](#)

All intermediate filament proteins feature a central alpha-helical rod domain and variable nonhelical domains located at the N-terminal (head) and C-terminal (tail). The central rod domain is the main driver of self-assembly into filaments, whereas the N- and C-terminal domains are involved in post-translational modifications and interactions with other proteins [[PUBMED:16710422](#)]. Type I and type II keratin form heteropolymeric intermediate filaments providing vital mechanical support in epithelia [[PUBMED:22705788](#)]. This entry represents the N-terminal domain (head) of type II keratins.

Summary**Sequence****Structures****TreeFam****Jump to...**

enter ID/acc

Go

SummaryThis is the summary of UniProt entry [K2C8_HUMAN](#) (P05787).**Description:** Keratin, type II cytoskeletal 8**Source organism:** [Homo sapiens \(Human\)](#) (NCBI taxonomy ID [9606](#))
[View Pfam proteome data](#).**Length:** 483 amino acids**Reference Proteome:** ✓

Please note: when we start each new Pfam data release, we take a copy of the UniProt sequence database. This snapshot of UniProt forms the basis of the overview that you see here. It is important to note that, although some UniProt entries may be removed after a Pfam release, these entries will not be removed from Pfam until the next Pfam data release.

Pfam domains

This image shows the arrangement of the Pfam domains that we found on this sequence. Clicking on a domain will take you to the page describing that Pfam entry. The table below gives the domain boundaries for each of the domains. [More...](#)

[Download](#) the data used to generate the domain graphic in JSON format.

Source	Domain	Start	End
Pfam	Keratin_2_head	6	87
disorder	n/a	18	26
low_complexity	n/a	23	48
disorder	n/a	31	34
disorder	n/a	36	38
low_complexity	n/a	41	61
disorder	n/a	89	91
Pfam	Filament	90	401
coiled_coil	n/a	95	115
coiled_coil	n/a	144	192
coiled_coil	n/a	200	234
disorder	n/a	297	298
disorder	n/a	300	309
coiled_coil	n/a	300	380
disorder	n/a	311	312
low_complexity	n/a	393	404
disorder	n/a	402	411
disorder	n/a	413	415
low_complexity	n/a	415	431
low_complexity	n/a	426	459
disorder	n/a	477	479

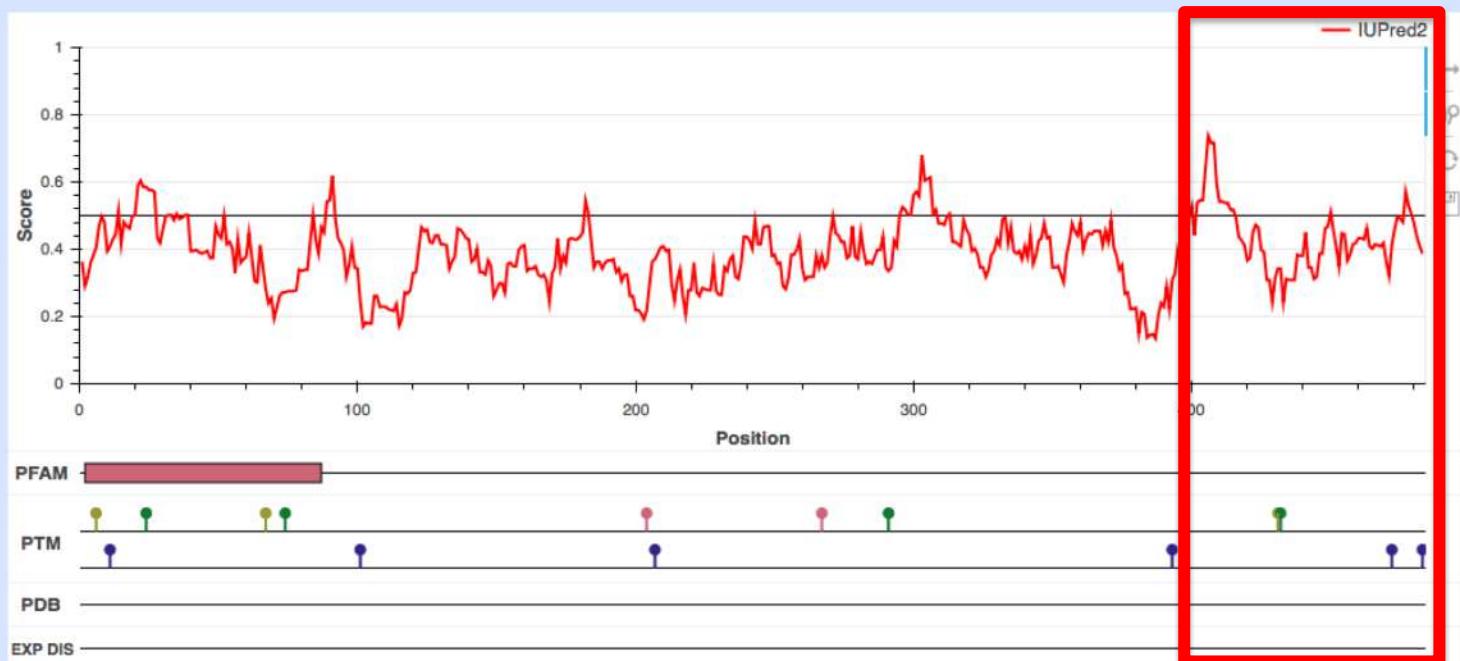
Tail domain region

→ Predictions of disorder and low complexity in some parts of the tail region

Prediction of Intrinsically Unstructured Proteins

[IUPred2A](#)[How to use](#)[Examples](#)[New features](#)[Statistics](#)[Related links](#)[Downloads](#)**NEW**
Just published in Science
Signaling: A review on
Degrons in Cancer

>sp|P05787|K2C8_HUMAN Keratin, type II cytoskeletal 8 OS=Homo sapiens GN=KRT8 PE=1 SV=7

[Download text](#)[Download JSON](#)



Classification of protein families

Home **Search** Browse Results Release notes Download Help About

Search InterPro

by sequence

by text

by domain architecture

Search families, domains, proteins, keywords or GO terms

P05787

e.g. IPR020422, kinase, O00167, PF02932, GO:0007165, 1t2v, UP000005640

Powered by EBI search

Search

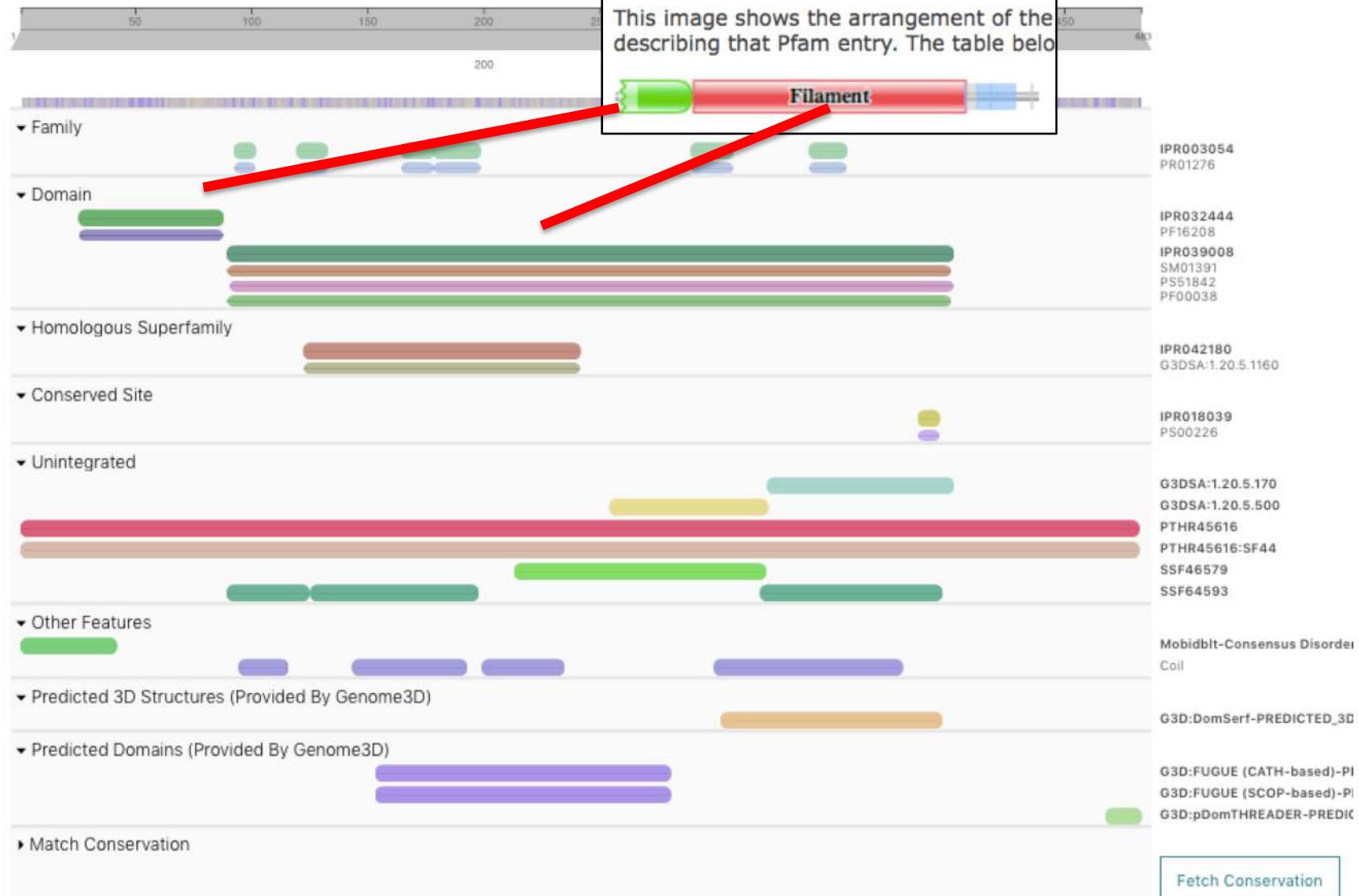
Clear

Found an exact match: [protein P05787](#)

Protein family membership

Keratin, type II (IPR003054)

Entry matches to this protein



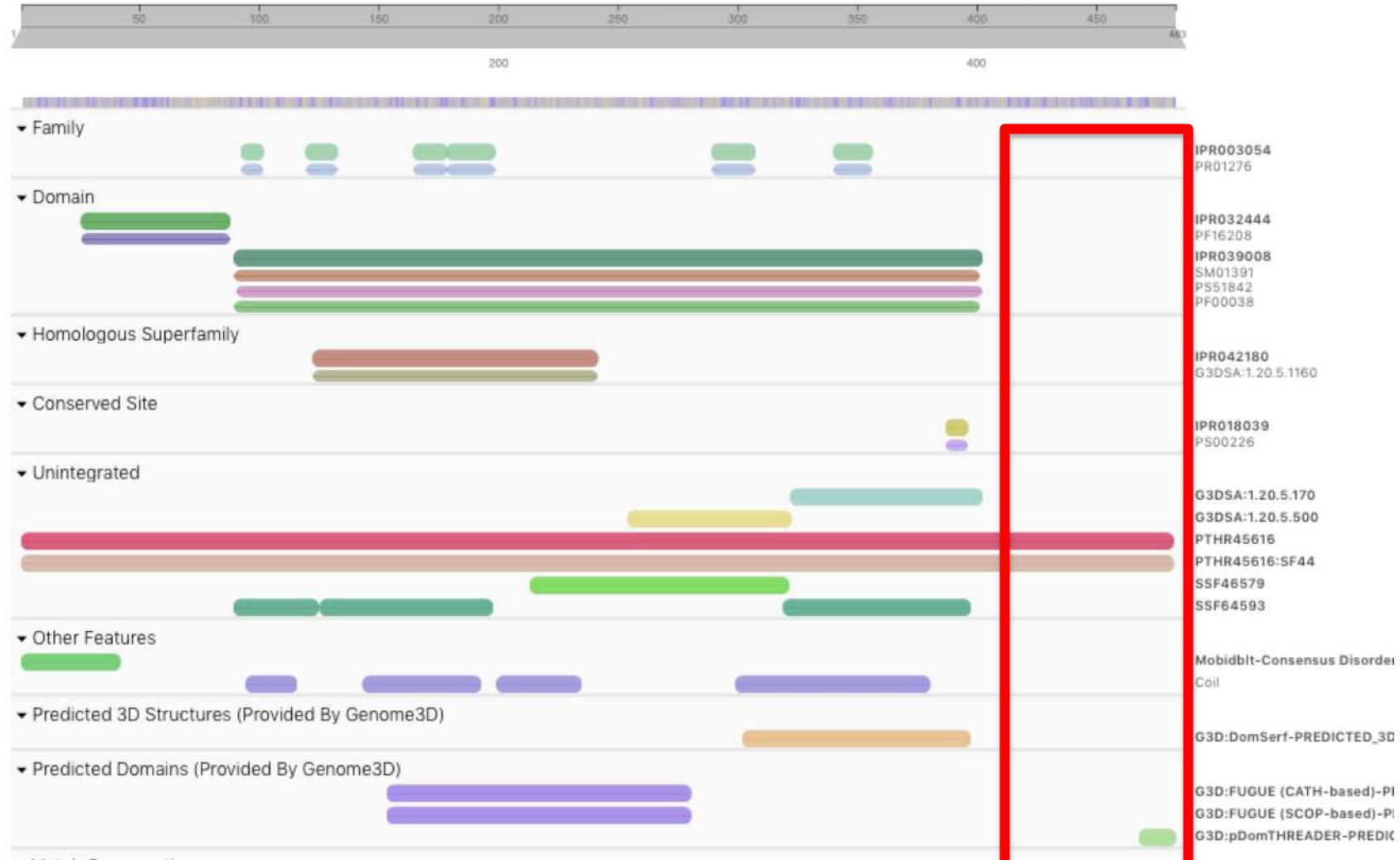
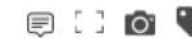
Protein family membership

F Keratin, type II (IPR003054)

Entry matches to this protein

Colour By: Accession

Collapse All



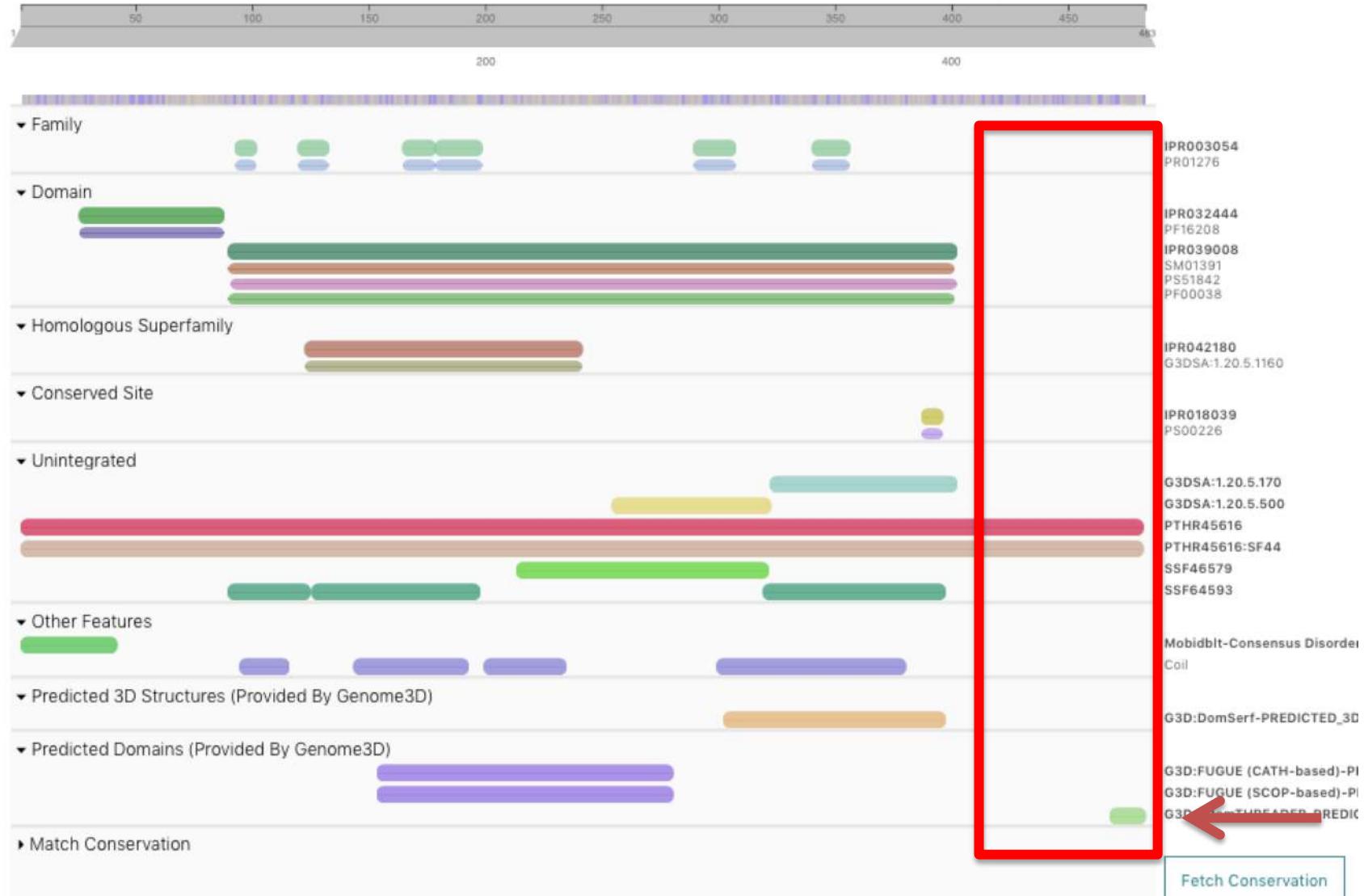
Protein family membership

F Keratin, type II (IPR003054)

Entry matches to this protein

Colour By: Accession

Collapse All



- We are going to look for homologs of KRT8's tail region and see if we can learn more about its function

>sp|P05787|K2C8_HUMAN Keratin, type II cytoskeletal 8 OS=Homo sapiens OX=9606 GN=KRT8 PE=1 SV=7
MSIRVTQKS YKVSTSGPRAFSSRSY TSGPGSRI SSSFSRVGSSNFRGGLGGGYGGASGM
GGITAVTVNQSLLSPLVLEVDPNIQAVRTQEKEQIKTLNNKFASFIDKVRFLEQQNKMLE
TKWSLLQQQKTARSNMDNMFESYINNLRRQLETLGQEKLKLEAELGNMQGLVEDFKNKYE
DEINKRTEMENEFVLIKKDVDEAYMNKVELESRLEGLTDEINFLRQLYEEEIRELQSQIS
DTSVVLSMDNSRSLDMDSIIAEVKAQYEDIANRSRAEAESMYQIKYEELQSLAGKHGDDL
RRTKTEISEMNRNISRLQAEIEGLKGQRASLEAAIADAEQRGE LAIKDANAKLSELEAAL
QRAKQDMARQLREYQELMVKLALDIEIATYRK LLEGEESRLES GMQNMSIHTKTTSGYA
GGLSSAYGGLTSPGLSY SLGSSFGSGAGSSFSRTSSRAVVVKIETRDGKLVSESSDV
LPK

>sp|P05787|K2C8_HUMAN Keratin, type II cytoskeletal 8 OS=Homo sapiens OX=9606 GN=KRT8 PE=1 SV=7
MSIRVTQKS YKVSTSGPRAFSSRSY TSGPGSRI SSSFSRVGSSNFRGGLGGGYGGASGM
GGITAVTVNQSLLSPLVLEVD PNIQAVRTQEKEQIKTLNNKFASFIDKVRFLEQQNKMLE
TKWSLLQQQKTARSNMDNMFESYINNLRRQLETLGQEKLKLEAELGNMQGLVEDFKNKYE
DEINKRTEMENE FVLIKKDVDEAYMNKVELESRLEGLTDEINFLRQLYEEEIRELQSQIS
DTSVVLSMDNSRSLDMD SIIAEVKAQYEDIANRSRAEAESMYQIKYEELQSLAGKHGDDL
RRTKTEISEMNRNISRLQAEIEGLKGQRASLEAAIAADEQRGE LAIKDANAKLSELEAAL
QRAKQDMARQLREYQELMNVKLALDIEIATYRK LLEGEES [REDACTED] IHTKTTSGYA
GGLSSAYGGLTSPGLSY SLGSSFGSGAGSSFSRTSSRAVVVKIETRDGKL VSESSDV
LPK

>sp|P05787|K2C8_HUMAN Keratin, type II cytoskeletal 8 OS=Homo sapiens OX=9606 GN=KRT8 PE=1 SV=7
MSIRVTQKS YKVSTSGPRAFSSRSY TSGPGSRI SSSFSRVGSSNFRGGLGGGYGGASGM
GGITAVTVNQS LLSPLVLEVDPNIQAVRTQEKEQIKTLNNKFASFIDKVRFL EQQNKMLE
TKWSLLQQQKTARSNMDNMFESYINNLRRQLETLGQEKLKLEAELGNMQGLVEDFKNKYE
DEINKRTEMENE FVLIKKDVDEAYMNKVELESRLEGLTDEINFLRQLYEEEIRELQSQIS
DTSVVLSMDNSRS LDMD SIIAEVKAQYEDIANRSRAEAESMYQIKYEELQSLAGKHGDDL
RRTKTEI SEMNRNISRLQAEIEGLKGQRASLEAAIAADEQRGE LAIKDANAKLSELEAAL
QRAKQDMARQLREYQELMNVKLALDIEIATYRK LLEGEES [REDACTED] IHTKTTSGYA
GGLSSAYGGLTSPGLSY SLGSSFGSGAGSSSF SRTSSRAVVVKIETRDGKL VSESSDV
LPK

KRT8-tail-long.fasta

Let's visit the HMMER website:

<https://www.ebi.ac.uk/Tools/hmmer/search>



HMMER

Biosequence analysis using profile hidden Markov Models

[Home](#) [Search](#) [Results](#) [Software](#) [Help](#) [About](#) [Contact](#)[phmm](#)[hmmscan](#)[hmmsearch](#)[jackhmmer](#)

protein sequence vs protein sequence database

[Paste a Sequence](#) | [Upload a File](#) | [Accession Search](#)

Paste in your sequence or use the [example](#)

[Submit](#) [Reset](#)

▼ [Sequence Database](#)

Frequently used databases: [Reference Proteomes](#) [UniProtKB](#) [SwissProt](#) [PDB](#) [Ensembl](#)

Current database selection:

[Reference Proteomes](#) ▾

► [Restrict by Taxonomy](#)

▼ [Cut-Offs](#)

E-value

Bit score

<https://www.ebi.ac.uk/Tools/hmmer/search>



HMMER

Biosequence analysis using profile hidden Markov Models

[Home](#) [Search](#) [Results](#) [Software](#) [Help](#) [About](#) [Contact](#)[phmmmer](#) [hmmscan](#) [hmmsearch](#) [jackhmmer](#)

protein sequence vs protein sequence database

[Paste a Sequence](#) | [Upload a File](#) | [Accession Search](#)Paste in your sequence or use the [example](#) **Click**[Submit](#) [Reset](#)**▼ Sequence Database** Frequently used databases: [Reference Proteomes](#) [UniProtKB](#) [SwissProt](#) [PDB](#) [Ensembl](#)

Current database selection:

[Reference Proteomes](#) **► Restrict by Taxonomy** **▼ Cut-Offs** E-value Bit score<https://www.ebi.ac.uk/Tools/hmmer/search>



HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

phmmmer hmmscan hmmsearch

jackhmmer

iterative search vs protein sequence database

Paste a Sequence or an Alignment | [Upload a File](#) | [Accession Search](#)

Paste in your sequence ([example](#)), HMM ([example](#)) or multiple sequence alignment ([example](#))

▼ Sequence Database

Frequently used databases: [Reference Proteomes](#) [UniProtKB](#) [SwissProt](#) [PDB](#) [Ensembl](#)

Current database selection:

Reference Proteomes

► Restrict by Taxonomy

▼ Cut-Offs

<https://www.ebi.ac.uk/Tools/hmmer/search>



HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

phmmmer hmmscan hmmsearch

jackhmmer

iterative search vs protein sequence database

Paste a Sequence or an Alignment | [Upload a File](#) | [Accession Search](#)

Paste in your sequence ([example](#)), HMM ([example](#)) or multiple sequence alignment ([example](#))

Click

Submit

Reset

▼ Sequence Database

KRT8-tail-long.fasta

Frequently used databases:

[Reference Proteomes](#)

