OPEN ACCESS
phytobiomesjournal.org
Phytobiomes Journal
A Transdisciplinary Journal of Sustainable Plant Productivity

RESEARCH

Non-Native Plant Viruses Prevalent in Remnant Natural Plant Communities Harm Native Perennial Hosts

Tessa M. Shates,[1] Marco Gebiola,[2] Penglin Sun,[1] Oaksoe Aung,[1] Amani Helo,[1] Jaimie R. Kenney,[1] Carolyn M. Malmstrom,[3] and Kerry E. Mauck[1,†]

[1] Department of Entomology, University of California-Riverside, Riverside, CA 92521
[2] Department of Agricultural Sciences, University of Naples Federico II, Italy
[3] Department of Plant Biology and Program in Ecology, Evolution, and Behavior, Michigan State University, East Lansing, MI 48824

Accepted for publication 26 September 2023.

ABSTRACT

Plant viruses are ubiquitous throughout plant communities, but research on viral impacts largely focuses on crops. Little is known about how viruses influence wild plants in their native habitats. To address this gap, we examined virus interactions with wild drought-tolerant perennials in California desert natural areas encroached upon by agriculture. We used metagenomics, targeted diagnostics, and phylogenetics to assess virus diversity and clade relationships, and experiments to investigate viral influence on hosts. We focused on three herbaceous perennials (*Cucurbita foetidissima*, *C. palmata*, and *Datura wrightii*) and tested the hypothesis that these wild species accumulate virus infections typically found in crops and transmitted by polyphagous insects. We predicted that such infections might be retained across seasons and potentially impair plant performance. Virome profiling revealed a rich community of previously characterized virus species (12 total), with virus community structure varying by site and host species. The dominant viruses in the wild hosts were non-native crop pathogens, including cucurbit aphid-borne yellows virus (CABYV) and cucurbit yellow stunting disorder virus (CYSDV). Targeted testing revealed that CABYV infected as many as 88% of sampled wild *Cucurbita* individuals, with dual CABYV–CYSDV infections common in natural areas adjacent to desert agriculture. CABYV infections reduced shoot and root production in greenhouse experiments with the two wild *Cucurbita* species. Phylogenetic analyses suggest that CABYV was introduced to California multiple times from other continents. Our findings provide concerning evidence of ways in which human activities can alter virus pressure on wild plants and potentially contribute to plant decline.

Keywords: cucurbit aphid-borne yellows virus, cucurbit yellow stunting disorder virus, virome, virus community, virus ecology

Viruses are present in all habitats that support cellular life, where they are important drivers of processes that shape and structure communities. Despite this, we know very little about how plant viruses influence the ecology and resilience of the undomesticated, long-lived perennial plants that dominate most native communities (Lefeuvre et al. 2019; Malmstrom and Alexander 2016). Lack of progress in this area reflects a historic focus on one-way virus transfer from wild hosts to managed crops, motivated by the desire to reduce viral economic impacts; in this context, wild hosts have been considered primarily as potential incubators and amplifiers of pests and pathogens that might threaten crops (Jones 2014). However, as natural land cover diminishes worldwide and agriculture become more extensive, it is important to consider how crops themselves shape local and regional virus abundance and diversity, and the reciprocal effects of viral exchange from crops to wild habitats. In today's human-altered landscapes, wild plants in remnant native communities likely experience notable exposure to viral fluxes from agriculture (Bernardo et al. 2018; Hochman Adler et al. 2014), representing both endemic native viruses and introduced non-native viral species.

Metagenomic surveys of virus diversity in wild plants underscore the need to examine the impacts of virus movement from crops to

†Corresponding author: K. E. Mauck; kerry.mauck@ucr.edu

Author contributions: T.M.S., K.E.M., and C.M.M. conceived the study. T.M.S., O.A., A.H., and J.R.K. selected sites, collected samples, and performed targeted diagnostics. P.S., M.G., and T.M.S. performed virome analyses, recombination analyses, and phylogenetic analyses. All authors worked together to write and edit the manuscript.

Funding: Support was provided by the Anza-Borrego Foundation Howie Wier Memorial Conservation Grant (T. M. Shates), the UCR Center for Conservation Biology via grants from the Shipley-Skinner Reserve Riverside County Endowment (T. M. Shates), Hatch project funds awarded to K. E. Mauck (CA-R-ENT-5144-H), and startup funds provided to K. E. Mauck by the University of California, Riverside. C. M. Malmstrom was supported by MSU AgBioResearch and USDA Hatch MICL 02382.

e-Xtra: Supplementary material is available online.

The author(s) declare no conflict of interest.

wild vegetation. For example, in a large metagenomic survey of virus diversity and prevalence across agricultural-to-wild transition zones, Bernardo et al. (2018) detected virus species that typically cause disease in agricultural crops in 59.5% of wild plant hosts sampled in uncultivated areas classified as "intact native habitat" or "degraded native habitat." These viruses have the potential to influence biological processes that affect the distribution, abundance, and performance of wild plant species, including growth and reproduction, defense responses, resource allocation, stress tolerance, and overall fitness (Hily et al. 2016; Malmstrom and Alexander 2016; Malmstrom et al. 2005, 2017; Xu et al. 2008; Zhao et al. 2016). In perennial plants, virus impacts are not necessarily confined to single growing seasons; viruses can be present in perennials as infectious agents for years, or even decades, during which hosts can experience cumulative fitness effects. For instance, Alexander et al. (2017) showed that common, aphid-vectored virus species that infect cereal crops can likewise reduce the multiyear fitness of a keystone grassland perennial (switchgrass); these same cereal-infecting virus species reduce the fitness and survival of threatened native perennial bunchgrasses in California (Malmstrom et al. 2005).

Viruses also compose viral communities within and among wild hosts (Sallinen et al. 2020; Susi et al. 2019). Untargeted sequencing has revealed that infections by multiple viruses within single host plants are the rule, rather than the exception (Schoelz and Stewart 2018; Thapa et al. 2012, 2015) and that most wild plants are infected with multiple viruses at any given time (Susi et al. 2019; Tollenaere et al. 2016). It is thus clear that viruses must be considered integral elements of natural plant communities. To evaluate their ecological influence requires interdisciplinary approaches that blend community ecology with virology, entomology, and plant biology. This work has been slowed by historical separations among these disciplines (Malmstrom et al. 2011) and by deficits in fundamental understanding of viral natural history. As a result, many key questions within virus ecology remain wide open for investigation: What factors shape virus communities within and across host species? How common are crop-associated viruses (those causing economic damage in crops) in wild plant communities, and are they genetically similar to viruses infecting crops? How long can such virus infections persist within wild perennial hosts? How do viruses, especially non-native, crop-associated species, influence plant health and fitness?

These research questions are logical extensions to viruses of the idea that microbes profoundly affect plants. However, where cellular microbes (e.g., fungi, bacteria) can be profiled using universal coding sequences targeting genomic DNA, no such universal coding sequences exist for viruses, which vary notably in genomic nucleic acid type and structure. Challenges in virus recovery and identification contributed to the slow start of virus ecology, but these challenges have diminished as high-throughput sequencing (HTS) has become more affordable and techniques for sequencing virus genomes in host material have improved (Edgar et al. 2022; Lacroix et al. 2016; Maclot et al. 2020; Roossinck 2017; Shates et al. 2019). As a result, we have a broad capability to study, with some limitations, the diversity and abundance of viruses in wild plant communities. Foundational work in this area (Roossinck 2011; Thapa et al. 2012, 2015) is a necessary first step to accelerating hypothesis generation and effective experimental design to quantify viral impacts on wild plant performance and fitness (Alexander et al. 2017; Bernardo et al. 2018; Malmstrom et al. 2017; Susi and Laine 2021; Susi et al. 2022).

Here, we examine virus diversity, prevalence, and impacts on summer-growing perennials in the unique drought-adapted grassland, sage scrub, and chaparral habitats of the California floristic province. These habitats house an incredible diversity of native species (Myers et al. 2000), but are shrinking and becoming more fragmented, creating new avenues for virus and vector incursions from agricultural and suburban landscapes (Brooks et al. 2002). Within these natural habitats, we focused on three herbaceous perennial angiosperms (*Cucurbita foetidissima* Kunth, *C. palmata* S. Watson, and *Datura wrightii* Regel; Fig. 1; Supplementary Table S1) that grow and bloom at the height of insect vector activity in crops (May through September) and when most annual plants in the area have senesced. This phenology results in the visual signature of large green plants against a dark soil background, which attracts generalist vector insects, including many common in crops (Döring and Chittka 2007; Döring et al. 2004; Perring et al. 2018). Based on their phenology, we hypothesized that (i) the focal native plant species are exposed to large numbers of vectors, and frequently support viruses previously detected in southwestern U.S. agricultural areas. We further hypothesized that (ii) the virus community structure here would vary with host species and site characteristics (habitat size, location) and that (iii) viruses typically found in agricultural crops might cause performance-reducing disease in the wild plant species.

MATERIALS AND METHODS

Field sampling and tissue handling. We collected plant tissue from *C. foetidissima*, *C. palmata*, and *D. wrightii* growing in three preserved natural areas (reserves) in southern California: Motte Rimrock Reserve (https://ucnrs.org/reserves/motte-rimrock-reserve/), Shipley-Skinner Multispecies Reserve, and Anza-Borrego Desert State Park (https://ucnrs.org/reserves/steeleburnand-anza-borrego-desert-research-center/; Fig. 1; Shates et al. 2019). For all collections, we sampled both leaf and stem tissue from each plant by inverting a clean plastic Ziplock bag over the tissue, pulling with force, and then sealing the bag. Approximately 10 g of leaf tissue was taken for double-stranded RNA (dsRNA) extraction, and 4 g of leaf plus stem tissue was carefully partitioned into 50-ml RNase-free Falcon tubes and stored at –80°C until further processing. In total, we collected tissue from 20 *C. foetidissima*, 14 *C. palmata*, and 9 *D. wrightii* for HTS from all three reserves.

We also collected additional samples to estimate the prevalence of select viruses identified by HTS within each reserve. We collected samples from a subset of the individuals at each location without regard to symptom expression (∼15 to 20% of the estimated total number of individuals in the location, based on visual surveys). Within Anza-Borrego, we collected an additional 20 *C. foetidissima*, 44 *C. palmata*, and 10 *D. wrightii* from each sampling location when we sampled hosts for HTS in 2019. For these collections, which required less tissue for total RNA extraction, we collected only one leaf per plant. At the Motte Rimrock and Shipley-Skinner reserves, we returned to the plant populations sampled in 2017 and collected leaves from ∼15 to 20% of the plants (estimated after initial counts early in the summer) for three additional years (2018, 2019, and 2020). Information about all surveyed plants, including the numbers of plants sampled for each analysis and associated metadata, is available in Supplementary Table S1.

GPS coordinates for sampled plants were recorded by photographing plants using smartphones with GPS capability and accessing photograph metadata. All three hosts are long-lived perennials with large tap roots, which are ideal for resampling. However, the cucurbit vines can grow many meters long and may overlap with other individuals. We attempted to return to the same individuals year to year, but because of variation in plant growth each year, only

a subset of sampled plants could be used conclusively for assessing year-to-year virus retention (numbers of sampled plants listed in Supplementary Table S2).

