



**HAL**  
open science

## **A New Prevalent Densovirus Discovered in Acari. Insight from Metagenomics in Viral Communities Associated with Two-Spotted Mite (*Tetranychus urticae*) Populations**

Sarah François, Doriane Mutuel, Alison Duncan, Leonor Rodrigues, Celya Danzelle, Sophie Lefevre, Inês Santos, Marie Frayssinet, Emmanuel Fernandez, Denis Filloux, et al.

### ► To cite this version:

Sarah François, Doriane Mutuel, Alison Duncan, Leonor Rodrigues, Celya Danzelle, et al.. A New Prevalent Densovirus Discovered in Acari. Insight from Metagenomics in Viral Communities Associated with Two-Spotted Mite (*Tetranychus urticae*) Populations: Metagenomics in Viral Communities Associated with Two-Spotted Mite (*Tetranychus urticae*) Populations. *Viruses*, 2019, 11 (3), pp.233. 10.3390/v11030233 . hal-02154545

**HAL Id: hal-02154545**

**<https://hal.science/hal-02154545>**

Submitted on 12 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.


L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Article

# A New Prevalent Densovirus Discovered in *Acari*. Insight from Metagenomics in Viral Communities Associated with Two-Spotted Mite (*Tetranychus urticae*) Populations

Sarah François <sup>1</sup>, Doriane Mutuel <sup>1</sup>, Alison B. Duncan <sup>2</sup>, Leonor R. Rodrigues <sup>3</sup>,  
Celya Danzelle <sup>1,2</sup>, Sophie Lefevre <sup>2</sup>, Inês Santos <sup>3</sup> , Marie Frayssinet <sup>1</sup>,  
Emmanuel Fernandez <sup>4,5</sup>, Denis Filloux <sup>4,5</sup>, Philippe Roumagnac <sup>4,5</sup>,  
Rémy Froissart <sup>6</sup> and Mylène Ogliastro <sup>1,\*</sup>

<sup>1</sup> DGIMI, Univ Montpellier, INRA, F-34095 Montpellier, France; sarah.francois.dgimi@gmail.com (S.F.); Doriane.mutuel@inra.fr (D.M.); celya.danzelle@icloud.com (C.D.); marie.frayssinet@inra.fr (M.F.)

<sup>2</sup> ISEM, Univ Montpellier, CNRS, IRD, EPHE, F-34095 Montpellier, France; alison.duncan@umontpellier.fr (A.B.D.); sophie.lefevre@umontpellier.fr (S.L.)

<sup>3</sup> Centre for Ecology, Evolution and Environmental Changes (cE3c), Univ Lisbon, Faculty of Science, P-1749016 Lisbon, Portugal; leonor.rodrigues89@gmail.com (L.R.R.); inesflsantos@gmail.com (I.S.)

<sup>4</sup> CIRAD, UMR BGPI, F-34398 Montpellier, France; emmanuel.fernandez@cirad.fr (E.F.); denis.filloux@cirad.fr (D.F.); philippe.roumagnac@cirad.fr (P.R.)

<sup>5</sup> BGPI, Univ Montpellier, INRA, CIRAD, Montpellier SupAgro, F-34398 Montpellier, France

<sup>6</sup> MIVEGEC, Univ Montpellier, CNRS, IRD, F-34394 Montpellier, France; remy.froissart@ird.fr

\* Correspondence: marie-helene.ogliastro@inra.fr

Received: 29 December 2018; Accepted: 28 February 2019; Published: 7 March 2019



**Abstract:** Viral metagenomics and high throughput sequence mining have revealed unexpected diversity, and the potential presence, of parvoviruses in animals from all phyla. Among arthropods, this diversity highlights the poor knowledge that we have regarding the evolutionary history of densoviruses. The aim of this study was to explore densovirus diversity in a small arthropod pest belonging to *Acari*, the two-spotted spider mite *Tetranychus urticae*, while using viral metagenomics based on virus-enrichment. Here, we present the viromes obtained from *T. urticae* laboratory populations made of contigs that are attributed to nine new potential viral species, including the complete sequence of a novel densovirus. The genome of this densovirus has an ambisens genomic organization and an unusually compact size with particularly small non-structural proteins and a predicted major capsid protein that lacks the typical PLA2 motif that is common to all ambidensoviruses described so far. In addition, we showed that this new densovirus had a wide prevalence across populations of mite species tested and a genomic diversity that likely correlates with the host phylogeny. In particular, we observed a low densovirus genomic diversity between the laboratory and natural populations, which suggests that virus within-species evolution is probably slower than initially thought. Lastly, we showed that this novel densovirus can be inoculated to the host plant following feeding by infected mites, and circulate through the plant vascular system. These findings offer new insights into densovirus prevalence, evolution, and ecology.

**Keywords:** parvovirus; viral metagenomics; virus diversity; virus phylogeny; agricultural pests; arthropod; mite; viral communities; viral ecology

## 1. Introduction

Arthropod-infecting parvoviruses, termed densoviruses, have been mostly discovered with a pathology-driven approach, which probably explains the relatively poor number of viral species (i.e., less than 60) that the International Committee for the Taxonomy of Viruses (ICTV) have referenced so far, and the strong bias that exists towards viruses infecting arthropods of health or economic importance [1].

Breakthrough techniques in viral metagenomics and mining high throughput sequencing (HGS) datasets have highlighted the extraordinary diversity and persistence of parvoviruses and parvovirus-related sequences in an unexpected array of animals, including invertebrates [2–4]. These discoveries support the now generally accepted view that viruses are integral components of the microbiome and participate in the functioning of ecosystems of differing size and complexity, i.e., an individual organism or a population of organisms. In the context of the general loss of biodiversity, which is particularly worrisome among arthropods, the question of the prevalence and diversity of their associated microorganisms, including viruses, is of particular interest [5].

Mites and ticks are small arachnids that belong to the *Acari* sub-class that are mostly known for their detrimental impact on human, animal, and plant health [6]. While ticks represent a relatively small number of taxa (around 900 species), all share a parasitic, blood-feeding alimentary regime; mites are extraordinarily diversified (more than 40,000 species) and exhibit a large diversity of lifestyles, including plant feeders, mite-predators, or arthropod ectoparasites, with only a few species presenting direct threats to plant or animal health. Among harmful mites, the two-spotted spider mite (*Tetranychus urticae*) is an agricultural pest that can cause significant damage to more than 1100 species of food-producing or flower cultures [7,8]. Treatment against *T. urticae* with synthetic acaricides has resulted in one of the highest incidences of pesticide resistance being recorded in arthropods [9]. In this context, the development of alternative solutions to chemicals is strongly encouraged; one promising way could be to diversify the use of spider mites' natural enemies, and particularly to include their pathogens.

Despite their environmental success and the threat that they represent for agriculture, due to their small size and the poor description of their associated pathologies, spider-mites have been long neglected by virologists. Only one non-occluded, rod-shaped virus, related to a baculovirus, has so far been described following infections in laboratory populations of mites [10,11]. More recently, viral metagenomics and database mining being applied to *Acari* was developed to explore the viral communities (so-called the viromes) that are associated with blood-feeding ticks [12–15]. These studies revealed an extraordinary wealth of viruses, including parvovirus related sequences, and they validated this approach for virus discovery in these small arthropods. A similar approach was also developed to analyze viruses that are associated with *Varroa destructor*, an ectoparasitic mite of honey bees that also transmits several viruses to the bees and are suspected to contribute to colony collapse [16,17]. This work revealed that several viruses from different families can be found in *Varroa* at the population level, including mite associated viruses that are common in arthropods (e.g., *Baculoviridae*, *Circoviridae*, *Dicistroviridae*, and *Iflaviridae*).

Herein, we used viral metagenomics and in depth sequencing of virion-associated nucleic acids (VANA) [18] to explore the viral communities that are associated with *T. urticae* from two laboratory populations of different geographic origins and rearing host plants. We found eight new putative virus species belonging to taxa that are associated with arthropods and one with fungi, with five of those viral sequences being shared by both mite populations. More specifically, we discovered that the most abundant reads found in both viromes correspond to a new densovirus and probably a new divergent species in the *Ambidensovirus* genus. RACE, PCR, and sequencing confirmed the genome of this new densovirus and the expression of the viral genes in mites. This new densovirus displayed a very compact genome, and the predicted sequence of its major capsid protein lacked the phospholipase A2 motif that is shared so far by all the members of this genus. This virus, initially named *Tetranychus urticae*-associated ambidensovirus (TuaDV), was also prevalent in different species of spider mites

from laboratory populations worldwide and sampled from natural environments. Comparing TuaDV sequence diversity reveals that the virus is highly conserved within each species, but that there is distinct variation between species that are indicative of some virus-host coevolution [19,20]. Finally, we analyzed TuaDV transmission and found that infection can potentially occur via horizontal and vertical routes. In particular, we showed that mites, when feeding, could transfer viral particles to host plants, which could then be further circulated *in planta*, probably through the vascular system. This mechanism, already shown for the aphid densovirus MpDV, suggests that the plant may be used to spread infection to conspecifics and/or to other species [21].

These discoveries expand our knowledge of densovirus prevalence and diversity in an arthropod taxon that has long been neglected by virologists.

## 2. Materials and Methods

### 2.1. Spider-Mite Populations and Rearing

Two laboratory *T. urticae* populations from Portugal and France were first analyzed using the VANA metagenomics-based approach to identify the potential virus species. The first population, the so-called the Portuguese (P) population corresponds to a mix of *T. cinnabarinus* and *T. urticae* populations established at the University of Lisbon. This mix comprised a *T. cinnabarinus* population collected in January 2014 in Spain (Almeria) on rose plants (so-called Almeria population), and two *T. urticae* strains; the “London strain”, originating from the Vineland region, Ontario, Canada [8] and the “EtoxR” strain originating from Japan and maintained for five years in the laboratory at Bayer CropScience [22]. All of the populations and strains were maintained on bean plants in the laboratory. It should be noted that this mix was established before the distinction was made between *T. urticae* and *T. cinnabarinus* as possibly being separate species and not two morphs of the same species. The second population, so-called the French (F) population was reared in Montpellier and it also constitutes a mix of populations. The majority of these originate from a *T. urticae* collection in May 1994 in the Netherlands (NL, Pijnacker) from cucumber, which were subsequently transferred to Montpellier in 2007, at which time they were split into two populations (still maintained on cucumber plants). In 2011, these two populations were mixed, and subsequently transferred to either tomato, bean, or kept on cucumber plants (four populations being created per host plant type). Note that two of these populations (HH4 and HH3) that were transferred to bean in 2011 were used to characterize TuaDV transmission. A second batch of mites (hereafter called “Mix-Tu”) that were used to characterize transmission contained a mix of 11 different mite populations; eight collected in Portugal, two in Spain (including the population collected in Almeria mentioned above), and one in France. All of the laboratory populations were maintained at 25 °C with an 8: 16 Light: Dark (L:D) cycle.

Once the TuaDV virus was identified in the mixes, we measured its prevalence in spider mite populations by analyzing a total of twenty-seven populations from three species of mites (*T. urticae*, *T. cinnabarinus*, and *T. evansi*), out of which TuaDV from eleven populations were partially or fully sequenced (Table S1). Specifically, the mites came from rearing facilities in France (Nice-Valbonne), Greece (Crete), the Netherlands (Santpoort), Belgium (Ghent), and Portugal (Lisbon), and from natural populations across Portugal sampled in the summer of 2017.

