



Potexvirus detection, taxonomy, and transmission as revealed by phylogenomics of cacti



Alexa Tyszka, Karolis Ramanauskas, and Boris Igić

Department of Biological Sciences, University of Illinois at Chicago

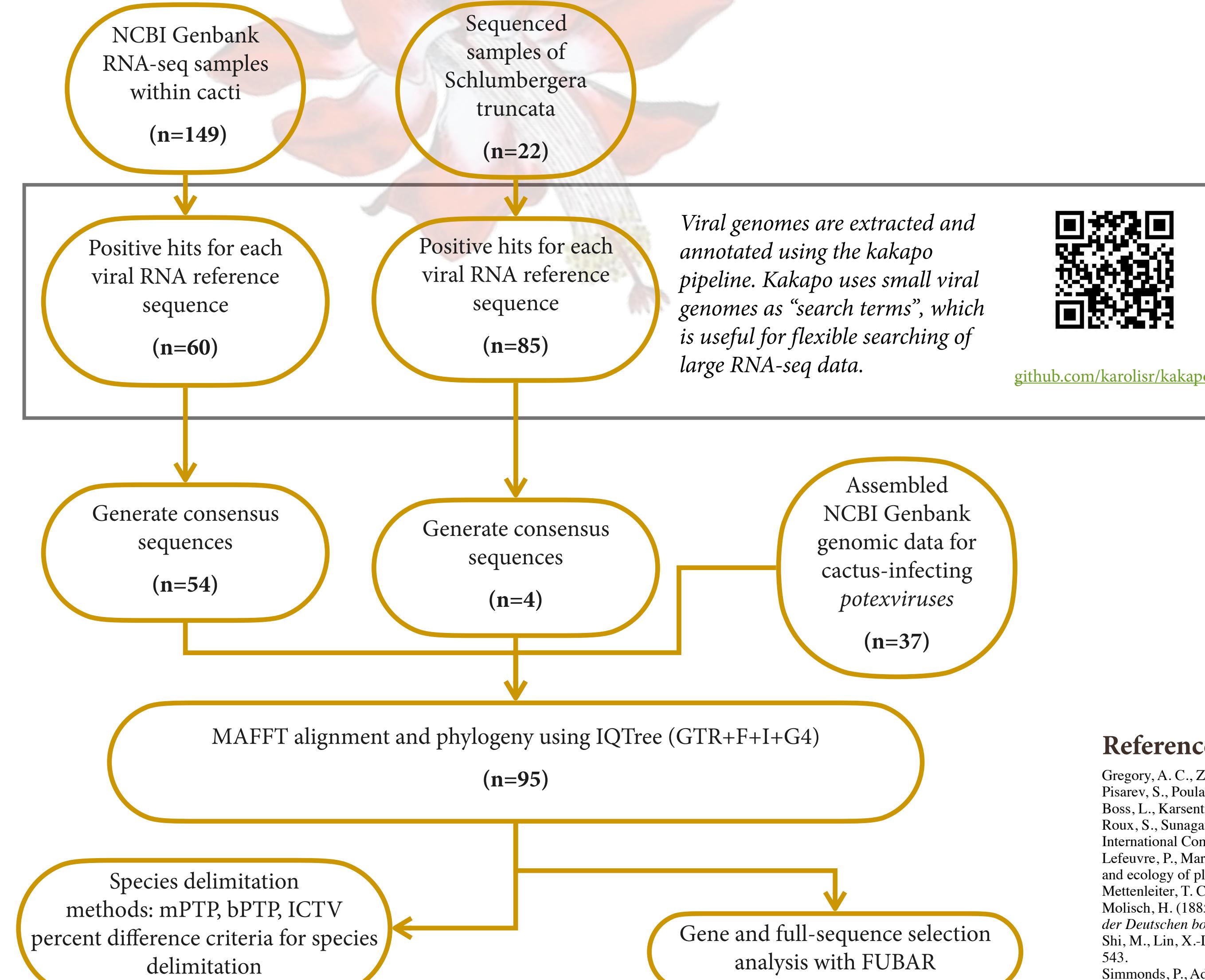
BACKGROUND

Molisch's (1885) discovery of "protein bodies" on several species of cacti was one of the first documented descriptions of viruses. For nearly a century, subsequent comparative study of viruses remained limited to direct observational data of gross morphology, augmented with clever experimental approaches, such as filtration and inoculation (Mettenleiter, 2017). A transformative advancement in virology—and all of biology—has been the advent of massively parallel DNA and RNA sequencing. The rapidly improving sequencing tools enable rapid identification of organisms from seemingly any sampled surface of the Earth. One common thread is that virtually every macro-organism genome study uncovers a micro-organismal metagenome, composed of both targeted host sequences and those from myriad co-existing organisms. Metagenomic studies have yielded an enormous number of genomes and have vastly expanded the global virome (Gregory et al., 2019; Lefevre et al., 2019; Shi et al., 2016). The unprecedented amount of data resulting from metagenomic studies has also caused significant policy changes and revisions by the International Committee on Taxonomy of Viruses (ICTV) policy (International Committee on Taxonomy of Viruses Executive Committee, 2020; Simmonds et al., 2017), but nearly all viruses remain named by their original description of host, location, and/or symptoms.

