Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi

- L. Nerva[a, c], M. Ciuffo[a], M. Vallino[a], P. Margaria[b], G.C. Varese[c], G. Gnavi[c], M. Turina[a,]

Show more

doi:10.1016/j.virusres.2015.10.028

Highlights

• First report of mycoviruses isolated from fungi from marine environment.
• Survey of mycoviruses using multiple approaches for both DNA and RNA genomes.
• RNAseq analysis is superior to sRNA for *de novo* assembly of mycoviruses.
• Twelve new virus species were characterized molecularly.
• Expression of a viral RNA from an endogenized cDNA segment.

Abstract

The number of reported mycoviruses is increasing exponentially due to the current ability to detect mycoviruses using next-generation sequencing (NGS) approaches, with a large number of viral genomes built *in-silico* using data from fungal transcriptome projects. We decided to screen a collection of fungi originating from a specific marine environment (associated with the seagrass *Posidonia oceanica*) for the presence of mycoviruses: our findings reveal a wealth of diversity among these symbionts and this complexity will require further studies to address their specific role in this ecological niche. In specific, we identified twelve new virus species belonging to nine distinct lineages: they are members of megabirnavirus, totivirus, chrysovirus, partitivirus and five still undefined clades. We showed evidence of an endogenized virus ORF, and evidence of accumulation of dsRNA from metaviridae retroviral elements. We applied different techniques for detecting the presence of mycoviruses including (i) dsRNA extraction and cDNA cloning, (ii) small and total RNA sequencing through NGS techniques, (iii) rolling circle amplification (RCA) and total DNA extraction analyses, (iv) virus purifications and electron microscopy. We tried also to critically evaluate the intrinsic value and limitations of each of these techniques. Based on the samples we could compare directly, RNAseq analysis is superior to sRNA for *de novo* assembly of mycoviruses. To our knowledge this is the first report on the virome of fungi isolated from marine environment.

The GenBank/eMBL/DDBJ accession numbers of the sequences reported in this paper are: KT601099–KT601110; KT601114–KT601120; KT592305; KT950836–KT950841.

Keywords

- Mycovirus;
- Taxonomy;
- *Posidonia oceanica*

---

1. Introduction

More than half a century has passed since the discovery of the first mycovirus, and the interest in these symbionts has grown steadily, particularly in the last decade as many new viral species are reported every year. Although initially reported as diseases of cultivated mushroom (Gandy, 1960 and Hollings, 1962), to date the majority of the described mycoviruses are studied for their ability to change the virulence (if the host is a pathogen) or more generally the phenotype of their hosts. Another interest in mycovirus research comes from early work on the use of fungi in biotechnology (yeast fermentation, antibiotic production, and interferon purification) and their possible interference with fermentation processes or their added value given by the presence of double-stranded RNA (dsRNA) in culture filtrates (Banks et al., 1968 and Woods and Bevan, 1968). Less is known about the ecological role of these fungal symbionts in nature and how they can change or drive the evolution of their hosts. From an ecological point of view, an interest for these symbionts has grown after the description of a three-way symbiotic system where a mycovirus harbored in the endophytic fungus *Curvularia protuberata* is able to confer thermal tolerance to *Dichanthelium lanuginosum*, a grass plant species (Marquez et al., 2007). Mycoviruses are widely distributed in all the major fungal taxonomic groups (Goeker et al., 2011), including saprophytes and symbionts. The majority of mycoviruses present in databases have dsRNA genomes and are transmitted among different individuals through cellular fusion (anastomosis) without an extracellular phase. However, in the last few years, mycoviruses possessing single-stranded RNA (ssRNA) positive (+) and negative (−) sense genomes as well as a circular ssDNA genome have been reported (Yu et al., 2010, Yu et al., 2013 and Du et al., 2014). In particular the characterized mycovirus with ssDNA genome seems to be acquired externally without the occurrence of fungal anastomoses (Du et al., 2014). The classic method for the detection of mycoviruses consists of the purification of dsRNA using CF11 cellulose: this easy and convenient technique is able to detect mostly virus with dsRNA genomes (and to some extent also dsRNA replication intermediates from ssRNA virus genomes) and this is likely the reason why the majority of the mycoviruses described so far have a dsRNA genome. Another common way to detect mycoviruses is through virus particle purification using differential centrifugation protocols combined with observation by transmission electron microscopy (TEM). This approach allows for the detection of viruses with different kinds of genomes, but requires traditional virological expertise and expensive equipment, which are not always available in fungal and molecular biology laboratories. Furthermore, this technique is unsatisfactory for “naked” viruses, a fairly common feature for mycoviruses. Therefore, identification of new mycovirus taxonomic groups requires new detection approaches.

