

Review

# Metagenomic Studies of Viruses in Weeds and Wild Plants: A Powerful Approach to Characterise Variable Virus Communities

Beata Hasiów-Jaroszewska <sup>1,\*</sup> , Dieke Boezen <sup>2</sup> and Mark P. Zwart <sup>2</sup> 

<sup>1</sup> Department of Virology and Bacteriology, Institute of Plant Protection-National Research Institute, Węgorza 20, 60-318 Poznań, Poland

<sup>2</sup> Department of Microbial Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Droevendaalsesteeg 10, 6708 PB Wageningen, The Netherlands; D.Boezen@nioo.knaw.nl (D.B.); M.Zwart@nioo.knaw.nl (M.P.Z.)

\* Correspondence: B.Hasiow@iorpib.poznan.pl; Tel.: +48-618649142

**Abstract:** High throughput sequencing (HTS) has revolutionised virus detection and discovery, allowing for the untargeted characterisation of whole viromes. Viral metagenomics studies have demonstrated the ubiquity of virus infection – often in the absence of disease symptoms – and tend to discover many novel viruses, highlighting the small fraction of virus biodiversity described to date. The majority of the studies using high-throughput sequencing to characterise plant viromes have focused on economically important crops, and only a small number of studies have considered weeds and wild plants. Characterising the viromes of wild plants is highly relevant, as these plants can affect disease dynamics in crops, often by acting as viral reservoirs. Moreover, the viruses in unmanaged systems may also have important effects on wild plant populations and communities. Here, we review metagenomic studies on weeds and wild plants to show the benefits and limitations of this approach and identify knowledge gaps. We consider key genomics developments that are likely to benefit the field in the near future. Although only a small number of HTS studies have been performed on weeds and wild plants, these studies have already discovered many novel viruses, demonstrated unexpected trends in virus distributions, and highlighted the potential of metagenomics as an approach.

**Keywords:** HTS; viromes; weeds; wild plants; plant virus



**Citation:** Hasiów-Jaroszewska, B.; Boezen, D.; Zwart, M.P. Metagenomic Studies of Viruses in Weeds and Wild Plants: A Powerful Approach to Characterise Variable Virus Communities. *Viruses* **2021**, *13*, 1939. <https://doi.org/10.3390/v13101939>

Academic Editor: Nuredin Habili

Received: 13 August 2021

Accepted: 23 September 2021

Published: 27 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Metagenomics is the analysis of environmental samples using high-throughput sequencing (HTS), with the intent of characterising microbial communities. HTS has the advantage of the identification and genomic characterisation of viruses in a target-independent manner, and it has proved to be a valuable tool for virus detection, discovery, or diversity studies. Classic virus detection methods, such as the enzyme-linked immunosorbent assay (ELISA) or polymerase chain reaction (PCR), usually target one or a few viral species and require prior knowledge of the virus to be detected, while HTS methods have been developed to look for virus-like sequences without the bias of only detecting known viruses. HTS technologies allow a generic approach to virus identification and can deliver species, or even strain-specific results [1]. There are several applications for metagenomics approaches in plant virus diagnostics: (i) identifying the causes of viral diseases in economically important crops; (ii) screening for specific viruses when their presence is suspected, (iii) detecting asymptomatic or cryptic viruses, and (iv) discovering novel viruses whose presence has not been reported before. To date, the majority of metagenomics studies have focused on economically important plant species, whereas the occurrence and population structure of viruses in weeds and wild plants have remained largely unstudied. Wild plants are important players in many biotopes—including environments with significant anthropogenic impact such as agroecosystems—and they have a great impact on virus epidemiology. Plant viruses can spill over in both directions between wild plants and

crops with potential adverse effects in both managed and natural ecosystems. Knowledge about the occurrence and diversity of viruses in wild plants can offer insight into factors that promote long-term coexistence between hosts and viruses in nature, as well as the evaluation of the disease emergence risk [2].

In this review, we focus on the role of weeds and wild plants in virus emergence, and discuss recent metagenomic studies of weeds which show the potential of this approach for untargeted virus detection in wild hosts. We review recent discoveries of plant viruses inhabiting various ecological niches and discuss “pros and cons” of using metagenomics for this purpose. Moreover, we show the potential of using HTS to explore the diversity and richness of the viruses associated with plant populations. Finally, we discuss the future directions of plant virus metagenomics in weeds and wild plants.

## 2. Role of Weeds and Wild Plants in Virus Emergence

In agroecosystems, crop plants often grow side by side with weeds and wild plants. These wild plants constitute potential reservoirs for viruses that may spread into cultivated plants, thereby leading to epidemics or to the emergence of novel viruses [3–6]. Many viruses are known to have wide host ranges encompassing both wild plants and crops [3,7,8]. Viruses can emerge in cultivated plants from wild hosts or, conversely, virus infections in crops may spill over and affect wild plant populations. Spread from reservoirs into a new environment, establishing productive infections, and effective between-host transmission are necessary steps for virus emergence to occur [4]. Most plant viruses are transmitted by insect vectors, which may increase the possibilities for virus transmission across landscapes and the distances over which viruses can be transmitted [9–11]. Knowledge of virus diversity, prevalence, and dynamics in wild plant populations is, therefore, relevant to a better understanding of virus epidemiology and emergence in crops. Conversely, viruses in natural systems and viruses spilling over from agroecosystems could have a major impact on natural ecosystems [12,13]. Viruses in wild plants have been reported to be diverse and often do not induce evident symptoms in their host [13,14]. Therefore, these viruses may have been overlooked when sampling strategies focused on plants displaying overt virus symptoms [5,15]. For example, only 2.3% plant samples collected in the Oklahoma Tallgrass Prairie Preserve displayed symptoms suggesting viral infection, but using metagenomics, viruses were discovered in 25% of samples [16]. Similarly, large scale metagenomic studies conducted in France, America and South Africa revealed the presence of virus-related sequences in 25–58% of samples where the majority of plant specimens did not have outward signs of viral infection [17]. Across ecosystems, virus prevalence is observed to be high in wild plants, regardless of the presence or absence of symptoms [2,8].

### 2.1. Mixed Infection in Wild Plants

In wild plants, mixed infection—that is, the coinfection of a single plant by two or more different viruses—is common. These infections often include both acute and persistent viruses [5]. Mixed infection leads to a variety of intra-host virus–virus interactions, many of which may result in the generation of novel genetic features and subsequently change the genetic structure of the viral population. Recombination can lead to the generation of new variants, and segmented or multipartite viruses can form new variants through reassortment [18]. Like the distribution of fitness effects for other mutation classes, the majority of recombinants will be deleterious [19] but occasionally a new variant may be fitter or more virulent than the parental virus [20]. The interactions between viruses infecting a single host plant can be neutral, synergistic, or antagonistic with respect to virus accumulation (i.e., titer), host response (disease symptoms), or transmission efficiencies by vectors [20,21]. From crop systems, we know that virus–virus interaction can impact the prevalence and characteristics of infection. One example of this is the interactions between Parsnip yellow fleck virus (PYFV), Anthriscus yellows virus (AYV) and aphid vectors including *Cavariella aegopodii*. PYFV can be transmitted to carrot (*Daucus carota*) by

aphids only when they have first acquired AYV. Carrot is not susceptible to infection by AYV, but introduction of the virus complex into carrot crops leads to scattered infections because individual aphids retain the transmission-dependent combination for only a few days [22], hereby increasing the effective host range of PYFV. The exchange of different virus components, including capsid proteins and helper components [9], could enhance transmission of viruses by mixed infection. Sharing of virus components in mixed infection may have far reaching and unexpected consequences. For example, a recent report suggests that Tobacco mosaic virus (TMV) infection enables a productive infection of tobacco plants by Cryphonectria hypovirus 1 (CHV1)—a fungal virus—because TMV's movement protein allows CHV1 to move from cell to cell [23]. While metagenomics studies have shown mixed infection is common in wild plants, most of the work on the phenotype resulting from the interaction between viruses has been carried out in crop systems rather than wild plants.

