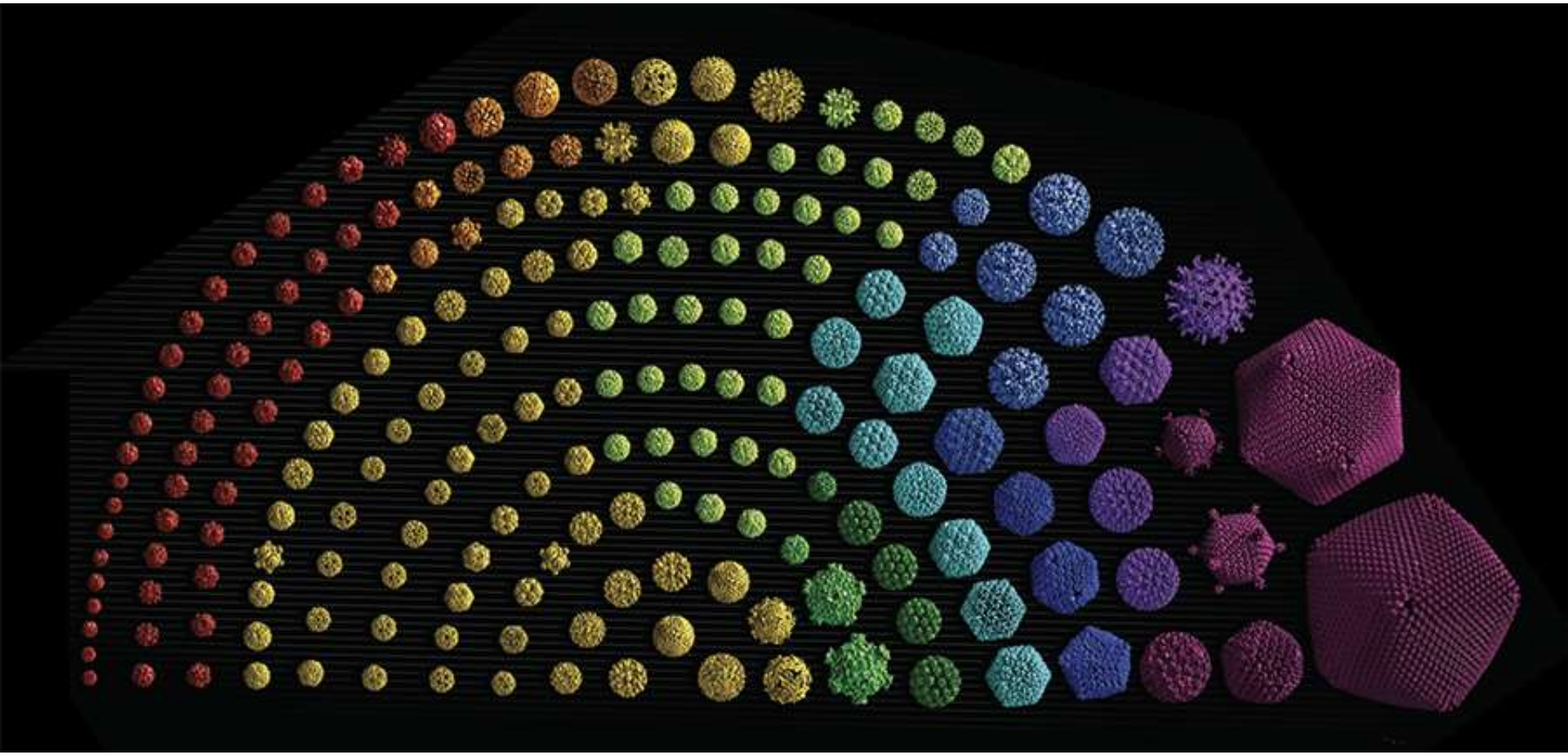
Capsidocentric view on virus diversity, origins and evolution



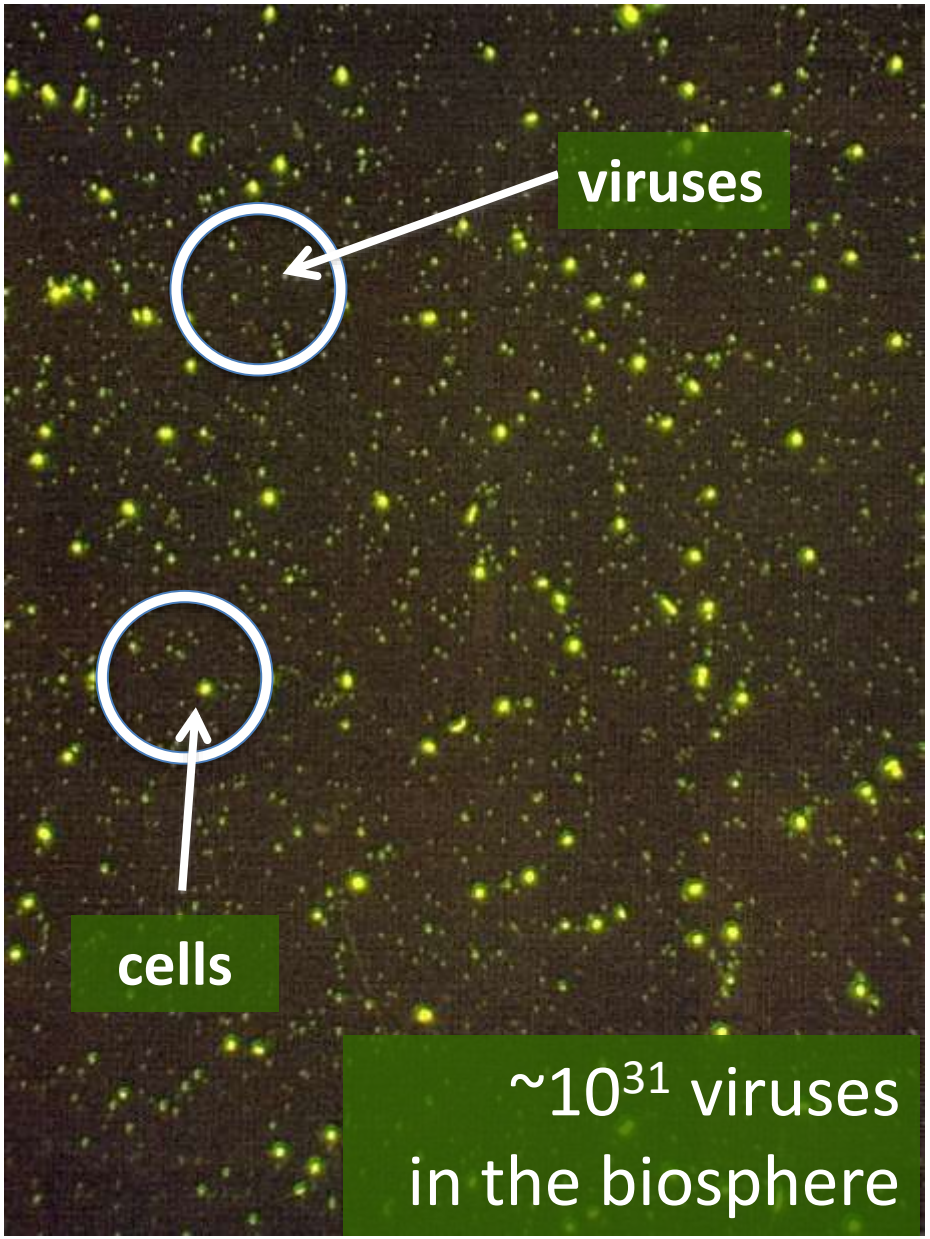
Mart Krupovic

Institut Pasteur

Evolution Symposium, Salzburg

05.07.2018

There are more viruses on Earth than there are stars in the Universe



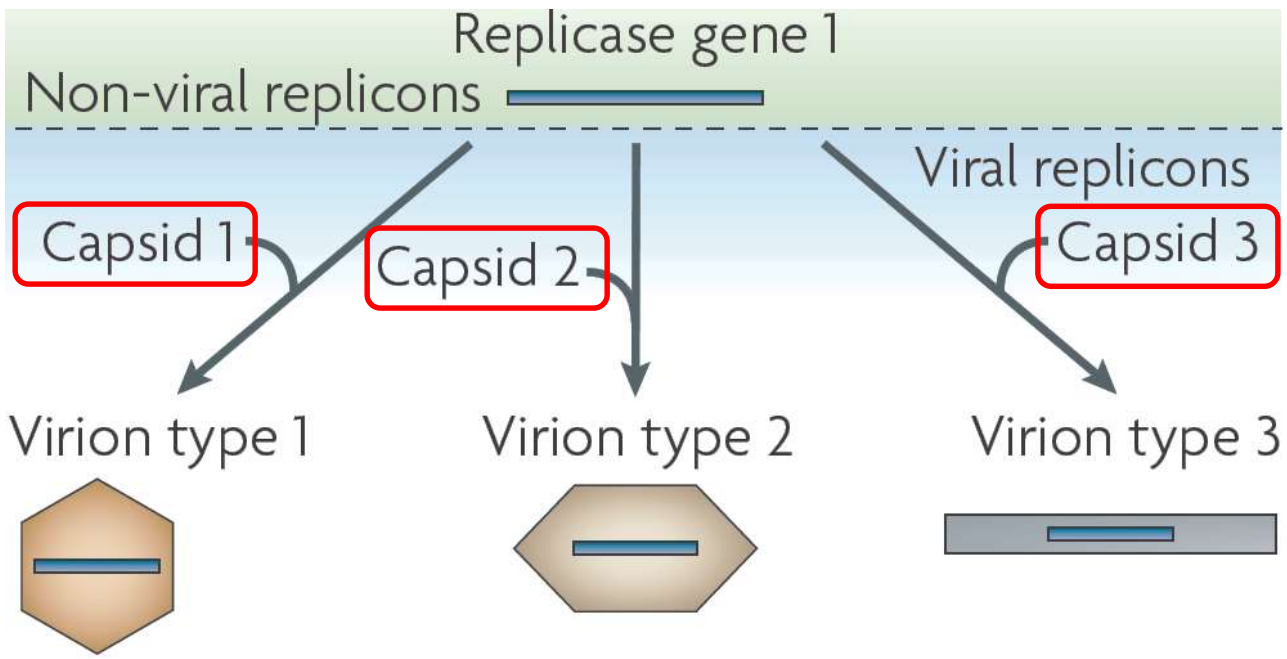
Adenoviridae, Albetovirus, Alloherpesviridae, Alphaflexiviridae, Alphatetraviridae, Alvernaviridae, Amalgaviridae, Ampullaviridae, Anelloviridae, Arenaviridae, Arteriviridae, Ascoviridae, Asfarviridae, Astroviridae, Aumaivirus, Bacilladnavirus, Bacillarnavirus, Baculoviridae, Barnaviridae, Benyviridae, Betaflexiviridae, Bicaudaviridae, Bidnaviridae, Birnaviridae, Bornaviridae, Botybirnavirus, Bromoviridae, Caliciviridae, Carmotetraviridae, Caulimoviridae, Chrysoviridae, Cilevirus, Circoviridae, Clavaviridae, Closteroviridae, Coronaviridae, Corticoviridae, Cystoviridae, Deltavirus, Dicistroviridae, Emaraviridae, Endornaviridae, Faustovirus, Feraviridae, Filoviridae, Flaviviridae, Fuselloviridae, Gammaflexiviridae, Geminiviridae, Genomoviridae, Globuloviridae, Guttaviridae, Hantaviridae, Hepadnaviridae, Hepeviridae, Herpesviridae, Higrevirus, Hypoviridae, Hytrosaviridae, Idaeovirus, Iflaviridae, Inoviridae, Iridoviridae, Jonviridae, Labyrnavirus, Lavidaviridae, Leviviridae, Lipothrixviridae, Luteoviridae, Malacoherpesviridae, Marnaviridae, Marseilleviridae, Megabirnaviridae, Mesoniviridae, Metaviridae, Microviridae, Mimiviridae, Mymonaviridae, Myoviridae, Nanoviridae, Nairoviridae, Narnaviridae, Nimaviridae, Nodaviridae, Nudiviridae, Nyamiviridae, Ophioviridae, Orthomyxoviridae, Ourmiavirus, Papanivirus, Papillomaviridae, Paramyxoviridae, Partitiviridae, Parvoviridae, Peribunyaviridae, Permutotetraviridae, Phenuiviridae, Phycodnaviridae, Picobirnaviridae, Picornaviridae, Phasmaviridae, Plasmaviridae, Pleolipoviridae, Pneumoviridae, Podoviridae, Polemovirus, Polydnaviridae, Polyomaviridae, Potyviridae, Poxviridae, Pseudoviridae, Quadriviridae, Reoviridae, Retroviridae, Rhabdoviridae, Roniviridae, Rudiviridae, Salterprovirus, Sarthroviridae, Secoviridae, Sinaivirus, Siphoviridae, Sobemovirus, Sphaerolipoviridae, Spiraviridae, Sunviridae, Tectiviridae, Togaviridae, Tombusviridae, Tospoviridae, Totiviridae, Tristromaviridae, Turriviridae, Tymoviridae, Virgaviridae, Virtovirus

