The saddest title is no title at all

Alexa Tyszka[1], Karolis Ramanauskas[1,2] and Boris Igić[1,3,4]

[1]Department of Biological Sciences
University of Illinois at Chicago
840 West Taylor St. MC067
Chicago, IL 60607, U.S.A.

[2]Email: <kraman2@uic.edu>
[3]Email: <boris@uic.edu>
[4]Corresponding author.

Abstract

Potexvirus species members are positive-sense single-stranded RNA viruses known to infect many species of flowering plants, including *Cactaceae*. The taxonomic naming schemes often employ outdated plant name synonyms, complicating taxonomic assignments. Also, the source of infections in cultivated plants is unclear, as is the distribution of and significance of infections in wild species of cacti. The lack of clarity is partly related to low sampling across the family. Here, we report results of original RNA-seq experiments and archived sequence deposits, aimed at detecting *Potexviruses* in cacti, assembling whole genomes, estimating their phylogenetic relationships, and delimiting viral species. The data suggests novel modes of transmission, based on expression analyses across tissues, particularly pollen. We also perform molecular evolutionary analyses to detect genomic regions under different modes of selection. Finally, we examine and discuss the implications of our analyses for the taxonomy of *Potexviruses* across cacti.

Introduction

The plant viriome represents a fundamentally complex evolutionary interaction between eukaryotic host and viral vector (Delwart 2007). Plant viruses were the earliest characterized viruses, beginning with Mayer's publication on his discovery of *Tobacco Mosaic Virus* in 1886 on tobacco plants (Mayer 1886) which followed Molisch's 1885 discovery of "protein bodies" on *Schlumbergera truncata* (previously *Epiphylum truncatumin*) (Molisch 1885). They may deserve to be called the first true purified viruses, but Molisch's "Proteinkörper" are absent from many reviews of historical virology (such as Lecoq (2001) and Lefeuvre et al. (2019)). Perhaps the most recent advancement in virology has been the development of faster, cheaper, higher-throughput environmental metagenomic techniques which have advanced many facets of evolutionary biology (Delwart 2007; Lefeuvre et al. 2019; Schulz et al. 2017). It has become evident through these discoveries that the greenhouse-raised and lab-grown organisms commonly analyzed in experiments actually represent a small fraction of living diversity. Metagenomic studies aim to sample hundreds of thousands of genomes and have vastly expanded both the cellular tree of life (Schulz et al. 2017; Hug et al. 2016) and the viral tree of life (Gregory et al. 2019; Lefeuvre et al. 2019; Shi et al. 2016). Viruses display impressive morphological diversity and adaptations, many of which allow them to infect plants (Delwart 2007; Lefeuvre et al. 2019). Metagenomic analysis, particularly understudied or non-crop plants, have both enriched our genetic knowledge of plants and uncovered novel insights on viral evolution, adaptation, and transmission (citations needed). A careful study of the plant viriome provides a view into underlying biological realities that are not currently understood.

The International Committee on Taxonomy of Viruses (ICTV) presently advises the taxonomy and approval of virus nomenclature (Simmonds et al. 2017; Lefkowitz et al. 2018; on Taxonomy of Viruses Executive Committee et al. 2020). The massive amounts of data resulting from metagenomic studies have caused significant revisions in ICTV policy (on Taxonomy of Viruses Executive Committee et al. 2020; Simmonds et al. 2017), but many viruses remain named by their host, location, or symptoms, all of which may cause confusion due to their overlap with other viruses. The Baltimore classification system attempts to standardize viral classification by intrinsic morphological characteristics of a virus' replication machinery and has been integrated into the ICTV guidelines to better reflect viral evolutionary relationships (on Taxonomy of Viruses Executive Committee et al. 2020). The study of plant viruses particularly suffers from poor nomenclature due to the practice of naming a virus after a first discovered host which is subject to reclassification or renaming. The term "plant virus" in itself is problematic since there is strong evidence to suggest that viruses frequently spill over from fungal or invertebrate hosts (Lefeuvre et al. 2019). Additionally, many plant viruses that infect agriculturally important species are named using the common name of a plant, which carries its own problems, for example: *Pitaya Virus X* is named for the common name "Pitaya" which can refer to as many as thirty-one species within the genus *Selenicereus* (Korotkova et al. 2017; Guerrero et al. 2019; Le Bellec and Vaillant 2011). To complicate matters further, one virus may infect many hosts, and one host may contain many viruses. A single-stranded RNA virus has a faster rate of evolution than a host plant, and a different mode of reproduction, making a direct assignment of viruses and their hosts difficult (citation needed). There is no guarantee that viral evolution and speciation follows linearly behind plant evolution and speciation—especially due to viral host-switching. These problems persist throughout the genus *Potexvirus* and are especially prominent in cactus-infecting *Potexvirus* species. We suggest a phylogeny-based approach to remedy some prominent taxonomic issues within this specific clade that cause naming inconsistencies.

