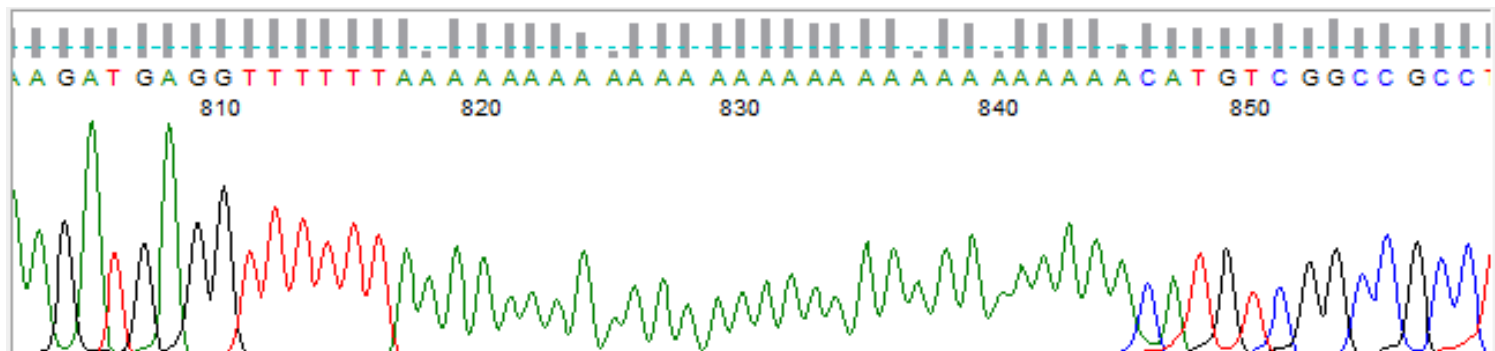
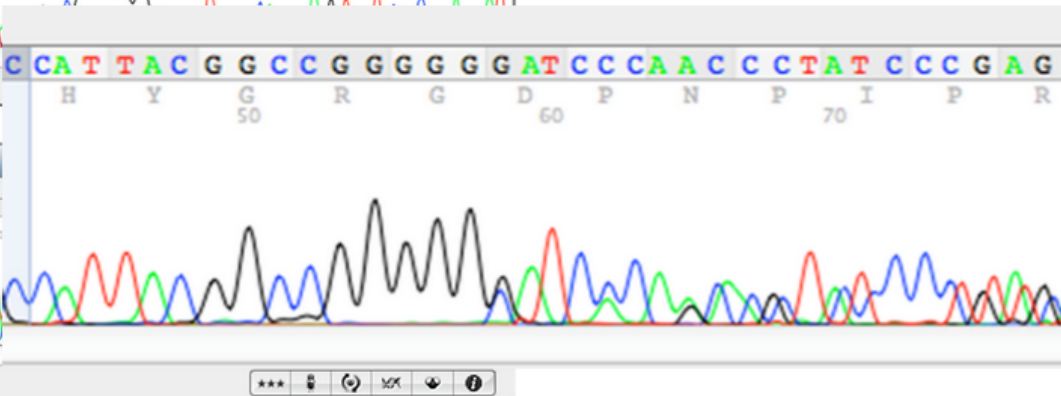
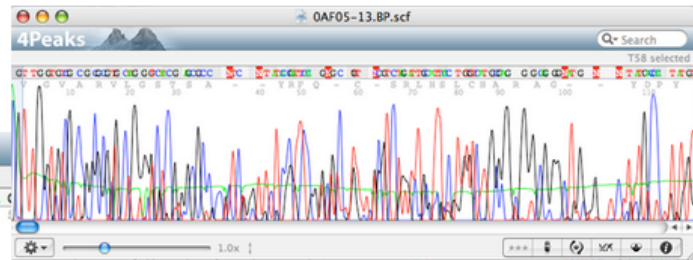
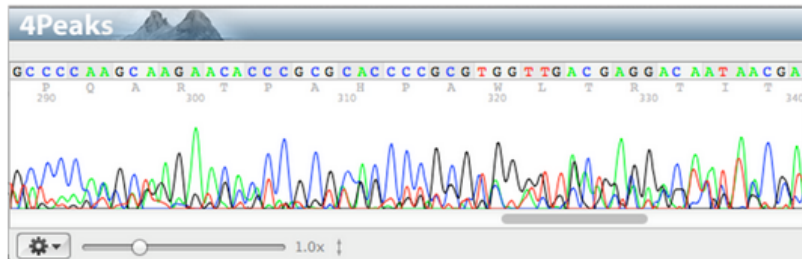
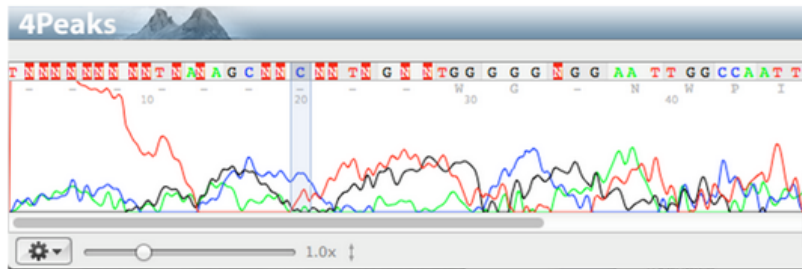




# DSAP PART 3: ANALYSIS



!?!?!?!?!?!?!?!?!?!?!?



