

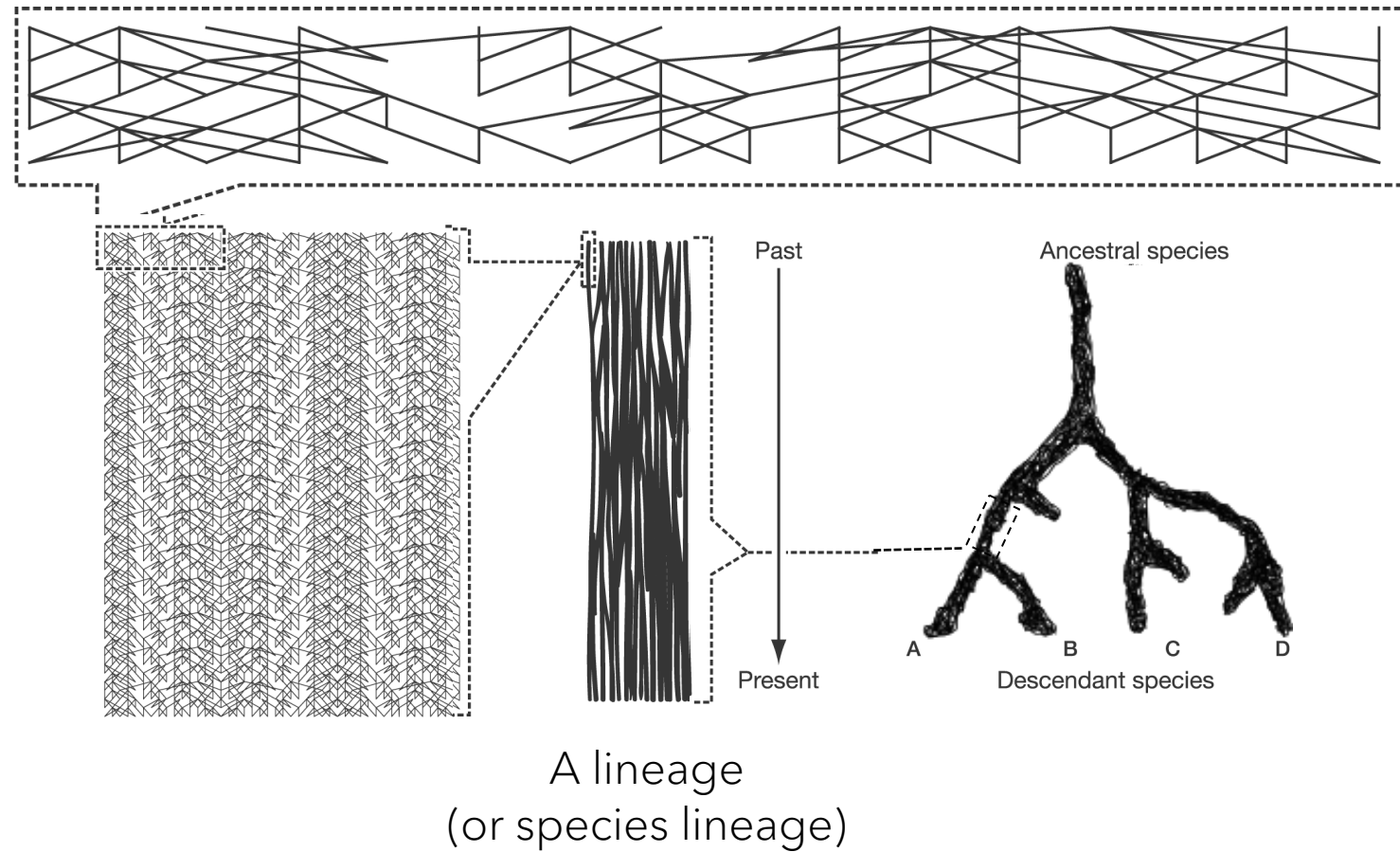
Phylogenomics

Lucas Moyer
and Olivia Steidl

How do we record reproductive history?

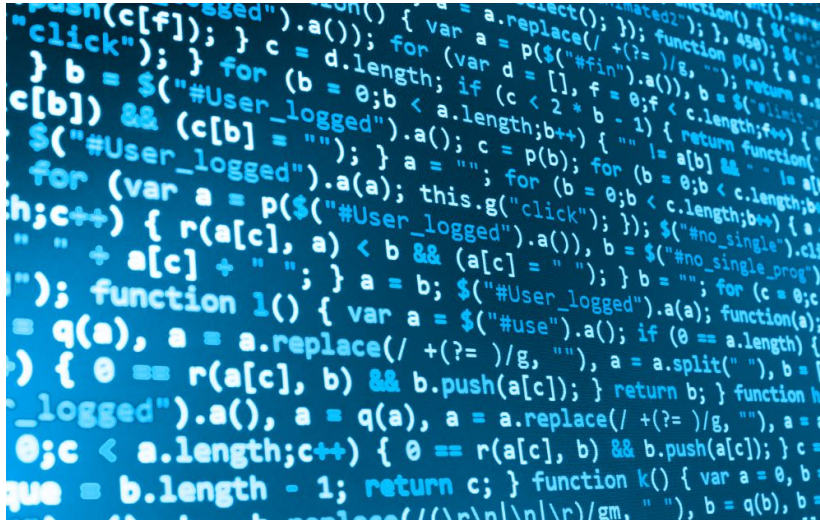
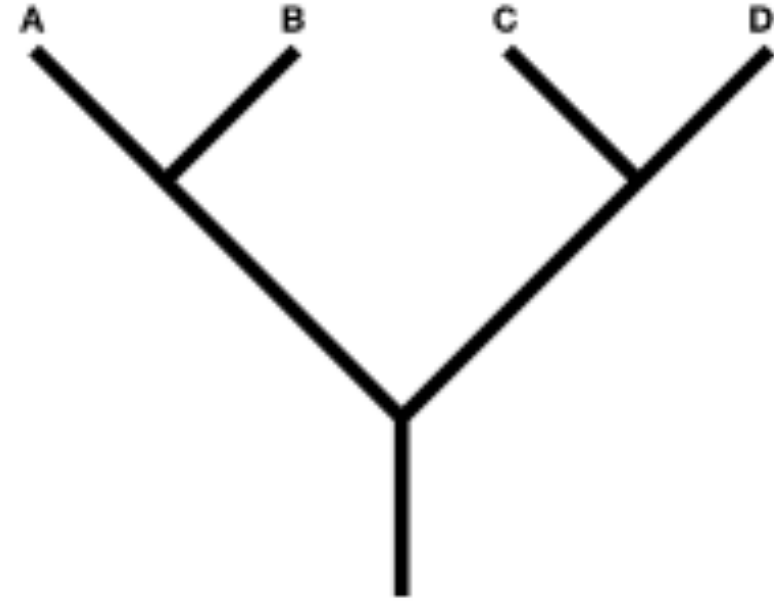
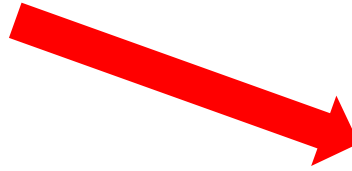
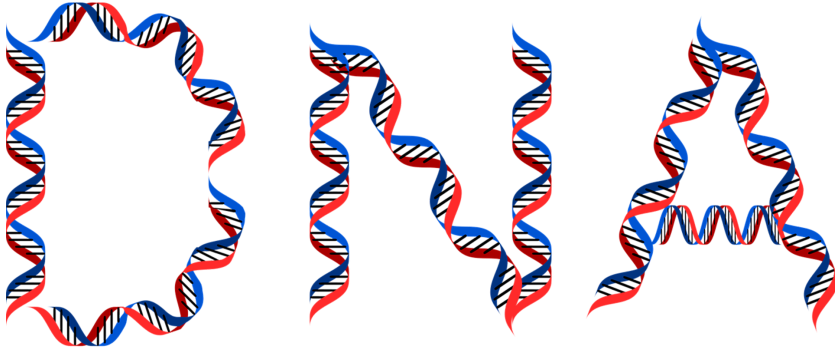


How does reproductive history get represented?



The dense network of lines become a branch

How is genetics used to build trees?



Estimating evolutionary past **based on DNA sequencing using algorithms**

What are some key terms you should know?

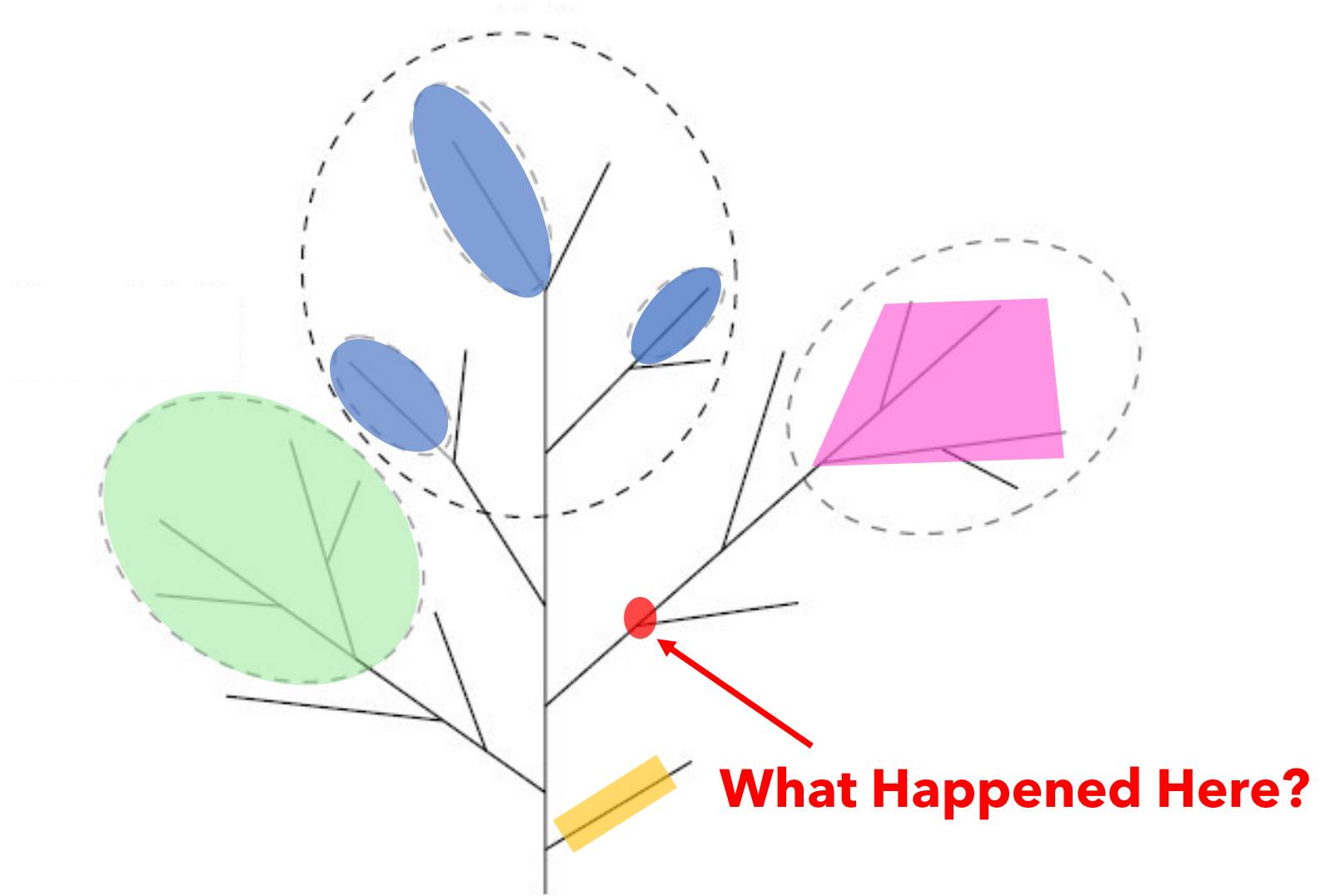
Branches

Nodes

**Monophyletic
Group/Clade**

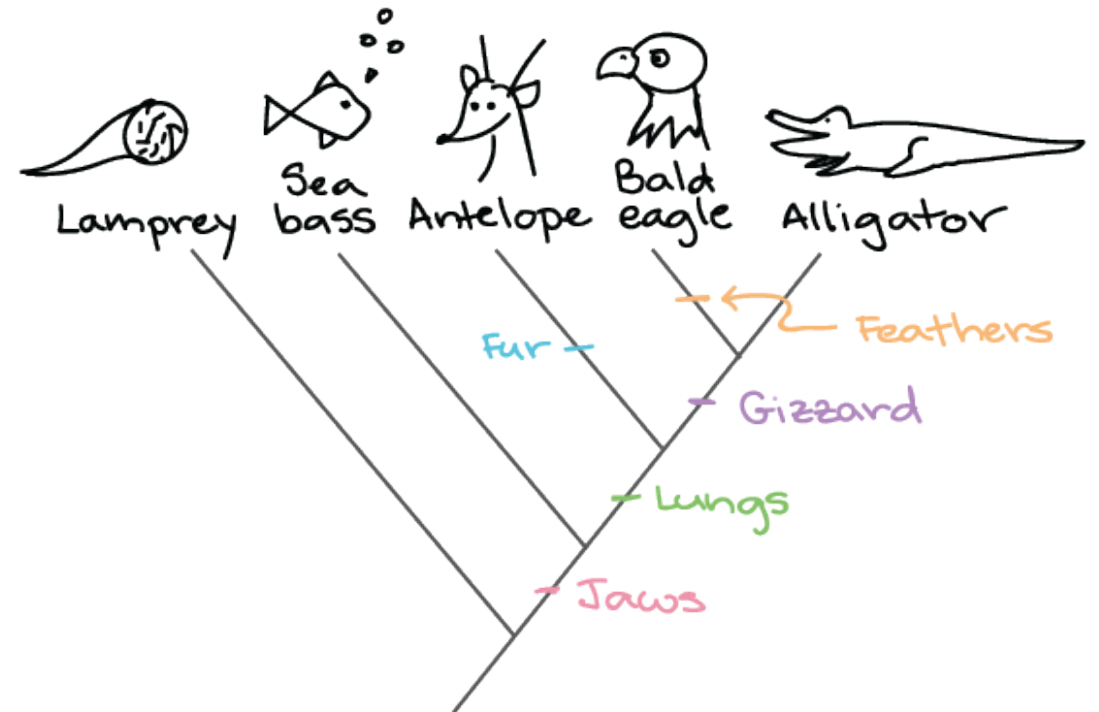
Paraphyletic Group

Polyphyletic "Group"



How do you make a tree from a group of organisms?

Feature	Lamprey	Antelope	Bald eagle	Alligator	Sea bass
Lungs	0	+	+	+	0
Jaws	0	+	+	+	+
Feathers	0	0	+	0	0
Gizzard	0	0	+	+	0
Fur	0	+	0	0	0



Identify characteristics and use *parsimony* and *maximum likelihood*

What can I use to build trees?



PHYML - Medium



Molecular Evolutionary
Genetics Analysis

MEGA2 - Easy



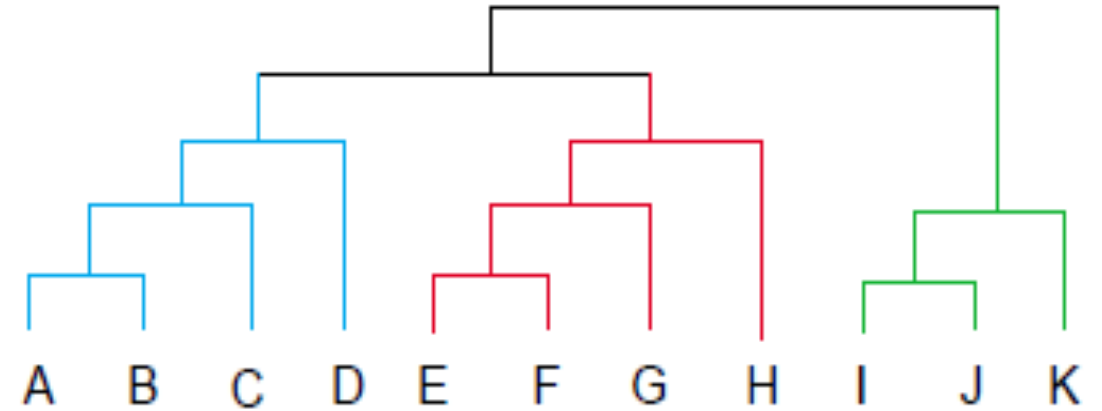
PAUP* - Hard

How does the algorithm build a tree?

