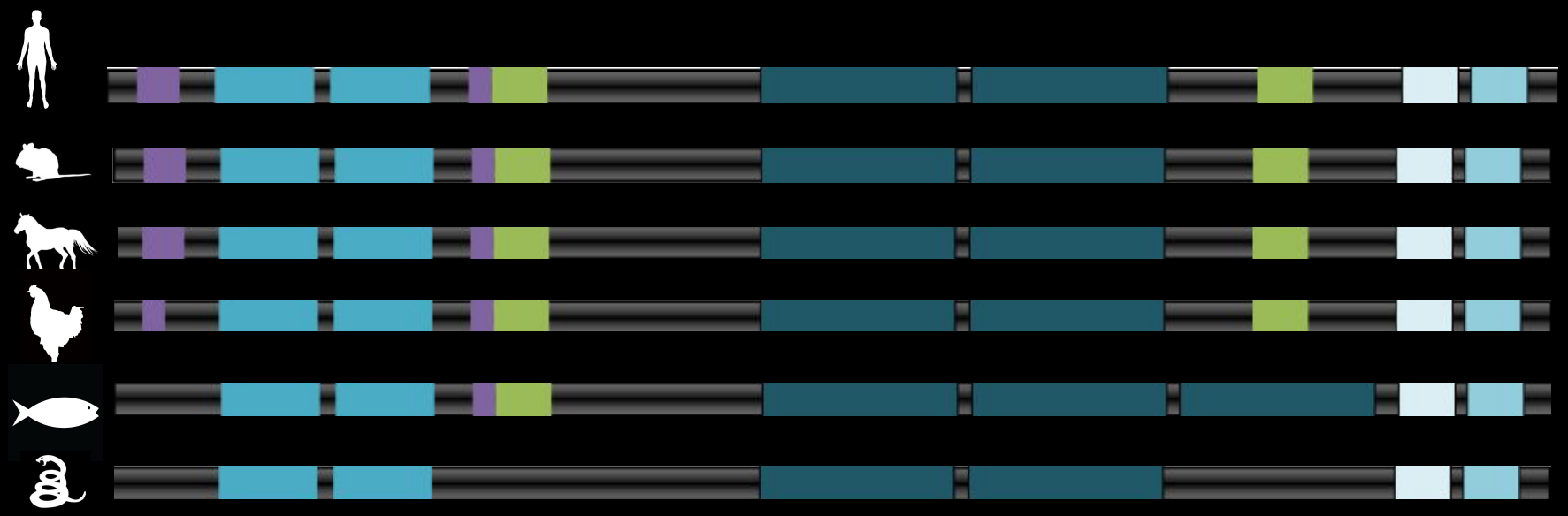
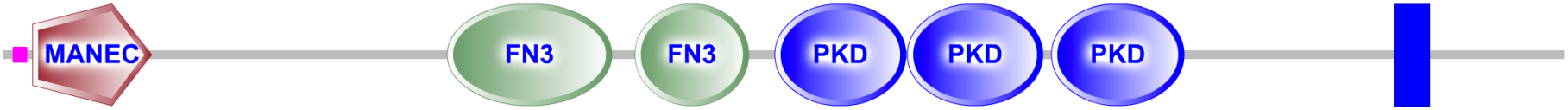
Domains, Motifs, and Pfam/SMART Analysis



Sara Grange and Katie Ness

http://kohlmanngen677s13.weebly.com/uploads/1/6/8/8/16889088/3541790_orig.jpg



Protein Domains/ DNA motifs

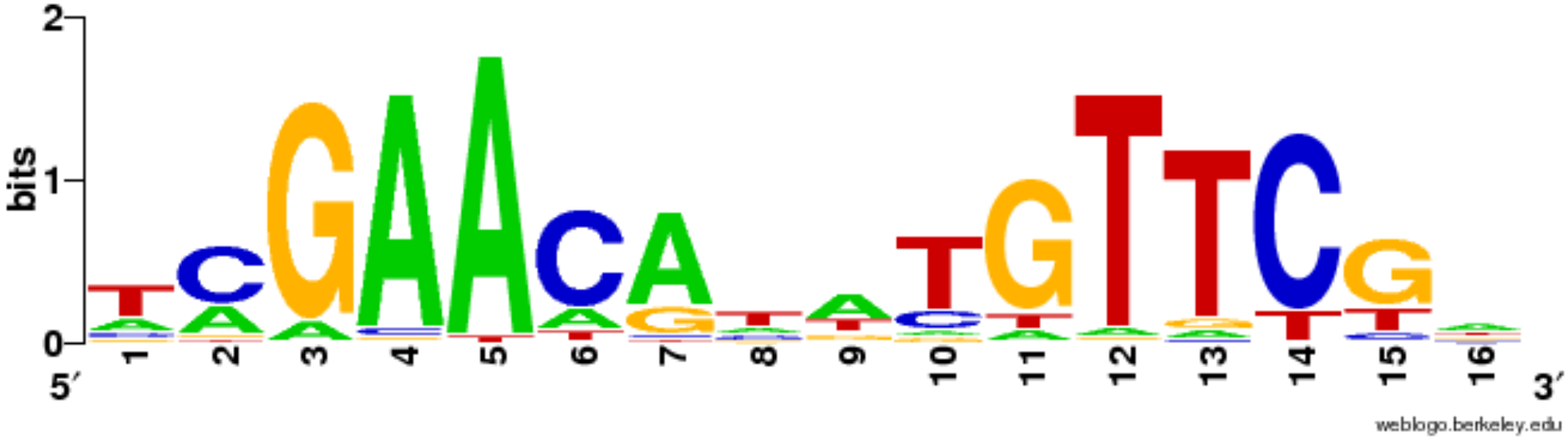
Why are they important?

How are they built?

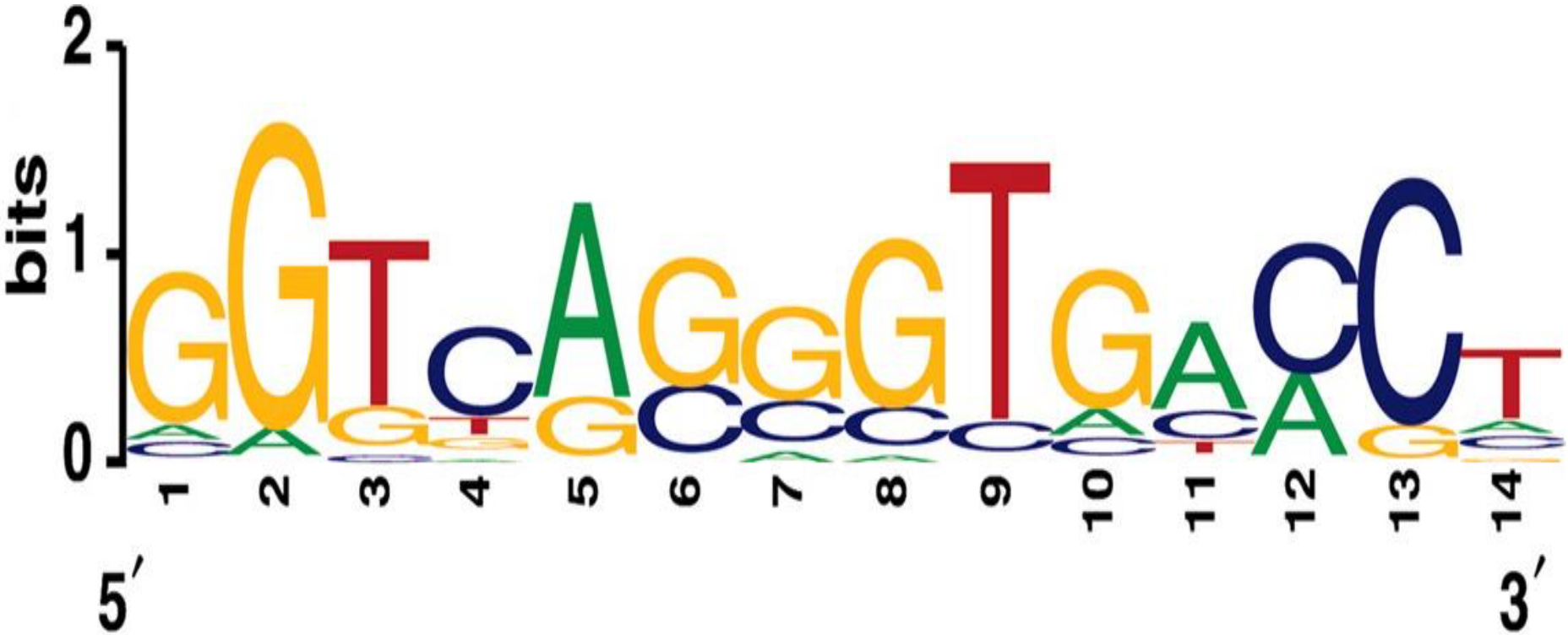
Analysis Techniques

How do they work?

Pfam and SMART databases



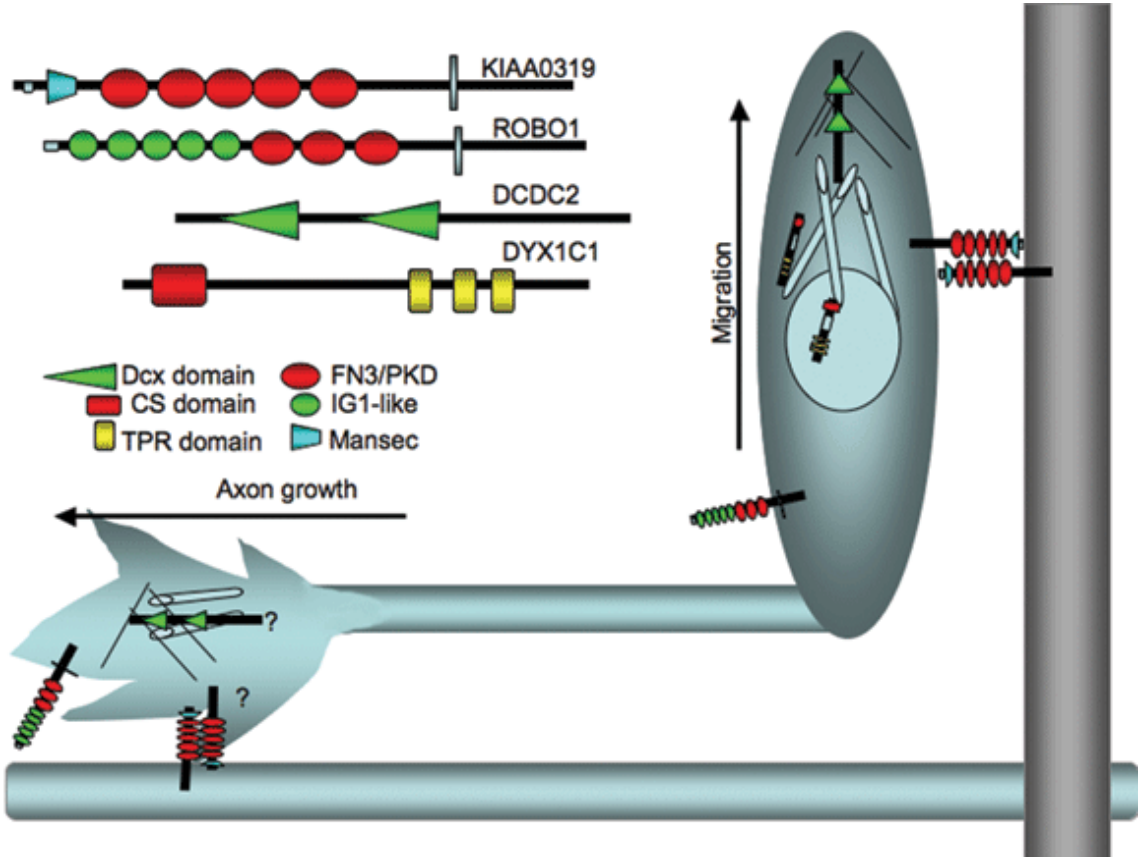
What is a DNA motif?



Short, conserved regions of DNA that can be found within domains

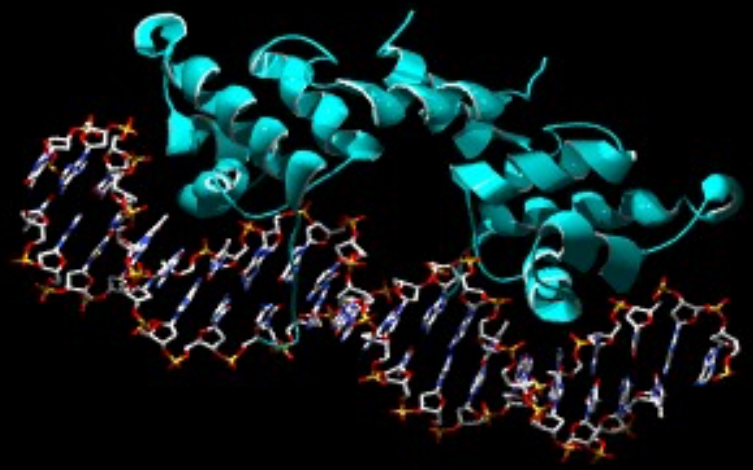
http://www.comp.nus.edu.sg/~bioinfo/LocalMotif/motiflogo1.JPG

What is a protein domain?

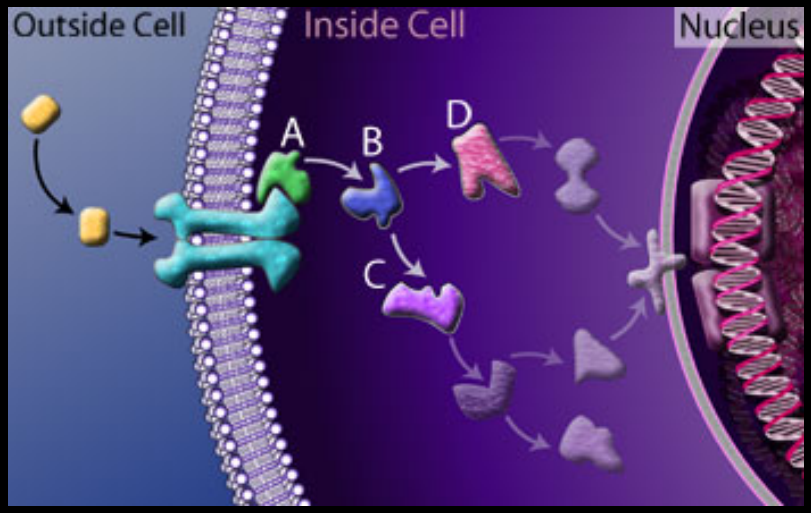


Conserved structural or functional protein segments

http://www.nature.com/neuro/journal/v9/n10/full/nn1772.html



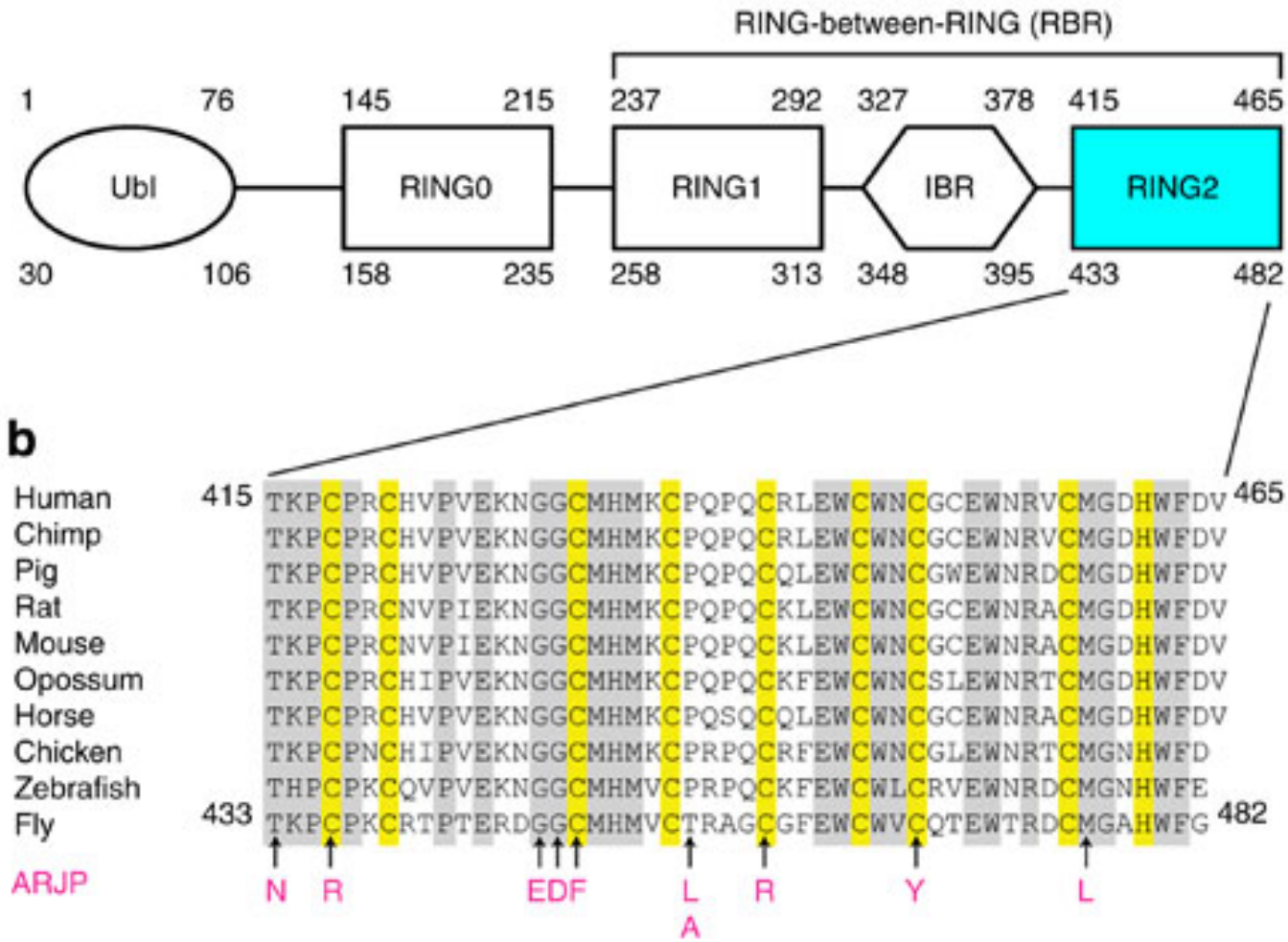
What other functions can proteins have that a protein domain can tell you?



http://www.wiley.com/college/pratt/0471393878/student/structure/dNA_binding_proteins/index.html

https://www.nsf.gov/news/news_summ.jsp?cntn_id=104085

How is a domain determined?



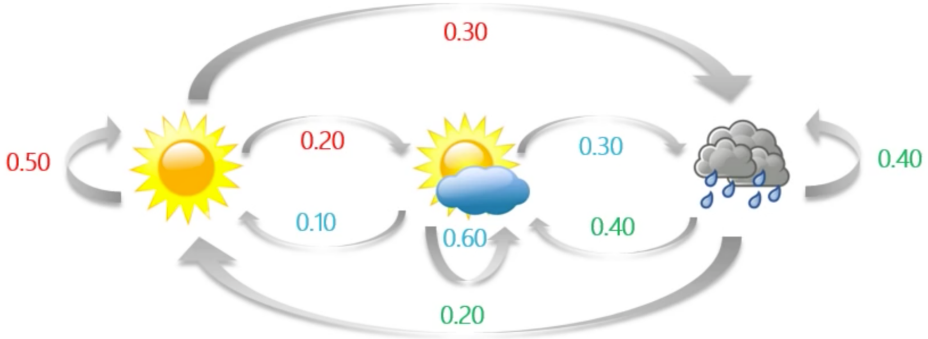
Multiple sequence alignments between homologues

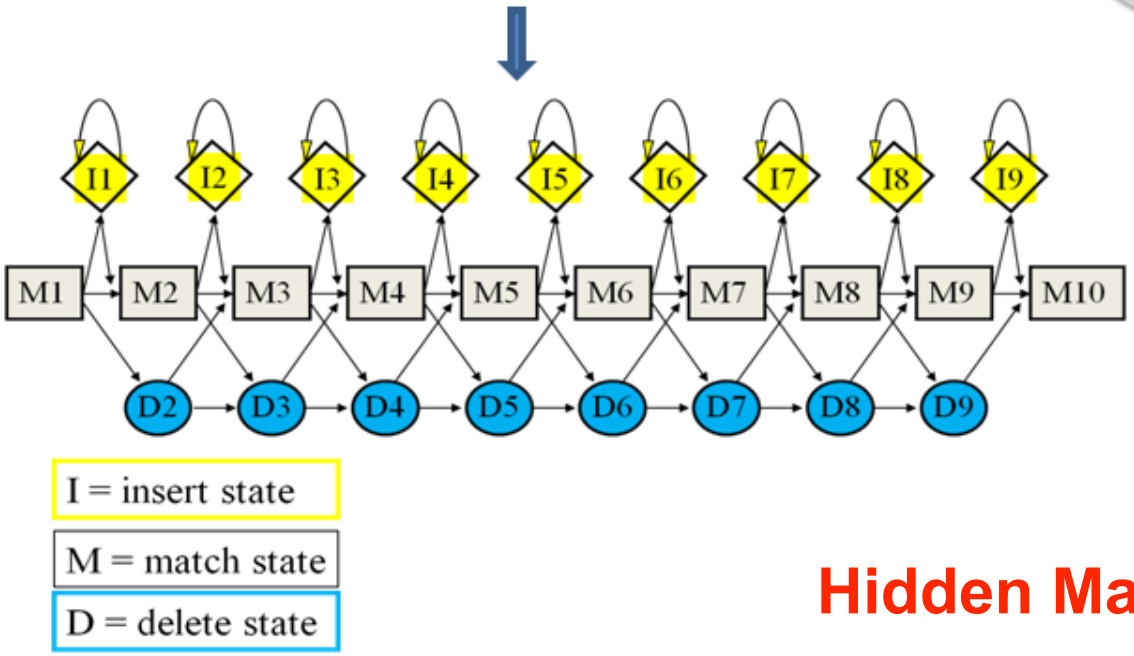
http://www.nature.com/ncomms/2013/130617/ncomms2983/fig_tab/ncomms2983_F1.html

How are multiple sequence alignment quantified?

