The first arriving virus shapes within-host viral diversity during natural epidemics

Jokinen, Maija ; Sallinen, Suvi ; Jones, Mirkka M ; Sirén, Jukka ; Guilbault, Emy ; Susi, Hanna ; Laine, Anna-Liisa

DOI: https://doi.org/10.1098/rspb.2023.1486

Title: The first arriving virus shapes within-host viral diversity during natural epidemics

Running title: Viral priority effects

Authors: Maija Jokinen[1], Suvi Sallinen[2], Mirkka M. Jones[2,3], Jukka Sirén[3], Emy Guilbault[2], Hanna Susi[2] & Anna-Liisa Laine[1,2]

[1] Department of Evolutionary Biology and Environmental Studies, University of Zürich, 8057 Zürich, Switzerland

[2] Organismal and Evolutionary Biology Research Programme, Faculty of Biological and Environmental Sciences, PO BOX 65 00014 University of Helsinki, Finland

[3] Institute of Biotechnology, HiLIFE-Helsinki Institute of Life Science, University of Helsinki, PO BOX 65 00014 University of Helsinki, Finland

Email: M.J. maija.jokinen@ieu.uzh.ch, S.S. suvi.sallinen@iki.fi, M.M.J. mirkka.jones@helsinki.fi, J.S. jukka.p.siren@helsinki.fi, E.G. emy.guilbault@helsinki.fi, H.S. hanna.susi@helsinki.fi, A-L.L. anna-liisa.laine@uzh.ch

Correspondence: Maija Jokinen, maija.jokinen@ieu.uzh.ch, Winterthurerstrasse 190, 8057 Zurich, Switzerland, +41787036272

Keywords: virus ecology, priority effects, sequential infection, co-infection

Abstract

Viral diversity has been discovered across scales from host individuals to populations. However, the drivers of viral community assembly are still largely unknown. Within-host viral communities are formed through coinfections, where the interval between the arrival times of viruses may vary. Priority effects describe the timing and order in which species arrive in an environment, and how early colonizers impact subsequent community assembly. To study the effect of the first-arriving virus on subsequent infection patterns of five focal viruses, we set up a field experiment using naïve *Plantago lanceolata* plants as sentinels during a seasonal virus epidemic. Using joint species distribution modelling, we find both positive and negative effects of early season viral infection on late season viral colonization patterns. The direction of the effect depends on both the host genotype and which virus colonized the host early in the season. It is well-established that co-occurring viruses may change the virulence and transmission of viral infections. However, our results show that priority effects may also play an important, previously unquantified role in viral community assembly. The assessment of these temporal dynamics within a community ecological framework will improve our ability to understand and predict viral diversity in natural systems.

Introduction

In recent years, largely due to advancements in metagenomic studies, there has been a rapid increase in the discovery of novel virus species. Metagenomic surveys have revealed high viral diversity in wild, uncultured habitats, and the tremendous complexity of viral communities [1–4]. Despite this increasing knowledge of natural viral diversity, the drivers of viral community assembly are still poorly understood [5,6]. Insights into the mechanisms of viral community assembly are key to an improved understanding of virus ecology and disease dynamics.

Community assembly is expected to be influenced by abiotic filters, spatial structure and biotic interactions [7], and the limited evidence available to date indicates that this also holds true for viruses [8–11]. In many pathogens, changes in temperature and humidity impact their ability to infect, reproduce and transmit within and between their hosts [12–14]. Hence, it is perhaps unsurprising that the prevalence and diversity of several pathogen species – including viruses – follow elevational and latitudinal gradients [15–17]. The effects of climate drivers on viral transmission and activity have been notably studied with viruses that cause clinically significant respiratory tract infections [18,19] and in marine environments [20,21]. In addition to natural environmental variation, human actions modify the environment, directly and indirectly, through climate change [22], habitat fragmentation [23,24] and altered nutrient cycling [25,26], all of which can result in changes in host biodiversity [27]. Biotic factors potentially affecting viral community assembly comprise their hosts and vectors as well as other pathogens, including other virus species and different virus strains. Given that the host acts as the immediate environment for viruses, the spatial distribution of hosts and host diversity – in terms of both interspecific and intraspecific variation – are critical determinants of viral diversity [10,28,29]. For viruses transmitted by vectors, landscape structure and possible disturbance in the vector community may affect vector distribution and thus also viral transmission [30,31].

Virus-virus interactions, whether direct or mediated by the host, are potentially important determinants of viral community composition [32–34]. For example, viral infection in plants can increase vector attraction or alter vector feeding behaviour, indirectly enhancing viral transmission and viral co-occurrences [35–38]. Coinfections by multiple viruses or other pathogens in the same host are common and may result from simultaneous exposure, for example, to a common vector [39,40], intermediate host [41], or from the consumption of prey carrying multiple viruses [42]. However, coinfections are more often thought to establish as a result of sequential infections, where the intervals between infections and the arrival times of pathogens vary [43–45]. Priority effects describe the timing and order in which species arrive in an environment, and how the first arriving species at a site impacts subsequent community development [46,47]. The study of priority effects has its roots in community ecology, where it has been used to understand how species' order of arrival contributes to defining the patterns and relative abundances of co-occurring species [48,49]. As the host delineates a clear habitat into which viral communities assemble [50], priority effects provide an intuitive framework for host-virus and host-pathogen research [51–54].

Inhibitory priority effects describe situations in which an early-arriving species alters the environment such that the establishment of later-arriving species is either suspended or entirely inhibited [55]. Early colonists may, for example, induce the host's immune response, thus inhibiting or delaying infection by species arriving later [56,57]. Inhibitory effects may also be caused by disease symptoms, such as necrosis of host tissue, which render the host environment unfavourable for colonization by other species [46]. In agriculture, cross-protection arising from the deliberate infection of crops with a milder virus strain in order to induce immunity to subsequent infection by more severe strains, is viewed as a promising plant protection tool [56]. In contrast, facilitative priority effects are characterized by the early-

arriving species modifying the host environment to be more favourable for later-arriving species. Defence against the first-arriving species can be costly for the host, leaving the host more susceptible to a secondary infection [58,59]. As an example of cross-kingdom pathogenic facilitative priority effects, respiratory tract infection and lung tissue damage caused by a virus often leads to secondary bacterial infection [60,61]. Furthermore, host immunity is likely an important mediator of priority effects, as recent research suggests that genotypes of the same species vary in their susceptibility to viruses and other pathogens that affect community structure [10,62,63]. By extension, different genotypes of the same host species may also vary in their resistance to sequential infections and in their tendency to harbour coinfections due to genetic differences in their immune responses [64,65].

Here, we investigate how viral infection at the beginning of seasonal epidemics affects subsequent viral community assembly. We utilized data on the early season viral infection status of host plants from a transplant experiment by Sallinen *et al*. (2020) [10] to select plant individuals that differ in their infection status thereby enabling us to assess how this affects subsequent virus community assembly. We focus on successive samples from these same host individuals collected later in the season that have not been previously analysed. Samples were analysed by PCR to detect five focal *Plantago lanceolata*-infecting viruses: *Plantago lanceolata latent virus* (PILV*)* [66]*, Plantago latent caulimovirus*, *Plantago lanceolata betapartitivirus*, *Plantago enamovirus* [67] and *Plantago closterovirus* [67]. The transplant experiment consisted of cloned naïve replicates of four *Plantago lanceolata* genotypes, which were placed into wild *P. lanceolata* populations during seasonal viral epidemics. In their study, Sallinen *et al.* (2020) [10] found both host genotype and local population context to explain a large proportion of variation in early season virus occurrences.

To date, priority effects have not been studied experimentally during naturally occurring viral epidemics In this study, we leverage the early season infection status data of Sallinen *et al* 2020 [10] to explore evidence of priority effects of early arrivals on the late season occurrences of the five focal viruses. We fitted a joint species distribution model (JSDM; [68,69]) to test for priority effects on viral community assembly, whilst controlling for host genotype and other host characteristics. Specifically, we ask: 1) Does the early infection status of the host shape subsequent viral occurrence patterns? 2) Can we identify inhibitory or facilitative priority effects depending on which virus infected the host first? and 3) Are viral priority effects associated with particular host genotypes?

Materials and Methods

Study species

The host species, *Plantago lanceolata*, a perennial herbaceous plant, is an obligate out-crosser capable of reproducing both clonally by side rosettes and sexually [70]. *Plantago lanceolata* occurs worldwide. In the Åland Islands, SW Finland, the plant forms a network consisting of c. 4000 small fragmented populations [71]. These populations have been monitored since 1990 for the presence of *Melitaea cinxia* butterflies and the epidemiological dynamics of the *P. lanceolata* infecting fungus *Podosphaera plantaginis* have been studied since 2001 [71,72]. Host-virus interactions in the system have been studied since 2017, revealing diverse viral communities [10,33,66,67].