Multiple Sequence Alignment

ClustalX, BioEdit

Guide Tree



Sequence
Addition
Order

Step 1

A + B

E + F

I + J

Step 2

AB + C

EF + G

IJ + K

Step 3

ABC + D

EFG + H

Step 4

ABCD + EFGH

Step 5

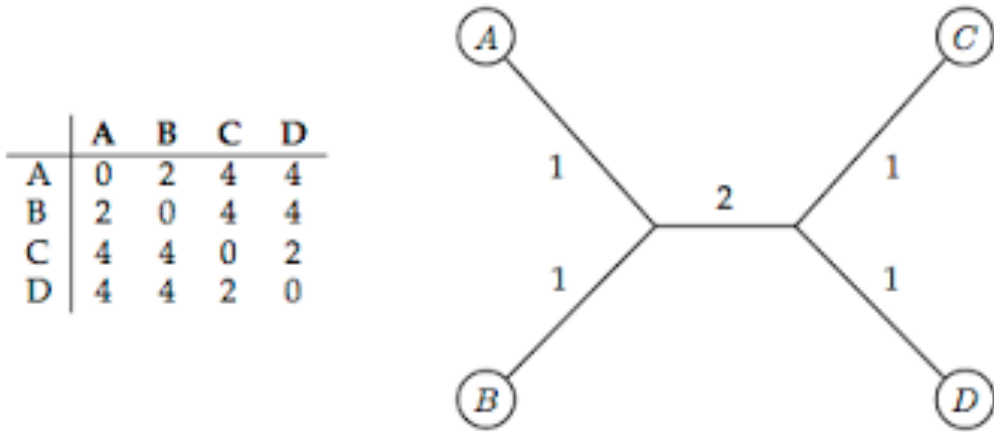
ABCDEFGH + IJK

How is sequencing data used?

taxon	10....20....30....40....50
Fu	Nosema.40928	QFGLFSPEEIRASSVAL--IRYPETLE--NGVPKESGLVCAGHFGHIELVK
Fu	Aspergillus.	QFGLFSPEEIKRMSVVH--VEYPETMDEQRQRPRTKGLECPGHFGHIELAT
Fu	Spombe.54881	QFGILSPEEIRSMSVAK--IEFPETMDESQQRPRVGGLDCPGHFGHIELAK
Ap	Plasmodium.3	ELGVLDPEI IKKISVCE--IVNVDIYK--DGFPREGGLYCPGHFGHIELAK
An	Cricetulus.2	QFGVLSPPDELKRMSVTEGGIKYPETTE--GGRPKLGGLECPGHFGHIELAK
An	Homo.7434727	QFGVLSPPDELKRMSVTEGGIKYPETTE--GGRPKLGGLECPGHFGHIELAK
An	Drosophila.9	QFGILSPDEIR RMSVTEGGVQFAETME--GGRPKLGGLECPGHFGHIDLAK
An	Celegans.133	QFGILGP E EIKRMSVAH--VEFP EVYE--NGKPKLGGLD CPGHFGHLELAK
Pl	Athaliana.40	QFGILSPDEIRQMSVIH--VEHSETTE--KGKPKVGGLECPGHFGYLELAK
My	Ddiscoideum.	-----EC PGHFGHIELAK
Rh	Porphyra.316	-----EC PGHFGFIELAK
Kt	Tbrucei.1021	QFEIFKERQIKSYAVCL--VEHAKSYA--NAADQSGEAEC PGHFGYIELAE
Kt	Leishmania.7	QFEVFKEAQIKAYAKCI--IEHAKSY--EHGQPVRGGIECPGHFGYVELAE

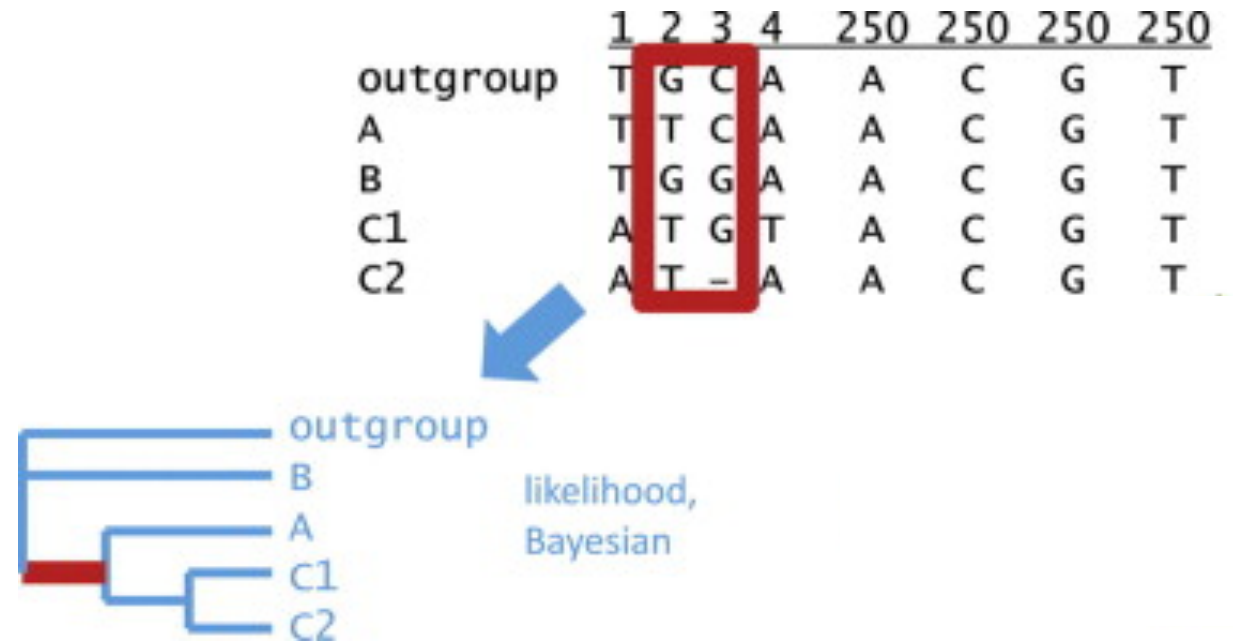
Minimize indels and *think logically*

How to decide where to put species?



Distance Matrix

Distance is essentially the percent difference between the sequences



Tree Searching

vs Includes Maximum likelihood and parsimony which examine each sequence column individually

Why should I use distance matrix or tree searching?

Distance Matrix



Fast and Simplistic

Tree Searching

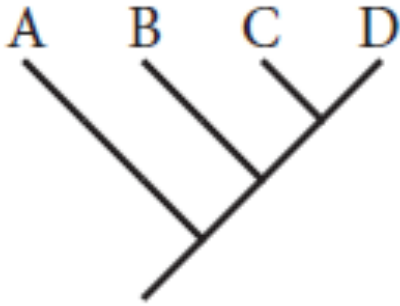
Slow and Computing Intensive



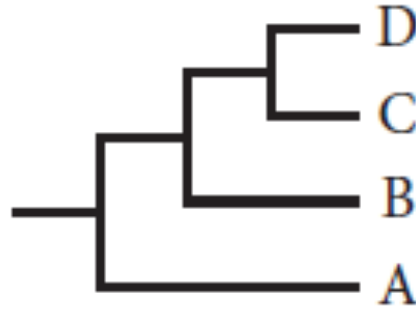
Only Yields a Tree

Data Rich

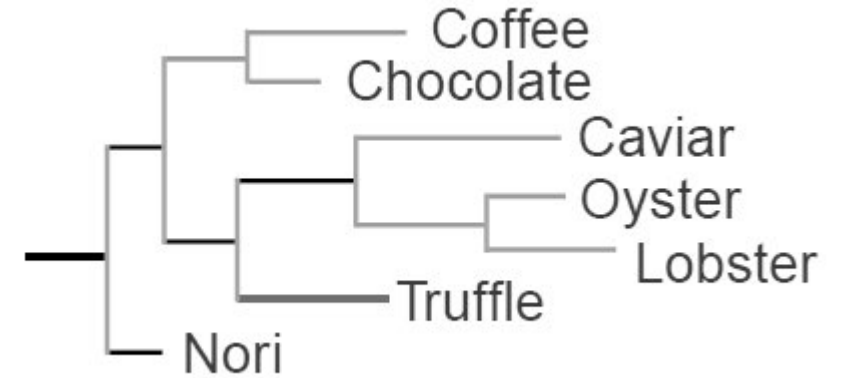
What tree should I use?



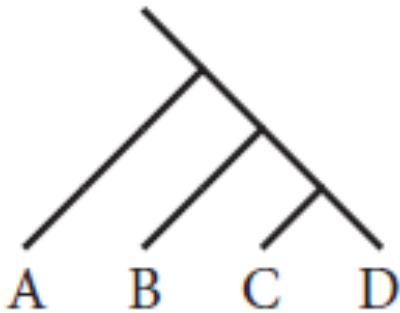
Diagonal-up



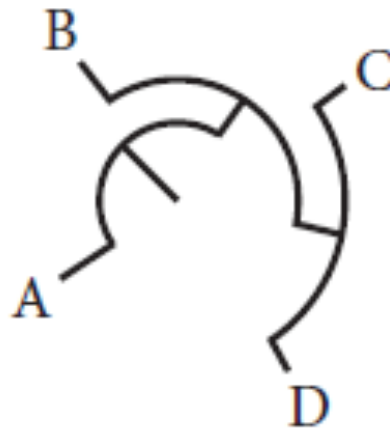
Rectangular-right



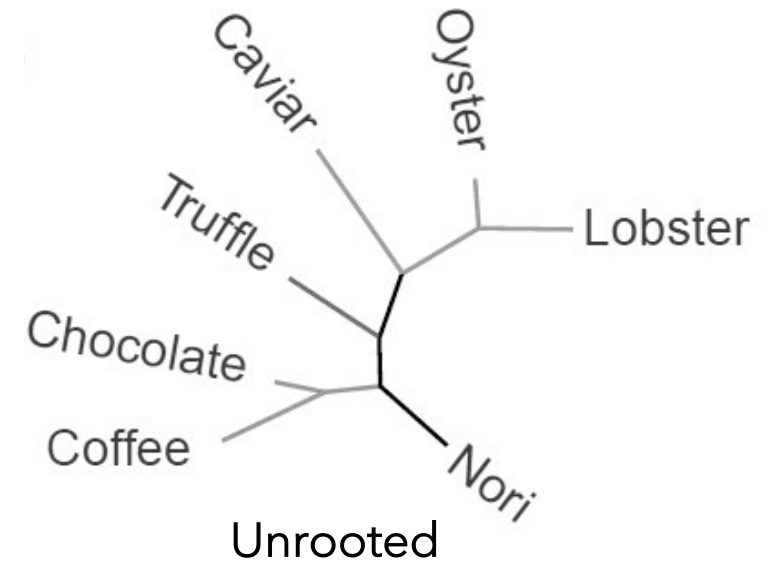
Rectangular-right



Diagonal-down

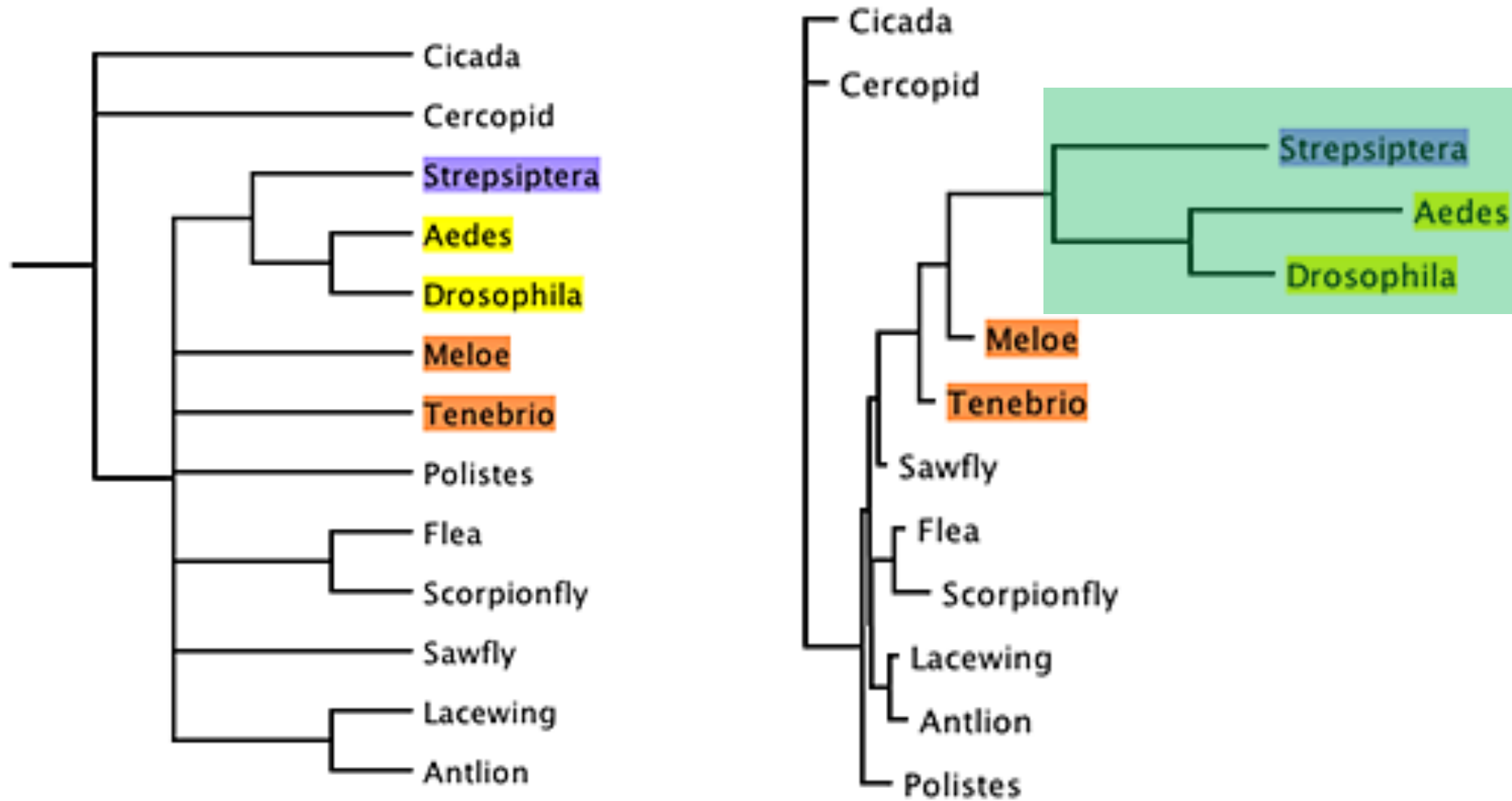


Circle



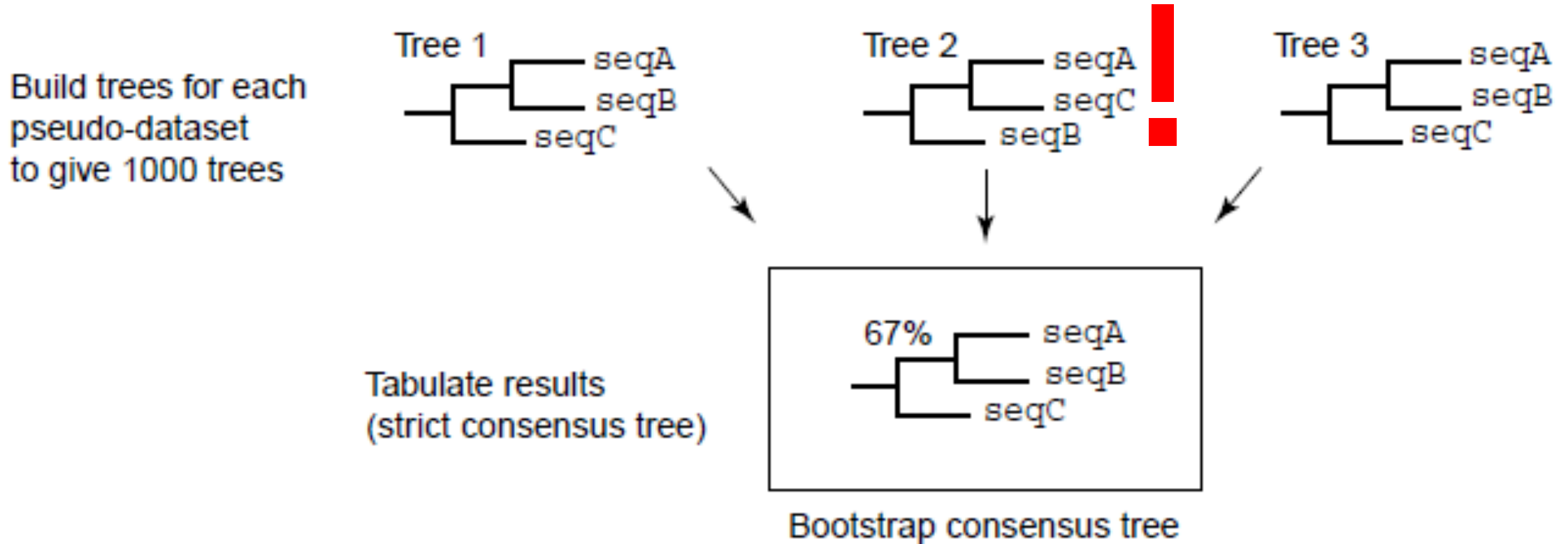
Unrooted

Where are potential sources of error?



Long branch attraction can cause erroneous groupings

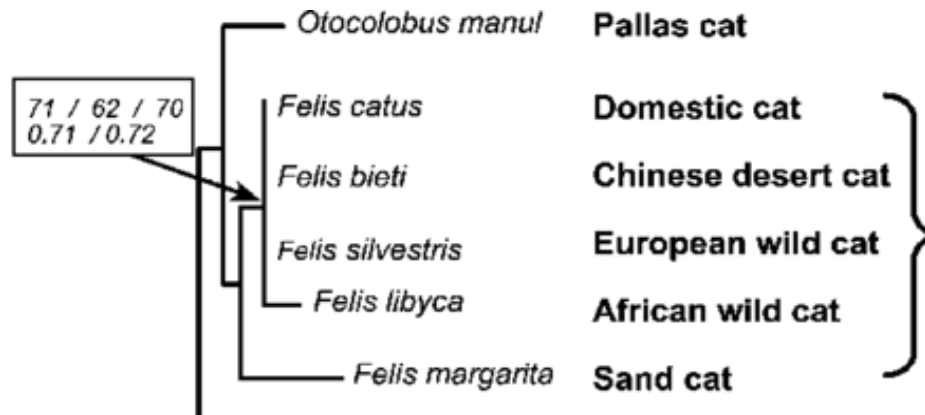
What is bootstrapping?



Check if Your Tree Statistically Acceptable

How information be included on the tree?