Multiple sequence alignment

Sequence 1:	F	K	L	L	S	H	C	L	L	V
Sequence 2:	F	K	A	F	G	Q	T	M	F	Q
Sequence 3:	Y	P	I	V	G	Q	E	L	L	G
Sequence 4:	F	P	V	V	K	E	A	I	L	K
Sequence 5:	F	K	V	L	A	A	V	I	A	D
Sequence 6:	L	E	F	I	S	E	C	I	I	Q
Sequence 7:	F	K	L	L	G	N	V	L	V	C



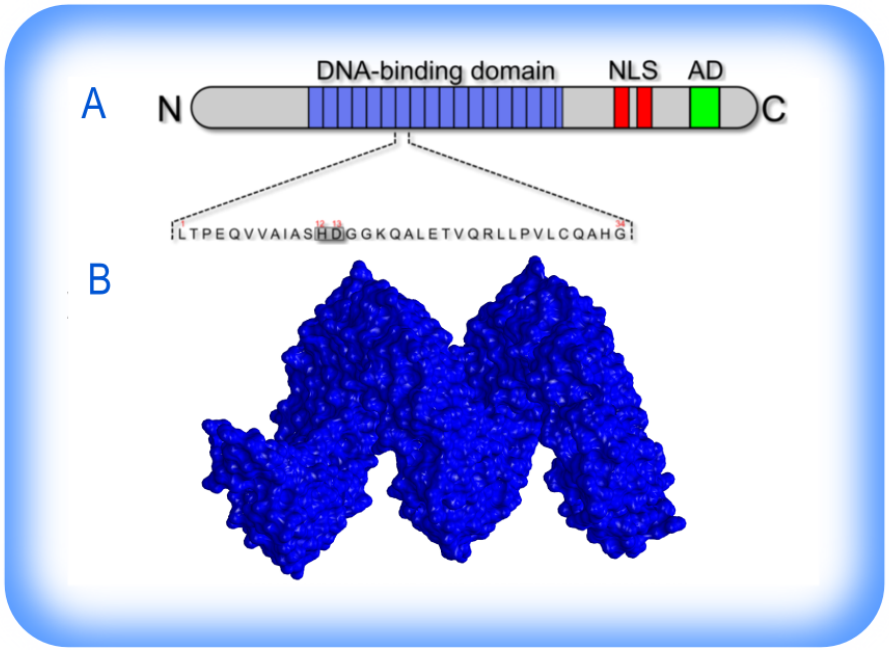


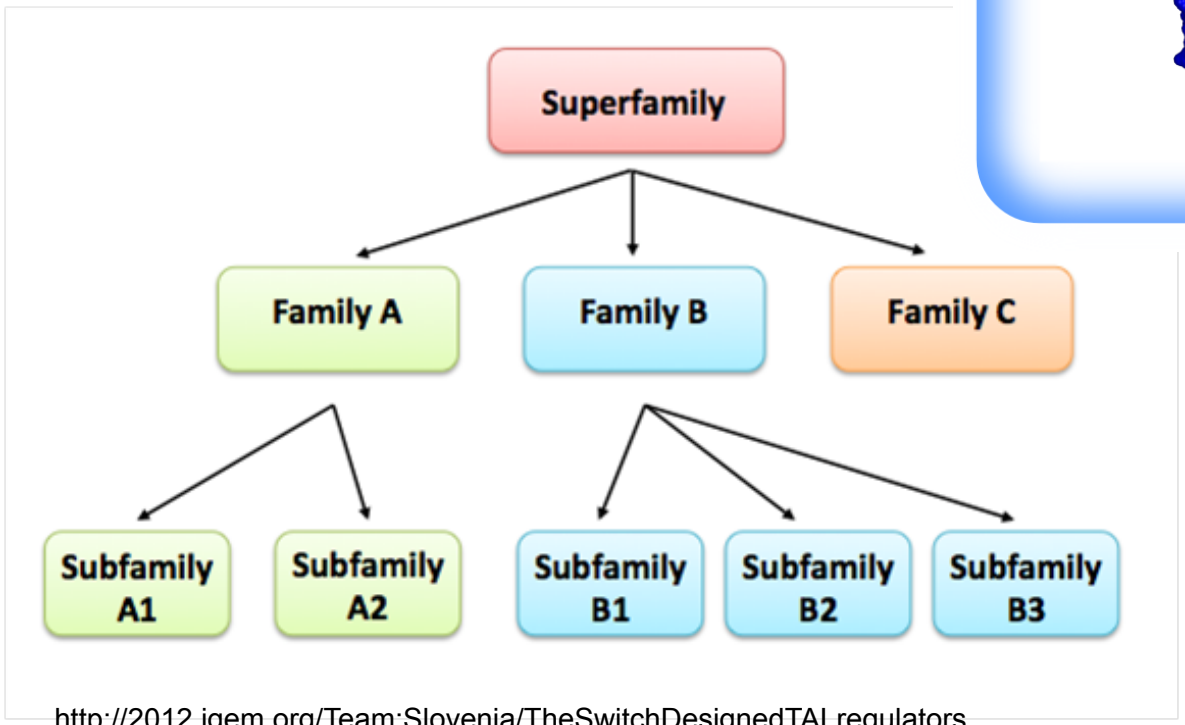
Hidden Markov Models (HMMs)

https://www.youtube.com/watch?v=j3r9a75zOvM

http://www.ebi.ac.uk/training/online/course/introduction-protein-classification-ebi/what-are-protein-signatures/signature-types/what-ar-1

Domain analysis classifies proteins into groups



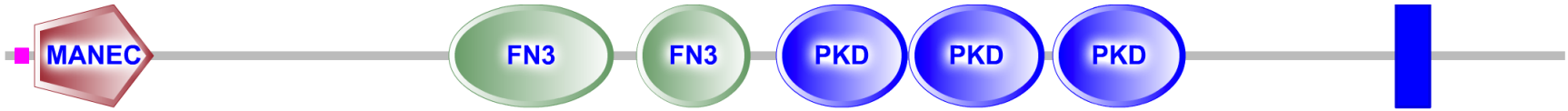


http://2012.igem.org/Team:Slovenia/TheSwitchDesignedTALregulators

http://www.ebi.ac.uk/training/online/course/introduction-protein-classification-ebi/protein-classification/what-are-protein-families

How do you analyze domains?

Pfam



SMART

Pfam

Advantages	Disadvantages
Identifies domains with wide spectrum of cellular functions Can browse clans for homologues	Less sensitive: more false positives and negatives

SMART

Advantages	Disadvantages
More accurate domain identification Domains extensively annotated >100 million protein domains	Less comprehensive: Mostly identifies domains in signaling, extracellular, and chromatin proteins

How do SMART results compare to Pfam?

Pfam

Homo sapiens

SMART

Homo sapiens

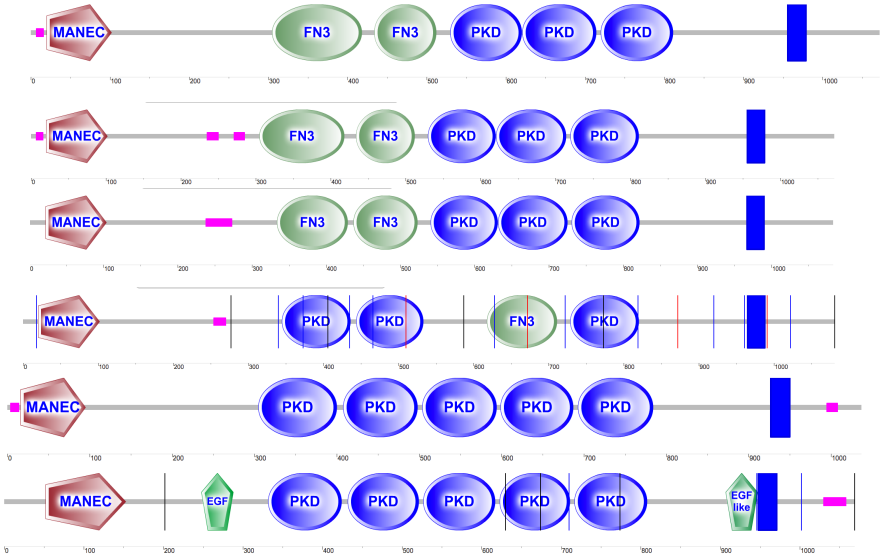
Macaca mulatta

Mus musculus

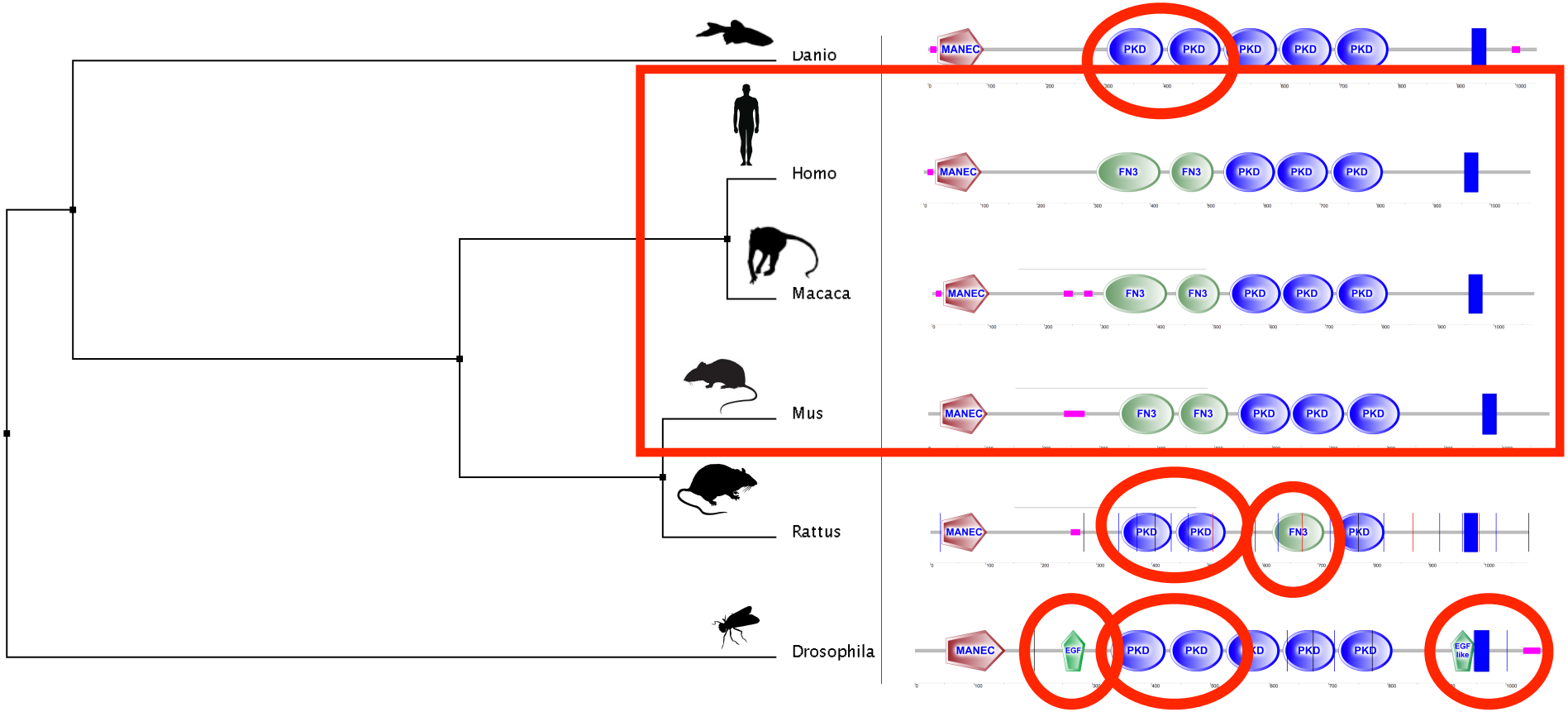
Rattus norvegicus

Danio rerio

Drosophila melanogaster

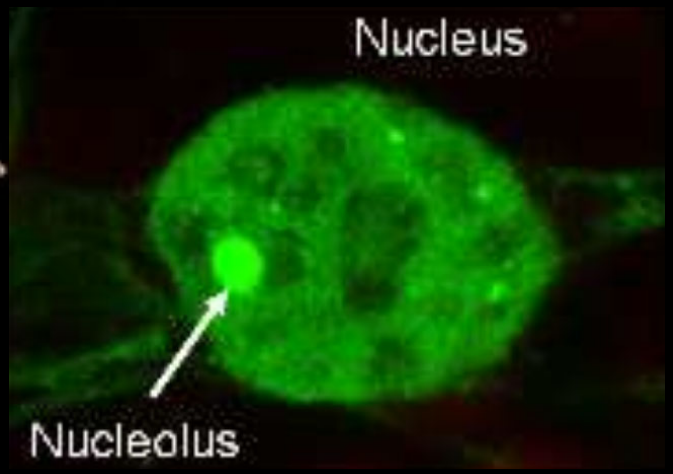


How can you combine phylogeny and domain analysis?



Clustal Omega and SMART

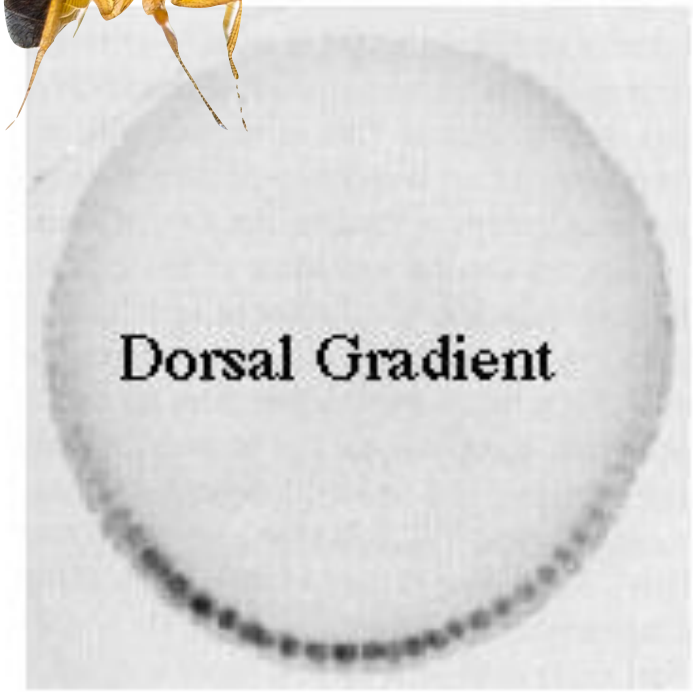
Domain Analysis of the *Nematostella vectenis* SNAIL ortholog reveals unique nucleolar localization that depends on zinc-finger domains

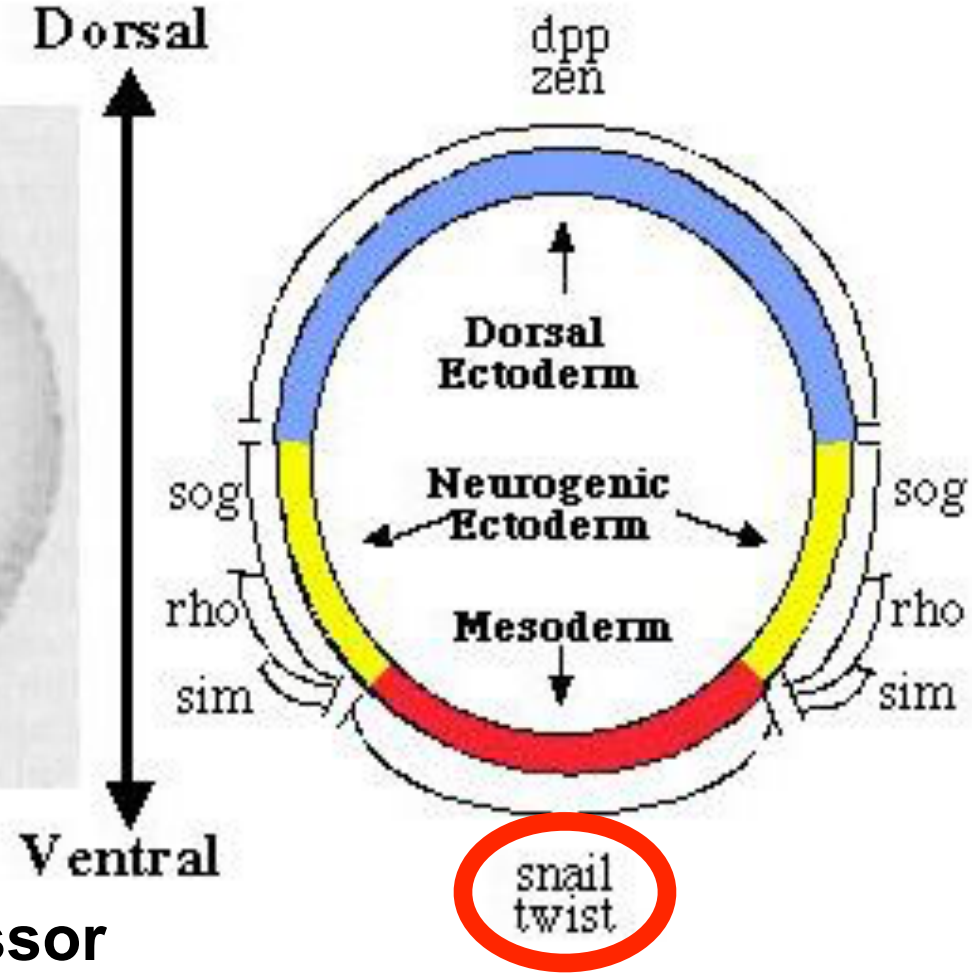


Ada A. Dattoli, Mark A. Hink, Timothy Q. DuBuc, Bram J. Teunisse, Joachim Goedhart, Eric Rottinger, Marten Postma

What do Cnidarians look like?

