Plant virus identification and virus-vector-host interactions

Yahya Zakaria Abdou Gaafar

ORCID

0000-0002-7833-1542

Plant virus identification and virus-vector-host interactions

Dissertation

for the award of the degree

"Doctor rerum naturalium" (Dr.rer.nat.)

"Doctor of Philosophy" Ph.D. Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

within the International PhD Programme for Agricultural Sciences in Göttingen (IPAG)

of the Graduate School Forest and Agricultural Sciences (GFA)

submitted by

Yahya Zakaria Abdou Gaafar

from Cairo, Egypt

Göttingen, 2019

Thesis committee

Prof. Dr. Stefan Vidal

Georg-August University Göttingen, Department for Crop Sciences, Agricultural Entomology

Prof. Dr. Edgar Maiss

Leibniz Universität Hannover, The Institute of Horticultural Production Systems, Section of Phytomedicine

Dr. Heiko Ziebell

Julius Kühn-Institut (JKI), Federal Research Institute for Cultivated Plants, Institute for Epidemiology and Plant Diagnostics

Members of the Examination Board Reviewer

Prof. Dr. Michael Rostás

Georg-August University Göttingen, Department for Crop Sciences, Agricultural Entomology

Plant virus identification and virus-vector-host interactions

Printed in Germany.

For information contact; yahyaz.a.gaafar@gmail.com

Book layout and cover were designed by the author.

November 2019

Dedicated to my family

Contents

Chapter 1: General introduction

1.1. Plant viruses

Plant viruses are the cause of many crop diseases worldwide, leading to both yield and quality losses e.g., reduction in growth, vigour and market value (Bos, 1982; Hull, 2014). The type and severity of the host reactions to virus infections are very variable (Hull, 2014). They depend greatly on the virus strains, sources of infection, the time of infection, the crop genotypes and also influenced by environmental conditions (Hull, 2014). Several plant viruses are highly contagious and their effects on plants are often drastic. The losses caused by plant virus infections can have severe financial implications or have socio-economic effects (Anderson *et al.*, 2004; Hull, 2009; Patil *et al.*, 2015; Pechinger *et al.*, 2019).

Virus infection can cause histological changes to the cells and the intracellular structure (Hull, 2009). The symptoms caused by virus infection vary from necrotic or chlorotic lesions on inoculated leaves to systemic e.g., mosaic, mottle, stunting and leaves distortions (Hull, 2014). Nevertheless, some virus infections cause mild or no symptoms. Infection agents e.g., viroids and phytoplasma can induce diseases that resemble virus infections (Hull, 2014). Also, virus associated nucleic acids can alter the disease symptoms (Roossinck *et al.*, 1992; Ziebell & Carr, 2010). Some virus-like symptoms e.g., yellowing and necrosis, can easily be confused with non-viral disorders. Additionally, complex or multiple infections in plants are very common (Bos, 1982; Al Rwahnih *et al.*, 2009; Carvajal-Yepes *et al.*, 2014). These can consist of different viruses, or viruses with other pathogens, pests or abiotic factors. Such infections generally alter the plant physiology and in consequence the susceptibility and sensitivity to other infecting agents in an additive or synergistic effect or non-additive effect resulting in changes in the displayed symptoms (Bos, 1982; Syller, 2012). Furthermore, there are no simple relationships between virus content within a plant or virus incidence within a crop and the yield losses (Bos, 1982). Thus, detecting and identifying the exact virus causing the disease can be difficult.

However, determining the exact disease causative agent (known viruses, new viruses or virus strains) is necessary to decide which management strategy (e.g., insect vector control, resistance breeding, provision of virus-free germplasm etc) could be applied. In addition, it is important to detect quarantine viruses and prevent them from entering a country and becoming established. Therefore, it helps also in deciding on monitoring and preventive strategies, and in the prediction of plant diseases in annual crops.

1.2. Plant virus diagnostics:

Plant virus diagnosis often starts with spotting suspicious plants with "virus-like" symptoms in a field or greenhouse and sending it to diagnostics laboratories for analysis.

A range of techniques are available to detect and confirm the aetiology of the disease (Hamilton *et al.*, 1981; Hull, 2009). Important factors are taken in consideration when choosing the detection methods i.e., the sensitivity of the method to small amounts of viruses, accuracy, reproducibility, cost, time required, level of expertise needed and ability to perform in field (Hull, 2009). Additionally, the choice of test will determine the outcome i.e., whether only a virus family can be determined or a virus species or if strain-specific detection is possible.

1.2.1. Conventional detection methods:

Conventional detection methods have been developed and successfully implemented for a long time as virus detection tools and are widely used in many laboratories. Virus diagnosis has been relying on experienced specialists who can recognize and describe the disease causal agent from the symptoms on hosts, complemented with methods e.g., bioassays on indicator plants and electron microscopy (Boonham *et al.*, 2014). Conventional methods include bioassays, electron microscopy (EM), enzyme-linked immunosorbent assay (ELISA), Western, Northern and Southern blotting and polymerase chain reaction (PCR)-based methods. They provide rapid and inexpensive diagnoses for known viruses and viroids (Wu *et al.*, 2015). The following are descriptions of the commonly used methods in virus/viroid diagnosis.

1.2.1.1. Bioassays

Indicator plants have been used from the early ages of plant virology for propagation of plant viruses. Based on symptoms and host range reactions, differential host plants were used for the identification and classification of a number of plant viruses (Kirby *et al.*, 2001; Hull, 2014). However, a correct diagnosis based on symptoms is not possible and these days indicator plants are mainly used for virus propagation and enrichment for the subsequent use in different tests (e.g., electron microscopy).

1.2.1.2. Electron microscopy

Due to the small size of plant viruses, EM is the only technology that can directly visualise virus particles. The high resolution power of EM provides direct images at nanometre scale for virus diagnosis and research (Richert-Pöggeler *et al.*, 2018). Transmission electron microscopy (TEM) can be used as initial step in virus diagnosis from crud plant extracts without the necessity of viral enrichment (Bawden & Nixon, 1951; Gentile & Gelderblom, 2014; Richert-Pöggeler *et al.*, 2018). A main advantage of EM for viral diagnosis is that it does not require virus-specific reagents thus it provides an open view on the sample (Goldsmith & Miller, 2009). EM may not be able to identify a virus beyond the family level, thus additional assays can be performed which require virus-specific reagents e.g., antibodies. Immunosorbent electron microscopy (ISEM), as an example, increases the sensitivity of EM by virus trapping (Debrick, 1973; Roberts & Harrison, 1979). If specific antibodies are available, they can be used to “decorate” the virus particles and therefore differentiate between different species depending on the specificity of these antibodies. However, not always are specific antibodies available for all viruses. Additionally, an enrichment step is required for phloem restricted and low titre viruses, viruses without virions and viroids e.g., ultracentrifugation (Richert-Pöggeler *et*

al., 2018). Moreover, EM requires expertise, and EM facilities are not widely available (Naidu & Hughes, 2003). Thus, other methods such as serological or molecular assays can be performed additionally for specific virus identification and characterisation e.g., ELISA and PCR-based methods.

1.2.1.3. Enzyme-linked immunosorbent assay

The establishment of the ELISA assay was a revolution in virus diagnostics by simplifying virus detection and shortening the time required to reach conclusive results (Clark & Adams, 1977; Torrance & Jones, 1981). ELISA assays target proteins e.g., viral coat and movement proteins by antibodies. As a routine virus diagnostics test, ELISA is easy to use, cost effective, robust and scalable (Casper & Meyer, 1981; Torrance & Jones, 1981; Koenig & Paul, 1982). However, it requires the production of high-quality antisera which requires viral protein purification and expertise which can be lengthy procedure (Boonham *et al.*, 2014). Moreover, the antisera are often lack the sensitivity to correctly identify closely related virus strains, and in several cases it is not possible to differentiate viruses from the same genus due to cross-reactivity of antisera (Boonham *et al.*, 2014). ELISA is also difficult to use for multiple-target detection from one sample, as several tests need to be set up (Boonham *et al.*, 2014).

1.2.1.4. PCR-based methods

PCR-based methods e.g., classical PCR, reverse-transcription (RT)-PCR and real-time (or quantitative RCR [qPCR]) have been used for the diagnosis of plant viruses and viroids, and many assays have been published (Boonham *et al.*, 2014). PCR-based methods target nucleic acid sequences by primers. They require a reliable nucleic acid extraction method and sequence information of the viruses for primer design. PCR-based methods can be sensitive, inexpensive and require minimal skill to be performed. They have been used for plant virus detection since early 1990s (Vunsh *et al.*, 1990). PCR multiplexing allows the detection of multiple species or strains in a single reaction by combining specific primers for different viruses (Webster *et al.*, 2004). The specificity of PCR-based methods depends on the design of proper primers that are unique to the target virus/viroid. Moreover, virus-specific primers or probes can detect virus up to limited sequence variation, however new viruses, strains or divergent isolates will not be detected.

In general, molecular or serological testing are targeted methods which means they are limited to detecting the knowns. Additionally, in cases of mixed infections, such methods would likely miss the other disease causal agent. Therefore, virology diagnosticians need additional tools for diagnosing the unknowns and the variants.

1.2.2. Sequencing

1.2.2.1. Sanger chain termination

The Sanger chain termination method was developed in 1977 (Sanger *et al.*, 1977). This method uses labelled dideoxynucleotide (ddATP, ddGTP, ddCTP, or ddTTP) in four separated sequencing reactions which terminate DNA synthesis upon incorporation. The

generated sequences are then denatured and visualised by gel electrophoresis. In virus diagnostics, Sanger sequencing is commonly used to sequence PCR amplicons directly (Bernad & Duran-Vila, 2006; Hoang *et al.*, 2011). When multiple PCR amplicons are present in one reaction or sequence variation within one amplicon is suspected, PCR products can be cloned in bacterial cells, followed by selection and propagation of bacterial colonies followed by plasmids purification and Sanger sequencing. These sequences can be used for bioinformatic analysis and comparison with sequences available in public databases. Another application when using generic primers for the identification of more than one virus from the same genus or family (Abraham *et al.*, 2007). For unknown virus identification, a random-PCR method (rPCR) to construct whole cDNA library from sample RNA or dsRNA extracts, or library from rolling circle amplification (RCA) for circular DNA and enzymatic fragmentation are performed (Dodds *et al.*, 1984; Haan *et al.*, 1989; Froussard, 1992; Johne *et al.*, 2009). These approaches helped in the detection of many virus including new ones.

For long time, a typical Sanger sequencing reaction included the use of radioisotopes and other harsh chemicals (Wu *et al.*, 2015). This sequencing method was a labour-intensive process and only determined few hundred nucleotides (nt) at a time (Kircher & Kelso, 2010). With the development of capillary array electrophoresis and other detection systems, the production of 96-channel capillary HTS sequencers was possible e.g., 3730xl DNA Analyzer (Applied Biosystems, Inc) (Kambara & Takahashi, 1993; Kircher & Kelso, 2010). Such sequencer yield 96 or 384 sequences of about 600 to 1,000 nt per run and a maximum of about 1.5 Mb sequences per day with single-pass error rate of 0.1 to 1% (Wu *et al.*, 2015).

1.2.2.2. High-throughput sequencing

Over 15 years ago high-throughput sequencing (HTS), formerly known as next-generation sequencing (NGS) appeared on the market and revolutionized sequencing capabilities (van Dijk *et al.*, 2014a). It enabled parallel sequencing of millions of nucleic acid sequences in short time for comparatively low cost. Since then, a lot of progress has been made in read length, speed, throughput, and in costs reduction (van Dijk *et al.*, 2014a). These advances paved the way for the development of novel HTS applications in life sciences such as in diagnostics and metagenomics (Roossinck *et al.*, 2010; Roossinck *et al.*, 2015).

HTS was first applied for plant virus identification in 2009 using different preparation approaches and different sequencing platforms (Adams *et al.*, 2009; Al Rwahnih *et al.*, 2009; Kreuze *et al.*, 2009). It has proven very successful for virus discovery to resolve the disease aetiology in many agricultural crops (Roossinck *et al.*, 2015). It enables the simultaneous sequencing of total nucleic acid content of a sample, and thus detection, of any organism present in this sample. HTS carries the promise of generic and routine tool for virus detection. However, several steps need to be taken into consideration when applying HTS technologies: a) nucleic acid extraction and virus sequences enrichment; b) library preparation; c) automated sequencing; d) data analysis.

a) Nucleic acid extraction and virus sequences enrichment:

Viruses possess different genomes i.e., DNA (single or double stranded, circular or linear) or RNA (single [negative or positive sense] or double stranded, circular or linear) (Hull, 2009). Additionally, viroids possess circular single stranded RNA. Total DNA, total RNA, double stranded RNA (dsRNA) or small RNA (sRNA) extraction can be used for HTS detection of plant viruses and viroids. Therefore, many extraction protocols are available depending on the target nucleic acid.

For plant virus detection by HTS, virus sequence enrichment is required prior to sequencing to relatively increase the virus sequences in comparison to the host sequences (Wu *et al.*, 2015). Although the sizes of the virus/viroid genomes are small in comparison to other organisms, to achieve full genome coverage, the number of virus sequence reads must be high in order to trace their sequences that can be undetectable within the host overrepresented sequences which is challenging for bioinformatic analysis (Wu *et al.*, 2015; Adams & Fox, 2016). Additionally, there is no single reference gene or marker sequence shared by all viruses that could be used for virus identification as in case of other organisms such as fungi or bacteria where the internal transcribed spacer (ITS) or 16S ribosomal RNA can be used for general detection of these pathogens (Leff *et al.*, 2017).

There are different methods available for virus enrichment with the most common ones dsRNA, ribosomal RNA (rRNA) depletion, rolling circle amplification (RCA) and sRNA enrichments (Dodds *et al.*, 1984; Kreuze *et al.*, 2009; Roossinck *et al.*, 2010; Idris *et al.*, 2014; Knierim *et al.*, 2017). Each enrichment method has its advantages and disadvantages i.e., some are time consuming, some require large amounts of sample starting material, others may cause bias in the detection (not suitable for all viruses) (Wu *et al.*, 2015; Visser *et al.*, 2016; Pecman *et al.*, 2017).

The extraction and analysis of dsRNA has been used for plant virus detection for a long time (Morris, 1979; Tzanetakis & Martin, 2008; Okada *et al.*, 2015). dsRNA is produced as an intermediate during replication of RNA viruses and viroids and not "naturally" occurring in plants (Wu *et al.*, 2015). Extraction of dsRNA molecules is relatively easy, and the molecules are quite stable. Al Rwahnih and colleague compared HTS sequences derived from dsRNA or total RNA extracts from the same plant samples and found that virus reads increased from 2% to 53% with dsRNA enrichment (Al Rwahnih *et al.*, 2009). However, not all RNA viruses accumulate high concentrations of dsRNA, and DNA viruses have a different mode of replication, but few DNA virus sequences could also be recovered (Wu *et al.*, 2015). Depending on the dsRNA extraction method, high amounts of plant material may be required, other protocols are time consuming, or requiring an extra amplification step (Roossinck *et al.*, 2010; Romanovskaya *et al.*, 2013; Blouin *et al.*, 2016).

rRNA depletion of total RNA is useful for reducing the host sequences by removing the majority of the rRNAs which are highly abundant from the plant sample before further preparations (Adams & Fox, 2016). This can result in a 10-fold virus RNA enrichment (Adams & Fox, 2016). An additional step may be applied by selecting poly (A) tailed RNAs

(Visser *et al.*, 2016). However, this additional step is not useful for the identification of viruses that do not possess poly (A) tails.

Plants produce virus-derived small interfering RNAs (vsiRNAs) in response to virus infection (Ding & Lu, 2011). Moreover, the dsRNA replicative intermediates produced during the replication of viruses and viriods can also be processed into small interfering RNA in plants (Ding, 2010). Thus, all replicating viruses and viroids in a diseased plant can be detected by sRNA extraction. sRNA extraction by gel purification requires long time (up to two weeks). Nevertheless, a new extraction method was developed which can be performed within a day (Li *et al.*, 2013).

RCA approach takes the advantage of Phi 29 DNA polymerase and amplify circular DNA molecules in a given sample (Dean *et al.*, 2001; Idris *et al.*, 2014). Thus, RNA and non-circular DNA viruses cannot be amplified with this approach. Moreover, the Phi 29 polymerase may induce errors at early stages of amplification.

It is therefore desired to identify a generic approach in which it is possible to detect and identify all viruses and viroids with different genomes that can be easily used in diagnostic laboratories.

b) Library preparation:

The quality of sequencing data depends on the quality of the sequenced material. Thus, the library preparation must guarantee low bias and high complexity in order to achieve the most genomic coverage (van Dijk *et al.*, 2014b). Many library preparation protocols are available and most of them are compatible with the Illumina system (described later) (van Dijk *et al.*, 2014a; van Dijk *et al.*, 2014b).

General steps for library preparation i.e., the fragmentation of the nucleic acid, reverse transcription of RNA and dscDNA synthesis (when RNA is the starting material), adaptors and barcodes/indexes ligation, and with or without size selection and amplification (van Dijk *et al.*, 2014b). The library design may allow sequencing of both strands of DNA which increases accuracy. For single-molecule sequencing platforms (described later), the library preparation is minimal, where it involves template fragmentation, adapters ligation with or without amplification step (van Dijk *et al.*, 2014a). The choice of the protocol depends on the platform on which the sequencing will be run, and the performed study.

c) Sequencing platforms:

Roche 454 system was the first commercially HTS platform (Liu *et al.*, 2012). This platform uses the pyrosequencing technology and emulsion PCR as amplification approach (Margulies *et al.*, 2005). It produced initially 100–150 bp of sequence lengths and improved to 1 kb but with relatively low throughput (700 Mb per run), high error rates and high reagent cost (Metzker, 2010). Now the production of 454 is shut down and no more supporting by the company since 2016 (van Dijk *et al.*, 2014a). Following 454, several platforms were developed including Solexa/Illumina, SOLiD, Ion Torrent, PacBio and MinION platforms.

The Solexa/Illumina GA platform adopted sequencing by synthesis (SBS) technology and generated larger number (1Gb per run) but shorter reads compared to 454 platform (Liu *et al.*, 2012). Illumina SBS uses a proprietary reversible terminator-based method that detects single bases as they are incorporated into DNA template strands. Prior to sequencing, clonal clusters of amplified DNA fragments are generated through bridge amplification. Sequencing then starts with sequencing reagents containing a polymerase and fluorescently labelled nucleotides. Each nucleotide base is added, and the flow cell is photographed and the emission from each cluster is recorded. Each base is identified by the intensity and wavelength of the emission. This cycle is repeated to create the read length specified. Illumina SBS technology results in highly accurate base-by-base sequencing compared to other technologies. With the improvements in polymerases, buffers, flowcells, and software, several platforms from Illumina are currently available e.g., HiSeq, MiSeq and NextSeq. The size of the generated reads, their quality and amounts vary from device to another (Liu *et al.*, 2012; Reuter *et al.*, 2015).

The HiSeq and MiSeq platforms are the most established sequencers (Reuter *et al.*, 2015). MiSeq is a fast, benchtop sequencer, generates reads of 300 bp in length and up to 15 Gb per run (Illumina). HiSeq produces maximum read length of 150 bp and over 1 Tb per run (Illumina). NextSeq produces up 120 Gb per run and maximum read length of 150 bp (Illumina). Their error rate of Illumina platforms is <1% and the quality of the generated reads decreases with increasing the read length (Liu *et al.*, 2012; Reuter *et al.*, 2015). In general, Illumina platforms offer the highest throughput and the lowest cost per-base compare to other platforms (Liu *et al.*, 2012).

Sequencing by Oligo Ligation Detection (SOLiD) by Applied Biosystems uses Ligation-based sequencing technology and emulsion PCR for amplification (Mardis, 2008). On a SOLiD flowcell, the libraries are sequenced by 8 base-probes ligation which contains ligation site, cleavage site and fluorescent dyes. The fluorescent signals are recorded during the probes binding to the template strand then cleaved (Mardis, 2008). The cycle is repeated four times using ladder primer sets. SOLiD initially generated average read length of 25-35 bp (3 Gb data per run) and could reach a high accuracy of 99.85% after filtering (Mardis, 2008). Later other SOLiD sequencing systems were released with improvements in read length, data output and accuracy (Liu *et al.*, 2012). SOLiD 5500xl, for example, generates reads with 85 bp length (30 Gb per run) and 99.99% accuracy (Liu *et al.*, 2012).

Ion Torrent Personal Genome Machine (PGM) uses semiconductor sequencing technology (Flusberg *et al.*, 2010). PGM detects the changes in pH induced by the release of hydrogen ions with the incorporation of each nucleotide (Rothberg *et al.*, 2011). It does not require fluorescence and camera scanning, resulting in higher speed, lower cost, and smaller instrument size (Liu *et al.*, 2012). PGM generates reads with up to 400 bp length (1 Gb per run) with high error rates in homopolymers and insertions and deletions (Liu *et al.*, 2012; van Dijk *et al.*, 2014b). Ion Proton, the second Ion Torrent platform, increased the output compared to the PGM with 10 Gb per run and maximum read length was 200 bp (Reuter *et al.*, 2015).

Pacific Bioscience (PacBio) produced the single-molecule real-time (SMRT) sequencing platform. In this technology the clonal amplification is avoided, allowing direct sequencing of the DNA (Reuter *et al.*, 2015). The sequencing is performed with continuous polymerisation of the template in a zero-mode waveguides (ZMWs) with a single polymerase is positioned at its bottom and the presence of labelled nucleotides. With each incorporated base, fluorescent signals are captured in a video in real-time (Levene *et al.*, 2003; Eid *et al.*, 2009). It produces very long reads (60 kb; with >14 kb average read lengths) about 50k reads (up to 1 Gb of data in 4 hr). However, it has high error rates (about 11%) which are dominated by indels (Reuter *et al.*, 2015).

MinION nanopore sequencer is a single-molecule sequencing platform produced by Oxford Nanopore Technologies (ONT) (Ip *et al.*, 2015). On the membrane with nanopores "transmembrane proteins with nanoscale pore" incorporated, a voltage can be applied to drive DNA through the pore and an ion current flow can be measured. When a DNA molecule passes through the nanopore, a change of the current in pattern or magnitude can be observed and characterised. The nanopore can then discriminate individual nucleotides by measuring the change in electrical conductivity as DNA molecules pass through the pore (Lu *et al.*, 2016). The current in the nanopore is measured by a sensor several thousand times per second, and the data streams are passed to a microchip called the application-specific integrated circuit (ASIC) (Lu *et al.*, 2016). Finally, data processing is carried out by the MinKNOW software, which deals with data acquisition and analysis (Lu *et al.*, 2016).

The MinION device attracted considerable interest by scientists particularly in the fields of pathogen surveillance and diagnostics applications. MinION is the smallest sequencing device available, portable, and can produce sequence data in real-time (theoretically) (Lu *et al.*, 2016). It can be powered by the Universal Serial Bus (USB) ports of a computer with low hardware requirements and simple configurations but portable devices to operate MinION are now being offered by the ONT company. The main advantage of MinION sequencing is that it can generate longer reads of up to 100 Kb (Lu *et al.*, 2016). Additionally, the device has a low capital cost and its sequencing is cheap (Lu *et al.*, 2016). However, it has rather higher error rates (up 45%) (Lu *et al.*, 2016).

d) Data analysis:

Bioinformatic analysis is a crucial step in virus detection using HTS. There are several commercial software e.g., Geneious Prime and CLC genomic workbench, and open platforms such as Galaxy which can be used for HTS data analysis (Massart *et al.*, 2014). Automated bioinformatic pipelines for viruses detection are also available e.g., VirusDetect and Virtool (Rott *et al.*, 2017; Zheng *et al.*, 2017). The high efficiency and sensitivity of the different analysis pipelines in detecting known and novel virus are variable. Several available tools are designed for the detection of certain viruses using mapping to reference or for certain enrichment approach or for certain sequencing platforms e.g., Paparazzi (Vodovar *et al.*, 2011).

In general, the raw data generated from the sequencing platforms are then subjected to quality trimming to remove the adaptor sequences and low-quality nucleotides (Fig. 1) (Ho & Tzanetakis, 2014). This is followed by two ways based on the experimental design and the reason for sequencing. If the virus in the sample is known, the reference genome of the virus can be used for mapping the virus reads (Ho & Tzanetakis, 2014). However, when the cause of the disease is unknown, a *de novo* assembly is required (Zheng *et al.*, 2017). This is followed by searching the nucleotide and protein databases using Basic Local Alignment Search Tool (BLAST) to detect the closest sequences to the searched reads. Removing the host sequences by mapping to the host genome can be used to reduce the number of reads for the following analyses (Zheng *et al.*, 2017).

Massart and colleagues identified four crucial factors influencing the sensitivity of the bioinformatic analysis for virus/viroid detection i.e., the abundance of virus reads and the novelty of the virus, the assembly and annotation parameters, the completeness of the reference databases and the expertise in results interpretation (Massart *et al.*, 2019). The bioinformatic tools for HTS data analysis are under constant development and these efforts will improve, facilitate and speed up the application of HTS as a diagnostics tool in laboratories.

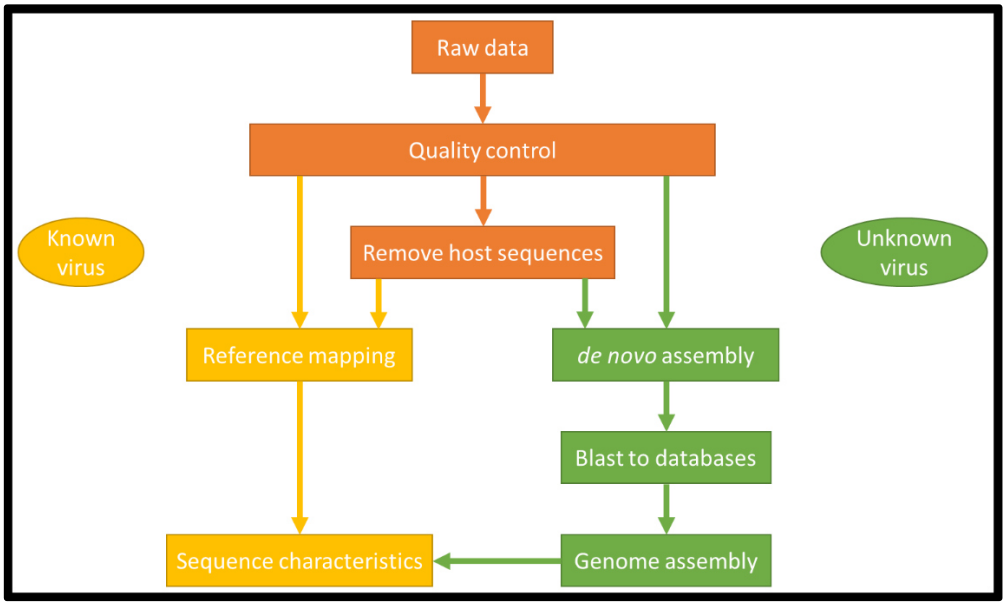

Figure 1: A diagram of two general strategies for HTS data analysis for known and unknown viruses

1.3. Plant virus transmission

Plant viruses possess many routes of transmission; they can be transmitted mechanically, by pollen, seeds or vectors such as nematodes, aphids and other insects. Aphids (*Hemiptera*, *Aphididae*) are the most common vector of plant viruses with fifty percent of the insect-vectored plant viruses are transmitted by aphids (Nault, 1997). The successful transport and transmission of viruses may be relying on aphids' high reproduction rates and their ability to spread for long distances.

Plant viruses are transmitted by aphids via different transmission modes depending on the aphid and the virus species: circulative or non-circulative, persistent, semi-persistent or non-persistent, propagative or non-propagative manners (Watson & Roberts, 1939; Sylvester, 1956; Harris, 1977; Ng & Perry, 2004; Ng & Falk, 2006). Depending on the mode of transmission, different strategies for preventing the spread of plant viruses need to be taken into consideration.

Many of the aphid-plant interactions take place at the cellular and tissue levels, especially as a virus vector (Collar *et al.*, 1997). As aphids are phloem feeders, acceptance of the phloem sap is an is required for aphid plant selection (Tjallingii, 1994). Changes in the chemical composition and physical structure of the plant tissues can affect the probing and feeding behaviours of the aphids (Guo *et al.*, 2014). Studying the aphid probing and feeding will reveal more details on the events involved in the transmission of plant viruses.

Electrical penetration graphs (EPG) have contributed substantially to the current knowledge of stylet penetration events in the plant tissue (Gabrys *et al.*, 1997; Pescod *et al.*, 2007; Brunissen *et al.*, 2009). Studying the aphid stylet penetration by electrical recording started with the AC system by Mclean and Kinsey and was further developed by Tjallingii (Mclean & Kinsey, 1964, 1965; Tjallingii, 1978, 1988). By attaching a gold wire electrode to the dorsum of the aphid and inserting the other electrode in the soil near to the plant root, once the aphid stylet penetrated the plant tissue the electrical circuit is completed (Tjallingii, 1985). The activity of the stylet in the plant tissue can then be recorded as waveforms. Each waveform is correlated to a certain stylet activity (Tjallingii, 1985).

EPG contributed to our understanding of virus transmission. Studies showed that plant virus infection can affect the behaviours and fitness of their vector (Castle & Berger, 1993; Eigenbrode *et al.*, 2002; Ziebell *et al.*, 2011). These effects are suggested to be related to the virus mode of transmission. Potato leafroll virus, a persistent transmitted circulative virus, enhanced the feeding behaviour of their aphid vector *Myzus persicae* Sulzer. Only after PLRV-infected potatoes showed visual symptoms, the aphids displayed a lower number of short probing periods before the first phloem activity and lower number and shorter total duration of derailed stylet mechanic (Alvarez *et al.*, 2007). Moreover, PLRV infection improved the aphid fitness by enhancing the aphids acceptance to the infected plant, making it more preferred and attractive for the aphid (Castle & Berger, 1993; Srinivasan & Alvarez, 2007). No effects were observed in case of potato virus

X, which is a mechanically transmissible and transmitted independently from insect vectors (Castle *et al.*, 1998; Eigenbrode *et al.*, 2002; Alvarez *et al.*, 2007; Srinivasan & Alvarez, 2007). In case of non-persistent virus (less vector-dependent) i.e., potato virus Y, it induced various effects on the behaviour of the vector depending on the species e.g., it enhanced the growth of *M. persicae* whereas it had no effect on *Macrosiphum euphorbiae* (Srinivasan & Alvarez, 2007; Boquel *et al.*, 2011).

In this thesis, the model system nanovirus-*Acyrthosiphon pisum*-*Vicia faba* was investigated further. Nanoviruses (members of the genus *Nanovirus*; family *Nanoviridae*) are multipartite viruses with at least eight circular ssDNA (Vetten *et al.*, 2011). They are considered a threat to important crops such as legumes (Vetten *et al.*, 2011). They are transmitted in circulative, non-propagative manner by various aphid species e.g., *A. pisum* and *Aphis craccivora* (Vetten *et al.*, 2011). The interactions between nanoviruses and their vectors is reviewed in detailed in chapter 5.

Nanoviruses are interesting model viruses as each genomic component encodes one protein, some of which with unknown functions (Vetten *et al.*, 2011). The availability of infectious clones for each component helped in understanding the role of each virus protein (Timchenko *et al.*, 1999; Timchenko *et al.*, 2000; Timchenko *et al.*, 2006; Grigoras *et al.*, 2009). Grigoras and colleagues found that the nuclear shuttle protein (NSP) encoded by component DNA-N is essential for aphid transmission (Grigoras *et al.*, 2018). The exact role of NSP in transmission is currently unknown. Nevertheless, this protein was found to be interacting with the stress granule component G3BP, the master replicase encoded by DNA-R and interacts with other NSPs in infected plants (Krapp *et al.*, 2017; Krenz *et al.*, 2017).

For faba bean necrotic yellows virus, a different nanovirus, a shorter plant access time is required for the inoculation minimum inoculation access period than for the acquisition ranging from 5 and 15min and 15 to 30 minutes, respectively, by both *A. craccivora* and *A. pisum* (Franz *et al.*, 1998). Phloem restricted plant viruses, such as nanoviruses, need to be ingested from the phloem cells in order to be transmitted by translocated through the gut, haemolymph and to the saliva for new infections. We expect a close relationship between the E2 waveforms (correlated to ingestion) in EPGs and nanoviruses acquisition. This could be experimentally confirmed for another circulative virus i.e., barely yellow dwarf virus (Prado & Tjallingii, 1994). The amount of acquired virus particles and proteins is expected to increase with the increase the E2 time. Whether the presence of nanovirus changes the behaviour as it is expected that nanovirus interacts with the aphids' body upon acquisition, or to change the host making it favourable by the vector, is currently unknown. Whether the absence of the NSP may also affect the behaviour is also unknown.

1.4. Aim and scope of the thesis:

- To identify the aetiology of several plant diseases using conventional and HTS tools (to be addressed in Chapter 2)
- To describe and characterise the viruses causing the disease (to be addressed in Chapter 2)
- To identify the virome of German peas and the spatio-temporal distribution of these viruses (to be addressed in Chapter 3)
- To compare three different viral enrichment approaches for HTS plant viruses/viroids detection (to be addressed in Chapter 4)
- To investigate the effect of nanoviruses infection on the feeding and probing behaviours of aphids (to be addressed in Chapter 6)
- To investigate the effect of FBNSV-NSP on the feeding and probing behaviours of aphids (to be addressed in Chapter 6)

Part one: Plant virus identification

Chapter 2: Plant disease aetiology

2.1. Characterisation of a novel nucleorhabdovirus infecting alfalfa (*Medicago sativa*)

Yahya Zakaria Abdou Gaafar, K. R. Richert-Pöggeler, C. Maaß, H.-Josef Vetten and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar YZA, Richert-Pöggeler KR, Maaß C, 2019. Characterisation of a novel nucleorhabdovirus infecting alfalfa (*Medicago sativa*). *Virology Journal* **16**, 113. doi: 10.1186/s12985-019-1147-3.

2.1.1. Abstract

2.1.1.1. Background

Nucleorhabdoviruses possess bacilliform particles which contain a single-stranded negative-sense RNA genome. They replicate and mature in the nucleus of infected cells. Together with viruses of three other genera of the family *Rhabdoviridae*, they are known to infect plants and can be transmitted by arthropod vectors, during vegetative propagation, or by mechanical means. In 2010, an alfalfa (*Medicago sativa*) plant showing virus-like symptoms was collected from Stadl-Paura, Austria and sent to Julius Kühn Institute for analysis.

2.1.1.2. Methods

Electron microscopy (EM) of leaf extracts from infected plants revealed the presence of rhabdovirus-like particles and was further used for ultrastructural analyses of infected plant tissue. Partially-purified preparations of rhabdovirus nucleocapsids were used for raising an antiserum. To determine the virus genome sequence, high throughput sequencing (HTS) was performed. RT-PCR primers were designed to confirm virus infection and to be used as a diagnostic tool.

2.1.1.3. Results

EM revealed bacilliform virions resembling those of plant-infecting rhabdoviruses. HTS of ribosomal RNA-depleted total RNA extracts revealed a consensus sequence consisting of 13,875 nucleotides (nt) and containing seven open reading frames (ORFs). Homology and phylogenetic analyses suggest that this virus isolate represents a new species of the genus *Nucleorhabdovirus* (family *Rhabdoviridae*). Since the virus originated from an alfalfa plant in Austria, the name alfalfa-associated nucleorhabdovirus (AaNV) is proposed. Viroplasms (Vp) and budding virions were observed in the nuclei of infected cells by EM, thus confirming its taxonomic assignment based on sequence data.

2.1.1.4. Conclusions

In this study, we identified and characterised a new nucleorhabdovirus from alfalfa. It shared only 39.8% nucleotide sequence identity with its closest known relative, black currant-associated rhabdovirus 1. The virus contains an additional open reading frame (accessory gene) with unknown function, located between the matrix protein and the glycoprotein genes. Serological and molecular diagnostic assays were designed for future screening of field samples. Further studies are needed to identify other natural hosts and potential vectors.

Keywords

electron microscopy; high throughput sequencing; Lucerne; rhabdovirus; alfalfa-associated nucleorhabdovirus

2.1.2. Background

Alfalfa or lucerne (*Medicago sativa* L.), a member of the *Fabaceae* family, is used as perennial forage crop which is important as fodder for livestock, as green manure for soil fertility, and can be used as food and medicine for humans (Marston *et al.*, 1943; Douglas *et al.*, 1995; Gray & Flatt, 1997; Peoples *et al.*, 2001). It is grown worldwide in temperate zones. Similar to other legumes, alfalfa is susceptible to a range of pests and pathogens (Samac *et al.*, 2016). Alfalfa can be infected by a large number of viruses such as alfalfa mosaic virus (AMV) (family: *Bromoviridae*) and two rhabdoviruses (alfalfa dwarf virus (ADV) and lucerne enation virus (LEV)) (Hull, 1969; Alliot & Signoret, 1972; Bejerman *et al.*, 2011; Bejerman *et al.*, 2015).

Members of the *Rhabdoviridae* family (order *Mononegavirales*) infect humans, invertebrates, vertebrates and plants (Augusto Lopez *et al.*, 1992; Longdon *et al.*, 2010; Bejerman *et al.*, 2011; Galinier *et al.*, 2012). Typically, their virions are bacilliform or bullet-shaped, composed of a helical nucleocapsid coated by a matrix layer and a lipid envelope while some have non-enveloped filamentous virions. The family has eighteen genera including 135 assigned species (Amarasinghe *et al.*, 2018). Sixteen genera have a monopartite genome while two are bipartite. Their genomes are linear and consist of negative-sense, single-stranded RNA (−ssRNA) (11–16 kb in length) and can comprise up to ten or more genes. They have five canonical genes that may be overprinted, overlapped or interspersed with additional accessory genes (Fu, 2005; Walker *et al.*, 2011; Walker *et al.*, 2015). Viruses of the genera *Cytorhabdovirus*, *Dichorhavirus*, *Nucleorhabdovirus* and *Varicosavirus* are known to infect plants (Walker *et al.*, 2018).

The genus *Nucleorhabdovirus* has currently ten assigned species. Nucleorhabdoviruses are known to be transmitted by leafhoppers (*Cicadellidae*), planthoppers (*Delphacidae*) and aphids (*Aphididae*) (Sylvester & Richardson, 1992; Nault, 1997; Walker *et al.*, 2018). Additionally, some can also be transmitted during vegetative propagation or by mechanical means. They can replicate in both plants and insect vectors (Goodin & Min, 2012). In plant cells, they replicate in the nucleus which becomes enlarged and develops large granular nuclear inclusions. They have non-segmented genomes, and like other rhabdoviruses they have highly conserved regulatory regions separating their genes, and complementary 3′ leader (l) and 5′ trailer (t) sequences. The 3′l and 5′t complementary sequence has the ability to form a putative panhandle structure suggested to be involved in genome replication (Jackson *et al.*, 2005).

With the advances in molecular techniques and bioinformatic tools, several new members of the *Rhabdoviridae* have been identified recently (Dilcher *et al.*, 2015; Axén *et al.*, 2017; Liu *et al.*, 2018; Økland *et al.*, 2018; Wu *et al.*, 2018). In this study, we succeeded in sap transmission of a rhabdovirus from *M. sativa* to *Nicotiana benthamiana* and identified it as a hitherto undescribed nucleorhabdovirus for which we propose the tentative name alfalfa-associated nucleorhabdovirus (AaNV).

2.1.3. Methods

2.1.3.1. Sample source and virus isolates used

During a survey in Stadl-Paura (Austria) in May 2010, a sample was collected by Dr. Herbert Huss from an alfalfa plant showing virus-like symptoms (symptoms were not recorded at the time) and sent to Julius Kühn Institute (JKI) for analysis. In initial attempts at virus isolation by sap transmission, the putative virus was transmitted to *N. benthamiana* seedlings as described below for further analysis and virus propagation (JKI ID 24249). For comparative studies, physostegia chlorotic mottle virus (PhCMoV; JKI ID 26372) and eggplant mottled dwarf virus (EMDV; JKI ID 29094) were maintained on *N. benthamiana* under greenhouse conditions by serial mechanical transmission.

2.1.3.2. Electron microscopy

For electron microscopy, small pieces (ca. 5 mm in diameter) of symptomatic leaves from *N. benthamiana* (5 to 7 weeks post inoculation) were directly homogenized in 2–5-fold volume of negative stain solution. This consisted of 2% ammonium molybdate, pH 6.5, with one drop of 0.5% bovine serum albumin (BSA) added. Viral particles were adsorbed by floating a pioloform carbon-coated copper grid for 5 min on the crude sap preparation. Finally, grids were rinsed with 5 drops of 2% ammonium molybdate and dried. The preparations were used for size measurements of virions including spikes.

Immunosorbent electron microscopy (ISEM) and immunoelectron microscopy (IEM) decoration experiments targeting the viral nucleocapsid protein were done as described in (Milne, 1984; Milne & Lesemann, 1984), using the JKI-1607 antiserum to AaNV. Fragments (ca. 2 mm in diameter) of a younger frizzy leaf from systemically infected *N. benthamiana* were embedded in Epon 812 after consecutive fixation of samples with 2.5% glutaraldehyde and 0.5% osmium tetroxide.

Ultrathin sections of 70 nm were prepared with an ultramicrotome UC7 (Leica, Germany) using a DiATOME diamond knife (Switzerland) and were placed on 75 mesh pioloform carbon-coated nickel grids. The grids were stained with 1% uranyl acetate for 30 min and grids were examined in a Tecnai G2 Spirit electron microscope at 80 kV. Images were taken with a 2 K Veleta camera. Brightness and contrast were adjusted when necessary using Adobe Photoshop CS6.

2.1.3.3. Purification of rhabdovirus nucleocapsids

Isolation of rhabdovirus nucleocapsids was performed using a modification of a method described by Roggero et al. (Roggero *et al.*, 2000; Verbeek *et al.*, 2013). Briefly, 100 g infected leaf materials of *N. benthamiana* were blended for 1 min in 500 ml homogenisation buffer consisting of 100 mM Tris-HCl, pH 8, containing 20 mM sodium sulfite, 10 mM Na-DIECA and 5 mM Na-EDTA. The homogenate was filtered through cheesecloth and centrifuged at 3000 rpm for 10 min in a GSA rotor (Sorvall). The supernatant was centrifuged at 25,000 rpm for 30 min in a 45 Ti fixed-angle rotor (Beckman Coulter), and the pellets were resuspended in 180 ml homogenisation buffer plus 2% (w/v) lauryl sulfobetaine and stirred for 1 h at 4 °C, followed by centrifugation at

9000 rpm for 10 min in a GSA rotor (Sorvall). The supernatant was placed onto a 20% sucrose cushion in homogenisation buffer (3.5 ml/tube) and ultracentrifuged at 25,000 rpm for 2.5 h in a SW 28 Ti rotor (Beckman Coulter). Then, the pellets were resuspended in 1 ml 10 mM Tris-HCl, pH 8, and centrifuged at 14,000 rpm in a MiniSpin centrifuge (Eppendorf). The supernatant was then placed onto preformed cesium sulfate-gradients (260, 405 and 575 mg/ml [w/v], respectively) in 10 mM Tris-HCl, pH 8, and ultracentrifuged at 35,000 rpm for 20 h in a SW 55 Ti rotor (Beckman Coulter). Opalescent bands were collected with a peristaltic pump, diluted to 25 ml with 10 mM Tris-HCl, pH 8, and ultracentrifuged at 40,000 rpm for 3 h in a 70 Ti rotor (Beckman Coulter). The resulting pellet was resuspended in 5.5 ml 10 mM Tris-HCl, pH 8, and used for nucleocapsid quantification by UV spectroscopy, for EM examination and antiserum production.

2.1.3.4. Antibody production and serological detection

For production of an antiserum to AaNV (designated JKI-1607), a purified nucleocapsid preparation (approximately 250 µg/ml in 0.01 M Tris-HCl, pH 8.0) was mixed with Freund's complete adjuvant (1:1) and injected directly into the hindleg muscles (IM) of a cross-bred rabbit. Such injections were repeated two times using Freund's incomplete adjuvant after 1 week and after 9 weeks. One week after the last injection, the rabbit was bled at weekly intervals for 1 month. Immunoglobulin G (IgG) isolation and conjugate production were performed according to (Clark & Adams, 1977). The specificity of the AaNV IgGs was tested at a dilution of 1:1000 [v/v] in a DAS-ELISA format using extracts from EMDV-, PhCMoV- and AaNV-infected *N. benthamiana (Clark & Adams, 1977)*. In reciprocal DAS-ELISA experiments, antisera to EMDV (JKI-1073) and PhCMoV (JKI-2051) were tested against extracts from AaNV-inoculated plants (upper, non-inoculated leaves). DAS-ELISA was also performed to confirm the presence of AaNV in plants inoculated for the (limited) host range study. The calculation of cut-off values for each ELISA plate carried out according to the Technical Information by Bioreba (Bioreba, 2014).

2.1.3.5. Whole genome sequencing

Total RNA (totRNA) was extracted from *N. benthamiana* infected leaf material using innuPREP RNA Mini Kit (Analytik Jena AG, Jena, Germany) following the manufacturer's protocol. Ribosomal RNA (rRNA) was depleted using RiboMinus Plant kit (Invitrogen, Carlsbad, CA, USA) according the manufacturer's protocol. Random cDNA was synthesized using ProtoScript II Reverse Transcriptase (New England Biolabs, Beverly, MA, USA) and 8 N random primers. The second strand was synthesized with NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module kit (New England Biolabs (NEB), Beverly, MA, USA). A library was prepared using Nextera XT Library kit (Illumina) and subsequently run on a MiSeq v3 platform as pair-end reads (2 × 301).

2.1.3.6. Sequencing of 5' and 3' ends

To obtain the 5' and 3' ends of the full-length AaNV sequence, RNA ligase mediated amplification of cDNA ends (RLM-RACE) (Liu & Gorovsky, 1993; Coutts & Livieratos, 2003; Li *et al.*, 2005) and RNA poly A tailing were used, respectively.

For the 5′ end, cDNA was produced using a virus specific primer (HZ-454 5′ ACT CTT GGT ACA GCA ACT CGT 3′) located 461 bases from the end. The resulting cDNA was purified using the DNA Clean & Concentrator kit (Zymo Research, Orange, CA, USA). An adaptor (Oligo1rev 5′ PO_4-GAT CCA CTA GTT CTA GAG CGG C-AminoC3 cordycepin 3′ adapted from (Coutts & Livieratos, 2003)) was ligated to the cDNA using T4 RNA ligase 1 (NEB) and the ligated cDNA was purified. PCR amplification of the 5′ end was performed using a primer (Oligo2for 5′ GCC GCT CTA GAA CTA GTG GAT C 3′) complementary to the ligated adaptor and a virus specific primer (HZ-452 5′ TCC ACA AGT TGC AAG CAG GT 3′) 397 bases from the genome end. A band of approximately 400 bases was obtained.

For obtaining the 3′ end, totRNA was poly-A tailed with the A-Plus™ Poly(A) Polymerase Tailing kit (Cellscript, Madison, WI, USA) and cDNA was synthesized using a primer (HZ-413 5′ GGA CAT TGT CCG GAT GGT CT 3′) binding 361 bases from the 3′ end of the RNA. The 3′ end was amplified by PCR using HZ-413 and oligo(d)T primer (5′ CCT CGG GCA GTC CTT TTT TTT TTT TTT T 3′) (Fletcher *et al.*, 2016).

The PCR products of both ends were cleaned using the Zymoclean Gel DNA Recovery (Zymo Research) and cloned with NEB PCR Cloning Kit (NEB). Cloning and plasmid amplification were carried out according to the manufacturer's instructions. Purification of plasmids was carried out using the NucleoSpin Plasmid EasyPure Kit (Macherey-Nagel, Düren, Germany); sequencing (ten colonies in both directions) was carried out at Macrogen (Seoul, Korea) and Eurofins Genomics (Ebersberg, Germany).

2.1.3.7. Sequence analysis

The reads produced from the MiSeq platform were analysed with Geneious software (v 11.0.4) (Biomatters Limited, Auckland, New Zealand). The raw reads were quality trimmed (error limit = 0.05), size filtered > 99 nt, error corrected and normalised using BBNorm (v. 37.64) tool, followed by de novo assembly with Geneious assembler. Assembled contigs were then used to search for similar sequences by BLASTn and BLASTx using the National Centre for Biotechnology Information (NCBI) GenBank non-redundant nucleotide and protein databases, respectively. Mapping of the clean reads to the complete viral genome sequence as a reference was performed using the mapping to reference tool in Geneious. Open reading frames were identified by Find ORF tool and were used to find similar sequences and conserved domains in BLASTp.

Sequence alignments were done with clustalW and phylogenetic trees (Neighbour-Joining algorithm, 1000 bootstrap replications) were created using MEGA 7.0.26 (Larkin *et al.*, 2007; Kumar *et al.*, 2016). The full genome of the virus was submitted to GenBank using Sequin application (v 15.50). Importin-dependent nuclear localisation signals were predicted using cNLS Mapper (Kosugi *et al.*, 2009) and nuclear export signals (NES) were predicted using NetNES 1.1 (La Cour *et al.*, 2004).

2.1.3.8. Reverse transcription polymerase chain reaction (RT-PCR) for detection and confirmation

Two primers (HZ-408 5′ GCA CGA TAA AGG CTG CAT CG 3′ and HZ-409 5′ TTG TGC ATC CTC TGT CGG AC 3′) were designed (Geneious design new primer tool) to confirm the virus presence by RT-PCR. The primers were designed to amplify a 971 bp fragment of the RNA-dependent RNA polymerase gene.

Extraction of totRNA was done from leaf tissues as described above, and cDNA was produced using HZ-409 primer. The cDNA product was used for PCR using OneTaq DNA Polymerase kit (NEB) (35 cycles of 30 s at 94 °C, 45 s at 52 °C, 1 min at 68 °C and a final elongation step for 4 min at 68 °C). The amplified PCR products were subject to electrophoresis on a 1.0% (w/v) agarose gel stained with ethidium bromide. The specificity of the designed primers was confirmed by testing EMDV- and PhCMoV-infected plants.

2.1.3.9. Infectivity assays

AaNV-infected *N. benthamiana* leaves were used to inoculate *N. benthamiana*, *M. sativa*, *M. lupulina*, *Pisum sativum* and *Vicia faba* mechanically. Briefly, symptomatic leaves were homogenized in Norit inoculation buffer (50 mM phosphate buffer, pH 7, containing 1 mM ethylenediaminetetraacetic acid (Na-EDTA), 20 mM sodium diethyldithiocarbamic acid (Na-DIECA), 5 mM thioglycolic acid, 0.75% activated charcoal and 30 mg Celite). Using a glass spatula, the homogenate was gently rubbed onto the leaves which were then rinsed with water. The inoculated plants were kept under greenhouse conditions (at 22 °C; photoperiod of 16 h light [natural daylight with additional growth light Phillips IP65, 400 W] and 8 h dark) and regularly inspected for symptoms for at least three weeks after inoculation.

2.1.4. Results

2.1.4.1. Virus transmission and maintenance

Upon receiving the infected alfalfa sample, the virus was mechanically inoculated onto standard indicator plants including *N. benthamiana* which were inspected for symptoms twice weekly. In *N. benthamiana*, chlorotic lesions appeared on inoculated leaves followed by systemic leaf rolling, mottling and yellowing in week three or four post inoculation. The virus was maintained continuously on *N. benthamiana* by regular mechanical passage onto young seedlings.

2.1.4.2. Virus morphology and cellular localisation

To elucidate the aetiology of the alfalfa disease, transmission electron microscopy (TEM) was performed on infected *N. benthamiana* plants following mechanical inoculation. Bacilliform-shaped virus particles were observed (Fig. 1). Using ammonium molybdate instead of uranyl acetate as negative stain was less disruptive on particle appearance. Only few mature virions displaying various degrees of disruption were detected in adsorption preparates. Preliminary measurements obtained from $n = 40$ revealed virion sizes ranging from 180 to 200 nm in length and 85–95 nm in diameter. The outer surface of virions is preserved comprising the lipid bilayer carrying the spikes, likely

glycoproteins. Virions shown in Fig. 1 are less disrupted with matrix proteins and envelope mostly intact. About 15% of the measured particles like those depicted in Fig. 1 were of shorter size (average length 167 nm) and may indicate defective particles not comprising the complete viral genome.

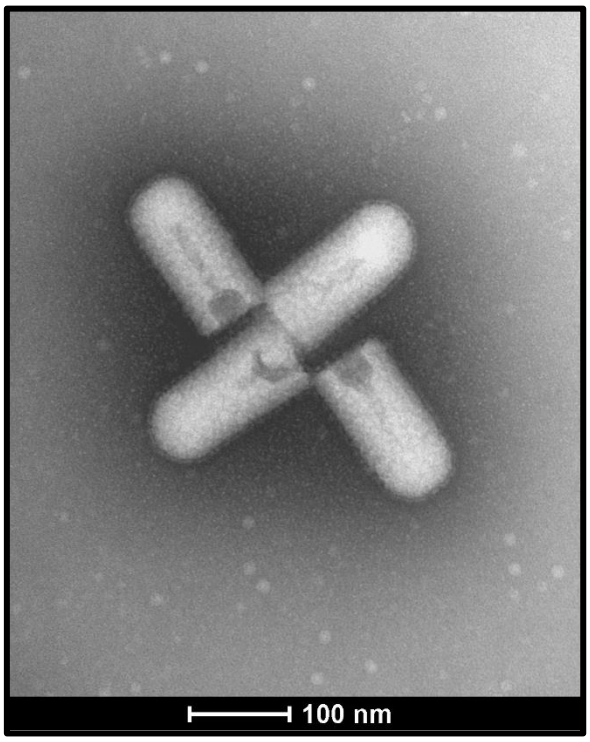

Figure 1: Electron micrograph of crude plant sap preparations of AaNV-infected *N. benthamiana* leaves. Four shorter mature bacilliform virions with average sizes of 167 nm in length and 86 nm in diameter

When ultrathin sections of embedded symptomatic *N. benthamiana* leaf tissue were analysed, very few virus particles were found in the cytoplasm only. Figure 2a shows two virus particles in epidermal cells. The transversely cut particle seems to be complete with attachment of glycoproteins visible (lower arrow, left hand side). Figure 2a (upper arrow, right hand side) seems to show two longitudinally particles appearing blunt end to blunt end and thus looking like a larger particle. Both epidermal and mesophyl cells were infected.

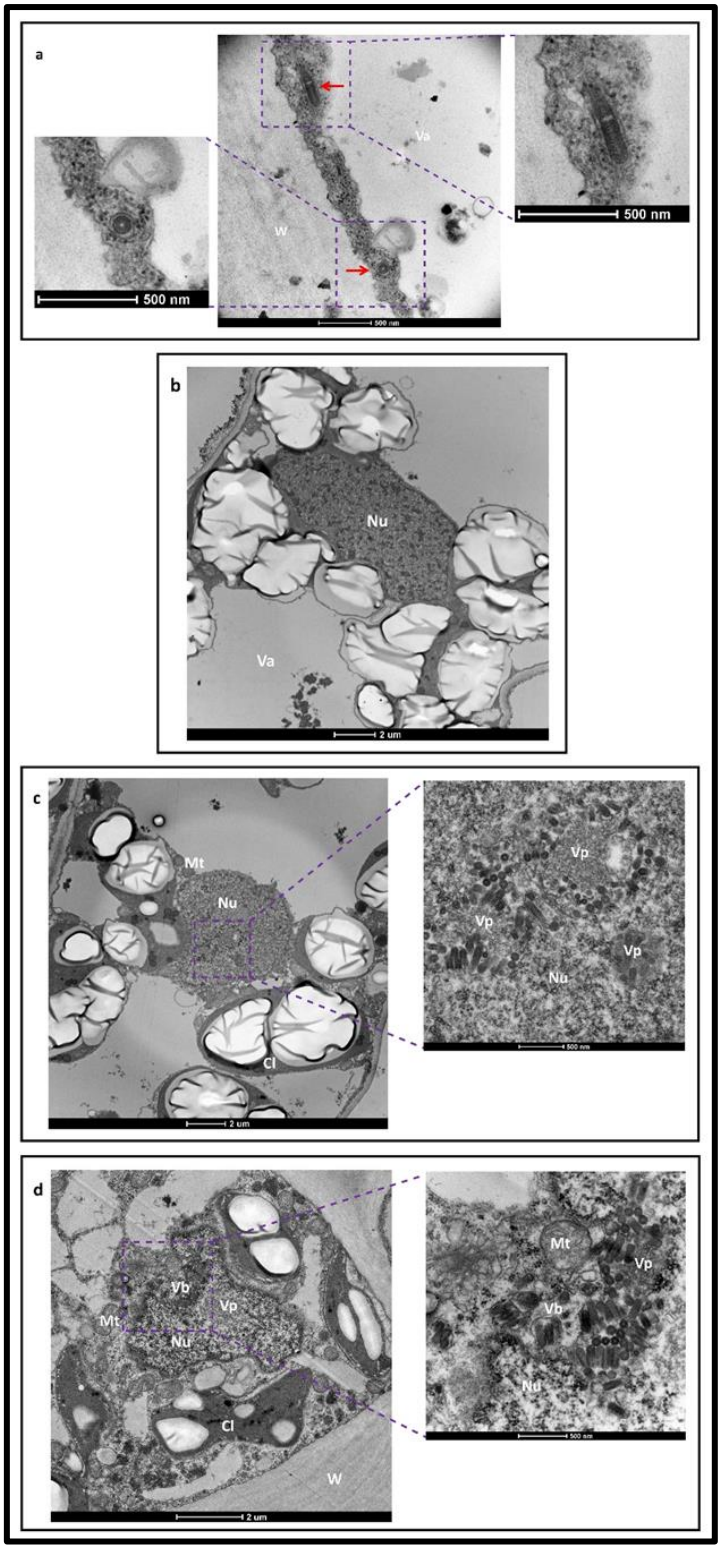

Figure 2: Electron micrographs of thin sections of AaNV-infected *N. benthamiana* cells. (**a)** Arrows indicate transversely (bottom) and longitudinally (upper part) cut particles in the cytoplasm located between the cell wall (W) and the vacuole (Va) of an epidermal cell; **(b)** Non-infected nucleus with heterochromatin and homogenous nuclear matrix; **(c)** Electron-dense granular areas throughout the nucleus are thought to represent viroplasm (Vp) with virions placed next to it; (**d)** arrays of mature virions budding (Vb) into the perinuclear space surrounded by the

nuclear membrane. The cell wall (W), chloroplast with starch granules (Cl), nucleus (Nu), mitochondrion (Mt), vacuole (Va), virus budding (Vb) and viroplasm (Vp) are indicated

In infected cells, the shape of the nucleus can change to a more condensed circular or even a distorted shape (Fig. 2c-d) compared to the ellipsoidal form present in healthy cells (Fig. 2b). In heavily infected cells, not only the nuclear compartments were affected but also chloroplasts were deformed (Fig. 2d). In the nuclei, granular areas distinct from heterochromatin were found representing putative virus replication sites known as viroplasms (Vp). Adjacent to them virus particles could be found (Fig. 2c and d). In Fig. 2c, there are few virions around the Vp and no virions were observed in the cytoplasm. In addition, vesicles or virus buddings (Vb) containing one or more complete viruses were visualized around the nucleus and in the cytoplasm of infected cells (Fig. 2d). Figure 2d also shows virus particles budding from the inner nuclear envelope in the perinuclear space.

2.1.4.3. Sequence analysis

A total of 1,561,664 reads were generated from the MiSeq sequencing. After quality trimming and size filtering, 1,141,662 quality-filtered reads were used for normalisation and de novo assembly. From the 23,180 assembled contigs, a contig of 13,854 nucleotides showed 66.9% identity (7% coverage and 3e-50 E-value) to black currant-associated rhabdovirus 1 (BCaRV-1) (MF543022), 66.2% (6% coverage and 2e-45 E-value) to datura yellow vein virus (DYVV) (NC_028231) and 66.2% (9% coverage and 5e-41 E-value) to sonchus yellow net virus (SYNV) (NC_001615) using BLASTn. Using BLASTx, the contig shared 44.9% (34% coverage and zero E-value) identity to DYVV (YP_009176977), 43.62% identity (35% coverage and zero E-value) to SYNV (NP_042286) and 43.5% (34% coverage and zero E-value) to BCaRV-1 (AUW36419). To determine the 5’ and 3’ends, primers were designed to anneal close to the assembled contig ends. The sequences of the two ends were assembled to the contig and the full-length genome sequence was determined as 13,875 bases in length with 29,727 mapped reads, 40.6% G + C content and mean coverage of 586X (GenBank accession number MG948563). The sequencing dataset generated in this study is available from the corresponding author upon request.

A pairwise nucleotide sequence alignment of the novel genome to selected rhabdoviruses and a phylogenetic tree was generated. ClustalW pairwise analysis showed that the AaNV sequence has 39.8% nt identity to BCaRV-1, 38.8% to DYVV and SYNV (Supplementary Fig. S1a). Moreover, the AaNV sequence falls within the genus *Nucleorhabdovirus* in a clade with SYNV, BCaRV-1 and DYVV (Supplementary Fig. S1b). This clustering was supported by a neighbour joining tree of the L protein amino acid sequences of selected members of the family *Rhabdoviridae* (Fig. 3).

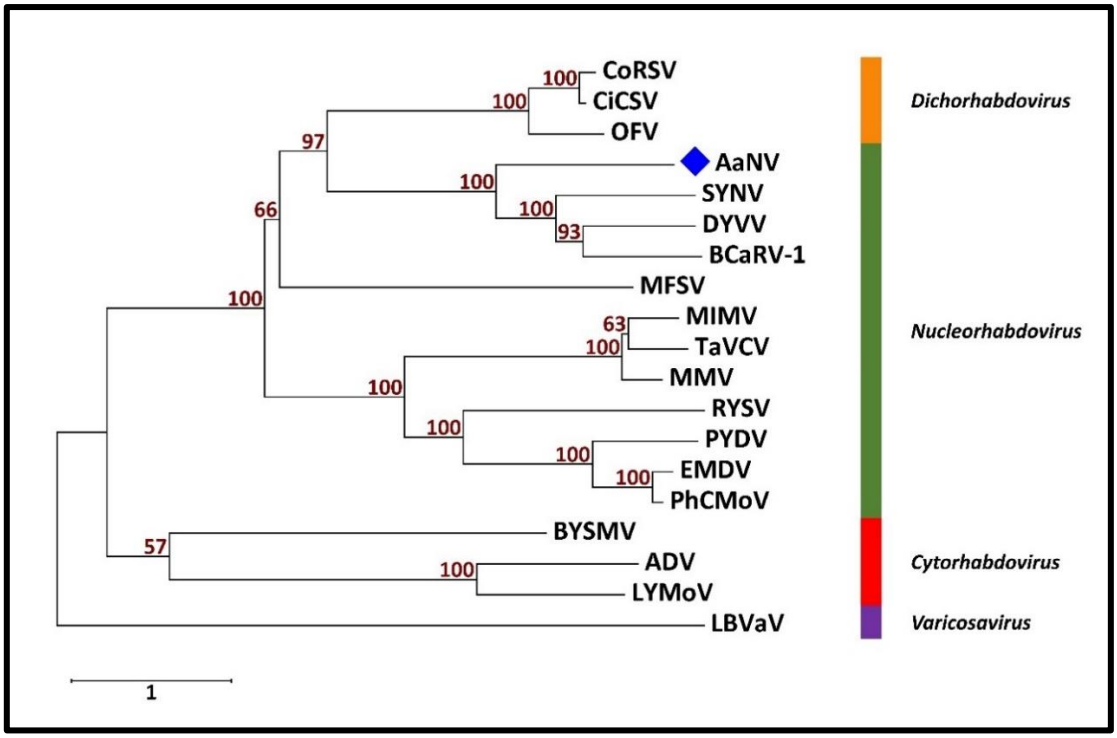

Figure 3: Unrooted neighbour-joining phylogenetic tree [Genetic distance model (Jones-Taylor-Thornton (JTT) model) and 1000 bootstrap replications] based on the amino acid sequence alignment of the L protein of selected members of different genera of the family *Rhabdoviridae*. AaNV indicated by a blue solid diamond shape. The bootstrap values above 50% are indicated for each node. The names and the accession numbers of the viruses are as follow: *Nucleorhabdovirus* (green): alfalfa-associated nucleorhabdovirus (AaNV; QAB45076), black currant-associated rhabdovirus 1 (BCaRV-1; AUW36419), datura yellow vein virus (DYVV; YP_009176977), eggplant mottled dwarf virus (EMDV; YP_009094358), maize fine streak virus (MFSV; YP_052849), maize Iranian mosaic virus (MIMV; YP_009444713), maize mosaic virus (MMV; YP_052855), physostegia chlorotic mottle virus (PhCMoV; ARU77002), potato yellow dwarf virus (PYDV; YP_004927971), rice yellow stunt virus (RYSV; NP_620502), sonchus yellow net virus (SYNV; NP_042286) and taro vein chlorosis virus (TaVCV; YP_224083). *Cytorhabdovirus* (red): alfalfa dwarf virus (ADV; YP_009177021), barley yellow striate mosaic virus (BYSMV; YP_009177231) and lettuce yellow mottle virus (LYMoV; YP_002308376). *Dichorhabdovirus* (orange): citrus chlorotic spot virus (CiCSV; ARJ35804), coffee ringspot virus (CoRSV; YP_009507905), orchid fleck virus (OFV; YP_001294929). *Varicosavirus* (violet): lettuce big-vein associated virus (LBVaV; YP_002308576)

2.1.4.4. The genome organisation of AaNV

Six putative open reading frames (ORFs) were identified in the antigenomic sense based on the genome organisation described for other nucleorhabdoviruses; nucleocapsid (N), phosphoprotein (P), putative cell to cell movement protein (P3), matrix protein (M), glycoprotein (G) and RNA-dependent RNA polymerase (L). Highly conserved regulatory regions separating the genes were identified. At the junctions between the genes, the consensus motif is (3' UAA AUU CUU UUU GGU UG 5'), which slightly differs between the 3' leader and N gene, and between the L gene end and the 5' trailer (Fig. 4a). The presence of a seventh ORF with unknown function (U), between M and G was

identified as it is flanked by the intergenic region consensus motif. Moreover, the 3′ leader (l) and the 5′ trailer (t) have complementary sequences of 43.1% nt identity (Fig. 4b). Therefore, the complete genome organisation was determined as 3′ l–N–P–P3–M–U–G–L–t 5′ (Fig. 4c). Additionally, comparisons between the consensus sequence of the intergenic conserved sequences, the 3′ and the 5′ ends, and the genome organisation of AaNV and selective members of the *Nucleorhabdovirus* genus are shown in Supplementary Figure S2.

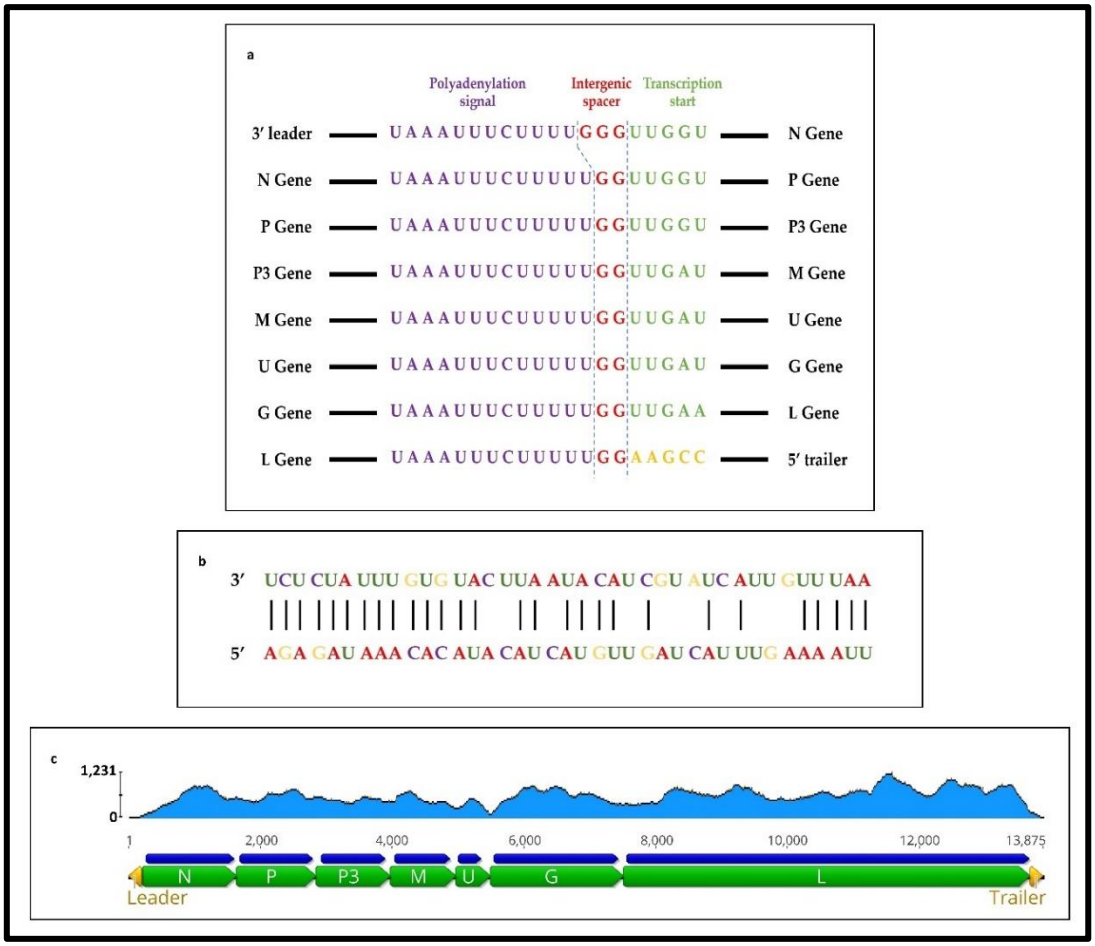

Figure 4: (**a)** The intergenic regions of the alfalfa-associated nucleorhabdovirus (AaNV) genome; the polyadenylation signal, the intergenic spacer and the transcription start site; (**b)** Alignment of ends of AaNV 3′ leader (l) and 5′ trailer (t) sequences (complementary nucleotides are indicated by vertical black lines); **c** Schematic representation of the full sequence of AaNV and its genome organisation. The open reading frames N, P, P3, M, U, G, L with their CDS are indicated as green and blue block arrows, respectively. The yellow block arrows represent the 3′ leader (l) and the 5′ trailer (t). The read map distribution is shown in light blue over the genome

2.1.4.5. Predicted protein features in silico

The AaNV protein sizes range from 113 amino acid (aa) for the U protein to 2038 aa for the L protein with molecular masses of 12.4 kDa and 234.8 kDa, respectively (Table 1). The N and G proteins have neutral isoelectric points (IEP) of 7.1 and 7.3. U and P are acidic proteins while P3, M and L are basic proteins (Table 1). Comparing the protein sequences of AaNV with those of BCaRV-1 and DYVV showed that the predicted proteins of AaNV are more closely related to those of DYVV except for the glycoprotein, which is more closely related to that of BCaRV-1 (Table 1). The aa sequences identities were between 11.5 and 35.8% compared to BCaRV-1 and between 14 and 33.7% for DYVV (Table 1). Additionally, the nuclear localisation signals (NLS, or a karyophilic domain) and the nuclear export signals (NES) of the proteins were predicted (Table 1). The highest cNLS mapper scores were for N, P and L (12.7, 10 and 10, respectively), followed by P3 and G with scores of 7 and 6. The M protein had the lowest score with 4.3 while the U ORF did not score any NLS. The cNLS scores predicted an exclusive nuclear localisation for N, P and L proteins, a partial nuclear localisation for P3 and G proteins, and a nuclear and cytoplasmic localisation for M protein (Table 1). Moreover, four of these proteins have a detectable NES (Table 1).

Table 1: Characteristics of alfalfa-associated nucleorhabdovirus (AaNV) encoded proteins [sizes in amino acids (aa), molecular masses (MW) in kilo Dalton (kDa), the isoelectric points (IEP), predicted cell nuclear localisation signals (cNLS) and nuclear export signals (NES)]

Putative gene function	Gene	Size (aa)	MW (kDa)	pairwise aa sequence identity (%)		IEP	Predicted NLS					NES position
				BCaRV-1	DYVV		Position	Partite	aa sequence	cNLS Mapper Score	Predicted location	
Nucleocapsid protein	N	443	50.3	33.5	36.6	7	408	Bi[c]	RAGIKRQAGDHETQGTKRARTS	12.7	eN	351
Phosphoprotein	P	363	41.1	11.5	14	6	146	Mono[d]	RGNKRRKRSD	10	eN	ND
Putative cell-to-cell movement protein	P3	322	37.2	14.6	22.4	9	17	Bi	PTKKRTSQDKYNFRSTESLYAEPYNKIIRTK	7	pN	ND
Matrix protein	M	277	31.4	14.9	16.5	8	213	Bi	RSSVKITGKQMRARSSSRSRSPYKVSLSSNKRTYLD	4.3	N/Cp	136
Unknown protein	U	113	12.4	NA[a]	NA	4	ND[b]	ND	ND	ND	NA	31
Glycoprotein	G	629	71.6	26.4	24.6	7	443	Bi	KSAYKKKLPYEVTAWNGDKIMSEYPYKNIVVE	6	N/Cp	ND
RNA-dependent RNA polymerase	L	2038	234.8	35.8	36.7	8	780	Bi	EKTAIKRRMRAFRDDLGQKMKKR	10	eN	88,90 and 149

The amino acid sequence identities between the putative gene products of AaNV and those of black currant-associated rhabdovirus 1 (BCaRV-1; MF543022) and datura yellow vein virus (DYVV; NC_028231). The predicted protein cell nuclear localisation signal (cNLS; http://nls-mapper.iab.keio.ac.jp/cgi-bin/NLS_Mapper_form.cgi) Mapper score and Nuclear export signals (NES) are also mentioned. The cut-off values of the cNLS mapper scores: 8, 9, or 10 = the protein is predicted to be exclusively localised to the nucleus (eN), 6, 7 or 8 = partially localised to the nucleus (pN), 3, 4, or 5 = localised to both the nucleus and the cytoplasm (N/Cp), and 1 or 2 = localised to the cytoplasm (Cp). [a]NA: not applicable. [b]ND: not detectable. [c]Bi: Predicted bipartite NLS[d] Mono: Predicted monopartite NLS

2.1.4.6. Antiserum production, serological specificity and diagnostic RT-PCR

The specificity of the AaNV antiserum was confirmed by DAS-ELISA using plant material infected with either AaNV or two other rhabdoviruses (EMDV and PhCMoV). The AaNV antibodies reacted only with plant material infected with AaNV but neither with non-inoculated nor non-infected plant material nor with plants infected with EMDV or PhCMoV (Table 1). In reciprocal tests, antisera to EMDV and PhCMoV did not react with AaNV-infected plant tissue. Diagnostic primers were designed to confirm the presence of AaNV by RT-PCR resulting in a 971 bp amplicon. The primers were specific to the partial sequence of the L-ORF of AaNV and did not amplify other nucleorhabdoviruses tested, i.e., EMDV and PhCMoV.

The specific recognition of nucleocapsids by the AaNV antibodies were demonstrated using IEM. Only nucleocapsid structures reacted with antibodies but not complete virions, see Supplementary Fig. S3a displaying enriched but undecorated nucleocapsids from the crude sap samples after the preincubation with antiserum (immunosorbent step), and Supplementary Fig. S3b showing nucleocapsids covered with antibodies after the decoration step. With antiserum against EMDV, neither enrichment nor decoration of nucleocapsids were obtained using IEM (data not shown).

2.1.4.7. Infectivity tests

In a limited host range study, the virus was successfully transmitted to *N. benthamiana*, *P. sativum*, and *V. faba*. Mechanically inoculated *N. benthamiana* plants showed systemic infection. Systemic symptoms consisted of leaf mottling, yellowing and curling at approximately 4 weeks after inoculation (Fig. 5). However, inoculated *P. sativum* and *V. faba* showed either no symptoms or a slight leaf mottling and the infection rate was low on these hosts (only 4 out of 36 *V. faba* and 1 out of 30 *P. sativum* plants). To confirm the infections, DAS-ELISA and RT-PCR were performed. Use of the AaNV antiserum in DAS-ELISA confirmed AaNV infections at high titres in *N. benthamiana* and at lower titres in *P. sativum* and *V. faba* and the absence of detectable virus from AaNV-inoculated *M. lupulina*, *M. sativa*, *T. pratense*, *T. repens* and *C. quinoa* (Table 2). Infections were also confirmed by RT-PCR. It was not possible to transmit AaNV mechanically to *M. lupulina* and *M. sativa* (21 and 18 plants tested, respectively). Plants remained symptomless and all the samples tested negative in DAS-ELISA and RT-PCR. Additionally, mechanical inoculation using fresh material from AaNV-infected *V. faba* and *P. sativum* plants failed to induce infection in *V. faba*, *P. sativum*, *M. lupulina* and *M. sativa*.