## 2.2. Novel Viruses and Viral Diversity in Wild Plant Populations

Metagenomics has revolutionised plant virology by enabling the untargeted, simultaneous detection of multiple viruses regardless of their genomic nature [24]. A number of HTS-based approaches for identifying viral genomes have been developed using a wide range of nucleic acid targets. Workflows that characterise total nucleic acid, double-stranded RNA (dsRNA), virion-associated nucleic acids (VANA) purified from virus-like particles, ribosomal-RNA-depleted RNA and messenger RNA have been developed for this purpose [25]. In addition, the sequencing of small RNAs from the host has also been used for virus detection [17]. The small RNA approach works well for detecting viruses in a range of hosts, as one important immune response against viruses is RNA silencing, where virus-specific double-stranded RNA that is generated in most virus infections triggers a plethora of host responses that result in the sequence-specific degradation of RNA [26]. Using this method made it possible to discover novel viruses from both insects and plants [27,28]. An important limitation of small RNA sequencing is that it fails to detect viruses that either do not trigger silencing responses or that induce effective silencing suppression (e.g., persistent viruses) [17], and sequence assembly can result in chimeric genomes. Although more approaches and tools for metagenomics are becoming available, the number of studies addressing viruses occurring in weeds and wild hosts is still limited.

The majority of viruses found in natural systems have not been previously described [8,13]. Novel viruses are being discovered readily in wild hosts in diverse natural ecosystems [29], giving insight into virus biodiversity and the distribution of plant viruses across the globe [2,8,13]. Though the number of metagenomic studies in weeds and wild plants is increasing, the proportion of novel viruses uncovered in these plants remains high, indicating that there are many viruses yet to be identified and described in wild plant ecosystems. Besides the many new viruses detected, virus metagenomic studies have also revealed considerable variation in viral prevalence and community composition [2,6,8]. It has been suggested that particular virus families may be associated with crops, whereas other virus families would be associated with native vegetation. Thus, virus families characterised from crops may be quite different from those in wild plants [13]. These differences raise questions about the evolutionary forces and drivers acting on these different virus populations, their effects on viral adaptation to different hosts, and ultimately their effect on the assembly of viral communities. Metagenomic studies are now being used to answer these questions and analyse (1) the spatial distribution and genetic diversity of viruses in different landscapes; (2) the contribution of changes in plant populations and the effects of ecosystem simplification on viral pathogenicity or emergence, and (3) the dynamics of spillover from reservoirs to other hosts, an important component of the intensity and directionality of fluxes of viruses between crops and wild plants [6,8,30].

An overview of studies using metagenomic approaches for virus detection and diversity in weeds and wild plants is presented in Table 1.

**Table 1.** Studies using metagenomic approaches for virus detection and diversity in weeds and wild plants.

Sample Collection Year	Country	Host Plant	HTS Platform	HTS Template	Main Conclusions	Reference
2005–2008	USA	<i>Ambrosia psilostachya</i> <i>Vernonia baldwinii</i> <i>Asclepias viridis</i> <i>Ruellia humilis</i> <i>Panicum virgatum</i> <i>Sorghastrum nutans</i>	454	VANA and dsRNA	<ul style="list-style-type: none"> <li>• Many unknown viruses identified in asymptomatic plants.</li> <li>• More viruses detected by dsRNA than by VANA.</li> <li>• Plant virus composition is highly variable. Of the factors investigated, host plant explains most but only a small fraction of this variation.</li> </ul>	[31]
2010–2012	South Africa, France	Cultivated and un-cultivated plants along the interface of natural and agricultural ecosystems.	454	VANA	<ul style="list-style-type: none"> <li>• 94 unknown viruses were identified, primarily from uncultivated plants.</li> <li>• Highest prevalence and identified diversity of plant viruses found in agricultural ecosystems.</li> </ul>	[8]
2012–2013	South Africa, New Zealand	<i>Arctopus echinatus</i> <i>Lolium perenne</i> , <i>Panicum ecklonii</i> , <i>Stipagrostis sp</i>	454, Illumina	VANA, and enrichment of circular DNA	<ul style="list-style-type: none"> <li>• 5 novel, circular DNA viruses were identified.</li> </ul>	[32]
2013	Finland	<i>Plantago lanceolata</i>	Illumina	Small RNA of 20–30 nucleotides, mainly composed of small interfering RNAs.	<ul style="list-style-type: none"> <li>• Virus prevalence and diversity varied among the sampled host populations.</li> <li>• Potentially novel virus species belonging to Caulimovirus, Betapartitivirus, Enamovirus, and Closterovirus genera were identified.</li> <li>• In a PCR-based follow-up study, viral prevalence and diversity were lower in natural ecosystems, suggesting spillover from agricultural to natural ecosystems.</li> </ul>	[2]
2017	France	200 plants (representing 80 species) collected for each of 4 sites, with 2 sites being managed and 2 being unmanaged sites.	Illumina	dsRNA on both leaves and leaf-associated, cultured fungi	<ul style="list-style-type: none"> <li>• 161 fungal families and 18 viral families were identified.</li> <li>• A large viral diversity dominated by novel dsRNA viruses was detected.</li> <li>• At the family level, more diversity in the managed plots, as some ssRNA viruses are only present there.</li> <li>• Leaf associated fungal communities were more diverse in unmanaged systems.</li> </ul>	[33]

Table 1. Cont.

Sample Collection Year	Country	Host Plant	HTS Platform	HTS Template	Main Conclusions	Reference
2017–2018	France	<i>Solanum nigrum</i>	Illumina	dsRNA	<ul style="list-style-type: none"> <li>•A total of 20 viral families were discovered.</li> <li>•The novel ilarvirus, here named <i>Solanum nigrum ilarvirus 1</i> (SnIV1), was detected in both tomato and nightshade samples.</li> <li>•Virome richness was highly variable. Tomato and nightshade shared only 17.9% OTUs, indicating different viromes for close relatives at the same location.</li> </ul>	[6]
2018–2019	Poland	<i>Verbena officinalis</i> L., <i>Silene latifolia</i> Poir. <i>Rorippa × prostrata</i> <i>Robinia pseudoacacia</i> L	Illumina	Total RNA	<ul style="list-style-type: none"> <li>•Detection of two viruses that had not been previously found in Poland: Clover yellow mosaic virus (CIYMV) and melandrium yellow fleck virus (MYFV).</li> </ul>	[34]

Since 2010, studies using HTS to identify plant viruses have been conducted in 20 African countries, reviewed by Ibaba et al. [35]. At least 29 host plants, including various economically important crops, ornamentals, and medicinal plants representing 18 different families have been used in HTS for virus detection and diagnostics [32,35]. The studies resulted in the detection of previously known and novel viruses from almost any host, confirming the wide distribution of plant viruses in different ecosystems and suggesting the importance of knowledge on the diversity, prevalence, and spatial distribution of viruses. For example, four novel circular replication-associated protein (Rep)-encoding single-stranded (CRESS) DNA viruses were detected in wild and medicinal plant samples from *Poaceae* and *Apiaceae* species collected in Darling in the Western Cape region of South Africa [32].

Recently, HTS has also been applied by Minicka et al. [34] for the detection and identification of different viral species occurring in single and mixed infections in plants in Poland. The authors analysed 50 samples collected from different plant species during 2018–2019, including cultivated and ornamental plants, weeds, and trees. An interesting finding was made for weeds sampled from the areas adjacent to cultivated fields. HTS allowed for the identification of Clover yellow mosaic virus (CIYMV) (*Potexvirus* genus, Alphaflexiviridae family) on *Verbena officinalis* plants, and Melandrium yellow fleck virus (MYFV) (*Bromovirus* genus, Bromoviridae family) on *Silene latifolia* plants, whose presence has not been reported in Poland before. CIYMV infects broad bean, pea, alfalfa, fat hen, chickweed, and tulips [36,37]. The occurrence of Turnip mosaic virus (TuMV, *Potyvirus* genus, Potyviridae family) was confirmed in *Rorippa prostrata*. TuMV infects a wide range of cruciferous plants worldwide including economically important vegetable, oilseed, biofuel, forage and ornamental crops [38,39]. It is transmitted non-persistently by more than 40 different aphid species and has a wide natural host range including a wide range of weed species [40]. The occurrence of different viral communities in weeds growing nearby the agriculture landscape indicate the possibility of a spillover from the reservoirs to other hosts. This study confirmed that weeds can harbour viruses between seasons, raising questions about how to mitigate the risk that weeds contribute to disease outbreaks in crops. These questions are highly relevant, as some contemporary management strategies (i.e., organic farming, and practices aimed at increasing biodiversity in and around farmlands) potentially may also affect the intensity of viral spillover between crops and wild plants. Interestingly, the presence of Tomato black ring virus (TBRV) (*Nepovirus* genus, Secoviridae family) was also confirmed using HTS in *Sambucus nigra* plants growing nearby

zucchini crops. The genomic RNAs of TBRV were associated with subviral satellite RNAs. The occurrence of TBRV in zucchini crops has been previously reported [41]; however, there is no evidence for a direct association between the infection of wild plants and zucchini crops. The origin of the virus within these zucchini crops remains unclear; there are no data regarding seed transmission, and the possibility of transmission by nematode vectors has not been verified yet.