Nucleic acid extractions, sequencing, and virus detection. We extracted double-stranded RNA (dsRNA) for nontargeted HTS from all samples from which we had collected 10 g of tissue (Supplementary Table S1; Kesanakurti et al. 2016; Ma et al. 2020; Shates et al. 2019). For quality control, we spiked one leaf punch of 'California Wonder' bell pepper into each sample before extraction. This tissue was infected with bell pepper endornavirus (BPEV, *Endornaviridae*), and we assessed successful dsRNA extraction and sequencing based on recovery of BPEV. For example, before library preparation, we performed reverse transcription-PCR (RT-PCR) for BPEV to evaluate virus recovery. To construct libraries, we denatured an aliquot of 5 µl of dsRNA solution from each extraction by incubating at 99°C for 5 min, placed the solution on ice, and used the NEBNext UltraTM II Directional RNA Library Prep Kit for Illumina (New England Biolabs) following the manufacturer's protocol. The average insert size was 250 to 300 bp. Sequencing was done at the UC Riverside Genomics Core on the Illumina NextSeq platform on a Mid Output flow cell (v2) with a paired-end 75-bp read length, which can produce 260M paired-end reads. Adaptors were trimmed by Core staff before we processed the output using a workflow we previously designed for the Galaxy Platform (Afgan et al. 2018; Shates et al. 2019; The Galaxy Community 2022). We filtered out host genomes, performed de novo assembly using Trinity, and used nucleotide BLAST to find putative virus identities from the NCBI GenBank plant virus database as in Shates et al. (2019). Sequencing results for Motte Rimrock Reserve and Shipley-Skinner Reserve were previously published in Shates et al. (2019) (GenBank accession SRP149013) and results for collections from Anza-Borrego are reported for the first time here.

For targeted detection of select viruses identified through HTS, we extracted total RNA from the leaf tissue of plants subjected to HTS, and then sampled additional plants to confirm HTS identities and estimate virus prevalence in each location. We homogenized samples under cryogenic conditions using a GenoGrinder (SPEX) and then used TRI Reagent (Sigma-Aldrich) to extract total RNA following the manufacturer's protocol. After TRI Reagent extractions, RNA was purified according to the protocol of Hughes (2019) to remove excess polysaccharides and other inhibitors that are frequent contaminants in nucleic acid extraction from wild plants. We had previously confirmed putative virus identities for the viruses present in plants from Motte Rimrock and Shipley-Skinner reserves. Here, we repeated these steps to validate dominant viruses (see section below) in the plants from Anza-Borrego. We used targeted RT-PCR with virus-specific PCR primers for the coat protein for known virus species to confirm virus identities from dsRNA extracts (primers [IDT DNA] in Supplementary Table S3). All reverse transcription and PCR were conducted according to the protocol in the Supplementary Material, Appendix I on a Bio-Rad T100TM thermal cycler (Bio-Rad Laboratories). For total RNA samples collected to determine dominant virus prevalence, we followed the same RT-PCR steps apart from the amount of RNA solution used

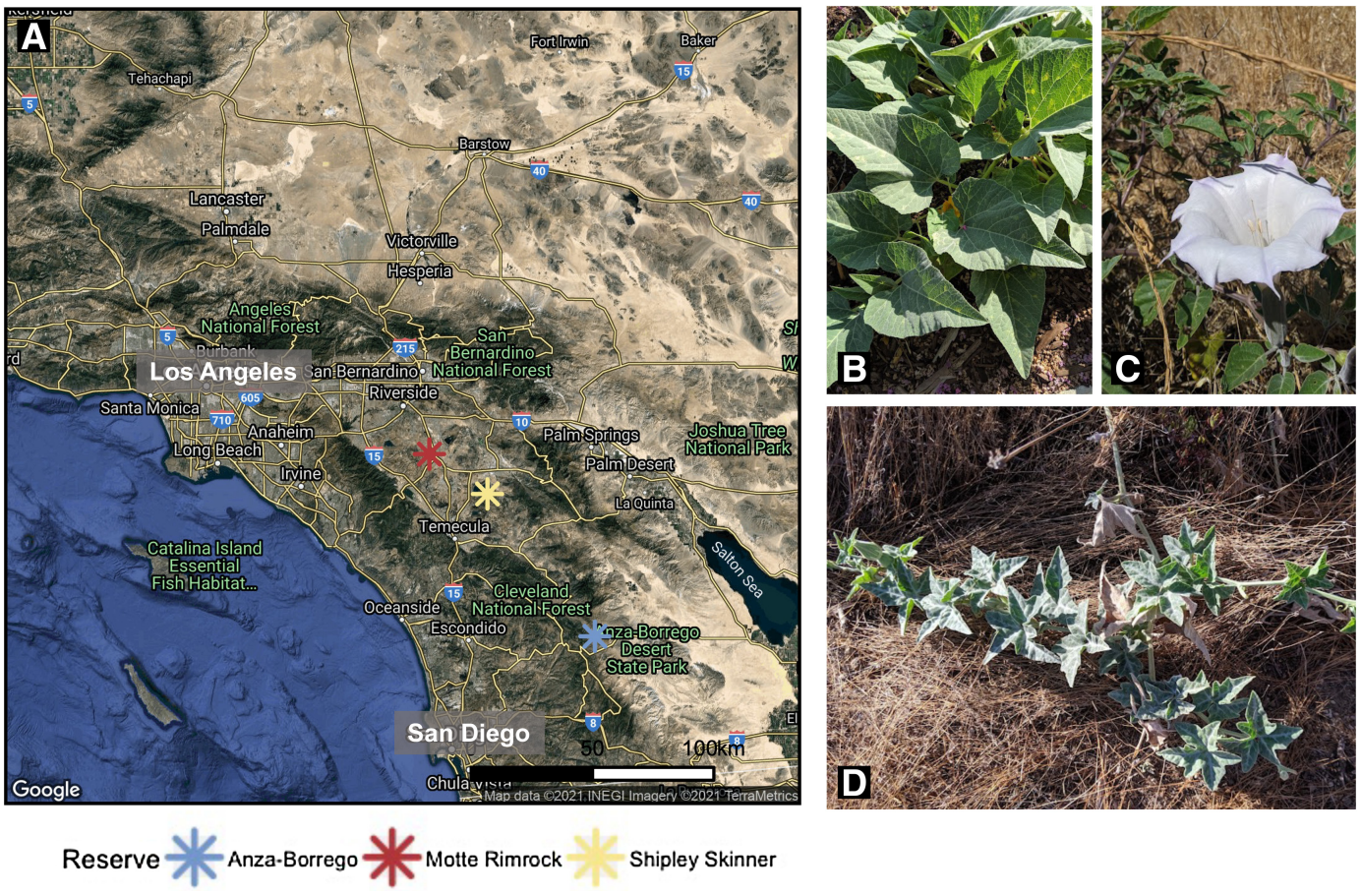

Fig. 1. A, Map of field sites in Riverside County and surrounding areas in southern California, constructed using the *ggmap* package in R (Kahle and Wickham 2013). **B,** *Cucurbita foetidissima* Kunth (buffalo gourd, Cucurbitaceae); **C,** *Datura wrightii* Regel (sacred thorn-apple, Solanaceae); and **D,** *C. palmata* S. Watson (coyote melon, Cucurbitaceae).

in each reaction (approximately 0.5 to 10 µl from each RNA solution, depending on concentration). We then placed the solution on ice and performed reverse transcription. For Sanger sequencing of PCR products, amplicons were purified using a Zymoclean Gel DNA recovery kit (Zymo Research). Sanger sequencing was performed by Retrogen.

Analysis of virus communities by site and host species. To determine likely drivers for the assembly of virus communities and identify key differences between regions, we calculated species richness and diversity metrics for viruses and hosts and then performed permutational multivariate analysis of variance (PERMANOVA), indicator species analysis (ISA), redundancy analyses (RDA), and distance-based Moran's Eigenvector Maps (dbMEM). PERMANOVA is a nonparametric multivariate analysis of variance commonly used to test the association of species composition with variables of interest. ISA uses the relative abundance of a given organism and its frequency of occurrence to estimate the degree of association with particular groups, as well as the probability of finding this organism in association with a given group (Severns and Sykes 2020). RDA is a type of constrained ordination analysis (similar to principal component analysis) that models linear relationships among nonnumerical predictors (e.g., host or site) and response variables (Capblancq and Forester 2021). dbMEM is a method for analysis of spatially explicit multivariate data and is used to explore geographic (spatial) influence on species assemblages (in this case, virus communities in target host species and sites; Legendre and Gauthier 2014). For all of these analyses, we removed plant individuals in which only previously uncharacterized (potentially novel) viruses were detected (Shipley-Skinner DW1, DW4, CF1B; Motte Rimrock CP4, CF1A) because our analyses focused on previously characterized virus species and many of these analyses do not tolerate zero values. Additionally, six undescribed viruses (<70% nucleotide identity match to a known virus species) found across all samples were removed from the virus community data because their identities could not be confirmed and full characterization of these potentially novel species was beyond the scope of this study. Alpha diversity was calculated as species richness (number of species per site) and species evenness was calculated as Pilou's *J* for both plant species and site factors. We chose Sorenson's dissimilarity metric for beta diversity calculations because the virus species data are presence–absence data, and this metric is more robust to extreme differences between and among communities. We calculated beta diversity using the *vegan* and *ade4* packages using R (R Core Team 2021) and RStudio (RStudio Team 2020). Gamma diversity is simply the total number of virus species found across all the hosts.

To test for differences in community composition, we used a PERMANOVA test that employs distance matrices (R package *PERMANOVA*; Vicente-Gonzalez and Vicente-Villardon 2021). This procedure tests for similarities among objects within and across groups in the context of environmental variables. The overall PERMANOVA tests for location effects of differences among communities in vector space. We also evaluated pairwise differences between groups. We used the R packages *ade4*, *vegan*, and *devtools* (pairwiseAdonis; Oksanen et al. 2017; Thioulouse et al. 2018; Wickham et al. 2022). To determine differences among virus communities by reserve and plant host species, we used indicator species analysis (ISA; R package *indicspecies*; De Cáceres and Legendre 2009; Severns and Sykes 2020). This analysis computes a *P* value using permutation and gives an ISA value of zero to one. One is the strongest indicator value and indicates a species that is exclusive to one group, occurs in all of the samples taken from that group, and has high relative abundance within that group (Severns and Sykes 2020). In contrast, a value of zero indicates a species that is rare or nearly uniform in distribution across groups. We then used dbMEM to calculate distance-based Moran's eigenvector maps to explore whether geographic distance between communities contributes to differences in community composition. To do this, we first plotted spatial correlograms to find evidence for spatial correlations based on diversity (R package *vegan*). Next, we calculated Euclidean distances among plots, constructed dbMEM eigenfunctions, and used an RDA for the forward selection process to isolate the positive and significant spatial variables (R package *adespatial*; Dray 2019).

Prevalence and retention of dominant viruses. HTS and indicator species analyses revealed that two virus species typically associated with crops dominated plant communities: cucurbit aphid-borne yellows virus (CABYV), which was present at all sites, and cucurbit yellow stunting disorder virus (CYSDV), which was present at two sites and dominant at one site. These two species were selected as targets of RT-PCR in additional tissue samples to estimate prevalence. We calculated the proportion of detections for each year, at each reserve, for each virus × host species combination. To explore differences in virus retention by host species across time, the proportion of positive detections per plant species was compared between distinct plants sampled in 2019 and resampled in 2020 using a two-tailed *Z*-test of two proportions calculated in Excel.

CABYV isolation and *Cucurbita* spp. culture. To determine the impact of CABYV on the two cucurbit species, we isolated the virus, collected seeds from hosts without regard to infection status, and carried out controlled inoculation experiments under greenhouse conditions. CABYV is not seed-transmitted or mechanically transmissible, so inoculations were performed with the vector *Aphis gossypii.* The *A. gossypii* colonies used originated from a collection from cultivated squash about a decade ago near Reedley, CA and were maintained on melon plants (*Cucumis melo* 'Iroquois'). We isolated CABYV in summer 2019 from wild-growing *C. palmata* plants at Motte Rimrock Reserve (https://ucnrs.org/reserves/motte-rimrock-reserve/) through aphid transmission and maintained the isolate in cultivated squash (*C. pepo* 'Dixie') through subsequent transmission by *A. gossypii*. Both aphid and CABYV cultures were maintained at 25 ± 2°C in mesh cages with supplemental lighting with a photoperiod of 16 h of light and 8 h of darkness (16L/8D).