# DSAP Review



**BLAST®** Basic Local Alignment Search Tool

Home Recent Results Saved Strategies Help

NCBI BLAST/ blastn suite/ Formatting Results - VFVE9MD8014

Edit and Resubmit Save Search Strategies Formatting options Download You

**Nucleotide Sequence (708 letters)**

RID VFVE9MD8014 (Expires on 07-06 20:21 pm)

Query ID Id|9135 Database Name nr  
Description None Description Nucleotide  
Molecule type nucleic acid  
Query Length 708 Program BLASTN

Other reports: Search Summary Taxonomy reports Distance tree of results

**Graphic Summary**

Distribution of 114 Blast Hits on the Query Sequence

Mouse over to see the define, click to show alignments

Color key for alignment scores

Query 1 100 200 300 400 500 600

<40 40-50 50-80 80-200

**Descriptions**

Sequences producing significant alignments:

Select: All None Selected:0

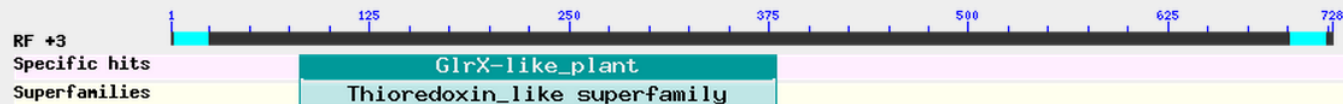
Alignments Download GenBank Graphics Distance tree of results

	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	PREDICTED: Setaria italica dynein light chain LC6, flagellar outer arm-like (LOC101770043). mRNA	244	244	39%	8e-61	79%	<a href="#">XM_004969977.1</a>
<input type="checkbox"/>	Zea mays dynein light chain LC6, flagellar outer arm (LOC100284150). mRNA	221	221	39%	1e-53	77%	<a href="#">NM_001157047.1</a>
<input type="checkbox"/>	Triticum aestivum cDNA, clone: SET5_E05, cultivar: Chinese Spring	219	219	38%	3e-53	78%	<a href="#">AK330768.1</a>
<input type="checkbox"/>	PREDICTED: Oryza brachyantha dynein light chain LC6, flagellar outer arm-like (LOC102702052). mRNA	217	217	38%	1e-52	77%	<a href="#">XM_006644677.1</a>
<input type="checkbox"/>	Sorghum bicolor hypothetical protein, mRNA	210	210	40%	2e-50	77%	<a href="#">XM_002453627.1</a>
<input type="checkbox"/>	Diospyros kaki microsatellite DKES78 sequence	206	206	39%	2e-49	76%	<a href="#">KC425431.1</a>
<input type="checkbox"/>	PREDICTED: Oryza brachyantha dynein light chain LC6, flagellar outer arm-like (LOC102719174). mRNA	203	203	39%	3e-48	77%	<a href="#">XM_006647084.1</a>
<input type="checkbox"/>	Sorghum bicolor hypothetical protein, mRNA	201	201	34%	9e-48	78%	<a href="#">XM_002456341.1</a>
<input type="checkbox"/>	Hordeum vulgare subsp. vulgare cDNA clone: FLba18024, mRNA sequence	201	201	38%	9e-48	76%	<a href="#">AK252996.1</a>
<input type="checkbox"/>	Brachypodium distachyon mRNA, clone: PL016C01-A-022_P10	199	199	40%	3e-47	75%	<a href="#">AK427785.1</a>
<input type="checkbox"/>	PREDICTED: Setaria italica dynein light chain LC6, flagellar outer arm-like (LOC101780022). mRNA	199	199	39%	3e-47	76%	<a href="#">XM_004951174.1</a>
<input type="checkbox"/>	PREDICTED: Brachypodium distachyon dynein light chain LC6, flagellar outer arm-like (LOC1008245)	199	199	40%	3e-47	75%	<a href="#">XM_003569813.1</a>
<input type="checkbox"/>	Brachypodium distachyon mRNA, clone: PL016C01-A-092_I18	196	196	39%	4e-46	76%	<a href="#">AK438403.1</a>
<input type="checkbox"/>	Oryza rufipogon (W1943) cDNA clone: ORW1943C107A12, full insert sequence	194	194	39%	1e-45	76%	<a href="#">CT841868.1</a>
<input type="checkbox"/>	Oryza sativa Japonica Group cDNA, clone: J075067H10, full insert sequence	194	194	39%	1e-45	76%	<a href="#">AK287966.1</a>
<input type="checkbox"/>	Oryza sativa (indica cultivar-group) cDNA clone: OSIGCSN016L02, full insert sequence	194	194	39%	1e-45	76%	<a href="#">CT828841.1</a>