To investigate priority effects in viral community assembly during seasonal viral epidemics we focused on five common, recently characterized viruses from the Åland Islands:

DNA viruses *Plantago lanceolata latent virus* (PlLV) in the genus *Capulavirus* [66] and *Plantago latent caulimovirus* [67] in the genus *Caulimovirus*, and RNA viruses *Plantago betapartitivirus* in the genus *Betapartitivirus*, *Plantago enamovirus* in the genus *Enamovirus*, and *Plantago closterovirus* in the genus *Closterovirus* [66,67]. For clarity, we will refer to these viruses hereafter as PlLV*, Caulimovirus*, *Closterovirus*, *Betapartivirus* and *Enamovirus.* The five focal viruses are among the most prevalent *P. lanceolata*-infecting viruses in the Åland system [33] and have been characterized by small RNA (sRNA) sequencing, which targets the small interfering RNA (siRNA), a plant immune response to viral infection [73]. Viruses from the genera *Capulavirus*, *Caulimovirus*, *Enamovirus* and *Closterovirus* are transmitted via aphid vectors [74–77] and are considered plant-specific, whereas viruses in the genus *Betapartitivirus* can infect both plants and fungi [78,79]. Symptoms associated with the five focal viruses are unclear, but can include yellowing, redness, curliness and necrotic lesions [67]. Susi *et al.* [67] linked PlLV infection with yellowing of the leaf.

Preparation of sentinel plant material and field experiment

To study priority effects during seasonal viral epidemics, we placed a total of 320 sentinel plants into four wild populations of *P. lanceolata* in the Åland Islands, Finland. To ensure genetically homogenous host material, 80 replicate sentinel plants were cloned from each of four greenhouse-grown maternal plants. The cloned individuals were assumed to represent four distinct *P. lanceolata* genotypes as the maternal plants originated from populations 7 to 40 kilometers apart (genotype IDs: 609_19; 4_13; 511_14; 2929_6). In short, four-week-old maternal plants potted with a 1:1 proportion of potting soil and sand were placed on top of 11 cm × 11 cm pots filled with vermiculate. The plants and pots were then placed on trays filled with water. When roots had grown through to the bottom pot, the roots were cut and allowed to sprout within the bottom pot. Shoots from the cut roots were individually planted into new 10 cm × 10 cm pots and grown for two additional months in the greenhouse. For more details on cloning of the host material, see Sallinen *et al.* (2020). To exclude possible seed-derived viral infection, the maternal plants were confirmed to be free of the five target viruses by PCR.

In the last week of May 2017, the cloned individuals were placed into four wild *P. lanceolata* populations in the Åland Islands. The populations (ID:s 877, 9031, 433, 3302) are located in different parts of the Åland Islands and do not include the populations of origin of the maternal plants. We placed 80 cloned plants into each of the four populations, 20 replicates of each of the four maternal genotypes. The experiment hence included a total of 320 sentinel plant individuals. The plants were randomly placed within natural vegetation in the chosen meadows and separated from the local soil by placing the plant pots inside plastic boxes (approximately 13 cm × 11 cm). To minimize within-population spatial effects, the placement of the plants among the plastic boxes was shuffled three times per week. The plants were watered when needed.

Two weeks after introducing the sentinel plants into the wild populations, we recorded signs of herbivory (holes, bite marks, and thrip damage), counted the number of leaves and measured the width and length of the largest leaf. Based on these measurements we calculated plant size as $n \times A$, where n is the number of leaves and $A = \pi ab$, where a is the half axis of the width of the largest leaf, and b is the half axis of the length of the largest leaf. For RNA and DNA extraction, 3 cm^2 and 1 cm^2 pieces of leaf tissue were collected, respectively. The samples for DNA extraction were kept in a cold bag and the sample for RNA extraction was snap frozen with liquid nitrogen on site. After the first sampling date, hereafter the early season timepoint, we repeated DNA and RNA sample collection at two-week intervals until the last week of July 2017 (sampling timepoints 2-4, hereafter T2-T4).

Nucleic acid extraction and virus detection in samples from the field experiment

To detect DNA or RNA of the target viruses from the leaf tissue samples, we first assigned the samples to nucleic acid extractions. Total DNA was extracted from a 1 cm^2 leaf tissue sample with E.Z.N.A. Plant Kit (Omega Biotek, USA) at the Institute of Biotechnology, University of Helsinki, Finland, following the manufacturer's instructions. The leaf tissue prior to extraction and the extracted total DNA were stored at -20 °C. For total RNA extraction, we used a modified protocol from Chang et al. (1993) with additional acid phenol (pH 4.5) clean-up steps [80]. In short, a 3 cm^2 piece of leaf tissue was ground into a fine powder using liquid nitrogen, then 800 µl of warm extraction buffer was added, and the sample was mixed thoroughly. After this, 800 µl of acid phenol-chloroform-IAA solution (with ratios: 25:24:1, respectively) was added and centrifuged at 13 500 rpm for 15 minutes in RT. The supernatant was transferred to a new tube and the clean-up step with acid phenol-chloroform-IAA solution was repeated. At the end of the extraction, the purified RNA pellet was resuspended in 25 µl of nuclease-free water. The leaf tissue prior to RNA extraction and the extracted RNA were stored at -80 °C. The extracted RNA was translated into cDNA before PCR. For cDNA translation, 2 ng of RNA was combined with 2 µl of 50 µM random hexamer primers (Promega Corporation, USA), nuclease-free water was added to a final volume of 17.125 µl and the reaction was incubated at 70 °C for 5 min. For the reverse transcription reaction, 1 µl of Moloney Murine Leukemia Virus Reverse Transcriptase (M-MLV RT; Promega Corporation, USA), 5 µl of M-MLV RT 5x buffer (Promega Corporation, USA), 1.25 of 10 mM dNTP mix (Thermo Fischer) and 0.625 µl of RiboLock RNase inhibitor (Thermo Fischer Scientific, USA) were added to the reaction mix and incubated for 60 min at 37 °C.

To detect PlLV and *Caulimovirus* DNA and *Betapartitivirus, Enamovirus* and *Closterovirus* RNA from the samples, we used previously described primers for each virus [9, 63, 64]. Primer sequences can be found in the Supplement (Supplementary Table 1). The PCR reaction consisted of 500 nmol of the corresponding reverse and forward primers, 1 µl of template DNA or cDNA, 5 µl of GoTaq Green® 5x Mastermix (Promega Corporation, USA) and nuclease-free water to a final reaction volume of 10 µl. The PCR conditions were initial denaturation at 95 °C for 2 min, followed by 35 cycles at 95 °C for 2 min, 50-60 °C for 40 s and 72 °C for 1 min. The final extension was done at 72 °C for 5 min. A positive control and a water control were included in each PCR run. The sizes of the amplicons were analysed on 1.5 % agarose gel, stained with GelRed (Biotium, USA) and visualized using the Bio-Rad Gel Doc XR+ imaging system (Bio-Rad Laboratories, USA).

Statistical analysis

To investigate how host early season characteristics affect subsequent viral occurrence patterns, we pooled the detected occurrences of all five viruses from the three late season sampling dates (T2-T4) into a single late season timeperiod for the statistical analysis. Data pooling was done to obtain an observed viral community for each host individual from the late season and to maximise sample size, and hence statistical power over the late season timeperiod. To explore the effects of early season viral colonization and host genotype on late season viral occurrences in hosts, we ran a JSDM using the hierarchical modelling of species communities (HMSC; [68,81]) framework. HMSC is a Bayesian Hierarchical Generalised Linear Mixed Model, in which the responses of species to ecological variables are modelled

with a combination of community- and species-level parameters [68]. In addition to the pooled model, we fitted models with the viral response data from each of the late season sampling timepoints (T2-T4) separately, to test whether the pooled model results were representative of these, or whether, for example, there is evidence that priority effects set in motion in the early summer either weaken or strengthen over time.

The response variables in our HMSC model were vectors of the late season occurrences of each of the five focal viruses recorded in the host individuals. As fixed effect predictors, we included 1) binary data on the early season occurrences of each of the three focal viruses detected at that timepoint (i.e., *Closterovirus*, *Betapartitivirus* and PlLV, see Results for details), 2) host genotype, 3) log-transformed host plant size, 4) the presence/absence of herbivory and 5) the number of failed RNA samples, if any. The latter variable was included to account for minor differences among host plants in the likelihood of recording the late season presence of RNA viruses, arising from the fact that data on RNA virus presence (*Closterovirus*, *Betapartitivirus* and *Enamovirus*) was unavailable for 29 late season samples. As random effect predictors, we included 1) transplant population identity and 2) individual host identity. Host individual identity was included as a random effect to allow us to quantify any residual (unexplained) structure in late season viral occurrences at the plant level. To account for any spatial dependence of observations in each of the four populations, we included population identity as a random effect. We sampled the posterior distribution with four independent Markov chain Monte Carlo (MCMC) chains, each run for 1,875,000 iterations, and discarded the first 625,000 as burn-in. The remaining iterations were thinned by 5000 to yield 250 posterior samples per chain, and thus 1000 posterior samples in total.