[UniProtKB](#)

[SwissProt](#)

Current database selection:

Reference Proteomes

► Restrict by Taxonomy

▼ Cut-Offs

<https://www.ebi.ac.uk/Tools/hmmer/search>



HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

phmmer hmmscan hmmsearch

jackhmmer

Iterative search vs protein sequence database

Paste a Sequence or an Alignment | [Upload a File](#) | [Accession Search](#)

Paste in your sequence ([example](#)), HMM ([example](#)) or multiple sequence alignment ([example](#))



or Copy and Paste Sequence

Submit

Reset

▼ Sequence Database

Frequently used databases: [Reference Proteomes](#) [UniProtKB](#) [SwissProt](#) [PDB](#) [Ensembl](#)

Current database selection:

Reference Proteomes

► Restrict by Taxonomy

▼ Cut-Offs

<https://www.ebi.ac.uk/Tools/hmmer/search>



HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

phmmmer hmmscan hmmsearch jackhmmer

iterative search vs protein sequence database

Paste a Sequence or an Alignment | Upload a File | Accession Search

Paste in your sequence ([example](#)), HMM ([example](#)) or multiple sequence alignment ([example](#)) ⓘ

```
>KRT8_tail_long
IHTKTTSGYAGGLSSAYGGLTSPGLSYSLGSSFGSGAGSSSFSRTSSRAVVKKIETRDGKLVSESSDVLPK
```

Submit **Reset**

▼ Sequence Database ⓘ

Frequently used databases: [Reference Proteomes](#) [UniProtKB](#) [SwissProt](#) [PDB](#) [Ensembl](#)

Current database selection:

Reference Proteomes

► Restrict by Taxonomy ⓘ

▼ Cut-Offs ⓘ

<https://www.ebi.ac.uk/Tools/hmmer/search>

▼ Sequence Database

Frequently used databases:

[Reference Proteomes](#)[UniProtKB](#)[SwissProt](#)[PDB](#)[Ensembl](#)

Current database selection:

[Reference Proteomes](#)

► Restrict by Taxonomy

▼ Cut-Offs

 E-value Bit score

Significance E-values:

Sequence

Hit

Report E-values:

Sequence

Hit



HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

phmmmer hmmscan hmmsearch

jackhmmer

iterative search vs protein sequence database

Paste a Sequence or an Alignment | Upload a File | Accession Search

Paste in your sequence ([example](#)), HMM ([example](#)) or multiple sequence alignment ([example](#)) ⓘ

```
>KRT8_tail_long
IHTKTTSGYAGGLSSAYGGLTSPGLSYSLGSSFGSGAGSSSFSRTSSRAVVKKIETRDGKLVSESSDVLPK
```

Submit

Click Submit

▼ Sequence Database ⓘ

Frequently used databases: [Reference Proteomes](#) [UniProtKB](#) [SwissProt](#) [PDB](#) [Ensembl](#)

Current database selection:

Reference Proteomes

► Restrict by Taxonomy ⓘ

▼ Cut-Offs ⓘ

<https://www.ebi.ac.uk/Tools/hmmer/search>

Score Taxonomy Domain Download

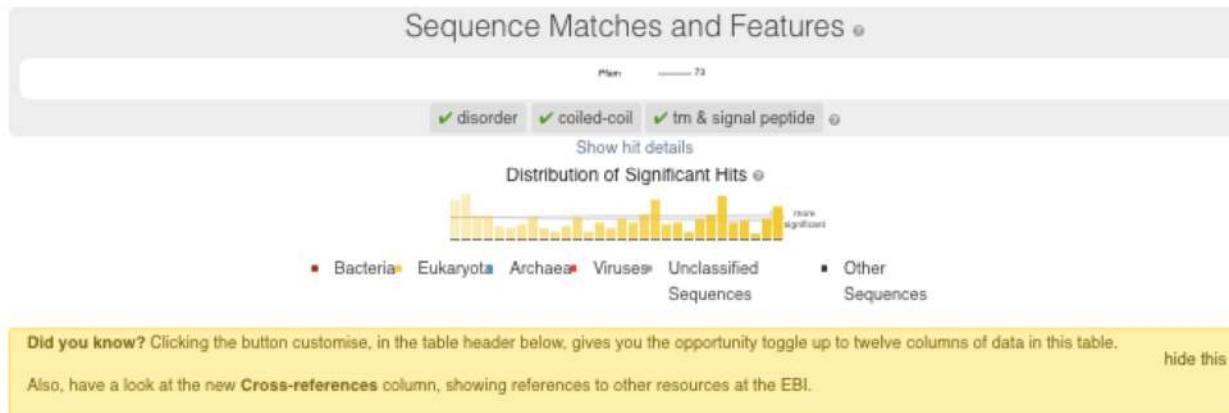
Iteration 1

[Return to the Results Summary](#) | [Jump to threshold](#)

Sequence selection

- above threshold
- unselect all

[Continue the search](#) **Start iteration 2**



« First « Previous Page 1 of 7 Next » Last »

Significant Query Matches (301) in *uniprotrefprot* (v.2019_09)

Customise

Target	Description	Species	Cross-references	E-value
> K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens		1.6e-35
> K2C8_HUMAN	Isoform 2 of Keratin, type II cytoskeletal 8	Homo sapiens		1.7e-35
> A0A096NQU2_PAPAN	IF rod domain-containing protein	Papio anubis		3.2e-35
> G1S7N4_NOMLE	IF rod domain-containing protein	Nomascus leucogenys		4.5e-35
> A0A2K5LZ8_CERAT	IF rod domain-containing protein	Cercopithecus alys		5.9e-35
> A0A2K6NVV1_RHIRO	IF rod domain-containing protein	Rhinopithecus roxellana		6.2e-35
> F5H0A2_MACMU	Keratin, type II cytoskeletal 8	Moschus moschiferus		6.2e-35

Score Taxonomy Domain Download

Iteration 1

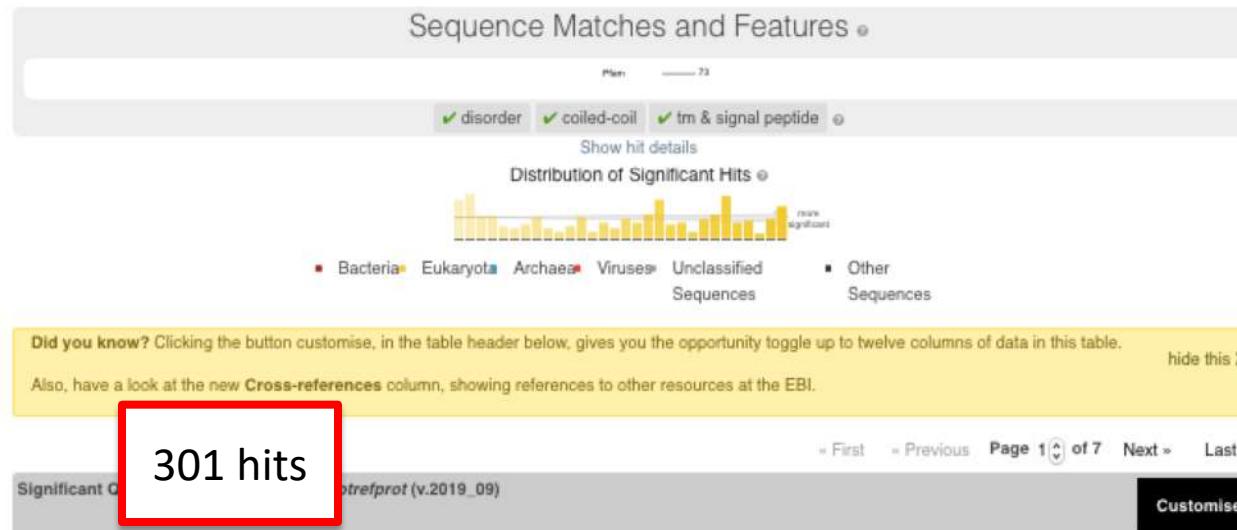
Return to the Results Summary | Jump to threshold

Sequence selection

above threshold

unselect all

Continue the search Start iteration 2



Target	Description	Species	Cross-references	E-value
> K2C8_HUMAN#	Keratin, type II cytoskeletal 8	Homo sapiens#		1.6e-35
> K2C8_HUMAN#	Isoform 2 of Keratin, type II cytoskeletal 8	Homo sapiens#		1.7e-35
> A0A096NQU2_PAPAN#	IF rod domain-containing protein	Papio anubis#		3.2e-35
> G1S7N4_NOMLE#	IF rod domain-containing protein	Nomascus leucogenys#		4.5e-35
> A0A2K5L5Z8_CERAT#	IF rod domain-containing protein	Cercopithecus alys#		5.9e-35
> A0A2K6NVV1_RHIRO#	IF rod domain-containing protein	Rhinopithecus roxellana#		6.2e-35
> ESU0A2_MACMU#	Keratin, type II cytoskeletal 8	Macaca mulatta#		6.2e-35

Score Taxonomy Domain Download

Iteration 1

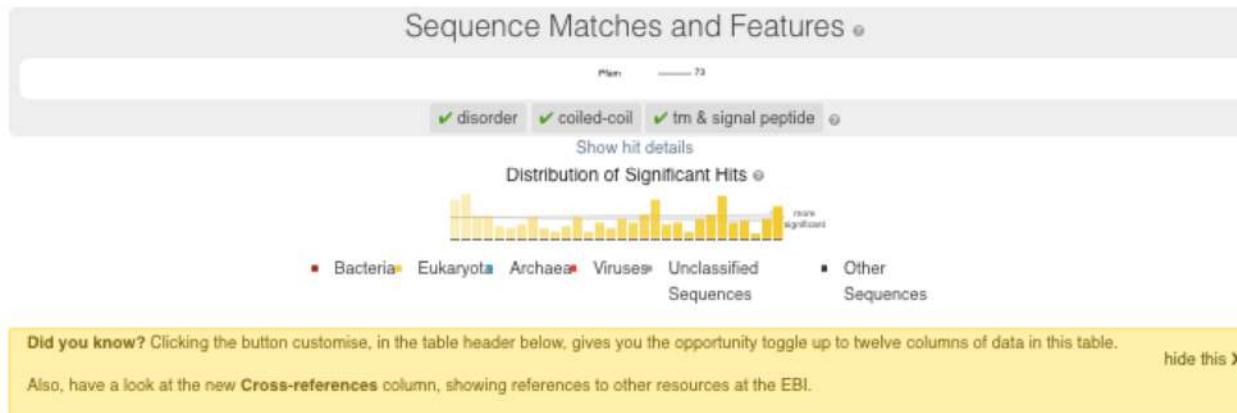
Return to the Results Summary | Jump to threshold

Sequence selection

above threshold

unselect all

Continue the search Start iteration 2



KRT8 is
our top
hit



Significant Query Matches (301) in *uniprotrefprot* (v.2019_09)

Customise

Target	Description	Species	Cross-references	E-value
> K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens		1.6e-35
> K2C8_HUMAN	Isoform 2 of Keratin, type II cytoskeletal 8	Homo sapiens		1.7e-35
> A0A096NQU2_PAPAN	IF rod domain-containing protein	Papio anubis		3.2e-35
> G1S7N4_NOMLE	IF rod domain-containing protein	Nomascus leucogenys		4.5e-35
> A0A2K5L5Z8_CERAT	IF rod domain-containing protein	Cercopithecus alys		5.9e-35
> A0A2K6NVV1_RHIRO	IF rod domain-containing protein	Rhinopithecus roxellana		6.2e-35
> F5H0A2_MACMU	Keratin, type II cytoskeletal 8	Moschus moschiferus		6.2e-35

Click on Taxonomy



Score Taxonomy Iteration 1

Return to the Results Summary | Jump to threshold

Sequence selection

- above threshold
- unselect all

Continue the search Start iteration 2

Sequence Matches and Features

Plots: 73

disorder coiled-coil tm & signal peptide

Show hit details

Distribution of Significant Hits



more significant

- Bacteria
- Eukaryota
- Archaea
- Viruses
- Unclassified Sequences
- Other Sequences

Did you know? Clicking the button customise, in the table header below, gives you the opportunity toggle up to twelve columns of data in this table. [hide this X](#)

Also, have a look at the new Cross-references column, showing references to other resources at the EBI.

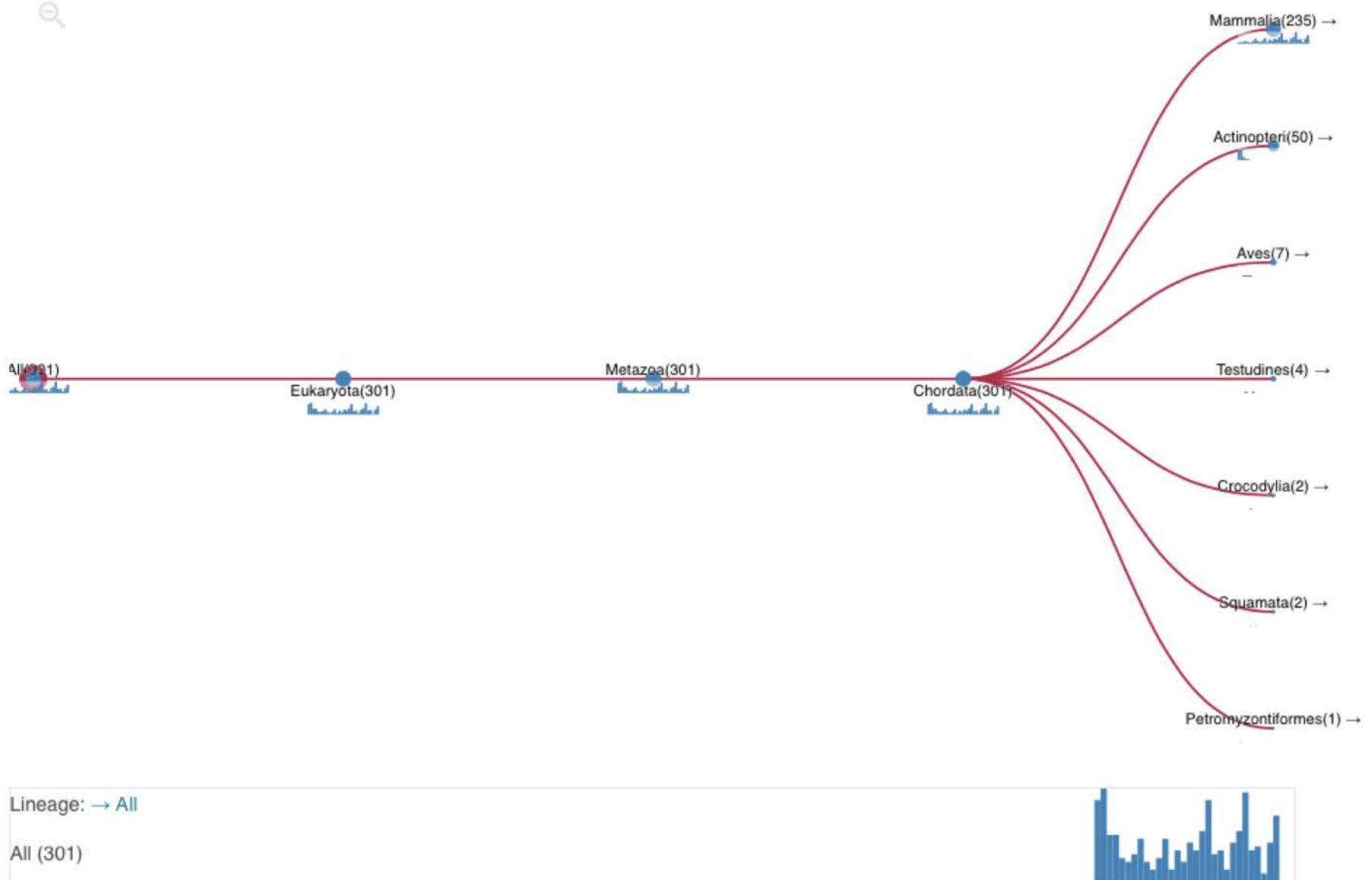
« First « Previous Page 1 of 7 Next » Last »

Significant Query Matches (301) in *uniprotrefprot* (v.2019_09)

Customise

Target	Description	Species	Cross-references	E-value
> K2C8_HUMAN#	Keratin, type II cytoskeletal 8	Homo sapiens#		1.6e-35
> K2C8_HUMAN#	Isoform 2 of Keratin, type II cytoskeletal 8	Homo sapiens#		1.7e-35
> A0A096NQU2_PAPAN#	IF rod domain-containing protein	Papio anubis#		3.2e-35
> G1S7N4_NOMLE#	IF rod domain-containing protein	Nomascus leucogenys#		4.5e-35
> A0A2K5LZ8_CERAT#	IF rod domain-containing protein	Cercopithecus alys#		5.9e-35
> A0A2K6NVV1_RHIRO#	IF rod domain-containing protein	Rhinopithecus roxellana#		6.2e-35
> F5U0A2_MACMU#	Keratin, type II cytoskeletal 8	Macaca mulatta#		6.2e-35

Taxonomic distribution of all search hits



Click on Domain

←

Score Taxonomy Domain

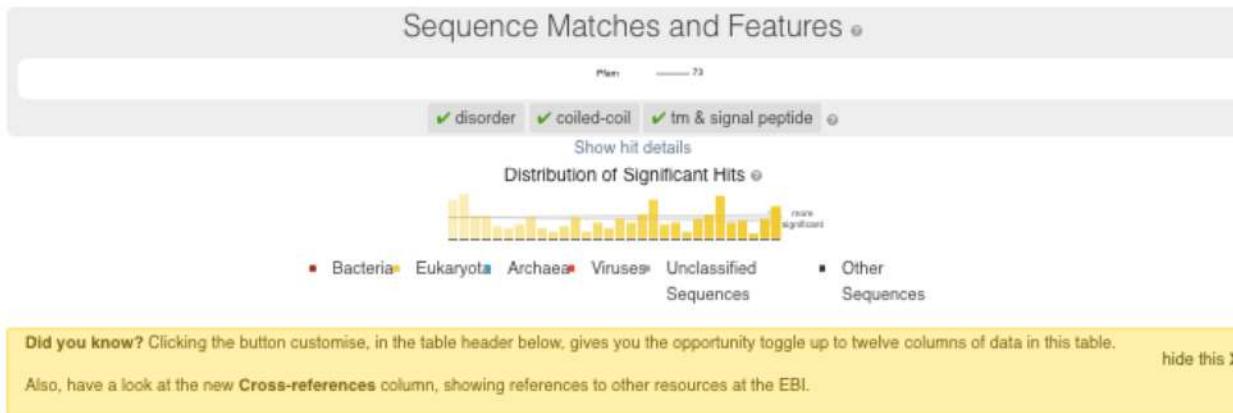
Iteration 1

Return to the Results Summary | Jump to threshold

Sequence selection

- above threshold
- unselect all

Continue the search Start iteration 2



« First « Previous Page 1 of 7 Next » Last »

Significant Query Matches (301) in *uniprotrefprot* (v.2019_09)

Customise

Target	Description	Species	Cross-references	E-value
> K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens		1.6e-35
> K2C8_HUMAN	Isoform 2 of Keratin, type II cytoskeletal 8	Homo sapiens		1.7e-35
> A0A096NQU2_PAPAN	IF rod domain-containing protein	Papio anubis		3.2e-35
> G1S7N4_NOMLE	IF rod domain-containing protein	Nomascus leucogenys		4.5e-35
> A0A2K5LZ8_CERAT	IF rod domain-containing protein	Cercopithecus alys		5.9e-35
> A0A2K6NVV1_RHIRO	IF rod domain-containing protein	Rhinopithecus roxellana		6.2e-35
> F8U0A2_MACMU	Keratin, type II cytoskeletal 8	Macaca mulatta		6.2e-35

Domain Architectures

199

SEQUENCES

with domain architecture: **Keratin_2_head, Filament**, example: K2C8_HUMAN 

[View Scores](#)



[Show All](#)

52

SEQUENCES

with domain architecture: **Filament**, example: AOA096NQU2_PAPAN 

[View Scores](#)



[Show All](#)

25

SEQUENCES

with domain architecture: **Keratin_2_head, Filament, Filament**, example: F7GPG9_MACMU 

[View Scores](#)



[Show All](#)

10

SEQUENCES

with domain architecture: **Pkinase, Keratin_2_head, Filament, Filament**, example: G3REJ4_GORGO 

[View Scores](#)



[Show All](#)

9

SEQUENCES

with domain architecture: **Filament, Filament**, example: AOA2K5M2T8_CERAT 

[View Scores](#)



[Show All](#)

3

SEQUENCES

with no domain architecture, example: AOA402E9P6_9SAUR 

[View Scores](#)



[Show All](#)

2

SEQUENCES

with domain architecture: **Keratin_2_head, Filament, Filament, Filament**, example: A0A2I3RRX0_PANTR 

[View Scores](#)



[Show All](#)

1

SEQUENCE

with domain architecture: **Keratin_2_head**, example: H0XMA4_OTOGA 

[View Scores](#)



Domain Architectures ?

199
SEQUENCES

with domain architecture: Keratin_2_head, Filament, example: K2C_HUMAN

Show All

52
SEQUENCES

with domain architecture: Filament, example: AOA096NQU2_PAPAN

Show All

25
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, example: F7GPG9_MACMU

Show All

10
SEQUENCES

with domain architecture: Pkinase, Keratin_2_head, Filament, Filament, example: G3REJ4_GORGO

Show All

9
SEQUENCES

with domain architecture: Filament, Filament, example: A0A2K5M2T8_CERAT

Show All

3
SEQUENCES

with no domain architecture, example: A0A402E9P6_9SAUR

Show All

2
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, Filament, example: A0A2I3RRX0_PANTR

Show All

1
SEQUENCE

with domain architecture: Keratin_2_head, example: H0XMA4_OTOGA

Show All

Pfam domains

This image shows the arrangement of the describing that Pfam entry. The table below



[View Scores](#)

Domain Architectures

199
SEQUENCES

with domain architecture: Keratin_2_head, Filament, example: K2C8_HUMAN

Show All

52
SEQUENCES

with domain architecture: Filament, example: AOA096NQU2_PAPAN

Show All

25
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, example: F7GPG9_MACMU

Show All

10
SEQUENCES

with domain architecture: Pkinase, Keratin_2_head, Filament, Filament, example: G3REJ4_GORGO

Show All

9
SEQUENCES

with domain architecture: Filament, Filament, example: A0A2K5M2T8_CERAT

Show All

3
SEQUENCES

with no domain architecture, example: A0A402E9P6_9SAUR

Show All

2
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, Filament, example: A0A2I3RRX0_PANTR

Show All

1
SEQUENCE

with domain architecture: Keratin_2_head, example: H0XMA4_OTOGA

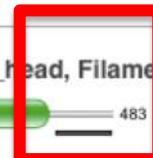
Show All

Pfam domains

This image shows the arrangement of the describing that Pfam entry. The table below



[View Scores](#)



Domain Architectures

199
SEQUENCES

with domain architecture: Keratin_2_head, Filament, example: K2C8_HUMAN

Show All

52
SEQUENCES

with domain architecture: Filament, example: AOA096NQU2_PAPAN

Show All

25
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, example: F7GPG9_MACMU

Show All

10
SEQUENCES

with domain architecture: Pkinase, Keratin_2_head, Filament, Filament, example: G3REJ4_GORGO

Show All

9
SEQUENCES

with domain architecture: Filament, Filament, example: AOA2K5M2T8_CERAT

Show All

3
SEQUENCES

with no domain architecture, example: AOA402E9P6_9SAUR

Show All

2
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, Filament, example: AOA2I3RRX0_PANTR

Show All

1
SEQUENCE

with domain architecture: Keratin_2_head, example: H0XMA4_OTOGA

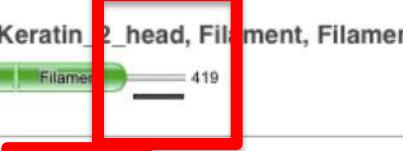
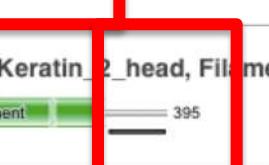
Show All

Pfam domains

This image shows the arrangement of the describing that Pfam entry. The table below



[View Scores](#)



[View Scores](#)

Domain Architectures

199
SEQUENCES

with domain architecture: Keratin_2_head, Filament, example: K2C8_HUMAN

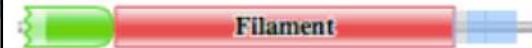
Sequence Features  483

Show All

[View Scores](#)

Pfam domains

This image shows the arrangement of the domains described in the Pfam entry. The table below lists the domain architectures found in the sequences.