HTS has had a major impact on the analysis of diversity and evolution of virus populations [42–44]. In this growing field, a lot of new data have been generated that shed light on genetic variability, evolutionary dynamics, and connections between virus evolution and the host evolution and ecology [45]. Susi et al. [2] used small RNA sequencing to identify viruses in *Plantago lanceolata* populations in Southwest Finland. Aside from the previously reported *Plantago lanceolata* latent virus (PLLV), the authors found sequences representative of four new virus species: *Plantago* latent caulimovirus, *Plantago* betapartitivirus, *Plantago* enamovirus, and *Plantago* closterovirus. Moreover, they discovered variations in virus prevalence across *P. lanceolata* populations and found that symptomatic plants were more likely to be infected than asymptomatic ones (84 vs. 44% infected). A metagenomics survey performed in 115 wild species growing in wild or anthropic habitats in central Spain revealed an infection caused by geminiviruses concentrated on 25 species growing in habitats next to crops and within them [46]. Geminiviruses constitute a large family of plant-infecting DNA viruses which have dramatically impacted agricultural yields over the past 50 years [47]. Geminiviruses are a major threat to the food security of developing countries in the tropical and subtropical regions of the world [48,49]. New, potentially divergent geminivirus species were also detected from uncultivated plants by Bernardo et al. [8] and further characterised by Claverie et al. [50].

Bernardo et al. [8] used the metagenomics approach to analyse the prevalence and diversity of virus communities in cultivated and uncultivated areas in two Mediterranean climate regions: the Rhone delta in France and the Western Cape in South Africa. Overall, 1725 plant samples were collected in crops and wild plants in areas with different levels of human management and viral genomes were sequenced using 454 pyrosequencing. Virus prevalence and diversity were evaluated only at the family level and, on the basis of pairwise sequence comparisons, similarities related to virus sequences were grouped into operational taxonomic units (OTUs) at the family level for deeper analysis. Surprisingly, this study found that both virus prevalence and diversity were higher in agricultural areas. In this study, 94 previously unknown virus species were reported. The majority of these unknown viruses originated from uncultivated plants, supporting the notion that the presently known crop-infecting viruses represent just a relatively small fraction of all plant-infecting virus species. Interestingly, six OTUs from this study potentially represent highly divergent geminiviruses. Further studies conducted by Claverie et al. (2018) [50] identified six completely novel geminiviruses from four different South African uncultivated endemic plant species (*Euphorbia caput-medusae*, *Limeum africanum*, *Exomis microphylla*, and *Polygala garcinia*) and two from France (*Juncus maritimus* and cultivated species *Medicago satvia*), thus revealing an unexpected degree of geminivirus diversity.

A high level of genetic variability was also shown in the study of Ma et al. [6] who simultaneously explored diversity of virus populations in a cultivated crop plant, tomato (*Solanum lycopersicum*), and a botanically related weed, *Solanum nigrum* (European black nightshade). The samples were collected from tomato and black nightshade growing side by side and HTS was used to determine composition and richness of the virome. A total of 20 viral families were discovered. Virome richness was highly variable, with tomato and nightshade sharing only 17.9% of OTUs. Overall, the results provide evidence for limited viral spillover between tomato and wild nightshade populations, with some exceptions. For example, *Broad bean wilt virus 1* (BBWV1, *Fabavirus* genus, Secoviridae family) was detected in six of the samples from black nightshade populations (out of 11 populations sampled), but it was not detected in any of the seven sampled tomato populations. Moreover, BBWV1 showed a large and unexpected level of genetic diversity, suggesting frequent reassortment

between RNA1 and RNA2. As BBMWV1 has a relatively wide host range—including many crops such as pea, spinach, lettuce, pepper, and tomato—and is aphid-transmitted, its absence in these tomato populations suggests the existence of biological or epidemiological barriers that limited its spillover potential [6]. These studies highlight the importance of characterising the viromes of both crops and wild plants to understand the drivers shaping viral emergence and disease dynamics [8,51,52].

All of the above-mentioned studies explore the spatial variation between viral communities in weeds and wild plants; however, information on the temporal variation in wild plant virus communities is still very limited. Thapa et al. [31] analysed the effects of host identity, location and sampling year on the taxonomic composition of plant viruses in six wild plant species [*Ambrosia psilostachya* (Asteraceae), *Vernonia baldwinii* (Asteraceae), *Asclepias viridis* (Asclepiadaceae), *Ruellia humilis* (Acanthaceae), *Panicum virgatum* (Poaceae) and *Sorghastrum nutans* (Poaceae)] comprising 400 specimens of the target hosts collected from twenty sites in the Tallgrass Prairie Preserve in northeastern Oklahoma over four years. Samples were tested for the presence of plant viruses applying VANA and double-stranded RNA enrichment methods, allowing for comparisons of community composition based on the host species, location, or time of collection. This large-scale metagenomic study classified viral sequences into OTUs for a better assessment of virus diversity. Results indicated that host species had a significant effect on virome composition compared to the location and sampling time. The majority of plant viruses identified during the study were novel viral species that had not previously been reported.

HTS techniques can also be used to study different host–pathogen interactions [33]. By combining metagenomic and culturomic approaches, Ma et al. [33] assessed the richness, diversity, and composition of leaf-associated fungal and viral communities from pools of herbaceous wild plants. Samples were collected in four sites in southwest France corresponding to cultivated or natural ecosystems. Overall, 161 fungal families and 18 viral families were identified. The community composition of both fungal and viral populations showed strong site specificity. Ecosystem management had a significant effect on the microbial communities, as indicated by higher fungal community richness in natural ecosystems and a higher viral family richness in cultivated ecosystems, suggesting that leaf-associated fungal and viral communities are under the influence of different ecological drivers.

We have provided an overview of different studies that use HTS methods on wild plants and weeds in Table 1. When we examine the results of these different studies, are there any trends that become apparent? Firstly, all of these studies discovered novel viruses, highlighting the power of HTS methods for virus discovery and demonstrating how little virus biodiversity has been catalogued to date. Secondly, plant viruses were regularly found in healthy plants, often at high prevalence. These observations confirm that many plant virus infections are asymptomatic in weeds and wild plants, but also underline the importance of understanding both the economic and ecological implications of this type of infection. Moreover, studies on viral prevalence should take the asymptomatic nature of many infections into account and sample plants accordingly. Thirdly, there has been little evaluation and comparison of the different methods for meta-genomics described so far. We think there is a need to systematically compare methods and identify their limitations, as this will impact the interpretation of results. For the studies we considered, only Thapa et al. [31] compared methods extensively, and found marked differences between the viromes obtained with dsRNA to VANA-based methods. In contrast, other studies do report confirmation of HTS identification with other methods [2,34]. Due to publication bias, studies with negative results may be underrepresented, so these confirmations should be interpreted conservatively. Fourth, three studies [8,30,33] found that virus diversity was higher in agricultural than wild ecosystems. The two studies that estimated prevalence [8,30] also found a higher prevalence of plant virus infections in agricultural systems. On one hand, agricultural systems with low diversity and plants selected for crop yield, at the expense of other traits such as resistance to pathogens, might be expected to amplify viruses. On the other hand, one might expect wild ecosystems to harbour more

viral diversity due to the higher host diversity present, if many infections are asymptomatic and have a low prevalence due to higher host diversity. As more metagenomics data become available, and as we learn more about the biases and sensitivity of different nucleic extraction and sequencing approaches, it will be extremely interesting to see whether this trend is confirmed and the extent to which it is found across different ecosystems. Finally, we note that the vast majority of studies have sampled only one or two times, and hence there is a pressing need for longitudinal studies. Here it is interesting to note that PCR-based detection of viruses over multiple years in wild *Arabidopsis halleri* populations shows seasonal variation in accumulation, whilst the prevalence and genome of the virus populations are surprisingly consistent [53].