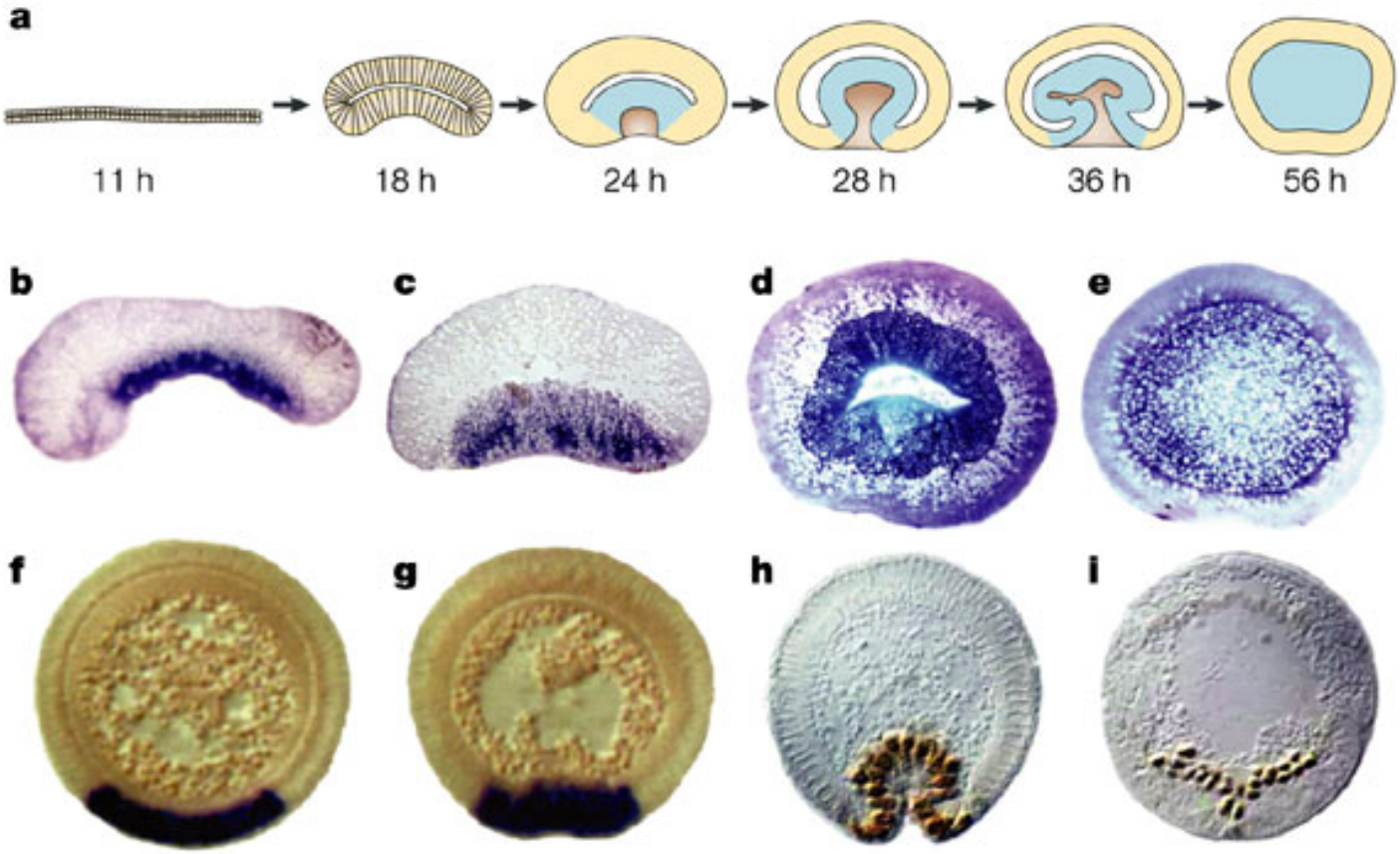
How did the SNAIL protein get it's name?





Transcriptional repressor
First identified in *Drosophila*
Required for mesoderm formation (gastrulation)

Why are they studying Cnidarians?

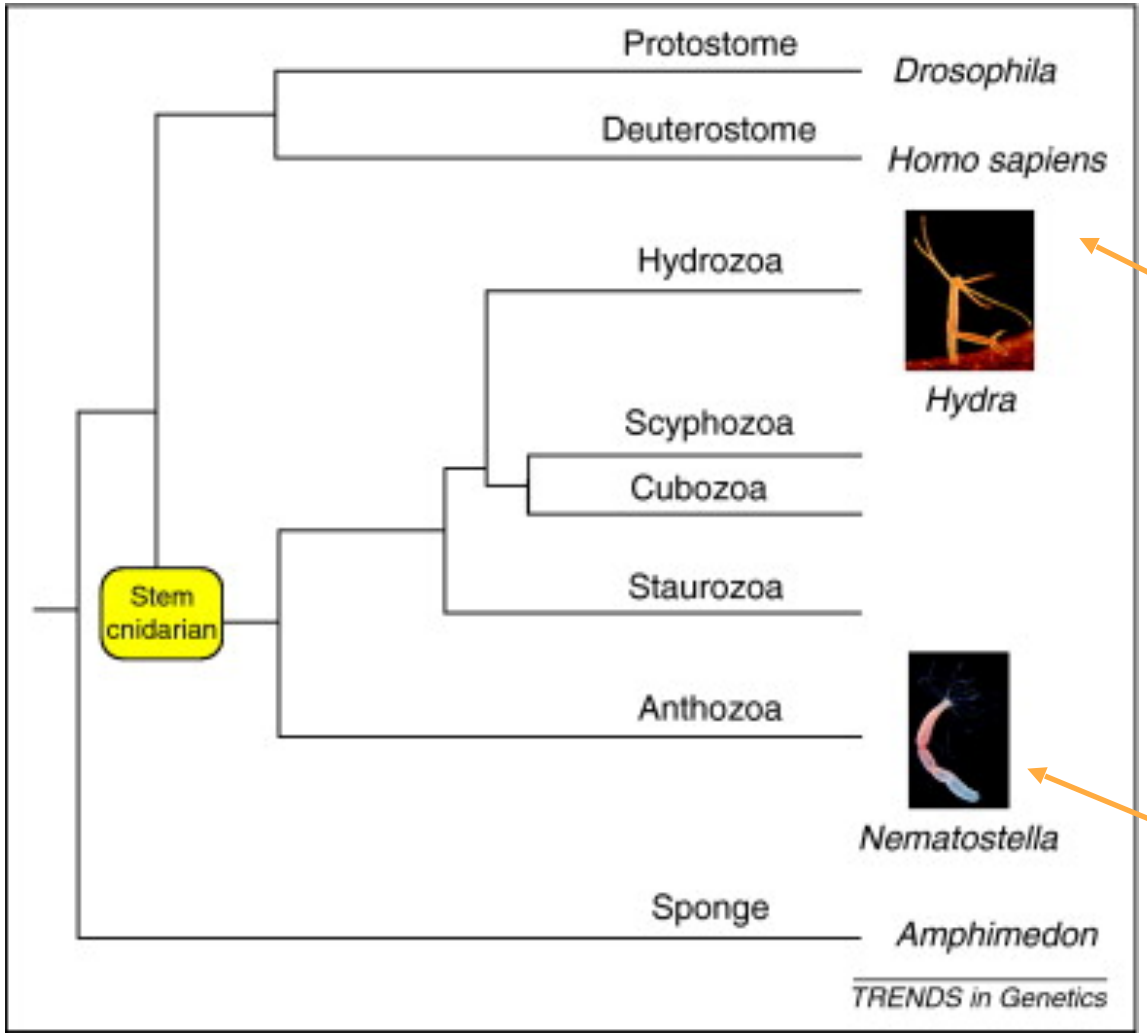


Nature Reviews | Genetics

Cnidarians don’t have a mesoderm layer

http://www.nature.com/nrg/journal/v5/n8/fig_tab/nrg1402_F2.html

Which SNAIL Proteins were studied?



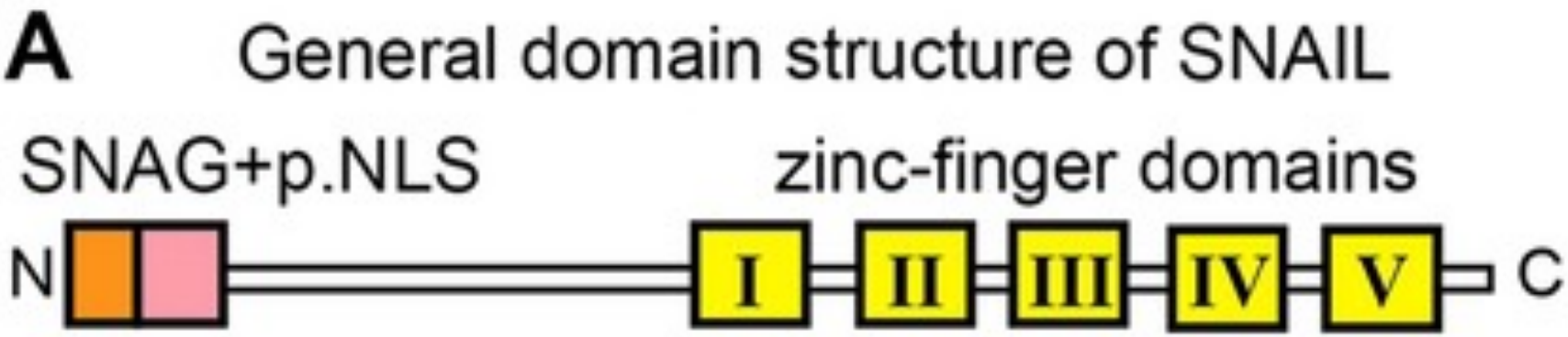
Human SNAIL proteins→ *Hs*SNAIL1

***Hs*SNAIL2**

Cnidarian SNAIL proteins→ *Nv*SNAILA

***Nv*SNAILB**

Fig 1a: What does the SNAIL protein look like?

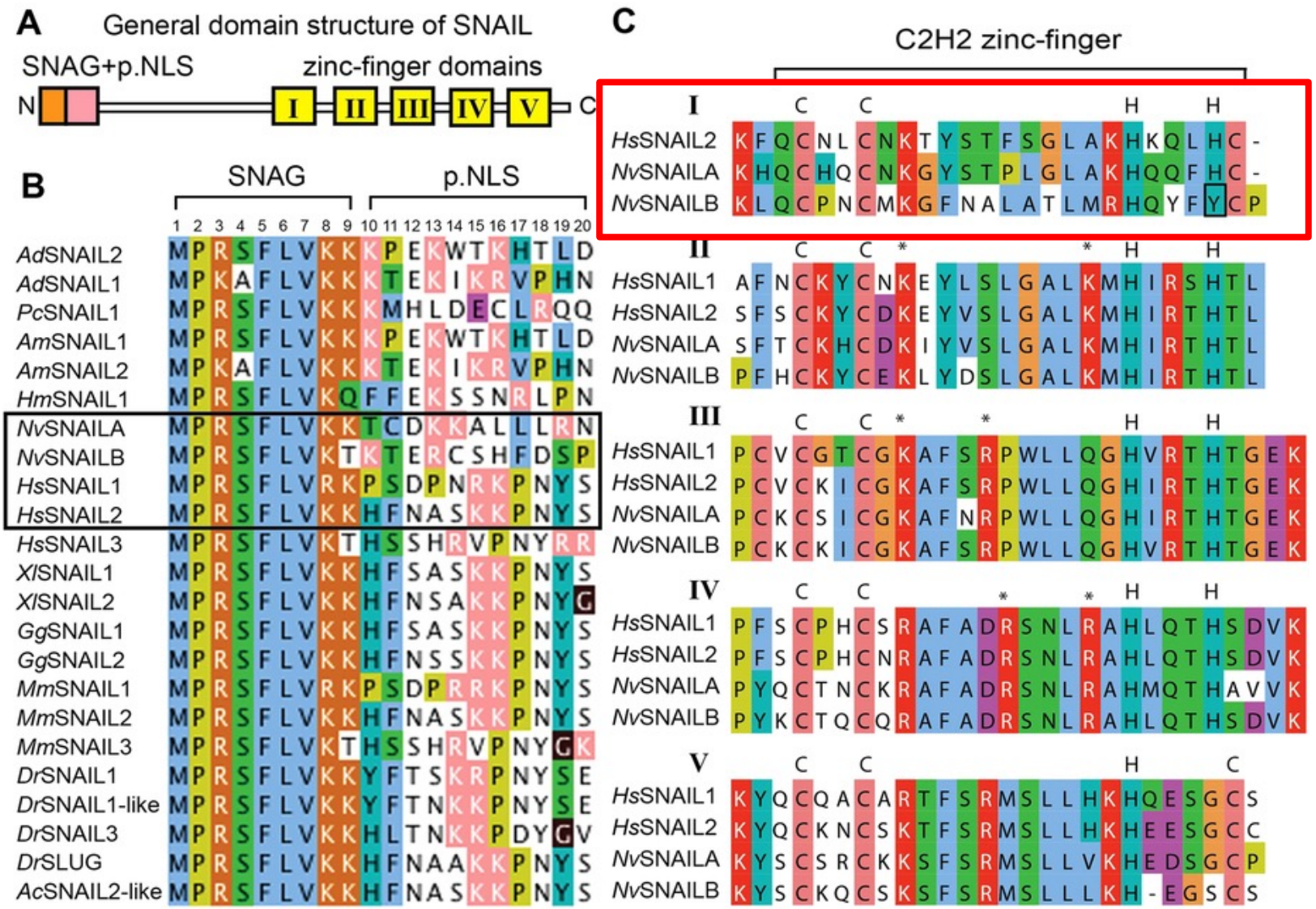


SNAG

Zn Finger

NLS

Fig 1: How were the SNAIL homologues determined?



*Nv*SNAILB has an incomplete <u>first</u> Zn Finger Domain

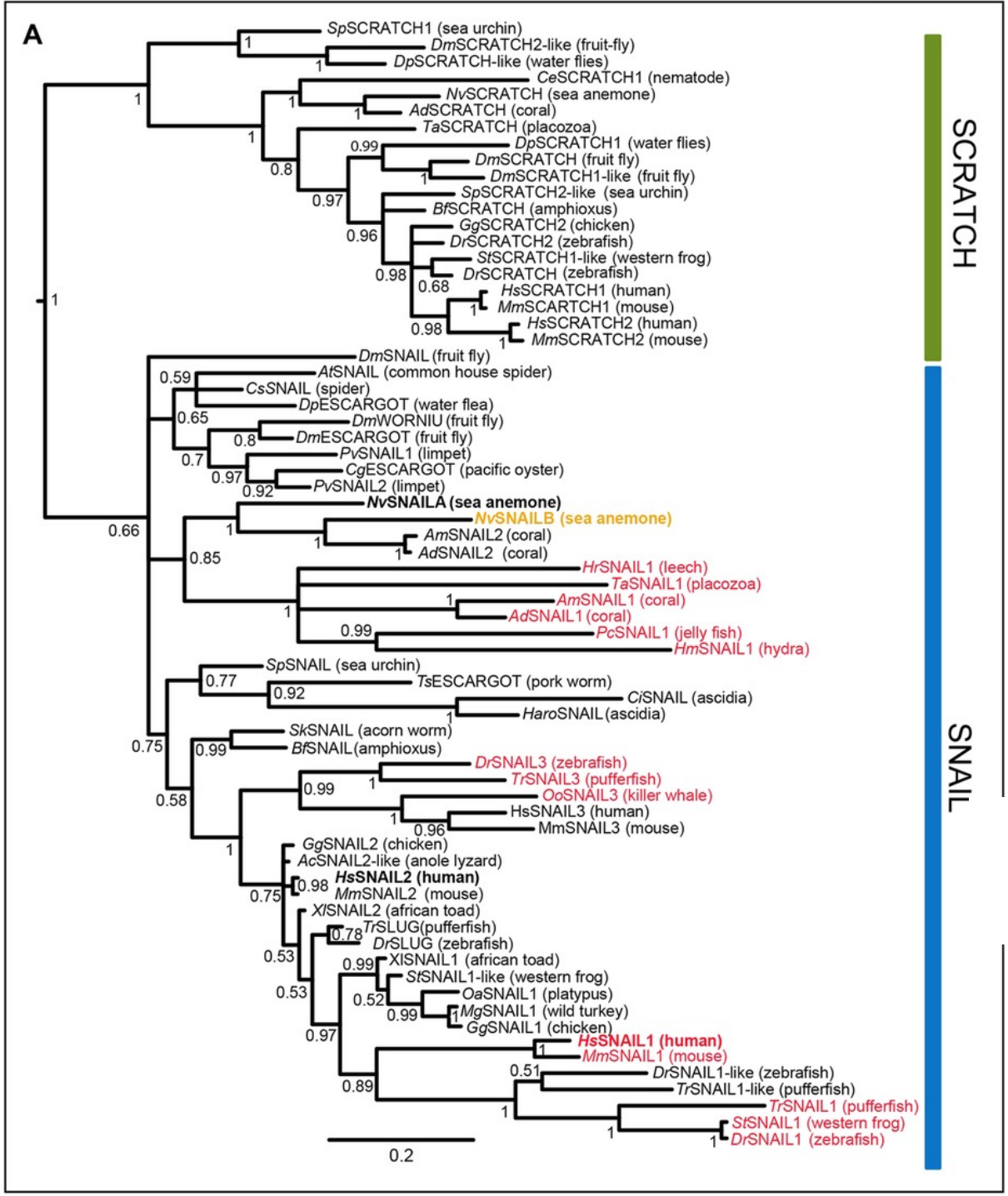
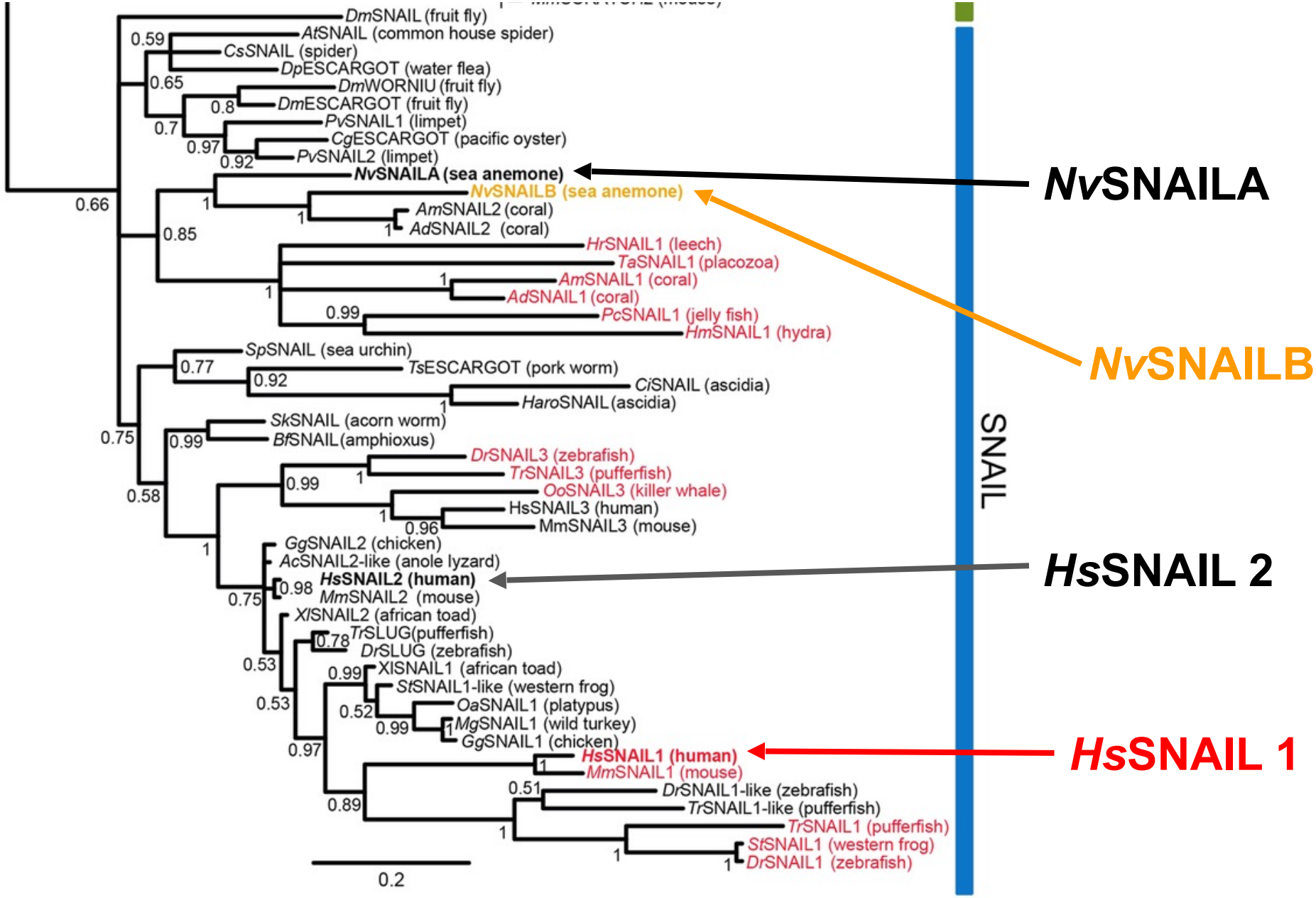
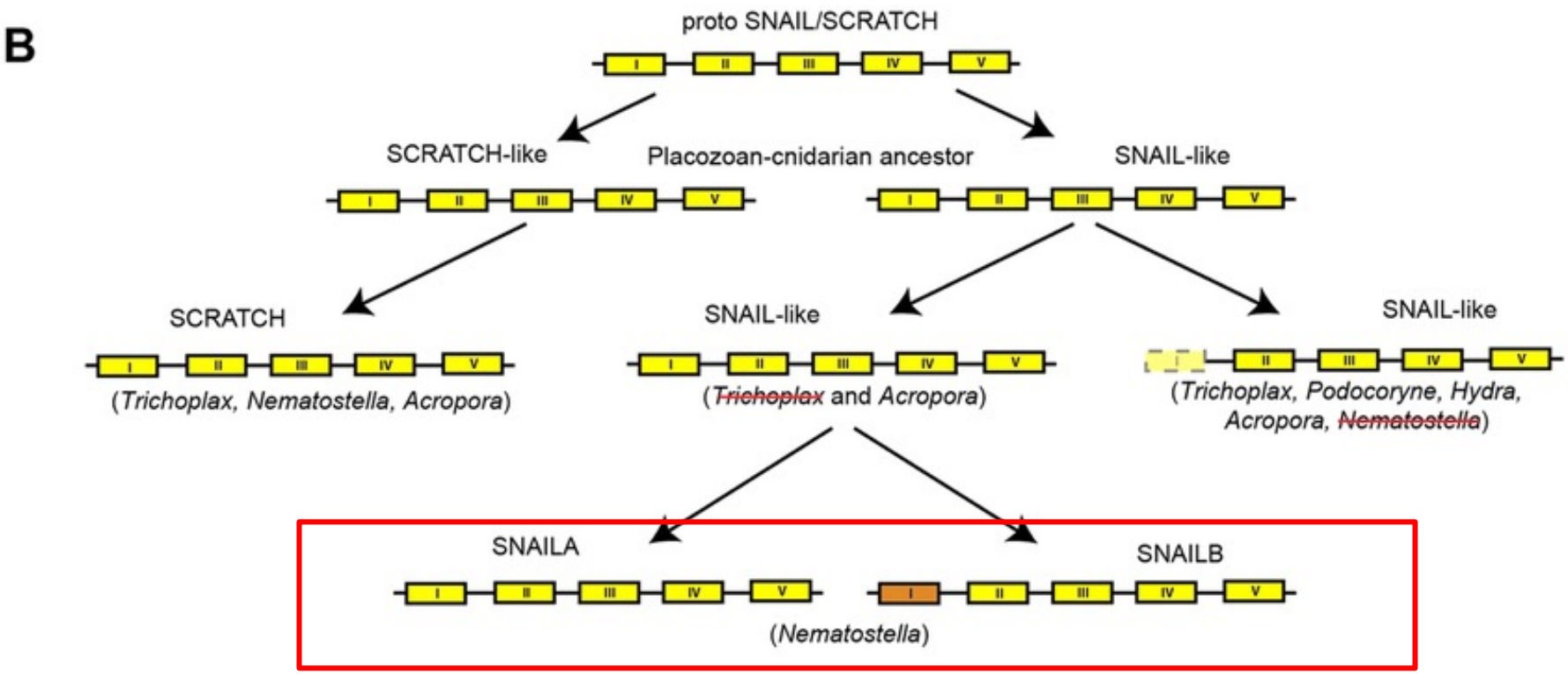


Fig 2a: How did the SNAIL superfamily evolve?