52
SEQUENCES

with domain architecture: Filament, example: AOA096NQU2_PAPAN

Sequence Features  266

[View Scores](#)

Show All

25
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, example: F7GPG9_MACMU

Sequence Features  395

[View Scores](#)

Show All

10
SEQUENCES

with domain architecture: Pkinase, Keratin_2_head, Filament, Filament, example: G3REJ4_GORGO

Sequence Features  1649

[View Scores](#)

Show All

9
SEQUENCES

with domain architecture: Filament, Filament, example: AOA2K5M2T8_CERAT

Sequence Features  258

[View Scores](#)

Show All

3
SEQUENCES

with no domain architecture, example: AOA402E9P6_9SAUR

Sequence Features  90

[View Scores](#)

Show All

2
SEQUENCES

with domain architecture: Keratin_2_head, Filament, Filament, Filament, example: A0A2I3RRX0_PANTR

Sequence Features  419

[View Scores](#)

Show All

1
SEQUENCE

with domain architecture: Keratin_2_head, example: H0XMA4_OTOGA

Sequence Features  347

[View Scores](#)

Score

**Click on Score**

Iteration 1

[Return to the Results Summary](#) | [Jump to threshold](#)

Sequence Matches and Features

Pfam — 73

 disorder coiled-coil tm & signal peptide [?](#)[Show hit details](#)[Jump to the exact match for your query architecture](#)

Domain Architectures

**199
SEQUENCES**with domain architecture: **Keratin_2_head, Filament**, example: [K2C8_HUMAN](#)[View Scores](#)Sequence Features [Show All](#)**52
SEQUENCES**with domain architecture: **Filament**, example: [A0A096NQU2_PAPAN](#)[View Scores](#)Sequence Features [Show All](#)**25
SEQUENCES**with domain architecture: **Keratin_2_head, Filament, Filament**, example: [F7GPG9_MACMU](#)[View Scores](#)Sequence Features [Show All](#)



HMMER

Biosequence analysis using profile hidden Markov Models

[Home](#) | [Search](#) | **Results** | [Software](#) | [Help](#) | [About](#) | [Contact](#)

Next release within a week, think about downloading your results

JACKHMMER Results

[Search Again](#)

Score

Taxonomy

Domain

[Download](#)

Iteration 1

[Return to the Results Summary](#) | [Jump to threshold](#)

Sequence selection

above threshold

unselect all

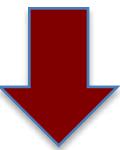
[Continue the search](#)

[Start iteration 2](#)

Distribution of Significant Hits



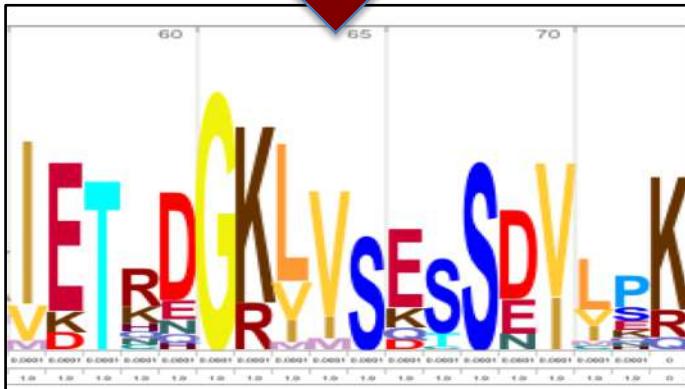
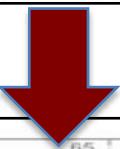
>KRT8_tail_long
IHTKTTSGYAGGLSSAYGGLTSPGLSYSLGSSFGSGA
GSSSFSRTSSRAVVVKKIETRDGKLVSESSDVLPK



We run a sequence to sequence search

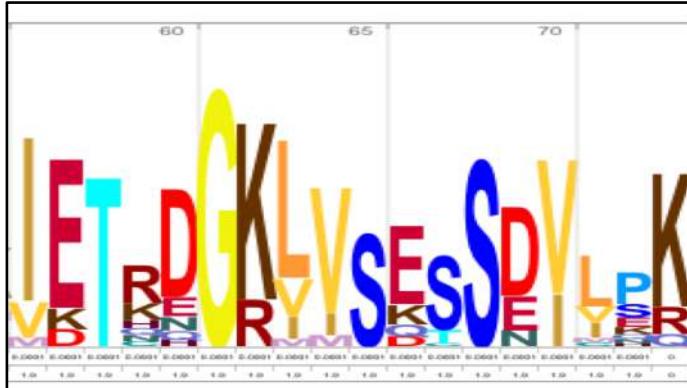
	10	20	30
AOA2K5IDQ5_COLAP/435-463	T S S T R A M V V I E I E T R D G K - - -	L M S E S S D V L P K	
G31MB9_CRIGR/460-498	Y S R T K A V V V V K K I E T R D G K - - -	L V S E S S D I L P K	
AOA2BHC78_NOMLE/392-420	T S S T R A M V V V K K I E T R D G K - - -	L C G H L G N L L P K	
AOA2DGSIN8_ICTPU/473-499	T Q S K K N I V I V I M I E T K D G R - - -	V V S E S S E I V - -	
AOA1UTSF53_ALIME/452-479	V T K S K A I V I V I K I E T R D G K - - -	L V S E S S D I L S -	
AOA087X5N0_POEFO/465-493	T T K K M N V I V I V T I E T N D G K - - -	L V S E S S S N I I E K	
AOA2B6XSR9_CAVPC/225-253	T R T K S A V V V V K I E T R D G K - - -	L V S E T S D V L P K	
HOXA11_OTOGA/235-263	S H T S A R A K V T K I E T H D R K - - -	L V S E S S D I L P K	
G1MBMMS_ALIME/375-401	V S S S X A L A V N T I E T C S G K - - -	L V S K P S T L L - -	
M7BMIMI_CHEMY/444-471	L A R S K A I V I V K I E T K D G R - - -	L V S E T S N V L S -	
G10220_MYOLU/360-385	T S S S X T V V V V E I E T L Q G K - - -	L V S E F S D V L - -	
G580G7_MET6A/184-212	L T H T K S A M V Y I V I E T H D G K - - -	V V S E T S D V L H K	
AOA1LIRW09_CHICK/24-47	T H S K R P T L L I V I V T V E T R D G O - - -	V I N E T S - - -	
F62XN6_COIN/482-507	T V T T K K V V V V K I E T K D G K - - -	V V S Q T E D V - -	
G1LPV9_ALIME/434-462	G S S F X T V V V V V K I K T C A G R - - -	L V S E S P A V L S K	
AOA153QMP6_SALSA/221-244	T T M S X K V L V I V T I E T R D G O - - -	V I N E S M - - -	
HOXMAA_OTOGA/319-347	P A P Q L A I V E K V I I E T H H G K - - -	L T A Q T S D V L S K	
H8BE29_LATCH/486-512	P S K R N L V V V V V K I E A Q Q G K - - -	I V S E T S N Y L - -	
OTZXU4_XENLA/478-506	K T S K P S I L V V V V K I E T K D G R - - -	V L S E S S D V F S K	
V8NZIM9_OPHMA/388-413	S S L K R P L M I N T I E T K D G K - - -	I I S E A S H D V F S K	
AOA1U7T2W5_TARSY/126-153	T G F S R A Y - V E K I I E T R H G K - - -	L V S E A S H A L P K	
G1L6P0_ALIME/439-466	A S S N O D A M A A K I K I G T C H G K - - -	L A S E S Y D V L S -	
H8BAT2_LATCH/492-456	V H T K K T V V M I K T I E T I D G K E E K V N E - - -		
G1M206_ALIME/450-475	V S F S K A A V A K R M E T H D G K - - -	L M L E S F V V - -	
M8XA98_FELCA/439-466	G R G S T S L S C K - I E I R D G K - - -	L V S E S S D V L S K	

We find close homologs of our protein



From the MSA we build a family profile

From the MSA we build
a family profile

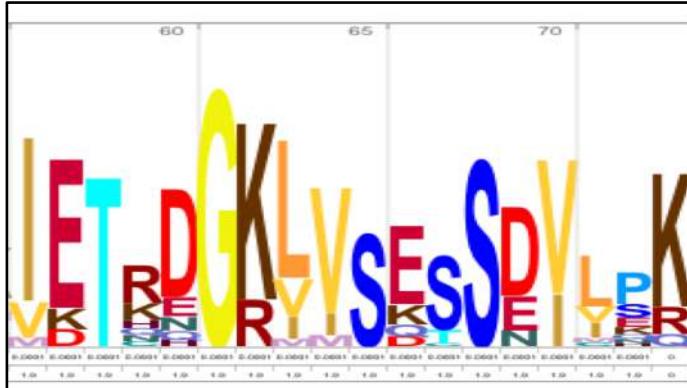


We run a profile
to sequence search

	10	20	30
ADA2K5J_DQ6_COLAP/495-493	T S S T R A M V V K E I E T R D G K --	L M S E S S D V L P K	
G3I195_CIGR/460-488	Y S R T K A V V V K K I E T R D G K --	L V S E S S D I L L P K	
F6YKZ7_HORSE/457-485	I G P P S A I V V K K I E T R D G K --	L V S E S S D I L L P K	
ADA2BHC78_NOMLE/393-420	T S S T R A M V V K K I E T R D G K --	L C G H L G N L L P K	
ADA2D0SIM8_ICTPU/473-499	T Q S K K N I V I K M I E T K D G R --	V V S E S S E I V --	
ADA1U7SF33_ALLS/452-479	V T K S K A I V I K K I E T R D G K --	L V S E S S D I L S -	
ADA087XSN0_POEFO/465-493	T T T K K N V I V K T I E T N D G K --	L V S E S S S N I I E K	
ADA286XSR9_CAVPO/225-253	T R T K S A V V V K K I E T R D G K --	L V S E T S D V L P K	
H2L7B8_ORYLA/463-489	T Q S K K S V V V I K T I E T K D G R --	V V S E S S E L I --	
HDXIAL_OTOGA/235-263	S H T S A R A K V T K I E T H D R K --	L V S E S S S D I L P K	
G1M6M5_AILME/375-401	V S S S K A L A V N T I E T C S G K --	L V S K P S T L L --	
M7BMIM1_CHEMY/444-471	L A R S K A I V I K K I E T K D G R --	L V S E T S N V L S -	
K7FDH1_PELS/460-487	V S K S K T V V I K K I E T R D G R --	L V S E T S D I L S -	
G1Q220_MYOLU/360-385	T S S S K T V V V V K E I E T L Q G K --	L V S E F S D V L --	
G5BQG7_HETGA/184-212	L T H T K S A M V V K I E T H D G K --	V V S E T S D V L H K	
W5UL23_ICTPU/446-470	I H S K K T V L I K T I E T R D G E -	S K V V S E S -	
B3JR8_ORENI/488-514	T Q T K K N V V I K M I E T K D G K --	V V S E S S E V I --	
H2RK76_TAKRU/432-455	M H S K K T V L I K T I E T R D G E --	I I T R T S -	
ADA1L1RW09_CHICK/24-47	T H S K R T L L I K T V E T R D G Q --	V I N E T S -	
F6ZXN6_CIOIN/482-507	T V T T K K V V V K T I E T K D G K --	V V S Q T E D V -	
ADA087XRH0_POEFO/456-478	T H G K K T V V V I K T V E T R D G E --	V V K E S -	
ADA091EIM5_FUKDA/466-494	T P T K S A V V V K K I K T C D R K --	L V S Q T S D V L P K	
G1LPY9_AILME/434-462	G S S F K T V V V K K I K T C A G R --	L V S E S P A V L S K	
ADA1S3QMPB_SALSA/221-244	T T M S I K V L I K T I E T R D G Q --	V I N E S M -	
HDXMM4_OTOGA/319-347	P A P Q L A I V E K K I E T H H G K --	L T A Q T S D V L S K	
U3IS84_ANAPI/402-424	A H B K R T I V V K T V E T R D G E -	V I K E S -	
H8BE29_LATCH/486-512	P S K R N L V V V K K I E A Q Q G K --	I V S E T S N Y L -	
H2ZA20_CIOSA/191-213	T K T T K V I V K T I E T R D G K --	I V V S S -	
07ZXU4_XENLA/478-506	K T S K R S I L V K T V E T K D G R --	V L S E S S S D V F S K	
Q502Q7_DANRE/409-431	P A P G K K V V I K T V E I R D G E --	V V K E S -	
V8NZM9_OPHHA/388-413	S S L K R P L M I K T I E T K D G K --	I I S E A S H F --	
ADA219MNDS_PAPAN/415-438	L V F R K T L L I K T V E T R D G Q --	V I N E T S -	
ADA1U7T2W5_TARSY/126-152	T G F S R A V - V E K I E T R H G K --	L V S E A S H A L P K	
G1L6P0_AILME/439-466	A S S N D A M A A K K I G T C H G K --	L A S E S Y D V L S -	
H8BATT_LATCH/432-456	V H T K K T V M I K T I E T D G K E E K V V N E -	- - - - -	

What do we
expect to find?

From the MSA we build
a family profile



We run a profile
to sequence search

	10	20	30
ADA2K5J_DQ6_COLAP/495-493	T S S T R A M V V K E I E T R D G K - - - L M S E S S D V L P K		
G3I192_CIGR/460-488	Y S R T K A V V V K K I E T R D G K - - - L V S E S S D I L L P K		
F6YKZ7_HORSE/457-485	I G P P S A I V V K K I E T R D G K - - - L V S E S S D I L L P K		
ADA2BHC78_NOMLE/393-420	T S S T R A M V V K K I E T R D G K - - - L C G H L G N L L P K		
ADA2D0SIM8_ICTPU/473-499	T Q S K K N I V I K M I E T K D G R - - - V V S E S S E I V - -		
ADA1U7SF33_ALLS/452-479	V T K S K A I V I K K I E T R D G K - - - L V S E S S D I L S -		
ADA087XNO_PEOF/465-493	T T T K K N V I V K T I E T N D G K - - - L V S E S S N I I E K		
ADA286XSR9_CAVPO/225-253	T R T K S A V V V K K I E T R D G K - - - L V S E T S D V L P K		
H2L7B8_ORYLA/463-489	T Q S K K S V V V I K T I E T K D G R - - - V V S E S S E L I - -		
HDXIAL_OTOGA/235-263	S H T S A R A K V T K I E T H D R K - - - L V S E S S D I L L P K		
G1M6M5_AILME/375-401	V S S S K A L A V N T I E T C S G K - - - L V S K P S T L L - -		
M7BMIM1_CHEMY/444-471	L A R S K A I V I K K I E T K D G R - - - L V S E T S N V L S -		
K7FDH1_PELS/460-487	V S K S K T V V I K K I E T R D G R - - - L V S E T S D I L S -		
G1Q220_MYOLU/360-385	T S S S K T V V V V K E I E T L L Q G K - - - L V S E T S D V L - -		
G5BQG7_HETGA/184-212	L T H T K S A M V V K I E T H D G K - - - V V S E T S D V L H K		
W5UL23_ICTPU/446-470	I H S K K T V L I K T I E T R D G E - - S K V V S E S - - -		
B3R18_ORENI/488-514	T Q T K K N V V I K M I E T K D G K - - - V V S E S S E V I - -		
H2RK76_TAKRU/432-455	M H S K K T V L I K T I E T R D G E - - - I I T R T S - - -		
ADA1L1RW09_CHICK/24-47	T H S K R T L L I K T V E T R D G Q - - - V I N E T S - - -		
F6ZXN6_CIOIN/482-507	T V T T K K V V V V K T I E T K D G K - - - V V S Q T E D V - -		
ADA087XRH9_PEOF/456-478	T H G K K T V V V I K T V E T R D G E - - - V V K E S - - -		
ADA091EIM8_FUKDA/466-494	T P T K S A V V V K K I K T C D R K - - - L V S Q T S D V L P K		
G1LPY9_AILME/434-462	G S S F K T V V V V K K I K T C A G R - - - L V S E S P A V L S K		
ADA1S3QMPE_SALSA/221-244	T T M S I K V L I K T I E T R D G Q - - - V I N E S M - - -		
HDXMM4_OTOGA/319-347	P A P Q L A I V E K K I E T H H G K - - - L T A Q T S D V L S K		
U3IS84_ANAPI/402-424	A H B K R T I V V V K T V E T R D G E - - - V I K E S - - -		
H8BE29_LATCH/486-512	P S K R N L V V V V K K I E A Q Q G K - - - I V S E T S N Y L - -		
H2ZA20_CIOSA/191-213	T K T T K V V V K T I E T R D G K - - - I V V S S - - -		
072XU4_XENLA/478-506	K T S K R S I L V K T V E T K D G R - - - V L S E S S D V F S K		
Q502Q7_DANRE/409-431	P A P G K K V V I K T V E I R D G E - - - V V K E S - - -		
V8NZM9_OPHHA/388-413	S S L K R P L M I K T I E T K D G K - - - I I S E A S H F - - -		
ADA215MND8_PAPAN/415-438	L V F R K T L L I K T V E T R D G Q - - - V I N E T S - - -		
ADA1U7T2W5_TARSY/126-152	T G F S R A V - V E K I E T R H G K - - - L V S E A S H A L P K		
G1L6P0_AILME/439-466	A S S N C A M A A K K I G T C H G K - - - L A S E S Y D V L S -		
H8BAT7_LATCH/432-456	V H T K K T V M I K T I E T D G K E E K V V N E - - - - -		

Previous
homologs +
more remote
homologs



HMMER

Biosequence analysis using profile hidden Markov Models

Home | Search | Results | Software | Help | About | Contact

Next release within a week, think about downloading your results

JACKHMMER Results

[Search Again](#)

Score

Taxonomy

Domain

[Download](#)

Iteration 1

[Return to the Results Summary](#) | [Jump to threshold](#)

Sequence selection ?

above threshold ?

unselect all ?

[Continue the search](#)

[Start iteration 2](#)



Click on Start Iteration 2

Distribution of Significant Hits ?





Biosequence analysis using profile hidden Markov Models

Home | Search | **Results** | Software | Help | About | Contact

Next release within a week, think about downloading your results

JACKHMMER Results

[Search Again](#)

Jackhmmer Summary

Iteration	Results	Hits			
		New	Lost <small>?</small>	Dropped <small>?</small>	Total
1	464126AA-03EA-11EA-A13A-4471F75AEC3D.1	+301	-	-	301
2	464126AA-03EA-11EA-A13A-4471F75AEC3D.2	+211	-	-	512

[Start Next Iteration](#)

HmmWeb version 2.40.0

Next release within a week, think about downloading your results

Comments or questions about the site? Click here to use our contact form



[Follow @hmm3r](#)



Biosequence analysis using profile hidden Markov Models

Home | Search | **Results** | Software | Help | About | Contact

Next release within a week, think about downloading your results

JACKHMMER Results

[Search Again](#)

Jackhmmer Summary

Iteration	Results	Hits			
		New	Lost <small>?</small>	Dropped <small>?</small>	Total
1	464126AA-03EA-11EA-A13A-4471F75AEC3D.1	+301	-	-	301
2	464126AA-03EA-11EA-A13A-4471F75AEC3D.2	+211	-	-	512

Click



[Start Next Iteration](#)

HmmWeb version 2.40.0

Next release within a week, think about downloading your results

Comments or questions about the site? Click here to use our contact form



[Follow @hmm3r](#)



HMMER

Biosequence analysis using profile hidden Markov Models

[Home](#)[Search](#)[Results](#)[Software](#)[Help](#)[About](#)[Contact](#)

Next release within a week, think about downloading your results

JACKHMMER Results

[Score](#)[Taxonomy](#)[Domain](#)[Download](#)

Iteration 2

[Return to the Results Summary](#) | [Jump to first new match](#) | [Jump to threshold](#)

Sequence selection [?](#)

above threshold [?](#)

unselect all [?](#)

[Continue the search](#)

[Start iteration 3](#)

[« previous iteration](#)

Distribution of Significant Hits [?](#)

	SEARCHED	IF rod domain-containing protein	GAMMA JACCHUS	AAA	BBB	CCC	ECC	FCC	GCC	HCC	I
>	A0A2K5Q3B3_CEBCA	IF rod domain-containing protein	Cebus capucinus imitator	XXX	BBB	CCC	ECC	FCC	GCC	HCC	✓
>	A0A2R9BVP0_PANPA	IF rod domain-containing protein	Pan paniscus	XXX	BBB	CCC	ECC	FCC	GCC	HCC	✓

(show all) alignments Your search took: 3.72 secs showing rows 1 - 50 of 738

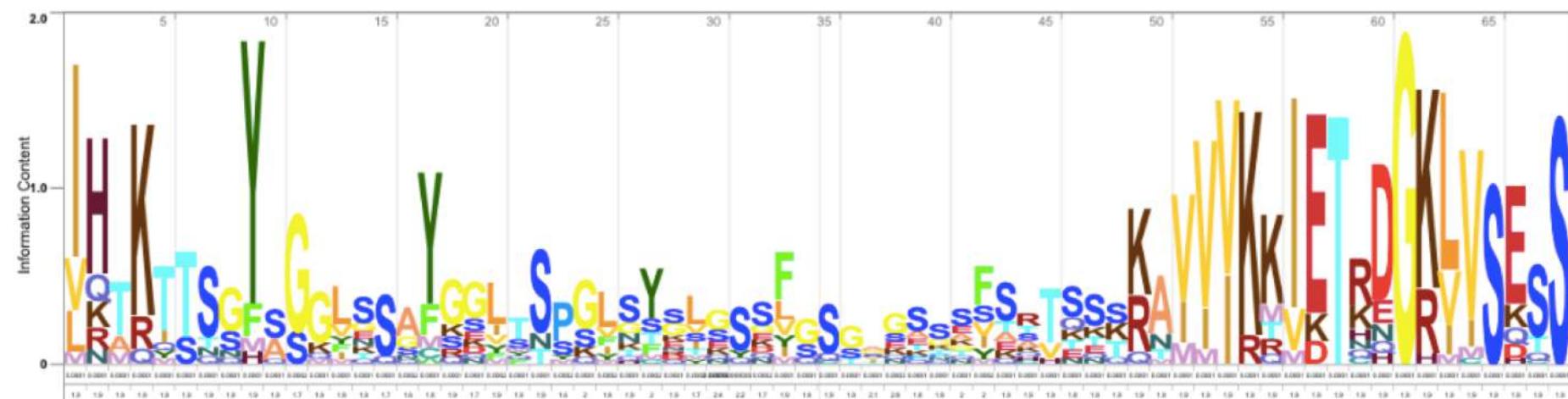
Continue the search

Start iteration 3

Search Details

« First « Previous Page 10 of 15 Next » Last »

Model Position



	GO:000220_0000000	IF rod domain-containing protein	Gambusia jacchus	AAA	BBB	CCC	4.0e-35	✓
>	A0A2K5Q3B3_CEBCA	IF rod domain-containing protein	Cebus capucinus imitator	XXX	BBB	CCC	4.9e-35	✓
>	A0A2R9BVP0_PANPA	IF rod domain-containing protein	Pan paniscus	XXX	BBB	CCC	5.4e-35	✓

(show all) alignments Your search took: 3.72 secs showing rows 1 - 50 of 738

Continue the search

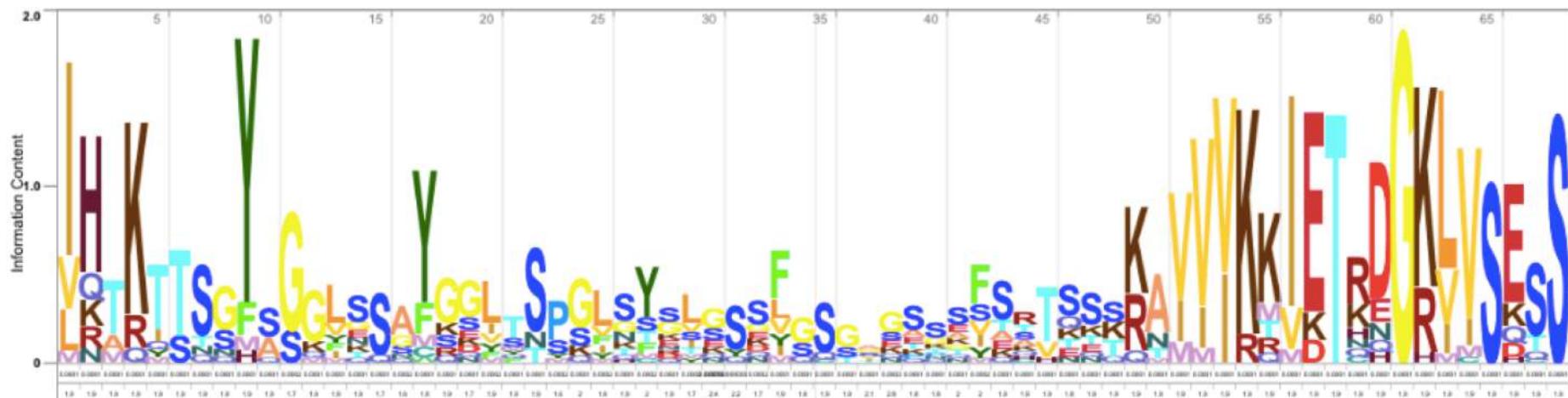
Start iteration 3

Click

Search Details

« First « Previous Page 10 of 15 Next » Last »

Model Position





Biosequence analysis using profile hidden Markov Models

Home Search **Results** Software Help About Contact

Next release within a week, think about downloading your results

JACKHMMER Results

[Search Again](#)

Jackhmmer Summary

Iteration	Results	Hits			
		New	Lost <small>?</small>	Dropped <small>?</small>	Total
1	464126AA-03EA-11EA-A13A-4471F75AEC3D.1	+301	-	-	301
2	464126AA-03EA-11EA-A13A-4471F75AEC3D.2	+211	-	-	512
3	464126AA-03EA-11EA-A13A-4471F75AEC3D.3	+438	-	-	950

[Start Next Iteration](#)

HmmerWeb version 2.40.0

Next release within a week, think about downloading your results

Comments or questions about the site? [Click here](#) to use our contact form.