# DSAP Review

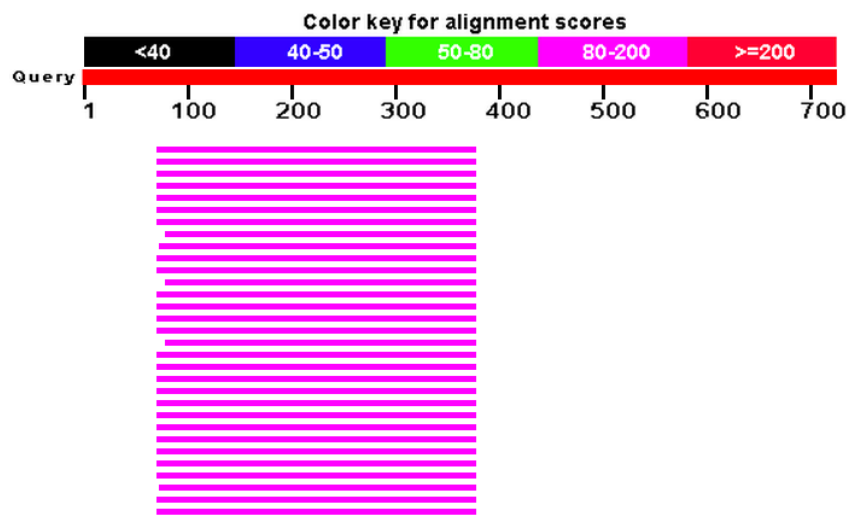


Putative conserved domains have been detected, click on the image below for detailed results.