The species *Cactus Virus X*, *Zygocactus Virus X*, *Schlumbergera Virus X*, *Pitaya Virus X*, and *Opuntia Virus X* are all *Potexviruses* (family Alphaflexiviridae) that are grouped broadly by their infections of certain cacti: *Selenicereus undatus* and *S. polyrhizus* (Li et al. 2015; Peng et al. 2016); *Opuntia spp.* especially *O. tuna* (Koenig et al. 2004; Duarte et al. 2008) and *O. monacantha* (Attathom et al. 1978) Sammons 1961 Duarte 2008; *Schlumbergera* (previously *Zygocactus*) *truncata* and *S. bridgesii* (Duarte et al. 2008; Koenig et al. 2004), *Parodia* (previously *Notocactus*) *leninghausii* (Park et al. 2018), *Echinopsis chamaecereus f.*

cristata, *E. pectinatus f. cristata*, *E. jusbertii*, and *E. macrogona* (Maliarenko and Mudrak 2013); *Mammillaria elongata f. cristata* (Maliarenko and Mudrak 2013); and multiple other species within many genera in the family Cactaceae (Evallo et al. 2021).

Of these viruses, only *Cactus Virus X* (CVX) has been reported on wild *Ferocactus cylindraceus* (previously *Ferrocactus acanthodes*) (Attathom et al. 1978) although this report predates DNA records confirming the viral identity. Additionally, although they are originally found on cacti, the viruses are frequently manipulated with serological experiments and have been found to produce lesions (which indicate infection) on: *Chenopodium murale L.* (Maliarenko and Mudrak 2013) and C. quinoa (Attathom 1978; Attathom et al. 1978; Brandes and Bercks 1963); Nicotiana alata Link el. Otto (Maliarenko and Mudrak 2013); Four species of Amaranthaceae (Attathom 1978); Escobaria vivipara (Attathom 1978); and other Cactaceae (Attathom 1978).

All cactus-infecting *Potexviruses* consist of roughly 6,600 bp of positive-sense single-stranded RNA. They have similar rod-shaped filamentous virions and share the same division of five primary open reading frames (ORFs): Replicase (Rep), Triple gene block (TGB), Coat protein (CP), coded in the 5' direction as well as two smaller overlapping ORFs coded in the 3' direction: ORF6 and ORF7 (Evallo et al. 2021; Liou et al. 2004; Martelli et al. 2007). They are closely related to other *Potexviruses* such as *Alternantha Mosaic Virus* and *Papaya Mosaic Virus* (Martelli et al. 2007; Park et al. 2018; Liou et al. 2004). These viruses produce a wide range of symptomatic and damaging infections in cacti. Reports of symptomatic plants range from 5.5 percent of wild *Ferocactus cylindraceus* (Attathom et al. 1978) up to 44 percent of crop plants on Hainan Island, China (Peng et al. 2016). However, many infected plants do not show external signs of viral infection (Liou et al. 2004; Bos 1977). The most commonly recognized symptoms of disease are mosaic, mottling, stunted growth and distortion (Maliarenko and Mudrak 2013; Peng et al. 2016; Attathom et al. 1978).