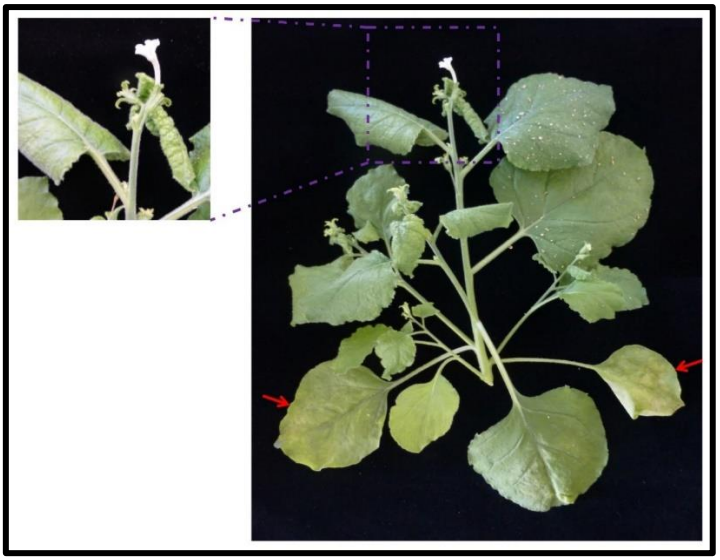

Figure 5: Photo of AaNV-infected *N. benthamiana* plant. The plant shows systemic leaf rolling, mottling, yellowing and curling, and chlorotic lesions on inoculated leaves at approximately 4 weeks post inoculation. Red arrows indicating inoculated leave

Table 2: DAS-ELISA reactions of various antisera raised against different plant nucleorhabdoviruses and limited host range study

Host species	Inoculated virus	Antisera		
		AaNV (JKI-1607)	EMDV (JKI-1073)	PhCMoV (JKI-2051)
N. benthamiana	AaNV	+++[a]	–	–
P. sativum	AaNV	+	NT[b]	NT
V. faba	AaNV	+	NT	NT
M. sativa	AaNV	–	NT	NT
M. lupulina	AaNV	–	NT	NT
T. repens	AaNV	–	NT	NT
T. pratense	AaNV	–	NT	NT
C. quinoa	AaNV	–	NT	NT
N. benthamiana	EMDV	–	+++	–
N. benthamiana	PhCMoV	–	–	+++
Buffer		–	–	–
N. benthamiana	Non-inoculated	–	–	–
P. sativum	Non-inoculated	–	NT	NT
V. faba	Non-inoculated	–	NT	NT
M. sativa	Non-inoculated	–	NT	NT

[a] Following a substrate incubation for 1 h, DAS-ELISA reactions were classed as follows: negative reaction (−): ≤ cut-off point (= OD_{A405} 0.025); weak reaction (+): cut-off point to 1.0, intermediate reaction (++): 1.0 to 2.0, strong reaction (+++): > 2.0). [b]NT = not tested

2.1.5. Discussion

Using EM and HTS technologies, the presence of a novel nucleorhabdovirus in alfalfa was established. The bacilliform appearance of the viral particles observed in infected *N. benthamiana* tissues is consistent with observations on previously described plant rhabdoviruses. Preliminary measurements indicated particle sizes within the range of the known nucleorhabdoviruses (130 to 300 nm × 45 to 100 nm in diameter (Goodin & Jackson, 2002)). The virions of AaNV had an average length of 180–200 nm and measured 85–95 nm in diameter. The observed ultra-cellular deformations of nuclei and chloroplast in epidermis and mesophyl cells are in accordance with the distorted phenotype of systemically infected *N. benthamiana* plants showing leaf rolling, mottling and yellowing.

The species demarcation criteria for the genus *Nucleorhabdovirus* state that a new species should have three characteristics (Walker *et al.*, 2018); a new species should have a minimum nucleotide divergence of 50% in cognate genes, can be clearly distinguished in serological tests or by nucleic acid hybridisation, and should occupy a different ecological niche (differences in hosts and/or vectors). The AaNV genome shares 39.8% nucleotide identity with BCaRV-1, its closest relative in the genus *Nucleorhabdovirus*. Moreover, all its ORFs have less than 40% amino acid sequence identity with their most closely related sequences in other rhabdoviruses. In addition, the AaNV antiserum reacted specifically with AaNV-infected plant tissue while antisera to EMDV and PhCMoV, two other nucleorhabdoviruses, failed to react with AaNV infected plant tissues in DAS-ELISA. Furthermore, the primers for RT-PCR are specific for AaNV. As for the third demarcation criterium, AaNV was originally identified in *Medicago sativa*, an important legume crop. However, the mode of transmission and/or potential vectors have not yet been identified. As a consequence, AaNV should be considered as a new virus species in the *Nucleorhabdovirus* genus.

As with all rhabdoviruses, the genome of AaNV has highly conserved regulatory regions (intergenic regions) separating its ORFs and complementary 3' leader and 5' trailer sequences. The intergenic regions of AaNV are closely related to those of DYVV, BCaRV-1 and SYNV (Dietzgen *et al.*, 2015; Wu *et al.*, 2018). The predicted features of AaNV proteins are similar to those of related nucleorhabdoviruses. The individual proteins of AaNV are similar in size to their homologs in DYVV and BCaRV-1. The predicted isoelectric point (IEP) of N protein of AaNV is the same as that of DYVV (Dietzgen *et al.*, 2015). Similar to DYVV, P3, M and L are basic proteins and P is an acidic protein.

The only difference is the G protein which is neutral in case of AaNV and acidic for DYVV. In addition to the six main nucleorhabdovirus proteins (N, P, P3, M, G and L), a new ORF (U) with unknown function was identified. Its predicted protein has an acidic IEP. All

the seven transcription units and the leader are predicted to be polyadenylated, but its functionality still needs to be proven.

Nucleorhabdoviruses are known to establish virus replication factories in the nuclei of infected plant cells (Jackson *et al.*, 2005). All AaNV proteins except U, display predicted mono- or bipartite nuclear localisation sequences (NLS) suggesting their independent importation into the nucleus. The presence of both the NLSs and the leucine-rich nuclear export signals (NESs) in N, M and L proteins indicates the ability of these proteins to shuttle between the nucleus and the cytoplasm through coordination of these transport signals. Although the unknown protein (U) seems to lack a NLS, the observation that it has an NES suggests its ability to be exported out of the nucleus.

AaNV was mechanically transmitted to *N. benthaminana*, *P. sativum* and *V. faba*. Although it did not show any noticeable or only slight mottling symptoms on *P. sativum* and *V. faba*, low infection rates were confirmed by DAS-ELISA. Interestingly, the virus could not be mechanically transmitted to *M. sativa* nor *M. lupulina*. It is not known if this is due to the serial passaging on *N. benthamiana*for propagation purposes and therefore a host adaption effect. The biological impact of the observed smaller sized particles of 167 nm length for mechanical transmission and host interactions awaits further investigation. As a (insect) vector has not been identified yet, it is unclear how the transmission from *M. sativa* to *M. sativa* would occur naturally or if *P. sativum* and *V. faba* crops or some weed species could act as natural alternative reservoirs for AaNV. It is also unknown if this virus still occurs naturally in alfalfa in the area it was originally found, or elsewhere in Europe. As no sequence data nor serological data are available for LEV, it is unclear whether these “historic” findings are related to AaNV.

2.1.6. Conclusions

In the present study, we describe a novel nucleorhabdovirus originating from infected *M. sativa* from Austria. Using HTS, we were able to determine the full-length sequence of this virus which was tentatively named AaNV. Since the sequence identity to BCaRV-1, its closest known relative, was only 39.8%, AaNV represents a new species according to the species demarcation criteria set by the International Committee on Taxonomy of Viruses (ICTV) (Walker *et al.*, 2018). The site of virus maturation was observed by EM in the nucleus of infected cells thus confirming the phylogenetic assignment. It was possible to transfer AaNV experimentally using mechanical inoculation to *N. benthaminana* as well as other members of the *Fabaceae* family, i.e., *P. sativum* and *V. faba*. Along with ADV and LEV, this is the third rhabdovirus and the second nucleorhabdovirus known to infect *M. sativa* in nature. However, it was not possible to transfer AaNV back to alfalfa by mechanical inoculation. Thus, further research is needed to identify natural vectors of this virus as well as other alternative host plants. The serological and molecular biological assays developed may aid larger surveys addressing these questions.

2.1.7. Supplementary

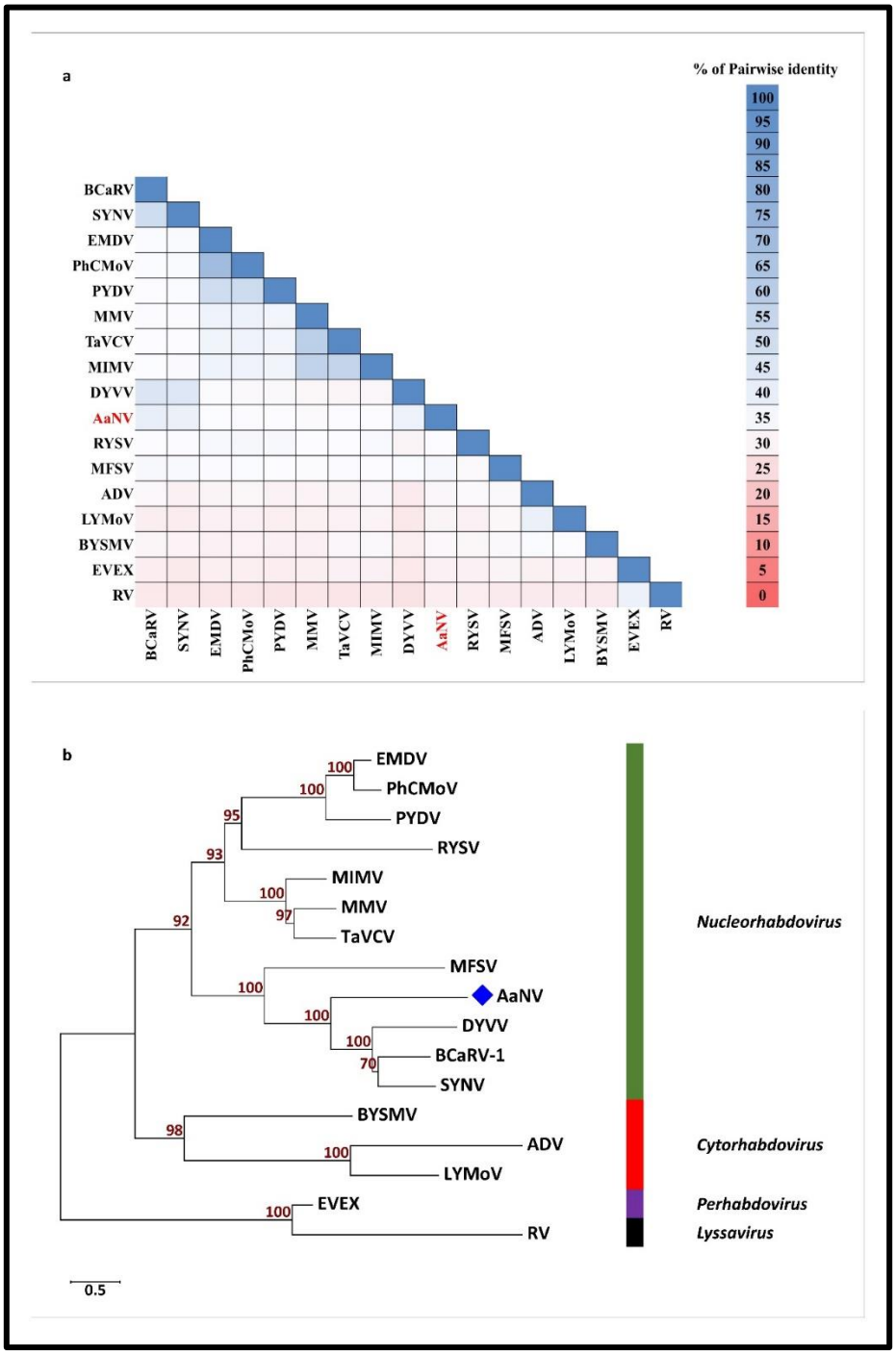

Figure S1. (**a**) Pairwise identity matrix of the whole genome sequences of AaNV with selected members of the family *Rhabdoviridae* (ClustalW 2.1); (**b**) Unrooted neighbour-joining phylogenetic tree [Genetic distance model (Jukes-Cantor) and 1,000 bootstrap replications] based on the nucleotide alignment of the whole genomes of AaNV and selected members of different genera of the family *Rhabdoviridae*. AaNV indicated by a blue solid diamond shape.

The names and the accession numbers of the viruses are as follow: *Nucleorhabdovirus* (green): alfalfa-associated nucleorhabdovirus (AaNV; MG948563), black currant-associated rhabdovirus 1 (BCaRV-1; MF543022), datura yellow vein virus (DYVV; NC_028231), eggplant mottled dwarf virus (EMDV; NC_025389), maize fine streak virus (MFSV; NC_005974), maize Iranian mosaic virus

(MIMV; NC_036390), maize mosaic virus (MMV; NC_005975), physostegia chlorotic mottle virus (PhCMoV; KY859866), potato yellow dwarf virus (PYDV; NC_016136), rice yellow stunt virus (RYSV; NC_003746), sonchus yellow net virus (SYNV; NC_001615) and taro vein chlorosis virus (TaVCV; NC_006942). *Cytorhabdovirus* (red): alfalfa dwarf virus (ADV; NC_028237), barley yellow striate mosaic virus (BYSMV; NC_028244) and lettuce yellow mottle virus (LYMoV; NC_011532). *Lyssavirus* (black): rabies virus (RV; NC_001542). *Perhabdovirus* (violet): eel virus European X (EVEX; NC_022581).

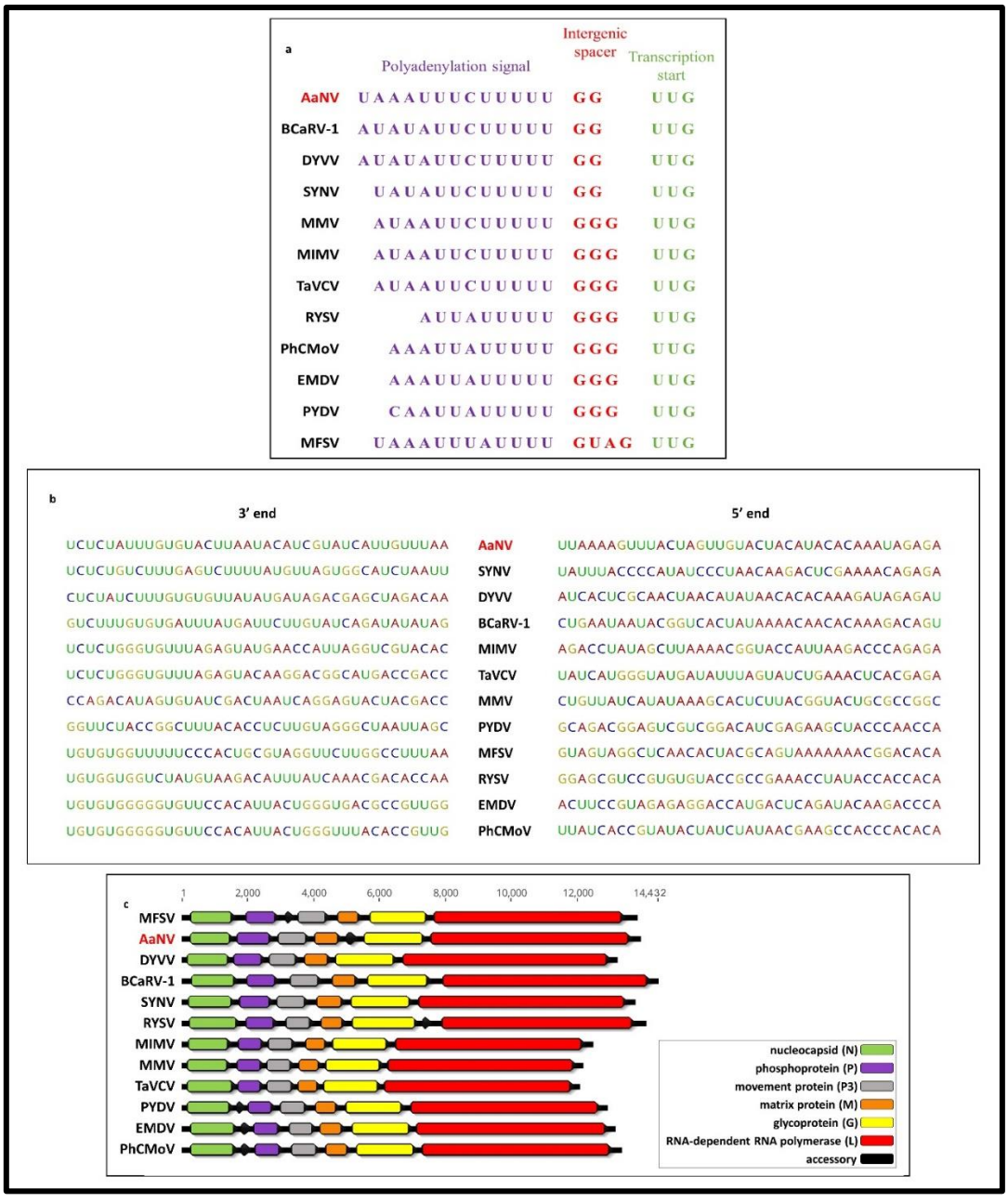

Figure S2. Comparisons between AaNV and selective members of the *Nucleorhabdovirus* genus. The consensus sequence of the intergenic conserved sequences **(a)**, the 3' and 5' ends **(b)**, and the genome organisation **(c)**. The names and the accession numbers of the viruses can be found under figure S1.

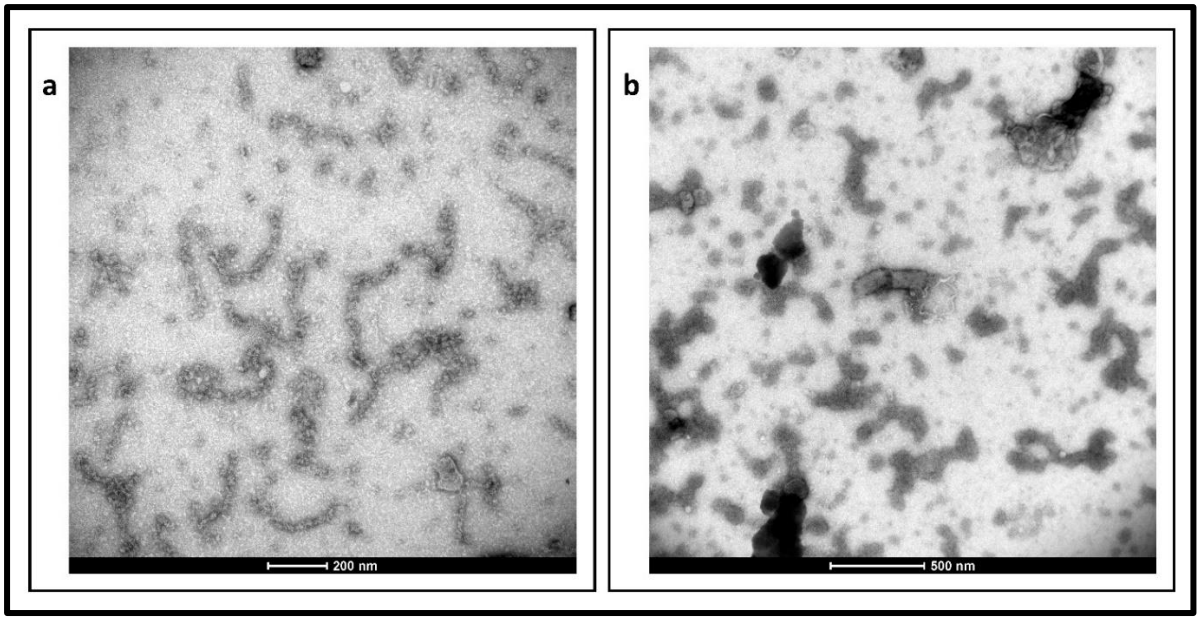

Figure S3. Electron micrograph of the JKI-1607 reacting with AaNV ribonucleoprotein (RNP). (a) Enriched nucleocapsids after immunosorbent step; (**b**) Enriched nucleocapsids but not virions are covered (decorated) with antibodies.

3.

3.1. Caraway yellows virus, a novel nepovirus from *Carum carvi*

Yahya Zakaria Abdou Gaafar, K. R. Richert-Pöggeler, A. Sieg-Müller, P. Lüddecke, K. Herz, J. Hartrick, C. Maaß, R. Ulrich and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar YZA, Richert-Pöggeler KR, Sieg-Müller A *et al.*, 2019. Caraway yellows virus, a novel nepovirus from *Carum carvi*. *Virology Journal* **16**, 529. doi: 10.1186/s12985-019-1181-1.

3.1.1. Abstract

A novel nepovirus was identified and characterised from caraway, and tentatively named caraway yellows virus (CawYV). Tubular structures with isomeric virus particles typical for nepoviruses were observed in infected tissues by electron microscopy. The whole genome of CawYV was identified by high throughput sequencing (HTS). It consists of two segments with 8026 nt for RNA1 and 6405 nt for RNA2, excluding the poly(A) tails. CawYV-RNA1 shared closest nt identity to peach rosette mosaic virus (PRMV) with 63%, while RNA2 shared 41.5% with blueberry latent spherical virus (BLSV). The amino acid sequences of the CawYV protease-polymerase (Pro-Pol) and capsid protein (CP) regions share the highest identities with those of the subgroup C nepoviruses. The Pro-Pol region shared highest aa identity with PRMV (80.1%), while the CP region shared 39.6% to soybean latent spherical virus. Phylogenetic analysis of the CawYV-Pro-Pol and -CP aa sequences provided additional evidence of their association with nepoviruses subgroup C. Based on particle morphology, genomic organization and phylogenetic analyses, we propose CawYV as a novel species within the genus *Nepovirus* subgroup C.

Keywords

Caraway; High throughput sequencing; Bipartite genome; Tubular structures; Nepovirus subgroup C

3.1.2. Main text

Viruses from the genus *Nepovirus* in the subfamily *Comovirinae* of the *Secoviridae* family possess a bipartite genome consisting of two positive ssRNAs with a 5' viral protein genome-linked (VPg) and a 3' poly(A) tail (Thompson *et al.*, 2017). The RNA1 segment encodes the helicase, protease and its cofactor, replicase and the viral protein genome linked whereas the RNA2 segment encodes the movement and coat proteins (Fuchs *et al.*, 2017; Thompson *et al.*, 2017). Nepoviruses are the only members of the family *Secoviridae* known to have a single CP (Fuchs *et al.*, 2017). Each of the two RNA segments are encapsidated separately in a non-enveloped icosahedral virion of 25-30 nm in diameter (Thompson *et al.*, 2017). Nepoviruses can be transmitted non-persistently and non-circulatively by nematodes, mite and thrips (Bergeson *et al.*, 1964; Thompson *et al.*, 2017). Seed and pollen transmissions are well-documented (Fuchs *et al.*, 2017; Thompson *et al.*, 2017). In herbaceous plants, the symptoms induced are often transient with symptom recovery being a common outcome (Fuchs *et al.*, 2017).

Caraway (*Carum carvi* L.) is an aromatic biennial plant in the *Apiaceae* family (Iacobellis *et al.*, 2005). It is native to Europe, north Africa and western Asia (Bailer *et al.*, 2001; Eddouks *et al.*, 2004). Caraway is used as a food flavour, fragrance additive, and for medical purposes as an antibacterial agent with antispasmodic, carminative, and appetite stimulant properties (Iacobellis *et al.*, 2005). In 2016, an organic caraway field showed

crop losses. A caraway plant sample with systemic yellowing was sent to Julius Kuehn-Institute (JKI) for analysis (Fig. 1a). The sample tested positive by DAS-ELISA using antiserum JKI 1283 developed against an uncharacterised nepovirus from carrot which is likely a strain of cherry leaf roll virus (CLRV) (unpublished data). The virus was mechanically transmitted to *Nicotiana benthamiana* and chlorotic local lesions were observed on inoculated leaves followed by systemic chlorosis and necrosis. Symptom recovery was not observed. The virus was provisionally named "caraway yellows virus" (CawYV).

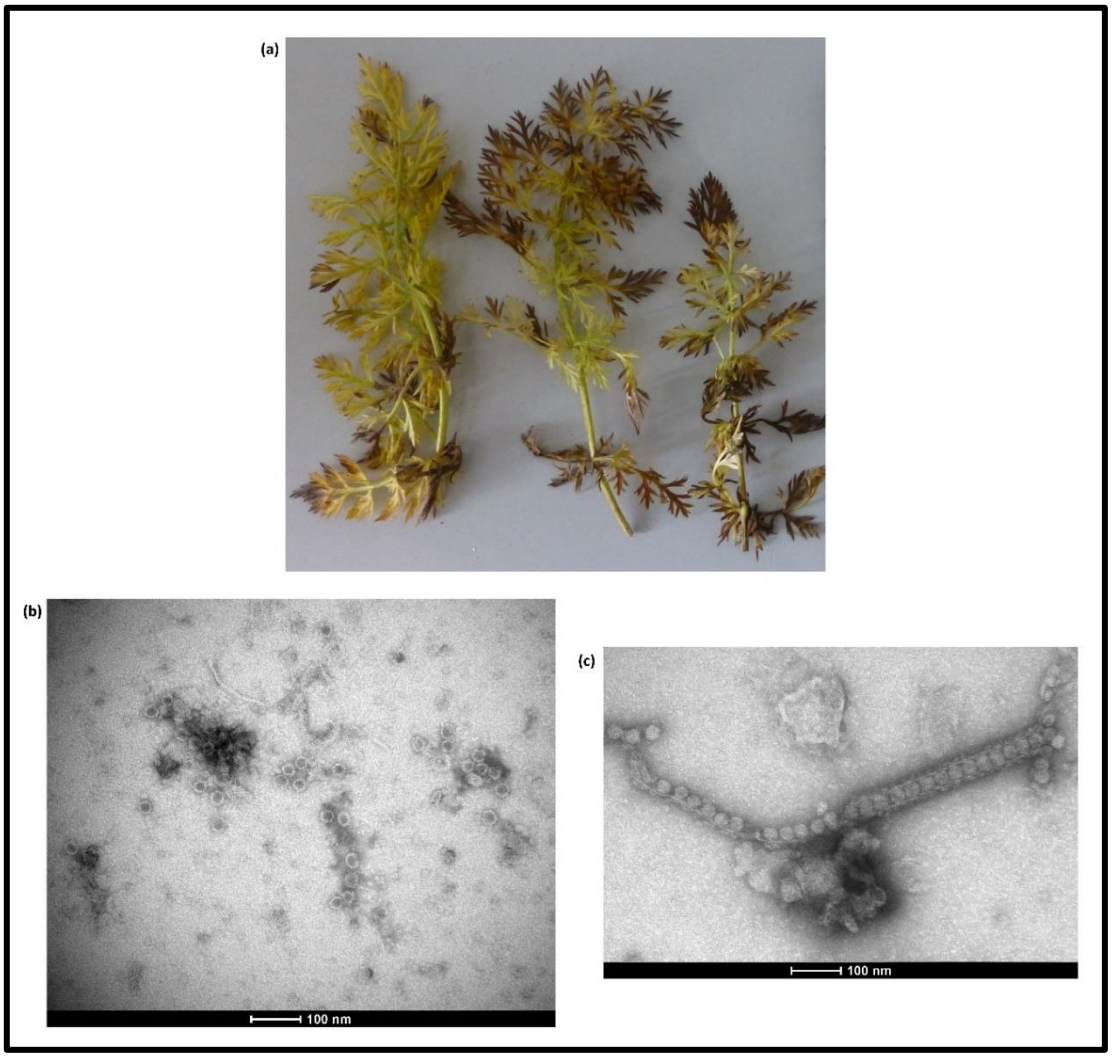

Figure 1. (**a**) Leaf symptoms of caraway yellows virus (CawYV) on caraway plants; (**b**) Electron micrograph of CawYV particles from the original infected caraway sample; (**c**) Electron micrograph showing tubular structure containing virus particles of CawYV in tissue homogenate of infected *Nicotiana benthamiana*

Electron microscopy (EM) revealed the presence of isomeric virus particles of about 30 nm in diameter in preparations made from the original infected caraway sample (Fig. 1b), indicating the presence of a nepovirus. Additionally, tubules containing virus-like particles in tissue homogenate of *N. benthamiana* infected with the nepovirus were also

observed by EM (Fig. 1c). This has also been shown for other nepoviruses e.g., grapevine fanleaf virus, where the movement and the capsid proteins act as components of tubular structures (required for cell to cell movement) that traverse the cell wall with the virus particles (Laporte *et al.*, 2003; Thompson *et al.*, 2017).

To obtain the full genome of CawYV, double stranded RNA (dsRNA) was extracted from infected *N. benthamiana* using Double-RNA Viral dsRNA Extraction Mini Kit for Plant Tissue (iNtRON) following the manufacturer's instructions. The extracted dsRNA was sent for library preparation and high throughput sequencing (HTS) at Eurofins GATC Biotech GmbH. The dsRNA was fragmented, strand specific cDNA was synthesized using random primers (the dsRNA was denatured at 99 °C for 2 min), followed by adapter ligation and adapter specific PCR amplification then sequencing on Illumina NovaSeq 6000 platform (2 × 150).

Using Geneious Prime (v. 2019.0.4), the raw reads (15,468,416) were quality trimmed, filtered, normalized, and error corrected followed by *de novo* assembly using Geneious assembler (Medium sensitivity/Fast setting). 36,634 contigs of lengths between 100 and 23,141 nt were generated. A BLASTn search of the contigs against a local database for viruses and viroids downloaded from NCBI showed that two contigs of 7180 and 6341 nt had 72% identity (73% coverage and zero E-value) to peach rosette mosaic virus (PRMV) and 79% (16% coverage and 1e-90 E-value) to blueberry latent spherical virus (BLSV), respectively. The 5' ends of both RNA segments were confirmed using RNA ligase-mediated amplification of cDNA ends (RLM-RACE) (Coutts & Livieratos, 2003). The 3' ends of the two RNA segments were determined by using an oligo(d)T primer for cDNA synthesis followed by PCR using virus specific primers and the oligo(d)T primer. The primers used for the 5' and 3' ends confirmation are listed in Supplementary Table S1. The PCR products were cloned, sequenced and the resulting sequences were assembled using the map to reference tool and the original assembled contigs as references. 72,977 of the quality trimmed reads were assembled to CawYV complete genome. The assembled genome of CawYV was 8026 nt for RNA segment 1 and 6405 nt for RNA segment 2 (excluding poly(A) tails). The sequences were deposited in the GenBank database under accessions MK492273 and MK492274. For diagnostic purposes and to confirm the presence of CawYV in symptomatic leaf tissue, a primer pair was designed using Primer 3 tool in Geneious (HZ-636 5' TGA AGA TCC GGG AAA GGC AC 3' and HZ-637 5' ACG CTT TCC ACT CTC ACC TG 3') (Untergasser *et al.*, 2012). The presence of CawYV was confirmed in the infected plants by RT-PCR using OneTaq One-Step RT-PCR Kit (NEB) resulting in amplicons of 481 bp (data not shown).

Further analyses of the CawYV sequence confirmed its identity as a nepovirus. In analogy to other nepoviruses, CawYV RNA1 contains an open reading frame (nt position 92 to 6733) encoding a polyprotein of 2213 aa in length. Pairwise comparisons of nt and aa sequences of this ORF to its homologues of the other nepoviruses were performed using ClustalW (Larkin *et al.*, 2007). The results show that the highest similarity was shared with PRMV at 66.1% on nt and 68.1% aa levels, respectively (Table 1). By searching for the different nepovirus motifs using the motif searching tool in Geneious, the locations of the

putative protease cofactor (Pro-cof), the NTP-binding helicase domain (Hel), the serine protease domain (Pro), and the RNA-dependent RNA polymerase (RdRp) core domain were found in the RNA1-encoded polyprotein (Isogai *et al.*, 2012). The putative viral protease cofactor motif ($FX_{27}WX_{11}LX_{21}LXE$) was located at aa residues 438–502. The typical NTP-binding helicase motif A (GX_4GKS), motif B (D), and motif C (N) were found at aa 752-759, 803, and 852, respectively. A serine protease motif was found at aa 1280-1449 ($HX_{40}EX_{106}$**S**$GX_8GX_5GXHX_2G$) of the CawYV RNA1 polyprotein sequence (Fig. 2a). The serine at this position is unusual for members of the *Picornavirales*(where cysteine is usually encoded) but was described for some members of genus *Nepovirus* subgroup C i.e., BLSV, CLRV, PRMV and soybean latent spherical virus (SLSV). The RNA-dependent RNA polymerase (RdRp) core domain was located at aa 1774-1880 ($DX_4DX_{56}GX_3TX_3NX_{33}GDD$). Pairwise analysis of the protease-polymerase (Pro-Pol) region aa sequences showed a closest identity to PRMV Pro-Pol with 80.1% (Table 1).

Table 1. Characteristics and pairwise nucleotide (nt) and amino acid (aa) alignments of the different regions of caraway yellows virus (CawYV) and selected members of subgroup C of the genus *Nepovirus* i.e., blueberry latent spherical virus (BLSV), blackcurrant reversion virus (BCRV), cherry leaf roll virus (CLRV), grapevine Bulgarian latent virus (GBLV), peach rosette mosaic virus (PRMV), soybean latent spherical virus (SLSV) and tomato ringspot virus (ToRSV)

Virus			CawYV	PRMV	BLSV	SLSV	BCRV	GBLV	ToRSV	CLRV
nt length	RNA 1	Accession no.	MK492273	NC_034214	NC_038764	NC_032270	NC_003509	NC_015492	NC_003840	NC_015414
		Complete -poly(A)	8026	8014	7960	8170	7711	7452	8214	7918
		5'UTR	91	41	61	13	66	87	77	11
		ORF	6642	6504	6519	6588	6285	6288	6594	6339
		3'UTR	1293	1469	1380	1569	1360	1077	1543	1568
	RNA 2	Accession no.	MK492274	NC_034215	NC_038763	NC_032271	NC_003502	NC_015493	NC_003839	NC_015415
		Complete -poly(A)	6405	5956	6344	5776	6405	5821	7271	6360
		5'UTR	94	47	55	23	161	189	75	11
		ORF	5022	4425	4896	4197	4881	4500	5649	4770
		3'UTR	1289	1484	1393	1556	1363	1132	1547	1579
Pairwise identity %										
nt %	RNA 1	Complete -poly(A)		63	60.8	49.6	38.5	37	36	33.8
		5'UTR		56.1	55.7	30.8	40.9	31.6	52	63.6
		ORF		66.1	62.4	53.5	40.4	39.2	36.7	36.3
		3'UTR		51.4	53.8	34	40.5	36.5	27	30.9
	RNA 2	Complete -poly(A)		41.3	41.5	37.7	35.9	30.9	38.6	35.6
		5'UTR		37.2	30.9	34.8	35.9	36.2	41.3	63.6
		ORF		39.9	38.5	39	37	31.8	40.9	37.7
		3'UTR		51	55.2	33.4	36.1	37.1	29.4	30.6
aa %	RNA 1	ORF		68.1	62.5	48.1	24.5	22.8	23.1	22.6
		X1		27.9	29.7	22.9	12.2	10.6	10.8	10.8
		X2 Pro-cof		52.9	53.2	33.9	25.8	21.4	15.6	18.5
		Hel		82.5	75.6	59.3	28	24.1	22.6	23.2
		VPg		75	57.6	56.2	6.1	27.3	25.9	37.9
		Pro		79	66.8	51.3	27.6	23.5	24.7	23.1
		RdRp		75.5	68.6	53.9	33.8	33.6	33.7	30.9
		Pro-Pol		80.1	70.2	54.9	7.1	35.3	36.5	34.6
	RNA 2	ORF		22.9	20.1	19	13.1	12.2	22.3	19
		HP		23.3	13.1	6.5	7.3	7.6	8.2	7
		MP		10	8.4	7.5	8.5	4.1	52.2	54.6
		CP		36.5	34.7	39.6	24.3	24.8	26.1	20

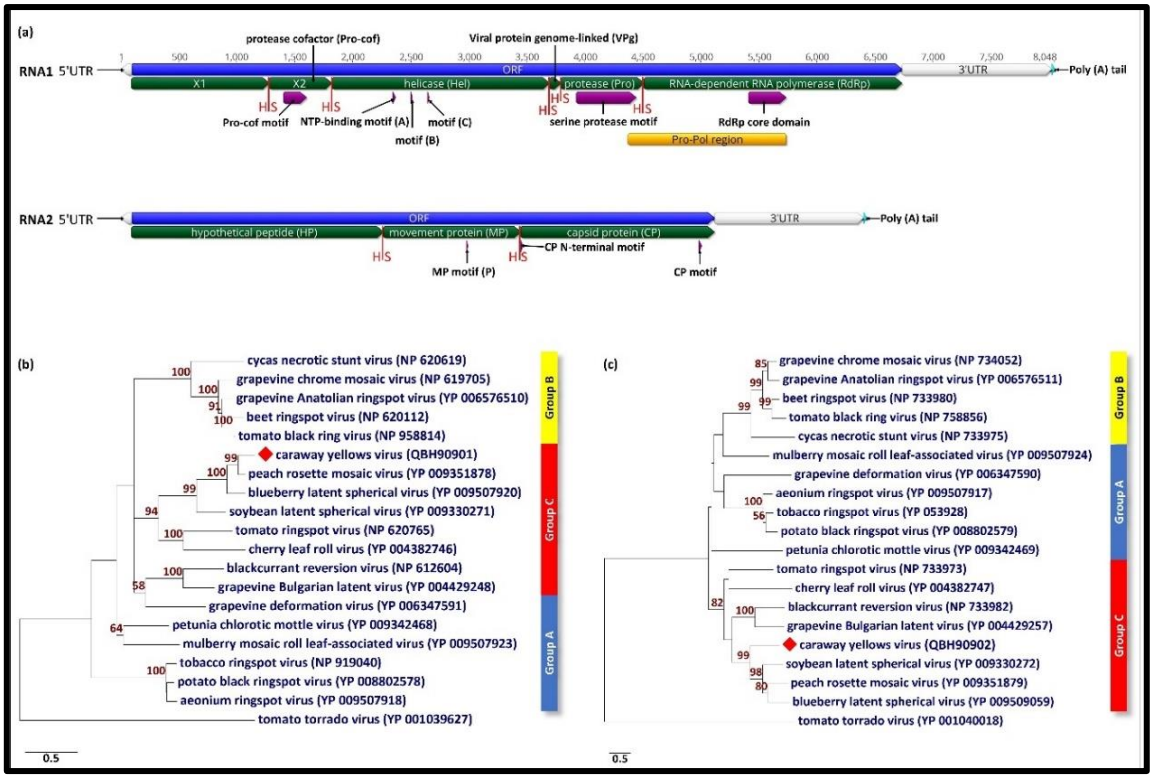

Figure 2. (**a**) Genome organization of CawYV-RNA1 and -RNA2. Each of RNAs 1 and 2 contain a single large open reading frame (in blue). The predicted putative peptides are shown in green, separated by the predicted cleavage site (H/S) (red). The RNA1 polyprotein contains X1 protein, protease cofactor (Pro-cof/X2), helicase (Hel), viral protein genome-linked (VPg), protease (Pro) and RNA-dependent RNA polymerase (RdRP). RNA2 encodes for a polyprotein with hypothetical protein (HP), movement protein (MP) and capsid protein (CP). The conserved nepovirus sequences (domains and motifs) are shown in violet. The protease-polymerase (Pro-Pol) region of CawYV-RNA1 starts with the serine (S) of the protease motif and ends with the (GDD) motif of the polymerase (shown in gold). (**b**) Maximum-likelihood (ML) phylogenetic trees showing the relationships between CawYV and members of the genus *Nepovirus* based on aa alignments of the Pro-Pol region and (**c**) the capsid protein (CP) region. Numbers on branches indicate the bootstrap percentages (1000 replicates, ≥50% are shown). Tomato torrado virus (genus *Torradovirus*, family *Secoviridae*) is an outgroup

CawYV-RNA2 contains an open reading frame (nt position 95 to 5116) encoding a polyprotein of 1673 aa in length. Pairwise comparisons of RNA2-ORF nt and aa sequences to the homologues of other nepoviruses showed the highest similarity with PRMV with 31% nt and 21.3% aa identities (Table 1). The conserved movement protein motif (P) was found at aa position 962 (Mushegian, 1994). The CP N-terminal five amino acid residues (SGLEE) together with an alternate capsid protein (CP) motif (FXFYGWS) were located at aa positions 1119–1122 and 1631–1637 (Le Gall *et al.*, 1995; Isogai *et al.*, 2012). Pairwise analysis of the CP region showed that it shares highest aa identity to SLSV (39.6%, Table 1).

Each of the CawYV polypeptides is predicted to be proteolytically cleaved into putative peptides by the virus-encoded protease. Sequence alignment of all nepovirus ORF

aa sequences suggest a putative proteolytic cleavage sites at dipeptides (H/S). This potential cleavage site was not identified before in the *Secoviridae* members. The conserved histidine in the substrate-binding pocket of the protease is known for members of the subgroup C, however the known cleavage sites are Q/G, Q/S or D/S (confirmed experimentally) (Thompson *et al.*, 2017). The H/S dipeptide is also found in SLSV, BLSV and PRMV. Although the VPg motif was not confirmed in the polyprotein of RNA1, the location of the putative VPg domain could be determined by the H/S dipeptides between the NTP-binding helicase and the protease using sequence alignment (Fig. 2a). Additionally, the X1 putative protein was identified at the N terminus of RNA1 polyprotein by the presence of a H/S dipeptide potential cleavage site before the protease cofactor motif of X2 (Fig. 2a). The 5' untranslated regions (UTR) of the two RNAs were 91 nt for RNA1 and 94 nt for RNA2 and shared 61.3% nt identity to each other. The 3'UTRs (1293 and 1289 nt for RNA 1 and 2 respectively, excluding the poly(A) tail) are 98.4% identical.

A maximum likelihood tree using MEGA7 software (v 7.0.26) based on the aa alignments of the Pro-Pol and the CP regions were additional evidence for the relatedness of CawYV to the *Nepovirus* subgroup C (Fig. 2b and c) (Kumar *et al.*, 2016).

The International Committee on Taxonomy of Viruses (ICTV) suggests the following criteria for species demarcation (Thompson *et al.*, 2017): distinct host range; distinct vector specificity; absence of cross-protection; differences in antigenic reactions; absence of reassortment between RNA1 and RNA2; Pro-Pol region aa < 80% and CP region aa < 75% identities. Although the host range was not studied, the closest relatives of CawYV, i.e., PRMV and BLSV, are not known to infect members of the *Apiaceae* family. The serological cross-reactivity is well known for members of the same genus in the family *Secoviridae (Thompson et al., 2017)*. This might explain why our antiserum raised against an uncharacterised strain of CLRV reacted with CawYV. Further investigations are necessary to test the antiserum against other nepoviruses, and attempts are currently underway to develop a CawYV-specific antiserum. When compared to the closest relative PRMV, the Pro-Pol region of CawYV is slightly above the species demarcation limit by 0.1%. However, this was also observed for other nepoviruses e.g., beet ringspot virus (BRSV) and tomato black ring virus (TBRV) that share 89% aa identity at the Pro-Pol but are yet classified as distinct species (Thompson *et al.*, 2017). However, the caraway virus-CP region is very different to other nepoviruses sharing only 39.6% aa identity with SLSV. Based on these results, we propose the assignment of CawYV as a new virus species within the subgroup C of the genus *Nepovirus*. Further studies are needed to investigate the natural mode of transmission and the biological characteristics of CawYV.

3.1.3. Supplementary

Table S1: List of the primers used for caraway yellows virus 5' and 3' ends confirmation.

Location	Virus Specific Primers		
	Name	Sequence	nt position
RNA1-5' end	HZ-648	5' GCT TGT TTA GTA GCG GCT GC 3'	504-485
	HZ-649	5' GCA ATC TGC AAA TAT CGT GGC T 3'	324-303
RNA2-5' end	HZ-644	5' CAA TGC CCA CAA GCT TAG CG 3'	464-445
	HZ-645	5' ACT TTG TCA TAG CGC TCG GC 3'	321-302
Adaptor	HZ-481	5' PO4-GAT CCA CTA GTT CTA GAG CGG C-AminoC3 3'	NA
Adaptor complement	HZ-482	5' GCC GCT CTA GAA CTA GTG GAT C 3'	NA
RNA1-3' end	HZ-670	5' GGG AGA CAT AGC ACC TCT TCT 3'	6625-6645
	HZ-671	5' GAC ATG TCT CCA GAC CTA TTT TCT 3'	6666-6689
RNA2-3' end	HZ-672	5' ACC CCA GCA GCT TTC ACT AC 3'	5024-5043
	HZ-673	5' CTA AGC CGA GAG AGG AAC GC 3'	5073-5092
Poly(T)18	HZ-012	5' CCT CGG GCA GTC CTT TTT TTT TTT TTT TTT T 3'	NA

NA: not applicable.

3.2. A divergent strain of melon chlorotic spot virus isolated from black medic (*Medicago lupulina*) in Austria

Yahya Zakaria Abdou Gaafar, K. R. Richert-Pöggeler, A. Sieg-Müller, P. Lüddecke, K. Herz, J. Hartrick, Y. Seide, H.-Josef Vetten and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar YZA, Richert-Pöggeler KR, Sieg-Müller A *et al.*, 2019b. A divergent strain of melon chlorotic spot virus isolated from black medic (*Medicago lupulina*) in Austria. *Virology Journal* **16**, 297. doi: 10.1186/s12985-019-1195-8.

3.2.1. Abstract

A tenuivirus, referred to here as JKI 29327, was isolated from a black medic (*Medicago lupulina*) plant collected in Austria. The virus was mechanically transmitted to *Nicotiana benthamiana*, *M. lupulina*, *M. sativa*, *Pisum sativum* and *Vicia faba*. The complete genome was determined by high throughput sequencing. The genome of JKI 29327 consists of eight RNA segments closely related to those of melon chlorotic spot virus (MeCSV) isolate E11–018 from France. Since segments RNA 7 and 8 of JKI 29327 are shorter, its genome is slightly smaller (by 247 nts) than that of E11–018. Pairwise comparisons between the predicted virus proteins of JKI 29327 and their homologues in E11–018 showed aa identities ranging from 80.6 to 97.2%. Plants infected with E11–081 gave intermediate DAS-ELISA reactions with polyclonal antibodies to JKI 29327. Since JKI 29327 and E11–018 appear to be closely related both serologically and genetically, we propose to regard JKI 29327 as the black medic strain of MeCSV. To our knowledge, JKI 29327 represents the second tenuivirus identified from a dicotyledonous plant. Serological and molecular diagnostic methods were developed for future detection.

Keywords

High throughput sequencing; Melon chlorotic spot virus; Segmented virus; *Medicago sativa*; *Pisum sativum*; *Vicia faba*

3.2.2. Main text

Members of the genus *Tenuivirus*, family *Phenuiviridae*, are plant viruses that possess non-enveloped filamentous particles and a genome consisting of four to eight single-stranded RNA segments with negative or ambisense polarity. The thin filamentous particles consist of ribonucleoprotein (RNP) complexes, measuring 3–10 nm in diameter and with lengths proportional to the sizes of the RNAs they contain. Based on the RNA sizes, the particles may appear as small, large or even branched circles (Shirako *et al.*, 2012; Lecoq *et al.*, 2019). Tenuivirus RNAs are neither capped at their 5′ end nor polyadenylated at the 3′ end. The nucleotide sequences of the 5′ and 3′ ends of each segment are complementary (Shirako *et al.*, 2012). Tenuiviruses are known to be transmitted by planthoppers or by mechanical means albeit with difficulty (Shirako *et al.*, 2012). According to the International Committee on Taxonomy of Viruses (ICTV), seven virus species are currently assigned to the genus *Tenuivirus*: *Echinochloa hoja blanca virus* (EHBV), *Iranian wheat stripe virus* (IWSV), *Maize stripe virus* (MSpV), *Rice grassy stunt virus* (RGSV), *Rice hoja blanca virus* (RHBV), *Rice stripe virus* (RSV) and *Urochloa hoja blanca virus* (UHBV). In addition, three more species have been proposed and are pending recognition by ICTV: melon chlorotic spot virus (MeCSV), Ramu stunt virus (RmSV) and wheat yellow head virus (WYHV) (Seifers *et al.*, 2005; Mollov *et al.*, 2016; Lecoq *et al.*, 2019). The natural host range of tenuiviruses is typically restricted to monocots of the *Poaceae* family causing yield losses in important food crops such as rice (*Oryza sativa* L.)

and maize (*Zea mays* L.) (Falk & Tsai, 1998). The recent identification of MeCSV from melon (*Cucumis melo*) in France represents the first report of a tenuivirus naturally infecting a dicotyledonous plant (Lecoq *et al.*, 2019).

In 2011, a black medic (*Medicago lupulina* L.) plant showing virus-like symptoms was collected in Stadl-Paura, Austria, but the symptoms were not recorded at the time. The sample was sent to Julius Kuehn Institute for analysis. Electron microscopy revealed the presence of RNP that appeared to resemble disassembled rhabdovirus particles (Iseni *et al.*, 1998; Gaafar *et al.*, 2019d). However, polyclonal antibodies JKI-1607 raised against alfalfa-associated nucleorhabdovirus (AaNV) (Gaafar *et al.*, 2019d) failed to react with this virus in DAS-ELISA. The virus was transmitted mechanically as described in (Gaafar *et al.*, 2019d) to *Nicotiana benthamiana*, *M. lupulina*, *M. sativa*, *Pisum sativum* and *Vicia faba*, and was referred to as JKI 29327. Three weeks post inoculation, *N. benthamiana* plants showed systemic mottling, slight vein clearing and top leaf curling, whilst *M. lupulina* and *M. sativa* plants showed systemic vein clearing. *P. sativum* plants showed systemic vein clearing and severe yellowing and *Vicia faba* showed systemic mottling, yellowing and leaf rolling (Fig. 1). The virus particles of JKI 29327 were partially purified from infected *N. benthamiana* and used for antiserum production as described before (Gaafar *et al.*, 2019d). The antibodies (JKI-1608) were used for DAS-ELISA analysis of sap-inoculated plants and confirmed infection of symptomatic plants.

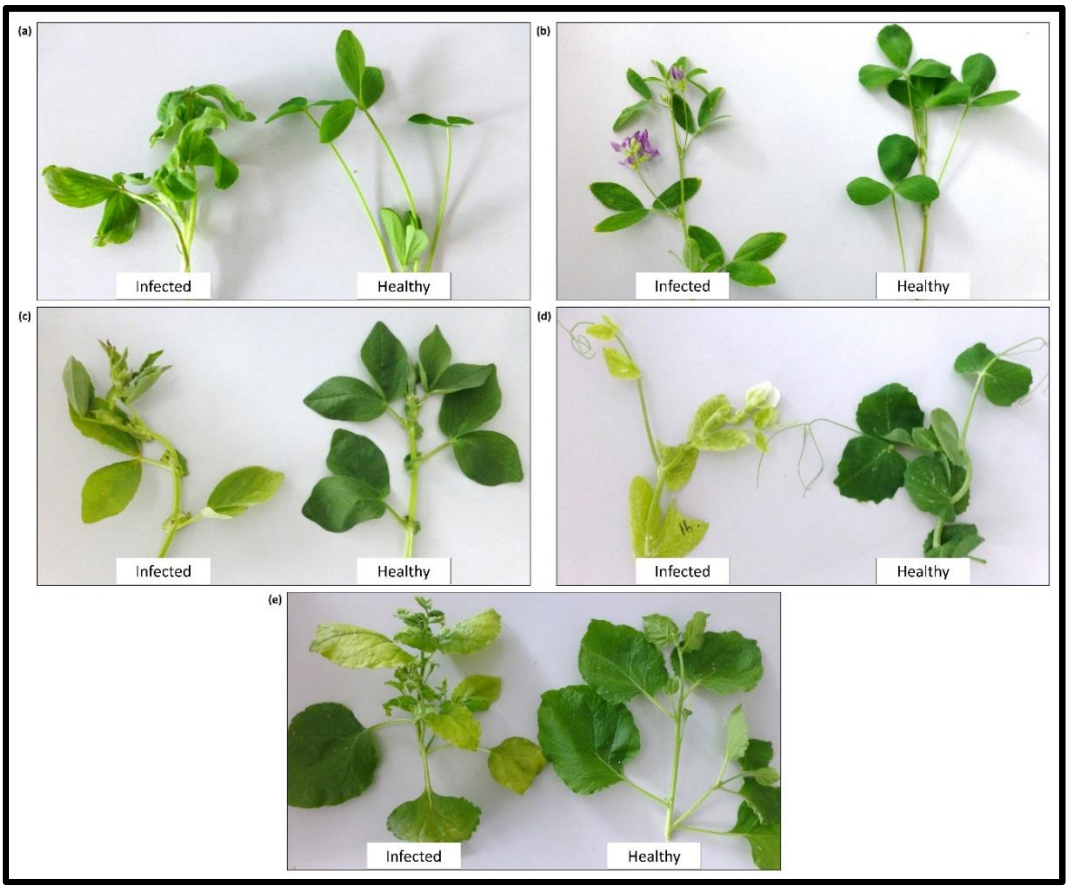

Figure 1. Plants infected with the black medic tenuivirus (JKI 29327): (**a**) *Medicago lupulina*, (**b**) *M. sativa*, (**c**) *Vicia faba*, (**d**) *Pisum sativum* and (**e**) *Nicotiana benthamiana*

For determination of the complete genome sequence of JKI 29327, total RNA was extracted from infected *N. benthamiana* using innuPREP RNA Mini Kit (Analytik Jena AG) followed by ribosomal RNA depletion using the RiboMinus Plant kit (Invitrogen). The ribo-depleted RNA was used for high throughput sequencing (HTS) on a MiSeq (v3) platform (2 × 301) as described before (Gaafar *et al.*, 2019d). A total of 2,056,956 reads were obtained. The raw reads were quality trimmed, and size filtered using Geneious Prime (v. 2019.0.3) (Biomatters Limited). The reads were then de novo assembled using Geneious assembler. A total of 53,651 contigs were generated and used for Blastn and Blastx search using virus/viroid databases on NCBI. Fifty-eight contigs shared nucleotide (nt) sequence identities (from 73.5 to 90.6%) and amino acid (aa) sequence identities from 63.8 to 97.2% to MeCSV. No other virus sequences were detected. The reference sequences of MeCSV (NC_040448 to NC_040455) were used to map the black medic tenuivirus sequences. The complete genome sequence of JKI 29327 (containing eight segments (Fig. 2a)) was assembled (19,805 nt; accession nos. MK450511 to MK450518) but segment RNA7 and RNA8 were 94 nt and 177 nt shorter than the genome of the isolate E11–018 of MeCSV. Analysis of each segment showed the presence of conserved nt sequences which can also be observed in other tenuiviruses (ACA CAA AGU C at the 5' end with its complementary sequence UGU GUU UCA G at the 3' end). Eight primers pairs were designed using Primer 3 (2.3.7) tool in Geneious (Table 1) to confirm the physical presence of all eight viral segments using RT-PCR (OneTaq One-Step RT-PCR Kit; NEB) (Untergasser *et al.*, 2012) on fresh RNA extracts from *N. benthamiana*. The amplicons were gel-purified using Zymoclean Gel DNA Recovery Kit (Zymo Research) and Sanger sequenced; sequence analyses of these amplicons showed that they were 100% identical to the corresponding segment sequences obtained by the HTS analysis and thus confirmed the presence of each individual viral segment.

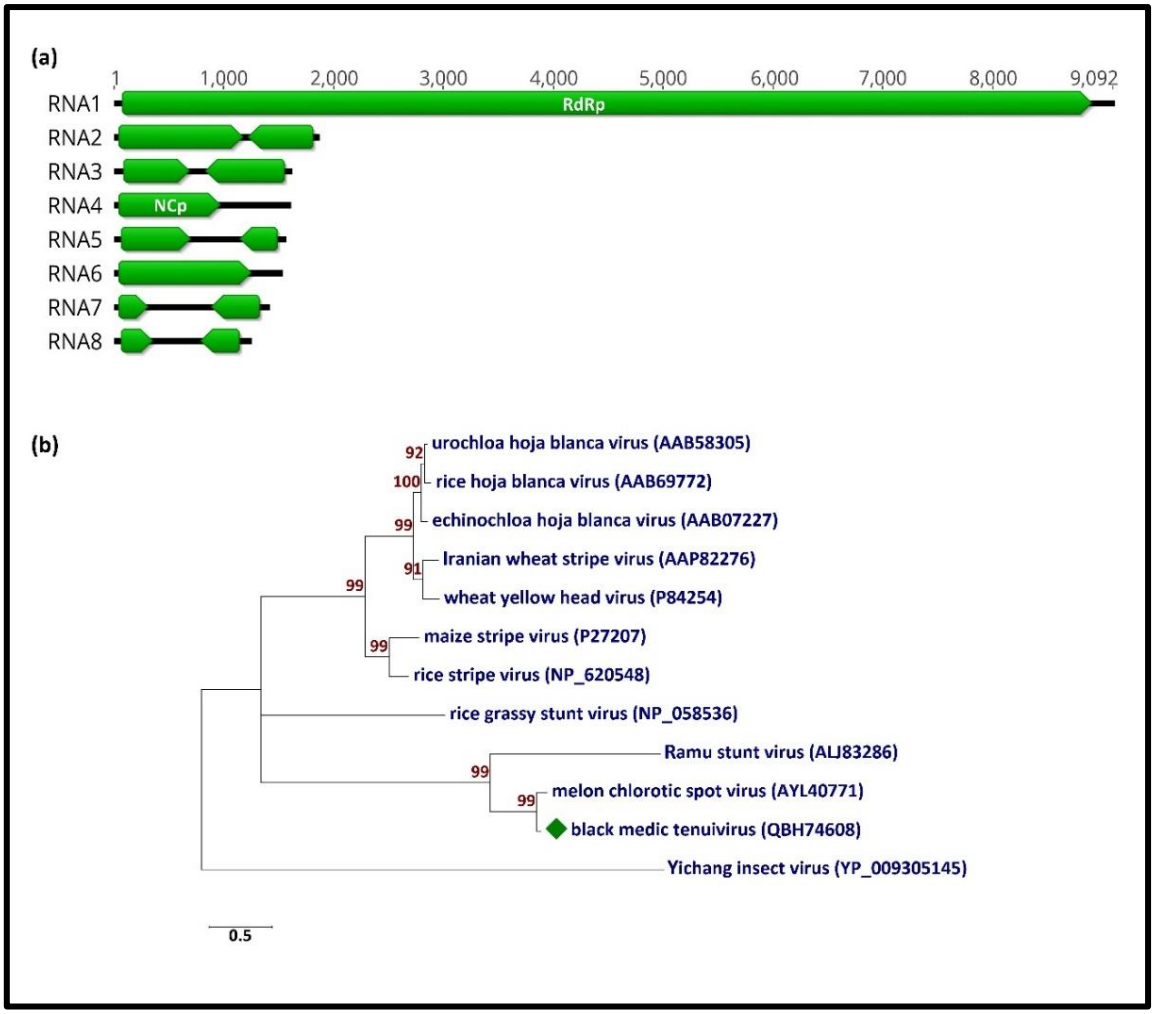

Figure 2. (**a**) Graphical representation of the genome of the black medic tenuivirus isolate JKI 29327. (**b**) Maximum-likelihood (ML) phylogenetic tree (using Jones-Taylor-Thornton (JTT) model) based on the amino acid sequence alignments of the nucleocapsid proteins (NCp) of JKI 29327 and members of the *Tenuivirus* genus. The GenBank accession nos. are in brackets. Yichang insect virus (genus *Goukovirus*) was used as an outgroup sequence. Numbers on branches indicate the bootstrap percentages (1000 replicates, only values ≥50% are shown) and the scale bar represents a genetic distance of 0.5

Table 1. The genome characteristics of melon chlorotic spot virus (MeCSV) isolate (JKI 29327) from Austria; nt and aa sequence identities of the 8 RNA segments compared to the respective homologous regions in the genome of MeCSV isolate E11–018 from France and the list of primers used for segment identification

Genome segments					Intergenic region (IR)		Predicted proteins				
Type	Length (nt)	Percent nt identity to MeCSV E11-018	PCR primers used for JKI 29327		Length (nt)	Percent nt identity to MeCSV E11-018	ORF	Putative functions	Length (aa)	Size (kDa)	Percent aa identity to MeCSV E11-018
			Name	Sequence							
RNA1	9092	82.4	HZ-603	5' ACA GAA GTG GAA TGG GCT GG 3'	NA	NA	ORF1	RNA-dependent RNA polymerase	2940	340	92.1
			HZ-604	5' GCA ACA CCC TCA TCA CTC CA 3'							
RNA2	1847	84.6	HZ-605	5' AGC TCA GTA ACC GGA ACT GC 3'	50	100	ORF2a	no match	373	43.6	87.7
			HZ-606	5' CGC AAT AGC AGG GTC CAG AT 3'			ORF2b	no match	196	43.7	87.3
RNA3	1598	85.8	HZ-607	5' TGG TGC CAG AAG GAA AGG AC 3'	157	72.6	ORF3a	no match	195	23.2	93.3
			HZ-608	5' GGC AAT GCC TCA CAA TCG TC 3'			ORF3b	no match	235	27.6	89.8
RNA4	1591	78.3	HZ-609	5' AAG TAA GGG CAG GCT GAA CC 3'	NA	NA	ORF4	nucleocapsid protein	305	33.6	88.2
			HZ-610	5' AGG CTT TCT GCT AAG TGG GC 3'							
RNA5	1547	81.7	HZ-611	5' GAA CTG TAC CGC TGA TGG GT 3'	446	68.3	ORF5a	no match	210	23.2	93.3
			HZ-612	5' CTT TGG TCT GGA GCT GTG CT 3'			ORF5b	no match	114	13	95.6
RNA6	1509	84.7	HZ-613	5' CGC ATC CTG AAT CCC ATC TCT 3'	NA	NA	ORF6	no match	400	45.5	91.2
			HZ-614	5' GCT GGC ATC ACT AGA CGG AT 3'							
RNA7	1392	73.2	HZ-615	5' ATC AGG TGT TAG CTG GCC AC 3'	586	55.6	ORF7a	no match	84	9.8	90.5
			HZ-616	5' TAA CCA CCT TCC CTG CTG TG 3'			ORF7b	no match	143	15.9	80.6
RNA8	1229	68.7	HZ-617	5' ACC CTA AGT GGA TCC GAG GT 3'	446	58.3	ORF8a	no match	93	10.9	94.8
			HZ-618	5' AGT TCC AAG TTG CCC TGC TT 3'			ORF8b	no match	115	12.9	97.2

To predict the open reading frames' (ORF) functions, the translation of each ORF was used to search for conserved domains on NCBI's conserved domain database (CDD v 3.16) (Marchler-Bauer *et al.*, 2017). Only two ORFs matched with entries in the database, i.e., RNA1-ORF1 with *Bunyavirus* RNA-dependent RNA polymerase (accession no.: cl20265) and RNA4-ORF1 with *Tenuivirus*/*Phlebovirus* nucleocapsid protein (accession no.: cl05345) (Table 1). Pairwise alignments for the different regions of each segment of JKI 29327 were performed with their homologous sequences of MeCSV using CLUSTALW (Table 1) (Larkin *et al.*, 2007). The genome components of JKI 29327 shared nt identities ranging from 68.7 to 85.8% with those of the MeCSV isolate E11–0188 (Table 1). The proteins potentially encoded by JKI 29327 and E11–018 shared aa sequence identities ranging from 80.6 to 97.2% (Table 1). A maximum-likelihood (ML) phylogenetic tree was generated using MEGA7 (7.0.26) (Jones-Taylor-Thornton (JTT) model) (Kumar *et al.*, 2016) for comparing the aa sequence of nucleocapsid proteins of JKI 29327 and other tenuiviruses. This showed a tight clustering of JKI 29327 with the MeCSV nucleoprotein (Fig. 2b). Additionally, RNA segments 7 and 8 respectively have shorter intergenic regions (IR) (586 and 446 nt) compared with those (680 and 623 nt) on the homologous RNAs of E11–018. The nt sequence identities between these IR regions of RNA7 and RNA8 are 55.6 and 58.3%, respectively. The results indicate that JKI 29327 is closely related to but distinct from the MeCSV isolate E11–018.

For additional confirmation, purified RNP preparations of JKI 29327 were again examined by electron microscopy and shown to contain tenuivirus-like circular filamentous particles representing the individual genome segments (Fig. 3).

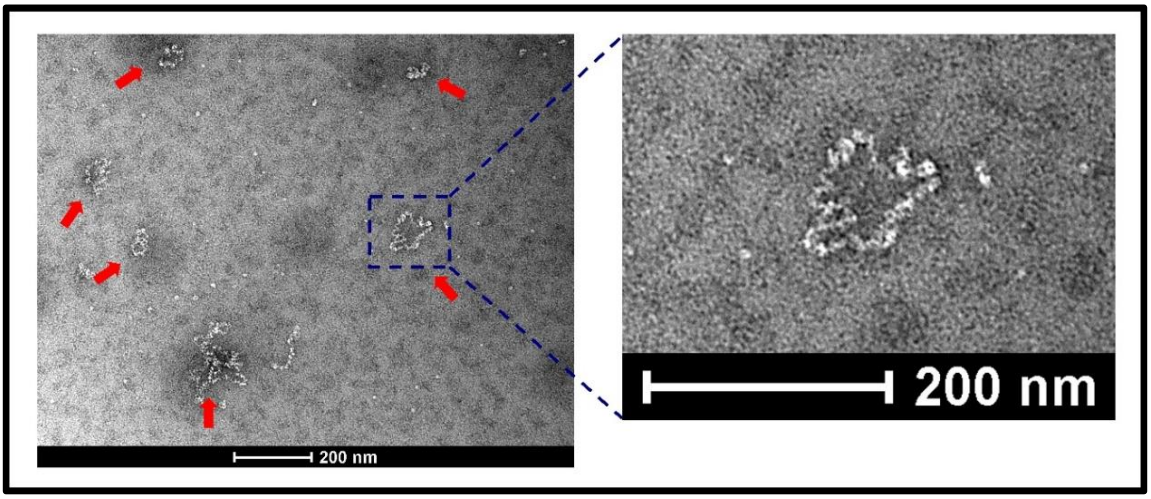

Figure 3. (**a**) Electron micrograph of a purified ribonucleoprotein preparation showing tenuivirus-like circular filamentous ribonucleoproteins (RNP) of different sizes (red arrows)

To assess the serological relationship between JKI 29327 and E11–018, *N. benthamiana* and *Physalis floridana* leaves infected with the MeCSV isolate E11–018 (kindly provided by Dr. C. Desbiez) were tested in DAS-ELISA using the JKI 1608 antibodies to JKI 29327. The latter gave strong ($A_{405\ nm}$ values: > 2.0) reactions with JKI 29327 (in four

different plant spp.) and intermediate reactions ($A_{405\ nm}$: 1.0 to 2.0) with E11–018, indicating that the serological relationship between these two isolates is close. Additionally, JKI 29327 was mechanically inoculated to melon cv. Védrantais (kindly provided by Dr. C. Desbiez). The plants showed chlorotic spots only on inoculated leaves and tested positive in DAS-ELISA with the JKI 1608 antibodies. Whilst JKI 29327 could be detected in inoculated leaves, no systemic infection was observed (data not shown).

The species demarcation criteria of ICTV for the genus *Tenuivirus* suggest that a new species should be considered when (i) the aa sequence identities between any corresponding gene products is below 85%; (ii) the nt sequence identities between corresponding IRs is below 60%; (iii) there are different sizes and/or numbers of genomic components; (iv) there are differences in host range; (v) the vectors are different (Shirako *et al.*, 2012). For certain tenuiviruses, it has been difficult to decide whether they belong to the same or different species because all the five criteria are not always met (Shirako *et al.*, 2012). For example, RHBV, EHBV and UHBV have different vectors, different hosts, different sizes and numbers of RNA segments and the nt sequence identity of their IR is < 60%. Yet, the four protein homologs on their RNA3 and RNA4 are 90% identical in aa sequences.

The black medic tenuivirus isolate JKI 29327 fulfils three out of these five criteria. Firstly, its ORF2 of RNA7 shares 80.6% aa identity with its homologue in the E11–081 genome. Secondly, the IRs of both RNA7 and RNA8 share < 60% nt identities with those of E11–081. Thirdly, the overall genome size of JKI 29327 is 247 nt shorter than that of E11–081. Based on these three criteria, the black medic virus should be considered a new species. However, although the host range was not studied in full detail, both JKI 29327 and E11–081 infected members of the *Fabaceae*, the *Cucurbitaceae* and the *Solanaceae* families under experimental conditions. Moreover, these two isolates appear to be serologically closely related when tested with the JKI 1608 antibodies. Small differences in size, particularly in the intergenic regions, are common and can be observed between isolates of RSV (Wei *et al.*, 2009; Lu *et al.*, 2018). Also, segment RNA 7 of MeCSV E11–018 was shown to present size heterogeneity due to indels in the intergenic region (Lecoq *et al.*, 2019). Furthermore, only one protein out of 13 was below the 85% identity threshold. Therefore, we propose that the black medic isolate from Austria is a strain of MeCSV and is referred to accordingly as black medic strain of MeCSV. Further studies are required to identify possible natural hosts and insects that may act as vectors of both JKI 29327 and E11–081. Moreover, there is a need to compare the experimental and natural host ranges of the two MeCSV strains. The antiserum obtained in this study will help to monitor prevalence and geographic distribution of MeCSV as well as its agronomic impact on crop plants (e.g., melons, legumes). Furthermore, it is important to study the function of the virus proteins that have been predicted in silico.

3.3. Complete genome sequence of highly divergent carrot torradovirus 1 strain from *Apium graveolens*

Yahya Zakaria Abdou Gaafar and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar YZA, Ziebell H, 2019. Complete genome sequence of a highly divergent carrot torradovirus 1 strain from *Apium graveolens*. *Archives of virology*. doi: 10.1007/s00705-019-04272-3.

3.3.1. Abstract

A new virus was identified in a celery plant showing chlorotic rings, mosaic and strong yellowing symptoms, and its complete genome sequence was determined. The genomic organization of this novel virus is analogous to that of known members of the genus *Torradovirus*, consisting of two single-stranded RNAs of 6,823 (RNA1) and 4,263 nucleotides (RNA2), excluding the poly(A) tails. BLAST searches against the nucleotide and protein databases showed that this virus is closely related to but different from carrot torradovirus 1 (CaTV1). Comparisons between the two viruses demonstrated relatively low levels of nucleotide and amino acid similarity in different parts of their genomes, as well as considerable differences in the sizes of their two genomic RNAs. However, the protease-polymerase (Pro-Pol) and capsid protein (CP) regions of this virus share >80% amino acid identity with the corresponding regions of CaTV1. Therefore, based on the current ICTV species demarcation criteria for the family *Secoviridae*, the virus from celery is a divergent strain of CaTV1, named "CaTV1-celery". Nevertheless, differences between CaTV1 and CaTV1-celery in genome size, as well as in biological and epidemiological features, may warrant their separation into two distinct species in the future.

3.3.2. Main text

A celery plant (*Apium graveolens* variant *graveolens*) was collected in August 2017 in the state of Hesse in Germany. The sample displayed mosaic symptoms with chlorotic rings and strong yellowing. To identify the possible cause of the disease, leaf material was examined by electron microscopy, but no virus particles were observed. However, mechanical inoculation of *Nicotiana benthamiana, N. clevelandii*, *N. occidentalis*-P1 and *Coriandrum sativum* with the sap of the celery plant resulted in systemic infections three weeks after inoculation, with symptoms consisting of chlorosis and necrosis. No virus particles could be observed in samples from symptomatic test plants. Attempts to transmit the pathogen mechanically to the original host species *A. graveolens* or to *Ammi majus*, *Anethum graveolens*, *Daucus carota* and *Petroselinum crispum* were not successful.

To characterise the genome of the suspected virus, total RNA was extracted from symptomatic *N. benthamiana* leaves using an innuPREP RNA Mini Kit (Analytik Jena AG). The ribosomal RNAs were depleted using a RiboMinus Plant Kit (Invitrogen) and the ribo-depleted RNA was used for library preparation using a Nextera XT Library Kit (Illumina). The library was subjected to high-throughput sequencing (HTS) on a MiSeq v3 platform (2x301). The reads were quality trimmed and filtered using Geneious software (version 11.1.3) (Biomatters Limited). The high-quality reads were assembled using the Geneious *de novo* assembly tool. In BLASTn searches, two assembled contigs of 6,727 and 4,106 nt shared 71.7% and 70.5% identical nucleotides with the two genomic RNAs of carrot torradovirus 1 (CaTV1) (NC_025479 and NC_025480).

The 5′ ends of both RNA segments were confirmed using RNA-ligase-mediated amplification of cDNA ends (RLM-RACE) (Coutts & Livieratos, 2003). The 3′ ends of the two RNA segments were determined via RT-PCR using a virus-specific primer and an oligo(d)T primer. The PCR products were cloned and sequenced, and the resulting sequences were assembled. The assembled full-length sequences of the two RNA segments were 6,823 nt (RNA1) and 4,263 nt (RNA2) in length, excluding their poly(A) tails. The complete genome sequences of RNA1 and RNA2 were deposited in the GenBank database (accession nos. MK063924 and MK063925, respectively). Diagnostic primers (HZ-539 5′ TGT TAG CAG AGC TAC GTC CTC 3′ and HZ-568 5′ CCT GAA TCT GCC CAC GAC TT 3′) were designed using the Primer3 v. 2.3.7 tool to amplify a partial sequence of RNA2-ORF1 (730 nt) to confirm the presence of this virus in infected plants (Untergasser *et al.*, 2012).

According to the species demarcation criteria proposed by the ICTV *Secoviridae* Study Group, viruses belonging to different species share less than 80% aa sequence identity in the protease-polymerase (Pro-Pol) region of the RNA1 polyprotein and less than 75% aa sequence identity in the coat protein (CP) region (Sanfaçon *et al.*, 2009). The celery virus shares 86.4% aa sequence identity in the Pro-Pol region and 80.3% aa sequence identity in the CP region with CaTV1 (Table 1). Based on these criteria, the celery virus should be considered a new strain of CaTV1, for which we propose the name “CaTV1-celery” (isolate JKI-29346). However, further comparison of the CaTV1-celery genome with the reference sequences of CaTV1 revealed considerable differences.