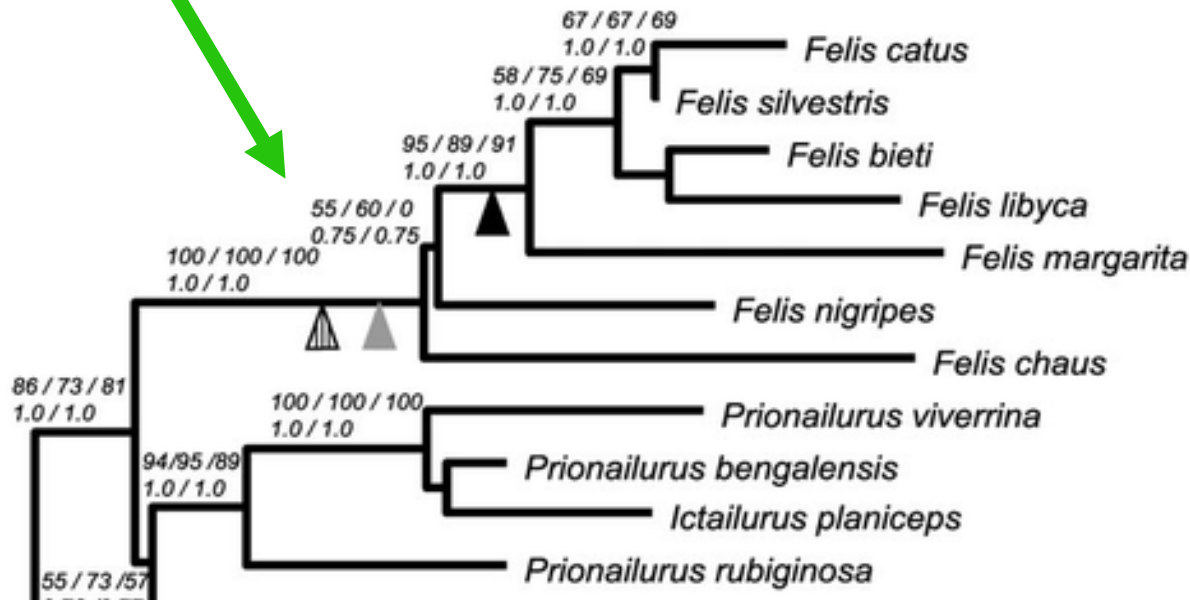
Include
Bootstrapping
Values > 50



UBE1Y SINE

SMCY SINE
SN327@1073

Names should
be Meaningful



Species

Lineage

Domestic cat
European wild cat
Chinese desert cat
African wild cat
Sand cat
Black-footed cat
Jungle cat

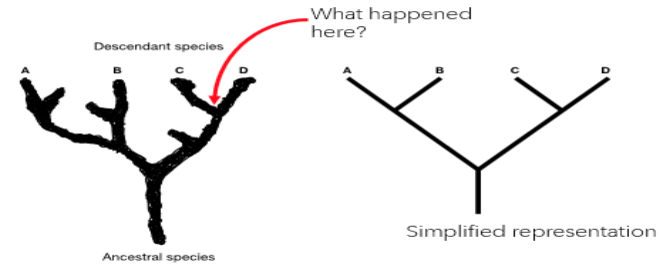
Domestic
Cat

Fishing cat
Asian leopard cat
Fishing cat
Rusty-spotted cat

Asian
Leopard
Cat

Summary

What is a Phylogeny?



A representation of evolutionary past

Phylogenies are representations of the evolutionary past of a group of organisms



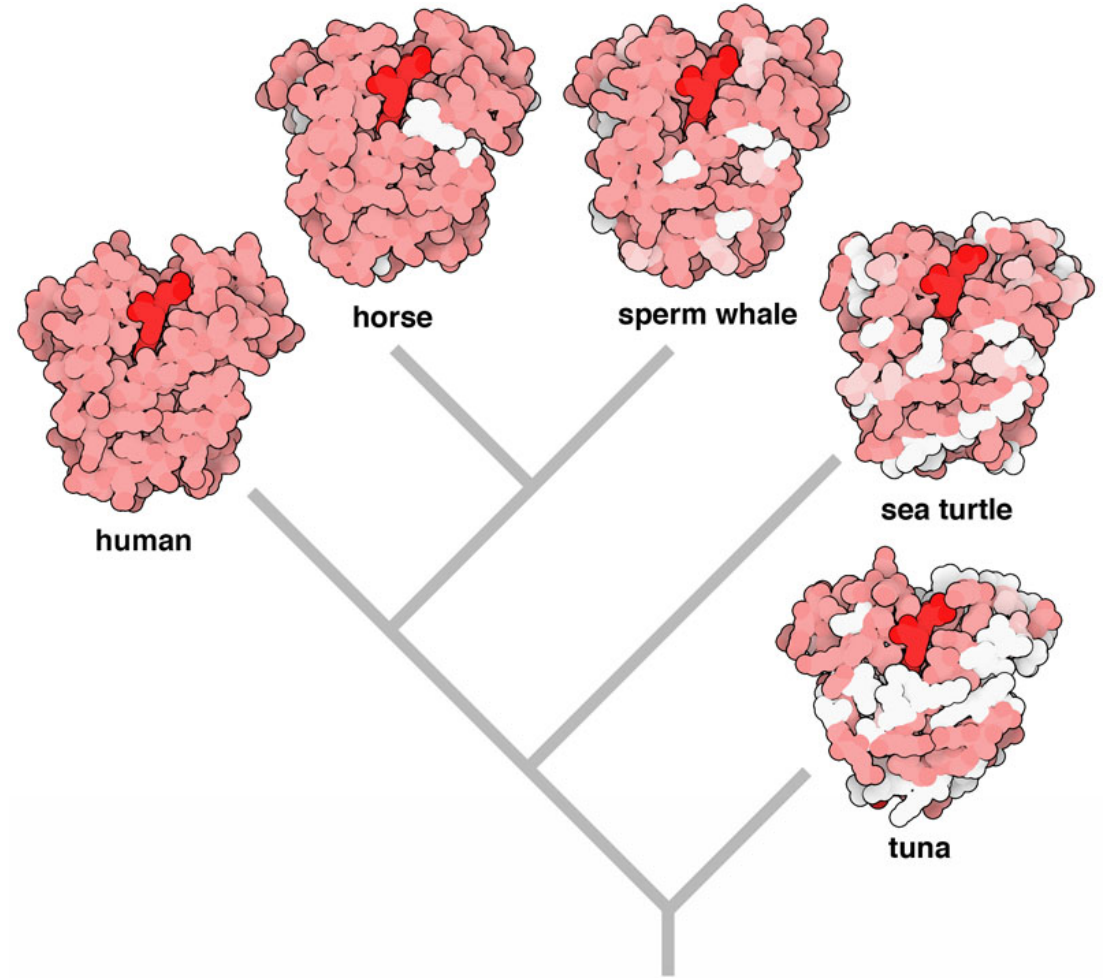
Several programs are available to help with building a tree

Why Should I Use Distance or Tree Searching?

Distance Matrix		Tree Searching	
+	Fast and Simplistic		Slow and Computing Intensive
-	Only Yields a Tree		Data Rich

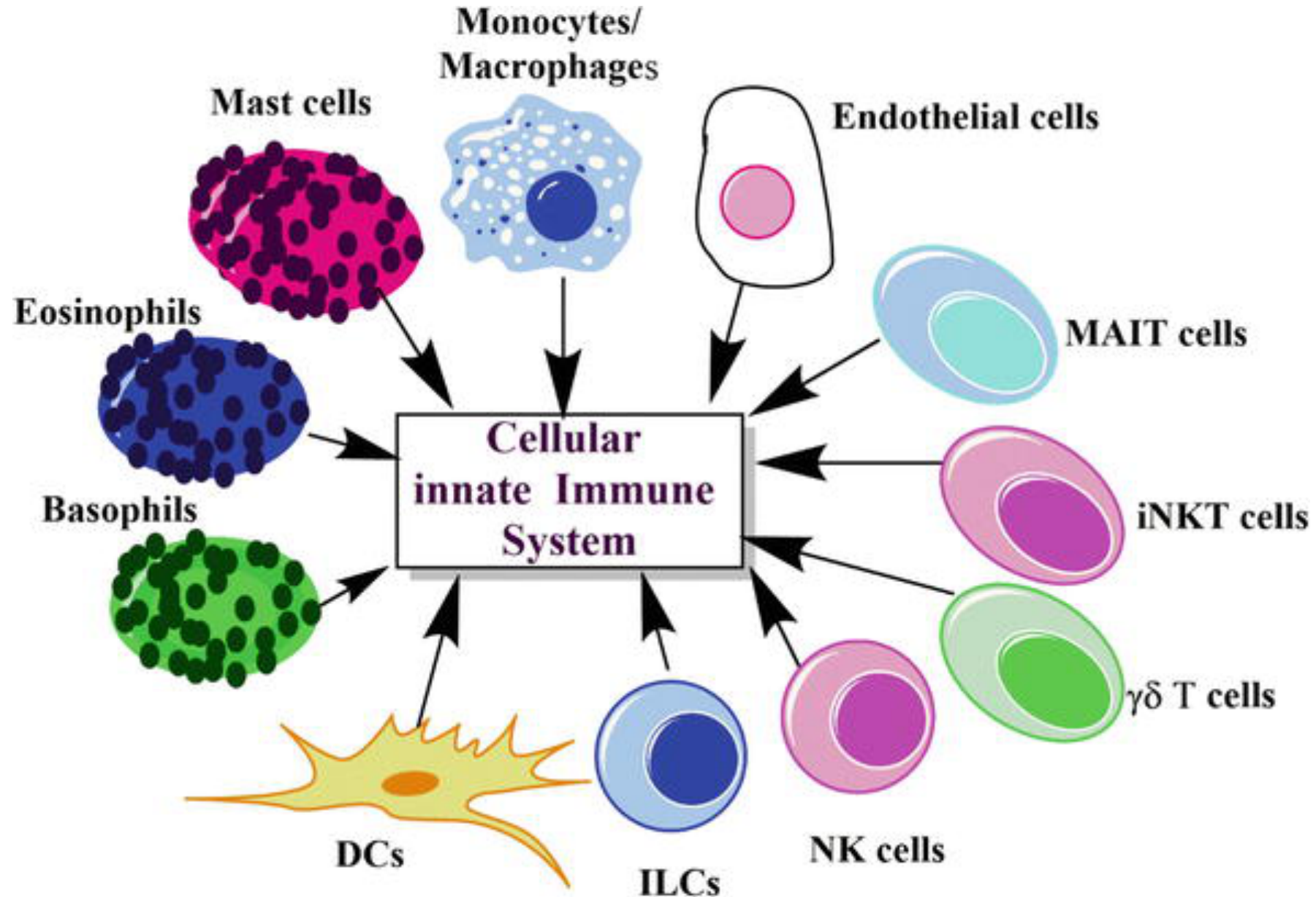
A variety of algorithms and programs contribute to creating statistically reliable, robust trees

What are some applications?



Phylogenetics can be used to understand how a protein or mechanism has evolved over time

Can phylogenomics be used in immunity research?



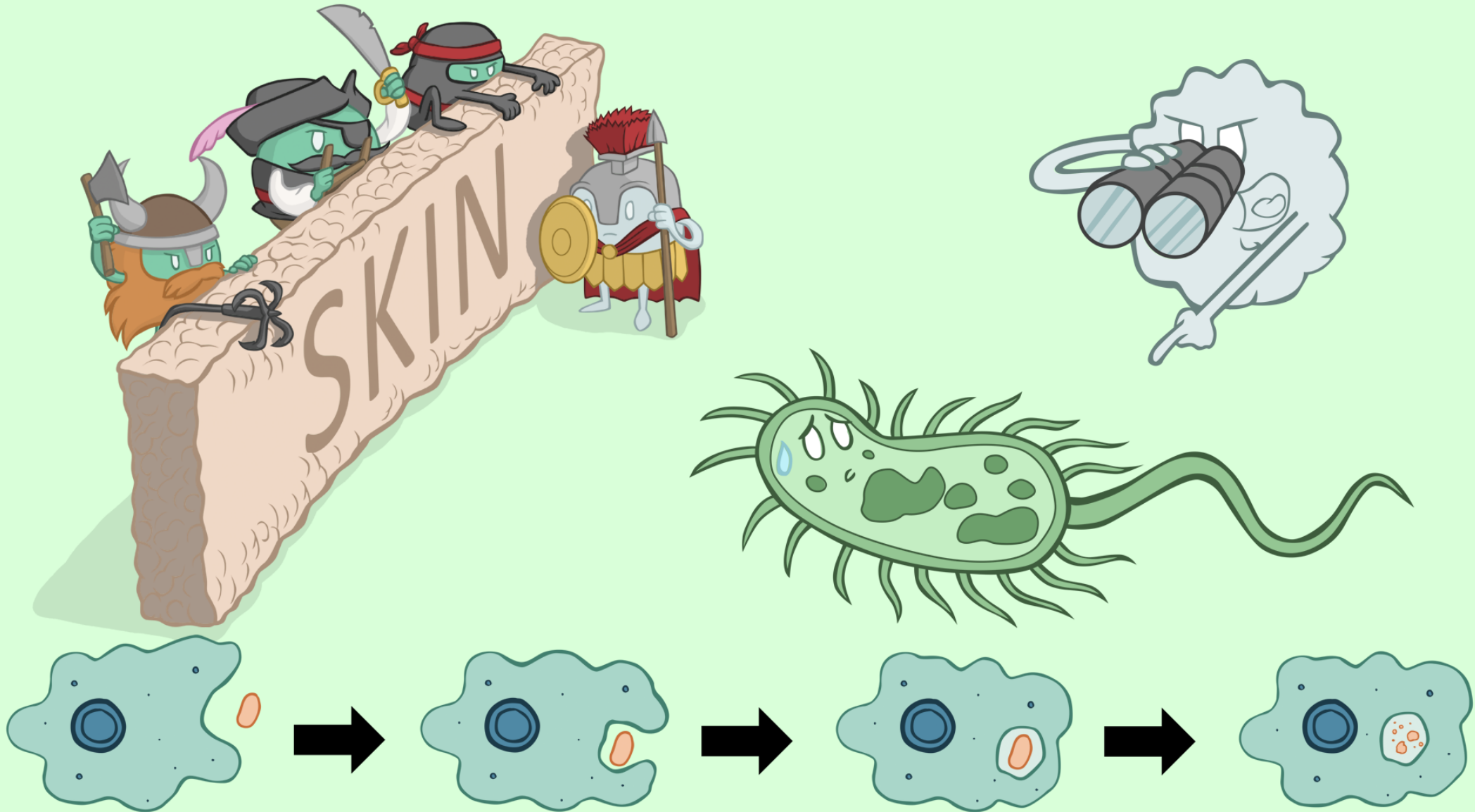
Special Issue: Innate Sensing Across Kingdoms

Review

Evolutionary Origins of cGAS-STING Signaling

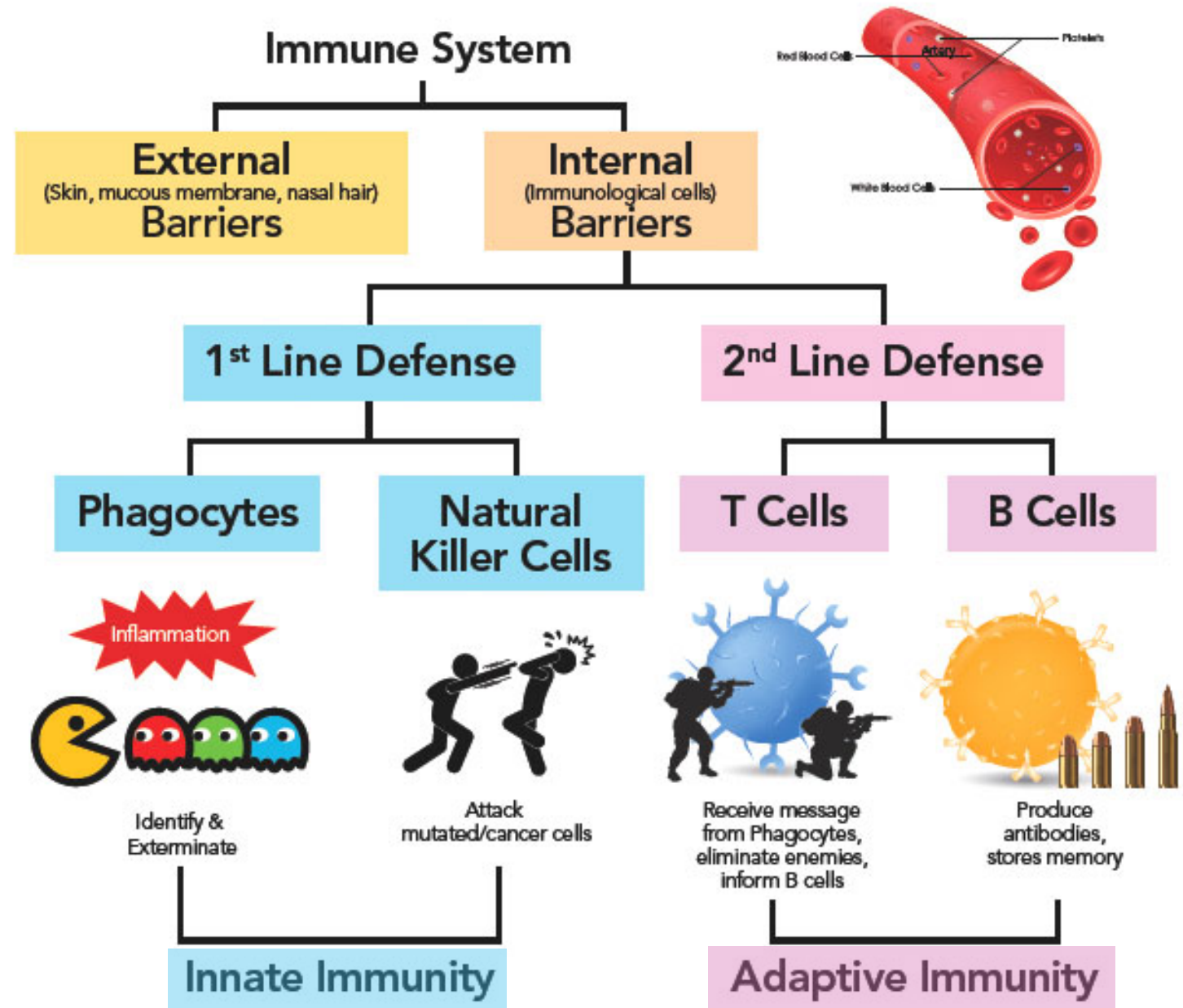
Shally R. Margolis,¹ Stephen C. Wilson,^{1,3} and
Russell E. Vance^{1,2,*}

What is the innate immune response?