C. foetidissima and *C. palmata* seeds used in experiments originated from four individuals (two of each species) growing wild at Motte Rimrock Reserve in Riverside County, California. *C. palmata* seeds were soaked in water and *C. foetidissima* seeds in aerated water, both for 72 h at room temperature. After pretreatments, seeds were placed in Petri dishes with damp filter paper for 48 h, planted in 5.08 × 5.08 × 5.08 cm pots, and maintained on a 16L/8D photoperiod and at a temperature of 25 to 28°C. Approximately 2 weeks postgermination, seedlings were moved to a temperature-controlled walk-in growth chamber kept at 23 ± 1°C, 60 ± 5% RH, and a 16L/8D photoperiod. To generate CABYV-infected and sham-inoculated treatments for the greenhouse experiment, 20 aphids from a clean aphid culture (sham inoculation) or 20 aphids from a CABYV culture (inoculation) were placed on each plant. We allowed aphids to feed for 72 h and then eliminated them by treating plants with Dinotefuran (Venom insecticide from Keystone Pest Solutions) at the label rate for cucurbits (from the label: 5 oz/acre delivered in 20 gallons of fluid, adjusted for soil delivery of 10 ml of fluid per pot). Two weeks later, plants were repotted into 6 × 5.75-in. large round pots and arranged in a greenhouse for the rest of the experiment. Every other row of plants was a different treatment group containing a mixture of *C. palmata* and *C. foetidissima*. Plants were watered daily and sprayed twice with miticide (Spear-T) and fungicide (Draconil), both provided by the Agricultural Operations facility at University of California, Riverside. In total, the experiment in-

cluded 109 plants: 35 CABYV-inoculated and 26 sham-inoculated *C. foetidissima*, and 26 CABYV-inoculated and 22 sham-inoculated *C. palmata*.

Impact of CABYV on plant performance metrics. We recorded the number of leaves on each plant every 2 weeks, from 2 to 8 weeks postinoculation (WPI). To quantify symptom expression and severity, we scored each fully expanded leaf with a foliar disease scale developed by (Takeshita et al. 2013), and subsequently refined for use in our laboratory (Kenney et al. 2020). Leaf ratings were combined to calculate whole-plant symptom severity with the formula $D = [(di)/n/4]$, where D is symptom severity, di is the symptom rating, n is the total number of leaves, and 4 represents the 4 levels of the rating. For mottling and yellowing symptoms, we used the following scale: 1 = slight mottling (tiny light spots visible), 2 = 10 to 20% leaf area yellow (but bigger, distinctly yellow spots, rather than tiny light spots), 3 = 21 to 50% leaf area yellow, 4 = ≥50% leaf area yellow. For leaf distortion, we used this scale: 1 = slight distortion (leaf shape is irregular, or surface slightly bumpy), 2 = 10 to 20% leaf area is curled, has large bumps, or is twisted, 3 = 21 to 50% leaf area is curled, has large bumps, or is twisted, 4 = ≥50% is curled, twisted, and/or not typically "leaf-shaped." At 8 WPI, we collected leaf punches from sham-inoculated and virus-inoculated individuals and stored them at –80°C before RNA extraction and virus detection with RT-PCR to confirm infections. Aboveground (foliar) tissue was weighed after drying for 72 h. Belowground (root) tissue was gently washed to clean as much soil as possible while retaining fine root tissue attached to the main tap root, and then dried for 72 h and weighed.

For data collected at the conclusion of the experiment (root and shoot masses), we used the Wilcoxon rank sum test for nonparametric data to compare CABYV and sham-inoculated treatments within each host species. For metrics repeated throughout the experiment, we used a repeated-measures generalized linear mixed effects model with the symptom or metric measured as a response, treatment × weeks postinoculation as a factor, and plant individual as the repeated factor, using a Poisson distribution with Log specified as the link function in the model. We used the R library *lme4* (Bates et al. 2015) to run the model and used the R library *car* (Fox and Weisberg 2019) to test the model for significance with an analysis of deviance table. Post hoc tests were performed with the R package *emmeans* (Lenth 2016) with a false discovery rate *P* value adjustment. Plots were created using the R packages *ggplot2* (Wickham 2016) and *ggpubr* (Kassambara 2020).

Phylogenetic analysis of CABYV sequences from wild plants. To detect recombination and for our main phylogenetic analyses, we selected GenBank accessions for which both coat protein (CP) and RNA-dependent RNA polymerase (RdRp) sequences were available. CP and RdRp sequences from California sampling efforts were extracted from HTS data (GenBank accessions OQ833203 to OQ833236). We created a dataset that consisted of 77 concatenated RdRp and CP sequences (open reading frame [ORF] 1 to 2 and 3; detailed in Supplementary Table S5). Sequences for outgroup taxa included melon aphid-borne yellows virus (MABYV), pepo aphid-borne yellows virus (PABYV), and suakwa aphid-borne yellows virus (SABYV; Supplementary Table S5). The original accessions submitted as MABYV in 2009 (EU091150, EU091151, and EU091149; Xiang et al. 2008) were mislabeled as CABYV in GenBank. In Supplementary Table S5 under the "true identity" column we designated these sequences as MABYV. To detect recombination, we used the program RDP version 5.34 (Martin et al. 2021). First, we performed a pairwise homoplasy index (PHI) test for overall evidence of recombination. This test gave strong evidence of recombination ($P < 0.000001$). Therefore, we followed this with a full exploratory recombination scan using nine methods (RDP, GENECONV, BootScan, MaxChi, Chimaera, SiScan, PhylPro, LARD, and 3Seq) and a Bonferroni-corrected *P* value cutoff of 0.01. To be conservative and avoid false detection of recombination, only signals with a minimum of 5 out of 9 supportive tests were included in the results (Rabadán et al. 2021). We excluded false positive signals of recombination events that could have been caused by processes other than recombination. After detecting recombination events, we reran phylogenetic analyses (see below) to account for the possible impact of recombinant regions, using an edited version of the input alignment where regions of sequences derived through recombination were removed by RDP5.

We then used phylogenetics to explore regional groupings of sequences originating from our sampling of wild plants in California. To supplement this analysis and provide more geographic resolution, we also acquired 209 CP sequences from GenBank (Supplementary Table S5). We produced phylogenies based on the concatenated dataset and on full CP sequences by aligning nucleotide sequences using Qiagen CLC Main Workbench (v. 20.0.4) and then using ModelFinder Plus (Kalyaanamoorthy et al. 2017) to find the best evolutionary model for the data. All datasets (concatenated, CP, and concatenated without recombinant regions) were assigned the TIM2 + G + I model by the Bayesian information criterion. We then generated maximum likelihood (ML) phylogenies on the IQTree2 executable (Hoang et al. 2018) with 1,000 ultrafast bootstrap replicates to assess statistical support.

RESULTS

Target wild hosts are infected with diverse crop-associated viruses. Using HTS, we detected 12 virus species known to infect crops across all sampled plant species at our three remnant native plant community sites (reserves; Fig. 2), as well as at least six previously uncharacterized viruses. The previously characterized viruses all have at least 90% nucleotide identity match to a known virus species and are all insect-transmitted by sap-feeding insects in the order Hemiptera, family Aphididae (aphids) or Aleyrodidae (whiteflies), which are present on sampled plants in our sites (Supplementary Table S6). The uncharacterized viruses were not considered in subsequent analyses, because each of these potentially novel viruses requires further studies that are beyond the scope of the present work. Out of 37 sampled plants, most were infected with one to three previously characterized virus species; one plant was infected with four (Fig. 3). Five samples were removed before species richness analysis (Fig. 3) because no previously characterized viruses were detected. We did not detect significant differences in beta diversity values based on reserve or host species (Table 1; ANOVA, *P* values 0.95 and 0.08 by reserve and species, respectively).

Virus communities in wild plants are shaped by spatial distributions and host species. We used PERMANOVA to test virus community differences by reserve, plant species, and the interaction of reserve with plant species. This analysis was followed by a pairwise PERMANOVA for each site or species pair. The overall PERMANOVA and model fit were significant, but pairwise tests showed that some reserve and host virus communities are more similar than others. Motte Rimrock and Shipley-Skinner reserve virus communities did not differ from each other, but both differed from the virus community at Anza-Borrego State Park (Table 2). The virus communities of *Cucurbita* spp. were more similar to each other than to the *Datura* virus community (Table 2). The PERMANOVA suggests that both geography and host species are important determinants of virus community composition. This finding is supported by a dbMEM analysis testing whether distance between communities (latitude, longitude) is correlated with community-level differences.

From this analysis, we found one positive spatial correlation and evidence for one spatial variable—"1MEM" (Supplementary Fig. S2). An RDA forward selection model with an adjusted R^2 value revealed that 20% of the variation in the virus communities is explained by this spatial variable.

The indicator species analysis suggests that there are four virus species differentiating reserves and three virus species differentiating host species (Table 3). Cucurbit yellow stunting disorder virus (CYSDV, genus *Crinivirus*, family *Closteroviridae*) is an indicator species for Anza-Borrego, papaya ringspot virus (PRSV, genus *Potyvirus*, family *Potyviridae*) is an indicator species for Motte Rimrock, and tomato chlorosis virus (ToCV, genus *Crinivirus*, family *Closteroviridae*) is an indicator species for Shipley-Skinner. Cucurbit aphid-borne yellows virus (CABYV, genus *Polerovirus*, family *Solemoviridae*), the most prevalent virus, is an indicator for both Motte and Shipley. At the host species level, CYSDV was identified as an indicator species for *C. palmata* and ToCV and tomato necrotic dwarf virus (ToNDV, genus *Torradovirus,* family *Secoviridae*) as indicator species for *D. wrightii.*

CABYV is highly prevalent in cucurbit hosts across all three sites. Diversity metrics and indicator species analyses identified CABYV and, to a lesser extent, CYSDV as key virus species in target host populations. CABYV was present across multiple host species at three sites (detected in 22 of 37 HTS samples) and was an indicator species for both the Motte Rimrock and Shipley-Skinner locations. CYSDV was the second most frequently detected virus (15 of 37 HTS samples) and was identified as an indicator species for the Anza-Borrego site and *C. palmata* hosts. Targeted screening of more individuals from each population therefore focused on these two virus species. In 2018 to 2020, we detected CABYV at all field sites surveyed (Fig. 4; Supplementary Table S8) with infection rates ranging from 10 to 88%. In 2019 and 2020, we detected CYSDV in 25 to 40% of plant samples from Anza-Borrego. CYSDV was detected in 8% of *C. palmata* plants, with co-occurring infections of CABYV in 2019 from Motte Rimrock. In 2020, 2 and 3% of *C. foetidissima* and *C. palmata* plants, respectively, had CABYV–CYSDV dual infections detected, and the single *D. wrightii* plant sampled was the only plant with a single CYSDV infection. CYSDV