Distribution of 102 Blast Hits on the Query Sequence

Mouse over to see the defline, click to show alignments



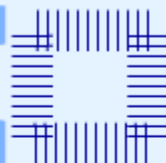
# DSAP Review



## WSSP Translator Toolbox

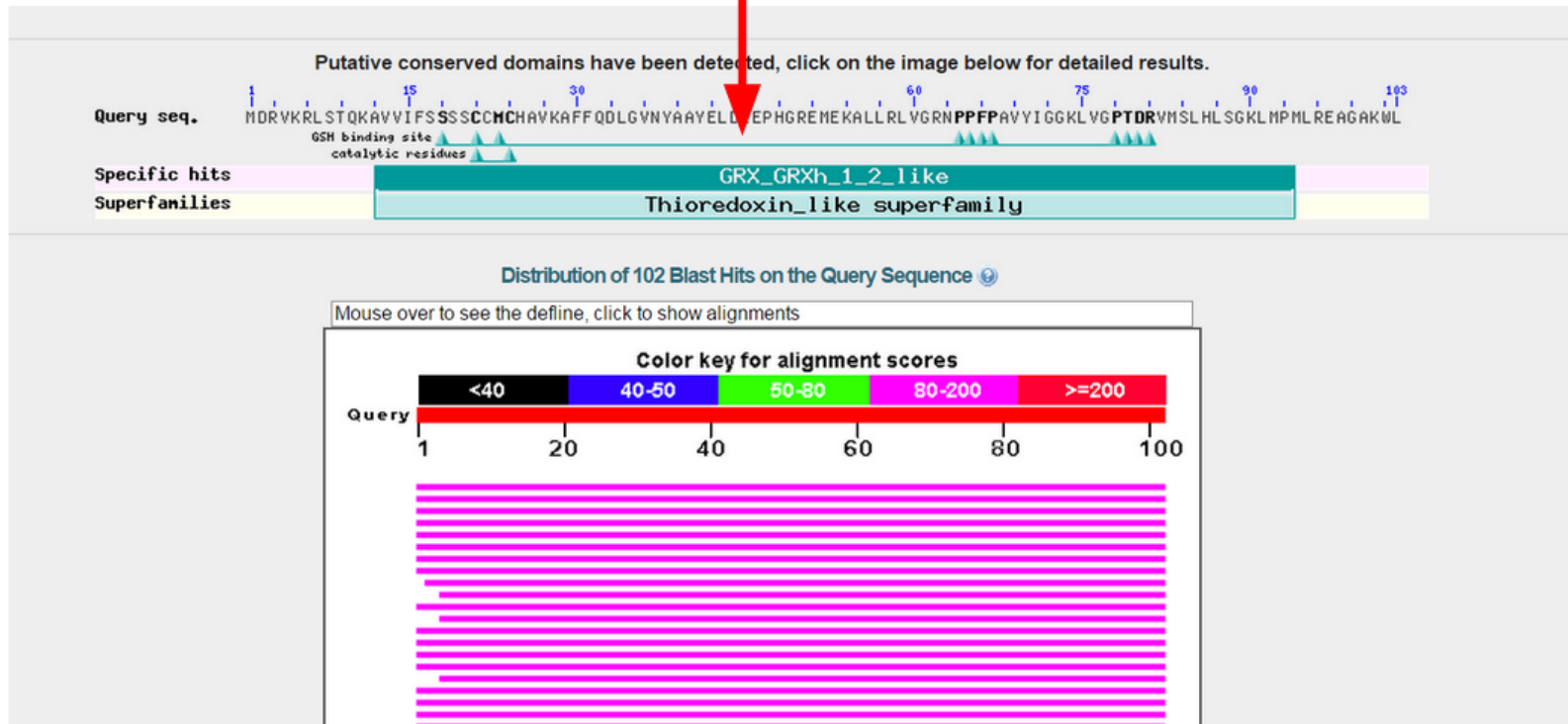
### Reading Frame 3

Total codons = 242. ORF #1 is longest with 126 amino acids.  
Select a run of amino acids to see the corresponding DNA sequence.



EKKKEEEKQSTSSFLTLHSAGEFMDRVKRLSTQKAVVIFSSSSCCMCHAVKAFFQ  
DLGVNYAAYELDEEPHGREMEKALLRLVGRNPPFPAVYIGGKLVGPTDRVMSLHL  
SGKLMPMLREAGAKWL\*SGSLRNPNASGSRLTCVDKWAALCSRALKWAWARAPFH  
LRFSPKSKSVR\*SRTWGNRQTRGSSSVNRNRPDIPRGLFLNGLLTSARLSDRRALL  
CTRAV\*LIYN\*LTKKKKKKKK-

# DSAP Review



# Now What?



## Analyze

In this step you will further analyze your results.



1

What is the function of your protein? List the sites where you found this information. (books, Google, Wikipedia, PubMed)

Reticulon and reticulon-like proteins are found on endoplasmic reticulum that promotes membrane curvature. In addition reticulons may play a role in nuclear pore complex formation, vesicle formation, and other processes yet to be defined. there

2

Was there a similar protein in humans? If so, what does this tell you about the evolution of this protein?

Yes. This protein formed very early in the evolution of life and plays a vital role in all cells.



# But first, “review”



## Review

```
CGTATCGCCGCGCGTTCCTCTCTCCAGCTCCCGGAGTTACTTCCAGCTCCGTCCGCAGCT  
CCCGCCTCCATTTCTCGGCCATGAAGCCCTCGGCGGAGGCGCAGCTGCCGACATAATCCTG  
TGGAGGAGAAGGAATCCGGCTATGGTGATCTGACCGGGCGACTGCCGCTTGGTTCCTCTT
```

Number of characters: 144    Cursor is on base: M1    Range marked:

Your predicted protein sequence: (taken from "[Define ORF](#)" step)

```
MAAAAADDPAGEPVSPPRSSLQPPGVTSSSVRSSRLHFSGHEALGGGAAADIILWRRRNPAMVI  
LTGATAAWFLFEIAGYSFLSLLASVLFLLVSILFLWARSASLLNRPLPPLPNLEIPDRVAEKIA  
DEARVWINRSCPSLER
```

Number of characters:    Cursor is on residue:    Range marked:

Mark the sections of the DNA sequence to determine the following:

- 1    5' UTR    Start:     End:     ☐ Not Applicable
- 2    ORF    Start:     End:   
(Taken from "Define ORF" - [Click here to return to Define ORF](#))
- 3    3' UTR    Start:     End:     ☐ Not Applicable  
(For non-coding clones enter range of entire sequence here.)