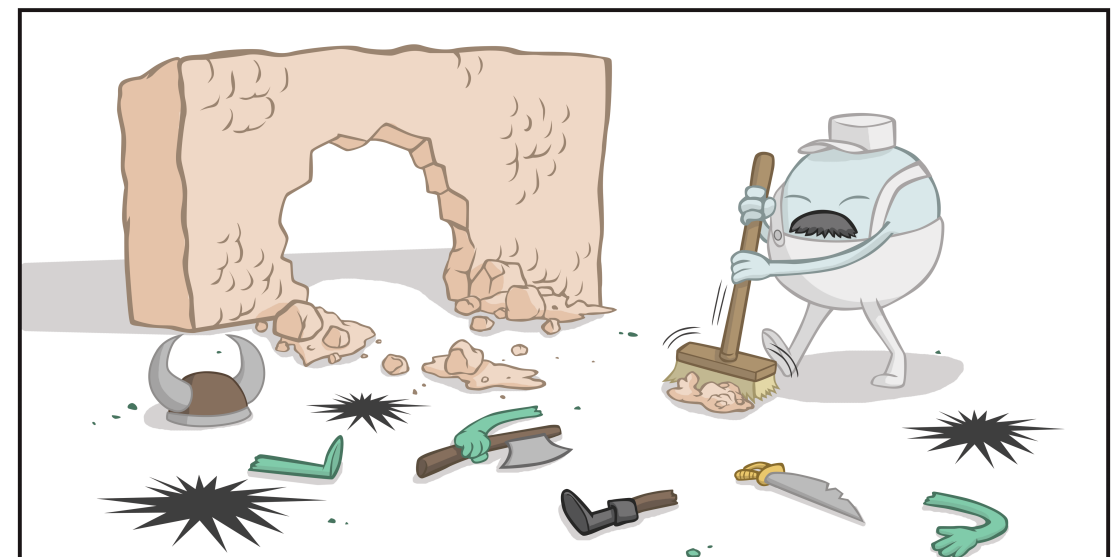
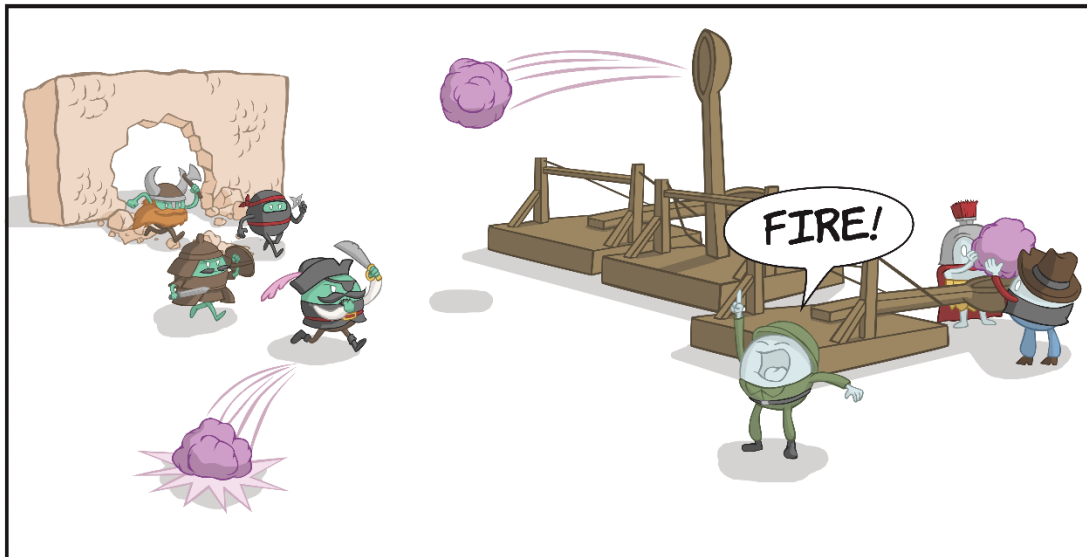
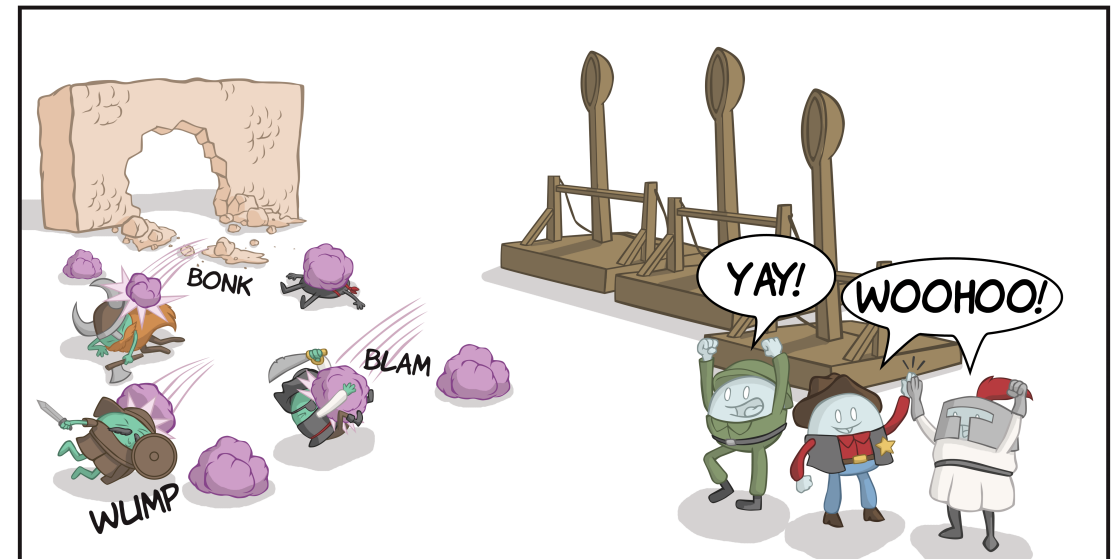
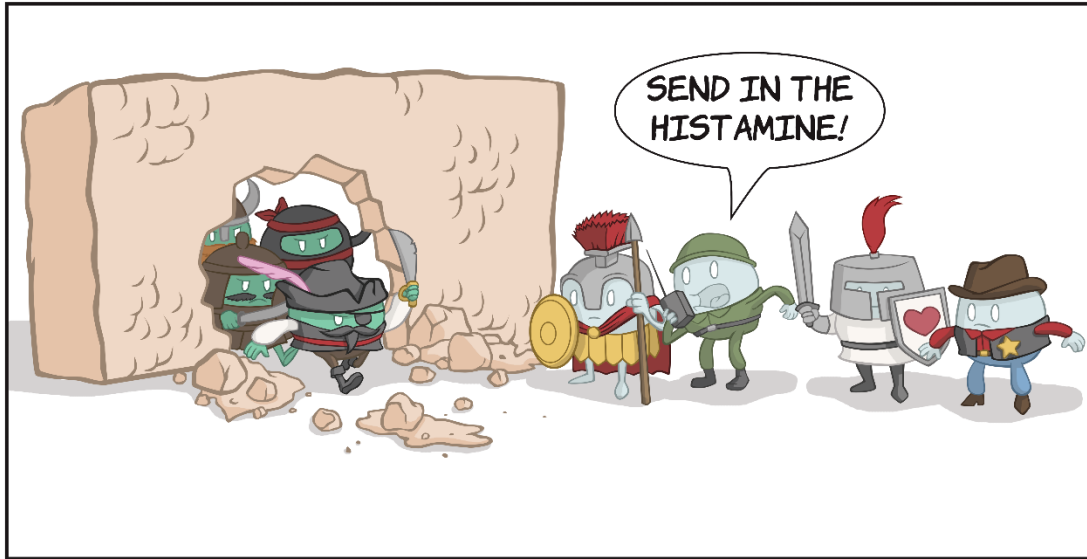


What is the difference between **innate immunity** and **adaptive immunity**?

Rapid and non-specific
VS.
slow and targeted



Why is the innate immune response important?

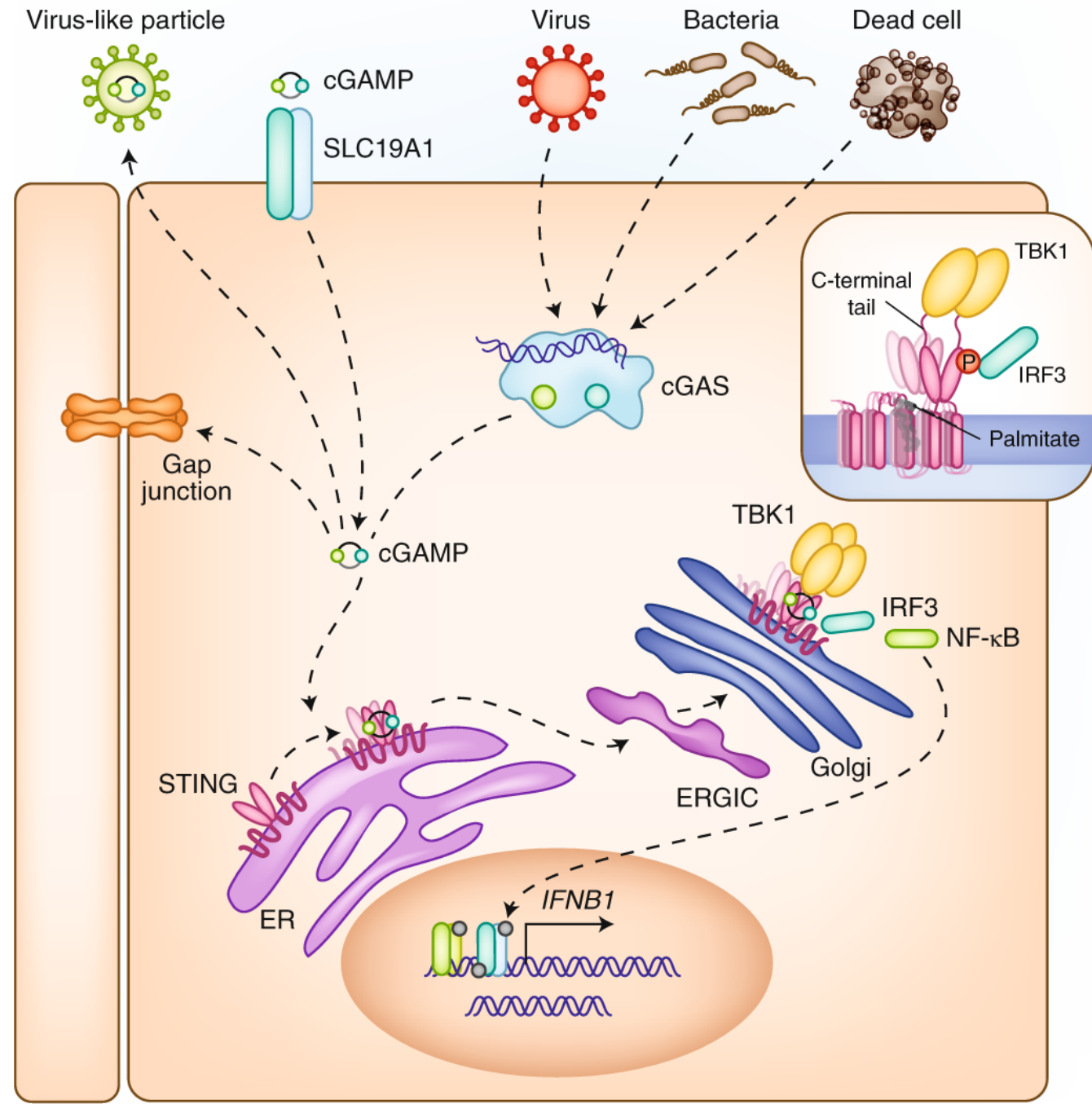


What is **cGAS** and what is its role in innate immunity?

Senses **DNA**

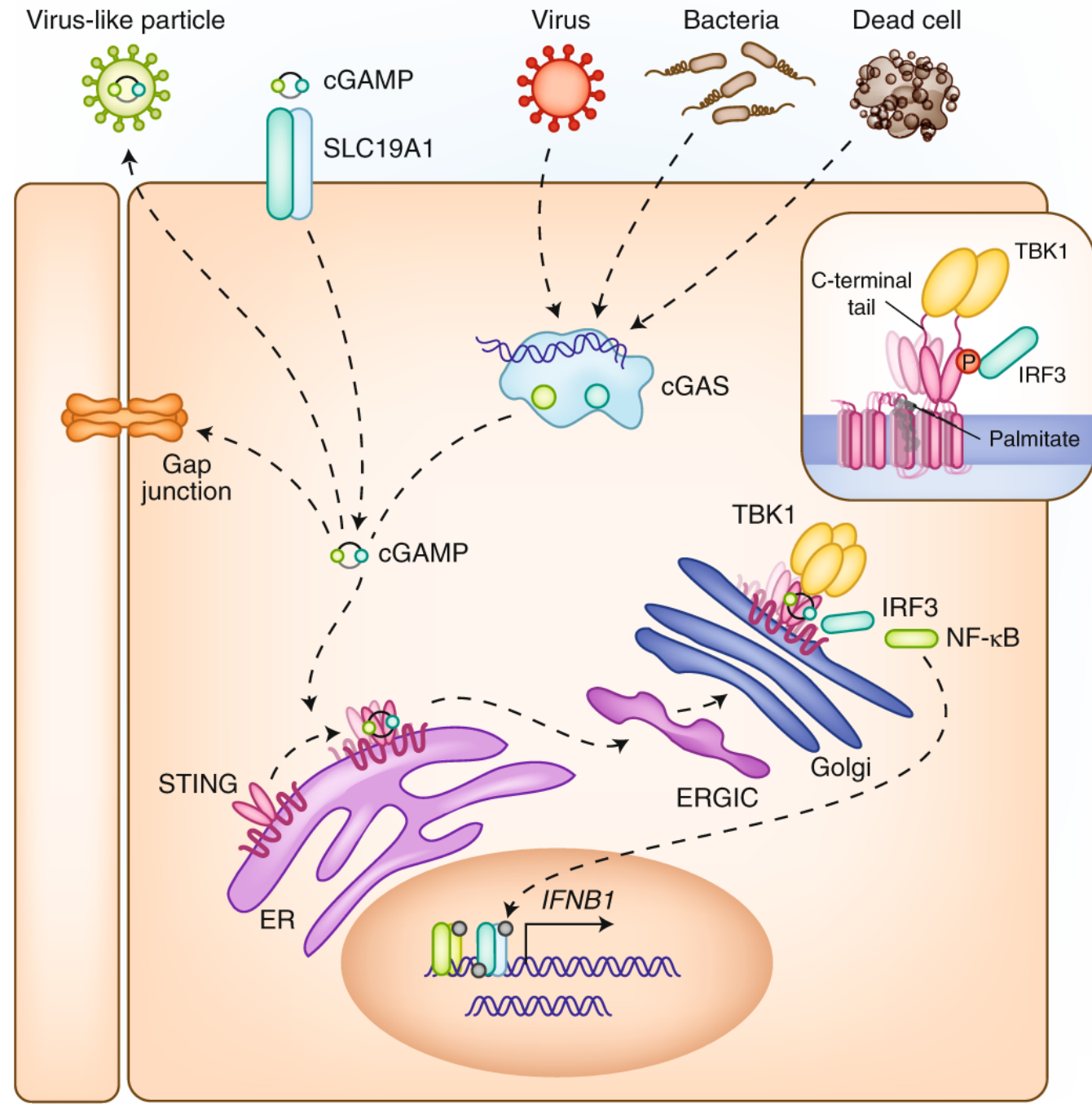
Produces cyclic dinucleotides, like **cGAMP**

Activates **STING**

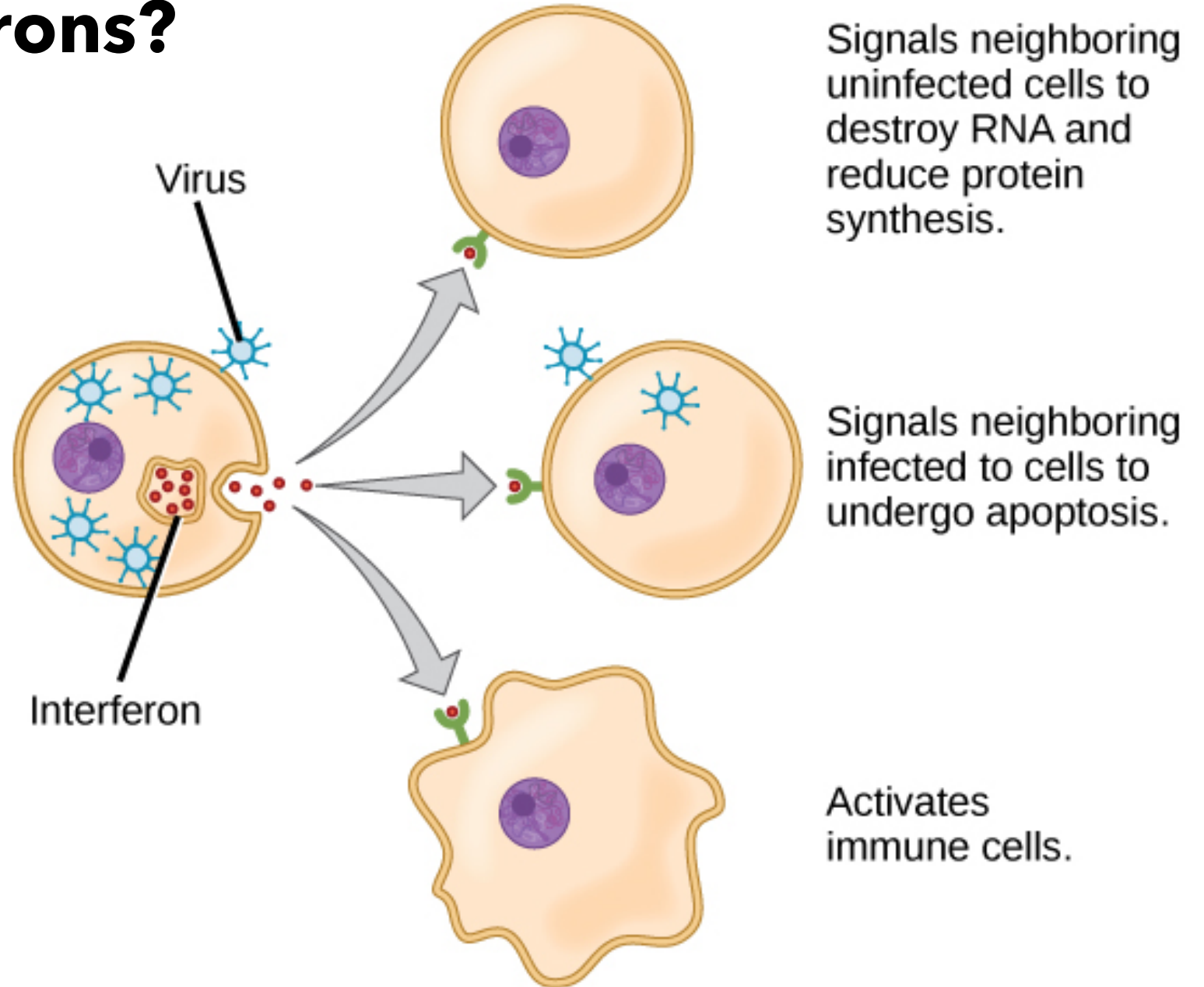


What is **STING**?