It is yet unclear what the method of transmission from infected plant to new host is. Some reports specify that cactus-infecting *Potexviruses* can only be transmitted through grafting (Duarte et al. 2008; Martelli et al. 2007) but most agree that transmission can occur through other mechanical contact such as sap inoculation (Liou et al. 2004; Maliarenko and Mudrak 2013; Park et al. 2018) and external tissue contact. Grafting is a primary means of propagation among crop cacti (Park et al. 2018), and *Selenicereus* is a commonly chosen graft stock. However, there are reports of other members within the family *Alphaflexiviridae* transmitting via insect and seed vectors (Martelli et al. 2007), and pre-DNA studies tentatively suggest that *CVX* may transmit via pollen in the wild (Attathom et al. 1978).

Knowledge about cactus-infecting *Potexviruses* contributes to a growing yet biased study of plant viruses. The evolutionary history of these viruses is obscured due to human-assisted dispersal, grafting, and cultivation, which parallels the disproportionate sampling representation of plants raised in greenhouses or for agricultural production. However, *Cactus Virus X* and associated viruses seem restricted to cactaceous hosts for unknown reasons—every sample of CVX or CVX-related viruses has come from cacti. A wild origin has not been definitively identified, and the few studies that have investigated wild *Potexviruses* of cacti predate DNA methods. Recent sequencing efforts have revealed multiple inconsistent virus-host pairs on cacti. Although many metagenomic studies capture environmental genetic information that allows for virus identification, these may be biased due to tissue type and expression rates of viruses (Lacroix et al. 2016). The pursuit of wild cactus-infecting *Potexviruses* serves to expand our evolutionary knowledge of viral evolution, host selection, and transmission mechanics. The relationships of the virus can be investigated with a thorough phylogenetic approach, using available virus samples. In this study we present the largest to date phylogeny of cactus-infecting *Potexviruses*. We attempt to use this expanded phylogeny to answer relevant questions about Potexvirus evolutionary relationships as well as revisiting the utility of decades-old taxonomy in current virus research.

Materials and Methods

Study Organism

RNA Sequencing

Tissues were collected and immediately submerged in 1.5 ml of RNA*later*™ solution (Invitrogen). Submerged samples were generally held at room temperature for thirty minutes and then stored at -80°C. Approximately 100 mg of tissue was ground with a cooled mortar in 1.5 ml tubes. Total RNA was isolated using Total RNA Mini Kit (Plant kit; IBI Scientific, Cat. No. IB47341) following manufacturer's instructions. We assessed RNA concentration and purity with a NanoDrop™ Lite Spectrophotometer (Thermo Scientific). The twenty-three samples used in this study were sequenced as part of a larger sequencing effort which consisted of four separate sequencing runs and included additional samples from other plant species. Sequencing libraries were prepared using the KAPA Stranded mRNA-Seq (Roche), and these libraries were sequenced on a single lane of Illumina HiSeq 4000 or Illumina NovaSeq 6000 platform (paired-end 150 bp reads) at the Duke University Center for Genomic and Computational Biology. The number of resulting read pairs (for the twenty-three samples presented here) ranged from 4,148,932 to 9,618,084 with a median of 6,363,556 and average of 6,293,553 (Table S1).

RNAseq Assemblies

Raw paired-end Illumina reads were first processed using Rcorrector v1.0.4 to correct for random sequencing errors. Then, reads were trimmed with Trimmomatic v0.39 to remove any read containing bases with Phred scores lower than 20, low quality reads less than 50 bp long, and any adapter or other Illumina-specific sequences that were still present. The remaining reads were filtered with Kraken 2 to remove Small and Large Subunit ribosomal RNA (SILVA database) and contaminating reads (minikraken2_v2 database). Additionally, we used custom-built databases, derived from RefSeq libraries: UniVec_Core, viral, mitochondrion, plastid, plasmid, archaea, bacteria, protozoa, human, and fungi to minimize the number of contaminating and non-nuclear reads. Only paired reads were used for transcriptome assemblies. *Schlumbergera truncata* filtered reads were combined across all samples into a single RNA-seq data set. We conducted a *de novo* transcriptome assembly using Trinity v2.8.5 to generate a single reference transcriptome assembly for *Schlumbergera truncata*.