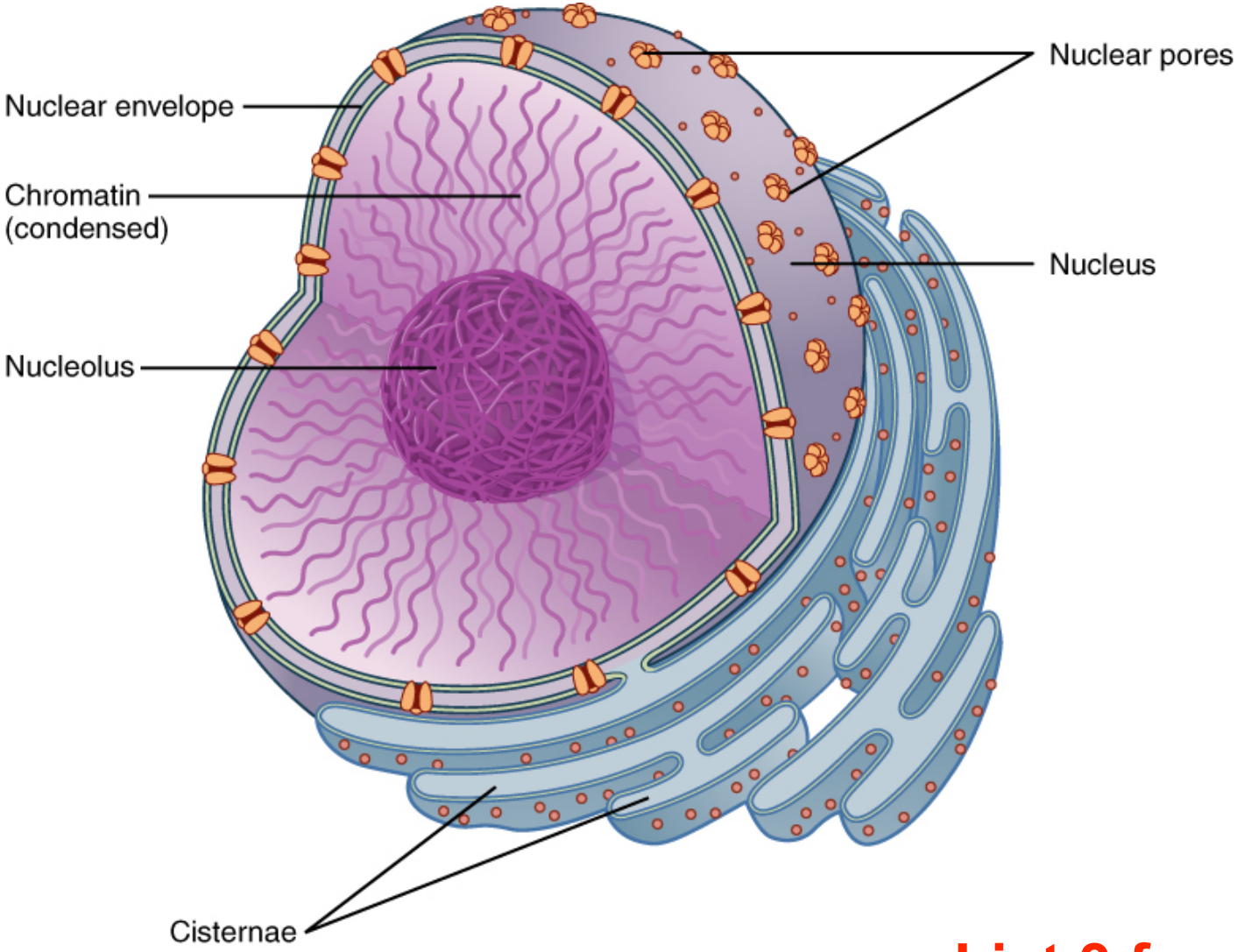


2A: How are the SNAIL proteins related?

Figure 2B: How did the two SNAIL protein types arise?

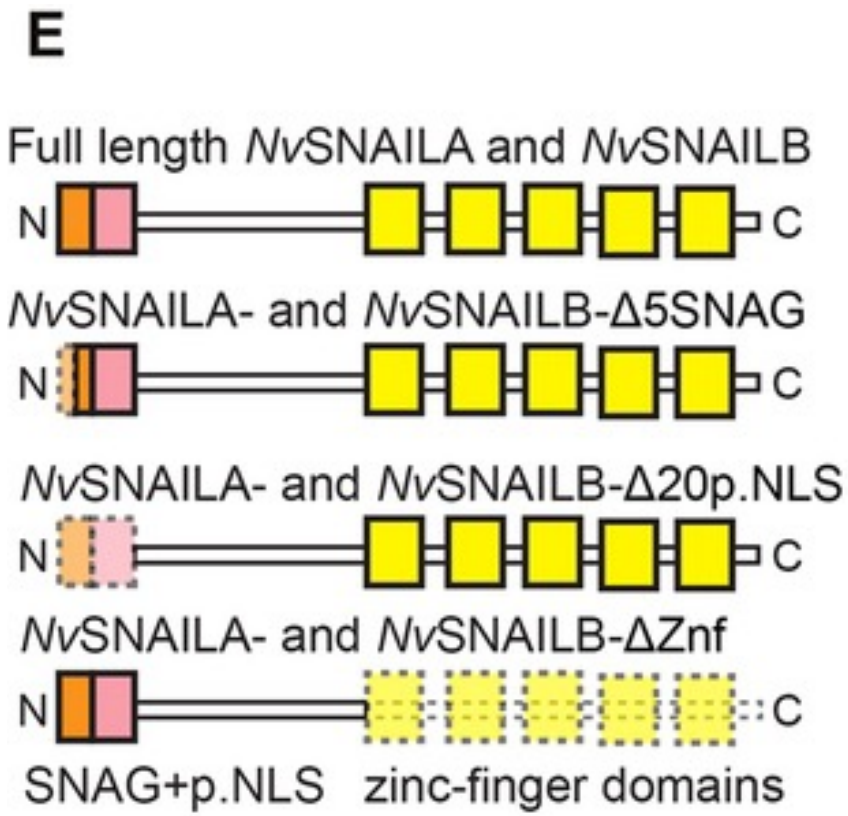


What is a nucleolus?



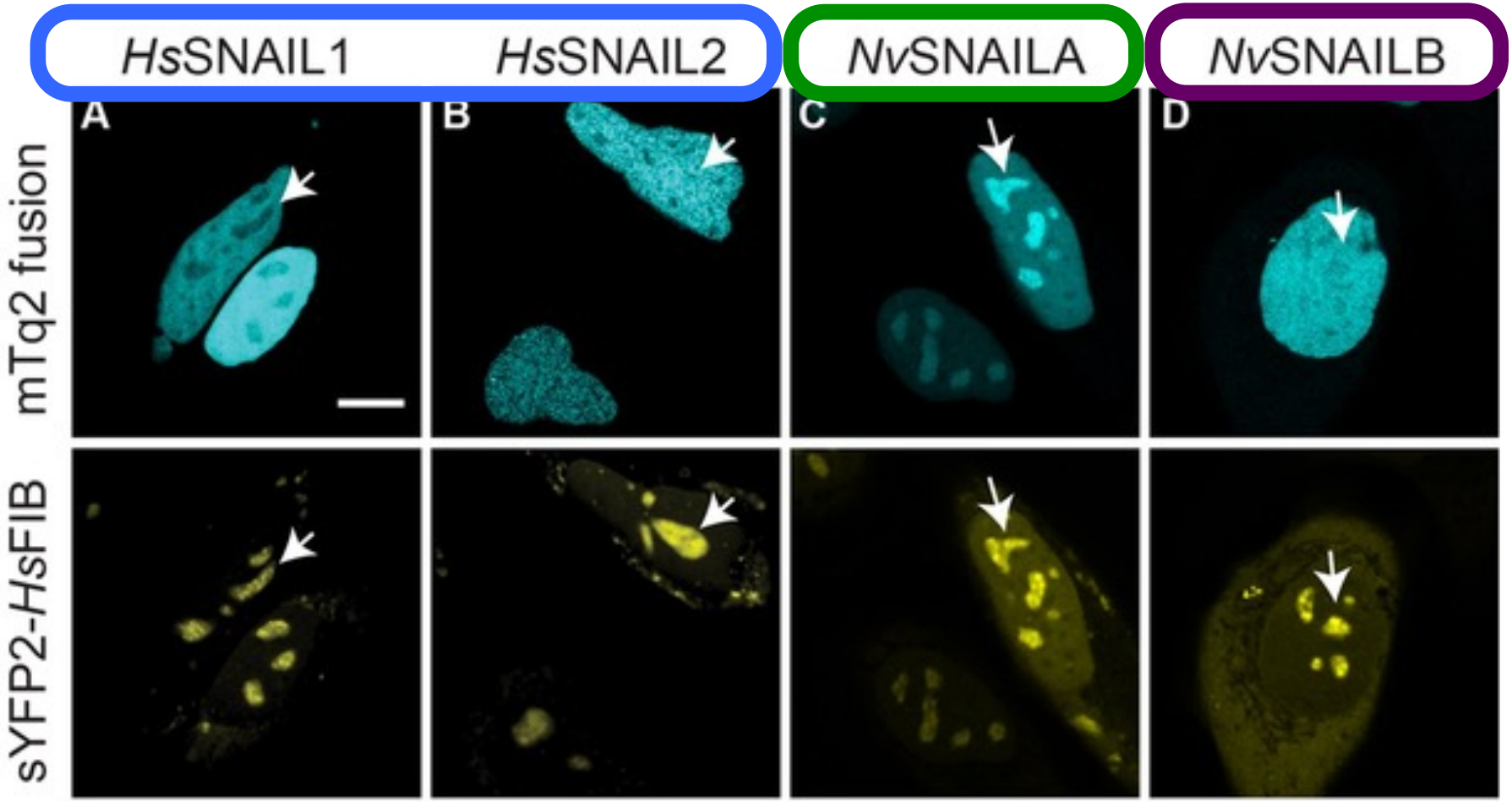
List 3 functions

Fig 3E: What kinds of mutants were constructed?



Deletion Type (Δ)
Full Length
First five AAs deleted from SNAG Domain
SNAG domain and putative NLS
All of the Zn-finger domains

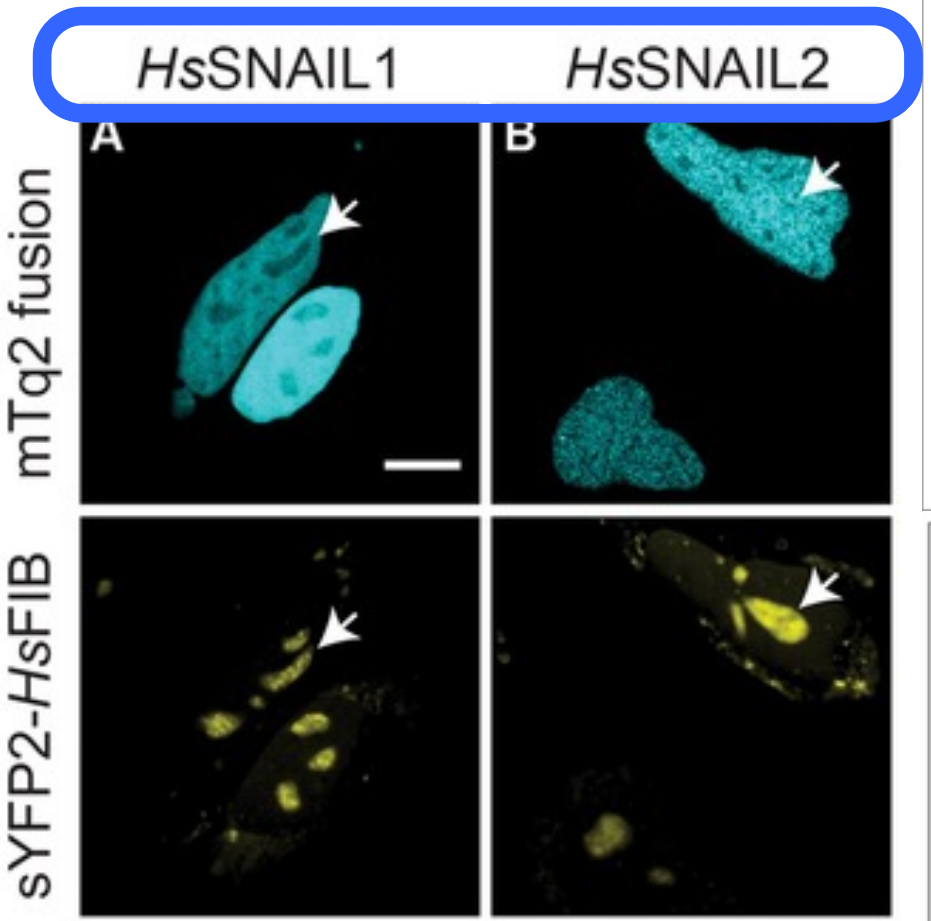
Fig 3A-D: Where does *Hs*SNAIL1,2 and *Nv*SNAILA,B localize?



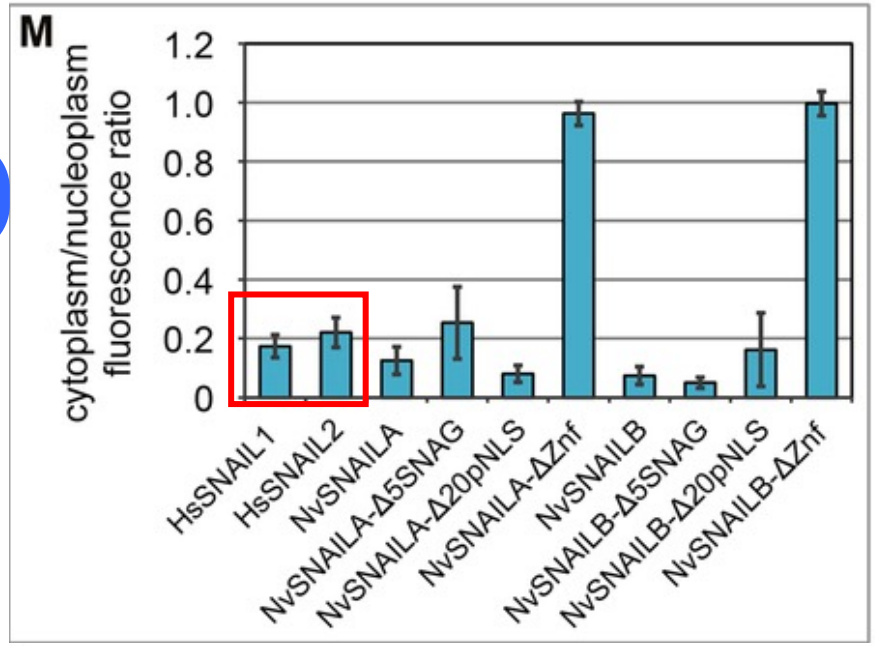
All exhibit strong <u>nuclear</u> localization

nuclear
nucleolar

Fig 3A-B: How do *Hs*SNAIL1 and 2 normally Localize?