135 different virus families

QUESTIONS:

- Extent of virus diversity (fundamentally different virus groups)
- (Macro)evolutionary connections in the virosphere
- Origins of viruses

Origin of capsid genes = origin of (true) viruses



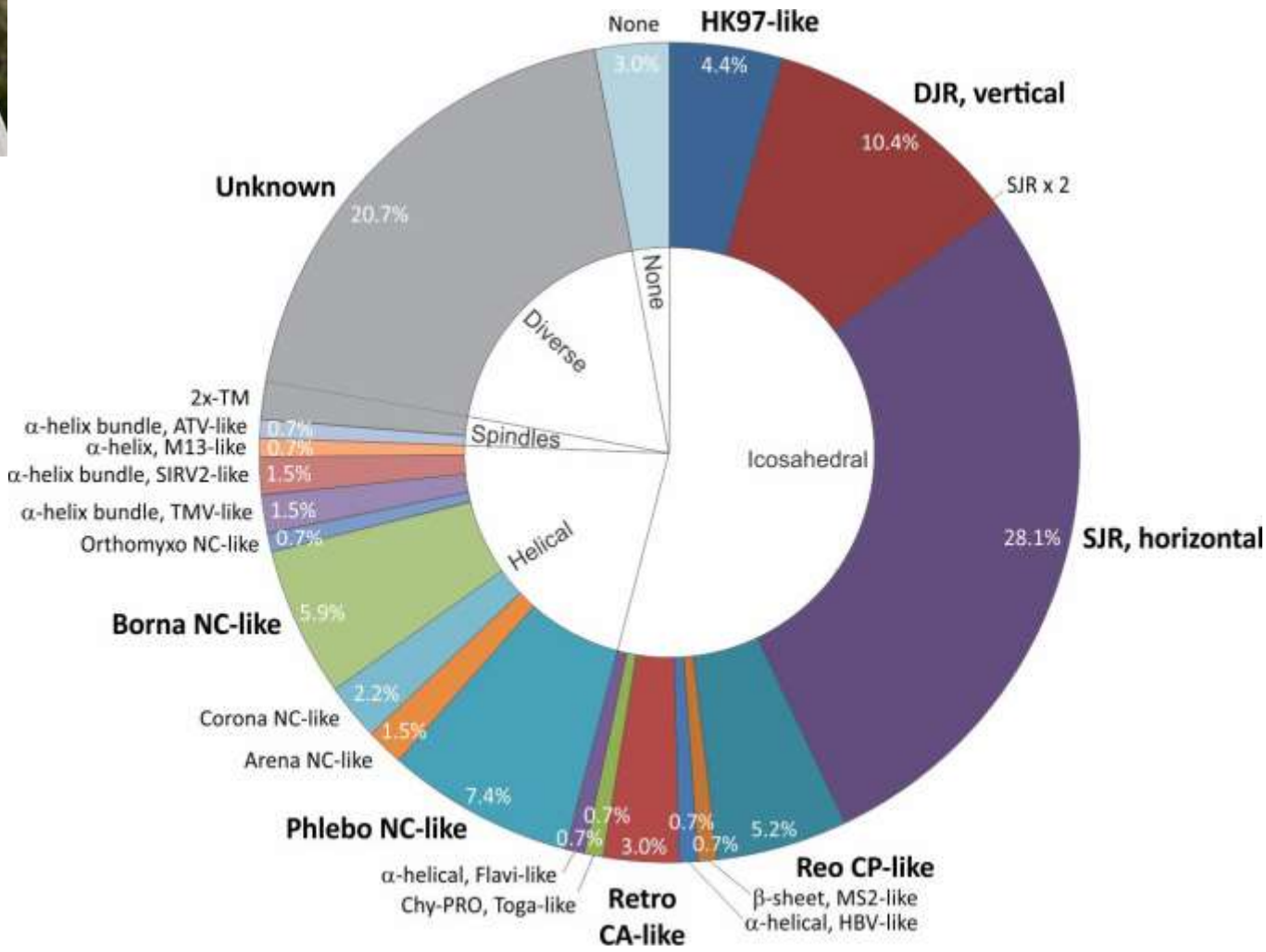
Krupovic & Bamford, *Nat Rev Microbiol* 2009

Adenoviridae, Albetovirus, Alloherpesviridae, Alphaflexiviridae, Alphatetraviridae, Alvernaviridae, Amalgaviridae, Ampullaviridae, Anelloviridae, Arenaviridae, Arteriviridae, Ascoviridae, Asfarviridae, Astroviridae, Aumaivirus, Bacilladnavirus, Bacillarnavirus, Baculoviridae, Barnaviridae, Benyviridae, Betaflexiviridae, Bicaudaviridae, Bidnaviridae, Birnaviridae, Bornaviridae, Botybirnavirus, Bromoviridae, Caliciviridae, Carmotetraviridae, Caulimoviridae, Chrysoviridae, Cilevirus, Circoviridae, Clavaviridae, Closteroviridae, Coronaviridae, Corticoviridae, Cystoviridae, Deltavirus, Dicistroviridae, Emaraviridae, Endornaviridae, Faustovirus, Feraviridae, Filoviridae, Flaviviridae, Fuselloviridae, Gammaflexiviridae, Geminiviridae, Genomoviridae, Globuloviridae, Guttaviridae, Hantaviridae, Hepadnaviridae, Hepeviridae, Herpesviridae, Higrevirus, Hypoviridae, Hytrosaviridae, Idaeovirus, Iflaviridae, Inoviridae, Iridoviridae, Jonviridae, Labyrnavirus, Lavidaviridae, Leviviridae, Lipothrixviridae, Luteoviridae, Malacoherpesviridae, Marnaviridae, Marseilleviridae, Megabirnaviridae, Mesoniviridae, Metaviridae, Microviridae, Mimiviridae, Mymonaviridae, Myoviridae, Nanoviridae, Nairoviridae, Narnaviridae, Nimaviridae, Nodaviridae, Nudiviridae, Nyamiviridae, Ophioviridae, Orthomyxoviridae, Ourmiavirus, Papanivirus, Papillomaviridae, Paramyxoviridae, Partitiviridae, Parvoviridae, Peribunyaviridae, Permutotetraviridae, Phenuiviridae, Phycodnaviridae, Picobirnaviridae, Picornaviridae, Phasmaviridae, Plasmaviridae, Pleolipoviridae, Pneumoviridae, Podoviridae, Polemovirus, Polydnaviridae, Polyomaviridae, Potyviridae, Poxviridae, Pseudoviridae, Quadriviridae, Reoviridae, Retroviridae, Rhabdoviridae, Roniviridae, Rudiviridae, Salterprovirus, Sarthroviridae, Secoviridae, Sinaivirus, Siphoviridae, Sobemovirus, Sphaerolipoviridae, Spiraviridae, Sunviridae, Tectiviridae, Togaviridae, Tombusviridae, Tospoviridae, Totiviridae, Tristromaviridae, Turriviridae, Tymoviridae, Virgaviridae, Virtovirus