### 2.2. Preparation of Viromes and Sequencing

The samples were made from pools of around 200 individual mites, from which viral particles were purified using the method that was described by Francois et al. [18]. Briefly, the mites were ground in HBSS buffer with beads using a tissue homogenizer. The homogenized extracts were filtered through a 0.45 µm filter and centrifuged at  $148,000 \times g$  for 2.5 h at 4 °C to concentrate viral particles. Next, non-encapsidated nucleic acids were eliminated by DNase and RNase digestion for 1.5 h at 37 °C, although some of the non-encapsidated nucleic acids and cellular contamination may remain. Encapsidated DNA and RNA were then extracted using Nucleospin 96 virus Core Kit (Macherey Nagel,

Düren, Germany) and RNA was converted to cDNA using a 26 nt primer (Dodeca Linker) composed of a 14 nt linker linked at the 3' end to N<sub>12</sub>. Double-stranded DNA was synthesized from single-stranded DNA using Large (Klenow) fragment DNA polymerase and the Dodeca Linker. Double-stranded DNA was further amplified using one 24 nt PCR multiplex identifier primer that was composed of the 14 nt linker used during the RT step linked at the 5' end to a 10 nt tag that allowed for sample identification. The PCR products were cleaned with QIAquick PCR Purification kit (Qiagen, Courtaboeuf, France) and the libraries were sequenced on Illumina MiSeq pair-end 300 nt (Genewiz, South Plainfield, NJ, USA).

### 2.3. Bioinformatics Analyses and Database Screening

Demultiplexing was done with the `agrep` command-line tool to assign reads to the samples from which they originated [23]. Adaptors were removed and the reads were filtered for quality (q30 quality and read length >45 nt) using `Cutadapt 1.9` [24]. The cleaned reads were assembled de novo into contigs using `SPAdes 3.6.2` (k-mer lengths 21,33,55,77,125) [25] and mapping was performed on contigs of aligned clean reads with `Bowtie 2.1.0` (options local very sensitive) [26]. Taxonomic assignment was achieved through searches against the NCBI RefSeq viral database and against the non-redundant (nr) GenBank database using `BLASTx` with an e-value cutoff of  $<10^{-3}$  [27]. Viral contigs were classified as viral operational taxonomic units (vOTU). The most abundant vOTUs were subsequently characterized using an arbitrary abundance cutoff of < 0.1% that was applied for each vOTU in all of the samples to remove the inter-sample contamination that occurred during the preparation and the sequencing of the samples. This abundance threshold was chosen to be twice above to the most abundant virus taxa that were found in the negative control (HBSS buffer). The predicted sequence of full-length viral proteins were aligned and compared with their closest related viruses (found in GenBank database) using `MUSCLE 3.7` (16 iterations) [28], according to the species demarcation thresholds that were recommended within the online reports of the ICTV ([www.ictv.global/report/parvoviridae](http://www.ictv.global/report/parvoviridae)).

For database screening, sequences from all viral contigs were used as queries, as well as the genomes contained in NCBI Viral genome database to perform `BLASTn` searches within the *T. urticae* genome: RefSeq genomic database (GCF\_000239435.1, 641 sequences), WGS (CAEY000000000.1, 2035 sequences) and transcriptomes: EST (txid32264, 80855 sequences) and TSA (BioProject 78685, 9614 sequences; BioProject 78689, 17739 sequences; BioProject 6829 sequences), with and an e-value cutoff of  $<10^{-3}$ .

### 2.4. Validation of the TuaDV Full-Length Genomic Sequence

To confirm the presence of TuaDV in mites and confirm its complete coding sequence, pools of mites (n~100) from several populations were tested using PCR. Total DNA was extracted with Wizard Genomic kit (Promega Corp., Madison, WI, USA) and then recovered in a final volume 100 µL (100–135 ng/µL). PCR was performed from 100 ng DNA with specific overlapping sets of primers covering the full length of the genome (Table S2) and while using the GoTaq reaction mix (Promega). The amplicons were sequenced with the Sanger method. To determine the extremities of the TuaDV genome, we performed 5'/3' RACE with specifically designed primers (Table S2). Total RNA was extracted from pools of mites from the HH4 population with the RNeasy minikit (Qiagen). 3' and 5'RACE PCR were performed using the 5'/3' RACE kit, 2nd Generation (Roche, Germany), according to the manufacturer's instructions.

### 2.5. Phylogenetic Analyses

The putative amino acid sequences of *Tetranychus* associated vOTUs were used for phylogenetic analyses. All of the ORFs were translated in silico using the ORF finder (cut off >300 nt, ATG start codon) on Geneious 1.7 [29] and aligned with the corresponding protein fragments of related viruses being deposited on the GenBank nr database using `MUSCLE 3.7` (16 iterations) with default settings [28]. The aligned sequences were manually edited to remove gaps. Maximum likelihood phylogenetic trees were produced from these alignments using `PhyML 3.1` [30,31] with substitution models being chosen

as the best-fit using Prottest 2.4 [32]. One-thousand bootstrap replicates were used to assign the strength of support for branches. Trees were visualized with FigTree 1.4 (<http://tree.bio.ed.ac.uk/software/20/figtree/>). Outgroups were used when possible, otherwise the trees were mid-point rooted.

### 2.6. *Densovirus Presence and Transmission*

The prevalence of TuaDV was assayed from pools of mites from the different populations using PCR, as described above. PCR was run using the following conditions: 95 °C for 2 min, 25 cycles of 95 °C for 45 sec, 57 °C for 45 sec, 72 °C for 1 min, and then 72 °C for 5 min and two sets of specific primers (61F and 693R (NS) and 1125 and 1793R (VP)). Amplicons were run in an agarose gel and stained with ethidium bromide. We also tested for the within-population prevalence of TuaDV in the laboratory population HH4. Extracting DNA from 20 individual mites from this population achieved this and running PCRs, as described above, to test for the presence of the virus.

When considering the universal presence of the virus, a separate experiment tested whether *T. urticae* associated densovirus was vertically transmitted from mothers to their offspring. To do this, 10 females from the HH4 population were isolated on individual bean leaves placed on water saturated cotton, and left to lay eggs for two days at 25 °C with a 8: 16 L: D cycle. Each female laid eggs that were recovered individually with a clean pair of tweezers and transferred to a new individual leaf patch that was placed on water saturated cotton, allowed to hatch, and offspring to develop into adult. Offspring ( $n = 10$ ) that became adults and their mother were individually tested for viral infection. Total DNA was extracted from each individual mite, as described above, and recovered in 25  $\mu\text{L}$  (8–12 ng/ $\mu\text{L}$  per mite). PCR reactions were set as above and run with 40 cycles. Amplicons were run in an agarose gel and stained with ethidium bromide. The detection limit of the method is estimated  $\sim 10\text{ng}$ .

To test for TuaDV inoculation and circulation in the host plant, we placed 50 infected mites on an individual leaf of a whole bean plant (8–10 leaves stage). Paraffin jelly was placed at the base of the petiole of each leaf in order to prevent mites from dispersing from the infested leaf to other leaves on the plant. Five days later, the mites were carefully removed from the infested leaf (named L0) and two uninfested leaves from the same plant, one  $\sim 8$  cm and the other  $\sim 20$  cm away from the infested leaf, were collected (named L8 and L20, respectively). We performed six independent transmission experiments with the HH4, HH3, and Mix-Tu populations. Three leaves from six non-infested plants were similarly taken and processed as negative controls. Total DNA was extracted from infested plants and negative controls using DNeasy blood and tissue kit (Qiagen) and recovered in 100  $\mu\text{L}$  (40–60 ng/ $\mu\text{L}$  for leaves) and 25  $\mu\text{L}$  (15–25 ng/ $\mu\text{L}$  for mites). The TuaDV inoculated to plant leaves was assessed by PCR, as described above, while viral loads in mites and leaves was achieved by qPCR with a reaction mixture containing 15–20 ng and 100 ng of mites or plant DNA, respectively, and a pair of TuaDV specific primers (1125F and 1230R, Table S2) using the Sensifast PCR kit (Bioline, UK). qPCR was run on a Roche LC480 cycler (qPHD facility, Univ Montpellier) using the following conditions: 95 °C for 2 min, 25 cycles of 95 °C for 30 sec, 60 °C for 30 sec, 72 °C for 50 sec, and then 72 °C for 5 min, which resulted in a 78 bp amplicon from ORF2 (NS). As the viral genome of TuaDV remains to be cloned, PCR Standardization was achieved with a standard curve that was made with the serial dilution of DNA of the *Junonia coenia* ambidensovirus (JcDV) as a proxy.

Statistical analyses were performed with the JMP®Version 11 software (SAS Institute Inc., Cary, NC, USA). The analysis of viral load (ng  $\mu\text{L}^{-1}$ ) within each leaf was performed using a general linear model with leaves (infested, uninfested 8cm or 20 cm away) and mite populations (HH3, HH4 and Mix-Tu) being included in the model as fixed factors. The plant from where the leaves were taken was included in the model as a random factor nested within population.

Accession numbers: The virus sequences reported herein have been deposited in the GenBank database under accession numbers MK533146 to MK533158 and MK543949.

### 3. Results

#### 3.1. Overview of the Spider Mite Virome

To explore the viral diversity of *T. urticae* populations, we prepared viral particles according to the VANA protocol, from two laboratory populations from France (Montpellier, F population) and Portugal (Lisbon, P population), having different geographical origins and rearing history. Pools of mites from both of the populations were processed for virus-enrichment (VANA) metagenomics and high throughput sequencing. A total of 680,600 cleaned reads were obtained, including 219,982 reads from the French (F) population and 460,618 reads from the Portuguese (P) population.

*De novo* assembly of the VANA reads in both populations obtained fourteen viral contigs (1.6 to 8.6 kb in length) (Table 1). Thirteen out of the fourteen contigs were mostly related to non-enveloped DNA and RNA viruses belonging to clades infecting arthropods, including *Dicistroviridae*, *Parvoviridae*, *Birnaviridae*, *Nodaviridae*, and unclassified picornavirales. In addition, one viral contig (2481 nt) was similar to yeast and fungi-infecting viruses of the *Narnaviridae* family, which might come from environmental contamination (e.g., food), even though we cannot exclude its replication in spider mites [33]. While nine viral contigs were isolated in the P population only, five viral contigs were common to both of the populations (Figure 1), which was surprising given their different origins and rearing history.

Among the five viral contigs that are common to both viromes, the one that is assigned to the *Parvoviridae* family was largely dominant in terms of reads abundance (>28% of viral reads in both populations), with the others being found at much lower frequencies (0.1% to 10.6%) (Figure 1). Sequence analysis and open reading frame (ORF) prediction showed that this parvovirus has an ambisense genomic organization with a predicted small Non Structural protein-1 (NS1) that shared 39% aa identity with the NS1 of its closest relative, the Lupine feces-associated densovirus 2, which was discovered from viral gut metagenomics of the Iberian wolf and has no associated arthropod host so far (accession number: ASM93489). Accordingly, the phylogenetic tree that is based on the NS1 protein clustered this densovirus with Lupine feces-associated densovirus 2 among the *Ambidensovirus* genus (Figure 2). According to the new species demarcation threshold in the *Densovirinae* sub-family that was proposed by the ICTV (i.e., < 85% related by NS1 amino acid sequence identity [34]), this new densovirus might represent a new divergent species in the *Ambidensovirus* genus and the first densovirus isolated from *Acari* (Figure 2, Table 1). Hereafter, this virus is referred to as *Tetranychus urticae*-associated ambidensovirus (TuaDV).

Although the *Birnaviridae* family has been poorly investigated, the phylogenetic analyses showed that the polymerase protein of *Tetranychus urticae*-associated birnavirus that was found in *T. urticae* viromes (representing 3.8% of reads) clustered with *Drosophila melanogaster* birnavirus (Figure 3) within an unclassified lineage. The polymerase and capsid proteins of this putative novel entomobirnavirus also shared 30% aa identity with the Infectious bursal disease virus (accession number AAS10174.1) and 33% aa identity with the Blotched snakehead virus (accession number YP\_052864.1), respectively. Therefore, this putative novel birnavirus could represent the first birnavirus isolated from Arachnids (Table 1).

**Table 1.** Description of the fourteen virus contigs found in *T. urticae* viromes and protein identity comparison with their closest relatives.