nuclear
nucleolar



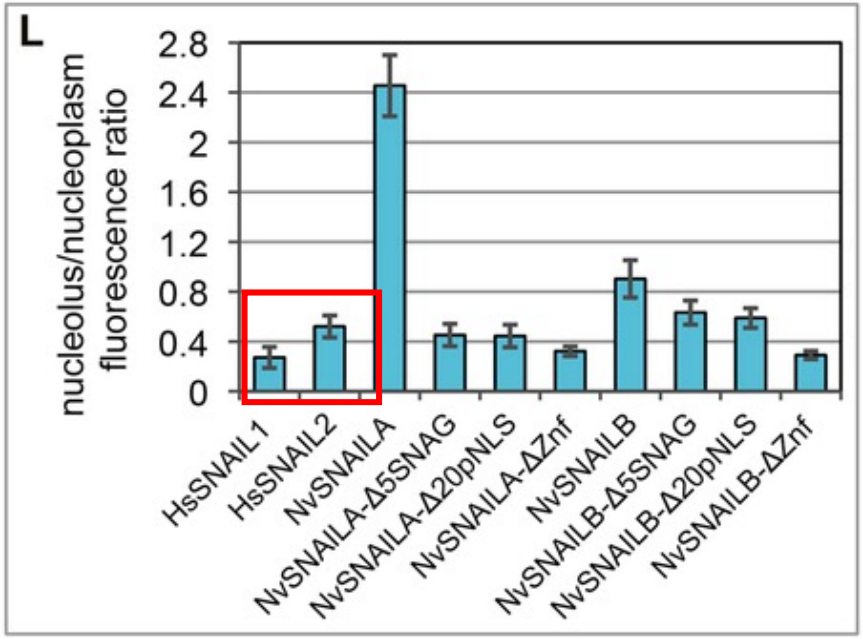
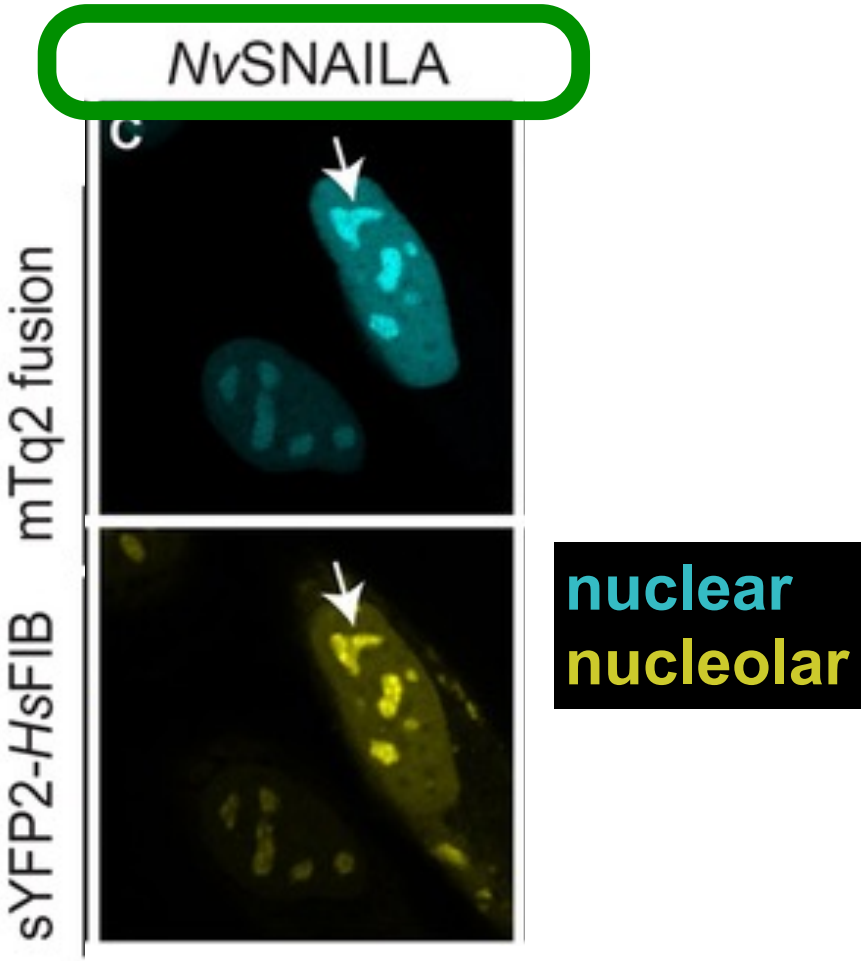
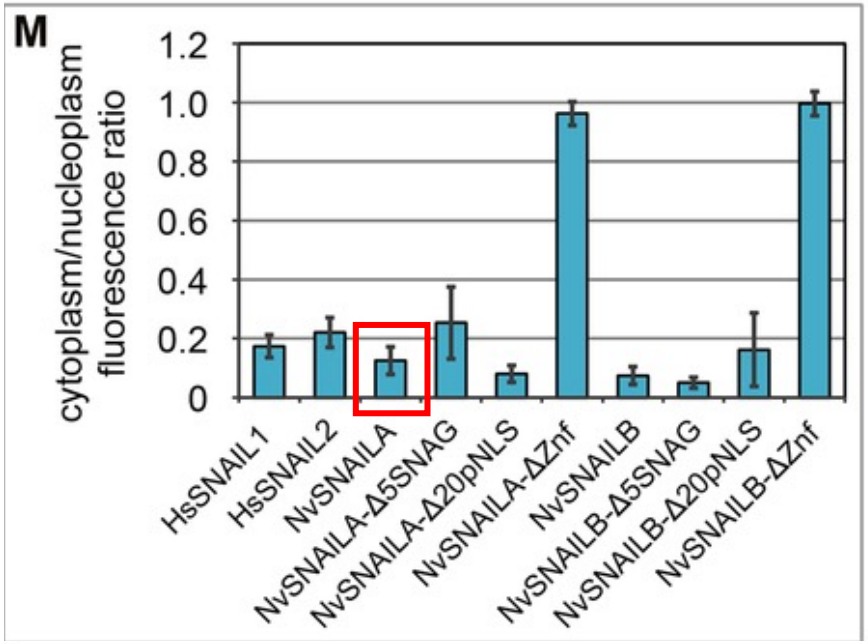
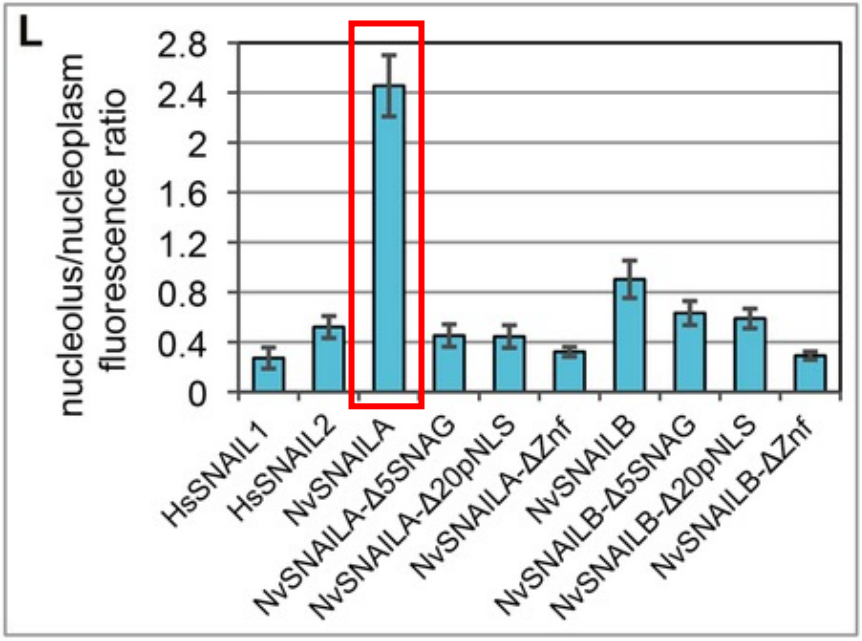
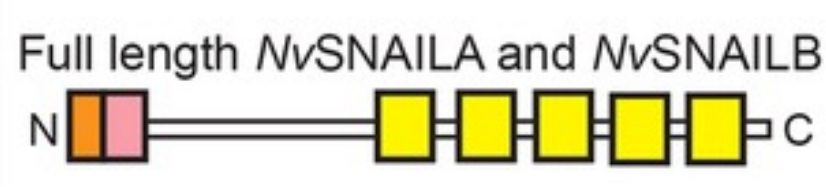


Fig 3C: How does *NvSNAILA* normally localize?



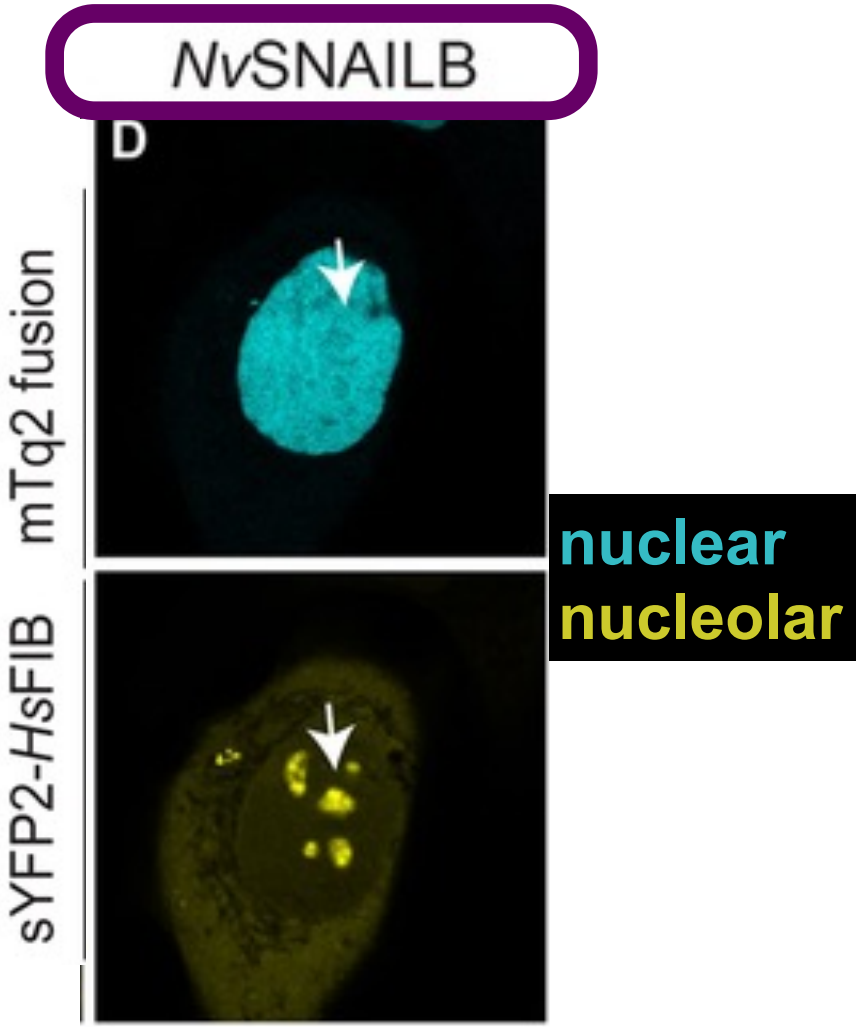


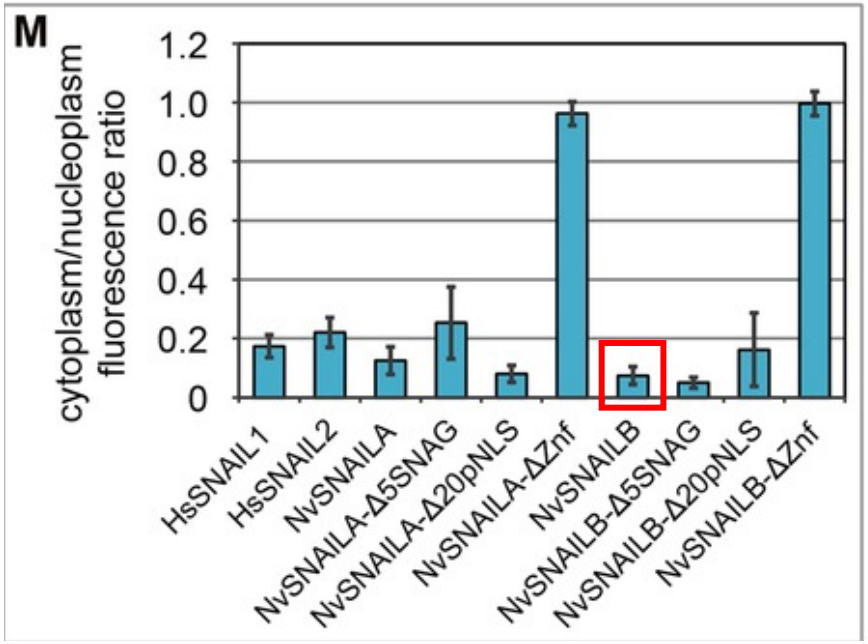


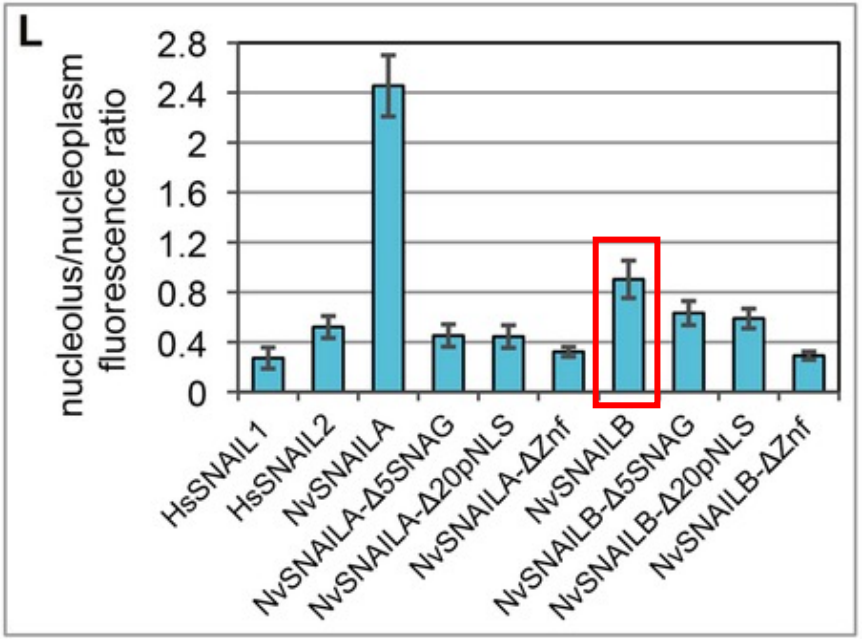


*Nv*SNAILA shows strong **Nucleolar** Localization

Fig 3D: How does *Nv*SNAILB normally localize?







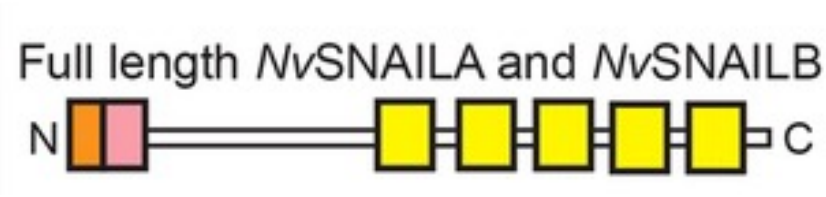
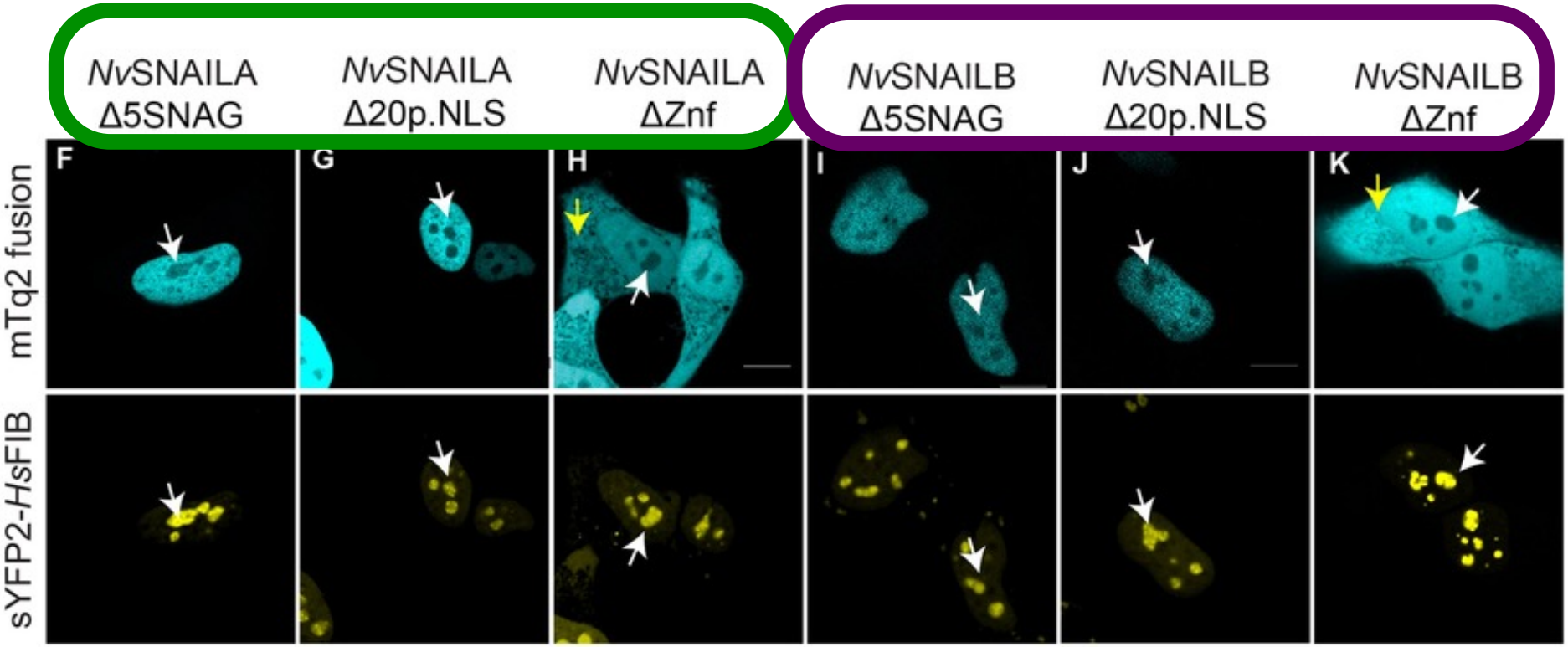
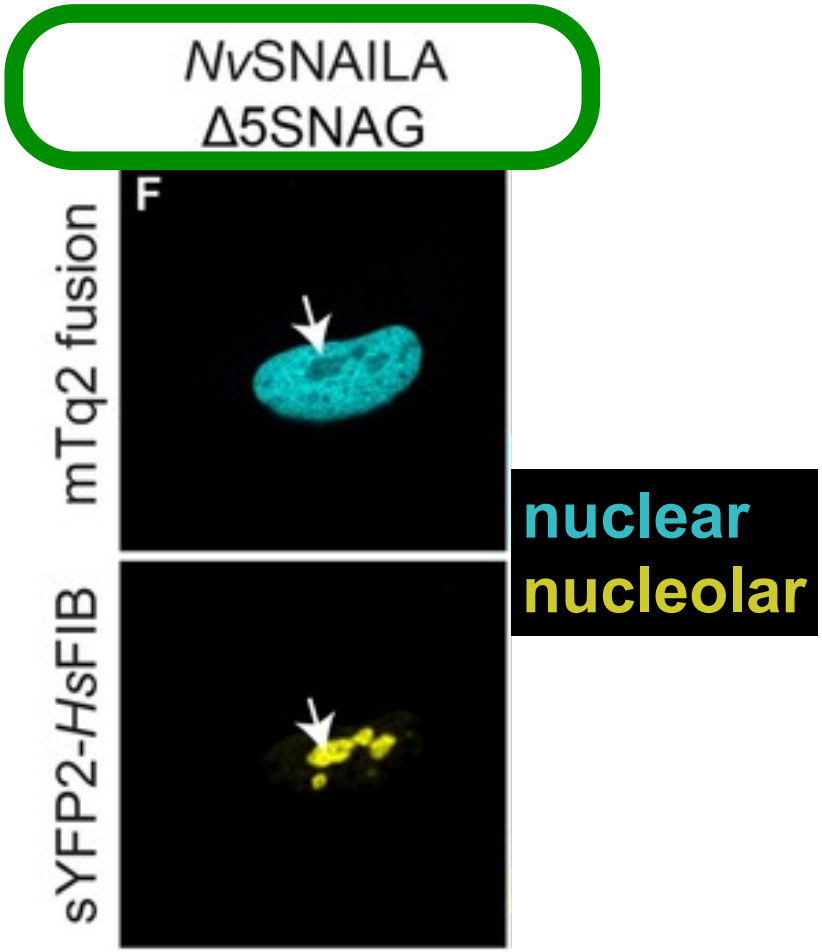


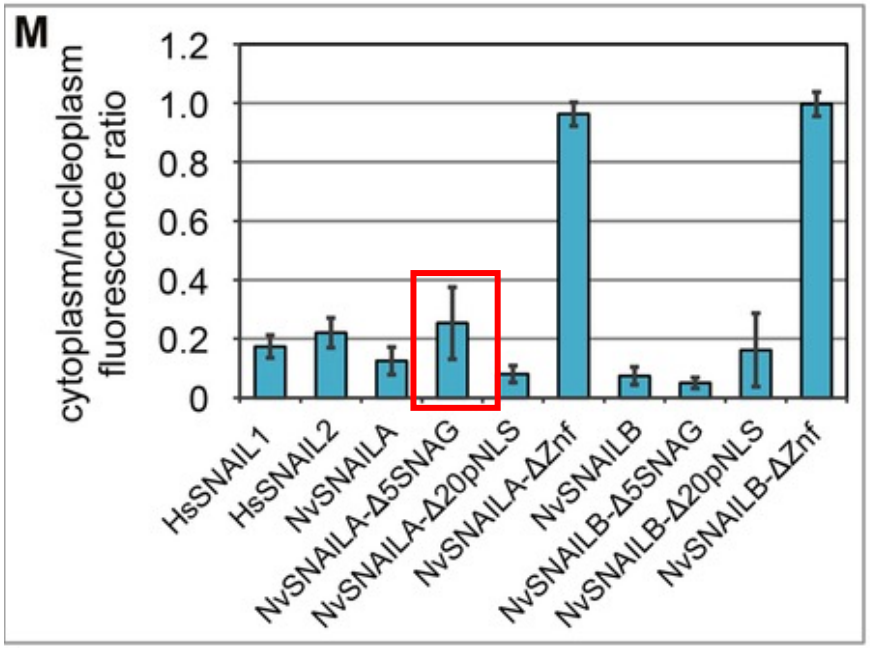
Fig 3F-K: How do the SNAIL mutants localize?