Table 1: Nucleotide (nt) and amino acid (aa) sequence comparisons between carrot torradovirus 1-celery (CaTV1-celery) and previously characterized torradoviruses

Virus			CaTV1-celery	CaTV1	LNLCV	MYMoV	SCLSV	ToChSV	ToChV	ToMarV	ToNDV	ToTV
Nt length	RNA 1	Accession no.	MK063924	NC_025479	NC_035214	NC_035218	NC_035221	NC_013075	FJ560489	NC_010987	NC_027926	NC_009013
		Whole -poly (A)	6,823	6,917	7,576	7,068	7,045	7,473	7,474	7,221	7,236	7,793
		5'UTR	146	126	170	135	231	138	135	140	150	106
		ORF1	6,228	6,579	6,672	6,576	6,624	6,468	6,468	6,456	6,453	6,477
		3'UTR	449	212	743	357	190	867	871	624	633	1,210
	RNA 2	Accession no.	MK063925	NC_025480	NC_035219	NC_035220	NC_035215	NC_013076	FJ560490	NC_010988	NC_027927	NC_009032
		Whole -poly (A)	4,263	4,969	5,290	4,963	4,730	5,093	5,695	4,898	4,896	5,389
		5'UTR	148	611	326	406	83	171	181	138	138	181
		ORF1	606	609	636	618	702	555	570	573	570	564
		ORF2	3,174	3,504	3,669	3,588	3,165	3,594	3,579	3,576	3,573	3,597
		3'UTR	294	301	766	461	899	814	1,406	652	656	1,105
Pairwise identity %												
Nt %	RNA 1	5'UTR		67.5	39.1	39.1	30.8	44.4	43.5	41.3	37.4	43.7
		Whole		70.1	53.3	54.4	42.3	43.1	43.1	42.5	42.9	42.5
		ORF1		71.2	55.2	55.9	43.6	45.2	45	44.9	45.3	44.6
		3'UTR		68.4	24.7	26.7	31.4	29.9	30.1	28.6	26.6	24.7
	RNA 2	5'UTR		45	32.3	33.9	46.3	39.8	43.8	47.2	39.6	49.5
		Whole		63.5	46.4	48	39.3	38.2	39.2	38.6	39.1	39.6
		ORF1		65.7	46.9	49.8	36.7	41.1	37.5	37.9	38.7	38.7
		ORF2		65.6	49.1	50.8	43.8	40	40.2	40.2	40.8	40.7
		3'UTR		55.5	30.4	31.3	33.1	34.2	34.2	30.9	33.9	30.5
aa %	RNA 1	ORF1		76.1	50.9	51.9	33.5	33.8	33.7	33.9	33.6	33.2
		Hel		99	85.4	85.4	57.3	61.2	61.2	61.2	61.2	59.2
		Pro-Pol		86.4	67.1	71	57.4	58.8	60.1	59.9	59.3	57.3
	RNA 2	ORF1		66.8	40.8	48.5	25.7	27	24.3	27.2	26.7	26.2
		ORF2		71.9	44.9	45.9	30.7	31.1	30.7	30.7	31.2	30.8
		MP		59.1	30.9	31.7	23.3	22.4	22.1	22.1	21.6	22.1
		CP		80.3	54.2	55.9	36.6	40	39.3	39.6	40.4	39.7
		Vp35		73.9	47.7	45.3	31.7	33.1	34.6	33.1	34.3	32.9
		Vp26		83	60.7	68.3	41.9	45.9	42.9	45.5	45.9	45
		Vp23		84.1	54.3	54.4	37.9	41.7	41.3	40.4	41.3	41.7

The viruses used for comparison and their respective abbreviations are as follows: carrot torradovirus 1 (CaTV1), lettuce necrotic leaf curl virus (LNLCV), motherworth yellow mottle virus (MYMoV), squash chlorotic leaf spot virus (SCLSV), tomato chocolate spot virus (ToChSV), tomato chocolate virus (ToChV), tomato marchitez virus (ToMarV), tomato necrotic dwarf virus (ToNDV) and tomato torrado virus (ToTV)

The genomic organization of CaTV1-celery is analogous to that of other torradoviruses (Thompson *et al.*, 2017). Accordingly, RNA1 contains a single open reading frame (ORF1) encoding a polyprotein of 2,076 aa (MW: 237 kDa). A search of the Conserved Domain Database (CDD) of NCBI identified the presence of two conserved domains: the RNA helicase (Hel) (RNA_helicase; pfam00910) and RNA-dependent RNA polymerase (RdRp) (RNA_dep_RNAP; cd01699) in the RNA1 polyprotein. An additional motif, characteristic of cysteine proteases (3C), was identified using the Geneious search tool (Fig. 1a) (Argos *et al.*, 1984; Gorbalenya *et al.*, 1989; Gorbalenya & Snijder, 1996). RNA2 is bicistronic, with RNA2 ORF1 encoding a 201-aa protein (MW: 22.0 kDa) and RNA2 ORF2 encoding a predicted polyprotein of 1,057 aa (MW: 117.1 kDa). The protein encoded by RNA2 ORF1 did not match any domain in the CDD database. The RNA2 ORF2 polyprotein contains two conserved domains: the 3A movement protein (MP) family domain (3A superfamily; cl02970) and picornavirus capsid protein (CP)-like domain (rhv_like superfamily; cl13999). A motif search identified the position of the MP conserved motif (LxxPxL) in the RNA2 ORF2 polyprotein (Mushegian, 1994; Verbeek *et al.*, 2007). In addition, the MP and the three mature CP subunit peptides were determined based on homologies to those of other torradoviruses (Fig. 1a).

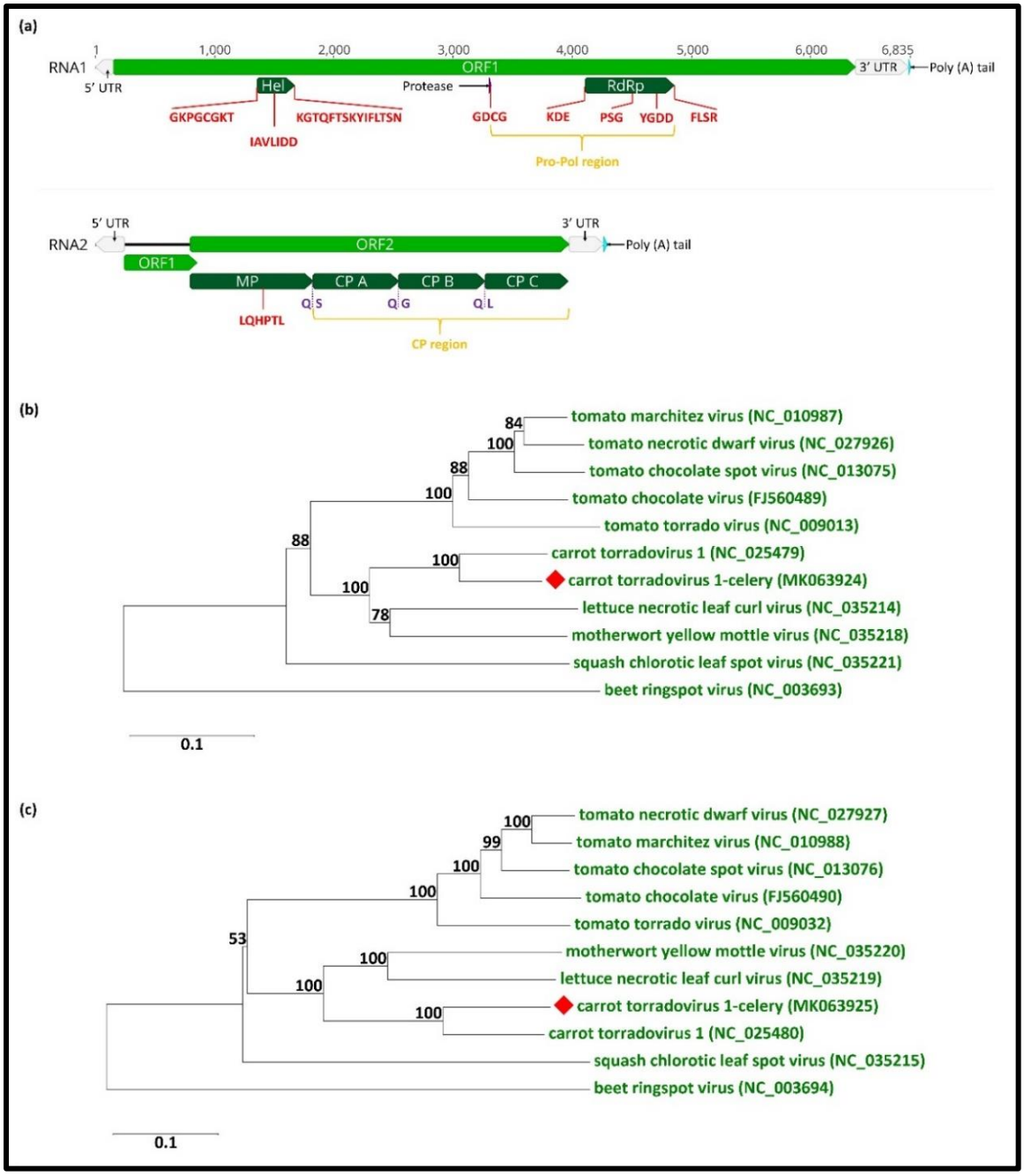

Figure 1: (a) Schematic representation of the carrot torradovirus 1-celery (isolate JKI 29346) genome. RNA1 encodes a polyprotein containing the helicase (Hel), protease and RNA-dependent RNA polymerase (RdRp). RNA2 has an ORF1 with unknown function and ORF2 encoding a polyprotein containing the movement protein (MP) and three coat protein subunits (CP A = Vp35, CP B = Vp26 and CP C = Vp23). Both strands are flanked by 5' and 3' untranslated regions (UTR) and are polyadenylated at the 3' end. The aa sequences of the conserved motifs in Hel, RdRp and MP are shown in red. The predicted cleavage sites are shown in purple. Taxonomically relevant protein segments are highlighted in yellow **(b and c)**. Neighbour-joining trees based on amino acid sequence alignments of the Pro-Pol **(b)** and CP **(c)** of CaTV1-celery (red diamond) with those of members of the genus *Torradovirus*, with beet ringspot virus (genus *Nepovirus*) as an outgroup. The GenBank accession numbers are shown in brackets. Bootstrap values above 50% (1000 replicates) are indicated for each node, and the scale bar represents a genetic distance of 0.1

The percentages of identity in a ClustalW 2.1 pairwise alignment between CaTV1-celery RNAs, ORFs and regions with their cognates in other torradoviruses are listed in Table 1. Comparisons of the CP region sequences showed that CaTV1-celery shares between 36.6 and 80.3% aa sequence identity with torradoviruses. Additionally, the Pro-Pol region of CaTV1-celery shares between 57.3 and 86.4% aa sequence identity. Neighbour-joining trees based on the aa sequence alignments of the Pro-Pol and CP regions showed that the celery virus clusters together with CaTV1 within the genus *Torradovirus* (Fig. 1b and c) (Larkin *et al.*, 2007; Kumar *et al.*, 2016).

Despite the observed similarities, the lengths of the 5' and 3' untranslated regions (UTR) of RNA1 are 146 and 449 nt long, respectively, and those of RNA2 are 245 and 294 nt long. These values differ from the 5' and 3' UTRs of CaTV1, which are 126 and 212 nt long, respectively, in RNA1 and 611 nt and 301 nt long in RNA2. Additionally, pairwise alignment of the UTRs with those of other torradoviruses, including CaTV1, showed low nt sequence identity values between 24.7% and 68.4% (Table 1). The predicted protein encoded by ORF1 of RNA2 of CaTV1-celery shares only 66.8% aa sequence identity with its homolog in CaTV1, and the 3'UTRs of both segments share ≤ 68.4% nt sequence identity with those of CaTV1. Furthermore, RNAs 1 and 2 of CaTV1-celery are shorter by 96 and 706 nt, respectively, than their CaTV1 counterparts, making the CaTV1-celery genome the shortest torradovirus genome identified so far, with a total size of 11,086 nt, excluding the poly (A) tail (Table 1).

Another criterion that is considered for species demarcation in the family *Secoviridae* is vector specificity. The close evolutionary relationships between CaTV1, CaTV1-celery and lettuce necrotic leaf curl virus (LNLCV) suggest that CaTV1-celery is also an aphid-borne torradovirus (Rozado-Aguirre *et al.*, 2016; Verbeek *et al.*, 2017). To test this, *Myzus persicae* aphids from a single laboratory clone were reared on CaTV1-celery-infected *N. benthamiana* for seven days, after which ten aphids were transferred to three groups of healthy plants (10 *N. benthamiana*, 10 *A. graveolens* and 10 *D. carota*). After an inoculation access period of seven days, the plants were treated with the systemic insecticide PIRIMOR (Deutsche ICI) according to the manufacturer's instructions. The plants were incubated under greenhouse conditions (at 22 °C; photoperiod of 16 h light and 8 h dark) for two months, but no symptoms were observed. The absence of virus infections in acceptor plants was additionally confirmed by negative RT-PCR results. Aphid transmission experiments were repeated three times.

Although CaTV1-celery is considered a divergent strain of CaTV1 based on their aa sequence similarity in the Pro-Pol and CP regions, the data suggest that it might also be useful to consider other molecular criteria for species demarcation, i.e., the total genome size and the size and degree of sequence similarity of the 5' and 3' UTR. Taking these criteria into consideration, CaTV1-celery might be accepted in the future as a member of a novel species. Indeed, Sanfaçon and colleagues have already suggested that the current demarcation criteria could be revisited and modified as more viruses become characterized (Sanfaçon *et al.*, 2009). Furthermore, Verbeek and colleagues have proposed additional criteria, i.e., that the aa sequence identity of the RNA2 ORF1

should be less than 75% and that the conservation level in the 3'UTR of both RNAs should be less than 85% (Verbeek *et al.*, 2010).

Further studies are needed to investigate possible vectors of CaTV1-celery and its potential impact on celery production. Currently, we are developing antibodies against CaTV1-celery as an additional tool for future diagnostic tests and the determination of serological relationships of different torradoviruses.

3.4. First report of natural infection of beetroot with beet soil-borne virus

Yahya Zakaria Abdou Gaafar, A. Sieg-Müller, P. Lüddecke, J. Hartrick, Y. Seide, Jürgen Müller, C. Maaß, S. Schuhmann, K.R. Richert-Pöggeler, A.G. Blouin, S. Massart and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar Y, Sieg-Müller A, Lüddecke, Hartrick J, Seide Y, Müller J, Maaß C, Schuhmann S, Richert-Pöggeler KR, Blouin AG, Massart S, Ziebell H, 2019. First report of natural infection of beetroot with *Beet soil-borne virus*. *New Disease Reports* 40, 5. doi: 10.5197/j.2044-0588.2019.040.005.

Beetroot (*Beta vulgaris* subsp. *vulgaris*) is becoming increasingly popular in Germany with an increase in field production from 1,205 ha in 2013 to 1,826 ha in 2018 (Behr, 2019). It is estimated that EU-wide 24,000 ha of beetroot were produced in 2018 (Behr, 2019). In contrast, sugarbeet was produced on a substantially larger scale with 413,900 ha in Germany in 2018 (Kemper *et al.*, 2019).

A symptomatic beetroot sample was collected in October 2018 from a small field (c. 200 m²) in Rhineland-Palatinate, Germany, where several plants displayed virus-like symptoms. The sample submitted displayed necrosis, reduced size and in particular root proliferation (bearding) resembling the symptoms of rhizomania (Fig. 1). However, immunosorbent electron microscopy (ISEM) examination using various antibodies raised against the following beet viruses was not successful in identifying any causal agent: beet black scorch virus, beet curly top virus, beet necrotic yellow vein virus, beet mosaic virus, beet oak-leaf virus, beet soil-borne virus (BSBV), beet soil-borne mosaic virus, beet virus Q, beet western yellows virus, beet yellows virus and tobacco rattle virus.

Figure 1: Beetroot sample infected with beet soil-borne virus and beet cryptic virus-2 showing leaf necrosis, reduced size and root bearding.

Total RNA was extracted from the infected beetroot sample using innuPREP RNA Mini Kit (Analytik Jena AG, Germany). Ribosomal RNA was depleted using RiboMinus Plant Kit for RNA-Seq (ThermoFisher Scientific, USA). A library was prepared using TrueSeq Stranded mRNA kit (Illumina, USA). The sequencing was done on a NextSeq 500 platform

(2×150). The generated data was analysed on Geneious Prime (2019.1.1). The raw reads were quality-trimmed then *de novo* assembled using SPAdes assembler (3.10.1) (Bankevich *et al.*, 2012). The contigs were compared against a local virus reference database using tBlastx. Twenty-one contigs showed only similarities to BSBV and Beet cryptic virus-2 (BCV-2), respectively; no other virus sequences were found in the data.

BSBV is a member of the *Pomovirus* genus (family *Virgaviridae*) (Adams *et al.*, 2017). The virus is widely distributed in sugar beet growing areas causing yield losses. BCV-2, a member of the *Deltapartitivirus* genus (family *Partitiviridae*), is a symptomless virus that is also common in sugar beet (Antoniw *et al.*, 1990; Vainio *et al.*, 2018). For confirmation of BSBV infection, total RNA was re-extracted from the infected beetroot sample and RT-PCR was performed using two specific primer pairs targeting the RNA-dependent RNA polymerase and movement protein regions of BSBV, respectively (HZ772 5'-GTTGGTGTTGGTCAGTTGGC-3' / HZ773 5'-TGGTCAACGGCGAAATCAGA-3' and HZ774 5'-GAGGGGTAAGACACAGCGAC-3' / HZ775 5'-CACTTCGTCCTCCTGGTCAC-3'). Two bands of the expected sizes (923 and 766 bp, respectively) were produced.

The almost complete genomes of BSBV and BCV-2 were assembled by Geneious mapping using reference sequences (BSBV: Genbank Accession Nos NC_003518-NC_003520 and BCV-2: NC_038845-NC_038847). The sequences of the beetroot BSBV and BCV-2 isolates were submitted to Genbank (MK731954-MK731959). Sequence analysis revealed that the BSBV isolate shares 97.3-98.5% nucleotide identity with the reference genome (German isolate NC_003518-NC_003520). The BCV-2 isolate shares 98.7-99.7% nt nucleotide identity with the reference genome (Hungarian isolate NC_038845-NC_038847).

This work provides the first suggestion that BSBV naturally infects beetroot. This identification exposes the limit of diagnostic methods such as ISEM, possibly due to low titre or the existence of a divergent virus isolate, and also highlights the strength of high throughput sequencing to rapidly and accurately diagnose plant viruses. Furthermore, these findings demonstrate that high value crops such as beetroot might be affected by pathogens of major crops and therefore should be considered in crop rotations.

3.5. First report of physostegia chlorotic mottle virus on tomato (*Solanum lycopersicum*) in Germany

Yahya Zakaria Abdou Gaafar, M. A. M. Abdelgalil, D. Knierim, K. R. Richert-Pöggeler, W. Menzel, S. Winter and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar YZA, Abdelgalil MAM, Knierim D et al., 2018. First report of physostegia chlorotic mottle virus on tomato (*Solanum lycopersicum*) in Germany. *Plant Disease* 102, 255. https://doi.org/10.1094/PDIS-05-17-0737-PDN

In September 2015, a tomato sample collected in the German state of Hesse was sent to the Julius Kühn-Institut for analysis. While the fruits showed marbling and discoloration, the leaf samples from this plant did not show any obvious symptoms (Fig. 1). Transmission electron microscopy (TEM) revealed the presence of bullet-shaped virus particles indicating the presence of a rhabdovirus (Fig. 2). However, immunosorbent electron microscopy using antiserum JKI-1073 for *Eggplant mottled dwarf virus* (EMDV) could not confirm EMDV infection. The virus was mechanically transmitted to *Nicotiana benthamiana*, *N. clevelandii*, and *Chenopodium quinoa* inducing yellowing and leaf deformation, while mechanical transmission to *N. occidentalis* (P1 and 37b) failed. Extraction of double stranded-RNA (dsRNA) followed by random-PCR (Froussard, 1992), cloning of PCR products, and sequencing failed to reveal any virus sequences.

Figure 1: Tomato fruit showing marbling, discoloration and leaf distortion whereas leaf symptoms consist of mild yellow spots.

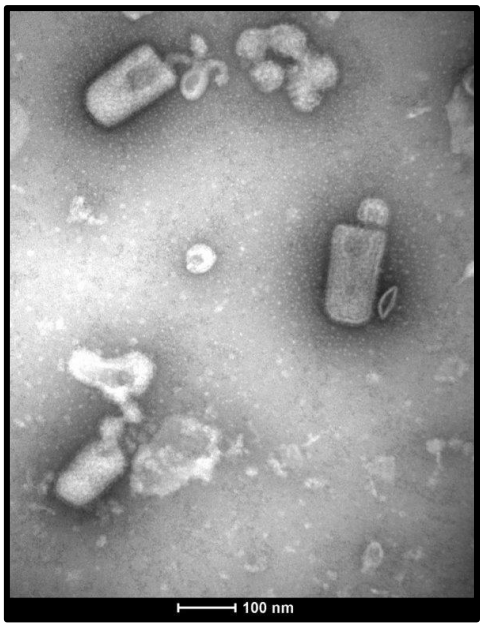

Figure 2: Electron microscopy photograph of physostegia chlorotic mottle virus.

Total RNA was extracted from infected *N. benthamiana*, followed by ribo-depletion, library preparation and submission for next-generation sequencing (NGS) using an Illumina MiSeq platform as described by (Knierim *et al.*, 2017). De novo assembly of the trimmed reads was done with Geneious v 10.1.3 (Biomatters LTD, NZ). Using MEGA BLAST, 13 contigs showed between 95.6 and 98.5% similarity with physostegia chlorotic mottle virus (PhCMoV) isolate PV-1182 (accession no. KX636164). The complete PhCMoV genome (13,321 nt length) was assembled by mapping reads to this reference genome and used to design PhCMoV-specific RT-PCR primers for detection (HZ-343 5' CGG TGA GTG GGG CAA CTA AT 3' and HZ-344 5' AGC GAT GGG GTC TAG TGT CT 3'). RT-PCR confirmed the presence of PhCMoV in the test plants resulting in amplicons of approximately 875 bp.

In August 2016, similar symptoms on tomato fruits were observed by a different grower in Hesse. The presence of PhCMoV was confirmed by TEM and RT-PCR. Additionally, the PCR products were sequenced and showed 97% identity to KX636164. Surprisingly, reanalysis of a tomato sample from 2003 that was infected by a hitherto unknown rhabdovirus using NGS also confirmed infection with PhCMoV. This sample also originated from Hesse although the original grower is unknown. The complete genome of the 2003 PhCMoV sample was assembled following the same methods described above. Pairwise comparison between the genomes of 2015 and 2003 isolates resulted in 99.7% nucleotide identity and 96.9% when compared with KX636164.

These findings indicate the presence of PhCMoV in tomato in Germany for a long time albeit isolated occurrences in different production areas. PhCMoV was recently identified from *Physostegia virginiana* plants showing leaf deformation and severe chlorotic and mottle symptoms in Austria (Menzel *et al.*, 2016). However, it is not known if there is a link between PhCMoV isolates infecting *P. virginiana* and tomato as the routes of transmission and dissemination are currently unknown. The sequences from this report were deposited in GenBank (accession nos. KY706238 and KY859866 [full-length sequences], KY882263 and KY882264 [partial sequences]). To our knowledge, this is the first host record of PhCMoV in tomato and a new country record for Germany.

3.6. First report of turnip crinkle virus infecting garlic mustard (*Alliaria petiolata*) in Germany

Yahya Zakaria Abdou Gaafar, A. Sieg-Müller, P. Lüddecke, K. Herz, J. Hartrick, C. Maaß, S. Schuhmann, K. R. Richert-Pöggeler and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar Y, Sieg-Müller A, Lüddecke P, Herz K, Hartrick J, Maaß C, Schuhmann S, Richert-Pöggeler KR, Ziebell H, 2019. First report of *Turnip crinkle virus* infecting garlic mustard (*Alliaria petiolata*) in Germany. *New Disease Reports* 39, 9. http://dx.doi.org/10.5197/j.2044-0588.2019.039.009

In May 2018, three samples of wild garlic mustard (*Alliaria petiolata, Brassicaceae*) were collected from a private garden in Koenigslutter, Germany. While sample EPV_18_002 was asymptomatic apart from slight yellowing, samples EPV_18_003 and EPV_18_004 showed stunting, yellowing, necrosis and severe crinkling (Figs. 1-3). It was possible to mechanically transmit the suspected virus from all three samples to *Nicotiana benthamiana*; the same systemic symptoms of leaf crinkling, rolling and yellowing appeared seven days post infection on all plants, and the plants died in the second week. Infected *N. benthamiana* leaves were analysed by electron microscopy and icosahedral particles of c. 30 nm in diameter were observed that reacted with antibodies (Julius Kuehn Institute, reference number JKI-1177) raised against a UK isolate of turnip crinkle virus (TCV) (Fig. 4).

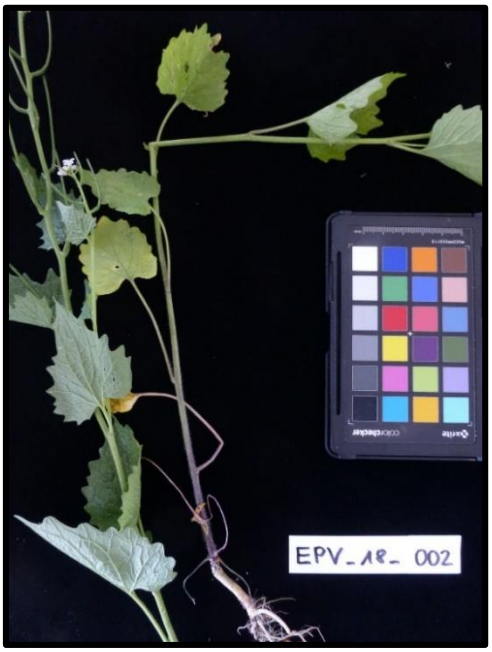

Figure 1: Sample EPV_18_002 of garlic mustard infected with turnip crinkle virus.

Figure 2: Sample EPV_18_003 of garlic mustard infected with turnip crinkle virus.

Figure 3: Sample EPV_18_004 of garlic mustard infected with turnip crinkle virus.

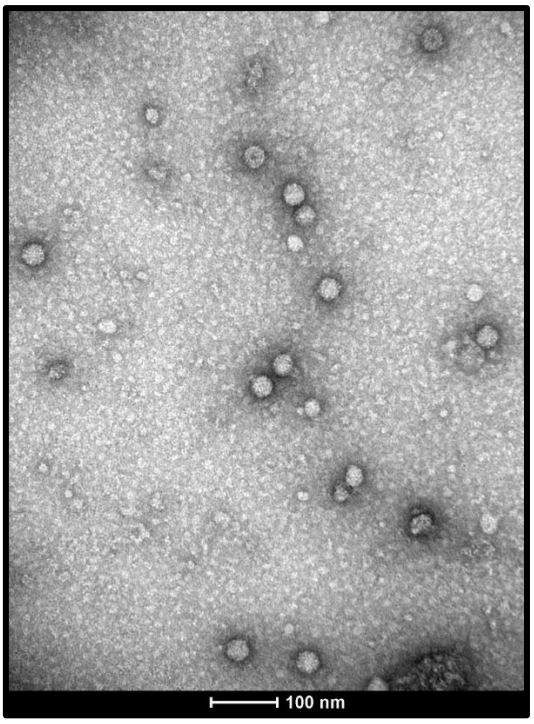

Figure 4: Electron microscope image of turnip crinkle virus (TCV) from infected *Nicotiana benthamiana* leaves. Particles were decorated using a TCV-specific antiserum (JKI-1177).

To obtain the full viral genome, dsRNA was extracted from symptomatic *N. benthamiana* leaves (inoculated from sample EPV_18_002) using a Viral dsRNA Extraction Mini Kit for Plant Tissue (iNtRON, South Korea) and sent for library preparation and sequencing on Illumina NovaSeq 6000 platform (2x150 bp) at Eurofins GATC Biotech GmbH, Germany. Using Geneious Prime (2019.0.4), the raw reads were quality trimmed,

filtered, error corrected and normalised, followed by *de novo* assembly. A BLASTn search of the generated contigs confirmed the presence of eight contigs (131 to 4,057nt) with nucleotide (nt) identities of 82.3% to 86.7% to TCV (NC_003821); there was no indication of the presence of satellite RNAs. The full genome of the German TCV isolate (JKI ID 29306) was assembled using this reference genome. The complete genome of TCV-JKI-29306 was 4,061 nt (accession no. MK301398). The presence of TCV was confirmed in all three original samples by RT-PCR using OneTaq One-Step RT-PCR Kit (NEB, USA) with the primer pair (HZ632 5' AAA GGC AAA ACT GGG TGG GA 3' and HZ633 5' TAA AGT TTG CGG CTA GGG G 3') generating a 339 bp fragment.

In further comparisons using MUSCLE (3.8.425, (Edgar, 2004)), the German TCV isolate shared 82.8% nt identity to NC_003821 and 82.6% to AY312063 (Table 1). Additionally, the protein sequences of the different TCV genes shared identities of 75.0 to 90.7% to the proteins of NC_003821 and AY312063 (Table 1). According to ICTV criteria, these data indicate the presence of a novel TCV strain (Rochon *et al.*, 2012).

Table 1. Pairwise identities of the turnip crinkle virus (TCV) strain from garlic mustard in Germany with whole TCV genomes from the UK using a multiple sequence alignment tool, MUSCLE.

Accession no.	Genome nucleotide identity (%)	Amino acid identity (%)				
		p88 RP	p28 ARP	p8 MP	p9 MP	p38 CP
NC_003821	82.8	90.5	84.4	75.0	89.4	85.5
AY312063	82.6	90.7	84.8	76.4	88.2	85.2

TCV belongs to the genus *Betacarmovirus* (family *Tombusviridae*). It infects members of the *Brassicaceae* family causing crop losses (Broadbent & Heathcote, 1958; Lister, 1958). To our knowledge, this is the first report of a TCV strain from Germany occurring on garlic mustard. Although TCV is one of the model organisms in plant virology and therefore well studied, there is surprisingly little literature available on occurrence, host range and impact of this virus. As the original host plants appeared as weeds in a private garden (approximately 40 plants with a minority displaying symptoms), the impact of TCV in Germany and other countries on cultivated crop plants is currently unknown.

3.7. First report of southern tomato virus in German tomatoes

Yahya Zakaria Abdou Gaafar, P. Lüddecke, C. Heidler, J. Hartrick, A. Sieg-Müller, C. Hübert, A. Wichura and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar Y, Lüddecke P, Heidler C et al., 2019. First report of *Southern tomato virus* in German tomatoes. *New Disease Reports* 40, 1. doi: 10.5197/j.2044-0588.2019.040.001.

Southern tomato virus (STV) is a member of the genus *Amalgavirus* (family: *Amalgaviridae*). It has been identified in tomatoes (*Solanum lycopersicum*) in several countries in Asia, Europe, and North and South America (Sabanadzovic *et al.*, 2009; Candresse *et al.*, 2013; Padmanabhan *et al.*, 2015). Its genome is composed of a dsRNA of ~3.5kb. STV is known to be transmitted through seed at high rates (Sabanadzovic *et al.*, 2009).

In 2019, greenhouse tomatoes from Lower Saxony, Germany, showed symptoms consisting of mottling, yellowing and/or chlorotic spots (Figs. 1-2). Eight samples were sent to the Julius Kuehn Institute for analysis. As an infection with Tomato brown fruit rugose virus was suspected, the samples were analysed by high throughput sequencing on a MinION sequencer (Oxford Nanopore Technologies, UK) for rapid diagnosis. Briefly, two samples were pooled, and dsRNA extracted from 100 mg leaf material using the Viral dsRNA Extraction Mini Kit for Plant Tissue (iNtRON, South Korea). Random cDNA was synthesised using ProtoScript II Reverse Transcriptase (NEB, USA) and 8N random primers preceded by a denaturation step at 99°C for two minutes. Second strand synthesis was done using a NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module kit (NEB). The samples were end-repaired using the NEBNext End Repair module (NEB), dA tailed with NEBNext dA-tailing module (NEB) and the four pools barcoded by native barcoding followed by adaptor ligation (Oxford Nanopore Technologies, UK) according to the manufacturers' instructions. All purification steps were performed using a Mag-Bind TotalPure NGS kit (Omega Bio-Tek, USA). The libraries were mixed and loaded to a MinION flow cell and sequenced for 16 hours using a MinION sequencer connected to a computer with MinION software (r18.12.9; ONT).

Figure 1: Chlorotic lesions observed on diseased tomato plants grown in Lower Saxony.

Figure 2: Severe chlorotic lesions and mottling of leaves on diseased tomato plants grown in Lower Saxony.

The reads were basecalled and barcode-splitting was done using the Guppy toolkit (v2.3.7; ONT). *De novo* assembly of reads was done using Canu (v1.8) (Koren *et al.*, 2017). The unassembled reads and assembled contigs were Blastn searched against a local GenBank nt database using Blast+ (v2.9.0) and visualised with Blast Viewer (v5.2.0) (Durand *et al.*, 1997; Camacho *et al.*, 2009). STV sequences were detected in two of the pools, and no other virus sequences were detected. The full genomes were assembled by mapping to STV reference (Genbank Accession no. NC_011591) using mapping to reference tool on Geneious Prime (v2019.1.3). The sequences had 99.9% nt identity to each other and to STV isolate CH_bpo 163 from Switzerland (MF422618).

To confirm the findings in the original samples, total RNA was extracted for the four samples in the positive pools using an innuPREP RNA MiniKit (Analytik Jena AG, Jena, Germany), and RT-PCR was performed using a primer pair (HZ782 5'-CAAGTGGGCCGTTTCTTTGG-3' and HZ783 5'-TGAAGACCGCCTGGAAAGTC-3'). STV infection was confirmed in three samples. The RT-PCR products were purified, and Sanger sequenced at Eurofins Genomics (Germany). The sequences had 100% identity to the sequences from the MinION. The genomes of the two pools were submitted to Genbank (MK948544 and MK948545). To our knowledge, this is the first report of STV infecting tomato in Germany. The study also shows the potential to use MinIon technology for rapid detection and identification of virus sequences.

3.8. Two divergent isolates of turnip yellows virus from pea and rapeseed and first report of turnip yellows virus-associated RNA in Germany

Yahya Zakaria Abdou Gaafar and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar YZA, Ziebell H, 2019. Two divergent isolates of turnip yellows virus from pea and rapeseed and first report of turnip yellows virus-associated RNA in Germany. *Microbiology Resource Announcements* **8**, 2254. doi: 10.1128/MRA.00214-19.

3.8.1. Abstract

Two divergent isolates of turnip yellows virus (TuYV) were identified in pea and rapeseed. The nearly complete genome sequences of the virus isolates share 93.3% nucleotide identity with each other and 89.7% and 92.9% with their closest isolate from South Africa. Additionally, a turnip yellows virus-associated RNA was identified.

3.8.2. Main text

Turnip yellows virus (TuYV), the non-sugar beet-infecting strain of beet western yellows virus (BWYV), is a polerovirus (family *Luteoviridae*) (Mayo, 2002; Graichen & Rabenstein, 1996). TuYV can infect a wide range of crops, predominantly members of the *Brassicaceae* and *Fabaceae* families.

Two plant samples (pea [*Pisum sativum*] and oilseed rape [*Brassica napus*]) displaying yellowing symptoms were collected in Germany (in 2016 and 2006, respectively). The samples were tested with a triple antibody sandwich enzyme-linked immunosorbent assay (TAS-ELISA) for the presence of luteoviruses/poleroviruses using monoclonal antibodies 5G4 and 6G4 (Katul, 1992), as described (Abraham *et al.*, 2006; Gaafar *et al.*, 2016). Both samples tested positive for polerovirus infection but showed different titers, which prompted us to determine the genomic sequences of these isolates. The viruses were maintained on pea (isolate identifier [ID] JKI 29344) and radish (JKI 29345) by aphid transmission using *Myzus persicae*. Nonviruliferous aphids were left to feed for 3 days on the infected leaves, and then 10 aphids were transferred to healthy plants for 3 days (3 cycles for the pea isolate and 33 cycles for the oilseed rape isolate). Four weeks post-aphid inoculation, polerovirus infection of plants was confirmed using TAS-ELISA and reverse transcription-PCR (RT-PCR) with generic polerovirus primers (S2 and AS3) (Abraham *et al.*, 2007). The RT-PCR bands were Sanger sequenced, and a BLASTn search resulted in the highest hits, with 99% (pea isolate) and 100% (oilseed rape isolate) nucleotide identities to the partial coat protein sequences of BWYV (GenBank accession no. L39976) and TuYV (GenBank accession no. KU198395).

For genome sequencing, total RNAs were extracted with the innuPREP RNA minikit (Analytik Jena AG), followed by rRNA depletion with the RiboMinus plant kit (Invitrogen). cDNAs were synthesized using ProtoScript II reverse transcriptase (NEB) and random octanucleotide primers, followed by second-strand synthesis with the NEBNext Ultra II nondirectional RNA second-strand synthesis module kit (NEB). The libraries were prepared using a Nextera XT library kit (Illumina) and submitted for high-throughput sequencing (HTS) on the MiSeq version 3 platform (2 × 301). The raw reads (total reads, 1,640,360 for JKI 29344 and 1,648,784 for JKI 29345) were analyzed using the Geneious software (11.1.4). The reads were quality trimmed (error limit, 0.05) and size filtered to >99 nucleotides (nt), followed by *de novo* assembly using the Geneious

assembler (parameter, medium sensitivity/fast). The assembled contigs were used to search the NCBI database using BLASTn. A contig of about 5.6 kb in each sample showed 90.7% and 93.1% nt identities to TuYV (GenBank accession no. KU198395). Contig extension using Geneious mapping to the reference tool (parameter, medium sensitivity/fast) resulted in the complete coding sequences and almost-full-genomic sequences for both TuYV isolates; they shared 93.3% nt identity to each other while sharing 89.7% and 92.9%, respectively, to the most closely related isolate, KU198395.

Pairwise comparisons of the amino acid sequences using MUSCLE (3.8.425) showed that some open reading frames (ORFs) are also highly divergent (P0 and P1), whereas others are not (CP and MP) (Table 1) (Edgar, 2004). The pea and the oilseed rape isolates' ORFs shared between 80.9% and 99.5% amino acid (aa) identities to each other and between 74.1% and 95.6% in comparison to their homologues of KU198395.

Table 1: Pairwise amino acid comparisons between the predicted proteins of the German TuYV isolates and their homologues of KU198395 using MUSCLE 3.8.425

Protein source (isolate)	Amino acid identity (%) by open reading frame(s)						Reference isolate or accession no.
	P0	P1-P2	P1	P3-P5	CP	MP	
Pea (JKI 29344)	80.9	90.6	86.5	97.8	99.5	99.4	JKI 29345
	74.1	89.8	85.3	94.1	93.1	90.9	KU198395
Oilseed rape (JKI 29345)	85.9	95.6	94.2	93.7	92.6	90.3	KU198395

An additional contig of about 2.8 kb was found in the oilseed rape sample that shared 98% nt identity with the partial sequence of beet western yellows virus-associated RNA (BWYVaRNA) from the United Kingdom (GenBank accession number KF533709) (Adams *et al.*, 2014). Polerovirus-associated RNAs are single-stranded RNAs (ssRNAs) of ~2.8 to 3 kb and have two major ORFs. They replicate autonomously and appear to depend on a helper virus for aphid transmission by encapsidating within the virus coat protein (Briddon *et al.*, 2012). They may increase the severity of disease symptoms. The full genome of 2,841 nt was assembled by mapping to the reference sequence with NCBI RefSeq accession no. NC_004045 (Chin *et al.*, 1993), and we propose the name "turnip yellows virus-associated RNA" (TuYVaRNA) for this RNA. While TuYVaRNA shares 93% nucleotide identity with NC_004045 and 98% with the partial sequence from the United Kingdom, its genomic organization was similar to that of the other polerovirus-associated RNAs, containing three ORFs, with the first one containing an amber readthrough ORF.

To our knowledge, these are the first complete coding sequences of TuYV and the first report of TuYVaRNA from Germany.

3.8.3. Data availability

The complete coding sequences of the two German TuYV isolates and the full sequence of TuYVaRNA can be found in NCBI GenBank under accession numbers MK450519, MK450520, and MK450521. Raw sequence data are available in the Sequence Read Archive (SRA) under BioProject accession number PRJNA524397 and under BioSample accession numbers SAMN11026350 and SAMN11026351.

3.9. *Vicia faba*, *V. sativa* and *Lens culinaris* as new hosts for pea necrotic yellow dwarf virus in Germany and Austria

Yahya Zakaria Abdou Gaafar, S. Grausgruber-Gröger and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar Y, Grausgruber-Gröger S, Ziebell H, 2016. *Vicia faba*, *V. sativa* and *Lens culinaris* as new hosts for *Pea necrotic yellow dwarf virus* in Germany and Austria. New Disease Reports 34, 28. doi: 10.5197/j.2044-0588.2016.034.028

Pea necrotic yellow dwarf virus (PNYDV) was identified in green peas (*Pisum sativum*) in Germany in 2009 (Grigoras *et al.*, 2010a). In subsequent years, sampling of symptomatic green peas showed that PNYDV was restricted to Saxony and Saxony-Anhalt (Ziebell, 2015). In Austria, PNYDV was detected in 2010, also in *P. sativum* (Grigoras *et al.*, 2014).

A countrywide outbreak of virus-like disease symptoms on faba beans (*Vicia faba*) was reported in Germany in 2016. Many fields had large patches of yellowish and dwarfed plants (Fig. 1). More than 460 samples of *P. sativum* (green and protein peas) and *V. faba* showing virus-like symptoms (Fig. 2) were analysed using ELISA for *Alfalfa mosaic virus*, *Cucumber mosaic virus*, *Pea enation mosaic virus* (PEMV), *Red clover vein mosaic virus*-like carlaviruses, and luteo-/poleroviruses, nanoviruses and potyviruses. PEMV was the predominant virus found (70.5% of samples) but infection with luteo-/poleroviruses (26.7%), potyviruses (4.6%) and carlaviruses (0.9%) was confirmed. More importantly, 54.7% of samples tested positive with an ELISA designed for broad detection of nanoviruses (Grigoras *et al.*, 2010a; Abraham *et al.*, 2012). The same samples did not react with an ELISA designed to detect only *Faba bean necrotic stunt virus* and *Faba bean necrotic yellows virus*, indicating infection with PNYDV. Using PCR with PNYDV-specific primers priPeaSdir (5' AAC CTC CGG ATA TCA CCA GAT 3') and priPeaSrev (5' CCG GAG GTT TTA TTT CAA AAC CAA C 3') targeting the coat protein encoding component S of the genome (T. Timchenko, pers. comm.), PNYDV infection was confirmed for a subset of 18 samples. Sequencing of amplicons showed 98.7 to 99.9% nucleotide identity with PNYDV (GenBank accession no. JN133279). Three lentil (*Lens culinaris*) samples from a field trial in central Germany also tested positive for PNYDV using differentiating monoclonal antibodies with confirmation by PCR and sequencing. Sequences from this study can be accessed under accession numbers KY191024 - KY191044.

Figure 1: Faba bean crop with a typical patch of virus-infected plants. In particular, yellowing, dwarfing and leaf deformation on the upper parts of the plants indicate nanovirus infection.

Figure 2: Faba bean samples with typical symptoms: top yellowing of leaves, leaf rolling and dwarfed appearance of leaves and top part of the plant.

In Austria, nanovirus symptoms appeared first in *P. sativum* in early June 2016 and shortly after in faba bean. In mid-late June, nearly every *V. faba* crop showed typical symptoms of nanovirus infection. In many faba bean and pea crops infection caused significant yield losses (Fig. 3). Typical symptoms of stunted growth, chlorosis and poorly developed pods were also found in lentils and vetch (*V. sativa*). Thirty-two samples of *L. culinaris*, *P. sativum*, *V. faba* and *V. sativa* from Burgenland, Styria and Upper and Lower Austria were tested for nanovirus infection using PCR primers designed by (Kumari *et al.*, 2010). The samples consisted of leaves pooled from several symptomatic plants from each field. Twenty-seven samples were positive for nanovirus infection. Representative amplicons from faba beans, lentils, peas, and vetch were sequenced (KY191009 - KY191023) and had 99.6 to 100% identity to PNYDV (KC979043).

Figure 3: Faba bean crop heavily infected with nanoviruses.

This is the first report of *L. culinaris*, *V. faba* and *V. sativa* as natural hosts of PNYDV in Austria and Germany. Due to changes in government policy, the area of legumes grown in Germany doubled from 2012 to 2015 (Table 1) with further increases expected. However, limited host range experiments on peas and faba beans have not identified PNYDV-resistant accessions in Austria or Germany suggesting that legume production in central Europe is threatened by PNYDV infection.

Table 1: Acreages (ha) of selected legumes grown in Germany and Austria in 2012 and 2015.

Crop	Germany [a]		Austria [b]	
	2012	2015	2012	2015
Glycine max	5,000	11,000	36,955	56,867
Pisum sativum (protein peas only)	44,800	79,100	10,700	7,183
Vicia faba	15,800	37,700	6,854	10,822
Total	65,600	127,700	54,509	74,872

[a] Anonymous, 2016. Eiweisspflanzenstrategie. http://www.bmel.de/DE/Landwirtschaft/Pflanzenbau/Ackerbau/_Texte/Eiweisspflanzenstrategie.html, (Accessed 07 August 2016)
Burghardt B, Schaack D, Von Schenck W, 2016. *AMI Markt Bilanz Getreide Ölsaaten Futtermittel 2016*. In. Bonn: Agrarmarkt Informations-Gesellschaft mbH, 225 pages. www.AMI-informiert.de (Accessed 07.08.2016)
[b] Agrar Markt Austria, 2016. *Flächenauswertung der Mehrfachanträge (MFA)*. In. Wien: Agrar Markt Austria. www.ama.at/Marktinformationen/Getreide-und-Olsaaten/Aktuelle-Informationen/2016/AMA-Flaechenauswertung-2016 (Accessed 07 August 2016)

3.10. First report of pea necrotic yellow dwarf virus in The Netherlands

Yahya Zakaria Abdou Gaafar, T. Timchenko and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar Y, Timchenko T, Ziebell H, 2017. First report of *Pea necrotic yellow dwarf virus* in The Netherlands. *New Disease Reports* 35, 23. doi: 10.5197/j.2044-0588.2017.035.023

Pea necrotic yellow dwarf virus (PNYDV) is a nanovirus that was first detected in pea crops (*Pisum sativum*) in Saxony-Anhalt, Germany in 2009 (Grigoras *et al.*, 2010a). In 2016, PNYDV was detected countrywide in both Germany and Austria not only on pea but also on faba bean (*Vicia faba*), vetch (*V. sativa*) and lentil (*Lens culinaris*) causing severe yield losses (Gaafar *et al.*, 2016).

During a routine survey of twelve green pea crops in the Province of Flevoland (The Netherlands), plants with virus-like symptoms were noticed (Fig. 1). Symptomatic plant material was pooled from each field and analysed by ELISA for typical pea viruses: *Alfalfa mosaic virus*, *Cucumber mosaic virus*, luteo-/poleroviruses, *Pea enation mosaic virus* (PEMV), potyviruses, and *Red clover vein mosaic virus*-like carlaviruses, and nanoviruses. PEMV was detected in all fields while luteo-/poleroviruses were found in one field. Two samples each from different pea fields reacted positively using a broad nanovirus monoclonal antibody mixture (Gaafar *et al.*, 2016). The lack of reaction with a monoclonal antibody mixture designed to detect only *Faba bean necrotic stunt virus* and *Faba bean necrotic yellows virus* suggested infection with *Pea necrotic yellow dwarf virus* (PNYDV). This was confirmed by PCR using PNYDV specific primers targeting the eight PNYDV components producing bands of approximately 1 kb (Table 1). All PCR products were cloned using the NEB PCR cloning kit (New England Biolabs, Germany) and at least four clones for each component were sequenced in both directions. The sequences of the eight components of the two Dutch isolates (NL HZ16-186 and NL HZ16-189) had between 96.7 and 99.9% identity with the equivalent PNYDV components of an isolate from Germany and between 96.7 and 99.8% with an Austrian isolate (Table 1). The sequences of the Dutch PNYDV isolates have been deposited in GenBank (KY593279 - KY593294).

Figure 1: Non-infected (left) and virus-infected (right) pea plants. Typical Pea necrotic yellow dwarf virus symptoms included severe stunting and dwarfing of plants, yellowing and leaf-rolling.

Table 1: List of the primers used for *Pea necrotic yellow dwarf virus* identification and pairwise comparisons between the sequences of the Dutch isolates (NL HZ16-186 and NL HZ16-189) and isolates from Austria (GenBank Accession No. KC979043 - KC979050) and Germany (GU553134 and JN133279 - JN133285).

PNYDV component	Primer name	Primer sequence	Pairwise comparisons	References	NL HZ16-186	NL HZ16-189
					Identity %	
DNA-C	priPeaCdir	5' GCC GGA AGC TTG CCG GAC TGA CGG AG 3'		KC979045	99.2	99
	priPeaCrev	5' AGC TTC CGG CAA GAC GCA GTA ATT G 3'		JN133280	99.5	99.1
DNA-M	priPeaMdir	5' TAC CTG AAC GTC CTG TAT CTT 3'		KC979046	98.7	98.3
	priPeaMrev	5' TCA GGT ACT GAA TTA CTT GCC 3'		JN133281	98.3	97.4
DNA-N	priPeaNdir	5' GAA GAA CCC AGG AAG GTG TTG C 3'		KC979047	99.4	98.9
	priPeaNrev	5' GGT TCT TCC AAT TTA CCT TTC ATG G 3'		JN133282	99.9	99.2
DNA-R	priPeaRdir	5' GGA ATA CCA AGG TGA GTT ACG G 3'		KC979043	99.8	99.7
	priPeaRrev	5' TAT TCC CTG AGA GTC CCG GAC 3'		GU553134	99.8	99.5
DNA-S	priPeaSdir	5' AAC CTC CGG ATA TCA CCA GAT 3'		KC979044	99.3	98.7
	priPeaSrev	5' CCG GAG GTT TTA TTT CAA AAC CAA C 3'		JN133279	99.3	99.8
DNA-U1	priPeaU1dir	5' TGG TGA AGA AAT TGC AGG TGA T 3'		KC979048	98	98.7
	priPeaU1rev	5' TTC ACC AGT TTC TCG TCA GAA C 3'		JN133283	98.3	98.8
DNA-U2	priPeaU2dir	5' GAT CAA GAA CAA GGT TAG TTT ATG 3'		KC979049	98.2	96.7
	priPeaU2rev	5' TCT TGA TCG GAG ACG AAC TGG A 3'		JN133284	98.2	96.7
DNA-U4	priPeaU4dir	5' ATC AAG TCT GAA GAT GAT ACG 3'		KC979050	99.1	99.3
	priPeaU4rev	5' GAC TTG ATT TCA ACA TCT CTT TCA C 3'		JN133285	99.8	99.7

To our knowledge, this is the first report of PNYDV in The Netherlands. This indicates that nanoviruses are far more spread throughout Europe than previously thought (Grigoras *et al.*, 2014). As PNYDV is aphid-transmitted in a circulative, non-propagative manner, it is expected that more nanovirus diseases will occur in the future as changes in climatic conditions (especially milder winters in Central Europe) favour aphid survival thus facilitating the spread of these viruses (Ziebell, 2017).

3.11. Molecular characterisation of the first occurrence of pea necrotic yellow dwarf virus in Denmark

Yahya Zakaria Abdou Gaafar, G. Cordsen Nielsen and H. Ziebell

This article has been published in a slightly modified version as:

Gaafar Y, Cordsen Nielsen G, Ziebell H, 2018. Molecular characterisation of the first occurrence of *Pea necrotic yellow dwarf virus* in Denmark. *New Disease Reports* 37, 16. doi: 10.5197/j.2044-0588.2018.037.016

Pea necrotic yellow dwarf virus (PNYDV), a member of the *Nanovirus* genus, has been reported from numerous European countries causing yield loss in peas, faba beans, vetches and lentils (Grigoras *et al.*, 2010a; Grigoras *et al.*, 2014; Gaafar *et al.*, 2016; Gaafar *et al.*, 2017). In July and August 2016, five faba bean (*Vicia faba*) samples were received from five fields in different regions of Denmark, in which 2 to 25% plants were diseased. Affected plants displayed leaf-rolling, yellowing, and symptoms of severe stunting (Figs. 1-2).

Figure 1: Symptoms of *Pea necrotic yellow dwarf virus* in a faba bean field near Åbenrå, Denmark (photograph courtesy Morten Steg).

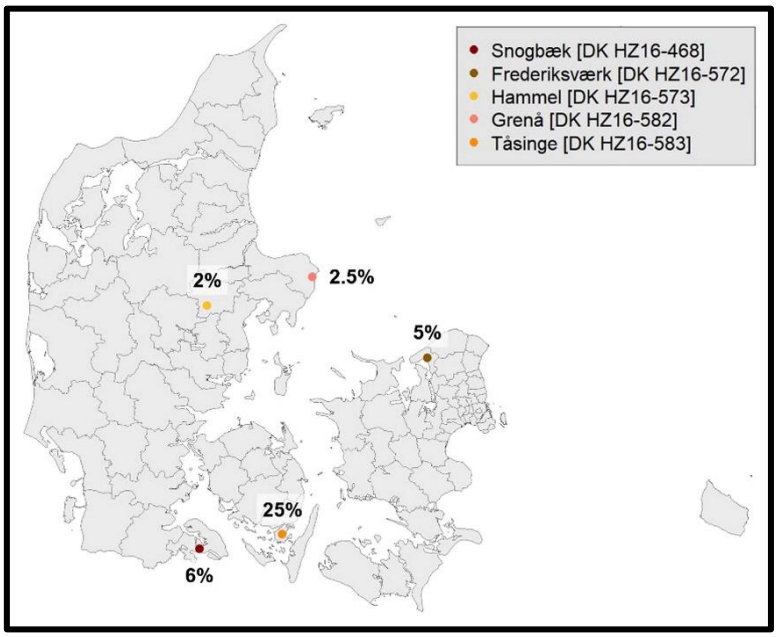

Figure 2: Map of *Pea necrotic yellow dwarf virus*-infected faba bean fields in Denmark in 2016. The coloured circles indicate the locations of the fields. The percentages represent the proportion of symptomatic plants within each field. The map was generated using MapDK (Barfort, 2015. mapDK: Maps of Denmark. R Package Version 0.3.0) and ggplot2 (Wickham, 2009. ggplot2: Elegant

Graphics for Data Analysis. New York, USA: Springer-Verlag) on R 3.3.2 (R Project for Statistical Computing; https://www.r-project.org).

The samples reacted positively when analysed with ELISA using antibodies against PNYDV (Gaafar *et al.*, 2016). DNA was extracted and PNYDV infection was confirmed by PCR using primer pairs specific for components R and S (Gaafar *et al.*, 2017). Additionally, the full genomes of these Danish PNYDV isolates were determined. Rolling circle amplification was performed using a TempliPhiTM 100 Amplification Kit (GE Healthcare Limited, UK). Libraries were prepared from the products and submitted for high-throughput sequencing on two Illumina platforms, three samples on HiSeq (2*150) and two on MiSeq (2*301). The paired-end reads were analysed using Geneious 11.0.4 software (Biomatters Limited, NZ). The reads were trimmed, error corrected and normalised, then used for de novo assembly. Assembled contigs were compared with the GenBank non-redundant nucleotide and protein databases using BLASTn and BLASTx, respectively.

The results confirmed the presence of all eight PNYDV components in all samples. Additionally, an alphasatellite 3 sequence was found associated with the DK HZ16-572 and DK HZ16-582 isolates, and a satellite DNA-X sequence was found associated with DK HZ16-572. The PNYDV genome Drohndorf-15 (Germany; GenBank Accession Nos. NC_023154 to NC_023161) was used to assemble the eight components of the Danish isolates and Austrian sequences (Gross-Enzernsdorf_1) for alphasatellite 3 and satellite DNA-X (KC979052 and KC979053, respectively) by mapping to reference. The sequences of the Danish PNYDV isolates can be accessed on GenBank (MH000227 to MH000269).

A neighbour-joining tree (Jukes-Cantor genetic distance model and 1000 bootstrap replicates) was created using Geneious Tree Builder (Fig. 3) and the alignment (ClustalW 2.1) of the concatenated genome components (DNA-R, -S, -C, -M, -N, -U1, -U2 and -U4) from the Danish PNYDV isolates as well as the Austrian, Dutch and German isolates. Additionally, a pairwise comparison between the associated alphasatellite 3 of DK HZ16-572 and DK HZ16-582 showed 99% nucleotide identity to each other, and 99.1% and 99.3% identity when compared with the Gross-Enzersdorf_1 isolate, respectively. Finally, DNA-X of DK HZ16-572 shared 98.8% identity with Gross-Enzersdorf_1. The tree and the sequence alignments indicate that the Danish isolates are very closely related to other European isolates of PNYDV.

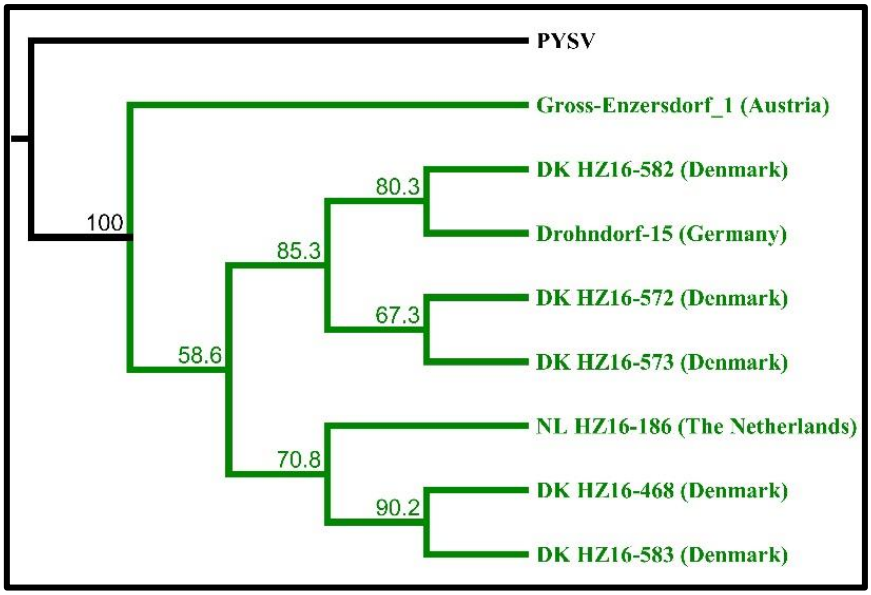

Figure 3: Phylogenetic neighbour-joining tree representing the relationship between the concatenated genome components (DNA-R, -S, -C, -M, -N, -U1, -U2 and -U4) of Danish Pea necrotic yellow dwarf virus isolates and those from Austria (KC979043 - KC979050), Germany (NC_023154-NC_023161) and The Netherlands (KY593279 - KY593286). Pea yellow stunt virus (PYSV) (NC_023296 - NC_023298, NC_023303 and NC_023308 - NC_023311) was used as an outgroup.

To our knowledge, this is the first report of a nanovirus disease in Denmark demonstrating that PNYDV is also widespread throughout Denmark. This is further evidence that PNYDV (and possible other nanoviruses) is an emerging threat for legumes not only in Denmark which has an increasing area of legume production (Table 1) but also in other European countries.

Table 1: Areas of cultivated legumes (hectares) in Denmark from 2014 to 2017 (Statistics Denmark, 2017. https://www.statbank.dk/. Accessed on 7-3-2018).

Year	2014	2015	2016	2017
Total legumes	8,400	11,900	15,700	20,600
Faba bean	3,900	6,900	10,700	14,800
Pea	4,200	4,700	4,800	5,400

Chapter 3: Investigating the pea virome in Germany – old friends and new players in the field(s)

Yahya Zakaria Abdou Gaafar, K. Herz, J. Hartrick, J. Fletcher, A. Blouin, R. MacDiarmid and H. Ziebell

3.1. Abstract

Peas are important legumes for human and animal consumption and are also being used as green manure or intermediate crops to sustain and improve soil condition. Pea production faces different constraints by fungi, bacteria pests and viral diseases. We investigated the virome of German pea crops over the course of three successive seasons in different regions of pea production in order to get an overview of the existing viruses. Pools from 540 plants randomly selected from symptomatic and asymptomatic pea, and non-crop plants surrounding the pea fields were used for ribosomal RNA-depleted total RNA extraction followed by high-throughput sequencing and RT-PCR confirmation. Thirty-five different viruses were detected in addition to eight associated nucleic acids. From these viruses, 25 are classified as either new viruses, novel strains or viruses that have not been reported previously from Germany. Pea enation mosaic virus (PEMV) was the most prevalent virus detected in the pea crops followed by turnip yellows virus (TuYV) which was also found in the surrounding non-legume weeds. Moreover, a new emaravirus was detected in peas in one region for two successive seasons. The results revealed a high virodiversity in the German pea fields that poses new challenges to diagnosticians, researchers, risk assessors and policy makers as the impact of the new findings are currently unknown.

3.2. Introduction

Green peas (*Pisum sativum* L.) are popular vegetables in Germany. The production of green peas increased from 4,444 ha in 2010 to 5,488 ha in 2018 (Behr, 2015, 2019). In addition, due to the "Protein Strategy" of the Federal Government of Germany, the production areas of protein peas used as animal fodder, green manure or as intermittent crops, increased from 57,200 ha in 2010 to 70,700 ha in 2018 (BMEL, 2019). However, depending on the intended use of the crop, pea production in Germany is highly regionalised. The main green pea production areas are located in Saxony due to the nearby frozen foods processing facilities. Seed production of peas is predominantly carried out in Saxony-Anhalt. By contrast, green pea production for the fresh market or protein pea production for animal fodder/green manure are scattered around the country often associated with small scale or organic farming.

Pea plants are known to be hosts to several viruses from different families e.g. *Luteoviridae*, *Nanoviridae* and *Potyviridae*, that often occur in mixed infections (Musil, 1966; Kraft, 2008; Bos *et al.*, 1988; Gaafar *et al.*, 2016; Gaafar & Ziebell, 2019b). Due to the emergence of novel pea-infected viruses such as pea necrotic yellow dwarf virus and their subsequent detection across Germany and within neighbouring countries (Grigoras *et al.*, 2010a; Gaafar *et al.*, 2016; Gaafar *et al.*, 2017; Gaafar *et al.*, 2018a), we were interested to know whether more unknown and or previously undetected viruses were present in this high value crop growing in Germany.

Conventional virus diagnostics generally depend on serological or molecular methods based on prior knowledge of the target virus. These specific tests do not address the potential presence of other viruses that may be present or contribute to the aetiology of a disease. In recent years, high-throughput sequencing (HTS) enabled identification of numerous new viruses from domesticated and wild plants (Roossinck *et al.*, 2015; Gaafar *et al.*, 2019f; Gaafar *et al.*, 2019e). HTS allows sequencing of all the genetic material in a given sample, therefore there is no need for prior knowledge of the infectious agent (Adams *et al.*, 2009; Roossinck *et al.*, 2015; Maree *et al.*, 2018). Improvements of HTS technologies and bioinformatic tools have helped to identify the virus community or virome of several crops (Coetzee *et al.*, 2010; Czotter *et al.*, 2018; Jo *et al.*, 2018b) The generated data permits the description of plant virus biodiversity, discovery of new viruses and viroids, identification of genomic variants of the viral species and aids the development of specific and sensitive diagnostics (Coetzee *et al.*, 2010; Gaafar *et al.*, 2019c; Li *et al.*, 2012; Gaafar *et al.*, 2019d). However, the vast number of new viruses identified by HTS results in challenges for diagnosticians, pest risk assessors and policy makers as the risk to crop plants and alternative hosts by these new viruses need to be evaluated and diagnostic protocols developed and validated or adapted to these findings (MacDiarmid *et al.*, 2013; Massart *et al.*, 2017; Rott *et al.*, 2017; Maree *et al.*, 2018). Nevertheless, metagenomics data and biological studies will help us understand the viral ecology and evolution as well as epidemiology.

In this study, we were interested in the spatio-temporal description and changes of the pea virome in selected German regions with different pea production aims (fresh produce, frozen produce, seed production, etc.) over a period of three years. Furthermore, we investigated potential alternative virus reservoirs in terms of legume and non-legume groundcover that were associated with the production sites. To our knowledge this is the first metagenomics study of a crop plant that takes spatio-temporal changes into consideration.

3.3. Material and Methods

3.3.1. Sampling

Six green pea production regions in Germany were chosen for sampling as they included sites for human consumption, protein pea production sites, seed production sites as well as experimental trial sites. The regions named by their county designation were Salzlandkreis-1: pea seed production, Salzlandkreis-2: trial site heritage material, Münster: trial site pea breeding, Kreis Stormarn: protein pea production, Landkreis Rostock: trial site (green manure mixtures) and Landkreis Meißen: green pea production frozen produce for human consumption.

In each region, one typical production field was chosen randomly. Over three successive seasons (2016, 2017 and 2018 between June and July; Figure 1) these regions were sampled. From each field, ten symptomatic (showing virus-like symptoms) pea

plants (SP) and ten asymptomatic pea plants (aSP), five surrounding non-crop legume plants (sL) and five non-legume plants (snL) were collected where possible. The metadata were recorded, i.e., the symptoms of each plant sampled, the region, and the average temperature of the season for each region (Supplementary Table S1). Any deviations from the sampling strategy (i.e., in cases where no non-symptomatic peas or no surrounding legume plants could be detected) were also noted in (Supplementary Table S1).

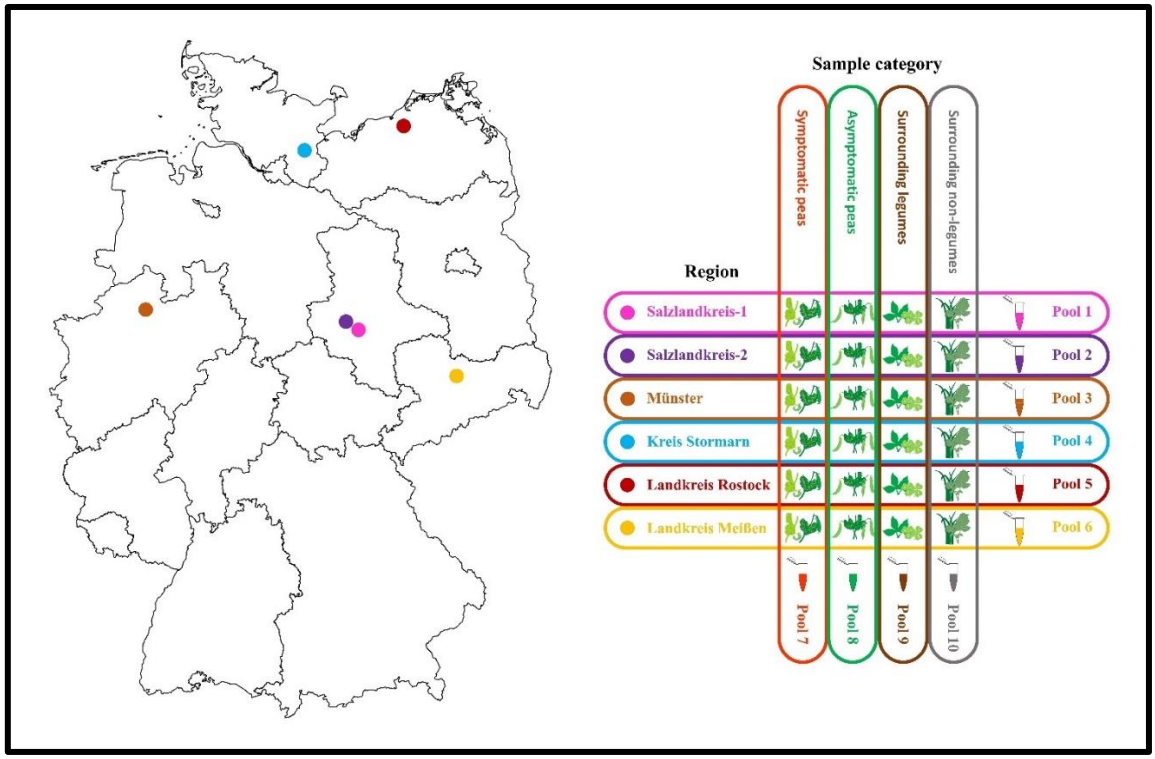

Figure 1: Map of the six different pea crop sampling regions in Germany and the sample pooling strategy used for each of three successive seasons 2016, 2017 and 2018.

Ten sample pools were prepared for each season i.e., six separate pools containing material from each of the six regions and four separate pools comprising material from all SP, aSP, sL or snL samples, respectively, from all regions (Figure 1). From each plant sample, 100 mg of fresh tissue was added to each pool. The samples were mixed and ground using a mortar and pestle under liquid nitrogen then collected into 50 ml Falcon tubes. The pools were stored at -20°C until RNA extraction. Additionally, the remaining material of fresh plant samples were stored at -20°C for further analysis.

3.3.2. RNA extraction and high throughput sequencing

Total RNA was extracted using innuPREP Plant RNA Kit (Analytik Jena) from a subsample of 300 µg from each pool. Ribosomal RNA (rRNA) was depleted using RiboMinus™ Plant Kit for RNA-Seq (Invitrogen). cDNA was synthesized using ProtoScript II Reverse Transcriptase (NEB) and random octanucleotide primers (8N), followed by second strand synthesis using NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module (NEB). The libraries were prepared from the double-stranded cDNA using Nextera XT Library Prep Kit (Illumina). The sequencing was performed on an Illumina MiSeq platform (301 x 2).

3.3.3. Bioinformatic analysis

Bioinformatic analysis was performed using Geneious Prime software (version 2019.1.1). The reads were quality trimmed and normalised. *De-novo* assembly was performed and the resulting contigs were searched against a local database of viruses and viroids sequences downloaded from NCBI using Blastn and Blastx (downloaded 13 August 2018). The generated consensus sequences were based on the highest quality threshold. Primers for virus validation were designed using a modified version of Primer3 (2.3.7) tool in Geneious Prime (Untergasser *et al.*, 2012). Pairwise alignments were performed using Clustal W tool (v 2.1) in Geneious (Larkin *et al.*, 2007). Neighbour joining phylogenetic trees were constructed using MEGA X software (Kumar *et al.*, 2018b). The phylogenetic relationships were established according to the species demarcation criteria set by International Committee on Taxonomy of Viruses (ICTV), using the nucleotide sequences or the amino acid sequences of the capsid protein (CP) or the RNA dependent RNA polymerase (RdRP) for the respective families. The isolates were named by region number and season e.g., R1_16 stands for region one and the season 2016. The assembled virus sequences can be accessed in GenBank under accession nos. (MN314973, MN399680-MN399748, MN412725-MN412751 and MN497793-MN497846).

3.3.4. RT-PCR confirmations

For virus confirmation, total RNA was re-extracted from each pool as described above, followed by RT-PCR with the primers listed in Supplementary Table S2 and using the OneTaq One-Step RT-PCR Kit (NEB). The products were purified using Zymoclean Gel DNA Recovery Kit (Zymo Research) and Sanger sequenced using both RT-PCR primers.

3.3.5. Statistical analysis

Statistical analysis was performed using scripts written on R (version 3.5.3) (R Core Team, 2019) The virus relative abundance (VRA) was calculated for each virus. VRA = the sum of the virus detections in each region in all season divided by the total of possible detection (3 seasons x 6 regions x 4 categories =72). The virus relative abundance was 0 for not detected to 1 for detected in all categories in all region in all seasons. Additionally,

the virus incidence was calculated for the regions and the seasons. The Venn diagram was generated using Venny tool (Oliveros, 2015).

3.4. Results

3.4.1. Collected metadata and HTS raw data

General symptoms observed on peas were typical of PEMV such as translucent spots and leaf enations, stunting of top leaves, dwarfed plants, severe yellowing, mottling and leaf rolling. Interestingly, in 2016 it was not possible to find legumes in fields surrounding Salzlandkreis-1 and no asymptomatic peas were found in Salzlandkreis-2. Moreover, no surrounding legumes were collected in Landkreis Meißen 2018. Therefore, they were not included in the analyses. The details of the raw data generated from the HTS MiSeq platform are in Supplementary Table S3.

3.4.2. Assignment of identified virus families

A total of thirty-five viruses were detected by HTS and confirmed by RT-PCR in the different pools over the three seasons representing 14 different families in addition to several unassigned viruses (Figure 2a and b). The family *Luteoviridae* was represented by seven species members, followed by the *Secoviridae* with six species members, then the *Potyviridae* with five members. During the three seasons, the highest virus diversity was in the SP pools with sixteen different virus species present, while the sL contained twelve different virus species. The snL pools included eleven virus species and the aSP had seven virus species (Figure 2 a and b).

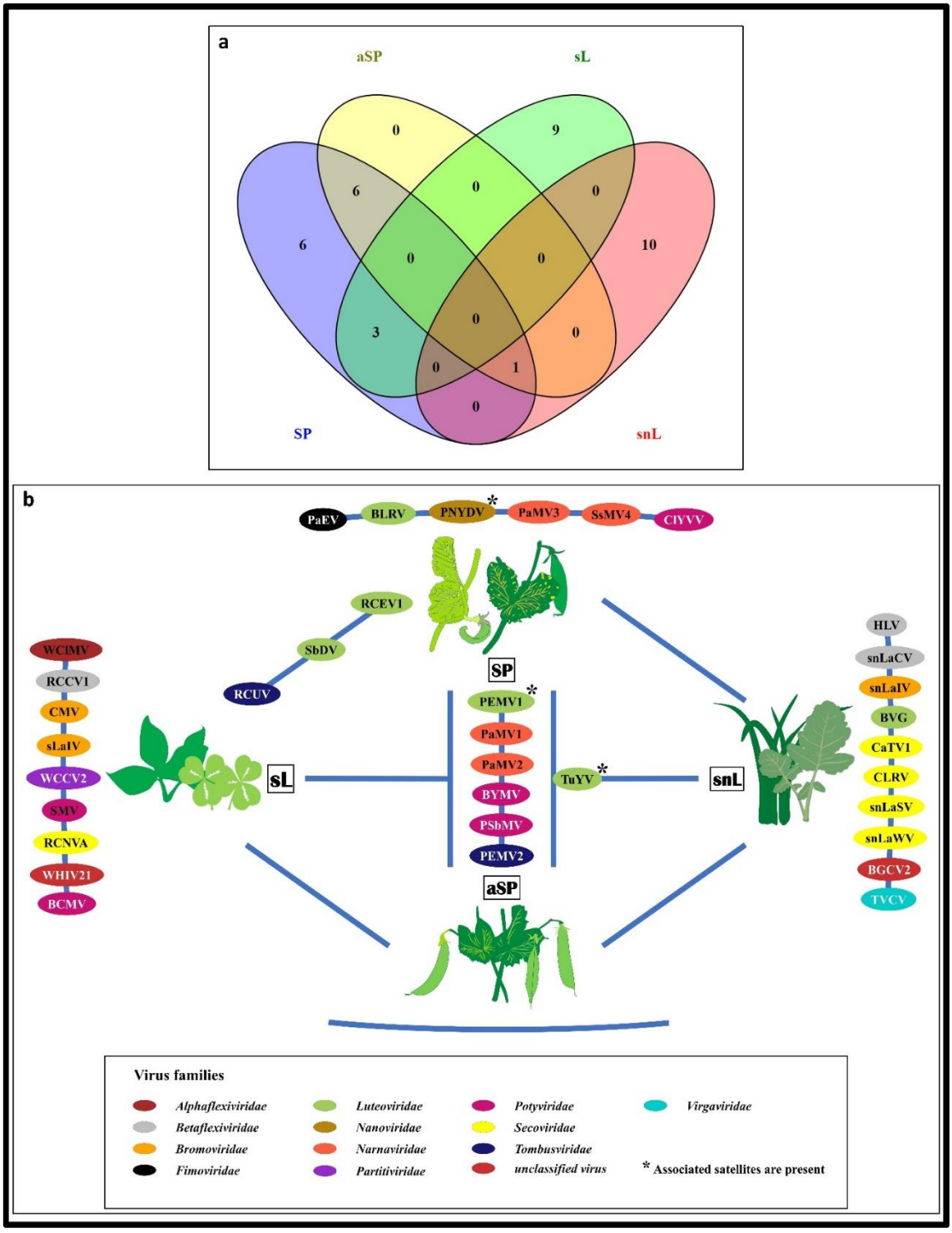

Figure 2: The virus species found in the four sampling categories symptomatic peas (SP), asymptomatic peas (aSP), surrounding legumes (sL) and surrounding non-legumes (snL). Samples were collected in six pea growing regions in Germany over three successive seasons 2016, 2017 and 2018. a) Venn diagram representing the number of virus species to each category and shared between the four categories. b) network illustration of virus species of each category and shared viruses between them. Background colours of virus acronyms correspond to the background colours of the respective virus family. The virus names are: BVG: barley virus G, BCMV: bean common mosaic virus, BLRV: bean leafroll virus, BYMV: bean yellow mosaic virus, BGCV2: black grass cryptic virus 2, CaTV1: carrot torradovirus 1, CLRV: cherry leaf roll virus, ClYVV: clover

yellow vein virus, CMV: cucumber mosaic virus, HLV: Heracleum latent virus , PaMV1: pea associated mitovirus 1, PaMV2: pea associated mitovirus 2, PaMV3: pea associated mitovirus 3, PaEV: pea associatted emaravirus, PEMV1: pea enation mosaic virus 1, PEMV2: pea enation mosaic virus 2, PNYDV: pea necrotic yellow dwarf virus, PSbMV: pea seed-borne mosaic virus, RCCV1: red clover carlavirus 1, RCEV1: red clover enamovirus 1, RCNVA: red clover nepovirus a, RCUV: red clover umbravirus, SsMV4: *Sclerotinia sclerotiorum* mitovirus 4, sLaIV: sL associated ilarvirus, snLaCV: snL associated chordovirus, snLaIV: snL associated ilarvirus, snLaSV: snL associated secoviridae, snLaWV: snL associated waikavirus, SbDV: soybean dwarf virus, SMV: soybean mosaic virus, TVCV: turnip vein-clearing virus, TuYV: turnip yellows virus, WCCV2: white clover cryptic virus 2, WClMV: white clover mosaic virus, WHIV21: Wuhan insect virus 21.