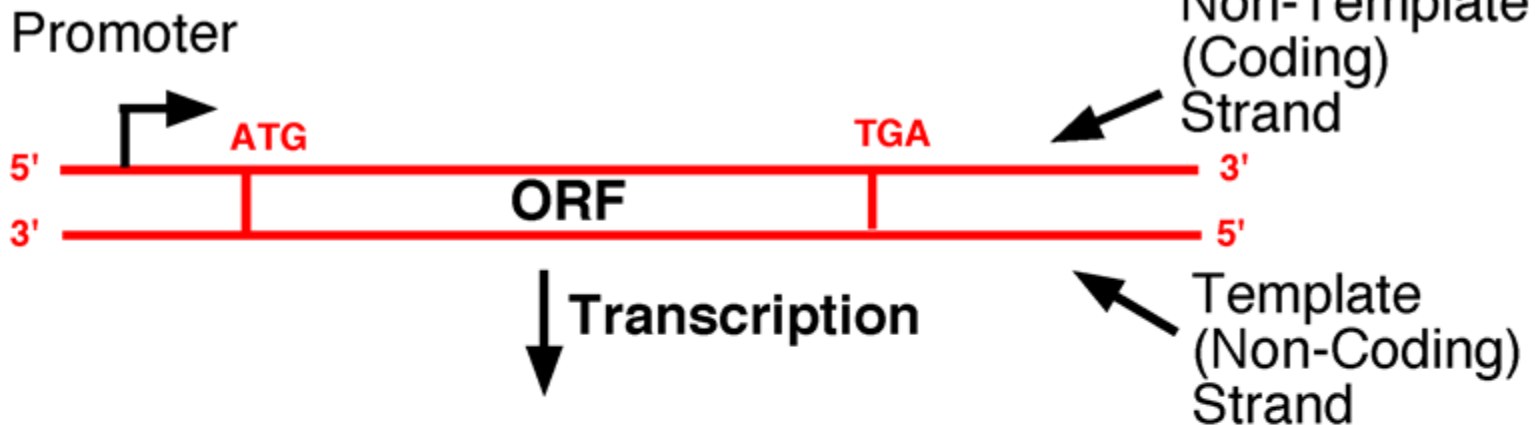
<< Back

Save Now

Next >>

☐ Ignore Errors

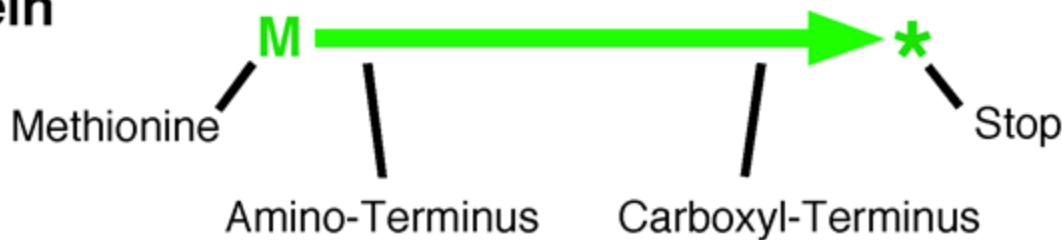
DNA



mRNA



Protein



Courtesy of WSSP  
2014

# Analysis Step 1: RESEARCH!



1

What is the function of your protein? List the sites where you found this information. (books, Google, Wikipedia, PubMed)

is evidence that they influence endoplasmic reticulum-Golgi trafficking, vesicle formation and membrane morphogenesis.  
<https://en.wikipedia.org/wiki/Reticulon>  
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2246256/>

Other sources: Pubmed, Research papers/journals, Textbooks.  
ONE from wiki and AT LEAST ONE from another!

# Step 2: Related in Humans?

And if so, what does it mean?

Download ▾ GenPept Graphics ▴ Next ▴ Previous ▴ Descriptions

dynein light chain LC6, flagellar outer arm [Zea mays]  
Sequence ID: [ref|NP\\_001150519.1|](#) Length: 93 Number of Matches: 1  
[▶ See 2 more title\(s\)](#)

Range 1: 1 to 93 [GenPept](#) [Graphics](#) ▾ Next Match ▴ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
159 bits(402)	3e-48	Compositional matrix adjust.	73/93(78%)	83/93(89%)	0/93(0%)

Query 1 MLEGARVVEDTDMFRKQAEANGAASHALDFDVADCKSLAAHIKKEFDKIYGPQWQCVV 60  
MLEG+A VEDTDMF EMQA+AM+AAS ALD FDV DC+S+A+HIKKEFD I+GPQWQCVV  
Sbjct 1 MLEGKAVVEDTDMFAKMQAQMSAASRALRFDVLCRSIASHIKKEFDALHNGPQWQCVV 60

Query 61 GSSFGCFTHKSGSFYFRLETLHFLIFKGA 93  
GS FGC+ TH KGSFIYFRLE+L FL+FKGA  
Sbjct 61 GSGFGCYITHSGSFYFRLESRLFLVFKGA 93

**Related Information**  
[Gene](#) - associated gene details  
[UniGene](#) - clustered expressed sequence tags  
[Identical Proteins](#) - Proteins identical to the subject

## Identical Proteins

NCBI Resources ▾ How To ▾

Protein  [Advanced](#)

[Show additional filters](#)

Species  
Plants  
More ...

Source  
databases  
GenBank  
RefSeq

Sequence length  
Custom range...

Molecular weight  
Custom range...

Release date  
Custom range...

Revision date  
Custom range...