We examined model fit by evaluating both explanatory performance and predictive performance, assessed via ten-fold cross-validation, as a function of Tjur R^2 (Tjur's coefficient of determination) [82] and AUC (Area Under the Curve) [83]. We ran our analyses using the R-package Hmsc 3.0-13. [69]. For all analyses we used R version 4.3.0. The statistical analysis pipeline and input data are available on GitHub (https://github.com/mirkkajones/ViralCommHmsc).

We addressed our first research question by exploring how much variation (%) in the late season occurrences of each five focal viruses in the host was explained by early season viral infection status using an extended variance partitioning summary from HMSC (mean and 95% credible interval). We addressed our second research question by exploring the strength and direction of the predicted effects of each early season viral covariate on late season viral occurrences, based on the beta parameters of the regression models. Positive and negative responses to the early season viral covariates with strong (>95%) posterior support were considered as evidence of positive (facilitative) and negative (inhibitory) priority effects, respectively. We addressed the third question by exploring whether the predicted late season colonization patterns of each virus as a function of early season viral infections differed among the four host plant genotypes.

Results

Virus detection from the host individuals by PCR

In the early season timepoint, 39% of the plants (n=320) were infected by a single virus, 14% harboured a coinfection and 47% were uninfected (more detailed description of all viral infections in the early season timepoint in Supplementary Table 2). To investigate how priority

effects shape the late season viral community, we selected a total of 110 plant individuals with a single virus infection by PlLV (n = 31), *Closterovirus* (n = 38) or *Betapartitivirus* (n = 41; Fig. 1) at the early season timepoint and, for comparison, 103 plant individuals with no viral infection at that time. Hence, our study included a total of 213 plant individuals, which were furthermore selected to represent the four plant genotypes as equally as possible (genotypes; 511_14: n = 45, 609_19: n = 53, 2818_6: n = 54, 4_13: n = 61). Among the pooled late season samples (n=639 pooled viral occurrences across the three late season sampling dates), 33% of the sentinel plants were uninfected, 35% had a single infection and 31% were coinfected by multiple viruses (Fig. 2). We detected all five focal viruses in the late season timeperiod; PlLV was the most abundant virus, accounting for 35% of all infections in sentinel plants, whereas *Enamovirus* was least abundant, with an 8% infection frequency. The late season viral communities were variable in composition, ranging from single infections to coinfections with as many as four viruses. We detected 21 different virus combinations in total (Fig. 2B).

Analysis of viral priority effects

Late season viral occurrences were well predicted, although model performance varied among the five focal viruses. PlLV occurrences were most reliably predicted and *Betapartitivirus* occurrences were least reliably predicted (Table 1). The mean explanatory power of our models in terms of Tjur R^2 was 0.23 (range among viruses 0.12 – 0.56) and the mean AUC was 0.86 (range 0.80 – 0.92). Model predictive power based on ten-fold cross-validation was slightly lower, with a mean Tjur R^2 of 0.16 (range among viruses 0.05 – 0.52) and a mean AUC of 0.71 (range 0.62 – 0.84) (Table 1). Lastly, the explanatory power of our models from each of the late season sampling timepoints (T2-T4) in terms of Tjur R^2 were lower than when compared to the pooled model and varied among the five focal viruses (see details in Supplementary Table 3). Mean AUC values resembled those of the pooled model (T2 = 0.89, T3 = 0.92 and T4 = 0.89; Supplementary Table 3).

In terms of contributions to explained variation, host plant genotype was the strongest determinant of mean late season viral occurrences explaining 39% of variance (Fig. 3, Table 2). This was followed by the early season infection status of the host, which explained 18% of the variance in late season viral occurrences. Plant size and the presence/absence of herbivory also contributed to explained variance (8% and 2% respectively, Fig. 3, Table 2). Finally, the number of failed RNA samples during the late season explained 13% of variance. The results for the models of each sampling timepoint individually (T2-T4) showed similar results, where the host genotype explained most of the variation, however, there were some differences between timepoints. Model results from sampling timepoint 4 were very similar to those of the pooled model (see Supplementary Fig. 1 showing explained variance of the total variance explained by all variables in the T2-T4 models).

The effects of host genotype on late season viral occurrences varied strongly among viruses. Genotype effects were pronounced for the two DNA viruses, PlLV and *Caulimovirus* (88% and 62% of variance explained by genotype, respectively; Fig. 3, Table 2), both of which were favored by host genotype 609_19 (Fig. 4). Host genotype played a less pronounced role in defining the colonization patterns of the three RNA viruses, *Betapartitivirus*, *Closterovirus* and *Enamovirus* (Fig. 3, Table 2). However, genotype 4_13 was more likely to host *Betapartitivirus* infections than the other genotypes (Fig. 4).

The five focal viruses also differed in their sensitivity to prior viral infection, as indicated by the summed proportion of variance explained by prior infection by any early

season virus (in decreasing order: *Enamovirus* 36%, *Closterovirus* 29%, *Betaparitivirus* 14%, *Caulimovirus* 7% and PlLV 6%; Fig. 3, Table 2). Infection by *Betapartitivirus* at the beginning of the season had the greatest impact on subsequent colonization dynamics as revealed by strongly supported ($P > 0.95$) negative and positive effects on late season viral colonization. Plant individuals colonized early by *Betapartitivirus* were less likely to host *Closterovirus* or *Enamovirus* in the late season, and were more likely to host PlLV (Fig. 4; see also estimated beta-parameter values and their credible intervals in Supplementary Table 4 and Supplementary Fig. 2). Furthermore, the presence of PlLV in the early season increased the probability of detecting *Enamovirus* in the late season (Fig. 4, Supplementary Table 4).

To illustrate the predicted effects of early season infection status on late season viral prevalence per host plant genotype, we generated marginal effects prediction plots based on the HMSC model, where all covariates apart from the focal early season virus and host genotype were set at their mean value in the dataset (Fig. 5, see Supplementary Fig. 3 for the predicted effects of the other covariates). As expected, based on the predicted responses to each viral covariate described above (Fig. 3) initial infection by *Betapartitivirus* is predicted to facilitate sequential PlLV infection. In addition, inhibitory priority effects of *Betapartitivirus* on both *Closterovirus* and *Enamovirus* are predicted. Early colonization by PlLV, in turn, is predicted to facilitate sequential infection by *Enamovirus*. Plant genotype furthermore affects the absolute predicted prevalence of each virus, for example, host genotype 609_19 is predicted to harbour PlLV at a very high prevalence.

Host plant size explained 8% of variance (Fig. 3, Table 2) and was predicted to have a positive effect on the late summer occurrence probability of *Closterovirus* (Fig. 4). Herbivory explained least variance in our model, 2%, and was not connected to any specific host genotype (Fig. 4). Lastly, the number of failed RNA samples during the late summer explained 13% of variance, and was predicted to have a negative effect – as anticipated – on the detection of late season occurrence probabilities of two of the three RNA viruses, *Enamovirus* and *Betapartitivirus* (Fig. 4).

We also detected some residual structure in late season viral occurrences that was unexplained by the fixed effect covariates, both among plant individuals and among the four populations, mean variance explained 10% and 9%, respectively (Fig. 3, Table 2). This unexplained structure could reflect the effects of 1) missing ecological drivers, or 2) structure arising from direct positive or negative interactions between the viruses that we did not capture with our three early season virus variables. However, we do not see any strongly supported positive or negative residual correlations between the late season occurrences of the viruses that would back the second interpretation.

Discussion

Prior research has demonstrated that priority effects – defined by the sequence of arrival of species – shape the trajectory of community assembly in many systems [51,84,85]. Much like free-living organisms, viruses, and other pathogens, can form diverse and complex assemblages but this assembly process takes place within their hosts. Our results suggest that priority effects can arise in viral assemblages as a result of sequential host infection. By combining a transplant experiment with a hierarchical joint model of viral occurrences, we were able to detect both facilitative and inhibitory priority effects among viruses, depending on which virus first infected the host early in the season. Furthermore, by transplanting naïve host clones, we were able to control for host genotypic variation and most importantly the timing of initial viral infection. However, we acknowledge that our study does not address the

effects of potential later season co-infections on viral colonization dynamics. Our findings nonetheless highlight the importance of biotic interactions – encompassing both the arrival order of viruses and host genotype – in determining viral community assembly.