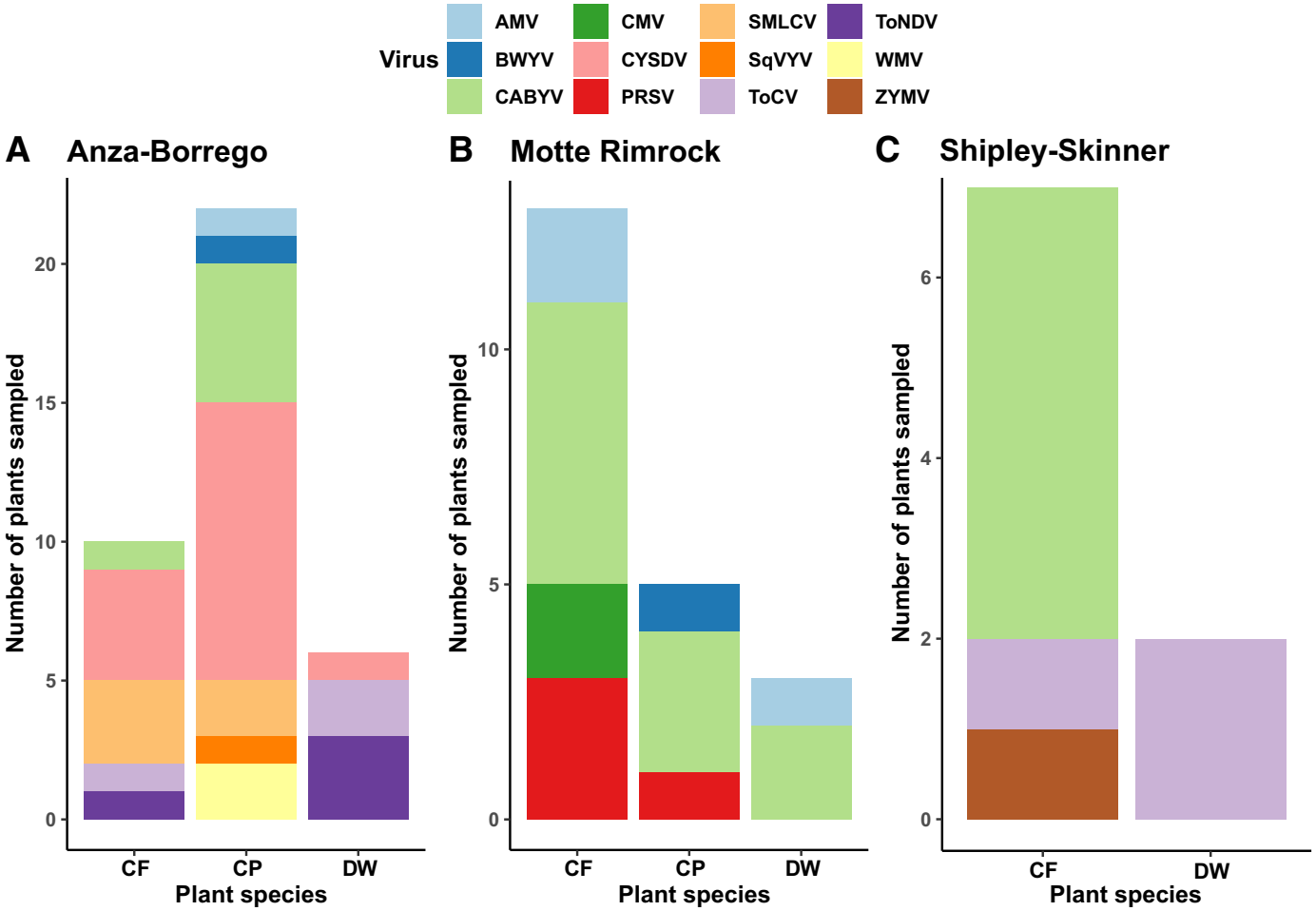

Fig. 2. Crop associated viruses detected in sampled plants of each species through high-throughput sequencing (γ diversity). Virus acronyms correspond to the following species: AMV = alfalfa mosaic virus (genus *Alfamovirus*, family Bromoviridae), BWYV = beet western yellows virus (genus *Polerovirus*, family *Solemoviridae*), CABYV = cucurbit aphid-borne yellows virus (genus *Polerovirus*, family *Solemoviridae*), CMV = cucumber mosaic virus (genus *Cucumovirus*, family *Bromoviridae*), CYSDV = cucurbit yellow stunting disorder virus (genus *Crinivirus*, family *Closteroviridae*), PRSV = papaya ringspot virus (genus *Potyvirus*, family *Potyviridae*), SMLCV = squash mild leaf curl virus (genus *Begomovirus*, family *Geminiviridae*), SqVYV = squash vein yellowing virus (genus *Ipomovirus*, family *Potyviridae*), ToCV = tomato chlorosis virus (genus *Crinivirus*, family *Closteroviridae*), ToNDV = tomato necrotic dwarf virus (genus *Torradovirus,* family *Secoviridae*), WMV = watermelon mosaic virus (genus *Potyvirus*, family *Potyviridae*), ZYMV = zucchini yellow mosaic virus (genus *Potyvirus*, family *Potyviridae*). **A,** Anza-Borrego; **B,** Motte Rimrock; and **C,** Shipley-Skinner. Plant codes are CP = *Cucurbita palmata*, CF = *C. foetidissima*, and DW = *Datura wrightii*. *C. palmata* was not sampled from Shipley-Skinner because it did not occur at this site.

was not detected at the Shipley-Skinner reserve (Fig. 4; Supplementary Table S8). At Anza-Borrego, multiple infections with CABYV and CYSDV were more common than single infections for both cucurbit species, but not for *D. wrightii*. At Motte Rimrock, where both viruses were present, CYSDV was found only in individuals also infected with CABYV, which was highly prevalent. One *D. wrightii* plant was sampled at Motte Rimrock in 2020, and it tested positive for CYSDV. The distributions of infections by year, plant species, and site are shown overlaid on reserve maps in the Supplementary Material.

CABYV infections persist across seasons in perennial cucurbits. A subset of *C. foetidissima* and *C. palmata* individuals were repeatedly sampled over multiple seasons to determine year-to-year retention of the most prevalent virus (CABYV) at Motte Rimrock Reserve (Fig. 5). In 2019, 69% (20 of 29) of sampled *C. foetidissima* plants tested positive for CABYV; 90% (26 of 29) of these same individuals tested positive when resampled in 2020, representing a marginally significant increase in prevalence across the years (*Z* test: z score –1.94, $P = 0.052$). In parallel, 85% (22 of 26) of sampled *C. palmata* plants tested positive for CABYV in 2019 and 92% (24 of 26 plants) in 2020 (between-year difference not significant; z score –0.87 and $P = 0.38$). Overall, eight individuals without detectable CABYV in year one of sampling tested positive in resampling years, while only two tested negative in subsequent years after testing positive in year one.

To explore within-plant virus reservoirs, we also dug up dormant root portions from one *C. foetidissima* plant that had tested CABYV-positive via leaf sampling in 2018. Foliar tissue sprouted from the roots in the greenhouse tested positive for CABYV. CYSDV prevalence was too low to facilitate a robust test of infection persistence between years. However, one of the *C. palmata* plants that tested positive for CYSDV in 2019 was also positive in 2020.

CABYV infection reduces plant vigor. A summary of the statistical analyses of greenhouse experiments is reported in the Supplementary Material, Appendix II, and in Figure 6. CABYV symptoms were apparent in the experimentally inoculated cucurbits as early as 2 weeks after inoculation. CABYV-related leaf distortion progressed similarly in both host species and was evident for *C. palmata* at 2 and 4 WPI and for *C. foetidissima* at 2, 4, and 8 WPI (Fig. 6A and B). In contrast, mottling (Fig. 6C and D) and yellowing (Fig. 6E and F) were evident in *C. palmata* at all time points but in *C. foetidissima* only at 2 WPI (mottling) and 8 WPI

TABLE 1
Virus beta diversity values based on Sorenson matrix by reserve and by plant species

Type	Factor	Beta[a]
Reserve	Anza-Borrego	0.45
Reserve	Motte Rimrock	0.337
Reserve	Shipley-Skinner	0.408
Plant species	*Cucurbita foetidissima*	0.482
Plant species	*C. palmata*	0.352
Plant species	*Datura wrightii*	0.498

[a] Beta diversity is calculated on a normalized scale from 0 to 1. A higher index indicates low levels of similarity for the indicated nonnumerical explanatory factor relative to the other factors of that type.

TABLE 2
Overall permutational multivariate analysis of variance (PERMANOVA) and pairwise PERMANOVA R^2 values and *P* values[a]

Group	Test	R^2	*P* value
Reserve	PERMANOVA	0.39	0.001
Plant species	PERMANOVA	0.13	0.001
Reserve: plant	PERMANOVA	0.13	0.001
Overall model	Model fit	0.66	0.001
Anza versus Motte	Pairwise	0.35	0.003
Anza versus Shipley	Pairwise	0.28	0.003
Motte versus Shipley	Pairwise	0.14	0.102
CP versus CF	Pairwise	0.42	0.126
CP versus DW	Pairwise	0.001	0.003
CF versus DW	Pairwise	0.011	0.033

[a] Significant and nonsignificant interactions are reported. CP: *Cucurbita palmata*; CF: *C. foetidissima*; and DW: *Datura wrightii*.

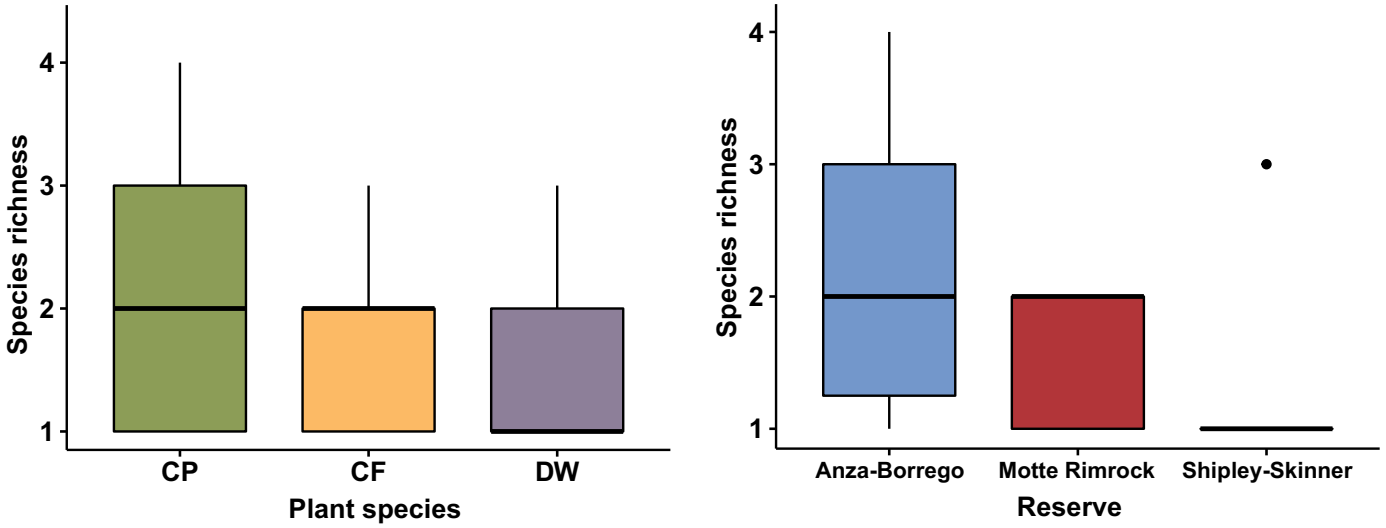

Fig. 3. Viral species richness of each site visualized by plant host species and reserve. Plant codes are CP = *Cucurbita palmata*, CF = *C. foetidissima*, and DW = *Datura wrightii*.