Baltimore Classification	Viral Family	Viral Contig	Contig Length (nt)	Putative Protein	Best Hit (BLASTx)	Accession Number	Protein Identity (%)
ssDNA	<i>Parvoviridae</i>	Tetranychus urticae-associated ambidensovirus	3411	polymerase (NS1) capsid	<i>Lupine feces-associated densovirus 2</i>	ASM93489.1	39%
						ASM93488.1	30%
dsRNA	<i>Birnaviridae</i>	Tetranychus urticae-associated birnavirus Segment A	2899	polymerase polyprotein	<i>Infectious bursal disease virus</i> <i>Blotched snakehead virus</i>	AAS10174.1	30%
		Tetranychus urticae-associated birnavirus Segment B	2501			YP_052864.1	33%
ss+RNA	<i>Nodaviridae</i>	Tetranychus urticae-associated nodavirus Segment A 1	3230	polymerase	<i>Hubei noda-like virus 9</i>	APG76321.1	58%
		Tetranychus urticae-associated nodavirus Segment A 2	2538	polymerase		YP_009337881.1	62%
		Tetranychus urticae-associated nodavirus Segment B 1	1661	capsid	<i>Hubei noda-like virus 9</i>	YP_009337880.1	21%
		Tetranychus urticae-associated nodavirus Segment B 2	1658	capsid			53%
		Tetranychus urticae-associated nodavirus Segment B 3	1570	capsid			58%
	<i>Dicistroviridae</i>	Tetranychus urticae-associated dicistrovirus 1	8290	polymerase capsid	<i>Beihai picorna-like virus 70</i>	APG78061.1	24%
		Tetranychus urticae-associated dicistrovirus 2	8449	polymerase capsid		APG78062.1	20%
	Unclassified <i>Picornavirales</i>	Aphis glycines virus 1	8592	polymerase capsid	<i>Aphis glycines virus 1</i>	AHC72013.1	96%
		Tetranychus urticae-associated picorna-like virus 1	8151	polymerase capsid		AHC72012.1	99%
		Tetranychus urticae-associated picorna-like virus 2	6432	polymerase capsid		AHC72013.1	71%
			6432	polyprotein		AHC72012.1	75%
<i>Narnaviridae</i>	Tetranychus urticae-associated narnavirus	2481	polymerase	<i>Hubei picorna-like virus 80</i>	YP_009337381.1	53%	
				<i>Hubei narna-like virus 3</i>	YP_009337787.1	45%	



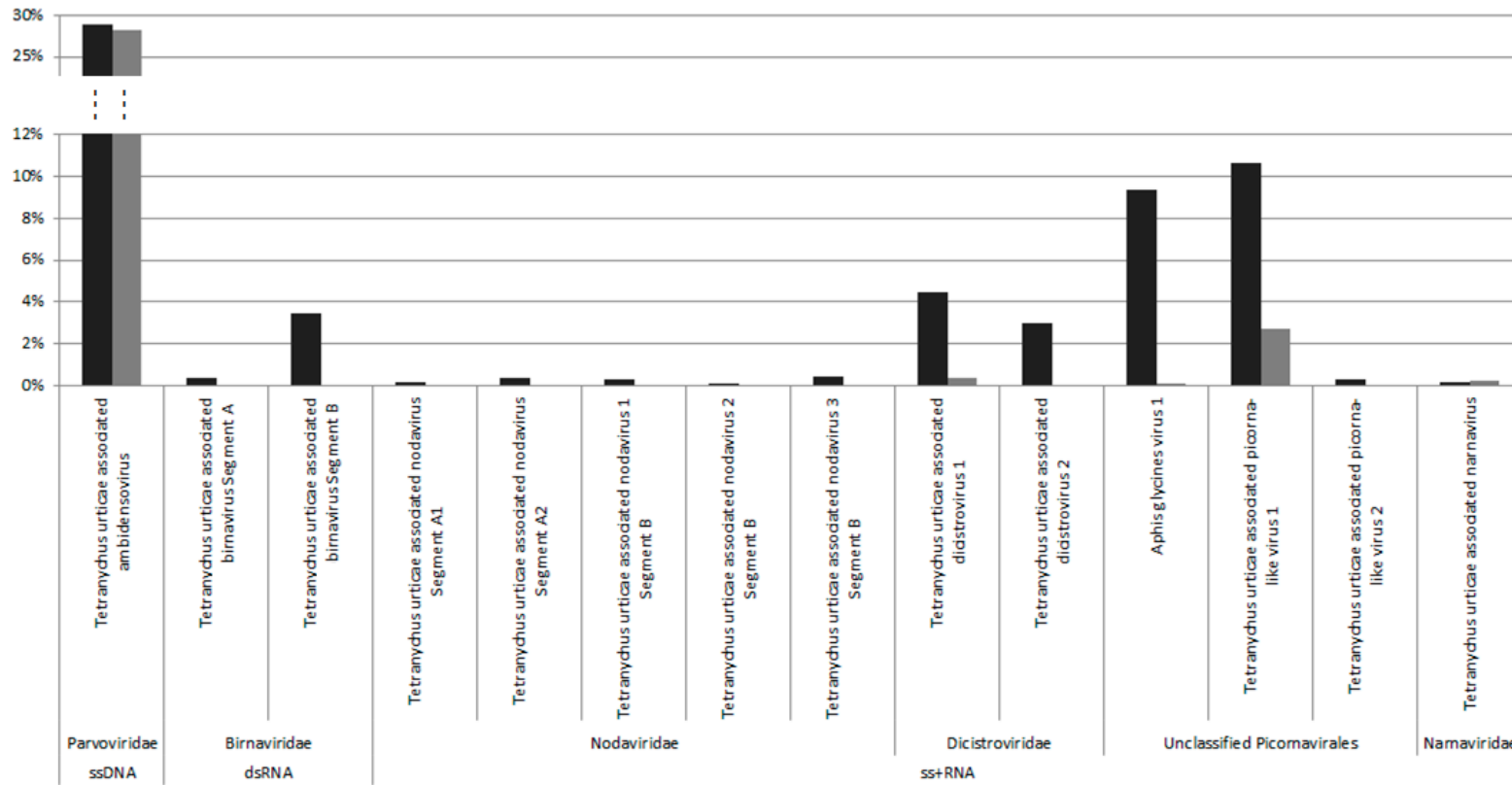


Figure 1. Relative abundance of putative viral species in the viromes of Portuguese (black) and French (grey) *T. urticae* colonies.

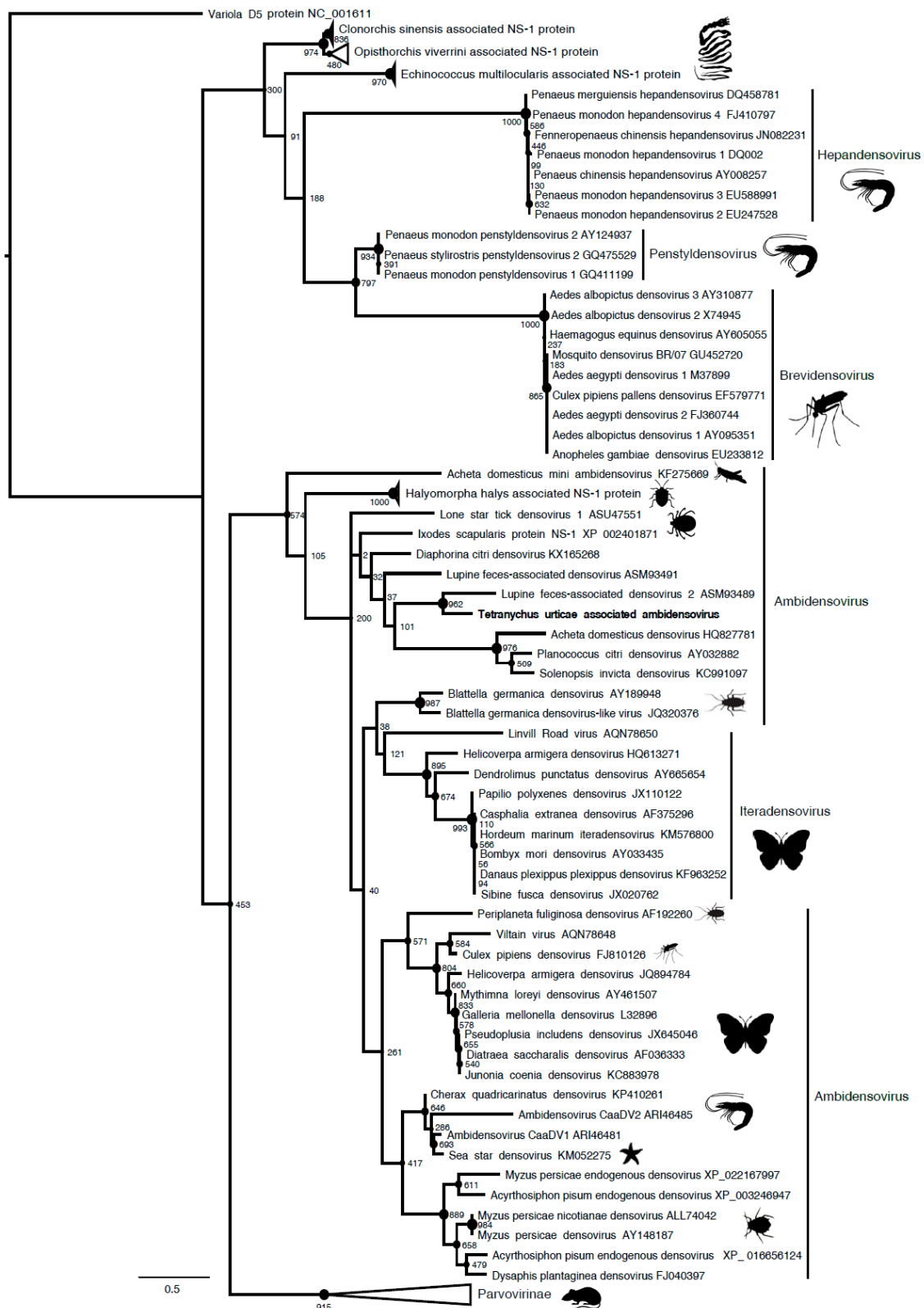
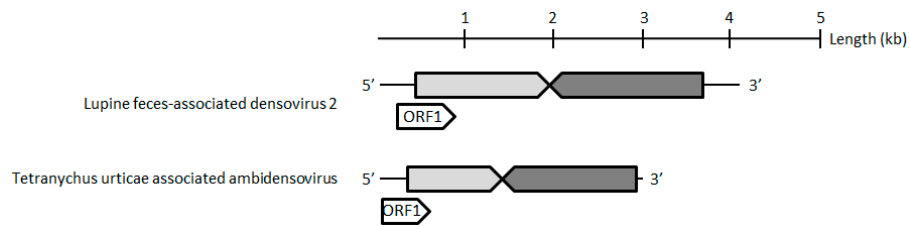
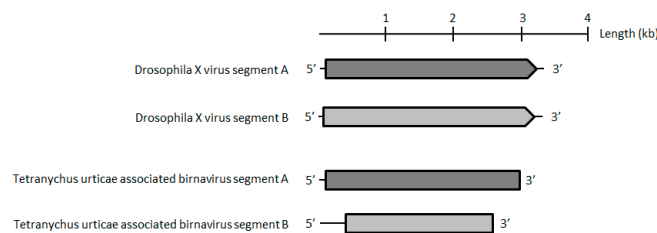
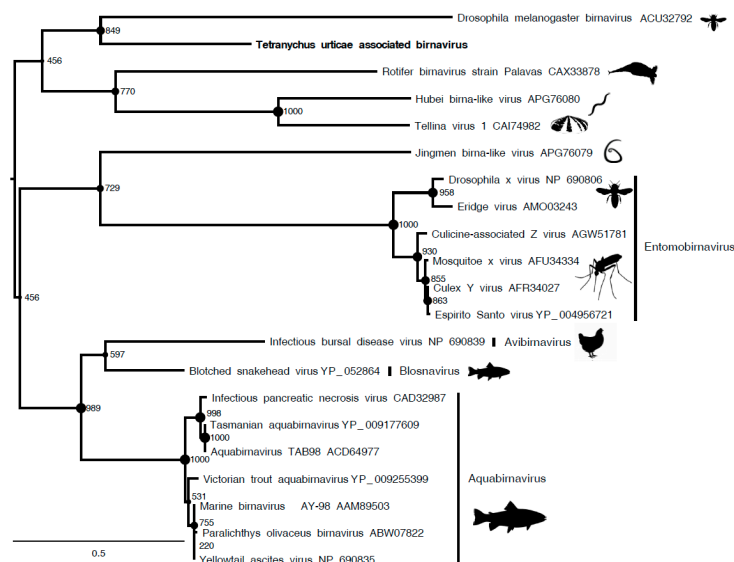


Figure 2. Cont.