Our study focused on five *P. lanceolata* infecting viruses that previous studies have shown to be among the most abundant viruses in this study system [68]. The transplanted sentinel plants harboured relatively high infection rates; 53% of the plants individuals were infected by one or more viruses early in the season and 66% in the late season time period. The high virus prevalence in this wild plant system resembles that of agricultural settings, where virus-free plants acquire virus infections during the growing season, highlighting the importance of seasonal transmission for viruses [86,87]. Our model results indicate that, after accounting for the effect of host genotype, the most important factor explaining viral colonization patterns in the late season was the early season viral infection status of the host. We found that early infection with *Betapartitivirus* inhibited subsequent colonization by *Closterovirus* and *Enamovirus* but facilitated PlLV infection. Moreover, PlLV colonization early in the season facilitated *Enamovirus* infection later in the season. Interestingly, we did not find clear indications that individual viruses are consistently facilitative or inhibitory. *Betapartitivirus* was involved both in inhibitory and facilitative priority effects. *Betapartitivirus* belongs to the family Partitiviridae, which is known to cause persistent viral infections [88] and to co-occur with other dsRNA viruses [89], suggesting that the role of *Betapartitivirus* in community assembly is often likely to be facilitative. Viruses from the Partitiviridae family are generally believed to be transmitted vertically through the germline [88]. While our cloned sentinel plants were initially virus-free, we detected *Betapartitivirus* infections in the late season, suggesting a non-vertical transmission route. In addition to plants, *Betapartitivirus* can also infect fungi and protozoa, and hence, a possible transmission route could be from fungi to plants [78,90]. PlLV is more comprehensively characterized than the other four viruses studied here and, based on our previous research, is one of the most prevalent viruses infecting *P. lanceolata* in the Åland Islands. Given the overall prevalence of PlLV in this system [24,66,67], and its tendency for facilitation, PILV may be an important biotic promotor of within-host viral diversity. By focusing on the five focal *P. lanceolata* infecting viruses for which we have developed PCR primers, we acknowledge that some viruses were likely excluded from the scope of this study. The effects of other unidentified viruses, as well as other micro-organisms, could of course contribute to some of the unexplained variation in our model.

In addition to viral priority effects, we detected differences among *P. lanceolata* genotypes in their probability of late season viral infection which are in line with earlier findings in this system [10]. Plant genotype was clearly the strongest determinant of late season viral colonization patterns, explaining on average 39% of the variation in our HMSC model (Fig. 3). We found similar genotype effect in our models for each sampling timepoint individually (Supplementary Fig. 1). Furthermore, the magnitude and direction of viral facilitative or inhibitory priority effects differed among genotypes (Fig. 4 and 5). The effect of genotype was most pronounced for genotypes 609_19 and 4_13, which were susceptible to PlLV, *Caulimovirus* and *Betapartitivirus*. Similar host genotypic effects on infection probabilities by viruses and other parasites have been detected across a range of host species [10,91–93]. By using cloned individuals, we were able to account for genotypic variation and to ensure homogenous host material. Previous studies from the Åland Islands system have shown *P. lanceolata* genotypes to harbour extensive diversity in resistance to powdery mildew *P. plantaginis* [29,94,95]. This observation, together with our new results, suggests that genotypic variation in resistance in *P. lanceolata* also holds true for viruses. As the sentinel plants were placed within natural *P. lanceolata* populations, viral prevalence in the transplant

populations may also have influenced the observed viral occurrences [96,97]. There is a growing body of evidence that insect vectors may prefer hosts that are already infected, enhancing viral spread and promoting co-infections [36–38,98–100]. Also here, the effect of plant genotype may partially arise from the plant virus vectors preferring specific host genotypes.

Viruses have been traditionally thought of only as infective agents causing disease, and indeed viruses cause severe diseases in humans as well as in domestic animals and crops [101,102]. Viral disease outbreaks can result in humanitarian crises and significant economic losses [103,104]. However, recent advances in *virome* research suggest that, in fact, the role of viruses in ecosystems is multifaceted [57]. For example, viruses contribute to important ecosystem services such as nutrient cycling [106]. Furthermore, viral communities harbour a tremendous pool of genetic variety and, through infection, maintain variation in host organisms as well [107]. It is well established that co-infecting viruses may strongly impact viral infection progression within hosts, as well as transmission among hosts [32,108]. Hence, understanding the drivers of viral community assembly can help us to understand virus ecology and, more specifically, the role of viral communities in health and in disease [55, 56]. To our knowledge, our study is the first field experiment to study viral priority effects in the wild. Our results highlight the importance of the infection history and genotype of the host individual in shaping viral community assembly, but also the importance of studying virus ecology at the community level. Our study shows that past infections may, in some cases, determine the course of future infections, bringing us one step closer to understanding the dynamics of viral communities.

Acknowledgements

We thank Mikko Jalo, Pauliina Hyttinen, Mikko Immonen and Vanja Milenkovic for helping with data collection in the field. We thank Pauliina Hyttinen and Laura Häkkinen for help with RNA extractions. We thank Krista Raveala for help in cloning and caring for the plants. The Institute of Biotechnology at University of Helsinki is acknowledged for carrying out DNA extractions. The work was funded grants from the European Research Council (Consolidator Grant RESISTANCE 724508), SNSF (310030_192770/1) and Academy of Finland (334276) to A.-L.L., Academy of Finland (321441) to H.S., Luova Doctoral Programme Fellowship to S.S. and Jane and Aatos Erkko Foundation, Ministyry of Environment, Finland (through the BIOMON research program) and research Council of Finland (317255) to E.G.. M.M.J. and J.S. were supported by the Academy of Finland's 'Thriving Nature' research profiling action. The CSC – IT Center for Science, Finland is acknowledged for computational resources.

Author Contributions

S.S., A.-L.L., and H.S. designed the study. S.S. performed the experiment and data collection. M.J. and S.S. performed the molecular lab work. M.M.J. and J.S. performed the statistical analysis. E.G. developed the code for the full posterior distribution of the variance partitioning and its summary. M.J., M.M.J. and A-L.L. wrote the first draft of the paper and all authors contributed to the final version.

Conflict of Interest

448 The authors declare no conflict of interests.

449

450 **Data Availability Statement**

451 The data and R scripts used in this study have been submitted to GitHub
452 (https://github.com/mirkkajones/ViralCommHmsc) and to Dryad (DOI:
453 https://doi.org/10.5061/dryad.rr4xgxdf2).

454

455 **References**

456 1. Breitbart M, Salamon P, Andresen B, Mahaffy JM, Segall AM, Mead D, Azam F,
457 Rohwer F. 2002 Genomic analysis of uncultured marine viral communities. *Proc. Natl.*
458 *Acad. Sci. U. S. A.* **99**, 14250–14255. (doi:10.1073/pnas.202488399)

459 2. Breitbart M, Hewson I, Felts B, Mahaffy JM, Nulton J, Salamon P, Rohwer F. 2003
460 Metagenomic analyses of an uncultured viral community from human feces.
461 *J. Bacteriol.* **185**, 6220–6223. (doi:10.1128/jb.185.20.6220-6223.2003)

462 3. Mushegian AR. 2020 Are There 1031 Virus Particles on Earth, or More, or Fewer?
463 *J. Bacteriol.* **202**. (doi:10.1128/jb.00052-20)

464 4. Edwards RA, Rohwer F. 2005 Viral metagenomics. *Nat. Rev. Microbiol.* **3**, 504–510.
465 (doi:10.1038/NRMICRO1163)

466 5. Cobián Güemes AG, Youle M, Cantú VA, Felts B, Nulton J, Rohwer F. 2016 Viruses
467 as Winners in the Game of Life. *Annu. Rev. Virol.* **3**, 197–214.
468 (doi:10.1146/ANNUREV-VIROLOGY-100114-054952)

469 6. Liang G, Bushman FD. 2021 The human virome: assembly, composition and host
470 interactions. *Nat. Rev. Microbiol.* **19**, 514–527. (doi:10.1038/s41579-021-00536-5)

471 7. Nemergut DR *et al.* 2013 Patterns and Processes of Microbial Community Assembly.
472 *Microbiol. Mol. Biol. Rev.* **77**, 342–356. (doi:10.1128/mmbr.00051-12)

8. Johnson PTJ, De Roode JC, Fenton A. 2015 Why infectious disease research needs community ecology. *Science (80-.).* **349**. (doi:10.1126/science.1259504)

9. Wille M, Eden JS, Shi M, Klaassen M, Hurt AC, Holmes EC. 2018 Virus–virus interactions and host ecology are associated with RNA virome structure in wild birds. *Mol. Ecol.* **27**, 5263–5278. (doi:10.1111/MEC.14918)

10. Sallinen S, Norberg A, Susi H, Laine AL. 2020 Intraspecific host variation plays a key role in virus community assembly. *Nat. Commun.* **11**, 1–11. (doi:10.1038/s41467-020-19273-z)