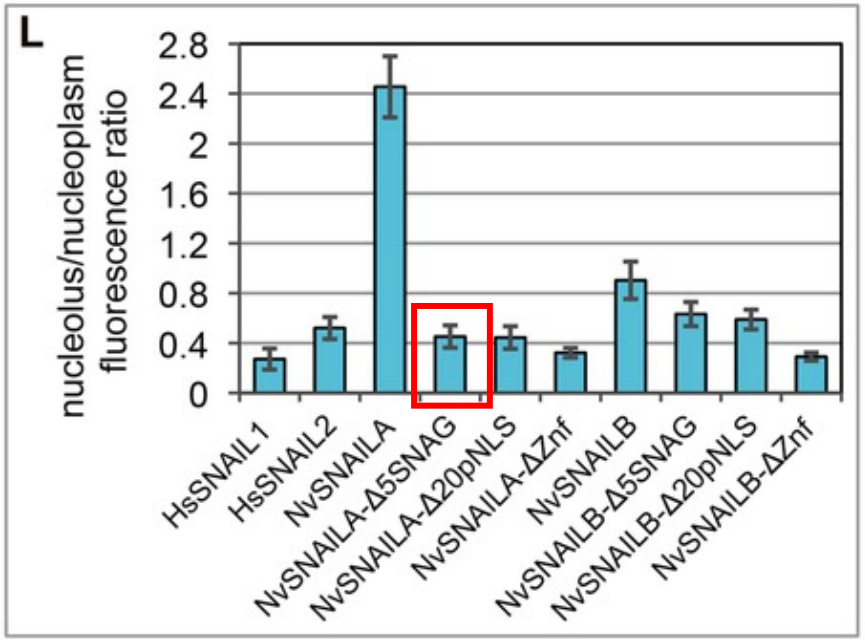


nuclear
nucleolar

Fig 3F: How does the *Nv*SNAILA-Δ5SNAG mutant localize?







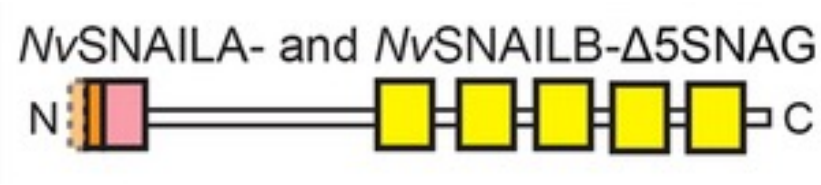
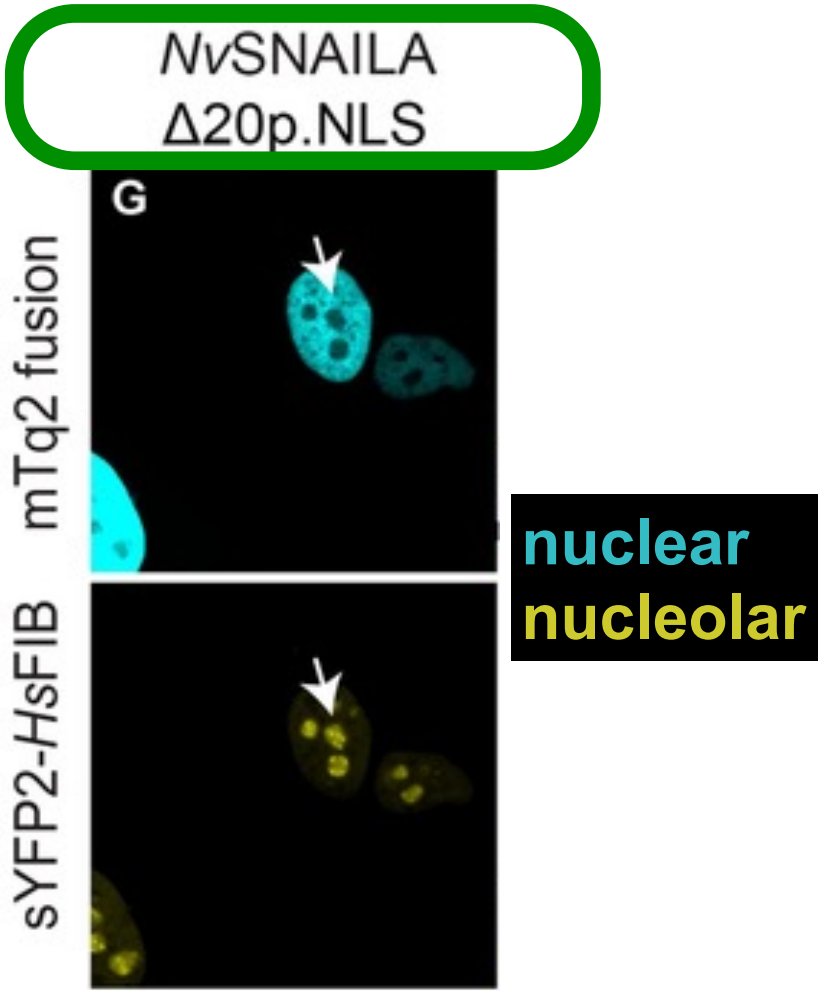
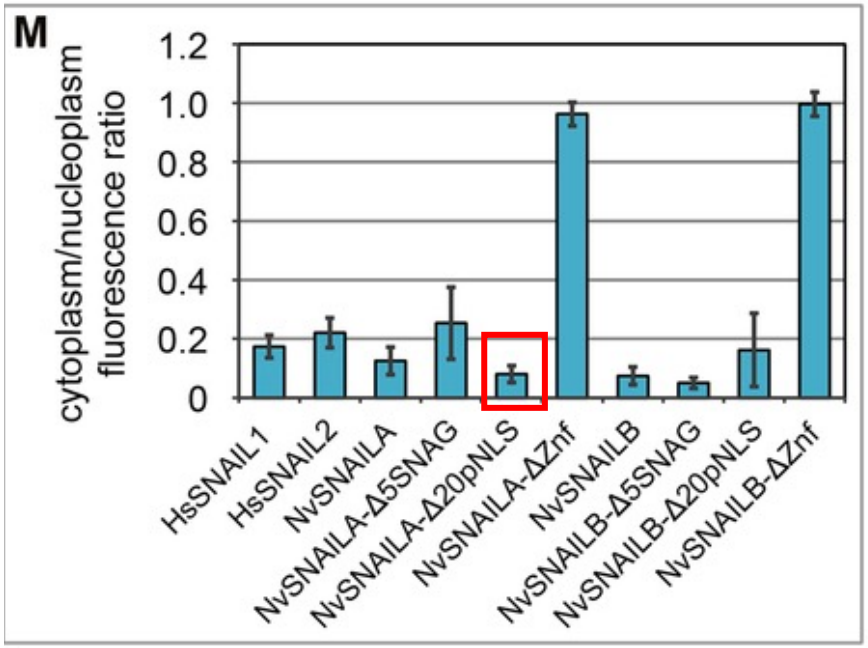
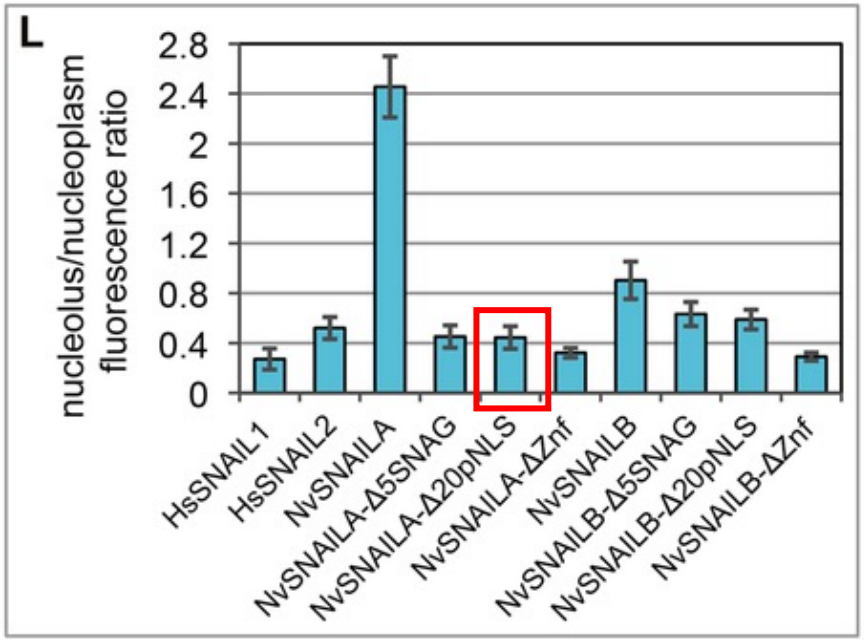


Fig 3G: How does the *Nv*SNAILA-Δ20p.NLS mutants localize?







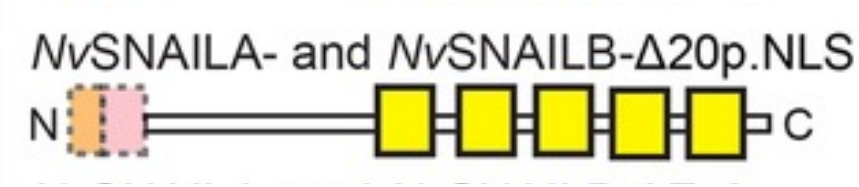
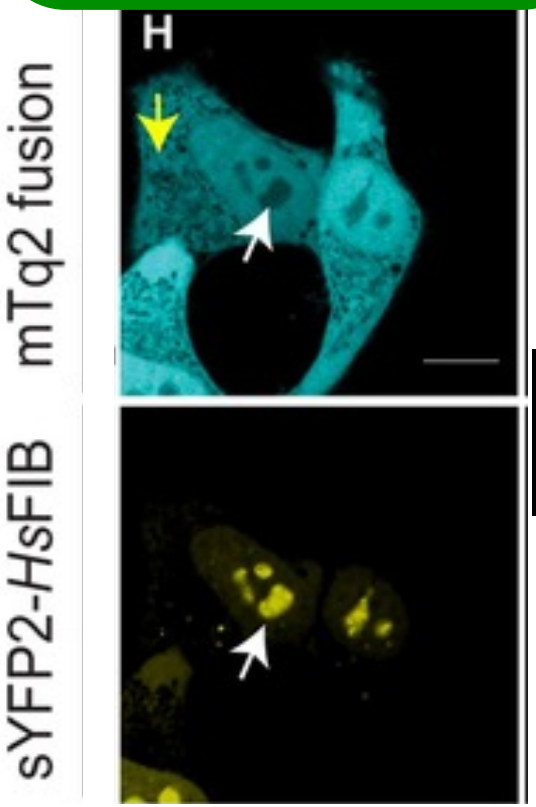
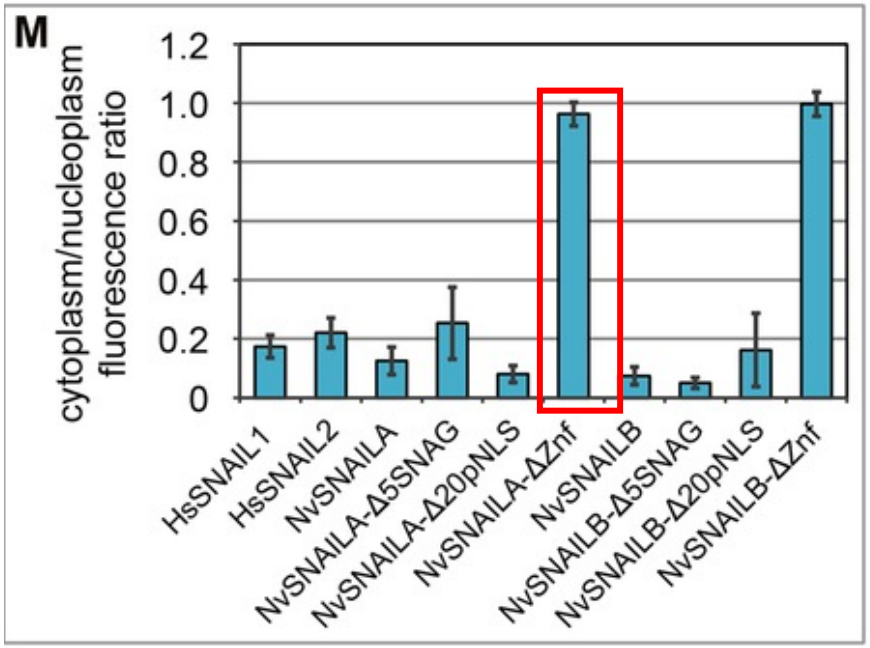


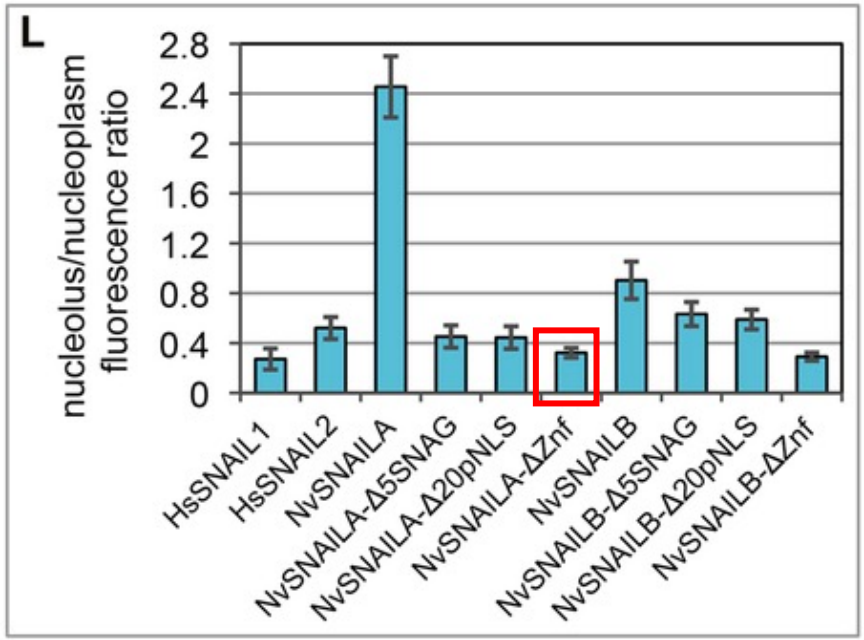
Fig 3H: How does the *Nv*SNAILA-ΔZnF mutant localize?

*Nv*SNAILA
ΔZnf



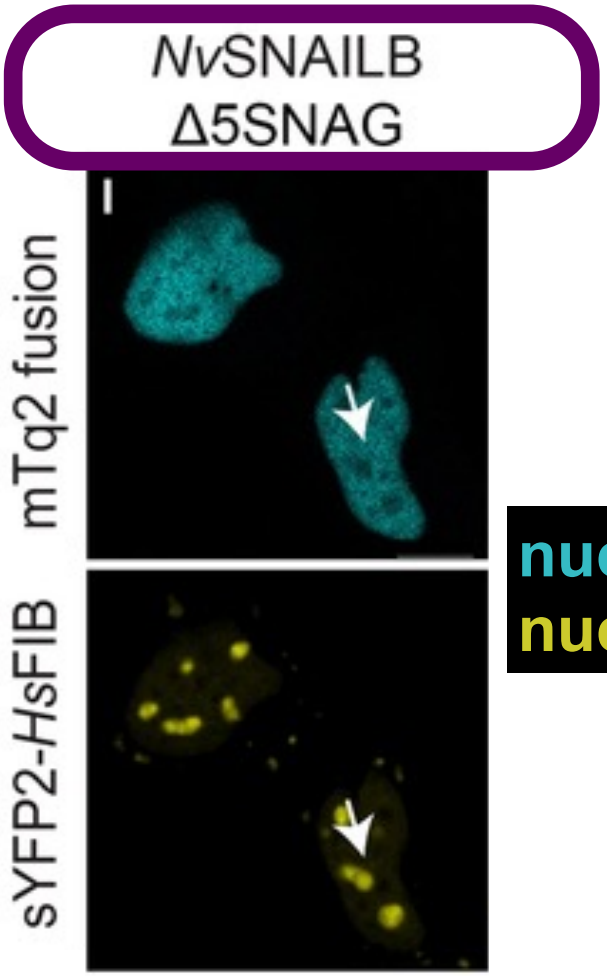
nuclear
nucleolar



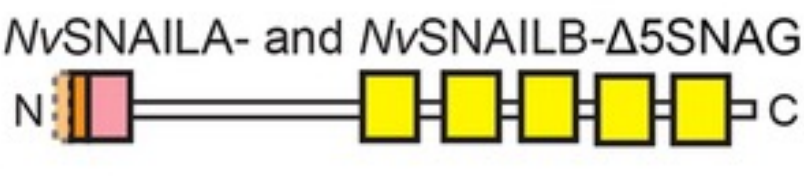


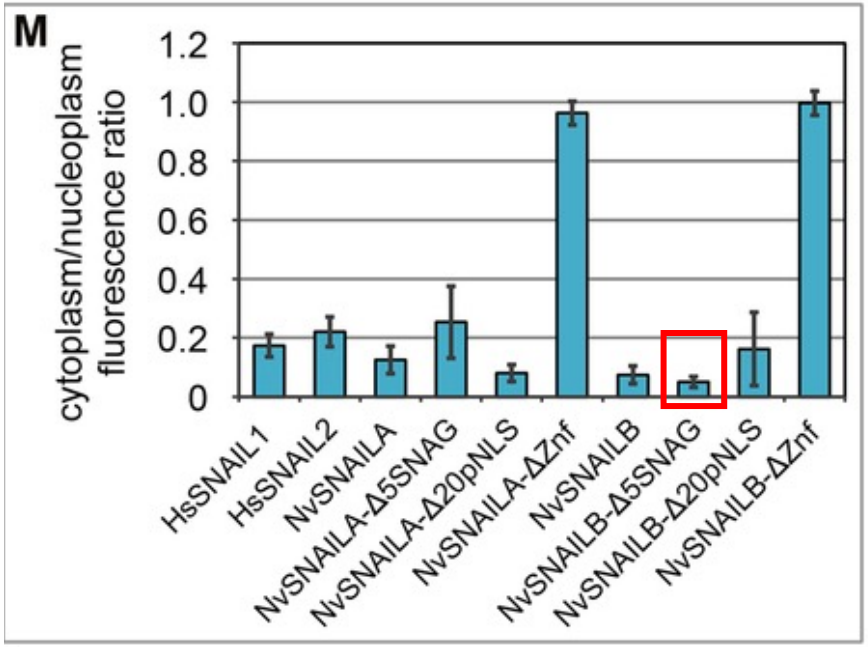
*Nv*SNAILA- and *Nv*SNAILB-ΔZnf

Fig 3I: How does the *Nv*SNAILB-Δ5SNAG mutant localize?