**Figure 2.** Maximum *Parvoviridae* likelihood phylogenetic tree based on a part of the NS1 protein containing the SF3 domain, including 105 parvovirus sequences and *Tetranychus urticae*-associated ambidensovirus (in bold). The alignment of 155 amino acids in length was produced using MUSCLE 3.7 (16 iterations) and was ungapped by hand. The tree was rooted with the SF3 domain of the Variola virus D5 protein. Bootstrap values are indicated at each node. Scale bar corresponds to amino acid substitutions per site. Genera of the *Parvoviridae* family are indicated in brackets. Genomic organization of *Tetranychus urticae*-associated ambidensovirus is also indicated. Grey arrows and rectangles: predicted open reading frames (ORF), Light grey: putative NS; dark grey: putative capsid protein (CP or VP). Arrow: complete ORF.



**Figure 3.** Maximum *Birnaviridae* likelihood phylogenetic tree based on part of the polyprotein, including 20 birnavirus species and *Tetranychus urticae*-associated birnavirus (in bold). The alignment of 602 amino acids in length was produced using MUSCLE 3.7 (16 iterations) and was ungapped by hand. The tree was mid-point rooted. Bootstrap values are indicated at each node. Scale bar corresponds to amino acid substitutions per site. Genera of the *Birnaviridae* family are indicated in brackets. Genomic organization of *Tetranychus urticae*-associated birnavirus is also indicated. Grey arrows and rectangles: predicted open reading frames (ORF), Light grey: putative NS; dark grey: putative capsid protein (CP). Arrow: complete ORF; Rectangle: truncated ORF.

Regarding two of the three viral contigs clustering within the *Nodaviridae* family (Figure 4, Table 1), coding for the capsid protein, and found in the P population only (representing 1.35% of reads), they share up to 58% aa identity with the Hubei noda-like virus 9 capsid protein (accession number:

YP\_009337880.1), an unclassified RNA virus. According to the ICTV species demarcation threshold (<87% of capsid protein aa identity) and to the position of polymerase proteins in the phylogenetic tree, the nodaviruses that were found in *T. urticae* viromes may correspond to novel species-level lineages in a novel genus-level lineage in the *Nodaviridae* family (Figure 4, Table 1).

Five contigs were assigned to the *Picornavirales* order. While two contigs clustered in the *Dicistroviridae* family (hereafter referred to as *Tetranychus urticae*-associated dicistrovirus 1 and 2 (Tuad1 and Tuad2), the other three clustered with the unclassified Picorna-like viruses. The capsids of Tuad1 and Tuad2 both share 20% aa identity with Beihai picorna-like virus 70 (accession number APG78062.1; Table 1). Based on the current species demarcation criteria that are used by the ICTV *Dicistroviridae* study group (<90% of capsid protein identity with closest relatives) and the phylogenetic analyses based on the conserved polymerase protein (Figure 5, Table 1), it is likely that both of the contigs could belong to two novel species-level lineages of the *Dicistroviridae* family. Interestingly, the CP of one of the three unclassified picorna-like viruses shared >99% aa identity with the unclassified picorna-like *Aphis glycines* virus 1 (Table 1). In addition, the proteins of the two unclassified picorna-like viruses (hereafter referred to as *Tetranychus urticae*, being associated picorna-like virus 1 and 2) share 53% to 75% aa identity with *Aphis glycines* virus 1 and Hubei picorna-like virus 80 (Table 1). Their phylogenetic trees showed that they might belong to a highly divergent lineage within the *Picornavirales* order (Figure 5). Moreover, *Tetranychus urticae*-associated picorna-like virus 1 and 2 could represent a new species according to species demarcation criteria that are defined by the ICTV (<90% of capsid protein identity with closest relatives) (Table 1).

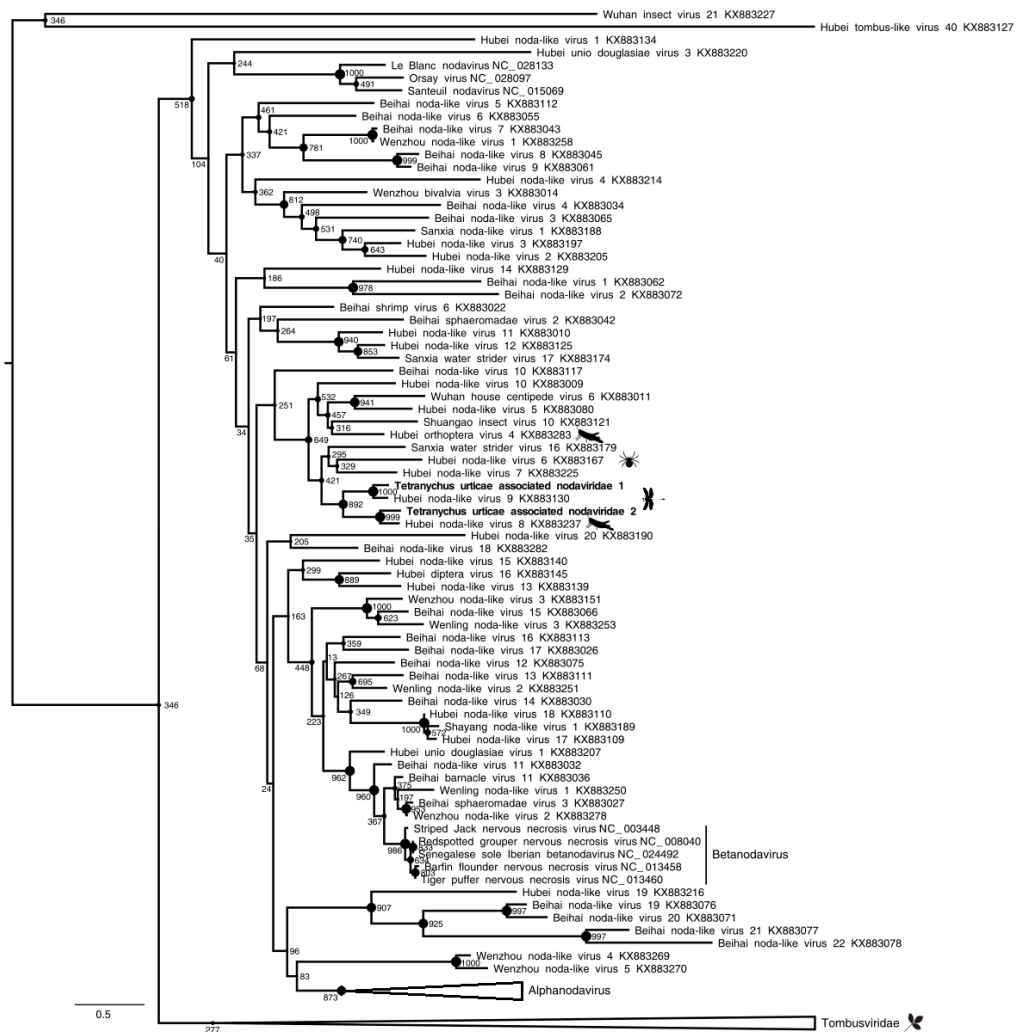
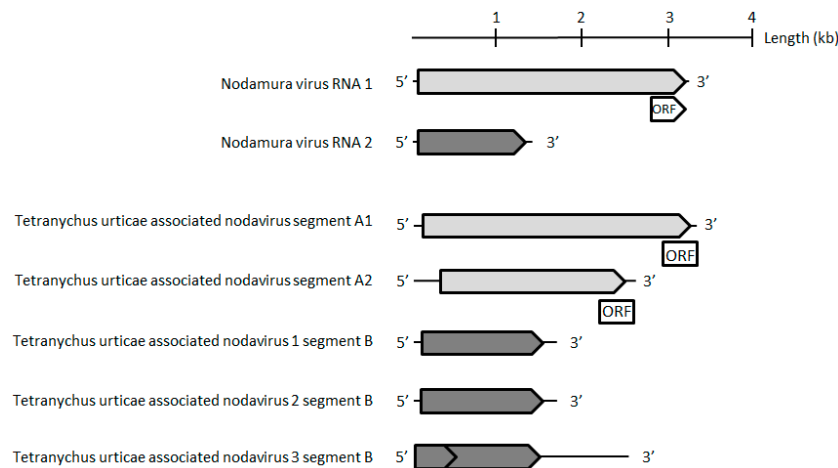
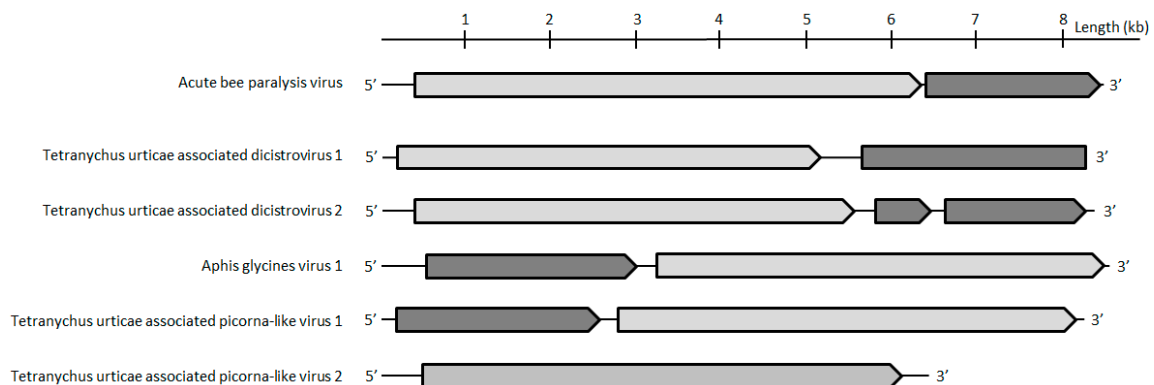


Figure 4. Cont.



**Figure 4.** Maximum *Nodaviridae* likelihood phylogenetic tree based on part of the capsid protein, including 239 species and *Tetranychus urticae*-associated nodaviruses (in bold). The alignment of 229 amino acids in length was produced using MUSCLE 3.7 (16 iterations) and was ungapped by hand. The tree was mid-point rooted. Bootstrap values are indicated at each node. Scale bar corresponds to amino acid substitutions per site. Genomic organization of *Tetranychus urticae*-associated nodaviruses is also indicated. Grey arrows and rectangles: predicted open reading frames (ORF), Light grey: putative NS; dark grey: putative capsid protein (CP). Arrow: complete ORF; Rectangle: truncated ORF.

Finally, the *Tetranychus urticae*-associated narnavirus, present in 0.15% and 0.2% of reads in P and F populations, respectively, was the only one found in this study that historically belongs to a family of fungi-infecting viruses. Although very little is known regarding these viruses, the phylogenetic position of its polymerase protein within a clade including associated narnaviruses, as well as members of the *Ourmiavirus* genus, suggests that it might represent a new species-level lineage according to the ICTV species demarcation threshold (<50% of protein sequence identity as compared to the closest relative) (Figure 6, Table 1).



**Figure 5.** Maximum *Picornavirales* likelihood phylogenetic tree based on part of the polymerase protein, including 504 sequences and *Tetranychus urticae*-associated picornaviruses (in bold). The alignment of 386 amino acids in length was produced using MUSCLE 3.7 (16 iterations) and was ungapped by hand. The tree was mid-point rooted. Bootstrap values are indicated at each node, values <50% were discarded to not overload the tree. Scale bar corresponds to amino acid substitutions per site. Families of the *Picornavirales* order and genera of the *Dicistroviridae* family are indicated in brackets. Genomic organization of *Tetranychus urticae*-associated picornaviruses is also indicated. Grey arrows and rectangles: predicted open reading frames (ORF), Light grey: putative NS; dark grey: putative CP. Arrows: complete ORF.