11. Andersson AF, Banfield JF. 2008 Virus population dynamics and acquired virus resistance in natural microbial communities. *Science (80-.).* **320**, 1047–1050. (doi:10.1126/SCIENCE.1157358/SUPPL_FILE/ANDERSSON_SOM.PDF)

12. Garren M, Son K, Tout J, Seymour JR, Stocker R. 2015 Temperature-induced behavioral switches in a bacterial coral pathogen. *ISME J. 2015 106* **10**, 1363–1372. (doi:10.1038/ismej.2015.216)

13. Deyle ER, Maher MC, Hernandez RD, Basu S, Sugihara G. 2016 Global environmental drivers of influenza. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 13081–13086. (doi:10.1073/PNAS.1607747113/SUPPL_FILE/PNAS.1607747113.SD01.TXT)

14. Romero F, Cazzato S, Walder F, Vogelgsang S, Bender SF, van der Heijden MGA. 2022 Humidity and high temperature are important for predicting fungal disease outbreaks worldwide. *New Phytol.* **234**, 1553–1556. (doi:10.1111/NPH.17340)

15. Bergner LM, Orton RJ, Benavides JA, Becker DJ, Tello C, Biek R, Streicker DG. 2020 Demographic and environmental drivers of metagenomic viral diversity in vampire bats. *Mol. Ecol.* **29**, 26–39. (doi:10.1111/mec.15250)

16. Guernier V, Hochberg ME, Guégan JF. 2004 Ecology Drives the Worldwide Distribution of Human Diseases. *PLOS Biol.* **2**, e141. (doi:10.1371/JOURNAL.PBIO.0020141)

17. Jones KE, Patel NG, Levy MA, Storeygard A, Balk D, Gittleman JL, Daszak P. 2008 Global trends in emerging infectious diseases. *Nature* **451**, 990–993. (doi:10.1038/nature06536)

18. Baumgartner EA *et al.* 2012 Seasonality, Timing, and Climate Drivers of Influenza Activity Worldwide. *J. Infect. Dis.* **206**, 838–846. (doi:10.1093/INFDIS/JIS467)

19. Lowen AC, Mubareka S, Steel J, Palese P. 2007 Influenza Virus Transmission Is Dependent on Relative Humidity and Temperature. *PLOS Pathog.* **3**, e151. (doi:10.1371/JOURNAL.PPAT.0030151)

20. Danovaro R, Corinaldesi C, Dell'Anno A, Fuhrman JA, Middelburg JJ, Noble RT, Suttle CA. 2011 Marine viruses and global climate change. *FEMS Microbiol. Rev.* **35**, 993–1034. (doi:10.1111/J.1574-6976.2010.00258.X)

21. Suttle CA. 2005 Viruses in the sea. *Nature* **437**, 356–361. (doi:10.1038/NATURE04160)

22. Laufkötter C, Zscheischler J, Frölicher TL. 2020 High-impact marine heatwaves attributable to human-induced global warming. *Science (80-.).* **369**. (doi:10.1126/SCIENCE.ABA0690/SUPPL_FILE/ABA0690_LAUFKOTTER_SM.PDF)

23. Alexander HM, Mauck KE, Whitfield AE, Garrett KA, Malmstrom CM. 2014 Plant-virus interactions and the agro-ecological interface. *Eur. J. Plant Pathol.* **138**, 529–547. (doi:10.1007/s10658-013-0317-1)

519 24. Susi H, Laine AL. 2020 Agricultural land use disrupts biodiversity mediation of virus
520 infections in wild plant populations. *New Phytol.* **230**, 2447–2458.
521 (doi:10.1111/nph.17156)

522 25. Hungate BA, Dukes JS, Shaw MR, Luo Y, Field CB. 2003 Nitrogen and Climate
523 Change. *Science (80-.).* **302**, 1512–1513.
524 (doi:10.1126/SCIENCE.1091390/SUPPL_FILE/HUNGATE.SOM.PDF)

525 26. Swann GEA *et al.* 2020 Changing nutrient cycling in Lake Baikal, the world's oldest
526 lake. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 27211–27217.
527 (doi:10.1073/PNAS.2013181117/SUPPL_FILE/PNAS.2013181117.SD04.XLSX)

528 27. Civitello DJ *et al.* 2015 Biodiversity inhibits parasites: Broad evidence for the dilution
529 effect. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 8667–8671.
530 (doi:10.1073/PNAS.1506279112)

531 28. Thrall PH, Burdon JJ. 2000 Effect of resistance variation in a natural plant host–
532 pathogen metapopulation on disease dynamics. *Plant Pathol.* **49**, 767–773.
533 (doi:10.1046/J.1365-3059.2000.00523.X)

534 29. Susi H, Laine AL. 2015 The effectiveness and costs of pathogen resistance strategies
535 in a perennial plant. *J. Ecol.* **103**, 303–315. (doi:10.1111/1365-2745.12373)

536 30. Lizarazo E, Vincenti-Gonzalez M, Grillet ME, Bethencourt S, Diaz O, Ojeda N, Ochoa
537 H, Rangel MA, Tami A. 2019 Spatial Dynamics of Chikungunya Virus,
538 Venezuela,2014. *Emerg. Infect. Dis.* **25**, 672. (doi:10.3201/EID2504.172121)

539 31. Whitfield AE, Falk BW, Rotenberg D. 2015 Insect vector-mediated transmission of
540 plant viruses. *Virology*. **479**–**480**, 278–289. (doi:10.1016/j.virol.2015.03.026)

541 32. Tollenaere C, Susi H, Laine AL. 2016 Evolutionary and Epidemiological Implications

of Multiple Infection in Plants. *Trends Plant Sci.* **21**, 80–90. (doi:10.1016/j.tplants.2015.10.014)

33. Norberg A, Susi H, Sallinen S, Safdari P, Clark N, Laine A-L. 2021 Direct and indirect viral associations predict coexistence in wild plant virus communities. *Res. Sq.* , 10.21203/rs.3.rs-741379/v1. (doi:10.21203/rs.3.rs-741379/v1)

34. DaPalma T, Doonan BP, Trager NM, Kasman LM. 2010 A systematic approach to virus–virus interactions. *Virus Res.* **149**, 1–9. (doi:10.1016/J.VIRUSRES.2010.01.002)

35. Mauck KE, De Moraes CM, Mescher MC. 2014 Biochemical and physiological mechanisms underlying effects of Cucumber mosaic virus on host-plant traits that mediate transmission by aphid vectors. *Plant, Cell Environ.* **37**, 1427–1439. (doi:10.1111/PCE.12249/SUPINFO)

36. Peñaflor MFGV, Mauck KE, Alves KJ, De Moraes CM, Mescher MC. 2016 Effects of single and mixed infections of Bean pod mottle virus and Soybean mosaic virus on host-plant chemistry and host–vector interactions. *Funct. Ecol.* **30**, 1648–1659. (doi:10.1111/1365-2435.12649)

37. Salvaudon L, De Moraes CM, Mescher MC. 2013 Outcomes of co-infection by two potyviruses: Implications for the evolution of manipulative strategies. *Proc. R. Soc. B Biol. Sci.* **280**. (doi:10.1098/rspb.2012.2959)

38. Stafford CA, Walker GP, Ullman DE. 2011 Infection with a plant virus modifies vector feeding behavior. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 9350–9355. (doi:10.1073/PNAS.1100773108/SUPPL_FILE/SM03.WMV)

39. Swanson SJ, Neitzel D, Reed KD, Belongia EA. 2006 Coinfections acquired from Ixodes ticks. *Clin. Microbiol. Rev.* **19**, 708–727. (doi:10.1128/CMR.00011-06)

40. Carrillo-Hernández MY, Ruiz-Saenz J, Villamizar LJ, Gómez-Rangel SY, Martínez-Gutierrez M. 2018 Co-circulation and simultaneous co-infection of dengue, chikungunya, and zika viruses in patients with febrile syndrome at the Colombian-Venezuelan border. *BMC Infect. Dis.* **18**, 1–12. (doi:10.1186/S12879-018-2976-1/TABLES/3)

41. Louhi KR, Karvonen A, Rellstab C, Louhi R, Jokela J. 2013 Prevalence of infection as a predictor of multiple genotype infection frequency in parasites with multiple-host life cycle. *J. Anim. Ecol.* **82**, 191–200. (doi:10.1111/j.1365-2656.2012.02028.x)

42. Malmberg JL, White LA, VandeWoude S. 2021 Bioaccumulation of Pathogen Exposure in Top Predators. *Trends Ecol. Evol.* **36**, 411–420. (doi:10.1016/J.TREE.2021.01.008)

43. Marchetto KM, Power AG. 2018 Coinfection Timing Drives Host Population Dynamics through Changes in Virulence. *Am. Nat.* **191**, 173–183. (doi:10.1086/695316)