Are there any general features of plant-associated viromes that can be tentatively identified, based on the studies reviewed here? First, there is the observation that diverse dsRNA viruses, a very large proportion of which appear to be novel, usually dominate the plant-associated viromes [8,31]. A second observation is that the richness in DNA viruses tends to be lower than that of RNA viruses, observing the balance between these two groups in the currently recognised taxonomy approved by International Committee on Taxonomy of Viruses [8].

### 2.3. Effect of Host Population, Climate Changes, and Land Use on Virus Emergence and Biodiversity

In natural ecosystems, host populations typically have greater genetic diversity and more complex age structures than populations in agroecosystems, particularly in regions where crops are grown as monocultures with low genetic diversity. In wild plant communities, by contrast, plant genomes are under natural selection. This may favour the persistence of multiple resistance alleles or result in selection for novel resistance alleles. Understanding the role of plant genetic diversity in controlling patterns of virus emergence is a long-standing goal of biological research, and it is central for the management of current infectious diseases and developing long-term strategies to prevent the emergence of new ones [54]. Increased use of nature areas for agriculture has resulted in significant losses in biodiversity and ecosystem simplification. Virus emergence in crops is a complex process that involves interactions among wild and crop hosts, insects, and other virus vectors, and changes in ecosystems. Although the data are quite limited, it is thought that reduced biodiversity in agroecosystems in terms of both plant species richness and the genetic diversity within species will lead to increases in the incidence of viral disease [5,54]. The loss of biodiversity would increase host abundance and density, which, in turn, would lead to higher transmission rates and a higher virus prevalence [55,56]. These conditions are likely to result in the evolution of higher virulence, as increasing the density of susceptible hosts is an important driver of the evolution of increased virulence [57]. Understanding whether there is a general relationship between biodiversity and disease risk is critical for projecting and reducing the impacts of future disease outbreaks. The relationship between biodiversity and disease risk has been analysed in Cereal yellow dwarf virus (CYDV) and Barley yellow dwarf virus (BYDV). CYDV and BYDV belong to the *Luteovirus* genus (Tombusviridae family) and infect a wide range of grass species [7]. CYDV and BYDV are generalists; however, they show partial host specialisation, as they accumulate at different levels in different grass species. Experimental studies on B/CYDVs in grasslands of the western US reveal that the prevalence and community compositions of B/CYDV is affected by interactions of host plants, vectors, vertebrate herbivores, as well as biotic and abiotic factors [7]. The interactions between B/CYDV, their aphid vectors, and host plants explain the effect of species composition on infection risk and underline the complexities of virus ecology and epidemiology [7,12,58,59]. These results also suggest that ecosystem simplification and biodiversity losses favour higher virus prevalence.

The opposite scenario has been shown in the study of Susi and Laine [30], who investigated the impact of land use on five virus species detected in 27 *P. lanceolata* populations across the Aland Islands in Finland. They investigated whether the structure of virus communities differed between natural and agriculture edge settings. The study revealed



that virus species richness declined with increasing plant diversity in a natural ecosystem, while in cultivated areas virus species richness was moderately higher, and not associated with plant richness. This difference was not explained by changes in host richness between these two habitats, suggesting potential virus spillover and increased transmission of viruses across the agroecological interface [30]. The effect of host richness on infection prevalence was negative and not altered by agricultural land use.

Some viruses capable of devastating crops and harvested products have become more active and damaging because their geographical ranges have expanded due to climate change. A changing climate can contribute to the successful spread of newly introduced viruses or their vectors, and the establishment of these organisms in new habitats that were previously unfavourable. Non-native plants and insects may also expand their geographic ranges and hereby invade communities where they were previously not present. These species might constitute reservoir hosts or vectors of plant viruses capable of causing epidemics in nearby crops [60]. In terrestrial ecosystems, studies have increasingly found evidence of virus infection in a broad range of wild plants [2,30]. Global climate change is one of the main reasons for the increased economic impact of aphids in temperate regions, which can adapt to new environmental conditions rather quickly [61,62]. The global temperature over the past century has been increasing by approximately 0.8 °C and is expected to rise between 0.9 and 3.5 °C by 2100. Such changes will not only have a great impact on the growth and cultivation of different crops but will also affect the reproduction, spread, and severity of many plant viruses [63]. Since pests and pathogens are currently responsible for more than 25% of crop and post-harvest losses, any range expansions or altered host–pathogen interactions that exacerbate damage and loss will have very serious implications for the security of the world’s food supplies [64]. Aphids generate both winged and wingless adult morphs, whereas whiteflies are all winged. Winged individuals are the ones that contribute the most to virus spread because of their ability to fly over long distances. The short developmental times and great capacity of reproduction of aphids make them especially suited to adapt to climate change [62,65]. Climate change can have various effects on vectors, such as the modification of a vector’s phenology, over-wintering, density, migration, and interactions with its natural enemies. Migration potential and long-distance dispersal of virus vectors may also be affected by changes in weather patterns. Increased winter temperature induces earlier starts to the aphids’ annual cycle, increasing the proportion of winged adults and stimulating their flight activity [66]. Furthermore, plant viruses are able to influence the behaviour and fitness of their vectors in such a way that changes in plant–virus–vector interactions can affect their transmission, although it remains to be seen that such changes are adaptive under natural conditions. It has been shown that aphid-transmitted Cucurbit aphid-borne yellows virus (CABYV; *Polerovirus* genus, Luteoviridae family) can induce changes in the alighting, settling, and probing behaviour of its vector, the cotton aphid *Aphis gossypii* [67]. CABYV infects cucumber, melon, squash, and watermelon and has also been detected in many weed species, which may be efficient reservoirs [68].

### 3. Problems to Be Solved and Future Directions

The efficacy of viral metagenomics to detect variable virus communities in weeds and wild plants relies on sampling, nucleic acid extraction, high-throughput sequencing, and bioinformatic analysis strategies. Each of these steps will pose its own challenges and may affect estimates of virus community composition.

#### 3.1. Challenges in Virus Enrichment and Nucleic Acid Extraction

The choice between protocols using viral particle enrichment or total nucleic acid extraction protocol will affect the relative abundance of viral reads, as reviewed by Maclot et al. [69]. It is important to be aware of the bias that the chosen method will introduce prior to metagenomic sequencing. For example, we expect differences in the sensitivity of VANA and dsRNA-based approaches. VANA protocols enrich virus particles,

and will thus not recover non-encapsidated viruses, viruses with low titers, or viruses with an unstable particle morphology. dsRNA-based approaches can also detect single-stranded RNA (ssRNA) and DNA viruses. dsRNA is a replication intermediate for ssRNA viruses, whereas for DNA viruses overlapping transcripts from ORFs in opposite orientations can hybridise to generate dsRNA [70,71]. However, for some ssRNA viruses, low amounts of dsRNA can be produced [72] and for DNA viruses successful detection will depend on the organisation of viral transcripts, hence this approach may not be suited for detecting many viruses that do not have dsRNA genomes. Ma [73] compared these two viral sequence enrichment approaches: dsRNA and VANA for plant virome analysis in complex pools representative of the most prevalent plant species in unmanaged and cultivated ecosystems. Additionally, a novel bioinformatic approach was implemented and virus richness was assessed by determining OTUs following the clustering of conserved viral domains [8]. To analyse virome diversity, six different sites were selected in southwest France and for each site, a total of 200 individual plant samples were collected in spring 2016. Analysis using either dsRNA or VANA metagenomic approaches revealed a high viral diversity in all sites, including the presence of novel dsRNA viruses. Generally, both approaches recovered largely the same viral families, however, DNA viruses were not efficiently recovered using dsRNA strategy. On the other hand, for the VANA approach, a low efficiency of detection was observed for virus families with low titer and/or less stable particles. Therefore, for researchers mainly interested in RNA viruses, a dsRNA-based approach was recommended as it provided a more comprehensive description of analysed (phyto)viromes [51,73].

