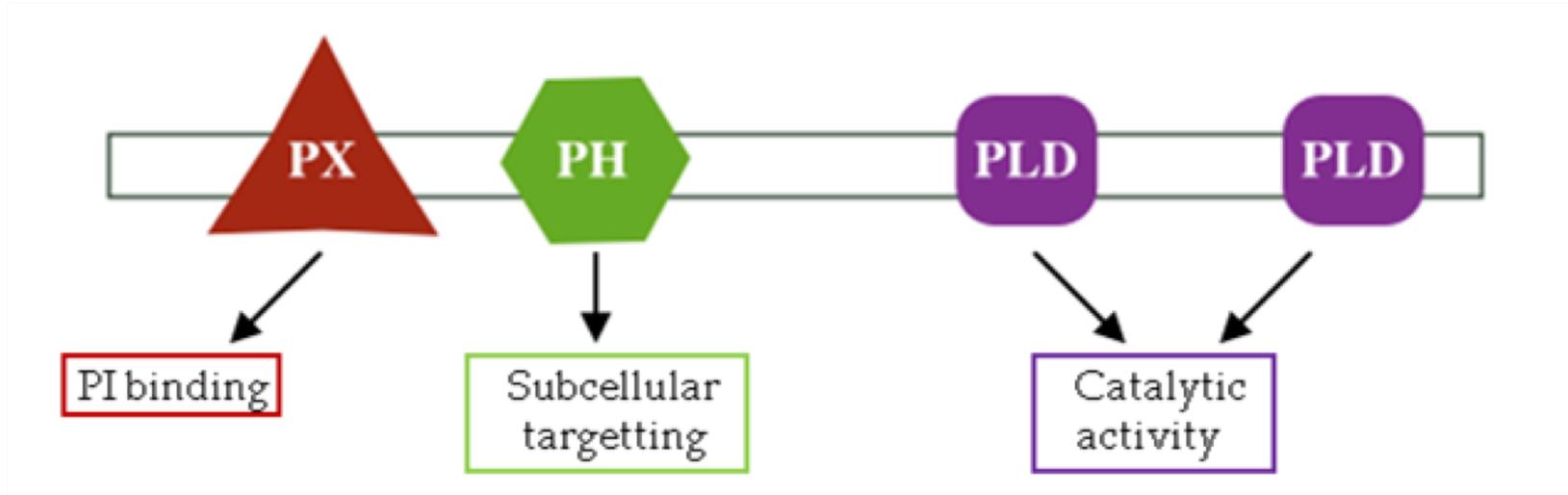


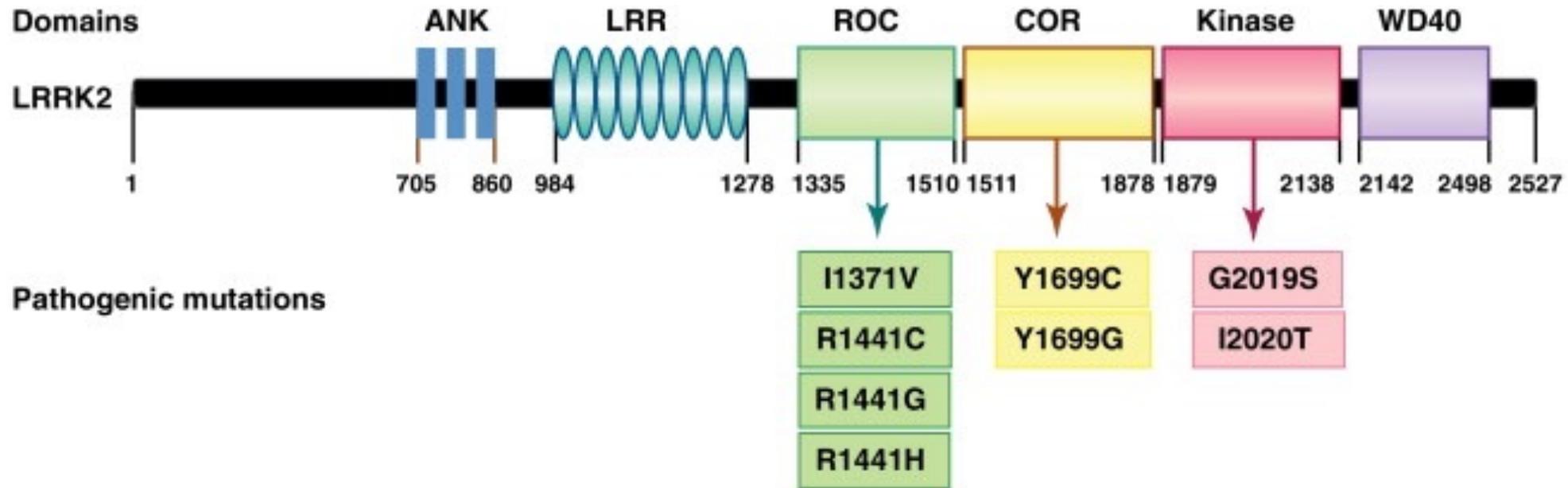
Protein domain analysis



Jared Akers
Zyann Furseth

Domains, alignments, and analysis (Jared)
Retrotransposons (Jared)
Paper review (Zyann)
Applications and future directions (Jared)

What is a protein domain?

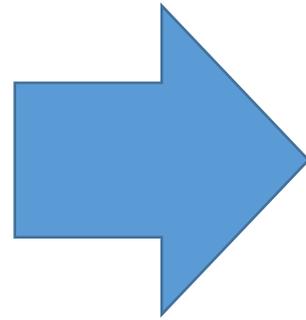


TRENDS in Pharmacological Sciences

An independently folded polypeptide unit, which may or may not have a characterized biological function, and is capable of existing outside of the protein.

Why is protein domain analysis useful?

RTDCYGNVNRI DTTG
ASCKTAKPEGLSYCG
VSASKKI AERDLQAM
DRYKTI I KKVGEKLC
VEPAVI AGI I SRESH
AGKVLKNGWDRGNG
FGLMQVDKRS HKPQG
TWNGEVHI TQGTTL
INFIKTI QKKFPSWT
KDQQLKGGI SAYNAG
AGNVRSYARM DIGTT
HDDYANDVVARAQYY
KQHGY

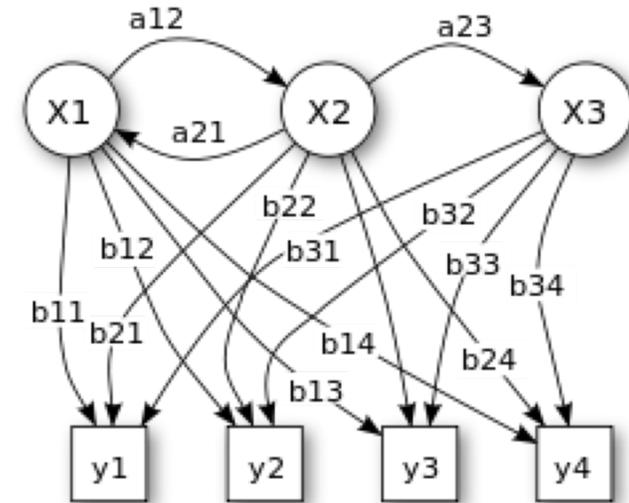


To learn a protein domain from a sequence.

How do you determine protein domains?

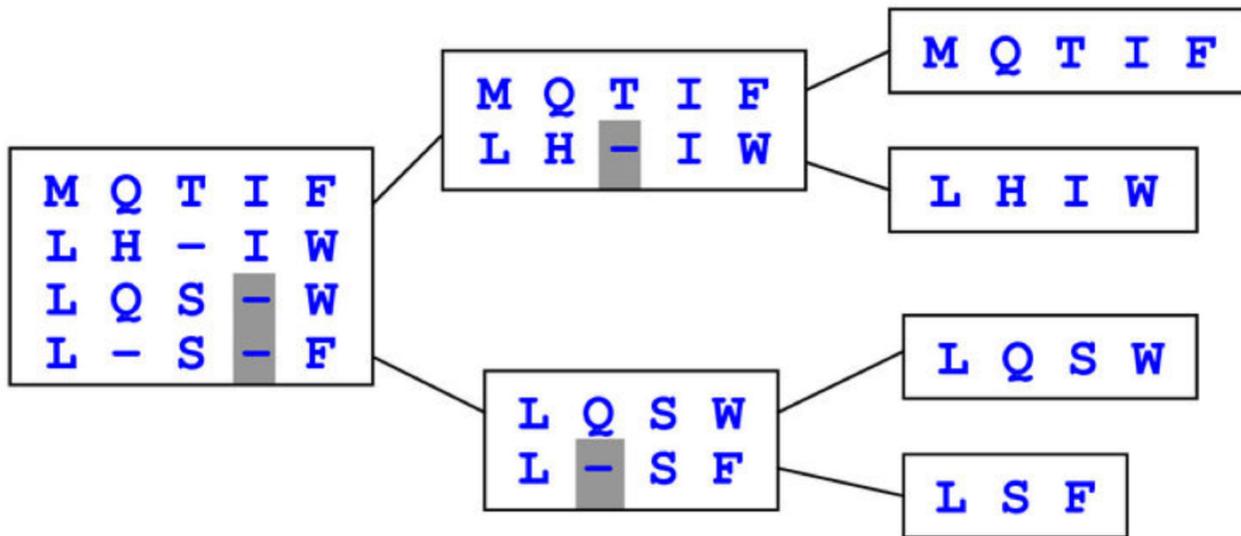
```
RLA0_METVA  --MIDAKSEHKIAPWKIEEVNALKELLKSANVIALIDMMEVPAVQLQEIIRDK
RLA0_METJA  ---METKVKAHVAPWKIEEVKTLKGLIKSKPVVAIVDMMDVPAPQLQEIIRDK
RLA0_PYRAB  -----MAHVAEWKKKEVEELANLIKSYPVIALVDVSSMPAYPLSQMRRL
RLA0_PYRHO  -----MAHVAEWKKKEVEELAKLIKSYPVIALVDVSSMPAYPLSQMRRL
RLA0_PYRFU  -----MAHVAEWKKKEVEELANLIKSYPVVALVDVSSMPAYPLSQMRRL
RLA0_PYRKO  -----MAHVAEWKKKEVEELANIIKSYPVIALVDVAGVPAYPLSKMRDK
RLA0_HALMA  MSAESERKTETIPEWKQEEVDAIVEMIESYESVGVVNIAGIPSRQLQDMRRD
RLA0_HALVO  MSESEVRQTEVIPQWKREEVDELVDVDFIESYESVGVVGVAGIPSRQLQSMRRE
RLA0_HALSA  MSAAEQRTTEEVPEWKRQEVAEVLDLLETYDSVGVVNVGTGIPSKQLQDMRRG
RLA0_THEAC  -----MKEVSQOKKELVNEITQRIKASRSVAIVDTAGIRTRQIQDIRGK
RLA0_THEVO  -----MRKINPKKKEIVSELAQDITKSKAVAIVDIKGVTRTROMQDIRAK
RLA0_PICTO  -----MTEPAQWKIDFVKNLENEINSRKVAAIVSIKGLRNNEFQKIRNS
```

Multiple sequence alignment



Hidden Markov models

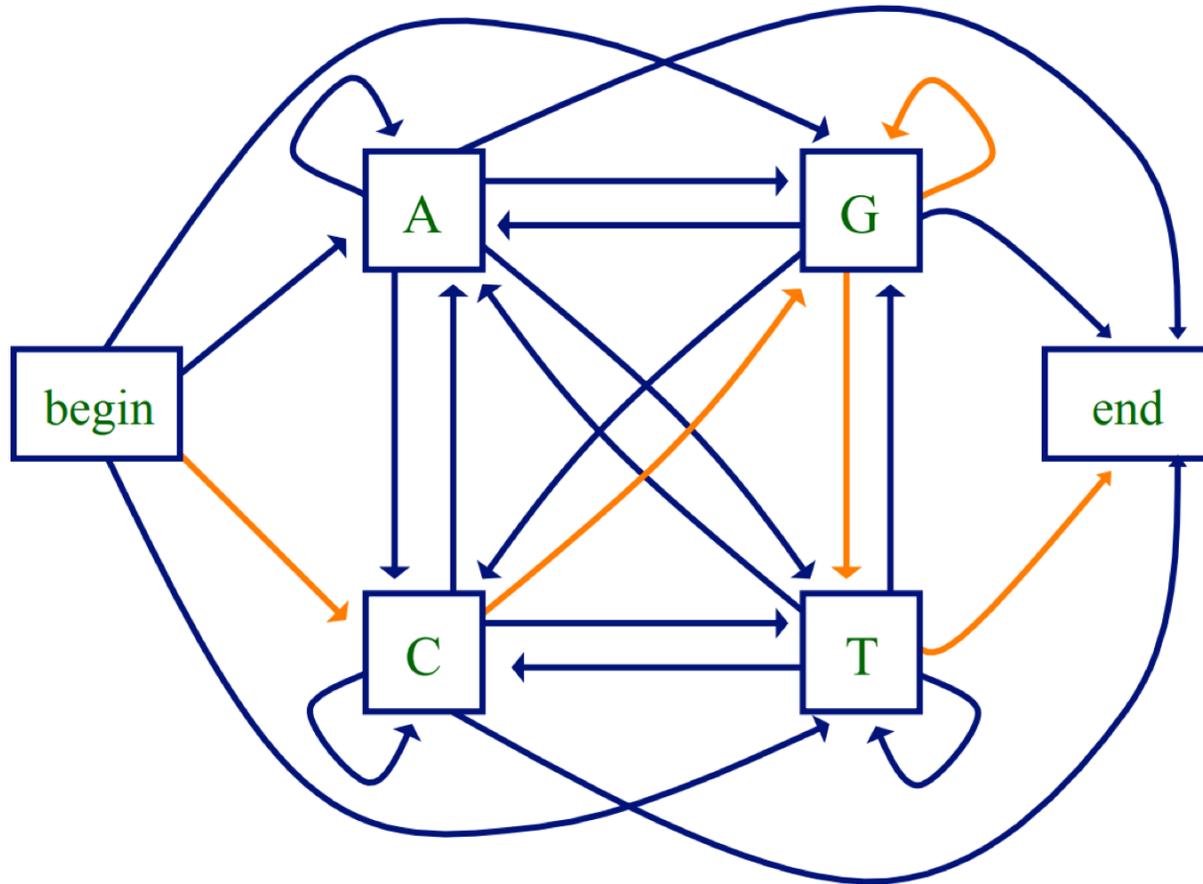
How does MUSCLE work?



MUSCLE is the multiple sequence alignment algorithm used in our paper.

**Progressive alignments
Iterative**

What is a Markov Model?