44. Natsopoulou ME, McMahon DP, Doublet V, Bryden J, Paxton RJ. 2015 Interspecific competition in honeybee intracellular gut parasites is asymmetric and favours the spread of an emerging infectious disease. *Proc. R. Soc. B Biol. Sci.* **282**. (doi:10.1098/RSPB.2014.1896)

45. Karvonen A, Jokela J, Laine AL. 2019 Importance of Sequence and Timing in Parasite Coinfections. *Trends Parasitol.* **35**, 109–118. (doi:10.1016/J.PT.2018.11.007)

46. Debray R, Herbert RA, Jaffe AL, Crits-Christoph A, Power ME, Koskella B. 2022 Priority effects in microbiome assembly. *Nat. Rev. Microbiol.* **20**, 109–121. (doi:10.1038/s41579-021-00604-w)

588 47. Fukami T. 2015 Historical Contingency in Community Assembly: Integrating Niches,
589 Species Pools, and Priority Effects. *Annu. Rev. Ecol. Evol. Syst.* **46**, 1–23.
590 (doi:10.1146/ANNUREV-ECOLSYS-110411-160340)

591 48. RA Alford HW. 1985 Priority effects in experimental pond communities: competition
592 between Bufo and Rana. *Ecology* **66**, 1097–1105.

593 49. Almany GR. 2003 PRIORITY EFFECTS IN CORAL REEF FISH COMMUNITIES.
594 *Ecology* **84**, 1920–1935. (doi:10.1890/0012-9658)

595 50. Bushman F, Liang G. 2021 Assembly of the virome in newborn human infants. *Curr.*
596 *Opin. Virol.* **48**, 17–22. (doi:10.1016/J.COVIRO.2021.03.004)

597 51. Hoverman JT, Hoye BJ, Johnson PTJ. 2013 Does timing matter? How priority effects
598 influence the outcome of parasite interactions within hosts. *Oecologia* **173**, 1471–
599 1480. (doi:10.1007/S00442-013-2692-X/FIGURES/3)

600 52. Karvonen A, Seppälä O, Tellervo Valtonen E. 2009 Host immunization shapes
601 interspecific associations in trematode parasites. *J. Anim. Ecol.* **78**, 945–952.
602 (doi:10.1111/j.1365-2656.2009.01562.x)

603 53. Natsopoulou ME, McMahon DP, Doublet V, Bryden J, Paxton RJ. 2015 Interspecific
604 competition in honeybee intracellular gut parasites is asymmetric and favours the
605 spread of an emerging infectious disease. *Proc. R. Soc. B Biol. Sci.* **282**.
606 (doi:10.1098/RSPB.2014.1896)

607 54. Chase JM. 2003 Community assembly: When should history matter? *Oecologia* **136**,
608 489–498. (doi:10.1007/S00442-003-1311-7/FIGURES/4)

609 55. O'Keeffe KR, Halliday FW, Jones CD, Carbone I, Mitchell CE. 2021 Parasites, niche
610 modification and the host microbiome: A field survey of multiple parasites. *Mol. Ecol.*

611 **30**, 2404–2416. (doi:10.1111/MEC.15892)

612 56. Ziebell H, Carr JP. 2010 Cross-Protection: A Century of Mystery. **76**, 211–264.
613 (doi:10.1016/S0065-3527(10)76006-1)

614 57. Mauch-Mani B, Baccelli I, Luna E, Flors V. 2017 Defense Priming: An Adaptive Part
615 of Induced Resistance. *https://doi.org/10.1146/annurev-arplant-042916-041132* **68**,
616 485–512. (doi:10.1146/ANNUREV-ARPLANT-042916-041132)

617 58. Spoel SH, Johnson JS, Dong X. 2007 Regulation of tradeoffs between plant defenses
618 against pathogens with different lifestyles. *Proc. Natl. Acad. Sci.* **104**, 18842–18847.
619 (doi:10.1073/PNAS.0708139104)

620 59. Thomma BPHJ, Eggermont K, Penninckx IAMA, Mauch-Mani B, Vogelsang R,
621 Cammue BPA, Broekaert WF. 1998 Separate jasmonate-dependent and salicylate-
622 dependent defense-response pathways in arabidopsis are essential for resistance to
623 distinct microbial pathogens. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 15107–15111.
624 (doi:10.1073/PNAS.95.25.15107/ASSET/E332270E-1444-40A9-ABD1-
625 8951824DB019/ASSETS/GRAPHIC/PQ2481720005.JPEG)

626 60. Morris DE, Cleary DW, Clarke SC. 2017 Secondary bacterial infections associated
627 with influenza pandemics. *Front. Microbiol.* **8**, 1041.
628 (doi:10.3389/FMICB.2017.01041/BIBTEX)

629 61. Wang H, Anthony D, Selemidis S, Vlahos R, Bozinovski S. 2018 Resolving viral-
630 induced secondary bacterial infection in COPD: A concise review. *Front. Immunol.* **9**,
631 2345. (doi:10.3389/FIMMU.2018.02345/BIBTEX)

632 62. Laine AL, Burdon JJ, Dodds PN, Thrall PH. 2011 Spatial variation in disease
633 resistance: from molecules to metapopulations. *J. Ecol.* **99**, 96–112.

(doi:10.1111/J.1365-2745.2010.01738.X)

63. Halliday FW, Penczykowski RM, Barrès B, Eck JL, Numminen E, Laine AL. 2020 Facilitative priority effects drive parasite assembly under coinfection. *Nat. Ecol. Evol.* **4**, 1510–1521. (doi:10.1038/s41559-020-01289-9)

64. Carius HJ, Little TJ, Ebert D. 2001 Genetic variation in a host-parasite association: Potential for coevolution and frequency-dependent selection. *Evolution (N. Y).* **55**, 1136–1145. (doi:10.1111/J.0014-3820.2001.TB00633.X)

65. Navarro R, Ambrós S, Butković A, Carrasco JL, González R, Martínez F, Wu B, Elena SF. 2022 Defects in plant immunity modulate the rates and patterns of RNA virus evolution. *Virus Evol.* **8**, 1–14. (doi:10.1093/VE/VEAC059)

66. Susi H *et al.* 2017 Genome sequences of a capulavirus infecting Plantago lanceolata in the Åland archipelago of Finland. *Arch. Virol.* **162**, 2041–2045. (doi:10.1007/s00705-017-3298-0)

67. Susi H, Filloux D, Frilander MJ, Roumagnac P, Laine AL. 2019 Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ* **2019**. (doi:10.7717/peerj.6140)

68. Ovaskainen O, Abrego N. 2020 *Joint Species Distribution Modelling*. Cambridge University Press. (doi:10.1017/9781108591720)

69. Tikhonov G, Opedal ØH, Abrego N, Lehikoinen A, de Jonge MMJ, Oksanen J, Ovaskainen O. 2020 Joint species distribution modelling with the r-package Hmsc. *Methods Ecol. Evol.* **11**, 442–447. (doi:10.1111/2041-210X.13345)

70. Sagar GR, Harper JL. 1964 Plantago Major L., P. Media L. and P. Lanceolata L. *J. Ecol.* **52**, 189. (doi:10.2307/2257792)

71. Ojanen SP, Nieminen M, Meyke E, Pöyry J, Hanski I. 2013 Long-term metapopulation study of the Glanville fritillary butterfly (Melitaea cinxia): Survey methods, data management, and long-term population trends. *Ecol. Evol.* **3**, 3713–3737. (doi:10.1002/ece3.733)

72. Laine AL, Hanski I. 2006 Large-scale spatial dynamics of a specialist plant pathogen in a fragmented landscape. *J. Ecol.* **94**, 217–226. (doi:10.1111/j.1365-2745.2005.01075.x)

73. Calil IP, Fontes EPB. 2017 Plant immunity against viruses: antiviral immune receptors in focus. *Ann. Bot.* **119**, 711–723. (doi:10.1093/AOB/MCW200)

74. Gray S, Gildow FE. 2003 Luteovirus-Aphid Interactions. *Annu. Rev. Phytopathol.* **41**, 539–566. (doi:10.1146/annurev.phyto.41.012203.105815)

75. Uzest M, Gargani D, Drucker M, Hébrard E, Garzo E, Candresse T, Fereres A, Blanc S. 2007 A protein key to plant virus transmission at the tip of the insect vector stylet. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 17959–17964. (doi:10.1073/pnas.0706608104)

76. Roumagnac P *et al.* 2015 Alfalfa Leaf Curl Virus: an Aphid-Transmitted Geminivirus. *J. Virol.* **89**, 9683–9688. (doi:10.1128/jvi.00453-15)