This is cool!



HMMER

Biosequence analysis using profile hidden Markov Models

Home | Search | Results | Software | Help | About | Contact

JACKHMMER Results

[Search Again](#)

Jackhmmer Summary

Iteration	Results	Hits			
		New	Lost	Dropped	Total
1	86F38F56-DB84-11E8-9267-8E60E976C163.1	+255	-	-	255
2	86F38F56-DB84-11E8-9267-8E60E976C163.2	+75	-	-	330
3	86F38F56-DB84-11E8-9267-8E60E976C163.3	+236	-	-	566

[Start Next Iteration](#)



Click up to 8th iteration

Let's take a 5 mins break



Jackhammer Summary

Iteration	Results	Hits				Total
		New	Lost	Dropped		
1	464126AA-03EA-11EA-A13A-4471F75AEC3D.1	+301	-	-		301
2	464126AA-03EA-11EA-A13A-4471F75AEC3D.2	+211	-	-		512
3	464126AA-03EA-11EA-A13A-4471F75AEC3D.3	+438	-	-		950
4	464126AA-03EA-11EA-A13A-4471F75AEC3D.4	+220	-	-		1170
5	464126AA-03EA-11EA-A13A-4471F75AEC3D.5	+193	-	3		1360
6	464126AA-03EA-11EA-A13A-4471F75AEC3D.6	+191	-	-		1551
7	464126AA-03EA-11EA-A13A-4471F75AEC3D.7	+1129	-	3		2677
8	464126AA-03EA-11EA-A13A-4471F75AEC3D.8	+2742	-16	34		5369
9	464126AA-03EA-11EA-A13A-4471F75AEC3D.9	+1875	-46	73		7125
10	464126AA-03EA-11EA-A13A-4471F75AEC3D.10	+834	-28	78		7853
11	464126AA-03EA-11EA-A13A-4471F75AEC3D.11	+559	-17	27		8368
12	464126AA-03EA-11EA-A13A-4471F75AEC3D.12	+320	-3	50		8635
13	464126AA-03EA-11EA-A13A-4471F75AEC3D.13	+162	-4	38		8755
14	464126AA-03EA-11EA-A13A-4471F75AEC3D.14	+281	-6	29		9001
15	464126AA-03EA-11EA-A13A-4471F75AEC3D.15	+108	-3	33		9073
16	464126AA-03EA-11EA-A13A-4471F75AEC3D.16	+67	-13	41		9086
17	464126AA-03EA-11EA-A13A-4471F75AEC3D.17	+54	-4	25		9111
18	464126AA-03EA-11EA-A13A-4471F75AEC3D.18	+45	-9	24		9123

Loads of hits!!!

Start Next Iteration

Jackhammer Summary

Iteration	Results	Hits				Total
		New	Lost 	Dropped 		
1	464126AA-03EA-11EA-A13A-4471F75AEC3D.1	+301	-	-	-	301
2	464126AA-03EA-11EA-A13A-4471F75AEC3D.2	+211	-	-	-	512
3	464126AA-03EA-11EA-A13A-4471F75AEC3D.3	+438	-	-	-	950
4	464126AA-03EA-11EA-A13A-4471F75AEC3D.4	+220	-	-	-	1170
5	464126AA-03EA-11EA-A13A-4471F75AEC3D.5	+193	-	3	-	1360
6	464126AA-03EA-11EA-A13A-4471F75AEC3D.6	+191	-	-	-	1551
7	464126AA-03EA-11EA-A13A-4471F75AEC3D.7	+1129	-	3	-	2677
8	464126AA-03EA-11EA-A13A-4471F75AEC3D.8	+2742	-16	34	-	5369
9	464126AA-03EA-11EA-A13A-4471F75AEC3D.9	+1875	-46	73	-	7125
10	464126AA-03EA-11EA-A13A-4471F75AEC3D.10	+834	-28	78	-	7853
11	464126AA-03EA-11EA-A13A-4471F75AEC3D.11	+559	-17	27	-	8368
12	464126AA-03EA-11EA-A13A-4471F75AEC3D.12	+320	-3	50	-	8635
13	464126AA-03EA-11EA-A13A-4471F75AEC3D.13	+162	-4	38	-	8755
14	464126AA-03EA-11EA-A13A-4471F75AEC3D.14	+281	-6	29	-	9001
15	464126AA-03EA-11EA-A13A-4471F75AEC3D.15	+108	-3	33	-	9073
16	464126AA-03EA-11EA-A13A-4471F75AEC3D.16	+67	-13	41	-	9086
17	464126AA-03EA-11EA-A13A-4471F75AEC3D.17	+54	-4	25	-	9111
18	464126AA-03EA-11EA-A13A-4471F75AEC3D.18	+45	-9	24	-	9123



Let's have a look

Start Next Iteration

>	A0A3L8S0H6_CHLGU	Uncharacterized protein	Chloebia gouldiae				1.0e-24	<input checked="" type="checkbox"/>
>	A0A151PC95_ALLMI	Keratin, type I cytoskeletal 20	Alligator mississippiensis				1.0e-24	<input checked="" type="checkbox"/>
>	H2NHF0_PONAB	Keratin 8	Pongo abelii				1.3e-24	<input checked="" type="checkbox"/>
>	A0A2I2ZH07_GORGO	IF rod domain-containing protein	Gorilla gorilla gorilla				1.9e-24	<input checked="" type="checkbox"/>
>	K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens				2.3e-24	<input checked="" type="checkbox"/>
>	G1P1Y9_MYOLU	IF rod domain-containing protein	Myotis lucifugus				2.6e-24	<input checked="" type="checkbox"/>
>	K2C8_HUMAN	Isoform 2 of Keratin, type II cytoskeletal 8	Homo sapiens				2.8e-24	<input checked="" type="checkbox"/>

(show all) alignments

Your search took: 4.93 secs

showing rows 1 - 50 of 10077

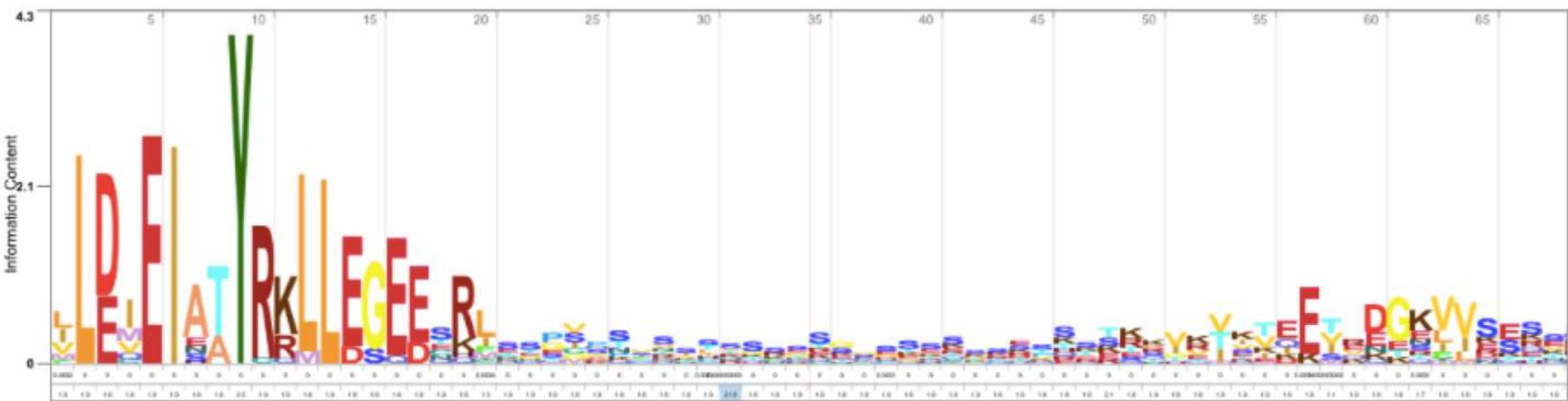
Continue the search

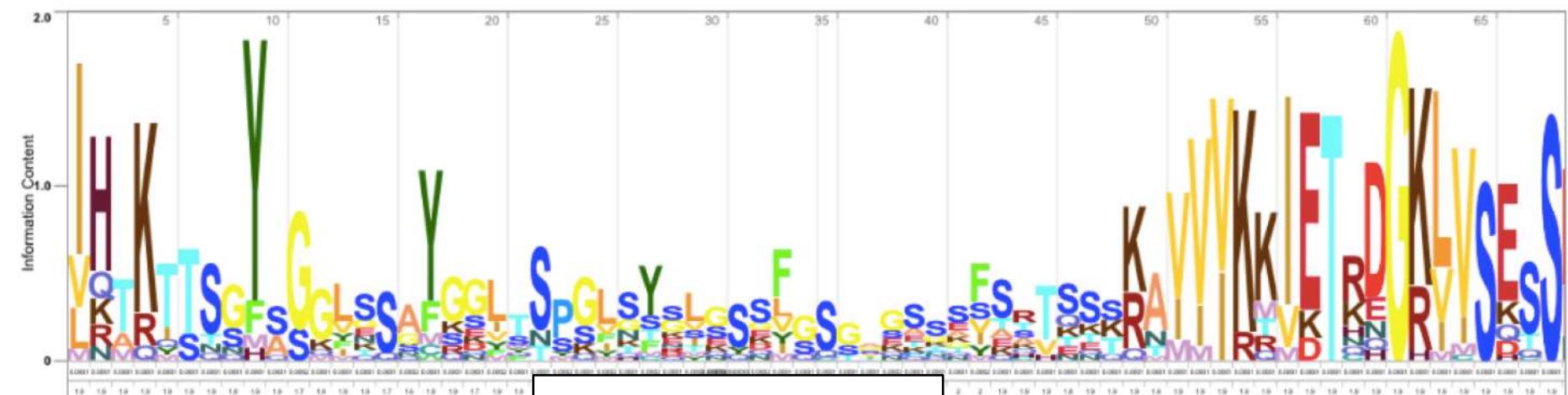
Start iteration 19

Search Details

« First « Previous Page 1 of 202 Next » Last »

Model Position

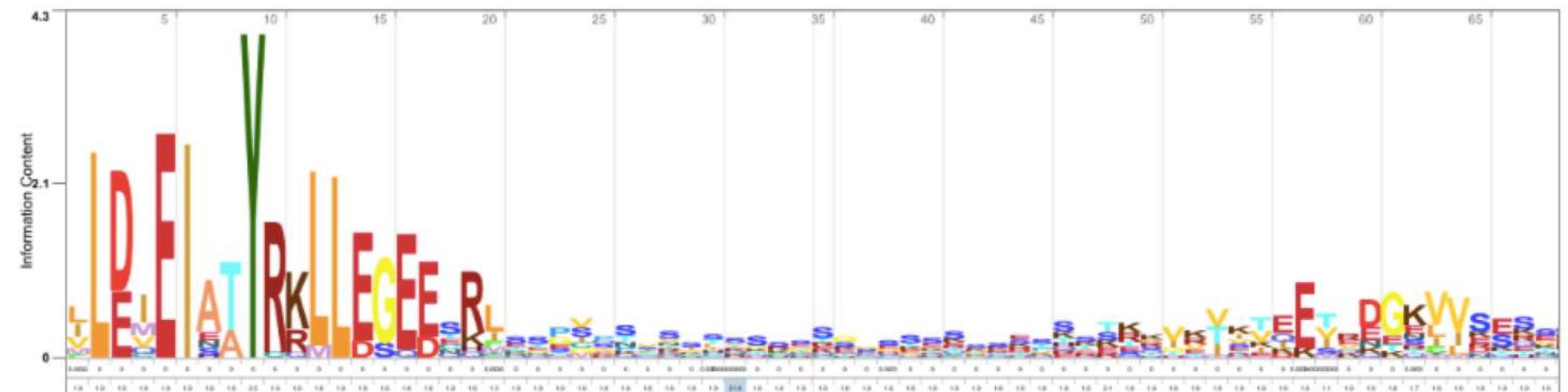




Column number:

1

Iteration 18



Column number:

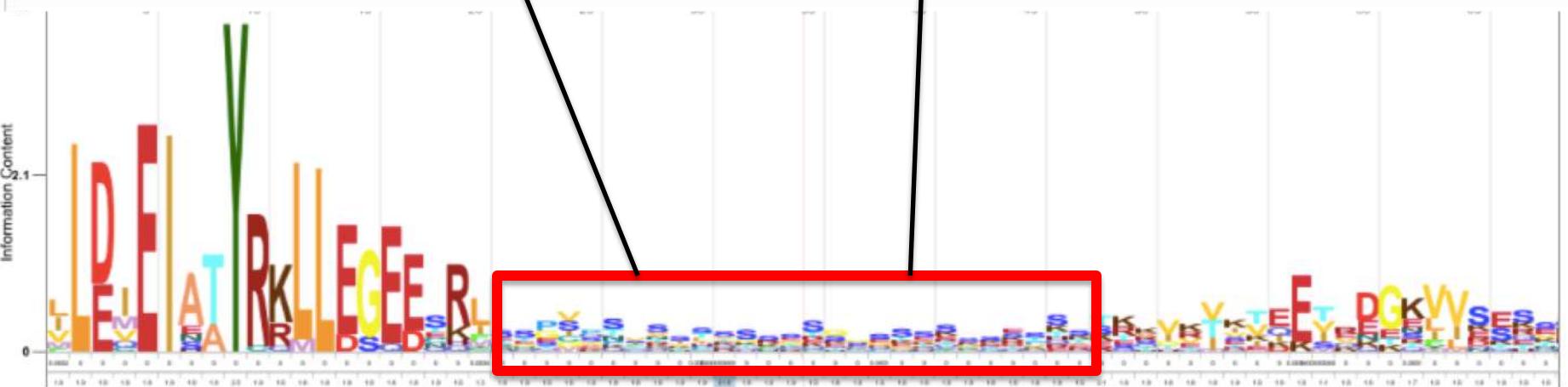
1

	Target		Description					Species		Cross-references		E-value		
V	A0A218UCX9_9PASE		Keratin, type II cytoskeletal 75					Lonchura striata domestica		xx	##	■	1.6e-72	✓
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)	Bit Score	E-value			
start	end	start	end	start	end						Ind.	Cond.		
2	36	484	549	485	519	0.02	0.76	54.3 (19)	80.0 (28)	40.7	8.6e-07	1.8e-10		
Query	2	LDiEIatYRkLLEGEESrlsspvpsvssssss...g 36					LDiEIatYR LLEGEESr++++ p ++ +s ++							
Target	485	LDIEIATYRTLLEGEESRICTGNPVSVAVVGCGT 519					PP 9*****9998876666543332							
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)	Bit Score	E-value			
start	end	start	end	start	end						Ind.	Cond.		
2	27	1177	1262	1178	1203	1.69	0.84	76.9 (20)	96.2 (25)	39.4	2.2e-06	4.4e-10		
Query	2	LDiEIatYRkLLEGEESrlsspvps... 27					LDiEIatYRkLLEGEESrls+++ ++							
Target	1178	LDIEIATYRKLLEGEESRLSGEGLNP 1203					PP 9*****6665543							
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)	Bit Score	E-value			
start	end	start	end	start	end						Ind.	Cond.		
2	32	1781	1857	1782	1821	0.08	0.76	74.2 (23)	90.3 (28)	47.9	4.8e-09	9.8e-13		
Query	2	LDiEIatYRkLLEGEESrlsspvps... 32					LDiEIatYRkLLEGEESrl ++vv vs+s+							
Target	1782	LDIEIATYRKLLEGEESRLaGEGVGAVSVSvsssgmgy 1821					PP 9*****999999999544444440							

V	K2C8_HUMAN			Keratin, type II cytoskeletal 8				Homo sapiens		xx	■	■	■	■	■	■	2.3e-24	✓
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)		% Similarity (count)		Bit Score	E-value					
start	end	start	end	start	end			% Identity (count)		% Similarity (count)			Ind.	Cond.				
2	73	383	483	384	483	5.90	0.99	65.3 (47)		98.6 (71)		97.0	2.3e-24	4.7e-28				
Query	2	LD <i>EIatYRkLLEGEEsrlsspppvsvssss</i>*********	53						
Target	384	LDIEIATYRK LLEGEEsrlsEGMQNMSIHT	ttsgyagglssaygglspglsyslg	s+tsg	+sss+s++ss	++v+v						463						
PP		9*****	*****	*****	*****	*****	*****	*****	*****	*****	*****							
Query	54	kteetrdgkvyseseeelpk	73															
Target	464	KKIETRDGKLVSESSDVLPK	483	k++etrdgk+vsess++lpk														
PP		*****	*****	*****	*****	*****	*****	*****	*****	*****	*****	997						

V	K2C8_HUMAN			Keratin, type II cytoskeletal 8				Homo sapiens		xx	■	■	■	■	■	2.3e-24	✓
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)		% Similarity (count)		Bit Score	E-value				
start	end	start	end	start	end			% Identity (count)		% Similarity (count)			Ind.	Cond.			
2	73	383	483	384	483	5.90	0.99	65.3 (47)		98.6 (71)		97.0	2.3e-24	4.7e-28			
Query	2	LDIEIatYRkLLEGEEs	lssspvpssvssss..****	sssggssssssessstkykv		53							
Target	384	LDIEIATYRKLLGEESRLESGMONMSIHTKttsgyagg	lssayggltspglsyslgsSFGSGAGSSFSRTSSSRAVVV****	s+tsg +tssst+s+st+v+v		463							
PP		9*****	*****	*****	*****	*****	*****	*****		*****	*****	*****	*****	*****	*****		
Low complexity																	
Query	54	kteetrdgkvvsesseelpk	73****	k++etrdgk+vsess++lpk									
Target	464	KKIETRDGKLVSESSDVLPK	483****	*****997									
PP		*****	*****	*****	*****	*****	*****	*****									

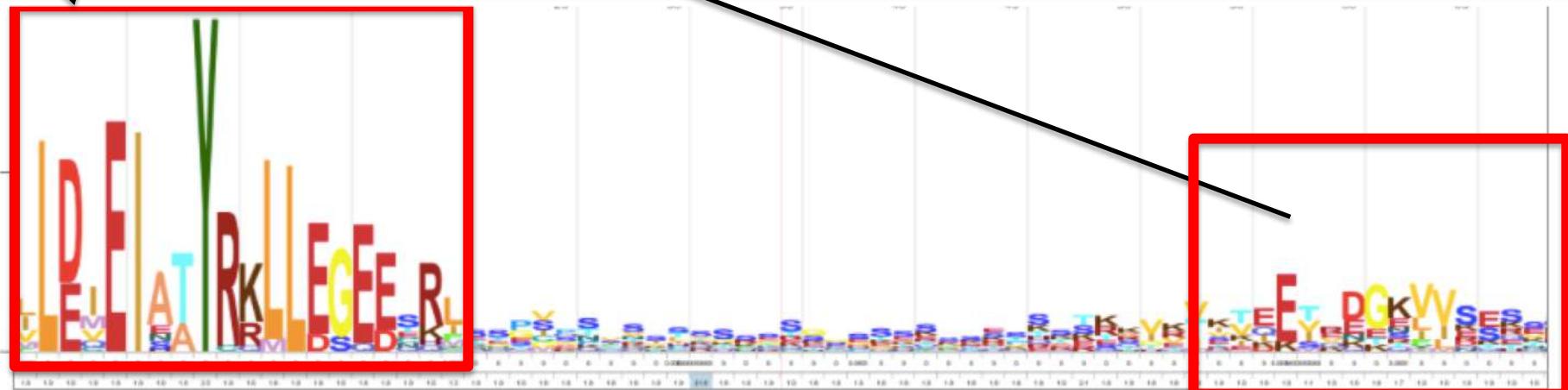
V	K2C8_HUMAN			Keratin, type II cytoskeletal 8				Homo sapiens							2.3e-24	<input checked="" type="checkbox"/>
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)		% Similarity (count)		Bit Score	E-value		Ind.	Cond.
start	end	start	end	start	end			% Identity (count)		% Similarity (count)			Ind.			
2	73	383	483	384	483	5.90	0.99	65.3 (47)		98.6 (71)		97.0	2.3e-24	4.7e-28		
Query	2	LD <i>EIatYRkLLEGEEs</i> :lssspvpsvssss..	*.....**		*	*						
		LD <i>EIatYRkLLEGEEs</i> :l+s+++++s+++						ssssggsssssssssst!kykv		53						
Target	384	LDIEIA <i>TYRkLLEGEEsRLESGMONMSIHTkttsgyagg</i> lssayggltspglsyslgssFGSGAGSSFSRTSSSRAVVV						s+tsga +sss+s+tss+s++v+v		463						
PP		9*****						*****		*****						
Query	54	kteetrdgkvyseseeelpk	73**			Low complexity								
		k++etrdgk+vsess++lpk														
Target	464	KKIETRDGKLVSESSDVLPK				483		*****		997						
PP		*****						*****		*****						



V	K2C8_HUMAN			Keratin, type II cytoskeletal 8				Homo sapiens								2.3e-24	<input checked="" type="checkbox"/>
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)		% Similarity (count)		Bit Score	E-value		Ind.	Cond.	
start	end	start	end	start	end			65.3 (47)		98.6 (71)			2.3e-24				
2	73	383	483	384	483	5.90	0.99	65.3 (47)		98.6 (71)		97.0	2.3e-24		4.7e-28		
Query																	
Target	38																
PP																	
Query	5																
Target	46																
PP																	

Sequence alignment details:

- Query: LDI EI at YR kLLEGEE Esrl s pvp sv vssss..... *.....*
- Target: LDIEIATYRkLLEGEEEsrl+s+++++s+++ s++sg +sss+s++ss+++v+v
- PP: 9*****
- Query: kteetrdgkvy sesseelpk 73
- Target: KKIETRDGKLVSESSDVLPK 483
- PP: 999





Next release within a week, think about downloading your results

JACKHMMER Results

[Search Again](#)[Score](#)[Taxonomy](#)[Domain](#)[Download](#)

Perform a new search
• with new input
• with these results

Iteration 18

[Return to the Results Summary](#) | Show 9 completely lost matches. | 24 matches dropped below your threshold. | [Jump to first new match](#) | [Jump to threshold](#)

Sequence selection

- above threshold
 unselect all

[Continue the search](#)[Start iteration 19](#)[« previous iteration](#)

Distribution of Significant Hits



■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

[First](#) [« Previous](#) **Page 1** of 202 [Next »](#) [Last »](#)

Significant Query Matches (9123) in *uniprotrefprot* (v.2019_09)[Customise](#)

Target

Description

Species

Cross-references

E-value



Next release within a week, think about downloading your results

JACKHMMER Results

Score

Taxonomy

Dom

Iteration 18

Click on WITH THESE RESULTS

Search Again

Perform a new search

- with new input
- with these results

Return to the [Results Summary](#) | Show 9 completely lost matches. | 24 matches dropped below your threshold. | [Jump to first new match](#) | [Jump to threshold](#)

Sequence selection

- above threshold [?](#)
 unselect all [?](#)

Continue the search

Start iteration 19

« previous iteration

Distribution of Significant Hits [?](#)

■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

« First « Previous Page 1 of 202 Next » Last »

Significant Query Matches (9123) in *uniprotrefprot* (v.2019_09)

Customise

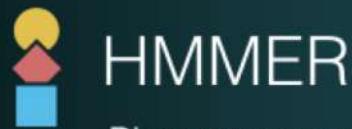
Target

Description

Species

Cross-references

E-value



Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

Next release within a week, think about downloading your results

phmmmer hmmscan **hmmssearch** jackhmmer

protein alignment/profile-HMM vs protein sequence database

Using Existing Search as Input



This search will be run with an HMM generated from the 9123 hits above threshold in your previous search [464126AA-03EA-11EA-A13A-4471F75AEC3D.18](#).