135 different virus families

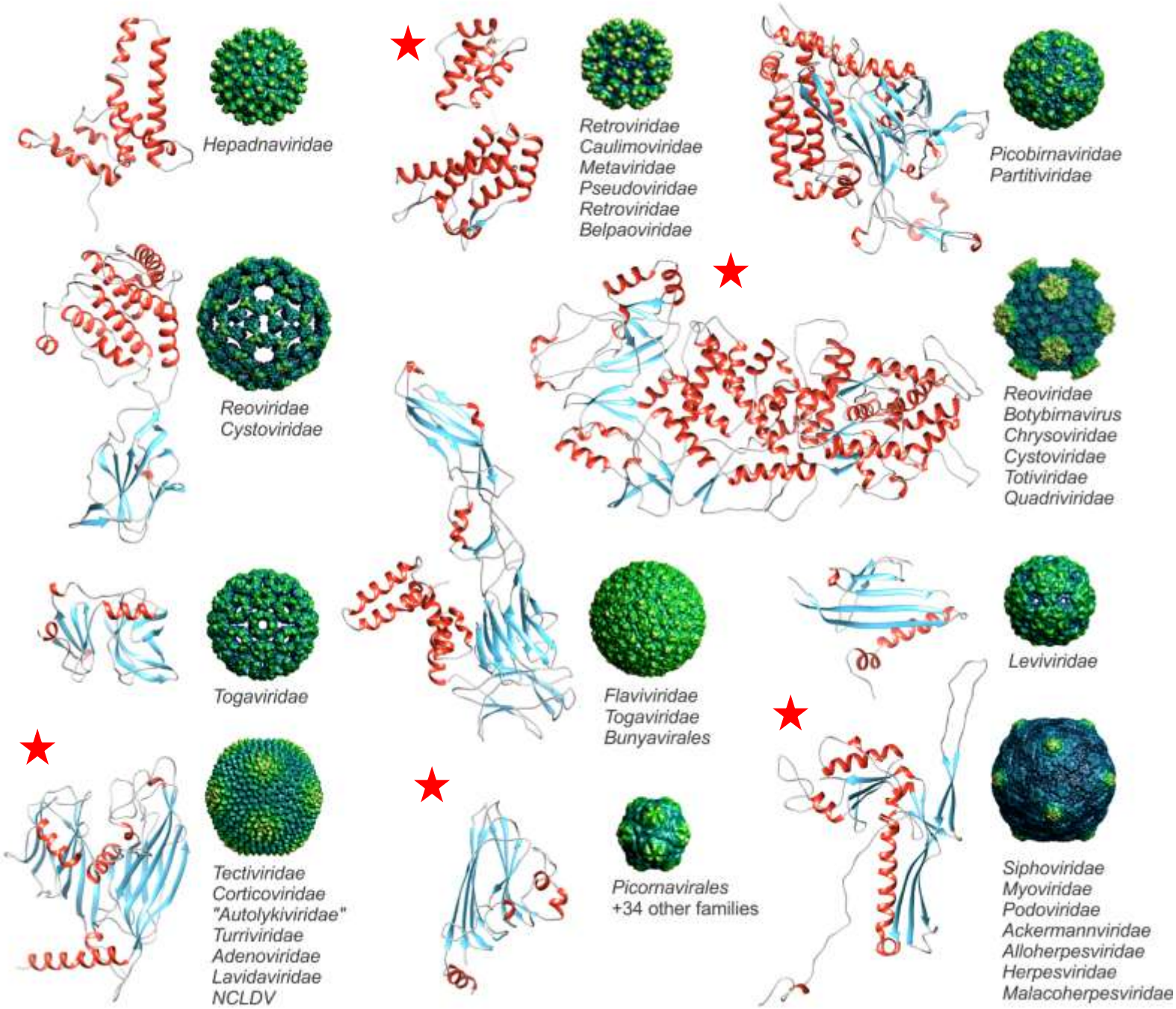
135 families

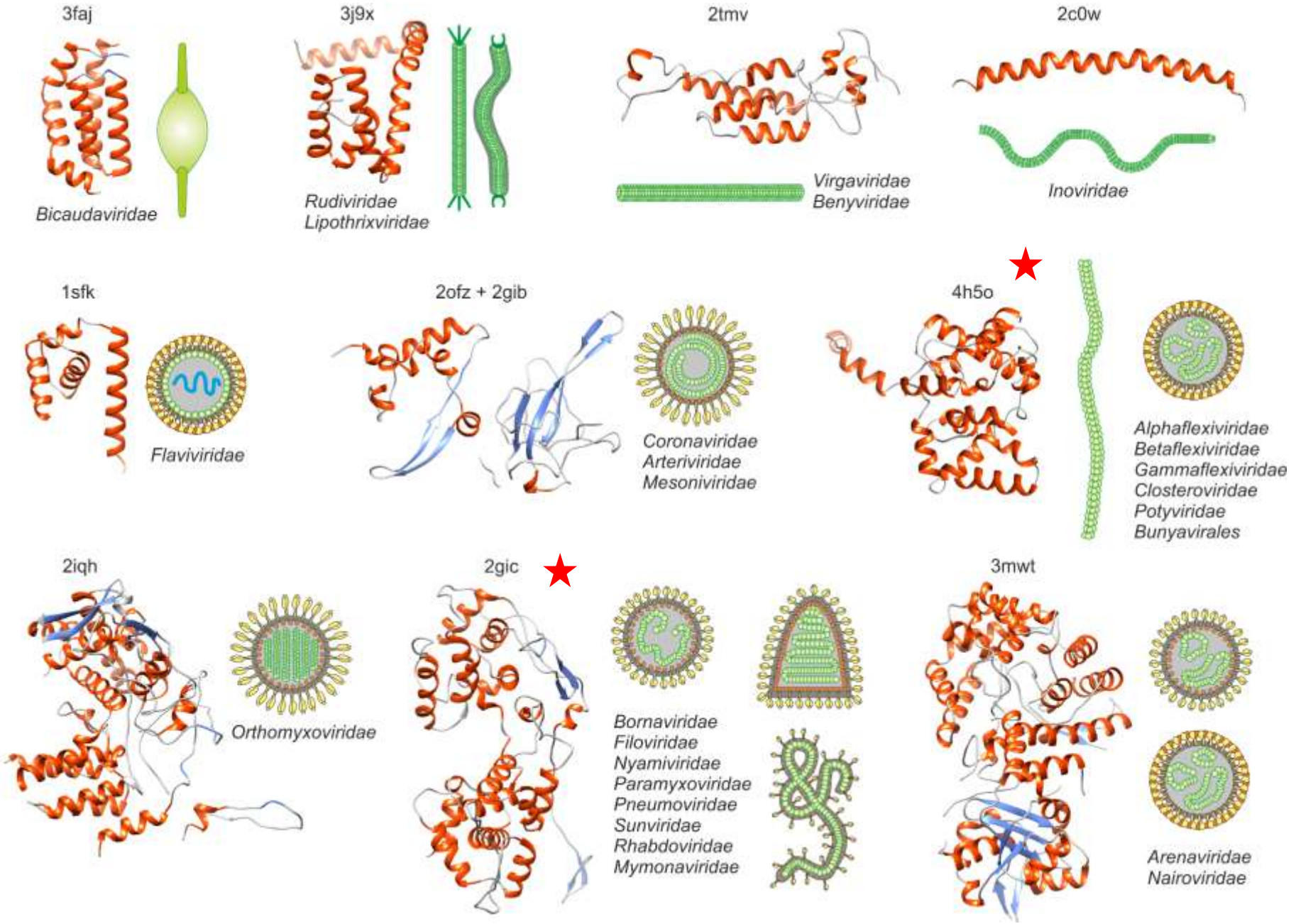
18 folds - 76% of viral families



Krupovic & Koonin, *PNAS* 2017

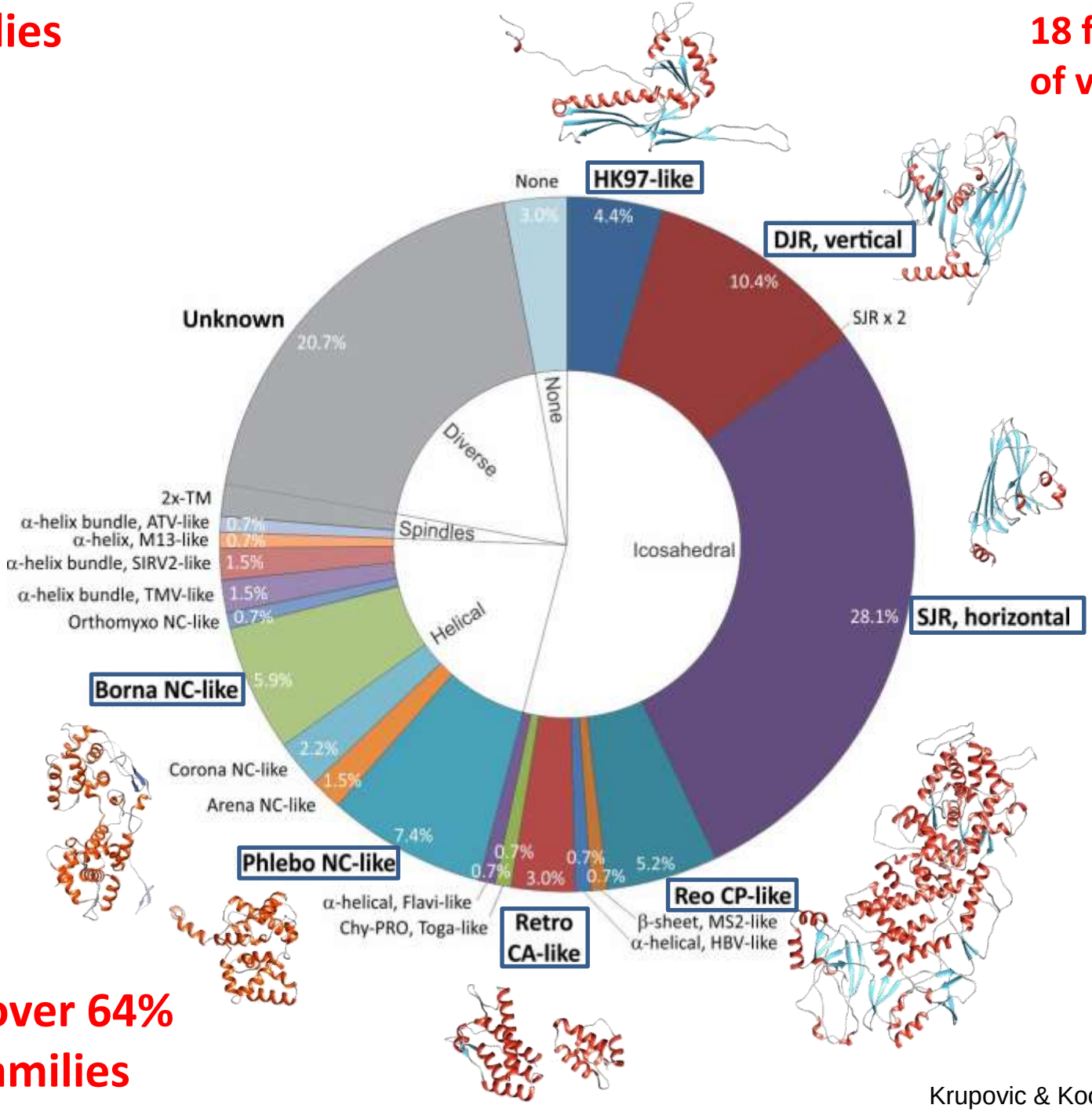
11 folds for icosahedral capsids




3faj
Bicaudaviridae
3j9x
Rudiviridae
Lipothrixviridae
2tmv
Virgaviridae
Benyviridae
2c0w
Inoviridae
1sfk
Flaviviridae
2ofz + 2gib
Coronaviridae
Arteriviridae
Mesoniviridae
4h5o
Alphaflexiviridae
Betaflexiviridae
Gammaflexiviridae
Closteroviridae
Potyviridae
Bunyavirales
2iqh
Orthomyxoviridae
2gic
Bornaviridae
Filoviridae
Nyamiviridae
Paramyxoviridae
Pneumoviridae
Sunviridae
Rhabdoviridae
Mymonaviridae
3mwt
Arenaviridae
Nairoviridae