77. Limburg DD, Mauk PA, Godfrey LD. 1997 Characteristics of beet yellows closterovirus transmission to sugar beets by Aphis fabae. *Phytopathology* **87**, 766–771. (doi:10.1094/PHYTO.1997.87.7.766)

78. Kanematsu S, Sasaki A, Onoue M, Oikawa Y, Ito T. 2010 Extending the fungal host range of a partitivirus and a mycoreovirus from rosellinia necatrix by inoculation of protoplasts with virus particles. *Phytopathology* **100**, 922–930. (doi:10.1094/PHYTO-100-9-0922)

680 79. Sasaki A, Miyanishi M, Ozaki K, Onoue M, Yoshida K. 2005 Molecular
681 characterization of a partitivirus from the plant pathogenic ascomycete Rosellinia
682 necatrix. *Arch. Virol. 2005 1506* **150**, 1069–1083. (doi:10.1007/S00705-005-0494-0)

683 80. Chang S, Puryear J, Cairney J. 1993 A simple and efficient method for isolating RNA
684 from pine trees. *Plant Mol. Biol. Report.* **11**, 113–116. (doi:10.1007/BF02670468)

685 81. Ovaskainen O, Tikhonov G, Norberg A, Guillaume Blanchet F, Duan L, Dunson D,
686 Roslin T, Abrego N. 2017 How to make more out of community data? A conceptual
687 framework and its implementation as models and software. *Ecol. Lett.* **20**, 561–576.
688 (doi:10.1111/ele.12757)

689 82. Tjur T. 2009 Coefficients of Determination in Logistic Regression Models—A New
690 Proposal: The Coefficient of Discrimination. *Am. Stat.* **63**, 366–372.

691 83. Pearce J, Ferrier S. 2000 Evaluating the predictive performance of habitat models
692 developed using logistic regression. *Ecol. Modell.* **133**, 225–245. (doi:10.1016/S0304-
693 3800(00)00322-7)

694 84. Halliday FW, Umbanhowar J, Mitchell CE. 2017 Interactions among symbionts
695 operate across scales to influence parasite epidemics. *Ecol. Lett.* **20**, 1285–1294.
696 (doi:10.1111/ele.12825)

697 85. Werner GDA, Kiers ET. 2015 Order of arrival structures arbuscular mycorrhizal
698 colonization of plants. *New Phytol.* **205**, 1515–1524. (doi:10.1111/NPH.13092)

699 86. Conci VC, Canavelli A, Lunello P. 2003 Yield Losses Associated with Virus-Infected
700 Garlic Plants During Five Successive Years.

701 87. Hall GS, Peters JS, Little DP, Power AG. 2010 Plant community diversity influences
702 vector behaviour and Barley yellow dwarf virus population structure. *Plant Pathol.* **59**,

1152–1158. (doi:10.1111/J.1365-3059.2010.02351.X)

88. Roossinck MJ. 2010 Lifestyles of plant viruses. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 1899. (doi:10.1098/RSTB.2010.0057)

89. Lesker T, Rabenstein F, Maiss E. 2013 Molecular characterization of five betacryptoviruses infecting four clover species and dill. *Arch. Virol.* **158**, 1943–1952. (doi:10.1007/S00705-013-1691-X/TABLES/2)

90. Vainio EJ, Chiba S, Ghabrial SA, Maiss E, Roossinck M, Sabanadzovic S, Suzuki N, Xie J, Nibert M. 2018 ICTV virus taxonomy profile: Partitiviridae. *J. Gen. Virol.* **99**, 17–18. (doi:10.1099/JGV.0.000985/CITE/REFWORKS)

91. Casanova JL, Abel L. 2007 Human genetics of infectious diseases: a unified theory. *EMBO J.* **26**, 915–922. (doi:10.1038/SJ.EMBOJ.7601558)

92. Liu R *et al.* 1996 Homozygous Defect in HIV-1 Coreceptor Accounts for Resistance of Some Multiply-Exposed Individuals to HIV-1 Infection. *Cell* **86**, 367–377. (doi:10.1016/S0092-8674(00)80110-5)

93. Burgess SC, Basaran BH, Davison TF. 2001 Resistance to Marek's disease herpesvirus-induced lymphoma is multiphasic and dependent on host genotype. *Vet. Pathol.* **38**, 129–142. (doi:10.1354/VP.38-2-129/ASSET/IMAGES/LARGE/10.1354_VP.38-2-129-FIG2.JPEG)

94. Laine AL. 2004 Resistance variation within and among host populations in a plant-pathogen metapopulation: Implications for regional pathogen dynamics. *J. Ecol.* **92**, 990–1000. (doi:10.1111/j.0022-0477.2004.00925.x)

95. Höckerstedt L, Numminen E, Ashby B, Boots M, Norberg A, Laine A-L. 2022 Spatially structured eco-evolutionary dynamics in a host-pathogen interaction render

isolated populations vulnerable to disease. *Nat. Commun. 2022 131* **13**, 1–11. (doi:10.1038/s41467-022-33665-3)

96. Mauck K, Bosque-Pérez NA, Eigenbrode SD, De Moraes CM, Mescher MC. 2012 Transmission mechanisms shape pathogen effects on host-vector interactions: Evidence from plant viruses. *Funct. Ecol.* **26**, 1162–1175. (doi:10.1111/j.1365-2435.2012.02026.x)

97. Raybould AF, Maskell LC, Edwards ML, Cooper JI, Gray AJ. 1999 The prevalence and spatial distribution of viruses in natural populations of Brassica oleracea. *New Phytol.* **141**, 265–275. (doi:10.1046/j.1469-8137.1999.00339.x)

98. Mauck KE, De Moraes CM, Mescher MC. 2014 Biochemical and physiological mechanisms underlying effects of Cucumber mosaic virus on host-plant traits that mediate transmission by aphid vectors. *Plant, Cell Environ.* **37**, 1427–1439. (doi:10.1111/pce.12249)

99. Moreno AB, López-Moya JJ. 2020 When viruses play team sports: Mixed infections in plants. *Phytopathology* **110**, 29–48. (doi:10.1094/PHYTO-07-19-0250-FI)

100. Shoemaker LG, Hayhurst E, Weiss-Lehman CP, Strauss AT, Porath-Krause A, Borer ET, Seabloom EW, Shaw AK. 2019 Pathogens manipulate the preference of vectors, slowing disease spread in a multi-host system. *Ecol. Lett.* **22**, 1115–1125. (doi:10.1111/ELE.13268)

101. Carroll D, Daszak P, Wolfe ND, Gao GF, Morel CM, Morzaria S, Pablos-Méndez A, Tomori O, Mazet JAK. 2018 The Global Virome Project. *Science (80-.).* **359**, 872–874. (doi:10.1126/SCIENCE.AAP7463/SUPPL_FILE/AAP7463-CARROLL-SM.PDF)

749 102. Lefkowitz EJ, Dempsey DM, Hendrickson RC, Orton RJ, Siddell SG, Smith DB. 2018
750 Virus taxonomy: the database of the International Committee on Taxonomy of Viruses
751 (ICTV). *Nucleic Acids Res.* **46**, D708. (doi:10.1093/NAR/GKX932)

752 103. Krammer F. 2020 SARS-CoV-2 vaccines in development. *Nature*. **586**, 516–527.
753 (doi:10.1038/s41586-020-2798-3)

754 104. Smith KM, Machalaba CC, Seifman R, Feferholtz Y, Karesh WB. 2019 Infectious
755 disease and economics: The case for considering multi-sectoral impacts. *One Heal.* **7**.
756 (doi:10.1016/J.ONEHLT.2018.100080)

757 105. Cadwell K. 2015 The Virome in Host Health and Disease. *Immunity* **42**, 805–813.
758 (doi:10.1016/J.IMMUNI.2015.05.003)

759 106. Fuhrman JA. 1999 Marine viruses and their biogeochemical and ecological effects.
760 *Nature*. **399**, 541–548. (doi:10.1038/21119)

761 107. Alcaide C, Rabadán MP, Moreno-Pérez MG, Gómez P. 2020 Implications of mixed
762 viral infections on plant disease ecology and evolution. *Adv. Virus Res.* **106**, 145–169.
763 (doi:10.1016/bs.aivir.2020.02.001)

764 108. Mclaughlin AA, Hanley-Bowdoin L, Kennedy GG, Jacobson AL. 123 ADVector
765 acquisition and co-inoculation of two plant viruses influences transmission, infection,
766 and replication in new hosts. *Sci. Reports* | **12**, 20355. (doi:10.1038/s41598-022-
767 24880-5)

768

769 Figure 1. Host genotypes grouped by their early season infection status and the locations of the
770 four transplant populations included in this study. A) *Plantago lanceolata* sentinel plants
771 grouped by their infection status, 1) uninfected 2) single infection by *Plantago closterovirus,*
772 *Plantago betapartitivirus* or *Plantago lanceolata latent virus* (n = 213), at the early season
773 timepoint. The colours represent different host genotypes. B) The locations of field

774 experimental sites in the Åland Islands, Finland. 'Unifected' refers to uninfected individuals,
775 *'Closterovirus'* to *Plantago closterovirus*, *'Betapartitivirus'* to *Plantago betapartitivirus* and
776 'PlLV' to *Plantago lanceolata latent virus*.