All identified viruses, the regions in which they were found and the pools from which they were identified are listed in Supplementary Table S4. The virus relative abundance (VRA) was calculated for each virus (Supplementary Table S4), showing that turnip yellows virus (TuYV) was the most abundant virus with a score of 0.472 followed by pea enation mosaic virus 2 (PEMV2) with a score of 0.444 and PEMV1 scoring 0.417. Pea necrotic yellow dwarf virus (PNYDV) was the fourth most abundant virus in the study with a score of 0.153 and pea seed borne mosaic virus (PSbMV) was the fifth most abundant one with 0.139.

Out of the viral reads identified from HTS, Figure 3 shows the percentage of each virus family in each region pool during the three seasons. Overall, the read percentage of the luteovirids was the highest, mainly because of high percentage of reads assigned to PEMV1 and TuYV. The second highest percentage was for the tombusvirids, due to the presence of PEMV2 reads, followed by the potyvirids, mainly because of PSbMV.

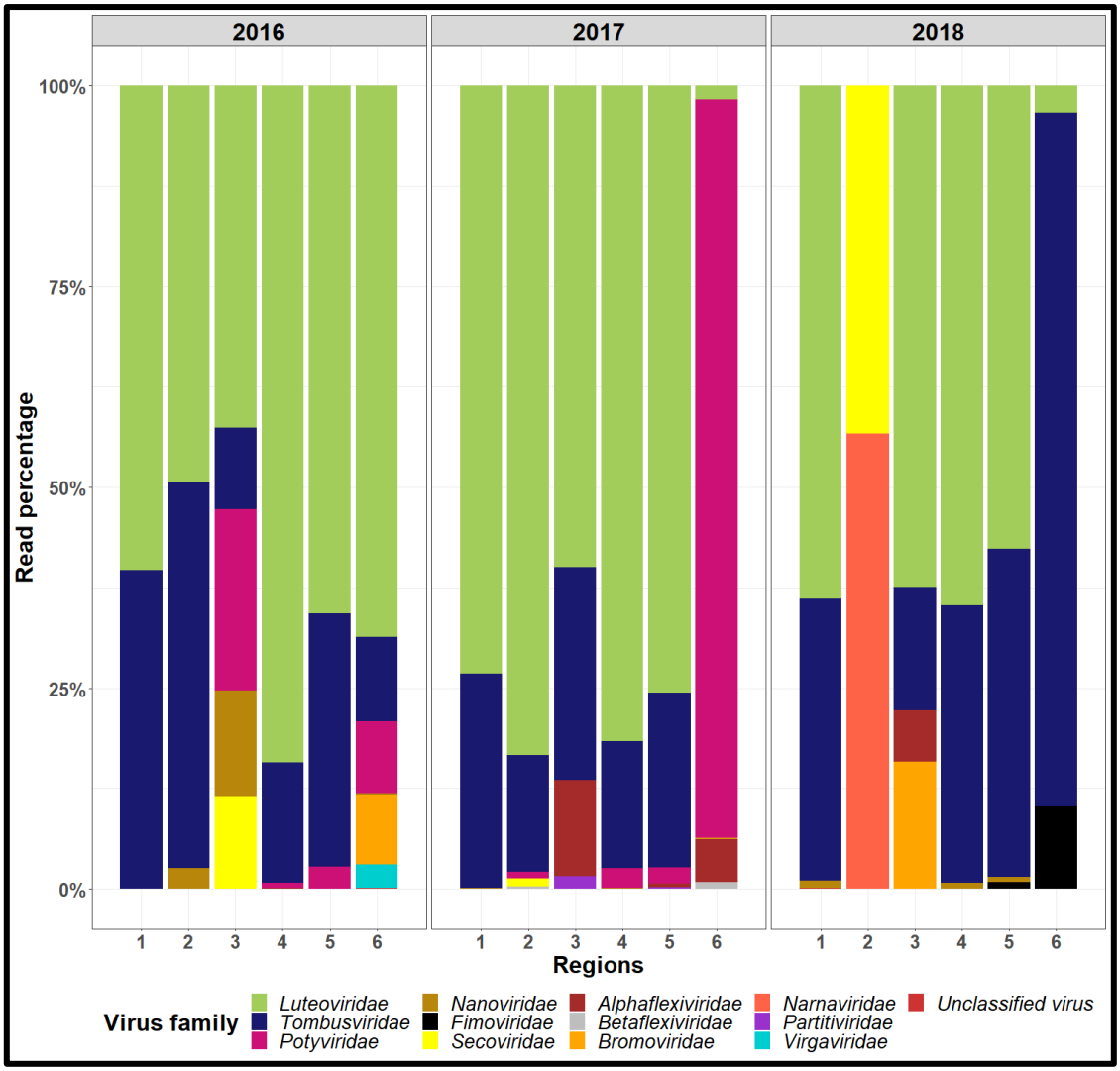

Figure 3: The read percentage of the different virus families identified in six different German pea growing regions over the three seasons 2016, 2017 and 2018.

3.4.3. Detailed description of the viruses detected in German pea fields

3.4.3.1. Viruses of the *Luteoviridae* family

Pea enation mosaic virus-1 (PEMV1) is a member of the genus: *Enamovirus*, family: *Luteoviridae*. PEMV1 was detected in all fields in almost all the three seasons in SP and aSP. The complete coding sequences of the PEMV1 isolates from each of the positive pools were assembled. The amino acid (aa) sequences of the RNA dependant RNA polymerases (RdRp) (used for taxonomic classification of luteovirids) shared between 96.9% to 98.4% aa identity with the RdRp of PEMV1 (NC_003629). They shared between 99.6% and 99.4% aa identity to each other.

In Münster 2016, a new strain of red clover enamovirus 1 (RCEV1) was identified in the sL samples. Its RdRp aa sequence showed 86.6% identity to a Czech strain (MG596229). Another isolate was detected in Landkreis Meißen in 2017, with 87.6% aa identity to the other German strain and 95.3% to the Czech strain. The viruses grouped together within the *Enamovirus* genus clade (Figure 4).

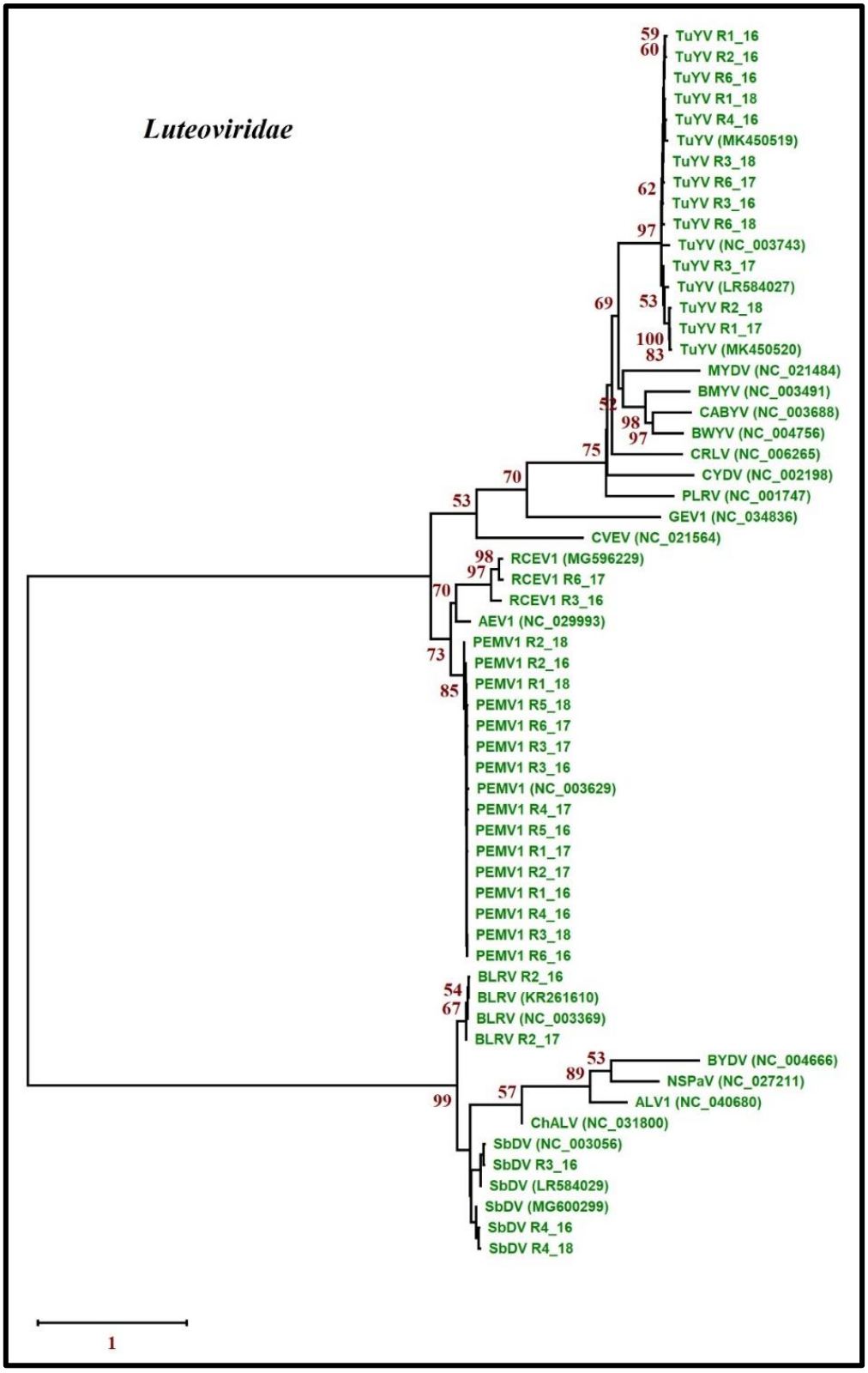

Figure 4: Neighbour joining trees (NJ) of virus isolates from the *Luteoviridae* family detected in German pea fields. The phylogenetic tree is based on amino acid sequence alignments of the readthrough RdRps and includes representatives of other viruses from the family *Luteoviridae*. Amino acid sequences were aligned with Clustal W and NJ trees constructed with MEGA X. The percentage of the bootstrap values above 50% (1,000 replications) are shown at the nodes. The names of the viruses are as follow: AEV1: alfalfa enamovirus 1, ALV1: apple luteovirus 1, BLRV:

bean leafroll virus, BMYV: beet mild yellowing virus, BYDV: barley yellow dwarf virus, BWYV: beet western yellows virus, CABYV: cucurbit aphid-borne yellows virus, ChALV: cherry associated luteovirus, CRLV: carrot red leaf virus, CVEV: citrus vein enation virus, CYDV: cereal yellow dwarf virus, GEV1: grapevine enamovirus-1, MYDV: maize yellow dwarf virus, NSPaV: nectarine stem pitting-associated virus, PEMV1: pea enation mosaic virus 1, PLRV: potato leafroll virus, RCEV1: red clover enamovirus 1, SbDV: soybean dwarf virus and TuYV: turnip yellows virus. The Genbank accession numbers are mentioned in brackets.

Bean leafroll virus (BLRV) was detected in Salzlandkreis-2 for two successive seasons (2016 and 2017) in SP in addition to the sL pool in 2017. The isolates of BLRV were closely related to strain Manfredi from Argentina (KR261610); the RdRp aa sequences of the two isolates shared 99.1% identity to each other and 99.7% to the Argentinian strain.

Turnip yellows virus (TuYV) was detected in SP, aSP and snL pools over the three seasons. Pairwise aa alignments for the RdRp of the TuYV isolates showed that eight isolates were 95.6% to 97.1% identical to TuYV isolate JKI 29345 (MK450519). However, two isolates from Salzlandkreis-2 and Salzlandkreis-1 showed higher identities with 98.8% and 99.1% to isolate JKI 29344 (MK450520). Moreover, the isolate from Münster 2017 was closely related to TuYV (LR584027) with 95.3% aa identity.

Soybean dwarf virus (SbDV) was detected in peas and surrounding legumes in two regions i.e., Münster and Kreis Stormarn. The Kreis Stormarn SbDV RdRp aa sequences found in the survey were 98% aa identical to each other and shared closest aa identity to SDV-HZ1 isolate from the Czech Republic (MG600299) with 98.9% and 99.1% identity. In contrast, the isolate from Münster showed the closest identity to a SbDV isolate (LR584029) from Australia with 99.1% and only shared 89.8% and 90.2% identities to the isolates from Kreis Stormarn.

In addition, PEMV satellite RNAs (PEMVSatRNAs), TuYV associated RNAs (TuYVaRNAs) and a potential new associated RNA (TuYVaRNA2) were detected. The PEMV satellite RNA isolates had 95.6% to 97.9% nt identity to PEMVSatRNA (NC_003854, from the USA) and between 95.5% and 98.7% to each other. The TuYVaRNA isolates were 97.8% to 98.2% nt identical to TuYVaRNA isolate JKI 29345 (MK450521). The TuYVaRNA2 isolates were closely related to cucurbit aphid borne virus associated RNA (CABYVaRNA; KM486094, from the USA) with 79.8% to 81.9% nt identity.

3.4.3.2. *Tombusviridae* (subfamily: *Calvusvirinae*)

PEMV2 was associated with PEMV1 in almost all fields during the three seasons. The aa sequences of the RdRp of the PEMV2 isolates grouped together in a clade within the genus *Umbravirus* (Figure 4b). They showed 97.1% to 99.7% identity to each other and 93.3% to 94.1% identity to PEMV2 isolate (NC_003853) from USA. A new strain of red clover umbravirus (RCUV), a novel umbravirus found in red clover from the Czech Republic (Koloniuk, pers. comm.), was detected in the sL in Münster 2016 and the SP in Landkreis

Meißen 2017. The complete coding sequence (CDS) shared 87.1% nt identity and the aa sequence of the RdRp shared to 91.6% aa identity to this Czech isolate (MG596234).

3.4.3.3. *Potyviridae*

Five potyviruses were present in the pea fields as well as in surrounding leguminous weeds. Bean common mosaic virus (BCMV) was identified in Salzlandkreis-2 in 2017. The isolate shared 98.9% nt identity with to BCMV strain NL1 from The Netherlands (AY112735). Bean yellow mosaic virus (BYMV) was identified in Salzlandkreis-2 and Salzlandkreis-1 in 2018. The isolates shared 83.9% aa identity to each other and 89.4% to BYMV strain MB4 (NC_003492) from Japan. Clover yellow vein virus (ClYVV) was identified in Salzlandkreis-1 and Münster in 2018 in SP. Their polyproteins shared 95% aa identity to each other and 95.5% and 93%, respectively, to ClYVV strain No.30 (NC_003536, probably from Japan). Pea seed-borne mosaic virus (PSbMV) was found in two regions in Germany i.e., Salzlandkreis-2 and Kreis Stormarn in all three seasons of the study. The virus was detected in SP and in aSP, respectively. The polyproteins of the isolates had 97.7% to 100% aa identity to each other and had 98% to 99.7% aa identity to PSbMV strain DPD1 (NC_001671) from Denmark. Another potyvirus, soybean mosaic virus (SMV), could be detected in Salzlandkreis-2 in 2017. The Salzlandkreis-2 isolate of SMV was very closely related to strain G4 from South Korea with 99.5% aa identity (FJ640979). Figure 5 shows a neighbour-joining tree representing the relationship between the polyproteins of the different potyviruses identified in the survey.

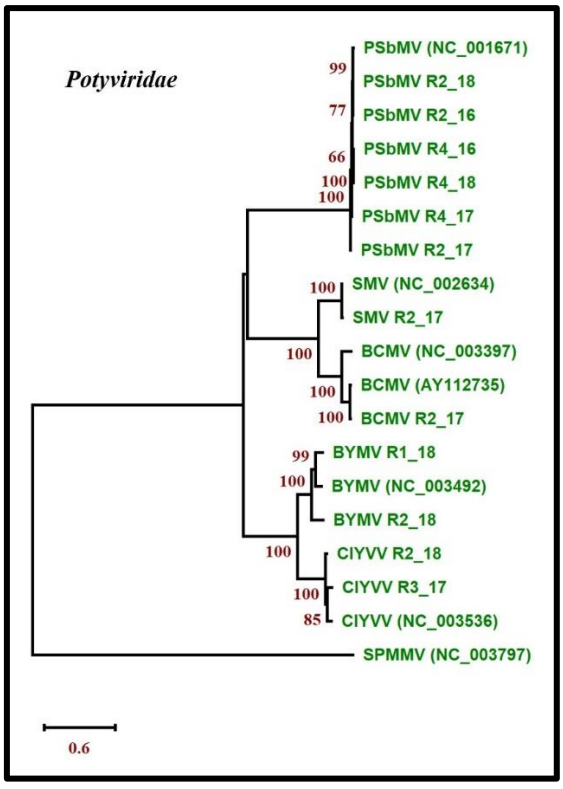

Figure 5: Neighbour joining trees (NJ) of *Potyviridae* virus isolates detected in German peas. The phylogenetic trees are based on amino acid sequence alignments of the polyproteins of the

potyviruses found in this survey and includes representative species from the family *Potyviridae*. Amino acid sequences were aligned using Clustal W and NJ trees were constructed with MEGA X. The percentage of the bootstrap values above 50% (1,000 replications) are shown at the nodes. The names of the viruses are as follow: BCMV: bean common mosaic virus, BYMV: bean yellow mosaic virus, ClYVV: clover yellow vein virus, PSbMV: pea seed-borne mosaic virus, SMV: soybean mosaic virus and SPMMV: sweet potato mild mottle virus. The Genbank accession numbers are mentioned in the brackets.

3.4.3.4. *Nanoviridae*: Pea necrotic yellow dwarf virus and its associated alphasatellites

Pea necrotic yellow dwarf virus (PNYDV) was only detected in symptomatic peas. There was no indication of the presence of other nanoviruses such as black medic leafroll virus or pea yellow stunt virus (include Grigoras et al. 2016 here). All DNA-S components shared between 97 and 99.9% nt identities with the on NCBI available PNYDV DNA-S sequences (originating from Germany and Austria). Additionally, six PNYDV alphasatellites (PNYDVαSat) were identified in the survey i.e., alphasatellites 1 and 3, and four new alphasatellites, tentatively named PNYDV-associated alphasatellites 4, 5, 6 and 7.

The PNYDVαSat1 isolate shared 99.6% nt identity with PNYDVαSat1 isolate Gross-Enzersdorf_1 (accession no. NC_038958) from Austria. The three isolates of PNYDVαSat3 shared between 99.3% and 99.4 % nt identities to PNYDVαSat3 isolates Gross-Enzersdorf_1 (Austria) and DK HZ16-582 from Denmark (NC_038959 and MH000253, respectively). The isolates of PNYDVαSat4, PNYDVαSat5, PNYDVαSat6 and PNYDVαSat7 were 1,030 nt, 991 nt, 1,037 nt and 1,015 nt in length, respectively. PNYDVαSat4 showed closest nt identities to faba bean necrotic yellows virus C7 alphasatellite (FBNYC7αSat; NC_003565) from Egypt and milk vetch dwarf virus C10 alphasatellite (MVDVC10αSat; NC_003647) from Japan with 82.3% to 82.9%. PNYDVαSat5 isolates shared the highest identities with 81.1% to 81.3% with parsley severe stunt virus alphasatellite 1 (PSSVαSat1; MK039138, from Germany). PNYDVαSat6 isolates had 82.8% identity to cow vetch latent virus alphasatellite (CvLVαSat; MF535455) from France and the PNYDVαSat7 isolate had 84.4% identity to the same CvLVαSat isolate. The phylogenetic relationships of those alphasatellites are shown in Figure 6a and b.

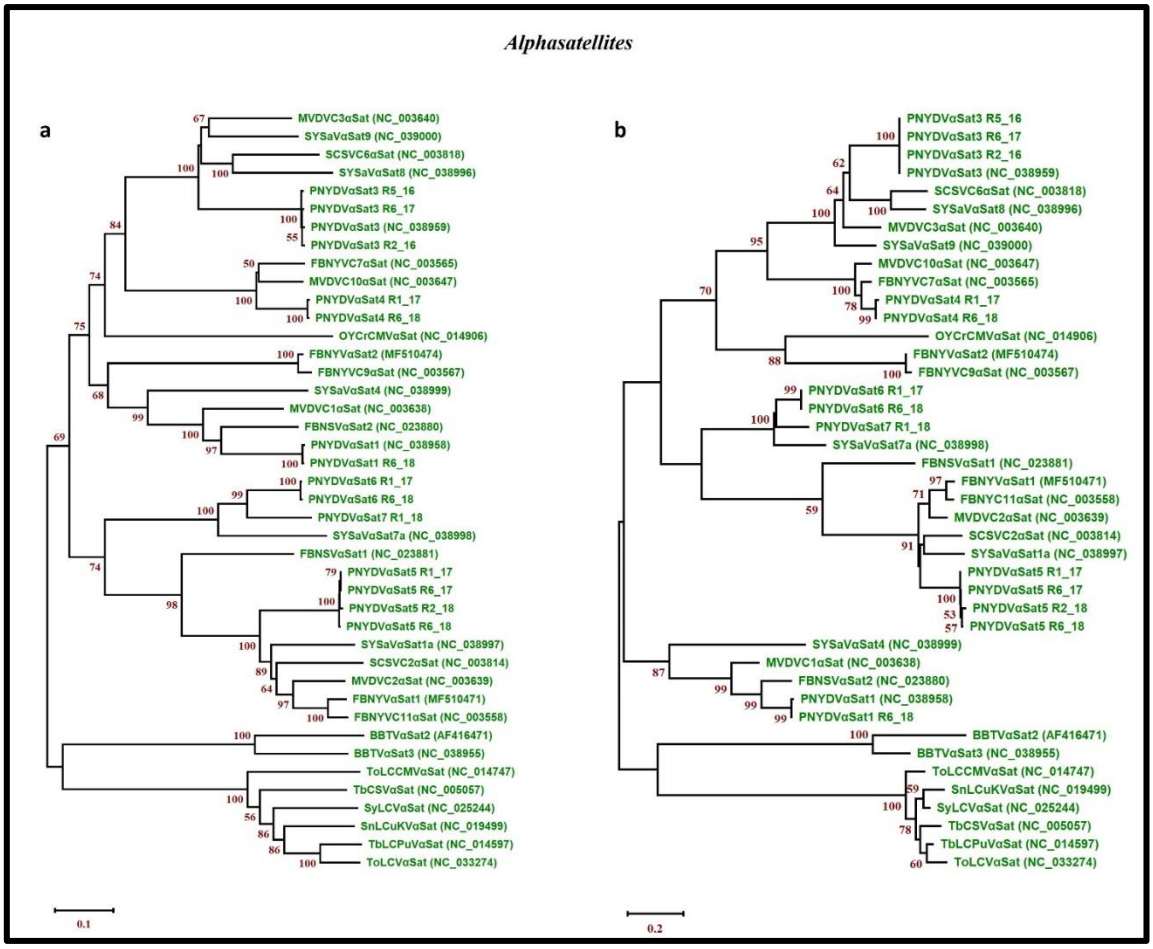

Figure 6: Neighbour joining trees (NJ) of PNYDV alphasatellites (PNYDVαSat) detected in German peas. The phylogenetic trees are based alignments of: **(a)** the nucleotide sequences **(b)** the amino acid sequences of the alphasatellite sequences identified in this study and include selected representatives of other alphasatellites. The sequences were aligned with Clustal W and NJ trees constructed with MEGA X. The percentage of the bootstrap values above 50% (1,000 replications) are shown at the nodes. The names of the alphasatellites are as follow: BBTVαSat2: banana bunchy top virus alphasatellite 2, BBTVαSat3: banana bunchy top virus alphasatellite 3, FBNSVαSat1: faba bean necrotic stunt virus alphasatellite 1, FBNSVαSat2: faba bean necrotic stunt virus alphasatellite 2, FBNYVαSat1: faba bean necrotic yellows virus alphasatellite 1, FBNYVαSat2: faba bean necrotic yellows virus alphasatellite 2, FBNYVC7αSat: faba bean necrotic yellows virus C7 alphasatellite, FBNYVC9αSat: faba bean necrotic yellows virus C9 alphasatellite, FBNYC11αSat: faba bean necrotic yellows virus C11 alphasatellite, MVDVC1αSat: milk vetch dwarf virus C1 alphasatellite, MVDVC2αSat: milk vetch dwarf virus C2 alphasatellite, MVDVC3αSat: milk vetch dwarf virus C3 alphasatellite, MVDVC10αSat: milk vetch dwarf virus C10 alphasatellite, OYCrCMVαSat: Okra yellow crinkle Cameroon virus alphasatellite, PNYDVαSat1: pea necrotic yellow dwarf virus alphasatellite 1, PNYDVαSat3: pea necrotic yellow dwarf virus alphasatellite 3, PNYDVαSat4: pea necrotic yellow dwarf virus alphasatellite 4, PNYDVαSat5: pea necrotic yellow dwarf virus alphasatellite 5, PNYDVαSat6: pea necrotic yellow dwarf virus alphasatellite 6, PNYDVαSat7: pea necrotic yellow dwarf virus alphasatellite 7, SCSVC2αSat: subterranean clover stunt virus C2 alphasatellite, SCSVC6αSat: subterranean clover stunt virus C6 alphasatellite, SnLCuKVαSat: alphasatellite, SyLCVαSat: Synedrella leaf curl virus alphasatellite, SYSaVαSat1a: Sophora yellow stunt associated virus alphasatellite 1a, SYSaVαSat4: Sophora yellow stunt associated virus alphasatellite 4, SYSaVαSat7a: Sophora yellow stunt associated virus

alphasatellite 7a, SYSaVαSat8: Sophora yellow stunt associated virus alphasatellite 8, SYSaVαSat9: Sophora yellow stunt associated virus alphasatellite 9, TbCSVαSat: tobacco curly shoot virus alphasatellite, TbLCPuVαSat: tobacco leaf curl Pusa virus alphasatellite, ToLCCMVαSat: tomato leaf curl Cameroon virus alphasatellite and ToLCVαSat: tomato leaf curl virus alphasatellite. The Genbank accession numbers are mentioned in the brackets.

3.4.3.5. *Fimoviridae*: Pea associated emaravirus (New virus)

A new emaravirus was identified in symptomatic peas in Landkreis Meißen for two successive season 2017 and 2018. The virus showed high similarity to other members of genus *Emaravirus* i.e., fig mosaic virus (FMV), pigeonpea sterility mosaic virus 2 (PPSMV2) and rose rosette virus (RRV). Members of the genus *Emaravirus* (family *Fimoviridae*, order *Bunyavirales*) have segmented, linear, single-stranded, negative-sense RNA genomes (Elbeaino *et al.*, 2018). Their genomes are composed of up to 8 segments. The partial sequences of segments RNA1 to 6 were assembled from the two isolates discovered from the two seasons. An NJ tree of the nucleocapsid protein (NP) aa sequences of the two isolates was constructed, and those isolates grouped together in a clade with PPSMV2 and FMV (Figure 7). The aa sequence of the NP shared the highest identity with PPSMV2 with 71.7%. According to the species demarcation of ICTV, a difference of 25% in the aa sequence of the NP indicates a new species (Elbeaino *et al.*, 2018). Therefore, this virus represents a new emaravirus species and was tentatively called pea associated emaravirus (PaEV).

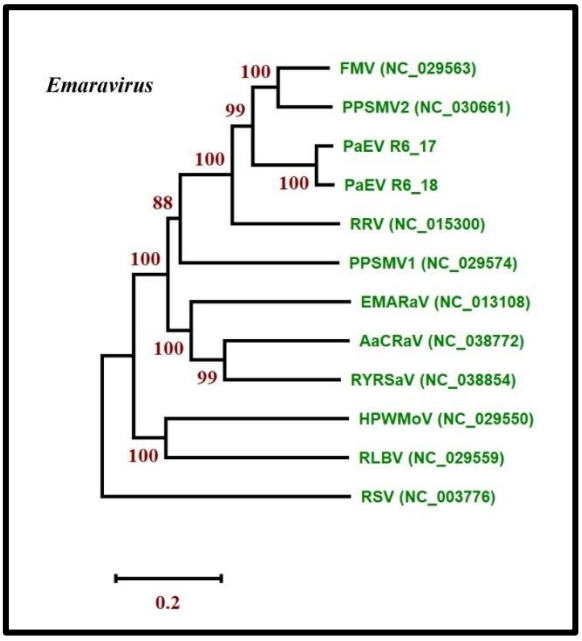

Figure 7: Neighbour joining trees (NJ) of the nucleocapsid protein (NP) of the two isolates of pea associated emaravirus (PaEV) and representative emaraviruses detected in German pea. The phylogenetic trees are based alignments of the amino acid sequences. The sequences were

aligned with Clustal W and NJ trees constructed with MEGA X. The percentage of the bootstrap values above 50% (1,000 replications) are shown at the nodes. The names of the viruses are as follow: AcCRaV: actinidia chlorotic ringspot-associated emaravirus, EMARaV: European mountain ash ringspot-associated emaravirus, FMV: fig mosaic emaravirus, HPWMEV: high plains wheat mosaic emaravirus, PaEV: pea associated emaravirus, PPSMV1: pigeonpea sterility mosaic emaravirus 1, PPSMV2: pigeonpea sterility mosaic emaravirus 2, RLBV: raspberry leaf blotch emaravirus, RRV: rose rosette emaravirus, RSV: rice stripe virus (*Tenuivirus, Phenuiviridae*) and RYRSaV: redbud yellow ringspot-associated emaravirus. The Genbank accession numbers are mentioned in the brackets.

3.4.3.6. *Narnaviridae:* Pea associated mitoviruses

In Landkreis Rostock, four mitoviruses (family: *Narnaviridae*) were detected in the SP and aSP in 2017. An isolate of Sclerotinia sclerotiorum mitovirus 4 (SsMV4) was also identified in the same region. The isolate shared 96.3% aa identity to SsMV4 from New Zealand (AGC24233). In addition, three mitoviruses provisionally called pea associated mitovirus 1, 2 and 3 (PaMV1, PaMV2 and PaMV3) were identified. The PaMV1-CDS shared 68.8% nt identity to its closest match Erysiphe necator mitovirus 3 (EnMV3; KY420040), that was identified from the grape powdery mildew fungus *Erysiphe necator* (Schwein.) (syn. *Uncinula necator* (Schw.)) that was described in the USA recently (Pandey *et al.*, 2018). Based on the RdRp region, PaMV1 shared 65.1% aa identity with EnMV3. PaMV2-CDS had 36% nt identity to Rhizoctonia solani mitovirus 6 (RsMV6; KP900915 from the USA). Based on the aa sequences, PaMV1 and RsMV6 share only 40.2% aa identity. PaMV3-CDS was closely related to Entomophthora muscae mitovirus 5 (EnmuMV5; MK682524) with 40% nt identity and 24.3% aa identity in the RdRp region.

3.4.3.7. *Secoviridae*

In Münster 2017, a new putative member of the family *Secoviridae* was identified. The virus was closely related to strawberry mottle virus (SMoV) and lettuce secovirus 1 (LSV1), two unassigned putative *Secoviridae* viruses. The virus was tentatively called surrounding non-Legume secovirus (snLSV). Based on the protease-polymerase region (Pro-Pol), this virus shared the closest aa identity to LSV1 with 67.3% identity while the CP region showed only 31.6% aa identity to LSV1. In addition, a new strain of carrot torradovirus 1 (CaTV1) was discovered that shared 95.9% aa identity based on the Pro-Pol region and 95.4% identity based on the CP region with the CaTV1 strain celery that was recently identified from Germany (MK063924 and MK063925).

A new strain of red clover nepovirus A (RCNVA) (genus: *Nepovirus*; subfamily: *Comovirinae*; family: *Secoviridae*) was detected in Landkreis Rostock 2017. This new strain had 96.5% identity based on the aa sequence of the Pro-Pol region with RCNVA-B46 from the Czech Republic (MG253828) and a CP aa identity of 83.2% (MG253829).

A divergent cherry leaf roll virus (CLRV) (genus: *Nepovirus*) was identified in Salzlandkreis-1 in 2018. The virus shared closest identity with CLRV isolates from New Zealand, with RNA1 sharing 82.4% nt identity to CLRV isolate KC937022 and RNA2 sharing

80% nt identity to KC937029. The aa sequence of the Pro-Pol region shared 97% identity to KC937022 while the CP region had only 89.8% aa identity with KC937029.

3.4.3.8. *Betaflexiviridae*

A divergent strain of red clover carlavirus 1 (RCCV1) (genus: *Carlavirus*; subfamily: *Quinvirinae*; family: *Betaflexiviridae*), was identified only once in 2018 in one location (Kreis Stormarn). The partial RdRp sequence shared 85.3% aa identity with RCCV1 (MG596238 and MG596239) from the Czech Republic. Interestingly, the complete coding sequence of an isolate of Heracleum latent virus (HLV) (genus: *Vitivirus*; subfamily: *Trivirinae*, family: *Betaflexiviridae*), was identified in snL of Münster 2017. Based on the CP sequence, this isolate shared 90.9% identity to HLV from Scotland (NC_039087) on the nt level and 96.4% identity based on the aa sequence. The RdRp region shared only 58.4% nt identity to grapevine virus B (GVB; MF991949) and 58.9% aa identity, respectively. A partial sequence of a chordovirus (subfamily: *Trivirinae*) was also detected in the snL of Münster 2017, tentatively called snL chordovirus (snLCV). The partial sequence shared 73.2% nt identity with carrot chordovirus 1 (CaChV1; NC_025469). Additionally, a partial waikavirus sequence was detected in the snL of Kreis Stormarn in 2018 with that showed 73.9% nt identity to bellflower vein chlorosis virus (BVCV; NC_027915) from South Korea.

3.4.3.9. *Bromoviridae*

Cucumber mosaic virus (CMV) (genus: *Cucumovirus*; family: *Bromoviridae*) was found in Salzlandkreis-2's sL in 2018. The three viral RNAs shared 99.1 to 99.5% nt identity to various CMV isolates (HE793685 from France, AF416900 from USA, EF202597 from China). Additionally, two new ilarviruses were identified in the snL samples of Trenthorst in 2016 and the sL of Landkreis Rostock in 2018. The putative CP aa sequence of the snL ilarvirus shared closest identity to asparagus virus 2 (AV2; NC_011807, from Mexico) with 81.1% aa identity. The sL ilarvirus CP was closely related to ageratum latent virus from Australia (AgLV; NC_022129) with 60.3% aa identity.

3.4.3.10. Other viruses

Black grass cryptic virus 2 (BGCV2) was detected in Salzlandkreis-1 in 2016; RNA1 shared 99% nt identity with an isolate probably from the UK (NC_026799) while its RdRp shared 100% aa identity. A turnip vein-clearing virus (TVCV), member of subgroup 3 of the genus *Tobamovirus* (family: *Virgaviridae*), was identified in Salzlandkreis-2's snL in 2018. The assembled complete CDS of Salzlandkreis-2 TVCV isolate shared 95.5% nt identity with TVCV strain NZ-587 from New Zealand (accession no. JN205074). The amino acid sequences shared between 97.7 and 99.4% with their homologs of strain NZ-587. White clover cryptic virus 2 (WCCV-2; genus *Betapartitivirus*) was identified in sL of Kreis Stormarn 2017 and Münster 2018, with 98.6% and 98.5% nt identity to WCCV-2 from Australia (MH427306). Six white clover mosaic virus (WClMV) isolates, genus: *Potexvirus*; family: *Alphaflexiviridae*, were identified in the sL namely in white clover (*Trifolium*

repens) and shared 99.2 to 99.3 to the WCMV-PV1 isolate from the Czech Republic (MG600296). Finally, a sequence with 81.3% nt identity to Wuhan insect virus 21 (WHIV21) (KX883227) from China was detected in sL of Salzlandkreis-1 2018.

3.4.4. Spatial and temporal differences in viral populations

The spatial and temporal compositions of the viromes in the different regions over a period of three growing seasons show, perhaps not unsurprisingly, similarities as well as fundamental differences. For example, PEMV1, PEMV2 and PNYDV were the viruses that could be located in all regions (but not in every season; Figure 8). The polerovirus TuYV was detected in all regions except for the Landkreis Rostock region (Figure 8). This location is close to the Baltic Sea with winds blown inland from the seaside; therefore, aphids as virus vectors are unlikely to carry viruses into the crops.

Interestingly, some viruses could only be located in very few regions. For example, WCMV was detected in four regions i.e., Salzlandkreis-2, Münster, Kreis Stormarn and Landkreis Rostock whereas several other viruses were only present in two different regions. BYMV was detected in two regions i.e., Salzlandkreis-1 and Salzlandkreis-2. BGCV2 was in Salzlandkreis-1 and Landkreis Rostock. ClYVV was in Salzlandkreis-2 and Münster. PSbMV was detected Salzlandkreis-2 and Kreis Stormarn. RCEV 1 and RCUV were detected in two regions i.e., Münster and Landkreis Meißen. SbDV and WCCV2 were detected Münster and Kreis Stormarn. Furthermore, BLRV and PaEV were detected only in one region (Salzlandkreis-2 and Landkreis Meißen, respectively).

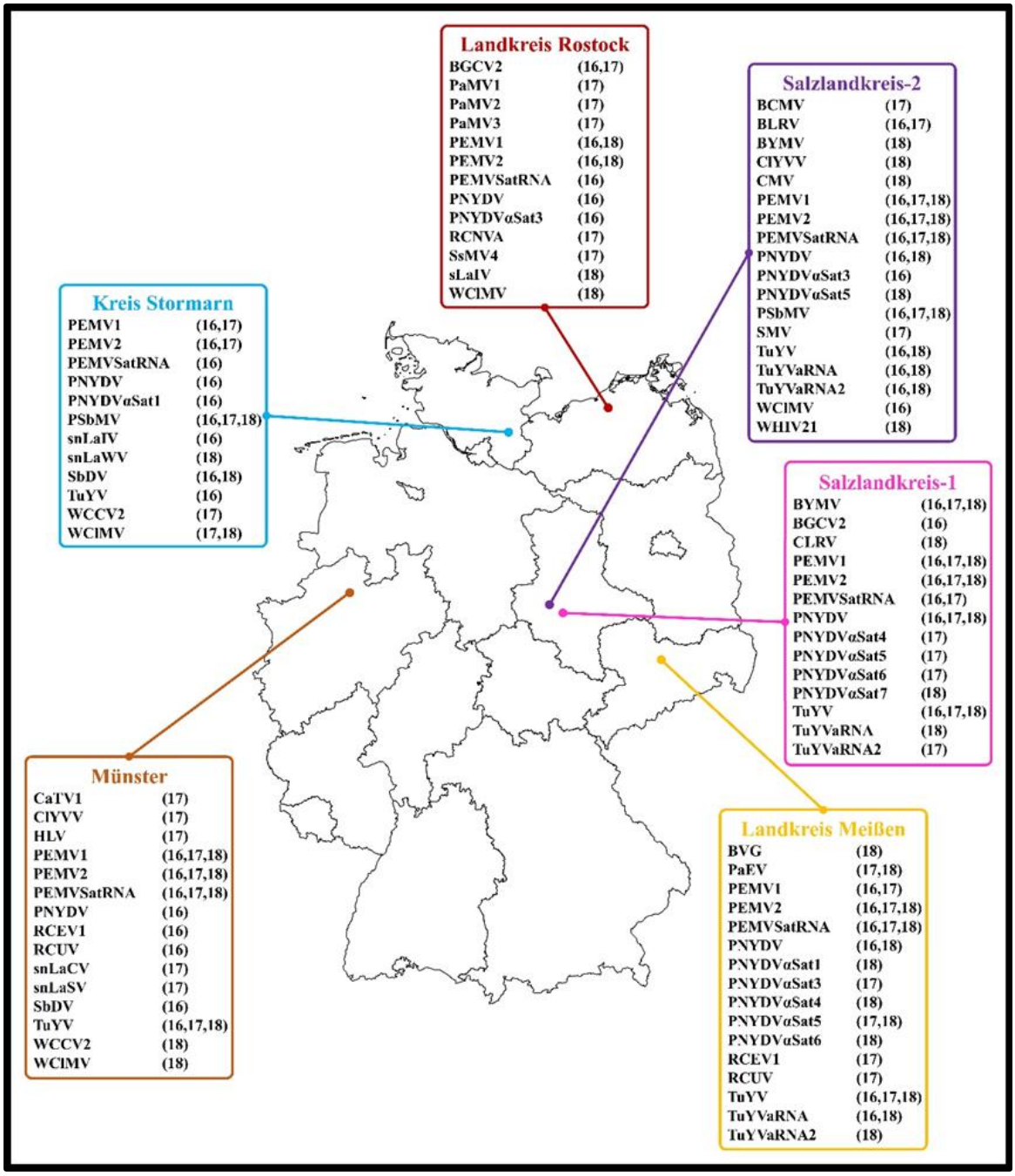

Figure 8: List of viruses and their associated nucleic acids detected in each of the six German pea growing regions sampled over three seasons 2016, 2017 and 2018. The seasons in which the viruses were detected are mentioned in brackets. The names of the viruses and their associated nucleic acids are: BVG: barley virus G, BCMV: bean common mosaic virus, BLRV: bean leafroll virus, BYMV: bean yellow mosaic virus, BGCV2: black grass cryptic virus 2, CaTV1: carrot torradovirus 1, CLRV: cherry leaf roll virus, ClYVV: clover yellow vein virus, CMV: cucumber mosaic virus, HLV: Heracleum latent virus , PaMV1: pea associated mitovirus 1, PaMV2: pea associated mitovirus 2, PaMV3: pea associated mitovirus 3, PaEV: pea associatted emaravirus, PEMV1: pea enation mosaic virus 1, PEMV2: pea enation mosaic virus 2, PEMVsatRNA: pea enation mosaic virus satellite RNA, PNYDVαSat1: pea necrotic yellow dwarf alphasatellite 1, PNYDVαSat3: pea necrotic yellow dwarf alphasatellite 3, PNYDVαSat4: pea necrotic yellow dwarf alphasatellite 4, PNYDVαSat5: pea necrotic yellow dwarf alphasatellite 5, PNYDVαSat6: pea necrotic yellow dwarf

alphasatellite 6, PNYDVαSat7: pea necrotic yellow dwarf alphasatellite 7, PNYDV: pea necrotic yellow dwarf virus, PSbMV: pea seed-borne mosaic virus, RCCV1: red clover carlavirus 1, RCEV1: red clover enamovirus 1, RCNVA: red clover nepovirus a, RCUV: red clover umbravirus, SsMV4: *Sclerotinia sclerotiorum* mitovirus 4, sLaIV: sL associated ilarvirus, snLaCV: snL associated chordovirus, snLaIV: snL associated ilarvirus, snLaSV: snL associated secoviridae, snLaWV: snL associated waikavirus, SbDV: soybean dwarf virus, SMV: soybean mosaic virus, TVCV: turnip vein-clearing virus, TuYV: turnip yellows virus, TuYVaRNA: turnip yellows virus associated RNA, TuYVaRNA2: turnip yellows virus associated RNA 2, WCCV2: white clover cryptic virus 2, WClMV: white clover mosaic virus, WHIV21: Wuhan insect virus 21.

Other viruses were found in only one season in one region i.e., BVG, BCMV, CaTV1, CLRV, CMV, HLV, RCCV1, RCNVA, SMV, TVCV, WHIV21 and all mitoviruses. So far, many of these viruses have not yet been included in standard monitoring programs of pea diseases, therefore the incidence and impact of these viruses on pea and other legume crops is currently unknown.

PEMV1, PEMV2 and their satellites were not detected in 2017 in Landkreis Rostock and in 2018 in Kreis Stormarn. In 2018, PEMV1 was not detected in Landkreis Rostock, however PEMV2 and the satellite RNA were present. TuYV was detected in all three seasons in Salzlandkreis-1, Münster and Landkreis Meißen, however after its detection in Kreis Stormarn in 2016, it could not be detected in 2017 and 2018. Meanwhile, TuYV could not be identified in Salzlandkreis-2 in 2017.

PNYDV, an emerging nanovirus in Germany, was detected in all the regions in 2016 (Figure 8). PNYDV was detected during in all three seasons in Salzlandkreis-1 and Landkreis Meißen, consistent with previous monitoring. PNYDV was not detected in the other four regions in 2017; however, in 2018, it was again detected in Salzlandkreis-2. In 2016 and 2017, WClMV could be found in Salzlandkreis-2 and Kreis Stormarn while in 2018 it was detected again in Kreis Stormarn with additional confirmations in Münster and Landkreis Rostock. PSbMV was detected in all seasons in Salzlandkreis-2 and Kreis Stormarn. Although BLRV in one region, it was detected for two successive season 2016 and 2017 but could not be found in 2018. PaEV was also detected for two seasons in 2017 and 2018.

Taken together, when we look at the spatial distribution of viruses, we can only observe minor differences in the viral community compositions. No major differences could be observed between the regions as e.g., a total of 36 virus incidences were detected in Salzlandkreis-2 followed by Münster with 33 virus incidences (over the three seasons). A total of 28 virus incidences were reported from Salzlandkreis-1, 25 for both Kreis Stormarn and Landkreis Meißen and the lowest rate for Landkreis Rostock (20).

The temporal virus incidences appear to be relatively stable over the successive seasons 2016, 2017 and 2018 with 53, 56 and 54 virus incidences, respectively (Supplementary Table S4).

3.5. Discussion

This is the first HTS-based study designed to describe the pea virome in Germany. In addition to the focus on viruses infecting pea crops, we also explored spatio-temporal aspects across six different production regions in Germany over three years. We also examined as potential virus reservoirs non-pea plants surrounding the pea crops, including those related to peas (belonging to the *Fabaceae*) and plants that were completely unrelated. This distinguishes our study from many metagenomics studies that focused on either just one crop plant, one production area, one season or combination thereof. We believe that this study demonstrates the importance of spatio-temporal aspects into metagenomics studies in order to draw a more complete picture of all the viruses present.

To discover viruses with different genomes and prevent bias we used a ribosomal RNA depleted total RNA (Pecman *et al.*, 2017). As demonstrated, this method detected RNA viruses with both a plus and negative sense genome, and viruses with a DNA genome, namely PNYDV and its associated satellites. We discovered a surprisingly high number of viruses in the different pools with some present in high abundance, i.e., CMV, PEMV1, PEMV2, PEMVSatRNA, TuYV, TuYVaRNA1 and TuYVaRNA2. The recovered reads were pool-dependent as well as dependent on the viral genome, virus titre and incidence within the pool (Supplementary Table S3, Figure 3). All these viruses are positive single-stranded viruses. Interestingly, dsRNA viruses, i.e., partitivirids, and the new negative sense RNA emaravirus were also detected using the ribo-depletion method, despite a low number of reads.

3.5.1. Detection of known viruses in the pea pools

As expected, we were able to detect and confirm by RT-PCR the presence of many known pea viruses that have been described for Germany previously, i.e., PEMV1, PEMV2, TuYV, PNYDV, BLRV, SbDV, PSbMV and CMV. This is in line with previous surveys and observations (data not shown). PEMV1 and PEMV2 are usually the most commonly found viruses in German pea crops (Ziebell, 2017). PEMV is associated with enation and mosaic symptoms on infected plants and can lead to severe yield losses (Clement, 2006). Furthermore, mixed infections of PEMV1 and PEMV2 are well documented in pea samples (Hagedorn & Khan, 1984; Brault *et al.*, 2010).

The second most prevalent virus detected was TuYV. It was found in peas and also in the snL pools. TuYV is known to infect peas but is of major concern for rapeseed in Germany (Graichen & Schliephake, 1999; Gaafar & Ziebell, 2019b). Although, we have no direct evidence that surrounding non-legumes are reservoirs for TuYV isolates that infect peas, in greenhouse experiments we demonstrated that TuYV isolates originating from peas can infect rapeseed and *vice versa* (data not shown). It is therefore very likely that rapeseed crops, other members of the *Brassicaceae* family as well as a large number

of common weeds and wild species host range are alternative hosts for pea-infecting poleroviruses such as TuYV (Stevens *et al.*, 1994).

BLRV is another luteovirus that has been known for a long time to infect peas in Germany (Quantz & Volk, 1954). It has been reported from many other countries e.g., Australia, Greece, India and USA and can infect various other legumes causing symptoms of stunting and leaf rolling (Reddy *et al.*, 1979; Vemulapati *et al.*, 2010; van Leur & Kumari, 2011; Chatzivassiliou *et al.*, 2016). Control of BLRV is best managed using host plant resistance (van Leur *et al.*, 2013), however, BLRV-resistant green and protein pea varieties may not be widely available at the present. As with other phloem-limited luteoviruses, BLRV is transmitted by aphids in a non-persistent manner with pea aphids (*Acyrthosiphon pisum*) being one of the most efficient vectors (Clement, 2006). Another luteovirus, SbDV, is also known to infect peas and other legumes in Germany (Grigoras *et al.*, 2010a; Abraham *et al.*, 2007). The virus was reported from many countries causing economic losses on soybean (*Glycine max* L.) (Phibbs *et al.*, 2004; Tamada *et al.*, 1969). The virus can cause several symptoms e.g., yellowing, dwarfing, downward curling, rugosity and reduction in growth (Tamada *et al.*, 1969; Abraham *et al.*, 2007) but it is unclear whether peas with BLRV resistance can be infected with other closely-related poleroviruses.

Ten years ago, PNYDV was first discovered in Germany (Grigoras *et al.*, 2010a). In the following years, PNYDV was only detected in two German regions in Saxony and Saxony-Anhalt, as well as in neighbouring Austria. However, 2016 saw the first country-wide outbreak of this virus in Germany and PNYDV was detected also in other European countries (e.g. Denmark and The Netherlands) (Gaafar *et al.*, 2016; Gaafar *et al.*, 2017; Ziebell, 2017; Gaafar *et al.*, 2018a). Effects on infected plants are severe and can cause high yield losses (Saucke *et al.*, 2019). PNYDV is an increasing threat to legume production in Europe as no PNYDV-resistant plant varieties have been identified yet (data not shown). Additionally, we believe the high mutation rate, reassortment and recombination rates of nanoviruses like PNYDV may lead to the appearance of novel strains (Grigoras *et al.*, 2010b; Grigoras *et al.*, 2014). An increasing number of nanovirus-associated single-stranded circular DNA alphasatellites have been reported in recent years in legumes such as *Sophora alopecuroides* L., *Vicia cracca* L. and *Apiaceae* members such as *Petroselinum crispum* (Mill.) Fuss, although their biological relevance is still unclear (Gallet *et al.*, 2018; Vetten *et al.*, 2019; Heydarnejad *et al.*, 2017).

The seed- and aphid-transmissible PSbMV has also been reported from many countries, including Germany (Khetarpal & Maury, 1987; Latham & Jones, 2001). However, in Germany, PSbMV is not seen as a major constraint of pea production as the provision of “clean” seed material and close surveillance of pea seed production sites has helped to reduce PSbMV incidence below those described for other countries. It is interesting that one of the two sites in which we detected PSbMV is a trial site for heritage material. The second region in which we detected PSbMV is in closer proximity to a protein pea breeding site in which we detected PSbMV in previous surveys (data not

shown) therefore viruliferous aphids may be vectors in this region. The symptoms of PSbMV are reported to be mild and transitory in pea resulting in limited detection of the virus in the field (Khetarpal & Maury, 1987). Other symptoms can include slight chlorosis, stunting, shortening, veins swelling and leaflets' downward rolling (Khetarpal & Maury, 1987). Seed symptoms of coat staining, coat splitting, and reduced seed size caused by PSbMV were observed in New Zealand, USA and Australia in the 1980s. As a result, a program to breed resistance was initiated leading to the release of pea varieties highly tolerant to PSbMV (Fletcher *et al.*, 1989; Russell, 1994). Similar breeding programs were also successfully developed in Australia (van Leur *et al.*, 2013).

The presence of CMV in the sL was not surprising, as CMV is a wide-spread virus, with an extremely wide host range; CMV can be transmitted by many aphid species, and cause severe losses in legumes (Palukaitis *et al.*, 1992; Fletcher, 1993; Garcıa-Arenal & Palukaitis, 2009; Tao *et al.*, 2002). CMV was reported before in German peas and other legumes but does not seem to be of major concern (Schmidt, 1981).

Although as being reported as present in Germany e.g., widespread on faba beans or clover plants, we could not find any detailed information on BYMV or ClYVV naturally infecting peas in Germany. BYMV has a wide host range compared to other potyviruses including legumes and ornamentals (Guyatt *et al.*, 1996; Nakazono-Nagaoka *et al.*, 2009). Additionally, it can be transmitted by more than 20 aphid species causing symptoms including mosaic, necrosis and yellowing resulting in severe yield losses (Guyatt *et al.*, 1996; Nakazono-Nagaoka *et al.*, 2009). The pathogenicity and serotypes of the BYMV differ from one strain to another (Bos, 1970; Barnett, 1987). Clover yellow vein virus (ClYVV) has a host range overlapping with BYMV and often confused with it as they are closed serologically (Barnett, 1987; Nakazono-Nagaoka *et al.*, 2009).

3.5.2. New players in German peas

Our study highlights that there are many pea viruses present in Germany that have neither been described previously and are not being monitored: RCEV1, RCUV, SsMV4, and two associated nucleic acids i.e., PEMVSatRNA and TuYVaRNA.

Interestingly, RCEV1 and RCUV were not only detected in peas but also in the sL (but not in the same location and not in the same season) which might indicate that surrounding, perennial legumes can be a virus reservoir for peas. The possibility of a mixed infection with both viruses was confirmed for red clover (*Trifolium pratense* L.) in the Czech Republic (Koloniuk et al. personal comm.). Such mixed infections between luteovirids and umbraviruses is common i.e., PEMV1 and PEMV2.

The mitovirus SsMV4 was found to infect *Sclerotinia sclerotiorum* (Lib.) de Bary, a widespread plant pathogenic fungus causing white mould disease especially in peas, lentils and beans (McKenzie & Morrall, 1975; Purdy, 1979; Bardin & Huang, 2001; Bolton *et al.*, 2006; Nibert *et al.*, 2018). . A study showed that SsMV4 in combination with two

other mitoviruses i.e., *Sclerotinia sclerotiorum* mitovirus 2 (SsMV2) and *Sclerotinia sclerotiorum* mitovirus 3 (SsMV3), reduced the in-vitro growth and virulence of *S. sclerotiorum* on cabbage, common bean, oilseed rape and tomato (Khalifa & Pearson, 2013).

For the first time in Germany we also discovered numerous virus associated nucleic acid sequences. PEMVSatRNA, a small linear single stranded RNA satellite, has also been extracted from peas in the USA (Demler & Zoeten, 1989). PEMVSatRNA appears not influence aphid transmission, particle morphology, or symptom expression in peas but can reduce the severity of symptoms on the indicator plant *Nicotiana benthamiana* (Demler & Zoeten, 1989; Demler *et al.*, 1994). Whether the PEMVSatRNA detected in Germany can modulate symptom expression of PEMV in its natural host *P. sativum* remains to be investigated. We also discovered PNYDV-associated alphasatellites 1 and 3. PNYDVαSat1 was previously detected in Austria, while PNYDVαSat3 was detected in both Austria and Denmark (Gaafar *et al.*, 2018a) In addition, four new alphasatellites were detected (discussed later). Finally, TuYVaRNA was also detected in association with TuYV. We discovered TuYVaRNA recently associated with TuYV from rapeseed in Germany (Gaafar & Ziebell, 2019b) but its role in symptom modulation, host range determination or vector transmission also remains to be investigated.

3.5.3. New viruses and associated satellites – what are the risks?

A new emaravirus i.e., PaEV, three new mitoviruses associated with peas i.e., PaMV1, PaMV2 and PaMV3, and novel associated nucleic acids in peas PNYDVαSat4, PNYDVαSat5, PNYDVαSat6 and TuYVaRNA2 were detected in this survey. PaEV was detected in German peas for two successive seasons in Landkreis Meißen, therefore we a quite confident that this virus is established in this region and might pose a risk. However, due to our pooling strategy, we are unable to attribute symptoms associated with PaEMV or to recover infectious virus material from the samples. Generally, emaraviruses can infect trees and deciduous shrubs. In the USA and Canada, the emaravirus rose rosette virus (RRV), mite-transmitted by *Phyllocoptes fructiphilus* Kiefer (Acari: Eriophyidae), causes extreme damage to roses leading to plant death within a short period of time (Olson *et al.*, 2017). RRV and its vector were placed on the A1 alert list by the European and Mediterranean Plant Protection Organisation (EPPO, 2019). Two other emaraviruses have been reported from legumes. On pigeonpea (*Cajanus cajan* L.; *Fabaceae*) sterility mosaic disease (SMD) is caused by pigeonpea sterility mosaic emaravirus 1 and pigeonpea sterility mosaic emaravirus 2 (Elbeaino *et al.*, 2014, 2015). These emaraviruses can also be transmitted by eriophyid mites (*Aceria cajani* Chann.) (Elbeaino *et al.*, 2015; Patil & Kumar, 2015). The total number of segments of our PaEV isolate are unknown as the virus full genome could not be recovered due to its low number of reads in the pools. In future monitoring programs, the distribution of PaEV will be evaluated to assess the risk that this virus might pose.

Three new mitoviruses were found in the pea pools i.e., PaMV1, PaMV2 and PaMV3. Mitoviruses are wide spread in plants and their infection to pathogenic fungi is often associated with virulence reduction (Bruenn *et al.*, 2015; Wu *et al.*, 2007; Xie & Ghabrial, 2012). The fungi which these new mitoviruses can infect are currently unknown and need more investigation. Also, it is important to understand the role of these mitoviruses as they may be used as biocontrol for fungal infections that in future could help to reduce the impact of fungal diseases in peas (powdery mildew, downy mildew, *Aphanomyces*, fusarium root rot).

In addition to the previously described PNYDV alphasatellites 1 and 3, we discovered four new PNYDV-associated alphasatellites. Alphasatellites rely on their helper virus for spread as they do not encode a coat protein (Briddon *et al.*, 2018). The presence of alphasatellites is associated with reduced infectivity of faba bean necrotic yellows virus (a nanovirus) or in the case of begomoviruses alphasatellites reduced or intensified symptoms and/or reduced virus genomic titre or betasatellite (Timchenko *et al.*, 2006; Paprotka *et al.*, 2010; Mar *et al.*, 2017; Idris *et al.*, 2011; Kon *et al.*, 2009). The Rep proteins encoded by begomoviruses' alphasatellites were found to suppress transcriptional gene silencing or post-transcriptional gene silencing (Nawaz-Ul-Rehman *et al.*, 2010; Abbas *et al.*, 2019). The role of these alphasatellites in unknown and their association with nanoviruses and other viruses e.g., babuviruses and begomoviruses require clarification.

Additional to the reported TuYVaRNA1, we detected a new TuYV associated RNA, TuYVaRNA2. The effects of these associated RNAs on TuYV transmission and infection need more investigation as previous studies showed that beet western yellows associated RNA (strain ST9) increases the severity of on beet western yellows virus and is not associated with the virions' escaping the phloem (Sanger *et al.*, 1994).

3.5.4. Non-crop plants: Reservoirs for viruses infected crops

In our study, we analysed plants (leguminous and non-leguminous plants) surrounding the pea fields to investigate potential virus reservoirs. We were able to detect various viruses in those plant pools that may pose a threat to peas as well as production of other crops.

3.5.4.1. In the surrounding legumes:

Not surprisingly, in surrounding legumes we were able to detect viruses which are known to infect other legumes i.e., BCMV, SMV, WCCV2 and WClMV in addition to CMV and SbDV (discussed above). BCMV is well known to infect *Phaseolus* beans causing common mosaic or black root disease depending on the host, virus strain and the environmental conditions (Drijfhout & Bos, 1977). BCMV is a seed-borne virus, aphid transmissible and can be transmitted mechanically. A previous study showed that BCMV NL1 strain could not infect peas (Drijfhout & Bos, 1977).

SMV is also distributed worldwide and is a main threat for soybean production in many countries e.g., Japan and USA (Hill & Whitham, 2014). Its symptoms including mosaic, leaf distortion, leaf deformation and seed damage. SMV reduces germination, seed size, oil content and nodulation, and total yield in soybean (Hill & Whitham, 2014; Ross, 1977). SMV can be seed transmitted or by aphid vectors (Li *et al.*, 2018). Besides soybean, it can infect other legumes including pea and was found in German faba beans before (Nandakishor *et al.*, 2017). WCCV2, a symptomless cryptic virus, has been previously detected in German in white clover (*Trifolium repens*) (Lesker *et al.*, 2013) and in many other countries like Japan, Korea, New Zealand and USA (Zhao *et al.*, 2016; Ido *et al.*, 2012; Park *et al.*, 2017; Nakabayashi *et al.*, 2002; Bos *et al.*, 1959). The virus can infect several legumes including pea, white clover, red clover (*Trifolium pratense*) and non-legumes e.g., garlic mustard (*Alliaria petiolata*) causing mosaic symptoms. Interestingly, this virus was also detected by HTS in Australian honey bee populations (Roberts *et al.*, 2018). WClMV is usually mechanically transmitted and may be seed transmitted with few records of insect transmission (Garrett, 1991).

In our survey, we identified new viruses in the surrounding legumes; i.e., RCCV1, RCNVA, WClMoV and WHIV21 in addition to RCEV1 and RCUV (discussed above). RCCV1 is avirus which was present in the sL pool and which may also infect peas. In a previous 2016 survey using antibodies developed for detection of red clover vein mosaic virus (RCVMV)-like carlaviruses, the presence of a carlavirus was suggested in several pea samples but the exact virus species was not determined (Ziebell, 2017). Koloniuk and colleagues recently identified the genome of RCCV1 and found that it shares the capsid protein sequence with RCVMV (Koloniuk et al., personal comm.). RCNVA is a new virus that was identified in red clover (*Trifolium pratense* L.) plant in the Czech Republic (Koloniuk *et al.*, 2018). It was detected only once in 2018 in the sL pool of Landkreis Rostock. The host range of RCNVA is currently unknown.

Based on the sequences available on NCBI, WClMoV was only detected in Korea in white clover, however no more information is available. The finding of WHIV21 in a plant is interesting as another strain (WHCCII13077) was only reported in China from insect tissues (Shi *et al.*, 2016a). The virus is taxonomically unassigned to any virus family and its role is currently unknown.

A new ilarvirus, which we named sLaIV was detected in the sL. Although the exact host of sLaIV is currently unknown due to our pooling strategy, this virus will be included in future surveys to investigate incidence and potential host plants. To our best knowledge, sLaIV is only the second ilarvirus to naturally infect legumes after tobacco streak virus (Kaiser, 1982).

3.5.4.2. In the surrounding non-legumes:

In the pools of surrounding non-legume plants, we were able to detect various viruses that were described previously for Germany i.e., CaTV1, CLRV and TuYV. CaTV1 was detected recently in celery plants exhibiting chlorotic ringspots, mosaic and strong

yellowing symptoms (Gaafar & Ziebell, 2019a) Similarly, the divergent strain of CLRV detected in this study is not new to Germany, as it has been previously detected in carrots (data not published). The presence of other strains of CLRV have been previously reported from German forest trees (beech and birch) (Jones *et al.*, 1990; Rebenstorf, 2005). This virus is widely distributed and has a wide natural host range including woody and herbaceous plants (Büttner *et al.*, 2011). CLRV may be transmitted by seed, pollen, as well mechanically including grafting. CLRV can cause a wide range of symptoms depending on the host and strain, for example in cherry trees it can cause a delay in leaf formation, upward leaf rolling, it can also delay flowering, reduce fruit production and can eventually kill the tree (Cropley, 1961) Finding CLRV in the surrounding non-legumes is not surprising as it was also detected previously in rhubarb (*Rheum rhaponticum*; family: *Polygonaceae*) in Britain (Tomlinson & Walkey, 1967). CLRV was also transmitted mechanically to peas and French beans in addition to many other economically important plants (Tomlinson & Walkey, 1967).

BVG, BGCV2, HLV and TVCV are new reports for Germany. BVG was detected recently in the Netherlands, a neighbouring country of Germany, and infects switchgrass (*Panicum virgatum* L.) (Kumar *et al.*, 2018a). The presence of this virus in Germany may be important for the production of barley and other cereal crops (Jo *et al.*, 2018a). Another cryptic virus i.e., BGCV2 was detected in the snL. BGCV2 is a recently identified virus sequence from black grass (*Alopecurus myosuroides* Huds.) based on NCBI data (no publication available).

Viruses belonging to the *Betaflexiviridae* family were also detected in the "non-legume weed pools". We were able to determine the complete coding sequence of HLV for the first time. This virus was only detected once before in Scotland and appears to be widely spread in wild hogweed (*Heracleum sphondylium* L.; *Apiaceae*), with no obvious leaf symptoms (Bem & Murant, 1979). However, it was possible to infect many other plants experimentally including carrot (*Daucus carota* L.), maize (*Zea mays* L.), and cowpea (*Vigna unguiculata* (L.) Walp.), another legume. HLV can be transmitted mechanically or by aphids (Bem & Murant, 1979) On experimental plants symptoms such as chlorosis, mottling, necrotic rings or spots or vein clearing could be observed (Bem & Murant, 1979). Although the virus did not show symptoms on most of the infected plants, the effect of the virus on plant production is currently unknown.

TVCV was detected one time in the snL of Salzlandkreis-2. Previously, TVCV was detected in several countries e.g., France, New Zealand and USA (Cardin *et al.*, 2009; Cohen *et al.*, 2012; Lockhart *et al.*, 2008). It infects different species including turnip (*Brassica rapa* L.), penstemon (*Penstemon digitalis* Nutt. ex Sims.) *Plantago major* L., and kidney weed (*Dichondra repens* L.) and the crop plant *Actinidia chinensis* Planch. (kiwifruit) causing different symptoms including vein clearing, red foliar ringspots, leaf deformation and plant stunting (Dorokhov *et al.*, 1994; Cohen *et al.*, 2012; Lartey *et al.*, 1994; Cardin *et al.*, 2009; Lockhart *et al.*, 2008; Blouin *et al.*, 2013).

We also detected several new viruses in the snL i.e., snLaCV, snLaIV, snLaSV and snLaWV. As the hosts of these viruses are not yet defined, the importance of these newly identified virus sequences i.e., sLaIV, snLaCV, snLaIV, snLaSV and snLaWV is unknown.

Our study identified several new viruses, virus strains and isolates that had not been reported before. Our findings help to improve pea virus surveys as the range of target viruses needs to be extended and the sequence data that we generated helps to improve our knowledge about virus variation thus improving taxonomy and finetuning of species demarcation criteria. The detection of another divergent CaTV1 (a torradovirus) as well as the divergent nepoviruses CLRV and RCNVA, indicate that the ICTV species demarcation criteria set for the family *Secoviridae* may require alteration. The criteria might also consider other genomic regions e.g., the full aa sequence of the poly proteins and the nt sequences of the untranslated regions as suggested by (Verbeek *et al.*, 2010; Gaafar & Ziebell, 2019a).

Our sequence data may also help to improve public databases as we were able to supply several full-length genomes. Since our BLAST results and phylogenetic trees indicate closest sequence matches from isolates that were described only from other continents more HTS studies are needed in Europe to improve local sequence databases and subsequent sequence analyses.

Whilst we acknowledge the strength of HTS in identifying known and unknown viruses of crops, our pooling strategy also has disadvantages. Firstly, we cannot obtain detailed information on the viruses infecting a single plant without back testing each specimen in the pool. Secondly, in some case it was not possible to recover the full-length viral sequence using this method. Thirdly, this approach does not allow us to link symptom data on individual plants with the viruses found in the pooled samples without back testing each specimen in the pool although subsequent testing of retained plant tissue will be insightful.

We believe that our survey data is of great interest for plant breeders (breeding for virus resistance), diagnosticians and pest risk analysist but it also demonstrates clearly the challenges of metagenomic HTS studies in the framework of virus diagnostics (laboratory and bioinformatics challenges, result interpretation, biological significance, pest risk analyses and data sharing) (Olmos *et al.*, 2018).

3.6. Conclusions

In conclusion, our method of using rRNA-depleted total RNA extracts from pooled plant tissue in combination with HTS, bioinformatic analysis and molecular confirmation has increased the speed and breadth of virus detection in one crop species in Germany over three seasons. This method enabled the detection of a range of viruses regardless of their genome type. After sequencing pea samples, we identified sequences representing thirty-five viruses, many of which were represented by nearly full genomes. As expected, well recognised pea viruses were detected in this study, including members of the

Luteoviridae, *Nanoviridae*, *Potyviridae* and *Tombusviridae* families. In addition, 25 new viruses associated with pea, non-crop legumes and non-legume plants were revealed, some unexpected and as yet unexplained. Much work is still needed to reveal the importance and context of these new host /virus associations.

We found PEMV1 and PEMV2 were the dominant virus species in pea which is consistent with what has been observed in the past. We also found some viruses had little similarity with known viruses and suggest they could be categorized as new viruses. For example, a new emaravirus was consistently detected in peas over two of the survey seasons. Other viruses were also detected in pea or in Germany first time and their importance to pea has yet to be determined. We believe the data from this study provides a comprehensive and improved overview of viruses present in German pea fields. For the newly detected viruses, further work is needed to determine the complete host range of these viruses, their effect on hosts and their likely vectors. It is also necessary to further investigate different locations and environments to increase our understanding of the virodiversity of these new viruses not only of pea but other legumes globally. The information produced on the long list of important viral pathogens and new virus species and strains present in the German pea fields should contribute to alert local governments and to establish sanitation measures to prevent viral transmission. Moreover, this study provides more evidence on the benefits of HTS and metagenomics in an important crop such as pea and the potential to develop similar virus databases in different crop fields.