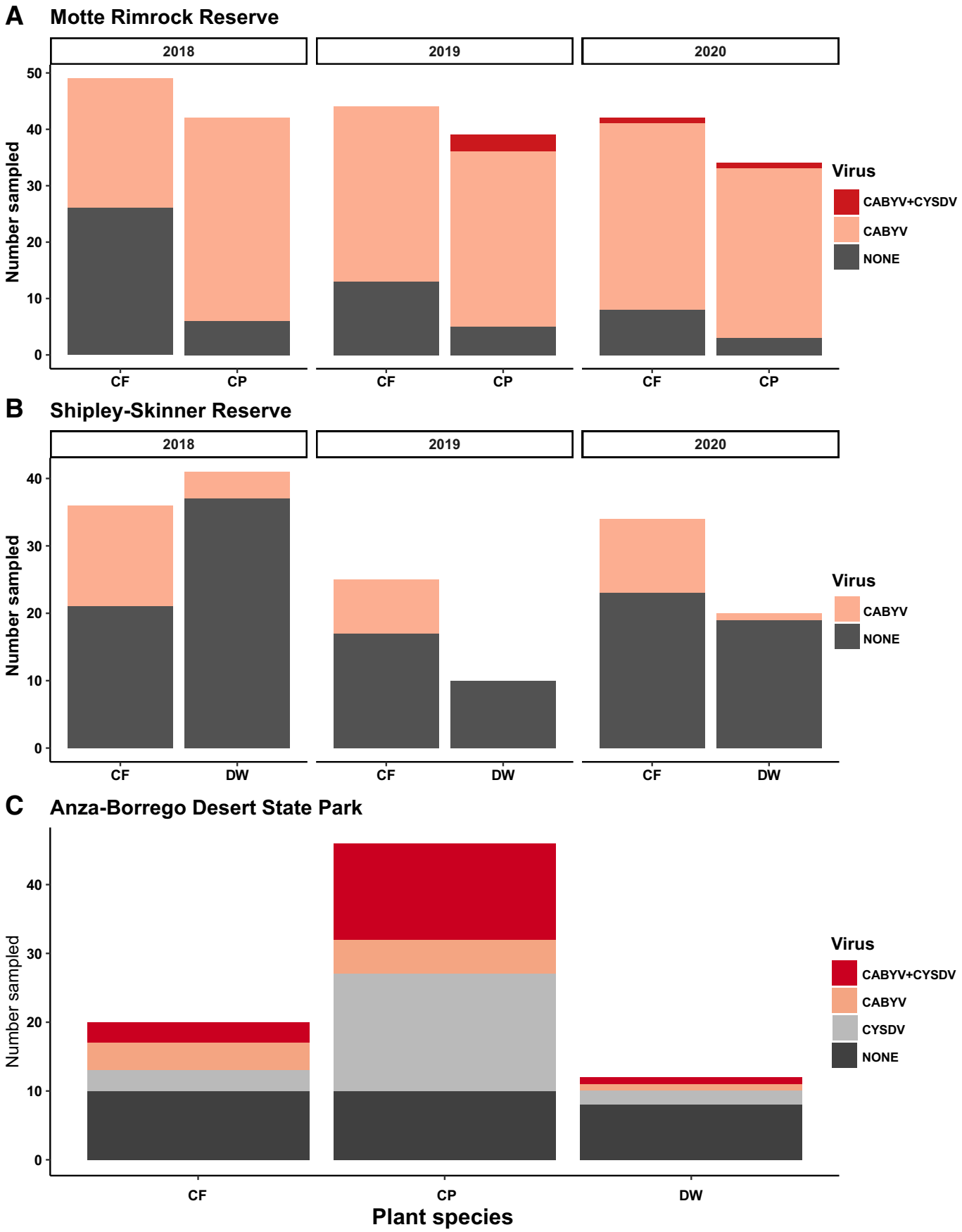

Fig. 4. Number of host plants sampled in which cucurbit aphid-borne yellows virus (CABYV), cucurbit yellow stunting disorder virus (CYSDV), or CABYV+CYSDV were detected. **A,** Motte Rimrock; **B,** Shipley-Skinner; and **C,** Anza-Borrego. CYSDV was only targeted for detection in 2019 and 2020 following discovery through high-throughput sequencing in Anza-Borrego samples and Anza-Borrego was only sampled during 2019. Plant species abbreviations are as follows: CF = *Cucurbita foetidissima*, CP = *C. palmata*, and DW = *Datura wrightii*.

(yellowing). CABYV infection did not alter total leaf number for either cucurbit (no statistically significant difference with control; Fig. 6G and H). However, infection significantly reduced the shoot mass of both cucurbit species by 20 to 30% (Fig. 7A), and significantly reduced the root mass of *C. foetidissima* by 20%. We also detected marginally significant reductions in root mass for *C. palmata* (~15%; $P = 0.059$; Fig. 7B).

California CABYV diversity is the product of multiple introductions. Recombination analysis using RDP5 detected 15 recombination events, 11 of which were supported by at least 5 of 9 different detection methods (Supplementary Table S7). Sequences from three California samples (MRCP6, MRCF3C, and SF2CF) were unambiguously identified as recombinants, all originating from both Californian parentals, and two of them (MRCP6 and SF2CF) were at the same time minor parentals in other recombination events. The remaining 14 California sequences were identified as nonrecombinants, along with 16 other sequences from different geographic areas (collected from GenBank). Our phylogenetic analysis showed that 13 California sequences clustered with sequences from Spain, France, and Pakistan, while four other sequences belonged to a clade with sequences from China, South Korea, Japan, Uzbekistan, the Philippines, and the United States. This tree also reveals much diversity within CABYV, with a sequence from Taiwan (of recombinant origin, based on recombination analysis) being sister to the CABYV assemblage and sequences from Taiwan and East Java forming another well-supported clade (Fig. 8). The phylogeny based on CP (ORF3) was less resolved overall but still informative, given the much broader geographical coverage represented by the coat protein sequences available. This analysis differentiated a clade containing sequences from Europe, the Mediterranean, and the Middle East from a second clade composed of sequences from Asia. California sequences clustered with both clades (Supplementary Fig. S6).

Taken together, the phylogenetic analyses suggest that there were at least two separate CABYV introductions to California, one from Europe and one from Asia. Both phylogenies provide support for CABYV sequences from Taiwan and India being distinct from these two major regional groups. The phylogeny based on a dataset without recombinant regions produced a similar topology overall (Supplementary Fig. S7) with all the major groups recovered, including the two major CABYV regional groups. Only a handful of taxa (MABYV and newly identified recombinants, including recombinants from California) are placed differently than in the original phylogeny, indicating a minor impact of recombination on phylogenetic inference.

TABLE 3
Indicator species analysis (ISA) values and *P* values[a]

Type	Group	Virus species	ISA value	*P* value
Reserve	Anza-Borrego	CYSDV	0.913	0.001
Reserve	Motte Rimrock	PRSV	0.555	0.027
Reserve	Shipley-Skinner	ToCV	0.555	0.048
Reserve	Motte + Shipley	CABYV	0.812	0.013
Plant species	*Cucurbita palmata*	CYSDV	0.722	0.009
Plant species	*Datura wrightii*	ToCV	0.692	0.002
Plant species	*D. wrightii*	ToNDV	0.616	0.009

[a] Only significant ($\alpha > 0.05$) species are reported. Two separate analyses were run: by reserve and by plant species. Both analyses are reported in this table. ISA value is the indicator species analysis value for each virus species in each group. It is the product of two proportions: the mean relative abundance of species in each group and the relative frequency of each species in each group. This value is then multiplied by 100 to yield an indicator value. In the table, type refers to the categorization of the group (geographic location or host plant species). Cucurbit yellow stunting disorder virus (CYSDV); papaya ringspot virus (PRSV); tomato chlorosis virus (ToCV); cucurbit aphid-borne yellows virus (CABYV); and tomato necrotic dwarf virus (ToNDV).

DISCUSSION

To investigate how virus species typically found in crops influence wild plants, we examined three wild drought-tolerant perennial species endemic to semiarid and arid habitats in California (semiarid grassland, sage scrub, and chaparral) that are encroached upon by agriculture and development. Across three natural reserves, we found that the target wild hosts (*Cucurbita* and *Datura* spp.) were infected with a rich community of 12 virus species typically found in crops. These virus communities varied with host species and geographic location. Across all locations, however, the dominant viruses were non-native crop pathogens originally introduced from outside of the United States, particularly CABYV (first detected in California in 1992) and, to a lesser extent, CYSDV (first detected in California in 2006; Kuo et al. 2007). Dual infections with both viruses were also identified, especially at Anza-Borrego Desert State Park, which is closest to the southern desert agricultural areas where CYSDV was first established in this region (Wintermantel et al. 2017). CABYV was the most prevalent virus overall, infecting up to 88% of sampled plants in some populations of wild hosts. Our monitoring suggests that CABYV infects the majority of plants in

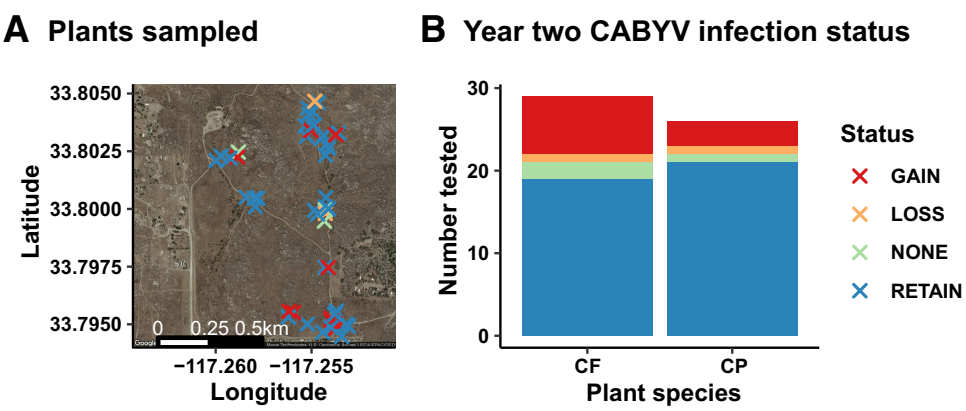

Fig. 5. A, Distribution of Motte Rimrock Reserve cucurbit plants tested for cucurbit aphid-borne yellows virus (CABYV) retention in 2019 and 2020. **B,** Numbers of plants that were newly positive (gain), negative 1 year after testing positive (loss), never tested positive (none), or remained positive (retain) for each host species (CF = *Cucurbita foetidissima* and CP = *C. palmata*).

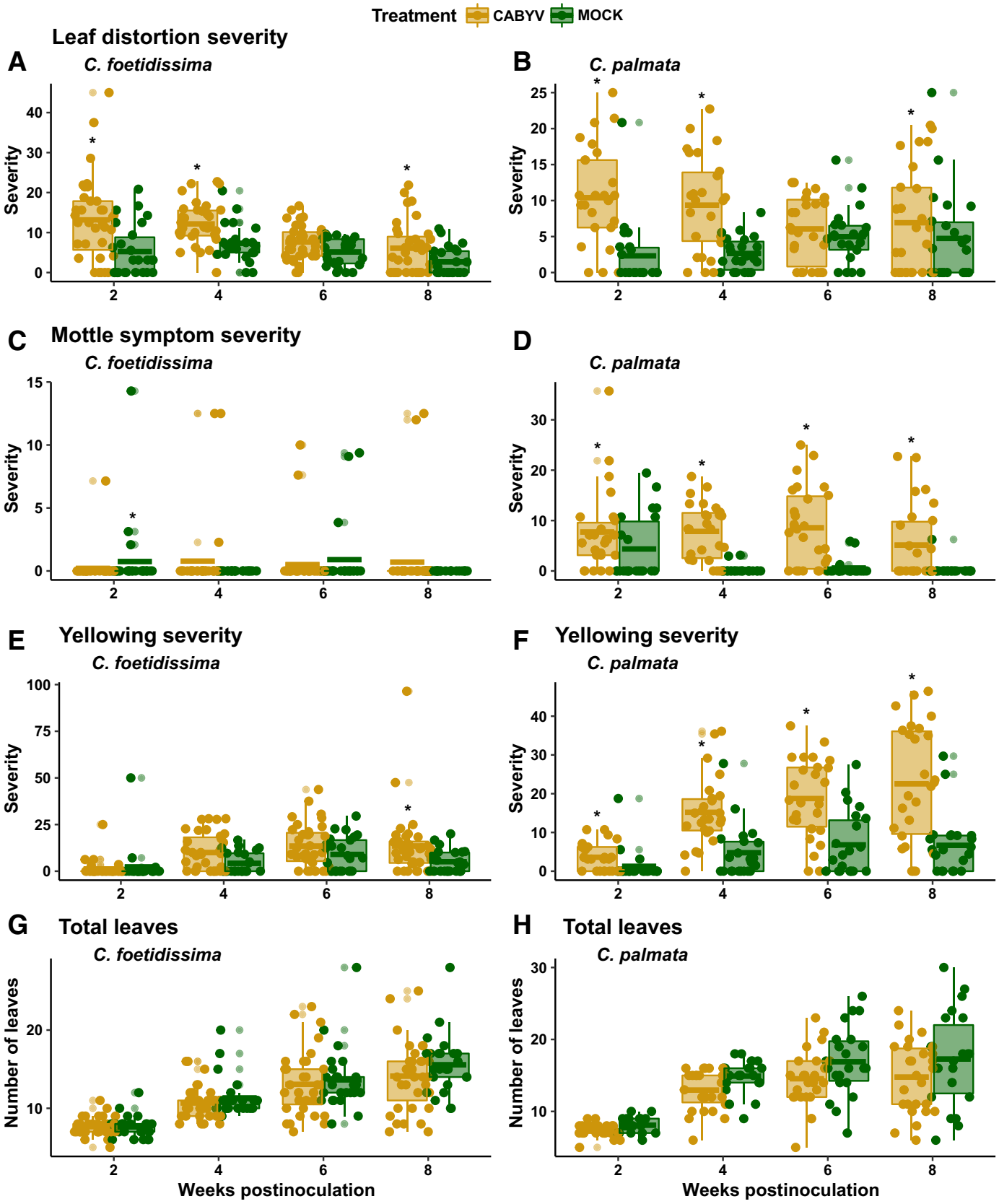

Fig. 6. Metrics of disease severity over time in the controlled inoculation greenhouse experiments for each cucurbit host species: leaf distortion severity for **A,** *Cucurbita foetidissima* and **B,** *C. palmata*; mottle symptom severity for **C,** *C. foetidissima* and **D,** *C. palmata*; yellowing severity for **E,** *C. foetidissima* and **F,** *C. palmata*; and total leaves for **G,** *C. foetidissima* and **H,** *C. palmata*. ∗ Indicates significant differences relative to the mock treatment at the given time point ($P < 0.05$). *P* values are in the Supplementary Materials, as well as model outputs for the full repeated measures analysis. CABYV, cucurbit aphid-borne yellows virus.

populations of the two cucurbit species, particularly the apparently more susceptible *C. palmata*. Symptoms of CABYV infection in cucurbit hosts were readily apparent in both the field and laboratory. Through a controlled inoculation experiment in the greenhouse, we discovered that CABYV infections cause measurable pathology in young plants, including reductions in shoot and root biomass. Our results demonstrate that viruses typically found in crops are prevalent in populations of long-lived, drought-tolerant perennials and that the most dominant virus significantly alters plant traits important for survival under challenging abiotic conditions.