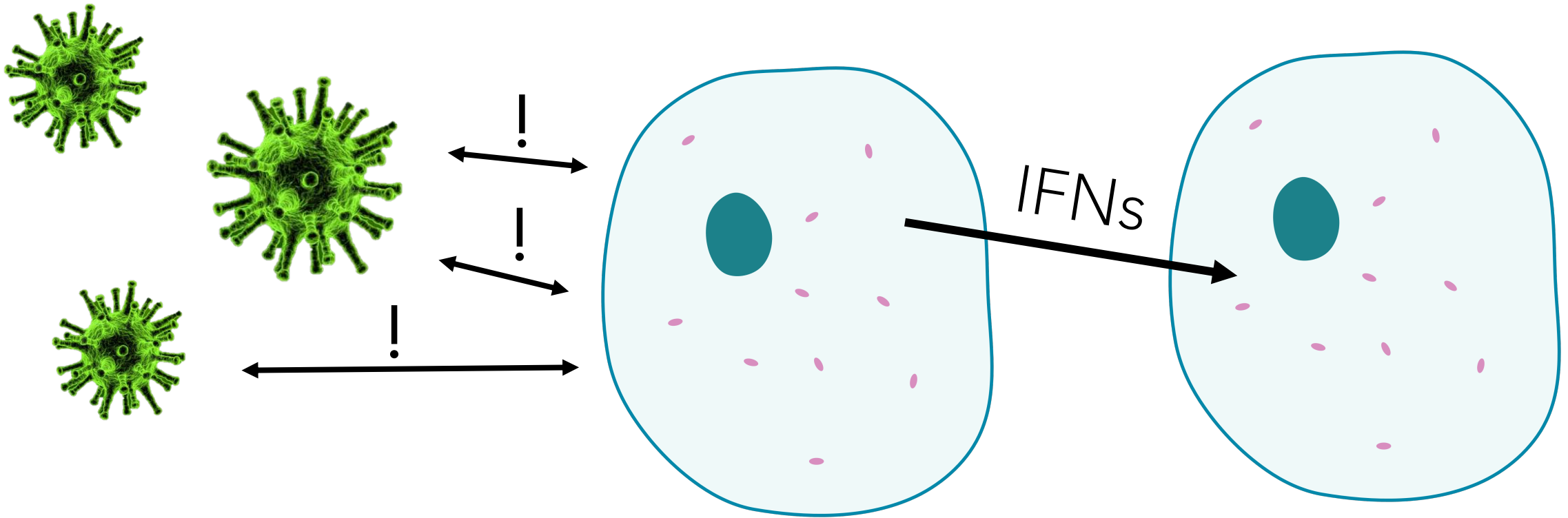
Activates transcription of
interferons



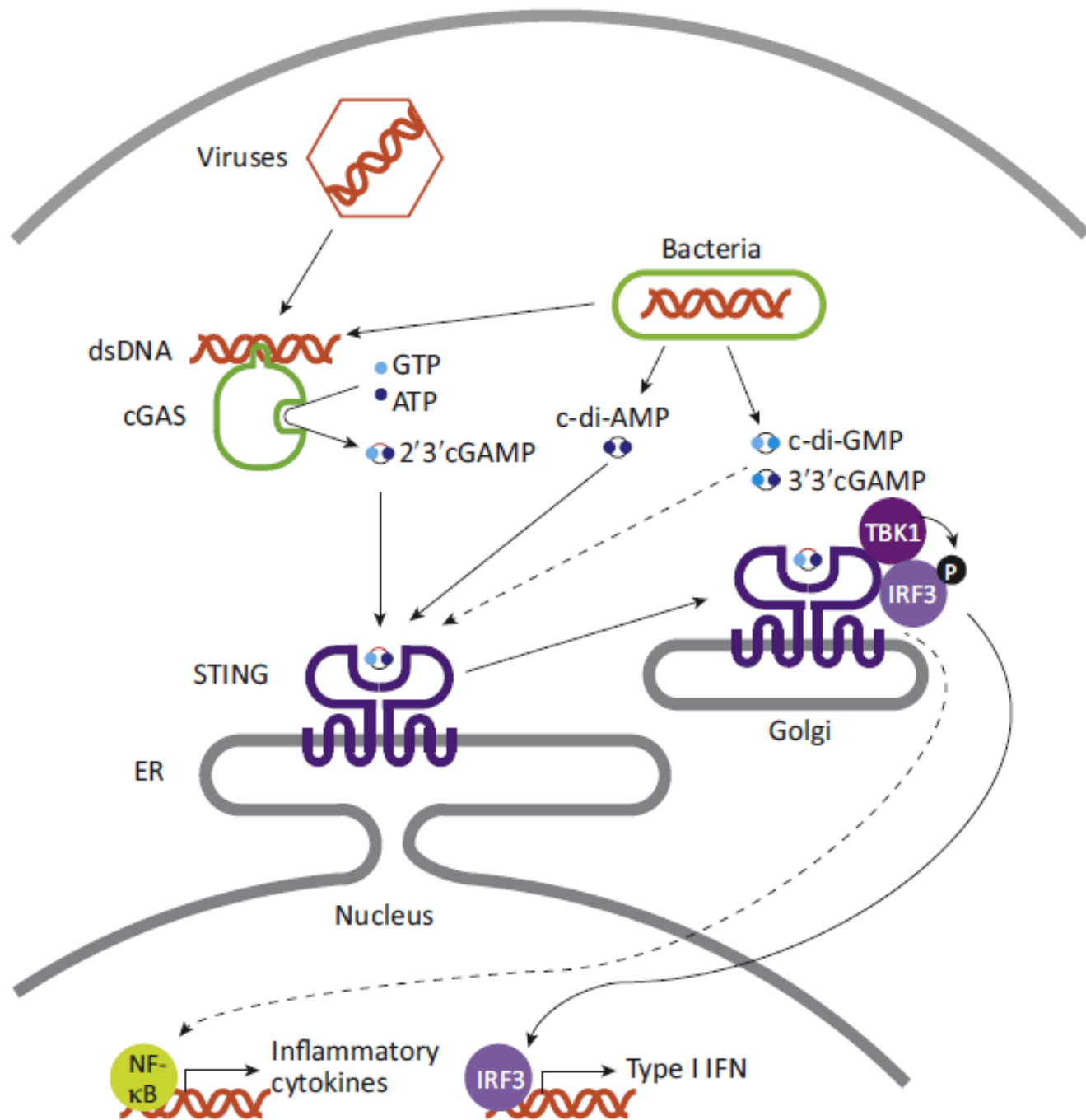
What are interferons?



What are type I interferons?

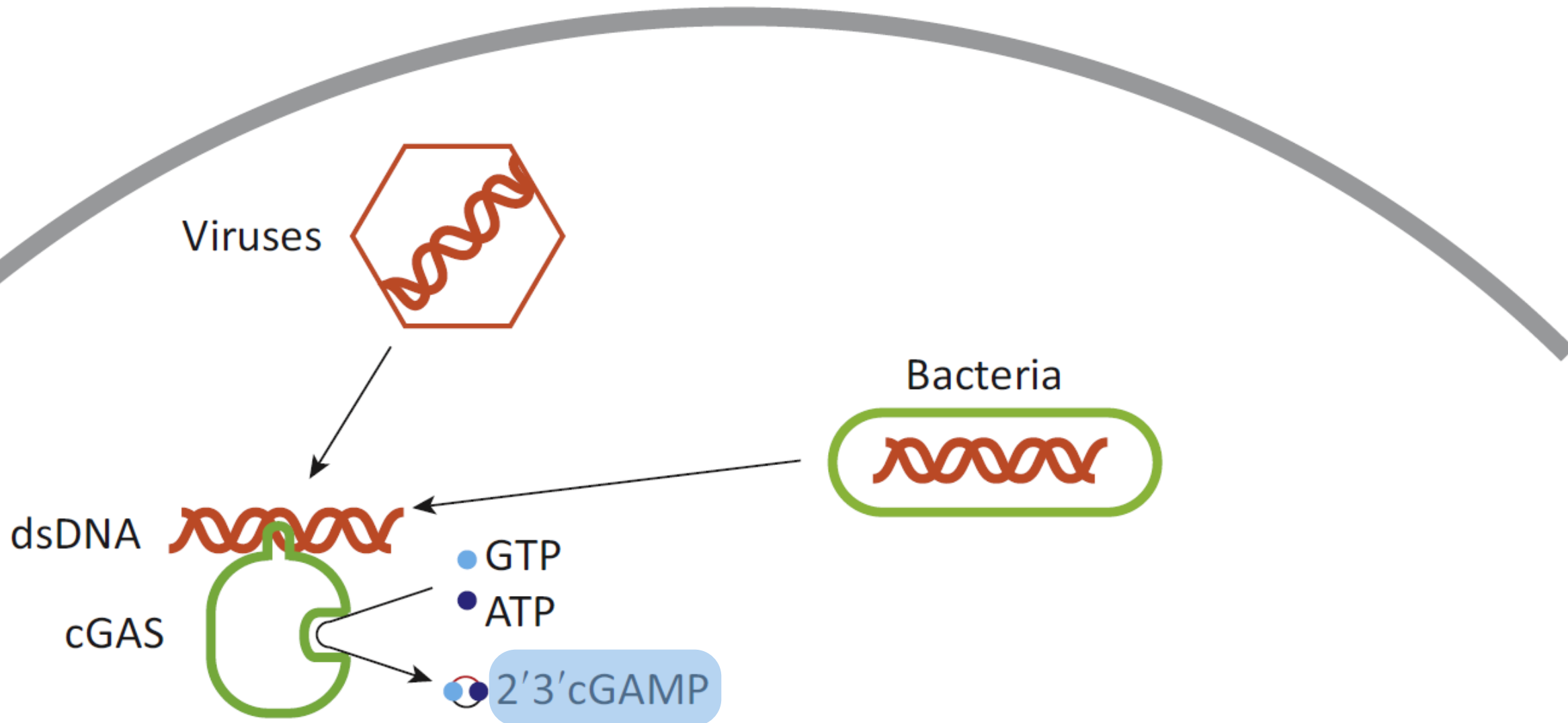


Antiviral cytokines triggered in humans by cGAS-STING response pathway

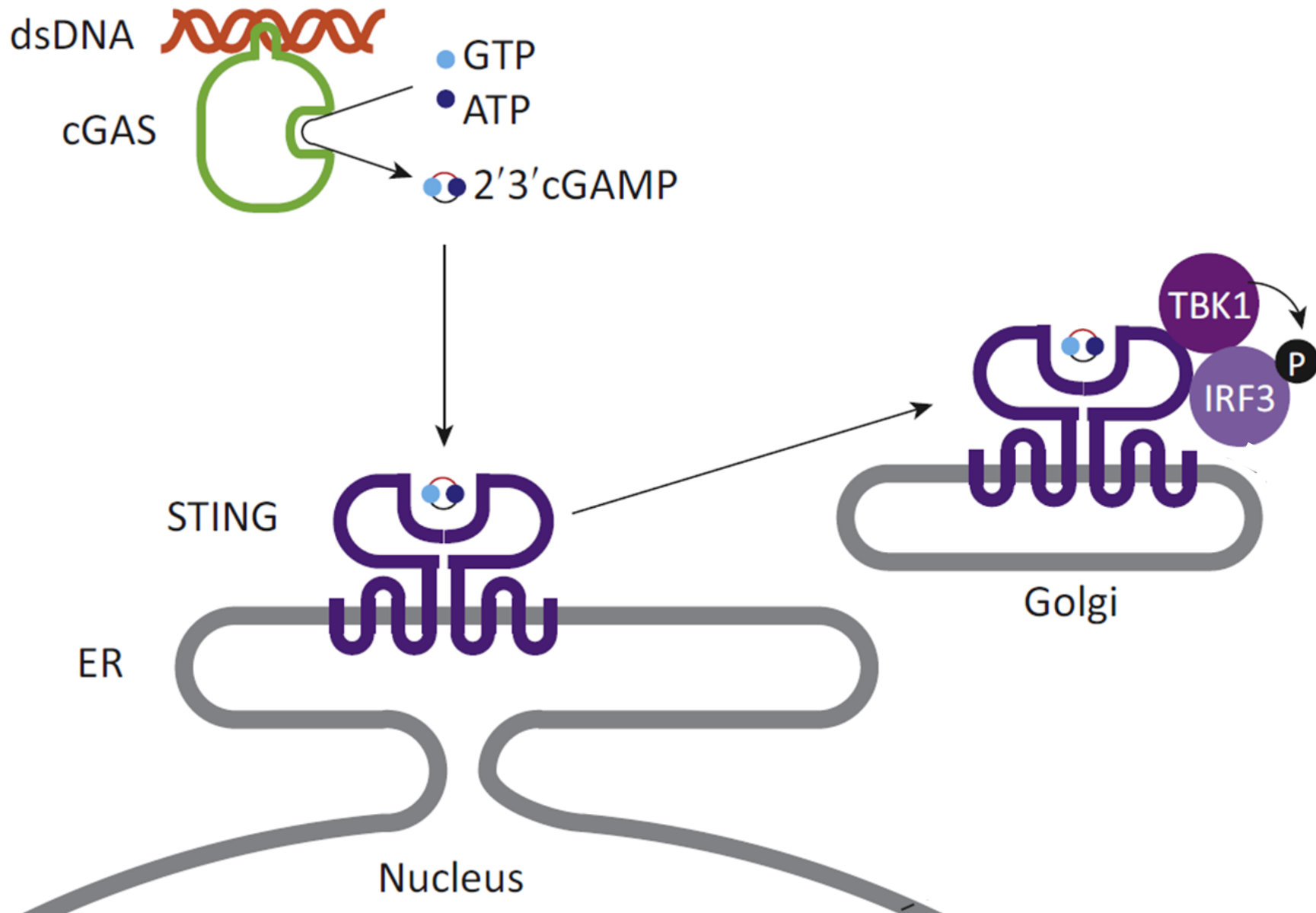


**Fig 1: How does the
cGAS-STING
pathway work?**

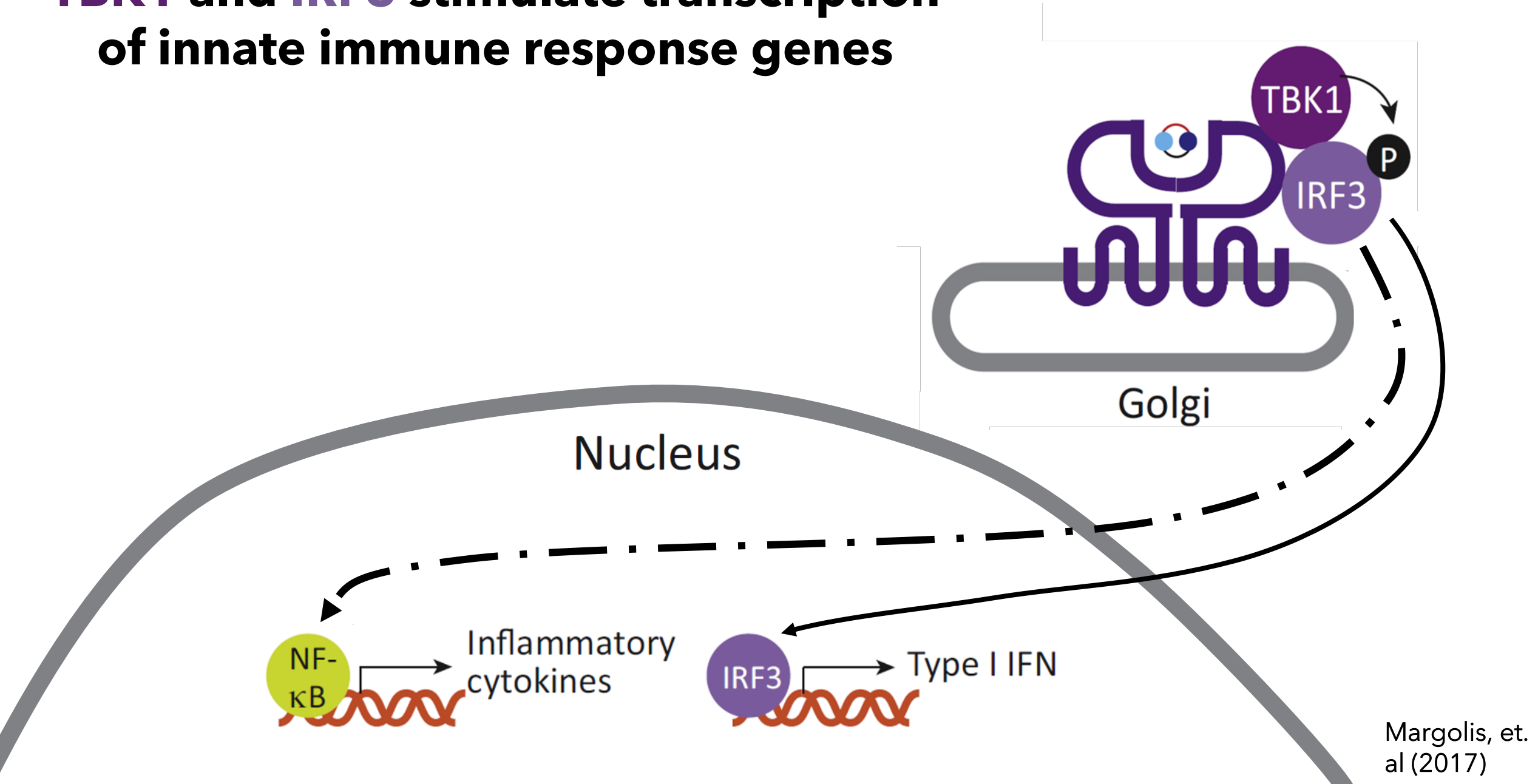
How does viral and bacterial dsDNA detection lead to 2'-3' cGAMP synthesis?