Submit

Reset

▼ Sequence Database

Frequently used databases:

Click on PDB

ssPfam PDB Ensembl

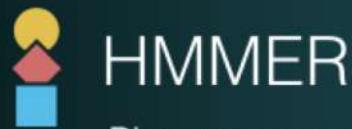
Current database selection:

Reference Proteomes

► Restrict by Taxonomy

▼ Cut-Offs

E-value Bit score



Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

Next release within a week, think about downloading your results

phmmmer hmmscan **hmmsearch** jackhmmer

protein alignment/profile-HMM vs protein sequence database

Using Existing Search as Input



This search will be run with an HMM generated from the 9123 hits above threshold in your previous search [464126AA-03EA-11EA-A13A-4471F75AEC3D.18](#).

Click on Submit

Submit

Reset

▼ Sequence Database

Frequently used databases: [Reference Proteomes](#) [UniProtKB](#) [SwissProt](#) [PDB](#) [Ensembl](#)

Current database selection:

Reference Proteomes

► Restrict by Taxonomy

▼ Cut-Offs

E-value Bit score



Next release within a week, think about downloading your results

HMMSEARCH Results

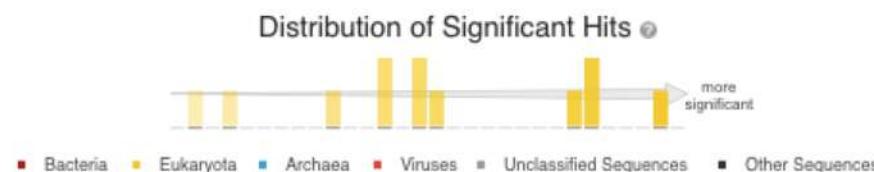
[Search Again](#)

Score

Taxonomy

Domain

Download

Significant Query Matches (12) in *pdB* (v.2019-09-20)[Customise](#)

Target	Description	Species	Cross-references	E-value
> 2xv5_A PDBe RCSB ⓘ	LAMIN-A/C	Homo sapiens ⓘ		2.1e-10
> 1gk6_A PDBe RCSB ⓘ	VIMENTIN	Saccharomyces cerevisiae S288C ⓘ		1.9e-09
> 1gk6_A PDBe RCSB ⓘ	VIMENTIN	Homo sapiens ⓘ		1.9e-09
> 1gk4_A PDBe RCSB ⓘ	VIMENTIN	Homo sapiens ⓘ		5.1e-09
> 3tyy_A PDBe RCSB ⓘ	Lamin-B1	Homo sapiens ⓘ		2.9e-07
> 1x8y_A PDBe RCSB ⓘ	Lamin A/C	Homo sapiens ⓘ		7.2e-07
> 3v4q_A PDBe RCSB ⓘ	Prelamin-A/C	Homo sapiens ⓘ		8.5e-07
> 3v5b_A PDBe RCSB ⓘ	Prelamin-A/C	Homo sapiens ⓘ		1.8e-06
> 3v4w_A PDBe RCSB ⓘ	Prelamin-A/C	Homo sapiens ⓘ		1.8e-06

V	1gk6_A PDB RCSB		VIMENTIN				Saccharomyces cerevisiae S288C						1.9e-09	
Target				Description				Species				Cross-references		E-value
start	end	start	end	start	end							Ind.	Cond.	
2	21	39	59	40	59	0.80	0.95	95.0 (19)		100.0 (20)		42.7	1.9e-09	7.3e-14

```

.....*.....*
Query      2 LDIEIATYRKLLGEEEsr1s 21
                  LDIEIATYRKLLGEEEsr+s
Target     40 LDIEIATYRKLLGEEESRIS 59
PP          9*****85

```

```

Query      2 LDIEIATYRKLLGEESRIS 21
          LDIEIATYRKLLGEESR+s
Target     40 LDIEIATYRKLLGEESRIS 59
PP         9*****85

```

V	1gk4_A PDB RCSB		VIMENTIN				Homo sapiens			aa	aa	B	S	C	O	G	5.1e-09
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)	Bit Score	E-value				Ind.	Cond.	
start	end	start	end	start	end						Ind.	Cond.					
2	20	65	84	66	84	1.21	0.96	94.7 (18)	100.0 (19)	40.2	1.1e-08	4.2e-13					

```

Query      2 LDIEIATYRkLLEGESrsl 20
           LDIEIATYRkLLEGESr+
Target     66 LDIEIATYRkLLEGESRI 84
PP          9*****9*****96

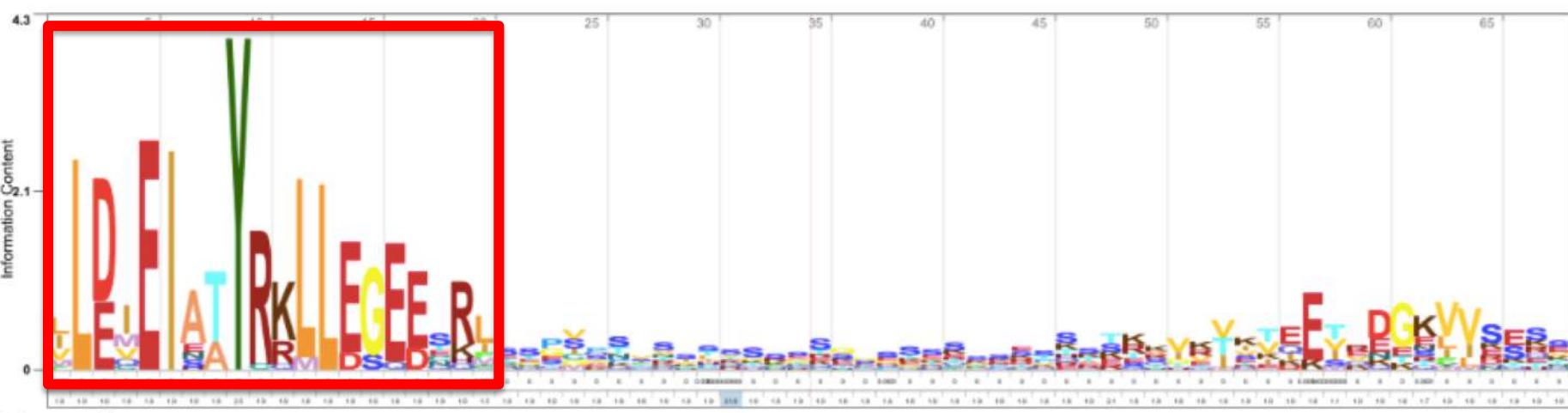
```

Insignificant matches

Query	34*..... sgggssssssssstkkvkvktestrd + + +++ +++++++ +++e +d	60
Target	7	HSSRNEFGQKESRACLERIQELED	33
PP		566677788899999999999999876	

v	1gk6_A PDBe RCSB®				VIMENTIN				Saccharomyces cerevisiae S288C®								1.9e-09		
Target				Description				Species				Cross-references				E-value			
start	end	start	end	start	end											Ind.	Cond.		
2	21	39	59	40	59	0.80	0.95		95.0 (19)		100.0 (20)		42.7	1.9e-09	7.3e-14				
<hr/>																			
Query	2	LDIEIlatYRkLLEGEESr+s 21				LDIEIlatYRkLLEGEESr+s				LDIEIATYRkLLEGEESRIS 59				LDIEIATYRkLLEGEESRIS 59					
Target	40																		
PP		9*****85																	
<hr/>																			
v	1gk6_A PDBe RCSB®				VIMENTIN				Homo sapiens®								1.9e-09		
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)		Bit Score		E-value						
start	end	start	end	start	end											Ind.	Cond.		
2	21	39	59	40	59	0.80	0.95	95.0 (19)		100.0 (20)		42.7	1.9e-09	7.3e-14					
<hr/>																			
Query	2	LDIEIlatYRkLLEGEESr+s 21				LDIEIlatYRkLLEGEESr+s				LDIEIATYRkLLEGEESRIS 59				LDIEIATYRkLLEGEESRIS 59					
Target	40																		
PP		9*****85																	
<hr/>																			
v	1gk4_A PDBe RCSB®				VIMENTIN				Homo sapiens®								5.1e-09		
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)		Bit Score		E-value						
start	end	start	end	start	end											Ind.	Cond.		
2	20	65	84	66	84	1.21	0.96	94.7 (18)		100.0 (19)		40.2	1.1e-08	4.2e-13					

Model Position





157530 Biological
Macromolecular Structures
Enabling Breakthroughs in
Research and Education

Search by PDB ID, author, macromolecule, sequence, or ligand

Go

[Advanced Search](#) | [Browse by Annotations](#)

PDB-101

WORLDWIDE
PROTEIN DATA BANK



EMDataResource
Unified Data Resource for IODEM

ndb
NUCLEIC ACID
DATABASE



Worldwide
Protein Data Bank
Foundation



Structure Summary

3D View

Annotations

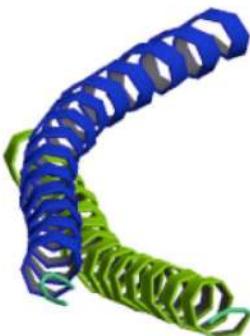
Sequence

Sequence Similarity

Structure Similarity

Experiment

Biological Assembly 1



Display Files ▾

Download Files ▾

2XV5

Human lamin A coil 2B fragment

DOI: [10.2210/pdb2XV5/pdb](https://doi.org/10.2210/pdb2XV5/pdb)

Classification: [STRUCTURAL PROTEIN](#)

Organism(s): [Homo sapiens](#)

Expression System: [Escherichia coli BL21\(DE3\)](#)

Deposited: 2010-10-22 Released: 2011-03-09

Deposition Author(s): [Kapinos, L.E.](#), [Burkhard, P.](#), [Aebi, U.](#), [Herrmann, H.](#), [Strelkov, S.V.](#)

Experimental Data Snapshot

Method: X-RAY DIFFRACTION

Resolution: 2.4 Å

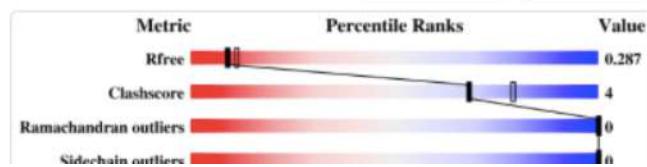
R-Value Free: 0.284

R-Value Work: 0.241

wwPDB Validation

3D Report

Full Report

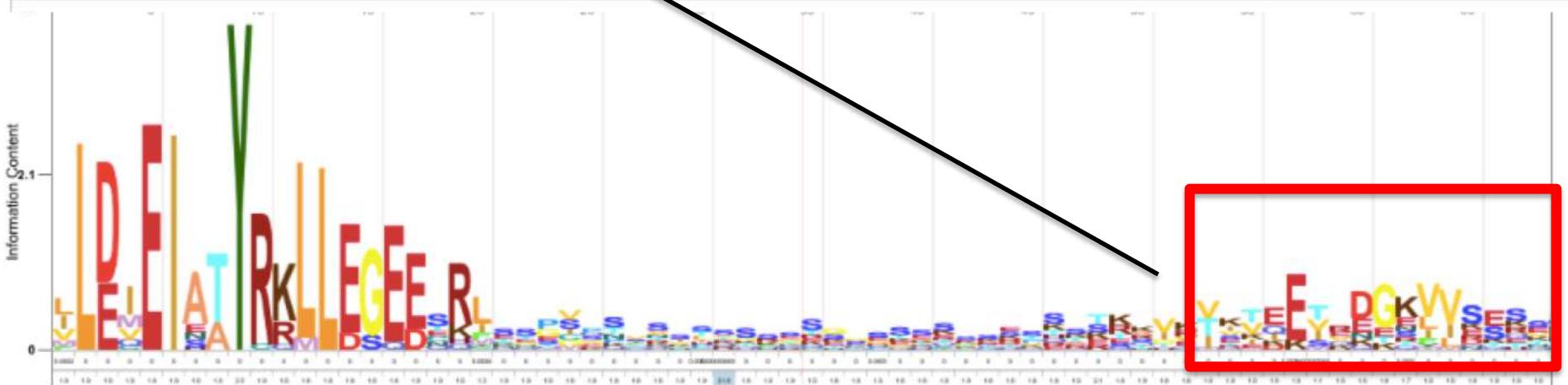


3D View: [Structure](#) | [Electron Density](#)

Standalone Viewers

V	K2C8_HUMAN				Keratin, type II cytoskeletal 8				Homo sapiens									2.3e-24	<input checked="" type="checkbox"/>
Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)		% Similarity (count)		Bit Score	E-value		Ind.	Cond.			
start	end	start	end	start	end			% Identity (count)		% Similarity (count)			Ind.						
2	73	383	483	384	483	5.90	0.99	65.3 (47)		98.6 (71)		97.0	2.3e-24		4.7e-28				

Query 2 LDI~~EIatYRkLLEGEEsrls~~spvpvs~~ssss~~..... ssssggssssssessstkkkyv 53
 LDI~~EIatYRkLLEGEEsrl+s+++++s++~~ s++sg +sss+s++ss+++v+v
 Target 384 LDIEIATYRK~~LLEGEEsrlESGMQNSIHTkttsgyagglssayggltspgl~~syslgsSFGSGAGSSFSRTSSSRAVVV 463
 PP 9*****
 Query 5 kteetrdgkvysesee~~lpk~~ 73
 k++etrdgk+vsess++1pk
 Target 46 KKIE~~TRDGKVSESSDVLPK~~ 483
 PP *****997



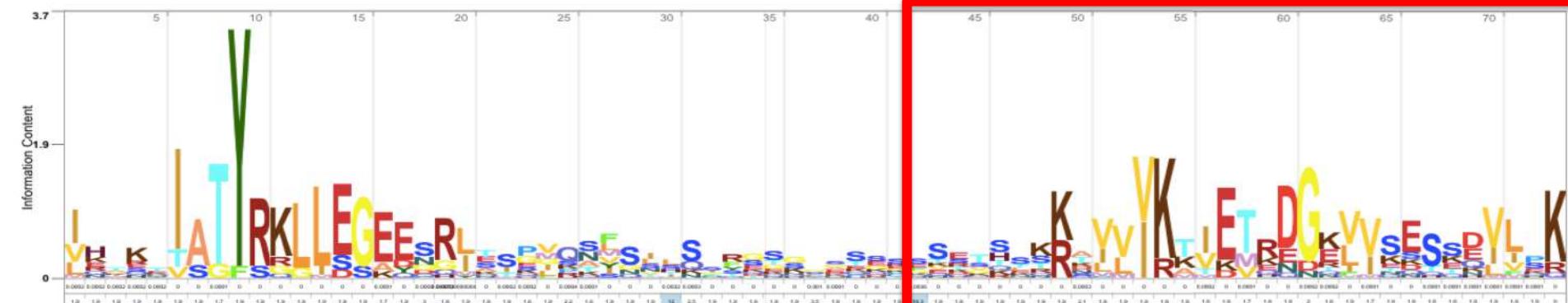
>sp|P05787|K2C8_HUMAN Keratin, type II cytoskeletal 8 OS=Homo sapiens OX=9606 GN=KRT8 PE=1 SV=7
MSIRVTQKS YKVSTSGPRAFSSRSY TSGPGSRI SSSFSRVGSSNFRGGLGGGYGGASGM
GGITAVTVNQS LLSPLVLEVDPNIQAVRTQEKEQIKTLNNKFASFIDKVRFLEQQNKMLE
TKWSLLQQQKTARSNMDNMFESYINNLRRQLETLGQEKLKLEAE LGNMQGLVEDFKNKYE
DEINKRTEMENE FVLIKKDVDEAYMNKVELESRLEGLTDEINFLRQLYEEEIRELQSQIS
DTSVVLSMDNSRS LDMD SIIAEVKAQYEDIANRSRAEAESMYQIKYEELQSLAGKHGDDL
RRTKTEISEMNRNISRLQAEIEGLKGQRASLEAAIAADEQRGELA IKDANAKLSELEAAL
QRAKQDMARQLREYQELMNVKLALDIEIATYRKLLGEES

FSRTSSSRAVVVKIETRDGKLVSESSDV

LPK

KRT8-tail-short.fasta

Model Position





HMMER

Biosequence analysis using profile hidden Markov Models

[Home](#) [Search](#) [Results](#) [Software](#) [Help](#) [About](#) [Contact](#)

Next release within a week, think about downloading your results

phmmmer hmmscan hmmsearch

jackhmmer

iterative search vs protein sequence database

[Paste a Sequence or an Alignment](#) | [Upload a File](#) | [Accession Search](#)

Upload a file

KRT8-tail-short.fasta

▼ Sequence Database ?

Frequently used databases:

Current database selection:

▾

[Return to the Results Summary](#) | [Jump to threshold](#)

Sequence selection

- above threshold
 unselect all

Continue the search

Start iteration 2

Sequence Matches and Features

Pfam == 32

disorder coiled-coil tm & signal peptide

Show hit details

Distribution of Significant Hits



■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

« First « Previous Page 1 of 7 Next » Last »

Target	Description	Species	Cross-references	E-value
> G1S7N4_NOMLE	IF rod domain-containing protein	Nomascus leucogenys		3.6e-11

Let's take a 5 mins break



JACKHMMER Results

[Search Again](#)

Jackhmmer Summary

Iteration	Results	Hits				Total
		New	Lost <small>?</small>	Dropped <small>?</small>		
1	D70C1804-0410-11EA-B37B-C972F75AEC3D.1	+212	-	-	-	212
2	D70C1804-0410-11EA-B37B-C972F75AEC3D.2	+159	-	-	-	371
3	D70C1804-0410-11EA-B37B-C972F75AEC3D.3	+354	-1	1	-	723
4	D70C1804-0410-11EA-B37B-C972F75AEC3D.4	+284	-	3	-	1004
5	D70C1804-0410-11EA-B37B-C972F75AEC3D.5	+45	-	-	-	1049
6	D70C1804-0410-11EA-B37B-C972F75AEC3D.6	+42	-	-	-	1091
7	D70C1804-0410-11EA-B37B-C972F75AEC3D.7	+7	-	-	-	1098
8	D70C1804-0410-11EA-B37B-C972F75AEC3D.8	+24	-	-	-	1122
9	D70C1804-0410-11EA-B37B-C972F75AEC3D.9	+2	-	-	-	1124
10	D70C1804-0410-11EA-B37B-C972F75AEC3D.10	-	-	-	-	1124

Your search has converged. No more iterations will be run.

JACKHMMER Results

Search Again

Jackhmmer Summary

Iteration	Results	Hits				Total
		New	Lost <small>?</small>	Dropped <small>?</small>		
1	D70C1804-0410-11EA-B37B-C972F75AEC3D.1	+212	-	-	-	212
2	D70C1804-0410-11EA-B37B-C972F75AEC3D.2	+159	-	-	-	371
3	D70C1804-0410-11EA-B37B-C972F75AEC3D.3	+354	-1	1	-	723
4	D70C1804-0410-11EA-B37B-C972F75AEC3D.4	+284	-	3	-	1004
5	D70C1804-0410-11EA-B37B-C972F75AEC3D.5	+45	-	-	-	1049
6	D70C1804-0410-11EA-B37B-C972F75AEC3D.6	+42	-	-	-	1091
7	D70C1804-0410-11EA-B37B-C972F75AEC3D.7	+7	-	-	-	1098
8	D70C1804-0410-11EA-B37B-C972F75AEC3D.8	+24	-	-	-	1122
9	D70C1804-0410-11EA-B37B-C972F75AEC3D.9	+2	-	-	-	1124
10	D70C1804-0410-11EA-B37B-C972F75AEC3D.10	-	-	-	-	1124

Your search has converged. No more iterations will be run.

JACKHMMER Results

Search Again

Jackhmmer Summary

Iteration	Results	Hits				Total
		New	Lost ⓘ	Dropped ⓘ		
1	D70C1804-0410-11EA-B37B-C972F75AEC3D.1	+212	-	-	-	212
2	D70C1804-0410-11EA-B37B-C972F75AEC3D.2	+159	-	-	-	371
3	D70C1804-0410-11EA-B37B-C972F75AEC3D.3	+354	-1	1	-	723
4	D70C1804-0410-11EA-B37B-C972F75AEC3D.4	+284	-	3	-	1004
5	D70C1804-0410-11EA-B37B-C972F75AEC3D.5	+45	-	-	-	1049
6	D70C1804-0410-11EA-B37B-C972F75AEC3D.6	+42	-	-	-	1091
7	D70C1804-0410-11EA-B37B-C972F75AEC3D.7	+7	-	-	-	1098
8	D70C1804-0410-11EA-B37B-C972F75AEC3D.8	+24	-	-	-	1122
9	D70C1804-0410-11EA-B37B-C972F75AEC3D.9	+2	-	-	-	1124
10	D70C1804-0410-11EA-B37B-C972F75AEC3D.10	-	-	-	-	1124

Your search has converged. No more iterations will be run.

Jackhmmer Summary

Iteration	Results	Hits				Total
		New	Lost	Dropped	Total	
1	D70C1804-0410-11EA-B37B-C972F75AEC3D.1	+212	-	-	212	
2	D70C1804-0410-11EA-B37B-C972F75AEC3D.2	+159	-	-	371	
3	D70C1804-0410-11EA-B37B-C972F75AEC3D.3	+354	-1	1	723	
4	D70C1804-0410-11EA-B37B-C972F75AEC3D.4	+284	-	3	1004	
5	D70C1804-0410-11EA-B37B-C972F75AEC3D.5	+45	-	-	1049	
6	D70C1804-0410-11EA-B37B-C972F75AEC3D.6	+42	-	-	1091	
7	D70C1804-0410-11EA-B37B-C972F75AEC3D.7	+7	-	-	1098	
8	D70C1804-0410-11EA-B37B-C972F75AEC3D.8	+24	-	-	1122	
9	D70C1804-0410-11EA-B37B-C972F75AEC3D.9	+2	-	-	1124	
10	D70C1804-0410-11EA-B37B-C972F75AEC3D.10	-	-	-	1124	

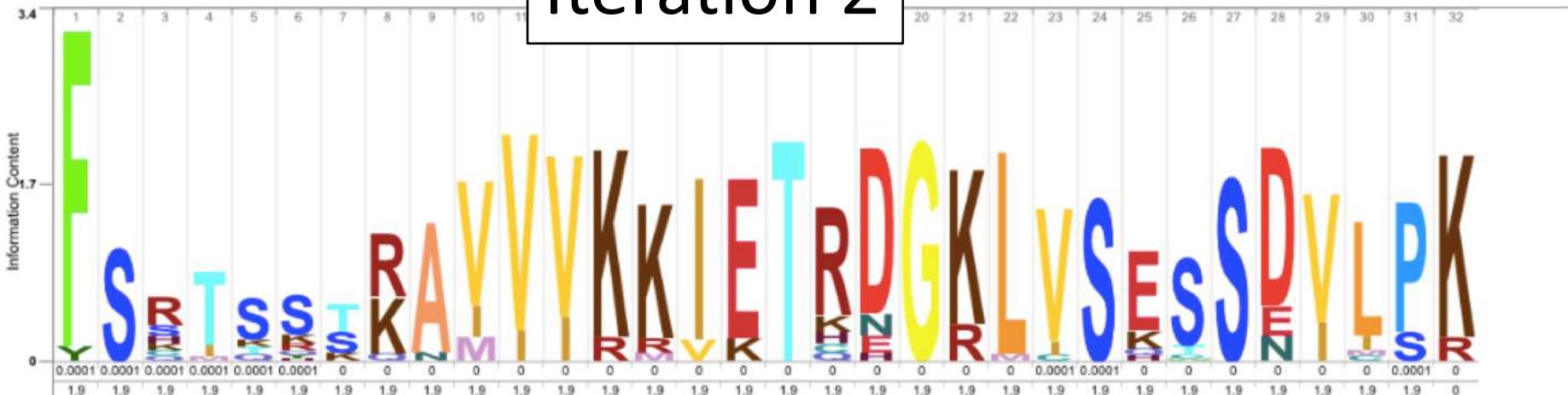


Let's have a look

Your search has converged. No more iterations will be run.