Virus community composition. From a plant conservation perspective, it is highly concerning that half of the virus species detected in our focal native plants were not native to the United States or the wild hosts' native range, indicating relatively recent changes in pathogen pressure on these natural populations. Among these non-native viruses are two species that are efficiently transmitted by *A. gossypii*, the main vector species that colonizes and reproduces on cucurbit hosts in our sites: zucchini yellow mosaic virus (ZYMV, family *Potyviridae*, genus *Potyvirus*) and CABYV. ZYMV is non-persistently transmitted by several aphid species and is thought to originate in the Mediterranean (Lisa et al. 1981). Phylogeographic evidence suggests that ZYMV diversity in the United States is likely a result of one introduction from Europe (Simmons et al. 2008). The virus has been present in the Southwest since the early 1980s (Nameth et al. 1983). ZYMV was detected at only one site (Shipley-Skinner Reserve). CABYV is persistently transmitted, mainly by *A. gossypii,* but also to a lesser extent by *Myzus persicae.* This virus was detected at all three sites, mostly in cucurbit hosts, but also in *D. wrightii* (which can support *M. persicae*). CABYV is thought to originate from the Mediterranean (Lecoq et al. 1992) and was first found in the United States infecting numerous cucurbit crops in the San Joaquin, Sacramento, and Salinas valleys of California in 1992 (Lemaire et al. 1993). The widespread nature of the infections detected in 1992, spanning three geographic regions, suggests that the virus may have already been well established in California by this date, including in wild hosts. Through expanded sampling and targeted diagnostics, we found that CABYV is certainly well established in the perennial wild cucurbits; CABYV was detected in up to 88% of sampled wild plants in some locations.

One reason for the high prevalence of CABYV in the natural reserves may be that its main vector, *A. gossypii*, reproduces readily not only on crop plants but also on the two wild cucurbit hosts. We observed *A. gossypii* populations reaching outbreak levels *within* the wild plant community, often on plants with CABYV infection. Since CABYV is transmitted in a persistent, circulative manner by *A. gossypii*, the large populations of vectors could create ample opportunities for within-community transmission. Transmission rates may also be increased by the presence of yellowing symptoms following CABYV infection, which we observed on wild cucurbit hosts in both the field and the laboratory (Supplementary Fig. S8). A study in cultivated cucurbits found that these symptoms are attractive to *A. gossypii* vectors (Carmo-Sousa et al. 2016). In the field, aphids are also attracted to hosts exhibiting yellowing symptoms, increasing the likelihood of vectors colonizing and reproducing on infected hosts (Döring and Chittka 2007). If similar dynamics occurs in the wild cucurbit populations, aphid preferences for infected hosts could partially explain the very high prevalence of CABYV observed at some sites.

CABYV–CYSDV dual infections. We documented high rates of dual infections with CABYV and CYSDV, a second non-native virus. This virus also originates from the Mediterranean region (Célix et al. 1996) and was first detected in the Rio Grande Valley of the United States in 2000 (Kao et al. 2000) and in the Low Desert region adjacent to the Anza Borrego site in 2006 (Kuo et al. 2007; Wintermantel et al. 2017). Experimental transmissions with CYSDV isolates from the Imperial Valley showed that it infects several common noncrop species, including *C. foetidissima* (Wintermantel et al. 2016). Fueled by the previous introduction and establishment of the non-native whitefly *Bemisia tabaci* MEAM1 (Bellows et al. 1994; Johnson et al. 1982) and abundance of domesticated cucurbit crop hosts, CYSDV rapidly established in the region, with nearly 100% infection levels in fall melon crops in the Imperial Valley becoming routine within a few years (Wintermantel et al. 2017). Our data suggest that CYSDV has spilled over from the Imperial Valley into wild cucurbit populations at least as far north as the Motte Rimrock reserve. CABYV and CYSDV often infect individual crop plants at the same time (Kassem et al. 2007), and a recent report documented frequent CYSDV–CABYV dual infections in plants of a wild cucurbit, *Echballium elatarium*, growing adjacent to melon fields in Spain (Maachi et al. 2022). CABYV–CYSDV dual infections appear to be common in wild cucurbits of the southwestern United States as well. In fact, in some locations

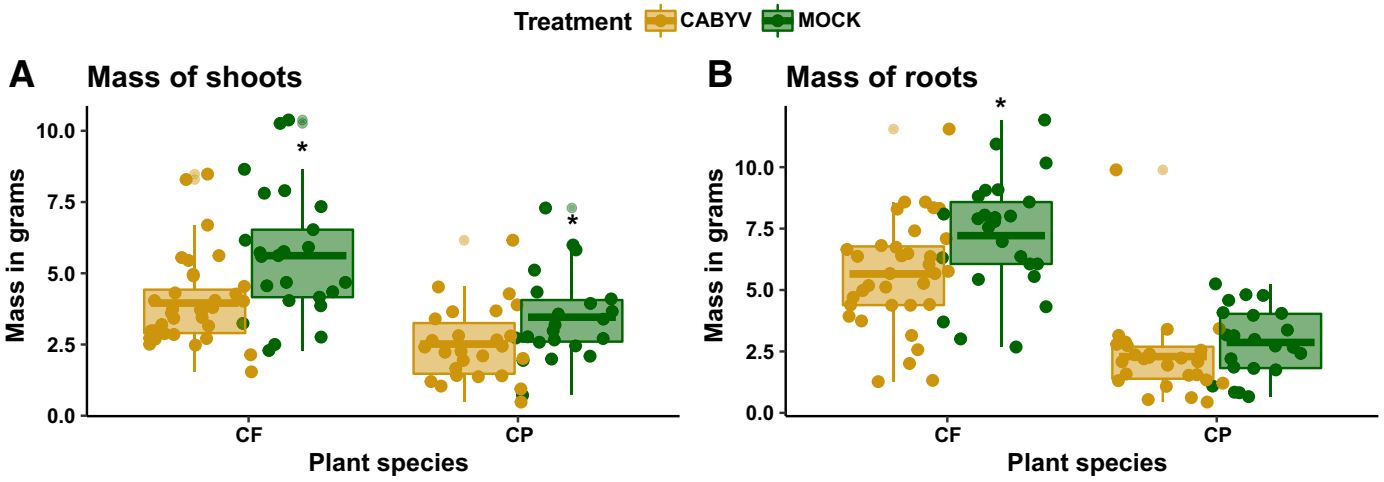

Fig. 7. End-point metrics of plant vigor for the controlled inoculation greenhouse experiment: **A,** mass of shoots; and **B,** mass of roots. ∗ Indicates significant differences relative to the mock treatment at the given time point ($P < 0.05$). CABYV, cucurbit aphid-borne yellows virus; CF, *Cucurbita foetidissima*; and CP, *C. palmata*.

and hosts in our study, dual infections are more common than single infections. At Anza-Borrego Desert State Park, 42% of CABYV-infected *C. foetidissima* and 73% of CABYV-infected *C. palmata* were also infected with CYSDV.

We speculate that CABYV may facilitate CYSDV invasion of the chaparral ecoregion by influencing vector–host interactions. In cultivated cucurbits, CYSDV symptoms are strongly attractive to whitefly vectors (Chesnais et al. 2022; Kenney et al. 2020). While we did not observe symptoms of CYSDV in wild cucurbits, CABYV does induce yellowing symptoms similar to those CYSDV triggers in crops (Kassem et al. 2007; Wintermantel et al. 2017; Supplementary Fig. S8). This may predispose CABYV-infected hosts to infections with CYSDV if immigrating whiteflies preferentially choose these hosts (see Supplementary Table S6 for evidence of whitefly immigration). For the same reason, CYSDV in dual infections with CABYV might also be transmitted more readily than CYSDV from single infections. Confirmation that these two non-native viruses facilitate each other's invasion of native plant communities will require additional behavioral studies with vectors, dual infection studies in the laboratory, and expanded field study.

CABYV effects on plant growth. CABYV reduces growth of the two wild cucurbit hosts during their first year. Symptom onset in wild hosts was rapid and occurred nearly as quickly as it does in crop hosts. Overall symptom severity was higher for *C. palmata*, but both hosts suffered reduced growth because of virus infection. For *C. foetidissima,* this included significantly reduced growth of the tap root, with marginally reduced tap root growth in *C. palmata* as well. Both *Cucurbita* species grow during the summer when little to no rainfall occurs in their sage scrub, chaparral, and grassland habitats. This is possible only because of the deep-reaching tap root and associated lateral roots. *Cucurbita* plants with larger root systems can maintain xylem pressure potential and photosynthetic activity even when other herbaceous flowering plants and grasses experience significant water stress (Knapp and Fahnestock