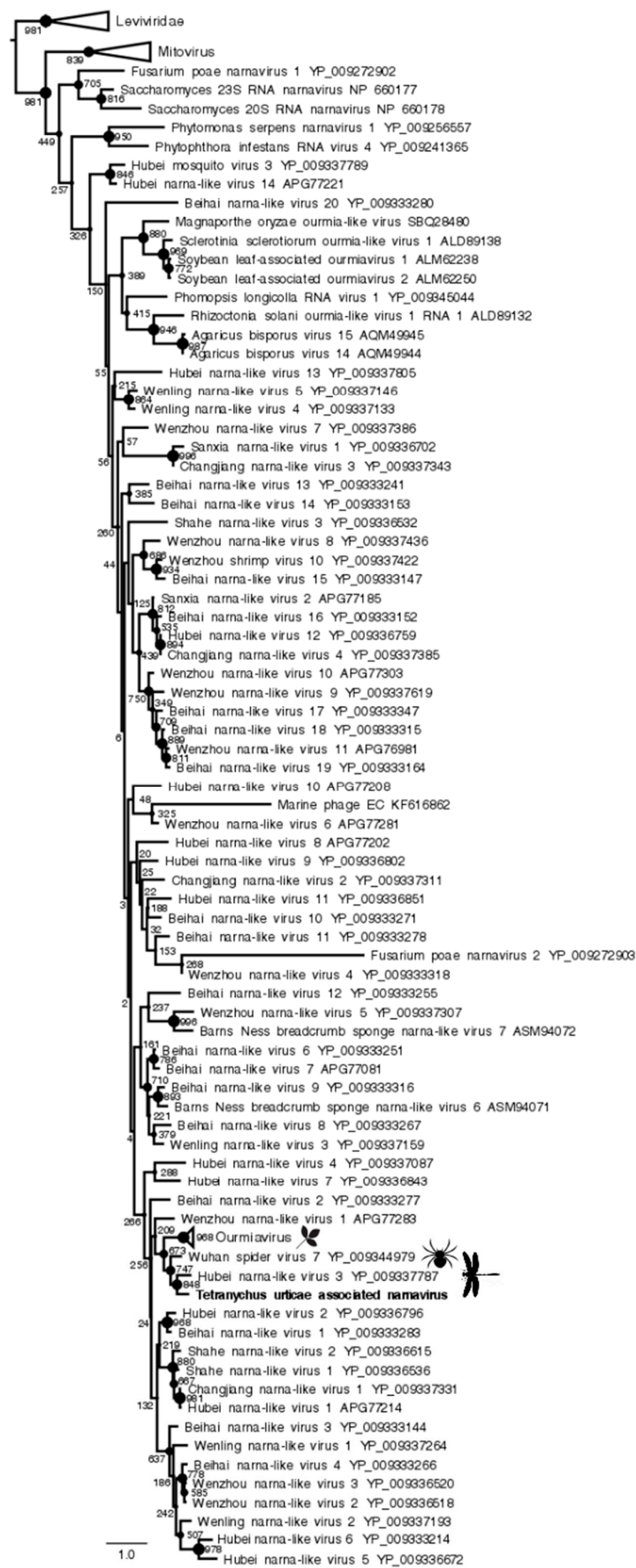
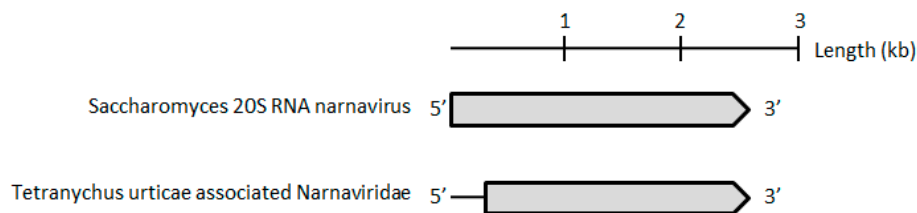


Figure 6. Cont.



**Figure 6.** Maximum *Narnaviridae* likelihood phylogenetic tree based on part of the polymerase protein, including 129 species and *Tetranychus urticae*-associated narnavirus (in bold). The alignment of 121 amino acids in length was produced using MUSCLE 3.7 (16 iterations) and was ungapped by hand. Bootstrap values are indicated at each node. Scale bar corresponds to amino acid substitutions per site. Genomic organization of *Tetranychus urticae*-associated narnavirus is also indicated. Grey arrows: putative NS.

### 3.2. Genomic and Transcriptomic Database Screening

To gain insights into the diversity and the distribution of viruses in *T. urticae*, we further screened mite genomic and transcriptomic datasets using all ten viruses that were identified in this study as queries. Our search highlighted that *T. urticae* transcriptomes contained sequences that displayed >95% of nucleotidic identity to *Aphis glycines* virus 1 and *Tetranychus urticae*-associated picorna-like virus 1 (Table 2). Interestingly, one sequence that was related to an ambidensovirus was found in the genome of *T. urticae*, but this sequence was different from TuaDV (70% of nucleotidic identity with a sequence of 536 nt length). No sequence corresponding to TuaDV was found in any of the transcriptomes that were analyzed, suggesting that this virus might correspond to a non-endogenized virus, whose origin remains to be clarified.

In addition to these viruses, we found in *T. urticae* genomes and transcriptomes, 16 sequences belonging to large viruses, including giant viruses (*Mimiviridae*, *Phycodnaviridae*, *Poxviridae*, and *Baculoviridae*). These sequences were not found in the viromes that we generated in this study (i.e., using VANA based method) and we cannot exclude that the sequences corresponding to this large dsDNA virus might have a cellular rather than a viral origin. Eight sequences that were related to *Rhabdoviridae* were found in the *T. urticae* genome, which all corresponded to the nucleoprotein N, and we also found these sequences in the EST transcriptome (six sequences > 95% of nucleotide identity). Last, we also found sequences that were related to plant viruses in the TSA transcriptome of *T. urticae* that might originate from diet contamination. One transcriptomic sequence (accession number GW017620.1) matched with the *Tetranychus urticae*-associated nodavirus segment B1, although its small size did not allow its for assignment with high confidence (Table 2).

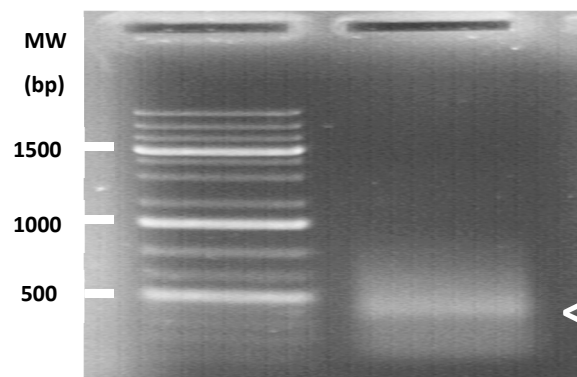
**Table 2.** Summary of the number of sequences found in the *Tetranychus urticae* genomic and/or transcriptomic databases that display homologies with viral species discovered in the viromes generated in this study.

Viral Contig	Contig Length (nt)	ref_seq	Genomics	EST	TSA
<i>Tetranychus urticae</i> -associated ambidensovirus	3411	0	0	0	0
<i>Tetranychus urticae</i> -associated birnavirus Segment A	2899	0	0	0	0
<i>Tetranychus urticae</i> -associated birnavirus Segment B	2501	0	0	0	0
<i>Tetranychus urticae</i> -associated nodavirus Segment A 1	3230	0	0	0	0
<i>Tetranychus urticae</i> -associated nodavirus Segment A 2	2538	0	0	0	0
<i>Tetranychus urticae</i> -associated nodavirus Segment B 1	1661	0	1	0	0
<i>Tetranychus urticae</i> -associated nodavirus Segment B 2	1658	0	0	0	0
<i>Tetranychus urticae</i> associated nodavirus Segment B 3	1570	0	0	0	0
<i>Tetranychus urticae</i> associated dicistrovirus 1	8290	0	0	0	0
<i>Tetranychus urticae</i> associated dicistrovirus 2	8449	0	0	0	0
<i>Aphis glycines</i> virus 1	8592	0	105	0	0
<i>Tetranychus urticae</i> associated picorna-like virus 1	8151	0	0	0	14
<i>Tetranychus urticae</i> associated picorna-like virus 2	6432	0	0	0	0
<i>Tetranychus urticae</i> associated narnavirus	2481	0	0	0	0

### 3.3. Characterization of a Novel Densovirus in Spider-Mites

As pointed out above, the size of the densovirus contig found in mites populations was about 3.4 kb, which was much smaller than genomes characterized so far in the *Ambidensovirus* genus, which, when excluding the terminal repeats (ITRs), is around 4 kb for the *Acheta domesticus* mini ambidensovirus and up to 5kb for Lepidopteran ambidensoviruses [34]. Ambidensoviruses usually display a single ORF encoding for one to four structural proteins (VP1–4) that are produced by splicing or leaky scanning, and three ORFs encoding for non-structural (NS) proteins. Viruses in the *Parvoviridae* family are characterized by two typical domains: i) a phospholipase A2 (PLA2) motif located in VP1 of most parvoviruses, including in all species described so far in the *Ambidensovirus* genus; and, ii) a Super Family 3 (SF3) helicase domain that is located in the NS1 protein and common to all parvoviruses [35].

The contig corresponding to the TuaDV genome predicted three to four open reading frames (ORFs), one encoding a typical VP protein (ORF1), and two to three putative ORFs encoding NS proteins (ORF2–4) (Figure 2). 5' and 3' RACEs further verified the 5' and 3' ends of the viral genome and we performed overlapping PCRs with specific primers and sequenced the amplicons using the Sanger method (Figure 7). These results further confirmed that both of the populations shared the same TuaDV genomic sequence and that both expressed viral genes.



**Figure 7.** Identification of the transcription starting site of ORF2 (*ns* gene) of the TuaDV by 5' RACE PCR. Agarose gel showed a major amplicon at ~500 bp (white arrowhead). The amplicon was sequenced to recover the 5' end sequence.

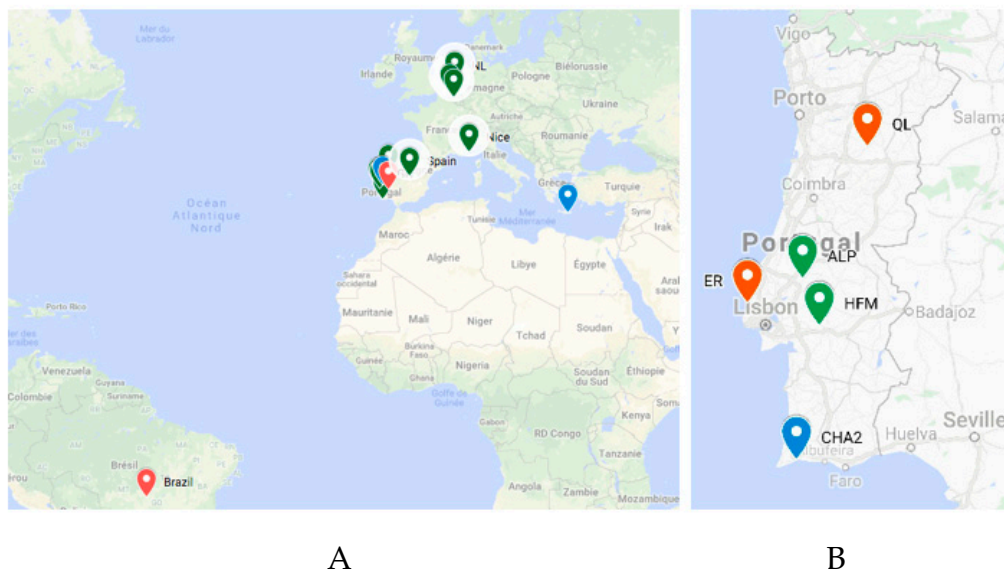
We did not obtain any complementary sequence for ORF1 (VP), which thus predicted a 505-amino acids (aa) protein of 55 kDa that lacked the typical PLA2 motif that is common to all ambidensoviruses described so far. Concerning the reverse strand, we obtained the complete coding sequence of the NS ORFs with a predicted ORF2 (NS1) sequence of 354 aa protein of 39 kDa, harboring the typical SF3 domain, while a predicted ORF3 (NS2) has a 273 aa protein and a molecular weight of 30 kDa. In addition, a small ORF4 (95 aa) was predicted and it might encode for a putative 10 kDa NS3 protein. With such a compact genome, this virus would be the smallest densovirus that is described so far, including the *Acheta domesticus* mini ambidensovirus, with both viruses having sizes that are comparable to species from the *Iteradensovirus* genus. The closest relative of TuaDV was the Lupine feces-associated densovirus 2 (accession number KY214445.1), which shares 30% identity with NS1 and 31% identity with VP [36]. The putative NS2 of these viruses shared 25% identity. If correct, these NS predicted ORFs would have an ORF3 initiating upstream ORF2, which is unusual among ambidensoviruses. Based on the current species demarcation criteria that were used by the ICTV, all of these features suggest that TuaDV is a new mini ambidensovirus species among the *Ambidensovirus* genus that shares little sequence identity with the *Acheta domesticus* mini ambidensovirus.