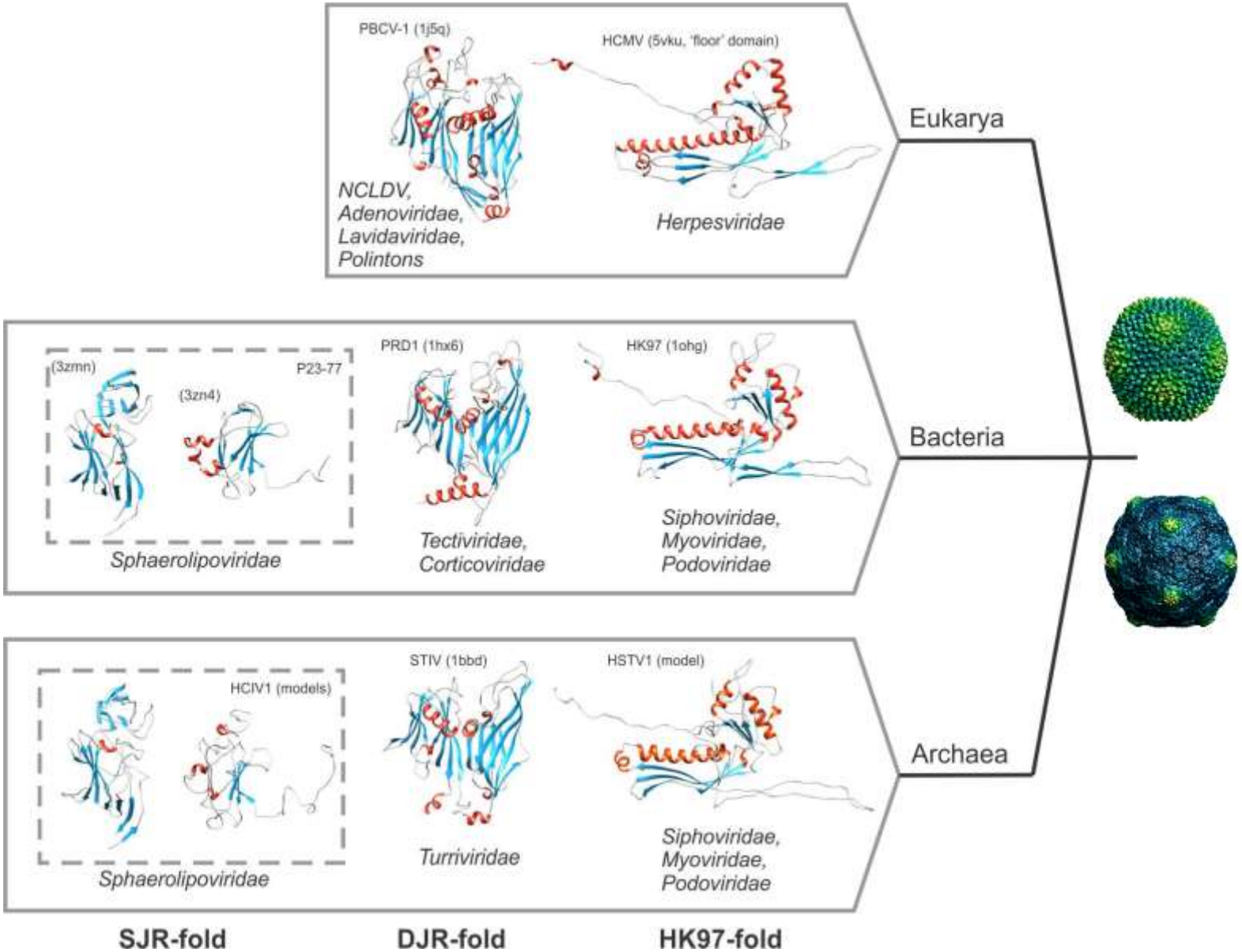
135 families

18 folds - 76% of viral families



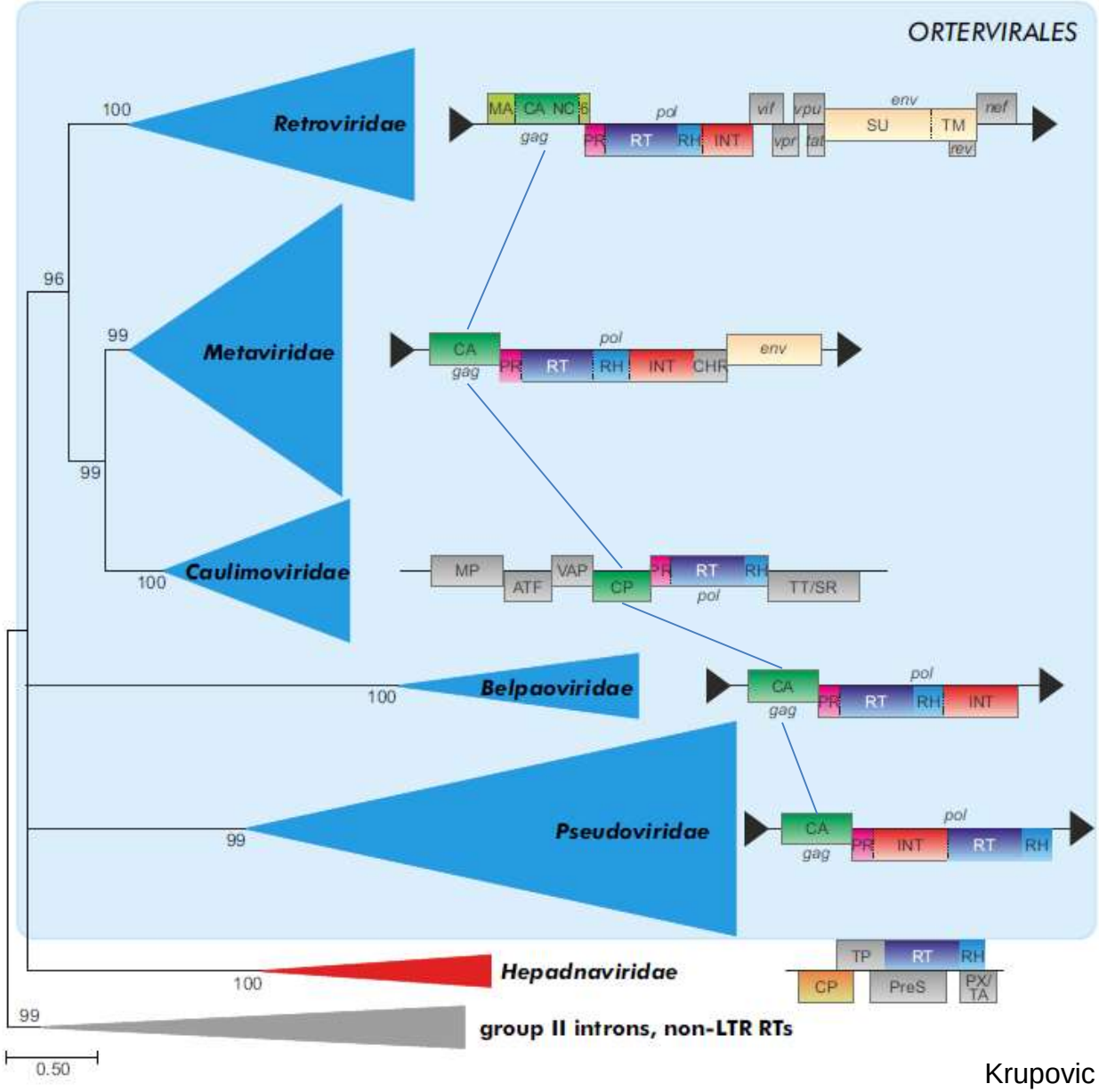
7 folds cover 64% of viral families

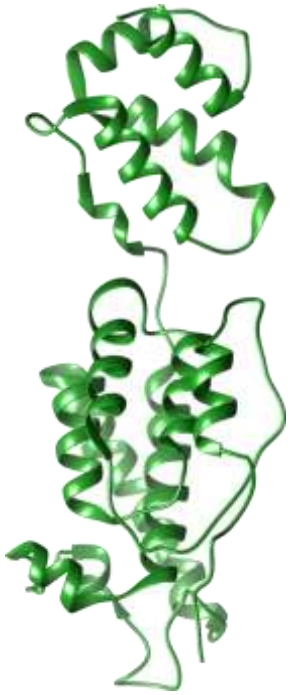
Krupovic & Koonin, *PNAS* 2017



Prangishvili, Bamford, Forterre, Iranzo, Koonin, Krupovic, *Nat Rev Microbiol* 2017

Ortervirales: New virus order unifying five families of reverse-transcribing viruses





Retro CA

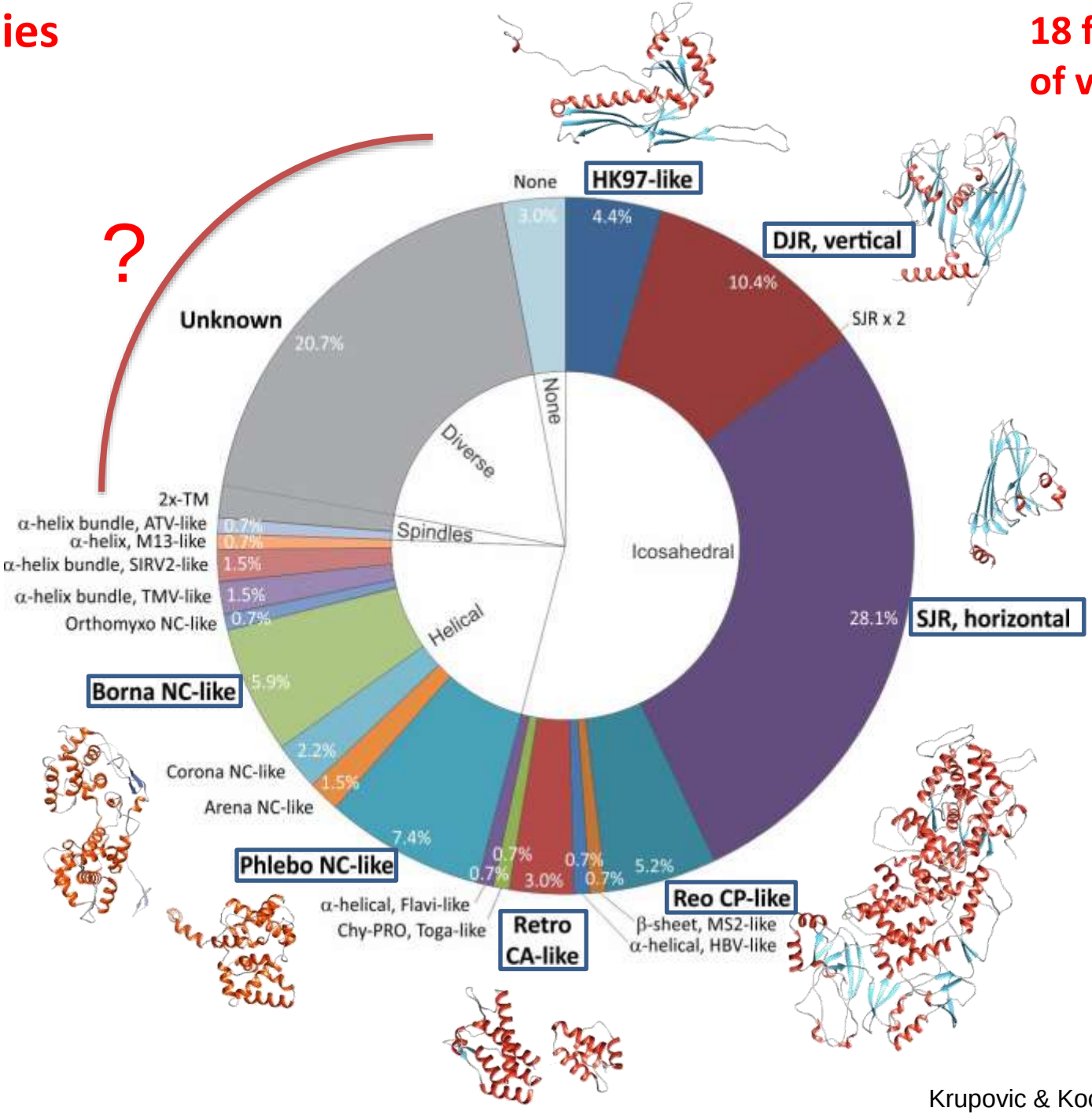
RT phylogeny

Krupovic & Koonin, *J Virol* 2017

Krupovic et al., *J Virol* 2018

135 families

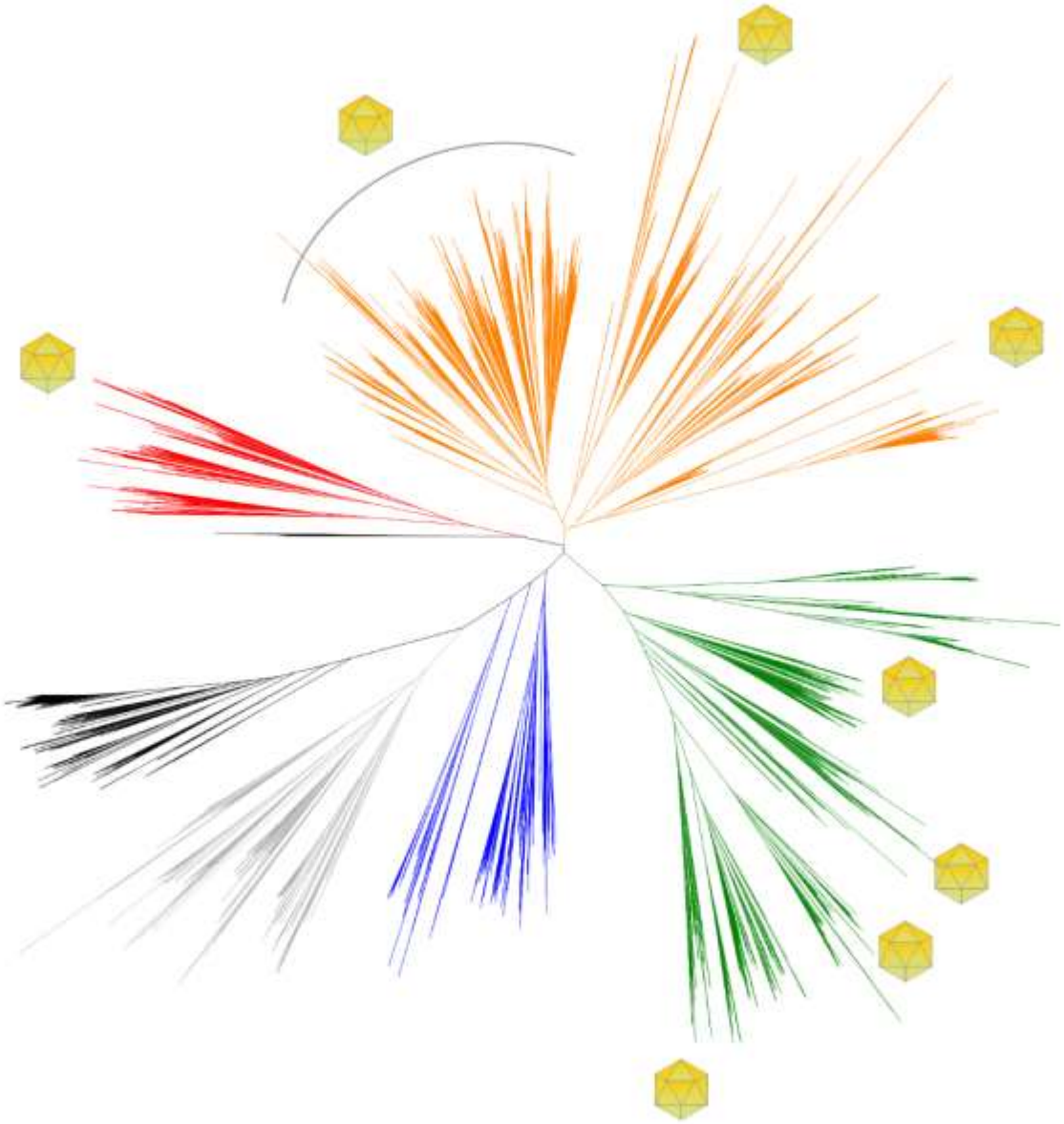
18 folds - 76% of viral families



Krupovic & Koonin, *PNAS* 2017

ORIGINS OF VIRAL CAPSID PROTEINS (& VIRUSES)