3.7. Supplementary

Table S1: The metadata records (the plant, symptoms, and the average temperature**) of the six German pea growing regions sampled over three seasons 2016, 2017 and 2018.**

Salzlandkreis-1 (seed production site)					
Sample	**Season**	**Category**	**Plant**	**Symptoms**	**Temperature**
R1-16-01	2016	SP	*Pisum sativum*	severe dwarfing, yellowing	19°C
R1-16-02	2016	SP	*Pisum sativum*	yellowing, top stunting, crippled pods	
R1-16-03	2016	SP	*Pisum sativum*	yellowing, top stunting, crippled pods	
R1-16-04	2016	SP	*Pisum sativum*	severe dwarfing, chlorotic spots	
R1-16-05	2016	SP	*Pisum sativum*	top yellowing, chlorotic spots	
R1-16-06	2016	SP	*Pisum sativum*	top yellowing, chlorotic spots, top stunting	
R1-16-07	2016	SP	*Pisum sativum*	top yellowing, vein clearing, top dwarfing	
R1-16-08	2016	SP	*Pisum sativum*	yellowing, top stunting	
R1-16-09	2016	SP	*Pisum sativum*	enation symptoms on top, top yellowing, crippled pods	
R1-16-10	2016	SP	*Pisum sativum*	severe dwarfing, yellowing, vein clearing	
R1-16-11	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-12	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-13	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-14	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-15	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-16	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-17	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-18	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-19	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-20	2016	aSP	*Pisum sativum*	no obvious symptoms	
R1-16-21	2016	sL	NA*	NA	
R1-16-22	2016	sL	NA	NA	
R1-16-23	2016	sL	NA	NA	
R1-16-24	2016	sL	NA	NA	
R1-16-25	2016	sL	NA	NA	
R1-16-26	2016	snL	*Euphorbia sp.*	no obvious symptoms	
R1-16-27	2016	snL	*Poacae sp.*	no obvious symptoms	
R1-16-28	2016	snL	*Viola sp.*	no obvious symptoms	
R1-16-29	2016	snL	Unknown	mottling, chlorotic spots	
R1-16-30	2016	snL	*Chenopodium sp.*	chlorotic spots, chlorotic rings	
R1-17-01	2017	SP	*Pisum sativum*	dwarfing, pod deformation	18°C
R1-17-02	2017	SP	*Pisum sativum*	top dwarfing, mottling, pod deformation	
R1-17-03	2017	SP	*Pisum sativum*	mottling	
R1-17-04	2017	SP	*Pisum sativum*	top yellowing and dwarfing	
R1-17-05	2017	SP	*Pisum sativum*	top yellowing and dwarfing, mottling	

R1-17-06	2017	SP	*Pisum sativum*	top yellowing and dwarfing, mottling	
R1-17-07	2017	SP	*Pisum sativum*	top dwarfing, mottling	
R1-17-08	2017	SP	*Pisum sativum*	top dwarfing, mottling, pod deformation	
R1-17-09	2017	SP	*Pisum sativum*	top yellowing and dwarfing	
R1-17-10	2017	SP	*Pisum sativum*	top dwarfing, yellowing	
R1-17-11	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-12	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-13	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-14	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-15	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-16	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-17	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-18	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-19	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-20	2017	aSP	*Pisum sativum*	no obvious symptoms	
R1-17-21	2017	sL	*Trifolium sp.*	no obvious symptoms	
R1-17-22	2017	sL	*Trifolium sp.*	no obvious symptoms	
R1-17-23	2017	sL	*Trifolium sp.*	no obvious symptoms	
R1-17-24	2017	sL	*Trifolium sp.*	no obvious symptoms	
R1-17-25	2017	sL	*Melilotus sp.*	no obvious symptoms	
R1-17-26	2017	snL	*Euphorbia sp.*	no obvious symptoms	
R1-17-27	2017	snL	*Chenopodium sp.*	no obvious symptoms	
R1-17-28	2017	snL	*Brassica napus*	chlorotic spots, reddening	
R1-17-29	2017	snL	*Poacae sp.*	no obvious symptoms	
R1-17-30	2017	snL	*Brassica napus*	chlorotic spots	
R1-18-01	2018	SP	*Pisum sativum*	top dwarfing	
R1-18-02	2018	SP	*Pisum sativum*	dwarfing	
R1-18-03	2018	SP	*Pisum sativum*	top dwarfing	
R1-18-04	2018	SP	*Pisum sativum*	top yellowing, pod deformation	
R1-18-05	2018	SP	*Pisum sativum*	short internodes	
R1-18-06	2018	SP	*Pisum sativum*	dwarfing	
R1-18-07	2018	SP	*Pisum sativum*	top dwarfing	
R1-18-08	2018	SP	*Pisum sativum*	enation, top yellowing	
R1-18-09	2018	SP	*Pisum sativum*	leaf rolling, stunting	
R1-18-10	2018	SP	*Pisum sativum*	short internodes	22°C
R1-18-11	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-12	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-13	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-14	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-15	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-16	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-17	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-18	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-19	2018	aSP	*Pisum sativum*	no obvious symptoms	

R1-18-20	2018	aSP	*Pisum sativum*	no obvious symptoms	
R1-18-21	2018	sL	*Chenopodium sp.*	no obvious symptoms	
R1-18-22	2018	sL	*Robinia sp.*	no obvious symptoms	
R1-18-23	2018	sL	*Robinia sp.*	no obvious symptoms	
R1-18-24	2018	sL	*Lathyrus sp.*	no obvious symptoms	
R1-18-25	2018	sL	*Robinia sp.*	no obvious symptoms	
R1-18-26	2018	snL	*Brassica napus*	no obvious symptoms	
R1-18-27	2018	snL	*Sambucus nigra*	mosaic	
R1-18-28	2018	snL	*Convolvulaceae sp.*	red ring spots	
R1-18-29	2018	snL	*Chenopodium sp.*	leaf reddening, mottling	
R1-18-30	2018	snL	Unknown	no obvious symptoms	
Salzlandkreis-2 (pea heritage collection site)					
Sample	**Season**	**Category**	**Plant**	**Symptoms**	**Temperature**
R2-16-01	2016	SP	*Pisum sativum*	yellowing, enation, dwarfing	19°C
R2-16-02	2016	SP	*Pisum sativum*	top yellowing, severe leaf rolling	
R2-16-03	2016	SP	*Pisum sativum*	yellowing, severe dwarfing	
R2-16-04	2016	SP	*Pisum sativum*	vein clearing, enation, leaf deformation	
R2-16-05	2016	SP	*Pisum sativum*	enation	
R2-16-06	2016	SP	*Pisum sativum*	top yellowing, leaf rolling, vein clearing	
R2-16-07	2016	SP	*Pisum sativum*	yellowing, severe dwarfing, necrosis	
R2-16-08	2016	SP	*Pisum sativum*	top yellowing, top dwarfing	
R2-16-09	2016	SP	*Pisum sativum*	leaf deformation, red edges, chlorotic spots	
R2-16-10	2016	SP	*Pisum sativum*	enation	
R2-16-11	2016	aSP	NA	NA	
R2-16-12	2016	aSP	NA	NA	
R2-16-13	2016	aSP	NA	NA	
R2-16-14	2016	aSP	NA	NA	
R2-16-15	2016	aSP	NA	NA	
R2-16-16	2016	aSP	NA	NA	
R2-16-17	2016	aSP	NA	NA	
R2-16-18	2016	aSP	NA	NA	
R2-16-19	2016	aSP	NA	NA	
R2-16-20	2016	aSP	NA	NA	
R2-16-21	2016	sL	*Medicago doliata*	no obvious symptoms	
R2-16-22	2016	sL	*Trifolium tomentosum*	no obvious symptoms	
R2-16-23	2016	sL	*Robinia pseudoacacia*	mottling, mosaic	
R2-16-24	2016	sL	*Trifolium sp.*	no obvious symptoms	
R2-16-25	2016	sL	*Trifolium squarrosum*	leaf reddening	
R2-16-26	2016	snL	*Amaranthus sp.*	no obvious symptoms	
R2-16-27	2016	snL	*Galium aparine*	red midrib	
R2-16-28	2016	snL	*Fumaria vaillantii*	yellowing, leaf reddening	
R2-16-29	2016	snL	*Chenopodium sp.*	chlorotic spots	

R2-16-30	2016	snL	*Taraxacum officinale*	no obvious symptoms	
R2-17-01	2017	SP	*Pisum sativum*	top dwarfing, chlorotic spots, pod deformation	
R2-17-02	2017	SP	*Pisum sativum*	leaf deformation	
R2-17-03	2017	SP	*Pisum sativum*	yellowing, leaf and pod deformantion	
R2-17-04	2017	SP	*Pisum sativum*	dwarfing	
R2-17-05	2017	SP	*Pisum sativum*	mosaic, top dwarfing	
R2-17-06	2017	SP	*Pisum sativum*	chlorotic spots	
R2-17-07	2017	SP	*Pisum sativum*	yellowing	
R2-17-08	2017	SP	*Pisum sativum*	yellowing, pod deformation	
R2-17-09	2017	SP	*Pisum sativum*	mosaic, pod deformation	
R2-17-10	2017	SP	*Pisum sativum*	severe yellowing, pod deformation	
R2-17-11	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-12	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-13	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-14	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-15	2017	aSP	*Pisum sativum*	no obvious symptoms	19°C
R2-17-16	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-17	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-18	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-19	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-20	2017	aSP	*Pisum sativum*	no obvious symptoms	
R2-17-21	2017	sL	*Vicia sp.*	yellowing, mottling	
R2-17-22	2017	sL	*Vicia sp.*	yellowing, mottling	
R2-17-23	2017	sL	*Trifolium sp.*	no obvious symptoms	
R2-17-24	2017	sL	*Trifolium sp.*	no obvious symptoms	
R2-17-25	2017	sL	*Vicia sp.*	no obvious symptoms	
R2-17-26	2017	snL	*Chenopodium sp.*	yellowing, mottling	
R2-17-27	2017	snL	*Polygonaceae sp.*	chlorotic spots	
R2-17-28	2017	snL	*Apiaceae sp.*	no obvious symptoms	
R2-17-29	2017	snL	*Apiaceae sp.*	no obvious symptoms	
R2-17-30	2017	snL	*Anethum sp.*	no obvious symptoms	
R3-18-01	2018	SP	*Pisum sativum*	enation	
R3-18-02	2018	SP	*Pisum sativum*	enation	
R3-18-03	2018	SP	*Pisum sativum*	enation	
R3-18-04	2018	SP	*Pisum sativum*	enation	
R3-18-05	2018	SP	*Pisum sativum*	chlorosis, yellowing	
R3-18-06	2018	SP	*Pisum sativum*	enation	22°C
R3-18-07	2018	SP	*Pisum sativum*	enation	
R3-18-08	2018	SP	*Pisum sativum*	enation	
R3-18-09	2018	SP	*Pisum sativum*	enation	
R3-18-10	2018	SP	*Pisum sativum*	enation	
R3-18-11	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-12	2018	aSP	*Pisum sativum*	no obvious symptoms	

R3-18-13	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-14	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-15	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-16	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-17	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-18	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-19	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-20	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-21	2018	sL	*Medicago doliata*	no obvious symptoms	
R3-18-22	2018	sL	*Trifolium incarnatum*	no obvious symptoms	
R3-18-23	2018	sL	*Trifolium pratense*	no obvious symptoms	
R3-18-24	2018	sL	*Trigonella caerulea*	no obvious symptoms	
R3-18-25	2018	sL	*Medicago muricoleptis*	no obvious symptoms	
R3-18-26	2018	snL	*Phacelia sp.*	no obvious symptoms	
R3-18-27	2018	snL	*Taraxacum officinalis*	leaf reddening	
R3-18-28	2018	snL	*Plantago major*	no obvious symptoms	
R3-18-29	2018	snL	*Brassica oleracea*	yellowing	
R3-18-30	2018	snL	*Brassica oleracea*	mottling, leaf reddening	
Münster (pea breeding site)					
Sample	**Season**	**Category**	**Plant**	**Symptoms**	**Temperature**
R3-16-01	2016	SP	*Pisum sativum*	top yellowing	
R3-16-02	2016	SP	*Pisum sativum*	severe chlorosis, top stunting, leaf deformation	
R3-16-03	2016	SP	*Pisum sativum*	chlorotic spots, top dwarfing, enation	
R3-16-04	2016	SP	*Pisum sativum*	curly pods, necrosis, top yellowing	
R3-16-05	2016	SP	*Pisum sativum*	enation, yellowing	
R3-16-06	2016	SP	*Pisum sativum*	vein clearing, deformation, necrosis, enation	
R3-16-07	2016	SP	*Pisum sativum*	vein clearing, chlorotic spots, necrosis, enation	
R3-16-08	2016	SP	*Pisum sativum*	enation, mosaic	
R3-16-09	2016	SP	*Pisum sativum*	severe dwarfing, yellowing, pod deformation, enation	
R3-16-10	2016	SP	*Pisum sativum*	severe stunting, yellowing	17°C
R3-16-11	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-12	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-13	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-14	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-15	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-16	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-17	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-18	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-19	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-20	2016	aSP	*Pisum sativum*	no obvious symptoms	
R3-16-21	2016	sL	*Vicia sp.*	no obvious symptoms	

R3-16-22	2016	sL	*Vicia sp.*	no obvious symptoms	
R3-16-23	2016	sL	*Trifolium pratense*	no obvious symptoms	
R3-16-24	2016	sL	*Trifolium repens*	no obvious symptoms	
R3-16-25	2016	sL	*Vicia sp.*	reddish pods	
R3-16-26	2016	snL	*Chenopodium sp.*	chlorotic spots	
R3-16-27	2016	snL	*Galinsoga parviflora*	no obvious symptoms	
R3-16-28	2016	snL	*Rumex sp.*	no obvious symptoms	
R3-16-29	2016	snL	*Capsella bursa-pastoris*	no obvious symptoms	
R3-16-30	2016	snL	*Matricaria chamomilla*	no obvious symptoms	
R3-17-01	2017	SP	*Pisum sativum*	yellowing, mottling	19°C
R3-17-02	2017	SP	*Pisum sativum*	enation, mottling, pod deformation	
R3-17-03	2017	SP	*Pisum sativum*	enation, yellowing	
R3-17-04	2017	SP	*Pisum sativum*	top dwarfing, pod deformation	
R3-17-05	2017	SP	*Pisum sativum*	enation	
R3-17-06	2017	SP	*Pisum sativum*	severe mottling	
R3-17-07	2017	SP	*Pisum sativum*	enation	
R3-17-08	2017	SP	*Pisum sativum*	top dwarfing and mottling	
R3-17-09	2017	SP	*Pisum sativum*	top dwarfing, pod deformation	
R3-17-10	2017	SP	*Pisum sativum*	yellowing, enation	
R3-17-11	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-12	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-13	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-14	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-15	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-16	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-17	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-18	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-19	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-20	2017	aSP	*Pisum sativum*	no obvious symptoms	
R3-17-21	2017	sL	*Trifolium sp.*	no obvious symptoms	
R3-17-22	2017	sL	*Melilotus sp.*	no obvious symptoms	
R3-17-23	2017	sL	*Trifolium sp.*	no obvious symptoms	
R3-17-24	2017	sL	*Trifolium sp.*	no obvious symptoms	
R3-17-25	2017	sL	*Vicia sp.*	no obvious symptoms	
R3-17-26	2017	snL	Unknown	yellowing, mottling	
R3-17-27	2017	snL	*Galium sp.*	chlorotic spots	
R3-17-28	2017	snL	*Geranium sp.*	yellowing	
R3-17-29	2017	snL	*Aegopodium sp.*	no obvious symptoms	
R3-17-30	2017	snL	*Chenopodium sp.*	chlorotic spots	
R3-18-01	2018	SP	*Pisum sativum*	enation, mosaic, leaf deformation	22°C
R3-18-02	2018	SP	*Pisum sativum*	enation, mosaic	
R3-18-03	2018	SP	*Pisum sativum*	enation, mosaic	
R3-18-04	2018	SP	*Pisum sativum*	yellowing, mottling, pod deformation	

R3-18-05	2018	SP	*Pisum sativum*	enation, mosaic, leaf deformation	
R3-18-06	2018	SP	*Pisum sativum*	enation, mosaic, leaf rolling	
R3-18-07	2018	SP	*Pisum sativum*	enation, mosaic	
R3-18-08	2018	SP	*Pisum sativum*	yellowing, mottling	
R3-18-09	2018	SP	*Pisum sativum*	chlorotic spots, mosaic	
R3-18-10	2018	SP	*Pisum sativum*	chlorotic spots, mosaic	
R3-18-11	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-12	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-13	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-14	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-15	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-16	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-17	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-18	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-19	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-20	2018	aSP	*Pisum sativum*	no obvious symptoms	
R3-18-21	2018	sL	*Oxalis stricta*	no obvious symptoms	
R3-18-22	2018	sL	*Trifolium repens*	no obvious symptoms	
R3-18-23	2018	sL	*Trifolium repens*	mottling	
R3-18-24	2018	sL	*Vicia sp.*	no obvious symptoms	
R3-18-25	2018	sL	*Vicia sp.*	chlorotic lesions	
R3-18-26	2018	snL	*Solanum nigrum*	no obvious symptoms	
R3-18-27	2018	snL	*Hypericum officinalis*	no obvious symptoms	
R3-18-28	2018	snL	Unknown	no obvious symptoms	
R3-18-29	2018	snL	*Chenopodium sp.*	chlorotic spots	
R3-18-30	2018	snL	*Euphorbia sp.*	no obvious symptoms	
Kreis Stormarn (organic farming site)					
Sample	**Season**	**Category**	**Plant**	**Symptoms**	**Temperature**
R4-16-01	2016	SP	*Pisum sativum*	top yellowing, leaf deformation	
R4-16-02	2016	SP	*Pisum sativum*	top mottling, leaf deformation	
R4-16-03	2016	SP	*Pisum sativum*	vein clearing	
R4-16-04	2016	SP	*Pisum sativum*	top yellowing, dwarfing, mottling	
R4-16-05	2016	SP	*Pisum sativum*	stunting, top yellowing, leaf deformation	
R4-16-06	2016	SP	*Pisum sativum*	yellowing, leaf rolling	
R4-16-07	2016	SP	*Pisum sativum*	stunting, top yellowing, leaf rolling	
R4-16-08	2016	SP	*Pisum sativum*	top yellowing, dwarfing	17°C
R4-16-09	2016	SP	*Pisum sativum*	leaf mottling, yellowing, brittle leaves	
R4-16-10	2016	SP	*Pisum sativum*	enation	
R4-16-11	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-12	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-13	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-14	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-15	2016	aSP	*Pisum sativum*	no obvious symptoms	

R4-16-16	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-17	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-18	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-19	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-20	2016	aSP	*Pisum sativum*	no obvious symptoms	
R4-16-21	2016	sL	*Trifolium repens*	no obvious symptoms	
R4-16-22	2016	sL	*Trifolium repens*	no obvious symptoms	
R4-16-23	2016	sL	*Trifolium repens*	no obvious symptoms	
R4-16-24	2016	sL	*Trifolium repens*	no obvious symptoms	
R4-16-25	2016	sL	*Trifolium repens*	no obvious symptoms	
R4-16-26	2016	snL	*Chenopodium sp.*	chlorotic spots	
R4-16-27	2016	snL	*Apiaceae sp.*	no obvious symptoms	
R4-16-28	2016	snL	*Apiaceae sp.*	no obvious symptoms	
R4-16-29	2016	snL	*Lamiaceae sp.*	no obvious symptoms	
R4-16-30	2016	snL	*Urtica sp.*	no obvious symptoms	
R4-17-01	2017	SP	*Pisum sativum*	top dwarfing, pod deformation	17°C
R4-17-02	2017	SP	*Pisum sativum*	mottling, pod deformation	
R4-17-03	2017	SP	*Pisum sativum*	yellowing	
R4-17-04	2017	SP	*Pisum sativum*	top dwarfing, pod deformation	
R4-17-05	2017	SP	*Pisum sativum*	yellowing	
R4-17-06	2017	SP	*Pisum sativum*	yellowing, pod deformation	
R4-17-07	2017	SP	*Pisum sativum*	mottling, pod deformation	
R4-17-08	2017	SP	*Pisum sativum*	yellowing, leaf hardening and rolling, pod deformation	
R4-17-09	2017	SP	*Pisum sativum*	yellowing, chlorotic spots	
R4-17-10	2017	SP	*Pisum sativum*	top dwarfing, pod deformation	
R4-17-11	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-12	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-13	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-14	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-15	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-16	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-17	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-18	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-19	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-20	2017	aSP	*Pisum sativum*	no obvious symptoms	
R4-17-21	2017	sL	*Trifolium sp.*	no obvious symptoms	
R4-17-22	2017	sL	*Trifolium sp.*	leaf reddening	
R4-17-23	2017	sL	*Trifolium sp.*	yellowing, leaf reddening	
R4-17-24	2017	sL	*Trifolium sp.*	no obvious symptoms	
R4-17-25	2017	sL	*Trifolium sp.*	yellowing, leaf reddening	
R4-17-26	2017	snL	*Rumex sp.*	red ring spots	
R4-17-27	2017	snL	Unknown	yellowing	
R4-17-28	2017	snL	Unknown	no obvious symptoms	
R4-17-29	2017	snL	*Aegopodium sp.*	red ring spots	

R4-17-30	2017	snL	*Aegopodium sp.*	severe yellowing	
R4-18-01	2018	SP	*Pisum sativum*	yellowing	20°C
R4-18-02	2018	SP	*Pisum sativum*	yellowing, leaf and pod deformation	
R4-18-03	2018	SP	*Pisum sativum*	yellowing, leaf and pod deformation	
R4-18-04	2018	SP	*Pisum sativum*	yellowing, mottling	
R4-18-05	2018	SP	*Pisum sativum*	yellowing, mottling, leaf rolling	
R4-18-06	2018	SP	*Pisum sativum*	severe yellowing, pod deformation	
R4-18-07	2018	SP	*Pisum sativum*	yellowing, mottling, top dwarfing	
R4-18-08	2018	SP	*Pisum sativum*	yellowing, mottling	
R4-18-09	2018	SP	*Pisum sativum*	chlorotic spots	
R4-18-10	2018	SP	*Pisum sativum*	yellowing, mottling, leaf rolling	
R4-18-11	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-12	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-13	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-14	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-15	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-16	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-17	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-18	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-19	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-20	2018	aSP	*Pisum sativum*	no obvious symptoms	
R4-18-21	2018	sL	*Trifolium pratense*	no obvious symptoms	
R4-18-22	2018	sL	*Trifolium pratense*	no obvious symptoms	
R4-18-23	2018	sL	*Trifolium pratense*	no obvious symptoms	
R4-18-24	2018	sL	*Trifolium repens*	chlorotic spots	
R4-18-25	2018	sL	*Trifolium pratense*	no obvious symptoms	
R4-18-26	2018	snL	*Poacae sp.*	no obvious symptoms	
R4-18-27	2018	snL	*Chenopodium sp.*	no obvious symptoms	
R4-18-28	2018	snL	*Brassica napus*	leaf reddening	
R4-18-29	2018	snL	*Matricaria chamomilla*	no obvious symptoms	
R4-18-30	2018	snL	*Polygonaceae sp.*	mottling	
Landkreis Rostock (experimental field station, peas as green manure)					
Sample	**Season**	**Category**	**Plant**	**Symptoms**	**Temperature**
R5-16-01	2016	SP	*Pisum sativum*	top yellowing, pod deformation, dwarfing	18°C
R5-16-02	2016	SP	*Pisum sativum*	top yellowing	
R5-16-03	2016	SP	*Pisum sativum*	top yellowing	
R5-16-04	2016	SP	*Pisum sativum*	top yellowing	
R5-16-05	2016	SP	*Pisum sativum*	top yellowing	
R5-16-06	2016	SP	*Pisum sativum*	severe yellowing	
R5-16-07	2016	SP	*Pisum sativum*	enation, vein clearing	
R5-16-08	2016	SP	*Pisum sativum*	top deformation	
R5-16-09	2016	SP	*Pisum sativum*	severe dwarfing	
R5-16-10	2016	SP	*Pisum sativum*	top yellowing	

R5-16-11	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-12	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-13	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-14	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-15	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-16	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-17	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-18	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-19	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-20	2016	aSP	*Pisum sativum*	no obvious symptoms	
R5-16-21	2016	sL	*Lupinus albus*	no obvious symptoms	
R5-16-22	2016	sL	*Lupinus albus*	no obvious symptoms	
R5-16-23	2016	sL	*Lupinus angustifolius*	no obvious symptoms	
R5-16-24	2016	sL	*Lupinus angustifolius*	no obvious symptoms	
R5-16-25	2016	sL	*Lupinus albus*	no obvious symptoms	
R5-16-26	2016	snL	Unknown	no obvious symptoms	
R5-16-27	2016	snL	*Phacelia sp.*	no obvious symptoms	
R5-16-28	2016	snL	*Convolvulaceae sp.*	no obvious symptoms	
R5-16-29	2016	snL	*Matricaria chamomilla*	no obvious symptoms	
R5-16-30	2016	snL	*Chenopodium sp.*	no obvious symptoms	
R5-17-01	2017	SP	*Pisum sativum*	dwarfing	
R5-17-02	2017	SP	*Pisum sativum*	dwarfing	
R5-17-03	2017	SP	*Pisum sativum*	dwarfing	
R5-17-04	2017	SP	*Pisum sativum*	dwarfing	
R5-17-05	2017	SP	*Pisum sativum*	dwarfing	
R5-17-06	2017	SP	*Pisum sativum*	dwarfing	
R5-17-07	2017	SP	*Pisum sativum*	dwarfing	
R5-17-08	2017	SP	*Pisum sativum*	dwarfing	
R5-17-09	2017	SP	*Pisum sativum*	dwarfing	
R5-17-10	2017	SP	*Pisum sativum*	dwarfing	
R5-17-11	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-12	2017	aSP	*Pisum sativum*	no obvious symptoms	16°C
R5-17-13	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-14	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-15	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-16	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-17	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-18	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-19	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-20	2017	aSP	*Pisum sativum*	no obvious symptoms	
R5-17-21	2017	sL	*Vicia faba*	leaf deformation and rolling	
R5-17-22	2017	sL	*Trifolium sp.*	no obvious symptoms	
R5-17-23	2017	sL	*Lupinus sp.*	no obvious symptoms	

R5-17-24	2017	sL	*Vicia sp.*	no obvious symptoms	
R5-17-25	2017	sL	*Trifolium sp.*	no obvious symptoms	
R5-17-26	2017	snL	*Borago officinalis*	no obvious symptoms	
R5-17-27	2017	snL	*Malva sp.*	no obvious symptoms	
R5-17-28	2017	snL	*Phacelia sp.*	no obvious symptoms	
R5-17-29	2017	snL	Unknown	no obvious symptoms	
R5-17-30	2017	snL	*Matricaria chamomilla*	no obvious symptoms	
R5-18-01	2018	SP	*Pisum sativum*	mottling	20°C
R5-18-02	2018	SP	*Pisum sativum*	top yellowing	
R5-18-03	2018	SP	*Pisum sativum*	top yellowing	
R5-18-04	2018	SP	*Pisum sativum*	yellowing, mottling	
R5-18-05	2018	SP	*Pisum sativum*	yellowing, mottling	
R5-18-06	2018	SP	*Pisum sativum*	yellowing, mottling	
R5-18-07	2018	SP	*Pisum sativum*	yellowing, mottling	
R5-18-08	2018	SP	*Pisum sativum*	yellowing, mottling	
R5-18-09	2018	SP	*Pisum sativum*	yellowing, mottling	
R5-18-10	2018	SP	*Pisum sativum*	yellowing, mottling	
R5-18-11	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-12	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-13	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-14	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-15	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-16	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-17	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-18	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-19	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-20	2018	aSP	*Pisum sativum*	no obvious symptoms	
R5-18-21	2018	sL	*Trifolium incarnatum*	chlorotic spots	
R5-18-22	2018	sL	*Trifolium repens*	yellowing, mottling	
R5-18-23	2018	sL	*Vicia sp.*	chlorotic spots	
R5-18-24	2018	sL	*Trifolium sp.*	yellowing, mottling	
R5-18-25	2018	sL	*Trifolium repens*	yellowing, mottling, leaf reddening	
R5-18-26	2018	snL	*Chenopodium sp.*	chlorotic spots	
R5-18-27	2018	snL	*Solanum sp.*	yellowing, mottling	
R5-18-28	2018	snL	*Chenopodium sp.*	leaf reddening, red ring spots	
R5-18-29	2018	snL	*Brassicaceae sp.*	yellowing, mottling	
R5-18-30	2018	snL	Unknown	no obvious symptoms	
Landkreis Meißen (green pea production site)					
Sample	**Season**	**Category**	**Plant**	**Symptoms**	**Temperature**
R6-16-01	2016	SP	*Pisum sativum*	yellowing	18°C
R6-16-02	2016	SP	*Pisum sativum*	enation	
R6-16-03	2016	SP	*Pisum sativum*	mottling, mosaic, yellowing, vein clearing	
R6-16-04	2016	SP	*Pisum sativum*	mottling, yellowing	

R6-16-05	2016	SP	*Pisum sativum*	yellowing, pod deformation	
R6-16-06	2016	SP	*Pisum sativum*	enation, pod deformation, mottling	
R6-16-07	2016	SP	*Pisum sativum*	enation, rusting	
R6-16-08	2016	SP	*Pisum sativum*	severe yellowing	
R6-16-09	2016	SP	*Pisum sativum*	top dwarfing, leaf deformation, yellowing, vein clearing, mottling	
R6-16-10	2016	SP	*Pisum sativum*	severe yellowing, top deformation, mottling	
R6-16-11	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-12	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-13	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-14	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-15	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-16	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-17	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-18	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-19	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-20	2016	aSP	*Pisum sativum*	no obvious symptoms	
R6-16-21	2016	sL	*Trifolium sp.*	mottling	
R6-16-22	2016	sL	*Vicia sp.*	no obvious symptoms	
R6-16-23	2016	sL	*Trifolium sp.*	no obvious symptoms	
R6-16-24	2016	sL	*Vicia sp.*	no obvious symptoms	
R6-16-25	2016	sL	*Vicia sp.*	no obvious symptoms	
R6-16-26	2016	snL	*Solanum sp.*	no obvious symptoms	
R6-16-27	2016	snL	*Poacae sp.*	no obvious symptoms	
R6-16-28	2016	snL	*Triticum sp.*	no obvious symptoms	
R6-16-29	2016	snL	*Brassicaceae sp.*	no obvious symptoms	
R6-16-30	2016	snL	Unknown	no obvious symptoms	
R6-17-01	2017	SP	*Pisum sativum*	mosaic, mottling	
R6-17-02	2017	SP	*Pisum sativum*	pod deformation	
R6-17-03	2017	SP	*Pisum sativum*	dwarfing	
R6-17-04	2017	SP	*Pisum sativum*	mottling, pod deformation, dwarfing	
R6-17-05	2017	SP	*Pisum sativum*	dwarfing	
R6-17-06	2017	SP	*Pisum sativum*	severe yellowing, mosaic	
R6-17-07	2017	SP	*Pisum sativum*	yellowing	
R6-17-08	2017	SP	*Pisum sativum*	yellowing, mottling	
R6-17-09	2017	SP	*Pisum sativum*	yellowing, mosaic, leaf rolling	18°C
R6-17-10	2017	SP	*Pisum sativum*	pod deformation	
R6-17-11	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-12	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-13	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-14	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-15	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-16	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-17	2017	aSP	*Pisum sativum*	no obvious symptoms	

R6-17-18	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-19	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-20	2017	aSP	*Pisum sativum*	no obvious symptoms	
R6-17-21	2017	sL	*Trifolium repens*	no obvious symptoms	
R6-17-22	2017	sL	*Trifolium sp.*	no obvious symptoms	
R6-17-23	2017	sL	*Trifolium sp.*	no obvious symptoms	
R6-17-24	2017	sL	*Melilotus sp.*	no obvious symptoms	
R6-17-25	2017	sL	*Trifolium repens*	no obvious symptoms	
R6-17-26	2017	snL	*Chenopodium sp.*	no obvious symptoms	
R6-17-27	2017	snL	*Brassica napus*	chlorotic spots, leaf deformation	
R6-17-28	2017	snL	Chenopodium sp.	no obvious symptoms	
R6-17-29	2017	snL	*Poacae sp.*	no obvious symptoms	
R6-17-30	2017	snL	*Apiaceae sp.*	no obvious symptoms	
R6-18-01	2018	SP	*Pisum sativum*	chlorosis	21°C
R6-18-02	2018	SP	*Pisum sativum*	dwarfing	
R6-18-03	2018	SP	*Pisum sativum*	chlorosis	
R6-18-04	2018	SP	*Pisum sativum*	chlorotic spots	
R6-18-05	2018	SP	*Pisum sativum*	chlorosis	
R6-18-06	2018	SP	*Pisum sativum*	chlorosis, mosaic	
R6-18-07	2018	SP	*Pisum sativum*	chlorosis, mosaic	
R6-18-08	2018	SP	*Pisum sativum*	chlorosis, leaf rolling	
R6-18-09	2018	SP	*Pisum sativum*	yellowing, dwarfing of side shoots	
R6-18-10	2018	SP	*Pisum sativum*	severe yellowing	
R6-18-11	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-12	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-13	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-14	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-15	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-16	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-17	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-18	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-19	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-20	2018	aSP	*Pisum sativum*	no obvious symptoms	
R6-18-21	2018	sL	*Polygonaceae sp.*	no obvious symptoms	
R6-18-22	2018	sL	*Polygonaceae sp.*	no obvious symptoms	
R6-18-23	2018	sL	*Polygonaceae sp.*	no obvious symptoms	
R6-18-24	2018	sL	*Polygonaceae sp.*	no obvious symptoms	
R6-18-25	2018	sL	*Polygonaceae sp.*	no obvious symptoms	
R6-18-26	2018	snL	*Poacae sp.*	chlorotic spots, leaf reddening	
R6-18-27	2018	snL	*Brassica napus*	chlorotic spots	
R6-18-28	2018	snL	*Helianthus annuus*	chlorotic spots	
R6-18-29	2018	snL	*Chenopodium sp.*	chlorotic spots, mottling	
R6-18-30	2018	snL	*Phacelia sp.*	no obvious symptoms	

*NA = no plants from the category were found at sampling.

Table S2: List of primers used to confirm each virus detected by HTS.

	Virus	Primer		
		Name	Sequence	Reference
1	Barley virus G	HZ-714	TGAGTCTCGCCAAACTCCAC	This article
		HZ-715	GATTGGGATCCTCGTAGCGG	
2	Bean common mosaic virus	HZ-483	TGCAACATGGCACTTGAAGC	This article
		HZ-484	ACGCATTCTGAGTGTGACGT	
3	Bean leafroll virus	S2	ATCACITTCGGGCCGWSTCTATCAGA	(Abraham *et al.*, 2007)
		AS3	CACGCGTCIACCTATTTIGGRTTITG	
4	Bean yellow mosaic virus	HZ-704	TGATGGATGTTGCGACAGCT	This article
		HZ-705	GCCATTGCCGATCCAAATCC	
5	Black grass cryptic virus 2	HZ-791	CGCTGATTGGTCCGAATTCG	This article
		HZ-792	AACGTCCCATTAGTGAGGCG	
6	Carrot torradovirus 1	HZ-682	TGCTAGCACACAAGGACAGG	This article
		HZ-683	AGAGGCTGGGGAAAAAGTGG	
7	Cherry leaf roll virus	HZ-710	GCTGAATTGATGCGAGCCTG	This article
		HZ-711	TCGGGAGTGTCAATCCAAGC	
8	Clover yellow vein virus	HZ-702	CCGGGTTTAGTGACTCCAGC	This article
		HZ-703	ACTAGCCCAGAGAGACGAGC	
9	Cucumber mosaic virus	CMV 5'CP	CTCGAATTCGGATCCGCTTCTCCGCGAG	(Rizos *et al.*, 1992)
		CMV 3'CP	GGCGAATTCGAGCTCGCCGTAAGCTGGATGGAC	
10	Heracleum latent virus	HZ-696	AGTTACTACCCCCAGAGCGT	This article
		HZ-697	TGGGTACTTCGAATAGGCGC	
11	Pea associatted emaravirus	HZ-692	TGGTCTCTTGCATGTTGCCT	This article
		HZ-693	TAAAGCAACCTCAGCTGGCA	
12	Pea associated mitovirus 1	HZ-674	AGGTTTACCGCGATGGCTAG	This article
		HZ-675	CAGGGGCGTGAACAAGAGAT	
13	Pea associated mitovirus 2	HZ-803	ACCGTCTTGTCTCAATCCCG	This article
		HZ-804	TGACCTTTTTCGGGGCCAAT	
14	Pea associated mitovirus 3	HZ-805	AAGCCGAATGAGGTGGGAAG	This article
		HZ-806	GCATACGATCTGTTGCTGCG	
15	Pea enation mosaic virus 1	HZ-355	TCAGAAATGACGCCGGAACA	Chapter 4
		HZ-356	GCGGAACAACCTGTCTCTGA	
16	Pea enation mosaic virus 2	HZ-363	GTTGTGCGTCCTCTTGGAGA	Chapter 4
		HZ-366	CCCAAGGAGGTGTCCATGTC	
17	Pea enation mosaic virus satellite RNA	HZ-698	CCACGTTGAGATACCTCGCA	This article
		HZ-699	CCCCTGACACAATGCCATCT	
18	Pea necrotic yellow dwarf alphasatellite 1	HZ-799	TCTCTGGCGATACCCCTCTT	This article
		HZ-800	CCTCCACGCGTGTAGAAGAA	

19	**Pea necrotic yellow dwarf alphasatellite 3**	HZ-801	TTGTCCTTGGATACGCGTGT	This article
		HZ-802	CTTCACCACCATTAGGGCCA	
20	**Pea necrotic yellow dwarf alphasatellite 4**	HZ-740	TGAAAGAGCTCCTTCCAGGC	This article
		HZ-741	TGTCGATTTCCCTCCAGCTG	
21	**Pea necrotic yellow dwarf alphasatellite 5**	HZ-742	TGGCGAGACGGTATTGTTTCA	This article
		HZ-743	ATCTGGGTTTTGTCGCGGAT	
22	**Pea necrotic yellow dwarf alphasatellite 6**	HZ-744	GGGCCAGAAGTCCAATATGCT	This article
		HZ-745	TGCATCCGCCATTAGAGCTT	
23	**Pea necrotic yellow dwarf alphasatellite 7**	HZ-824	ATGAAGATGGAGGACCCCGA	This article
		HZ-825	GCGCTACAGTTTGTCCGTTG	
24	**Pea necrotic yellow dwarf virus**	priPeaRdir	GGAATACCAAGGTGAGTTACGG	This article
		priPeaRrev	TATTCCCTGAGAGTCCCGGAC	
25	**Pea seed-borne mosaic virus**	PSbMV P1 forward	GCTTCATGGTTGGAACTATTAAATG	(Giakountis *et al.*, 2015)
		PSbMV P1 reverse	AAAGTTACTTGTTTTGCATGCTTTC	
26	**Red clover carlavirus 1**	HZ-708	ACAGCATGGGTGGGAATGAG	This article
		HZ-709	ACACTCCGTCGCGCTTATAG	
27	**Red clover enamovirus 1**	HZ-686	CGTTTTCGGCTCTATGCAGC	This article
		HZ-687	GGAGACTTTCTTGCCTCGCT	
28	**Red clover nepovirus A**	HZ-676	GTCGCTGTCAGGAGTGGAAA	This article
		HZ-677	CCGTCAAATTGTGCAGCACA	
29	**Red clover umbravirus**	HZ-688	CTTTTGGTGTGCCAGGAACG	This article
		HZ-689	TGATAGCAGAGGCAGGGACT	
30	**Sclerotinia sclerotiorum mitovirus 4**	HZ-690	AGCCGCCTTTACCATATCGC	This article
		HZ-691	TGCTTCAGACACCATTCCTCC	
31	**sL associated Ilarvirus**	HZ-724	AGACGAGCTTCCCTGGTTTG	This article
		HZ-725	TTCCTCACACCACGCCTTAC	
32	**snL associated Chordovirus**	HZ-718	TGGAGAGCATGACAGGCTTG	This article
		HZ-719	TTGACGGCCATCCAGAAAGT	
33	**snL associated Ilarvirus**	HZ-730	CACCCAATAATGCCCCGACT	This article
		HZ-731	GCGCAGTACTTCCCCTTCTT	
34	**snL associated Secoviridae**	HZ-694	AGCTGCACATCCGAAACTGA	This article
		HZ-695	CGCCATTTCAGCAAAACCCA	
35	**snL associated Waikavirus**	HZ-738	ATCTTGGAAGGCTGTGTCCC	This article
		HZ-739	AACAATGCCTGGCTCTAGCA	
36	**Soybean dwarf virus**	SbDVf	GTCTACCTAAAAATTTCAAAGAATCTG	(Abraham *et al.*, 2007)
		SbDVas	CGGACCCGGTTCTCCGTCTA	
37	**Soybean mosaic virus**	HZ-493	GACAAGTGGGTTTGCGTTCC	This article
		HZ-494	TAAGCCTGGATTTGCGCTCA	
38		HZ-716	ATTGCTCGCCATGAAGGACA	This article

	Turnip vein-clearing virus	HZ-717	TGGGTGTAATTGAGCGTGCT	
39	**Turnip yellows virus**	HZ-809	ACCGTGGGTGGGTAGAAGAT	This article
		HZ-810	ACTTTTCTGAACGCCCGGAT	
40	**Turnip yellows virus associated RNA**	HZ-654	TGGACCGATACTTGCCCCTA	This article
		HZ-655	AAGTGGGTATGCTGGAGTGC	
41	**Turnip yellows virus associated RNA 2**	HZ-807	CCCGTCTGCTTCAAAGGACT	This article
		HZ-808	CTCGTGGACCGGTTCTTCAA	
42	**White clover cryptic virus 2**	HZ-797	CCATCCCTGAAGATGCTGCT	This article
		HZ-798	AGCGGAAGATAAGGCTGAGC	
43	**White clover mosaic virus**	HZ-793	TGATTGGTTACCAGTGGCCC	This article
		HZ-794	GGTGTATTTCAGGGCACGGA	
44	**Wuhan insect virus 21**	HZ-732	AGATCGACGCGTCAGACATC	This article
		HZ-733	TGGTTCCCTGTCGTACGTTG	

Table S3: Statistics of the raw data of the generated reads from each pool of the six German pea growing regions sampled over three seasons 2016, 2017 and 2018.

Pool	Season	Raw reads	Quality >Q30	Quality controlled reads	Generated contigs
Region 1	2016	831076	90.0%	689540	10457
Region 2	2016	814314	87.4%	650346	18207
Region 3	2016	863642	89.6%	700264	21660
Region 4	2016	900214	89.7%	716702	21342
Region 5	2016	687850	89.0%	502260	14335
Region 6	2016	885666	88.0%	697544	16865
Symptomatic pea	2016	931992	88.5%	741390	16095
Asymptomatic pea	2016	886412	88.0%	699910	4258
surrounding legumes	2016	1346156	87.4%	1029344	23577
surrounding non-legumes	2016	1102776	87.4%	819312	33895
Region 1	2017	2032192	76.1%	1103590	39489
Region 2	2017	2106164	76.5%	1178880	30649
Region 3	2017	1616044	84.7%	894074	38079
Region 4	2017	200210	59.3%	91166	2297
Region 5	2017	2177380	78.2%	1083988	47335
Region 6	2017	1123864	77.8%	576750	24920
Symptomatic pea	2017	1979682	79.0%	1157966	36645
Asymptomatic pea	2017	1931294	77.8%	1045932	30680
surrounding legumes	2017	1215960	82.2%	684066	29305
surrounding non-legumes	2017	1455624	77.7%	754550	28433
Region 1	2018	1877930	84.2%	815626	47637
Region 2	2018	2414122	86.8%	1326580	67186
Region 3	2018	2511364	85.4%	1354858	31197
Region 4	2018	3056130	83.9%	1565932	113465
Region 5	2018	3404406	86.0%	1794174	42290
Region 6	2018	2739384	84.7%	1410600	31100
Symptomatic pea	2018	2203570	88.4%	1220234	80298
Asymptomatic pea	2018	2125576	84.4%	1264694	87403
surrounding legumes	2018	2393184	88.9%	1192422	11814
surrounding non-legumes	2018	3299964	84.5%	1832160	112211

Table S4: List of the viruses and their associated nucleic acid satellites detected by HTS and confirmed by RT-PCR in each pool of the six German pea growing regions sampled over three seasons 2016, 2017 and 2018.

Virus	Acronym	Genus	Family	Salzlandkreis-1			Salzlandkreis-2			Münster			Kreis Stormarn			Landkreis Rostock			Landkreis Meißen			Virus relative abundance
				2016	2017	2018	2016	2017	2018	2016	2017	2018	2016	2017	2018	2016	2017	2018	2016	2017	2018	
Barley virus G	BVG	*Polerovirus*	*Luteoviridae*																		snL	0.014
Bean common mosaic virus	BCMV	*Potyvirus*	*Potyviridae*					sL														0.014
Bean leafroll virus	BLRV	*Luteovirus*	*Luteoviridae*				SP	SP														0.028
Bean yellow mosaic virus	BYMV	*Potyvirus*	*Potyviridae*			SP			SP, aSP													0.042
Black grass cryptic virus 2	BGCV2	Unclassified virus		snL												snL	snL					0.014
Carrot torradovirus 1	CaTV1	*Torradovirus*	*Secoviridae*								snL											0.014
Cherry leaf roll virus	CLRV	*Nepovirus*	*Secoviridae*			snL																0.014
Clover yellow vein virus	ClYVV	*Potyvirus*	*Potyviridae*						SP		SP											0.028
Cucumber mosaic virus	CMV	*Cucumovirus*	*Bromoviridae*						sL													0.014
Heracleum latent virus	HLV	*Vitivirus*	*Betaflexiviridae*								snL											0.014
Pea associated mitovirus 1	PaMV1	*Mitovirus*	*Narnaviridae*														SP, aSP					0.028
Pea associated mitovirus 2	PaMV2	*Mitovirus*	*Narnaviridae*														SP, aSP					0.028
Pea associated mitovirus 3	PaMV3	*Mitovirus*	*Narnaviridae*														SP					0.014
Pea associatted emaravirus	PaEV	*Emaravirus*	*Fimoviridae*																	SP	SP	0.028
Pea enation mosaic virus 1	PEMV1	*Enamovirus*	*Luteoviridae*	SP, aSP	SP, aSP	SP, aSP	SP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP		SP, aSP		SP, aSP	SP, aSP	SP, aSP		0.417
Pea enation mosaic virus 2	PEMV2	*Umbravirus*	*Tombusviridae*	SP, aSP	SP, aSP	SP, aSP	SP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP		SP, aSP		SP, aSP	SP, aSP	SP, aSP	SP, aSP	0.444
Pea enation mosaic virus satellite RNA	PEMVsatRNA	Small linear single stranded RNA satellites		SP, aSP	SP, aSP		SP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP	SP, aSP			SP, aSP			SP, aSP	SP, aSP	SP, aSP	0.361
Pea necrotic yellow dwarf alphasatellite 1	PNYDVαSat1	*alphasatellite*											SP								SP	0.028
Pea necrotic yellow dwarf alphasatellite 3	PNYDVαSat3	*alphasatellite*					SP									SP				SP		0.042
Pea necrotic yellow dwarf alphasatellite 4	PNYDVαSat4	*alphasatellite*			SP																SP	0.028
Pea necrotic yellow dwarf alphasatellite 5	PNYDVαSat5	*alphasatellite*			SP				SP											SP	SP	0.056
Pea necrotic yellow dwarf alphasatellite 6	PNYDVαSat6	*alphasatellite*			SP	SP															SP	0.042
Pea necrotic yellow dwarf alphasatellite 7	PNYDVαSat7	*alphasatellite*				SP																0.014
Pea necrotic yellow dwarf virus	PNYDV	*Nanovirus*	*Nanoviridae*	SP	SP	SP	SP		SP	SP			SP			SP			SP	SP	SP	0.153
Pea seed-borne mosaic virus	PSbMV	*Potyvirus*	*Potyviridae*				SP	SP, aSP	SP, aSP				SP	SP, aSP	SP, aSP							0.139
Red clover carlavirus 1	RCCV1	*Carlavirus*	*Betaflexiviridae*												sL							0.014
Red clover enamovirus 1	RCEV1	*Enamovirus*	*Luteoviridae*							sL										SP		0.028
Red clover nepovirus A	RCNVA	*Nepovirus*	*Secoviridae*														sL					0.014
Red clover umbravirus	RCUV	*Umbravirus*	*Tombusviridae*							sL										SP		0.014
Sclerotinia sclerotiorum mitovirus 4	SsMV4	*Mitovirus*	*Narnaviridae*														SP					0.014
sL associated Ilarvirus	sLalV	*Ilarvirus*	*Bromoviridae*															sL				0.014
snL associated Chordovirus	snLaCV	*Chordovirus*	*Betaflexiviridae*								snL											0.014
snL associated Ilarvirus	snLalV	*Ilarvirus*	*Bromoviridae*										snL									0.014
snL associated Secoviridae	snLaSV		*Secoviridae*								snL											0.014
snL associated Waikavirus	snLaWV	*Waikavirus*	*Secoviridae*												snL							0.014
Soybean dwarf virus	SbDV	*Luteovirus*	*Luteoviridae*							SP			sL		sL							0.042
Soybean mosaic virus	SMV	*Potyvirus*	*Potyviridae*					sL														0.014
Turnip vein-clearing virus	TVCV	*Tobamovirus*	*Virgaviridae*						snL													0.014
Turnip yellows virus	TuYV	*Polerovirus*	*Luteoviridae*	SP, aSP, snL	SP, aSP, snL	SP, aSP, snL	SP, snL		SP, aSP, snL	SP, aSP, snL	SP, aSP, snL	SP, aSP, snL	SP, aSP, snL						SP, aSP, snL	SP, aSP, snL	snL	0.472
Turnip yellows virus associated RNA	TuYVaRNA	Unclassified ssRNA positive-strand viruses				snL	SP, snL		snL										snL		snL	0.083
Turnip yellows virus associated RNA 2	TuYVaRNA2	Unclassified ssRNA positive-strand viruses			snL		SP		SP, snL												SP, snL	0.083
White clover cryptic virus 2	WCCV2	*Betapartitivirus*	*Partitiviridae*									sL		sL								0.028
White clover mosaic virus	WClMV	*Potexvirus*	*Alphaflexiviridae*				sL					sL		sL	sL			sL				0.069
Wuhan insect virus 21	WHIV21	Unclassified virus							sL													0.014

Chapter 4: Comparative study on three viral enrichment procedures based on RNA extraction for plant viruses/viroids detection using high throughput sequencing

Yahya Zakaria Abdou Gaafar and H. Ziebell

4.1. Abstract

High throughput sequencing (HTS) has become increasingly popular as virus diagnostic tool. It was used to detect and identify plant viruses and viroids in different types of samples. A virus sequence enrichment method prior to HTS is required to increase the viral reads in the generated data. In this study, we compared the sensitivity of three viral enrichment approaches i.e., double stranded RNA (dsRNA), ribosomal RNA depleted total RNA (ribo-depleted totRNA) and small RNA (sRNA) for plant virus/viroid detection. The three viral enrichment approaches used here enabled the detection of all known viruses/viroid, in the study, with different amounts of viral/viroid nucleotides and depths. Interestingly, both dsRNA and ribo-depleted totRNA approaches detected a divergent strain of Wuhan aphid virus 2 as well. Moreover, Vicia cryptic virus was detected in the data of dsRNA and sRNA approaches only. These results support the ability of HTS for the detection of plant viruses/viroids using RNA extracts from different plant samples. The dsRNA approach used here detected all viruses/viroid, consumed less time, is lower in cost, and required less starting material. This study can serve as guidelines for starting virus diagnostics laboratories.

4.2. Introduction

Viruses and viroids are one of the major emerging threats to agricultural and horticultural production (Anderson *et al.*, 2004). Climate change and increasing global trade are only two factors accelerating the dispersal of plant viruses by vectors into new geographic areas where they can potentially cause greater damage thus threating food supplying for humans and animals (Canto *et al.*, 2009; Hulme, 2017; Jones, 2016; Myers *et al.*, 2017). Correct identification of the underlaying causes of the diseases is important for correct management procedures, such as switching to virus-resistant cultivars (where available), quarantine or eradication measures or vector control. In the past, traditional detection methods such as serological (i.e., ELISA, Tissue blot-ELISA) or nucleic acid-based (PCR, probe-based methods) detection methods required *priori* knowledge of the pathogen that needed to be detected, such as previously purified virus particles used for raising antibodies or nucleic acid sequences for the design of specific primers or target probes (Ward *et al.*, 2004). However, divergent sequences or virus variants with different antigenic epitopes on the virion surface would not be detected using these methods. Additionally, using electron microscopy often failed to detect low titre viruses or viruses that are phloem-based; disease-causing entities without protein shells such as viroids or satellite RNAs would not be detected at all.

The development and evolution of novel high throughput sequencing (HTS) technologies has revolutionised virus discovery, plant virus diagnostics as well as metagenomic, evolutionary and community studies in recent years (Roossinck, 2017; Villamor *et al.*, 2019). Unlike the methods described above, no prior knowledge about the pathogen is needed for HTS since all the nucleic acid (viral or non-viral) in the sample can

be sequenced. Many new viruses and viroids have been discovered using HTS and the number is increasing (Elbeaino *et al.*, 2015; Chen *et al.*, 2016; Villamor *et al.*, 2017). However, HTS-based approaches for virus/viroid detection is facing several challenges in order to be validated for routine diagnostic laboratories (Maree *et al.*, 2018). As HTS can sequence all nucleic acids within the given sample, no matter if of plant or pathogen origin, suitable enrichment strategies should be used to minimise the generated host reads that reduces the pathogen reads and may interfere with subsequent bioinformatic analyses. Additionally, there is no universal sequence that can be used for the analysis of virus/viroid communities in contrast to fungi or bacteria where the internal transcribed spacer (ITS) or 16S ribosomal RNA can be used in amplicon sequencing manner for the general detection of these pathogens (Leff *et al.*, 2017). Thus, in case of unknown virus/viroid infections, studies relay on untargeted identification approaches utilizing random primers for detection (Gaafar *et al.*, 2018b).

As viruses/viroids have different genome types i.e., single- or double-stranded DNA, single- (positive [+ve] or negative [-ve]) or double-stranded RNA, circular or non-circular nucleic acids, there are various extraction methods and enrichment strategies available for these molecules. For example, the extraction of dsRNA (the replicative form of most plant viruses) has been used for a long time to generate sequence information from plant viruses and can be used for the enrichment of viral sequences for HTS (Morris, 1979; Bar-Joseph *et al.*, 1983; Roossinck *et al.*, 2010; Gaafar *et al.*, 2019c).

More recently, virus-derived small RNAs (vsRNAs) or ribo-depleted total RNA extracts have been used to prepare samples for HTS (Kreuze *et al.*, 2009; Gaafar *et al.*, 2019d). Alternatively, rolling circle amplification and subsequent sequencing worked well for viruses with circular DNA genomes (Wyant *et al.*, 2012; Gaafar *et al.*, 2018a). Few studies have directly compared different enrichment strategies and their ability to detect plant virus sequences through HTS (Visser *et al.*, 2016; Pecman *et al.*, 2017). In this study, we compare the virus/viroid detection, their genome coverage recovery and depth from the reads produced using three RNA based-enrichment strategies, i.e., dsRNA extraction, ribo-depleted totRNA and sRNA extraction. We included viruses with different genomes ((+ve) ssRNA, (-ve) ssRNA and ssDNA) as well as a viroid.

4.3. Material and methods

4.3.1. Plant cultivation

Four plant species (*Nicotiana benthamiana* [cultivar: JKI-Wild], *Pisum sativum* [cultivar: Rainier], *Solanum lycopersicum* [cultivar: Linda] and *Vicia faba* [cultivar: Tattoo]) were used (Table 1). In addition, *Phaseolus vulgaris* (cultivar: Black Turtle) seed infected with the cryptic virus phaseolus vulgaris alphaendornavirus 1 and 2 (PvEV1 and PvEV2 JKI ID 31403) were also sown to spike the samples during extraction (kindly provided by Dr. Mike Rott). The plants were kept under greenhouse conditions (at 22°C; photoperiod of 16 h light [natural daylight with additional growth light Phillips IP65, 400 Watt] and 8 h dark).

Table 1: List of plant viruses and the viroid used in this research

Sample	Name	Acronym	JKI ID number	Genome nature			Family	Genus	Host
				Type	Organization	Size (bp)			
1	**pea enation mosaic virus 1**	PEMV1	31399	+ssRNA	Linear/ Non-segmented	5,706	*Luteoviridae*	*Enamovirus*	*Pisum sativum*
2	**pea necrotic yellow dwarf virus**	PNYDV	31400	ssDNA	Circular/ segmented	7,896	*Nanoviridae*	*Nanovirus*	*Vicia faba*
3	**physostegia chlorotic mottle virus**	PhCMoV	31401	-ssRNA	Linear/ Non-segmented	13,321	*Rhabdoviridae*	*Nucleorhabdovirus*	*Nicotiana benthamiana*
4	**potato spindle tuber viroid**	PSTVd	31402	viroid	Circular/ Non-segmented	360	*Pospiviroidae*	*Pospiviroid*	*Solanum lycopersicum*
Internal controls	**phaseolus vulgaris alphaendornavirus 1**	PvEV1	31403	dsRNA	Linear/ Non-segmented	~14,072	*Endornaviridae*	*Alphaendornavirus*	*Phaseolus vulgaris*
	phaseolus vulgaris alphaendornavirus 2	PvEV2		dsRNA	Linear/ Non-segmented	~14,817	*Endornaviridae*	*Alphaendornavirus*	*Phaseolus vulgaris*

4.3.2. Viruses and viroid isolates

Four viruses with different genomes and one viroid were used in this research (Table 1). Pea enation mosaic virus 1 (PEMV1) originally from a *P. sativum* plant showing enation symptoms, collected in 2011 from Hondeghem, northern France. Pea necrotic yellow dwarf virus (PNYDV), Elbtal isolate, originally from infected *P. sativum* sample showing top leaves dwarfing and yellowing and leaf rolling symptoms, from Saxony, Germany in 2011. The original physostegia chlorotic mottle virus (PhCMoV), HZ16-558 isolate, was from infected *S. lycopersicum* plant, collected from Hesse state in Germany with fruits marbling and discoloration symptoms in 2016 (Gaafar *et al.*, 2018a). The potato spindle tuber viroid (PSTVd), isolate PV-0950, was kindly provided by DSMZ (German collection of microorganisms and cell cultures) in the form of lyophilized infected *S. lycopersicum* plant leaves in 2014.

4.3.3. Virus maintenance

PEMV1 and PNYDV were maintained by aphid transmission using *Acyrthosiphon pisum*. The aphids were reared for five days on infected *P. sativum* and *V. faba,* respectively, and ten viruliferous aphids were transferred onto healthy plants. The inoculation access period was five days. The aphids were killed using a non-systemic insecticide (Spruzit Schädlingsfrei, Neudorff GmbH KG, Germany).

PhCMoV and PSTVd were maintained by mechanical transmission to *N. benthamiana* and *S. lycopersicum*, respectively. For mechanical transmission, 100 mg of infected plant material was ground in Norit buffer (0.05M phosphate buffer [pH 7.0]; 0.001M ethylenediaminetetraacetic acid; 0.02M sodium diethyldithiocarbamic acid; 0.005M thioglycolic acid; 0.75% activated charcoal [Norit]) and 30 mg of diatomaceous earth (Celite) was added. The homogenate was rubbed gently on healthy plants' leaves using glass rods. The inoculated leaves were rinsed with water within 5 min. After inoculation, all plants were kept under greenhouse conditions for four weeks until virus symptoms were observed (except for PSTVd; no symptoms).

4.3.4. Confirmation of infection by DAS-ELISA and/or PCR/RT-PCR

To confirm infection, all plants were tested by ELISA and/or RT-PCR or PCR (Table 2), ELISA tests were performed, except for PSTVd infected plants, using antibodies mentioned in Table 2 as described in (Clark & Adams, 1977; Fletcher *et al.*, 2016). Additionally, to confirm infections with PEMV1, PEMV2, PhCMoV and PSTVd, RT-PCR was performed using total RNA extracted with innuPREP Plant RNA Kit (Analytik Jena, Germany), following the manufacturer's instructions. cDNA was synthesised using ProtoScript II Reverse Transcriptase (New England Biolabs) using the reverse primer of the primer pairs mentioned for each virus (Table 2). PNYDV infection was confirmed by DNA extraction according to Edward's method for plant DNA extraction with 0.1%

Mercaptoethanol added to the extraction buffer followed by PCR using PNYDV specific primers (Edwards *et al.*, 1991).

Table 2: Virus/viroid infection confirmation. Symptoms, ELISA and RT-PCR or PCR including antibodies and primer sets used for confirmation, and the reference sequences used for mapping

Virus	Symptoms	ELISA	RT-PCR or PCR			Reference sequence accession no.
			Primers			
			Name	Sequence	Source	
PEMV1	Enation and mosaic	JKI-1841	HZ-355 r	5′-TCA GAA ATG ACG CCG GAA CA-3′	This study	NC_003629
			HZ-356 f	5′-GCG GAA CAA CCT GTC TCT GA-3′		
PEMV2			HZ-363 r	5′-GTT GTG CGT CCT CTT GGA GA-3′	This study	NC_003853
			HZ-366 f	5′-CCC AAG GAG GTG TCC ATG TC-3′		
WHAV2	NA	NA	HZ-576 r	5′-AGC GGA CAT CGA GAA AGC TT-3′	This study	NC_028382, NC_028383, NC_028386 and NC_028387
			HZ-577 f	5′-GAG AGC GTC AAT GGT GGT GA-3′		
PNYDV	Yellowing, dwarfing and leaf rolling	JKI-1604	priPeaSdir	5′-AAC CTC CGG ATA TCA CCA GAT-3′	(Gaafar et al., 2016; Gaafar et al., 2017)	NC_023154 to NC_023161
			priPeaSrev	5′-CCG GAG GTT TTA TTT CAA AAC CAA C-3′		
VCV	NA	NA	HZ-658 r	5′-GCG AGT TCA TCC GAA TGC AC-3′	This study	NC_007241 and NC_007242
			HZ-659 f	5′-GCC TCC AGT GAA GGT TTC GA-3′		
PhCMoV	Discoloration, mild yellow spots, distortion and fruit marbling	JKI-2051	HZ-343 r	5′-CGG TGA GTG GGG CAA CTA AT-3′	(Gaafar *et al.*, 2018b)	KY859866
			HZ-344 f	5′-AGC GAT GGG GTC TAG TGT CT-3′		
PSTVd	Symptomless	NA	primer A	5′-CCC TGA AGC GCT CCT CCG AG-3′	(Weideman n & Buchta, 1998)	NC_002030
			primer B	5′-ATC CCC GGG GAA ACC TGG AGC GAA C-3′		

NA: Not applicable.
r: reverse.
f: forward.

4.3.5. Nucleic acid extraction and virus/viroid enrichment

Five grams of leaf tissue from each infected plant were ground in liquid nitrogen and stored at -80°C until further extraction. For extraction, 100 mg leaf materials were mixed 20 mg leaf disc from *P. vulgaris* infected with PvEV1 and PvEV2. The mix was used for three different RNA extraction methods (Fig 1):

a) Double-stranded RNA extraction (dsRNA):

dsRNA was extracted using Double-RNA Viral dsRNA Extraction Mini Kit (iNtRON Biotechnology, USA) according the manufacturer's protocol.

b) Total RNA extraction followed by ribo-depletion (ribo-depleted totRNA):

Total RNA extraction was performed using innuPREP Plant RNA Kit as described by the manufacturers' instructions. The ribosomal RNA (rRNA) was depleted using the RiboMinus™ Plant Kit for RNA-Seq (Invitrogen) according the manufacturers' protocol.

c) Total RNA extraction followed by small RNA extraction (sRNA):

Total RNA was extracted as described above, then DNase treated using innuPREP DNase I Digest Kit (Analytik Jena AG) according the manufacturers' protocol. sRNA was extracted using polyacrylamide gel selection at Fasteris Life Sciences SA (Plan-les-Ouates, Switzerland).

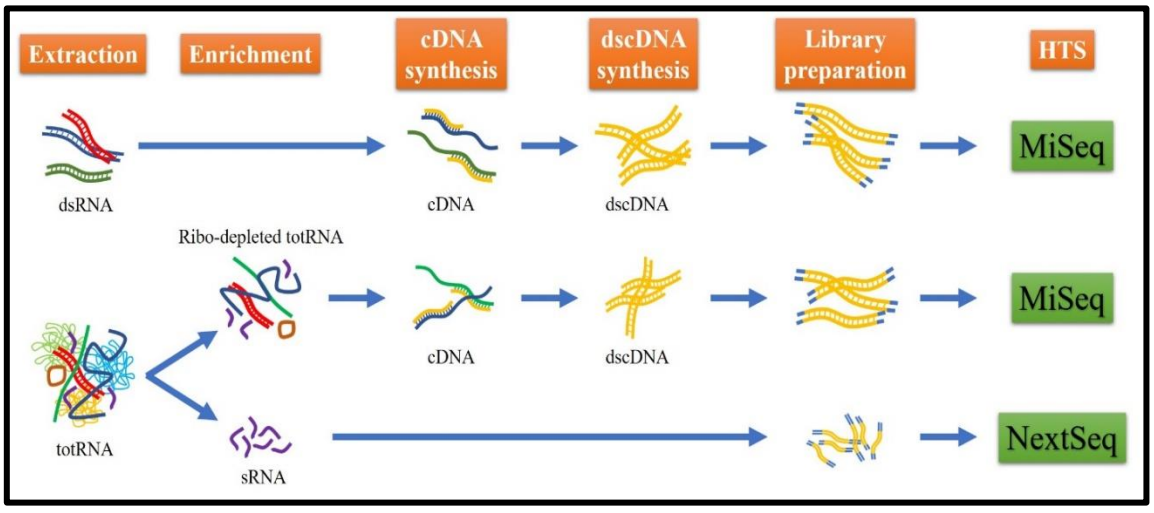

Figure 1: Graphical representation of the three RNA approaches used in this research i.e., dsRNA, ribo-depleted totRNA and sRNA. The steps are mentioned in orange boxes. The sequencing Illumina platforms are in green.

Additionally, DNA extraction followed by rolling circle amplification (RCA) was carried out for the nanovirus infected plants. Genomic DNA was extracted as described

before. The extracted DNA was treated by RNase A followed by RCA using TempliPhiTM 100 Amplification Kit (GE Healthcare Limited, UK).

4.3.6. Nucleic acid preparations for HTS

For dsRNA and ribo-depleted totRNA, random cDNA was synthesized using ProtoScript II Reverse Transcriptase and random octamer primers (8N). A denaturation step of 99°C for 2 min for the dsRNA and 65°C for 5 min for the ribo-depleted totRNA. ds-cDNA was synthesized using NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module (New England Biolabs). Libraries were prepared using Nextera DNA Library Prep Kit (Illumina) following the manufacturer protocol. The quantification was done using Qubit dsDNA HS Assay Kit (Life Technologies) and quality analysis was done using High Sensitivity DNA Chips on Agilent 2100 Bioanalyzer (Agilent Technologies) following the manufacturers' protocols. Subsequently, the libraries were sequenced on a MiSeq Illumine platform v.3 pair-end reads (2x301) at DSMZ, Germany. For the sRNA, libraries were prepared from sRNA extracted using TruSeq small RNA kit (Illumina) at Fasteris Life Sciences SA (Plan-les-Ouates, Switzerland) and sequenced on a NextSeq Illumine platform single-end reads (1x50). For the RCA products, the library was also prepared using Nextera DNA Library Prep Kit and run on a NextSeq Illumine platform (2x151) at DSMZ.

4.3.7. Bioinformatic data analysis

The data analysis was performed using Geneious (version 11.1.5) (Biomatters Limited, Auckland, New Zealand). The adaptors and low-quality nucleotides were trimmed from the raw reads (quality score set to 0.05), then the trimmed reads were filtered by length (100 to 301nt for dsRNA and ribo-depleted totRNA; 20 to 24nt for sRNA). The filtered trimmed reads were *de novo* assembled using Geneious (parameters; Medium Sensitivity/Fast). Moreover, sRNA reads were also assembled using Velvet (*kmer* = 13/ minimum contig length = 30) (Zerbino & Birney, 2008).

The filtered-quality trimmed reads were also *kmer* normalised using BBNorm tool 37.64 (Brian Bushnell within Geneious) (parameters: Minimum depth = 5/ Target coverage level = 40 for MiSeq reads, and Minimum depth = 5/ Target coverage level = 100 for NextSeq reads). After that, the reads were *de novo* assembled as described above.

Assembled contigs were compared against a local database for viruses and viroids reference sequences using BLASTn (maximum E-value: 1e-5) downloaded 18 August 2018. To confirm the virus/viroid presence in each sample, the contigs were mapped to references (accession no. in Table 2). Additionally, filtered and trimmed reads were mapped to the reference sequences (Geneious; Medium Sensitivity/Fast and 5 iterations for dsRNA and ribo-depleted totRNA, and Medium-low Sensitivity/Fast and 5 iterations for sRNA). A cut-off for virus/viroid detection was set at ≥ 40% recovery of the reference sequence (for viruses) and ≥ 80% (for viroids). The consensus sequences were generated

from the quality trimmed reads by mapping to reference sequences. The results were manually inspected to refine the ends of the genomes and consensus sequences were generated based on the quality of the nucleotides.

Pairwise nucleotide alignments were performed with ClustalW 2.1 (Cost matrix: CLUSTALW/ Gap open cost: 15/ Gap extend cost: 6.66) on Geneious. While for protein alignments Clustal W 2.1 with parameters (Cost matrix: BLOSUM/ Gap open cost: 10/ Gap extend cost: 0.1).

4.3.8. Comparing the three RNA-based approaches

The quality-controlled reads of each dataset were randomly subsampled (10 replicates) into the same number of reads (equal to approximately the same number of nucleotides 1, 10, 20, 30, 40 and 50 million nt). Resulting in a total of 720 subsets, each was used for *de novo* assembly and mapping to its reciprocal consensus sequence generated from the total reads. The number of nucleotides matched the references, percentage of the reference sequence recovered, and mean depth were calculated for each. Furthermore, *de novo* assembly (Geneious parameters; Medium Sensitivity/Fast) was performed for each subset and the resulting contigs were mapped to the corresponding consensus virus/viroid sequence and the percentages of whole genome was generated.

4.3.9. Statistical analysis

The generated data from the bioinformatic analysis was statistically analysed using R version 3.5.1 (R Core Team, 2014). The number of nucleotides matched the references, percentage of the reference sequence recovered by both reads and *de novo* assembled contigs, and mean depth were statistically compared. The data were visualized by ggplot2 and VennDiagram packages (Chen & Boutros, 2011; Ginestet, 2011).

4.4. Results

4.4.1. Raw data

The statistics of the raw data generated from the HTS platforms of the three RNA-based approaches are mentioned in S1 Table. As we used restricted quality and length of the reads, part of the generated datasets was less quality than expected, these can be shown by the number of reads after trimming and filtering and by the mean read length. Nevertheless, the datasets were still used for bioinformatic analysis. Furthermore, the spiked internal control viruses PvEV1 and PvEV2 were detected in all the samples with all the different approaches in different amounts of reads and depths (S2 Table).

The costs of each approach on were calculated on average based on our experience and the prices until early 2019 (S3 Table). The average cost of the dsRNA

approach was about 307 Euros, the ribo-depleted totRNA cost 380 Euro, while sRNA approach cost about 348 Euro on average.

4.4.2. Virus/viroid detection

Using BLASTn search, all three known viruses and the viroid in this study were detected with all three RNA approaches i.e., dsRNA, ribo-depleted totRNA and sRNA. Surprisingly, other viruses were detected in sample 1 and sample 2. In sample 1, in addition to PEMV1, a divergent isolate of pea enation mosaic virus 2 (PEMV2) (Genus: *Umbravirus*/ Family: *Tombusviridae*) was detected using all three approaches, and a divergent strain of Wuhan aphid virus 2 (WHAV2) was detected by dsRNA and ribo-depleted totRNA approaches (Fig 2). In sample 2, in addition to PNYDV, Vicia cryptic virus (VCV) (Genus: *Alphacryptovirus*/ Family: *Partitiviridae*) was also detected by dsRNA and sRNA approaches (Fig 2). These results showed that all viruses (Known and unknown) were detected by the dsRNA approach while ribo-depleted RNA and sRNA approaches detected either WHAV2 or VCV, respectively. No associated RNAs or alphasatellites DNAs were identified in the different samples. The presence of PEMV2, WHAV2 and VCV were confirmed using RT-PCR and virus specific primers as listed in Table 2.

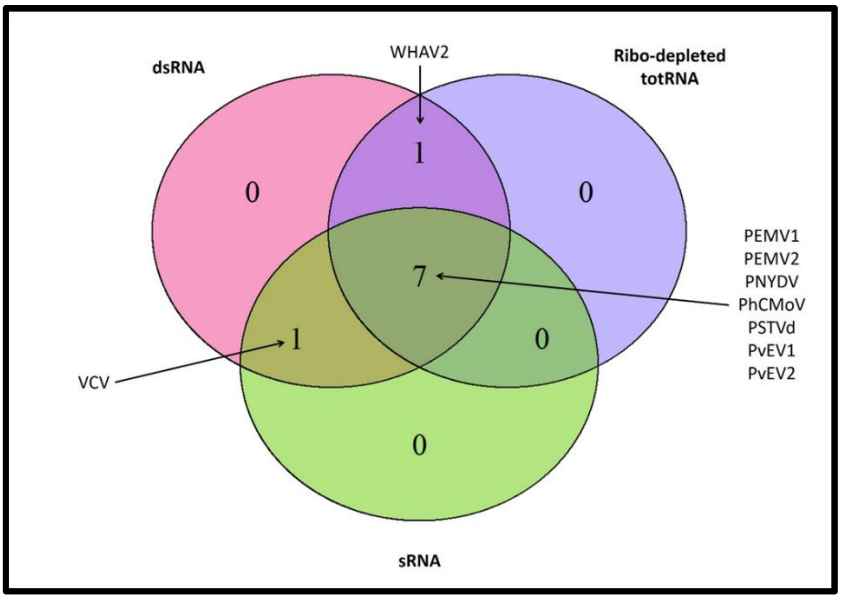

Figure 2: Venn diagram showing the viruses/viroid detected in all samples using different viral enrichment approaches (dsRNA, ribo-depleted totRNA and sRNA). The overlapping regions correspond to the number of viruses/viroid detected by more than one approach. The detected viruses were PEMV1: pea enation mosaic virus 1, PEMV2: pea enation mosaic virus 2, WHAV2: Wuhan aphid virus 2, PNYDV: pea necrotic yellow dwarf virus, VCV: Vicia cryptic virus and PhCMoV: physostegia chlorotic mottle virus, PSTVd: potato spindle tuber viroid, PvEV1: phaseolus vulgaris alphaendornavirus 1, and PvEV2: phaseolus vulgaris alphaendornavirus 2.

4.4.3. Virus/viroid recovery

The total filter and trimmed reads of each dataset were mapped to the different reference genome of the nine viruses in the samples (six viruses, a viroid and the two spiked internal control viruses). The total mapped reads and the percentage of the reference coverage can be found in Table 3. The numbers of mapped reads are different from one approach to another and from one virus to another. The full genomes of PEMV2, PhCMoV and PSTVd was recovered in the datasets by all three viral enrichment approaches. While the almost complete genomes of PEMV1 (missing the ends [not confirmed]) was detected by the three approaches. The almost complete genome of WHAV2 (missing the ends of some segments [not confirmed]) was identified by both dsRNA and ribo-depleted totRNA approaches. For VCV, less than 90% of the genome was assembled by ribo-depleted totRNA and sRNA approaches, but for PNYDV, less than 80% of the genome was assembled by the three approaches (Table 3). Therefore, to obtain the full genome of PNYDV, RCA enrichment was used. All virus sequences were submitted to GenBank (accession no. MK948524 to MK948543).

Table 3: Viruses detected in each sample by the different approaches (dsRNA, ribo-depleted totRNA and sRNA) in the total quality filter and trimmed data. Total number of mapped reads and nucleotides (nt), percentages (%) of recovered reference sequence and the mean depth.

Sample	Virus	dsRNA				Ribo-depleted totRNA				sRNA				Accession no.
		Total no.		% of ref.	Mean depth	Total no.		% of ref.	Mean depth	Total no.		% of ref.	Mean depth	
		reads	nt			reads	nt			reads	nt			
1	PEMV1	20,238	4,512,493	99.9	758.7	11,042	2,024,078	99.1	347.4	102,696	2,224,442	99.9	394.1	MK948533
	PEMV2	105,191	24,014,812	100	5,570.7	4,485	797,036	100	186.2	273,161	5,830,566	100	1,370.2	MK948534
	WHAV2	5,081	1,163,873	98.4	105.3	701	130,809	93.8	12	0	0	0	0	MK948535 to MK948538
2	PNYDV	156	39,512	77.2	4.9	214	38,587	72	4.6	9,929	215,596	63.5	27.1	MK948525 to MK948532
	VCV	200	51,391	89.4	13	0	0	0	0	3,157	67,266	87.5	17.4	MK948539 and MK948540
3	PhCMoV	43,724	9,737,744	100	725.3	210,559	37,885,237	100	3,750.4	369,715	8,081,770	100	606.8	MK948541
4	PSTVd	44	11,378	100	17.6	1,411	259,548	100	521.6	1,510,547	32,356,230	100	92,360.3	MK948524

4.4.4. Virus/viroid sequences characterisation

Pairwise nucleotide comparison between the sequences of PEMV1, PNYDV, PhCMoV and PSTVd showed nt identities range between 95.7% and 100% to their closest sequences. The French isolates of PEMV1 resulted in 95.7% identity with the ID isolate from Idaho USA (accession no. HM439775). The 8 segments of PNYDV Elbtal isolate shared between 97.9% to 99.9% nt identities to their closest sequences on NCBI. Segments DNA-N, -R, -S, -U4 shared 99.6% to 99.9% nt identities to the German isolate 110726 (accession no. KY810776 to KY810778 and KY810781). While segments DNA-C, -U1 and -U2 were close to the Danish isolate DK HZ16-582 by 97.9% to 99.4% (accession no. MH000257, MH000258 and MH000260), and segment DNA-M is closer to the Danish isolate DK HZ16-573 by 98.6% (accession no. MH000250). PhCMoV HZ16-558 isolate shared 99.6% nt identity to HZ15-192 isolate (accession no. KY859866). The PSTVd isolate shared 100% nt identity with isolates 6718566 from Netherlands and 07087900 from Belgium (accession no. KX370618 and FM998548, respectively). The sequences of PEMV2 and WHAV2 isolates were divergent from the reference sequences (S4 and S5 Tables). PEMV2 predicted proteins shared between 92.5% to 97.6% aa identities to their analogues of the closest isolate from the UK (S4 Table) whereas WHAV2 proteins shared 90.2% to 96.8% aa identities to the Chinese strain WHYC-2 predicted proteins (S5 Table). PvEV1 shared 99.7% nt identity to the Mexican isolate INIFAP CG1 (accession no. MG640415) and PvEV2 shared closest nt identity with a Brazilian isolate with 99.4% (accession no. AB719398).

4.4.5. The sensitivity of the three approaches (virus nucleotides and average depth)

The sensitivity of each approach for the detection of the viruses/viroid in this study was analysed using normalised subsamples (sizes: 1M, 10M, 20M, 30M, 40M and 50M nt). Each of the normalised subsamples from the different sizes was mapped to the consensus sequences of the detected viruses/viroid. The percentages of recovered references (here the consensus sequences) by the viral/viroid nt of each subsample are shown in Fig 3. With the dsRNA approach, the recovered percentages of the genomes reached 100% at 10M nt for PEMV2 and PhCMoV, whereas PEMV1 recovered at 20M nt. In cases of PSTVd, the full genome was recovered in some replicates from 20M to 50M nt, but the means were lower than 80%. However, for the segmented viruses WHAV2, PNYDV and VCV, the whole genomes could not be recovered even when using 50M nt. Using ribo-depleted totRNA, the full genomes of PEMV1, PEMV2 and PhCMoV recovered at 10M nt (Fig 3). The full genome of PSTVd was recovered at 1M nt. Similar to dsRNA approach, the genomes of WHAV2 and PNYDV were not recovered in all the subsamples. sRNA behaved similar to dsRNA approach with the viruses they both detected. PSTVd recovered at 1M nt, PEMV2 and PhCMoV at 10M, PEMV1 at 20M, and PNYDV reached about 60% of the genomes at 50M nt. Additionally, VCV also reached about 60% of the genomes at 50M nt.

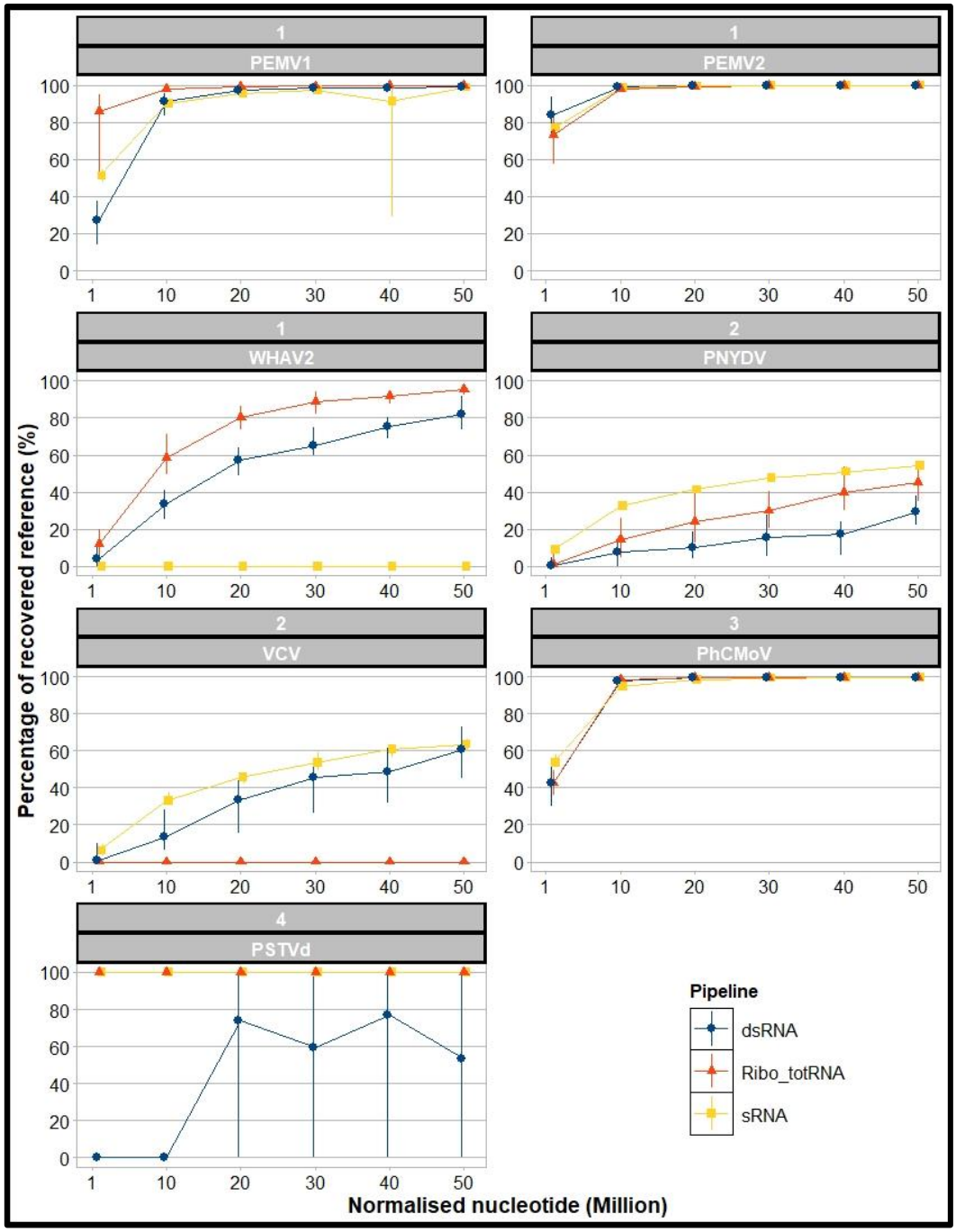

Figure 3: Percentage of reference sequences recovered by the reads of the RNA-based approaches on each of the normalised nucleotide subsamples (sizes: 1M, 10M, 20M, 30M, 40M and 50M nt). The means of each approach are shown as (blue circle: dsRNA, red triangle: ribo-depleted totRNA and yellow square: sRNA). The means are joined by lines with same colours. The vertical lines represent the standard deviation of the ten replicates. The strips over each graph are divided into two parts (upper: sample number, lower: virus/viroid acronym).