777

778 Figure 2. Late season viral occurrences grouped according to the early season infection status
779 of host plants. A) Prevalences of the five focal viruses during the late season timeperiod in
780 transplanted sentinel *Plantago lanceolata* plants. Viral prevalences are grouped by the host's
781 early season infection status. B) Co-occurrences of the five focal viruses during the late season
782 timeperiod, grouped by the host's early season infection status. *'En'* refers to *Plantago*
783 *enamovirus, 'Ca'* to *Plantago latent caulimovirus*, *'Clo'* or *'Closterovirus'* to *Plantago*
784 *closterovirus,* PlLV to *Plantago lanceolata latent virus* and *'Be'* or *'Betapartitivirus'* to
785 *Plantago betapartitivirus*.

786

787 Table 1. Explanatory performance and predictive performance, based on 10-fold cross
788 validation, of late season occupancy models of five focal viruses in terms of Tjur R^2 and AUC.

Model	**Model explanatory performance**		**Model predictive performance with 10-fold cross validation (cv)**	
Response variable*	Tjur R2	AUC	Tjur R2 (cv)	AUC (cv)
Closterovirus	0.19	0.84	0.07	0.62
Enamovirus	0.14	0.9	0.08	0.75
Betapartitivirus	0.12	0.8	0.05	0.64
PlLV	0.56	0.92	0.52	0.84
Caulimovirus	0.15	0.83	0.1	0.71

789

790

791 Figure 3. Proportion of variance explained (out of the total variation explained by the model)
792 by the fixed and random effects in the HMSC model of the late season occurrences of the five
793 focal viruses in the transplant experiment. The nine variables explaining the late season
794 occurrences of each virus (in columns) were transplant population and plant individual ID
795 (random effects), host plant genotype, host plant size, the presence/absence of herbivory, the
796 early summer occurrences of three of the viruses (*Plantago closterovirus*, *Plantago*
797 *betapartitivirus* and *Plantago lanceolata latent virus*) and differences in RNA viral sampling
798 efficiency. Here, *'Enamovirus'* refers to *Plantago enamovirus, 'Caulimovirus'* to *Plantago*
799 *latent caulimovirus*, *'Closterovirus'* to *Plantago closterovirus,* PlLV to *Plantago lanceolata*
800 *latent virus* and *'Betapartitivirus'* to *Plantago betapartitivirus*.

Table 2. Posterior mean and 95% credible intervals for the model variance partitioning. Response variables in the model were the late season occurrences of the five focal viruses and the parameters were the fixed and random effect covariates included in the model. Fixed effects in the model were: host plant genotype, early season infection status of the host (*Plantago closterovirus*, *Plantago betapartitivirus* or *Plantago lanceolata latent virus* single infection), signs of herbivory, plant area and failed late season RNA sampling. Random effects were the plant individual ID and transplant population ID. In the table *'Enamovirus'* refers to *Plantago enamovirus, 'Caulimovirus'* to *Plantago latent caulimovirus*, *'Closterovirus'* to *Plantago closterovirus,* PlLV to *Plantago lanceolata latent virus* and *'Betapartitivirus'* to *Plantago betapartitivirus*.

Response (late season viral occurrence)	**Parameter**	**Mean**	**95% Credible Interval**	
			Lower 2.5%	**Upper 97.5%**
Closterovirus	Genotype	10.5	0.5	34
Closterovirus	Closterovirus early season	2.7	0	13.1
Closterovirus	Betapartitivirus early season	23.2	1.7	58.2
Closterovirus	PlLV early season	3.5	0	18.1
Closterovirus	Herbivory	2.1	0	10.8
Closterovirus	Plant area	20.7	1.1	52.4
Closterovirus	Failed late season RNA sample	4	0	20.1
Closterovirus	Pland individual ID	19.1	0.1	82.5
Closterovirus	Transplant population ID	14.2	0.1	51.6
Enamovirus	Genotype	12.3	1.1	33.6
Enamovirus	Closterovirus early season	3.2	0	15.7
Enamovirus	Betapartitivirus early season	21	1.1	57.3
Enamovirus	PlLV early season	11.3	0.4	34.9
Enamovirus	Herbivory	4	0	14.9
Enamovirus	Plant area	5.4	0	24.3
Enamovirus	Failed late season RNA sample	34.5	1.9	77.9
Enamovirus	Pland individual ID	4.6	0	25.2
Enamovirus	Transplant population ID	3.7	0	20.3
Betapartitivirus	Genotype	24.7	4	56.4
Betapartitivirus	Closterovirus early season	2.8	0	12.6
Betapartitivirus	Betapartitivirus early season	7.5	0.1	27.6
Betapartitivirus	PlLV early season	3.5	0	17.4
Betapartitivirus	Herbivory	2.9	0	12.5
Betapartitivirus	Plant area	5.4	0	22.3
Betapartitivirus	Failed late season RNA sample	22.7	1	57.9
Betapartitivirus	Pland individual ID	13.7	0.1	74.6
Betapartitivirus	Transplant population ID	16.8	0.2	56.2
PlLV	Genotype	88	74.2	95.7
PlLV	Closterovirus early season	1.6	0	6.3
PlLV	Betapartitivirus early season	3.4	0	10
PlLV	PlLV early season	1.2	0	5.4
PlLV	Herbivory	0.7	0	3.5
PlLV	Plant area	1.3	0	7.2
PlLV	Failed late season RNA sample	1	0	4.8
PlLV	Pland individual ID	1.8	0	10.9
PlLV	Transplant population ID	1	0	5.5
Caulimovirus	Genotype	61.7	29.3	85.3
Caulimovirus	Closterovirus early season	2.4	0	11.6
Caulimovirus	Betapartitivirus early season	2.6	0	12.7
Caulimovirus	PlLV early season	1.9	0	10.5
Caulimovirus	Herbivory	2.3	0	10.9
Caulimovirus	Plant area	6.3	0	23.4
Caulimovirus	Failed late season RNA sample	2.8	0	15.1
Caulimovirus	Pland individual ID	8.6	0.1	45.9
Caulimovirus	Transplant population ID	11.4	0.1	42.1

Figure 4. Predicted responses of the late season occurrences of five focal viruses to host genotype, early season infection status, herbivory, and size, as well as the number of failed late season RNA samples, in a transplant experiment on *Plantago lanceolata*. The colours indicate predicted positive (yellow, *'+'*) and negative (blue, *'-'*) late season responses of each virus to the fixed effect covariates in the model that received > 95% posterior support. Predicted responses to plant genotypes 609_19, 2818_6 and 4_13 are illustrated relative to baseline genotype 511_14. *'Enamovirus'* refers to *Plantago enamovirus, 'Caulimovirus'* to *Plantago latent caulimovirus*, *'Closterovirus'* to *Plantago closterovirus,* PlLV to *Plantago lanceolata latent virus* and *'Betapartitivirus'* to *Plantago betapartitivirus*. The mean values of the beta parameter estimates and their credible intervals are in Supplementary Table 4.

Figure 5. Marginal effect predictions of early season infection status on late season virus prevalence by host genotype. Predicted effect of early season infection status on late season prevalence of the five focal viruses on four genotypes of *Plantago lanceolata* in a transplant experiment. The predicted late season prevalences of the five focal viruses (response variables) per host plant genotype are illustrated as a function of early season viral infection status (in rows) by *Plantago betapartitivirus*, *Plantago closterovirus* or *Plantago lanceolata latent virus* (binary fixed effect predictors). Each distinct colour represents a plant host genotype, and within these, the lighter tone represents early season viral absence and the darker tone viral presence (colour codes: dark blue = genotype 2818_6/viral presence, light blue = genotype 2818_6/viral absence; dark green = 4_13/viral presence, light green = 4_13/viral absence; dark yellow = 511_14/viral presence, light yellow = 511_14/viral absence; dark purple = 609_19/viral presence, light purple=609_19/viral absence). The whiskers represent the bounds of 95 % credible intervals of the median prediction. Orange stars highlight cases in which there is strong posterior support for a directional effect of early season viral infection status on the late season prevalence of a focal virus. In the figure *'Enamovirus'* refers to *Plantago enamovirus, 'Caulimovirus'* to *Plantago latent caulimovirus*, *'Closterovirus'* to *Plantago closterovirus,* PlLV to *Plantago lanceolata latent virus* and *'Betapartitivirus'* to *Plantago betapartitivirus*.