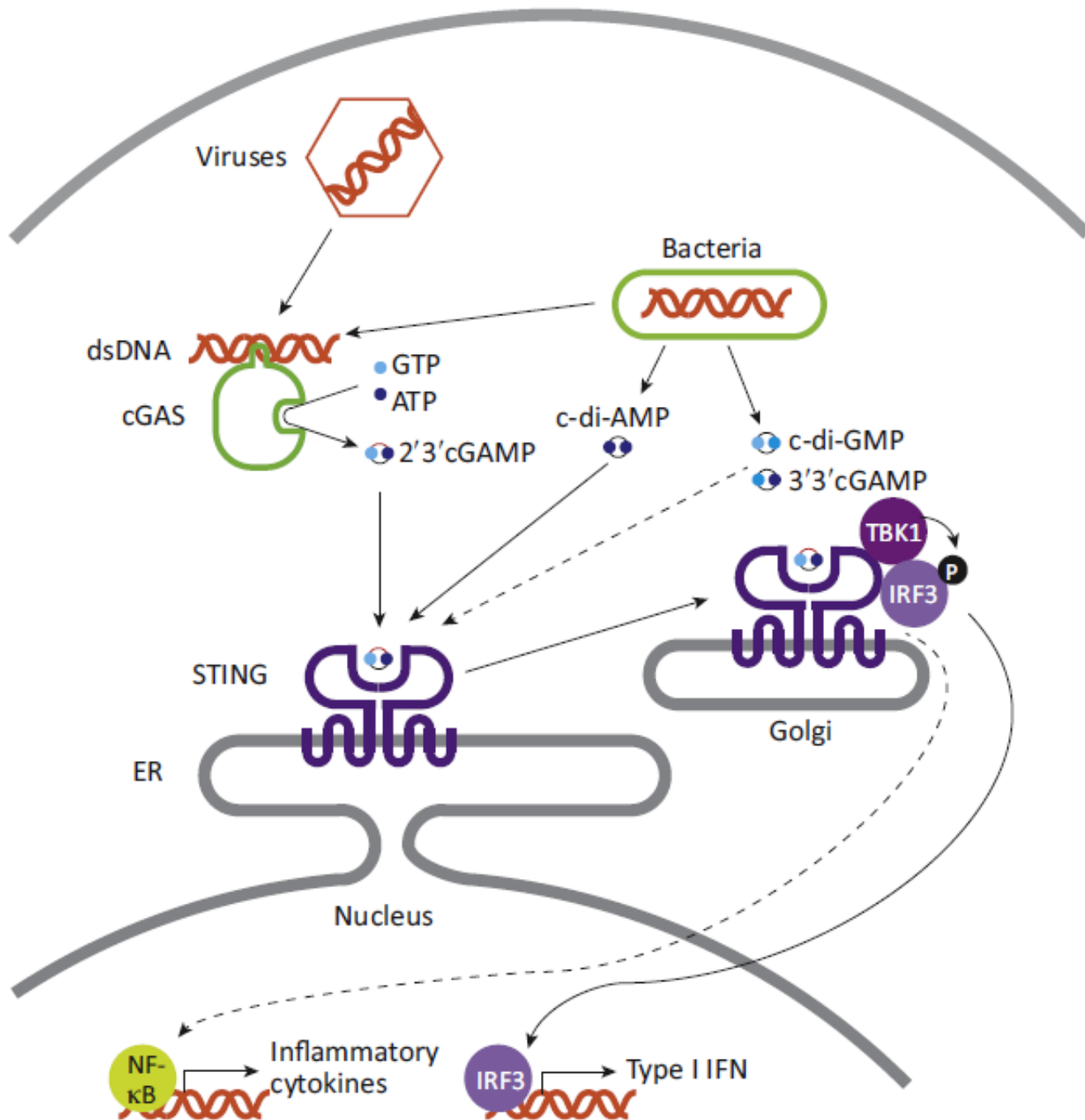


What happens after cGAMP binds to activate STING?



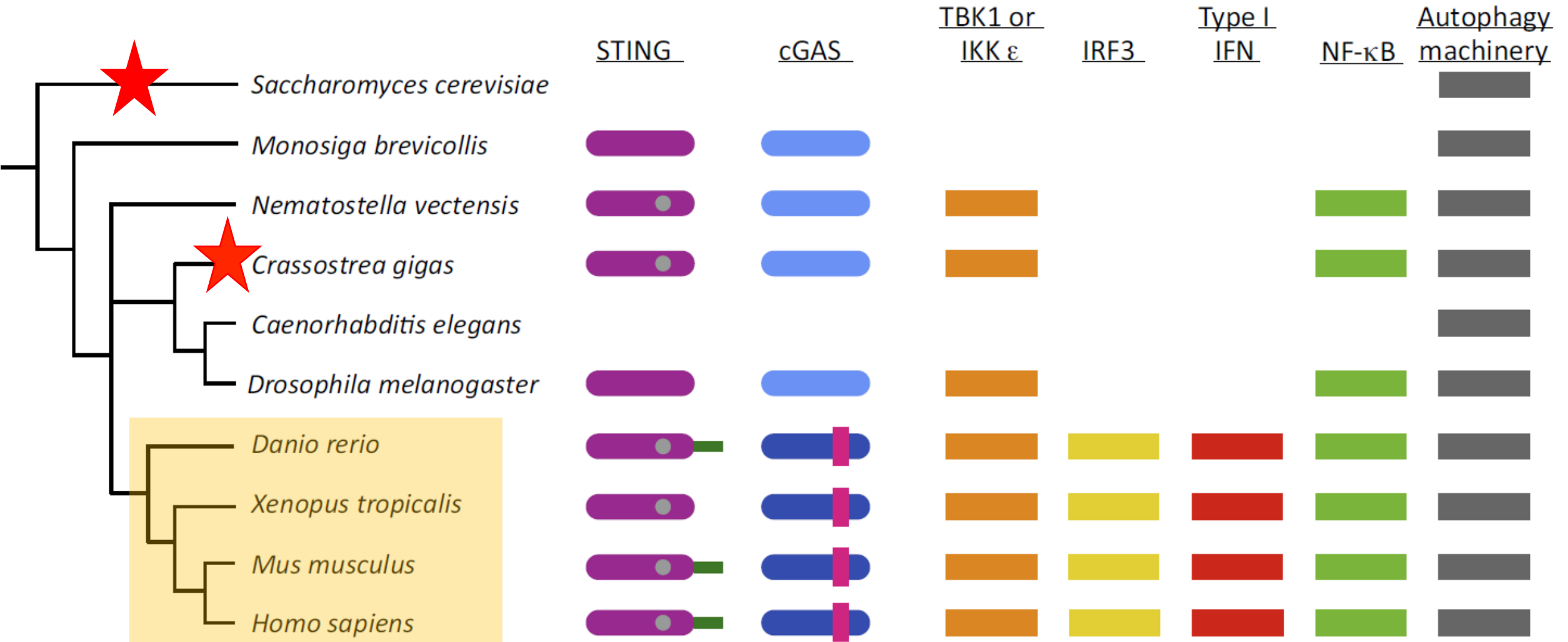
TBK1 and IRF3 stimulate transcription of innate immune response genes



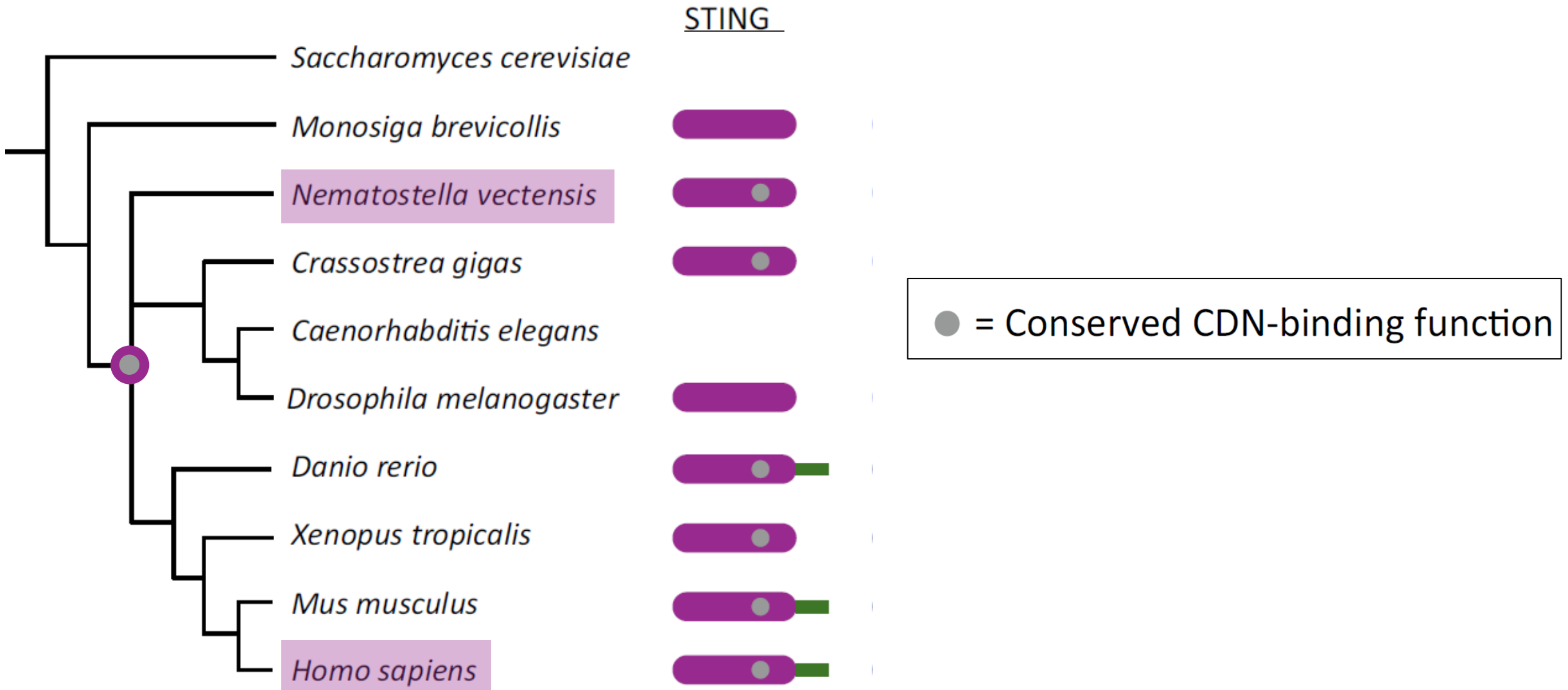


**How was phylogenetics
used to investigate the
evolution of this pathway?**

How have the **STING**-**cGAS** genes evolved?



How long ago did the **STING** CDN-binding function evolve?

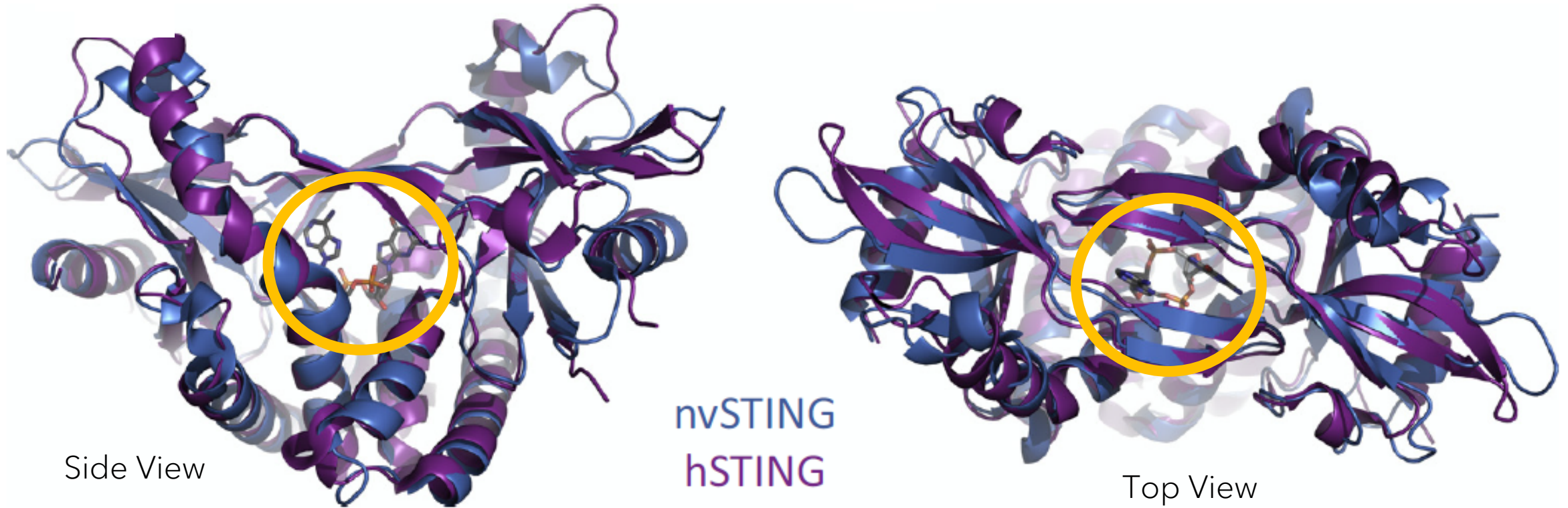


What is a Nematostella?



Most recent
common
ancestor
600 million
years ago

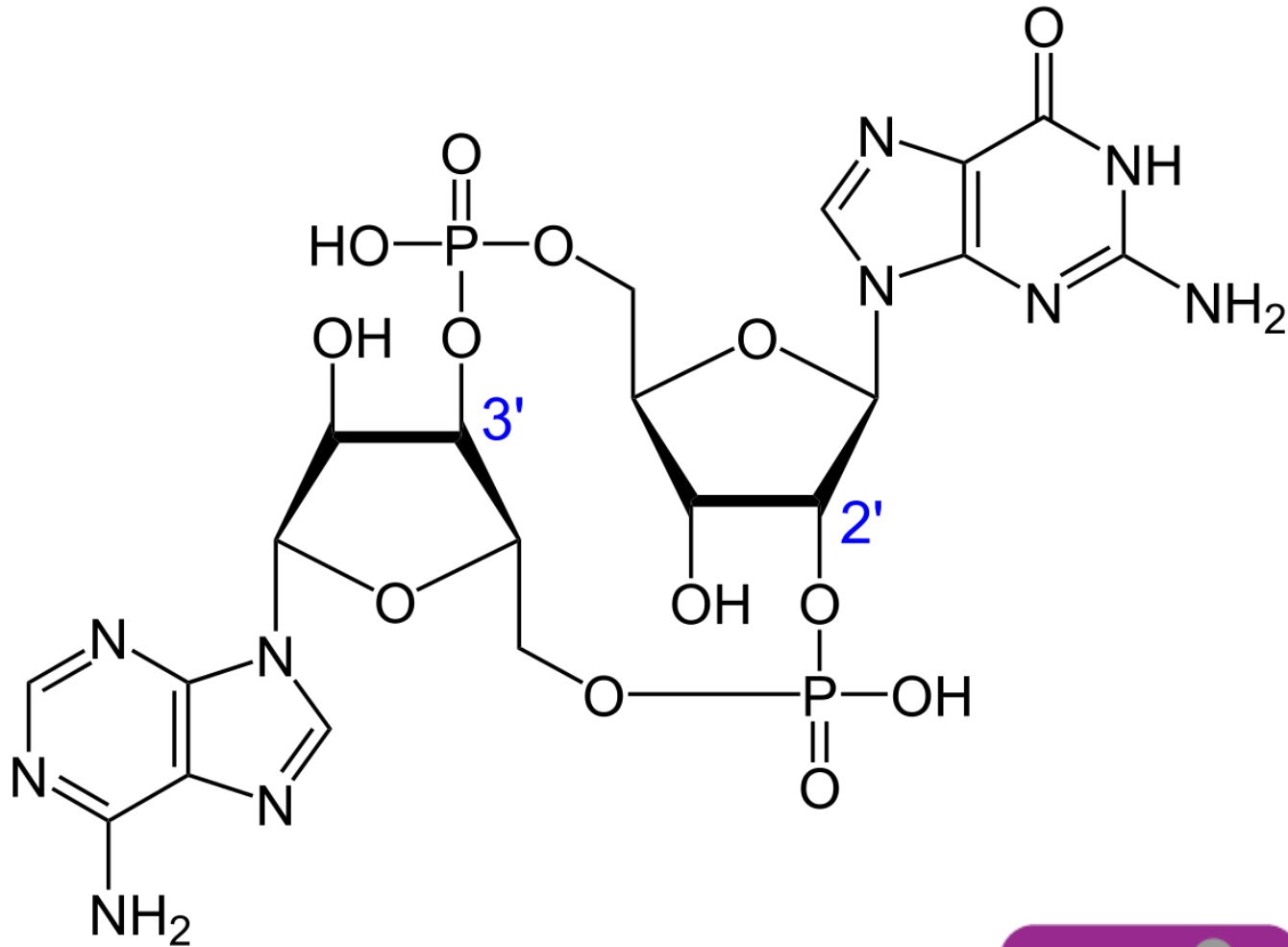
How do the CDN binding structures compare?