NCBI Data Collection and Compilation

We collected publicly available genomes, complete proteins, gene annotations, and available metadata from Potexviruses (NCBI:txid12176) (NCBI: www.ncbi.nlm.nih.gov/, accession numbers provided in Supplemental Data). The untranslated regions (UTRs) were trimmed from the sequences to provide consistency.

We also searched the NCBI Sequence Read Archive (SRA) database (www.ncbi.nlm.nih.gov/sra) for RNA-sequencing (RNA-seq) data within Caryophyllales (NCBI:txid3524) that had been sequenced using the Illumina library sequencing platform. For each identified SRA run accession (SRR), any viral RNA that matched sample cactus-infecting Potexvirus RNA (accession numbers provided in Supplemental Data) was identified, extracted, and assembled using the kakapo 0.7.3-dev pipeline (http://flightless.one) with Kraken2 viral filters disabled. The .sam files produced through kakapo were loaded through Geneious 11.1.5 along with the Schlumbergera reads. These sequences were annotated using the Geneious 11.1.5 "Find ORFs" function.

The complete dataset comprises: 37 existing Potexvirus genomes and proteins, 4 new viral sequences located within original Schlumbergera truncata RNA-seq data, and 52 viral sequences found within NCBI Caryophyllales RNA-seq data.

Sequence Alignment and Phylogenetic Analyses

Sequence alignments were performed through MAFFT v7.429 (Katoh (2002)) using the full dataset. The aligned sequences were divided by ORF using the annotations to produce five partial sequence alignments corresponding to each ORF to accompany the full-sequence alignment. The individual proteins were exported to .FASTA files, then gaps at the start of the sequence and stop codons were removed manually. Phylogenetic relationships and bootstrap values were inferred using IQtree v1.6.12 (Nguyen et al. (2015)), ModelFinder (Kalyaanamoorthy et al. (2017)), and UFBoot (Hoang et al. (2018)) for both the individual gene/protein alignments and the full sequence alignment. Trees were visualized in R version 4.0.3 using ggtree v2.4.2 (Yu et al. (2017)). Host information was obtained through reported metadata and mapped onto the phylogeny. Species groupings were determined using the existing species boundaries when compared to the phylogenetic branch lengths within the Potexvirus genus. This was generally consistent with most recent branch lengths over 0.1 subs/site and this value was therefore used as a cutoff. Pairwise distance analysis was conducted on the sequence alignments in R using the ape v5.5 dist.dna() function with a raw model. For each defined clade, nonzero pairwise distances between each possible combination of tips was averaged. Expanded phylogenetic trees and individual gene/protein trees are available in the Supplementary Data. Pistils (without ovaries), pollen, leaf, and root tissues were collected and immediately submerged in 1.5 ml of RNA*later*™ solution (Invitrogen). Submerged samples were held at room temperature for thirty minutes and then moved to a -80 C freezer for storage. Approximately 100 mg of tissue was ground to a fine powder in 1.5 ml tubes submerged in liquid nitrogen. Total RNA was isolated using Total RNA Mini Kit (Plant kit; IBI Scientific, Cat. No. IB47341) following manufacturer's instructions. We assessed RNA concentration and purity with a NanoDrop™ Lite Spectrophotometer (Thermo Scientific). The XX samples used in this study were sequenced as part of a larger sequencing effort which consisted of XXX separate sequencing runs and included additional samples from other plant species. Sequencing libraries were prepared using the KAPA Stranded mRNA-Seq (Roche), and these libraries were sequenced on a single lane of Illumina HiSeq 4000 or Illumina NovaSeq 6000 platform (paired-end 150 bp reads) at the Duke University Center for Genomic and Computational Biology. The number of resulting read pairs (for the XX samples presented here) ranged from X,XXX,XXX to X,XXX,XXX with a median of X,XXX,XXX and average of X,XXX,XXX (Table S1).