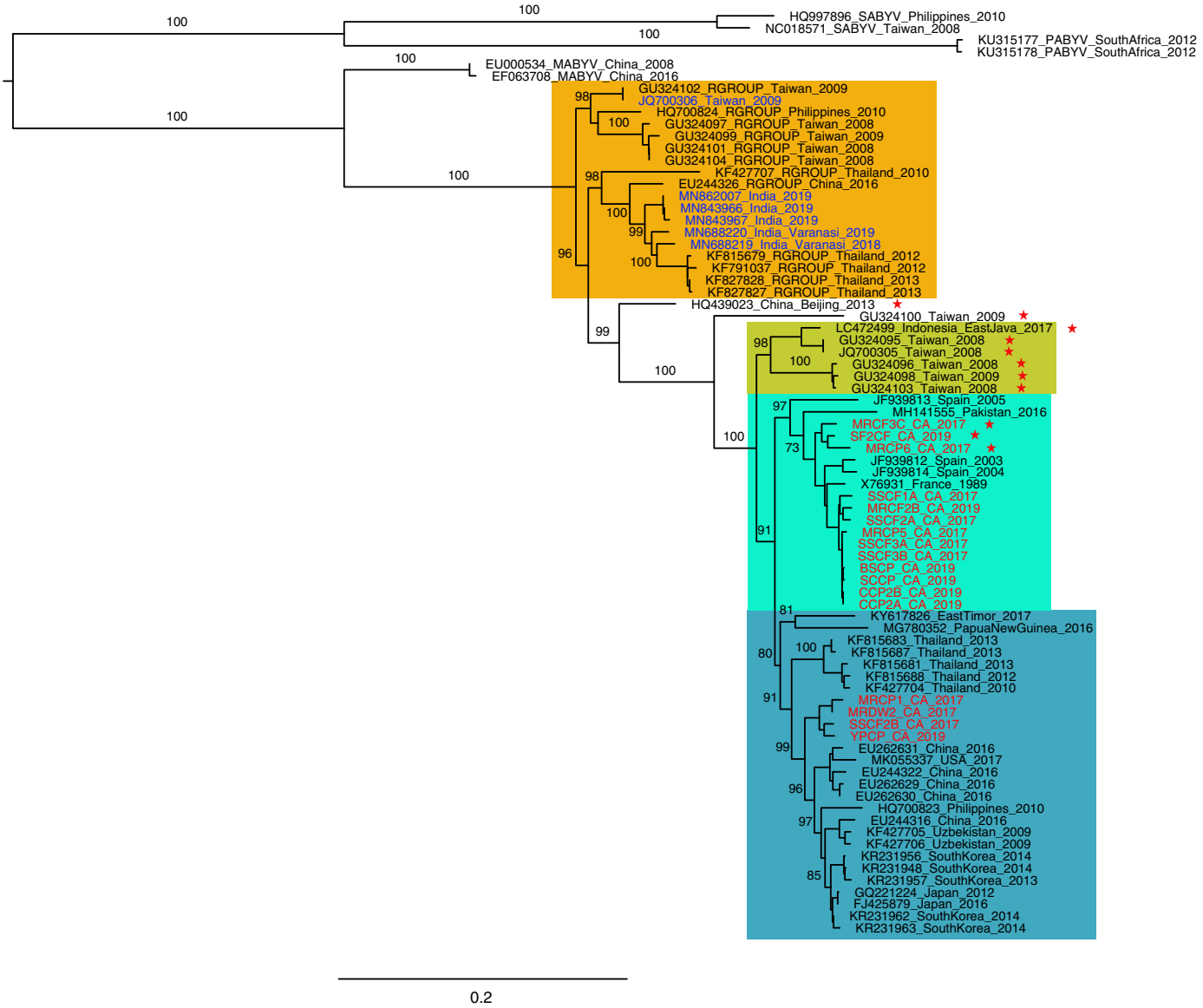

Fig. 8. Maximum likelihood phylogeny based on a concatenated open reading frame (ORF)1-2 (RNA-dependent RNA polymerase) + coat protein (CP; ORF3) dataset highlighting the major clades identified. Bootstrap values >70 are reported above branches. California sequences are in red. Blue text indicates sequences previously identified in GenBank as cucurbit aphid-borne yellows virus (CABYV), but that are actually derived from interspecific recombination between melon aphid-borne yellows virus (MABYV) and CABYV. Red stars denote recombinants other than MABYV and *R* group sequences, which originated from recombination between CABYV and *R* group/MABYV genomes.

1990). But *Cucurbita* plants with small root systems suffer from water stress and must reduce stomatal conductance during midday, sacrificing photosynthetic activity (Knapp and Fahnestock 1990). Our data suggest that CABYV infections in wild cucurbit hosts are maintained across multiple seasons, which might result in cumulative negative impacts on tap root size, photosynthesis, water use efficiency, and overall fitness. Such multiyear fitness impacts have been documented for infection of a native perennial grass species (switchgrass, *Panicum virgatum*) with virus species typically found in cereal crops using controlled inoculations in field-grown plants (Alexander et al. 2017). Similar field experiments with controlled CABYV inoculations could illuminate multiyear impacts of CABYV on wild *Cucurbita* hosts, rates of CABYV *s*pread within populations, and propensity to acquire dual infections over time relative to noninfected hosts in the same locations.

Possible origins of CABYV infecting wild perennials. We do not know when CABYV first infected California wild cucurbit populations. However, our phylogenetic analyses of concatenated RdRp + CP and CP sequences suggests that some California CABYV sequences recovered in this study and previously (Shates et al. 2019) originated from Asia and some from Europe. Thus, the exposure of wild cucurbits in California to CABYV is likely a relatively new development in their ecological history and merits further investigation. We also found that our California sequences do not group with sequences of an isolate recovered from cultivated pumpkin in Oklahoma, which is most closely related to CABYV sequences from China in a prior analysis (Khanal et al. 2021), as well as in our analyses. Unfortunately, there are currently no other publicly available sequences of CABYV from samples collected in the United States. We therefore cannot determine if the CABYV infecting target hosts is distinct from CABYV circulating in crops in the southwestern United States. In a study of CABYV infections in melon fields in Spain, the authors report that a wild cucurbit common in edge habitats, *E. elaterium,* harbors variants of CABYV that are distinct from those in immediately adjacent melon fields (Maachi et al. 2022). The populations sampled for our study do not directly abut melon fields, but rather inhabit remnant native plant communities of various sizes and proximities to crops known to support CABYV (Fig. 1). As discussed above, we also suspect that there is substantial secondary spread within the wild cucurbit populations, and that CABYV had been established in California long before the first official detection (Lemaire et al. 1993). These factors may favor the evolution of distinct variants endemic to the chaparral ecoregion host populations, a possibility that could be explored through expanded sequencing of CABYV samples from crop habitats in the Southwest, as well as from herbarium specimens of wild cucurbit hosts (Malmstrom et al. 2007).

Conclusions. Our study demonstrates that wild perennial plants growing within a region of intensive agricultural production (California) accumulate multiple infections by virus species typically found in crop habitats, particularly non-native pathogens previously introduced to the region from other continents. Both CABYV and CYSDV, the dominant pathogens in our natural sites, likely originated outside of the United States and were introduced into wild vegetation through human activities, where they have become widely prevalent, especially in wild cucurbits. The potential impact of these likely anthropogenic introductions on wild plant ecology is of concern. CABYV is retained by perennial wild cucurbits during dormancy periods and may negatively affect their growth, as established in greenhouse studies. Of increased concern is the suggestion that prior infection with either CABYV or CYSDV may predispose individual plants to secondary virus infections, perhaps by altering host resistance and/or other host traits that influence visitation by vector insects and subsequent virus exposure rates (Mauck et al. 2018). The ecological implications of both single and dual infections by crop-associated viruses in wild perennial hosts are not well understood, but our findings suggest this should be a focus of future studies on this and other systems. The host species studied here are endemic to an ecoregion that is currently experiencing prolonged drought conditions and record-breaking temperature extremes, likely because of human-induced climate change. Moving forward, it will be important to determine to what degree infections by the non-native viruses documented here will influence the long-term persistence of wild cucurbits in their native habitats, as well as of the organisms that depend on them for survival.

ACKNOWLEDGMENTS

We thank Robert Williams, Thomas Ash, and Kenneth Halama for providing access to research sites, assisting with identification of plant populations, and providing reserve statistics; Margarita López-Uribe for providing advice and protocols to improve seed germination for *Cucurbita* species; and two anonymous reviewers for their feedback.

LITERATURE CITED

Afgan, E., Baker, D., Batut, B., van den Beek, M., Bouvier, D., Čech, M., Chilton, J., Clements, D., Coraor, N., Grüning, B. A., Guerler, A., Hillman-Jackson, J., Hiltemann, S., Jalili, V., Rasche, H., Soranzo, N., Goecks, J., Taylor, J., Nekrutenko, A., and Blankenberg, D. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 46:W537-W544.

Alexander, H. M., Bruns, E., Schebor, H., and Malmstrom, C. M. 2017. Crop-associated virus infection in a native perennial grass: Reduction in plant fitness and dynamic patterns of virus detection. J. Ecol. 105:1021-1031.

Bates, D., Mächler, M., Bolker, B., and Walker, S. 2015. Fitting linear mixed-effects models using lme4. J. Stat. Software 67:1-48.

Bellows, T. S., Jr., Perring, T. M., Gill, R. J., and Headrick, D. H. 1994. Description of a species of *Bemisia* (Homoptera: Aleyrodidae). Ann. Entomol. Soc. Am. 87:195-206.

Bernardo, P., Charles-Dominique, T., Barakat, M., Ortet, P., Fernandez, E., Filloux, D., Hartnady, P., Rebelo, T. A., Cousins, S. R., Mesleard, F., Cohez, D., Yavercovski, N., Varsani, A., Harkins, G. W., Peterschmitt, M., Malmstrom, C. M., Martin, D. P., and Roumagnac, P. 2018. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. ISME J. 12:173-184.

Brooks, T. M., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., Rylands, A. B., Konstant, W. R., Flick, P., Pilgrim, J., Oldfield, S., Magin, G., and Hilton-Taylor, C. 2002. Habitat loss and extinction in the hotspots of biodiversity. Conserv. Biol. 16:909-923.

Capblancq, T., and Forester, B. R. 2021. Redundancy analysis: A Swiss Army Knife for landscape genomics. Methods Ecol. Evol. 12:2298-2309.

Carmo-Sousa, M., Moreno, A., Plaza, M., Garzo, E., and Fereres, A. 2016. *Cucurbit aphid-borne yellows virus* (CABYV) modifies the alighting, settling and probing behaviour of its vector *Aphis gossypii* favouring its own spread. Ann. Appl. Biol. 169:284-297.

Célix, A., López-Sesé, A., Almarza, N., Gómez-Guillamón, M. L., and Rodríguez-Cerezo, E. 1996. Characterization of cucurbit yellow stunting disorder virus, a *Bemisia tabaci*-transmitted closterovirus. Phytopathology 86:1370-1376.

Chesnais, Q., Sun, P., and Mauck, K. E. 2022. Advanced infections by cucurbit yellow stunting disorder virus encourage whitefly vector colonization while discouraging non-vector aphid competitors. J. Pest Sci. 95:231-247.

De Cáceres, M., and Legendre, P. 2009. Associations between species and groups of sites: Indices and statistical inference. Ecology 90:3566-3574.

Döring, T. F., and Chittka, L. 2007. Visual ecology of aphids—A critical review on the role of colours in host finding. Arthropod–Plant Interact. 1:3-16.

Döring, T. F., Kirchner, S. M., Kühne, S., and Saucke, H. 2004. Response of alate aphids to green targets on coloured backgrounds. Entomol. Exp. Appl. 113:53-61.

Dray, S. 2019. adespatial: Multivariate multiscale spatial analysis. https://cran.r-project.org/web/packages/adespatial/index.html

Edgar, R. C., Taylor, J., Lin, V., Altman, T., Barbera, P., Meleshko, D., Lohr, D., Novakovsky, G., Buchfink, B., Al-Shayeb, B., Banfield, J. F., de la Peña, M., Korobeynikov, A., Chikhi, R., and Babaian, A. 2022. Petabase-scale sequence alignment catalyses viral discovery. Nature 602:142-147.

Fox, J., and Weisberg, S. 2019. An R Companion to Applied Regression, 3rd ed. Sage, Thousand Oaks, CA.

Hily, J.-M., Poulicard, N., Mora, M.-Á., Pagán, I., and García-Arenal, F. 2016. Environment and host genotype determine the outcome of a plant-virus interaction: From antagonism to mutualism. New Phytol. 209:812-822.

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., and Vinh, L. S. 2018. UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol. 35:518-522.

Hochman Adler, V., Lubin, Y., and Coll, M. 2014. Spillover of crop herbivores into adjacent desert habitats. Agric. Ecosyst. Environ. 193:117-124.

Hughes, L. 2019. RNA precipitation. https://www.protocols.io/view/rna-precipitation-yxmvmx7p6l3p/v1

Johnson, M. W., Toscano, N. C., Reynolds, H. T., Sylvester, E. S., Kido, K., and Natwick, E. T. 1982. Whiteflies cause problems for southern California growers. Calif. Agric. 36:24-26.

Jones, R. A. C. 2014. Plant virus ecology and epidemiology: Historical perspectives, recent progress and future prospects. Ann. Appl. Biol. 164:320-347.

Kahle, D., and Wickham, H. 2013. ggmap: Spatial visualization with ggplot2. R J. 5:144-161.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., and Jermiin, L. S. 2017. ModelFinder: Fast model selection for accurate phylogenetic estimates. Nat. Methods 14:587-589.