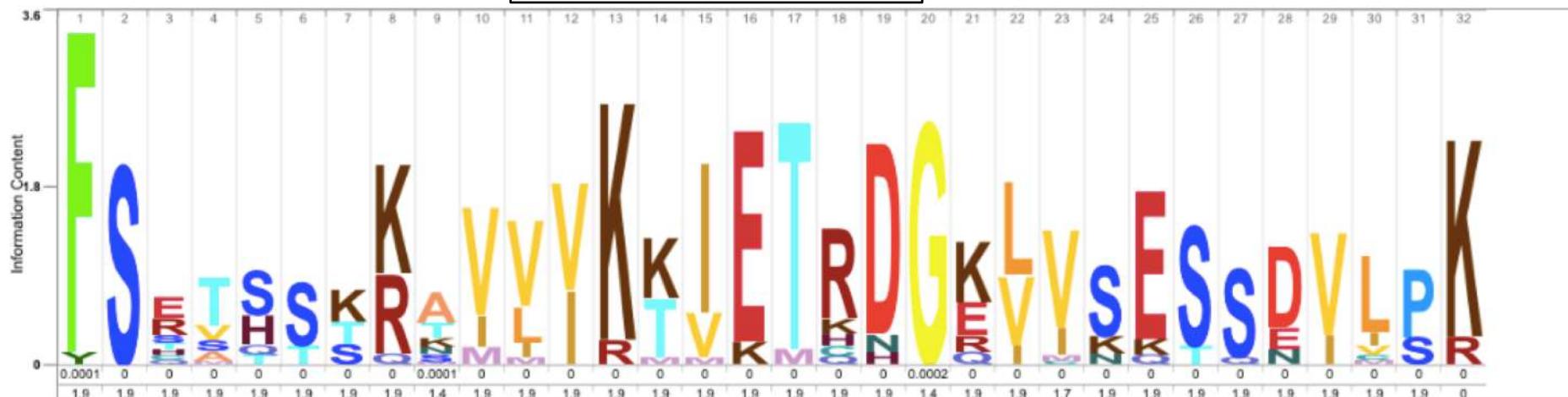
Target	Description	Species	Cross-references	E-value
> A0A2I3RRX0_PANTR	IF rod domain-containing protein	Pan troglodytes	xxx	4.2e-17
> A0A2I3T2G2_PANTR	IF rod domain-containing protein	Pan troglodytes	xxx	4.2e-17
> L9KZY6_TUPCH	Keratin, type II cytoskeletal 8	Tupaia chinensis	xxx	4.3e-17
> A0A096NMK7_PAPAN	IF rod domain-containing protein	Papio anubis	xxx	4.4e-17
> A0A2I3M7V3_PAPAN	IF rod domain-containing protein	Papio anubis	xxx	6.9e-17
> G1S7N4_NOMLE	IF rod domain-containing protein	Nomascus leucogenys	xxx	7.6e-17
> F7GPG9_MACMU	IF rod domain-containing protein	Macaca mulatta	xxx	7.6e-17
> G3S7K2_GORGO	IF rod domain-containing protein	Gorilla gorilla gorilla	xx	8.2e-17
> A0A2I2YMN1_GORGO	IF rod domain-containing protein	Gorilla gorilla gorilla	xx	8.2e-17
> A0A2R9AM79_PANPA	IF rod domain-containing protein	Pan paniscus	xxx	8.3e-17
> A0A2R9AK97_PANPA	IF rod domain-containing protein	Pan paniscus	xxx	8.4e-17
> A0A2K6E4R7_MACNE	IF rod domain-containing protein	Macaca nemestrina	xxx	9.0e-17
> G3SEY3_GORGO	Keratin 8	Gorilla gorilla gorilla	xxx	9.1e-17
> K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens	xxx	9.2e-17
> A0A2R9A6T6_PANPA	IF rod domain-containing protein	Pan paniscus	xxx	9.2e-17
> A0A2J8KDC3_PANTR	KRT8 isoform 11	Pan troglodytes	xxx	9.2e-17
> A0A2K6NVV1_RHIRO	IF rod domain-containing protein	Rhinopithecus roxellana	xxx	9.3e-17
> A0A2K6BEC7_MACNE	IF rod domain-containing protein	Macaca nemestrina	xxx	9.3e-17

Model

Iteration 2

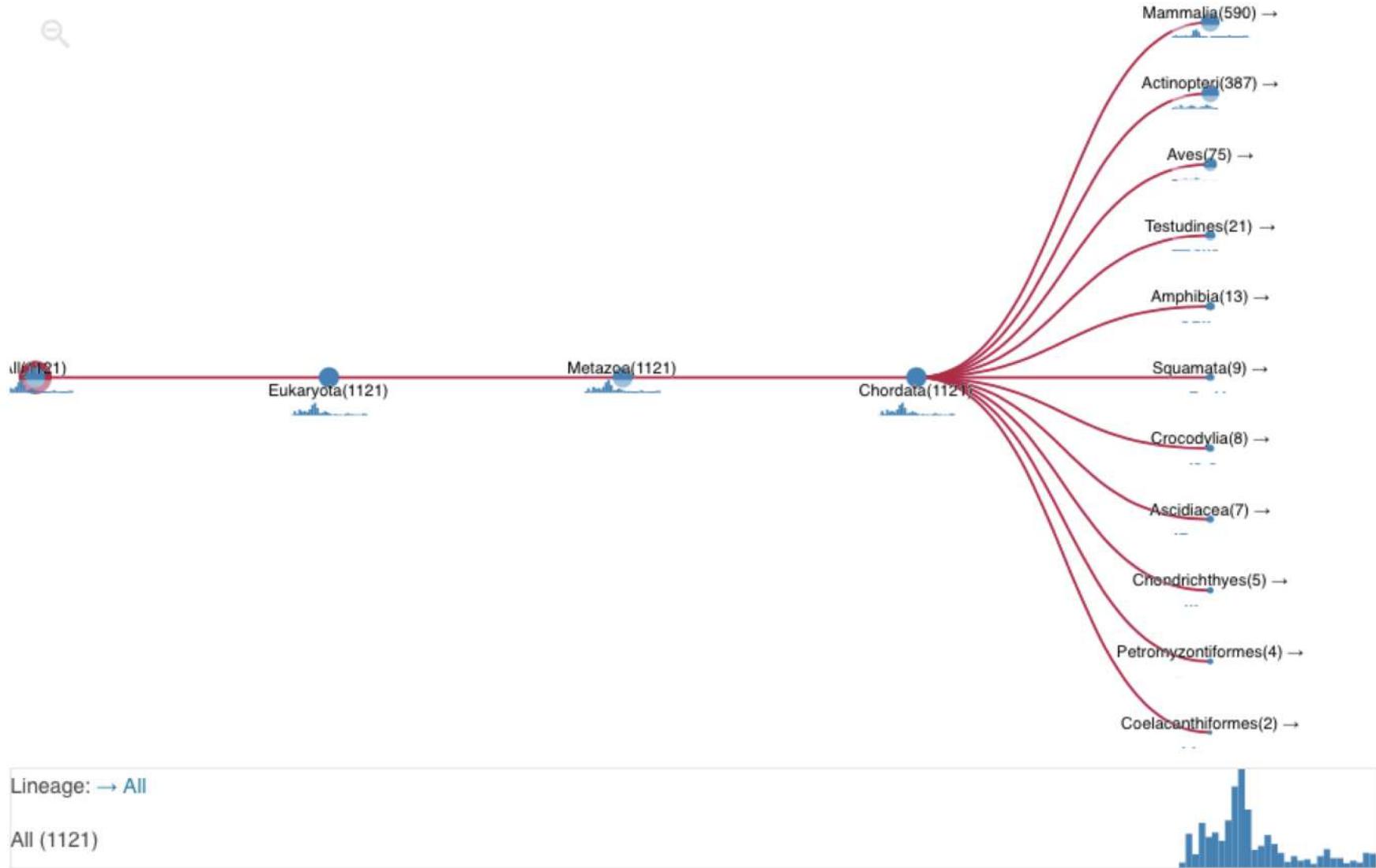
Column number:

Mod

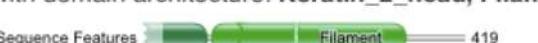
Iteration 10

Column number:

Taxonomic distribution of all search hits

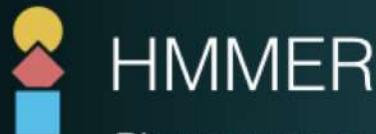


Domain Architectures

703 SEQUENCES	with domain architecture: Filament_head, Filament , example: A0A3B4EA97_PYGNA	View Scores
Show All		
217 SEQUENCES	with domain architecture: Keratin_2_head, Filament , example: G3SEY3_GORGO	View Scores
Show All		
116 SEQUENCES	with domain architecture: Filament , example: A0A096NMK7_PAPAN	View Scores
Show All		
25 SEQUENCES	with domain architecture: Keratin_2_head, Filament, Filament , example: A0A2I3M7V3_PAPAN	View Scores
Show All		
19 SEQUENCES	with domain architecture: Filament_head, Filament, Filament , example: A0A4U5U033_COLLU	View Scores
Show All		
12 SEQUENCES	with no domain architecture, example: A0A402E9P6_9SAUR	View Scores
Show All		
12 SEQUENCES	with domain architecture: Filament, Filament , example: L9KZY6_TUPCH	View Scores
Show All		
7 SEQUENCES	with domain architecture: Pkinase, Keratin_2_head, Filament, Filament , example: G3REJ4_GORGO	View Scores
Show All		
2 SEQUENCES	with domain architecture: Keratin_2_head, Filament, Filament, Filament , example: A0A2I3RRX0_PANTR	View Scores
Show All		

2 SEQUENCES	with domain architecture: Filament_head, Filament, Filament, Filament , example: A0A3L8S059_CHLGU	View Scores
	Sequence Features  499	
Show All		
1 SEQUENCE	with domain architecture: P34-Arc, Filament , example: M7BB29_CHEMY	View Scores
	Sequence Features  707	
1 SEQUENCE	with domain architecture: Keratin_2_head , example: H0XMA4_OTOGA	View Scores
	Sequence Features  347	
1 SEQUENCE	with domain architecture: Filament, I-set , example: M7AQ77_CHEMY	View Scores
	Sequence Features  207	
1 SEQUENCE	with domain architecture: Filament_head, Filament, I-set , example: A0A3M0KZ98_HIRRU	View Scores
	Sequence Features  570	
1 SEQUENCE	with domain architecture: Filament_head, Filament, Retrotrans_gag , example: A0A0P7UGQ3_SCLFO	View Scores
	Sequence Features  553	
1 SEQUENCE	with domain architecture: PKD, Filament_head, Filament , example: A0A2U9BGS4_SCOMX	View Scores
	Sequence Features  1216	

[Search Details](#)



Biosequence analysis using profile hidden Markov Models

Home | Search | Results | Software | Help | About | Contact

Next release within a week, think about downloading your results

JACKHMMER Results

[Search Again](#)

Score

Taxonomy

Domain

[Download](#)

Click on DOWNLOAD

Perform a new search:

- with new input
- with these results

Iteration 10 Converged

Your jackhmmer search has reached convergence. No further iterations will be performed.

[Return to the Results Summary](#) | [Jump to threshold](#)

[« previous iteration](#)

Distribution of Significant Hits



■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

[« First](#) [« Previous](#) **Page 1** of 24 [Next »](#) [Last »](#)

Significant Query Matches (1124) in *uniprotrefprot* (v.2019_09)

[Customise](#)

Target	Description	Species	Cross-references	E-value
> A0A2I3RRX0_PANTR	IF rod domain-containing protein	Pan troglodytes		4.2e-17



JACKHMMER Results

[Score](#)[Taxonomy](#)[Domain](#)[Download](#)

- **Job:** CE12EB7E-14BE-11EA-90B6-D5BAF75AEC3D.10
- **Started:** 2019-12-02 04:48:09
- **Algorithm:** hmmsearch
- **HMMER Options:** -E 1 --domE 1 --incE 0.01 --incdomE 0.03 --mx BLOSUM62 --pextend 0.4 --popen 0.02 --seqdb uniprotrefprot

▼ Format

Text

A plain text file containing the hit alignments and scores.



Tab Delimited

A tab delimited text file containing the hit information. No alignments.



XML

An XML file formatted for machine parsing of the data.



JSON

All the results information encoded as a single JSON string.



HMM (v3.0 format)

The profile HMM generated from the uploaded multiple sequence alignment



HMM (v3.1 format)

The profile HMM generated from the uploaded multiple sequence alignment.



FASTA

Download the significant hits from your search as a gzipped FASTA file.



Full length FASTA

A gzipped file containing the full length sequences for significant search hits.



Aligned FASTA

A gzipped file containing aligned significant search hits in FASTA format.



STOCKHOLM

Download an alignment of significant hits as a gzipped STOCKHOLM file.



ClustalW

Download an alignment of significant hits as a gzipped ClustalW file.



PSI-BLAST

Download an alignment of significant hits as a gzipped psiblast file.





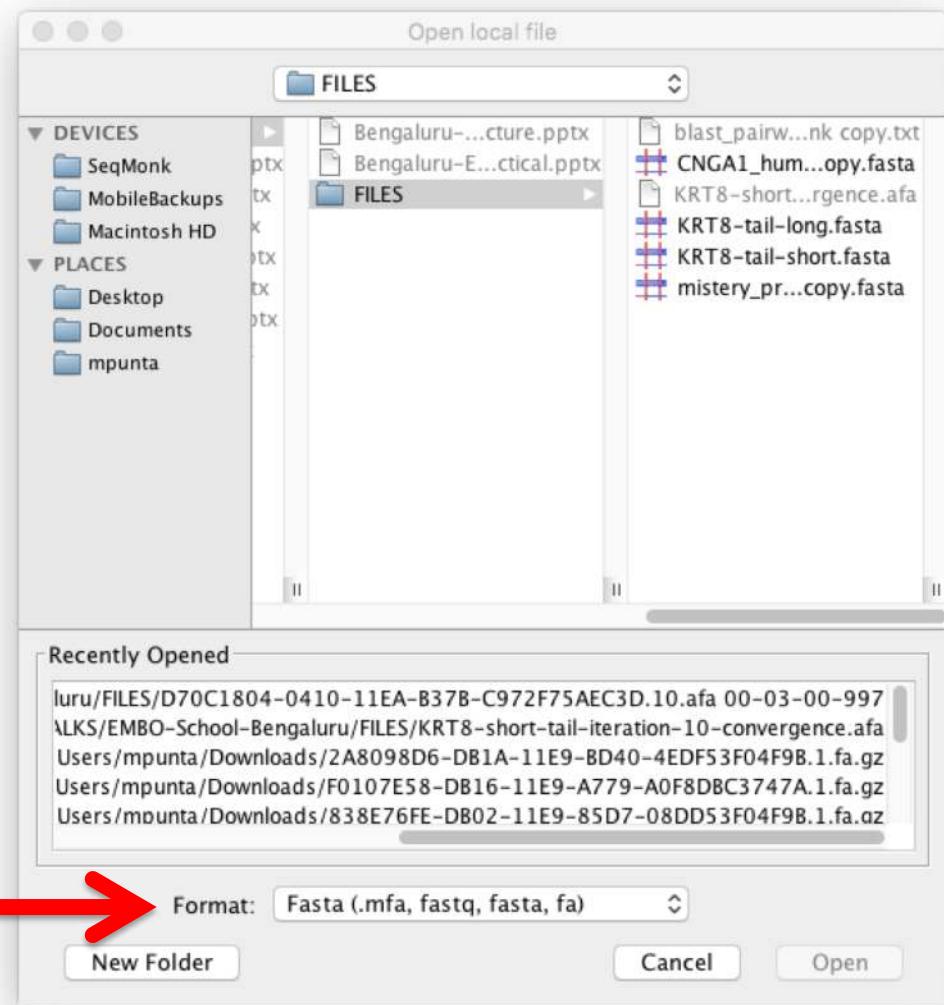
Jalview File Tools Vamsas Help Window



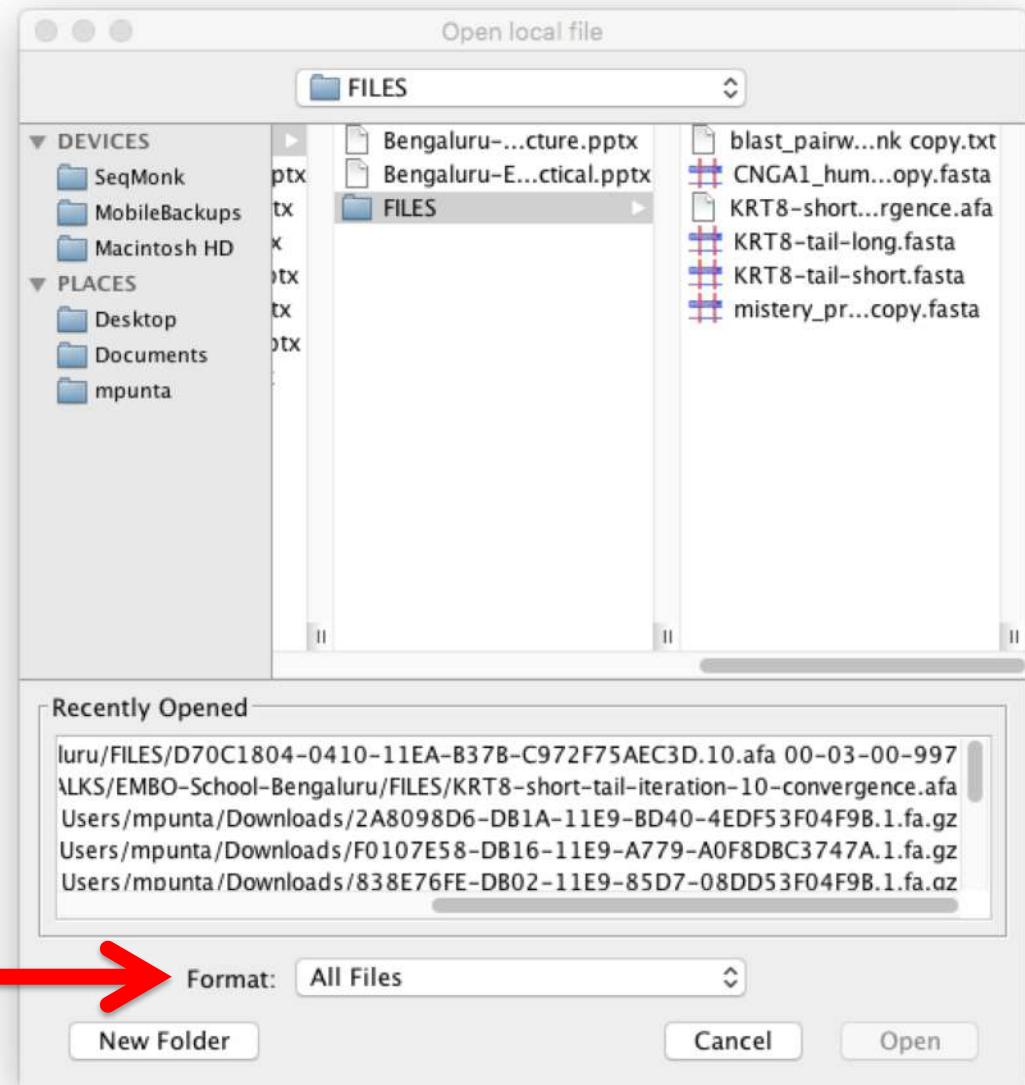
Jalview 2.10.1

Click: File-> Input Alignment -> From File

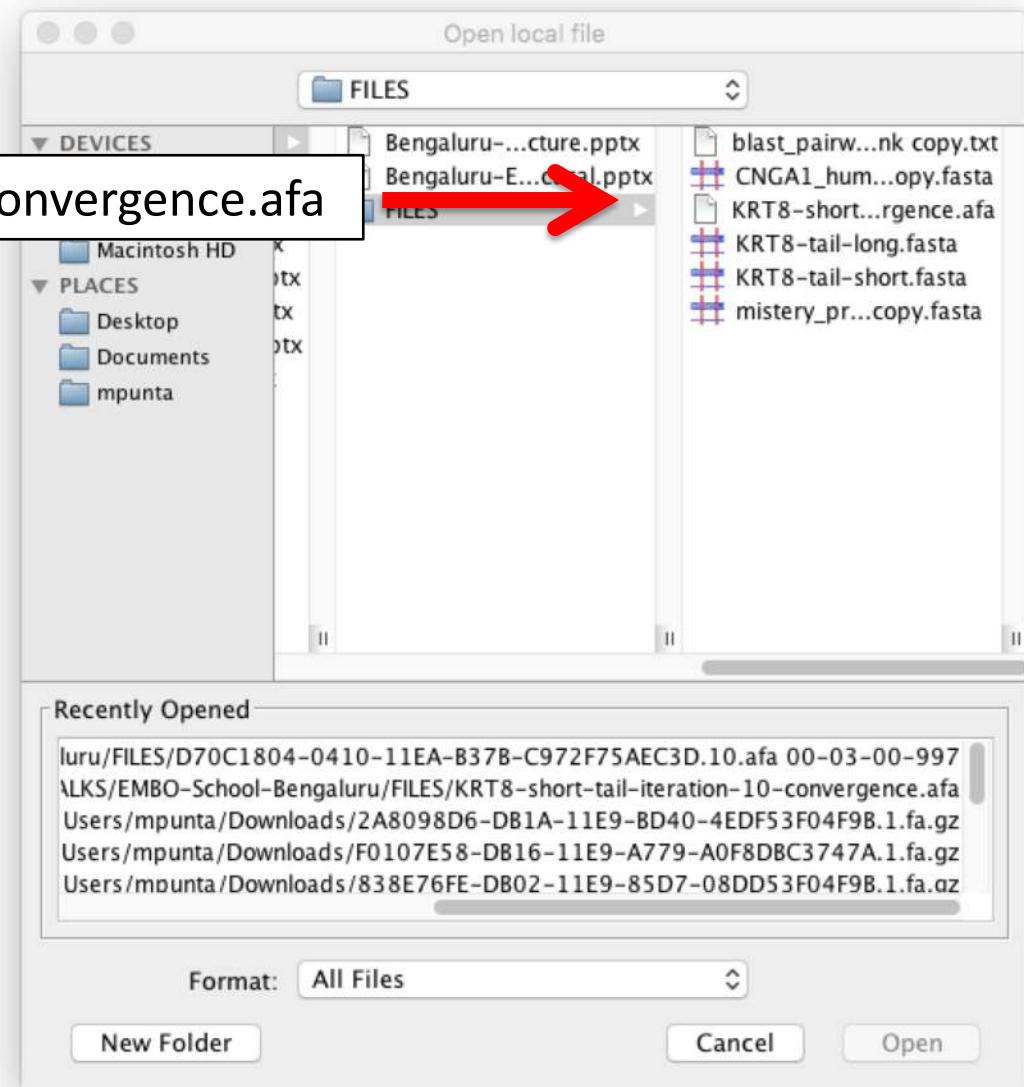
KRT8-short-tail-iteration-10-convergence.afa



Click



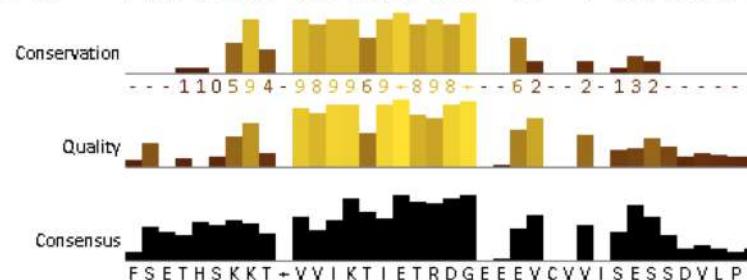
Select KRT8-short-tail-iteration-10-convergence.afa



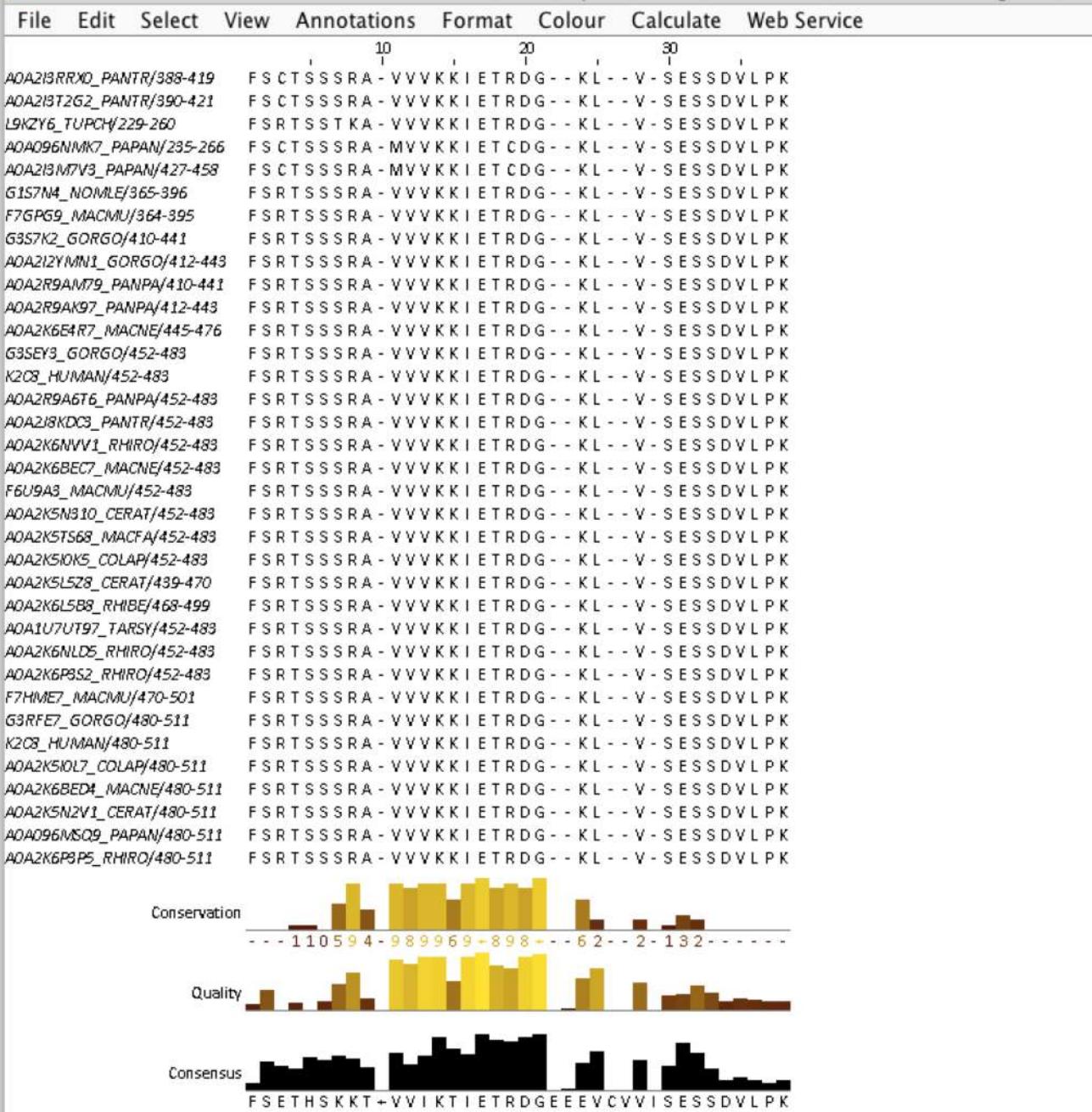
/Users/mpunta/TALKS/2019-TALKS/EMBO-School-Bengaluru/FILES/KRT8-short-tail-iterat...