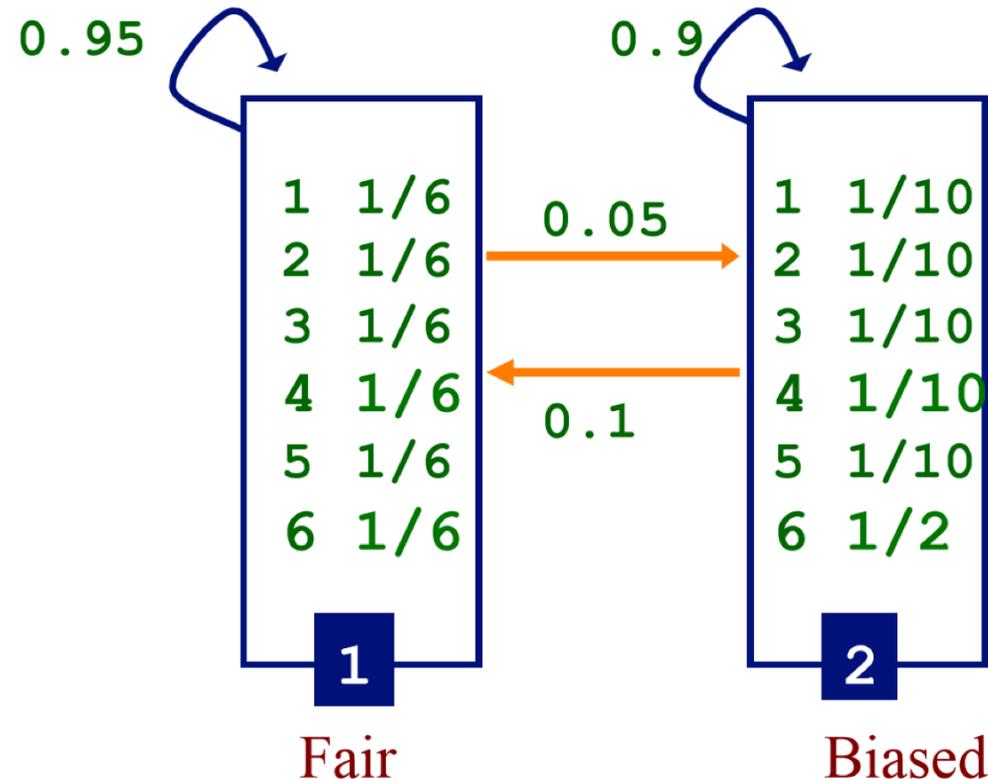


Modeling of dependencies in sequential data.

What is the probability of sequence (CGGT)?

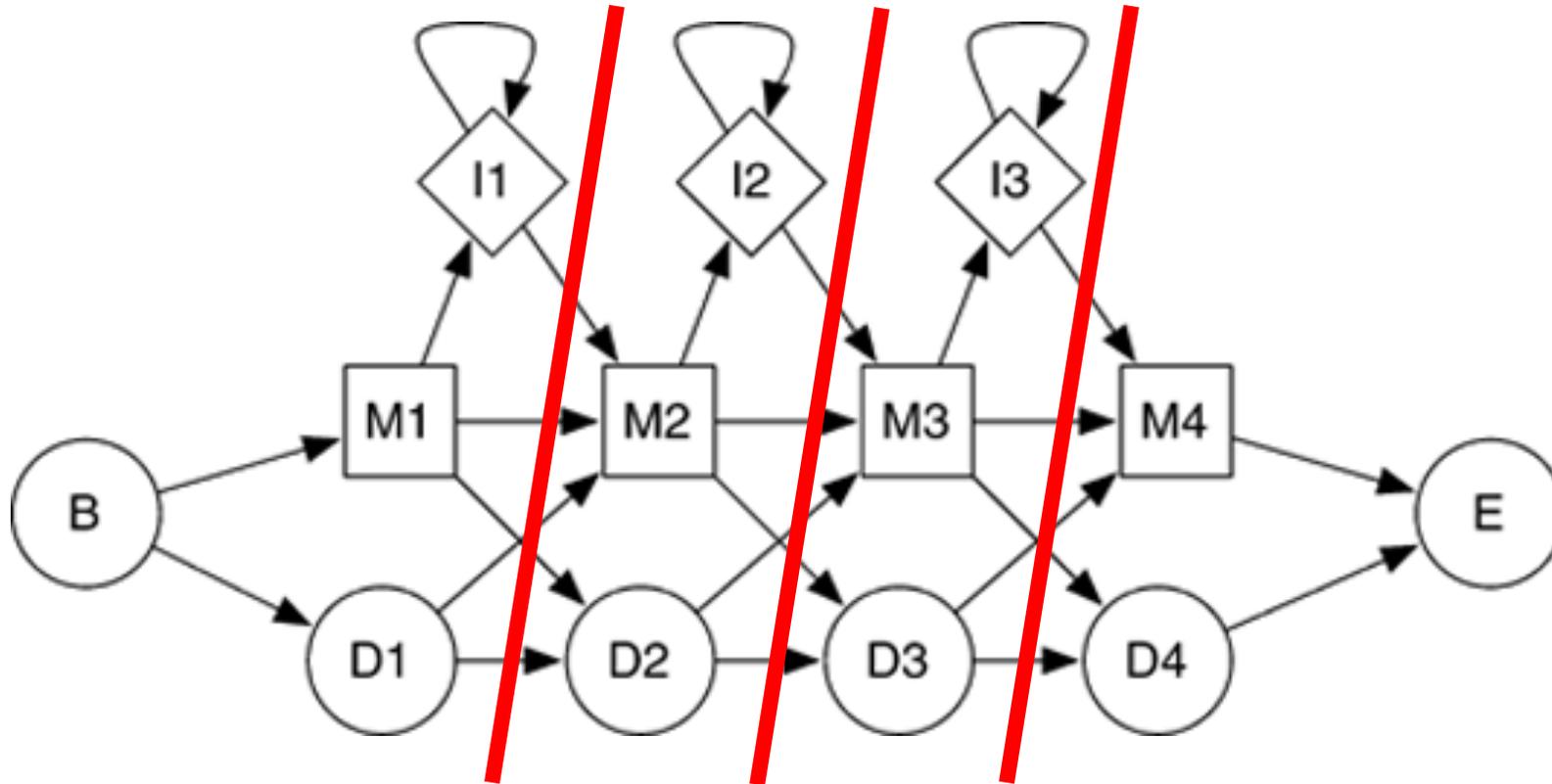
What is a hidden Markov model (HMM)?

“HMM for an Occasionally Dishonest Casino”



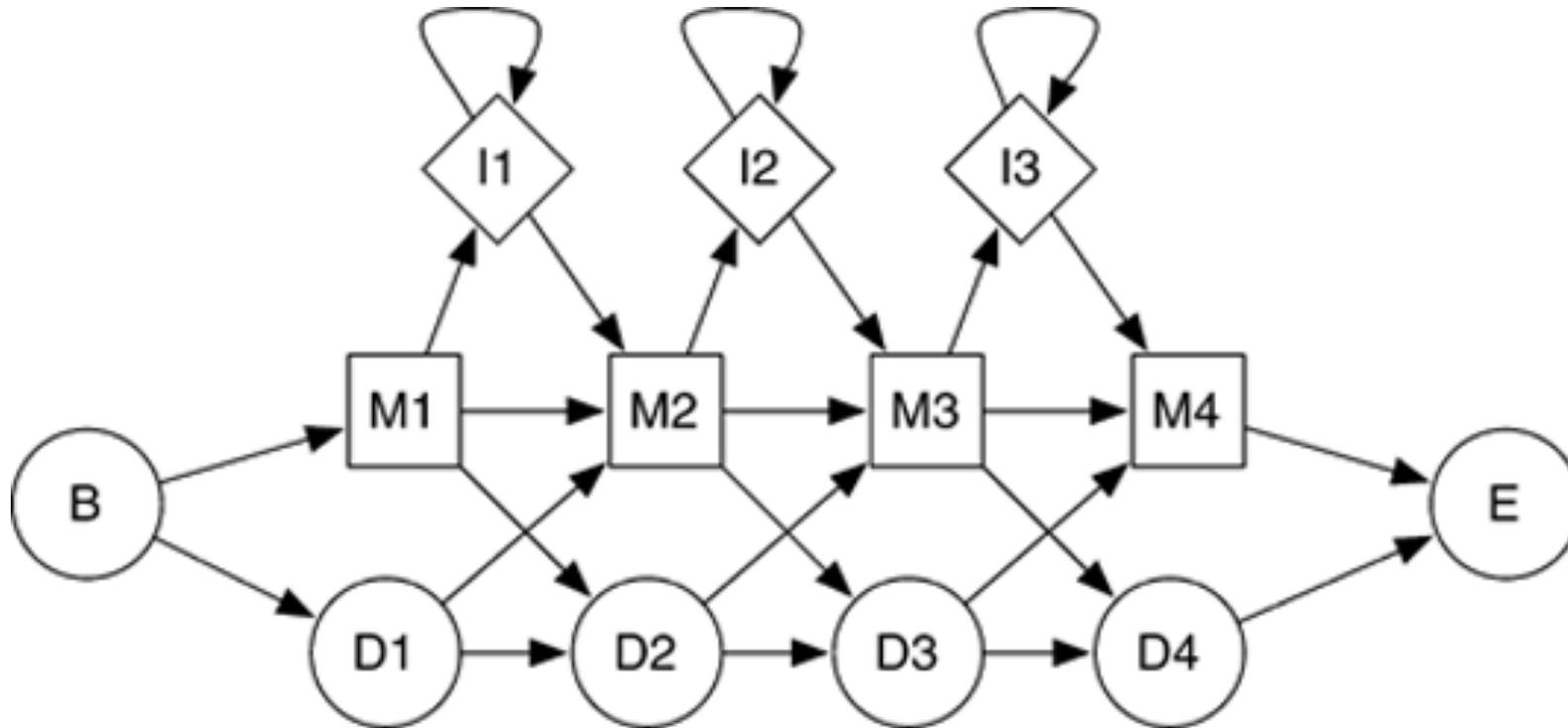
Hidden Markov models that also have a “black box” component, i.e. the current state is unknown.

What is I_1 , M_1 , and D_1 ?



Although all sequences represent the same domain, the sequences can be different.

What is a profile HMM?



Idea: Create a single profile based off a MSA.

What websites can be used to determine protein domains?

Pfam

EMBL-EBI HOME | SEARCH | BROWSE | FTP | HELP | ABOUT **Pfam**
keyword search

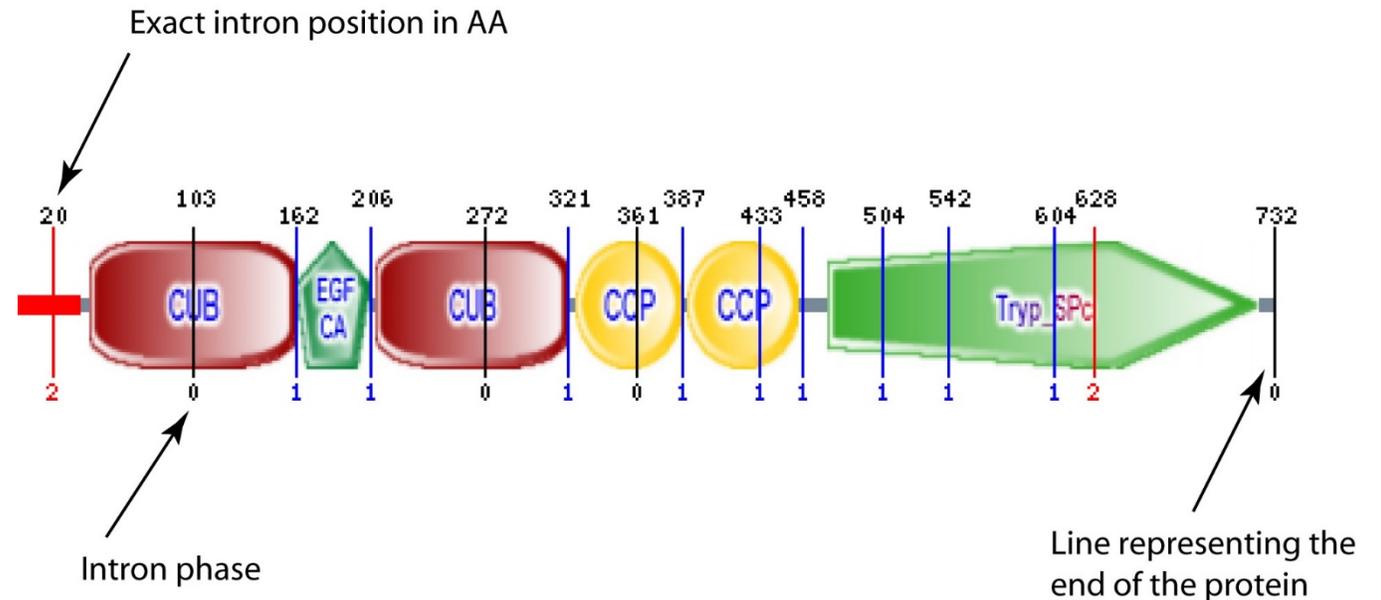
Sequence search results
[Show](#) the detailed description of this results page.
 We found **6** Pfam-A matches to your search sequence (all significant)

[Show](#) the search options and sequence that you submitted.
[Return](#) to the search form to look for Pfam domains on a new sequence.

Significant Pfam-A Matches
[Show](#) or [hide](#) all alignments.

Family	Description	Entry type	Clan	Envelope		Alignment		HMM		HMM length	Bit score	E-value	Predicted active sites	Show/hide alignment
				Start	End	Start	End	From	To					
PI-PLC-X	Phosphatidylinositol-specific phospholip ...	Family	CL0384	322	465	322	465	1	145	145	218.4	2.4e-65	n/a	Show
SH2	SH2 domain	Domain	CL0541	550	639	550	639	1	77	77	84.7	3.2e-24	n/a	Show
SH2	SH2 domain	Domain	CL0541	668	741	668	741	1	77	77	70.8	6.6e-20	n/a	Show
SH3_1	SH3 domain	Domain	CL0010	797	843	797	843	1	48	48	54.0	8e-15	n/a	Show
PI-PLC-Y	Phosphatidylinositol-specific phospholip ...	Family	CL0384	953	1069	953	1069	1	114	114	131.5	1.6e-38	n/a	Show
C2	C2 domain	Domain	CL0154	1088	1193	1090	1187	3	95	103	53.3	2.4e-14	n/a	Show

Comments or questions on the site? Send a mail to pfam-help@ebi.ac.uk.
 European Molecular Biology Laboratory



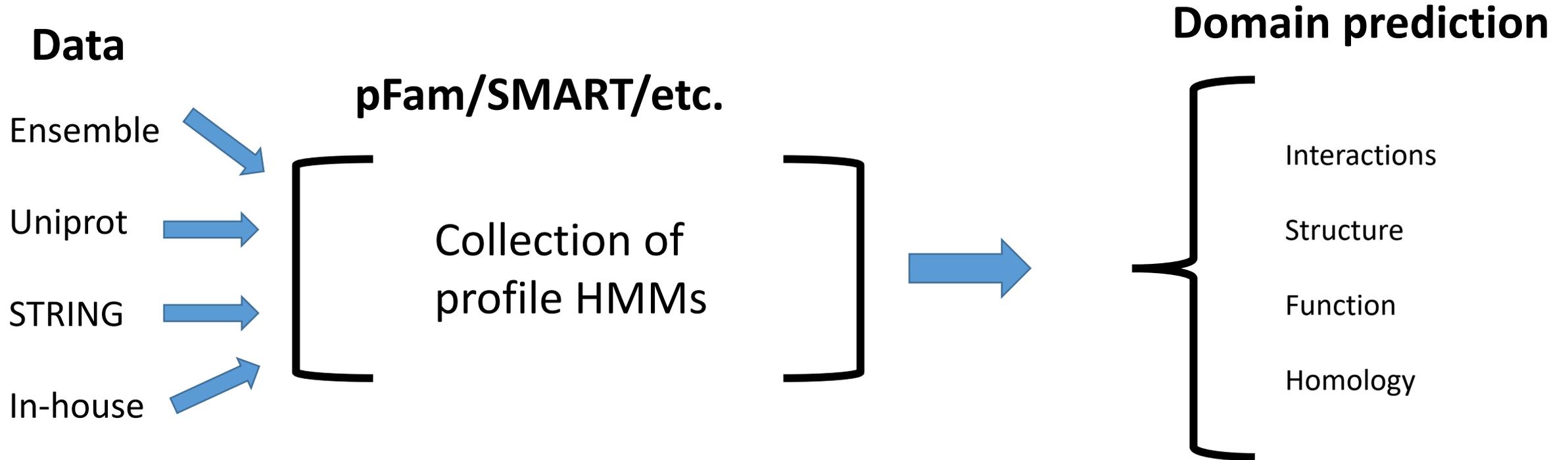
Pfam

Advantages	Disadvantages
<p data-bbox="512 329 1225 444">Identifies domains with wide spectrum of cellular functions</p> <p data-bbox="448 518 1284 572">Can browse clans for homologues</p>	<p data-bbox="1352 394 2211 508">Less sensitive: more false positives and negatives</p>

SMART

Advantages	Disadvantages
<p data-bbox="430 993 1299 1048">More accurate domain identification</p> <p data-bbox="486 1119 1243 1173">Domains extensively annotated</p> <p data-bbox="504 1245 1225 1299">200+ million protein domains</p>	<p data-bbox="1340 1059 2224 1230">Less comprehensive: Mostly identifies domains in signaling, extracellular, and chromatin proteins</p>

How do domain databases work?