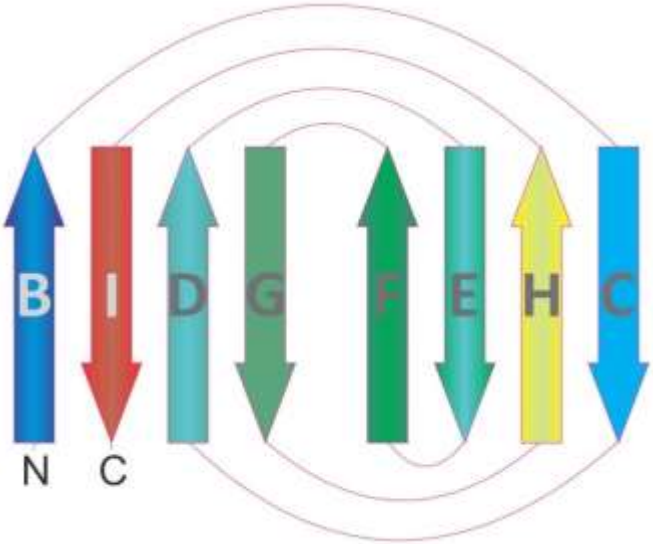
Jelly-roll fold: the most widespread CP fold (28.1%)



RdRp phylogeny

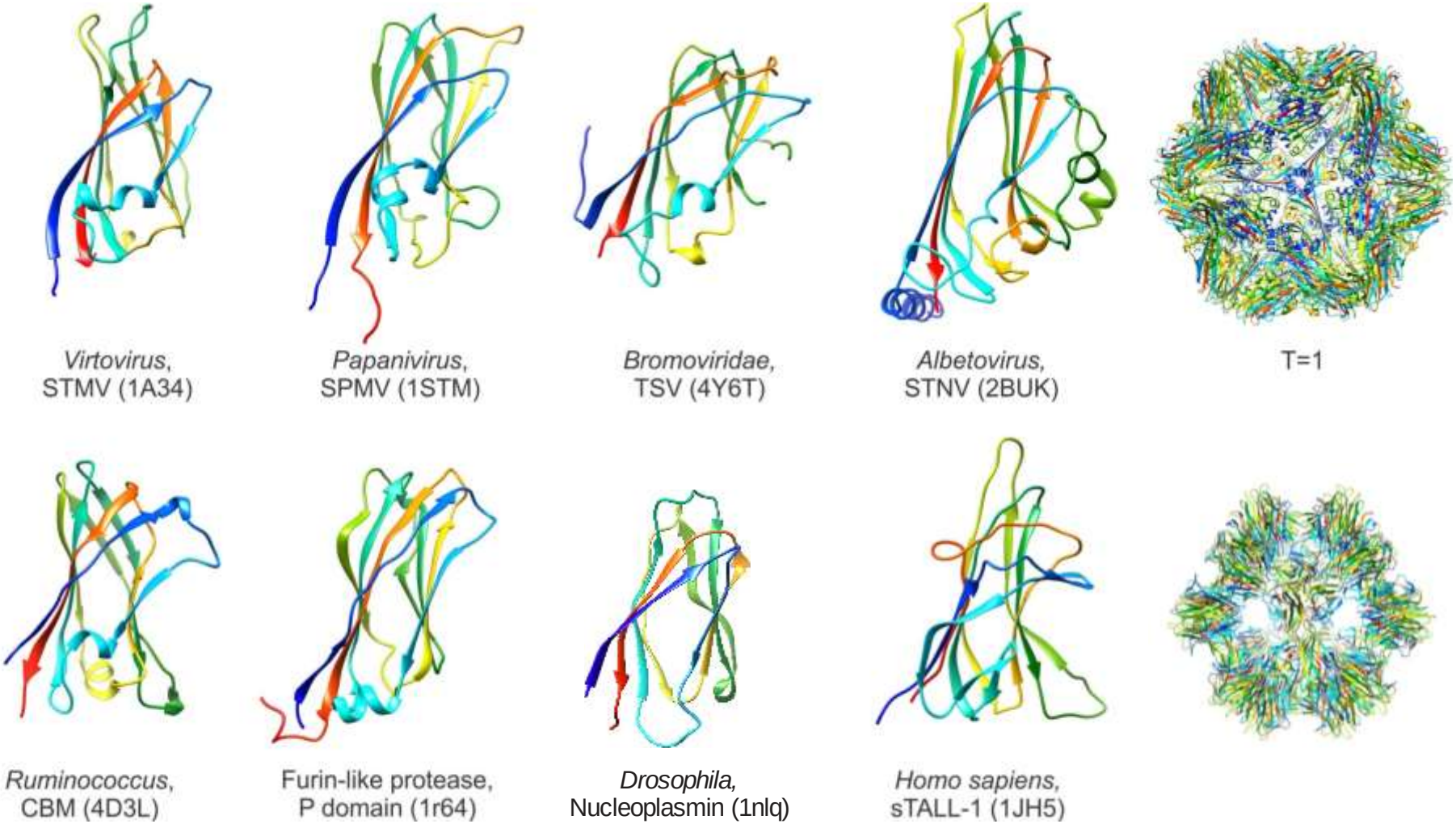
Jelly-roll fold: the most widespread CP fold (28.1%)