nuclear
nucleolar





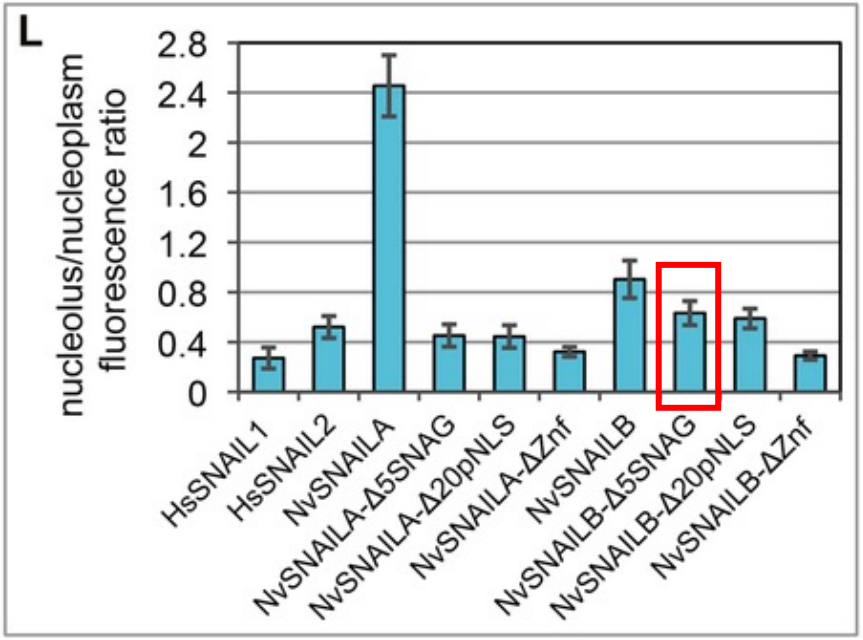
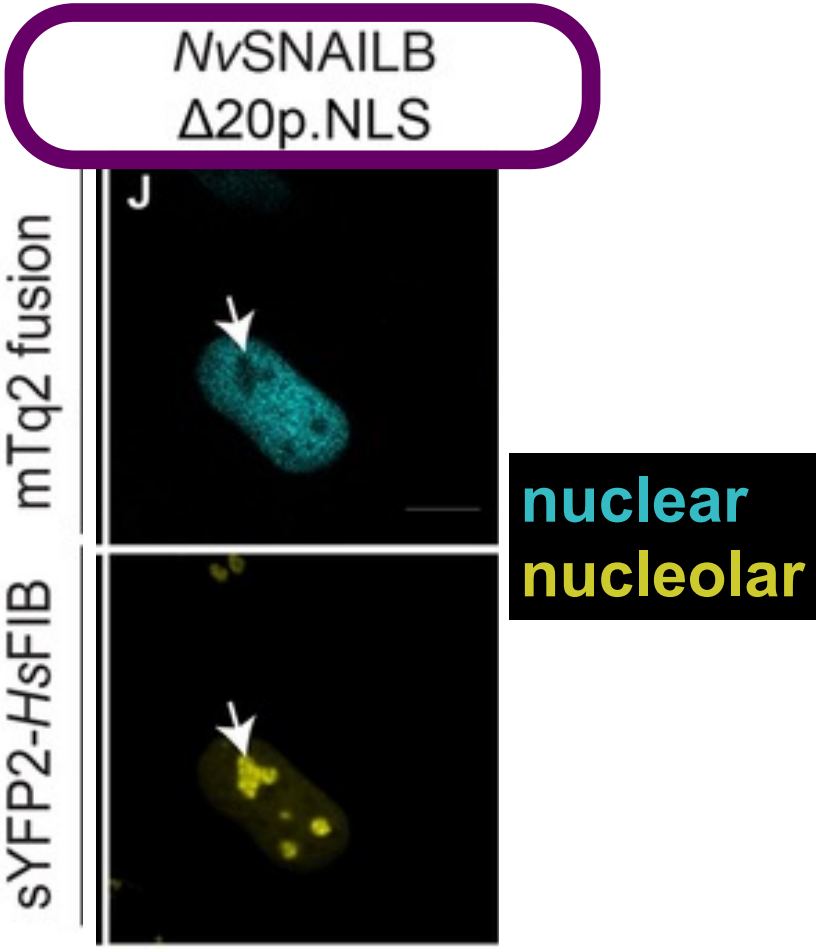
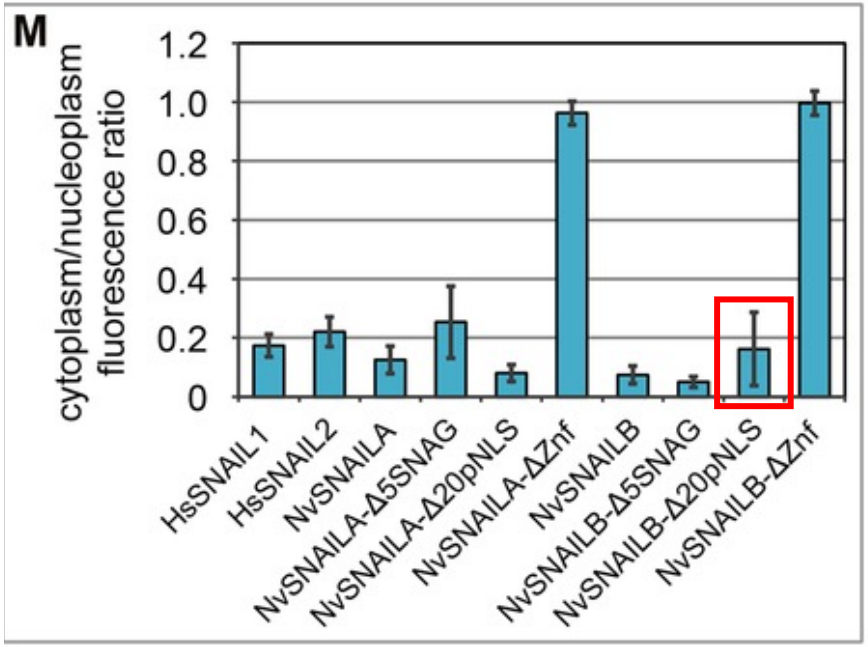
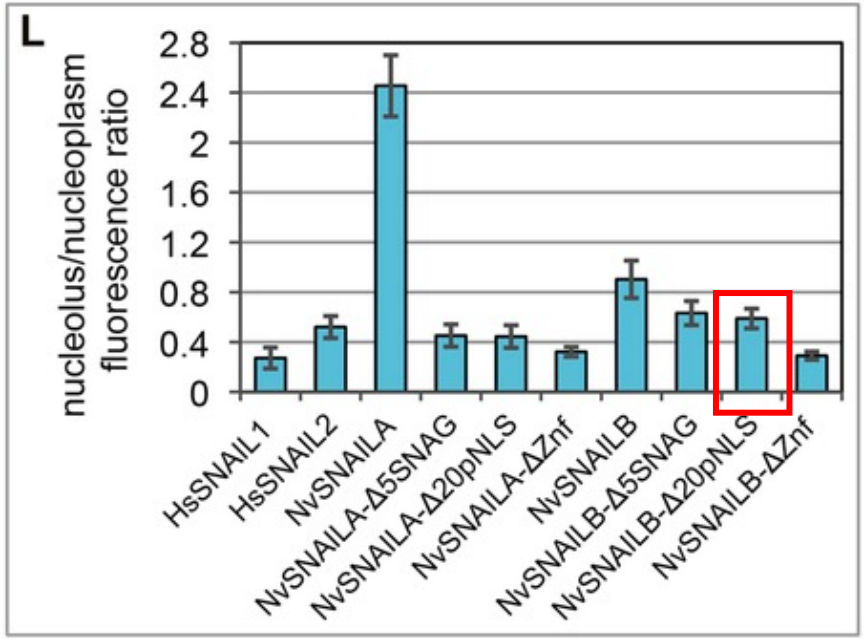


Fig 3J: How does the *Nv*SNAILB-Δ20p.NLS mutant localize?







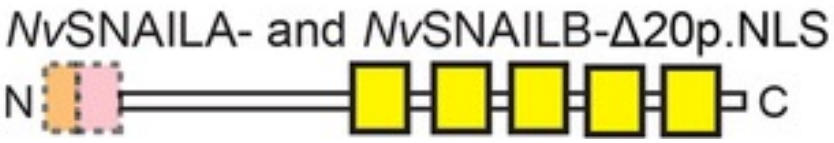
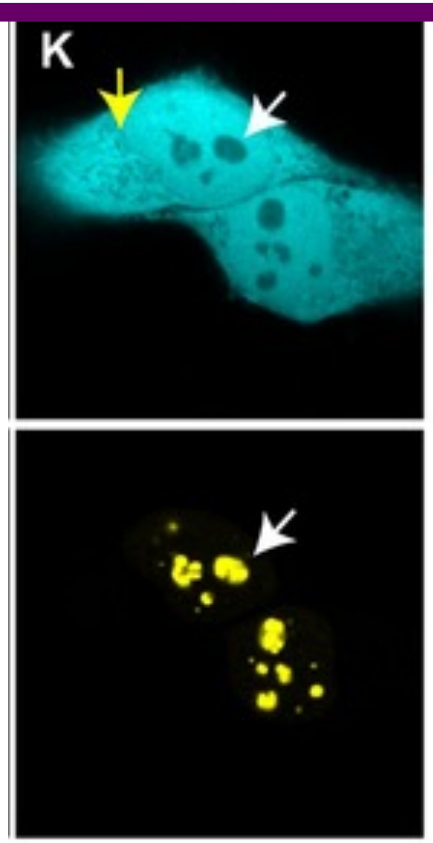
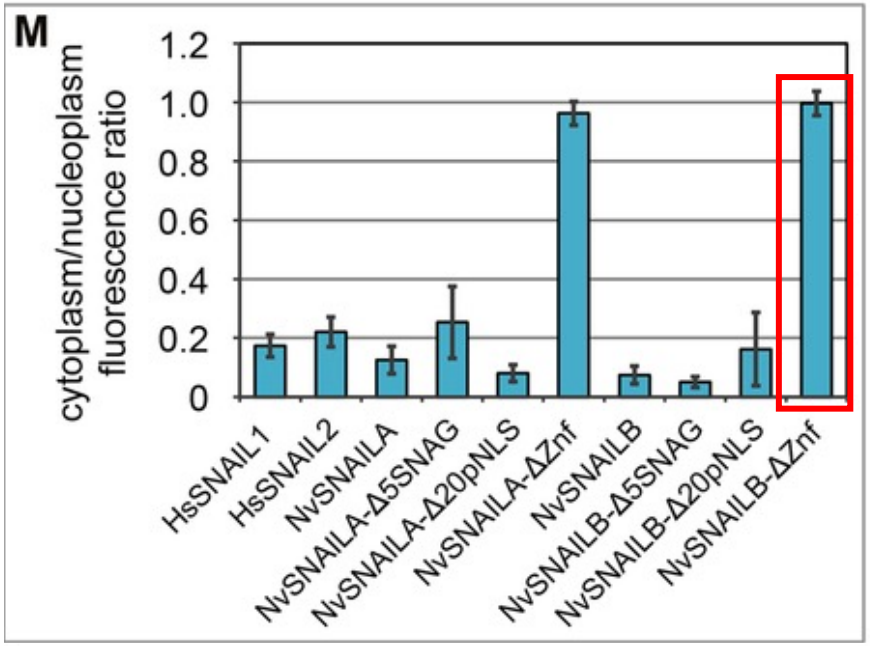


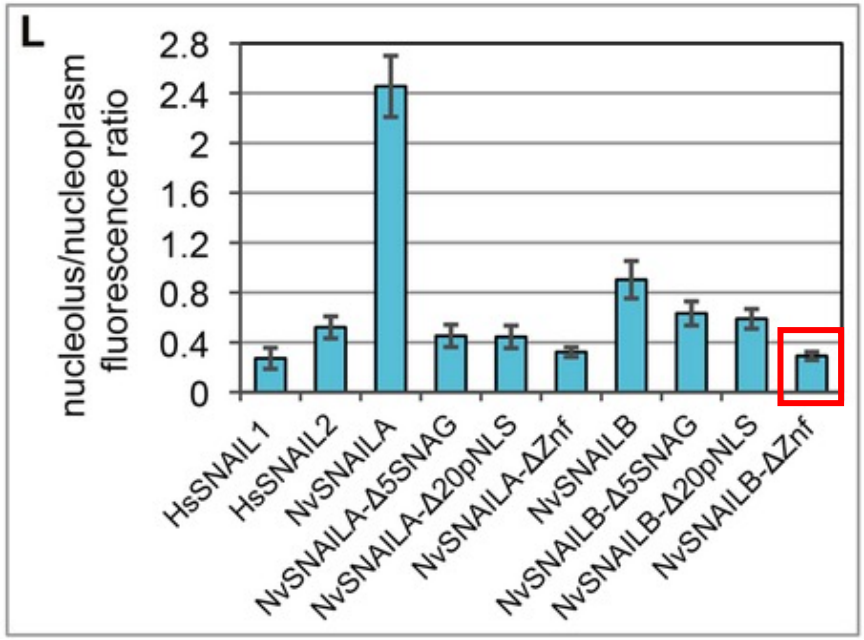
Fig 3K: How does the *Nv*SNAILB-ΔZnF mutant localize?

*Nv*SNAILB
ΔZnf



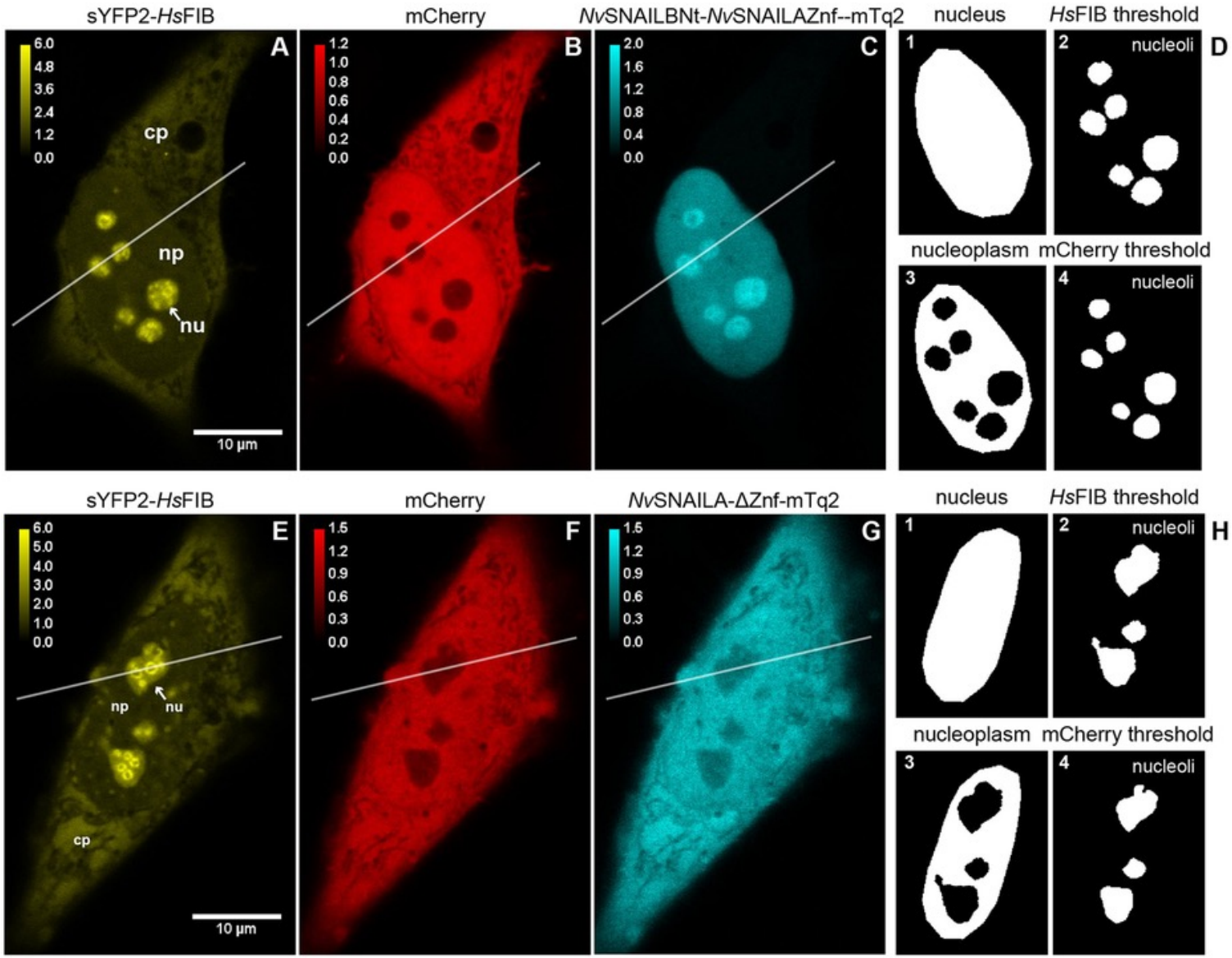
nuclear
nucleolar





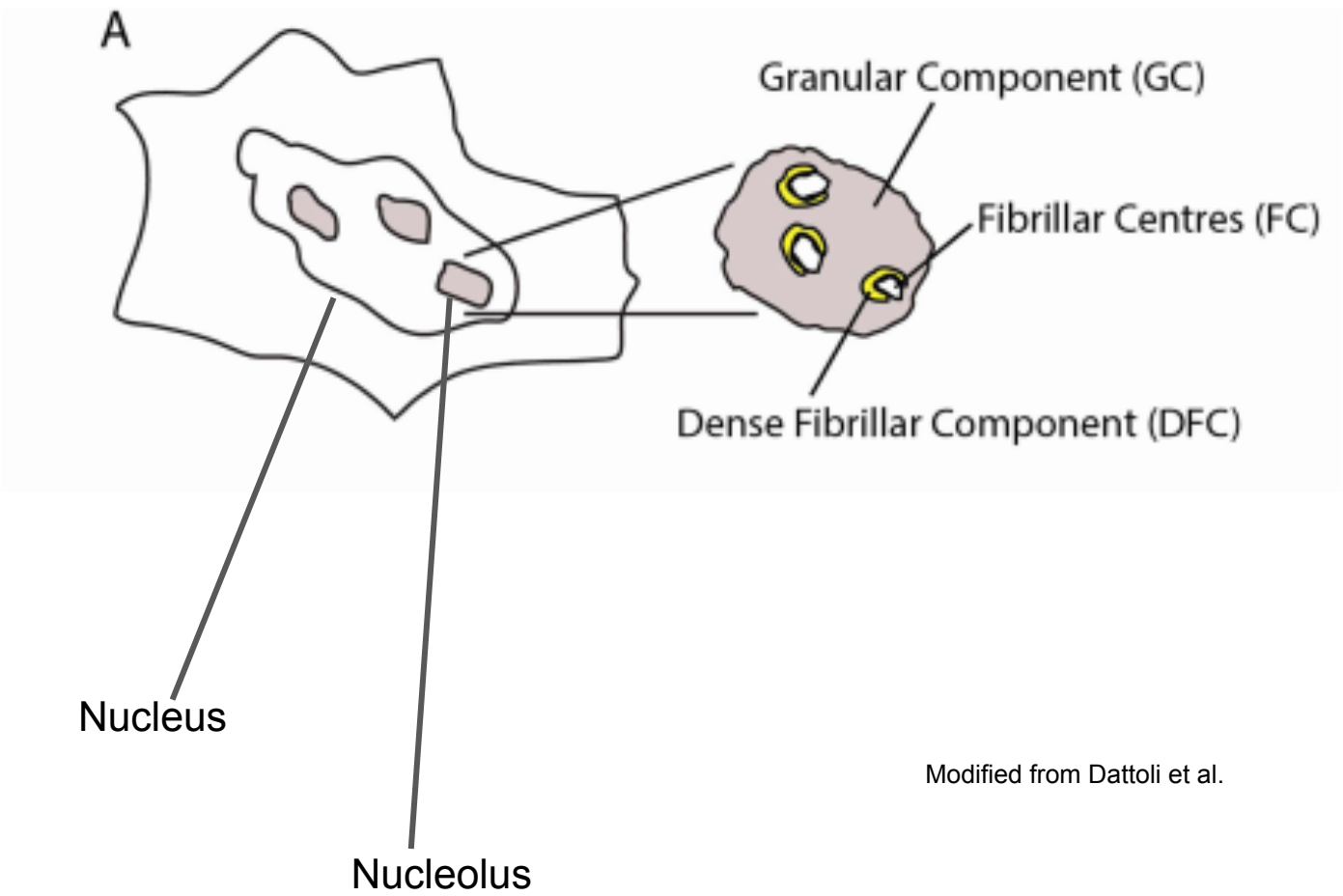
*Nv*SNAILA- and *Nv*SNAILB-ΔZnf

Fig 4A-G: How did they standardize fluorescence levels?



Images used to show normalized fluorescence intensity observed with the 3 different fluorophores

Fig 5A: Where specifically within the nucleolus does *Nv*SNAILA localize?



Modified from Dattoli et al.

Fig 5B-K: Where does *NvSNAILA* localize specifically within the nucleolus?

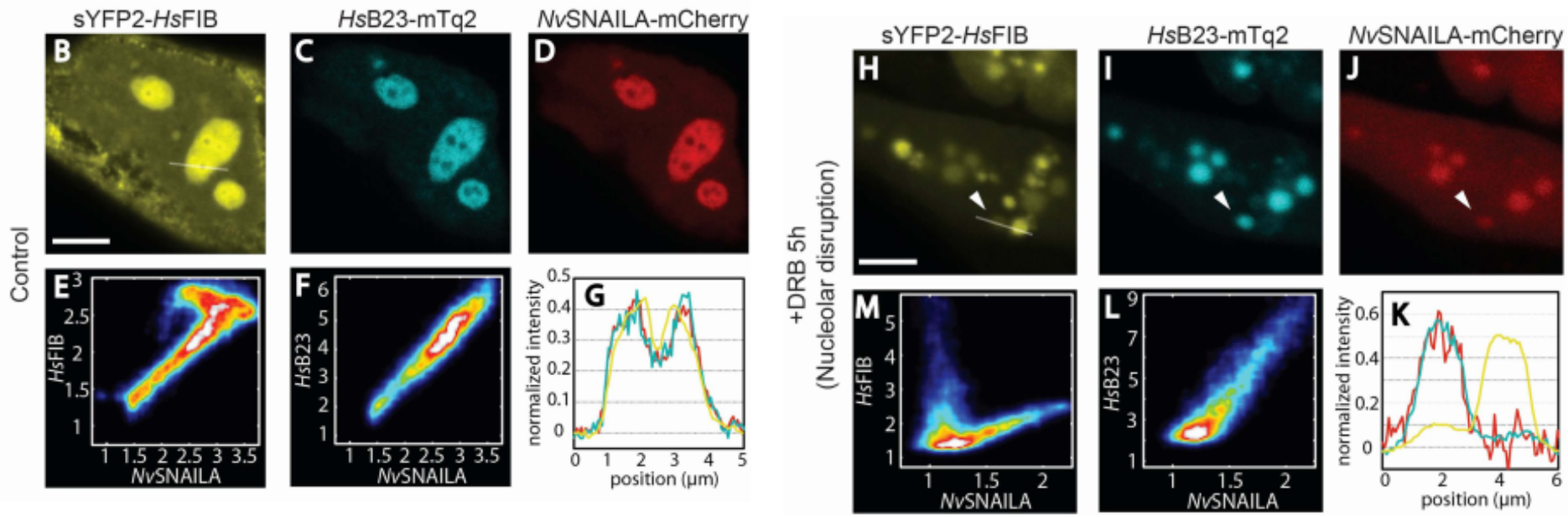


Fig 6A-D: Where does *Nv*SNAILA and B localize in *Nematostella vectenis* embryos?