Biological background knowledge

Transposable elements

Membrane vesicle trafficking

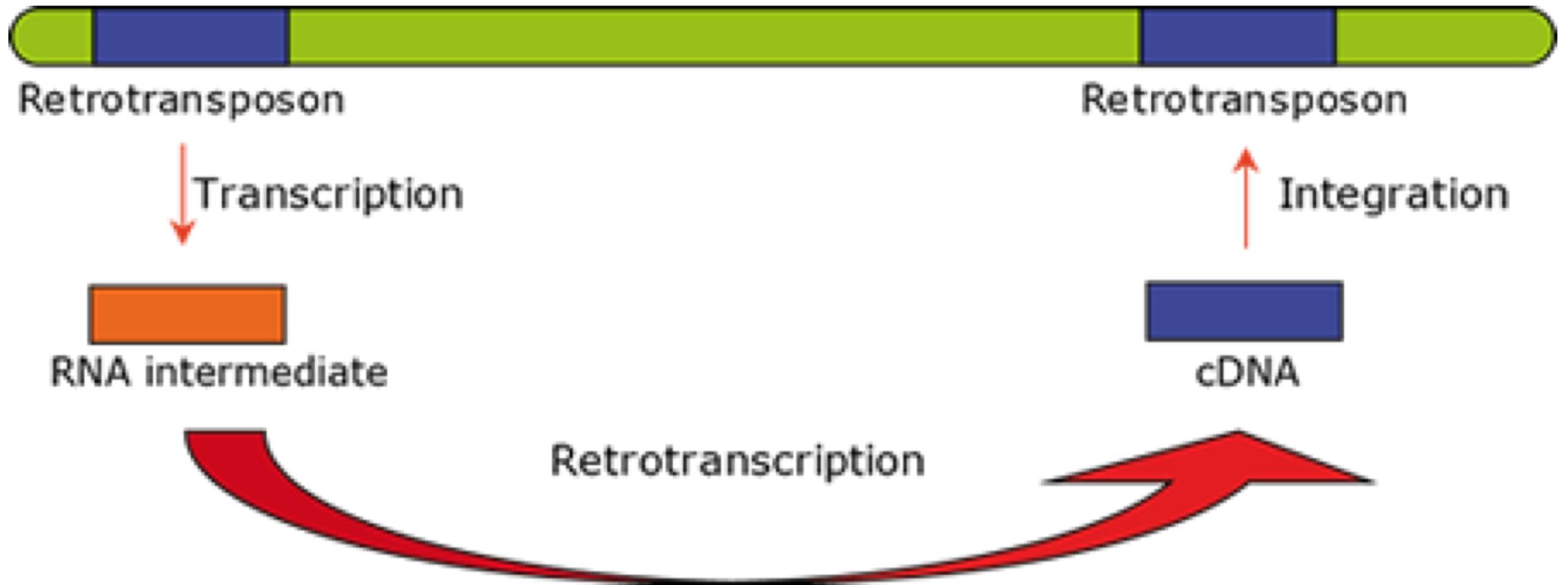


What is a transposon?

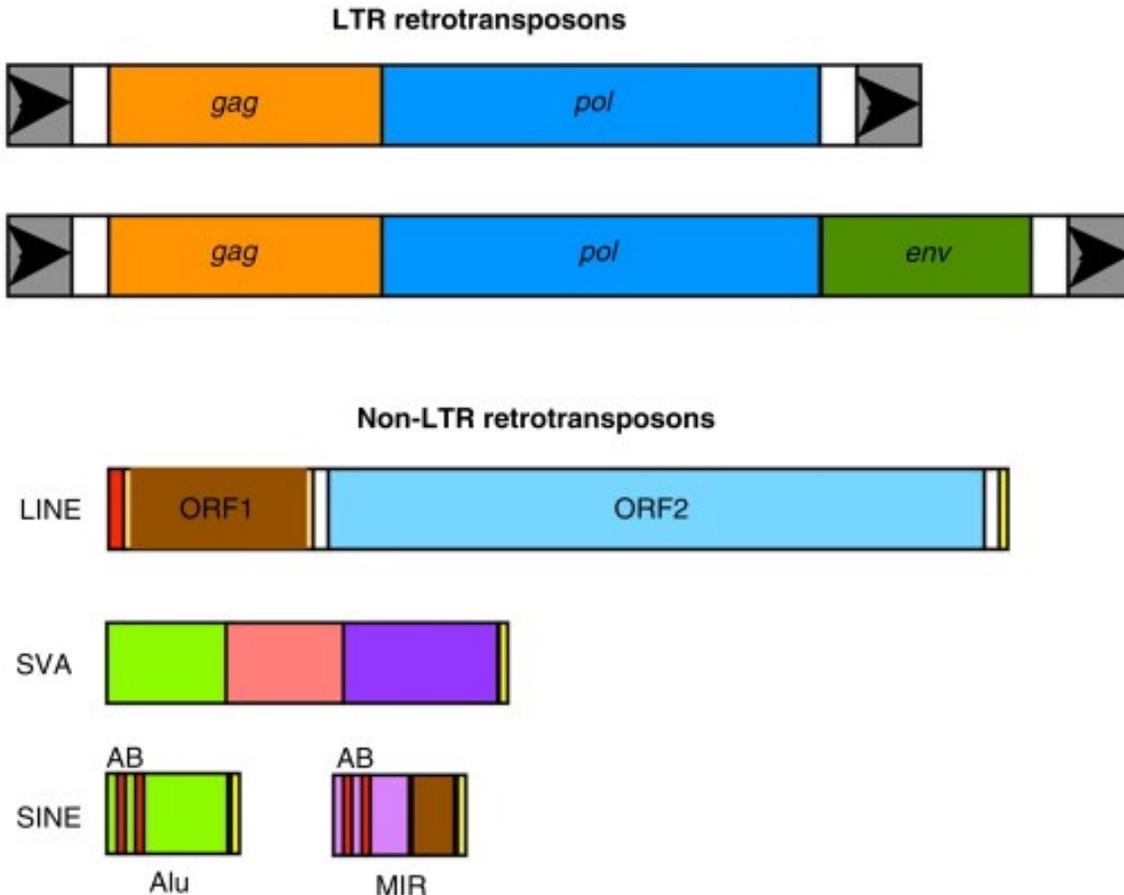


Genetic elements capable of changing position inside of the genome. Transposons have been referred to as “jumping genes”.

What is a retrotransposon?



How are retrotransposons structured?



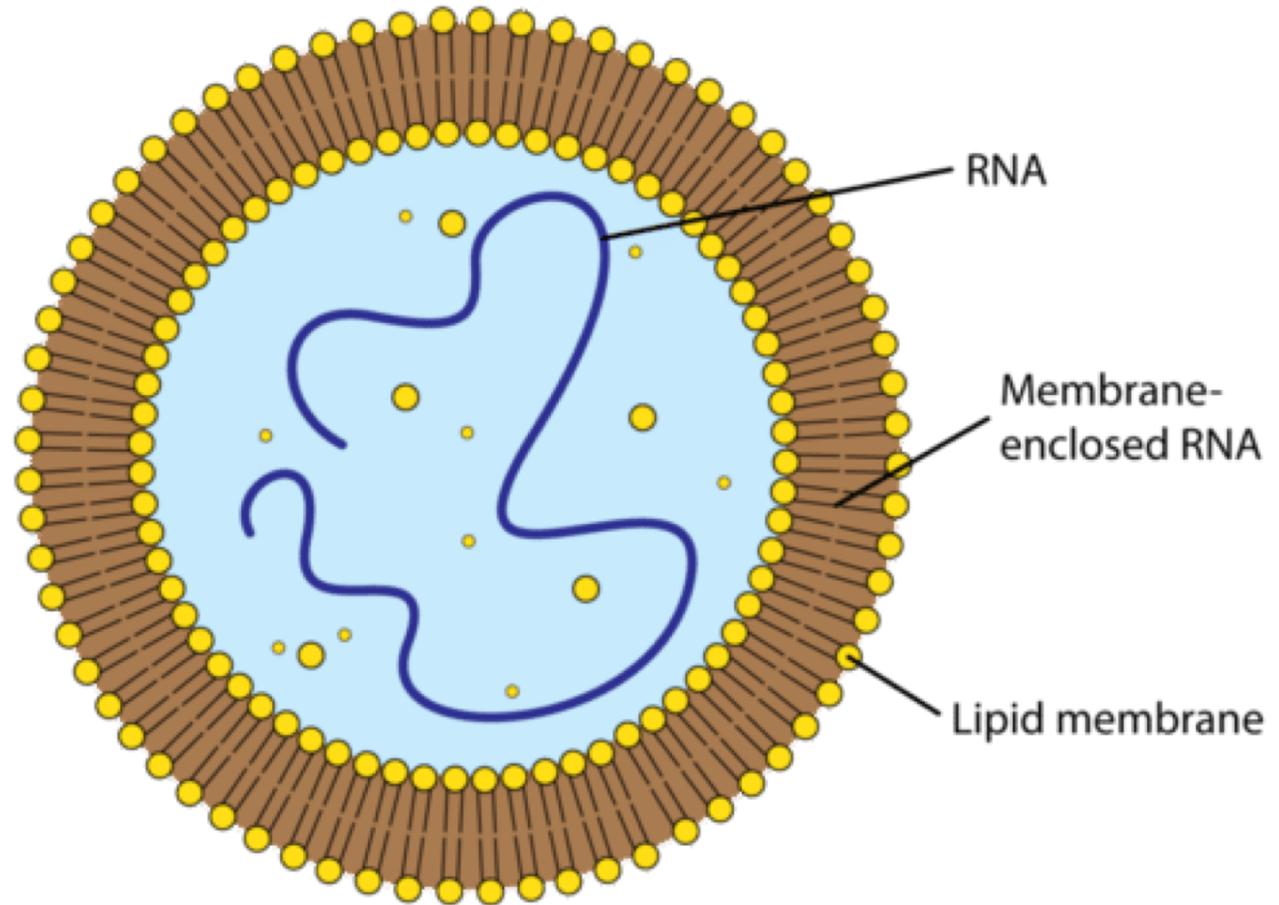
Gag : group antigens : RNA genome binding proteins, nucleoprotein core particle

Pol : reverse transcriptase

Env : envelope protein

What is our current perception of extracellular vesicles (EVs)?

Phospholipid bilayer particles which exit the cell carrying proteins and nucleic acids.



The background features several glowing, semi-transparent spheres of varying sizes. Inside these spheres, there are intricate, multi-colored (green, blue, purple, yellow) particle-like structures that resemble molecular models or complex crystalline lattices. The overall aesthetic is futuristic and scientific, with a dark, gradient background transitioning from purple to green.

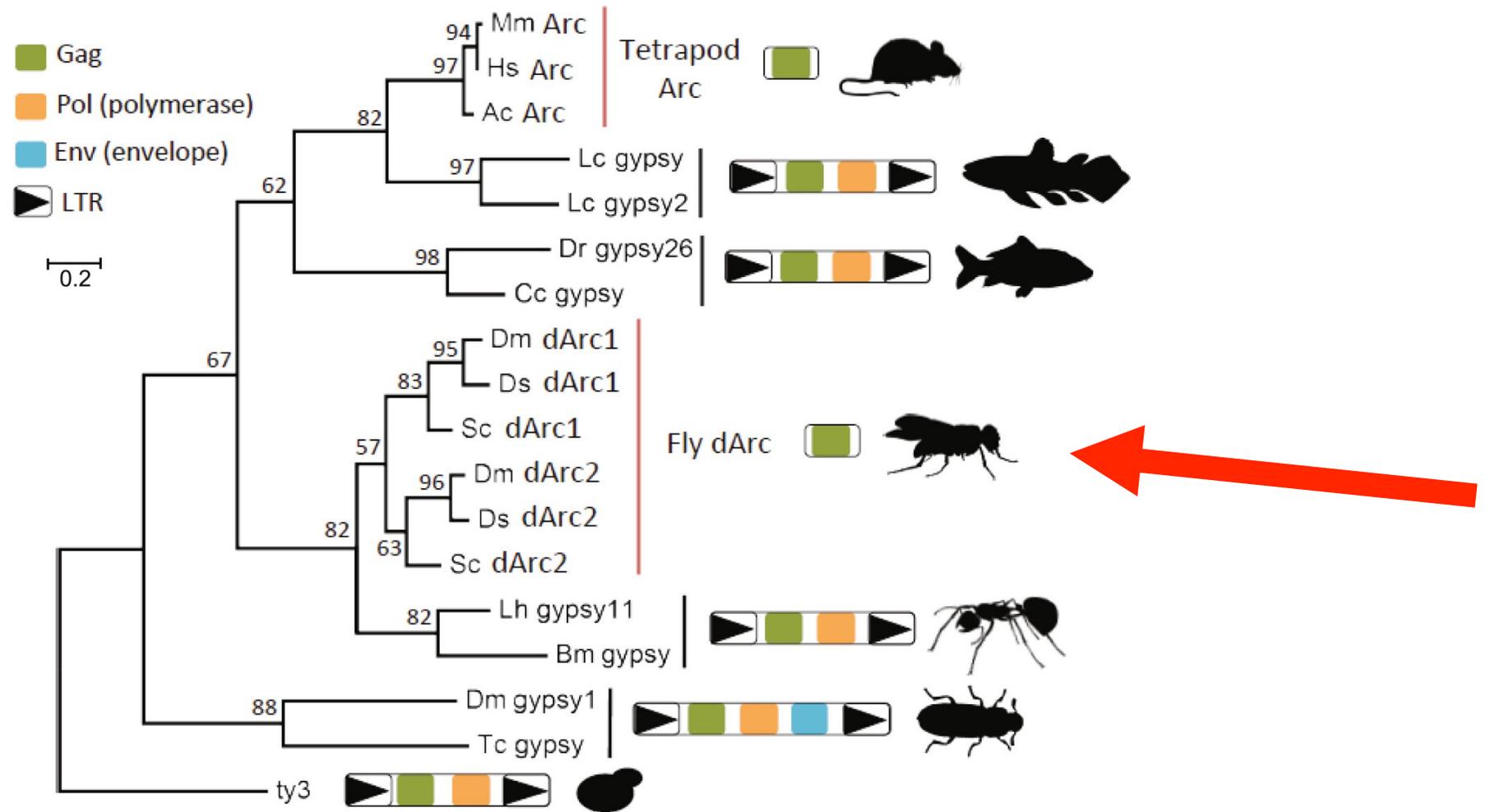
The Neuronal Gene Arc
Encodes a Repurposed
Retrotransposon Gag Protein
that Mediates Intercellular RNA
Transfer

Pastuzyn, et al 2018

Cell, [Volume 172, Issues 1-2](#), p275–288.e18, 11
January 2018

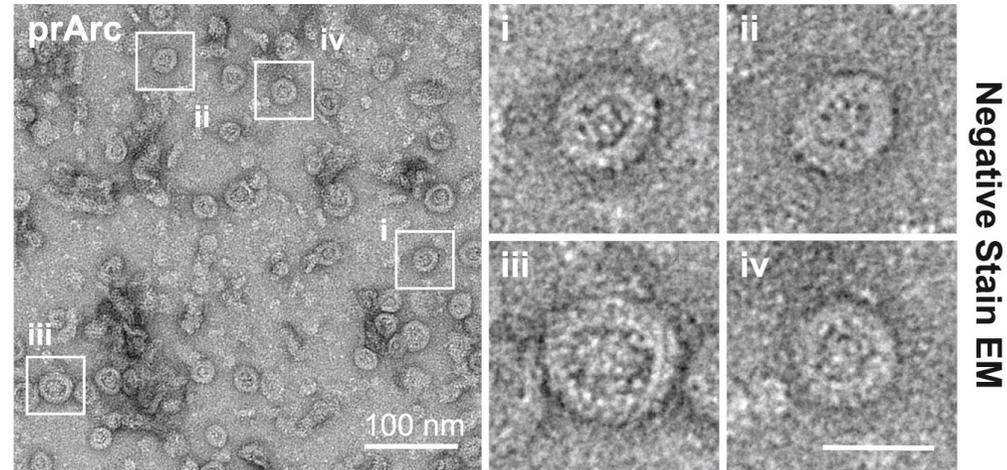
**Zyann Furseth & Jared Akers
Spring 2018**

Fig 1A: What was the maximum likelihood phylogeny of Arc?