In this study, we:

- interrogate cactus samples for the presence of a common pathogen, *Cactus Virus X*,
- estimate their taxonomic and phylogenetic relationships, and
- infer modes of infection.

METHODS



References:

- Gregory, A. C., Zayed, A. A., Concepcion-Neto, N., Temperton, B., Boldic, B., Alberti, A., Ardyna, M., Arkhipova, K., Carmichael, M., Crouaud, C., Dimier, C., Dominguez-Huerta, G., Ferland, J., Kandul, S., Liu, Y., Moree, C., Pesant, S., Picheral, M., Ross, L., Karsenti, E., Not, F., Ogata, H., Pesant, S., Poulin, N., Raes, J., Sarlet, C., Speich, S., Stiemmann, L., Sullivan, M. B., Sunagawa, S., Wincker, P., Babin, M., Bowler, C., Culley, A. I., de Vargas, C., Dutih, B. E., Juticane, D., Karp-Boss, L., Roux, S., Sunagawa, S., Wincker, P., and Sullivan, M. B. (2019). Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell*, 177(5):109-112.e14.
- Lefevre, P., Martin, D. P., Elena, S. F., Shepherd, D. N., Rounique, P., and Varsani, A. (2019). Evolution and diversification of plant viruses. *Nature Reviews Microbiology*, 17(10):632-644.
- Mettenleiter, T. C. (2017). The First "Virus Hunter". *Archives in Virus Research*, 99:1-16.
- Molisch, H. (1885). Über merkwürdig geformte Proteinkörper in den Zweigen des Epiphyllum. *Berichte der Deutschen botanischen Gesellschaft in Berlin*, 3:195-202.
- Si, M., Lin, X.-D., Tian, J.-H., Chen, L.-J., Chen, Li, C.-X., Qin, X.-C., Li, J., Cao, J.-P., Eden, J.-S., Buchmann, J., Wang, W., Xu, J., Holmes, E. C., and Zhang, Y.-Z. (2016). Redefining the invertebrate RNA virosphere. *Nature*, 540(7634):539-543.
- Simmonds, P., Adams, M. J., Benkő, M., Breitbarth, M., Brister, J. R., Carstens, E. B., Davison, A. J., Delwart, E., Gorbalenya, A. E., Harrach, B., et al. (2017). Virus taxonomy in the age of metagenomics. *Nature Reviews Microbiology*, 15(3):161-168.