File Edit Select View Annotations Format Colour Calculate Web Service

	10	20	30
ADA2I3RRX0_PANTR/388-419	F S C T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
ADA2I3T2G2_PANTR/390-421	F S C T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
L9KZY6_TUPCH/229-260	F S R T S S S T K A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
ADA096NMK7_PAPAN/235-266	F S C T S S S R A - M V V K K I E T C D G - - K L - - V - S E S S D V L P K		
ADA2I3M7V3_PAPAN/427-458	F S C T S S S R A - M V V K K I E T C D G - - K L - - V - S E S S D V L P K		
G15TM4_NOMLE/365-396	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
F7GPG9_MACMU/364-395	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
G35TK2_GORGO/410-441	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
ADA2I2YMN1_GORGO/412-443	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
ADA2R9AM79_PANPA/410-441	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
ADA2R9AK97_PANPA/412-443	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
ADA2K6E4R7_MACNE/445-476	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		
G3SEY3_GORGO/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K		



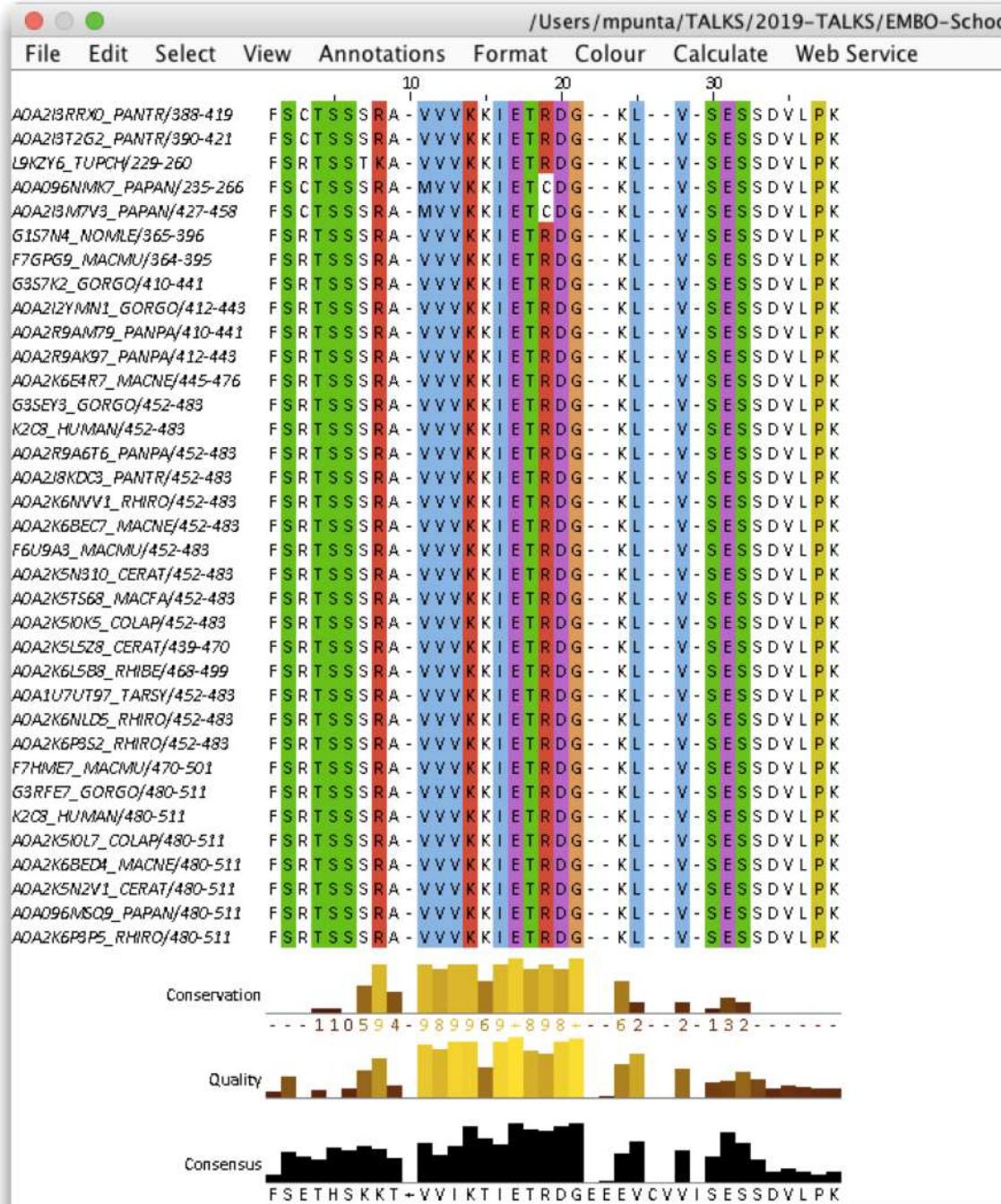
Sequence 1 ID: ADA2I3RRX0_PANTR Residue: THR (391)



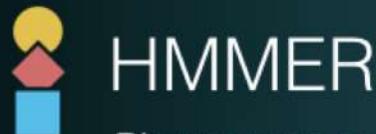
File Edit Select View Annotations Format Colour Calculate Web Service

		10	20	30	
AOA213RRX0_PANTR/388-419	F S C T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA213T2G2_PANTR/390-421	F S C T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
L9KZY6_TUPCH/229-260	F S R T S S S T K A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA096NMK7_PAPAN/235-266	F S C T S S S R A - M V V V K K I E T C D G - - K L - - V - S E S S D V L P K				
AOA213M7V3_PAPAN/427-458	F S C T S S S R A - M V V V K K I E T C D G - - K L - - V - S E S S D V L P K				
G157N4_NOMLE/365-396	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
F7GP69_MACMU/364-395	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
G357K2_GORGO/410-441	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA212YMN1_GORGO/412-443	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2R9AM79_PANPA/410-441	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2R9AK97_PANPA/412-443	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K6E4R7_MACNE/445-476	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
G3SEY3_GORGO/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
K203_HUMAN/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2R9A6T6_PANPA/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA218KDC3_PANTR/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K6NNV1_RHIRO/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K68EC7_MACNE/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
F6U9A3_MACMU/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K5N3I0_CERAT/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K5TS68_MACFA/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K5I0K5_COLAP/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K5L5Z8_CERAT/439-470	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K6L5B8_RHIBD/468-499	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA1U7UT97_TARSY/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K6NL05_RHIRO/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K6P6S2_RHIRO/452-483	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
F7HME7_MACMU/470-501	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
G3RFE7_GORGO/480-511	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
K203_HUMAN/480-511	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K5I0L7_COLAP/480-511	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K6BED4_MACNE/480-511	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K5N2V1_CERAT/480-511	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA096MSQ9_PAPAN/480-511	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				
AOA2K6P6P5_RHIRO/480-511	F S R T S S S R A - V V V K K I E T R D G - - K L - - V - S E S S D V L P K				





Sequence 11 ID: A0A2R9AK97_PANPA Residue: ASP (430)



Biosequence analysis using profile hidden Markov Models

Home | Search | Results | Software | Help | About | Contact

Next release within a week, think about downloading your results

JACKHMMER Results

Click on WITH THESE RESULTS

Score Taxonomy Domains

Iteration 10 Converged

Your jackhmmer search has reached convergence. No further iterations will be performed.

[Return to the Results Summary](#) | [Jump to threshold](#)

[« previous iteration](#)

Distribution of Significant Hits



■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

[« First](#) [« Previous](#) **Page 1** of 24 [Next »](#) [Last »](#)

Significant Query Matches (1124) in *uniprotrefprot* (v.2019_09)

[Customise](#)

Target	Description	Species	Cross-references	E-value
> A0A2I3RRX0_PANTR	IF rod domain-containing protein	Pan troglodytes		4.2e-17



HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

Next release within a week, think about downloading your results

phmmer hmmsearch **hmmer** jackhmmer

protein alignment/profile-HMM vs protein sequence database

Using Existing Search as Input



This search will be run with an HMM generated from the **1124** hits above threshold in your previous search D70C1804-0410-11EA-B37B-C972F75AEC3D.10.

Submit

Reset

▼ Sequence Database

Frequently used databases

Click on PDB



Via Prot PDB Ensembl

Current database selection:

PDB

► Restrict by Taxonomy



HMMER

Biosequence analysis using profile hidden Markov Models

[Home](#) [Search](#) [Results](#) [Software](#) [Help](#) [About](#) [Contact](#)

Next release within a week, think about downloading your results

[Search Again](#)

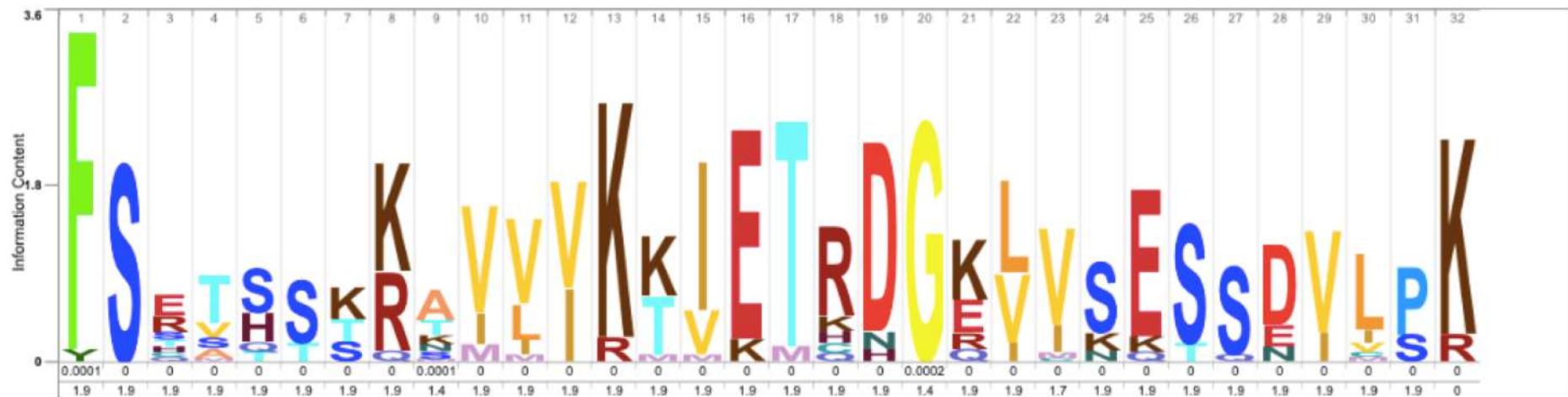
HMMSEARCH Results

[Score](#) [Taxonomy](#) [Domain](#) [Download](#)

No hits were found for your query.

Search Details

Model Position



Column number:



HMMER

Biosequence analysis using profile hidden Markov Models

Home Search Results Software Help About Contact

Next release within a week, think about downloading your results

phmmmer hmmscan **hmmsearch** jackhmmer

protein alignment/profile-HMM vs protein sequence database

Using Existing Search as Input



This search will be run with an HMM generated from the **1124** hits above threshold in your previous search [D70C1804-0410-11EA-B37B-C972F75AEC3D.10](#).

Submit

Reset

▼ Sequence Database ?

Frequent

Click on SwissProt

rotKB

SwissProt

DB

Ensembl

Current database selection:

SwissProt

► Restrict by Taxonomy ?

Next release within a week, think about downloading your results

HMMSEARCH Results

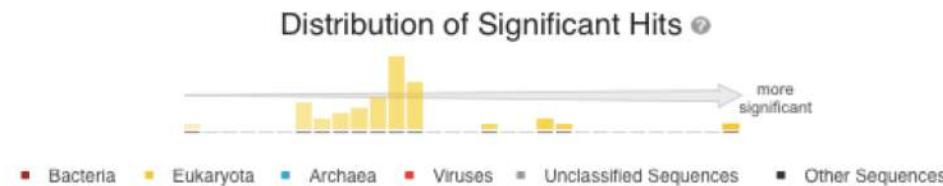
Search Again

Score

Taxonomy

Domain

Download



	Target	Description	Species	ⓘ Cross-references	E-value
>	K2C8_HUMAN ⓘ	Keratin, type II cytoskeletal 8	Homo sapiens ⓘ		1.0e-18
>	K2C8_BOVIN ⓘ	Keratin, type II cytoskeletal 8	Bos taurus ⓘ		4.1e-14
>	K2C8_RAT ⓘ	Keratin, type II cytoskeletal 8	Rattus norvegicus ⓘ		1.9e-13
>	K2C8_MOUSE ⓘ	Keratin, type II cytoskeletal 8	Mus musculus ⓘ		1.9e-13
>	VIM1_XENLA ⓘ	Vimentin-1/2	Xenopus laevis ⓘ		1.7e-11
>	DESM_MESAU ⓘ	Desmin	Mesocricetus auratus ⓘ		9.8e-10
>	DESM_XENLA ⓘ	Desmin	Xenopus laevis ⓘ		9.8e-10
>	DESM_HUMAN ⓘ	Desmin	Homo sapiens ⓘ		1.0e-09
>	DESM_RAT ⓘ	Desmin	Rattus norvegicus ⓘ		1.0e-09
>	DESM_BOVIN ⓘ	Desmin	Bos taurus ⓘ		1.0e-09
>	DESM_MOUSE ⓘ	Desmin	Mus musculus ⓘ		1.0e-09



■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

Significant Query Matches (49) in swissprot (v.2019_09)

Customise

	Target	Description	Species	Cross-references	E-value
>	K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens		1.0e-18
>	K2C8_BOVIN	Keratin, type II cytoskeletal 8	Bos taurus		4.1e-14
>	K2C8_RAT	Keratin, type II cytoskeletal 8	Rattus norvegicus		1.9e-13
>	K2C8_MOUSE	Keratin, type II cytoskeletal 8	Mus musculus		1.9e-13
>	VIM1_XENLA	Vimentin-1/2	Xenopus laevis		1.7e-11
>	DESM_MESAU	Desmin	Mesocricetus auratus		9.8e-10
>	DESM_XENLA	Desmin	Xenopus laevis		9.8e-10

Click on arrow next to DESM_HUMAN to show alignment

Query	Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)	Bit Score	E-value	
	start	end	start	end						Ind.	Cond.
	2	25	438	463	438	461	4.60	0.97	62.5 (15)	100.0 (24)	41.8
										2.1e-09	1.9e-13

Query 2 SetsskkaVvvKkIETrDGklVSE 25
Se+++kk+V++K+IETrDG++VSE
Target 438 SEVHTKKTVMIKTIETRDGEVVSE 461
PP 89*****9

>	DESM_RAT	Desmin	Rattus norvegicus		1.0e-09
---	----------	--------	-------------------	--	---------



■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

Significant Query Matches (49) in swissprot (v.2019_09)

Customise

	Target	Description	Species	Cross-references	E-value
>	K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens		1.0e-18
>	K2C8_BOVIN	Keratin, type II cytoskeletal 8	Bos taurus		4.1e-14
>	K2C8_RAT	Keratin, type II cytoskeletal 8	Rattus norvegicus		1.9e-13
>	K2C8_MOUSE	Keratin, type II cytoskeletal 8	Mus musculus		1.9e-13
>	VIM1_XENLA	Vimentin-1/2	Xenopus laevis		1.7e-11
>	DESM_MESAU	Desmin	Mesocricetus auratus		9.8e-10
>	DESM_XENLA	Desmin	Xenopus laevis		9.8e-10
>	DESM_HUMAN	Desmin			1.0e-09

Click on DESM_HUMAN to open UniProt page

Query	Target	Length	Start	End	Start	End	Bias	Accuracy	% Identity (count)	% Similarity (count)	Bit Score	E-value	
start	end	start	end	start	end						Ind.	Cond.	
2	25	438	463	438	461	4.60	0.97	62.5 (15)	100.0 (24)	41.8	2.1e-09	1.9e-13	

.....*.....*.....
 Query 2 SetsskkaVvVkkIETrDGklVSE 25
 Se+++kk+V++K+IETrDG++VSE
 Target 438 SEVHTKKTVMIKTIETRDGEVVSE 461
 PP 89*****9

>	DESM_RAT	Desmin	Rattus norvegicus		1.0e-09
---	----------	--------	-------------------	--	---------

UniProtKB - P17661 (DESM_HUMAN)

Display

[BLAST](#) [Align](#) [Format](#) [Add to basket](#) [History](#)

Entry

Publications

Feature viewer

Feature table

None
<input checked="" type="checkbox"/> Function
<input checked="" type="checkbox"/> Names & Taxonomy
<input checked="" type="checkbox"/> Subcellular location
<input checked="" type="checkbox"/> Pathology & Biotech
<input checked="" type="checkbox"/> PTM / Processing
<input checked="" type="checkbox"/> Expression
<input checked="" type="checkbox"/> Interaction
<input checked="" type="checkbox"/> Structure
<input checked="" type="checkbox"/> Family & Domains
<input checked="" type="checkbox"/> Sequence
<input checked="" type="checkbox"/> Similar proteins
<input checked="" type="checkbox"/> Cross-references
<input checked="" type="checkbox"/> Entry information
<input checked="" type="checkbox"/> Miscellaneous

▲ Top

Protein | Desmin

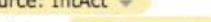
Gene | DES

Organism | Homo sapiens (Human)

Status |  Reviewed - Annotation score:  - Experimental evidence at protein levelⁱFunctionⁱ

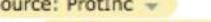
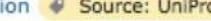
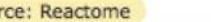
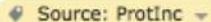
Muscle-specific type III intermediate filament essential for proper muscular structure and function. Plays a crucial role in maintaining thick disks and forming the myofibrils, linking them not only to the sarcolemmal cytoskeleton, but also to the nucleus and mitochondria, thus (PubMed:25358400). In adult striated muscle they form a fibrous network connecting myofibrils to each other and to the plasma membrane (PubMed:24200904, PubMed:25394388, PubMed:26724190). May act as a sarcomeric microtubule-anchoring protein: specifically associated with buckled microtubules and mechanical resistance to contraction. Contributes to the transcriptional regulation of the NKX2-5 gene in cardiac myogenesis and in cardiac side population stem cells in the adult. Plays a role in maintaining an optimal conformation of nebulin to recruit cardiac alpha-actin (By similarity).   

GO - Molecular functionⁱ

- cytoskeletal protein binding 
- identical protein binding 
- structural constituent of cytoskeleton 

Complete GO annotation on QuickGO ...

GO - Biological processⁱ

- cytoskeleton organization 
- intermediate filament organization 
- muscle contraction 
- muscle filament sliding 
- regulation of heart contraction 

Complete GO annotation on QuickGO ...