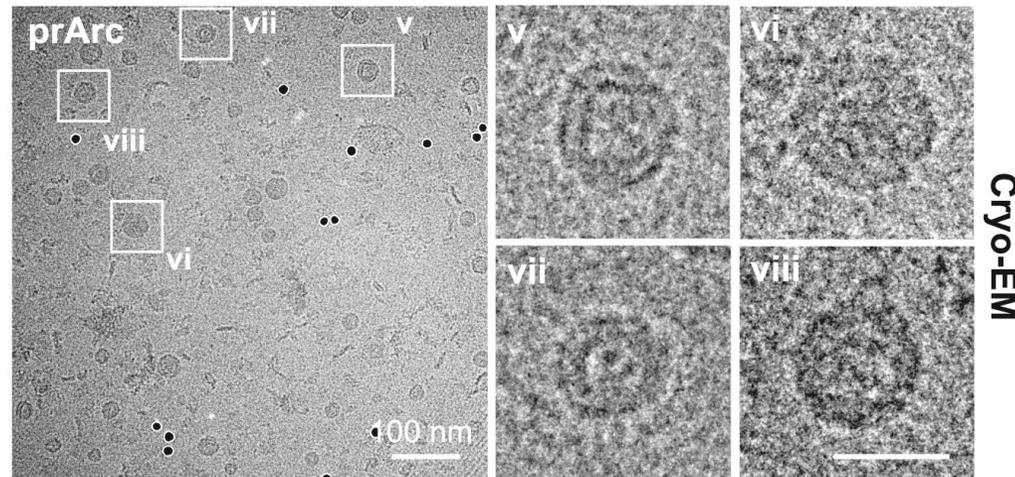


Closely Related To Ty3/gypsy Retrotransposons

Fig 1B: How are Arc capsids structured?

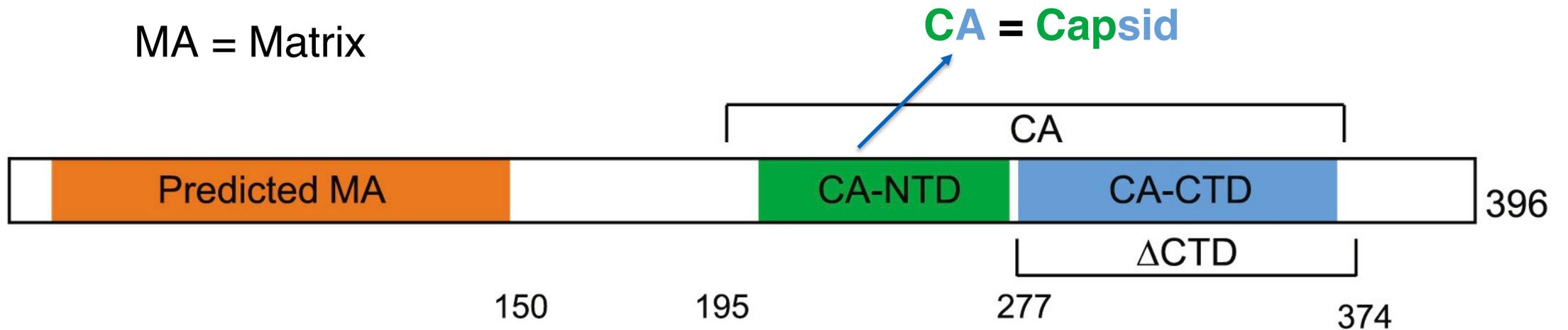


Overexpressed
rat Arc (prArc)



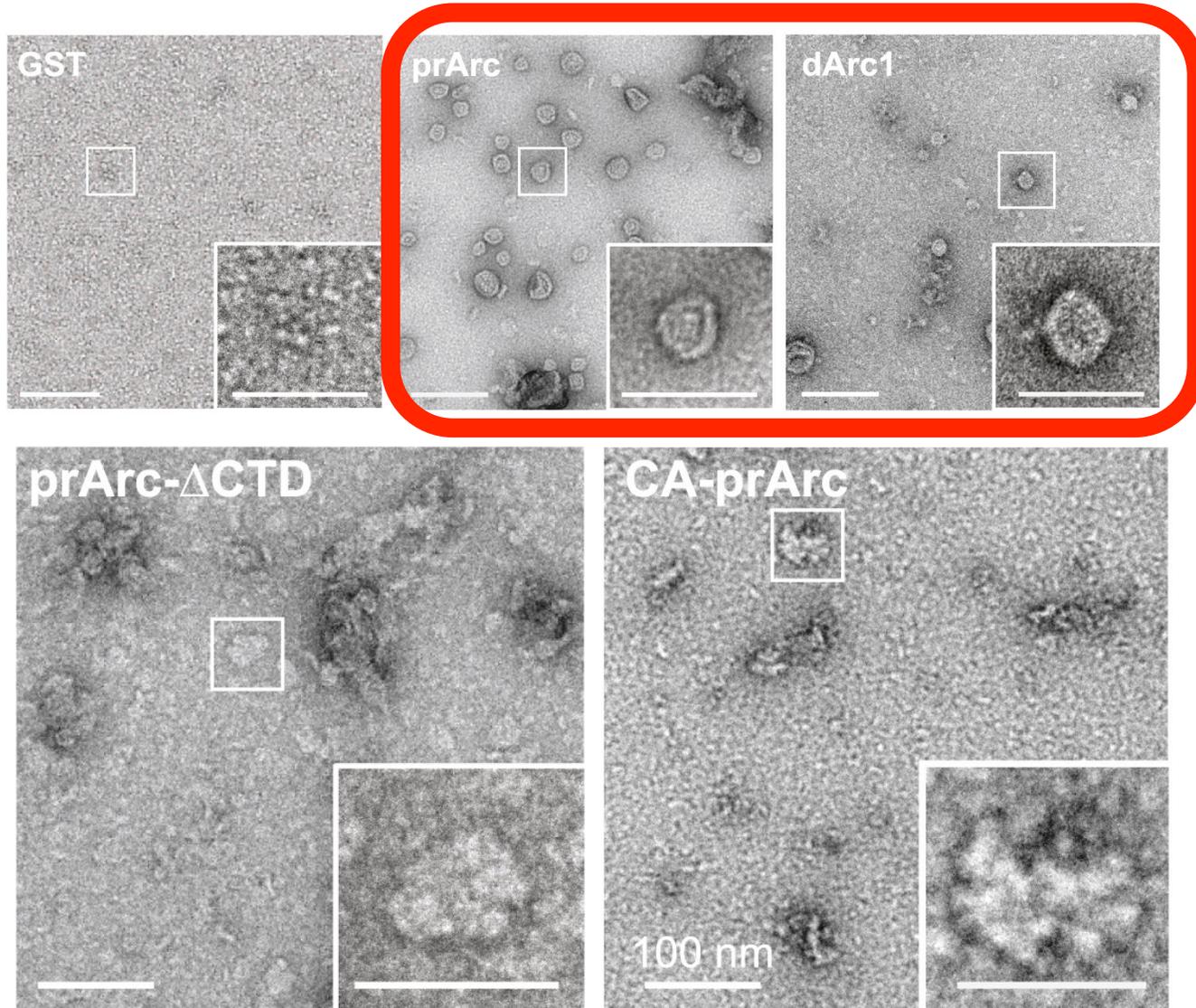
Arc capsids showing the double-layered capsid shell

Fig 1C: What does Arc protein look like?

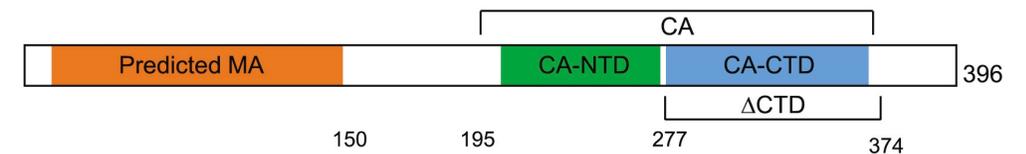


Deletion of the CTD domain

Fig 1C: What does Arc look like via EM?

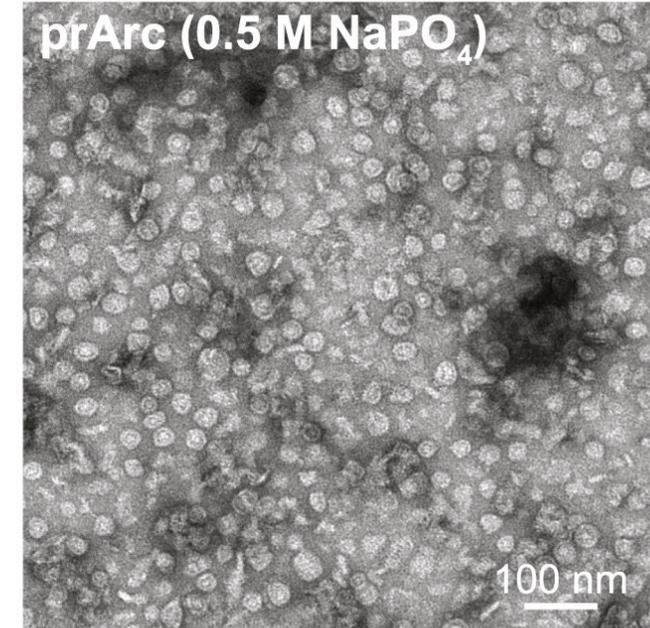
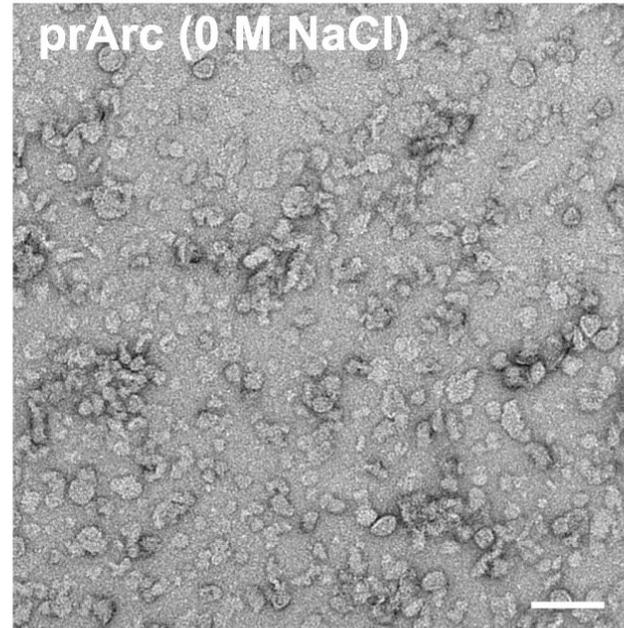
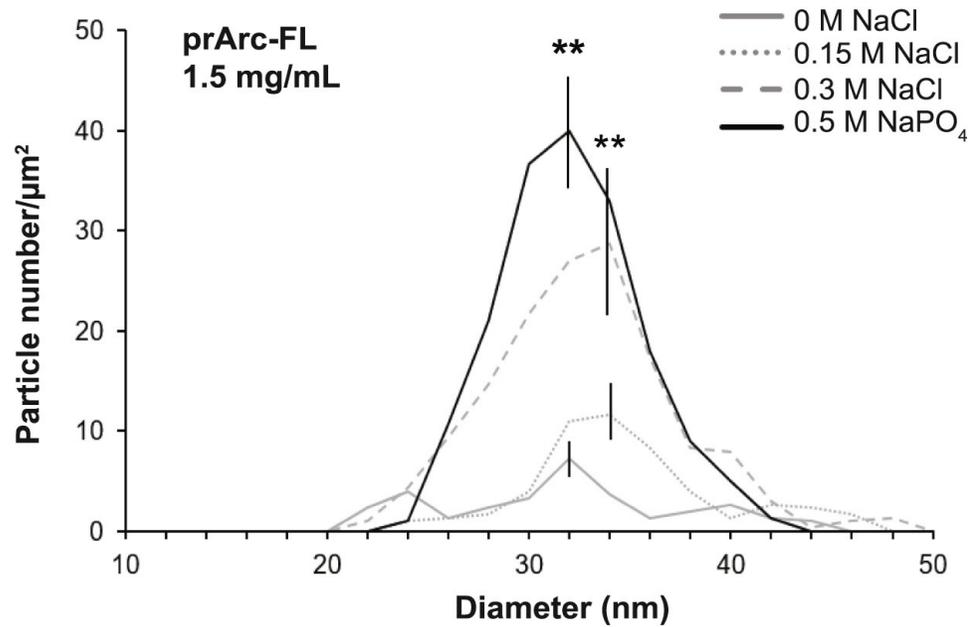


prArc and dArc over expression have clear double shells



Loss of the CA-CTD leads to loss of the capsid structure

Fig 1D: How stable are the capsids?