Results and Discussion

Characterization

The collection of *Schlumbergera* samples and thorough investigation of previously published data on Cactaceae resulted in the discovery of XX new virus lineages. The genome sizes of X,XXX - X,XXX bp were all consistent with published genomic *Potexvirus* data, which range from X.Xk - X.Xk bp. The newly discovered viruses from *Schlumbergera* were generated as consensus sequences of XX individual reads which reliably recovered XX percent of the published *Cactus Virus X* genome. For the newly discovered viruses with *Selenicereus* hosts, the XXX sample reads recovered XX percent of the *CVX* genome. All of the publicly available new viral lineages were discovered on *Selenicereus* hosts from SRA XXXXX. We annotated the open reading frames of the viruses to recover all seven Potexvirus proteins.

The *Schlumbergera*-infecting viruses were found in high amounts on pollen and style tissue. The viral loads of each *Selenicereus* sample that was found to have viruses ranged from XX-XX percent of all reads. This was a relatively high recovery rate, and in fact in some pollen tissues there existed more viral reads (XX,XXX reads) than *Schlumbergera* reads (XX,XXX reads).

Distribution of Genetic Distances

The extremely similar and well-clustered newly discovered viruses displayed very low diversity within the clusters. Therefore, these additions to the *Potexvirus* family tree do not drastically alter the tree structure. Since each sampled cactus in SRA XXXXX was located close to other sampled cacti, this low diversity

potentially represents a first example of background mutation among viruses incurred due to host infection. The average nonzero pairwise distance between the included subset of related *Potexviruses* was 0.256 (maximum = 0.492). When the outgroup (including *Plantago asiatica MV, Alternantha MV, Papaya MV,* etc.) was excluded from pairwise analysis, the average nonzero pairwise distance value was 0.177 (maximum = 0.326). When these cactus-infecting *Potexviruses* were subdivided into six groups of relatively recent diversification, the average nonzero pairwise distance for full-genome sequences among groups was always above 0.015. The newly discovered *Schlumbergera*-infecting viruses displayed XX percent similarity, and the *Selenicereus*-infecting viruses from existing cactus samples displayed XX percent similarity. The nearly identical nature of these viruses combined with their host group's physical closeness likely represents a case of th For the genes RNA-dependent RNA polymerase (RdRp) and Coat protein (CP), which the ICTV recommends be analyzed for species delimitation, the average nonzero pairwise distance was always above 0.02 (Figure 2). This correlates to roughly greater than 97.5 and 98 nucleotide identity.

The ICTV guidelines for *Potexviruses* indicate that less than 72 percent nucleotide sequence identity (or 80 percent amino acid identity) between the CP or Rep genes demarcates separate viral species. Because we compare closely related *Potexviruses*, it might be expected that members of the same putative species would have higher than 72 percent nucleotide identity and members of different putative species would have lower than 72 percent sequence identity. However, the low pairwise distances between *Potexviruses* cause very few cactus-infecting *Potexviruses* to be demarcated as separate species, even when only considering previously described species compared to each other. Examples here.

Phylogenetic Relationships

A well-supported phylogenetic tree was recovered including closely related non-cactus-infecting *Potexviruses*. The phylogenetic tree (Figure 1) places the new viral sequences from *Schlumbergera* and *Selenicereus* near existing viral species within *Potexvirus* (Figure 1). +Phylogenetic analysis recovered defined monophyletic groups corresponding to five or six major groups of cactus-infecting *Potexviruses*, with *Cactus Virus X* displaying two branching subgroups. *Opuntia Virus X* was the basal clade to the rest of the cactus-infecting *Potexviruses*. The *S. truncata* samples located within the *Cactus Virus X* clade appear to represent the first known discovery of *Cactus Virus X* on *Schlumbergera*.