2'-3'cGAMP



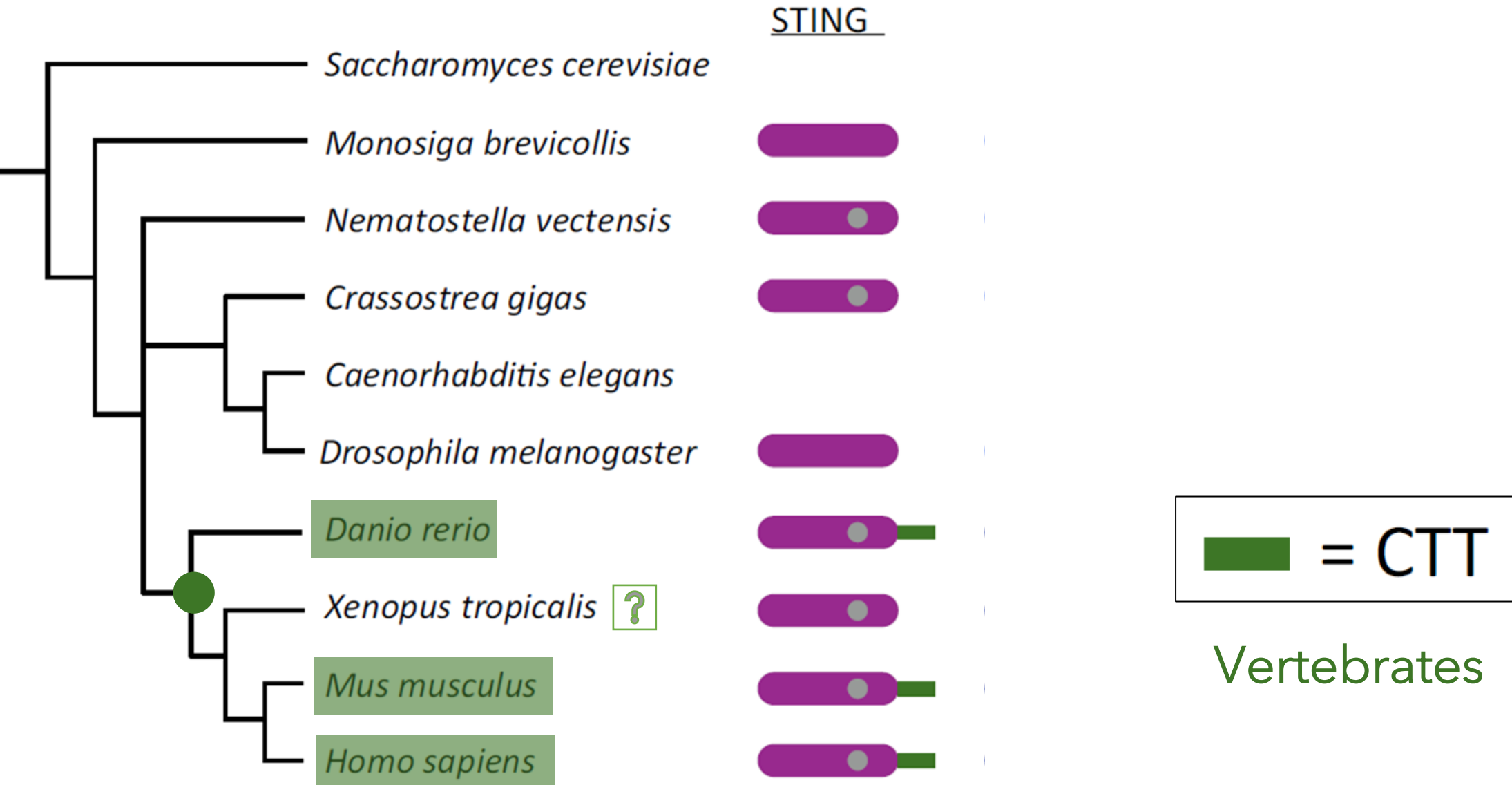
Why is the CDN binding similarity significant?



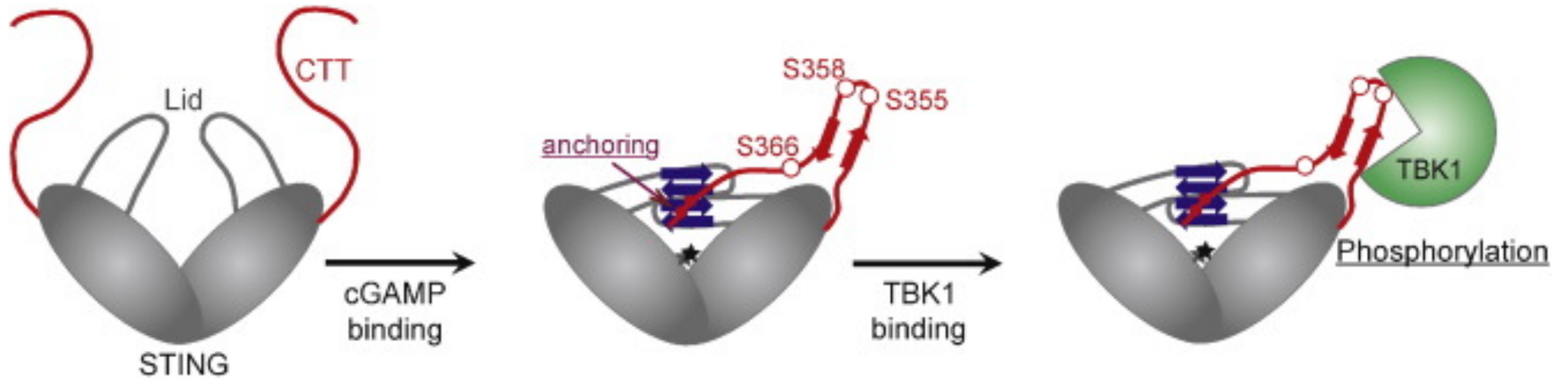
Preference for and
recognition of 2'-3'
cGAMP is not a recent
evolutionary innovation



How long ago did the **STING** **CTT** tail evolve?



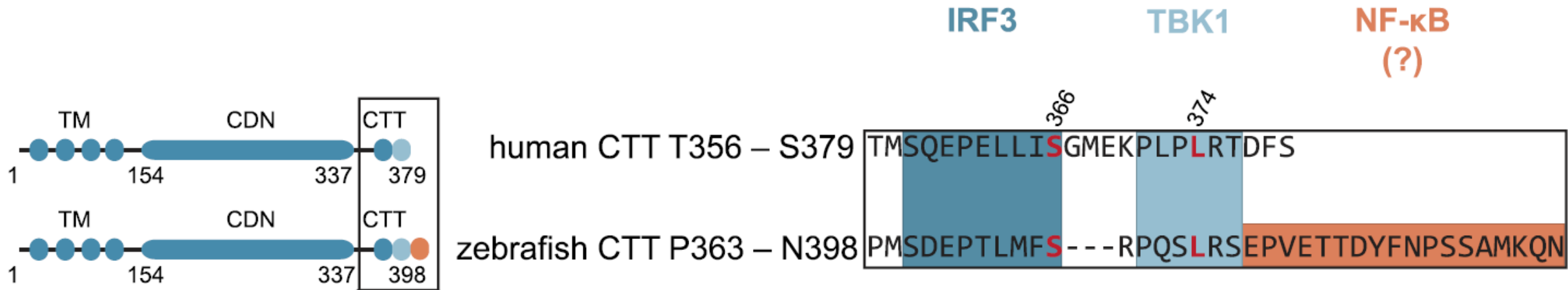
Why is the CTT tail important?



Aids in TBK1 recruitment and IRF3 phosphorylation when STING is activated



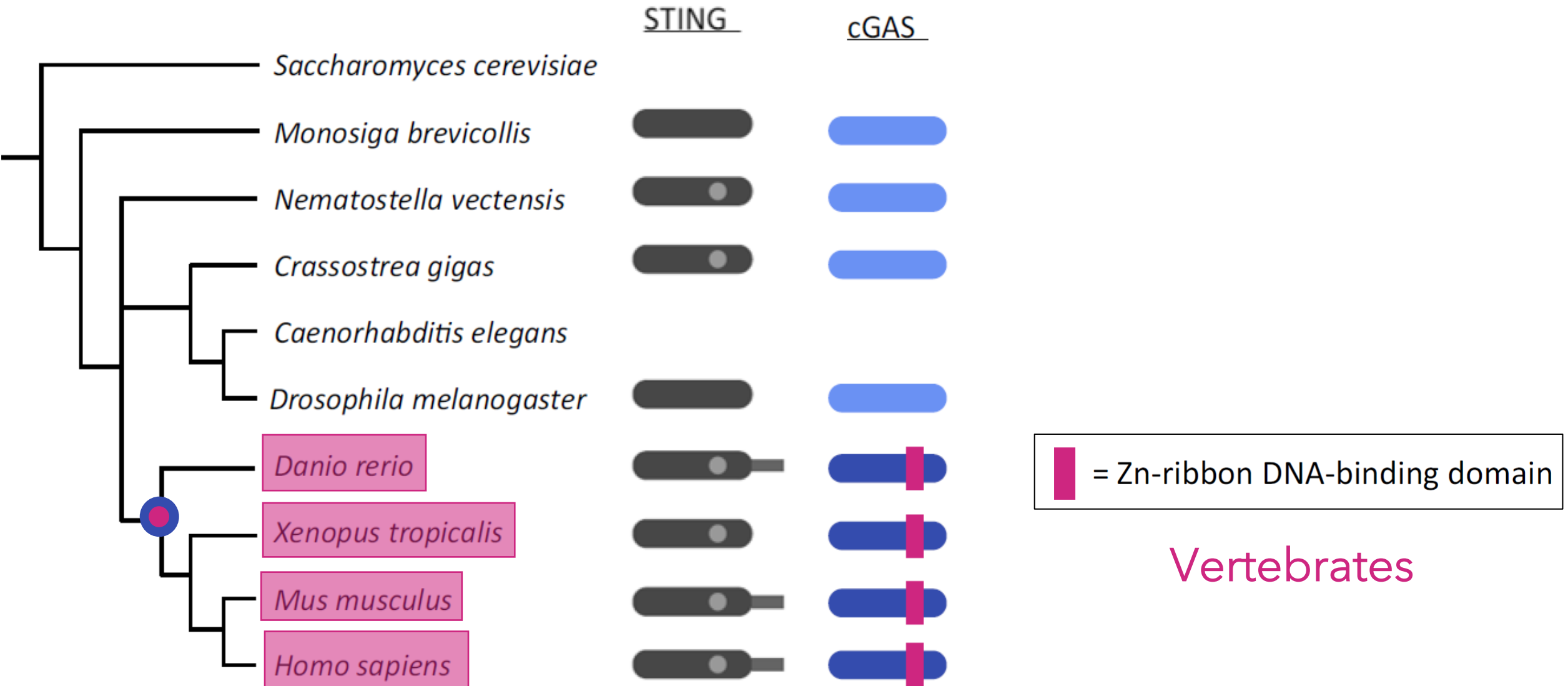
Why do only vertebrates have the CTT tail?



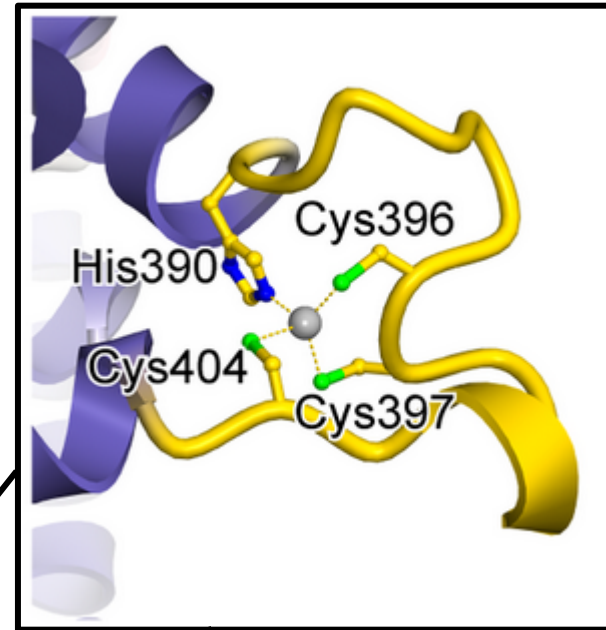
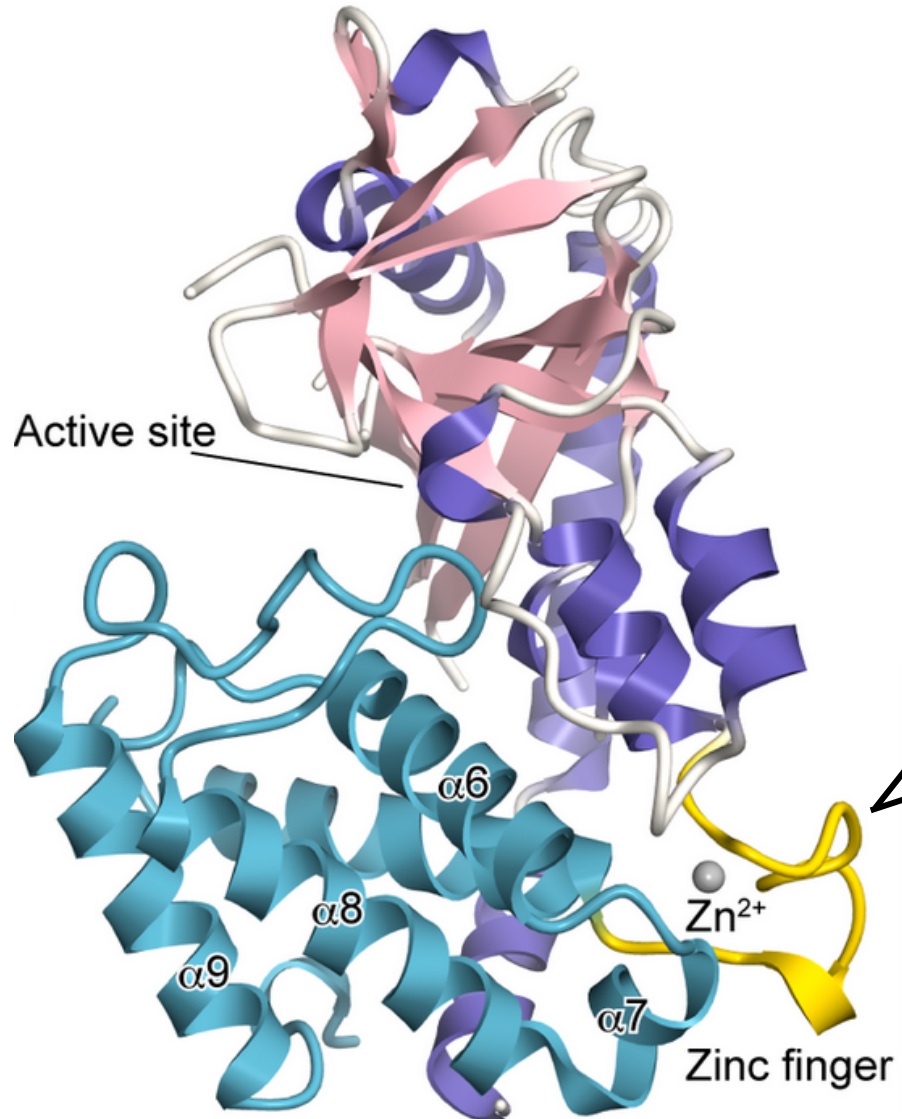
Some downstream functions of STING may not require the CTT tail and the tail has evolved through modularity



How long ago did the **cGAS** **Zn-ribbon** evolve?



Why is the **cGAS Zn ribbon** important?

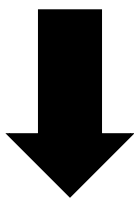


Required for cytosolic DNA binding and cGAMP synthesis

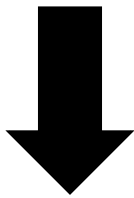


What is a possible mechanism for the evolution of the cGAS-STING pathway?