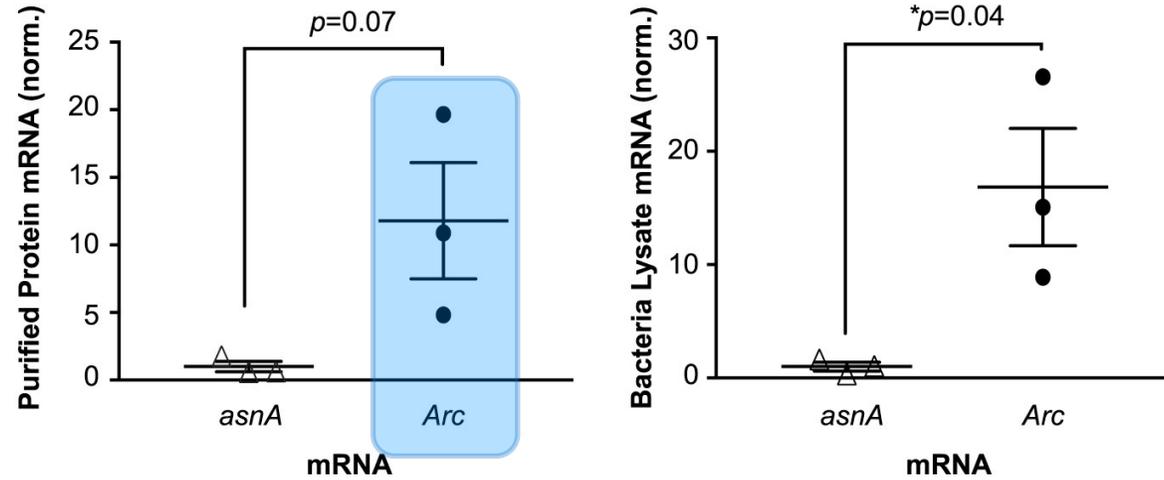


Higher molar concentrations of salt = Higher stability

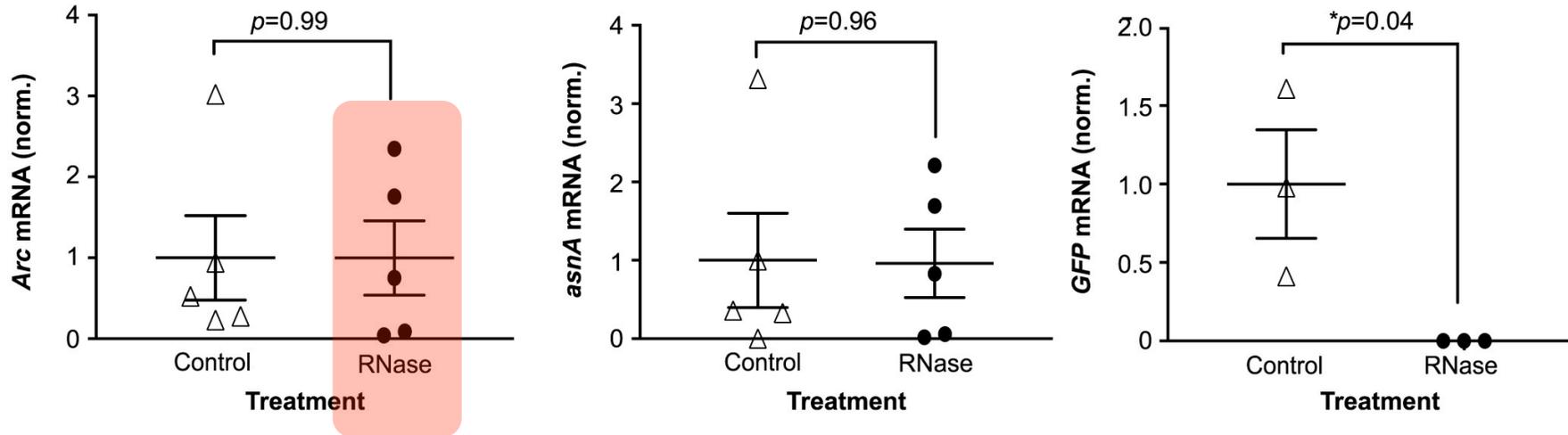
The prArc capsids are very stable, even in high salt

Fig 2A: Does Arc protein bind mRNA?

A

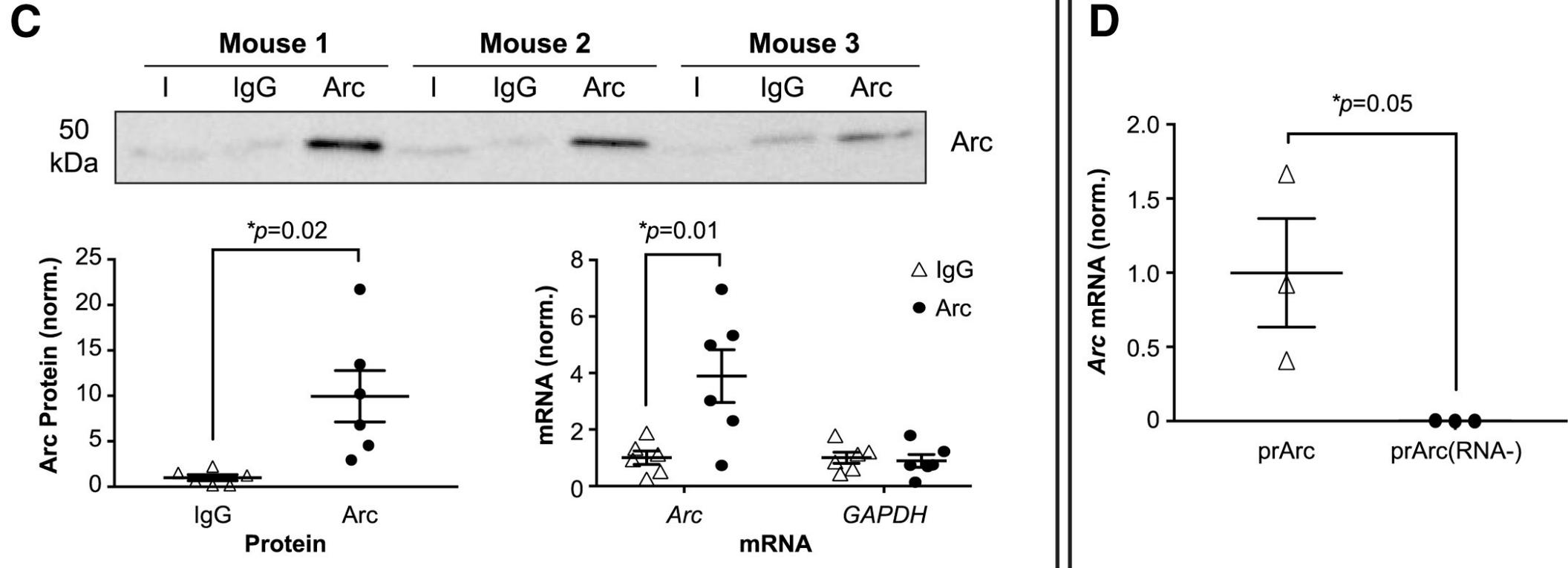


B



RNAase treatment did not affect Arc's ability to bind mRNA

Fig 2C/D: Does Arc protein bind Arc mRNA?



Arc mRNA selectively immunoprecipitates with Arc protein

Fig 2E: Is RNA necessary to assemble Arc capsids?

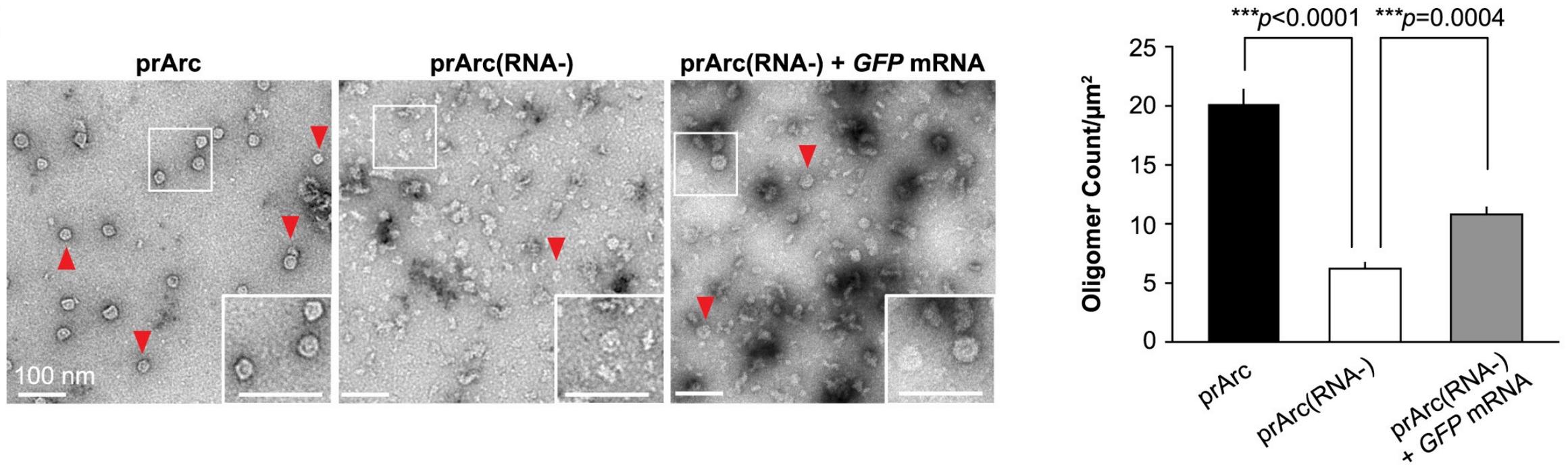
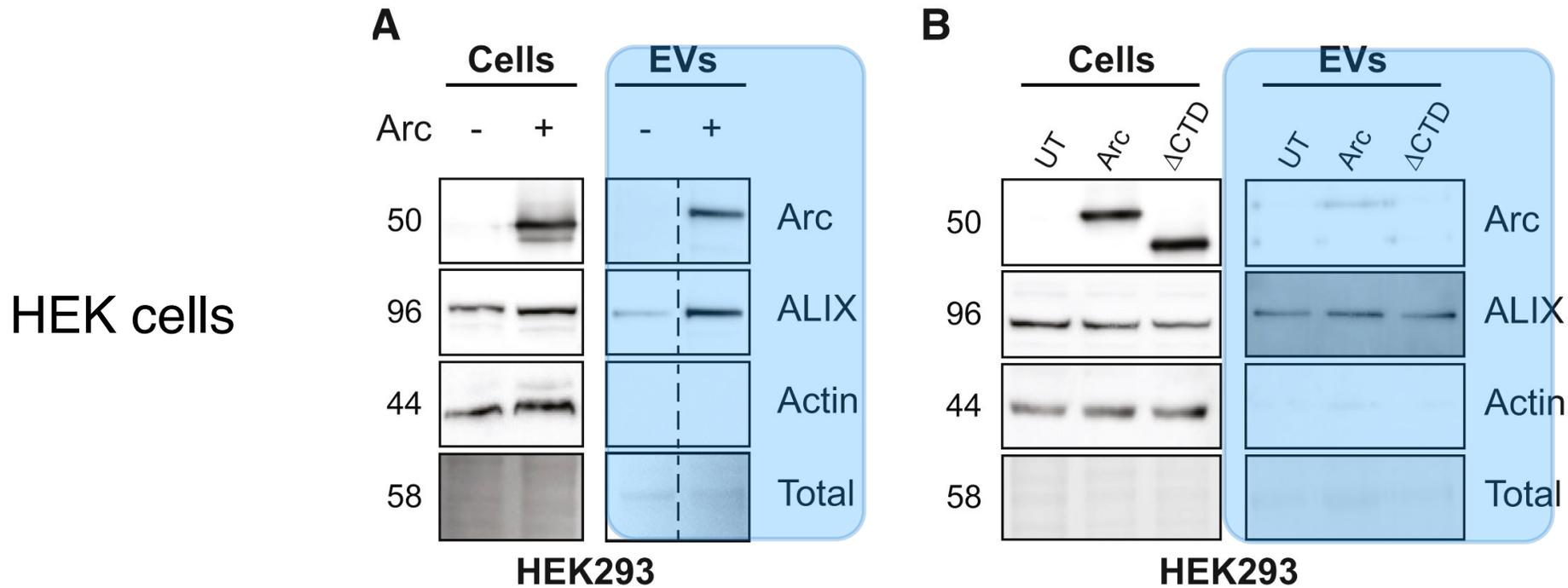
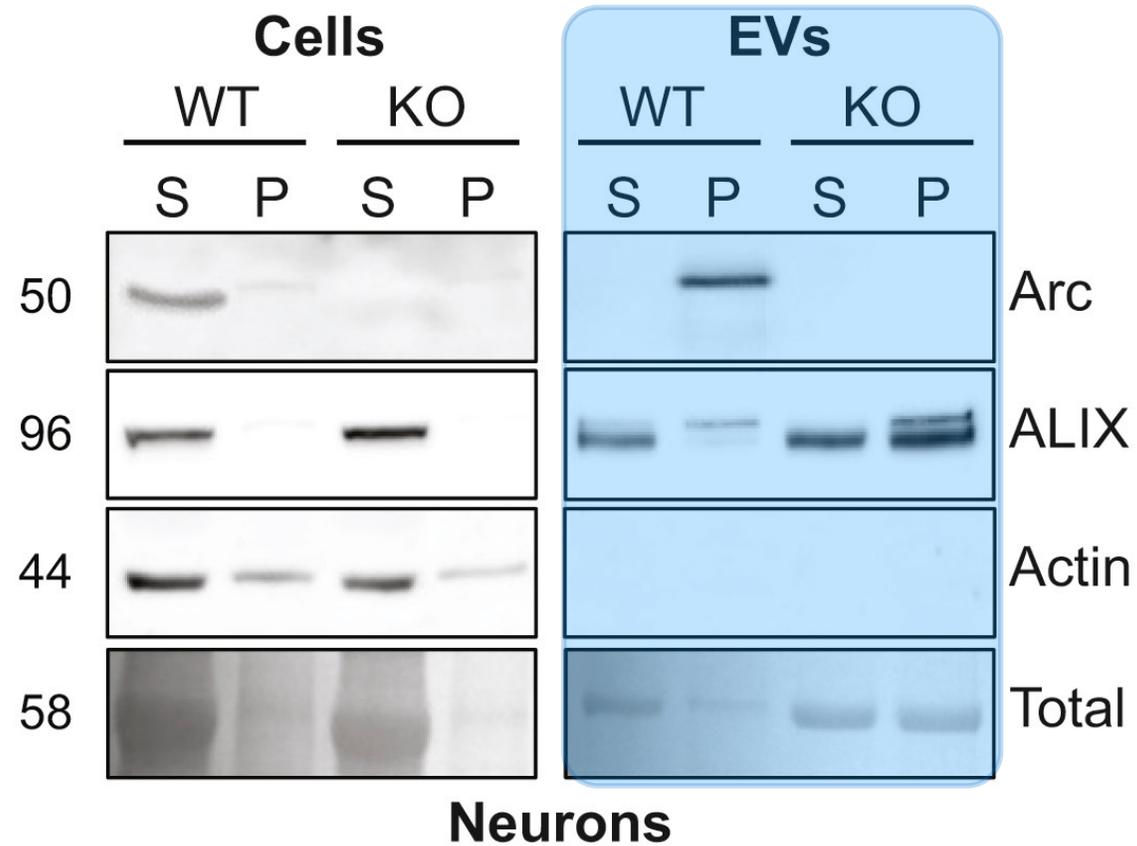


Fig 3A-B: Is Arc released from cells in **extracellular vesicles**?