Reported host genera are presented for each viral sample. The reported taxonomy of each existing sample aligns well with the tree structure, but this monophyly is not recovered for hosts. Putative viral species appear to infect as few as one genera, in the case of *Opuntia Virus X*, or as many as three in the case of *Schlumbergera Virus X*. Further, the three genera are evolutionarily distinct and polyphyletic in phylogenetic analyses of *Cactaceae*. This raises questions about a *Potexvirus*'s ability to switch hosts. All three genera that are reported hosts of *Schlumbergera Virus X* are ornamental crops: *Selenicereus, Schlumbergera*, and *Opuntia*. It is possible that extended greenhouse contact between the three genera has resulted in viral spillover of *Schlumbergera Virus X*. However, it is also important to note that the present viral taxonomy is not infallible. The putative species known as *Schlumbergera Virus X* may have been initially - incidentally - found on a spillover host rather than a member of an actual viral circulating population. Unfortunately there is no way to know the "true host" of any virus that possesses the mechanisms to infect multiple Cactaceaeous genera. Further testing of known host genera are necessary, as well as metagenomic sequencing of closely related genera to monitor potential opportunistic spillover.

Taxonomic descriptions must also be analyzed for accuracy and reflection of actual viral activity. A virus that is named for its first known host may not represent the evolutionary history of the virus. Although phylogenetic analysis is likely to produce monophyletic clades of viruses that have each been named in succession for the first known member, this produces inconsistent and confusing names. *Zygocactus Virus X* is a clear example: Although the name *Zygocactus* is outdated in reference to the host genera, now classified as *Schlumbergera*, the viral name remains. This is exacerbated by the presence of a separate viral species, *Schlumbergera Virus X*. The ICTV strongly opposes unnecessary name changes, but it is unclear what should

be done with outdated and confusing names that nonetheless describe a real clade of viruses. Mixed infections and the lack of a one-to-one correlation between virus and host complicate naming endeavors further, and more sampling will surely uncover novel hosts and mixed infections.

Recombination and selection analysis

Recombination events, how they were detected. Selection analysis goes here.

Grouping

The new species fall into already existing species groups. Pairwise distances between species were calculated for six groups, with a value of 0 indicating identical sequences and a value of 1 indicating completely dissimilar sequences.

Concluding Remarks

X new viruses were included as part of this study, which represents a multifaceted approach to viral discovery using metagenomic techniques for already available public data as well as newly collected data. Mixed infections, co-infection dynamics. Viral abundance, viral loads, tissue type, and diversity. Taxonomic recommendations.

Acknowledgments

Acknowledgments text.

References

Attathom, S., 1978. Identification and Characterization of a Potexvirus from California Barrel Cactus. Phytopathology 68:1401. URL http://www.apsnet.org/publications/phytopathology/backissues/Documents/1978Abstracts/Phyto68_1401.htm.

Attathom, S., L. G. Weathers, and D. J. Gumpf, 1978. Occurrence and distribution of a virus induced disease of barrel cactus in California. Plant disease reporter URL https://scholar.google.com/scholar_lookup?title=Occurrence+and+distribution+of+a+virus+induced+disease+of+barrel+cactus+in+California&author=Attathom%2C+S.&publication_year=1978.

Bos, L., 1977. SYMPTOMS OF VIRUS DISEASES IN PLANTS. Research Institute for Plant Protection.

Brandes, J. and R. Bercks, 1963. Untersuchungen zur Identifizierung und Klassifizierung des Kakteen-X-Virus (cactus virus X). Journal of Phytopathology 46:291–300. URL https://onlinelibrary.wiley.com/doi/10.1111/j.1439-0434.1963.tb02072.x.

Delwart, E. L., 2007. Viral metagenomics. Reviews in Medical Virology 17:115–131. URL https://onlinelibrary.wiley.com/doi/10.1002/rmv.532.

Duarte, L. M. L., M. A. V. Alexandre, E. B. Rivas, R. Harakava, S. R. Galleti, and M. M. Barradas, 2008. POTEXVIRUS DIVERSITY IN CACTACEAE FROM SÃO PAULO STATE IN BRAZIL. Journal of Plant Pathology P. 9.

Evallo, E., J. D. Taguiam, and M. A. Balendres, 2021. A brief review of plant diseases caused by Cactus virus X. Crop Protection 143:105566. URL https://linkinghub.elsevier.com/retrieve/pii/S0261219421000363.