RESULTS

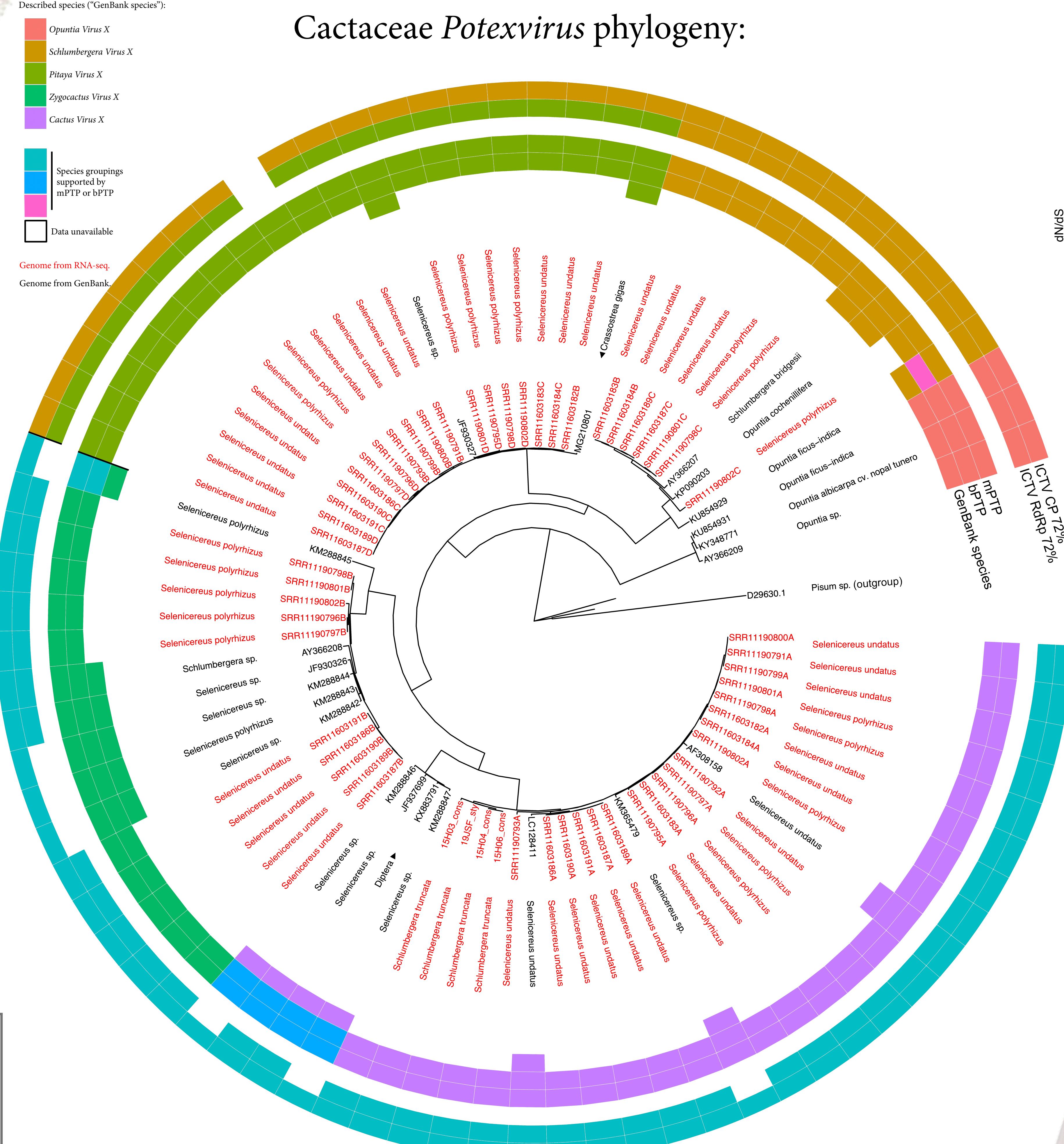


Figure 1. An updated cactus-infecting potexvirus phylogeny reflects distinct viral groupings which opportunistically and often asymptotically infect host plants. Relationships were inferred with a Maximum-Likelihood (GTR+F+I+G4) phylogenetic tree assembled from 94 virus sequences representing previously published full and partial cactus-infecting potexvirus sequences in addition to viral consensus sequences assembled by this study. Branch tips are labelled with International Nucleotide Sequence Database Collaboration (INSDC) accessions or sample unique identifiers and also labelled with the corresponding host plant. The occasional non-plant host is marked with ▶.

Positive selection across these genomes:

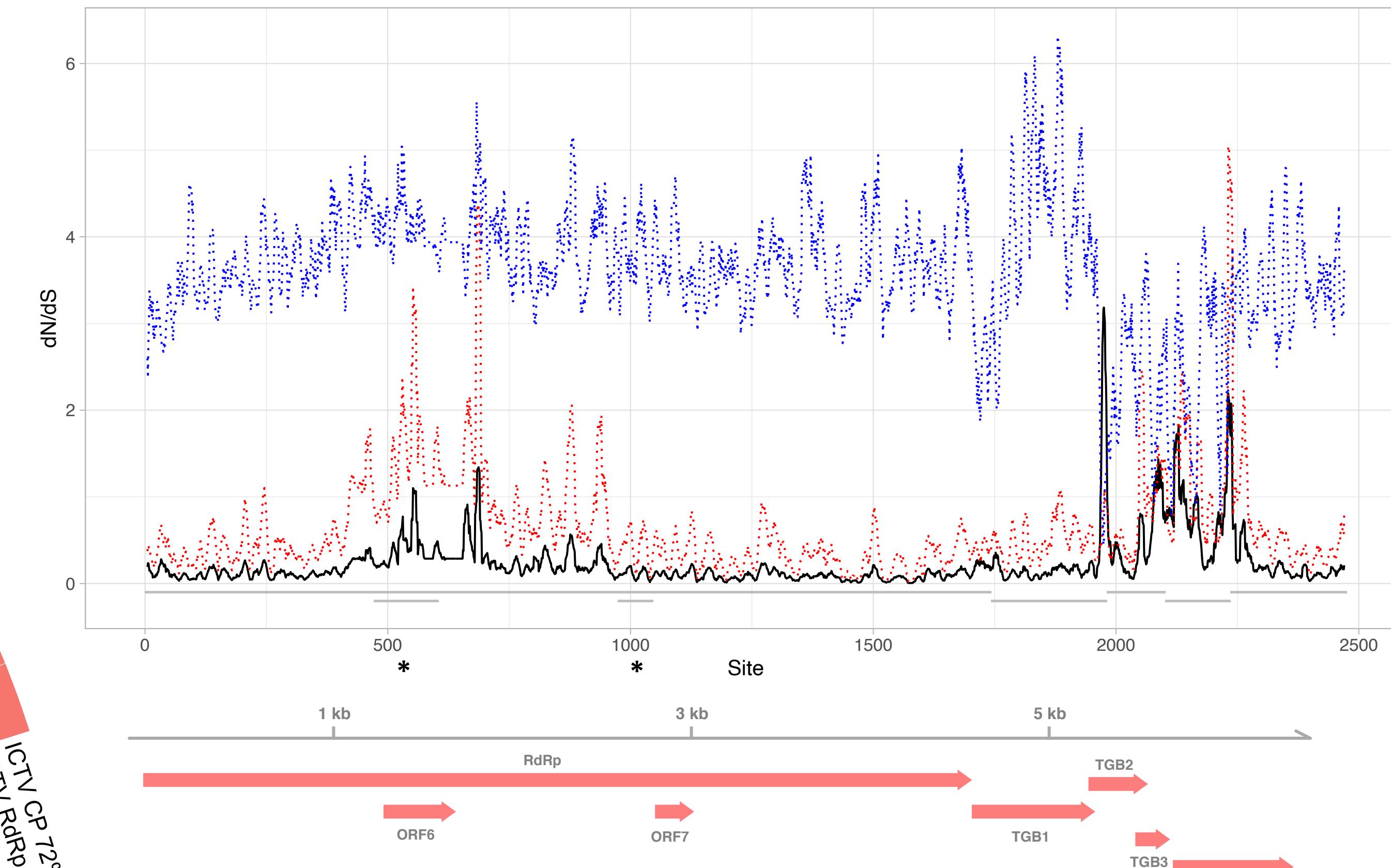


Figure 2. dN/dS as calculated over a ten-site rolling mean. dN is shown in red, dS is represented in blue, and dN/dS is represented in black. dN and dS were calculated for 95 genome sets using Fast Unconstrained Bayesian Approximation (FUBAR). Before analysis, genes were separated and stop codons were removed, so partially overlapping regions are represented twice. Note: asterisk * indicates ORF 6 and 7, which are protein coding regions located within the RdRp ORF.

DISCUSSION

We detected 58 viral genomes in 171 plant samples that were phylogenetically related to the currently described species within the genus *Potexvirus*: *Cactus Virus X*, *Schlumbergera Virus X*, *Pitaya Virus X*, and *Zygoactus Virus X*. There may be as many as eight species in this group.

Despite broad agreement among taxonomic delineation methods, there are some important differences. Using ICTV % identity criteria, GenBank taxonomy, mPTP, and bPTP, we find 4-5, 5, 7, and 8 species, respectively. We found that species as named in GenBank submissions sometimes differ from ICTV taxonomic guidance for delineation. This is particularly evident in the case of *Pitaya Virus X*, where ICTV species cutoffs for the CP and RdRp proteins disagree on species identity.

Analyses based on bPTP/mPTP also differ in their species delineation from both the designated GenBank species and the ICTV cutoffs. According to our bPTP/mPTP analysis, *Cactus Virus X* is comprised of two currently unrecognized cryptic species that appear to share a common ancestor. bPTP/mPTP analyses also recognize similar species divisions within *Zygoactus Virus X* and *Schlumbergera Virus X*.

These sequence-based analyses are on their own not entirely sufficient for a confident determination of taxonomy. For example, taxonomic descriptions may also be determined based on host specificity, but we only possess such data for a small subset. Moreover, a virus named for its first known host may not reflect the evolutionary history of the virus accurately. The natural host of any cactus-infecting potexvirus is yet unknown and may have been transmitted to another host in a glasshouse.

We find elevated dN/dS ratios, which implies positive selection, in some regions of the genome. In regions of overlap, this measure is invalid, but yields interesting results for further study. In CP, the elevated dN/dS may reflect interaction with plant immunity.

Finally, we find *Cactus Virus X* in pollen and pistil tissues, which may indicate sexual transmission of the virus.