Given the theoretical rapid evolution of ssDNA viruses [37], the fact that mite populations with different origins and rearing history share identical virus sequences more likely suggests that contamination occurred between laboratories, due to mites and material exchange, which is also supported by the introduction of the Almeria population in the Montpellier laboratory at the time.



### 3.4. TuaDV Prevalence and Diversity in Mite Populations

To better understand the origin of the TuaDV that infected the P and F reared populations (so-called here after TuaDV\_Rearing (R) to discriminate from \_Field populations), we investigated its prevalence and diversity in different *Tetranychus* species with different origins. Twenty-seven samples from three species of *Tetranychus* (*T. urticae*, *T. cinnabarinus*, and *T. evansi*) were collected from various fields in Portugal and from rearing facilities worldwide (France, Brazil, Netherlands, Belgium, and Crete), representing 19 and eight samples, respectively (Figure 8). Interestingly, the genome of the Lupine feces densovirus 2 was recently discovered by metagenomics from feces of Iberian wolves that were sampled in Portugal (South Douro region), which suggests that distant relatives of TuaDV are present in the environment, at least on the Iberian Peninsula.



**Figure 8.** Map representing the origin of spider mites' populations worldwide (A) and in Portugal (boxed) (B) where the prevalence of TuaDV has been assayed by PCR and sequenced (full or partial genome) and in Portugal (B). *T. urticae* (green), *T. cinnabarinus* (blue), and *T. evansi* (red). Maps were made from maps.google.com.

TuaDV presence and its genomic sequences in each population were confirmed by conventional PCRs using various sets of primers and sequencing (see Materials & Methods section). Note that each sample corresponds to a pool of mites from the same species that were collected in the same field/rearing. Amplicons were obtained with all sets of primers and they were of the expected size, for all samples, thus indicating 100% prevalence in the samples, and suggests a wide prevalence of TuaDV in spider mite populations, both in natural environments and from all of the rearing facilities tested. However, we cannot determine whether all of the individual mites were infected or few were highly infected.

To investigate the phylogenetic relationship between the viruses, we sequenced TuaDV amplicons from 11 selected samples, i.e., four *T. urticae* (a second sample from France, Belgium, Netherlands, and Portugal), four *T. cinnabarinus* (Portugal), and three *T. evansi* (Portugal and Brazil); for each species, one sample corresponds to mites that were collected from the field. We used various sets of primers in order to cover the full-length genome or the ORF2 (NS1 gene) only (Table S1).

We reconstructed the TuaDV full length genomes from three species that were recently collected in Portugal by sequence assembly and alignments. The recovered genomes displayed a high identity score (Table 3), with the TuaDV that was isolated from *T. evansi* being the most distantly related (81% identity with the original TuaDV\_R genome; 86% when considering only the identity of NS1); although, these viruses belong to the same species according to the species demarcation criteria of the ICTV.

The TuaDV genomes that were sequenced from *T. urticae* and *T. cinnabarinus* were both similar to the initial TuaDV\_R (99.8%; 100% identity of NS1), with only three and five synonymous substitutions between TuaDV\_R and TuaDV from *T. cinnabarinus* and *T. urticae* from natural populations, respectively; the virus from the two natural populations differed by five substitutions, which further suggests that TuaDV\_R originates from *T. cinnabarinus*. This relatedness is congruent with the phylogenetic position of *T. urticae* and *T. cinnabarinus*, often considered as two morphs of the same species, with *T. evansi* being more distantly related [19,20]. Moreover, it suggests that the infection of the population reared in Portugal originated from the Almeria population introduced in the rearing three years ago (see methods section).

**Table 3.** Percentage identity (DNA; aa) between the TuaDVs. *Upper panel.* VP1 coding sequence (CDS)/protein isolated from field populations. *Lower panel* NS1 CDS/protein isolated from field populations.

	<i>T. cinnabarinus</i>	<i>T. evansi</i>	<i>T. urticae</i> _Field	<i>T. urticae</i> _R (TuaDV)
<i>T. cinnabarinus</i>		(79.0; 79.6)%	(99.8; 99.4)%	(99.8; 100)%
<i>T. evansi</i>	(79.0; 79.6)%		(79.0; 79.6)%	(78.9; 79.6)%
<i>T. urticae</i> _Field	(99.8; 99.4)%	(79.0; 79.6)%		(99.6; 99.4)%
<i>T. urticae</i> _R (TuaDV)	(99.8; 100)%	(78.9; 79.6)%	(99.6; 99.4)%	

	<i>T. evansi</i>	<i>T. urticae</i> _Field	<i>T. urticae</i> _R (TuaDV)
<i>T. evansi</i>		(83.3; 86.0)%	(83.2; 86.5)%
<i>T. urticae</i> _Field	(83.3; 86.0)%		(99.9; 99.7)%
<i>T. urticae</i> _R (TuaDV)	(83.2; 86.5)%	(99.9; 99.7)%	

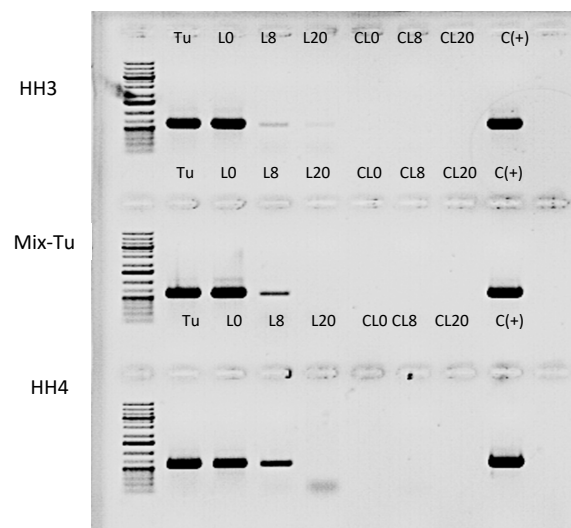
To better assess the prevalence and origin of the TuaDV, we partially sequenced ORF2 from 10 other spider mite samples that originated from natural and rearing populations worldwide. The results also revealed high identity scores between all of the populations, with the highest being within species (100% identity at the nucleotidic level) independently of their origin, i.e., from natural populations and rearings in Portugal.

Altogether, these results suggest that TuaDV is prevalent in natural populations of *Tetranychus* in Portugal and it displays a genetic diversity that likely correlates with the mite phylogeny. The poor genetic diversity that exists between the *T. urticae* laboratory and natural populations, with the exception of the Nice population, is striking and it suggests that infections of laboratory populations occurred in Portugal from wild populations and then spread between rearing facilities worldwide with material exchanges between laboratories. A broader geographical sampling of natural populations, including more *Tetranychus* species, needs to be performed in order to test this hypothesis and evaluate the diversity of TuaDV worldwide.

### 3.5. TuaDV Can be Vertically Transmitted and Circulate in Planta

In the search for the TuaDV route(s) of transmission, first we assayed TuaDV prevalence in individual mites using PCR. We found that 90% (18/20) of the individuals from the population HH4 tested were positive for TuaDV. Due to this high prevalence and the absence of an uninfected population, we did not test for horizontal transmission. We tested vertical transmission (from mothers to offspring), from 10 females from the HH4 population that all laid eggs and then tested 10 offspring per positive female. We found that seven females were positive to TuaDV, out of which four females had positive offspring, with levels of vertical transmission that greatly differed between females. Mean levels of vertical transmission was estimated at 28% ( $\pm 0.9$  SE), and it was highly variable between females with 8/10 for one female and 5/10 for the 3 others (versus 0/10 for females where no vertical transmission was detected). Such variability between offspring probably revealed quantitative variability in the vertical transmission of TuaDV. Although, we cannot exclude to have underestimated the vertical transmission due to the limit of the detection method.

Fecal-oral contamination is a common route of horizontal transmission. Interestingly, an aphid densovirus (MpDV) has been shown to be injected into plants by aphids and circulate *in planta*, which thus participate to its horizontal transmission to conspecifics [21]. Unlike aphids, mites used their stylets to pierce leaf mesophyll cells, where they then inject saliva and suck the cell cytoplasm. Mites are not considered to be vectors, although their feeding behavior injures their host, which can cause the transmission of viruses, in a way that is similar the mite *Varroa destructor* transmit viruses to bees [16]. To test whether TuaDV could be transmitted to the plant during mite feeding, we set up an infestation assay on bean plants (see method section) and assayed virus transmission to plant leaves by PCR. The results obtained from all pairs of primers showed that the leaves on which mites were reared on (L0) were all positive for the three populations tested. This result showed that infected mites have contaminated leaf surfaces, which further suggests that feces and/or saliva are a source of virus (Figure 9). Interestingly, the distant leaves (L8 and L20) were also positive, although the levels decreased with the distance from the primary leaf (L0), suggesting that the virus could circulate within the plant without replicating. Control leaves from the non-infested plants (no mites) and negative controls with no template were negative. We next wanted to quantify the likelihood of TuaDV transmission from the mite to the plant and between the leaves of different distances. Quantification of TuaDV from mites and leaves was performed using a pair of specific primers in ORF2 and a standard curve that was established from serial dilutions of a construct containing the *Junonia coenia* ambidensovirus genome as a proxy for TuaDV genome quantification. Negative controls (leaves with not mites) displayed Ct values that were similar to negative controls with no template. We estimated that the load of TuaDV per mite was  $7.10^6$  veg ( $\pm 8.10^5$  SE), and we found  $2.10^7$  veg ( $\pm 4.10^6$  SE) on L0, which represents the viral shedding from 50 mites for five days. Thus, viral shedding per mite with saliva and feces was estimated at  $9.10^4$  veg ( $\pm 1.10^4$  SE)/day, which represents 1.5% ( $\pm 0.37$  SE) of the total amount of virus estimated per mite. Concerning leaves at 8 (L8) and 20 cm (L20) distance from L0, we found  $2.10^4$  veg ( $\pm 7.10^3$  SE) and 3500 veg/leaf ( $\pm 3100$  SE) in L8 and L20, respectively, with an increasing variation between experiments when increasing the distance from L0. Based on these results, we estimated that 0.08 % ( $\pm 0.02$  SE) of TuaDV contaminating L0 was inoculated into the plant vascular system and recovered in L8, and representing 0.017% ( $\pm 0.016$  SE) in L20. Levels of virus declined with distance from the infested leaf ( $F_{2,10} = 124.08$ ,  $p < 0.0001$ ), but there was no different between the different populations ( $F_{2,3} = 3.28$ ,  $p = 0.1757$ ).



**Figure 9.** Transmission of TuaDV in plants. PCR from total DNA (100ng) extracted from mites, their isolated feeding leaf (L0) and two non-infested leaves at 8 and 20 cm distance (L8 and L20 respectively). Leaves from non-infested plants (CL0, CL8, and CL20) were used as negative controls, while mites were used as positive controls. Three independent F-populations were assayed (HH3, Mix-Tu, and HH4).