[Clear all](#)

[Show additional filters](#)

**Display Settings:** ☒ Summary, Sorted by Default order

**Results: 3**

☐ [TPA: putative dynein light chain type 1 family protein \[Zea mays\]](#)  
1. **93 aa protein**  
Accession: DAA57536.1 GI: 414880405  
[GenPept](#) [FASTA](#) [Graphics](#) [Related Sequences](#) [Identical Proteins](#)

☐ [dynein light chain LC6, flagellar outer arm \[Zea mays\]](#)  
2. **93 aa protein**  
Accession: NP\_001150519.1 GI: 226493894  
[GenPept](#) [FASTA](#) [Graphics](#) [Related Sequences](#) [Identical Proteins](#)

☐ [dynein light chain LC6, flagellar outer arm \[Zea mays\]](#)  
3. **93 aa protein**  
Accession: ACG39387.1 GI: 195639838  
[GenPept](#) [FASTA](#) [Graphics](#) [Related Sequences](#) [Identical Proteins](#)

**Display Settings:** ☒ Summary, Sorted by Default order

Select to go  
to Related  
Sequences

# (Extra) Organism Cheat Sheet



## Species Name

*Zea mays*

*Ricinus communis*

*Populus trichocarpa*

*Vitis vinifera*

*Glycine max*

*Arabidopsis thaliana*

*Oryza sativa Japonica*

*Saacharomyces cerevisiae*

*Strongylocentrotus purpuratus*

*Ixodes scapularis*

*Drosophila melanogaster*

*Caenorhabditis elegans*

*Bos taurus*

*Rattus norvegicus*

*Mus musculus*

*Homo sapiens*

## Common Name

corn

wheat

popular (tree)

grape

soy bean

mustard (model organism)

rice

Yeast (model organism)

Sea urchin

Deer tick

Fruit fly (model organism)

Nematode (model organism)

Cow

Rat

Mouse (model organism)

Human

# Related in Humans?

To the right of the results list you can organize results by species. Click “More”, then type in Homo sapiens to search humans.

The screenshot displays the NCBI Protein search results page. At the top, there are tabs for 'NCBI', 'Resources', and 'How To'. Below these is a 'Protein' search bar with a dropdown menu set to 'Protein' and an 'Advanced' link. To the right of the search bar is a list of filters: 'Species', 'Enzyme types', 'Source databases', 'Sequence length', 'Molecular weight', 'Release date', and 'Revision date'. Each filter has a 'Custom range...' option. A red arrow points from the 'More...' link under the 'Species' filter to a dropdown menu that is open, showing a list of species including 'Homo', 'Homo/Pan/Gorilla group', 'Homo sapiens', 'Pseudomonas RNA homology group II', 'Homobasidiomycetes', and 'Homoptera'. The 'Homo' option is selected. Below the dropdown menu, there are links for 'Show additional filters', 'Clear all', and 'Show additional filters'.

NCBI Resources How To

Protein Protein Advanced

Show additional filters

Species  
Animals  
Plants  
Fungi  
Protists  
More ...

Enzyme types  
Hydrolases

Source databases  
DDBJ  
EMBL  
GenBank  
PDB  
PIR  
RefSeq  
UniProtKB / Swiss-Prot

Sequence length  
Custom range...

Molecular weight  
Custom range...

Release date  
Custom range...

Revision date  
Custom range...

Clear all

Show additional filters

# The Fun Stuff!!!



## Analyze

3

Does your protein contain domains? If so, what are the roles of these domains?

4

Can you find the three dimensional structure of a protein that is similar to yours? If so, indicate which organism it is from and the accession number for the structure.

5

If there is a three dimensional structure of a homolog to your protein:

a) How many alpha helices are in the protein?

b) How many beta-strands are in the protein?

c) If there are beta-strands are they all parallel, antiparallel or a combination of both?

# Finding the 3-D Model



NCBI

Conserved Domains

HOME SEARCH GUIDE NewSearch Structure Home 3D Macromolecular Structures Conserved Domains Pubchem BioSystems

Conserved domains on [lcl|32016] View Standard Results ?

Local query sequence

Graphical summary show options ?

Query seq. Specific hits Non-specific hits Superfamilies

Dynein\_light PTZ00059 PLN03058 Dynein\_light superfamily

Search for similar domain architectures ? Refine search ?

List of domain hits

	Name	Accession	Description	Interval	E-value
[+]	Dynein_light	pfam01221	Dynein light chain type 1;	4-91	3.36e-37
[+]	PTZ00059	PTZ00059	dynein light chain; Provisional	1-89	6.31e-26
[+]	PLN03058	PLN03058	dynein light chain type 1 family protein; Provisional	7-91	9.46e-22

Blast search parameters

Data Source: Live blast search RID = VGKK8YH4014

User Options: Database: CDSEARCH/cdd v3.11 Low complexity filter: yes Composition Based Adjustment: no E-value threshold: 0.01 Maximum number of hits: 500



# Finding the 3-D Model



pfam01221: Dynein\_light, with user query added

Dynein light chain type 1

Links ?

Statistics ?

Structure ?

Solution structure of a

pf

## Sequence Alignment

Reformat

Format:

Compact Hypertext

Row Display:

up to

1Y03_C	17	.	SVVKNVDMTEEMQIDAIDCANQA
query	4	.	ARVEDTDMPRKMQAEAMNAASHA
gi 75158506	118	.	VKVMAADMSPFMQLHAFRCAKRS
gi 75180276	34	.	VRVRSADMPLPQQNRAFSLSREI
gi 75161711	4	.	AVMGDTDMKQTMKEDALSLASKA
gi 75165366	4	.	AMVEDTDMPVKMQLQAMSAAYKA
gi 75121811	4	.	ARVEDTDMFARMQAAATSAASRA
gi 6320632	7	.	PIVKASDITDKLKEDILTISKDA
gi 225466127	55	.	VRLRSADMSAAMQERAFRYARSL
gi 225456608	179	.	VKVVSVDMPFPMQIHAVDCARNA.