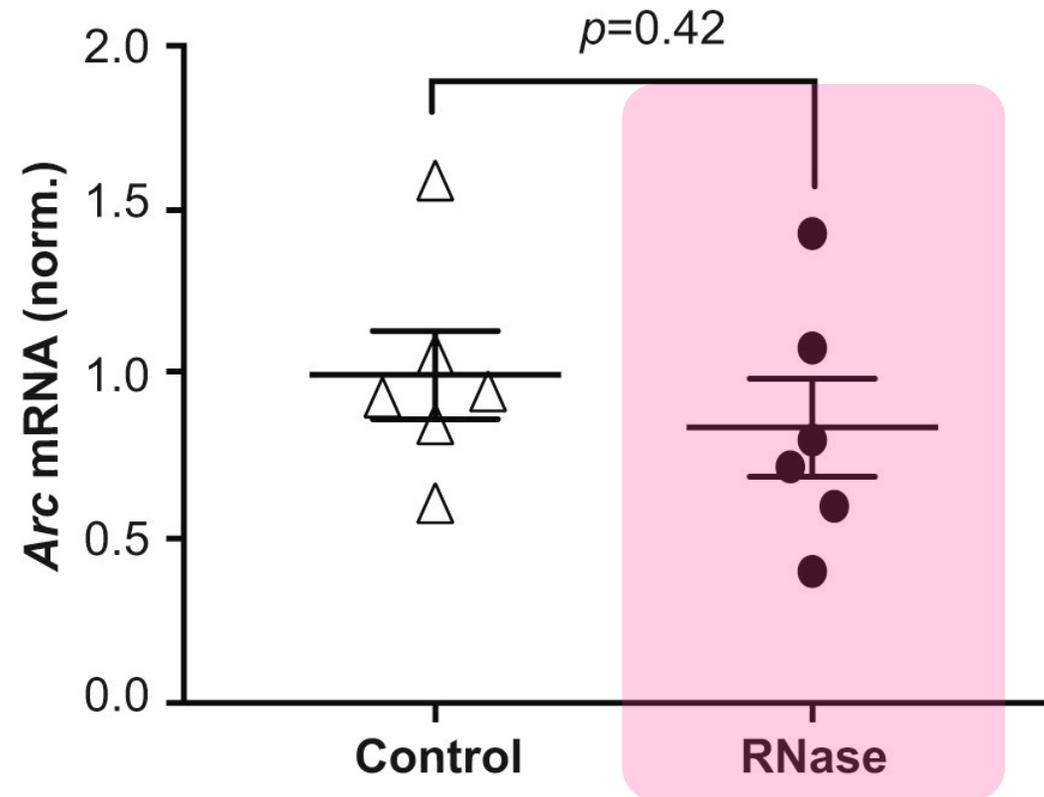
Yes Arc is found outside the cell in EVs!

Fig 3D: Is Arc released from NEURONS in extracellular vesicles?



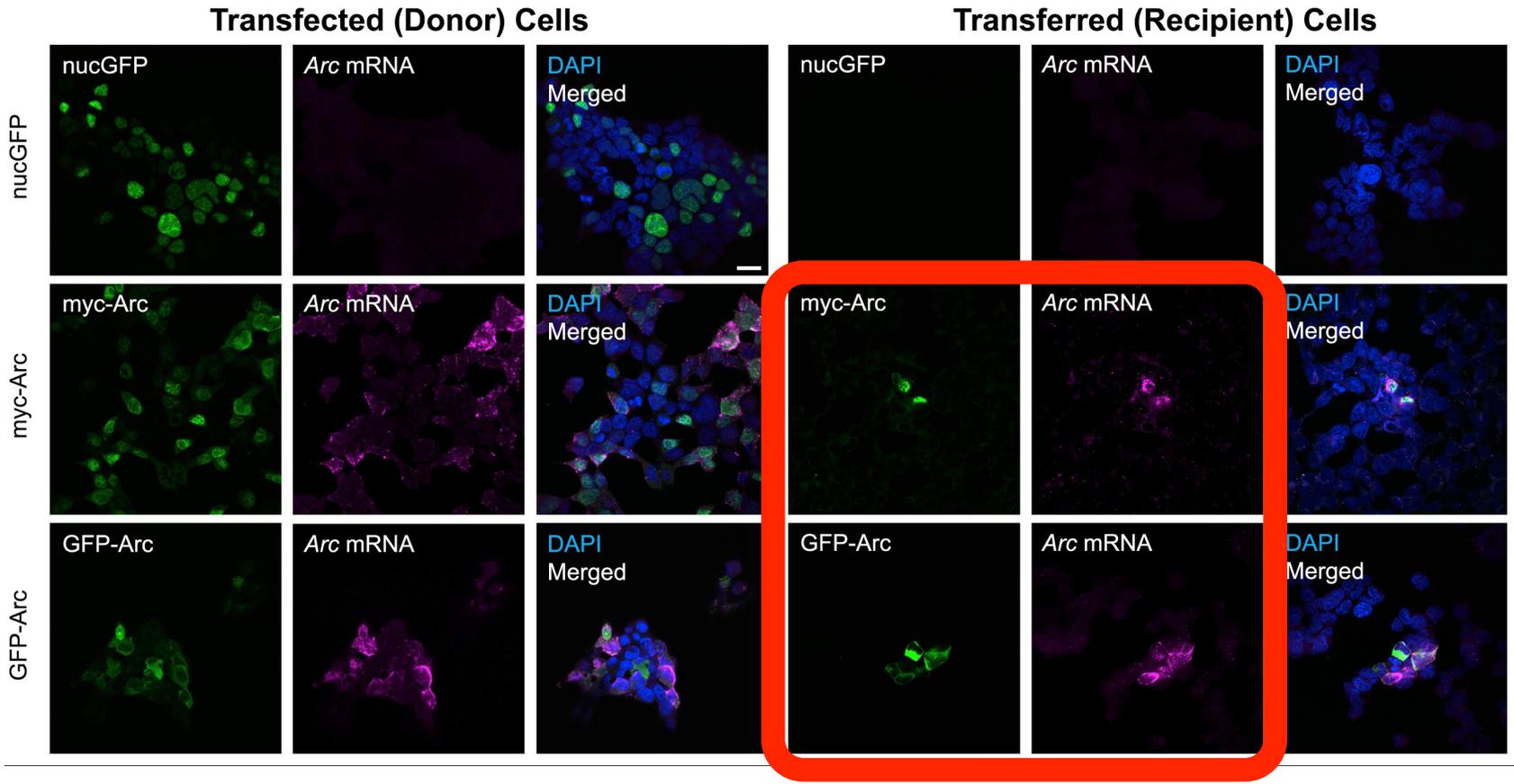
Yes Arc is found outside NEURONS in EVs!

Fig 3C: When these vesicles are released are they **degraded?**



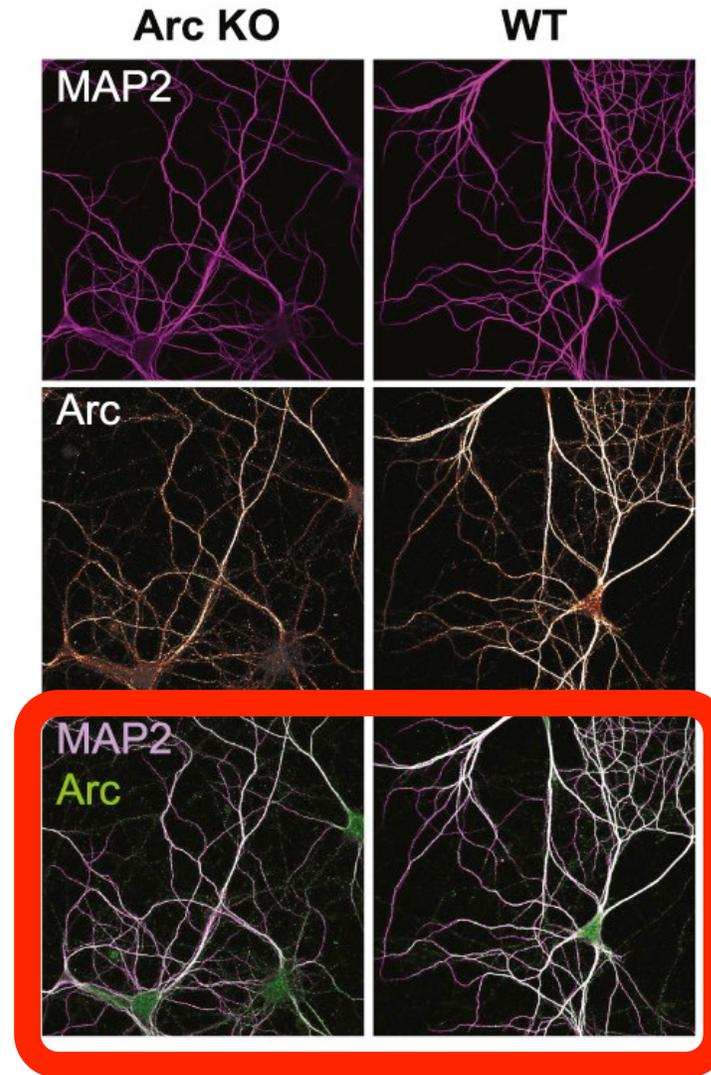
NO! Arc mRNA is stable outside cells!

Fig 4A: Can Arc extracellular vesicles transfer protein and mRNA to recipient HEK cells?



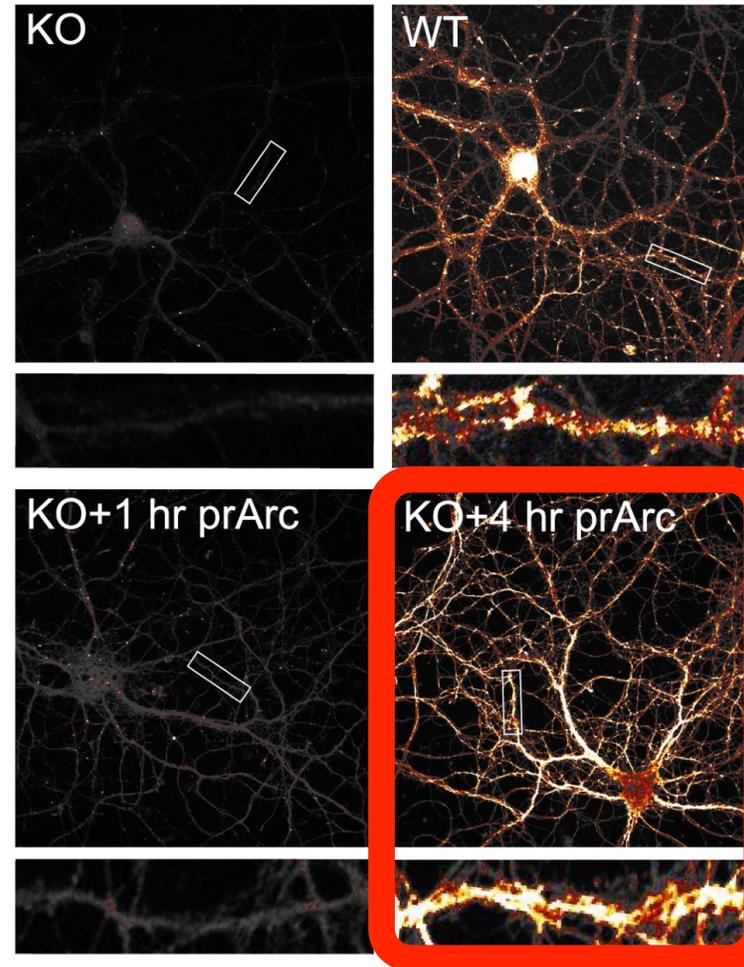
Yes! Both Arc Protein and mRNA can be transferred to recipient HEK cells!

Figure 4B : Can Arc protein be in knockout NEURONS?



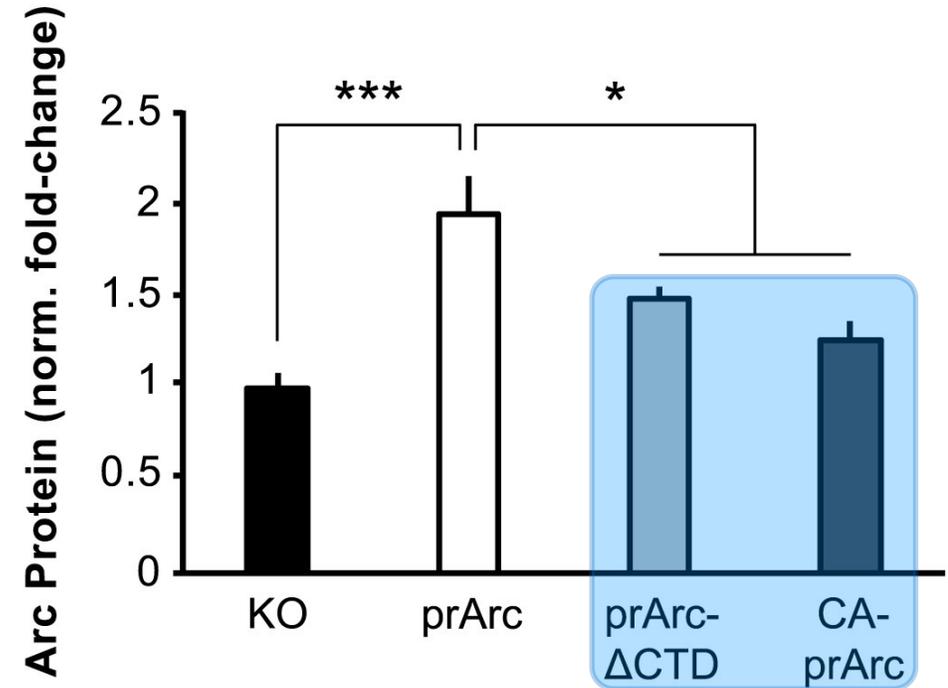
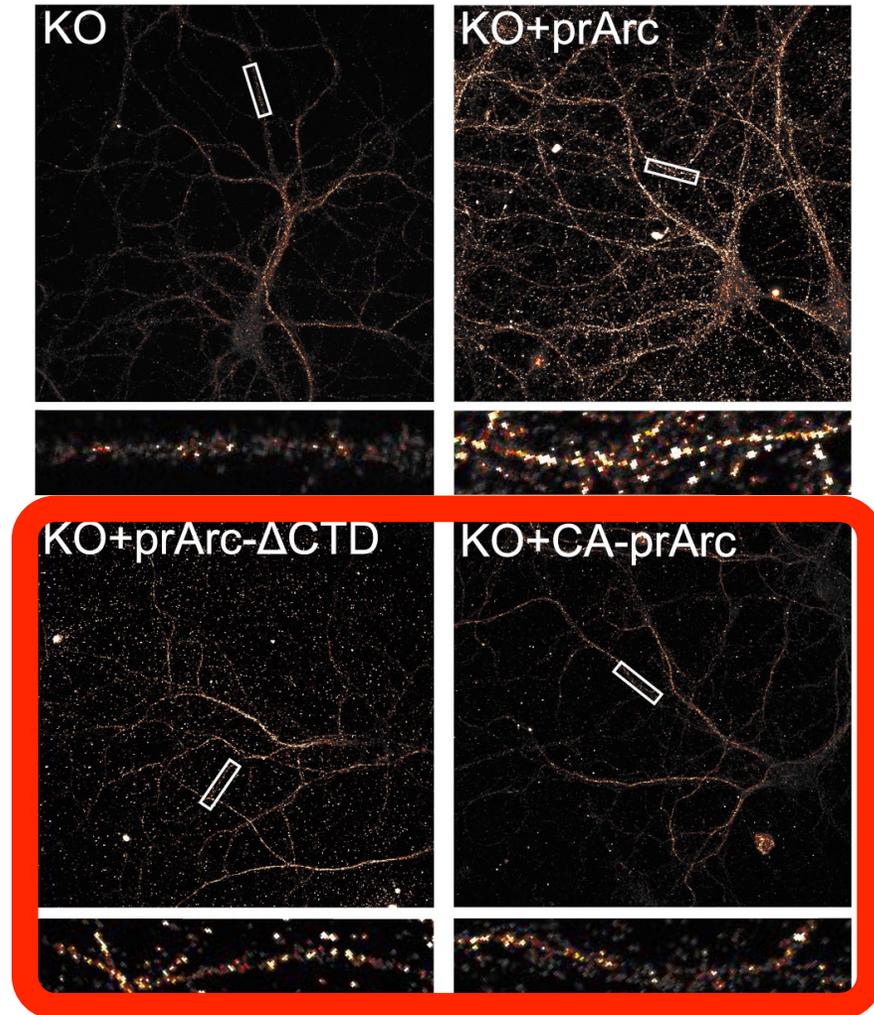
Yes, observed uptake of Arc protein into KO neurons

Fig 5B : Can Arc capsids transfer Arc mRNA into NEURONS?