Altogether, these results showed that *T. urticae* could vertically and horizontally transmit TuaDV. Vertical transmission varies greatly between females, suggesting an important variation in the amount of virus that can be transmitted. Horizontal transmission occurred through the contamination of food (i.e., plant tissues) from virus shedding with feces and/or saliva. In addition, mites inoculate TuaDV to the plant vascular system, which then can probably circulate from cell-to-cell through plasmodesmata because of its small size (~20nm), to systemically reach distant leaves.

#### 4. Discussion

In the work that is presented here, we used a viral metagenomic approach to explore the diversity of viruses that are associated with the spider mite *Tetranychus urticae*, a small arthropod that has long been neglected by virologists. We showed that several viruses simultaneously circulate in mites, including a panel of eight putative virus species that belong to small, non-enveloped viruses with a RNA genome, and one ssDNA virus with a predicted ambisense genomic organization and belonging to the *Densovirinae* sub-family [34]. This novel densovirus displays two uncommon features among ambidensovirus, i.e., the most compact genome identified so far, all densovirus included, and the absence of the PLA2 motif in the major capsid protein.

An unexpected result was that two mite laboratory populations with different origins shared a set of five viruses, representing 5 taxa. If we independently consider these viral taxa replicate, then the probability is low that five phylogenetically distant viruses were found together in two independent populations (although composed of the same host species) just by chance. Thus, the most parsimonious explanation is to consider that a cross contamination occurred between laboratories, especially when considering the absence of diversity between the common virus genotypes. One possibility is that the Almeria population of *T. cinnabarinus*, which was introduced three years ago in the Portuguese rearing, was at the origin of the contamination of the F population, as it was also maintained in Montpellier laboratory at the time the F-population was sampled.

A second unexpected result was to find that three years after we identified TuaDV (which represents more than 70 mites generations), nearly identical virus genotypes were still found to be circulating in both rearings, while considering the expected occurrence of sequence change in viruses due to error-prone replication (particularly high for RNA and ssDNA viruses [37]). One possibility could be that sequence change of virus genomes was slower than expected. In support of that hypothesis, we found that TuaDV genomic sequence from rearings differed by a few substitutions (3 to 5) from the sequences that were found recently in the natural populations of *T. urticae* and *T. cinnabarinus* sampled in Portugal and in different rearing facilities worldwide. Although the level of relatedness was high, we did observe slight differences, indicating that TuaDVs from the Netherlands (Santpoort) and Belgium (Ghent) were more closely related than they were to the samples from Portugal, and the TuaDV from Greece (Crete) was almost identical to the TuaDV from Portugal (99 to 100%). When considering that the Greek population is actually composed of a mix of populations from different origins, this suggests that a cross contamination probably occurred between the laboratories. More interestingly, we found that the TuaDV from the French population of *T. urticae* that originates from Nice-Valbonne was more distantly related to the other strains (88% with the ORF2 of TuaDV\_R versus 94 to 100% that are shared by the other *T. urticae* samples). This population was collected in the Nice area 3–5 years ago and then reared there since then, without any exchange with the other facilities of this study (M. Ferrero, personal communication). This suggests that several strains of TuaDV probably circulate in natural populations. We speculate that TuaDV might be under stabilizing selection when circulating in *T. urticae* and *T. cinnabarinus*, that is that most new mutations would be deleterious and thus be quickly removed from the viral populations [38,39]. Moreover, the high identity score of the TuaDV genome found in two "sister" species, *T. cinnabarinus* and *T. urticae*, and the distance with TuaDV from *T. evansi* further supports that divergence and evolution have co-occurred with their hosts' evolution.

The third unexpected result was the discovery in mites viromes of contigs corresponding to three unclassified picorna-like related to Aphis glycines virus 1, one of them sharing 99% nucleotidic identity with the virus that was discovered in aphids at the University of Illinois and published in Genbank in 2013 (#KF360262), the two others being more distantly related (Table 1). Like for TuaDV, such sequence conservation was striking when considering the time and the geographic distance between sampling but also the phylogenetic distance between aphids and mites. The constraints of the host immune system are among mechanisms conditioning virus changes for adaptation. Interestingly, mites and aphids share a genome that has lost canonical components of the immune system [40]. We speculate that such feature may contribute to provide tolerant environments to viruses. Although, the mechanism remains to be addressed.

Most of the virus genotypes that were found in this study were classified within arthropod-infecting taxa and/or their closest phylogenetic taxa were associated with arthropods. The phylogenetic trees that we obtained for each virus highlighted the poor knowledge that we have on arthropod viruses [41–43], with viruses often being located at the base of the phylogenetic trees of the largest taxa, such as for *Parvoviridae* and *Picornavirales*, or grouped into poorly documented viral families, which indicates that a wealth of viruses from these groups remain to be discovered. The set of viruses that we found in this study mostly corresponds to small, non-enveloped viruses, which contrasts with sequences that are found in nucleotidic databases and corresponding to large viruses. The under-representation of large viruses in viromes prepared with the VANA method may result from a technical bias. Indeed, the filtration step that is used in this method to prepare viral particles could eliminate large viruses [44]. This observation thus pinpoints the need to combine approaches to get an exhaustive view of virus diversity that is associated with arthropods.

In natural or experimental populations, mite density can vary due to variation in mortality levels, which can occur without the knowledge of the causal agent(s)/conditions. Whether and how TuaDV affects the mites phenotype remains to be experimentally addressed. As the level of vertical transmission is relatively low, the establishment of TuaDV-free lines to investigate the role of TuaDV by experimental infection is made possible. Interestingly, persistent densoviruses can be found in aphids where they have been shown to confer protection against secondary infections by other pathogens, including viruses and bacteria [45,46]. Unlike those observed for densoviruses that were discovered in aphids, we could not find evidence of virus genomic integration in the mite genome, suggesting that the TuaDV could correspond to an extant virus [4,47]. However, we cannot exclude that viral sequences were cleaned off during *T. urticae* genome assembly [8].

It is becoming clear that both multi-infections and persistent viruses should be considered in order to better understand the phenotypic outcome of infections and the pathogenicity of specific strains that can change their life cycle, from pathogenic to persistent [48]. Densovirus persistent infections are likely common in aphids, and several densoviruses have been described, with one being transmitted through the plant [21,45,47]. Although we need to improve the viral loads quantification method, our results suggest that potentially large amount of TuaDV can contaminate plants when considering the high densities that mites colonies can reach. In a previous study, we described a new densovirus found from viral metagenomics from sea barley (*Hordeum marinum*), probably contaminating/circulating in the plant as well, since the virus was recovered from different leaves and was not associated with any arthropods host [49]. We hypothesized that plants participate to TuaDV horizontal transmission by concentrating viral particles on leaves, thus favoring the spreading of infection in mite populations. Furthermore, the systemic inoculation of viral particles in the plant vascular system could also provide the virus a protective niche against abiotic factors (particularly ultraviolet radiation from the sun) that can be detrimental for viral particles. Interestingly, our results also suggest that the plant might also mediate infections with a panel of arthropod-infecting viruses.

We hypothesized that virus abundance and/or transmission routes might have been selected for virus communities that occur as a unit within their host populations. Indeed, we find that five viruses co-occur and can be maintained in two separate laboratory populations for three years, even if the source is due to a recent contamination. Our results suggest that transmission can occur by vertical and horizontal routes, including via the host plant, leading to a wide prevalence in their host populations. This begs the question as to whether all of these viruses can occur in a multiple infection within single individual hosts or if they never co-infect but separately circulate in mite populations (i.e., one virus preventing the infection by others). Further investigation is required to evaluate whether these viruses interact in a synergic, antagonist, or neutral manner.

In addition to sharing common ecosystems and similar immunologic environments, mites and aphids also share diverse endosymbionts that may provide protection in compensation [40,50]. Whether and how viruses could also provide protection against immunological stress remain to be addressed. Experimental mite colonies could provide a powerful system to combine descriptive and manipulative experiments to test for virus pathogenicity in individual hosts; and their dynamics (prevalence and persistence), and evolution in host populations.

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/1999-4915/11/3/233/s1>, Table S1: *Tetranychus* species screened for the presence of TuaDV, Table S2: Sequences of TuaDV primers.

**Author Contributions:** Conceived and designed the experiments S.F., A.B.D., P.R., R.F., M.O. Performed the experiments: S.F., D.M., L.R.R., S.L., D.F., I.S., C.D., E.F. Sampling: L.R.R., M.F., I.S. Analyzed the data: S.F., A.B.D., P.R., D.F., L.R.R. Contributed reagents/materials/analysis/tools: A.B.D., L.R.R., D.F., S.L., C.D., I.S. Wrote and review the paper: S.F., A.B.D., L.R.R., P.R., R.F., M.O.

**Funding:** This work was supported by a grant from the Plant Health and Environment INRA division (SPE) and a H2020-EU VIROPLANT grant (773567) to M.O. It was also supported by a PHC-PESSOA grant (38014YC) to A.D.B. and S. Magalhães.

**Acknowledgments:** S. F. was supported by a scholarship from the National Institute of Agronomical Research (INRA) and from the University of Montpellier (UM). This work is an ISEM contribution number 2019-042. We warmly thank T. Van Leeuwen (Ghent University, Belgium) for providing the LondonS and EtoxR lines, M. Ferrero (Bioline Agrosociences, France) for providing the *T. urticae* Nice-Valbonne strain, and J. M. Alba (University of Amsterdam, NL) for the Ghent strain. We also warmly thank M. van Munster, M. Uzeit, V. Brault and S. Blanc for all the insightful discussions and D. Sarkozy for its critical reading of the manuscript.

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

## References

1. Junglen, S.; Drosten, C. Virus discovery and recent insights into virus diversity in arthropods. *Curr. Opin. Microbiol.* **2013**, *16*, 507–513. [[CrossRef](#)] [[PubMed](#)]
2. Bovo, S.; Mazzoni, G.; Ribani, A.; Utzeri, V.J.; Bertolini, F.; Schiavo, G.; Fontanesi, L. A viral metagenomic approach on a nonmetagenomic experiment: Mining next generation sequencing datasets from pig DNA identified several porcine parvoviruses for a retrospective evaluation of viral infections. *PLoS ONE* **2017**, *12*, 1–17. [[CrossRef](#)] [[PubMed](#)]
3. Liu, H.; Fu, Y.; Xie, J.; Cheng, J.; Ghabrial, S.A.; Li, G.; Peng, Y.; Yi, X.; Jiang, D. Widespread Endogenization of Densoviruses and Parvoviruses in Animal and Human Genomes. *J. Virol.* **2011**, *85*, 9863–9876. [[CrossRef](#)] [[PubMed](#)]
4. François, S.; Filloux, D.; Roumagnac, P.; Bigot, D.; Gayral, P.; Martin, D.P.; Froissart, R.; Ogliastro, M. Discovery of parvovirus-related sequences in an unexpected broad range of animals. *Sci. Rep.* **2016**, *6*, 1–13. [[CrossRef](#)] [[PubMed](#)]
5. Lister, B.C.; Garcia, A. Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proc. Natl. Acad. Sci.* **2018**, *115*, E10397–E10406. [[CrossRef](#)] [[PubMed](#)]
6. Walter, D.E.; Proctor, H.C. *Mites: Ecology, Evolution & Behaviour*; Springer: Dordrecht, The Netherlands, 2013.
7. Jeppson, L.R.; Keifer, H.H.; Baker, E.W. *Mites Injurious to Economic Plants*; Univ of California Press: Berkeley, CA, USA, 1975.