Links ?

Statistics ?

Structure ?

Structure View

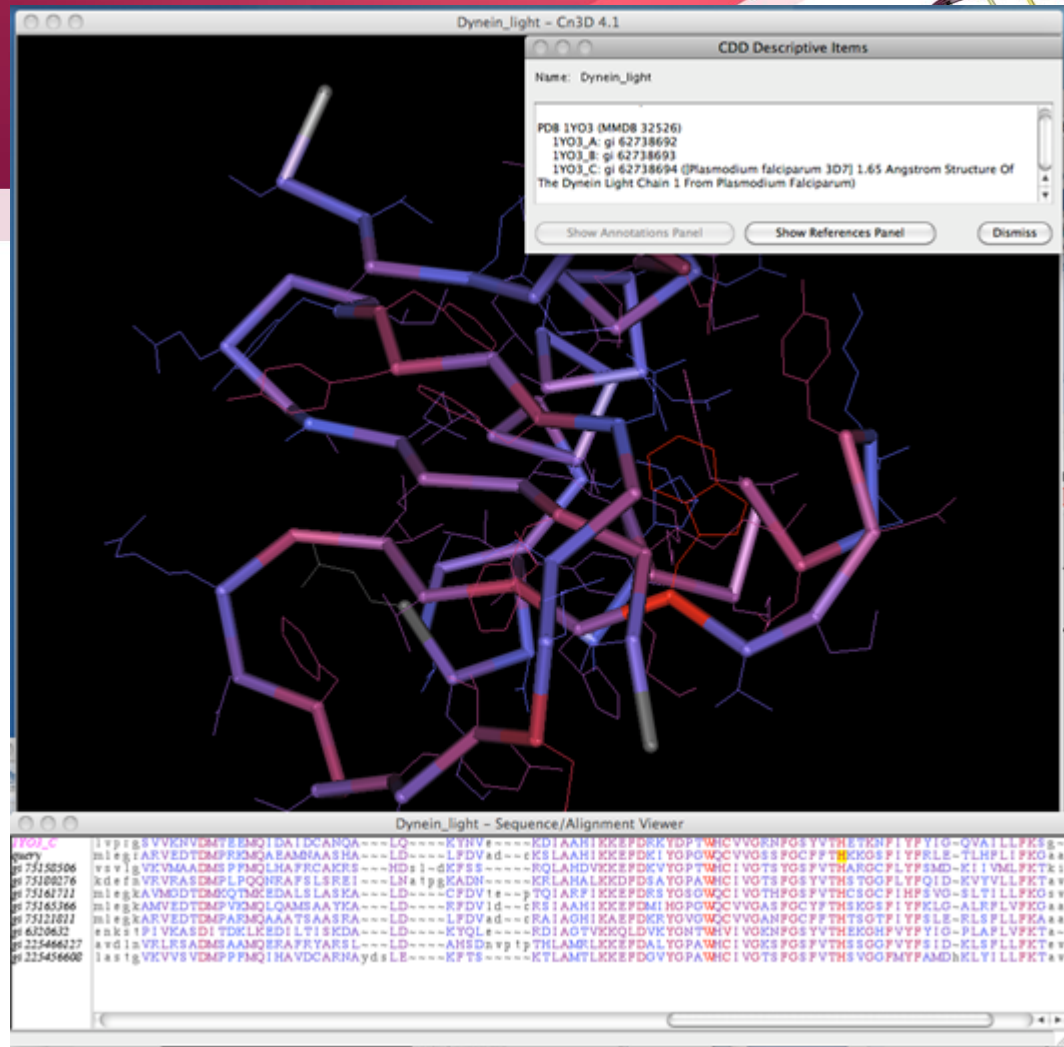
Program: Cn3D

Drawing: All Atoms

Aligned Rows: up to 10

Download Cn3D

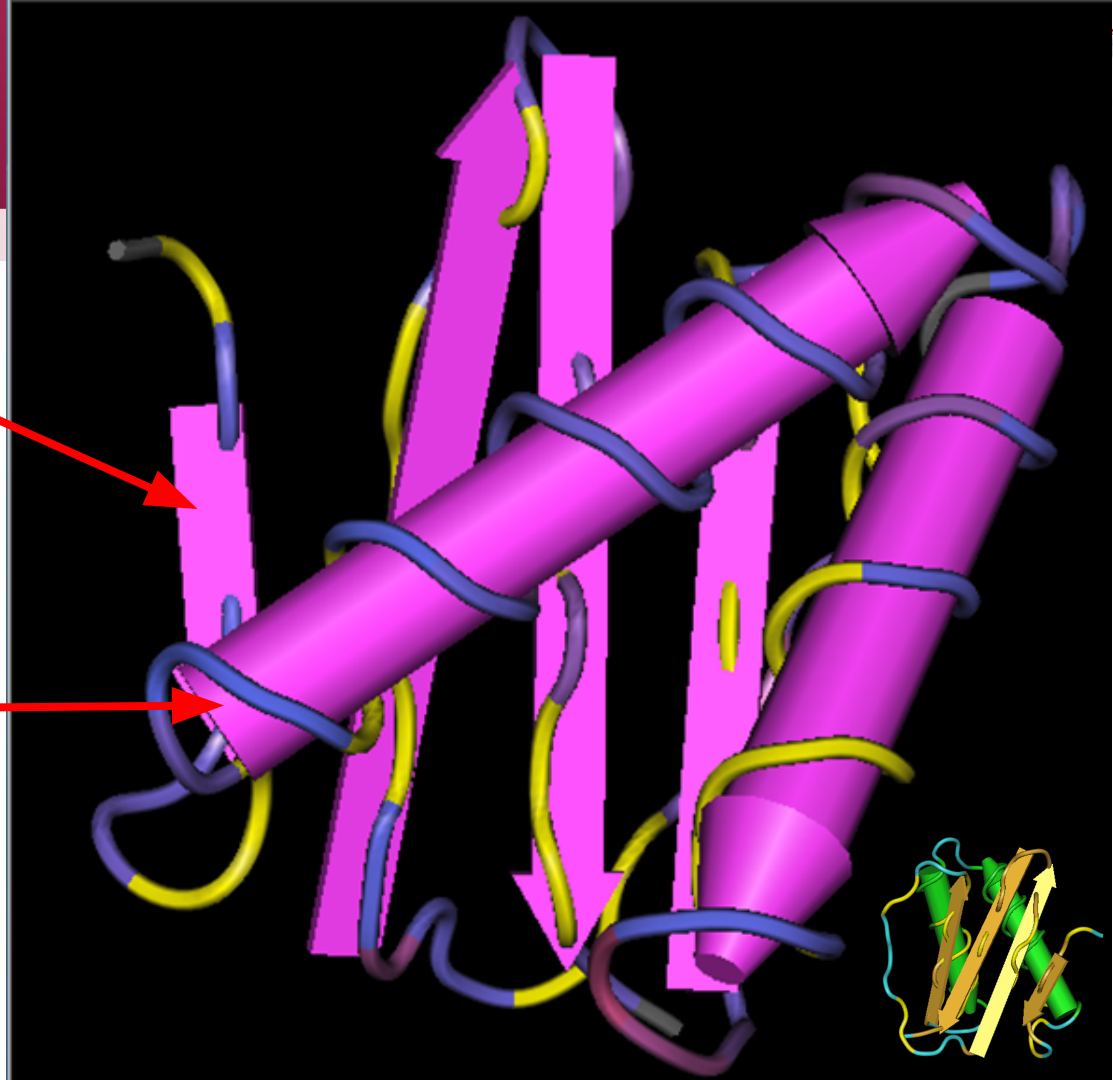
# Using Cn3D



# Using Cn3D

Beta pleated sheets

Alpha Helices



# THE MOST IMPORTANT PART!!!!!!



7

As a scientist investigating this protein in *Landoltia*, what experiments would you do next?  
Please be specific about the experiments and predict what results you would expect to see.

# THE OTHER IMPORTANT PART



Submit

1

This clone is marked as:

Coding

Describe the clone in one line:

Reticulum

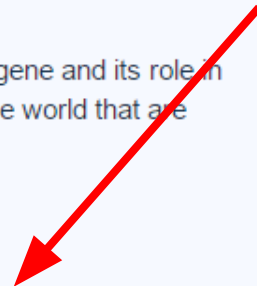
Congratulations!

You have just reached the final step of the clone analysis process.

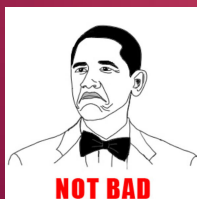
We hope that you have been able to determine the likely function of your gene and its role in *Wolffia*. This information may be valuable to other scientists throughout the world that are working on similar proteins.

Submit

PLEASE PRESS THIS  
BUTTON.

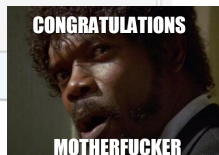


# UR DONE!!!!



## NOW UR PUBLISHED!!!!!!!!!!!!

Status ?
Being worked on by a student
Unreadable
Submitted to NCBI
Submitted to NCBI



Just like this  
smokin guy  
over here o  
dam



Yes you get this if you'  
re published.



# Future President of Waksman



“I’m running for President of Waksman in 2020.”