The dsRNA normalised subsamples had low variation for PEMV1, PEMV2, WHAV2 and PhCMoV, slight variation for PNYDV and VCV, and high variation in case of PSTVd (Fig 3). For ribo-depleted totRNA, all the replicates in all the viruses had low variation except

for PNYDV showed slight variation. sRNA had low variation except in one subsample of PEMV1 i.e., size 40M nt. The generated contigs by de novo assembly of each normalised subsamples showed the same results (S1 Fig). In Fig 4, the mean depth increased with the size of subsamples in all the three RNA approaches. Regarding the variation, it was the same as the percentages of recovered references of Fig 3.

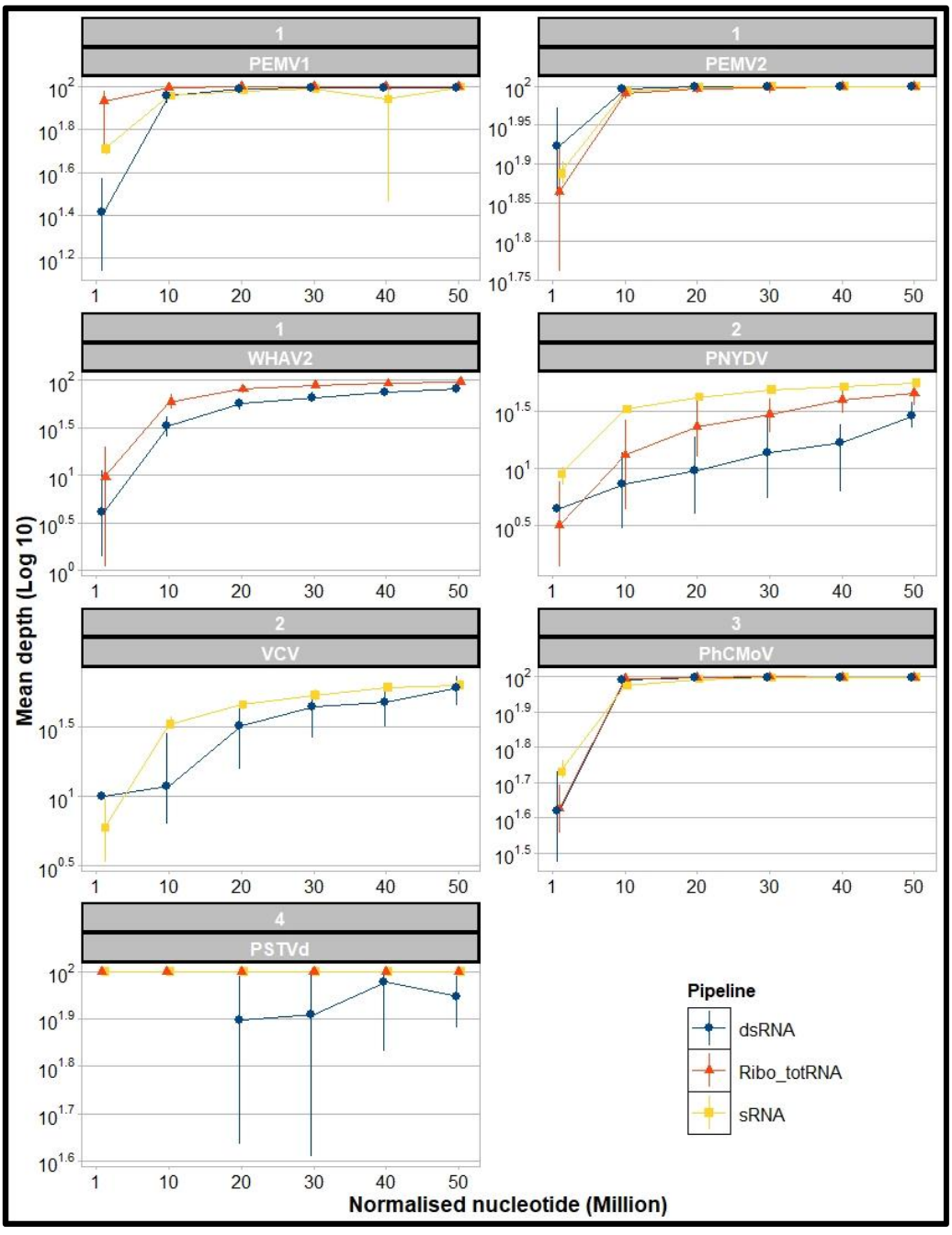

Figure 4: Mean depth of the RNA-based approaches for the detection of the viruses/viroid in this study on each of the normalised nucleotide depth (1M, 10M, 20M, 30M, 40M and 50M nt) for each subsample. The means of each approach are shown as (blue circle: dsRNA, red triangle: ribo-depleted totRNA and yellow square: sRNA). The means are joined by lines with same colours.

The vertical lines represent the standard deviation of the ten replicates. The strips over each graph are divided into two parts (upper: sample number, lower: virus/viroid acronym).

4.5. Discussion

The three viral enrichment approaches (dsRNA, ribo-depleted totRNA and sRNA) used here enabled the detection of the known and unknown plant viruses/viroid in the study. The efficiency of the approaches from extraction to analysis were confirmed by the detection of the spiked viruses (PvEV1 and PvEV2) in all samples by all approaches. dsRNA approach was more efficient than ribo-depleted totRNA and sRNA approaches in terms of virus detection. This is because all the eight viruses and the viroid in this study were detected by the dsRNA approach. While ribo-depleted totRNA and sRNA approaches, each detected seven viruses and the viroid. Three unknown viruses were detected i.e., PEMV2, VCV and WHAV2.

The detection of PEMV2 in mixed infection with PEMV1, is well documented (Doumayrou *et al.*, 2017). Both viruses' genome recoveries were high with the three approaches. VCV is a cryptic (symptomless) virus which occur at very low concentrations in infected tissues of several German varieties of *V. faba* (Blawid *et al.*, 2007). VCV was detected by the dsRNA and sRNA approaches. Furthermore, the ability of HTS to detect cryptic viruses was reported before in (Roossinck, 2011). Interestingly, the four segments of a divergent strain of WHAV2 was identified in sample 1 by dsRNA and ribo-depleted totRNA. WHAV2 was discovered in *Hyalopterus pruni* and *Aulacorthum magnoliae* aphids from Wuhan, Hubei province, China in 2013 (Li *et al.*, 2015; Shi *et al.*, 2016b). It has a segmented linear (+ve) ssRNA genome, its virion structure is unknown, and it is not assigned to a virus family yet. However, WHAV2 was phylogenetically grouped in the Jingmenviruses clade (Shi *et al.*, 2016b). Moreover, WHAV2 segments were also detected in three other pea samples collected from Germany and Austria in 2012 and 2013 (data not published). This considered the first detection of WHAV2 sequences in plant tissues and in Europe.

Concerning the recovered number of viral/viroid nt and the mean depth, in general, their amounts were different depends on the species, the sample and the approach. All viral/viroid nt were low in the data of the three approaches, this can be explained by a low virus titre in plant tissues at the time of sampling. As expected, the number of viral/viroid nt and the mean depth increased with the increase in the size of subsample. Same for the generated contigs, their genome coverage increased with the size of the subsamples.

Linear monopartite ssRNA genomes i.e., PEMV1, PEMV2 and PhCMoV, were detected by all approaches. When the viral nt were mapped to the single stranded linear non-segmented genomes (positive or negative), 10M to 20M nt were enough to recover the complete viral genomes. The reasons for the complete recovery at small amount of nt

that, dsRNA is generated by RNA viruses as an intermediate in their life cycle, additionally, the removal of ribosomal RNA increases the viral/viroid reads. Regarding viral small RNAs, they are known to be produced during RNA silencing defence of infected plants.

Segmented viruses, regardless of the genomic nature (nucleic acid type, and size and number of the segments), did not recover totally by the three approaches up to 50M nt. The DNA virus in this study, PNYDV, was not recovered totally by the three RNA approaches, with a slightly higher nt in case of sRNA. As PNYDV has a circular ssDNA multipartite genome, mRNA is synthesised for protein translation. Furthermore, a long dsRNA covering the full-genome is not produced as an intermediate replication by DNA viruses (Wu *et al.*, 2015). As the virus is a phloem restricted virus, a low titre is expected (Vetten *et al.*, 2011). These might be the reasons for the lower recovery of the virus nt. Same conclusion was suggested by (Pecman *et al.*, 2017). VCV was detected by the dsRNA and sRNA approaches. The number of viral nt in the two methods was not significantly different. The low concentration of cryptic viruses such as VCV and its dsRNA nature could be the main reasons that it was not recovered by ribo-depleted totRNA approach. Specially that the denaturation step of dsRNA was 65 ^{o}C for ribo-depleted totRNA approach. WHAV2 might be an aphid virus that is circulating in plant tissues as a vector, this may explain its low titre (Shi *et al.*, 2016b).

For the viroid PSTVd, as it has a short nt sequence (360nt), 1M nt was enough in cases of ribo-depleted totRNA and sRNA, but not for the dsRNA, it required more than 20M nt. Higher viroid titres in infected tissues may reflect a higher viroid ssRNA and dsRNA (produced during replication) concentrations. Both RNA forms can serve as templates for the Dicer of the RNA silencing (Markarian *et al.*, 2004). These may explain the very high PSTVd nt recovered by the sRNA approach than by dsRNA and ribo-depleted totRNA.

dsRNA extraction consumed less time (<1 hour) in compare to total RNA followed by ribodepletion (about 4 hours) and total RNA followed by sRNA extraction (about 6 hours). Furthermore, the costs of dsRNA extraction per sample was less than the other two approaches (S2 Table). The costs were calculated based on our experience and the prices until early 2019. The main difference between the costs was the cost of the enrichment step.

Additionally, the libraries were sequenced on Illumina platforms MiSeq for both dsRNA and ribo-depleted totRNA, while sRNA on a NextSeq platform. In general, the three approaches generate enough reads to detect the plant viruses and viroid in a given sample. Increasing the number of samples per lane will reduce the costs of platform runs as the generated read, this was also concluded by (Pecman *et al.*, 2017). Moreover, multiplexing by additional barcodes before library preparation can sequence more samples per run and consequently reduce the costs to a more comparable price (Roossinck *et al.*, 2010).

From these, we recommend sequencing dsRNA for plant virus/viroid detection by HTS, and at least 130 thousand of high-quality reads. The ability of dsRNA enrichment to detect ssRNA (positive and negative), dsRNA and DNA viruses, and viroids was reported before (Rott *et al.*, 2017; Rott *et al.*, 2018). Since most plant viruses produce an intermediate dsRNA or their genome consists of dsRNA and that dsRNA is very stable and can be easily purified, sequencing of dsRNA is therefore a very powerful method for detecting all virus types. There are different dsRNA extraction methods used for HTS. The reasons for using the dsRNA extraction kit mentioned here that it consumes less time (<1 hour), uses less milligrams (50 to 200) of plant tissue starting material, does not require PCR amplification in compare to other dsRNA methods extraction methods, in addition to its comparable cost per sample (Kesanakurti *et al.*, 2016; Yanagisawa *et al.*, 2016; Blouin *et al.*, 2016).

Regarding bioinformatic analysis, we recommend after quality control, reads normalisation, followed by *de novo* assembly of the reads, then BLAST search and mapping to reference based on the BLAST results. The results should be additionally confirmed by RT-PCR. These recommendations can be used as guidelines for viruses/viroids diagnostic.

The study further concludes the ability of HTS to detect known and unknown plant viruses and viroids. This study showed that the performance of the three RNA-based approaches is virus/viroid and sample dependent. We conclude that HTS generated data from the dsRNA approach outcompeted the ones generated from ribo-depleted totRNA and sRNA, and potentially can be used for the detection of all plant viruses and viroids. We also suggest comparing between the different available dsRNA extraction methods to reach the best method.

4.6. Supplementary

Table S1: Statistics of the raw data of the generated reads of the three RNA approaches in this study (dsRNA, ribo-depleted totRNA and sRNA) from the different samples

Approach		dsRNA				Ribo-depleted totRNA				sRNA			
Sample		1	2	3	4	1	2	3	4	1	2	3	4
Total raw reads		9,950,706	3,544,494	5,569,228	3,393,744	835,294	1,721,612	1,402,288	1,723,250	66,643,808	41,625,899	69,053,975	47,036,623
Raw data quality	≥Q20	90.1%	84.5%	91.6%	91.5%	96.20%	96.50%	96.10%	95.30%	94.60%	94.60%	94.80%	95.20%
	≥Q30	80.3%	71.7%	82.7%	81.9%	89.40%	89.90%	89.20%	87.20%	92.60%	92.80%	92.80%	93.50%
Quality and size filtered/trimmed reads		5,987,048	1,692,308	3,330,634	2,226,748	471,574	1,039,438	826,984	1,037,554	11,370,176	9,565,289	19,291,514	21,828,752
Read length Mean ± Std. Dev. (quality trimmed)		228.4 ± 65.0	258.7 ± 56.7	238.7 ± 63.5	255.0 ± 58.4	176.7 ± 49.7	180.1 ± 50.6	179.6 ± 50.2	187.7 ± 51.6	21.8 ± 1.3	22.0 ± 1.4	22.0 ± 1.4	21.7 ± 1.3

Table S2: Analysis of the spiked cryptic viruses phaseolus vulgaris alphaendornavirus 1 and 2 (PvEV1 and PvEV2) in each library of the four samples with the three RNA approaches

Virus	Sample	dsRNA				Ribo-depleted totRNA				sRNA				Reference
		No. of nt	no of reads	% of ref.	Mean depth	No. of nt	no of reads	% of ref.	Mean depth	No. of nt	no of reads	% of ref.	Mean depth	
PvEV1	1	5,082,751	22,050	99.1	356	1,745	10	6.6	0.1	24,455	1,122	43.3	1.7	NC_039217
	2	968,340	3,715	98.9	67.9	4,434	23	15.7	0.3	2,335	107	10.8	0.2	
	3	5,208,700	21,435	99.3	366.8	706	4	2.5	<0.1	72,613	3,345	41.6	5.16	
	4	2,724,640	18,048	99.5	322.9	645	4	2.2	<0.1	28,008	1,291	33.2	2	
PvEV2	1	2,881,894	12,654	99.9	190.5	596	4	2	<0.1	52,719	2,446	48.7	3.6	NC_038422
	2	451,820	1,733	99.2	29.5	1,902	10	5.3	0.1	4,264	197	16.6	0.3	
	3	2,925,625	12,124	100	193.1	2,611	6	4.4	0.1	188,153	8,749	55.4	12.7	
	4	4,689,607	10,491	99.7	179.3	1,258	8	4.9	<0.1	61,161	2,844	43.3	4.1	

Table S3: Pairwise comparisons of the nucleotide (nt) sequences and the amino acid (aa) sequence identities of Fr HZ11-065 isolates of PEMV1 and PEMV2 proteins with their closest known isolates

PEMV Fr HZ11-065	Nucleotides		Proteins			
	Identity	Reference	ORF		Identity	Reference
PEMV1	95.7%	HM439775	ORF1	hypothetical 34 kDa protein	96.0%	ADO86938
			ORF2	hypothetical protein	96.1%	ADO86939
			ORF3	RNA-dependent RNA polymerase	96.4%	ADO86940
			ORF4	coat protein	99.5%	ADO86941
			ORF5	aphid transmission protein	97.6%	ADO86942
PEMV2	93.5%	AY714213	ORF1	hypothetical protein	92.5%	ALP43778
			ORF2	RNA-dependent RNA polymerase	95.6%	AAU20330
			ORF3	phloem RNA movement protein	96.5%	AAU20331
			ORF4	cell-to-cell RNA movement protein	97.6%	AAU20332

Table S4: Pairwise comparisons of the four segments' nucleotide sequences and the amino acid (aa) sequences of the predicted proteins of the French isolate of WHAV2 (Fr HZ11-065)

WHAV2 Fr HZ11-065	Nucleotides		Proteins			
	Identity	Reference	ORF		Identity	Reference
Segment 1	86.4%	NC_028382	ORF1	NS5-like	96.8%	YP_009179378
Segment 2	90.7%	NC_028386	ORF1	VP4	96.9%	YP_009179384
			ORF2	VP1	90.2%	YP_009179385
Segment 3	85.2%	NC_028383	ORF1	NS3-like	95.5%	YP_009179379
Segment 4	85.6%	NC_028387	ORF1	VP2	92.0%	YP_009179386
			ORF2	VP3	96.6%	YP_009179387

Table S5: Average costs of the three approaches uses in the study. Costs presented in Euros.

Approach	dsRNA		Ribo-total RNA		sRNA	
	Method	Cost	Method	Cost	Method	Cost
Extraction kit	Double-RNA Viral dsRNA Extraction Mini Kit for Plant Tissue (iNtRON)	6.8	innuPREP RNA Mini Kit (Analytik Jena AG)	3.46	innuPREP RNA Mini Kit (Analytik Jena AG)	3.46
DNase treatment	NA	0	NA	0	innuPREP DNase I Digest Kit (Analytik Jena AG)	2.2
Enrichment	NA	0	Ribodepletion	77	sRNA polyacrylamide gel extraction	342.86
Kit	NA		RiboMinus Plant kit (Invitrogen)		Fastaris	
cDNA	ProtoScript II First Strand cDNA Synthesis Kit	4.55	ProtoScript II First Strand cDNA Synthesis Kit	4.55	NA	
dscDNA	NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module kit (NEB)	15.53	NEBNext Ultra II Non-Directional RNA Second Strand Synthesis Module kit (NEB)	15.53	NA	
Library	Nextera XT Library Kit (Illumina)/Eurofins GATC	280	Nextera XT Library Kit (Illumina)	280	TruSeq small RNA kit (Illumina)	
Sequencing platform	MiSeq (2x300 bp) No. of reads 2,000,000		MiSeq (2x300 bp) No. of reads 2,000,000		NextSeq (1x50 bp) No. of reads 50,000,000	
Total	306.88		380.54		348.52	

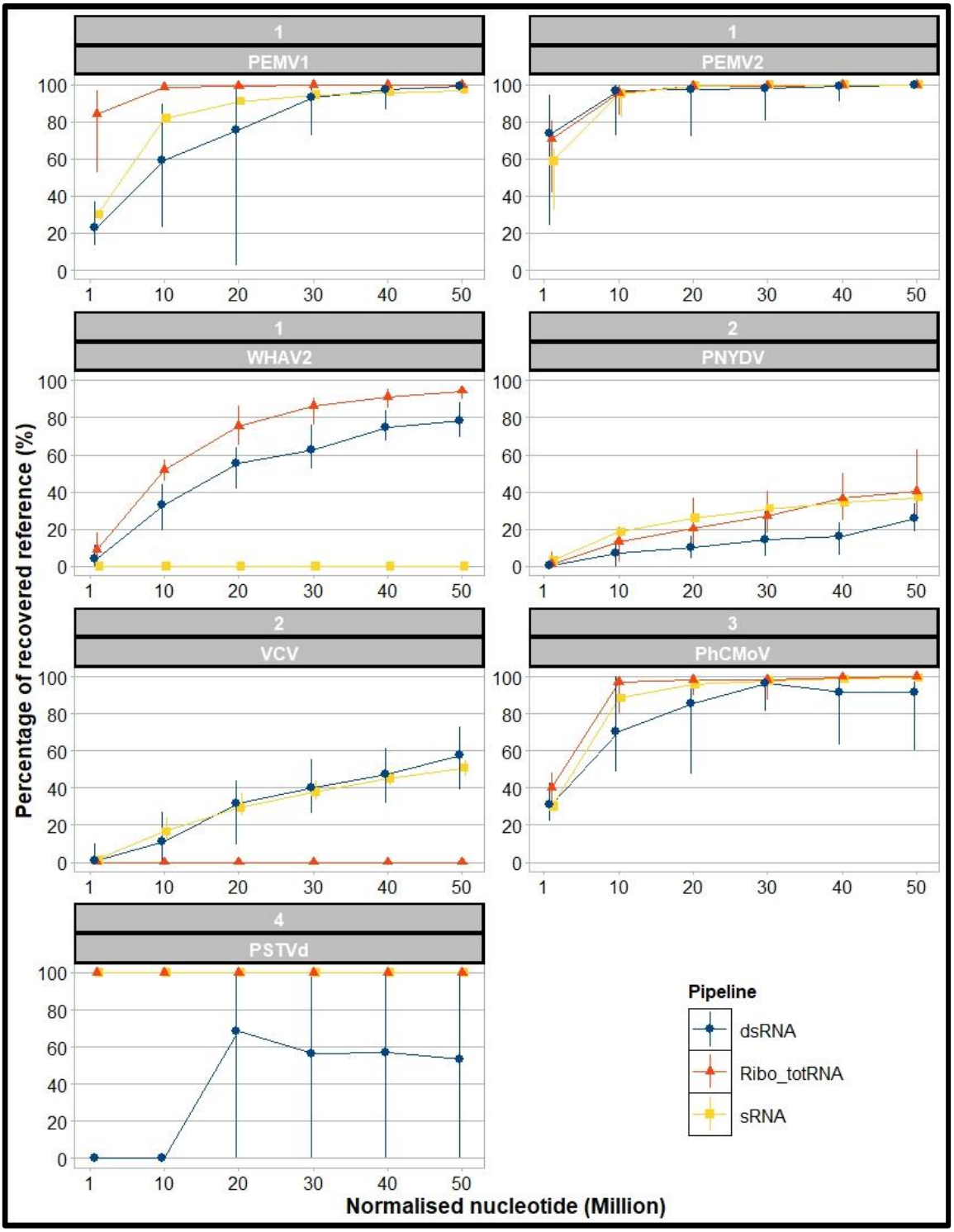

Figure S1: Percentage of reference sequences recovered by the produced contigs of the RNA-based approaches on each of the normalised nucleotide subsamples (sizes: 1M, 10M, 20M, 30M, 40M and 50M nt). The means of each approach are shown as (blue circle: dsRNA, red triangle: ribo-depleted totRNA and yellow square: sRNA). The means are joined by lines with same colours. The vertical lines represent the standard deviation of the ten replicates. The strips over each graph are divided into two parts (upper: sample number, lower: virus/viroid acronym).

Part two: Virus-vector-host interactions

Chapter 5: Aphid transmission of nanoviruses: a review

Yahya Zakaria Abdou Gaafar and H. Ziebell

5.1. Abstract

The genus *Nanovirus* is composed of plant viruses that predominantly infect legumes and can cause devastating crop losses. Nanoviruses are vectored by various aphid species. The transmission occurs in a circulative, non-propagative manner. It was long suspected that a virus-encoded helper factor would be needed for successful transmission by aphids. Recently, this helper factor was identified as the nanovirus-encoded nuclear shuttle protein (NSP). The mode of action of NSP is currently unknown – in contrast to other helper factors that e.g., facilitate binding of virus particles to receptors within the aphids' stylets. In this review, we are summarizing the current knowledge about nanovirus-aphid vector interactions.

5.2. Introduction

Viruses are one of the main constraints for successful and sustainable crop production. Viral infections can reduce both yield and quality of the produce but may also cause total crop losses. Plant viruses can be transmitted from plant to plant via different routes of transmission; they can be transmitted mechanically (i.e., handling of plants or harvesting of fruits, root contacts, nutrient solutions within closed cropping systems etc.), by seeds, pollen, or vectors such as fungi, mites, nematodes, aphids and other insects. Phloem-feeding insects such as aphids (Hemiptera, Aphididae) are the most common vectors of plant viruses (Hogenhout *et al.*, 2008). Aphids can be found worldwide, and many aphid species are able to act as virus vectors for one or more virus species (Stevens & Lacomme, 2017). Half of the insect-vectored known plant viruses depend on aphids for their transmission (Nault, 1997).

Traditionally, four different mechanisms are used to describe virus transmission by aphids (Hogenhout *et al.*, 2008): non-persistent transmission (NP), semi-persistent (SP) transmission, persistent circulative (non-propagative) transmission (PC) and persistent propagative (PP) transmission. Viruses that are transmitted NP are retained in the stylets and can only be transmitted for a few minutes after acquisition; the ability to transmit the viruses are lost within a few minutes and upon insect molting. Most SP viruses are retained in the foregut while some are retained in the stylets. and transmission can occur minutes to hours to a few days after acquisition; again, viruses are lost upon molting. However, some SP transmitted viruses are retained in stylets (Uzest *et al.*, 2007). No latent period (the time between acquisition access period [AAP] and inoculation access period [IAP]) is required for the transmission of NP or SP viruses contrary to persistently-transmitted viruses (Nault, 1997; Hogenhout *et al.*, 2008). Persistently transmitted viruses are classified as either circulative (and mainly non-propagative) viruses and propagative viruses. Upon acquisition, these viruses circulate through the insect vectors, moving from the gut lumen into the hemolymph and from there into the salivary glands from which the virus particles can be inoculated into new plants upon feeding. In addition, while non-

propagative viruses are assumed not to replicate within their insect vectors, viruses transmitted in a persistently propagative manner do.

Most knowledge about virus translocation pathways of PC transmitted viruses within aphids comes from research on members of the *Luteoviridae* family (Garret *et al.*, 1993; Gray & Gildow, 2003). It is suspected that luteovirids enter the gut epithelium by endocytosis before being released into the hemocoel by (Figure 1) (Garret *et al.*, 1993; Gray & Gildow, 2003). For some luteovirids, such as beet western yellows virus, a minor read-through protein (RTP) has been shown to be involved in aphid transmission together with the coat protein (CP); determining whether the virions can accumulate in the midguts or in both midguts and hindguts of aphids (Brault *et al.*, 1995; Brault *et al.*, 2000; Brault *et al.*, 2005).

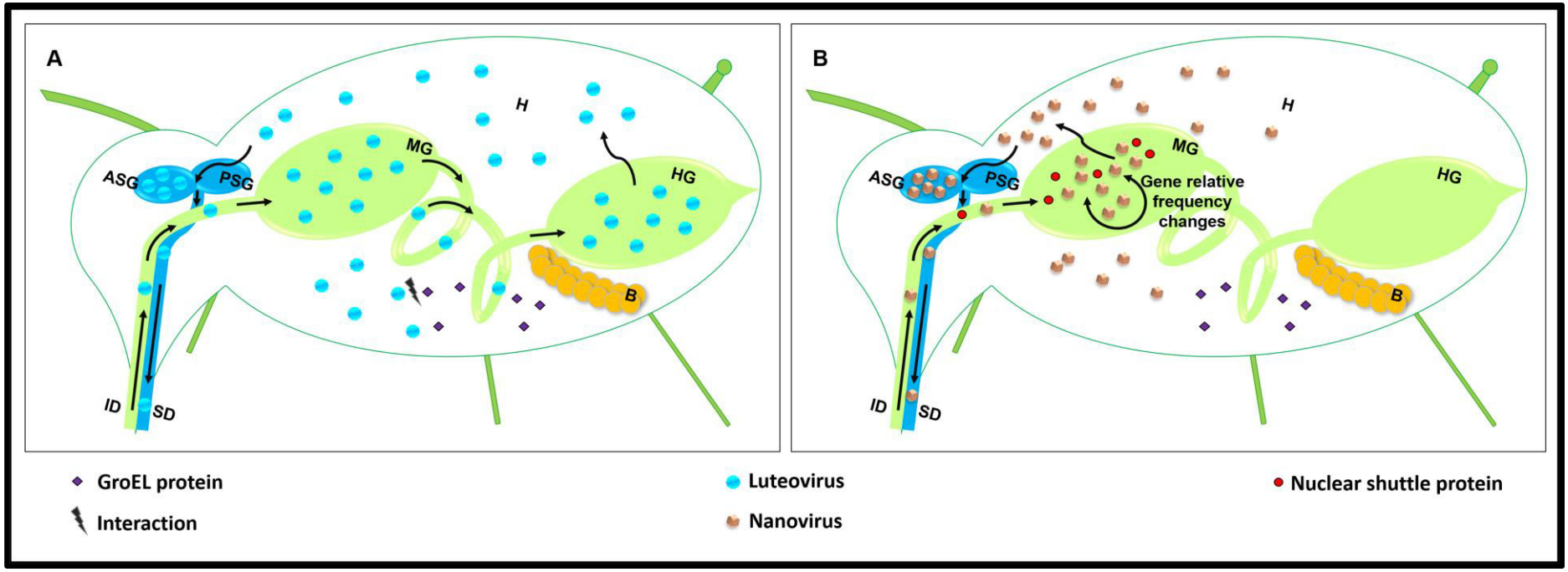

Figure 1. A representative diagram of luteoviruses (**A**) and nanoviruses (**B**) transmission within aphid's body. The arrows indicate the pathway of the virus particles. ASG: accessory salivary gland, B: bacteriocytes, H: haemocoel, HG: hindgut, ID: ingestion duct, MG: midgut, PSG: principle salivary gland, SD: salivary duct.

The virions are then acquired through into the hemocoel, regardless whether the aphid was a vector of particular species (Gildow, 1993; Gildow *et al.*, 2000). The virus particles are transported in vesicles through the cytoplasm of epithelial cells. Then the vesicles fuse with the basal plasmalemma and release particles into the interspace between the membrane and the basal lamina (Gray *et al.*, 2014). The virions move across the basal lamina into the hemocoel. Within four hours of the viroins' arrival in the alimentary canal lumen, they could be observed in the gut epithelial cells and within 8 hours they were detected in the hemolymph (Garret *et al.*, 1996). Luteovirids then move across the accessory salivary gland (ASG) cells in a pathway similar to that used to cross the gut (Gray & Gildow, 2003). In non-vector aphids, luteovirids are unable to penetrate the ASG suggesting that the basal lamina and the basal plasmalemma act as barriers to transmission (Gray *et al.*, 2014). With longer feeding periods, the number of epithelial cells containing the virions increased (Garret *et al.*, 1996). Mutational analysis of the CP and the RTP of turnip yellows virus (TuYV) and potato leafroll virus (PLRV), both poleroviruses, identified that they affected both aphid transmission and/or plant-virus-interactions (Bruyère *et al.*, 1997; Brault *et al.*, 2000; Lee *et al.*, 2005; Kaplan *et al.*, 2007; Peter *et al.*, 2008). CP was sufficient to transcytose virions of the gut to the hemocoel while it was suggested that the RTP facilitate the docking of the virions to the epithelial cells (Brault *et al.*, 2005; Chavez *et al.*, 2012). Moreover, the CP-RTP appeared to be required for interacting and passing through the ASG membranes (Bruyère *et al.*, 1997; Brault *et al.*, 2000; Peter *et al.*, 2008). Furthermore, Cilia and colleagues suggested critical virion-host protein interactions required for aphid transmission of cereal yellow dwarf virus (CYDV), or that the virus infection modulates phloem protein expression to favor aphid virus uptake (Cilia *et al.*, 2012). Investigating the particles of luteoviruses, which are transmitted in a circulative non-propagative manner, showed that they are transported across cells through membrane vesicles, preventing any contact between the virus and the cytoplasm of aphid's cells (Brault *et al.*, 2007). Yet there is no evidence that nanoviruses have the same manner.

5.3. Nanoviruses

The genus *Nanovirus* (family: *Nanoviridae*) currently comprises eight species accepted by the International Committee on the Taxonomy of Viruses (ICTV). Nanoviruses predominantly infect legumes which are important crops for human and animal consumption and are also used to improve soil health or as green manures (Johnstone & Mclean, 1987; Vetten *et al.*, 2011; Foyer *et al.*, 2016; Gaafar *et al.*, 2016). Two potentially new nanovirus species have been recently discovered from *Sophora alopecuroides* L. (also a legume) and *Petroselinum crispum* (Mill.) Fuss (belonging to the family of *Apicacea*) (Heydarnejad *et al.*, 2017; Vetten *et al.*, 2019). Nanovirus infection can cause a variety of different symptoms e.g., severe yellowing and dwarfing of plants or necrosis of plant tissues (Vetten *et al.*, 2011; Saucke *et al.*, 2019). In some cases, the infection can lead to early death of the host plants thus leading to crop failures (Grylls & Butler, 1956; Makkouk

et al., 1994; Makkouk *et al.*, 1998; Kumari & Makkouk, 2007; Saucke *et al.*, 2019). Nanoviruses have been found in Australia, Asia, Northern Africa and the Middle East (Grylls & Butler, 1956; Chu & Helms, 1988; Makkouk *et al.*, 1994; Sano *et al.*, 1998; Abraham *et al.*, 2012); a sporadic outbreak of faba bean necrotic yellows virus (FBNYV) has been reported in Spain (Ortiz *et al.*, 2006). In 2009, the nanovirus pea necrotic yellow dwarf virus was discovered for the first time in Central Europe (Germany) (Grigoras *et al.*, 2010a); in subsequent surveys this and other nanovirus species were discovered in Europe (Grigoras *et al.*, 2014; Gaafar *et al.*, 2016; Gaafar *et al.*, 2017; Gaafar *et al.*, 2018a; Vetten *et al.*, 2019).

Nanoviruses possess a multipartite single-stranded circular DNA genome of at least eight components (Vetten *et al.*, 2011). Each component is about 1kb nucleotides in size and encodes a single protein. Each DNA component is encapsidated separately in a single non-enveloped icosahedral virion (T=1 symmetry) (Vetten *et al.*, 2011). Several nanovirus-associated alphasatellite DNAs occur frequently with natural infection of nanoviruses; however, their biological function is unclear (do they impact infectivity and symptomology? Do they influence the host range or aphid transmissibility?) (Gaafar *et al.*, 2018a; Gallet *et al.*, 2018; Heydarnejad *et al.*, 2017). Nanoviruses can evolve rapidly by mutations, recombination and reassortment (Grigoras *et al.*, 2010b; Grigoras *et al.*, 2014) and are closely related to babuviruses, the second accepted genus in the *Nanoviridae* family (Vetten *et al.*, 2011). Only six genomic components are known for babuviruses, which infect banana and cardamom.

5.4. Aphid transmission of nanoviruses

Nanoviruses are restricted to the phloem of infected host plants and are therefore not transmissible by seeds or mechanical means (Vetten *et al.*, 2011). They are dependent on vectors for the transmission from plant to plant. In nature, they are transmitted by a range of aphid species, e.g., the pea aphid *Acyrthosiphon pisum* Harris, the cowpea aphid *Aphis craccivora* C.L. Koch or the bean aphid *A. fabae* Scopoli. It is unclear whether nanoviruses can be transmitted by other aphid species such as *Aphis gossypii* Glover or *Macrosiphum euphorbiae* Thomas (Vetten *et al.*, 2016). Nanoviruses are transmitted in a PC manner (Franz *et al.*, 1998; Vetten *et al.*, 2011). This means that aphids can only ingest the virus particles from infected plants when they are actively feeding on phloem sap. The viral particles need to cross the cellular barriers from the aphids' gut epithelium into the hemolymph; from the hemolymph, translocation to the salivary glands needs to occur (Figure 1) (Blanc *et al.*, 2014). The viruliferous aphid can then inject the particles with saliva during feeding on new plants.

For successful transmission of wild-type virus to new host plants, all eight genomic components need to be acquired, translocated within the vector and transmitted to new hosts. Interesting, Sicard and colleagues discovered that the relative frequencies of certain components of faba bean necrotic stunt virus (FBNSV) changed reproducibly, not

only depending on the host plant but also within different aphid vectors (Sicard *et al.*, 2013; Sicard *et al.*, 2015). It appears that these changes occur in early stages of the virus cycle in aphids but there were no further changes in the segment relative frequencies thereafter when the virions are translocated into the salivary glands (Sicard *et al.*, 2015). These frequencies were affected partially by the initial genomic formula from the source plant. These effects were observed in three FBNSV vector species *A. pisum*, *A. craccivora* and *Myzus persicae* Sulzer. Interestingly, similar amounts of FBNSV gene copy numbers (GCN) accumulated and changes in the FBNSV genomic formula were also observed in the non-vector species *A. gossypii* (Sicard *et al.*, 2015).

The most detailed transmission studies have been carried out with FBNYV. FBNYV infects several food and fodder legumes in west Asia and North Africa and Europe, and has caused huge epidemics in the past (Makkouk *et al.*, 1994; Kumari *et al.*, 2009; Ortiz *et al.*, 2006). Franz and colleagues investigated the minimum AAP of FBNYV for two vector species, *A. craccivora* and *A. pisum*. They found that the AAPs were similar for both species, ranging from 15 to 30 minutes. Additionally, the minimum IAP was determined for both aphid species and ranged between 5 and 15 minutes. However, longer AAPs and IAPs resulted in higher virus transmission rates. After 72 hours of feeding on infected plants, many aphids retained and transmitted FBNYV throughout their life in an erratic manner (Franz *et al.*, 1998). FBNYV particles were not lost during molts, i.e., the aphids remained viruliferous for up to one month, but FBNYV was not passed on to parthenogenetic offspring and no indication for viral replication within the aphids was found. Interestingly, nymphs were more efficient in transmitting FBNYV than adult aphids, a phenomenon also observed in other persistently transmitted viruses (Simons, 1953; Zhou & Rochow, 1984; Damsteegt & Hewings, 1986; van den J. M. Heuvel, 1991).

In further studies, Franz and colleagues have shown that no transmission occurred when using purified virions from artificial diets or directly microinjected into the vector's hemocoel would lead to successful transmission of FBNYV (Franz *et al.*, 1999) It was, however, possible to infect plants directly with purified virions using gold particle bombardment (Franz *et al.*, 1999). These bombarded plants acted as reservoir of FBNYV for further successful virus acquisition by aphid and FBNYV transmission to uninfected target plants (Franz *et al.*, 1999). Franz and colleagues (1999) suggested that a helper factor (HF) was required for successful aphid transmission and that the HF was either absent or non-functional in purified virus solution. Their hypothesis was supported by complementation experiments using two different FBNYV strains: when feeding on plants infected with one strain, acquisition and transmission of the second strain from artificial diets was possible (Franz *et al.*, 1999). Similarly, microinjection of a second, purified, strain into aphids that had fed on plants infected with the first strain led to successful transmission. The authors speculated that the HF would be an intermediate viral protein, attaching the virus capsid protein (CP) to receptors in the vector to facilitate virus

transport across the hemolymph - salivary gland interface (Franz *et al.*, 1999). However, at the time they were unable to identify the nature of the HF.

The availability of infectious clones for nanoviruses including FBNYV and FBNSV allowed to study the functions of the genomic components of nanoviruses in more details (Timchenko *et al.*, 2006; Grigoras *et al.*, 2009; Grigoras *et al.*, 2018). Using agroinoculation of all eight components of FBNSV, plants showed severe symptoms 10 to 14 days post inoculation (Grigoras *et al.*, 2018). Omitting one component at a time, no change in infectivity or symptomology was found for DNA-C (encoding a cell-cycle-link protein [C-link]) or DNA-U4 (encoding for a protein of unknown function). Omission of DNA-R (encoding the master replicase protein [M-Rep]), DNA-S (coat protein) or DNA-M (movement protein [MP]) resulted in the absence of infection on inoculated plants, therefore no symptoms could be observed. Omission of DNA-U2 or DNA-U1 (both encoding hypothetical proteins of unknown function) lead to reduced symptom severity, but virus transmission from inoculated plants by cowpea aphids was still possible. More strikingly, by omitting DNA-N (encoding for the nuclear shuttle protein (NSP)), plants became infected showing similar symptoms as when all eight viral components were inoculated, whereas the aphid transmission was completely abolished (Grigoras *et al.*, 2018). The virions that were produced within the plants inoculated with the seven components excluding DNA-N, accumulated to similar titers and were virtually morphologically and structurally indistinguishable from wild-type FBNSV particles. Introduction of a 13 amino acid tag at the carboxy-terminus of NSP also abolished aphid transmission. Interestingly, when the seven FBNSV components minus NSP were combined with DNA-N from a different nanovirus, pea necrotic yellow dwarf virus (PNYDV), the aphid transmission was restored indicating that this protein is the helper factor that has been previously proposed by Franz and colleagues (1999).

5.5. What do we know about NSP?

DNA-N of both nano- and babuviruses encodes the NSP. NSP of the babuvirus banana bunchy top virus (BBTV) was preferentially targeted to the nucleus of infected cells when expressed alone, but in the presence of the viral movement protein, NSP was relocalized to the cell periphery (Wanitchakorn *et al.*, 2000). Its function has been mainly inferred from comparisons with the homologue proteins of the closely related geminiviruses (Wanitchakorn *et al.*, 2000; Krapp *et al.*, 2017). Supposedly, NSP shuttles replicated viral DNA out of the nucleus of infected cell. NSPs of geminiviruses i.e., cabbage leaf curl virus (CaLCuV), tomato golden mosaic virus (TGMV) and tomato crinkle leaf yellows virus (TCrLYV) were found to interact with NSP-interacting kinases (NIKs) in vitro (Fontes *et al.*, 2004; Mariano *et al.*, 2004). NIKs are leucine-rich-repeat (LRR) receptor-like-kinases (RLKs) and are membrane localized protein. LRR-RLKs are involved in plant developmental processes and/or resistance response (Gómez-Gómez & Boller, 2000; Jinn *et al.*, 2000; Fontes *et al.*, 2004). Fontes and colleagues found a positive correlation between infection rate and loss of NIK function (Fontes *et al.*, 2004). *In vitro* binding

between the NSPs and NIKs inhibit NIK kinase activity and prevent the signal transduction pathway activation that would trigger an antiviral defense response (Fontes *et al.*, 2004; Carvalho *et al.*, 2008; Santos *et al.*, 2009; Santos *et al.*, 2010).

In infected faba beans, FBNSV-NSP could only be localized in the phloem-tissue where also FBNSV-CP was detected (Grigoras *et al.*, 2018). Using green fluorescence protein tagging and bimolecular fluorescence complementation (BiFC), Krenz and colleagues were able to localize NSP of PNYDV in the nucleus and in the cytoplasm of infected plant cells (Krenz *et al.*, 2017). Using BiFC, PNYDV-NSP was found to interact with the stress granule component G3BP which led to a speculation of the involvement of NSP in modulation of the plant stress response pathway (Krapp *et al.*, 2017). Additionally, NSP was found to interact with the M-Rep, encoded by DNA-R, which may affect the virus infection cycle (Krapp *et al.*, 2017). Krenz and colleagues hypothesized that NSP may regulate the virus replication by interacting with M-Rep as in the case of the geminivirus REn (Hanley-Bowdoin *et al.*, 2013; Krapp *et al.*, 2017). Furthermore, the NSP was found to self-interact in BiFC experiment in a yeast two-hybrid assay (Krenz *et al.*, 2017).

A recent study on the closely related BBTV found that NSP was located in both the nucleus and the cytoplasm of infected *Nicotiana benthamiana* Domin (Ji *et al.*, 2019). The presence of BBTV-NSP affected the cellular distribution of BBTV-CP in colocalization experiments in planta. Co-immunoprecipitation verified the interaction between the BBTV-NSP and BBTV-CP suggesting that BBTV-NSP relocates BBTV-CP in infected cells (Ji *et al.*, 2019). The direct role of NSP in nanovirus acquisition, translocation within aphids and/or inoculation of uninfected plants is currently unknown. It is also unclear how NSP does interact with nanovirus virions or other viral proteins *in planta*, as these interactions were only shown in model plants.

5.6. Virus translocation within aphids

Using immunofluorescence, Watanabe and colleagues showed that BBTV coat protein (CP), encoded by DNA-S, localizes in the anterior midgut cells of *Pentalonia nigronervosa* Coquerel aphids (Bressan & Watanabe, 2011). The labelling intensity suggested high concentration accumulation in epithelial cells, but no accumulation could be observed in the posterior midgut or hindgut of aphids. In addition, accumulation was observed in principal salivary glands but not in the accessory salivary glands (Bressan & Watanabe, 2011) In further studies, Bressan and Watanabe used PCR and immunofluorescent assays to examine possible translocation pathways over time (Watanabe & Bressan, 2013). They observed a progressive internalization of BBTV from the gut lumen to the anterior midgut, where accumulation occurred, followed by translocation into the principal salivary glands via the hemolymph (Watanabe & Bressan, 2013). However, they suggested also an alternative route whereby direct movement of BBTV from the anterior midgut to the principal salivary gland would be possible.

To address these possibilities, they used co-labelling assays of BBTV and cellular compartments of the aphid vectors (Watanabe *et al.*, 2016). These experiments suggest that an endosome-independent process is used by BBTV for internalization through the gut tissue. In contrast to the cellular translocation mechanisms used by luteovirids, BBTV appears to use endocytosis-independent processes for internalization that does not include endosomes, clathrin- and caveolae-mediated endocytosis, phagocytic uptake or raft-mediated cytosis (Watanabe *et al.*, 2016). In transmission electron microscopy studies, large numbers of vesicles were observed in the anterior midgut of BBTV-carrying aphids but not in aphids that were reared on healthy banana plants (Vetten *et al.*, 2016)However, it needed to be confirmed that these vesicles contain BBTV particles.

Circulative plant viruses such as luteovirids and geminiviruses were found to bind to GroEL proteins, produced by endosymbiotic bacteria (*Buchnera aphidicola*) inhabiting their vectors (Munson *et al.*, 1991; Kliot & Ghanim, 2013). Such interaction seems to protect the virus particles from degradation in the aphids' haemocoel. To date, there is no evidence that nanovirids interact with the GroEL proteins. Although GroEL proteins from *Buchnera* were detected in the hemolymph of *P. nigronervosa*, no interactions with BBTV virions could be observed using immunocapture PCR, dot blot and far-western blot analyses (Watanabe *et al.*, 2013). Nanovirids translocate in large clusters of virions which may protect individual virions from degradation and could prevent them from interacting with proteins such as GroEL (Vetten *et al.*, 2011; Watanabe *et al.*, 2013; Vetten *et al.*, 2016). It is also possible that NSP assists in preventing degradation in the hemocoel. To date, the interactions of BBTV-NSP and the translocation of virions through the aphid vector have not been investigated.

5.7. Concluding remarks

Even though we have seen a huge advancement in nanovirus research in the recent years, many questions regarding the interactions of nanoviruses with their hosts and vectors remain. As more and more nanoviruses and nanovirus-associated satellites are being discovered, it is necessary to understand these interactions in order to prevent nanovirus epidemics as we have seen in the past. As with all plant virus diseases, no curative methods are available once a plant is infected; preventive measures such as planting virus-resistant varieties (if available) or preventing spread by vectors by decreasing vector population for example. It appears that the mode of nanovirus transmission by aphids is more complicated than previously thought. It is noteworthy that NSP has been described as helper factor necessary for successful transmission by aphids. However, its role and mode of action during the transmission process currently remains a mystery. Additionally, it needs to be confirmed whether on the self-interaction between NSPs found in *in vitro* studies play a role within the aphids or within the transmission process. We also do not know if other viral proteins such as the products of DNA-U1, -U2 and -U4 or the associated alphasatellites influence aphid transmission or virus-vector interactions. It was shown that M-Rep interacts with both the CP and NSP and may

therefore also influence virus-aphid interactions. It is even more important to identify the motifs which interact with the aphid and host cells for trafficking.

It appears that the mode of nanoviruses aphid transmission not as trivial than it was suggested before. Although both CP and NSP are required for successful transmission, the observation that the GCN of the different segments are host-dependent and can change in aphid vectors raises the question whether these "virus formulas" are a type of host/vector adaptation and whether these changes are necessary for virus transmission or not. Additional studies on the relation of these changes and the erratic retention and transmission of nanoviruses by aphids are clearly required.

Furthermore, studying aphid probing and feeding activities on nanovirus-infected plants (before, during, or after virus acquisition) is crucial. In addition, we need data on aphid fitness on nanovirus-infected plants as it is currently not known if and how nanovirus infection influences its vectors, something which has been shown for other viruses (Alvarez *et al.*, 2007; Ziebell *et al.*, 2011). There are many more research questions that need to be addressed in this exciting topic of nanovirus-vector-host plant interactions.

Chapter 6: Probing and feeding behaviours of *Acyrthosiphon pisum* change on nanoviruses-infected faba beans

Yahya Zakaria Abdou Gaafar, Stefan Vidal and H. Ziebell

6.1. Abstract

The probing and feeding behaviours of the pea aphid, *Acyrthosiphon pisum* (Hemiptera: Aphididae) on nanoviruses- infected faba bean (*Vicia faba*) were investigated using electrical penetration graph (EPG). We assessed the behaviours of *A. pisum* on *V. faba* each infected with a different nanovirus *i.e.,* FBNYV and PNYDV. Moreover, we studied the effect of DNA-N, the aphid transmission helper component, on the aphid probing and feeding behaviours. The results showed that nanovirus infection changes the behaviours of *A. pisum*, by making the plants less attractive to the aphids. Moreover, the absence of DNA-N did not affect the behaviour of the aphids.

6.2. Introduction

Viruses of the genus *Nanovirus* (family: *Nanoviridae*) are multipartite plant DNA viruses that infect predominantly legumes and can cause crop losses (Vetten *et al.*, 2011; Makkouk *et al.*, 2014). Nanoviruses have been reported from Asia, Australia, Europe, and Africa (Gutierrez *et al.*, 1971; Makkouk *et al.*, 1994; Ortiz *et al.*, 2006; Kumari *et al.*, 2009). The genus *Nanovirus* has eight assigned species recognized by the International Committee on Taxonomy of Viruses i.e., *Black medic leaf roll virus* (BMLRV), *Faba bean necrotic stunt virus* (FBNSV), *Faba bean necrotic yellows virus* (FBNYV), *Faba bean yellow leaf virus* (FBYLV), *Milk vetch dwarf virus* (MVDV), *Pea necrotic yellow dwarf virus* (PNYDV), *Pea yellow stunt virus* (PYSV) and *Subterranean clover stunt virus* (SCSV). Their genomes consist of eight circular ssDNA components (DNA-C, -M, -N, -R, -S, -U1, -U2 and -U4) of about 1kb each (Vetten *et al.*, 2011). Each component is encapsidated separately in non-enveloped virions (Vetten *et al.*, 2011). Additionally, several associated alphasatellite DNAs can also be detected (Gaafar *et al.*, 2018a; Gallet *et al.*, 2018).

Nanoviruses are phloem-restricted and are transmitted by aphids e.g., *Aphis craccivora* and *Acyrthosiphon pisum* (Franz *et al.*, 1998; Vetten *et al.*, 2016). Nanovirus aphid transmission occurs in a circulative non-propagative manner, though a recent study showed a more complex manner (Sicard *et al.*, 2015). When the aphids feed on infected plants, the viruses are acquired with the phloem sap, then translocate from the gut to the haemolymph to the salivary glands in unknown mechanism without replicating or expressing their genes. Sicard and colleagues found reproducible changes in the relative frequencies of some of the FBNSV components in *A. pisum*, *A. craccivora*, and *Myzus persicae* aphids compared to the host plants they fed on (Sicard *et al.*, 2015). These changes occurred inside the gut but did not change later when the virus particles are in the salivary glands (Sicard *et al.*, 2015). Thus, they suggested that the canonical circulative non-propagative transmission might not fit with nanoviruses.

Franz and colleagues suggested that a virus helper factor (HF) is required for FBNYV aphid transmission (Franz *et al.*, 1999). They also suggested that the HF might attach the nanovirus coat protein to receptors in the aphids (Franz *et al.*, 1999). Grigoras

and colleagues found that the nuclear shuttle protein (NSP) encoded by DNA-N is the required HF for PNYDV aphid transmission (Grigoras *et al.*, 2018). However, it is currently unclear at which step of the transmission pathway NSP is involved.

Several factors can influence the virus-aphid-plant interactions, including the chemical composition and physical structure of the plant (Guo *et al.*, 2014). It has also been shown that virus infection can affect the feeding behaviour of their vectors and therefore the transmission (Ziebell *et al.*, 2011). Electrical penetration graph (EPG) is a powerful tool for investigating feeding behaviour of aphids (Tjallingii, 1978, 1985). It has been used to demonstrate that the feeding behaviour of non-viruliferous *Myzus persicae* was enhanced on potato leafroll virus (PLRV) infected potato plants (Alvarez *et al.*, 2007).

With the recent advancement of nanovirus research, we were interested whether nanovirus infection would alter the feeding behaviour of non-viruliferious aphids. We therefore aimed to study the effects of nanoviruses infection on aphid probing and feeding behaviours using the EPG technique, comparing the effects of two nanoviruses (FBNYV and PNYDV) on the feeding and probing behaviour of *A. pisum* on faba beans (*Vicia faba*). Moreover, to assess the effect of DNA-N absence on these behaviours.

6.3. Materials and methods

6.3.1. Plants and aphids:

V. faba plants (variety: Tattoo) were grown in a greenhouse at 16/8 h light/dark (natural daylight with additional growth light Phillips IP65, 400 W) and 22°C. In all experiments, 11 days old seedlings were used.

The pea aphids, *A. pisum* (Harris) (JKI clone), used in this study were reared continuously in an insect-proof cage on faba bean plants in a climate chamber at 22°C and 16/8 light/dark (sodium high pressure lamps). Fresh plants were added frequently to the cage to prevent overcrowding.

To synchronize the aphids for the experiments, 10 viviparous adults were placed on *V. faba* plants in a cage in greenhouse chamber at 22°C 16/8 h light/dark. After 24 hours, the adults were removed, and the newly born nymphs were left for about 13 days. This was repeated daily to have enough aphids of almost the same age (about 24 h difference) for each day of recording.

6.3.2. Endosymbiont detection and confirmation:

DNA was extracted from three sets of 10 aphids collected in a 2ml tube according to (Shahjahan *et al.*, 1995). PCR was performed using primers for *A. pisum* endosymbionts (Tsuchida *et al.*, 2002).

6.3.3. Virus isolates maintenance and confirmation:

Three nanoviruses were used in this study i.e., faba bean necrotic stunt virus (FBNSV) isolate JKI-2000 from Ethiopia, faba bean necrotic yellows virus (FBNYV) isolate AZ originally from Azerbaijan (Grigoras *et al.*, 2014), and pea necrotic yellow dwarf virus (PNYDV) isolate Drohndorf-15 originally from Germany (Grigoras *et al.*, 2010a). FBNYV and PNYDV were propagated by aphid transmission using *A. pisum* from infected plants. For FBNSV, agrobacteria containing infectious clones constructed as described in (Grigoras *et al.*, 2009) (kindly provided by Prof. Stephane Blanc) were agro-inoculated to *V. faba* plants as described in (Timchenko *et al.*, 2006).

For all three viruses, infection was confirmed by ELISA (Gaafar *et al.*, 2016; Gaafar *et al.*, 2017). However, for confirmation of FBNSV infection an additional PCR was carried out using FBNSV specific primers for DNA-R and DNA-N as described in (Grigoras *et al.*, 2018).

6.3.4. Electrical penetration graph (EPG) monitoring:

Adult apterae *A. pisum* were starved for about 1 hour before the recording. Each aphid was immobilized using a vacuum-operated plate under a binocular microscope. A gold wire (insect electrode), 18µm in diameter and about 3cm long, was glued to the aphid dorsum using a small drop of water-based silver glue and connected to the EPG probe (EPG Systems, Wageningen, The Netherlands). A copper wire, 0.2 cm in diameter and 10cm long, was inserted into soil near to the plant root (plant electrode). The two electrodes were connected to an eight-channel GIGA-8 direct current amplifier (EPG Systems). The wired insect was placed on the stabilised abaxial surface of a faba bean leaf. The signal was digitized (100 Hz) using a DI-710 board (Dataq Instruments, Akron, OH, USA) and analysed with Stylet+ d software (EPG Systems). The experiment was carried out in an electrically grounded Faraday cage at a room temperature maintained around 22°C. After starting the recording, the output voltage was optimized during the first probes for each channel by adjusting the plant voltage and gain (Giga manual; EPG Systems). The probing and feeding behaviours of individual aphids on faba bean plants was monitored for 9 hours.

6.3.5. EPG parameters:

The recorded waveforms were analysed with Assisted Analysis of Electrical Penetration Graph (A2EPG) software (Adasme-Carreño *et al.*, 2015). For the calculation of the EPG parameters, the Excel-VBA macro prepared by Prof. Edgar Schliephake was used (Schliephake *et al.*, 2013). Eighty-nine parameters were estimated and used for statistical analysis.

6.3.6. Statistical analysis:

The statistical analysis was carried out using scripts written on R software (version 3.5.1) (R Core Team, 2019). The EPG data were not normally distributed, thus non-

parametric Kruskal-Wallis test was used for analysis. For significant different parameters, Pairwise Wilcox test was used to know which pairs of groups are different (*P* value adjustment method: BH). Plotting was done using ggplot package.

6.3.7. Experimental design:

6.3.7.1. Effect of nanovirus infection on the behavioural responses of *A. pisum* during probing and feeding on faba beans:

Ten viruliferous aphids per plant were placed on 11-day old V. faba seedlings and let to feed on the plants for five days. The aphids were gently removed using water without any insecticides, and the plants were checked to confirm that no aphids (adult or nymph) were still present. Ten days after removing the aphid the plants were used for EPG experiments. Healthy control plants were sown the same time with the other plants and used for EPG experiments 26 days from sowing. At this stage, synchronised non-viruliferous adult aphids were used (with a maximum 24 hours age difference). After EPG recording, the exposed leaves were cut and used for ELISA confirmation tests. The recording was done for 9 hours, with all the four experimental parameters were present at the same time.

The four treatments in this experiment:

1- *A. pisum* feeding on FBNYV infected *V. faba* (I-FBNYV)

2- *A. pisum* feeding on PNYDV infected *V. faba* (I-PNYDV)

3- *A. pisum* feeding on *V. faba* pre-treated with *A. pisum* (H-Aphid)

4- *A. pisum* feeding on healthy *V. faba* without treatments (Healthy)

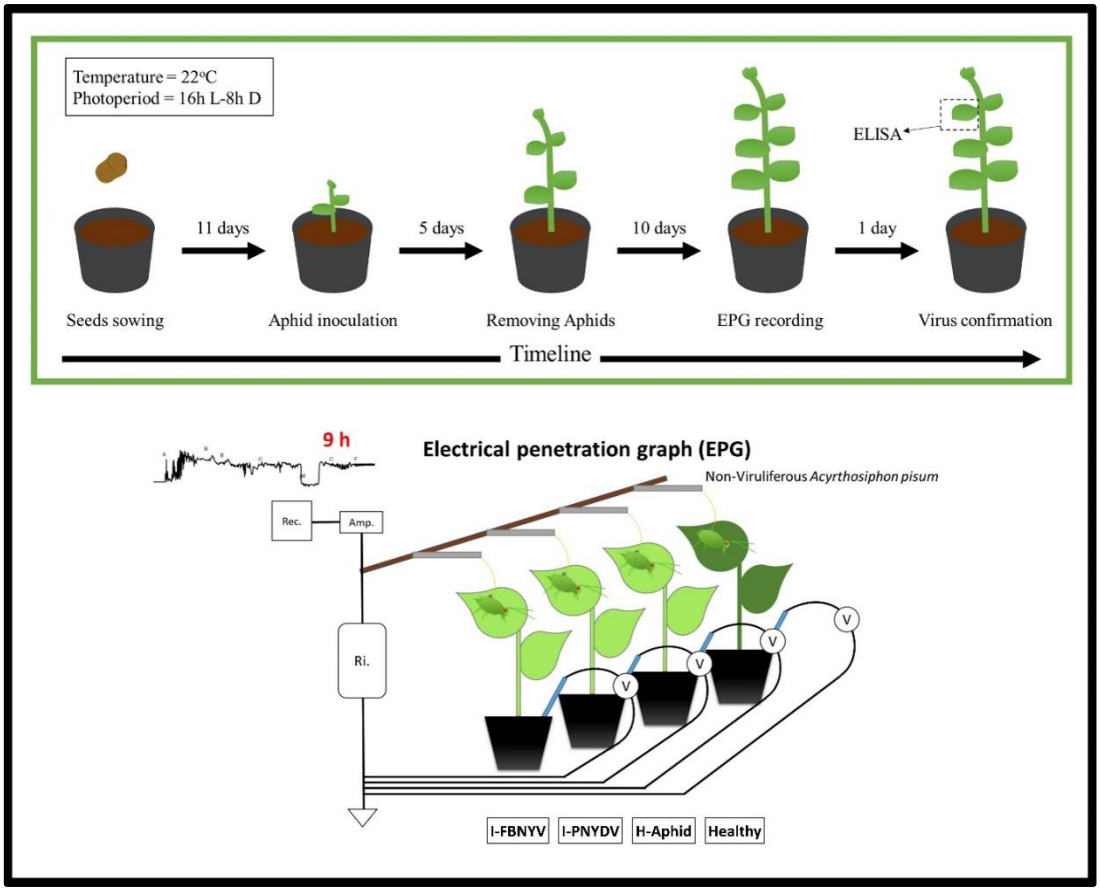

Figure 1: Experimental design to investigate the effect of nanovirus infection on probing and feeding behaviour of A. pisum on faba beans. I-FBNYV: FBNYV infected plants, I-PNYDV: plants infected with PNYDV, H-Aphid: plants pre-treated with aphids, and Healthy: plants without treatments.

6.3.7.2. Effect of the nuclear shuttle protein (NSP) of FBNSV on the behavioural responses of A. pisum during probing and feeding on faba bean:

After 11 days from seeds' sowing, the plants were agroinoculated as described before using these parameters: a) a mix of agrobacteria carrying all eight FBNSV components, b) a mix of agrobacteria carrying seven components of FBNSV (missing DNA-N), and c) agrobacteria not containing any FBNSV infectious clones. Fifteen days from agroinoculation, the plants were used for EPG. Healthy control plants were sown the same time with the other plants and used for EPG 26 days from sowing. Non-viruliferous adult aphids were used. After recording, the leaves where the aphid probed were cut and used for ELISA and PCR confirmation tests. The recording was done for 9 hours, with all the four experimental parameters present at the same time.

The four treatments in this experiment:

1- *A. pisum* feeding on *V. faba* agroinoculated with agrobacteria containing the infectious clones of the eight components of FBNSV (I-FBNSV)

2- *A. pisum* feeding on *V. faba* agroinoculated with agrobacteria containing the infectious clones of the seven components of FBNSV without component DNA-N (I-FBNSV/-N)

3- *A. pisum* feeding on *V. faba* agroinoculated with agrobacteria without any infectious clones (H-Agro)

4- *A. pisum* feeding on healthy *V. faba* without treatments (Healthy)

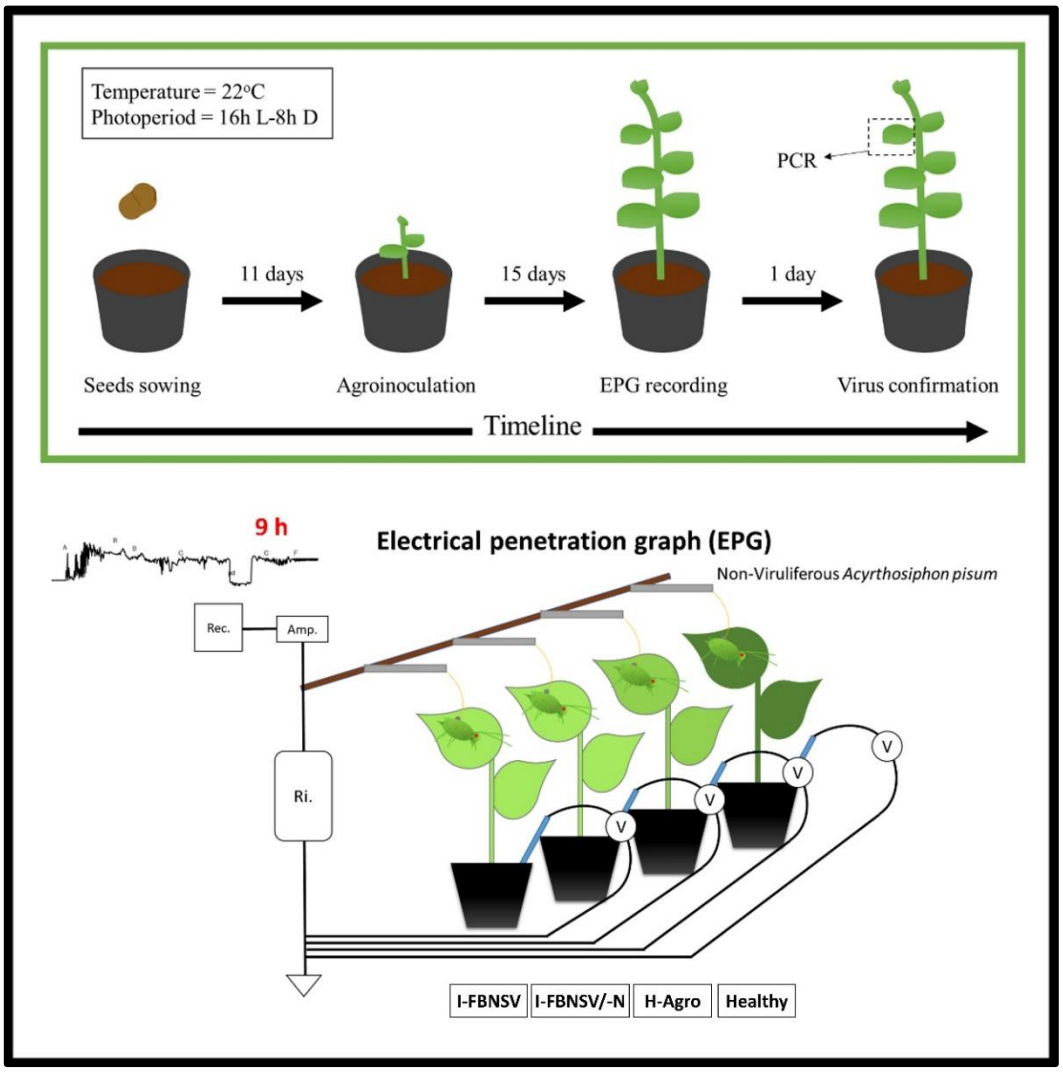

Figure 2: Graphical illustration of the experimental design for the effect of NSP absence experiment. I-FBNSV: plants agroinoculated with the eight components of FBNSV, I-FBNSV/-N: plants agroinoculated with seven components of FBNSV without DNA-N, H-Agro: plants agroinoculated with agrobacteria without infectious clones, and Healthy: plants without treatments.

6.4. Results

6.4.1. Endosymbiotic bacteria:

PCR analyses confirmed the presence of endosymbiotic bacteria in the JKI clone of the *A. pisum* colony. The bacteria belonged to the genera: *Buchnera*, *Hamiltonella* and *Spiroplasma*.

6.4.2. Virus-induced symptoms:

The faba bean plants infected with FBNYV isolate AZ showed strong symptoms of top leaf carling and hardening, and the plants were severely dwarfed compared to the controls at the time of the EPG recording. The plants infected with PNYDV Drohndorf-15 showed only slight yellowing at the time of recording. In addition, the agroinoculated plants with FBNSV and FBNSV/-N showed top leaf carling and the plants were more stunted than control plants. Plants that were agroinoculated without the virus infectious clones showed no obvious symptoms.

6.4.3. EPG probing and feeding behaviour:

6.4.3.1. Effect of nanovirus infection on the behavioural responses of *A. pisum* during probing and feeding on faba bean:

On the non-treated healthy controls (Healthy), *A. pisum* spent 93.1 % of the recorded time probing while the phloem activity was 67.5 % (Figure 3). The time to first probing was 4623.8 ± 776.7 sec (Table 1). The average number of pathways (n_C) was 25.4 ± 3.4 and the average number of intracellular punctures (n_pd) was 149.1 ± 21.8 (Table 1). The average total derailed stylet mechanics (s_F) was 1244.0 ± 535.2 sec (Table 1). The salivation period (s_E1) was 468.8 ± 160.7 sec and the time to start the first feeding (t_1E2) was 4011.9 ± 698.5 sec (Table 1). The average total duration of sustained (>10 min) phloem sap ingestion (s_sE2) was 21428.4 ± 1611.3 sec.

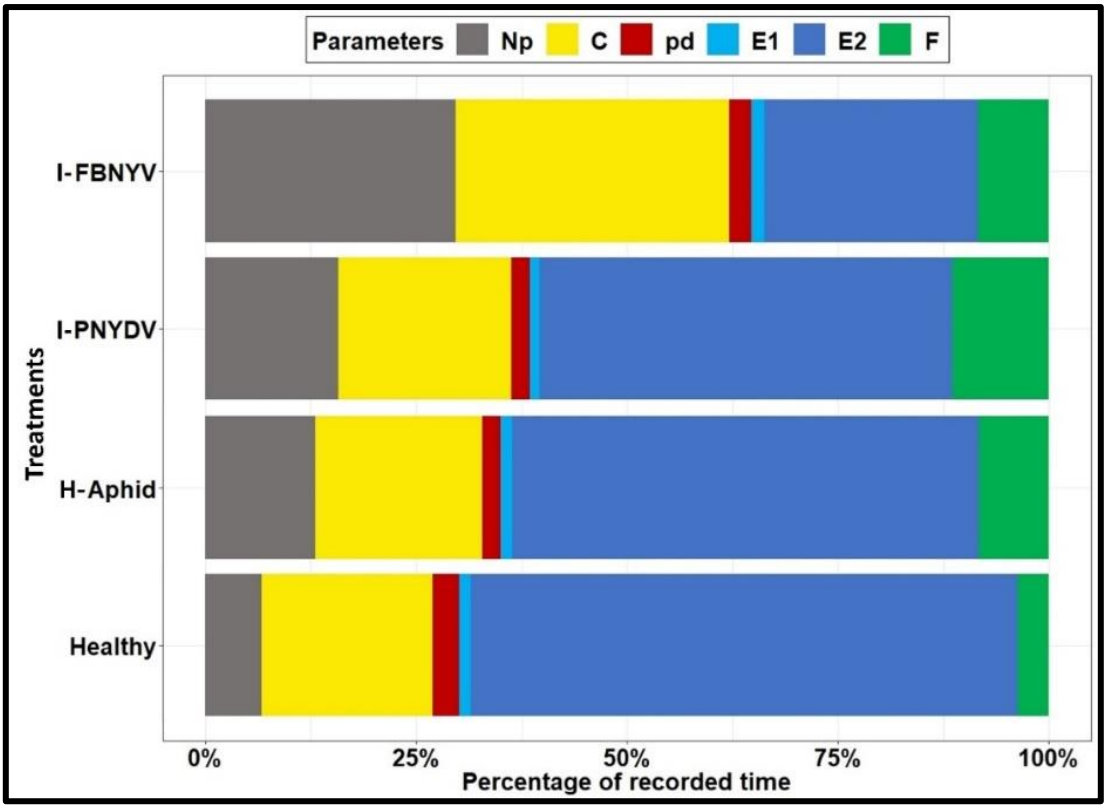

Figure 3: Percentages of the main parameters recorded for *A. pisum* feeding on FBNYV and PNYDV infected *V. faba* (I-FBNYV and I-PNYDV, respectively), aphid pre-treated healthy plants (H-Aphid) and healthy plants using EPG. The parameters are Np: non probing, C: pathway, pd: intracellular puncture, E1: salivation, E2: ingestion, F: derailed stylet mechanics.

Table 1: Statistical analysis of EPG parameters of *A. pisum* probing healthy faba bean plants and healthy plants pre-treated with *A. pisum* (H-Aphid) and faba bean plants inoculated with FBNYV and PNYDV (I-FBNYV and I-PNYDV, respectively) (experiment 1). The parameters are Np: non probing, C: pathway, pd: intracellular puncture, E1: salivation, E2: ingestion, and F: derailed stylet mechanics. n_ = the number of, s_ = the sum of (in seconds), and t_1 = time to first (in seconds).

Parameters	I-FBNYV		I-PNYDV		H-Aphid		Healthy		*P* value
	n=17		n=20		n=9		n=17		
	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error	
n_Np	54.9	6.5	34.1	4.1	33.8	10.4	21.6	2.8	0.00
s_Np	9878.4	1289.4	5250.7	819.2	4333.2	1101.3	2241.5	418.0	0.00
n_C	58.6	6.5	38.4	4.4	37.6	10.2	25.4	3.4	0.00
s_C	10791.1	1326.8	6768.5	763.1	6555.4	1735.4	6786.0	952.3	0.05
n_F	1.9	0.4	1.6	0.4	1.4	0.6	0.8	0.5	0.07
s_F	2794.2	673.3	3778.9	764.0	2747.0	916.5	1244.0	535.2	0.07
n_E1	3.1	0.5	4.1	0.4	4.4	1.2	4.7	0.6	0.25
s_E1	504.9	170.7	374.6	77.9	466.9	108.6	468.8	160.7	0.85
n_E2	1.6	0.4	3.0	0.3	3.0	0.6	3.6	0.6	0.01
s_E2	8431.2	2542.3	16226.3	1595.0	18293.8	2771.9	21656.9	1552.3	0.00
n_pd	167.7	20.9	115.6	11.4	112.2	24.5	149.1	21.8	0.20
s_pd	876.8	103.2	754.8	86.5	694.2	149.6	1049.9	169.6	0.34
t_1Pr	16020.0	2916.5	6097.3	836.0	6515.6	1010.8	4623.8	776.7	0.00
t_1E	106.9	41.0	65.0	8.1	54.0	10.5	88.0	16.8	0.41
t_1E2	11450.3	2250.2	5090.3	657.8	6369.7	987.6	4011.9	698.5	0.01
s_sE2	8269.8	2546.7	16100.2	1594.6	18128	2750.7	21428.4	1611.3	0.00

On plants pre-treated with aphids (H-Aphid), the EPG parameters were not significantly different from those on the Healthy plants. The probing time of *A. pisum* was 86.6% and phloem activity was recorded (57.5%) (Figure 3). The n_C was 37.6 ± 10.2, n_pd was 112.2 ± 24.5, and the s_F was 2747.0 ± 916.5 sec (Table 1). The s_E1 was 466.9 ± 108.6 sec and the t_1E2 took 6369.7 ± 987.6 sec on average (Table 1). The s_sE2 was 18128 ± 2750.7 sec.

When *A. pisum* feed on *V. faba* infected with FBNYV (I-FBNYV), the EPG parameters were significantly different in many cases than those recorded when feeding on healthy plants H-Aphid and the Healthy. The aphids probed and fed on I-FBNYV for 69.5% of the recording time (9 hours) on average, shorter than the other treatments and spent shorter time on the phloem, only 26.4% of the recording time (Figure 3). Once the stylet reached the phloem, the salivation was not significantly different from the other three treatments. Interestingly, the t_1E2 was significantly longer than the other treatments 11450.3 ± 2250.2 seconds (Table 1). Outside the phloem, the n_C was significantly different with the Healthy but not with H-Aphid (Table 1). The derailed stylet mechanics was also not significantly different between the different treatments. The s_sE2 (8269.8 ± 2546.7 sec) also significantly decreased.