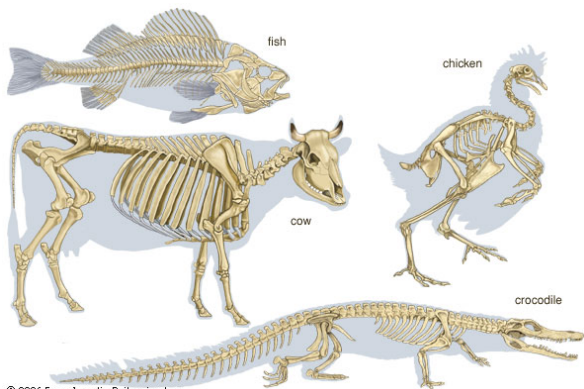
Autophagy + NF- κ B



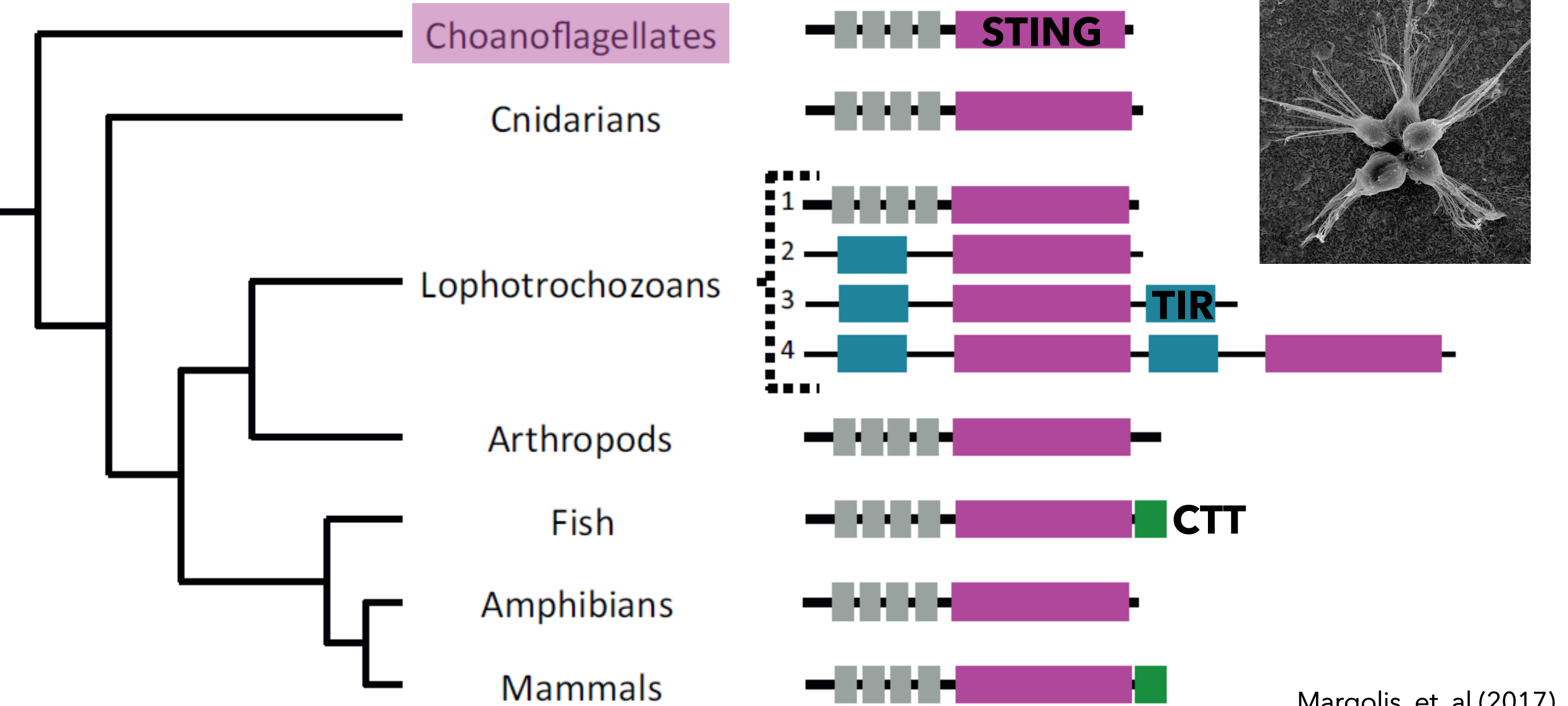
STING CDN-Binding Domain



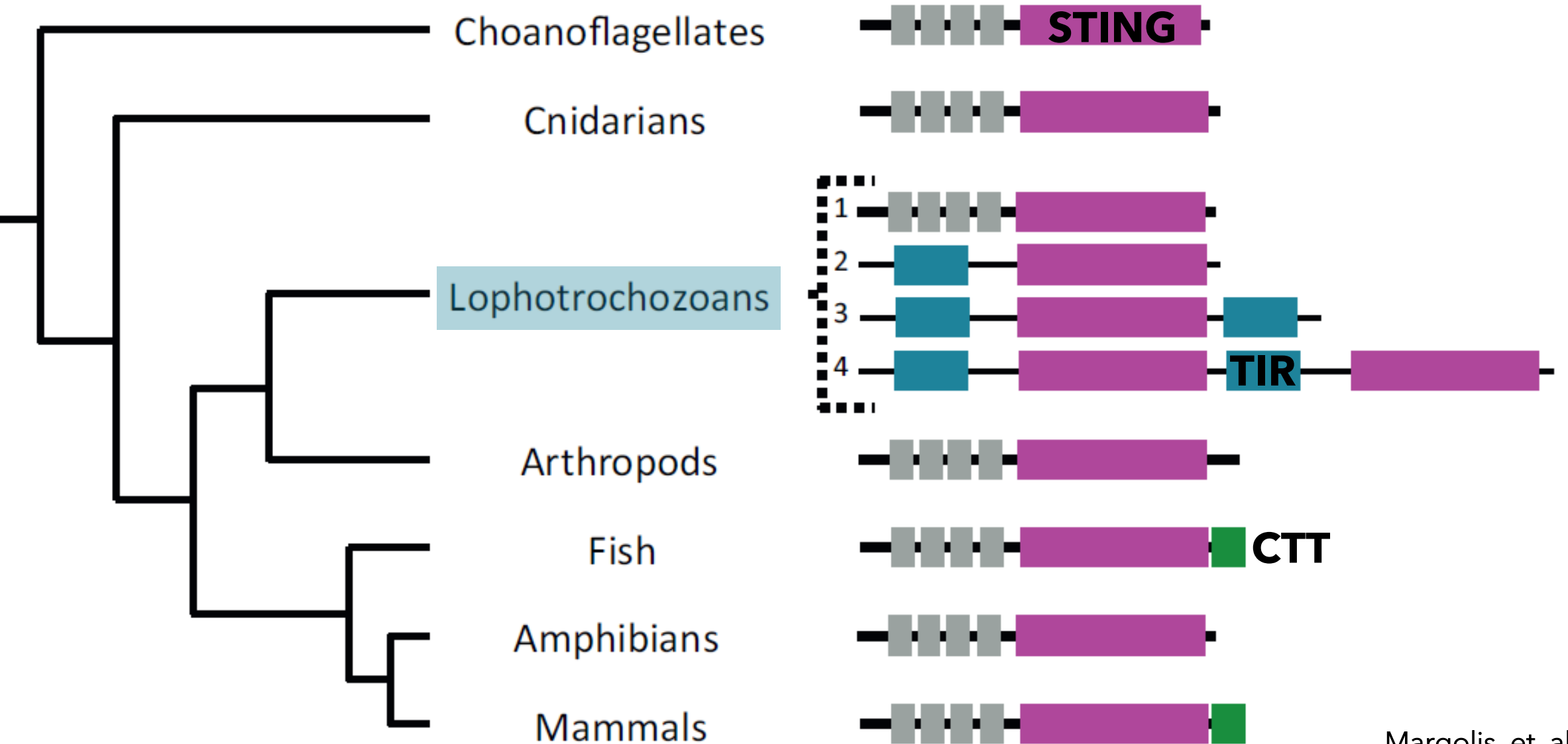
STING CTT + cGAS Zn-Ribbon +
Type I IFNs



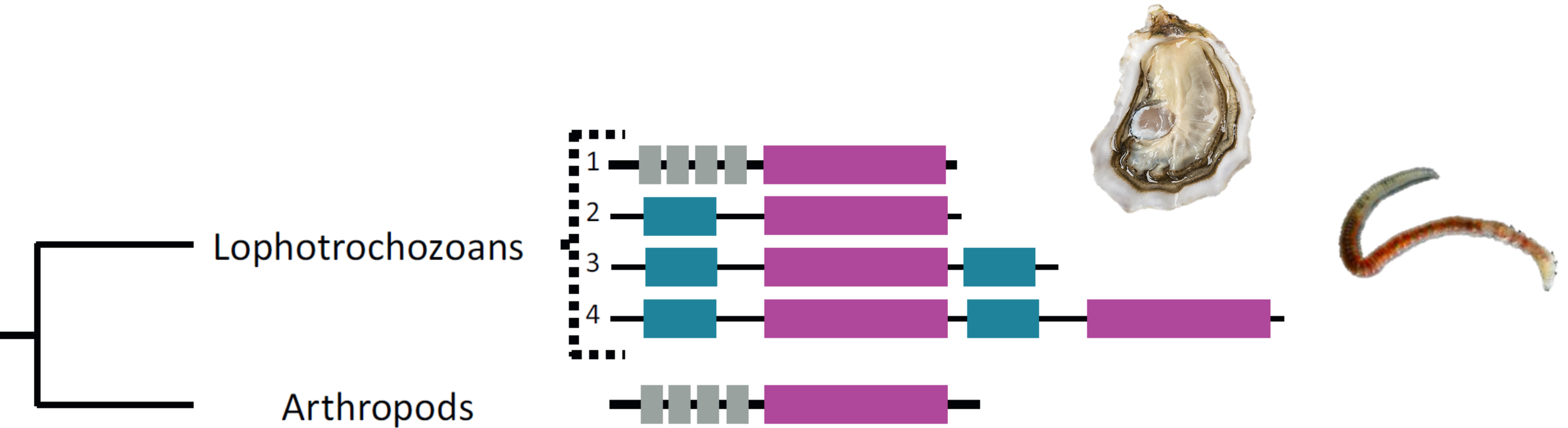
Is the cGAS-STING pathway ancient?



How divergent are **STING** domains?



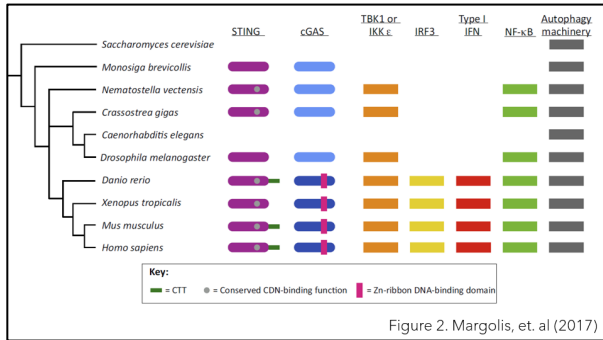
Why are **TIR domains** important?



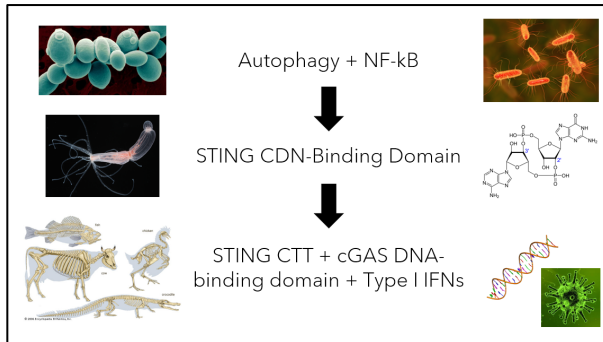
Proves evolutionary diversification of STING and TIRs are involved in innate immune signaling

Summary

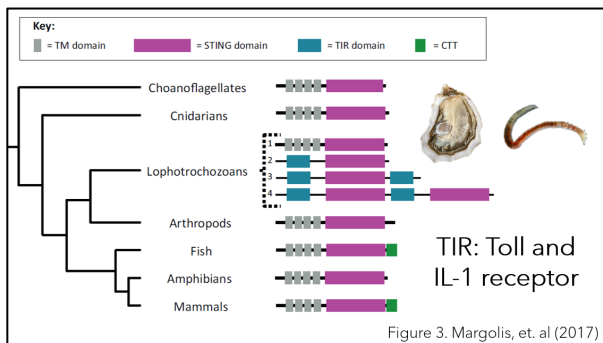
Some features, like STING CDN binding, are highly conserved



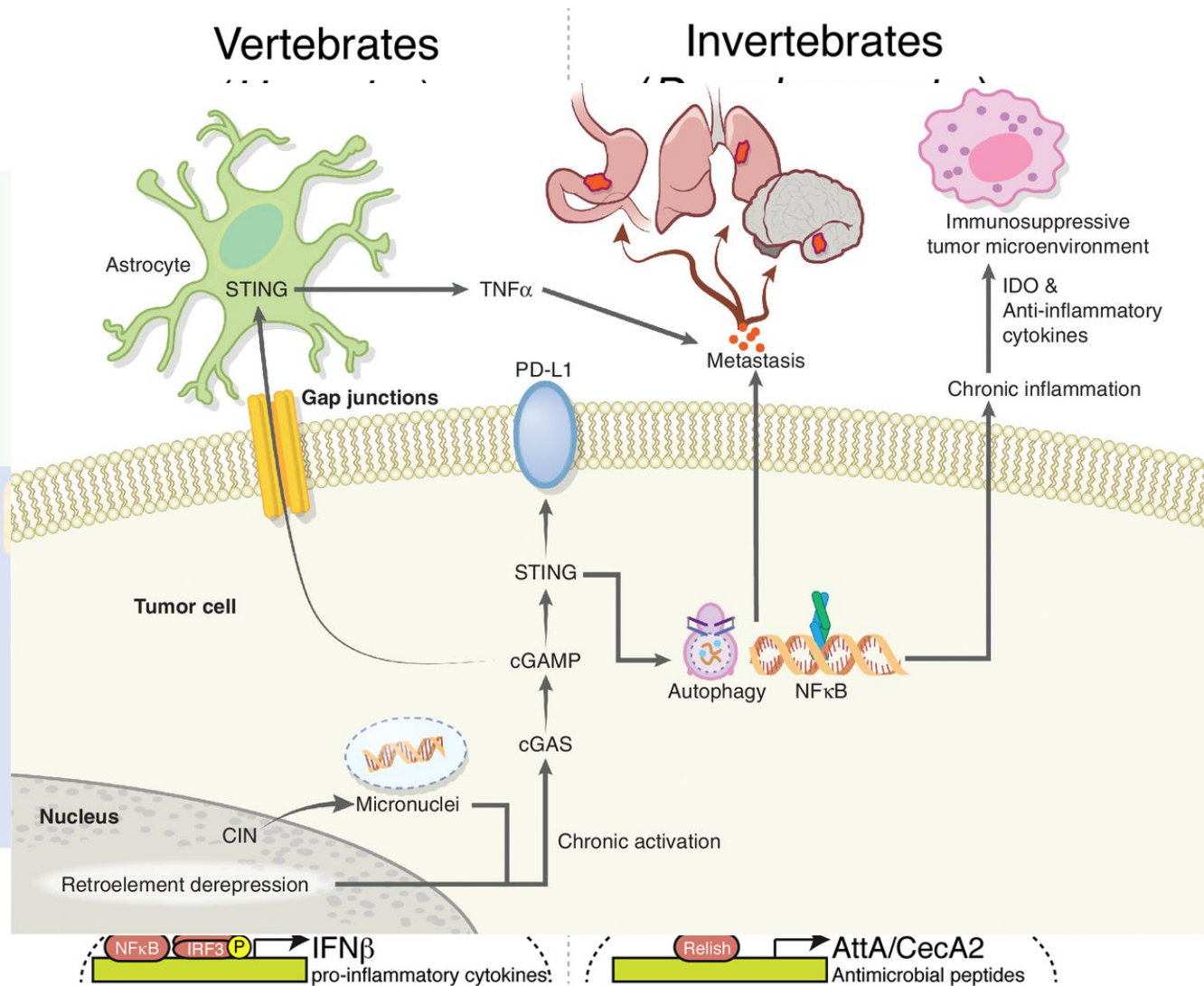
Vertebrate specific features may have evolved on top of ancestral STING pathway



Divergent STING complexes in other organisms may have adapted different features, like TIR



Future Directions



Function of divergent
STING homologs

Importance of NF- κ B
pathway and autophagy

Implications for health and
immune response

Questions?

Additional Images

Images from <https://www.khanacademy.org/science/biology/her/tree-of-life/a/building-an-evolutionary-tree>

Slide adapted from Bachinski, T. & Rogers, R.

Image 1: Baum & Smith, p. 48 Image 2: <https://www.sciencedirect.com/science/article/pii/S0168952503001124#FIG1>

Image 1: <https://www.cshl.edu/quiz/dna-and-pop-culture-quiz/>

Image 2: <https://www.rambus.com/blogs/algorithms-for-the-enterprise-2/>

Image 3: Adapted from Baum, D.

Image adapted from: <https://www.sciencedirect.com/science/article/pii/S0168952503001124#FIG1>

Image 1: <https://fuzzyatelin.github.io/bioanth-stats/module-24/module-24.html>

Image 2: <https://www.sciencedirect.com/science/article/pii/S1055790314003066>

Images from http://mesquiteproject.org/mesquiteArchives/mesquite2.75/Mesquite_Folder/docs/mesquite/studies/study002/index.html

Image 1: evolution.genetics.washington.edu/phylip.html

Image 2: https://en.wikipedia.org/wiki/Molecular_Evolutionary_Genetics_Analysis

Image 3: <https://www.sc.fsu.edu/software>

Image 1: <https://www.sciencedirect.com/science/article/pii/S0168952503001124#FIG1>

Image 1: http://clipart-library.com/clip-art/165-1657025_dna-svg-black-png-genetics-black-and-white.htm

Image 2: <https://pdb101.rcsb.org/motm/206>

Slide 35: Vertebrates: <https://cdn.britannica.com/14/93314-004-F9954BFA/vertebrate-skeletons.jpg>

Nematostella: https://en.wikipedia.org/wiki/Starlet_sea_anemone

References and Additional Images

Ablasser, Andrea, and Zhijian J. Chen. "Cgas In Action: Expanding Roles In Immunity And Inflammation". *Science*, vol 363, no. 6431, 2019. *American Association For The Advancement Of Science (AAAS)*, doi:10.1126/science.aat8657.

Baldauf SL. Phylogeny for the faint of heart: a tutorial. *Trends Genet*. 2003 Jun;19(6):345-51. PMID: 12801728

*Baum, D (2008): <https://www.nature.com/scitable/topicpage/reading-a-phylogenetic-tree-the-meaning-of-41956>

de Oliveira Mann, Carina C. et al. "Modular Architecture Of The STING C-Terminal Tail Allows Interferon And NF-Kb Signaling Adaptation". *Cell Reports*, vol 27, no. 4, 2019, pp. 1165-1175.e5. *Elsevier BV*, doi:10.1016/j.celrep.2019.03.098.

Hong, Christy et al. "The Cgas Paradox: Contrasting Roles For Cgas-STING Pathway In Chromosomal Instability". *Cells*, vol 8, no. 10, 2019, p. 1228. *MDPI AG*, doi:10.3390/cells8101228.

Kato, Kazuki et al. "Structural And Functional Analyses Of DNA-Sensing And Immune Activation By Human Cgas". *Plos ONE*, vol 8, no. 10, 2013, p. e76983. *Public Library Of Science (Plos)*, doi:10.1371/journal.pone.0076983.

Margolis SR1, Wilson SC2, Vance RE3. Evolutionary Origins of cGAS-STING Signaling. *Trends Immunol*. 2017 Oct;38(10):733-743. doi: 10.1016/j.it.2017.03.004. Epub 2017 Apr 14.