Kao, J., Jia, L., Tian, T., Rubio, L., and Falk, B. W. 2000. First report of *cucurbit yellow stunting disorder virus* (genus *Crinivirus*) in North America. Plant Dis. 84:101.

Kassambara, A. 2020. ggpubr: 'ggplot2' based publication ready plots. https://CRAN.R-project.org/package=ggpubr

Kassem, M. A., Sempere, R. N., Juárez, M., Aranda, M. A., and Truniger, V. 2007. *Cucurbit aphid-borne yellows virus* is prevalent in field-grown cucurbit crops of southeastern Spain. Plant Dis. 91:232-238.

Kenney, J. R., Grandmont, M.-E., and Mauck, K. E. 2020. Priming melon defenses with acibenzolar-*S*-methyl attenuates infections by phylogenetically distinct viruses and diminishes vector preferences for infected hosts. Viruses 12:257.

Kesanakurti, P., Belton, M., Saeed, H., Rast, H., Boyes, I., and Rott, M. 2016. Screening for plant viruses by next generation sequencing using a modified double strand RNA extraction protocol with an internal amplification control. J. Virol. Methods 236:35-40.

Khanal, V., Wells, H., and Ali, A. 2021. High prevalence of three potyviruses infecting cucurbits in Oklahoma and phylogenetic analysis of cucurbit aphid-borne yellows virus isolated from pumpkins. Pathogens 10:53.

Knapp, A. K., and Fahnestock, J. T. 1990. Influence of plant size on the carbon and water relations of *Cucurbita foetidissima* HBK. Funct. Ecol. 4: 789-797.

Kuo, Y.-W., Rojas, M. R., Gilbertson, R. L., and Wintermantel, W. M. 2007. First report of *cucurbit yellow stunting disorder virus* in California and Arizona, in association with *cucurbit leaf crumple virus* and *squash leaf curl virus*. Plant Dis. 91:330.

Lacroix, C., Renner, K., Cole, E., Seabloom, E. W., Borer, E. T., and Malmstrom, C. M. 2016. Methodological guidelines for accurate detection of viruses in wild plant species. Appl. Environ. Microbiol. 82:1966-1975.

Lecoq, H., Bourdin, D., Wipf-Scheibel, C., Bon, M., Lot, H., Lemaire, O., and Herrbach, E. 1992. A new yellowing disease of cucurbits caused by a luteovirus, cucurbit aphid-borne yellows virus. Plant Pathol. 41:749-761.

Lefeuvre, P., Martin, D. P., Elena, S. F., Shepherd, D. N., Roumagnac, P., and Varsani, A. 2019. Evolution and ecology of plant viruses. Nat. Rev. Microbiol. 17:632-644.

Legendre, P., and Gauthier, O. 2014. Statistical methods for temporal and space–time analysis of community composition data. Proc. R. Soc. B 281:20132728.

Lemaire, O. J., Gubler, W. D., Valencia, J., Lecoq, H., and Falk, B. W. 1993. First report of cucurbit aphid-borne yellows luteovirus in the United States. Plant Dis. 77:1169.

Lenth, R. V. 2016. Least-squares means: The R package *lsmeans*. J. Stat. Software 69:1-33.

Lisa, V., Boccardo, G., D'Agostino, G., Dellavalle, G., and d'Aquilio, M. 1981. Characterization of a potyvirus that causes zucchini yellow mosaic. Phytopathology 71:667-672.

Ma, Y., Marais, A., Lefebvre, M., Theil, S., Svanella-Dumas, L., Faure, C., and Candresse, T. 2020. Phytovirome analysis of wild plant populations: Comparison of double-stranded RNA and virion-associated nucleic acid metagenomic approaches. J. Virol. 94:e01462-19.

Maachi, A., Donaire, L., Hernando, Y., and Aranda, M. A. 2022. Genetic differentiation and migration fluxes of viruses from melon crops and crop edge weeds. J. Virol. 96:e00421-22.

Maclot, F., Candresse, T., Filloux, D., Malmstrom, C. M., Roumagnac, P., van der Vlugt, R., and Massart, S. 2020. Illuminating an ecological blackbox: Using high throughput sequencing to characterize the plant virome across scales. Front. Microbiol. 11:578064.

Malmstrom, C. M., and Alexander, H. M. 2016. Effects of crop viruses on wild plants. Curr. Opin. Virol. 19:30-36.

Malmstrom, C. M., Bigelow, P., Trębicki, P., Busch, A. K., Friel, C., Cole, E., Abdel-Azim, H., Phillippo, C., and Alexander, H. M. 2017. Crop-associated virus reduces the rooting depth of non-crop perennial native grass more than non-crop-associated virus with known viral suppressor of RNA silencing (VSR). Virus Res. 241:172-184.

Malmstrom, C. M., Hughes, C. C., Newton, L. A., and Stoner, C. J. 2005. Virus infection in remnant native bunchgrasses from invaded California grasslands. New Phytol. 168:217-230.

Malmstrom, C. M., Melcher, U., and Bosque-Pérez, N. A. 2011. The expanding field of plant virus ecology: Historical foundations, knowledge gaps, and research directions. Virus Res. 159:84-94.

Malmstrom, C. M., Shu, R., Linton, E. W., Newton, L. A., and Cook, M. A. 2007. Barley yellow dwarf viruses (BYDVs) preserved in herbarium specimens illuminate historical disease ecology of invasive and native grasses. J. Ecol. 95:1153-1166.

Martin, D. P., Varsani, A., Roumagnac, P., Botha, G., Maslamoney, S., Schwab, T., Kelz, Z., Kumar, V., and Murrell, B. 2021. RDP5: A computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. Virus Evol. 7:veaa087.

Mauck, K. E., Chesnais, Q., and Shapiro, L. R. 2018. Evolutionary determinants of host and vector manipulation by plant viruses. Adv. Virus Res. 101:189-250.

Myers, N., Mittermeier, R. A., Mittermeier, C. G., da Fonseca, G. A. B., and Kent, J. 2000. Biodiversity hotspots for conservation priorities. Nature 403: 853-858.

Nameth, S. T., Dodds, J. A., Paulus, A. O., and Kishaba, A. 1983. A New potyvirus associated with severe disease of cantaloupe (*Cucumis melo*) in southern California. (Abstr.). Phytopathology 73:793.

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Solymos, P., Stevens, M. H. H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Dorcard, D., Garvalho, G., Chirico, M., De Caceres, M., Durand, S., Evangelista, H. B. A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M. O., Lahti, L., McGlinn, D., Ouellette, M.-H., Cunha, E. R., Smith, T., Stier, A., Ter Braak, C. J. F., and Weedon, J. 2017. vegan: Community ecology package. https://CRAN.R-project.org/package=vegan

Perring, T. M., Stansly, P. A., Liu, T. X., Smith, H. A., and Andreason, S. A. 2018. Whiteflies: Biology, ecology, and management. Pages 73-110 in: Sustainable Management of Arthropod Pests of Tomato. W. Wakil, G. E. Brust, and T. M. Perring, eds. Academic Press, San Diego.

R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Rabadán, M. P., Juárez, M., De Moya-Ruiz, C., and Gómez, P. 2021. Aphid-borne viruses infecting cultivated watermelon and squash in Spain: Characterization of a variant of cucurbit aphid-borne yellows virus (CABYV). Plant Pathol. 70:1476-1485.

Roossinck, M. J. 2011. The big unknown: Plant virus biodiversity. Curr. Opin. Virol. 1:63-67.

Roossinck, M. J. 2017. Deep sequencing for discovery and evolutionary analysis of plant viruses. Virus Res. 239:82-86.

RStudio Team. 2020. RStudio: Integrated development for R. RStudio, PBC, Boston.

Sallinen, S., Norberg, A., Susi, H., and Laine, A.-L. 2020. Intraspecific host variation plays a key role in virus community assembly. Nat. Commun. 11:5610.

Schoelz, J. E., and Stewart, L. R. 2018. The role of viruses in the phytobiome. Annu. Rev. Virol. 5:93-111.

Severns, P. M., and Sykes, E. M. 2020. Indicator species analysis: A useful tool for plant disease studies. Phytopathology 110:1860-1862.

Shates, T. M., Sun, P., Malmstrom, C. M., Dominguez, C., and Mauck, K. E. 2019. Addressing research needs in the field of plant virus ecology by defining

knowledge gaps and developing wild dicot study systems. Front. Microbiol. 9:3305.
Simmons, H. E., Holmes, E. C., and Stephenson, A. G. 2008. Rapid evolutionary dynamics of zucchini yellow mosaic virus. J. Gen. Virol. 89:1081-1085.
Susi, H., Filloux, D., Frilander, M. J., Roumagnac, P., and Laine, A.-L. 2019. Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. PeerJ 7:e6140.
Susi, H., and Laine, A.-L. 2021. Agricultural land use disrupts biodiversity mediation of virus infections in wild plant populations. New Phytol. 230:2447-2458.
Susi, H., Sallinen, S., and Laine, A.-L. 2022. Coinfection with a virus constrains within-host infection load but increases transmission potential of a highly virulent fungal plant pathogen. Ecol. Evol. 12:e8673.
Takeshita, M., Okuda, M., Okuda, S., Hyodo, A., Hamano, K., Furuya, N., and Tsuchiya, K. 2013. Induction of antiviral responses by acibenzolar-*S*-methyl against *cucurbit chlorotic yellows virus* in melon. Phytopathology 103:960-965.
Thapa, V., McGlinn, D. J., Melcher, U., Palmer, M. W., and Roossinck, M. J. 2015. Determinants of taxonomic composition of plant viruses at the Nature Conservancy's Tallgrass Prairie Preserve, Oklahoma. Virus Evol. 1: vev007.
Thapa, V., Melcher, U., Wiley, G. B., Doust, A., Palmer, M. W., Roewe, K., Roe, B. A., Shen, G., Roossinck, M. J., Wang, Y. M., and Kamath, N. 2012. Detection of members of the *Secoviridae* in the Tallgrass Prairie Preserve, Osage County, Oklahoma, USA. Virus Res. 167:34-42.
The Galaxy Community. 2022. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. Nucleic Acids Res. 50:W345-W351.
Thioulouse, J., Dray, S., Dufour, A.-B., Siberchicot, A., Jombart, T., and Pavoine, S. 2018. Multivariate Analysis of Ecological Data with ade4. Springer, New York.
Tollenaere, C., Susi, H., and Laine, A.-L. 2016. Evolutionary and epidemiological implications of multiple infection in plants. Trends Plant Sci. 21:80-90.
Vicente-Gonzalez, L., and Vicente-Villardon, J. L. 2021. Package PERMANOVA: Multivariate analysis of variance based on distances and permutations. https://cran.rstudio.com/web/packages/PERMANOVA/PERMANOVA.pdf
Wickham, H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer, New York.
Wickham, H., Hester, J., Chang, W., and Bryan, J. 2022. Package "devtools": Tools to make developing R packages easier. https://cran.r-project.org/web/packages/devtools/devtools.pdf
Wintermantel, W. M., Gilbertson, R. L., McCreight, J. D., and Natwick, E. T. 2016. Host-specific relationship between virus titer and whitefly transmission of *cucurbit yellow stunting disorder virus*. Plant Dis. 100:92-98.
Wintermantel, W. M., Gilbertson, R. L., Natwick, E. T., and McCreight, J. D. 2017. Emergence and epidemiology of *cucurbit yellow stunting disorder virus* in the American Desert Southwest, and development of host plant resistance in melon. Virus Res. 241:213-219.
Xiang, H.-y., Shang, Q.-x., Han, C.-g., Li, D.-w., and Yu, J.-l. 2008. Complete sequence analysis reveals two distinct poleroviruses infecting cucurbits in China. Arch. Virol. 153:1155-1160.
Xu, P., Chen, F., Mannas, J. P., Feldman, T., Sumner, L. W., and Roossinck, M. J. 2008. Virus infection improves drought tolerance. New Phytol. 180:911-921.
Zhao, J., Zhang, X., Hong, Y., and Liu, Y. 2016. Chloroplast in plant–virus interaction. Front. Microbiol. 7:1565.