Jelly-roll fold

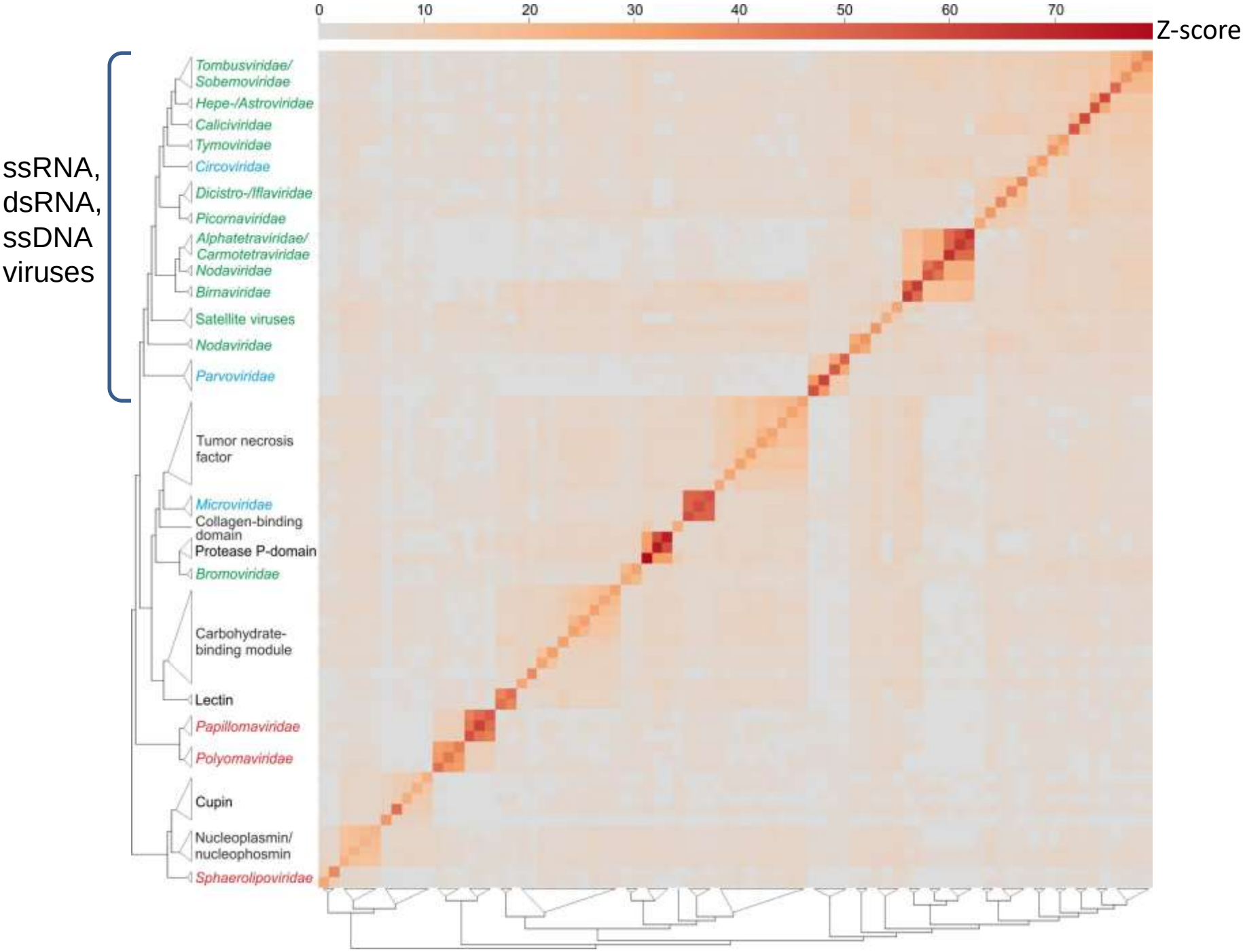


Jelly-roll cake

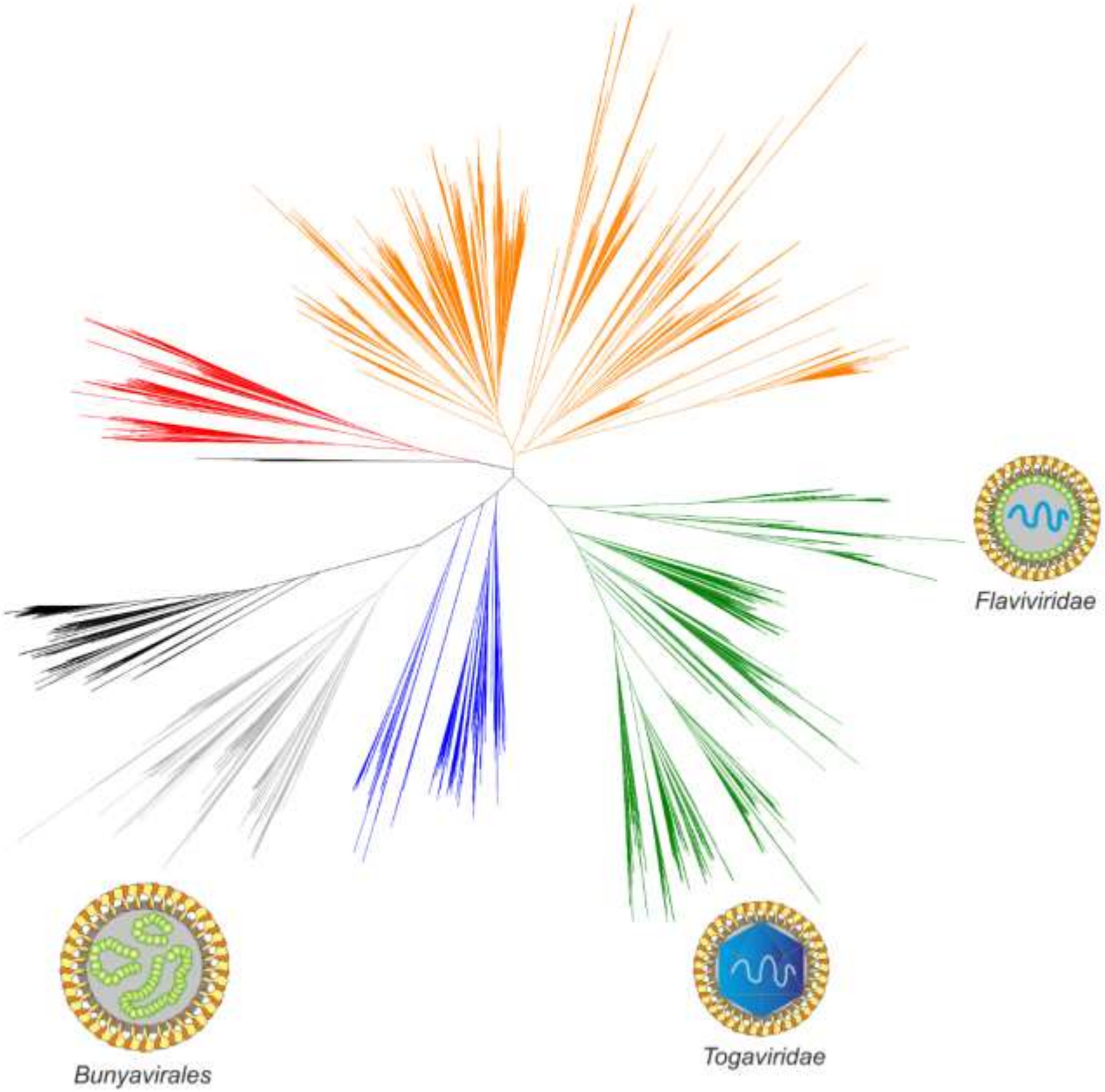
Jelly-roll fold: the most widespread CP fold (28.1%)



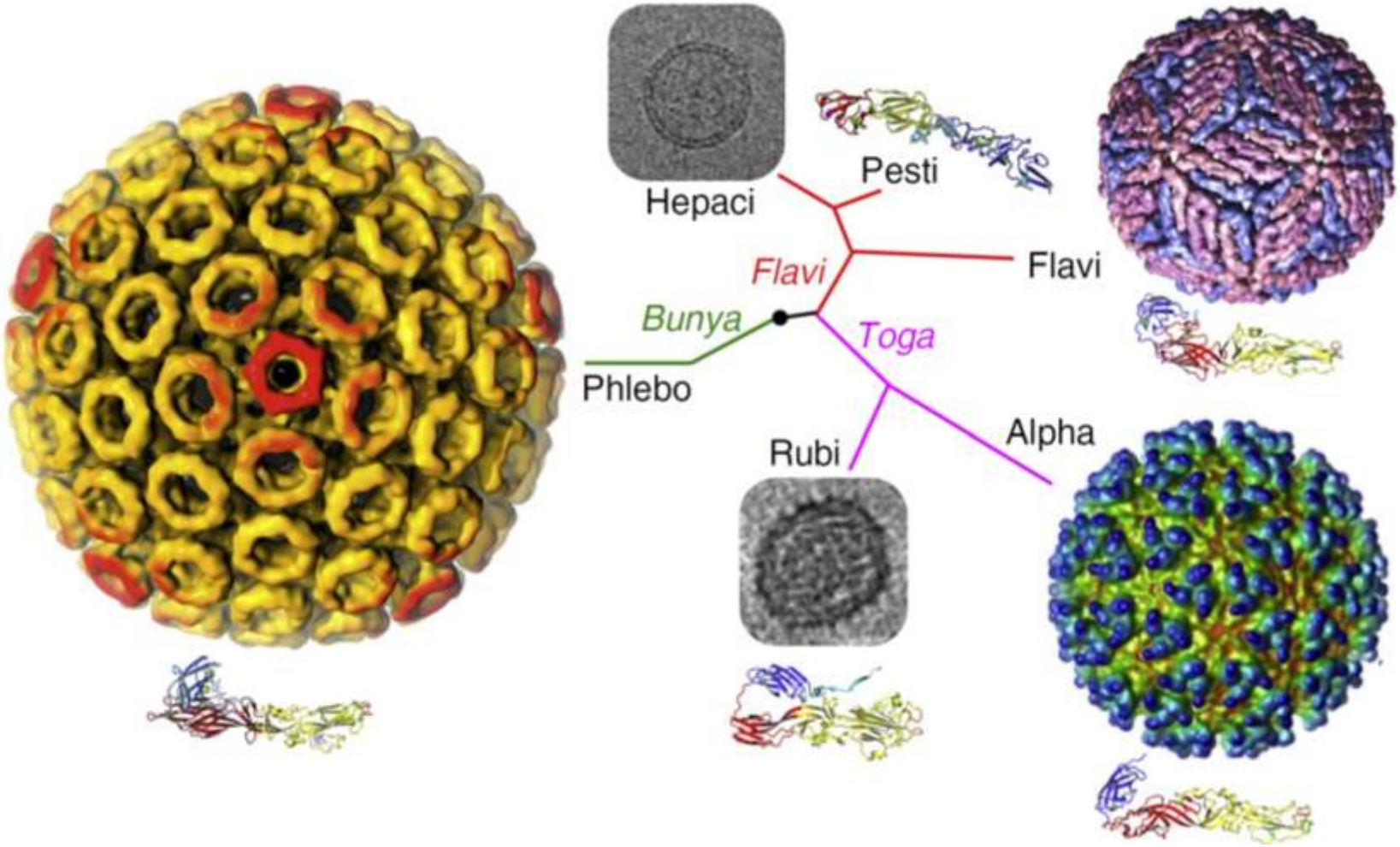
Krupovic & Koonin, *PNAS* 2017


0
10
20
30
40
50
60
70
Z-score
ssRNA, dsRNA, ssDNA viruses
Tombusviridae/ Sobemoviridae
Hepe-/Astroviridae
Caliciviridae
Tymoviridae
Circoviridae
Dicistro-/Iflaviridae
Picornaviridae
Alphatetraviridae/ Carmotetraviridae
Nodaviridae
Birnaviridae
Satellite viruses
Nodaviridae
Parvoviridae
Tumor necrosis factor
Microviridae
Collagen-binding domain
Protease P-domain
Bromoviridae
Carbohydrate-binding module
Lectin
Papillomaviridae
Polyomaviridae
Cupin
Nucleoplasmin/ nucleophosmin
Sphaerolipoviridae

Class II fusion proteins form icosahedral shells

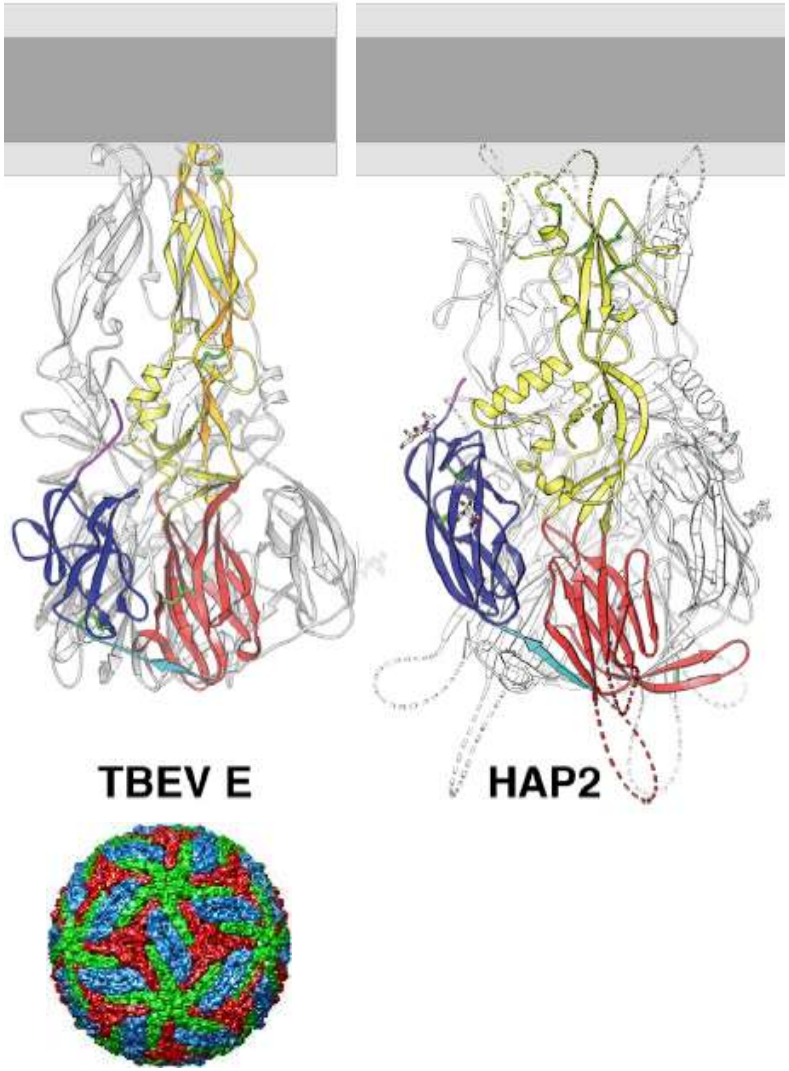


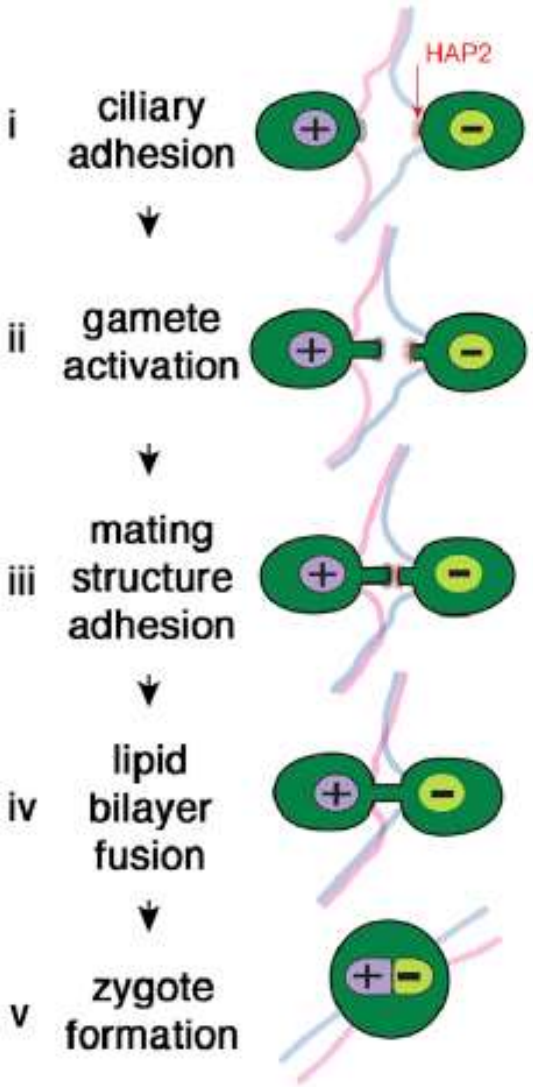
Class II fusion proteins form icosahedral shells