Secondly, another critical step in plant virus metagenomics is the preparation of high-quality nucleic acid extracts. Polysaccharides, high levels of RNases, polyphenols, and fibrous tissues are often encountered in plant samples, and fibrous tissues such as lignin (wood) are difficult to break up and remove, which makes extracting high-quality nucleic acids challenging for wild plant species. Numerous adaptations have been made to nucleic acid isolation methods—including modifications to tissue storage, maceration, preparation and extraction—to obtain sufficient, high-purity nucleic acids for non-model plant species, depending on the tissue type and the presence of secondary metabolites and inhibitors [74]. These experimental details can be of vital importance, but are often not reported in sufficient detail for plant virus metagenomics work to be fully reproducible. We hope these omissions become less common due to fewer constraints on publication format and the possibility of including a wide range of supplementary materials, aided by many contemporary initiatives to increase scientific reproducibility and transparency. Given the broad pallet of nucleic acid approaches and techniques available, as well as the challenge extractions of many wild species pose, selecting the most appropriate approach will remain a hurdle that must be carefully managed.

### 3.2. Challenges in HTS and Bioinformatics

Analysing virus metagenomics data in weeds and wild plants is challenging, both conceptually and practically. Currently, the HTS method commonly used for viral metagenomics is Illumina sequencing. Illumina HiSeq and NextSeq instruments generate short, paired end reads of  $2 \times 75$ –150 bp in length, read from the short fragments produced during the library preparation process. Illumina MiSeq offers a considerably lower volume of reads, but read length can be doubled up to  $2 \times 300$  bp paired end reads. From these short sequencing reads, longer virus genome sequences can be assembled using a variety of pipelines and software packages. After the assembly process, viral genomes can, in some cases, be incomplete (e.g., partial genomes or genome segments that are missing), or assembly can result in chimeric sequences (where reads from different genetic origins within a sample are similar and therefore erroneously assembled into one sequence). Chimeric assemblies are more likely to happen when metagenomics are performed on pooled samples, which is common in plant virus metagenomics studies. In the future work, the problem of chimeras may be solved by using HTS methods which generate longer reads, such as Oxford Nanopore sequencing, or PacBio SMRT, or HiFi sequencing. Due to the

small genome size of plant viruses, ranging from 4 to 30 kb [75], long read sequencing can capture entire genome segments in a single read. For plant viruses, long read sequencing has been used on geminiviruses [76], *tombusviridae* [77], and several potyviruses including Wheat streak mosaic virus [78], Plum pox virus [79] and Yam mild mosaic virus [80]. However, the host plants in these studies were all crop plants, with the exception of the geminiviruses found in *Medicago arborea*, a shrub in the *Fabaceae* family [76]. In these studies, a variety of DNA and RNA HTS library preparation protocols are used, and the choice of library preparation affects the virus detection sensitivity. Generally, viruses were reliably detected with long-read sequencing.

Although long read sequencing provides opportunities for plant virus metagenomics, there are a few drawbacks. Firstly, all library preparation protocols for RNA sequencing require poly-A tailed RNA as input, as sequencing adapters are designed to bind to this region. However, many plant virus species do not have a poly-A tail. This limitation can be circumvented by adding a poly-A tailing step prior to sequencing, but this reaction is not specific to viral RNA and will also affect any host RNA present in the sample [81]. Secondly, many of the RNA-based library preparation protocols do not allow for barcoding, lowering the throughput of samples that can be sequenced. For DNA viruses these constraints are not applicable. One key limitation for large-scale application of long read Nanopore sequencing is the requirement for higher quantity and purity of the nucleic acid input compared to other sequencing technologies such as Illumina, which is more difficult to achieve for wild plant species. Given that most of the long-read plant virus metagenomics have been performed in crops, it remains to be seen whether Nanopore sequencing is readily amenable to larger metagenomic surveys of plant viruses in weeds and wild plants.

While other microbial communities, such as bacteria or fungi, may be routinely characterised through a conserved gene (16S and ITS, respectively), viruses have no such conserved genomic regions due to their varied evolutionary history. Since there is no universal marker for viruses, viral genome segments are recognised through sequence similarity to other viruses. This results in several problems in identifying viral metagenomes from weeds. Firstly, virus genome databases are incomplete, and there is currently a bias towards plant viruses found in crops. This means that viruses that occur exclusively in wild plants may not be recognised due to a lack of sequence similarity with known viruses. The significant proportion of viral metagenomics data that has no detectable sequence similarity with any known biological entity has been referred to as “dark matter”, and its viral nature remains an open question [17,51]. The observations of many researchers indicate that even enriching specifically for viral nucleic acids associated with viruses by VANA or dsRNA purification leads to the detection of many viral sequences that cannot be readily assigned to species, either due to their novelty or the short length of the sequence [8,17]. This pervasive ambiguity indicates the need to create new and innovative algorithms for discovering viruses based not only on sequence similarity, but also to use conserved, informative genome regions or domains for the clustering of viral sequences or the functional protein motifs utilised by viruses for productive infection [82,83]. Nevertheless, the taxonomic assignment of the viral contigs identified from HTS data faces many challenges that have been reviewed in detail [82]. Given the large number of novel agents uncovered in metagenomics studies, the most widely used approach, BLAST-based annotation, generally provides unreliable results at the species and genus level and still has weaknesses at the family level [17,82]. These problems raise questions of whether and how metagenomic sequence data should be incorporated into the ICTV taxonomy. To potentially improve this situation, Simmonds et al. [82] proposed solutions for virus classification including creation of new virus species, assigning new species and genera to existing families, and improvement of the procedure for the classification of viruses. These problems are not unique to viruses, and perhaps inspiration can be derived from initiatives taken with other microorganisms. For example, many bacterial and archaeal species detected by metagenomics cannot be cultured. There is a convention in place on

how to taxonomically classify these candidate species and further improvements on how to recognise the value of these genome sequences have been suggested [84,85].

Lastly, the analysis of virus metagenomics data is a practical challenge for non-specialised teams. Working with HTS-data requires computing equipment capable of storing and processing large datasets, as well as expertise on genome assembly analysis pipelines and packages. However, the decrease in its basic operational costs, further development of simple, reliable, and user-friendly bioinformatics tools for data analysis and sequence mapping will make the utilisation of metagenomics affordable to a larger group of researchers and phytosanitary services. Although virus detection using HTS is a powerful tool, in case of limited coverage across a viral genome and low sequencing depth, confirmatory tests may be required.

#### 4. Concluding Remarks

In summary, metagenomics approaches have a huge potential for detecting viruses, characterising viromes and exploring virus emergence and epidemiology. Plant virus metagenomic studies have shown that viruses are common in weeds and wild plants, even in the absence of symptoms, and these studies have uncovered many new, uncharacterised viruses. These results highlight how powerful HTS approaches are for virus discovery, whilst also showing the spectacular amount of viral biodiversity harboured by wild plants. Nonetheless, HTS methods have also some pitfalls and several challenges remain in this field of study: choices regarding protocols for enrichment and sequencing of plant viromes will affect the virus species and viral genetic variation that can be uncovered. This can be achieved by the wider implementation of plant virus metagenomics studies, which have the capacity to detect viruses alone or in mixed infection and the potential to reveal the presence of novel or unsuspected viruses [17,86]. Key questions about the association between plant diversity and virus prevalence, as well as how the biodiversity–disease relationship changes in plant populations adjacent to agriculture areas, remain open. However, by combining HTS approaches with an ecological perspective on plant viruses, we can take important steps towards addressing these issues in earnest.