Keywordsⁱ

Molecular function

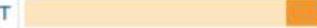
Muscle protein

Family & Domainsⁱ

Domains and Repeats

Feature key	Position(s)	Description	Actions	Graphical view	Length
Domain ⁱ	108 – 416	IF rod PROSITE-ProRule annotation	Add BLAST		309

Region

Feature key	Position(s)	Description	Actions	Graphical view	Length
Region ⁱ	1 – 108	Head	Add BLAST		108
Region ⁱ	109 – 141	Coil 1A	Add BLAST		33
Region ⁱ	142 – 151	Linker 1			10
Region ⁱ	152 – 252	Coil 1B	Add BLAST		101
Region ⁱ	253 – 268	Linker 12	Add BLAST		16
Region ⁱ	268 – 415	Interaction with NEB 1 Publication	Add BLAST		148
Region ⁱ	269 – 287	Coil 2A	Add BLAST		19
Region ⁱ	288 – 295	Linker 2			8
Region ⁱ	296 – 412	Coil 2B	Add BLAST		117
Region ⁱ	413 – 470	Tail	Add BLAST		58
Region ⁱ	438 – 453	Interaction with CRYAB 1 Publication	Add BLAST		16

Sequence similaritiesⁱ

Belongs to the intermediate filament family. PROSITE-ProRule annotation

Keywords - Domainⁱ

Coiled coil

Let's take a 5 mins break





■ Bacteria ■ Eukaryota ■ Archaea ■ Viruses ■ Unclassified Sequences ■ Other Sequences

Significant Query Matches (49) in swissprot (v.2019_09)

Customise

	Target	Description	Species	Cross-references		E-value
>	K2C8_HUMAN	Keratin, type II cytoskeletal 8	Homo sapiens			1.0e-18
>	K2C8_BOVIN	Keratin, type II cytoskeletal 8	Bos taurus			4.1e-14
>	K2C8_RAT	Keratin, type II cytoskeletal 8	Rattus norvegicus			1.9e-13
>	K2C8_MOUSE	Keratin, type II cytoskeletal 8	Mus musculus			1.9e-13
>	VIM1_XENLA	Vimentin-1/2	Xenopus laevis			1.7e-11
>	DESM_MESAU	Desmin	Mesocricetus auratus			9.8e-10
>	DESM_XENLA	Desmin	Xenopus laevis			9.8e-10
v	DESM_HUMAN	Desmin	Homo sapiens			1.0e-09

Query		Target Envelope		Target Alignment		Bias	Accuracy	% Identity (count)	% Similarity (count)	Bit Score	E-value	
start	end	start	end	start	end						Ind.	Cond.
2	25	438	463	438	461	4.60	0.97	62.5 (15)	100.0 (24)	41.8	2.1e-09	1.9e-13

Query	2	SetsskkaVvvKkIETrDGklVSE	25
		Se+++kk+V++K+IETrDG++VSE	
Target	438	SEVHTKKTVMIKTIETRDGEVVSE	461
PP		89*****	9

>	DESM_RAT	Desmin	Rattus norvegicus			1.0e-09
---	----------	--------	-------------------	--	--	---------

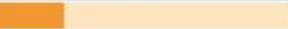
more

Family & Domainsⁱ

Domains and Repeats

Feature key	Position(s)	Description	Actions	Graphical view	Length
Domain ⁱ	108 – 416	IF rod PROSITE-ProRule annotation	Add BLAST		309

Region

Feature key	Position(s)	Description	Actions	Graphical view	Length
Region ⁱ	1 – 108	Head	Add BLAST		108
Region ⁱ	109 – 141	Coil 1A	Add BLAST		33
Region ⁱ	142 – 151	Linker 1			10
Region ⁱ	152 – 252	Coil 1B	Add BLAST		101
Region ⁱ	253 – 268	Linker 12	Add BLAST		16
Region ⁱ	268 – 415	Interaction with NEB 1 Publication	Add BLAST		148
Region ⁱ	269 – 287	Coil 2A	Add BLAST		19
Region ⁱ	288 – 295	Linker 2			8
Region ⁱ	296 – 412	Coil 2B	Add BLAST		117
Region ⁱ	413 – 470	Tail	Add BLAST		58
Region ⁱ	438 – 453	Interaction with CRYAB 1 Publication	Add BLAST		16

Sequence similaritiesⁱ

Belongs to the intermediate filament family. PROSITE-ProRule annotation

Keywords - Domainⁱ

Coiled coil

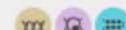
start	end	start	end	start	end					Ind.	Cond.
2	25	438	463	438	461	4.60	0.97	62.5 (15)	100.0 (24)	41.8	2.1e-09

Query	2 SetsskkaVvVvKkIETrDGklVSE 25
	Se+++kk+V++K+IETrDG++VSE
Target	438 SEVHTKKTVMIKTIETRDGEVVSE 461
PP	89*****9

> DESM_RAT

Desmin

Rattus norvegicus



1.0e-09

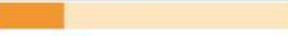
more

Family & Domainsⁱ

Domains and Repeats

Feature key	Position(s)	Description	Actions	Graphical view	Length
Domain ⁱ	108 – 416	IF rod PROSITE-ProRule annotation	Add BLAST		309

Region

Feature key	Position(s)	Description	Actions	Graphical view	Length
Region ⁱ	1 – 108	Head	Add BLAST		108
Region ⁱ	109 – 141	Coil 1A	Add BLAST		33
Region ⁱ	142 – 151	Linker 1			10
Region ⁱ	152 – 252	Coil 1B	Add BLAST		101
Region ⁱ	253 – 268	Linker 12	Add BLAST		16
Region ⁱ	268 – 415	Interaction with NEB 1 Publication	Add BLAST		148
Region ⁱ	269 – 287	Coil 2A	Add BLAST		19
Region ⁱ	288 – 295	Linker 2			8
Region ⁱ	296 – 412	Coil 2B	Add BLAST		117
Region ⁱ	413 – 429	Tail	Add BLAST		58
Region ⁱ	438 – 453	Interaction with CRYAB 1 Publication	Add BLAST		16

Sequence similaritiesⁱ

Belongs to the intermediate filament family. PROSITE-ProRule annotation

Keywords - Domainⁱ

Coiled coil

start	end	start	end	start	end				Ind.	Cond.		
2	25	438	463	438	461	4.60	0.97	62.5 (15)	100.0 (24)	41.8	2.1e-09	1.9e-13

Query	2 *.....*.....*
	SetsskkaVvVvKkIETrDGklVSE 25
	Se+++kk+V++K+IETrDG++VSE
Target	438 SEVHTKKTVMIKTIETRDGEVVSE 461
PP	89*****9

> DESM_RAT

Desmin

Rattus norvegicus



1.0e-09

Family & Domains

Domains and Repeats

Feature key	Position(s)	Description	Actions	Graphical view	Length
Domain ⁱ	108 – 416	IF rod PROSITE-ProRule annotation	Add BLAST		309

Region

Feature key	Position(s)	Description	Actions	Graphical view	Length	
Region ⁱ	1 – 108	Head	Add BLAST		108	
Region ⁱ	109 – 141	Coil 1A	Add BLAST		33	
Region ⁱ	142 – 151	Linker 1			10	
Region ⁱ	152 – 252	Coil 1B	Add BLAST		101	
Region ⁱ	253 – 268	Linker 12	Add BLAST		16	
Region ⁱ	268 – 415	Interaction with NEB 1 Publication	Add BLAST		148	
Region ⁱ	269 – 287	Coil 2A	Add BLAST		19	
Region ⁱ	288 – 295	Linker 2			8	
Region ⁱ	296 – 412	Coil 2B	Add BLAST		117	
Region ⁱ	413 – 470	Tail	Add BLAST		58	
Region ⁱ	438 – 453	Interaction with CRYAB	1 Publication	Add BLAST		16



Manual assertion based on experiment in:

"alphaB-crystallin is a sensor for assembly intermediates and for the subunit topology of desmin intermediate filaments."

Sharma S., Conover G.M., Elliott J.L., Der Perng M., Herrmann H., Quinlan R.A.

Cell Stress Chaperones 22:613-626(2017) [PubMed] [Europe PMC] [Abstract]

Cited for: CHARACTERIZATION OF VARIANTS MFM1 MET-451 AND TRP-454. INTERACTION WITH CRYAB.

Sequence similarities

Belongs to the intermediate filament family. PROSITE-ProRule

Keywords - Domain

Coiled coil

Phylogenomic databases

enNOMI: FN0G410TF71 Fukarvota

αB-crystallin is a sensor for assembly intermediates and for the subunit topology of desmin intermediate filaments.

Sharma S¹, Conover GM², Elliott JL³, Der Perng M⁴, Herrmann H^{1,5}, Quinlan RA⁶.

Author information

Abstract

Mutations in the small heat shock protein chaperone CRYAB (αB-crystallin/HSPB5) and the intermediate filament protein desmin, phenocopy each other causing cardiomyopathies. Whilst the binding sites for desmin on CRYAB have been determined, desmin epitopes responsible for CRYAB binding and also the parameters that determine CRYAB binding to desmin filaments are unknown. Using a combination of co-sedimentation centrifugation, viscometric assays and electron microscopy of negatively stained filaments to analyse the *in vitro* assembly of desmin filaments, we show that the binding of CRYAB to desmin is subject to its assembly status, to the subunit organization within filaments formed and to the integrity of the C-terminal tail domain of desmin. Our *in vitro* studies using a rapid assembly protocol, C-terminally truncated desmin and two disease-causing mutants (I451M and R454W) suggest that CRYAB is a sensor for the surface topology of the desmin filament. Our data also suggest that CRYAB performs an assembly chaperone role because the assembling filaments have different CRYAB-binding properties during the maturation process. We suggest that the capability of CRYAB to distinguish between filaments with different surface topologies due either to mutation (R454W) or assembly protocol is important to understanding the pathomechanism(s) of desmin-CRYAB myopathies.

Alignments

[Download](#) [Graphics](#)

sp|P17661|DESM_HUMAN Desmin OS=Homo sapiens OX=9606 GN=DES PE=1 SV=3
Sequence ID: Query_170373 Length: 470 Number of Matches: 1

“Our results indicate that CRYAB binding most likely occurs between the amino acid residues 442–453 of desmin.”

Sbjct	72	SR+G++	YG	+	A	VNQ	L+	RT	EK	+++	L	
SRLGGTTR----TPSSYAGAGELLDLDFSLADAVNQEFLT-----TRTNEKVELQEL												115

ii. Sensor Function during *in vitro* assembly

A. Coassembly of CRYAB with desmin – modulation of binding by mutants



Sbjct	293	E E W Y K S K V S D L T Q A A N K N N D A L R Q A K Q E M M E Y R H Q I Q S Y T C E I D A L K G T N D S L M R Q M R E L	352
-------	-----	---	-----

Query	339	E Q R G E L A I K D A N A K L S E L E A A L Q R A K Q D M A R Q L R E Y Q E I L M N V K L A L D I E I A T Y R K L L E G E	398
		E R ++ LE ++ K + M A R L R E Y Q + L + N V K + A L D + E I A T Y R K L L E G E	

Sbjct	353	E D R F A S E A S G Y Q D N I A R L E E E I R H L K D E M A R H L R E Y Q D L L N V K M A L D V E I A T Y R K L L E G E	412
-------	-----	---	-----

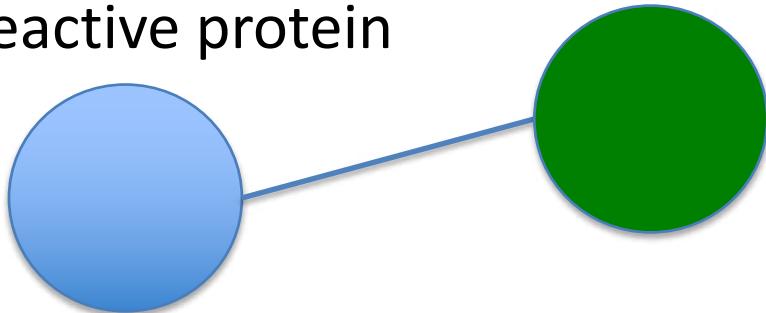
Query	399	E S R L E S G M Q N M S I H T K T T S G Y A G G L S S A Y G G L T S P G L S Y S L G S S F G S G A G S S S F S R T S S S	458
		E S R + N + I T Y L + F + S +	

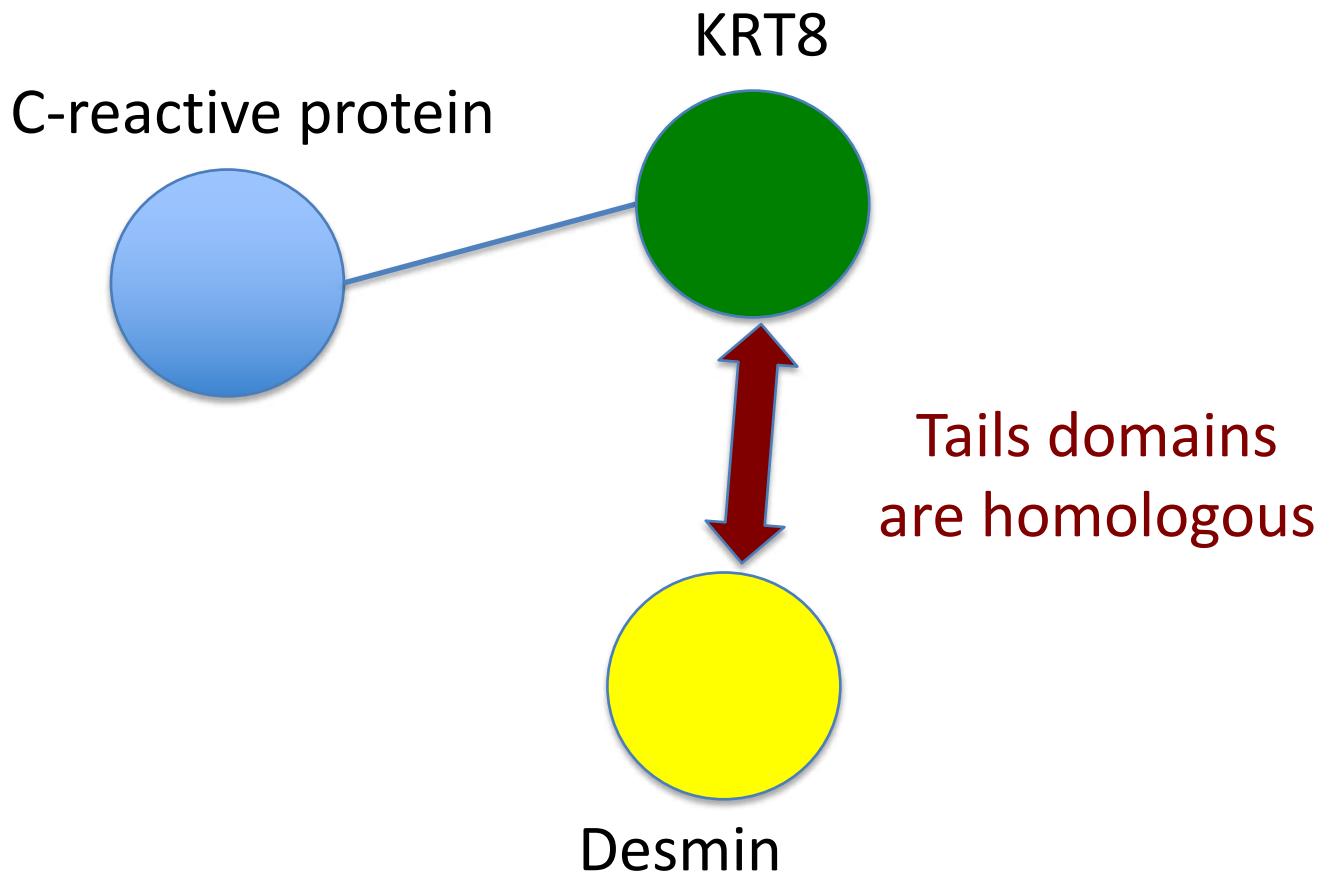
Sbjct	413	E S R I ----- N L P I Q T ----- Y S A L ----- N F R E T S P E Q R G S E V H T K	443
-------	-----	---	-----

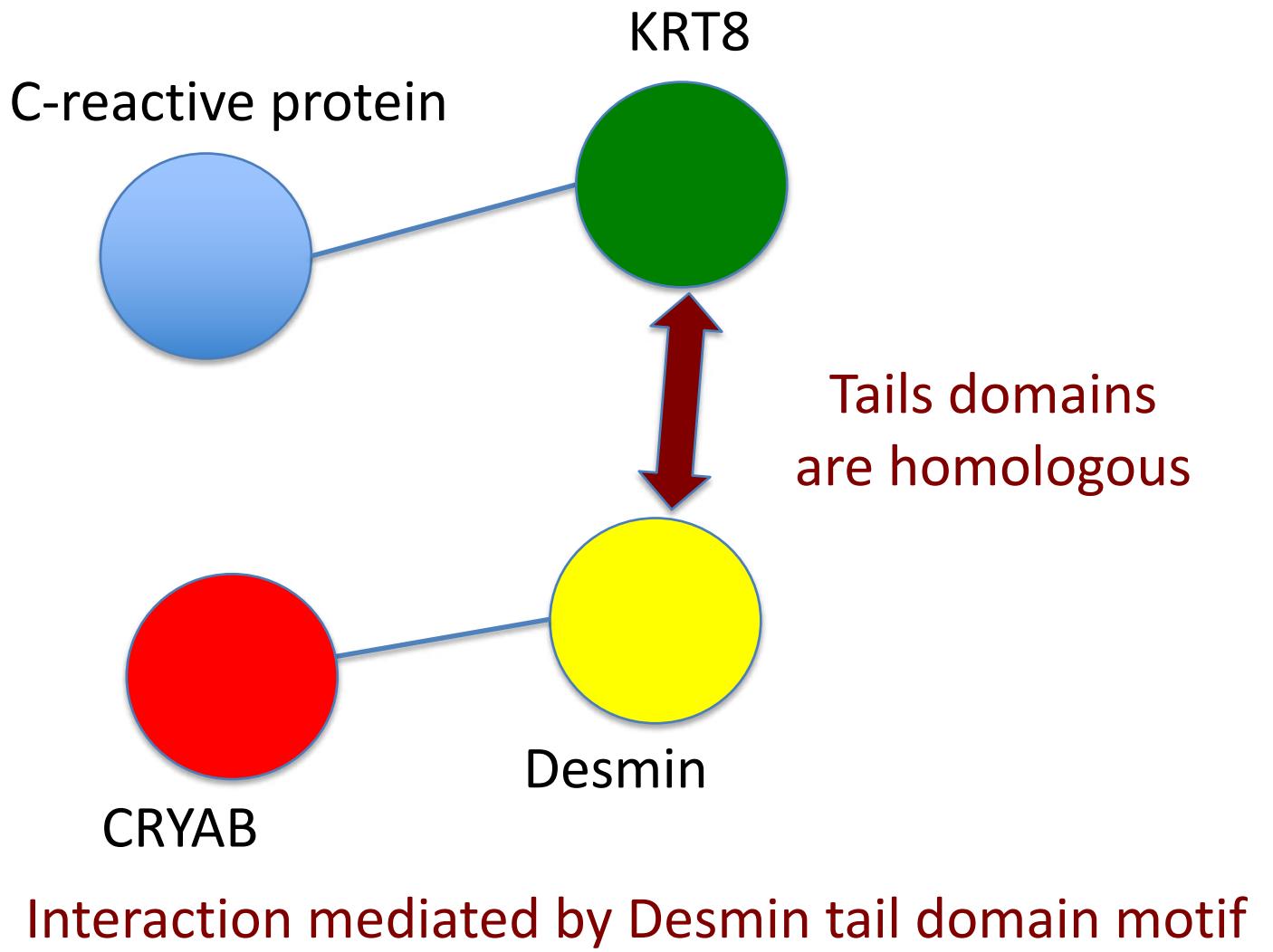
Query	459	R A V V V K K I E T R D G K L V S E S S	478
		+ V + + K I E T R D G + + V S E + +	

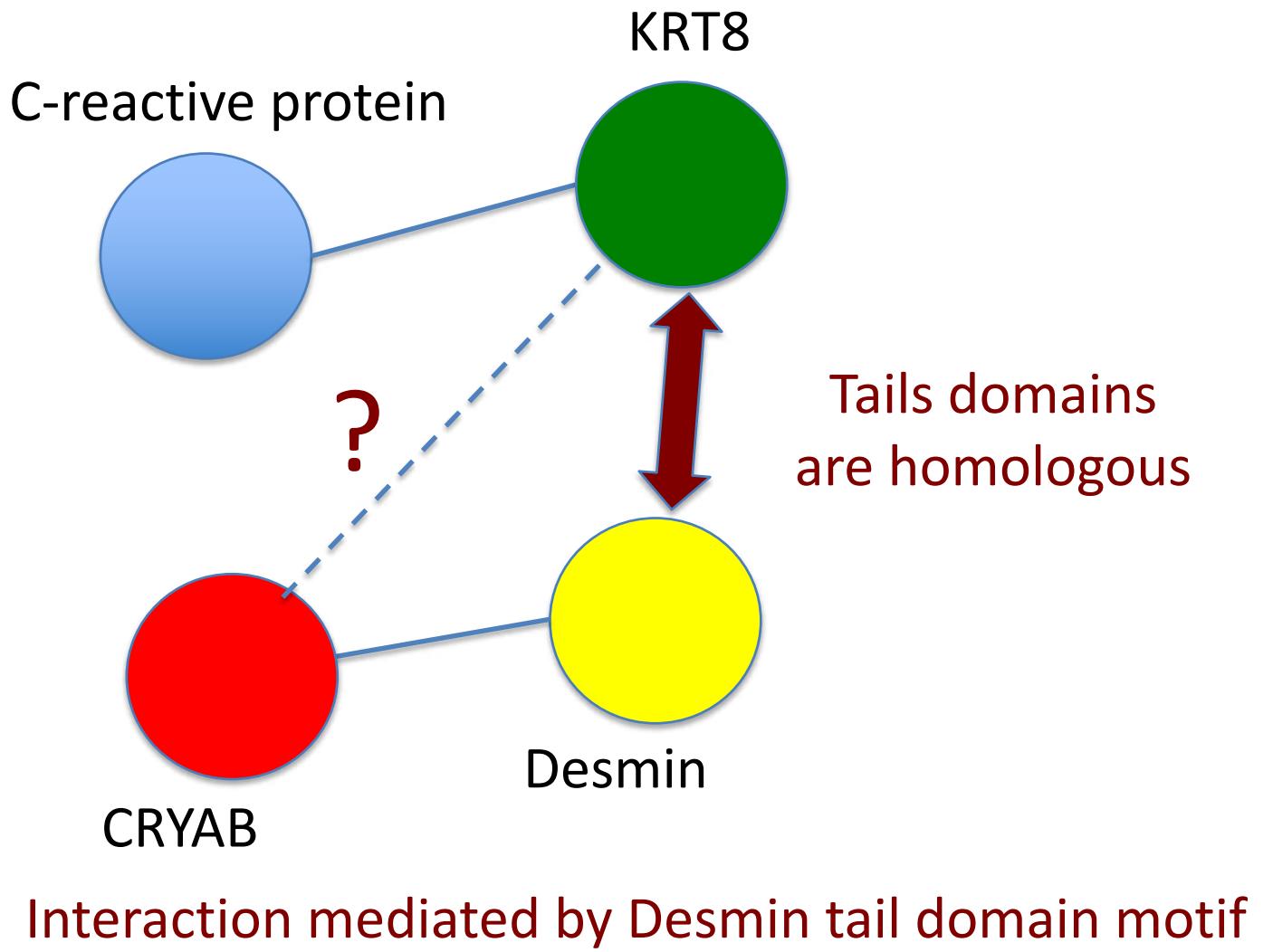
Sbjct	444	K T V M I K T I E T R D G E V V S E A T	463
-------	-----	---	-----

KRT8
C-reactive protein









αB-crystallin is a sensor for assembly intermediates and for the subunit topology of desmin intermediate filaments.

Sharma S¹, Conover GM², Elliott JL³, Der Perng M⁴, Herrmann H^{1,5}, Quinlan RA⁶.

Author information

Abstract

Mutations in the small heat shock protein chaperone CRYAB (αB-crystallin/HSPB5) and the intermediate filament protein desmin, phenocopy each other causing cardiomyopathies. Whilst the bind

CRYAB responsible for
unknown filaments are
electron says and
filament ly of desmin
subunit y status, to the
desmin tail domain of
two dise desmin and
surface or for the
assemb rms an
properti 3-binding
distingui 3 to
on (R454W)

For now it's just speculation: we need to understand more about the proteins and the interaction motif to see if it's a reasonable hypothesis that merits experimental validation

or assembly protocol is important to understanding the pathomechanism(s) of desmin-CRYAB myopathies.

Systemic augmentation of α B-crystallin provides therapeutic benefit twelve hours post-stroke onset via immune modulation

Ahmet Arac^{a,b,c,1}, Sara E. Brownell^{b,d,e,1}, Jonathan B. Rothbard^f, Charlene Chen^{b,c,e}, Rose M. Ko^f, Marta P. Pereira^{a,b,c}, Gregory W. Albers^{b,c,e}, Lawrence Steinman^{b,e,2,3}, and Gary K. Steinberg^{a,b,c,2,3}

^aDepartment of Neurosurgery, ^bStanford Institute for Neuro-Innovation and Translational Neurosciences, ^cStanford Stroke Center, ^dDepartments of Biology, ^eNeurology and Neurological Sciences, and ^fMedicine, Stanford University School of Medicine, Stanford, CA 94305

Edited by Marc Feldmann, Imperial College London, London, United Kingdom, and approved June 27, 2011 (received for review May 10, 2011)

Tissue plasminogen activator is the only treatment option for stroke victims; however, it has to be administered within 4.5 h after symptom onset, making its use very limited. This report describes a unique target for effective treatment of stroke, even 12 h after onset, by the administration of α B-crystallin (Cryab), an endogenous immunomodulatory neuroprotectant. In *Cryab*^{-/-} mice, there was increased lesion size and diminished neurologic function after stroke compared with wild-type mice. Increased plasma Cryab was detected after experimental stroke in mice and after stroke in human patients. Administration of Cryab even 12 h after experimental stroke reduced both stroke volume and inflammatory cytokines associated with stroke pathology. Cryab is

(TTC) staining, which stains the viable tissue in red due to the activity of the mitochondrial dehydrogenases (Fig. 1A). This difference remained at 7 d after stroke as assessed by silver stain (used to detect neuronal and axonal degeneration; Fig. 1B and *SI Appendix*, Fig. S1), indicating that the deficiency of Cryab affected both the early and delayed phases of ischemic damage. Functional outcome was assessed by a 28-point neurobehavioral scoring test (14, 15). The *Cryab*^{-/-} mice had significantly worse scores at both 2-d and 7-d time points compared with wild-type controls (Fig. 1C). No differences were seen in cerebral blood flow measured by laser Doppler flow meter immediately after the occlusion and at 15 and 30 min of reperfusion between the groups (*SI Appendix*, Fig. S2).

small heat shock protein after stroke. There are several potential biomarkers in plasma (e.g., C-reactive protein, S100B, and matrix metalloproteinase-9) that can be used for diagnosis or outcome prediction of stroke (30). Given that Cryab is secreted from human retinal pigment epithelial cells (31), whether it can be used as a biomarker after stroke requires more extensive studies. Variations with age of the patient must be considered in such studies.

THANK YOU!!