Modis, *Curr Opin Virol* 2014

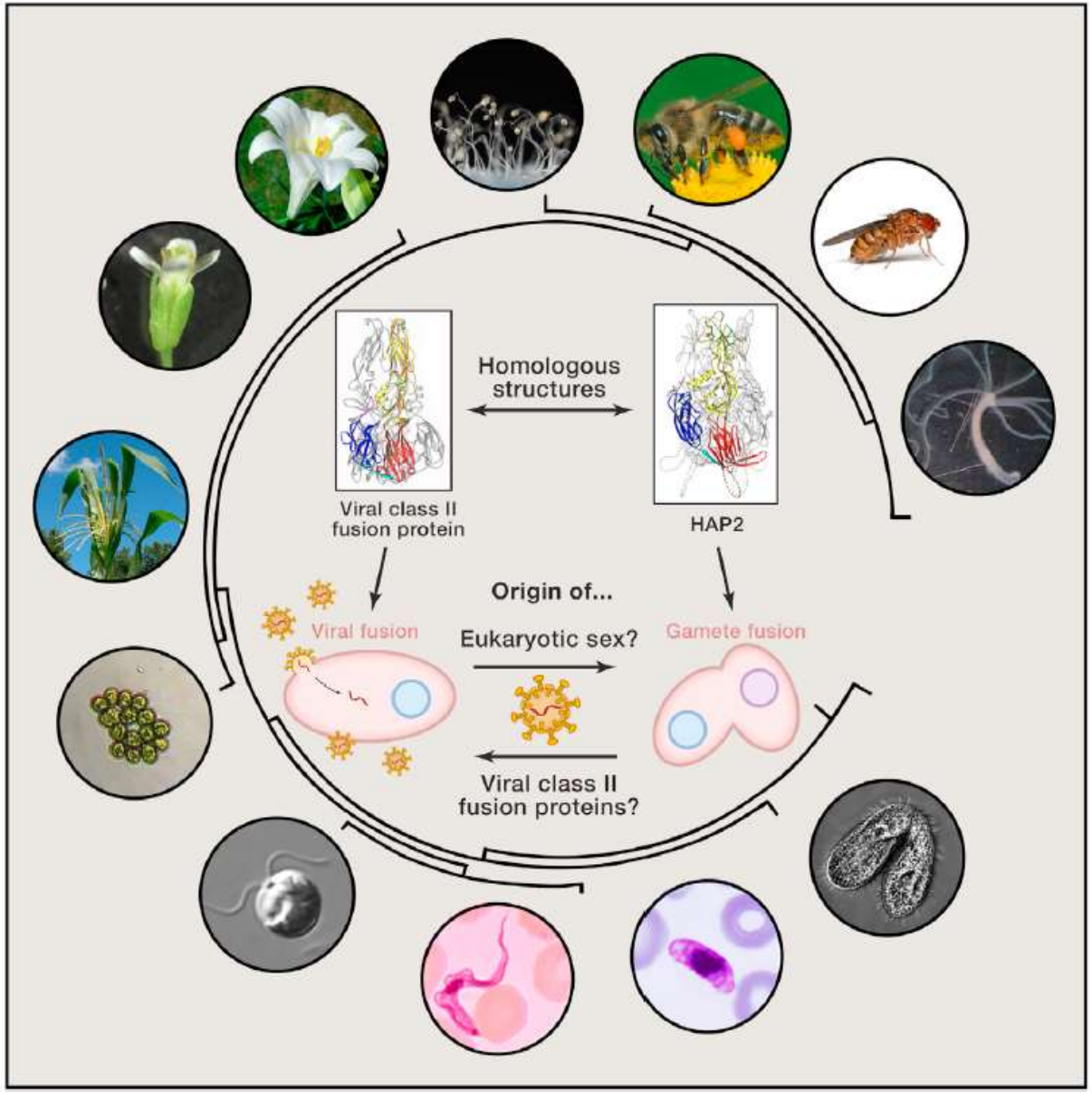
Viral class II fusion proteins are homologous to HAP2





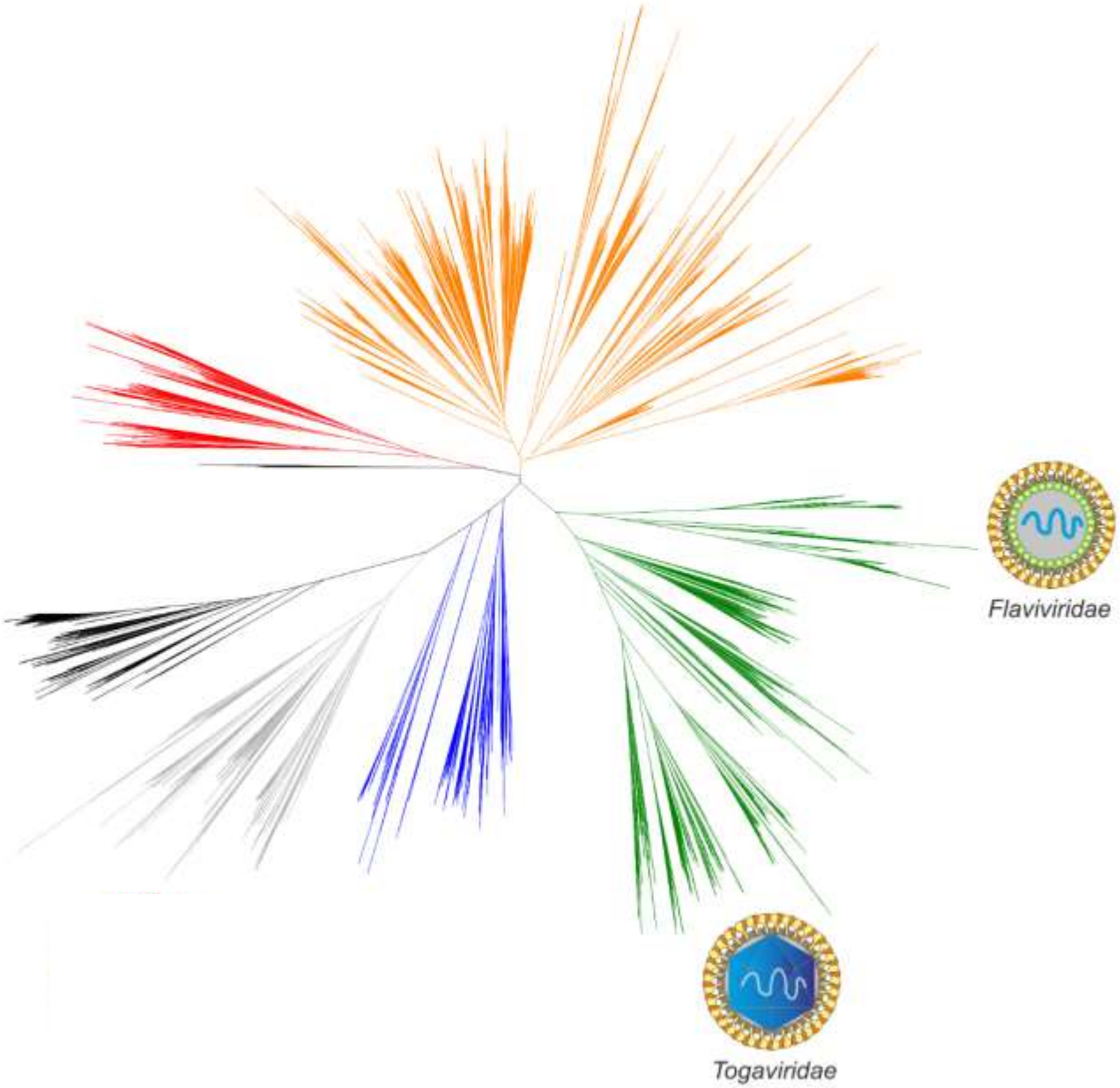
Fédry, Liu, Péhau-Arnaudet, Pei, Li, Tortorici, Traincard, Meola, Bricogne, Grishin, Snell, Rey, Krey *Cell* 2017

HAP2 is conserved in eukaryotes across kingdoms

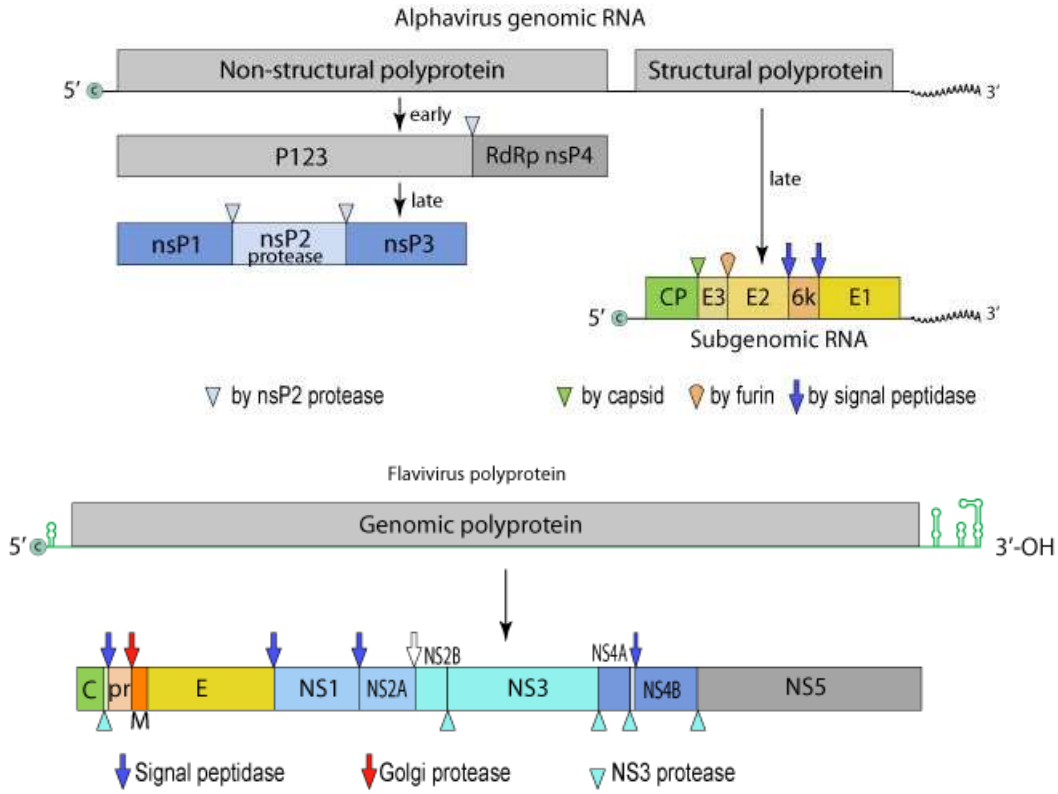


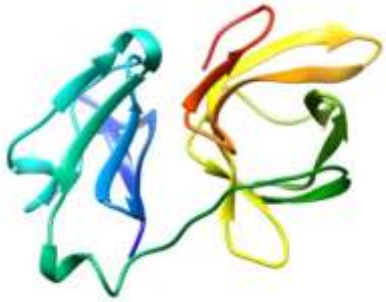
Doms, *Cell* 2017

Class II fusion proteins form icosahedral shells

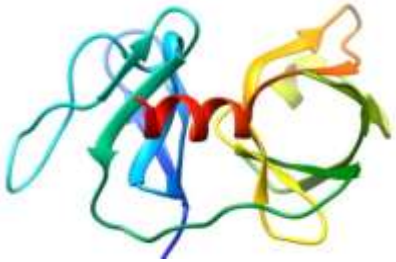


Alphavirus: capsid protein





SINV CP
Alphavirus
(Togaviridae)



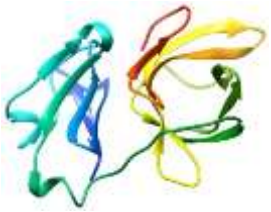
HCV NS3
(*Flaviviridae*)

HtrA PRO
(3nwu)

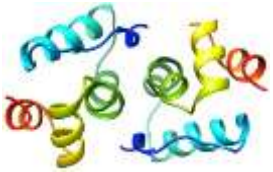
Many virion proteins have evolved from cellular ancestors