Gregory, A. C., A. A. Zayed, N. Conceição-Neto, B. Temperton, B. Bolduc, A. Alberti, M. Ardyna, K. Arkhipova, M. Carmichael, C. Cruaud, C. Dimier, G. Domínguez-Huerta, J. Ferland, S. Kandels, Y. Liu, C. Marec, S. Pesant, M. Picheral, S. Pisarev, J. Poulain, J.-É. Tremblay, D. Vik, S. G. Acinas, M. Babin, P. Bork, E. Boss, C. Bowler, G. Cochrane, C. de Vargas, M. Follows, G. Gorsky, N. Grimsley, L. Guidi, P. Hingamp, D. Iudicone, O. Jaillon, S. Kandels-Lewis, L. Karp-Boss, E. Karsenti, F. Not, H. Ogata, S. Pesant, N. Poulton, J. Raes, C. Sardet, S. Speich, L. Stemmann, M. B. Sullivan, S. Sunagawa, P. Wincker, M. Babin, C. Bowler, A. I. Culley, C. de Vargas, B. E. Dutilh, D. Iudicone, L. Karp-Boss, S. Roux, S. Sunagawa, P. Wincker, and M. B. Sullivan, 2019. Marine DNA Viral Macro- and Microdiversity from Pole to Pole. Cell 177:1109–1123.e14. URL https://www.sciencedirect.com/science/article/pii/S0092867419303411.

Guerrero, P. C., L. C. Majure, A. Cornejo-Romero, and T. Hernández-Hernández, 2019. Phylogenetic Relationships and Evolutionary Trends in the Cactus Family. Journal of Heredity 110:4–21. URL https://academic.oup.com/jhered/article/110/1/4/5205122.

Hoang, D. T., O. Chernomor, A. von Haeseler, B. Q. Minh, and L. S. Vinh, 2018. UFBoot2: Improving the Ultrafast Bootstrap Approximation. Molecular Biology and Evolution 35:518–522. URL https://academic.oup.com/mbe/article/35/2/518/4565479.

Hug, L. A., B. J. Baker, K. Anantharaman, C. T. Brown, A. J. Probst, C. J. Castelle, C. N. Butterfield, A. W. Hernsdorf, Y. Amano, K. Ise, Y. Suzuki, N. Dudek, D. A. Relman, K. M. Finstad, R. Amundson, B. C. Thomas, and J. F. Banfield, 2016. A new view of the tree of life. Nature Microbiology 1:16048. URL http://www.nature.com/articles/nmicrobiol201648.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin, 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature Methods 14:587–589. URL http://www.nature.com/articles/nmeth.4285.

Katoh, K., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Research 30:3059–3066. URL https://academic.oup.com/nar/article-lookup/doi/10.1093/nar/gkf436.

Koenig, R., H. Aust, J. Schiemann, C. W. A. Pleij, S. Loss, and W. Burgermeister, 2004. Molecular characterisation of potexviruses isolated from three different genera in the family Cactaceae. Archives of Virology 149:903–914. URL http://link.springer.com/10.1007/s00705-003-0268-5.

Korotkova, N., T. Borsch, and S. Arias, 2017. A phylogenetic framework for the Hylocereeae (Cactaceae) and implications for the circumscription of the genera. Phytotaxa 327:1. URL https://biotaxa.org/Phytotaxa/article/view/phytotaxa.327.1.1.

Lacroix, C., K. Renner, E. Cole, E. W. Seabloom, E. T. Borer, and C. M. Malmstrom, 2016. Methodological guidelines for accurate detection of viruses in wild plant species. Applied and environmental microbiology 82:1966–1975.

Le Bellec, F. and F. Vaillant, 2011. Pitahaya (pitaya) (Hylocereus spp.). Pp. 247–273e, *in* Postharvest Biology and Technology of Tropical and Subtropical Fruits. Elsevier. URL https://linkinghub.elsevier.com/retrieve/pii/B9780857090904500123.

Lecoq, H., 2001. Découverte du premier virus, le virus de la mosaïque du tabac : 1892 ou 1898 ? Comptes Rendus de l'Académie des Sciences - Series III - Sciences de la Vie 324:929–933. URL https://www.sciencedirect.com/science/article/pii/S0764446901013683.