On PNYDV-infected plants (I-PNYDV), the EPG parameters were significantly different in several phloem related activities compared to I-FBNYV and Healthy plants but did not differ from the H-Aphid. *A. pisum*'s probing time was 83.8% significantly higher than I-FBNYV but lower than the H plants, but not significantly different from H-Aphid plants (Figure 3). The phloem activity on the I-PNYDV was 50.7% of the recorded time (Figure 3) also significantly higher than the I-FBNYV, lower than the Healthy, and not different from H-Aphid. The salivation activity was not significantly different between I-PNYDV and the other treatments. The t_1E2 (5090.3 ± 657.8 seconds) was significantly shorter than I-FBNYV but not significantly different form the controls (Table 1). The n_C was significantly lower than I-FBNYV but not different from both controls. The s_sE2 was 16100.2 ± 1594.6 sec, significantly higher than I-FBNYV but not different from Healthy and H-Aphid.

6.4.3.2. Effect of the nuclear shuttle protein (NSP) of FBNSV on the behavioural responses of A. pisum during probing and feeding on faba bean:

In general, the EPG parameters for *A. pisum* feeding on the four different *V. faba* plants were not significantly different (Table 2). *A. pisum* probed and fed on average on Healthy plants for 81.1% of the recording time (Figure 4). The aphids spent 45.4% phloem feeding and 1.9% salivating on H plants (Figure 4). The n_C, n_pd and t_1E2 were 36.1 ± 5.6, 193.8 ± 26.2 and 9286.4 ± 1771.9 sec, respectively (Table 4). The n_E1 and n_E2 of *A. pisum* were 6.0 ± 0.9 and 3.4 ± 0.6, respectively. The n_F and s_F were 0.5 ± 0.3 and 981.4 ± 482.1 sec, respectively. The s_sE2 was 14395.3 ± 2037.3 sec.

Table 2: Statistical analysis of EPG parameters of *A. pisum* probing healthy faba bean plants and faba bean plants agroinoculated with agrobacteria without infectious clones (H-Agro), and agrobacteria containing FBNSV (I-FBNSV) and FBNSV without component DNA-N (I-FBNSV/-N) infectious clones (experiment 2). The parameters are Np: non probing, C: pathway, pd: intracellular puncture, E1: salivation, E2: ingestion, F: derailed stylet mechanics, and sE2: sustained ingestion. n_ = the number of, s_ = the sum of (in seconds), and t_1 = time to first (in seconds).

Parameters	I-FBNSV		I-FBNSV/-N		H-Agro		Healthy		*P* value
	n=10		n=15		n=10		n=16		
	Mean	Standard error	Mean	Standard error	Mean	Standard error	Mean	Standard error	
n_Np	47.8	9.9	50.8	6.8	51.9	6.3	31.6	5.3	0.09
s_Np	7802.6	1721.3	8896.3	1137.6	9452.2	1483.9	6129.1	1335.8	0.10
n_C	49.5	10.2	56.0	6.8	57.0	6.6	36.1	5.6	0.09
s_C	7553.9	1765.3	9584.1	1085.2	8500.3	1234.4	9971.2	1159.1	0.40
n_F	1.7	0.8	2.5	0.6	2.1	0.5	0.5	0.3	0.00
s_F	2135.6	1140.0	4277.0	929.0	5893.4	1165.8	981.4	482.1	0.00
n_E1	1.5	0.3	3.5	0.7	4.0	0.7	6.0	0.9	0.00
s_E1	210.6	81.6	319.7	63.8	316.8	77.8	615.1	130.3	0.03
n_E2	1.2	0.3	2.5	0.6	3.3	0.4	3.4	0.6	0.01
s_E2	14697.0	3858.0	9322.8	2070.9	8236.5	2040.5	14700.3	1999.7	0.19
n_pd	122.5	31.1	201.1	24.2	169.5	19.9	193.8	26.2	0.06
s_pd	697.9	152.0	1360.2	207.4	1251.7	175.8	1584.5	299.6	0.05
t_1Pr	77.6	19.0	74.6	18.0	95.7	24.3	434.4	255.9	0.80
t_1E	14084.8	3986.2	12569.5	2578.7	10249.7	2957.2	8201.8	1763.7	0.51
t_1E2	14267.7	3965.3	14824.4	3047.8	11986.6	3009.8	9286.4	1771.9	0.78
s_sE2	14669.1	3865.8	9074.4	2091.6	7827.2	2075.2	14395.3	2037.3	0.16

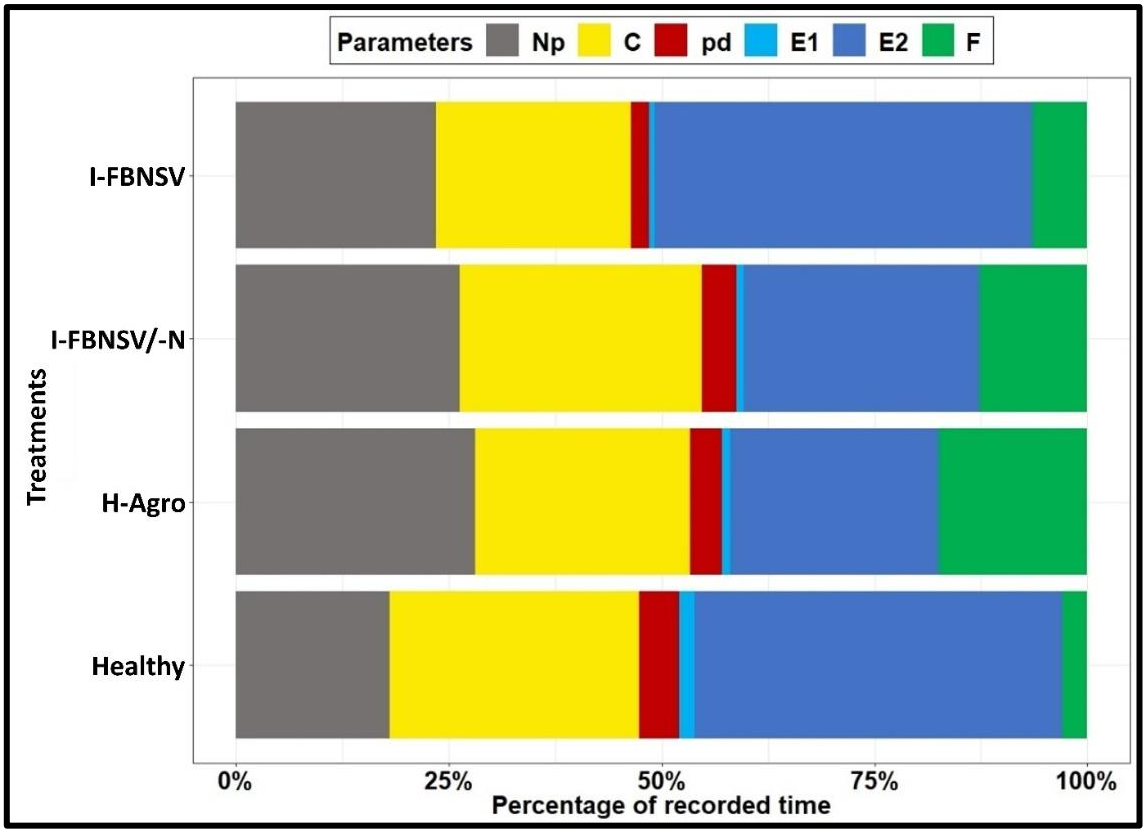

Figure 4: Percentages of the main parameters recorded for *A. pisum* feeding on I-FBNSV and I-FBNSV/-N infected *V. faba*, agrobacteria pre-treated healthy plants (H-Agro), and healthy plants using EPG. The parameters are Np: non probing, C: pathway, pd: intracellular puncture, E1: salivation, E2: ingestion, and F: derailed stylet mechanics.

The aphids spent 70.8% of the recorded time probing on agrobacteria treated plants (H-Agro), which was not significantly different from on Healthy control (Figure 4). *A. pisum* spent 25.4% phloem feeding on H-Agro plants and 0.9% salivation. The n_C was 57.0 ± 6.6, the n_pd was 169.5 ± 19.9 and t_1E2 was 11986.6 ± 3009.8 sec (Table 4). The n_E1 and n_E2 were 4.0 ± 0.7 and 3.3 ± 0.4, respectively. The n_F and s_F of *A. pisum* on H-Agro were significantly higher than from on Healthy plants with 2.1 ± 0.5 and 5893.4 ± 1165.8 sec, respectively (Table 4). The s_sE2 was 7827.2 ± 2075.2 sec, not significantly different from on the Healthy.

A. pisum probed and fed on FBNSV-infected plants (I-FBNSV) for 75.9% of the recording time on average. On average, the aphids spent less than 50% of the recording time (9 hours) on the phloem of I-FBNSV with 45.4% ingestion and 0.6% salivation (Figure 4). The n_C on I-FBNSV was 49.5 ± 10.2 and the n_pd was 122.5 ± 31.1 (Table 2). The n_F and s_F (2.5 ± 0.6 and 4277 ± 929, respectively) were significantly different from the H-Agro, but not from the Healthy. The t_1E2 was 14267.7 ± 3965.3 sec (Table 2). Although the s_E2 (14697 ± 3858 sec) was not significantly different from the other treatments, the n_E2 was 1.2 ± 0.3 significantly lower than on H-Agro and on Healthy plants. The s_sE2 was 14669.1 ± 3865.8 sec.

On I-FBNSV/-N plants, the aphids spent 72.5% of the recorded time probing, not significantly different from the other treatments (Figure 4). Aphids spent only 28.8% of the recording time phloem feeding and 0.9% salivating (Figure 4). The n_C was 56.0 ± 6.8, the n_pd was 201.1 ± 24.2 and the t_1E2 was 14824.4 ± 3047.8 sec (Table 4). None of the phloem activities were significantly different from the other treatments. The n_F and s_F (2.5 ± 0.6 and 4277 ± 929 sec, respectively) were significantly higher than on Healthy plants.

6.5. Discussion

6.5.1. Endosymbionts of the JKI *A. pisum* clone

There is evidence that microbial symbioses influence aphid-plant interactions (Frago *et al.*, 2012). It was therefore necessary to identify the symbionts in the *A. pisum* clone JKI used in this study. In the JKI *A. pisum* clone, the essential intracellular symbiotic bacterium *Buchnera* was detected. The primary symbiont (P-symbiont) *Buchnera* is present in almost all aphids in the cytoplasm of their mycetocytes (or bacteriocytes) in their abdomen (Buchner, 1965; Baumann *et al.*, 1995). *Buchnera* is tightly restricted to the aphid's body cavity. *Buchnera* symbionts and their aphids are intimately mutualistic (Houk & Griffiths, 1980; Ishikawa & Yamaji, 1985; Ohtaka, 1991). It was found that the *Buchnera* cannot survive when removed from their host cells, and that aphids suffer sterility or even death when deprived of *Buchnera* (Houk & Griffiths, 1980; Ishikawa & Yamaji, 1985; Ohtaka, 1991). Thus, it is not surprising to detect such intimately mutualistic symbionts in the JKI aphid clone.

In addition to *Buchnera*, two facultative secondary symbionts (S-symbiont) were detected in the aphids. These are a pea aphid *Bemisia*-type symbiont (PABS) "*Hamiltonella defensa*", and the *Spiroplasma* symbiont (Fukatsu *et al.*, 2001; Moran *et al.*, 2005). PABS was detected before in both the gut and the ovaries of aphids; in contrast with the primary symbiont *Buchnera* which is restricted to cytoplasm of the bacteriocytes (Munson *et al.*, 1991; Wilkinson & Douglas, 1998). The removal of *H. defensa* from *A. pisum* led to decrease in the fecundity rate by about 20% (McLean *et al.*, 2011). Moreover, *H. defensa*-infected *A. craccivora* exhibiting depressed probing behaviour when tested with EPG (Angelella *et al.*, 2018). *Spiroplasma* symbionts can negatively affect the growth, longevity and reproduction of *A. pisum* (Fukatsu *et al.*, 2001). No information about the interaction between *Spiroplasma* sp. and plant viruses is available. Therefore, the endosymbiotic organization of the JKI clone of *A. pisum* is *Buchnera* (P-symbiont) and both *Hamiltonella* and *Spiroplasma* (S-symbiont). The presence of other genera in the clone is still possible, thus for additional confirmation, we suggest using high throughput sequencing (HTS) to detect all the endosymbionts in the clone.

The importance of endosymbiotic activity for virus transmission has been described for few systems. For example, virus particles of potato leafroll virus (PLRV; genus: *Polerovirus*, family: *Luteoviridae*), also a circulative, persistently-transmitted virus,

bind to the chaperonin protein "GroEL" produced by *Buchnera aphidicola.* This binding seems to protect the virus particles from degradation in the aphid's haemocoel (van den Heuvel *et al.*, 1994). Feeding of antibiotics to *Myzus persicae* nymphs prior PLRV acquisition reduced the virus transmission by more than 70% (van den Heuvel *et al.*, 1994). A study on tomato yellow leaf curl virus (TYLCV) (genus: *Begomovirus*, family: *Geminiviridae*) using yeast two-hybrid and protein pulldown assays showed that GroEL protein produced by *Hamiltonella* interacts with the coat protein of TYLCV and facilitates the virus transmission (Gottlieb *et al.*, 2010). The virus particles binding to the GroEL produced by symbionts seems to protect the viruses' particles from degradation in the aphid's haemocoel (van den Heuvel *et al.*, 1994; Gottlieb *et al.*, 2010). In case of banana bunchy top virus (genus: *Babuvirus*, family *Nanoviridae*), there was no interaction of virus particles with *Buchnera* GroEL (Watanabe *et al.*, 2013) and there is currently no evidence that nanoviruses interact with their vectors' endosymbionts. It is suggested that nanovirids translocate in large clusters of virions in aphid (Watanabe & Bressan, 2013; Vetten *et al.*, 2016). This might protect the individual particle from degradation and might be the reason that prevents the virions interaction with the GroEL (Watanabe *et al.*, 2013). The biological effect of the S-symbionts and their interactions with nanoviruses are unknown and it is of interest to investigate these interactions.

6.5.2. Experiment 1: Effect of nanovirus infection on the behavioural responses of *A. pisum* during probing and feeding on faba bean:

In this study, we used EPG to evaluate whether the infection of nanoviruses to faba beans can change the feeding and probing behaviours of *A. pisum*.

6.5.2.1. The behaviour of *A. pisum* changed on FBNYV and PNYDV infected faba beans:

At the time of recording, FBNYV infected plants were dwarfed with top leaf carling and hardening. As the aphid feed on the lower surface of the top leaves with these severe symptoms, these cytological changes of the host plant might explain the change in the feeding and probing behaviour of *A. pisum*. The probing period on I-FBNYV was significantly shorter than on other treatments and aphids spent less time on phloem feeding (26.4% of the recording time). Persistently transmitted plant viruses such as nanoviruses are restricted to the phloem tissue. Thus, early phloem sap ingestion (E2) on virus-infected plants will enhance transmission of phloem-restricted viruses, as aphids will acquire the virus faster and potentially with longer acquisition times, therefore allowing the acquisition of all viral components and helper factors need for successful virus transmission. However, the time to start the first feeding was significantly longer than the other treatments and the total duration of sustained phloem sap ingestion also decreased. The period before the first phloem activity is determined by epidermal, mesophyll, general vascular, and early phloem factors (Schwarzkopf *et al.*, 2013). The

significantly longer t_1E2 on FBNYV-infected plants suggests increased resistance to aphids in these tissues. Additionally, an increased number of probing was detected. From these, the manipulation of FBNYV to the infected plants appeared to be less attractive to *A. pisum*.

In contrast to FBNYV infected plants, the symptoms induced by PNYDV on infected plants were mild (slight yellowing compared to the healthy plants). Yet, the probing time of *A. pisum* was significantly higher than I-FBNYV and lower than the Healthy plants and it was not different from H-Aphid plants. The same results were obtained for the phloem activity. We also observed at a later stage of the disease, that infected plants with PNYDV isolate Drohndorf-15 showed top leaf dwarfing, leaf curling and whole plant yellowing. We assume that at a later recording time with the disease development, the effect of PNYDV on the feeding and probing behaviours of *A. pisum* will increase.

These results contradict previous studies on other circulative viruses i.e., PLRV and tomato yellow leaf curl virus (TYLCV), where virus infections reduced the resistance and increased the feeding and probing behaviours (Alvarez *et al.*, 2007; Liu *et al.*, 2013).

6.5.2.2. Aphid behaviours on aphid pre-treated plants were not different from those on healthy and PNYDV infected plants:

Aphid infestation is known to induce local and systemic changes to plant consequently the plants develop defences to limit the damage caused by aphid (Moran & Thompson, 2001; Cooper & Goggin, 2005; Thompson & Goggin, 2006). Moreover, previous infestation could affect subsequent aphid performance positively or negatively. For example, *Aphis gossypii* pre-infestation of cotton plants decreased recolonization by aphids (Wool & Hales, 1996). On the other hand, *M. persicae* pre-infestation led to an improvement of subsequent same aphid species performance (Sauge *et al.*, 2002). To compare whether the aphid inoculation had effects on any change in behaviour of *A. pisum* on nanoviruses infected plants, additional control of plants pre-treated with aphid was also tested. There were no major differences between healthy untreated plants and aphid pre-treated plants. It was only different from I-FBNYV as other treatments with higher phloem activity and faster feeding starting time.

These behavioural changes can be correlated to the visual symptoms of the virus disease. This is in consistence with previous finding where *M. persicae* only exhibit differences in feeding behaviour on PLRV-infected potato plants after the disease symptoms were observed (Alvarez *et al.*, 2007). To confirm this hypothesis, additional studies are required to determine the structural and metabolic changes of nanoviruses infected plants during the development of the disease symptoms. Whether this is due to a change in the phloem sap quality or the structure of the plant tissues remains unclear as there is no study on the effect of nanovirus infection on the structure and metabolic components of the host plant.

6.5.3. Experiment 2: Effect of NSP on the behavioural responses of *A. pisum* during probing and feeding on faba bean:

In general, the EPG parameters for *A. pisum* feeding on the four different *V. faba* plants were not significantly different (Table 2). Our results showed no differences in the behaviours of *A. pisum*'s on FBNSV infected plants or FBNSV missing component DNA-N infected plants. Although the NSP produced by DNA-N is required for nanoviruses transmission, our results suggest that it does not affect the aphid feeding behaviour. A previous study showed that DNA-N is not essential for symptoms development, and the combination of the other seven DNAs was still able to cause the disease symptoms (Timchenko *et al.*, 2006). Surprisingly, the aphid behaviours on I-FBNSV were not different from those on H plants. The only difference was the number of salivations and ingestions where they were significantly lower on I-FBNSV.

The derailed stylet mechanics (F) on Healthy and I-FBNSV plants were significantly different from I-FBNSV/-N and H-Agro. The derailed stylet occurs when a stylet protruded much further than the other three stylets thus loses the bundle formation which occurs at low frequency (Tjallingii, 1988). It is suggested that the intercellular structural composition is responsible for such derailments (Tjallingii, 1988). Here, an increase in derailing of the stylet with the presence of agrobacteria, although in case of I-FBNSV the virus infection seems to reduce such effect. This could be due to the presence of the agrobacteria in the plants. As there is no study on the effect of agrobacteria on the aphid behaviours, suggesting adding an extra control with infiltration buffer and no agrobacteria, or using plants from successive aphid transmissions to reduce the effect of the agroinoculation.

It was important to identify and characterise the genomes of the viruses in the study, as the presence of different components and any of the satellite DNAs can affect the virus disease severity and the virus transmission. Previous study on FBNYV suggested that the presence of faba bean necrotic yellows C11 alphasatellite cause reduction in infected plants due to competition between the different components on the resources, encapsidation and movement (Timchenko *et al.*, 2006). The isolate of FBNYV in this study contains in addition to the eight components, an alphasatellite i.e., faba bean necrotic yellows C1 alphasatellite, while both PNYDV and FBNSV each contains only their eight components. Whether there is an effect "positive or negative" for the alphasatellite on the severity of the nanoviruses diseases or on their aphid vectors, and consequently the aphid behaviour, is currently unknown.

The effect of the plant variety and whether the aphid behaviour will be different from a variety to another. Comparing the Healthy plants in both experiments showed the probing time was significantly different between both experiments. The reason for that is unknown as we used the same growing conditions for both experiments. Other aphid biotypes should be considered in future testing. Additional studies will be required to

determine the structural and metabolic changes of nanoviruses infected plants during the development of the disease symptoms to confirm the relation between the change in the behaviour and symptoms. Suggesting studying the dynamic of aphid behaviour on plants with series of infection periods. Therefore, this study needs follow up with further investigations to confirm the findings.

Chapter 7: General discussion

7.1. Part 1: Plant virus identification

A prerequisite to the control of a plant viral disease is the proper detection and identification of its causal agent. Thus, a sensitive and reliable detection is crucial for plant protection. To identify the aetiology of a viral disease, there are different diagnostics tools used e.g., electron microscopy (EM), serology-based methods such as enzyme linked immunosorbent assays (ELISA), molecular biology-based methods such as polymerase chain reaction (PCR) and high throughput sequencing (HTS).

7.1.1. Disease aetiology of several plants using conventional and HTS methods

In a legume survey, several viruses were detected in many plant samples from Germany and Austria in 2016 using conventional methods (PCR and ELISA) ((Gaafar *et al.*, 2016), Chapter 2). We could identify pea enation mosaic virus (PEMV) as the predominant virus followed by pea necrotic yellow dwarf virus (PNYDV). Viruses from other genera e.g., potyviruses, poleroviruses and luteoviruses could be also detected by ELISA; however, the exact virus species could not be confirmed. The antibodies used for ELISA were preselected for the commonly known legume viruses. Thus, we could not confirm whether other viruses were present in the samples or not, especially when several symptomatic samples tested negative with ELISA.

Additionally, in two samples from the Netherlands, we could detect by ELISA and PCR the presence of PNYDV for the first time ((Gaafar *et al.*, 2017), Chapter 2). We could confirm the presence of the eight genomic components of PNYDV by PCR amplification of each component, cloning in *Escherichia coli*, extracting the plasmids followed by Sanger sequencing. This approach was laborious and time consuming. Moreover, we could not confirm this way whether there were PNYDV-associated alphasatellites present or not.

High throughput sequencing (HTS) gave us the ability to sequence all the nucleic acids present in a given sample (Adams & Fox, 2016). Using HTS, we were able to detect the full genomes of different PNYDV isolates from five samples from Denmark, and also we detected their associated alphasatellites ((Gaafar *et al.*, 2018a), Chapter 2).

Other legume and vegetable samples were also analysed by conventional methods and HTS. Using EM and ELISA, we were able to identify viruses from different families and genera. However, without the presence of species-specific antibodies, we could not confirm the virus species and in some cases the virus genus. For example, using immunosorbent electron microscopy (ISEM) and ELISA, we detected the presence of turnip crinkle virus (TCV) in wild garlic mustard ((Gaafar & Ziebell, 2019a), Chapter 2). Using HTS, a new strain of TCV was identified as being a highly divergent sequence (Rochon *et al.*, 2012). Such divergence could not be detected by ISEM or ELISA. In a similar

case we investigated a new nepovirus in a caraway sample where nepoviruses-specific antibodies developed against a divergent strain of cherry leafroll virus (CLRV) infecting carrot (unpublished) was used in ELISA and ISEM, and confirmed the presence of a nepovirus but could not identify the virus species ((Gaafar *et al.*, 2019f), Chapter 2). Using HTS, we identified the presence of a novel nepovirus, tentatively called caraway yellows virus (CawYV).

Also, we could detect the presence of rhabdovirus-like particles in tomato samples from Germany, alfalfa and black medic samples from Austria. Physostigia chlorotic mottle virus (PhCMoV; a nucleorhabdovirus) in the tomato sample and a novel nucleorhabdovirus, tentatively called alfalfa associated nucleorhabdovirus (AaNV) in alfalfa sample were identified using HTS, confirming the findings of EM ((Gaafar *et al.*, 2018b; Gaafar *et al.*, 2019d), Chapter 2). However, in the black medic sample, the presence of a rhabdovirus could not be confirmed; surprisingly a tenuivirus was detected (melon chlorotic spot virus; MeCSV) ((Gaafar *et al.*, 2019e), Chapter 2). The observed ribonucleoproteins (RNP) by EM in the original black medic sample were confused with the disassembled particles of rhabdoviruses. Nevertheless, these RNPs were the virions of the tenuivirus which are later confirmed in partially- purified virion preparations from infected plants.

The limitation of the conventional methods was further demonstrated with a beetroot sample with leaf necrosis, reduced size and root bearding: Using ISEM with various antibodies for the detection of various beetroot and sugar beet viruses, no virus could be detected ((Gaafar *et al.*, 2019b), Chapter 2). Using HTS however, the full-length sequences of beet soil-borne virus and beet cryptic virus 2 could be recovered although BSBV specific-antibodies were also used in the ISEM. A similar example is represented by a celery sample where no virus particles could be observed with EM, and ELISA tests with antibodies against carrot red leaf virus and CLRV were negative ((Gaafar & Ziebell, 2019a), Chapter 2). Nevertheless, we could detect a divergent strain of carrot torradovirus 1 (CaTV1) using HTS.

Furthermore, using MinION sequencer allowed us to rapidly identify the virus infecting several tomato samples i.e., Southern tomato virus ((Gaafar et al., 2019a), Chapter 2). Using dsRNA enrichment approach followed by MinION sequencing saved time allowing us to identify the virus in less than 24 hours. The speed at which this method was able to produce reads suggesting that it could be used in the future to as a laboratory diagnostic tool with a room for improvement in the quality of the produced raw reads and the costs of the platform and the sequencing kits.

7.1.2. Pea survey in Germany using HTS

In our survey on pea fields in Germany, HTS was used to detect viruses in peas and surrounding plants. In this survey, the plants were sampled from six different regions for three successive growing seasons. The samples were pooled together in ten pools i.e.,

region 1 to 6, and all symptomatic peas (SP), asymptomatic peas (aSP), surrounding legumes (sL) and surrounding non-legumes (snL) for each season. Thirty-five viruses could be detected in addition to associated satellite nucleic acids and each finding was confirmed by RT-PCR. In addition to eleven known pea and legume viruses, nine new viruses and thirteen viruses new to Germany were reported (Chapter 3). Interestingly, new emaravirus provisionally called pea associated emaravirus, with close relationship to rose rosette virus, a quarantine pathogen, was detected in symptomatic peas in two seasons. The pooling strategy minimised the number of samples for HTS, nevertheless could not assist in identifying the exact host sample for each virus. Due to low virus titre or biased amplification of other nucleic acids, not all complete genomes of the viruses could be recovered demonstrating some limitations of the HTS-based study. Nevertheless, the discovery of nine novel viruses on one crop and its surrounding weeds imposes new challenges for diagnosticians, risk assessors and policy makers. Specific detection tests would need to be developed and implemented in future pea surveys to assess the presence of these viruses. As required by law, for each newly described virus a pest risk analysis would have to be carried out and policy makers would need to decide how to deal with these new viruses with respect to quarantine regulations and import/export restrictions.

7.1.3. Virus characterisation

Each newly identified virus was characterised molecularly (Chapter 2, 3 and 4). This was necessary for taxonomic assignment of these viruses and to characterise their potential open reading frames (ORF), motifs and domains in the viral sequences. For example, the isolates of pea enation mosaic virus 2 (PEMV2) from France and Germany have slightly shorter ORF3 (encoding a putative long-distance movement protein) than the PEMV2 reference sequence from the USA. This is due to a nucleotide substitution “A to T” at nt position 2783 in the French isolate and at nt position 2755 in German isolates. The first ORF (encoding a putative RNA-dependent RNA polymerase) of TuYVaRNA has an amber codon similar to other poleroviruses-associated RNAs which was not detected before in the available TuYVaRNA sequences on NCBI. Another example, is the potential new protease cleavage site “HS” that was identified in both predicted polyproteins of CawYV ((Gaafar *et al.*, 2019f), Chapter 2).

These variations can contribute to the fitness and virulence of each virus. Such characteristics can be beneficial, neutral, deleterious or even lethal to the virus (Sanjuán, 2010). However, very little is known about the fitness and virulence costs of these characteristics. Therefore, it is important to confirm these molecular characteristics experimentally and to study their effects on virus infectivity, symptomology and fitness.

Moreover, EM helped us to study the virus particles morphology and with thin sections we could study the virion locations in infected cells. We could also confirm the relatedness of the novel viruses to their respective genera or family e.g., the presence of

the replication factories in the nucleus of cells infected with AaNV (a feature of nucleorhabdoviruses) and the presence of tubular structure containing virus particles in CawYV infected plants (a feature of nepoviruses).

7.1.4. Species demarcations set by ICTV

Many new viruses were discovered in recent years using HTS technology. An advantage of these discoveries is the improvement of virus taxonomy. For example, in the *Secoviridae* family, all the members of the different genera are following the same "suggested" demarcations i.e., the conserved protease-polymerase (Pro-Pol) and the capsid protein (CP) regions share aa sequence identity < 80% and < 75% respectively, differences in antigenic reactions, distinct host range and vector specificity, absence of cross-protection and absence of re-assortment between RNA-1 and RNA-2 (for bipartite genome). Sanfaçon and colleagues suggested that these criteria need to be improved with more viral sequences of this family (Sanfaçon *et al.*, 2009).

These criteria are challenging as not all of them are met at the same time. The novel CawYV ((Gaafar *et al.*, 2019f); Chapter 2) had a Pro-Pol region which shared 80.1% amino acid (aa) identity to the closest virus, i.e., the value is 0.1% above the threshold set for *Secoviridae* species criteria therefore suggesting the presence of a novel strain of an existing species. In contrast, the CP region shared only 39.6% aa identity to its closest virus therefore this isolate should have been considered as new species. Other researchers face similar challenges, e.g., the discovery of red clover nepovirus A (RCNVA) (Koloniuk *et al.*, 2018). The Pro-Pol region of RCNVA was 86% aa identical to artichoke Italian latent virus (AILV) and tomato black ring virus whereas the CP region was 64.4% aa to AILV. Moreover, the conserved regions i.e., the Pro-Pol and the CP regions of CLRV from carrot and CaTV1 (a torradovirus) from celery, both were detected again in the pea virome survey, also did not meet the *Secoviridae* species demarcations. However, they are divergent from their closest sequences in the other regions of the genomes e.g., genome length, whole ORFs and untranslated regions sequence identities (Chapter 2). This was also the same for other viruses e.g., tomato chocolate virus (torradovirus) (Verbeek *et al.*, 2010).

We suggest that the ICTV *Secoviridae* study group revises the species demarcation criteria of *Torradovirus* and *Nepovirus* genera. We suggested not only reconsider the conserved regions but adding other criteria e.g., the identity of the nucleotide sequences of the untranslated regions and the amino acid identity of the whole predicted products of the ORFs. Verbeek and colleagues also suggested newly proposed additional demarcation criteria for the genus *Torradovirus* (Verbeek *et al.*, 2010).

7.1.5. Which viral enrichment approach is the best?

Building on the success of HTS for plant virus and viroid detection, it is time to apply HTS for routine detection in laboratories. As with other new methods, HTS as a virus/viroid detection tool has some challenges. Firstly, due to their different genomic

characteristics, there is no universal extraction method to fit all viruses/viroids. Secondly, there are no universal sequence for all viruses/viroids that can be amplified and used for sequencing such as the internal transcribed spacer (ITS) or 16S ribosomal RNA can be used for general detection of pathogens i.e., fungi and bacteria (Leff *et al.*, 2017).Thirdly, viral enrichment is required during the sample preparation for HTS as the viral reads can be relatively very low compared to the host sequences (Adams & Fox, 2016).

There are different enrichment methods are being used for virus/viroid detection by HTS. Each can differ in its efficiency and can have specific strengths and weaknesses (Adams & Fox, 2016). In the performed studies (Chapter 2 and 3), different viral enrichment approaches have been used i.e., dsRNA extraction, ribosomal RNA depleted total RNA (ribo-depleted totRNA) extraction, small RNA extraction and rolling circle amplification (RCA). DsRNA extraction, ribo-depleted totRNA and small RNA (sRNA) viral enrichment approaches seem to be the most generic approaches recovering a range of plant viruses/viroids with different genome types and could be integrated in workflows in diagnostic laboratories.

A comparison between the three approaches was performed (Chapter 4). All known viruses were detected using the three viral enrichment methods. Surprisingly, two unknown viruses were detected by the dsRNA approach i.e., Vicia cryptic virus (VCV) and Wuhan aphid virus 2 (WHAV2), whereas only one of the two viruses was detected by the ribo-depleted totRNA and sRNA enrichment approaches, WHAV2 and VCV, respectively. The number of recovered viral reads were different from one method to another and were also virus dependent. The dsRNA approach used in this study detected all viruses (knowns and unknowns). It was cheaper, faster and easier compared to the other two approaches. Moreover, in comparison with other available dsRNA extraction methods, the dsRNA extraction kit used in this study is also faster, does not require a large amount of plant starting material and does not require an extra amplification step as in (Kesanakurti *et al.*, 2016; Yanagisawa *et al.*, 2016; Blouin *et al.*, 2016).

The main limitation of this approach is the low concentration of extracted dsRNA which can be challenging for some library preparations. It is unclear whether PNYDV sequences (a DNA virus) were recovered as RNA/DNA hybrid molecule and amplified by HTS or if contaminating DNA was carried over in the dsRNA preparations (Knierim *et al.*, 2019).

7.1.6. Data analysis

There are different software and tools used for bioinformatic analyses e.g., HTS data analysis and molecular characterisation of the findings (Bao *et al.*, 2014). The algorithms used for analysis are different from one bioinformatic tool to another (Miller *et al.*, 2010; Li & Homer, 2010; Narzisi & Mishra, 2011). Changing the parameters used in the analyses can be beneficial where unknown findings can be detected or can be incorrect thus some data might be concealed (Del Fabbro *et al.*, 2013; Massart *et al.*,

2019). For virus/viroid detection, there are different thresholds set by each laboratory and there are no generic criteria set for all. Some laboratories a very strict with the number of reads and coverage to consider a finding and thus confirm it. Such restrictions can cause the ignorance of some findings e.g., viruses with low titre. We confirmed all the findings even with low number of reads and coverage. This led to the identification and confirmation of many viruses (Chapter 2, 3 and 4).

The computer facility was a limitation with large data and with some tools i.e., the analysis of sRNA data required a strong hardware computer with larger memory and stronger processor. Thus, we tend to normalise the reads and remove duplicated reads. This facilitated the subsequent analyses e.g., *de novo* assembly on such low hardware specifications computer.

Additionally, in some of our analysed samples no viruses or viroids were detected. As our focus is plant viruses and viroids, our Basic Local Alignment Search Tool (BLAST) databases do not contain sequences of other pathogens e.g., phytoplasma or fungi.

7.2. Part 2: Virus-vector-host interactions

7.2.1. Nanoviruses infection change the probing and feeding behaviours of aphids

To investigate the effect of nanoviruses infection on the probing and feeding behaviours of pea aphids (*Acyrthosiphon pisum*) on faba beans (*Vicia faba*), electrical penetration graph (EPG) was used (Chapter 6). Our results showed that infections of *V. faba* plants with nanoviruses i.e., faba bean necrotic yellows virus (FBNYV) and PNYDV changed the feeding and probing behaviours of *A. pisum* where the aphids spent less probing time on infected plants compared to those on healthy plants. Furthermore, aphids spent less time phloem feeding compared to those on healthy plants. The FBNYV-infected plants displayed severe disease symptoms compared to PNYDV-infected plants and the changes in the behaviours of the aphids were higher on FBNYV-infected than those on PNYDV-infected. It is possible that these behavioural changes are correlated to the disease development (changes in the chemical composition and physical structure of infected plant tissue).

7.2.2. Absence of the nuclear shuttle protein does not affect the aphid probing and feeding behaviours

The nuclear shuttle protein (NSP) is encoded by DNA-N of nanoviruses and is required for the aphid transmission (Grigoras *et al.*, 2018). To investigate the effect of NSP on the probing and feeding behaviours of the aphids, plants were agroinoculated with infectious clones of all eight components of faba bean necrotic stunt virus (FBNSV) i.e., DNA-C, -M, -N, -R, -S, -U1, -U2 and -U4. In addition, other plants were agroinoculated with seven components omitting DNA-N (FBNSV/-N). Two additional controls were used i.e., plants inoculated with agrobacteria without the infectious clones and healthy plants without agroinoculation. No significant changes in feeding behaviour could be observed

when *A. pisum* fed on FBNSV/-N-infected plants compared to those fed on FBNSV-infected plants (Chapter 6). This can be due to that nanoviruses symptoms can be developed without DNA-N (Timchenko *et al.*, 2006). The only significant difference from the healthy control was derailed stylet mechanics (F) which was higher on FBNSV/-N infected plants. We also found significant effects for the agrobacteria on the behaviours where F was higher than on healthy plants "without agroinoculation". The intercellular structural composition is suggested to be responsible for these derailments (Tjallingii, 1988). The increase in derailing of the stylet with the presence of agrobacteria was reduced in case of FBNSV-infected plants. This suggests that the agrobacteria may have altered the effect of the virus.

7.3. General conclusion and future perspectives

I have demonstrated the strength of HTS as a rapid and accurate diagnostics tool for plant virus discovery, characterisation and for metagenomics studies. During the optimisation process I found out that viral enrichment based on dsRNA approach were better than the ribo-depleted totRNA and sRNA approaches for HTS virus/viroid detection. It is important to follow-up with a comparison of different dsRNA enrichment approaches to reach the most reliable viral enrichment approach. This will help fastening the application of HTS as a diagnostic tool in laboratories. Further research is needed to identify the natural vectors and mode of transmission of the new viruses as well as their host range. Moreover, it is important to test experimentally the molecular characteristics of the new viruses and their effect on virulence and fitness.

Aphid-borne nanoviruses are threatening legume production in different countries. Infection of nanoviruses of *V. faba* caused reduction in *A. pisum*'s probing and feeding behaviours which may be related to changes in the chemical composition and physical structure of plant tissues during disease development. Consequently, the aphids may not develop colonies on the infected plants and will spread wider within the field. It is recommended testing the dynamic of changes in these behaviours with the disease development and to relate them to the virus titres and to compare the behaviours on symptomatic and asymptomatic leaves. Furthermore, the absence of the NSP did not affect the feeding behaviour of *A. pisum*, however there was an effect of agroinoculation observed. Thus, we recommend experimenting with additional controls i.e., FBNSV-infected plants after successive aphid transmission to eliminate or reduce the effect of agrobacteria, plants inoculated with inoculation buffer without bacteria and plants with needle injection only.

Summary

The overall aim of this dissertation was the implementation of high-throughput sequencing (HTS) technologies for plant viruses' identification and characterisation. Different viral enrichment approaches were tested to optimise a tool that would detect a broad range of different viruses and that can be used as a generic, reliable and cost-effective strategy for virus diagnostics. Moreover, the interactions of aphid vectors of nanoviruses with infected plants were investigated and the effect of NSP on aphid feeding behaviour were closely observed.

Chapter 1 gives a general introduction to the concept of this research, and it also outlines the aims of this work. Currently used plant virus diagnostics methods are described with a focus on HTS for virus detection using different enrichment and platforms. It also covers the virus-vector-host interactions with focus on the model system nanoviruses-aphid-host used in this research.

In addition, this thesis is divided into two parts (part one and part two). **Part one** focuses on the plant virus identification and describes the discovery and characterisation of new viruses, divergent strains and isolates from new virus hosts or geographical locations. The virome of peas was investigated over a period of three growing seasons in six different German regions. Furthermore, three viral enrichment approaches for plant viruses and viroids detection by HTS were compared. Part one contains three chapters (2, 3 and 4). In **Chapter 2**, the discovery and characterisation of two new viruses i.e., alfalfa associated nucleorhabdovirus from *Medicago sativa* (Austria) and caraway yellows virus from *Carum carvi* (Germany)are described.

Moreover, the reports on the identification and characterisation of the causal agent of several diseases in legumes and vegetables. HTS in combination with conventional diagnostic methods (ELISA, EM, PCR-based methods and Sanger sequencing) were applied to various plant virus samples in the JKI selection. The following viruses were described:

- Divergent strains of:
  1. Melon chlorotic spot virus from *M. lupulina* from Austria
  2. Carrot torradovirus 1 from *Apium graveolens* from Germany
  3. Turnip crinkle virus from *Alliaria petiolata* from Germany
- Divergent isolates of turnip yellows virus (TuYV) from pea and rapeseed oil, in addition to a first report of turnip yellows associated RNA from Germany
- First reports of:
  1. Beet soil borne virus and beet cryptic virus 2 from *Beta vulgaris* subsp. *vulgaris*
  2. Physostegia chlorotic mottle virus from *Solanum lycopersicum* from Germany
  3. Southern tomato virus from *S. lycopersicum* from Germany (new MinION approach was described)

4. Pea necrotic yellow dwarf virus (PNYDV) was detected for the first time infecting *Vicia faba*, *V. sativa* and *Lens culinaris* in Austria and Germany, and for the first time in Denmark and the Netherlands.

The focus in **chapter 3** is set on the identification and characterisation of viruses in German pea fields. For three successive seasons, pea fields in six different regions were surveyed. Samples were taken from symptomatic pea and asymptomatic pea plants, in addition, samples from surrounding non-crop legume and non-legumes weeds were taken. The samples were analysed using HTS and the presence of all detected viruses was confirmed by RT-PCR. Thirty-five viruses in total were detected during this survey, including nine new viruses, thirteen viruses new for Germany. In addition to plant viruses, virus associated nucleic acids were also detected. Pea enation mosaic virus 1 and 2, TuYV, and PNYDV were the most common viruses in the German pea fields. Interestingly a new emaravirus was detected for two successive seasons in the same region, tentatively called pea associated emaravirus. Additionally, several new virus sequences and divergent sequences were detected in the surrounding legumes and non-legumes. By analysing the data spatially and temporally, no differences were observed.

Chapter 4 describes the comparison of three different viral enrichment approaches for virus discovery by HTS. Double stranded (dsRNA), ribosomal RNA depleted total RNA (ribo-depleted totRNA) and small RNA (sRNA) enrichment approaches were compared for the detection of viruses and a viroid representing different genomes i.e., ssRNA [(+ve) and (-ve) senses], DNA and a viroid. The dsRNA approach in this study was better compared to the other two approaches as it is faster, cheaper and all the known and unknown viruses/viroid in the study were detected. However, the number of reads from dsRNA approach were lower when compared to the other two methods. Interestingly, two additional unknown viruses were discovered with the dsRNA enrichment approach i.e., Vicia cryptic virus and Wuhan aphid virus 2 (WHAV2) but only one of them was discovered with sRNA or ribo-depleted totRNA, respectively. To our knowledge this is the first detection of WHAV2, a recently discovered virus from aphids in China, from plant tissues.

In the second part of this thesis (**part two**), the interactions of nanoviruses with their insect vectors and their hosts were studied. **Chapter 5** reviews the current knowledge about nanovirus-vector interactions. In **chapter 6**, the probing and feeding behaviours of *Acyrthosiphon pisum* on nanoviruses-infected faba beans were investigated using electrical penetration graph. The probing and feeding behaviours of *A. pisum* changed on faba bean necrotic yellows virus and PNYDV-infected plants. Additionally, the effect of the absence of the nuclear shuttle protein of FBNSV on the feeding and probing behaviours of *A. pisum* was tested. Additionally, few significant differences were detected, however no clear differences could be concluded.

A general discussion of the findings of this thesis is laid out in the final chapter (**7**), and a prospect for future research is given.

References

Abbas Q, Amin I, Mansoor S, 2019. The Rep proteins encoded by alphasatellites restore expression of a transcriptionally silenced green fluorescent protein transgene in *Nicotiana benthamiana*. *Virus Disease* **30**, 101–5. doi: 10.1007/s13337-017-0413-5.

Abraham AD, Menzel W, Lesemann D-E, 2006. Chickpea chlorotic stunt virus: a new polerovirus infecting cool-season food legumes in Ethiopia. *Phytopathology* **96**, 437–46. doi: 10.1094/PHYTO-96-0437.

Abraham AD, Menzel W, Vetten HJ, 2007. First report of *Soybean dwarf virus* (genus *Luteovirus*) infecting faba bean and clover in Germany. *Plant Disease* **91**, 1059. doi: 10.1094/PDIS-91-8-1059B.

Abraham AD, Varrelmann M, Josef Vetten H, 2012. Three distinct nanoviruses, one of which represents a new species, infect faba bean in Ethiopia. *Plant Disease* **96**, 1045–53. doi: 10.1094/PDIS-09-11-0734-RE.

Adams I, Fox A, 2016. Diagnosis of plant viruses using next-generation sequencing and metagenomic analysis. In: Wang Z, ed. *Current Research Topics in Plant*. Vol 10, 323–35.

Adams IP, Glover RH, Monger WA *et al.*, 2009. Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. *Molecular Plant Pathology* **10**, 537–45. doi: 10.1111/j.1364-3703.2009.00545.x.

Adams IP, Skelton A, Macarthur R *et al.*, 2014. *Carrot yellow leaf virus* is associated with carrot internal necrosis. *PLOS ONE* **9**, e109125. doi: 10.1371/journal.pone.0109125.

Adams MJ, Adkins S, Bragard C *et al.*, 2017. ICTV virus taxonomy profile: *Virgaviridae*. *The Journal of General Virology* **98**, 1999–2000. doi: 10.1099/jgv.0.000884.

Adasme-Carreño F, Muñoz-Gutiérrez C, Salinas-Cornejo J, 2015. A2EPG. A new software for the analysis of electrical penetration graphs to study plant probing behaviour of hemipteran insects. *Computers and Electronics in Agriculture* **113**, 128–35. doi: 10.1016/j.compag.2015.02.005.

Al Rwahnih M, Daubert S, Golino D, 2009. Deep sequencing analysis of RNAs from a grapevine showing Syrah decline symptoms reveals a multiple virus infection that includes a novel virus. *Virology* **387**, 395–401. doi: 10.1016/j.virol.2009.02.028.

Alliot B, Signoret PA, 1972. La «maladie àénations de la Luzerne», une maladie nouvelle pour la France. *Journal of Phytopathology* **74**, 69–73. doi: 10.1111/j.1439-0434.1972.tb04647.x.

Alvarez AE, Garzo E, Verbeek M, 2007. Infection of potato plants with potato leafroll virus changes attraction and feeding behaviour of *Myzus persicae*. *Entomologia Experimentalis et Applicata* **125**, 135–44. doi: 10.1111/j.1570-7458.2007.00607.x.

Amarasinghe GK, Aréchiga Ceballos NG, Banyard AC *et al.*, 2018. Taxonomy of the order *Mononegavirales*: update 2018. *Archives of Virology* **163**, 2283–94. doi: 10.1007/s00705-018-3814-x.

Anderson PK, Cunningham AA, Patel NG, 2004. Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *Trends in Ecology & Evolution* **19**, 535–44. doi: 10.1016/j.tree.2004.07.021.

Angelella G, Nalam V, Nachappa P, 2018. Endosymbionts differentially alter exploratory probing behavior of a nonpersistent plant virus vector. *Microbial Ecology* **76**, 453–8. doi: 10.1007/s00248-017-1133-5.

Antoniw JF, White RF, Xie W, 1990. Cryptic viruses of beet and other plants. In: Fraser RSS, ed. *Recognition and response in plant-virus interactions*, Springer: Berlin, Heidelberg, 273–86.

Argos P, Kamer G, Nicklin MJ, 1984. Similarity in gene organization and homology between proteins of animal picornaviruses and a plant comovirus suggest common ancestry of these virus families. *Nucleic Acids Research* **12**, 7251–67.

Augusto Lopez R, Percy Miranda P, Edgar Tejada V, 1992. Outbreak of human rabies in the Peruvian jungle. *The Lancet* **339**, 408–11. doi: 10.1016/0140-6736(92)90088-K.

Axén C, Hakhverdyan M, Boutrup TS *et al.*, 2017. Emergence of a new rhabdovirus associated with mass mortalities in eelpout (*Zoarces viviparous*) in the Baltic Sea. *Journal of Fish Diseases* **40**, 219–29. doi: 10.1111/jfd.12506.

Bailer J, Aichinger T, Hackl G, 2001. Essential oil content and composition in commercially available dill cultivars in comparison to caraway. *Industrial Crops and Products* **14**, 229–39. doi: 10.1016/S0926-6690(01)00088-7.

Bankevich A, Nurk S, Antipov D *et al.*, 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**, 455–77. doi: 10.1089/cmb.2012.0021.

Bao R, Huang L, Andrade J *et al.*, 2014. Review of current methods, applications, and data management for the bioinformatics analysis of whole exome sequencing. *Cancer Informatics* **13**, 67–82. doi: 10.4137/CIN.S13779.

Bardin SD, Huang HC, 2001. Research on biology and control of *Sclerotinia* diseases in Canada 1. *Canadian Journal of Plant Pathology* **23**, 88–98. doi: 10.1080/07060660109506914.

Bar-Joseph M, Rosner A, Moscovitz M, 1983. A simple procedure for the extraction of double-stranded RNA from virus-infected plants. *Journal of Virological Methods* **6**, 1–8. doi: 10.1016/0166-0934(83)90062-9.

Barnett OW, 1987. Relationships among Australian and North American isolates of the bean yellow mosaic potyvirus subgroup. *Phytopathology* **77**, 791. doi: 10.1094/Phyto-77-791.

Baumann P, Baumann L, Lai CY, 1995. Genetics, physiology, and evolutionary relationships of the genus *Buchnera*: intracellular symbionts of aphids. *Annual Review of Microbiology* **49**, 55–94. doi: 10.1146/annurev.mi.49.100195.000415.

Bawden FC, Nixon HL, 1951. The application of electron microscopy to the study of plant viruses in unpurified plant extracts. *Journal of General Microbiology* **5**, 104–9. doi: 10.1099/00221287-5-1-104.

Behr H-C, 2015. *AMI Markt Bilanz Gemüse 2015*. Bonn, Germany: Agrarmarkt Informations-Gesellschaft mbH.

Behr H-C, 2019. *AMI Markt Bilanz Gemüse 2019*. Bonn, Germany: Agrarmarkt Informations-Gesellschaft mbH.

Bejerman N, Giolitti F, Breuil S de *et al.*, 2015. Complete genome sequence and integrated protein localization and interaction map for alfalfa dwarf virus, which combines properties of both cytoplasmic and nuclear plant rhabdoviruses. *Virology* **483**, 275–83. doi: 10.1016/j.virol.2015.05.001.

Bejerman N, Nome C, Giolitti F *et al.*, 2011. First report of a rhabdovirus infecting alfalfa in Argentina. *Plant Disease* **95**, 771. doi: 10.1094/PDIS-10-10-0764.

Bem F, Murant AF, 1979. Host range, purification and serological properties of heracleum latent virus. *Annals of Applied Biology* **92**, 243–56. doi: 10.1111/j.1744-7348.1979.tb03870.x.

Bergeson GB, Athow KL, Laviolette FA, 1964. Transmission, movement, and vector relationships of tobacco ringspot virus in soybean. *Phytopathology* **54**, 723–8.

Bernad L, Duran-Vila N, 2006. A novel RT-PCR approach for detection and characterization of citrus viroids. *Molecular and Cellular Probes* **20**, 105–13. doi: 10.1016/j.mcp.2005.11.001.

Bioreba, 2014. ELISA data analysis. [http://www.bioreba.ch/saas/CustomUpload/374O357O340O370O356O369O350O321O360O366O369O356O353O352O350O320O326O/ELISA_Data_Analysis.pdf]. Accessed 25 October 2018.

Blanc S, Drucker M, Uzest M, 2014. Localizing viruses in their insect vectors. *Annual Review of Phytopathology* **52**, 403–25. doi: 10.1146/annurev-phyto-102313-045920.

Blawid R, Stephan D, Maiss E, 2007. Molecular characterization and detection of Vicia cryptic virus in different *Vicia faba* cultivars. *Archives of Virology* **152**, 1477–88. doi: 10.1007/s00705-007-0966-5.

Blouin AG, Pearson MN, Chavan RR *et al.*, 2013. Viruses of kiwifruit (*Actinidia* species). *Journal of Plant Pathology*, 221–35.

Blouin AG, Ross HA, Hobson-Peters J, 2016. A new virus discovered by immunocapture of double-stranded RNA, a rapid method for virus enrichment in metagenomic studies. *Molecular Ecology Resources* **16**, 1255–63. doi: 10.1111/1755-0998.12525.

BMEL, 2019. Eiweißpflanzenstrategie. [https://www.bmel.de/DE/Landwirtschaft/Pflanzenbau/Ackerbau/_Texte/Eiweisspflanzenstrategie.html]. Accessed 4 July 2019.

Bolton MD, Thomma BPHJ, Nelson BD, 2006. *Sclerotinia sclerotiorum* (Lib.) de Bary: biology and molecular traits of a cosmopolitan pathogen. *Molecular Plant Pathology* **7**, 1–16. doi: 10.1111/j.1364-3703.2005.00316.x.

Boonham N, Kreuze J, Winter S *et al.*, 2014. Methods in virus diagnostics: from ELISA to next generation sequencing. *Virus Research* **186**, 20–31. doi: 10.1016/j.virusres.2013.12.007.

Boquel S, Giordanengo P, Ameline A, 2011. Divergent effects of PVY-infected potato plant on aphids. *European Journal of Plant Pathology* **129**, 507–10. doi: 10.1007/s10658-010-9732-8.

Bos L, 1970. The identification of three new viruses isolated from *Wisteria* and *Pisum* in The Netherlands, and the problem of variation within the potato virus Y group. *Netherlands Journal of Plant Pathology* **76**, 8–46. doi: 10.1007/BF01976763.

Bos L, 1982. Crop losses caused by viruses. *Crop Protection* **1**, 263–82. doi: 10.1016/0261-2194(82)90002-3.

Bos L, Delević B, Want JPH, 1959. Investigations on white clover mosaic virus. *Netherlands Journal of Plant Pathology* **65**, 89–106. doi: 10.1007/BF02112372.

Bos L, Hampton RO, Makkouk KM, 1988. Viruses and virus diseases of pea, lentil, faba bean and chickpea. In: Summerfield RJ, ed. *World crops: Cool season food legumes. A global perspective of the problems and prospects for crop improvement in pea, lentil, faba bean and chickpea*, Springer Netherlands: Dordrecht. Vol 5, 591–615.

Brault V, Herrbach E, Reinbold C, 2007. Electron microscopy studies on luteovirid transmission by aphids. *Micron* **38**, 302–12. doi: 10.1016/j.micron.2006.04.005.

Brault V, Mutterer J, Scheidecker D *et al.*, 2000. Effects of point mutations in the readthrough domain of the beet western yellows virus minor capsid protein on virus accumulation in planta and on transmission by aphids. *Journal of Virology* **74**, 1140–8. doi: 10.1128/jvi.74.3.1140-1148.2000.

Brault V, Périgon S, Reinbold C *et al.*, 2005. The polerovirus minor capsid protein determines vector specificity and intestinal tropism in the aphid. *Journal of Virology* **79**, 9685–93. doi: 10.1128/JVI.79.15.9685-9693.2005.

Brault V, Tanguy S, Reinbold C *et al.*, 2010. Transcriptomic analysis of intestinal genes following acquisition of pea enation mosaic virus by the pea aphid *Acyrthosiphon pisum*. *The Journal of General Virology* **91**, 802–8. doi: 10.1099/vir.0.012856-0.

Brault V, van den Heuvel JF, Verbeek M *et al.*, 1995. Aphid transmission of beet western yellows luteovirus requires the minor capsid read-through protein P74. *The EMBO Journl* **14**, 650–9.

Bressan A, Watanabe S, 2011. Immunofluorescence localisation of *Banana bunchy top virus* (family *Nanoviridae*) within the aphid vector, Pentalonia nigronervosa, suggests a virus tropism distinct from aphid-transmitted luteoviruses. *Virus Research* **155**, 520–5. doi: 10.1016/j.virusres.2010.12.005.

Briddon RW, Ghabrial S, Lin NS, 2012. Satellites and other virus-dependent nucleic acids. In: King AMQ, Lefkowitz E, Adams MJ, Carstens EB, eds. *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, Elsevier academic press: Oxford, 1211–9.

Briddon RW, Martin DP, Roumagnac P *et al.*, 2018. *Alphasatellitidae*: a new family with two subfamilies for the classification of geminivirus- and nanovirus-associated alphasatellites. *Archives of Virology* **163**, 2587–600. doi: 10.1007/s00705-018-3854-2.

Broadbent L, Heathcote DG, 1958. Properties and host range of turnip crinkle, rosette and yellow mosaic viruses. *Annals of Applied Biology* **46**, 585–92. doi: 10.1111/j.1744-7348.1958.tb02242.x.

Bruenn JA, Warner BE, Yerramsetty P, 2015. Widespread mitovirus sequences in plant genomes. *PeerJ* **3**, e876. doi: 10.7717/peerj.876.

Brunissen L, Cherqui A, Pelletier Y, 2009. Host-plant mediated interactions between two aphid species. *Entomologia Experimentalis et Applicata* **132**, 30–8. doi: 10.1111/j.1570-7458.2009.00862.x.

Bruyère A, Brault V, Ziegler-Graff V *et al.*, 1997. Effects of mutations in the beet western yellows virus readthrough protein on its expression and packaging and on virus accumulation, symptoms, and aphid transmission. *Virology* **230**, 323–34. doi: 10.1006/viro.1997.8476.

Buchner P, 1965. *Endosymbiosis of animals with plant microorganisms*. New York: Interscience Publishers.

Büttner C, Bargen S von, Bandte M, 2011. *Cherry leaf roll virus*. In: Hadidi A, Barba M, Candresse T, Jelkmann W, eds. *Virus and virus-like diseases of pome and stone fruits*, APS Press/American Phytopathological Society: St. Paul, Minnesota, 119–25.

Camacho C, Coulouris G, Avagyan V *et al.*, 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421. doi: 10.1186/1471-2105-10-421.

Candresse T, Marais A, Faure C, 2013. First report of *Southern tomato virus* on tomatoes in Southwest France. *Plant Disease* **97**, 1124. doi: 10.1094/PDIS-01-13-0017-PDN.

Canto T, Aranda MA, Fereres A, 2009. Climate change effects on physiology and population processes of hosts and vectors that influence the spread of hemipteran-borne plant viruses. *Global Change Biology* **15**, 1884–94.

Cardin L, Delecolle B, Moury B, 2009. First report of cucumber mosaic virus and turnip vein-clearing virus in *Dichondra repens* in France, Italy, and China. *Plant Disease* **93**, 201. doi: 10.1094/PDIS-93-2-0201B.

Carvajal-Yepes M, Olaya C, Lozano I, 2014. Unraveling complex viral infections in cassava (*Manihot esculenta* Crantz) from Colombia. *Virus Research* **186**, 76–86. doi: 10.1016/j.virusres.2013.12.011.

Carvalho CM, Santos AA, Pires SR *et al.*, 2008. Regulated nuclear trafficking of rpL10A mediated by NIK1 represents a defense strategy of plant cells against virus. *PLOS Pathogens* **4**, e1000247. doi: 10.1371/journal.ppat.1000247.

Casper R, Meyer S, 1981. Anwendung des ELISA-Verfahrens zum Nachweis pflanzenpathogener Viren. *Nachrichtenblatt des Deutschen Pflanzenschutzdienstes* **33**, 49–54.

Castle SJ, Berger PH, 1993. Rates of growth and increase of *Myzus persicae* on virus-infected potatoes according to type of virus-vector relationship. *Entomologia Experimentalis et Applicata* **69**, 51–60. doi: 10.1111/j.1570-7458.1993.tb01727.x.

Castle SJ, Mowry TM, Berger PH, 1998. Differential settling by *Myzus persicae* (Homoptera: Aphididae) on various virus infected host plants. *Annals of the Entomological Society of America* **91**, 661–7. doi: 10.1093/aesa/91.5.661.

Chatzivassiliou EK, Giakountis A, Kumari SG, 2016. Viruses affecting lentil (*Lens culinaris* Medik.) in Greece; incidence and genetic variability of *Bean leafroll virus* and *Pea enation mosaic virus*. *Phytopathologia Mediterranea* **55**, 239–52.

Chavez JD, Cilia M, Weisbrod CR *et al.*, 2012. Cross-linking measurements of the Potato leafroll virus reveal protein interaction topologies required for virion stability, aphid transmission, and virus-plant interactions. *Journal of Proteome Research* **11**, 2968–81. doi: 10.1021/pr300041t.

Chen H, Boutros PC, 2011. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35. doi: 10.1186/1471-2105-12-35.

Chen S, Jiang G, Wu J, 2016. Characterization of a novel polerovirus infecting maize in China. *Viruses* **8**. doi: 10.3390/v8050120.

Chin LS, Foster JL, Falk BW, 1993. The beet western yellows virus ST9-associated RNA shares structural and nucleotide sequence homology with carmo-like viruses. *Virology* **192**, 473–82. doi: 10.1006/viro.1993.1063.

Chu PW, Helms K, 1988. Novel virus-like particles containing circular single-stranded DNAs associated with subterranean clover stunt disease. *Virology* **167**, 38–49.

Cilia M, Peter KA, Bereman MS *et al.*, 2012. Discovery and targeted LC-MS/MS of purified polerovirus reveals differences in the virus-host interactome associated with altered aphid transmission. *PLOS ONE* **7**, e48177. doi: 10.1371/journal.pone.0048177.

Clark MF, Adams AN, 1977. Characteristics of the microplate method of enzyme-linked immunosorbent assay for the detection of plant viruses. *The Journal of General Virology* **34**, 475–83. doi: 10.1099/0022-1317-34-3-475.

Clement SL, 2006. Pea aphid outbreaks and virus epidemics on peas in the US Pacific Northwest: histories, mysteries, and challenges. *Plant Health Progress* **7**, 34. doi: 10.1094/PHP-2006-1018-01-RV.

Coetzee B, Freeborough M-J, Maree HJ, 2010. Deep sequencing analysis of viruses infecting grapevines: Virome of a vineyard. *Virology* **400**, 157–63. doi: 10.1016/j.virol.2010.01.023.

Cohen D, Chavan RR, Blouin AG, 2012. First report of *Turnip vein clearing virus* and *Ribgrass mosaic virus* from New Zealand. *Australasian Plant Disease Notes* **7**, 67–9. doi: 10.1007/s13314-012-0050-1.

Collar JL, Avilla C, Fereres A, 1997. New correlations between aphid stylet paths and nonpersistent virus transmission. *Environmental Entomology* **26**, 537–44. doi: 10.1093/ee/26.3.537.

Cooper WR, Goggin FL, 2005. Effects of jasmonate-induced defenses in tomato on the potato aphid, *Macrosiphum euphorbiae*. *Entomologia Experimentalis et Applicata* **115**, 107–15. doi: 10.1111/j.1570-7458.2005.00289.x.

Coutts RHA, Livieratos IC, 2003. A rapid method for sequencing the 5'- and 3'-termini of double-stranded RNA viral templates using RLM-RACE. *Journal of Phytopathology* **151**, 525–7. doi: 10.1046/j.1439-0434.2003.00755.x.

Cropley R, 1961. Cherry leaf-roll virus. *Annals of Applied Biology* **49**, 524–9. doi: 10.1111/j.1744-7348.1961.tb03645.x.

Czotter N, Molnar J, Szabó E *et al.*, 2018. NGS of virus-derived small RNAs as a diagnostic method used to determine viromes of Hungarian vineyards. *Frontiers in Microbiology* **9**, 758. doi: 10.3389/fmicb.2018.00122.

Damsteegt VD, Hewings AD, 1986. Comparative transmission of soybean dwarf virus by three geographically diverse populations of *Aulacorthum* (= *Acyrthosiphon*) *solani*. *Annals of Applied Biology* **109**, 453–63. doi: 10.1111/j.1744-7348.1986.tb03202.x.

Dean FB, Nelson JR, Giesler TL, 2001. Rapid amplification of plasmid and phage DNA using Phi 29 DNA polymerase and multiply-primed rolling circle amplification. *Genome Research* **11**, 1095–9. doi: 10.1101/gr.180501.

Debrick KS, 1973. Quantitative assay for plant viruses using serologically specific electron microscopy. *Virology* **56**, 652–3. doi: 10.1016/0042-6822(73)90068-8.

Del Fabbro C, Scalabrin S, Morgante M, 2013. An extensive evaluation of read trimming effects on Illumina NGS data analysis. *PLOS ONE* **8**, e85024. doi: 10.1371/journal.pone.0085024.

Demler SA, Rucker DG, Nooruddin L, 1994. Replication of the satellite RNA of pea enation mosaic virus is controlled by RNA 2-encoded functions. *Journal of General Virology* **75 (Pt 6)**, 1399–406. doi: 10.1099/0022-1317-75-6-1399.

Demler SA, Zoeten GA de, 1989. Characterization of a satellite RNA associated with pea enation mosaic virus. *Journal of General Virology* **70**, 1075–84. doi: 10.1099/0022-1317-70-5-1075.

Dietzgen RG, Innes DJ, Bejerman N, 2015. Complete genome sequence and intracellular protein localization of datura yellow vein nucleorhabdovirus. *Virus Research* **205**, 7–11. doi: 10.1016/j.virusres.2015.05.001.

Dilcher M, Faye O, Faye O *et al.*, 2015. Zahedan rhabdovirus, a novel virus detected in ticks from Iran. *Virology Journal* **12**, 183. doi: 10.1186/s12985-015-0410-5.

Ding S-W, 2010. RNA-based antiviral immunity. *Nature Reviews Immunology* **10**, 632–44. doi: 10.1038/nri2824.

Ding S-W, Lu R, 2011. Virus-derived siRNAs and piRNAs in immunity and pathogenesis. *Current Opinion in Virology* **1**, 533–44. doi: 10.1016/j.coviro.2011.10.028.

Dodds JA, Morris TJ, Jordan RL, 1984. Plant viral double-stranded RNA. *Annual Review of Phytopathology* **22**, 151–68. doi: 10.1146/annurev.py.22.090184.001055.

Dorokhov YL, Ivanov PA, Novikov VK *et al.*, 1994. Complete nucleotide sequence and genome organization of a tobamovirus infecting cruciferae plants. *FEBS Letters* **350**, 5–8. doi: 10.1016/0014-5793(94)00721-7.

Douglas GB, Wang Y, Waghorn GC *et al.*, 1995. Liveweight gain and wool production of sheep grazing *Lotus corniculatus* and lucerne (*Medicago sativa*). *New Zealand Journal of Agricultural Research* **38**, 95–104. doi: 10.1080/00288233.1995.9513108.

Doumayrou J, Sheber M, Bonning BC, 2017. Quantification of *Pea enation mosaic virus* 1 and 2 during infection of *Pisum sativum* by one step real-time RT-PCR. *Journal of Virological Methods* **240**, 63–8. doi: 10.1016/j.jviromet.2016.11.013.

Drijfhout E, Bos L, 1977. The identification of two new strains of bean common mosaic virus. *Netherlands Journal of Plant Pathology* **83**, 13–25. doi: 10.1007/BF01976508.

Durand P, Canard L, Mornon JP, 1997. Visual BLAST and Visual FASTA: graphic workbenches for interactive analysis of full BLAST and FASTA outputs under Microsoft Windows 95/NT. *Bioinformatics* **13**, 407–13. doi: 10.1093/bioinformatics/13.4.407.

Eddouks M, Lemhadri A, Michel J-B, 2004. Caraway and caper: potential anti-hyperglycaemic plants in diabetic rats. *Journal of Ethnopharmacology* **94**, 143–8. doi: 10.1016/j.jep.2004.05.006.

Edgar RC, 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792–7. doi: 10.1093/nar/gkh340.

Edwards K, Johnstone C, Thompson C, 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research* **19**, 1349.

Eid J, Fehr A, Gray J *et al.*, 2009. Real-time DNA sequencing from single polymerase molecules. *Science (New York, N.Y.)* **323**, 133–8. doi: 10.1126/science.1162986.

Eigenbrode SD, Ding H, Shiel P, 2002. Volatiles from potato plants infected with potato leafroll virus attract and arrest the virus vector, *Myzus persicae* (Homoptera: Aphididae). *Proceedings: Biological Sciences* **269**, 455–60. doi: 10.1098/rspb.2001.1909.

Elbeaino T, Digiaro M, Mielke-Ehret N, 2018. ICTV Virus Taxonomy Profile: *Fimoviridae*. *The Journal of General Virology* **99**, 1478–9. doi: 10.1099/jgv.0.001143.

Elbeaino T, Digiaro M, Uppala M, 2014. Deep sequencing of pigeonpea sterility mosaic virus discloses five RNA segments related to emaraviruses. *Virus Research* **188**, 27–31. doi: 10.1016/j.virusres.2014.03.022.

Elbeaino T, Digiaro M, Uppala M, 2015. Deep sequencing of dsRNAs recovered from mosaic-diseased pigeonpea reveals the presence of a novel emaravirus: pigeonpea sterility mosaic virus 2. *Archives of Virology* **160**, 2019–29. doi: 10.1007/s00705-015-2479-y.

EPPO, 2019. EPPO A1 List. [https://www.eppo.int/ACTIVITIES/plant_quarantine/A1_list]. Accessed 30 August 2019.

Falk BW, Tsai JH, 1998. Biology and molecular biology of viruses in the genus *Tenuivirus*. *Annual Review of Phytopathology* **36**, 139–63. doi: 10.1146/annurev.phyto.36.1.139.

Fletcher J, Tang J, Blouin A, 2016. *Red clover vein mosaic virus* - a novel virus to New Zealand that is widespread in legumes. *Plant Disease* **100**, 890–5. doi: 10.1094/PDIS-04-15-0465-RE.

Fletcher JD, 1993. Surveys of virus diseases in pea, lentil, dwarf and broad bean crops in South Island, New Zealand. *New Zealand Journal of Crop and Horticultural Science* **21**, 45–52. doi: 10.1080/01140671.1993.9513745.

Fletcher JD, Goulden DS, Russell AC, 1989. Breeding for resistance to pea seed-borne mosaic virus. *Grain Legumes: National Symposium and Workshop. Agronomy Society of New Zealand Special Publication* **7**, 124.

Flusberg BA, Webster DR, Lee JH *et al.*, 2010. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nature Methods* **7**, 461–5. doi: 10.1038/nmeth.1459.

Fontes EPB, Santos AA, Luz DF, 2004. The geminivirus nuclear shuttle protein is a virulence factor that suppresses transmembrane receptor kinase activity. *Genes & Development* **18**, 2545–56. doi: 10.1101/gad.1245904.

Foyer CH, Lam H-M, Nguyen HT *et al.*, 2016. Neglecting legumes has compromised human health and sustainable food production. *Nature plants* **2**, 16112. doi: 10.1038/nplants.2016.112.

Frago E, Dicke M, Godfray HCJ, 2012. Insect symbionts as hidden players in insect-plant interactions. *Trends in Ecology & Evolution* **27**, 705–11. doi: 10.1016/j.tree.2012.08.013.

Franz A, Makkouk KM, Vetten HJ, 1998. Acquisition, retention and transmission of faba bean necrotic yellows virus by two of its aphid vectors, *Aphis craccivora* (Koch) and *Acyrthosiphon pisum* (Harris). *Journal of Phytopathology* **146**, 347–55. doi: 10.1111/j.1439-0434.1998.tb04703.x.

Franz AWE, van der Wilk F, Verbeek M, 1999. Faba bean necrotic yellows virus (genus *Nanovirus*) requires a helper factor for its aphid transmission. *Virology* **262**, 210–9. doi: 10.1006/viro.1999.9904.

Froussard P, 1992. A random-PCR method (rPCR) to construct whole cDNA library from low amounts of RNA. *Nucleic Acids Research* **20**, 2900. doi: 10.1093/nar/20.11.2900.

Fu ZF, 2005. Genetic comparison of the rhabdoviruses from animals and plants. *Current Topics in Microbiology and Immunology* **292**, 1–24.

Fuchs M, Schmitt-Keichinger C, Sanfaçon H, 2017. A renaissance in nepovirus research provides new insights into their molecular interface with hosts and vectors. In: Kielian M, Thomas CM, Marilyn JR, eds. *Advances in Virus Research*, Elsevier Science: Saint Louis. Vol 97, 61–105.

Fukatsu T, Tsuchida T, Nikoh N, 2001. *Spiroplasma* symbiont of the pea aphid, *Acyrthosiphon pisum* (Insecta: Homoptera). *Applied and Environmental Microbiology* **67**, 1284–91. doi: 10.1128/AEM.67.3.1284-1291.2001.

Gaafar Y, Cordsen Nielsen G, Ziebell H, 2018a. Molecular characterisation of the first occurrence of *Pea necrotic yellow dwarf virus* in Denmark. *New Disease Reports* **37**, 16. doi: 10.5197/j.2044-0588.2018.037.016.

Gaafar Y, Grausgruber-Gröger S, Ziebell H, 2016. *Vicia faba V. sativa* and *Lens culinaris* as new hosts for *Pea necrotic yellow dwarf virus* in Germany and Austria. *New Disease Reports* **34**, 28. doi: 10.5197/j.2044-0588.2016.034.028.

Gaafar Y, Lüddecke P, Heidler C *et al.*, 2019a. First report of *Southern tomato virus* in German tomatoes. *New Disease Reports* **40**, 1. doi: 10.5197/j.2044-0588.2019.040.001.

Gaafar Y, Sieg-Müller A, Lüddecke P *et al.*, 2019b. First report of natural infection of beetroot with *Beet soil-borne virus*. *New Disease Reports* **40**, 5. doi: 10.5197/j.2044-0588.2019.040.005.

Gaafar Y, Sieg-Müller A, Lüddecke P *et al.*, 2019c. First report of *Turnip crinkle virus* infecting garlic mustard (*Alliaria petiolata*) in Germany. *New Disease Reports* **39**, 9. doi: 10.5197/j.2044-0588.2019.039.009.

Gaafar Y, Timchenko T, Ziebell H, 2017. First report of *Pea necrotic yellow dwarf virus* in The Netherlands. *New Disease Reports* **35**, 23. doi: 10.5197/j.2044-0588.2017.035.023.

Gaafar YZA, Abdelgalil MAM, Knierim D *et al.*, 2018b. First report of physostegia chlorotic mottle virus on tomato (*Solanum lycopersicum*) in Germany. *Plant Disease* **102**, 255. doi: 10.1094/PDIS-05-17-0737-PDN.

Gaafar YZA, Richert-Pöggeler KR, Maaß C, 2019d. Characterisation of a novel nucleorhabdovirus infecting alfalfa (*Medicago sativa*). *Virology Journal* **16**, 113. doi: 10.1186/s12985-019-1147-3.

Gaafar YZA, Richert-Pöggeler KR, Sieg-Müller A *et al.*, 2019e. A divergent strain of melon chlorotic spot virus isolated from black medic (*Medicago lupulina*) in Austria. *Virology Journal* **16**, 297. doi: 10.1186/s12985-019-1195-8.

Gaafar YZA, Richert-Pöggeler KR, Sieg-Müller A *et al.*, 2019f. Caraway yellows virus, a novel nepovirus from *Carum carvi*. *Virology Journal* **16**, 529. doi: 10.1186/s12985-019-1181-1.

Gaafar YZA, Ziebell H, 2019a. Complete genome sequence of a highly divergent carrot torradovirus 1 strain from *Apium graveolens*. *Archives of Virology*. doi: 10.1007/s00705-019-04272-3.

Gaafar YZA, Ziebell H, 2019b. Two divergent isolates of turnip yellows virus from pea and rapeseed and first report of turnip yellows virus-associated RNA in Germany. *Microbiology Resource Announcements* **8**. doi: 10.1128/MRA.00214-19.

Gabrys B, Tjallingii WF, van Beek TA, 1997. Analysis of EPG recorded probing by cabbage aphid on host plant parts with different glucosinolate contents. *Journal of Chemical Ecology* **23**, 1661–73. doi: 10.1023/B:JOEC.0000006442.56544.1a.

Galinier R, van Beurden S, Amilhat E *et al.*, 2012. Complete genomic sequence and taxonomic position of eel virus European X (EVEX), a rhabdovirus of European eel. *Virus Research* **166**, 1–12. doi: 10.1016/j.virusres.2012.02.020.

Gallet R, Kraberger S, Filloux D *et al.*, 2018. Nanovirus-alphasatellite complex identified in *Vicia cracca* in the Rhône delta region of France. *Archives of Virology* **163**, 695–700. doi: 10.1007/s00705-017-3634-4.

Garcıa-Arenal F, Palukaitis P, 2009. Cucumber mosaic virus. In: van Regenmortel MHV, ed. *Desk Encyclopedia of Plant and Fungal Virology*, Elsevier professional: s.l., 171–6.