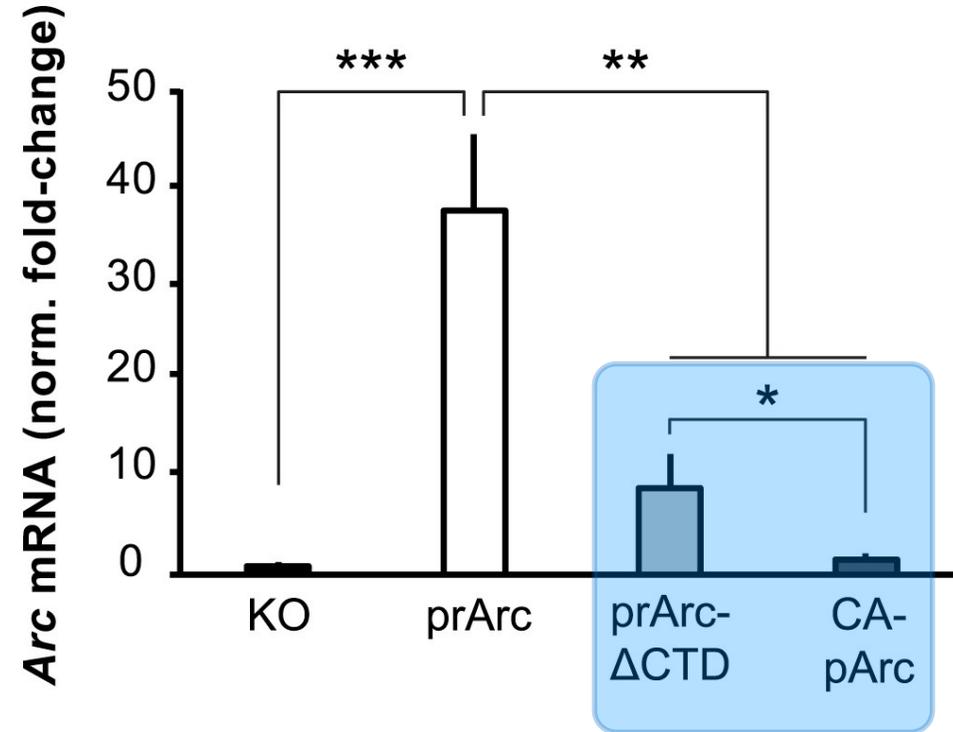
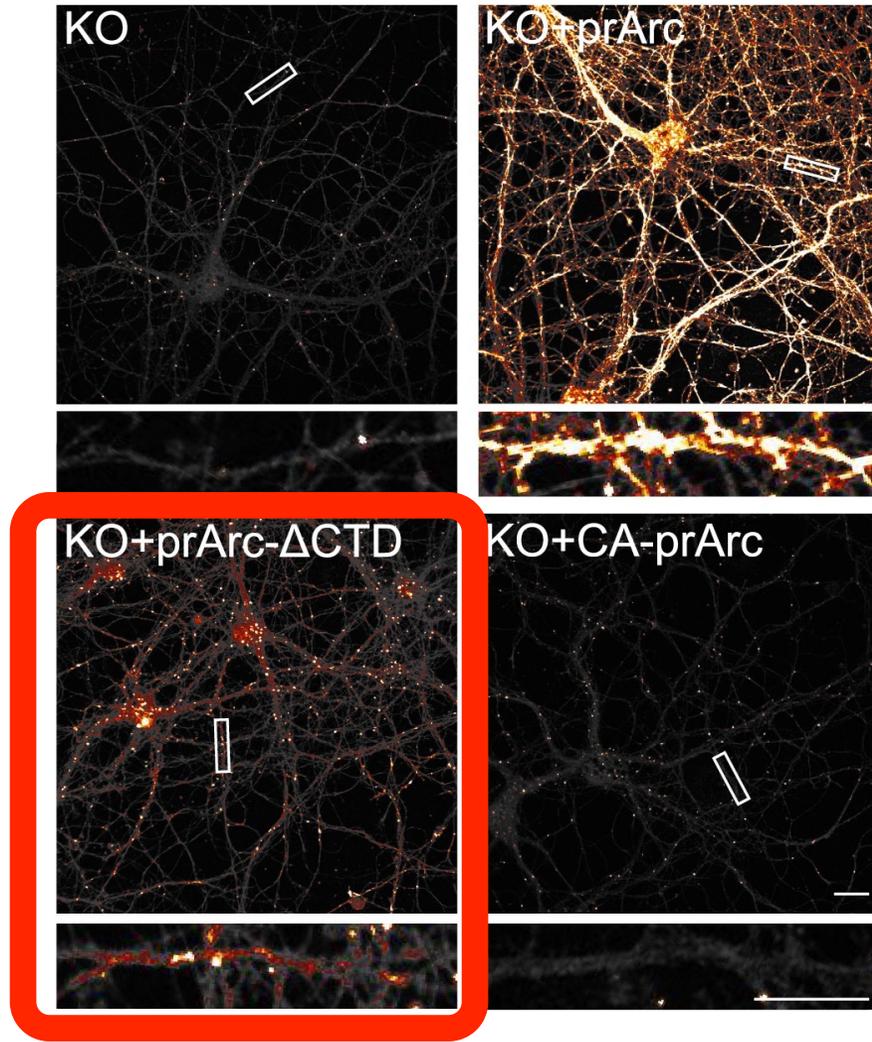
Showed high levels of transferred Arc mRNA

Fig 5C : Was capsid formation required to be taken in by NEURONS?



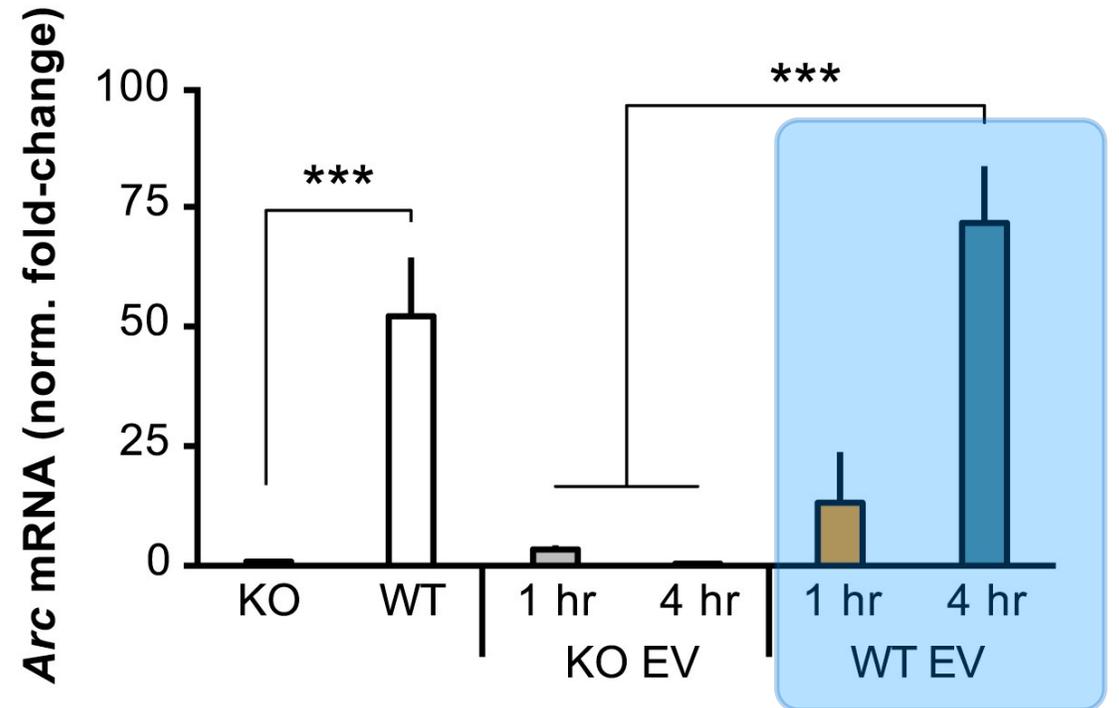
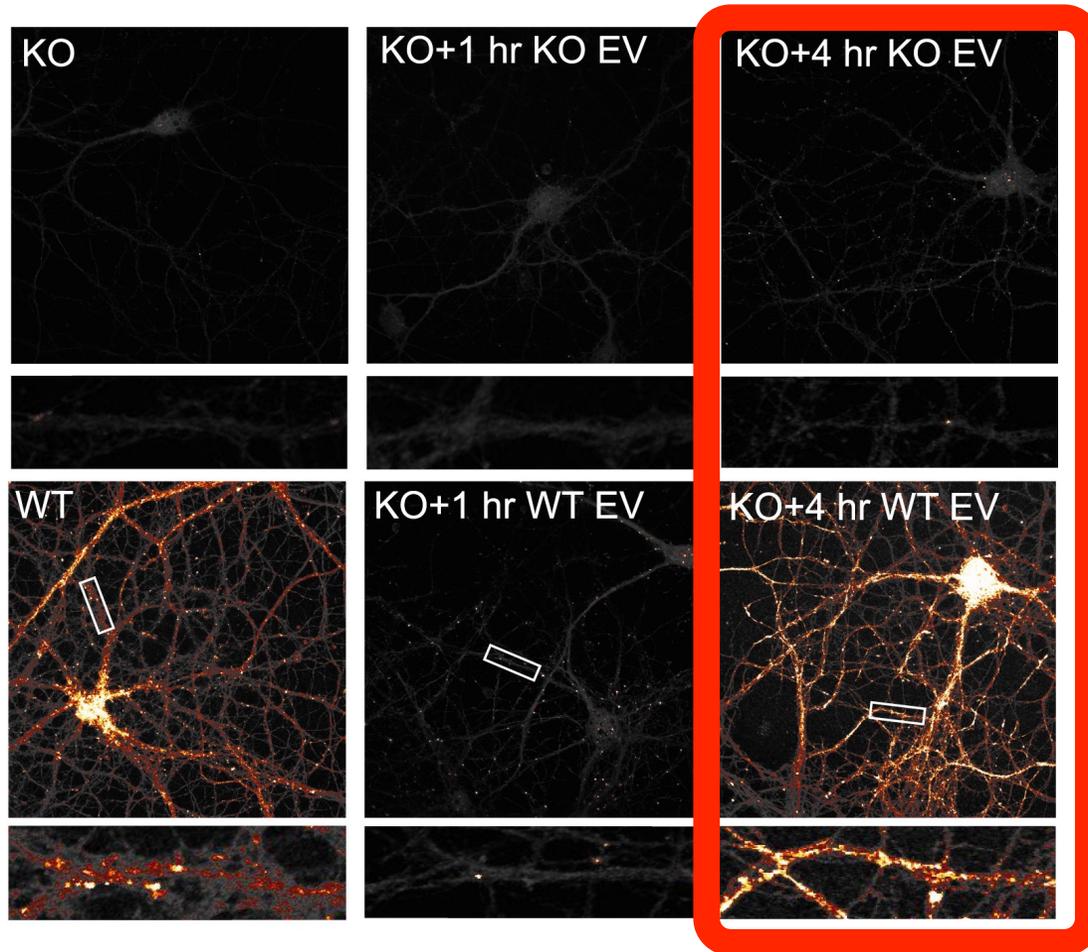
Yes! Capsid formation is required for uptake and transfer of mRNA into neurons

Fig 5D : How did capsid formation effect mRNA uptake into NEURONS?



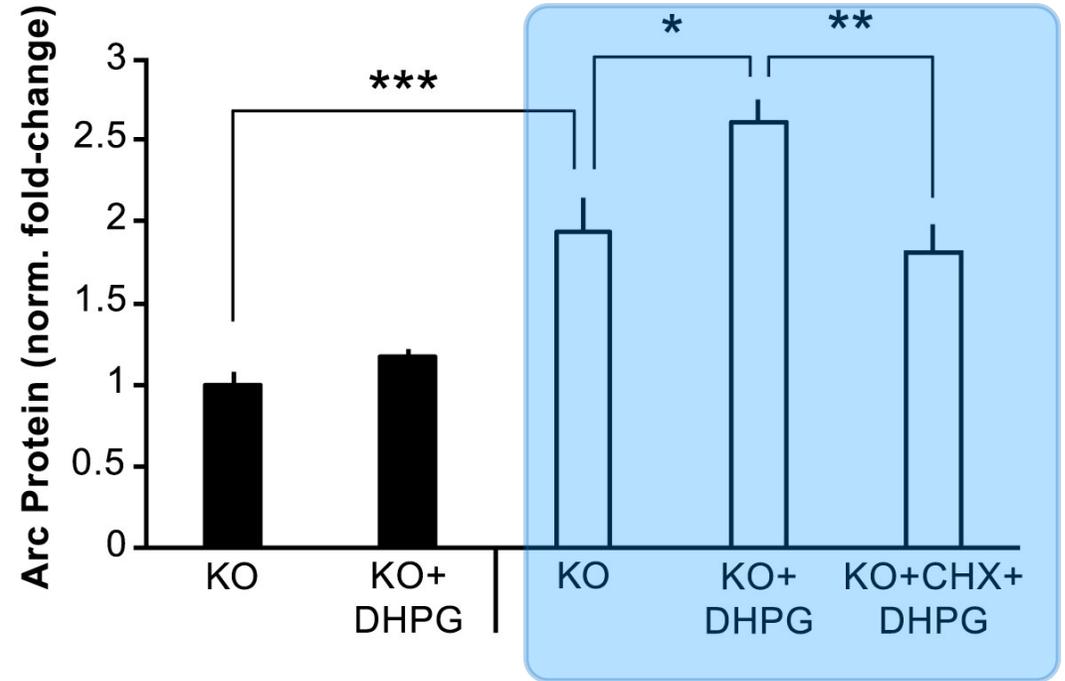
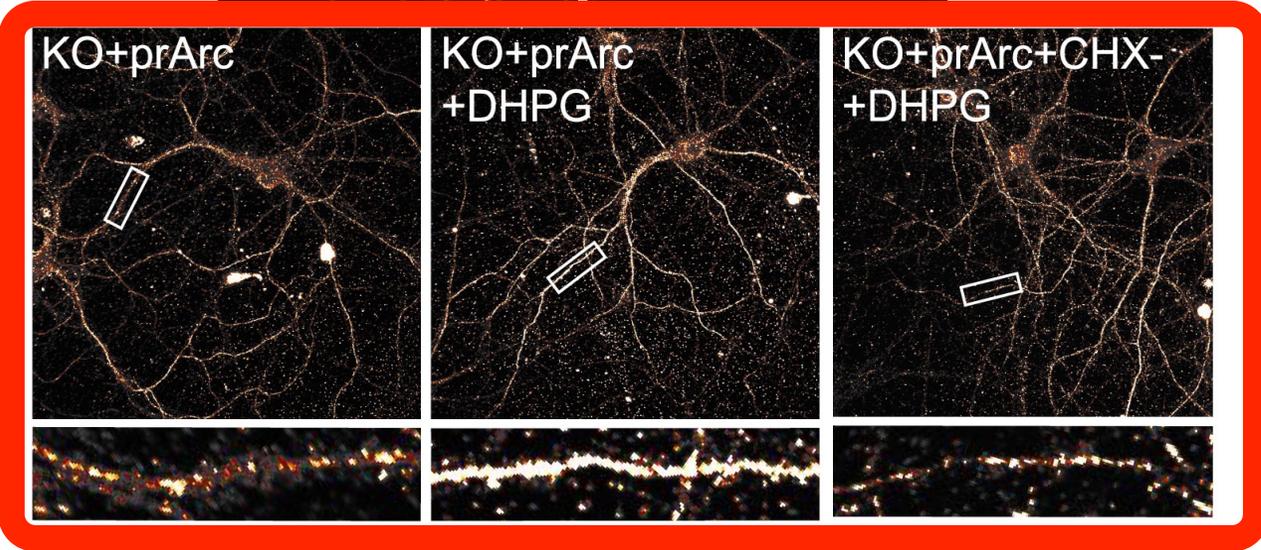
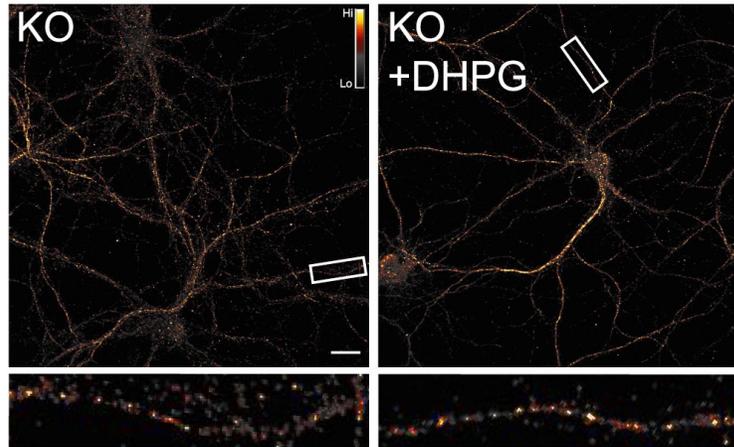
prArc(RNA-) was unable to be taken up, instead coated the outside of neurons

Fig 6B : Can Arc transfer mRNA into knockout extracellular vesicles?