**Author Contributions:** Conceptualisation, B.H.-J., D.B.; writing—draft preparation, B.H.-J., D.B., M.P.Z.; writing—review and editing, B.H.-J., D.B., M.P.Z. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by project 2017/25/B/NZ9/01715 (B.H.J.) from National Science Centre, Poland.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### References

1. Pecman, A.; Kutnjak, D.; Gutiérrez-Aguirre, I.; Adams, I.; Fox, A.; Boonham, N.; Ravnikar, M. Next Generation Sequencing for Detection and Discovery of Plant Viruses and Viroids: Comparison of Two Approaches. *Front. Microbiol.* **2017**, *8*, 1998. [[CrossRef](#)]
2. Susi, H.; Filloux, D.; Frilander, M.J.; Roumagnac, P.; Laine, A.L. Diverse and variable virus communities in wild plant populations revealed by metagenomic tools. *PeerJ* **2019**, *7*, e6140. [[CrossRef](#)] [[PubMed](#)]
3. Power, A.G.; Mitchell, C.E. Pathogen spillover in disease epidemics. *Am. Nat.* **2004**, *164*, 79–89. [[CrossRef](#)] [[PubMed](#)]
4. Elena, S.F.; Fraile, A.; García-Arenal, F. Evolution and emergence of plant viruses. *Adv. Virus Res.* **2014**, *88*, 161–191. [[CrossRef](#)] [[PubMed](#)]
5. Roossinck, M.J.; García-Arenal, F. Ecosystem simplification, biodiversity loss and plant virus emergence. *Curr. Opin. Virol.* **2015**, *10*, 56–62. [[CrossRef](#)]
6. Ma, Y.; Marais, A.; Lefebvre, M.; Faure, C.; Candresse, T. Metagenomic analysis of virome cross-talk between cultivated *Solanum lycopersicum* and wild *Solanum nigrum*. *Virology* **2020**, *540*, 38–44. [[CrossRef](#)]
7. Power, A.G.; Borer, E.T.; Hosseini, P.; Mitchell, C.E.; Seabloom, E.W. The community ecology of barley/cereal yellow dwarf viruses in Western US grasslands. *Virus Res.* **2011**, *159*, 95–100. [[CrossRef](#)]

8. Bernardo, P.; Charles-Dominique, T.; Barakat, M.; Ortet, P.; Fernandez, E.; Filloux, D.; Hartnady, P.; Rebelo, T.A.; Cousins, S.R.; Mesleard, F.; et al. Geometagenomics illuminates the impact of agriculture on the distribution and prevalence of plant viruses at the ecosystem scale. *ISME J.* **2018**, *12*, 173–184. [[CrossRef](#)] [[PubMed](#)]
9. Ng, J.C.; Falk, B.W. Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. *Annu. Rev. Phytopathol.* **2006**, *44*, 183–212. [[CrossRef](#)]
10. Hogenhout, S.A.; Ammar, E.D.; Whitfield, A.E.; Redinbaugh, M.G. Insect vector interactions with persistently transmitted viruses. *Annu. Rev. Phytopathol.* **2008**, *46*, 327–359. [[CrossRef](#)]
11. Pleydell, D.R.J.; Soubeyrand, S.; Dallot, S.; Labonne, G.; Chadœuf, J.; Jacquot, E.; Thébaud, G. Estimation of the dispersal distances of an aphid-borne virus in a patchy landscape. *PLoS Comput. Biol.* **2018**, *14*, e1006085. [[CrossRef](#)] [[PubMed](#)]
12. Malmstrom, C.M.; Hughes, C.C.; Newton, L.A.; Stoner, C.J. Virus infection in remnant native bunchgrasses from invaded California grasslands. *New Phytol.* **2005**, *168*, 217–230. [[CrossRef](#)] [[PubMed](#)]
13. Roossinck, M.J. Plant Virus Metagenomics: Biodiversity and Ecology. *Annu. Rev. Genet.* **2012**, *46*, 359–369. [[CrossRef](#)]
14. Prendeville, H.R.; Ye, X.; Morris, T.J.; Pilon, D. Virus infections in wild plant populations are both frequent and often unapparent. *Am. J. Bot.* **2012**, *99*, 1033–1042. [[CrossRef](#)]
15. Melcher, U.; Grover, V. Genomic approaches to discovery of viral species diversity of non-cultivated plants. In *Recent Advances in Plant*; Caranta, C., Aranda, M.A., Tepfer, M., López-Moya, J.J., Eds.; Virology, Caister Academic Press: Norfolk, UK, 2011; pp. 321–342. [[CrossRef](#)]
16. Muthukumar, V.; Melcher, U.; Pierce, M.; Wiley, G.B.; Roe, B.A.; Palmer, M.W.; Thapa, V.; Ali, A.; Ding, T. Non-cultivated plants of the Tallgrass Prairie Preserve of northeastern Oklahoma frequently contain virus-like sequences in particulate fractions. *Virus Res.* **2009**, *141*, 169–173. [[CrossRef](#)] [[PubMed](#)]
17. Roossinck, M.J.; Martin, D.P.; Roumagnac, P. Plant Virus Metagenomics: Advances in Virus Discovery. *Phytopathology* **2015**, *105*, 716–727. [[CrossRef](#)] [[PubMed](#)]
18. Roossinck, M.J. Mechanisms of plant virus evolution. *Annu. Rev. Phytopathol.* **1997**, *35*, 191–209. [[CrossRef](#)] [[PubMed](#)]
19. Martin, D.P.; van der Walt, E.; Posada, D.; Rybicki, E.P. The Evolutionary Value of Recombination Is Constrained by Genome Modularity. *PLoS Genet.* **2005**, *1*, e51. [[CrossRef](#)]
20. Syller, J. Facilitative and antagonistic interactions between plant viruses in mixed infections. *Mol. Plant Pathol.* **2012**, *13*, 204–216. [[CrossRef](#)]
21. Moreno, A.B.; López-Moya, J. When Viruses Play Team Sports: Mixed Infections in Plants. *Phytopathology* **2020**, *110*, 29–48. [[CrossRef](#)]
22. Cooper, I.; Jones, R.A. Wild plants and viruses: Under-investigated ecosystems. *Adv. Virus Res.* **2006**, *67*, 1–47. [[CrossRef](#)]
23. Bian, R.; Andika, I.B.; Pang, T.; Lian, Z.; Wei, S.; Niu, E.; Wu, Y.; Kondo, H.; Liu, X.; Sun, L. Facilitative and synergistic interactions between fungal and plant viruses. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 3779–3788. [[CrossRef](#)] [[PubMed](#)]
24. Jones, S.; Baizan-Edge, A.; Macfarlane, S.; Torrance, L. Viral Diagnostics in Plants Using Next Generation Sequencing: Computational Analysis in Practice. *Front. Plant Sci.* **2017**, *8*, 1770. [[CrossRef](#)] [[PubMed](#)]
25. Villamor, D.E.; Ho, T.; Al Rwahnih, M.; Martin, R.R.; Tzanetakis, I. High Throughput Sequencing for Plant Virus Detection and Discovery. *Phytopathology* **2019**, *109*, 716–725. [[CrossRef](#)] [[PubMed](#)]
26. Dunoyer, P.; Voinnet, O. The complex interplay between plant viruses and host RNA-silencing pathways. *Curr. Opin. Plant Biol.* **2005**, *8*, 415–423. [[CrossRef](#)] [[PubMed](#)]
27. Kreuze, J.F.; Perez, A.; Untiveros, M.; Quispe, D.; Fuentes, S.; Barker, I.; Simon, R. Complete Viral Genome Sequence and Discovery of Novel Viruses by Deep Sequencing of Small RNAs: A Generic Method for Diagnosis, Discovery and Sequencing of Viruses. *Virology* **2009**, *388*, 1–7. [[CrossRef](#)]
28. Wu, Q.; Luo, Y.; Lu, R.; Lau, N.; Lai, E.C.; Li, W.-X.; Ding, S.-W. Virus Discovery by Deep Sequencing and Assembly of Virus-Derived Small Silencing RNAs. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 1606–1611. [[CrossRef](#)]
29. Polischuk, V.; Budzanivska, I.; Shevchenko, T.; Oliynik, S. Evidence for plant viruses in the region of Argentina Islands, Antarctica. *FEMS Microbiol. Ecol.* **2007**, *59*, 409–417. [[CrossRef](#)] [[PubMed](#)]
30. Susi, H.; Laine, A.-L. Agricultural land use disrupts biodiversity mediation of virus infections in wild plant populations. *New Phytol.* **2021**, *230*, 2447–2458. [[CrossRef](#)] [[PubMed](#)]
31. Thapa, V.; McGlenn, D.J.; Melcher, U.; Palmer, M.W.; Roossinck, M.J. Determinants of taxonomic composition of plant viruses at the Nature Conservancy’s Tallgrass Prairie Preserve, Oklahoma. *Virus Evol.* **2015**, *1*, vev007. [[CrossRef](#)]
32. Richet, C.; Kraberger, S.; Filloux, D.; Bernardo, P.; Harkins, G.W.; Martin, D.P.; Roumagnac, P.; Varsani, A. Novel circular DNA viruses associated with *Apiaceae* and *Poaceae* from South Africa and New Zealand. *Arch. Virol.* **2019**, *164*, 237–242. [[CrossRef](#)]
33. Ma, Y.; Fort, T.; Marais, A.; Lefebvre, M.; Theil, S.; Vacher, C.; Candresse, T. Leaf-associated fungal and viral communities of wild plant populations differ between cultivated and natural ecosystems. *Plant-Environ. Inter.* **2021**, *2*, 87–99. [[CrossRef](#)]
34. Minicka, J.; Zarzyńska-Nowak, A.; Budzyńska, D.; Borodynko-Filas, N.; Hasiów-Jaroszewska, B. High-Throughput Sequencing Facilitates Discovery of New Plant Viruses in Poland. *Plants* **2020**, *9*, 820. [[CrossRef](#)] [[PubMed](#)]
35. Ibaba, J.D.; Gubba, A. High-Throughput Sequencing Application in the Diagnosis and Discovery of Plant-Infecting Viruses in Africa, A Decade Later. *Plants* **2020**, *9*, 1376. [[CrossRef](#)] [[PubMed](#)]
36. Adams, M.J.; Antoniw, J.F. DPVweb: An open access internet resource on plant viruses and virus diseases. *Outlooks Pest Manag.* **2005**, *16*, 268–270. [[CrossRef](#)]