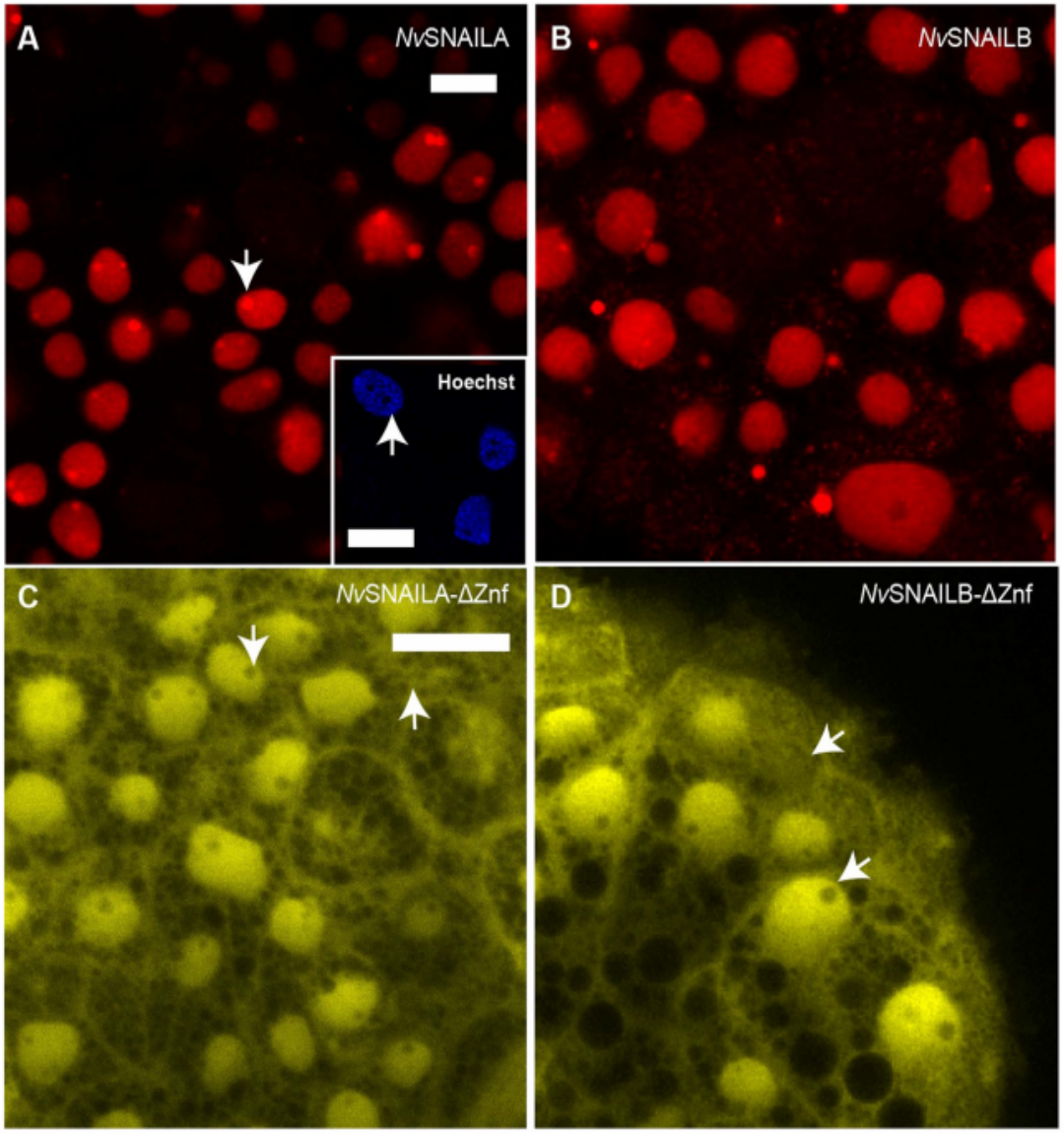
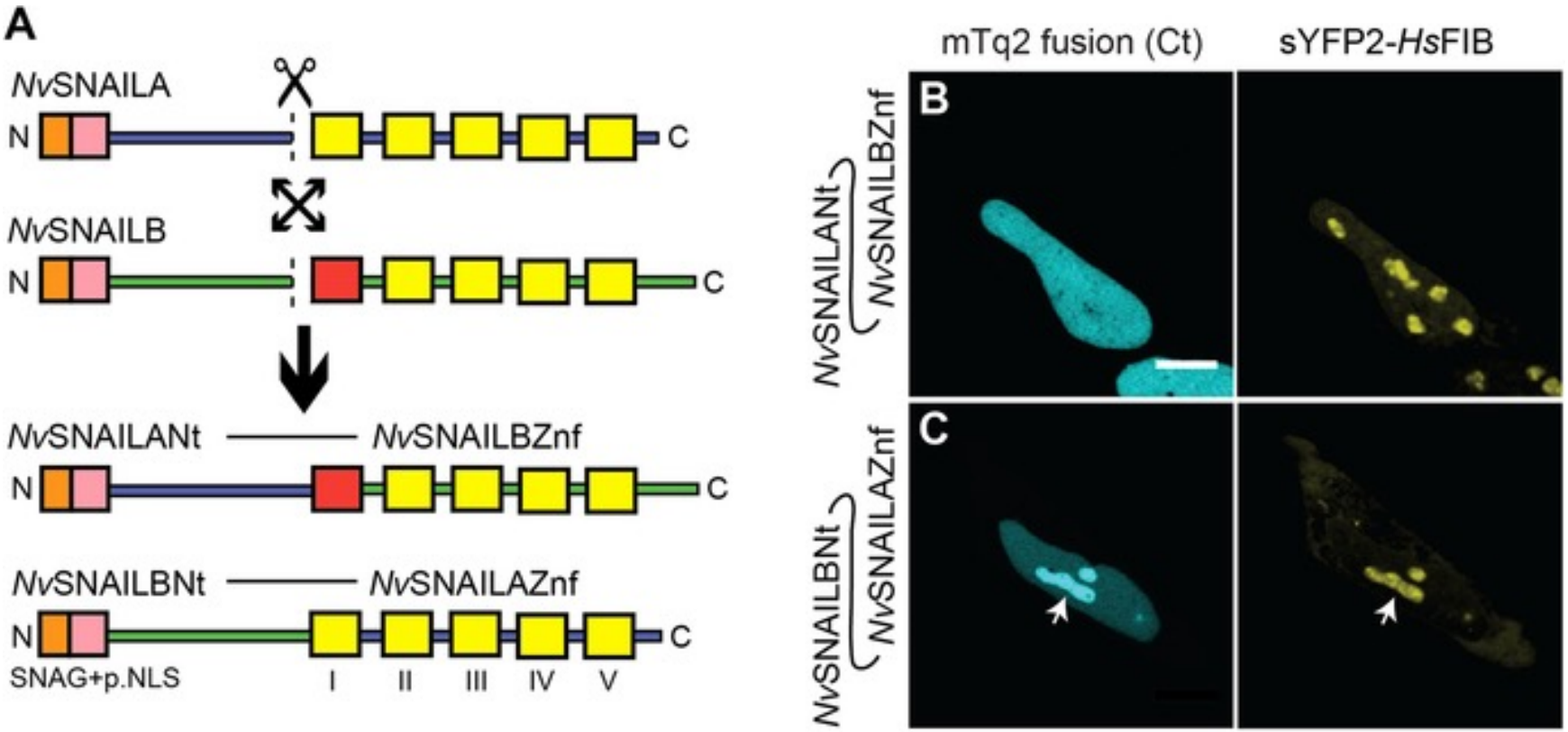
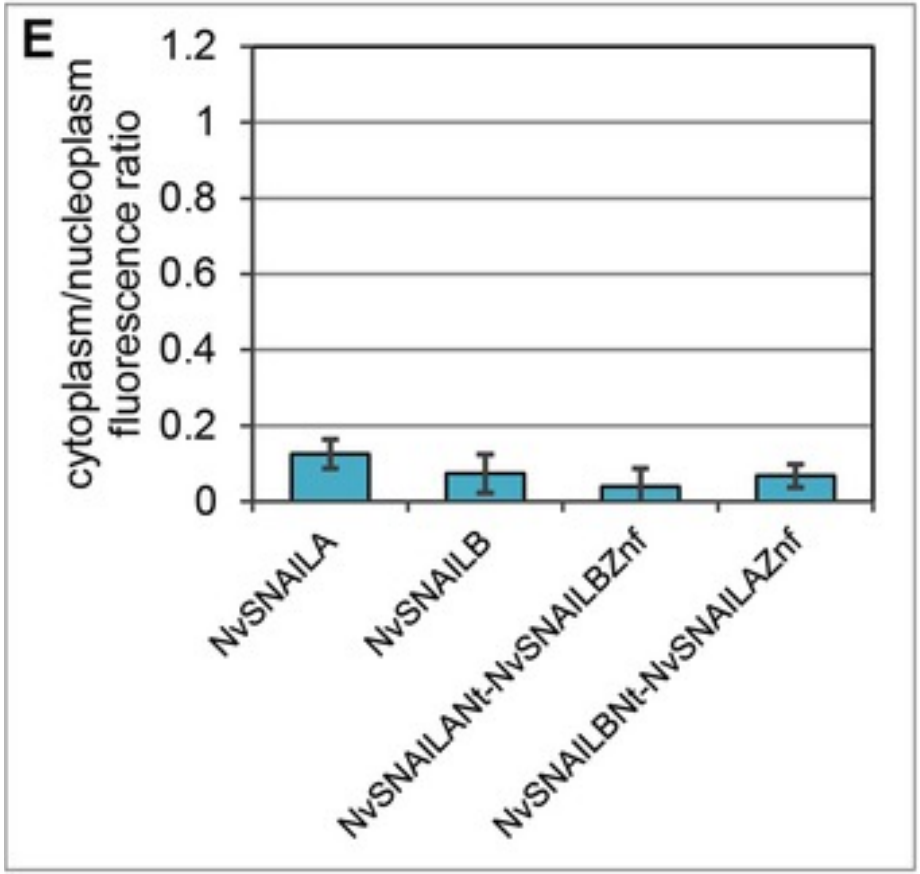


Fig 7A-C:How were they sure that *Nv*SNAILA's localization capacity was due to the Zinc-finger domains?



nuclear
nucleolar

Fig 7D-E: How were they sure that *Nv*SNAILA's localization capacity was due to the Zinc-finger domains?



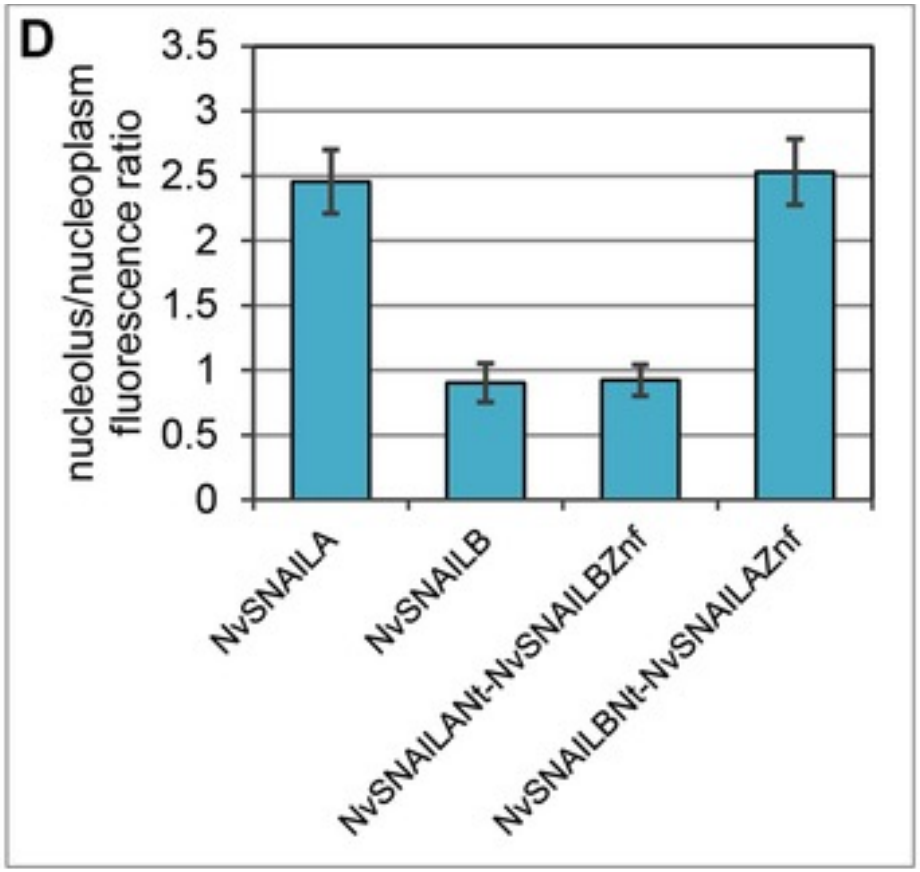
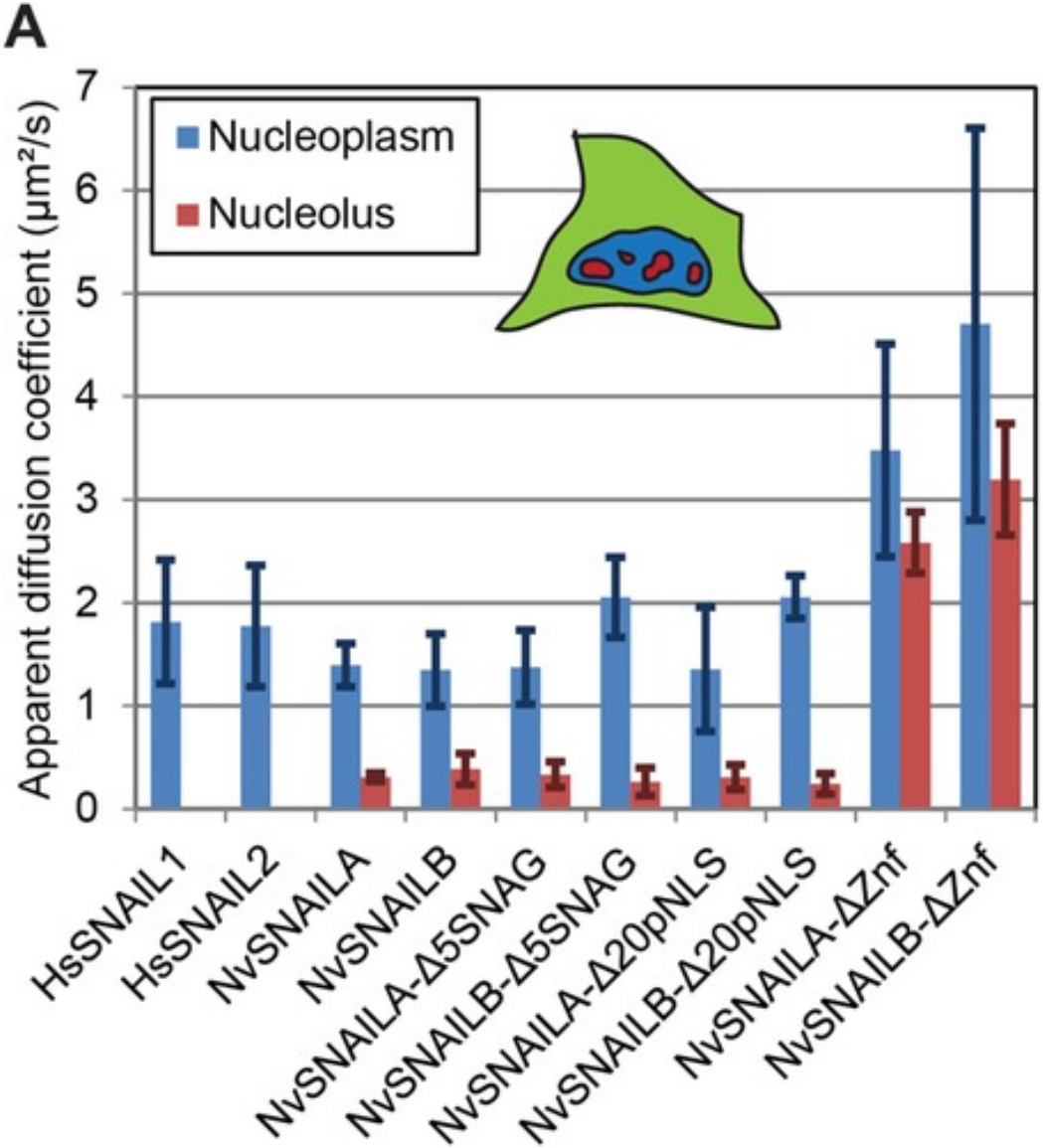
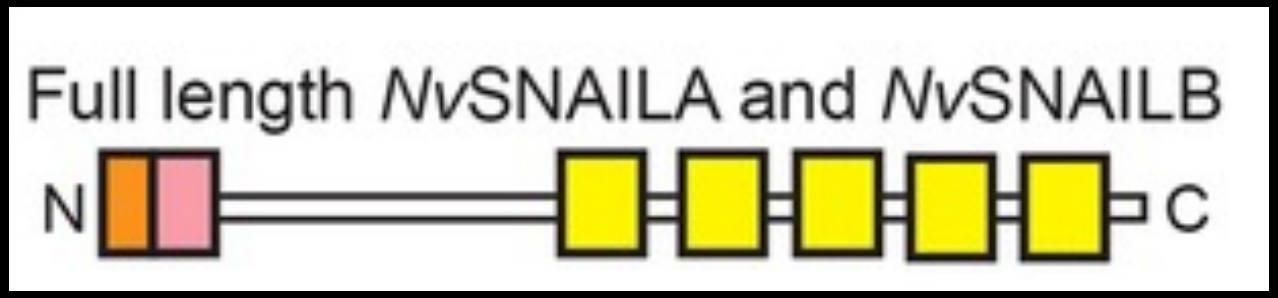


Fig 8A-B: How do the SNAIL protein domains confer function?



In summary

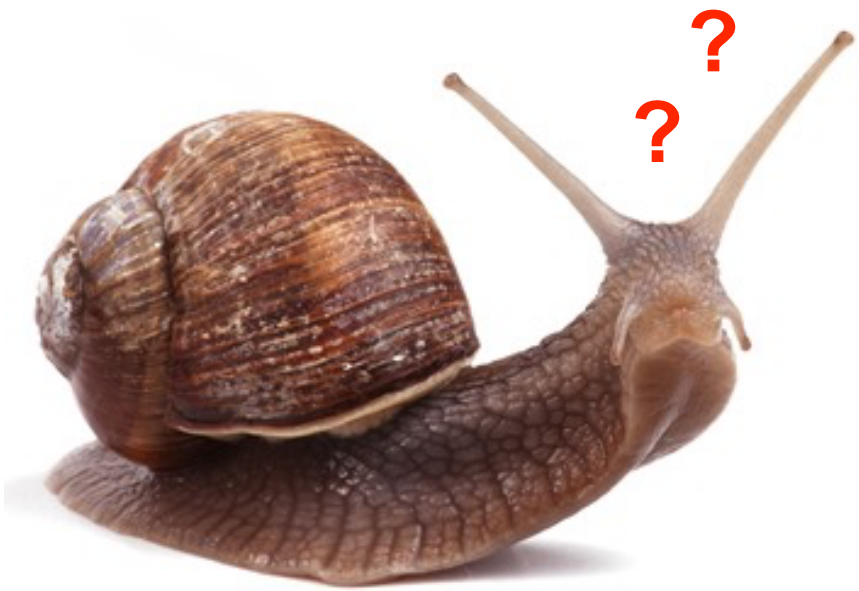


Protein domains are conserved structural and functional regions that can be analyzed with programs such as Pfam and SMART.

Zn Finger domain loss may have happened multiple times throughout evolutionary history.

The Zinc Finger domains play a role in nuclear localization, while putative NLS does not.

The first Zinc Finger Domain of *Nv*SNAILA confers strong nucleolar localization capability.



Further Questions?