8. Grbić, M.; Van Leeuwen, T.; Clark, R.M.; Rombauts, S.; Rouzé, P.; Grbić, V.; Osborne, E.J.; Dermauw, W.; Ngoc, P.C.T.; Ortego, F.; et al. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* **2011**, *479*, 487–492. [[CrossRef](#)] [[PubMed](#)]
9. Van Leeuwen, T.; Vontas, J.; Tsagkarakou, A.; Dermauw, W.; Tirry, L. Acaricide resistance mechanisms in the two-spotted spider mite *Tetranychus urticae* and other important Acari: A review. *Insect Biochem. Mol. Biol.* **2010**, *40*, 563–572. [[CrossRef](#)] [[PubMed](#)]
10. Beavers, J.B.; Reed, D.K. Susceptibility of seven tetranychids to the nonoccluded virus of the citrus red mite and the correlation of the carmine spider mite as a vector. *J. Invertebr. Pathol.* **1972**, *20*, 279–283. [[CrossRef](#)]
11. Reed, D.K.; Desjardins, P.R. Morphology of a non-occluded virus isolated from citrus red mite, *Panonychus citri*. *Experientia* **1982**, *38*, 468–469. [[CrossRef](#)]
12. Tokarz, R.; Williams, S.H.; Sameroff, S.; Sanchez Leon, M.; Jain, K.; Lipkin, W.I. Virome analysis of *Amblyomma americanum*, *Dermacentor variabilis*, and *Ixodes scapularis* ticks reveals novel highly divergent vertebrate and invertebrate viruses. *J. Virol.* **2014**, *88*, 11480–11492. [[CrossRef](#)] [[PubMed](#)]
13. Moutailler, S.; Popovici, I.; Devillers, E.; Vayssier-Taussat, M.; Eloit, M. Diversity of viruses in *Ixodes ricinus*, and characterization of a neurotropic strain of Eyach virus. *New Microbes New Infect.* **2016**, *11*, 71–81. [[CrossRef](#)] [[PubMed](#)]
14. Sakamoto, J.M.; Ng, T.F.F.; Suzuki, Y.; Tsujimoto, H.; Deng, X.; Delwart, E.; Rasgon, J.L. Bunyaviruses are common in male and female *Ixodes scapularis* ticks in central Pennsylvania. *PeerJ* **2016**, *4*, e2324. [[CrossRef](#)] [[PubMed](#)]
15. Xia, H.; Hu, C.; Zhang, D.; Tang, S.; Zhang, Z.; Kou, Z.; Fan, Z.; Bente, D.; Zeng, C.; Li, T. Metagenomic profile of the viral communities in rhipicephalus spp. ticks from Yunnan, China. *PLoS ONE* **2015**, *10*, 1–16. [[CrossRef](#)] [[PubMed](#)]
16. Levin, S.; Sela, N.; Chejanovsky, N. Two novel viruses associated with the *Apis mellifera* pathogenic mite *Varroa destructor*. *Sci. Rep.* **2016**, *6*, 1–9. [[CrossRef](#)] [[PubMed](#)]
17. Nazzi, F.; Brown, S.P.; Annoscia, D.; Del Piccolo, F.; Di Prisco, G.; Varricchio, P.; Della Vedova, G.; Cattonaro, F.; Caprio, E.; Pennacchio, F. Synergistic parasite-pathogen interactions mediated by host immunity can drive the collapse of honeybee colonies. *PLoS Pathog.* **2012**, *8*, e1002735. [[CrossRef](#)] [[PubMed](#)]
18. Palanga, E.; Filloux, D.; Martin, D.P.; Fernandez, E.; Bouda, Z.; Gargani, D.; Ferdinand, R.; Zabre, J.; Neya, B.; Sawadogo, M.; et al. Metagenomic-Based Screening and Molecular Characterization of Cowpea- Infecting Viruses in Burkina Faso. *PLoS ONE* **2016**, *11*, e0165188. [[CrossRef](#)] [[PubMed](#)]
19. Matsuda, T.; Morishita, M.; Hinomoto, N.; Gotoh, T. Phylogenetic Analysis of the Spider Mite Sub-Family Tetranychinae (Acari: Tetranychidae) Based on the Mitochondrial COI Gene and the 18S and the 59 End of the 28S rRNA Genes Indicates That Several Genera Are Polyphyletic. *PLoS ONE* **2014**, *9*, e108672. [[CrossRef](#)] [[PubMed](#)]
20. Matsuda, T.; Kozaki, T.; Ishii, K.; Gotoh, T. Phylogeny of the spider mite sub-family Tetranychinae (Acari: Tetranychidae) inferred from RNA-Seq data. *PLoS ONE* **2018**, *13*, e0203136. [[CrossRef](#)] [[PubMed](#)]
21. Van Munster, M.; Janssen, A.; Clériveret, A.; Van Den Heuvel, J. Can plants use an entomopathogenic virus as a defense against herbivores? *Oecologia* **2005**, *143*, 396–401. [[CrossRef](#)] [[PubMed](#)]
22. Uesugi, R.; Goka, K.; Osakabe, M. Genetic basis of resistances to chlorfenapyr and etoxazole in the two-spotted spider mite (Acari: Tetranychidae). *J. Econ. Entomol.* **2002**, *95*, 1267–1274. [[CrossRef](#)] [[PubMed](#)]
23. Wu, S.; Manber, U. Agrep—A fast approximate pattern-matching tool. In *Conference Proceedings: USENIX Winter 1992 Technical Conference*; USENIX Association: Berkeley, CA, USA, 1992; pp. 153–162.
24. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMB* **2011**, *17*, 10–12. [[CrossRef](#)]
25. Zerbino, D.R.; Birney, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* **2008**, *18*, 821–829. [[CrossRef](#)] [[PubMed](#)]
26. Langmead, B. Aligning short sequencing reads with Bowtie. *Curr. Protoc. Bioinforma.* **2010**, *32*, 11–17.
27. Altschul, S.F.; Gish, W.; Miller, W.; Myers, E.W.; Lipman, D.J. Basic local alignment search tool. *J. Mol. Biol.* **1990**, *215*, 403–410. [[CrossRef](#)]
28. Edgar, R.C. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **2004**, *32*, 1792–1797. [[CrossRef](#)] [[PubMed](#)]

29. Kearse, M.; Moir, R.; Wilson, A.; Stones-Havas, S.; Cheung, M.; Sturrock, S.; Buxton, S.; Cooper, A.; Markowitz, S.; Duran, C.; et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **2012**, *28*, 1647–1649. [[CrossRef](#)] [[PubMed](#)]
30. Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J.F.; Guindon, S.; Lefort, V.; Lescot, M.; et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **2008**, *36*, 465–469. [[CrossRef](#)] [[PubMed](#)]
31. Guindon, S.; Dufayard, J.-F.; Lefort, V.; Anisimova, M.; Hordijk, W.; Gascuel, O. New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **2010**, *59*, 307–321. [[CrossRef](#)] [[PubMed](#)]
32. Abascal, F.; Zardoya, R.; Posada, D. ProtTest: Selection of best-fit models of protein evolution. *Bioinformatics* **2005**, *21*, 2104–2105. [[CrossRef](#)] [[PubMed](#)]
33. Dolja, V.V.; Koonin, E.V. Metagenomics reshapes the concepts of RNA virus evolution by revealing extensive horizontal virus transfer. *Virus Res.* **2018**, *244*, 36–52. [[CrossRef](#)] [[PubMed](#)]
34. Cotmore, S.F.; Agbandje-McKenna, M.; Chiorini, J.A.; Mukha, D.V.; Pintel, D.J.; Qiu, J.; Soderlund-Venermo, M.; Tattersall, P.; Tijssen, P.; Gatherer, D.; et al. The family Parvoviridae. *Arch. Virol.* **2014**, *159*, 1239–1247. [[CrossRef](#)] [[PubMed](#)]
35. Cotmore, S.F.; Tattersall, P. Parvoviruses: Small Does Not Mean Simple. *Annu. Rev. Virol.* **2014**, *1*, 517–537. [[CrossRef](#)] [[PubMed](#)]
36. Conceição-Neto, N.; Godinho, R.; Álvares, F.; Yinda, C.K.; Deboutte, W.; Zeller, M.; Laenen, L.; Heylen, E.; Roque, S.; Petrucci-Fonseca, F.; et al. Viral gut metagenomics of sympatric wild and domestic canids, and monitoring of viruses: Insights from an endangered wolf population. *Ecol. Evol.* **2017**, *7*, 4135–4146. [[CrossRef](#)] [[PubMed](#)]
37. Duffy, S.; Shackelton, L.A.; Holmes, E.C. Rates of evolutionary change in viruses: Patterns and determinants. *Nat. Rev. Genet.* **2008**, *9*, 267–276. [[CrossRef](#)] [[PubMed](#)]
38. Duffy, S. Why are RNA virus mutation rates so damn high? *PLoS Biol.* **2018**, *16*, 1–6. [[CrossRef](#)] [[PubMed](#)]
39. Simmonds, P.; Aiweesakun, P.; Katzourakis, A. Prisoners of war—host adaptation and its constraints on virus evolution. *Nat. Rev. Microbiol.* **2018**. [[CrossRef](#)] [[PubMed](#)]
40. Santos-Matos, G.; Wybouw, N.; Martins, N.E.; Zélé, F.; Riga, M.; Leitão, A.B.; Vontas, J.; Grbić, M.; Van Leeuwen, T.; Magalhães, S.; et al. Tetranychus urticae mites do not mount an induced immune response against bacteria. *Proceedings. Biol. Sci.* **2017**, *284*, 20170401. [[CrossRef](#)] [[PubMed](#)]
41. Marklewitz, M.; Zirkel, F.; Kurth, A.; Drosten, C.; Junglen, S. Evolutionary and phenotypic analysis of live virus isolates suggests arthropod origin of a pathogenic RNA virus family. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 7536–7541. [[CrossRef](#)] [[PubMed](#)]
42. Li, C.-X.; Shi, M.; Tian, J.-H.; Lin, X.-D.; Kang, Y.-J.; Chen, L.-J.; Qin, X.-C.; Xu, J.; Holmes, E.C.; Zhang, Y.-Z. Unprecedented genomic diversity of RNA viruses in arthropods reveals the ancestry of negative-sense RNA viruses. *Elife* **2015**, *4*, 1–26. [[CrossRef](#)] [[PubMed](#)]
43. Shi, M.; Lin, X.; Vasilakis, N.; Tian, J.; Li, C.; Chen, L.; Eastwood, G.; Diao, X. Divergent Viruses Discovered in Arthropods and Vertebrates Revise the Evolutionary History of the Flaviviridae and Related Viruses. *J. Virol.* **2016**, *90*, 659–669. [[CrossRef](#)] [[PubMed](#)]
44. Halary, S.; Temmam, S.; Raoult, D.; Desnues, C. Viral metagenomics: are we missing the giants? *Curr. Opin. Microbiol.* **2016**, *31*, 34–43. [[CrossRef](#)] [[PubMed](#)]
45. Van Munster, M.; Dullemans, A.M.; Verbeek, M.; Van Den Heuvel, J.F.J.M.; Reinbold, C.; Brault, V.; Clérvet, A.; Van Der Wilk, F. Characterization of a new densovirus infecting the green peach aphid *Myzus persicae*. *J. Invertebr. Pathol.* **2003**, *84*, 6–14. [[CrossRef](#)]
46. Xu, P.; Liu, Y.; Graham, R.I.; Wilson, K.; Wu, K. Densovirus Is a Mutualistic Symbiont of a Global Crop Pest (*Helicoverpa armigera*) and Protects against a Baculovirus and Bt Biopesticide. *PLoS Pathog.* **2014**, *10*, 2–12. [[CrossRef](#)] [[PubMed](#)]
47. Clavijo, G.; Van Munster, M.; Monsion, B.; Bochet, N.; Brault, V. Transcription of densovirus endogenous sequences in the *Myzus persicae* genome. *J. Gen. Virol.* **2016**, *97*, 1000–1009. [[CrossRef](#)] [[PubMed](#)]
48. Alizon, S.; de Roode, J.C.; Michalakis, Y. Multiple infections and the evolution of virulence. *Ecol. Lett.* **2013**, *16*, 556–567. [[CrossRef](#)] [[PubMed](#)]



49. Francois, S.; Bernardo, P.; Filloux, D.; Roumagnac, P.; Yaverkovski, N.; Froissart, R.; Ogliastro, M. A Novel Itera-Like Denguevirus Isolated by Viral Metagenomics from the Sea Barley *Hordeum marinum*. *Genome Announc.* **2014**, *2*, e01196-14. [[CrossRef](#)] [[PubMed](#)]
50. Zélé, F.; Weill, M.; Magalhães, S. Identification of spider-mite species and their endosymbionts using multiplex PCR. *Exp. Appl. Acarol.* **2018**, *74*, 123–138. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).