37. Veerakone, S.; Tang, J.; Zheng, A.; Ward, L.I.; Mason, C. First Report of Clover yellow mosaic virus in Postentry Quarantine Tulipa gesneriana Plants Imported from the United States. *Plant Dis. Notes* **2018**, *102*, 464. [[CrossRef](#)]
38. Tomlinson, J.A. Epidemiology and control of virus disease of vegetables. *Ann. Appl. Biol.* **1987**, *110*, 661–681. [[CrossRef](#)]
39. Edwardson, J.R.; Christie, R.G. CRC Handbook of Viruses Infecting Legumes, 1st ed. CRC Press: Boca Raton, FL, USA, 1991; p. 546.
40. Guerret, M.G.L.; Nyalugwe, E.P.; Maina, S.; Barbetti, M.J.; van Leur, J.A.G.; Jones, R.A.C. Biological and Molecular Properties of a Turnip mosaic virus (TuMV) Strain that Breaks TuMV Resistances in Brassica napus. *Plant Dis.* **2017**, *101*, 674–683. [[CrossRef](#)]
41. Rymelska, N.; Borodynko, N.; Pospieszny, H.; Hasiów-Jaroszewska, B. Analysis of the biological and molecular variability of the Polish isolates of tomato black ring virus (TBRV). *Virus Genes* **2013**, *47*, 338–346. [[CrossRef](#)]
42. Venter, J.C.; Remington, K.; Heidelberg, J.F.; Halpern, A.L.; Rusch, D.; Eisen, J.A.; Wu, D.; Paulsen, I.; Nelson, K.E.; Nelson, W.; et al. Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **2004**, *304*, 66–74. [[CrossRef](#)] [[PubMed](#)]
43. Acevedo, A.; Brodsky, L.; Andino, R. Mutational and fitness landscapes of an RNA virus revealed through population sequencing. *Nature* **2014**, *505*, 686–690. [[CrossRef](#)]
44. Cuevas, J.M.; Willemsen, A.; Hillung, J.; Zwart, M.P.; Elena, S.F. Temporal Dynamics of Intrahost Molecular Evolution for a Plant RNA Virus. *Mol. Biol. Evol.* **2015**, *32*, 1132–1147. [[CrossRef](#)] [[PubMed](#)]
45. Dolja, V.V.; Koonin, E.V. Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res.* **2018**, *244*, 36–52. [[CrossRef](#)]
46. García-Arenal, F.; Zerbini, F.M. Life on the Edge: Geminiviruses at the Interface Between Crops and Wild Plant Hosts. *Annu. Rev. Virol.* **2019**, *6*, 411–433. [[CrossRef](#)] [[PubMed](#)]
47. Moffat, A.S. Plant pathology—Geminiviruses emerge as serious crop threat. *Science* **1999**, *286*, 1835. [[CrossRef](#)]
48. Rey, M.E.C.; Ndunguru, J.; Berrie, L.C.; Paximadis, M.; Berry, S.; Cossa, N.; Nuaila, V.N.; Mabasa, K.G.; Abraham, N.; Rybicki, E.P.; et al. Diversity of Dicotyledenous-Infecting Geminiviruses and Their Associated DNA Molecules in Southern Africa, Including the South-West Indian Ocean Islands. *Viruses* **2012**, *4*, 1753–1791. [[CrossRef](#)] [[PubMed](#)]
49. Rybicki, E.P.; Pietersen, G. Plant Virus Disease Problems in the Developing World. *Adv. Virus Res.* **1999**, *53*, 127–175. [[CrossRef](#)] [[PubMed](#)]
50. Claverie, S.; Bernardo, P.; Kraberger, S.; Hartnady, P.; Lefevre, P.; Lett, J.M.; Galzi, S.; Filloux, D.; Harkins, G.W.; Varsani, A.; et al. From Spatial Metagenomics to Molecular Characterization of Plant Viruses: A Geminivirus Case Study. *Adv. Virus Res.* **2018**, *101*, 55–83. [[CrossRef](#)] [[PubMed](#)]
51. Ma, Y.; Marais, A.; Lefebvre, M.; Theil, S.; Svanella-Dumas, L.; Faure, C.; Candresse, T. Phytoviroome Analysis of Wild Plant Populations: Comparison of Double-Stranded RNA and Virion-Associated Nucleic Acid Metagenomic Approaches. *J. Virol.* **2019**, *94*, e01462-19. [[CrossRef](#)] [[PubMed](#)]
52. Fraile, A.; García-Arenal, F. Environment and evolution modulate plant virus pathogenesis. *Curr. Opin. Virol.* **2016**, *17*, 50–56. [[CrossRef](#)]
53. Honjo, M.N.; Emura, N.; Kawagoe, T.; Sugisaka, J.; Kamitani, M.; Nagano, A.J.; Kudoh, H. Seasonality of Interactions between a Plant Virus and Its Host during Persistent Infection in a Natural Environment. *ISME J.* **2020**, *14*, 506–518. [[CrossRef](#)] [[PubMed](#)]
54. Pagán, I.; González-Jara, P.; Moreno-Letelier, A.; Rodelo-Urrego, M.; Fraile, A.; Piñero, D. Effect of Biodiversity Changes in Disease Risk: Exploring Disease Emergence in a Plant-Virus System. *PLoS Pathog.* **2012**, *8*, e1002796. [[CrossRef](#)] [[PubMed](#)]
55. Burdon, J.J.; Thrall, P.H. Pathogen Evolution across the Agro-Ecological Interface: Implications for Disease Management. *Evol. Appl.* **2008**, *1*, 57–65. [[CrossRef](#)] [[PubMed](#)]
56. Ostfeld, R.S.; Keesing, F. Effects of Host Diversity on Infectious Disease. *Annu. Rev. Ecol. Evol. Syst.* **2012**, *43*, 157–182. [[CrossRef](#)]
57. Cressler, C.E.; McLeod, D.V.; Rozins, C.; Van Den Hoogen, J.; Day, T. The adaptive evolution of virulence: A review of theoretical predictions and empirical tests. *Parasitology* **2016**, *143*, 915–930. [[CrossRef](#)]
58. Borer, E.T.; Adams, V.T.; Engler, G.A.; Adams, A.L.; Schumann, C.B.; Seabloom, E.W. Aphid fecundity and grassland invasion: Invader life history is the key. *Ecol. Appl.* **2009**, *19*, 1187–1196. [[CrossRef](#)]
59. Hall, G.S.; Peters, J.S.; Little, D.P.; Power, A.G. Plant community diversity influences vector behaviour and barley yellow dwarf virus population structure. *Plant Pathol.* **2010**, *59*, 1152–1158. [[CrossRef](#)]
60. Canto, T.; Aranda, M.A.; Fereres, A. Climate change effects on physiology and population processes of hosts and vectors that influence the spread of hemipteran-borne plant viruses. *Glob. Chang. Biol.* **2009**, *15*, 1884–1894. [[CrossRef](#)]
61. Hullé, M.; Cœur d’Acier, A.; Bankhead-Dronnet, S.; Harrington, R. Aphids in the face of global changes. *C. R. Biol.* **2010**, *333*, 497–503. [[CrossRef](#)]
62. Skendžić, S.; Zovko, M.; Živković, I.P.; Lešić, V.; Lemić, D. The Impact of Climate Change on Agricultural Insect Pests. *Insects* **2021**, *12*, 440. [[CrossRef](#)]
63. Das, T.; Hajong, M.; Majumdar, D.; Tombisana, R.; Devi, K.; Rajesh, T. Climate change impacts on plant diseases. *SAARC J. Agric.* **2017**, *14*, 200–209. [[CrossRef](#)]
64. Dixon, G.R. Climate change – impact on crop growth and food production, and plant pathogens. *Can. J. Plant Pathol.* **2012**, *34*, 362–379. [[CrossRef](#)]
65. Bezemer, T.M.; Jones, T.H.; Knight, K.J. Long term effects of elevated CO<sub>2</sub> and temperature on populations of the peach potato aphid *Myzus persicae* and its parasitoid *Aphidius matricariae*. *Oecologia* **1998**, *116*, 128–135. [[CrossRef](#)] [[PubMed](#)]
66. Jones, R.A.; Barbetti, M.J. Influence of climate change on plant disease infections and epidemics caused by viruses and bacteria. *CAB Rev. Perspect. Agric. Vet. Sci. Nutr. Nat. Resour.* **2012**, *7*, 1–31. [[CrossRef](#)]