Garret A, Kerlan C, Thomas D, 1993. The intestine is a site of passage for potato leafroll virus from the gut lumen into the haemocoel in the aphid vector, Myzus persicae Sulz. *Archives of Virology* **131**, 377–92. doi: 10.1007/bf01378639.

Garret A, Kerlan C, Thomas D, 1996. Ultrastructural study of acquisition and retention of potato leafroll luteovirus in the alimentary canal of its aphid vector, *Myzus persicae* Sulz. *Archives of Virology* **141**, 1279–92. doi: 10.1007/BF01718830.

Garrett RG, 1991. Impact of viruses on pasture legume productivity. In: *Proceedings of the Department of Agriculture, Victoria, white clover conference*, 50–7.

Gentile M, Gelderblom HR, 2014. Electron microscopy in rapid viral diagnosis: an update. *The New Microbiologica* **37**, 403–22.

Giakountis A, Skoufa A, Paplomatas EI, 2015. Molecular characterization and phylogenetic analysis of a Greek lentil isolate of *Pea seed-borne mosaic virus*. *Phytoparasitica* **43**, 615–28. doi: 10.1007/s12600-015-0495-9.

Gildow FE, 1993. Evidence for receptor-mediated endocytosis regulating luteovirus acquisition by aphids. *Phytopathology* **83**, 270. doi: 10.1094/Phyto-83-270.

Gildow FE, Reavy B, Mayo MA *et al.*, 2000. Aphid acquisition and cellular transport of potato leafroll virus-like particles lacking P5 readthrough protein. *Phytopathology* **90**, 1153–61. doi: 10.1094/PHYTO.2000.90.10.1153.

Ginestet C, 2011. ggplot2: elegant graphics for data analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **174**, 245–6. doi: 10.1111/j.1467-985X.2010.00676_9.x.

Goldsmith CS, Miller SE, 2009. Modern uses of electron microscopy for detection of viruses. *Clinical Microbiology Reviews* **22**, 552–63. doi: 10.1128/CMR.00027-09.

Gómez-Gómez L, Boller T, 2000. FLS2: An LRR receptor–like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Molecular Cell* **5**, 1003–11. doi: 10.1016/S1097-2765(00)80265-8.

Goodin MM, Jackson AO, 2002. *Nucleorhabdovirus*. In: Tidona CA, Darai G, Osmond CB, eds. *The Springer index of viruses*, Springer: Berlin, London, 1074–7.

Goodin MM, Min B-E, 2012. Virus-host protein interactions of plant-adapted rhabdoviruses. In: Dietzgen RG, Kuzmin IV, eds. *Rhabdoviruses. Molecular taxonomy, evolution, genomics, ecology, host-vector interactions, cytopathology and control*, Caister Academic Press: Norfolk, UK, 133–46.

Gorbalenya AE, Koonin EV, Donchenko AP, 1989. Two related superfamilies of putative helicases involved in replication, recombination, repair and expression of DNA and RNA genomes. *Nucleic Acids Research* **17**, 4713–30.

Gorbalenya AE, Snijder EJ, 1996. Viral cysteine proteinases. *Perspectives in Drug Discovery and Design* **6**, 64–86. doi: 10.1007/BF02174046.

Gottlieb Y, Zchori-Fein E, Mozes-Daube N *et al.*, 2010. The transmission efficiency of tomato yellow leaf curl virus by the whitefly *Bemisia tabaci* is correlated with the presence of a specific symbiotic bacterium species. *Journal of Virology* **84**, 9310–7. doi: 10.1128/JVI.00423-10.

Graichen K, Rabenstein F, 1996. European isolates of beet western yellows virus (BWYV) from oilseed rape (*Brassica napus* L. ssp. napus) are non-pathogenic on sugar beet (*Beta vulgaris* L var. altissima) but represent isolates of turnip yellows virus (TuYV). *Zeitschrift für Pflanzenkrankheiten und Pflanzenschutz / Journal of Plant Diseases and Protection* **103**, 233–45.

Graichen K, Schliephake E, 1999. Infestation of winter oilseed rape by turnip yellows luteovirus and its effect on yield in Germany. In: *10 th International Rapeseed Congress - New Horizons from an Old Crop*, 131–6.

Gray AM, Flatt PR, 1997. Pancreatic and extra-pancreatic effects of the traditional anti-diabetic plant, *Medicago sativa* (lucerne). *British Journal of Nutrition* **78**, 325. doi: 10.1079/BJN19970150.

Gray S, Cilia M, Ghanim M, 2014. Circulative, "nonpropagative" virus transmission: an orchestra of virus-, insect-, and plant-derived instruments. *Advances in Virus Research* **89**, 141–99. doi: 10.1016/B978-0-12-800172-1.00004-5.

Gray S, Gildow FE, 2003. Luteovirus-aphid interactions. *Annual Review of Phytopathology* **41**, 539–66. doi: 10.1146/annurev.phyto.41.012203.105815.

Grigoras I, Ginzo AIdC, Martin DP *et al.*, 2014. Genome diversity and evidence of recombination and reassortment in nanoviruses from Europe. *The Journal of General Virology* **95**, 1178–91. doi: 10.1099/vir.0.063115-0.

Grigoras I, Gronenborn B, Vetten HJ, 2010a. First report of a nanovirus disease of pea in Germany. *Plant Disease* **94**, 642. doi: 10.1094/PDIS-94-5-0642C.

Grigoras I, Timchenko T, Grande-Pérez A, 2010b. High variability and rapid evolution of a nanovirus. *Journal of Virology* **84**, 9105–17. doi: 10.1128/JVI.00607-10.

Grigoras I, Timchenko T, Katul L, 2009. Reconstitution of authentic nanovirus from multiple cloned DNAs. *Journal of Virology* **83**, 10778–87. doi: 10.1128/JVI.01212-09.

Grigoras I, Vetten H-J, Commandeur U, 2018. Nanovirus DNA-N encodes a protein mandatory for aphid transmission. *Virology* **522**, 281–91. doi: 10.1016/j.virol.2018.07.001.

Grylls NE, Butler FC, 1956. An aphid transmitted virus affecting subterranean clover. *Journal of the Australian Institute of Agricultural Science* **22**, 73–4.

Guo H, Sun Y, Li Y *et al.*, 2014. Elevated CO2 alters the feeding behaviour of the pea aphid by modifying the physical and chemical resistance of *Medicago truncatula*. *Plant, cell & environment* **37**, 2158–68. doi: 10.1111/pce.12306.

Gutierrez AP, Morgan DJ, Havenstein DE, 1971. The ecology of *Aphis craccivora* Koch and subterranean clover stunt virus. I. The phenology of aphid populations and the epidemiology of virus in pastures in South-East Australia. *The Journal of Applied Ecology* **8**, 699. doi: 10.2307/2402678.

Guyatt KJ, Proll DF, Menssen A, 1996. The complete nucleotide sequence of bean yellow mosaic potyvirus RNA. *Archives of Virology* **141**, 1231–46. doi: 10.1007/BF01718827.

Haan P de, Wagemakers L, Peters D, 1989. Molecular cloning and terminal sequence determination of the S and M RNAs of tomato spotted wilt virus. *The Journal of General Virology* **70 (Pt 12)**, 3469–73. doi: 10.1099/0022-1317-70-12-3469.

Hagedorn DJ, Khan TN, 1984. *Compendium of pea diseases*. American Phytopathological Society St Paul, MN.

Hamilton RI, Edwardson JR, Francki RIB *et al.*, 1981. Guidelines for the Identification and Characterization of Plant Viruses. *The Journal of General Virology* **54**, 223–41. doi: 10.1099/0022-1317-54-2-223.

Hanley-Bowdoin L, Bejarano ER, Robertson D, 2013. Geminiviruses: masters at redirecting and reprogramming plant processes. *Nature Reviews Microbiology* **11**, 777–88. doi: 10.1038/nrmicro3117.

Harris KF, 1977. An ingestion-egestion hypothesis of noncirculative virus transmission. In: Harris KF, Maramorosch K, eds. *Aphids as virus vectors*, Elsevier, 165–220.

Heydarnejad J, Kamali M, Massumi H *et al.*, 2017. Identification of a nanovirus-alphasatellite complex in *Sophora alopecuroides*. *Virus Research* **235**, 24–32.

Hill JH, Whitham SA, 2014. Control of virus diseases in soybeans. *Advances in Virus Research* **90**, 355–90. doi: 10.1016/B978-0-12-801246-8.00007-X.

Ho T, Tzanetakis IE, 2014. Development of a virus detection and discovery pipeline using next generation sequencing. *Virology* **471-473**, 54–60. doi: 10.1016/j.virol.2014.09.019.

Hoang AT, Zhang H-M, Yang J *et al.*, 2011. Identification, characterization, and distribution of Southern rice black-streaked dwarf virus in Vietnam. *Plant Disease* **95**, 1063–9. doi: 10.1094/PDIS-07-10-0535.

Hogenhout SA, Ammar E-D, Whitfield AE, 2008. Insect vector interactions with persistently transmitted viruses. *Annual Review of Phytopathology* **46**, 327–59. doi: 10.1146/annurev.phyto.022508.092135.

Houk EJ, Griffiths GW, 1980. Intracellular Symbiotes of the Homoptera. *Annual Review of Entomology* **25**, 161–87. doi: 10.1146/annurev.en.25.010180.001113.

Hull R, 1969. Alfalfa mosaic virus. In: Smith KM, Lauffer MA, Bang FB, eds. *Advances in Virus Research*, Academic Press: New York. Vol 15, 365–433.

Hull R (ed), 2009. *Comparative plant virology,* 2nd edn. Amsterdam, Boston: Elsevier/Academic Press.

Hull R, 2014. *Plant virology*. Amsterdam: Academic Press.

Hulme PE, 2017. Climate change and biological invasions: evidence, expectations, and response options. *Biological Reviews* **92**, 1297–313. doi: 10.1111/brv.12282.

Iacobellis NS, Lo Cantore P, Capasso F, 2005. Antibacterial activity of *Cuminum cyminum* L. and *Carum carvi* L. essential oils. *Journal of Agricultural and Food Chemistry* **53**, 57–61. doi: 10.1021/jf0487351.

Ido Y, Nakahara KS, Uyeda I, 2012. White clover mosaic virus-induced gene silencing in pea. *Journal of General Plant Pathology* **78**, 127–32. doi: 10.1007/s10327-012-0360-3.

Idris A, Al-Saleh M, Piatek MJ, 2014. Viral metagenomics: analysis of begomoviruses by illumina high-throughput sequencing. *Viruses* **6**, 1219–36. doi: 10.3390/v6031219.

Idris AM, Shahid MS, Briddon RW, 2011. An unusual alphasatellite associated with monopartite begomoviruses attenuates symptoms and reduces betasatellite accumulation. *The Journal of General Virology* **92**, 706–17. doi: 10.1099/vir.0.025288-0.

Illumina. Sequencing Platforms | Compare NGS platform applications & specifications. [https://www.illumina.com/systems/sequencing-platforms.html]. Accessed 10 August 2019.

Ip CLC, Loose M, Tyson JR *et al.*, 2015. MinION Analysis and Reference Consortium: Phase 1 data release and analysis. *F1000Research* **4**, 1075. doi: 10.12688/f1000research.7201.1.

Iseni F, Barge A, Baudin F, 1998. Characterization of rabies virus nucleocapsids and recombinant nucleocapsid-like structures. *The Journal of General Virology* **79 (Pt 12)**, 2909–19. doi: 10.1099/0022-1317-79-12-2909.

Ishikawa H, Yamaji M, 1985. Symbionin, an aphid endosymbiont-specific protein—I: Production of insects deficient in symbiont. *Insect Biochemistry* **15**, 155–63.

Isogai M, Tatuto N, Ujiie C, 2012. Identification and characterization of blueberry latent spherical virus, a new member of subgroup C in the genus *Nepovirus*. *Archives of Virology* **157**, 297–303. doi: 10.1007/s00705-011-1177-7.

Jackson AO, Dietzgen RG, Goodin MM, 2005. Biology of plant rhabdoviruses. *Annual Review of Phytopathology* **43**, 623–60. doi: 10.1146/annurev.phyto.43.011205.141136.

Ji X-L, Yu N-T, Qu L, 2019. Banana bunchy top virus (BBTV) nuclear shuttle protein interacts and re-distributes BBTV coat protein in *Nicotiana benthamiana*. *3 Biotech* **9**, 121. doi: 10.1007/s13205-019-1656-1.

Jinn T-L, Stone JM, Walker JC, 2000. HAESA, an Arabidopsis leucine-rich repeat receptor kinase, controls floral organ abscission. *Genes & Development* **14**, 108–17.

Jo Y, Bae J-Y, Kim S-M, 2018a. Barley RNA viromes in six different geographical regions in Korea. *Scientific Reports* **8**, 13237. doi: 10.1038/s41598-018-31671-4.

Jo Y, Lian S, Chu H *et al.*, 2018b. Peach RNA viromes in six different peach cultivars. *Scientific Reports* **8**, 1844. doi: 10.1038/s41598-018-20256-w.

Johne R, Müller H, Rector A, 2009. Rolling-circle amplification of viral DNA genomes using phi29 polymerase. *Trends in Microbiology* **17**, 205–11. doi: 10.1016/j.tim.2009.02.004.

Johnstone GR, Mclean GD, 1987. Virus diseases of subterranean clover. *Annals of Applied Biology* **110**, 421–40. doi: 10.1111/j.1744-7348.1987.tb03274.x.

Jones AT, Koenig R, Lesemann D-E, 1990. Serological comparison of isolates of cherry leaf roll virus from diseased beech and birch trees in a forest decline area in Germany with other isolates of the virus. *Journal of Phytopathology* **129**, 339–44. doi: 10.1111/j.1439-0434.1990.tb04311.x.

Jones RAC, 2016. Future scenarios for plant virus pathogens as climate change progresses. *Advances in Virus Research* **95**, 87–147. doi: 10.1016/bs.aivir.2016.02.004.

Kaiser WJ, 1982. Natural hosts and vectors of tobacco streak virus in Eastern Washington. *Phytopathology* **72**, 1508. doi: 10.1094/Phyto-72-1508.

Kambara H, Takahashi S, 1993. Multiple-sheathflow capillary array DNA analyser. *Nature* **361**, 565–6. doi: 10.1038/361565a0.

Kaplan IB, Lee L, Ripoll DR, 2007. Point mutations in the potato leafroll virus major capsid protein alter virion stability and aphid transmission. *The Journal of General Virology* **88**, 1821–30. doi: 10.1099/vir.0.82837-0.

Katul L, 1992. *Serological and molecular characterization of bean leaf roll virus (BLRV) and faba bean necrotic yellows virus (FBNYV)*. Ph. D. dissertation, University of Göttingen, Germany.

Kemper S, Schaack D, Schenck W von, 2019. AMI Markt Bilanz Getreide · Ölsaaten · Futtermittel 2019. [www.ami-informiert.de]. Accessed 8 July 2019.

Kesanakurti P, Belton M, Saeed H, 2016. Screening for plant viruses by next generation sequencing using a modified double strand RNA extraction protocol with an internal amplification control. *Journal of Virological Methods* **236**, 35–40. doi: 10.1016/j.jviromet.2016.07.001.

Khalifa ME, Pearson MN, 2013. Molecular characterization of three mitoviruses co-infecting a hypovirulent isolate of *Sclerotinia sclerotiorum* fungus. *Virology* **441**, 22–30. doi: 10.1016/j.virol.2013.03.002.

Khetarpal RK, Maury Y, 1987. Pea seed-borne mosaic virus: a review. *Agronomie* **7**, 215–24. doi: 10.1051/agro:19870401.

Kirby MJ, Guise CM, Adams AN, 2001. Comparison of bioassays and laboratory assays for apple stem grooving virus. *Journal of Virological Methods* **93**, 167–73. doi: 10.1016/S0166-0934(01)00271-3.

Kircher M, Kelso J, 2010. High-throughput DNA sequencing--concepts and limitations. *BioEssays* **32**, 524–36. doi: 10.1002/bies.200900181.

Kliot A, Ghanim M, 2013. The role of bacterial chaperones in the circulative transmission of plant viruses by insect vectors. *Viruses* **5**, 1516–35. doi: 10.3390/v5061516.

Knierim D, Menzel W, Winter S, 2017. Analysis of the complete genome sequence of euphorbia ringspot virus, an atypical member of the genus *Potyvirus*. *Archives of Virology* **162**, 291–3. doi: 10.1007/s00705-016-3087-1.

Knierim D, Menzel W, Winter S, 2019. Immunocapture of virions with virus-specific antibodies prior to high-throughput sequencing effectively enriches for virus-specific sequences. *PLOS ONE* **14**, e0216713. doi: 10.1371/journal.pone.0216713.

Koenig R, Paul HL, 1982. Variants of ELISA in plant virus diagnosis. *Journal of Virological Methods* **5**, 113–25.

Koloniuk I, Přibylová J, Fránová J, 2018. Molecular characterization and complete genome of a novel nepovirus from red clover. *Archives of Virology* **163**, 1387–9. doi: 10.1007/s00705-018-3742-9.

Kon T, Rojas MR, Abdourhamane IK, 2009. Roles and interactions of begomoviruses and satellite DNAs associated with okra leaf curl disease in Mali, West Africa. *Journal of General Virology* **90**, 1001–13. doi: 10.1099/vir.0.008102-0.

Koren S, Walenz BP, Berlin K, 2017. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research* **27**, 722–36. doi: 10.1101/gr.215087.116.

Kosugi S, Hasebe M, Tomita M, 2009. Systematic identification of cell cycle-dependent yeast nucleocytoplasmic shuttling proteins by prediction of composite motifs. *Proceedings of the National academy of Sciences of the United States of America* **106**, 10171–6. doi: 10.1073/pnas.0900604106.

Kraft JM (ed), 2008. *Compendium of pea diseases and pests,* 2nd edn. St. Paul, Minn.: APS Press.

Krapp S, Greiner E, Amin B, 2017. The stress granule component G3BP is a novel interaction partner for the nuclear shuttle proteins of the nanovirus pea necrotic yellow dwarf virus and geminivirus abutilon mosaic virus. *Virus Research* **227**, 6–14. doi: 10.1016/j.virusres.2016.09.021.

Krenz B, Schießl I, Greiner E, 2017. Analyses of pea necrotic yellow dwarf virus-encoded proteins. *Virus Genes* **53**, 454–63. doi: 10.1007/s11262-017-1439-x.

Kreuze JF, Perez A, Untiveros M *et al.*, 2009. 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* **388**, 1–7. doi: 10.1016/j.virol.2009.03.024.

Kumar LM, Foster JA, McFarland C, 2018a. First report of *Barley virus G* in switchgrass (*Panicum virgatum*). *Plant Disease* **102**, 466. doi: 10.1094/PDIS-09-17-1390-PDN.

Kumar S, Stecher G, Li M, 2018b. MEGA X: molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution* **35**, 1547–9. doi: 10.1093/molbev/msy096.

Kumar S, Stecher G, Tamura K, 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* **33**, 1870–4. doi: 10.1093/molbev/msw054.

Kumari SG, Attar N, Mustafayev E, 2009. First report of *Faba bean necrotic yellows virus* affecting legume crops in Azerbaijan. *Plant Disease* **93**, 1220. doi: 10.1094/PDIS-93-11-1220C.

Kumari SG, Makkouk KM, 2007. Virus diseases of faba bean *(Vicia faba* L.) in Asia and Africa. *Plant Viruses* **1**, 93–105.

Kumari SG, Rodoni B, Vetten H-J *et al.*, 2010. Detection and partial characterization of *Milk vetch dwarf virus* isolates from faba bean (*Vicia faba* L.) in Yunnan Province, China. *Journal of Phytopathology* **158**, 35–9. doi: 10.1111/j.1439-0434.2009.01572.x.

La Cour T, Kiemer L, Mølgaard A, 2004. Analysis and prediction of leucine-rich nuclear export signals. *Protein Engineering, Design & Selection* **17**, 527–36. doi: 10.1093/protein/gzh062.

Laporte C, Vetter G, Loudes A-M *et al.*, 2003. Involvement of the secretory pathway and the cytoskeleton in intracellular targeting and tubule assembly of *Grapevine fanleaf virus* movement protein in tobacco BY-2 cells. *The Plant Cell Online* **15**, 2058–75. doi: 10.1105/tpc.013896.

Larkin MA, Blackshields G, Brown NP *et al.*, 2007. Clustal W and Clustal X version 2.0. *Bioinformatics (Oxford, England)* **23**, 2947–8. doi: 10.1093/bioinformatics/btm404.

Lartey RT, Lane LC, Melcher U, 1994. Electron microscopic and molecular characterization of turnip vein-clearing virus. *Archives of Virology* **138**, 287–98. doi: 10.1007/BF01379132.

Latham LJ, Jones RAC, 2001. Alfalfa mosaic and pea seed-borne mosaic viruses in cool season crop, annual pasture, and forage legumes: susceptibility, sensitivity, and seed transmission. *Australian Journal of Agricultural Research* **52**, 771. doi: 10.1071/AR00165.

Le Gall O, Candresse T, Dunez J, 1995. A multiple alignment of the capsid protein sequences of nepoviruses and comoviruses suggests a common structure. *Archives of Virology* **140**, 2041–53. doi: 10.1007/BF01322691.

Lecoq H, Wipf-Scheibel C, Verdin E, 2019. Characterization of the first tenuivirus naturally infecting dicotyledonous plants. *Archives of Virology* **164**, 297–301. doi: 10.1007/s00705-018-4057-6.

Lee L, Kaplan IB, Ripoll DR, 2005. A surface loop of the potato leafroll virus coat protein is involved in virion assembly, systemic movement, and aphid transmission. *Journal of Virology* **79**, 1207–14. doi: 10.1128/JVI.79.2.1207-1214.2005.

Leff JW, Lynch RC, Kane NC, 2017. Plant domestication and the assembly of bacterial and fungal communities associated with strains of the common sunflower, *Helianthus annuus*. *The New Phytologist* **214**, 412–23. doi: 10.1111/nph.14323.

Lesker T, Rabenstein F, Maiss E, 2013. Molecular characterization of five betacryptoviruses infecting four clover species and dill. *Archives of Virology* **158**, 1943–52. doi: 10.1007/s00705-013-1691-x.

Levene MJ, Korlach J, Turner SW, 2003. Zero-mode waveguides for single-molecule analysis at high concentrations. *Science (New York, N.Y.)* **299**, 682–6. doi: 10.1126/science.1079700.

Li C-X, Shi M, Tian J-H *et al.*, 2015. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *eLife* **4**. doi: 10.7554/eLife.05378.

Li H, Homer N, 2010. A survey of sequence alignment algorithms for next-generation sequencing. *Briefings in Bioinformatics* **11**, 473–83. doi: 10.1093/bib/bbq015.

Li H, Liu X, Liu X *et al.*, 2018. Host plant Infection by soybean mosaic virus reduces the fitness of its vector, *Aphis glycines (Hemiptera: Aphididae*). *Journal of Economic Entomology* **111**, 2017–23. doi: 10.1093/jee/toy165.

Li R, Gao S, Hernandez AG, 2012. Deep sequencing of small RNAs in tomato for virus and viroid identification and strain differentiation. *PLOS ONE* **7**, e37127. doi: 10.1371/journal.pone.0037127.

Li Y, Lu J, Han Y, 2013. RNA interference functions as an antiviral immunity mechanism in mammals. *Science (New York, N.Y.)* **342**, 231–4. doi: 10.1126/science.1241911.

Li Z, Yu M, Zhang H, 2005. Improved rapid amplification of cDNA ends (RACE) for mapping both the 5' and 3' terminal sequences of paramyxovirus genomes. *Journal of Virological Methods* **130**, 154–6. doi: 10.1016/j.jviromet.2005.06.022.

Lister RM, 1958. Some turnip viruses in Scotland and their effect on yield. *Plant Pathology* **7**, 144–6. doi: 10.1111/j.1365-3059.1958.tb00854.x.

Liu B, Preisser EL, Chu D *et al.*, 2013. Multiple forms of vector manipulation by a plant-infecting virus: *Bemisia tabaci* and tomato yellow leaf curl virus. *Journal of Virology* **87**, 4929–37. doi: 10.1128/JVI.03571-12.

Liu L, Li Y, Li S *et al.*, 2012. Comparison of next-generation sequencing systems. *Journal of Biomedicine & Biotechnology* **2012**, 251364. doi: 10.1155/2012/251364.

Liu X, Gorovsky MA, 1993. Mapping the 5’ and 3’ ends of *Tetrahymena thermophila* mRNAs using RNA ligase mediated amplification of cDNA ends (RLM-RACE). *Nucleic Acids Research* **21**, 4954–60. doi: 10.1093/nar/21.21.4954.

Liu Y, Du Z, Wang H, 2018. Identification and characterization of wheat yellow striate virus, a novel leafhopper-transmitted nucleorhabdovirus infecting wheat. *Frontiers in Microbiology* **9**, 468. doi: 10.3389/fmicb.2018.00468.

Lockhart BE, Swenson AS, Olszewski NE, 2008. Characterization of a strain of *Turnip vein-clearing virus* causing red ringspot of penstemon. *Plant Disease* **92**, 725–9. doi: 10.1094/PDIS-92-5-0725.

Longdon B, Obbard DJ, Jiggins FM, 2010. Sigma viruses from three species of *Drosophila* form a major new clade in the rhabdovirus phylogeny. *Proceedings of the Royal Society of London B: Biological Sciences* **277**, 35–44. doi: 10.1098/rspb.2009.1472.

Lu H, Giordano F, Ning Z, 2016. Oxford Nanopore MinION sequencing and genome assembly. *Genomics, Proteomics & Bioinformatics* **14**, 265–79. doi: 10.1016/j.gpb.2016.05.004.

Lu L, Wu S, Jiang J, 2018. Whole genome deep sequencing revealed host impact on population structure, variation and evolution of *Rice stripe virus*. *Virology* **524**, 32–44. doi: 10.1016/j.virol.2018.08.005.

MacDiarmid R, Rodoni B, Melcher U, 2013. Biosecurity implications of new technology and discovery in plant virus research. *PLoS pathogens* **9**, e1003337. doi: 10.1371/journal.ppat.1003337.

Makkouk KM, Kumari SG, van Leur JAG, 2014. Control of plant virus diseases in cool-season grain legume crops. *Advances in Virus Research* **90**, 207–53. doi: 10.1016/B978-0-12-801246-8.00004-4.

Makkouk KM, Rizkallah L, Madkour M *et al.*, 1994. Survey of faba bean (*Vicia faba* L.) for viruses in Egypt. *Phytopathologia Mediterranea* **33**, 207–11.

Makkouk KM, Vetten HJ, Katul L, 1998. Epidemiology and control of faba bean necrotic yellows virus. In: Hadidi A, Khetarpal RK, Koganezawa H, eds. *Plant virus disease control*, APS Press: St. Paul, Minn., 534–40.

Mar TB, Mendes IR, Lau D *et al.*, 2017. Interaction between the new world begomovirus *Euphorbia yellow mosaic virus* and its associated alphasatellite: effects on infection and transmission by the whitefly *Bemisia tabaci*. *The Journal of General Virology* **98**, 1552–62. doi: 10.1099/jgv.0.000814.

Marchler-Bauer A, Bo Y, Han L *et al.*, 2017. CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research* **45**, D200-D203. doi: 10.1093/nar/gkw1129.

Mardis ER, 2008. The impact of next-generation sequencing technology on genetics. *Trends in Genetics* **24**, 133–41. doi: 10.1016/j.tig.2007.12.007.

Maree HJ, Fox A, Al Rwahnih M, 2018. Application of HTS for routine plant virus diagnostics: state of the art and challenges. *Frontiers in Plant Science* **9**, 1082. doi: 10.3389/fpls.2018.01082.

Margulies M, Egholm M, Altman WE *et al.*, 2005. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* **437**, 376–80. doi: 10.1038/nature03959.

Mariano AC, Andrade MO, Santos AA *et al.*, 2004. Identification of a novel receptor-like protein kinase that interacts with a geminivirus nuclear shuttle protein. *Virology* **318**, 24–31. doi: 10.1016/j.virol.2003.09.038.

Markarian N, Li HW, Ding SW, 2004. RNA silencing as related to viroid induced symptom expression. *Archives of Virology* **149**, 397–406. doi: 10.1007/s00705-003-0215-5.

Marston HR, Quinlan-Watson F, Dewey DW, 1943. The dehydration of lucerne (*Medicago sativa*) and its potentialities as a concentrated source of ascorbic acid and of carotene for human

consumption. *Journal of the Council for Scientific and Industrial Research, Australia* **16**, 113–24.

Massart S, Candresse T, Gil J *et al.*, 2017. A framework for the evaluation of biosecurity, commercial, regulatory, and scientific impacts of plant viruses and viroids identified by NGS technologies. *Frontiers in Microbiology* **8**. doi: 10.3389/fmicb.2017.00045.

Massart S, Chiumenti M, Jonghe K de *et al.*, 2019. Virus detection by high-throughput sequencing of small RNAs: large-scale performance testing of sequence analysis strategies. *Phytopathology* **109**, 488–97. doi: 10.1094/PHYTO-02-18-0067-R.

Massart S, Olmos A, Jijakli H, 2014. Current impact and future directions of high throughput sequencing in plant virus diagnostics. *Virus Research* **188**, 90–6. doi: 10.1016/j.virusres.2014.03.029.

Mayo MA, 2002. Virology division news: ICTV at the Paris ICV: Results of the plenary session and the binomial ballot. *Archives of Virology* **147**, 2254–60. doi: 10.1007/s007050200052.

McKenzie DL, Morrall RAA, 1975. Diseases of specialty crops in Saskatchewan: II. Notes on field pea in 1973-74 and on lentil in 1973. *Canadian Plant Disease Survey* **55**, 97–100.

Mclean DL, Kinsey MG, 1964. A technique for electronically recording aphid feeding and salivation. *Nature* **202**, 1358–9. doi: 10.1038/2021358a0.

Mclean DL, Kinsey MG, 1965. Identification of electrically recorded curve patterns associated with aphid salivation and ingestion. *Nature* **205**, 1130–1. doi: 10.1038/2051130a0.

McLean AHC, van Asch M, Ferrari J, 2011. Effects of bacterial secondary symbionts on host plant use in pea aphids. *Proceedings: Biological Sciences* **278**, 760–6. doi: 10.1098/rspb.2010.1654.

Menzel W, Knierim D, Richert-Pöggeler KR, 2016. Characterization of a nucleorhabdovirus from *Physostegia*. *Julius-Kühn-Archiv* **454**, 283–4. doi: 10.5073/JKA.2016.454.000.

Metzker ML, 2010. Sequencing technologies - the next generation. *Nature reviews. Genetics* **11**, 31–46. doi: 10.1038/nrg2626.

Miller JR, Koren S, Sutton G, 2010. Assembly algorithms for next-generation sequencing data. *Genomics* **95**, 315–27. doi: 10.1016/j.ygeno.2010.03.001.

Milne RG, 1984. Electron microscopy for the identification of plant viruses in in vitro preparations. In: Maramorosch K, Koprowski H, eds. *Methods in Virology*, Academic Press: New York, London. Vol 7, 87–120.

Milne RG, Lesemann D-E, 1984. Immunosorbent electron microscopy in plant virus studies. In: Maramorosch K, Koprowski H, eds. *Methods in Virology*, Academic Press: New York, London. Vol 8, 85–101.

Mollov D, Maroon-Lango C, Kuniata L, 2016. Detection by next generation sequencing of a multi-segmented viral genome from sugarcane associated with Ramu stunt disease. *Virus Genes* **52**, 152–5. doi: 10.1007/s11262-015-1279-5.

Moran NA, Russell JA, Koga R, 2005. Evolutionary relationships of three new species of *Enterobacteriaceae* living as symbionts of aphids and other insects. *Applied and Environmental Microbiology* **71**, 3302–10. doi: 10.1128/AEM.71.6.3302-3310.2005.

Moran PJ, Thompson GA, 2001. Molecular responses to aphid feeding in Arabidopsis in relation to plant defense pathways. *Plant Physiology* **125**, 1074–85. doi: 10.1104/pp.125.2.1074.

Morris TJ, 1979. Isolation and analysis of double-stranded RNA from virus-infected plant and fungal tissue. *Phytopathology* **69**, 854. doi: 10.1094/Phyto-69-854.

Munson MA, Baumann P, Kinsey MG, 1991. *Buchnera* gen. nov. and *Buchnera aphidicola* sp. nov., a taxon consisting of the mycetocyte-associated, primary endosymbionts of aphids. *International Journal of Systematic Bacteriology* **41**, 566–8. doi: 10.1099/00207713-41-4-566.

Mushegian AR, 1994. The putative movement domain encoded by nepovirus RNA-2 is conserved in all sequenced nepoviruses. *Archives of Virology* **135**, 437–41. doi: 10.1007/BF01310028.

Musil M, 1966. Uber das Vorkommen des Virus des Blattrollens der Erbse in der Siowakei. (On the occurrence of a leaf roll virus in pea in Slovakia.). *Biologia, Bratislava* **21**, 133–8.

Myers SS, Smith MR, Guth S *et al.*, 2017. Climate change and global food systems: potential impacts on food security and undernutrition. *Annual Review of Public Health* **38**, 259–77.

Naidu RA, Hughes JDA, 2003. Methods for the detection of plant virus diseases. *Plant Virology in Sub Saharan Africa*, 233–53.

Nakabayashi H, Yamaji Y, Kagiwada S, 2002. The complete nucleotide sequence of a Japanese isolate of white clover mosaic virus strain RC. *Journal of General Plant Pathology* **68**, 173–6. doi: 10.1007/PL00013072.

Nakazono-Nagaoka E, Takahashi T, Shimizu T *et al.*, 2009. Cross-protection against bean yellow mosaic virus (BYMV) and clover yellow vein virus by Attenuated BYMV isolate M11. *Phytopathology* **99**, 251–7. doi: 10.1094/PHYTO-99-3-0251.

Nandakishor HV, Kumaraswamy B, Mane SS, 2017. Host range studies of soybean mosaic virus. *International Journal of Current Microbiology and Applied Sciences* **6**, 304–8. doi: 10.20546/ijcmas.2017.607.035.

Narzisi G, Mishra B, 2011. Comparing de novo genome assembly: the long and short of it. *PLOS ONE* **6**, e19175. doi: 10.1371/journal.pone.0019175.

Nault LR, 1997. Arthropod transmission of plant viruses. A new synthesis. *Annals of the Entomological Society of America* **90**, 521–41. doi: 10.1093/aesa/90.5.521.

Nawaz-Ul-Rehman MS, Nahid N, Mansoor S, 2010. Post-transcriptional gene silencing suppressor activity of two non-pathogenic alphasatellites associated with a begomovirus. *Virology* **405**, 300–8. doi: 10.1016/j.virol.2010.06.024.

Ng JCK, Falk BW, 2006. Virus-vector interactions mediating nonpersistent and semipersistent transmission of plant viruses. *Annual Review of Phytopathology* **44**, 183–212. doi: 10.1146/annurev.phyto.44.070505.143325.

Ng JCK, Perry KL, 2004. Transmission of plant viruses by aphid vectors. *Molecular Plant Pathology* **5**, 505–11. doi: 10.1111/j.1364-3703.2004.00240.x.

Nibert ML, Vong M, Fugate KK, 2018. Evidence for contemporary plant mitoviruses. *Virology* **518**, 14–24. doi: 10.1016/j.virol.2018.02.005.

Ohtaka C, 1991. Effects of heat treatment on the symbiotic system of an aphid mycetocyte. *Symbiosis* **11**, 19–30.

Okada R, Kiyota E, Moriyama H, 2015. A simple and rapid method to purify viral dsRNA from plant and fungal tissue. *Journal of General Plant Pathology* **81**, 103–7. doi: 10.1007/s10327-014-0575-6.

Økland AL, Skoge RH, Nylund A, 2018. The complete genome sequence of CrRV-Ch01, a new member of the family *Rhabdoviridae* in the parasitic copepod *Caligus rogercresseyi* present on farmed Atlantic salmon (*Salmo salar*) in Chile. *Archives of Virology*, 1657–61. doi: 10.1007/s00705-018-3768-z.

Oliveros JC, 2015. Venny. An interactive tool for comparing lists with Venn's diagrams. [https://bioinfogp.cnb.csic.es/tools/venny/index.html].

Olmos A, Boonham N, Candresse T *et al.*, 2018. High-throughput sequencing technologies for plant pest diagnosis: challenges and opportunities. *EPPO Bulletin* **48**, 219–24. doi: 10.1111/epp.12472.

Olson J, Rebek E, Schnelle M, 2017. Rose rosette disease. [http://factsheets.okstate.edu/documents/epp-7329-rose-rosette-disease/]. Accessed 18 October 2019.

Ortiz V, Navarro E, Castro S, 2006. Incidence and transmission of faba bean necrotic yellows virus (FBNYV) in Spain. *Spanish Journal of Agricultural Research* **4**, 255. doi: 10.5424/sjar/2006043-200.

Padmanabhan C, Zheng Y, Li R, 2015. Complete Genome Sequence of Southern tomato virus Naturally Infecting Tomatoes in Bangladesh. *Genome Announcements* **3**. doi: 10.1128/genomeA.01522-15.

Palukaitis P, Roossinck MJ, Dietzgen RG, 1992. Cucumber mosaic virus. In: Maramorosch K, Murphy FA, Shatkin Aaron J., eds. *Advances in Virus Research*, Elsevier. Vol 41, 281–348.

Pandey B, Naidu RA, Grove GG, 2018. Detection and analysis of mycovirus-related RNA viruses from grape powdery mildew fungus *Erysiphe necator*. *Archives of Virology* **163**, 1019–30. doi: 10.1007/s00705-018-3714-0.

Paprotka T, Metzler V, Jeske H, 2010. The first DNA 1-like alpha satellites in association with New World begomoviruses in natural infections. *Virology* **404**, 148–57. doi: 10.1016/j.virol.2010.05.003.

Park CY, Lee S-H, Lim S, 2017. First report of white clover mosaic virus on white clover (*Trifolium repens*) in Korea. *Plant Disease* **101**, 1559. doi: 10.1094/PDIS-02-17-0256-PDN.

Patil BL, Kumar PL, 2015. Pigeonpea sterility mosaic virus: a legume-infecting Emaravirus from South Asia. *Molecular Plant Pathology* **16**, 775–86. doi: 10.1111/mpp.12238.

Patil BL, Legg JP, Kanju E, 2015. Cassava brown streak disease: a threat to food security in Africa. *The Journal of General Virology* **96**, 956–68. doi: 10.1099/vir.0.000014.

Pechinger K, Chooi KM, MacDiarmid RM, 2019. A new era for mild strain cross-protection. *Viruses* **11**. doi: 10.3390/v11070670.

Pecman A, Kutnjak D, Gutiérrez-Aguirre I *et al.*, 2017. Next generation sequencing for detection and discovery of plant viruses and viroids: comparison of two approaches. *Frontiers in Microbiology* **8**, 1998. doi: 10.3389/fmicb.2017.01998.

Peoples MB, Bowman AM, Gault RR *et al.*, 2001. Factors regulating the contributions of fixed nitrogen by pasture and crop legumes to different farming systems of eastern Australia. *Plant and Soil* **228**, 29–41. doi: 10.1023/A:1004799703040.

Pescod KV, Quick WP, Douglas AE, 2007. Aphid responses to plants with genetically manipulated phloem nutrient levels. *Physiological Entomology* **32**, 253–8. doi: 10.1111/j.1365-3032.2007.00577.x.

Peter KA, Liang D, Palukaitis P, 2008. Small deletions in the potato leafroll virus readthrough protein affect particle morphology, aphid transmission, virus movement and accumulation. *The Journal of General Virology* **89**, 2037–45. doi: 10.1099/vir.0.83625-0.

Phibbs A, Barta A, Domier LL, 2004. First report of *Soybean dwarf virus* on soybean in Wisconsin. *Plant Disease* **88**, 1285. doi: 10.1094/PDIS.2004.88.11.1285A.

Prado E, Tjallingii WF, 1994. Aphid activities during sieve element punctures. *Entomologia Experimentalis et Applicata* **72**, 157–65. doi: 10.1111/j.1570-7458.1994.tb01813.x.

Purdy LH, 1979. *Sclerotinia sclerotiorum*: History, diseases and symptomatology, host range, geographic distribution, and impact. *Phytopathology* **69**, 875. doi: 10.1094/Phyto-69-875.

Quantz L, Volk J, 1954. Die blattrollkrankheit der ackerbohne und erbse, eine neue viruskrankheit bei leguminosen. *Nachrichtenblatt des Deutschen Pflanzenschutzdienstes* **6**, 177–82.

R Core Team, 2014. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, Vienna, Austria.

R Core Team, 2019. *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, Vienna, Austria.

Rebenstorf K, 2005. *Untersuchungen zur Epidemiologie des Cherry leaf roll virus (CLRV) – genetische und serologische Diversität in Abhängigkeit von der Wirtspflanzenart und der geographischen Herkunft*. Humboldt-Universität zu Berlin. PhD thesis.

Reddy MV, Nene YL, Verma JP, 1979. Pea leaf roll virus causes chickpea stunt. *International Chickpea Newsletter* **1**, 8.

Reuter JA, Spacek DV, Snyder MP, 2015. High-throughput sequencing technologies. *Molecular Cell* **58**, 586–97. doi: 10.1016/j.molcel.2015.05.004.

Richert-Pöggeler KR, Franzke K, Hipp K, 2018. Electron microscopy methods for virus diagnosis and high resolution analysis of viruses. *Frontiers in Microbiology* **9**, 3255. doi: 10.3389/fmicb.2018.03255.

Rizos H, Gunn LV, Pares RD, 1992. Differentiation of cucumber mosaic virus isolates using the polymerase chain reaction. *Journal of General Virology* **73 (Pt 8)**, 2099–103. doi: 10.1099/0022-1317-73-8-2099.

Roberts IM, Harrison BD, 1979. Detection of potato leafroll and potato mop-top viruses by immunosorbent electron microscopy. *Annals of Applied Biology* **93**, 289–97. doi: 10.1111/j.1744-7348.1979.tb06544.x.

Roberts JMK, Ireland KB, Tay WT, 2018. Honey bee-assisted surveillance for early plant virus detection. *Annals of Applied Biology* **173**, 285–93. doi: 10.1111/aab.12461.

Rochon D, Lommel S, Martelli GP, 2012. Family *Tombusviridae*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, eds. *Virus taxonomy. Classification and nomenclature of viruses. Ninth report of the International Committee on Taxonomy of Viruses*, Elsevier academic press: London, Waltham MA, 1111–38.

Roggero P, Ciuffo M, Vaira AM, 2000. An *Ophiovirus* isolated from lettuce with big-vein symptoms. *Archives of Virology* **145**, 2629–42. doi: 10.1007/s007050070012.

Romanovskaya A, Sarin LP, Bamford DH, 2013. High-throughput purification of double-stranded RNA molecules using convective interaction media monolithic anion exchange columns. *Journal of Chromatography. A* **1278**, 54–60. doi: 10.1016/j.chroma.2012.12.050.

Roossinck MJ, 2011. The good viruses: viral mutualistic symbioses. *Nature reviews. Microbiology* **9**, 99–108. doi: 10.1038/nrmicro2491.

Roossinck MJ, 2017. Deep sequencing for discovery and evolutionary analysis of plant viruses. *Virus Research* **239**, 82–6. doi: 10.1016/j.virusres.2016.11.019.

Roossinck MJ, Martin DP, Roumagnac P, 2015. Plant virus metagenomics: advances in virus discovery. *Phytopathology* **105**, 716–27. doi: 10.1094/PHYTO-12-14-0356-RVW.

Roossinck MJ, Saha P, Wiley GB *et al.*, 2010. Ecogenomics. Using massively parallel pyrosequencing to understand virus ecology. *Molecular Ecology* **19 Suppl 1**, 81–8. doi: 10.1111/j.1365-294X.2009.04470.x.

Roossinck MJ, Sleat D, Palukaitis P, 1992. Satellite RNAs of plant viruses: structures and biological effects. *Microbiology and Molecular Biology Reviews* **56**, 265–79.

Ross JP, 1977. Effect of aphid-transmitted soybean mosaic virus on yields of closely related resistant and susceptible soybean lines. *Crop Science* **17**, 869. doi: 10.2135/cropsci1977.0011183X001700060014x.

Rothberg JM, Hinz W, Rearick TM *et al.*, 2011. An integrated semiconductor device enabling non-optical genome sequencing. *Nature* **475**, 348–52. doi: 10.1038/nature10242.

Rott M, Xiang Y, Boyes I *et al.*, 2017. Application of next generation sequencing for diagnostic testing of tree fruit viruses and viroids. *Plant Disease* **101**, 1489–99. doi: 10.1094/PDIS-03-17-0306-RE.

Rott ME, Kesanakurti P, Berwarth C *et al.*, 2018. Discovery of negative-sense RNA viruses in trees infected with apple rubbery wood disease by next-generation sequencing. *Plant Disease* **102**, 1254–63. doi: 10.1094/PDIS-06-17-0851-RE.

Rozado-Aguirre Z, Adams I, Collins L, 2016. Detection and transmission of carrot torrado virus, a novel putative member of the *Torradovirus* genus. *Journal of Virological Methods* **235**, 119–24. doi: 10.1016/j.jviromet.2016.05.018.

Russell AC, 1994. Three new pulse cultivars for New Zealand's arable industry. In: *Proceedings Agronomy Society of N.Z.*, 125–8.

Sabanadzovic S, Valverde RA, Brown JK, 2009. Southern tomato virus: The link between the families Totiviridae and Partitiviridae. *Virus Research* **140**, 130–7. doi: 10.1016/j.virusres.2008.11.018.

Samac DA, Rhodes LH, Lamp WO (eds), 2016. *Compendium of alfalfa diseases and pests*. St. Paul, Minnesota, USA: The American Phytopathological Society.

Sanfaçon H, Wellink J, Le Gall O, 2009. Secoviridae: a proposed family of plant viruses within the order *Picornavirales* that combines the families *Sequiviridae* and *Comoviridae*, the unassigned genera *Cheravirus* and *Sadwavirus*, and the proposed genus Torradovirus. *Archives of Virology* **154**, 899–907. doi: 10.1007/s00705-009-0367-z.

Sanger F, Nicklen S, Coulson AR, 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–7. doi: 10.1073/pnas.74.12.5463.

Sanger M, Passmore B, Falk BW, 1994. Symptom severity of beet western yellows virus strain ST9 is conferred by the ST9-associated RNA and is not associated with virus release from the phloem. *Virology* **200**, 48–55. doi: 10.1006/viro.1994.1161.

Sanjuán R, 2010. Mutational fitness effects in RNA and single-stranded DNA viruses: common patterns revealed by site-directed mutagenesis studies. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **365**, 1975–82. doi: 10.1098/rstb.2010.0063.

Sano Y, Wada M, Hashimoto Y, 1998. Sequences of ten circular ssDNA components associated with the milk vetch dwarf virus genome. *The Journal of General Virology* **79 (Pt 12)**, 3111–8. doi: 10.1099/0022-1317-79-12-3111.

Santos AA, Carvalho CM, Florentino LH, 2009. Conserved threonine residues within the A-loop of the receptor NIK differentially regulate the kinase function required for antiviral signaling. *PLOS ONE* **4**, e5781. doi: 10.1371/journal.pone.0005781.

Santos AA, Lopes KVG, Apfata JAC, 2010. NSP-interacting kinase, NIK: a transducer of plant defence signalling. *Journal of experimental botany* **61**, 3839–45. doi: 10.1093/jxb/erq219.

Saucke H, Uteau D, Brinkmann K, 2019. Symptomology and yield impact of pea necrotic yellow dwarf virus (PNYDV) in faba bean (*Vicia faba* L. minor). *European Journal of Plant Pathology* **153**, 1299–315. doi: 10.1007/s10658-018-01643-5.

Sauge M-H, Lacroze J-P, Poessel J-L, 2002. Induced resistance by *Myzus persicae* in the peach cultivar 'Rubira'. *Entomologia Experimentalis et Applicata* **102**, 29–37. doi: 10.1046/j.1570-7458.2002.00922.x.

Schliephake E, Habekuss A, Scholz M, 2013. Barley yellow dwarf virus transmission and feeding behaviour of *Rhopalosiphum padi* on *Hordeum bulbosum* clones. *Entomologia Experimentalis et Applicata* **146**, 347–56. doi: 10.1111/eea.12033.

Schmidt HE, 1981. The diagnosis of viruses of legume crops in German Democratic Republic as a prerequisite for breeding broad bean (*Vicia faba* L.), bean (*Phaseolus vulgaris* L.), pea (*Pisum*

sp.) and lupin (*Lupinus sp*.) for resistance to virus diseases. In: *Principles of plant resistance to diseases and pests*. Research Institute Plant Protection, 25–33.

Schwarzkopf A, Rosenberger D, Niebergall M, 2013. To feed or not to feed: plant factors located in the epidermis, mesophyll, and sieve elements influence pea aphid's ability to feed on legume species. *PLOS ONE* **8**, e75298. doi: 10.1371/journal.pone.0075298.

Seifers DL, Harvey TL, Martin TJ *et al.*, 2005. Association of a virus with wheat displaying yellow head disease symptoms in the great plains. *Plant Disease* **89**, 888–95. doi: 10.1094/PD-89-0888.

Shahjahan RM, Hughes KJ, Leopold RA, 1995. Lower incubation temperature increases yield of insect genomic DNA isolated by the CTAB method. *BioTechniques* **19**, 332–4.

Shi M, Lin X-D, Tian J-H *et al.*, 2016a. Redefining the invertebrate RNA virosphere. *Nature* **540**, 539–43. doi: 10.1038/nature20167.

Shi M, Lin X-D, Vasilakis N *et al.*, 2016b. Divergent viruses discovered in arthropods and vertebrates revise the evolutionary history of the *Flaviviridae* and related viruses. *Journal of Virology* **90**, 659–69. doi: 10.1128/JVI.02036-15.

Shirako Y, Falk BW, Haenni AL, 2012. Genus *Tenuivirus*. In: King AMQ, Adams MJ, Carstens EB, Lefkowitz EJ, eds. *Virus taxonomy: ninth report of the international committee on taxonomy of viruses*, Elsevier Academic Press San Diego, 771–6.

Sicard A, Yvon M, Timchenko T *et al.*, 2013. Gene copy number is differentially regulated in a multipartite virus. *Nature Communications* **4**, 2248. doi: 10.1038/ncomms3248.

Sicard A, Zeddam J-L, Yvon M, 2015. Circulative nonpropagative aphid transmission of nanoviruses: an oversimplified view. *Journal of Virology* **89**, 9719–26. doi: 10.1128/JVI.00780-15.

Simons JN, 1953. *Vector-virus relationships of pea-enation mosaic and the pea aphid Macrosiphum pisi (Kalt.)*. University of California, Berkeley.

Srinivasan R, Alvarez JM, 2007. Effect of Mixed Viral Infections (Potato Virus Y–Potato Leafroll Virus) on Biology and Preference of Vectors Myzus persicae and Macrosiphum euphorbiae (Hemiptera: Aphididae). *Journal of Economic Entomology* **100**, 646–55. doi: 10.1093/jee/100.3.646.

Stevens M, Lacomme C, 2017. Transmission of plant viruses. In: van Emden HF, Harrington R, eds. *Aphids as crop pests*, CABI: Wallingford, 323–61.

Stevens M, Smith HG, Hallsworth PB, 1994. The host range of beet yellowing viruses among common arable weed species. *Plant Pathology* **43**, 579–88. doi: 10.1111/j.1365-3059.1994.tb01593.x.

Syller J, 2012. Facilitative and antagonistic interactions between plant viruses in mixed infections. *Molecular Plant Pathology* **13**, 204–16. doi: 10.1111/j.1364-3703.2011.00734.x.

Sylvester ES, 1956. Beet yellows virus transmission by the green peach aphid. *Journal of Economic Entomology* **49**, 789–800. doi: 10.1093/jee/49.6.789.

Sylvester ES, Richardson J, 1992. Aphid-borne rhabdoviruses—relationships with their vectors. In: Harris KF, ed. *Advances in disease vector research*, Springer New York: New York, NY. Vol 9, 313–41.

Tamada T, Goto T, Chiba I, 1969. Soybean dwarf, a new virus disease. *Japanese Journal of Phytopathology* **35**, 282–5. doi: 10.3186/jjphytopath.35.282.

Tao X, Zhou X, Li G, 2002. The pathogenicity on legumes of *Cucumber mosaic virus* was determined by 243 nucleotides on 2a polymerase gene of viral RNA2. *Chinese Science Bulletin* **47**, 748. doi: 10.1360/02tb9169.

Thompson GA, Goggin FL, 2006. Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. *Journal of experimental botany* **57**, 755–66. doi: 10.1093/jxb/erj135.

Thompson JR, Dasgupta I, Fuchs M *et al.*, 2017. ICTV virus taxonomy profile: *Secoviridae*. *The Journal of General Virology* **98**, 529–31. doi: 10.1099/jgv.0.000779.

Timchenko T, Katul L, Aronson M *et al.*, 2006. Infectivity of nanovirus DNAs: induction of disease by cloned genome components of *Faba bean necrotic yellows virus*. *The Journal of General Virology* **87**, 1735–43. doi: 10.1099/vir.0.81753-0.

Timchenko T, Katul L, Sano Y, 2000. The master rep concept in nanovirus replication: identification of missing genome components and potential for natural genetic reassortment. *Virology* **274**, 189–95. doi: 10.1006/viro.2000.0439.

Timchenko T, Kouchkovsky F de, Katul L, 1999. A single Rep protein initiates replication of multiple genome components of faba bean necrotic yellows virus, a single-stranded DNA virus of plants. *Journal of Virology* **73**, 10173–82.

Tjallingii WF, 1978. Electronic recording of penetration behaviour by aphids. *Entomologia Experimentalis et Applicata* **24**, 721–30. doi: 10.1111/j.1570-7458.1978.tb02836.x.

Tjallingii WF, 1985. Electrical nature of recorded signals during stylet penetration by aphids. *Entomologia Experimentalis et Applicata* **38**, 177–86. doi: 10.1111/j.1570-7458.1985.tb03516.x.

Tjallingii WF, 1988. Electrical recording of stylet penetration activities. In: Minks AK, Harrewijn P, Helle W, eds. *Aphids, their biology, natural enemies and control*, Elsevier Science Publishers, 95–108.

Tjallingii WF, 1994. Sieve element acceptance by aphids. *European Journal of Entomology* **91**, 47–52.

Tomlinson JA, Walkey DGA, 1967. The isolation and identification of rhubarb viruses occurring in Britain. *Annals of Applied Biology* **59**, 415–27. doi: 10.1111/j.1744-7348.1967.tb04458.x.

Torrance L, Jones RAC, 1981. Recent developments in serological methods suited for use in routine testing for plant viruses. *Plant Pathology* **30**, 1–24. doi: 10.1111/j.1365-3059.1981.tb01218.x.

Tsuchida T, Koga R, Shibao H, 2002. Diversity and geographic distribution of secondary endosymbiotic bacteria in natural populations of the pea aphid, *Acyrthosiphon pisum*. *Molecular Ecology* **11**, 2123–35. doi: 10.1046/j.1365-294X.2002.01606.x.

Tzanetakis IE, Martin RR, 2008. A new method for extraction of double-stranded RNA from plants. *Journal of Virological Methods* **149**, 167–70. doi: 10.1016/j.jviromet.2008.01.014.

Untergasser A, Cutcutache I, Koressaar T *et al.*, 2012. Primer3--new capabilities and interfaces. *Nucleic Acids Research* **40**, e115. doi: 10.1093/nar/gks596.

Uzest M, Gargani D, Drucker M *et al.*, 2007. A protein key to plant virus transmission at the tip of the insect vector stylet. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 17959–64. doi: 10.1073/pnas.0706608104.

Vainio EJ, Chiba S, Ghabrial SA *et al.*, 2018. ICTV virus taxonomy profile: *Partitiviridae*. *The Journal of General Virology* **99**, 17–8. doi: 10.1099/jgv.0.000985.

van den Heuvel JF, Verbeek M, van der Wilk F, 1994. Endosymbiotic bacteria associated with circulative transmission of potato leafroll virus by *Myzus persicae*. *The Journal of General Virology* **75 (Pt 10)**, 2559–65. doi: 10.1099/0022-1317-75-10-2559.

van den J. M. Heuvel JF, 1991. Transmission of potato leafroll virus from plants and artificial diets by *Myzus persicae*. *Phytopathology* **81**, 150. doi: 10.1094/Phyto-81-150.

van Dijk EL, Auger H, Jaszczyszyn Y, 2014a. Ten years of next-generation sequencing technology. *Trends in Genetics* **30**, 418–26. doi: 10.1016/j.tig.2014.07.001.

van Dijk EL, Jaszczyszyn Y, Thermes C, 2014b. Library preparation methods for next-generation sequencing: tone down the bias. *Experimental Cell Research* **322**, 12–20. doi: 10.1016/j.yexcr.2014.01.008.

van Leur JAG, Kumari SG, 2011. A survey of lucerne in northern New South Wales for viruses of importance to the winter legume industry. *Australasian Plant Pathology* **40**, 180–6. doi: 10.1007/s13313-011-0028-z.

van Leur JAG, Kumari SG, Aftab M, 2013. Virus resistance of Australian pea (*Pisum sativum*) varieties. *New Zealand Journal of Crop and Horticultural Science* **41**, 86–101. doi: 10.1080/01140671.2013.781039.

Vemulapati B, Druffel KL, Eigenbrode SD, 2010. Molecular characterization of pea enation mosaic virus and bean leafroll virus from the Pacific Northwest, USA. *Archives of Virology* **155**, 1713–5. doi: 10.1007/s00705-010-0767-0.

Verbeek M, Dullemans A, van den Heuvel H, 2010. Tomato chocolàte virus: a new plant virus infecting tomato and a proposed member of the genus *Torradovirus*. *Archives of Virology* **155**, 751–5. doi: 10.1007/s00705-010-0640-1.

Verbeek M, Dullemans AM, van Bekkum PJ, 2013. Evidence for lettuce big-vein associated virus as the causal agent of a syndrome of necrotic rings and spots in lettuce. *Plant Pathology* **62**, 444–51. doi: 10.1111/j.1365-3059.2012.02645.x.

Verbeek M, Dullemans AM, van den Heuvel JFJM, 2007. Identification and characterisation of tomato torrado virus, a new plant picorna-like virus from tomato. *Archives of Virology* **152**, 881–90. doi: 10.1007/s00705-006-0917-6.

Verbeek M, Dullemans AM, van der Vlugt RAA, 2017. Aphid transmission of *Lettuce necrotic leaf curl virus*, a member of a tentative new subgroup within the genus *Torradovirus*. *Virus Research* **241**, 125–30. doi: 10.1016/j.virusres.2017.02.008.

Vetten HJ, Dale JL, Grigoras I, 2011. Family: *Nanoviridae*. In: King AMQ, Lefkowitz E, Adams MJ, Carstens EB, eds. *Virus taxonomy: ninth report of the International Committee on Taxonomy of Viruses*, Elsevier, 395–404.

Vetten HJ, Gronenborn B, Bressan A, 2016. Aphid transmission of viruses of the family *Nanoviridae*. In: Brown JK, ed. *Vector-mediated transmission of plant pathogens*, APS Press: St. Paul Minnesota, 453–62.

Vetten HJ, Knierim D, Rakoski MS *et al.*, 2019. Identification of a novel nanovirus in parsley. *Archives of Virology* **164**, 1883–7. doi: 10.1007/s00705-019-04280-3.

Villamor DEV, Ho T, Rwahnih M al, 2019. High throughput sequencing for plant virus detection and discovery. *Phytopathology* **109**, 716–25. doi: 10.1094/PHYTO-07-18-0257-RVW.

Villamor DEV, Pillai SS, Eastwell KC, 2017. High throughput sequencing reveals a novel fabavirus infecting sweet cherry. *Archives of Virology* **162**, 811–6. doi: 10.1007/s00705-016-3141-z.

Visser M, Bester R, Burger JT, 2016. Next-generation sequencing for virus detection: covering all the bases. *Virology Journal* **13**, 85. doi: 10.1186/s12985-016-0539-x.

Vodovar N, Goic B, Blanc H, 2011. *In silico* reconstruction of viral genomes from small RNAs improves virus-derived small interfering RNA profiling. *Journal of Virology* **85**, 11016–21. doi: 10.1128/JVI.05647-11.

Vunsh RON, Rosner A, Stein A, 1990. The use of the polymerase chain reaction (PCR) for the detection of bean yellow mosaic virus in gladiolus. *Annals of Applied Biology* **117**, 561–9. doi: 10.1111/j.1744-7348.1990.tb04822.x.

Walker PJ, Blasdell KR, Calisher CH *et al.*, 2018. ICTV virus taxonomy profile: *Rhabdoviridae*. *The Journal of General Virology* **99**, 447–8. doi: 10.1099/jgv.0.001020.

Walker PJ, Dietzgen RG, Joubert DA, 2011. Rhabdovirus accessory genes. *Virus Research* **162**, 110–25. doi: 10.1016/j.virusres.2011.09.004.

Walker PJ, Firth C, Widen SG *et al.*, 2015. Evolution of genome size and complexity in the *Rhabdoviridae*. *PLOS Pathogens* **11**, e1004664. doi: 10.1371/journal.ppat.1004664.

Wanitchakorn R, Hafner GJ, Harding RM, 2000. Functional analysis of proteins encoded by banana bunchy top virus DNA-4 to -6. *The Journal of General Virology* **81**, 299–306. doi: 10.1099/0022-1317-81-1-299.

Ward E, Foster SJ, Fraaije BA, 2004. Plant pathogen diagnostics: immunological and nucleic acid-based approaches. *Annals of Applied Biology* **145**, 1–16. doi: 10.1111/j.1744-7348.2004.tb00354.x.

Watanabe S, Borthakur D, Bressan A, 2013. Lack of evidence for an interaction between *Buchnera* GroEL and banana bunchy top virus (*Nanoviridae*). *Virus Research* **177**, 98–102. doi: 10.1016/j.virusres.2013.06.002.

Watanabe S, Borthakur D, Bressan A, 2016. Localization of *Banana bunchy top virus* and cellular compartments in gut and salivary gland tissues of the aphid vector *Pentalonia nigronervosa*. *Insect Science* **23**, 591–602. doi: 10.1111/1744-7917.12211.

Watanabe S, Bressan A, 2013. Tropism, compartmentalization and retention of banana bunchy top virus (*Nanoviridae*) in the aphid vector *Pentalonia nigronervosa*. *The Journal of General Virology* **94**, 209–19. doi: 10.1099/vir.0.047308-0.

Watson MA, Roberts FM, 1939. A comparative study of the transmission of hyoscyamus virus 3, potato virus Y and cucumber virus 1 by the vectors *Myzus persicae* (Sulz), *M. circumflexus* (Buckton), and *Macrosiphum gei* (Koch). *Proceedings of the Royal Society of London. Series B - Biological Sciences* **127**, 543–76. doi: 10.1098/rspb.1939.0039.

Webster CG, Wylie SJ, Jones MGK, 2004. Diagnosis of plant viral pathogens. *Current Science* **86**, 1604–7.

Wei T-Y, Yang J-G, Liao F-R *et al.*, 2009. Genetic diversity and population structure of rice stripe virus in China. *The Journal of General Virology* **90**, 1025–34. doi: 10.1099/vir.0.006858-0.

Wilkinson TL, Douglas AE, 1998. Host cell allometry and regulation of the symbiosis between pea aphids, *Acyrthosiphon pisum*, and bacteria, *Buchnera*. *Journal of Insect Physiology* **44**, 629–35. doi: 10.1016/S0022-1910(98)00030-4.

Wool D, Hales DF, 1996. Previous infestation affects recolonization of cotton by *Aphis gossypii*: induced resistance or plant damage? *Phytoparasitica* **24**, 39–48. doi: 10.1007/BF02981452.

Wu L-P, Yang T, Liu H-W, 2018. Molecular characterization of a novel rhabdovirus infecting blackcurrant identified by high-throughput sequencing. *Archives of Virology*, 1363–6. doi: 10.1007/s00705-018-3709-x.

Wu MD, Zhang L, Li GQ, 2007. Hypovirulence and double-stranded RNA in *Botrytis cinerea*. *Phytopathology* **97**, 1590–9. doi: 10.1094/PHYTO-97-12-1590.

Wu Q, Ding S-W, Zhang Y, 2015. Identification of viruses and viroids by next-generation sequencing and homology-dependent and homology-independent algorithms. *Annual Review of Phytopathology* **53**, 425–44. doi: 10.1146/annurev-phyto-080614-120030.

Wyant PS, Strohmeier S, Schäfer B *et al.*, 2012. Circular DNA genomics (circomics) exemplified for geminiviruses in bean crops and weeds of northeastern Brazil. *Virology* **427**, 151–7. doi: 10.1016/j.virol.2012.02.007.

Xie J, Ghabrial SA, 2012. Molecular characterization of two mitoviruses co-infecting a hypovirulent isolate of the plant pathogenic fungus *Sclerotinia sclerotiorum* corrected. *Virology* **428**, 77–85. doi: 10.1016/j.virol.2012.03.015.

Yanagisawa H, Tomita R, Katsu K *et al.*, 2016. Combined DECS analysis and next-generation sequencing enable efficient detection of novel plant RNA viruses. *Viruses* **8**, 70. doi: 10.3390/v8030070.

Zerbino DR, Birney E, 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18**, 821–9. doi: 10.1101/gr.074492.107.

Zhao K, Margaria P, Rosa C, 2016. First report of white clover mosaic virus and turnip mosaic virus mixed infection on garlic mustard in Pennsylvania. *Plant Disease* **100**, 866. doi: 10.1094/PDIS-09-15-1083-PDN.

Zheng Y, Gao S, Padmanabhan C *et al.*, 2017. VirusDetect: an automated pipeline for efficient virus discovery using deep sequencing of small RNAs. *Virology* **500**, 130–8. doi: 10.1016/j.virol.2016.10.017.

Zhou GH, Rochow WF, 1984. Differences among five stages of *Schizaphis graminum* in transmission of a barley yellow dwarf luteovirus. *Phytopathology* **74**, 1450–3.

Ziebell H, 2015. Kleines Virus - großer Schaden. *ForschungsReport spezial Ökologischer Landbau* **4**, 6–7.

Ziebell H, 2017. Die Virusepidemie an Leguminosen 2016 - eine Folge des Klimawandels? *Journal für Kulturpflanzen* **69**, 64–8. doi: 10.1399/JfK.2017.02.09.

Ziebell H, Carr JP, 2010. Cross-protection. In: Carr JP, Loebenstein G, eds. *Natural and Engineered Resistance to Plant Viruses, Part II*, Elsevier. Vol 76, 211–64.

Ziebell H, Murphy AM, Groen SC *et al.*, 2011. Cucumber mosaic virus and its 2b RNA silencing suppressor modify plant-aphid interactions in tobacco. *Scientific Reports* **1**, 187. doi: 10.1038/srep00187.

Acknowledgments

“Alhamdulillah” All praise is due to Allah.

Allah said:

بِسْمِ اللَّهِ الرَّحْمَـٰنِ الرَّحِيمِ

لَا تَحْسَبَنَّ الَّذِينَ يَفْرَحُونَ بِمَا أَتَوا وَّيُحِبُّونَ أَن يُحْمَدُوا بِمَا لَمْ يَفْعَلُوا فَلَا تَحْسَبَنَّهُم بِمَفَازَةٍ مِّنَ الْعَذَابِ ۖ وَلَهُمْ عَذَابٌ أَلِيمٌ (188)

سورة آل عمران

In the name of Allah, the Entirely Merciful, the Especially Merciful

And never think that those who rejoice in what they have perpetrated and like to be praised for what they did not do - never think them [to be] in safety from the punishment, and for them is a painful punishment (188)

(Sura Aal-i-Imraan)

Moreover, following the guidance of prophet Muhammad (peace and blessings of Allah be upon him) who said, لاَ يَشْكُرُ اللَّهَ مَنْ لاَ يَشْكُرُ النَّاسَ “He has not thanked Allah who has not thanked the people.” (reported by Abu Huraira, source Sunan Abī Dāwūd: 4811 (authentic)). Thus, this part of the thesis is very important as every person mentioned here for me.

Although the PhD is an individual work, it is not indefinitely on your own. Many people contributed significantly to this work. Special thanks to Dr. Heiko Ziebell, Prof. Dr. Stefan Vidal and Prof. Dr. Edgar Maiss for accepting me as a PhD student and for their confidence in me. Thank you for your mentoring, guidance and patience. Dr. Heiko, you supported and stimulated me, and gave me many opportunities to improve my skills and knowledge. You are a great supervisor who stands beside his student and fully support him. To Prof. Dr. Stefan Vidal for your positive critics and valuable discussions during my work and all support specially the administrative work. I would like to thank Prof. Dr. Edgar Maiss for his advice and his critical reading for the thesis. I would like to thank Prof. Michael Rostás for accepting to be in the examination committee.

I am also grateful to Angelika Sieg-Müller, Petra Lüddecke, Kerstin Herz and Jonas Hartrick for their excellent technical assistance and for the great working atmosphere during my research period. Special thanks to all the co-author of the different publications. Many thanks to Prof. Dr. Edger Schliephake who taught me the EPG method, Dr. Katja Richert-Pöggeler for teaching me the EM for virus detection and for the valuable discussions, and to Prof. Dr. Stefan Winter and Dr. Dennis Knierim for training me to the ribodepletion approach for HTS virus detection. To Prof. Thomas Kühne and Dr. Wolfgang Maier for the valuable discussions.

My friend Amjad Zia, sincerely you are a great support for me on personal level during master's degree and specially during my PhD studies. The PhD period was full of stressful moments, and our friendly meetings were a relief "second after prayers". I cannot forget the valuable comments and suggestions to improve my work. Thank you, brother.

My father Dr. Zakaria Abdou Gaafar and my mother, thank you for raising me and giving me all the opportunities to be successful. My brothers, sisters and brother in law for your emotional support and your prayers. My grandmother (may Allah surround you with his mercy), you mean a lot to me, sorry I was not there beside you. See you in a better place my dear.

I was supported financially by a German Egyptian long-term scholarship (GERLS), thanks to all the staff involved in this scholarship. I was also partially supported by scholarships from the Gemeinschaft der Förderer und Freunde des Julius Kühn-Instituts e. V. (GFF) and Stegemann Stiftung, thank you very much for your financial support.

This research was funded by several agencies: a grant from the Federal Office of Food and Agriculture within the Euphresco network "2015-F-172: The application of Next-Generation Sequencing technology for the detection and diagnosis of non-culturable organisms: viruses and viroids"; a German-New Zealand cooperation grant from the German Federal Ministry of Food and Agriculture (BMEL) through the Federal Office for Agriculture and Food (BLE), Germany and The Royal Society of New Zealand. Additional support was from the GFF and Stegemann Stiftung.

Thank you all, I wish you all the best

Yahya Zakaria Abdou Gaafar

Curriculum vitae

List of publications

2019

1. Gaafar, Y.; Lüddecke, P.; Heidler C.; Hartrick, J.; Sieg-Müller, A; Hübert, C.; Wichura, A.; Ziebell, H. (2019): First report of southern tomato virus in German tomatoes. In *New Dis. Rep*. 40, 1. DOI: 10.5197/j.2044-0588.2019.040.001.
2. Gaafar, Y.; Sieg-Müller, A.; Lüddecke, P.; Hartrick, J.; Seide, Y.; Müller, J.; Maaß, C.; Schuhmann, S.; Richert-Pöggeler, K. R.; Blouin, A.; Massart, S.; Ziebell, H. (2019): First report of natural infection of beetroot with beet soil-borne virus. In *New Dis. Rep*. DOI: 10.5197/j.2044-0588.2019.040.005.
3. Gaafar, Y. Z. A.; Richert-Pöggeler, K. R.; Sieg-Müller, A.; Lüddecke, P.; Herz, K.; Hartrick, J.; Seide, Y.; Vetten, H.-J.; Ziebell, H. (2019): A divergent strain of melon chlorotic spot virus isolated from black medic (*Medicago lupulina*) in Austria. In *Virol. J*. 16, 297. DOI: 10.1186/s12985-019-1195-8.
4. Gaafar, Y. Z. A.; Richert-Pöggeler, K. R.; Sieg-Müller, A.; Lüddecke, P.; Herz, K.; Hartrick, J.; Maaß, C.; Ulrich, R.; Ziebell, H. (2019): Caraway yellows virus, a novel nepovirus from *Carum carvi*. In *Virol. J.* 16, 529. DOI: 10.1186/s12985-019-1181-1.
5. Gaafar, Y. Z. A.; Ziebell, H. (2019): Complete genome sequence of highly divergent carrot torradovirus 1 strain from *Apium graveolens*. In *Arch. Virol.* DOI: 10.1007/s00705-019-04272-3.
6. Gaafar, Y. Z. A.; Richert-Pöggeler, K. R.; Maaß, C.; Vetten, H.-J.; Ziebell, H. (2019): Characterisation of a novel nucleorhabdovirus infecting alfalfa (*Medicago sativa*). In *Virol. J.* 16 (1), p. 113. DOI: 10.1186/s12985-019-1147-3.
7. Gaafar, Y. Z. A.; Ziebell, H. (2019): Two divergent isolates of turnip yellows virus from pea and rapeseed and first report of turnip yellows virus-associated RNA in Germany. In *Microbiol. Resour. Announc.* 8 (17), p. 2254. DOI: 10.1128/MRA.00214-19.
8. Gaafar, Y.; Sieg-Müller, A.; Lüddecke, P.; Herz, K.; Hartrick, J.; Maaß, C. Schuhmann, S.; Richert-Pöggeler, K. R.; Ziebell, H. (2019): First report of turnip crinkle virus infecting garlic mustard (*Alliaria petiolata*) in Germany. In *New Dis. Rep*. 39, p. 9. DOI: 10.5197/j.2044-0588.2019.039.009.

2018

9. Gaafar, Y.; Cordsen Nielsen, G.; Ziebell, H. (2018): Molecular characterisation of the first occurrence of pea necrotic yellow dwarf virus in Denmark. In *New Dis. Rep.* 37, p. 16. DOI: 10.5197/j.2044-0588.2018.037.016.
10. Gaafar, Y. Z. A.; Abdelgalil, M. A. M.; Knierim, D.; Richert-Pöggeler, K. R.; Menzel, W.; Winter, S.; Ziebell, H. (2018): First report of physostegia chlorotic mottle virus on tomato (*Solanum lycopersicum*) in Germany. In *Plant Dis.* 102 (1), p. 255. DOI: 10.1094/PDIS-05-17-0737-PDN.

2017

11. Gaafar, Y.; Timchenko, T.; Ziebell, H. (2017): First report of pea necrotic yellow dwarf virus in The Netherlands. In *New Dis. Rep.* 35, p. 23. DOI: 10.5197/j.2044-0588.2017.035.023.

2016

12. Gaafar, Y.; Grausgruber-Gröger, S.; Ziebell, H. (2016): *Vicia faba*, *V. sativa* and *Lens culinaris* as new hosts for pea necrotic yellow dwarf virus in Germany and Austria. In *New Dis. Rep.* 34, p. 28. DOI: 10.5197/j.2044-0588.2016.034.028.

Conferences and meetings

Jun 2019	Poster	Legume viruses	EP-colloquium, JKI Braunschweig, Germany
Mar 2019	Presentation	Plant viruses identified using high throughput sequencing	DPG meeting of plant virus diseases working group, Göttingen, Germany
Sep 2018	Presentation	Comparative study on three RNA-based pipelines for plant virus/viroid detection using high throughput sequencing	Euphresco meeting, NPPO Wageningen, The Netherlands
Sep 2018	Presentation	EUPHRESCO project (2015-F-172) report	Euphresco meeting, NPPO Wageningen, The Netherlands
Aug 2018	Poster	Plant virus detection and identification, and virus-vector-host interactions	EP-Minisymposium, JKI braunschweig, Germany
Aug 2018	Poster	Legume viruses	EP-Minisymposium, JKI Braunschweig, Germany
Mar 2017	Presentation	Legume viruses in Germany 2016	Dutch and German Plant Virologists meeting, Bonn, Germany
Nov 2016	Presentation	Insights into the Nanovirus-legume-aphid interactions	Young Scientists Meeting (YSM) JKI Quedlinburg, Germany
Oct 2016	Presentation	dsRNA immunocapturing	Euphresco meeting, JKI Braunschweig, Germany
Sep 2016	Poster	The current status of legume viruses in Germany	Advances in Virology conference, Association of Applied Biologists, University of Greenwich London, The UK