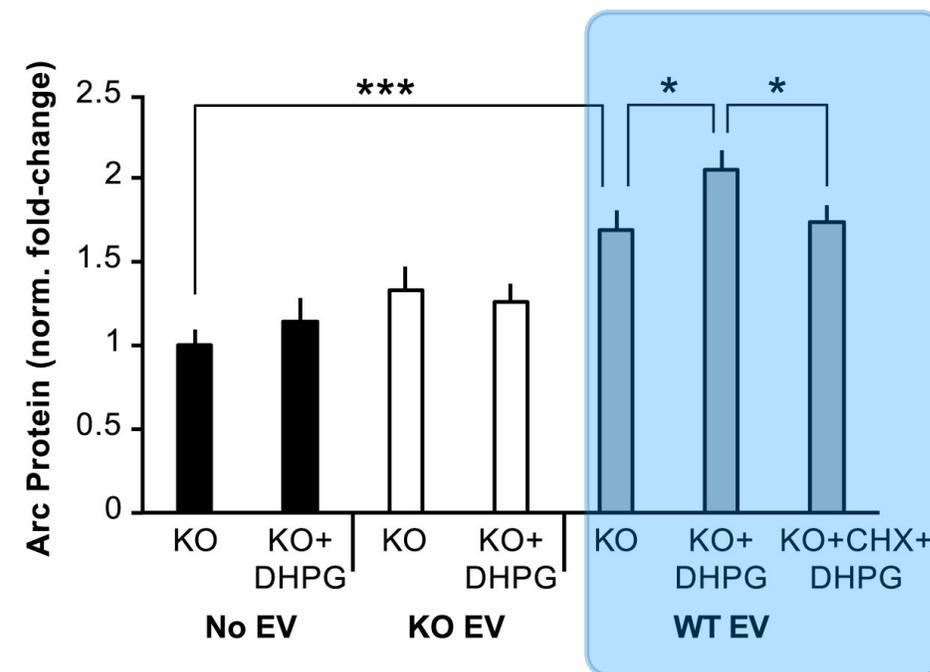
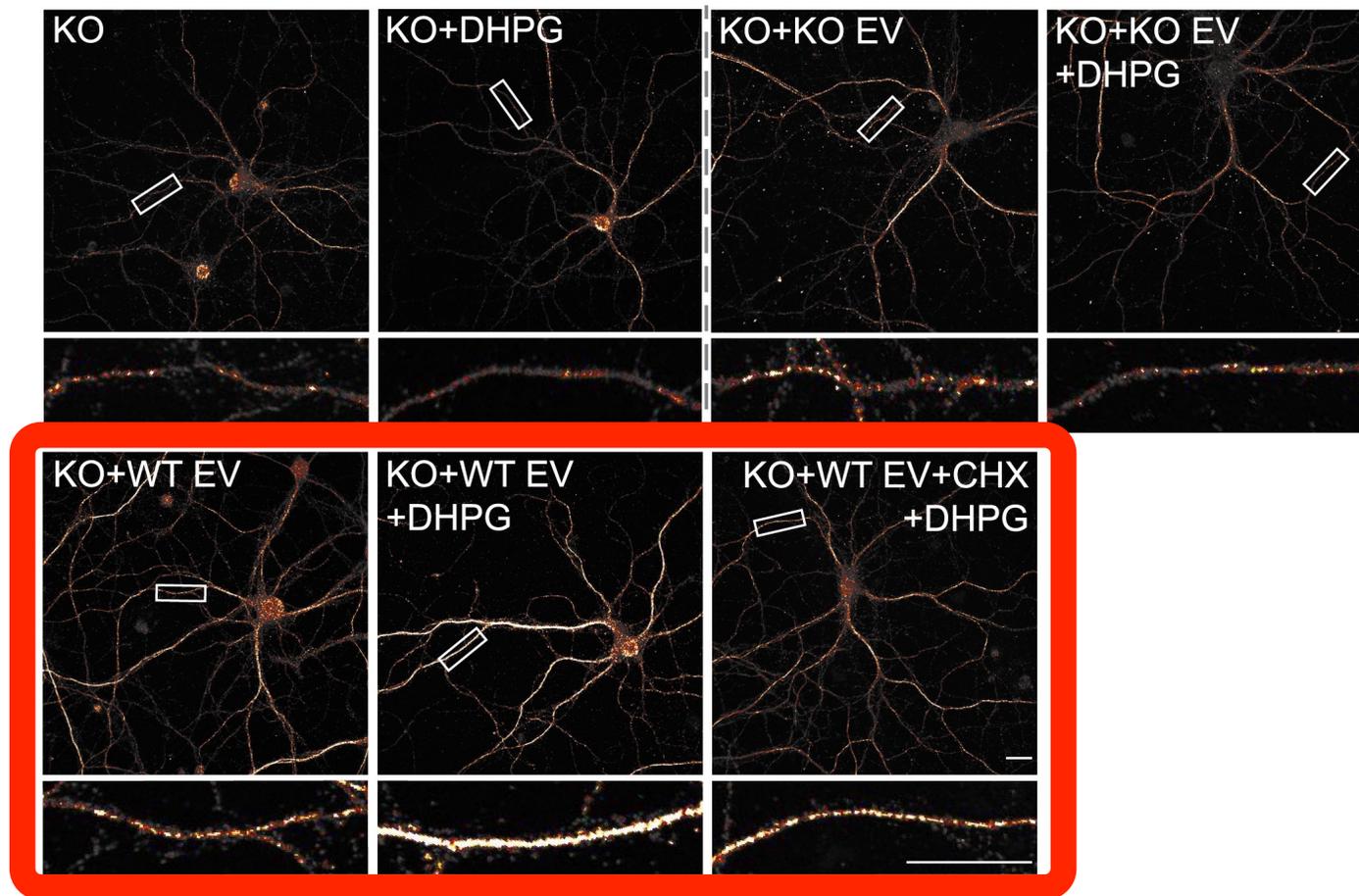
Arc mRNA in WT EVs was transferred into KO neurons

Fig 7A : Is transition occurring after Arc capsids & mRNAs are transferred to NEURONS?



Yes! Once the mRNA is transferred the Arc mRNA message can be translated into a protein in the NEURON

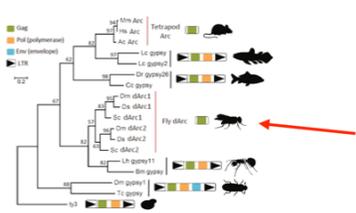
Fig 7B : Is translation occurring after EVs with Arc mRNAs are transferred to NEURONS?



Yes! Arc mRNA can be transferred and translated via EVs in NEURONS

Summary of Paper and Images

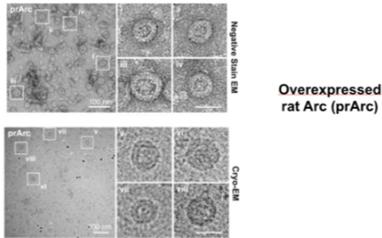
Fig 1A: What was the maximum likelihood phylogeny of Arc?



Closely Related To Tv3/cvovsv Retrotransposons

Arc contains Gag proteins and is found in most eukaryotes

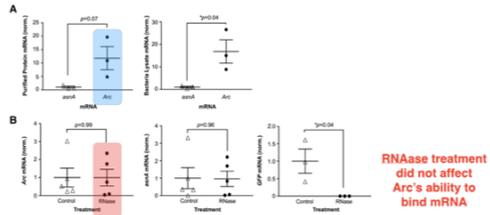
Fig 1B: How are Arc capsids structured?



Arc capsids showing the double-layered capsid shell

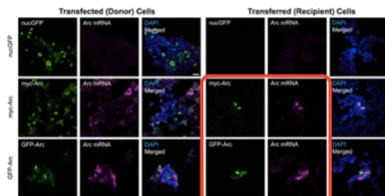
Arc forms capsids, a double-layer shell

Fig 2A: Does Arc protein bind mRNA?



Arc is able to store and transfer mRNA

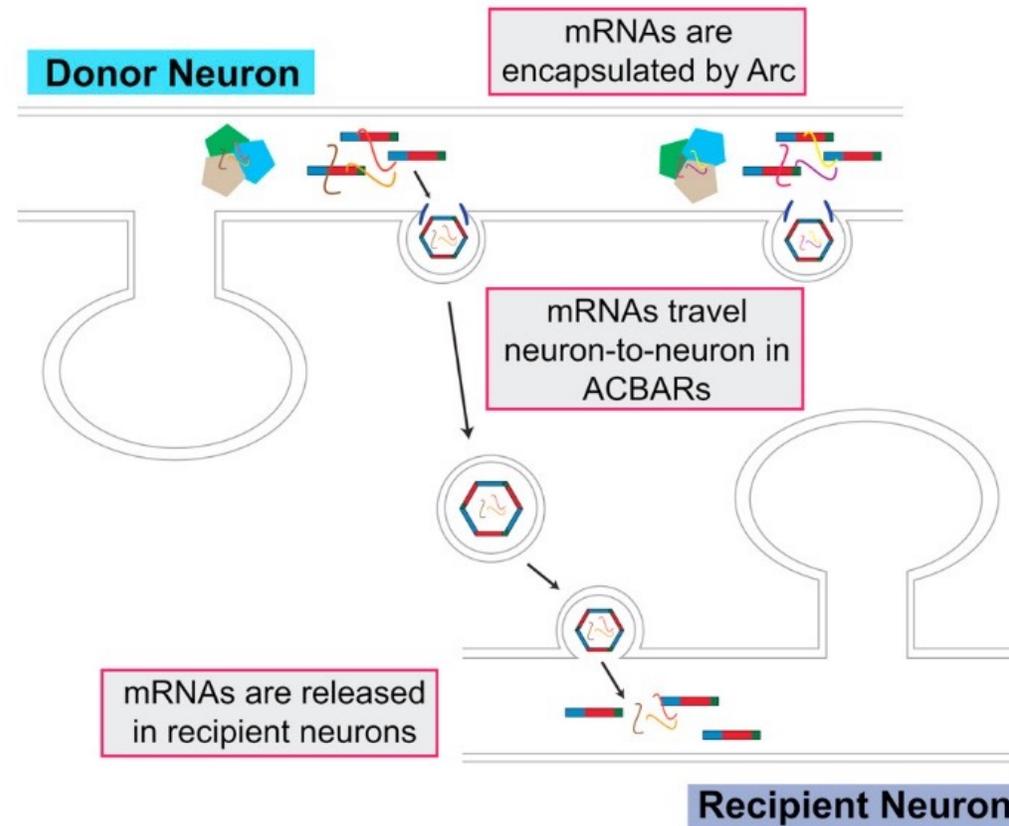
Fig 4A: Can Arc extracellular vesicles transfer protein and mRNA to recipient HEK cells?



Yes! Both Arc Protein and mRNA can be transferred to recipient HEK cells!

Arc mRNA can signal between all types of cells

What are applications of this discovery?



Utilizing this “naturally occurring mechanism of RNA transfer, we believe that this pathway may allow new means of genetic engineering or RNA delivery into cells”.

What questions need answered?

What cellular factors affect capsid biogenesis and packing?

What other cargo do these capsids contain?

What are the docking mechanisms for these capsids?



Additional information

<https://www.biostat.wisc.edu/bmi576/syllabus.html>

- Sequence alignment and HMM

<https://www.biostat.wisc.edu/bmi776/>

- Motif discovery

Profile HMM

- *Biological Sequence Analysis: Probabilistic Models of Proteins and Nucleic Acids*, Chapter 5

Questions?

Work cited

- 1.) <http://genetics564.weebly.com/motif--domain-discovery.html>
- 2.) http://www.spring8.or.jp/en/news_publications/press_release/2010/100317/
- 3.) <http://carrot.mcb.uconn.edu/~olgazh/bioinf2010/class12.html>
- 4.) Rogers, K. (n.d.). What Is the Difference Between a Peptide and a Protein? Retrieved February 17, 2018, from <https://www.britannica.com/story/what-is-the-difference-between-a-peptide-and-a-protein>
- 5.) <https://www.biologicscorp.com/blog/protein-folding-and-protein-structure-prediction-problem/#.WoepG6inGM8>
- 6.) https://www.flaticon.com/free-icon/idea-hand-drawn-symbol-of-a-side-head-with-a-lightbulb-inside_35497
- 7.) <https://blog.dellemc.com/en-us/long-term-retention-its-in-our-dna/>
- 8.) <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-5-113>
- 9.) Ong, Irene. (2017). Markov models [PDF]. <https://www.biostat.wisc.edu/bmi576/syllabus.html>
- 10.) Ong, Irene. (2017). hmms [PDF]. <https://www.biostat.wisc.edu/bmi576/syllabus.html>
- 11.) <https://bibiserv.cebitec.uni-bielefeld.de/sadr2/databasesearch/hmmer/index.html>
- 12.) Grange, Sara. Ness, Katie. (2017) Domains, Motifs and Pfam/SMART Analysis [PDF]. http://genetics564.weebly.com/uploads/8/6/5/7/865764/saragrangekatieness2-16-16final_pres_gen564.pdf
- 13.) <http://gydb.org/index.php/Intro>
- 14.) <https://www.sciencedirect.com/science/article/pii/S0960982212004459>
- 15.) Nolan Lab, Department of Molecular Pharmacology, Stanford University. *The Gag, Pol, and Env Proteins*. https://web.stanford.edu/group/nolan/OldWebsite/tutorials/ret_3_maj_prot.html
- 16.) <https://geresidents.wordpress.com/2016/11/20/aqm-and-future-directions-community-forum/>
- 17.) <https://hubpages.com/education/Genetics-What-Are-Transposons>

Work cited

<http://www.wisegeek.com/what-is-a-cell.htm>

<http://www.abcam.com/primary-antibodies/extracellular-vesicles-an-introduction>

[https://bio.libretexts.org/TextMaps/Map%3A_Introductory_Biology_\(K-12\)/5%3A_Evolution/5.4%3A_First_Cells](https://bio.libretexts.org/TextMaps/Map%3A_Introductory_Biology_(K-12)/5%3A_Evolution/5.4%3A_First_Cells)

The Neuronal Gene Arc Encodes a Repurposed Retrotransposon Gag Protein that Mediates Intercellular RNA Transfer. *Cell*. 2018 Jan 11;172(1-2):275-288.e18. doi: 10.1016/j.cell.2017.12.024.