67. Carmo-Sousa, M.; Moreno, A.; Plaza, M.; Garzo, E.; Fereres, A. Cucurbit aphid-borne yellows virus (CABYV) modifies the alighting, settling and probing behaviour of its vector *Aphis gossypii* favoring its own spread. *Ann. Appl. Biol.* **2016**, *169*, 284–297. [[CrossRef](#)]
68. Lecoq, H.; Desbiez, C. Cucurbit viruses in the Mediterranean region: An ever-changing picture. *Adv. Virus Res.* **2012**, *84*, 67–126.
69. Maclot, F.; Candresse, T.; Filloux, D.; Malmstrom, C.M.; Roumagnac, P.; van der Vlugt, R.; Massart, S. Illuminating an Ecological Blackbox: Using High Throughput Sequencing to Characterize the Plant Virome Across Scales. *Front. Microbiol.* **2020**, *11*, 2575. [[CrossRef](#)]
70. Colby, C.; Jurale, C.; Kates, J.R. Mechanism of synthesis of vaccinia virus double-stranded ribonucleic acid in vivo and in vitro. *J. Virol.* **1971**, *7*, 71–76. [[CrossRef](#)]
71. Son, K.N.; Liang, Z.; Lipton, H.L. Double-Stranded RNA Is Detected by Immunofluorescence Analysis in RNA and DNA Virus Infections, Including Those by Negative-Stranded RNA Viruses. *J. Virol.* **2015**, *89*, 9383–9392. [[CrossRef](#)]
72. Weber, F.; Wagner, V.; Rasmussen, S.B.; Hartmann, R.; Paludan, S.R. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J. Virol.* **2006**, *80*, 5059–5064. [[CrossRef](#)]
73. Ma, Y. Wild Plant Species Associated Viromes: Towards Improved Characterization Strategies and Variability in Various Ecological Environments. In *Human Health and Pathology*; English NNT: 2019BORD0134ff. fftel-0242625; Université de Bordeaux: Bordeaux, France, 2019.
74. Yockteng, R.; Almeida, A.M.R.; Yee, S.; Andre, T.; Hill, C.; Specht, C.D. A Method for Extracting High-Quality RNA from Diverse Plants for Next-Generation Sequencing and Gene Expression Analyses. *Appl. Plant Sci.* **2013**, *1*, 1300070. [[CrossRef](#)] [[PubMed](#)]
75. Holmes, E.C. *The Evolution and Emergence of RNA Viruses*; Oxford University Press: Oxford, UK, 2009; ISBN 9780199211128.
76. Ben Chehida, S.; Filloux, D.; Fernandez, E.; Moubset, O.; Hoareau, M.; Julian, C.; Blondin, L.; Lett, J.-M.; Roumagnac, P.; Lefeuvre, P. Nanopore Sequencing Is a Credible Alternative to Recover Complete Genomes of Geminiviruses. *Microorganisms* **2021**, *9*, 903. [[CrossRef](#)]
77. Adams, I.P.; Braidwood, L.A.; Stomeo, F.; Phiri, N.; Uwumukiza, B.; Feyissa, B.; Mahuku, G.; Wangi, A.; Smith, J.; Mumford, R.; et al. Characterising Maize Viruses Associated with Maize Lethal Necrosis Symptoms in Sub Saharan Africa. *bioRxiv* **2017**. [[CrossRef](#)]
78. Fellers, J.P.; Webb, C.; Fellers, M.C.; Shoup Rupp, J.; De Wolf, E. Wheat Virus Identification Within Infected Tissue Using Nanopore Sequencing Technology. *Plant Dis.* **2019**, *103*, 2199–2203. [[CrossRef](#)] [[PubMed](#)]
79. Phannareth, T.; Nunziata, S.O.; Stulberg, M.J.; Galvez, M.E.; Rivera, Y. Comparison of Nanopore Sequencing Protocols and Real-Time Analysis for Phytopathogen Diagnostics. *Plant Health Prog.* **2021**, *22*, 31–36. [[CrossRef](#)]
80. Filloux, D.; Fernandez, E.; Loire, E.; Claude, L.; Galzi, S.; Candresse, T.; Winter, S.; Jeeva, M.L.; Makesh Kumar, T.; Martin, D.P.; et al. Nanopore-Based Detection and Characterization of Yam Viruses. *Sci. Rep.* **2018**, *8*, 17879. [[CrossRef](#)]
81. Wongsurawat, T.; Jenjaroenpun, P.; Taylor, M.K.; Lee, J.; Tolardo, A.L.; Parvathareddy, J.; Kandel, S.; Wadley, T.D.; Kaewnapan, B.; Athipanyasilp, N.; et al. Rapid Sequencing of Multiple RNA Viruses in Their Native Form. *Front. Microbiol.* **2019**, *10*, 260. [[CrossRef](#)]
82. Simmonds, P.; Adams, M.; Benkő, M.; Breitbart, M.; Brister, J.; Carstens, E.; Davison, A.; Delwart, E.; Gorbalenya, A.; Harrach, B.; et al. Virus taxonomy in the age of metagenomics. *Nat. Rev. Microbiol.* **2017**, *15*, 161–168. [[CrossRef](#)]
83. Sobhy, H.A. A review of functional motifs utilized by viruses. *Proteomes* **2016**, *4*, 3. [[CrossRef](#)]
84. Konstantinidis, K.T.; Rosselló-Móra, R.; Amann, R. Uncultivated microbes in need of their own taxonomy. *ISME J.* **2017**, *11*, 2399–2406. [[CrossRef](#)]
85. Chuvochina, M.; Rinke, C.; Parks, D.H.; Rappé, M.S.; Tyson, G.W.; Yilmaz, P.; Whitman, W.B.; Hugenholtz, P. The importance of designating type material for uncultured taxa. *Syst. Appl. Microbiol.* **2019**, *42*, 15–21. [[CrossRef](#)] [[PubMed](#)]
86. MacDiarmid, R.; Rodoni, B.; Melcher, U.; Ochoa-Corona, F.; Roossinck, M. Biosecurity implications of new technology and discovery in plant virus research. *PLoS Pathog.* **2013**, *9*, e1003337. [[CrossRef](#)] [[PubMed](#)]