VIRAL STRUCTURAL PROTEINS


STMV CP

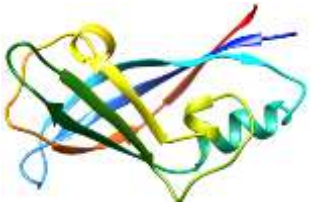

SINV CP


MMTV MA

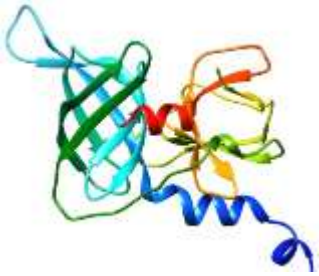

HIV-1 CA


HIV-1 NC

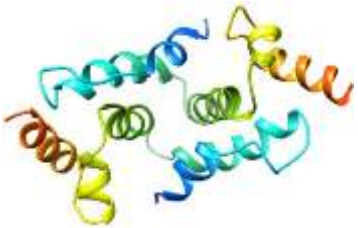

LASV Z


EBOV M


CBM

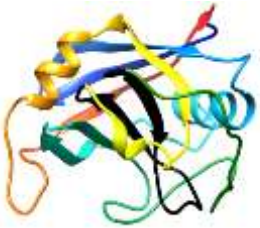

HtrA PRO


Y-integrase NTD


SCAN


Lin28


RING E3

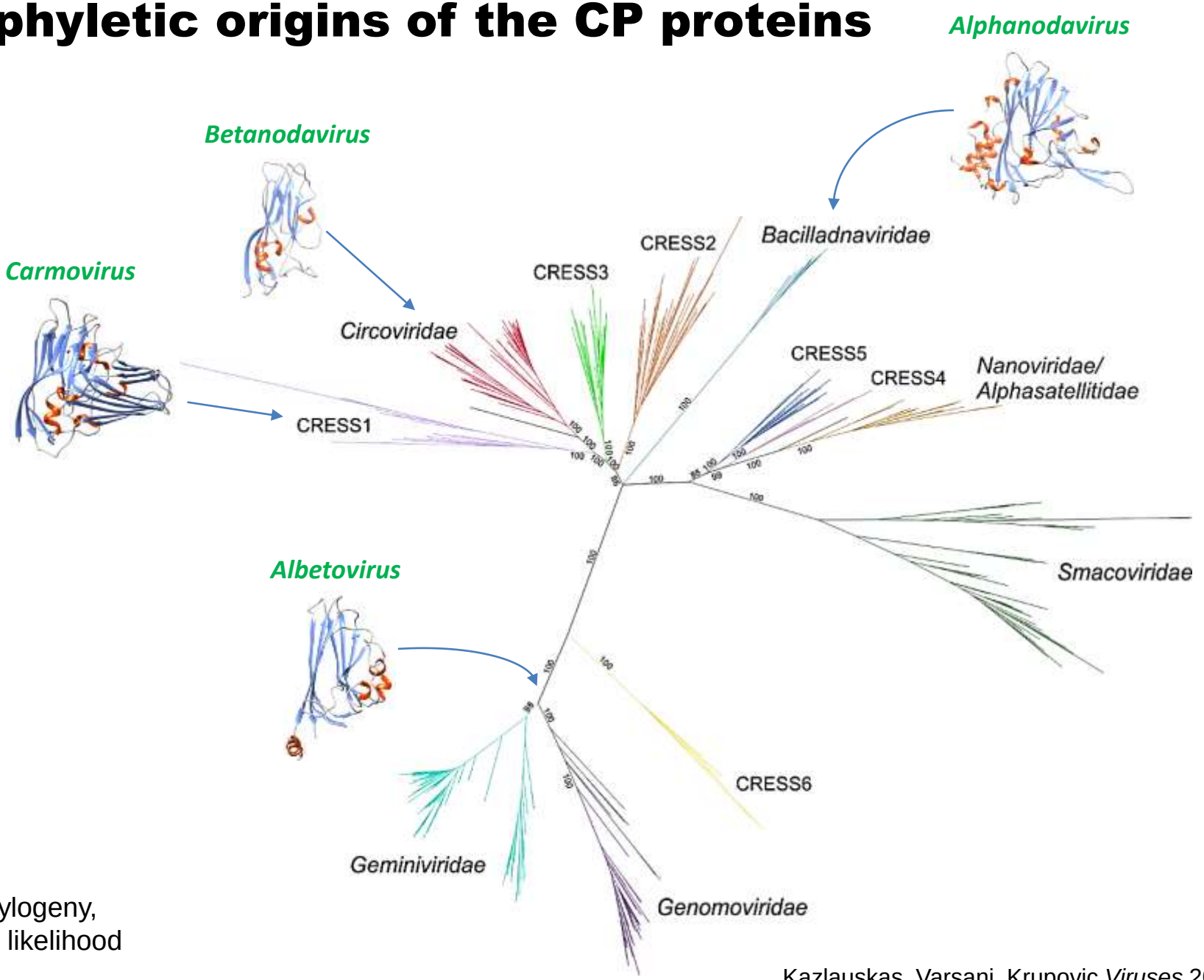

Cyclophilin

CELLULAR PROTEINS

Krupovic & Koonin, *PNAS* 2017

ORIGINS OF VIRAL REPLICATION PROTEINS

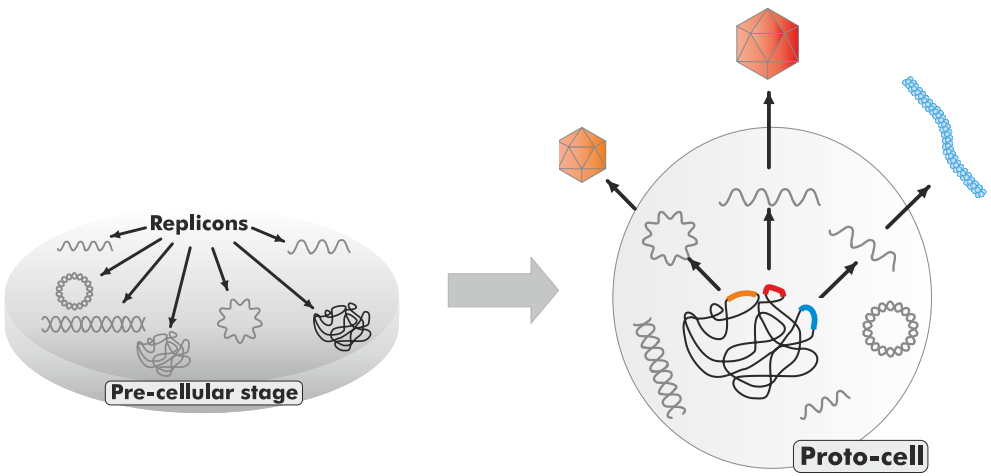
Polyphyletic origins of the CP proteins

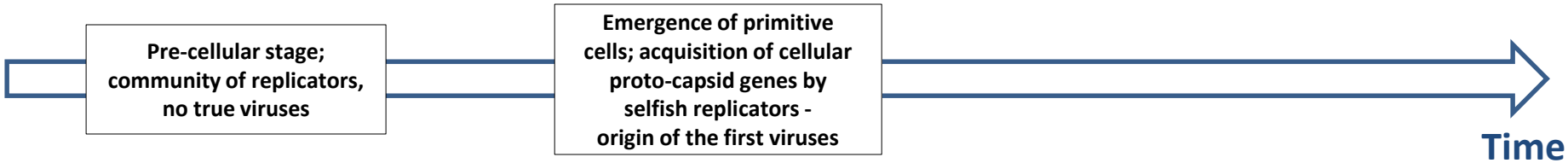


RCRE phylogeny,
maximum likelihood

Kazlauskas, Varsani, Krupovic *Viruses* 2018

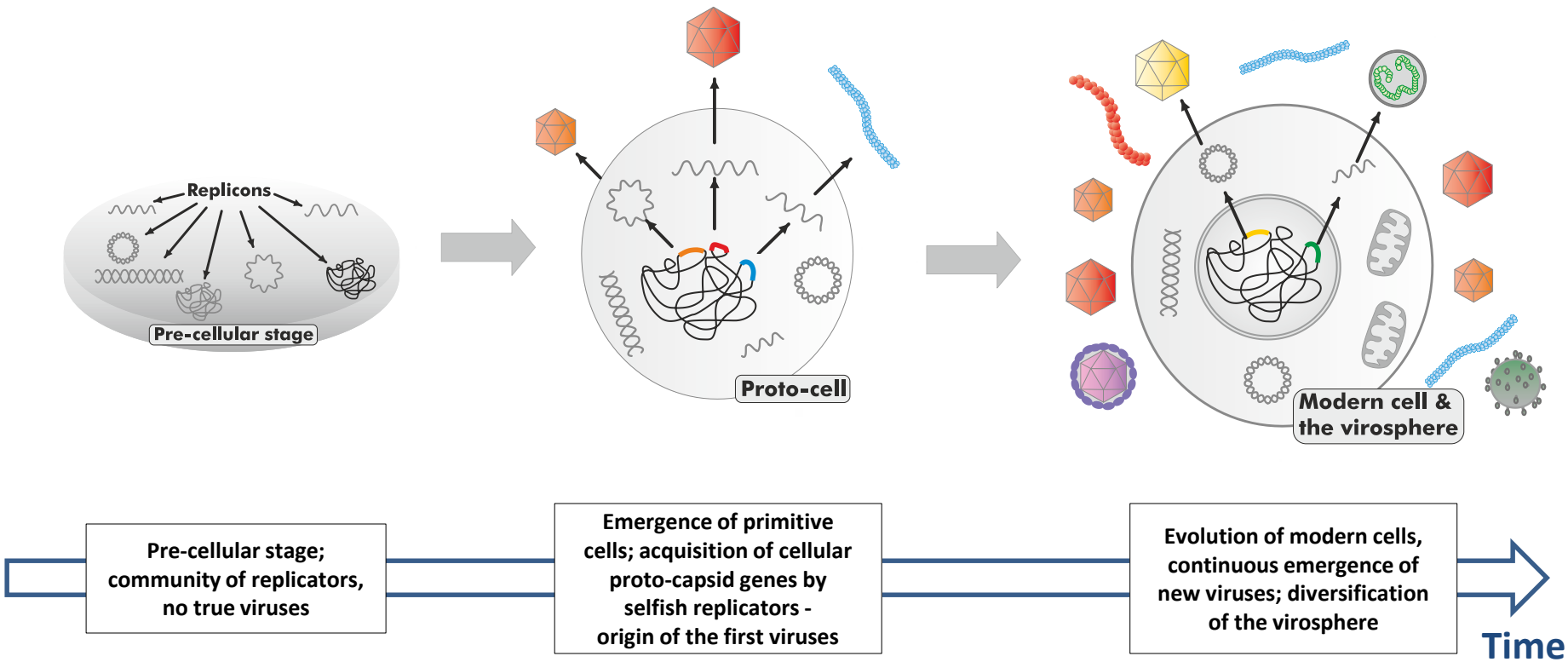
Grand scenario for the origin of viruses





Krupovic & Koonin, *PNAS* 2017

Grand scenario for the origin of viruses



Krupovic & Koonin, *PNAS* 2017

Summary

- Origin of capsids/virions is integral to the origin of true viruses;
- The diversity of viral capsid proteins is great but not limitless:
  - 7 folds cover 64% of viral families,
  - 20% of viral families are structurally uncharacterized;
- Many of the viral structural proteins could evolve from cellular ancestors;
- Evolution of the structural and replication modules is often disconnected (shuffling by recombination);
- Virions have likely evolved on multiple occasions at different stages of evolution;
- Complete picture emerges only through exploration of the evolutionary history of both structural and replication modules.

ACKNOWLEDGEMENTS

Eugene Koonin
Jaime Iranzo
Yuri Wolf

Valerian Dolja

Arvind Varsani

Darius Kazlauskas

FUNDING