Lefeuvre, P., D. P. Martin, S. F. Elena, D. N. Shepherd, P. Roumagnac, and A. Varsani, 2019. Evolution and ecology of plant viruses. Nature Reviews Microbiology 17:632–644. URL http://www.nature.com/articles/s41579-019-0232-3.

Lefkowitz, E. J., D. M. Dempsey, R. C. Hendrickson, R. J. Orton, S. G. Siddell, and D. B. Smith, 2018. Virus taxonomy: the database of the international committee on taxonomy of viruses (ictv). Nucleic acids research 46:D708–D717.

Li, Y.-S., C.-H. Mao, T.-Y. Kuo, and Y.-C. Chang, 2015. VIRAL DISEASES OF PITAYA AND OTHER CACTACEAE PLANTS. Improving Pitaya Production and Marketing P. 9.

Liou, M. R., Y. R. Chen, and R. F. Liou, 2004. Complete nucleotide sequence and genome organization of a Cactus virus X strain from Hylocereus undatus (Cactaceae). Archives of Virology 149:1037–1043. URL http://link.springer.com/10.1007/s00705-003-0251-1.

Maliarenko, V. M. and T. P. Mudrak, 2013. Cactus viruses in fasciated plants. Biologija 59. URL http://lmaleidykla.lt/ojs/index.php/biologija/article/view/2754.

Martelli, G. P., M. J. Adams, J. F. Kreuze, and V. V. Dolja, 2007. Family *Flexiviridae* : A Case Study in Virion and Genome Plasticity. Annual Review of Phytopathology 45:73–100. URL http://www.annualreviews.org/doi/10.1146/annurev.phyto.45.062806.094401.

Mayer, A., 1886. Über die mosaikkrankheit des tabaks. Die Landwirtschaftliche Versuchs-stationen 32:451–467.

Molisch, H., 1885. Über merkwürdige geformte Proteinkörper in den Zweigen von Epiphylium.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh, 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution 32:268–274. URL https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msu300.

Park, C. H., E. G. Song, and K. H. Ryu, 2018. Detection of Co-Infection of Notocactus leninghausii f. cristatus with Six Virus Species in South Korea. The Plant Pathology Journal 34:65–70. URL http://ppjonline.org/journal/view.php?doi=10.5423/PPJ.NT.08.2017.0187.

Peng, C., N. T. Yu, Z. W. Luo, H. Y. Fan, F. He, X. H. Li, Z. L. Zhang, and Z. X. Liu, 2016. Molecular Identification of Cactus virus X Infecting Hylocereus polyrhizus (Cactaceae) in Hainan Island, China. Plant Disease 100:1956. URL https://apsjournals.apsnet.org/doi/10.1094/PDIS-01-16-0048-PDN.

Schulz, F., E. A. Eloe-Fadrosh, R. M. Bowers, J. Jarett, T. Nielsen, N. N. Ivanova, N. C. Kyrpides, and

T. Woyke, 2017. Towards a balanced view of the bacterial tree of life. Microbiome 5:140. URL https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-017-0360-9.

Shi, M., X.-D. Lin, J.-H. Tian, L.-J. Chen, X. Chen, C.-X. Li, X.-C. Qin, J. Li, J.-P. Cao, J.-S. Eden, J. Buchmann, W. Wang, J. Xu, E. C. Holmes, and Y.-Z. Zhang, 2016. Redefining the invertebrate RNA virosphere. Nature 540:539–543. URL http://www.nature.com/articles/nature20167.

Simmonds, P., M. J. Adams, M. Benkő, M. Breitbart, J. R. Brister, E. B. Carstens, A. J. Davison, E. Delwart, A. E. Gorbalenya, B. Harrach, et al., 2017. Virus taxonomy in the age of metagenomics. Nature Reviews Microbiology 15:161–168.

on Taxonomy of Viruses Executive Committee, I. C. et al., 2020. The new scope of virus taxonomy: partitioning the virosphere into 15 hierarchical ranks. Nature Microbiology 5:668.

Yu, G., D. K. Smith, H. Zhu, Y. Guan, and T. T. Lam, 2017. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution 8:28–36. URL https://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12628.

Figures