A new impulse to mycovirus research comes from the wide availability of next-generation sequencing (NGS) protocols. As demonstrated in a large number of scientific works, NGS techniques are able to detect the presence of viral sequences in samples where the viral titer is very low. An example of the usefulness of these techniques comes from two previous works (Al Rwahnih et al., 2011 and Espach, 2013), where the authors analyzed a plant total RNA extract using NGS techniques and not only were they able to detect plant viruses but they also detected the complex virome of the fungal endophytes. NGS applied to total RNA extracts allows for the

detection of all the different viral taxa because, independently from the DNA or RNA genome, all viruses must produce mRNA during their replication cycle.

Another way to detect viral sequences is to apply the NGS technique to sequencing small RNA (sRNA) libraries. This idea comes from the work of Kreuze et al. 2009 where the authors showed the possibility to assemble new and unknown viral sequences from sRNA libraries obtained from plants. To our knowledge, such an approach has only very recently been used for *de novo* assembly of viral genomes from fungi (Vainio et al., 2015).

After the description of a circular ssDNA mycoviruses with a strong impact on host virulence (Du et al., 2014), a relatively new branch of mycovirology (the one dealing with DNA genome viruses) seems to be born. A general protocol used for environmental samples and for higher eukaryotes to specifically detect ssDNA virus (the "circuloma" approach) relies on the aid of a new technique called rolling circle amplification (RCA) (Ali et al., 2014). RCA has the potential to detect also plasmid DNA and since the 1970s it has been shown that fungi can host DNA plasmids in their cytoplasm or in their mitochondria (Stahl et al., 1978). These plasmids seem to confer beneficial phenotypes to the fungi such as hypovirulence (Monteiro-Vitorello et al., 2000), drug resistance or resistance to abiotic stress (Griffiths, 1995).

The marine environment is characterized by the presence of high salinity and a recent study showed that the number of viruses in the sea is at least ten times higher than the number of any other microorganism (Breitbart, 2012 and Suttle, 2007). Specifically, most marine viruses are cyanophages and have been shown to be major players in biogeochemical cycles and drivers of evolution of their algal hosts (Brussaard et al., 2008 and Fuhrman, 1999) by influencing microbial population size through their lytic capacity, altering their metabolic output and providing an immensely diverse pool of genetic material available for horizontal gene transfer (Brum et al., 2015). Till now, no virus associated to marine fungi has been reported, probably because the presence and ecological role of fungi in marine environment has been neglected till few years ago. Hence, we decided to investigate viruses infecting fungi in marine environment, so far a much overlooked aspect of ocean microbiology (Richards et al., 2012).

Water samples are not fully representative of the marine environment, and fungi could actually play a major role associated to algae and marine plants. We decided, therefore, to screen for the presence of virus and plasmid symbionts in a collection of marine fungi isolated from the seagrass *Posidonia oceanica* (Panno et al., 2013 and Gnavi et al., 2014). *P. oceanica* is a protected species of utmost importance for sea ecology as exemplified by specific European legislation addressing this issue (EEC 92/43, 1992). Characterizing the microbiological niche associated with this plant could help in monitoring and managing the health of our sea environment. We selected a subset of 91 fungal isolates representative of the biological diversity of the fungi isolated from *P. oceanica* and we applied the different techniques for mycovirus detection described above to compare the positive and negative aspects of each methodological approach.

2. Materials and methods

2.1. Fungal isolates

The fungal isolates used in this work come from a collection of marine fungi of the Mycoteca Universitatis Taurinensis (MUT) (Supplementary Table 1). The analyzed strains were isolated from different districts of the seagrass *P. oceanica* (Panno et al., 2013). All the fungi were grown in liquid cultures at 24 °C in malt extract broth with 3% of sea salts (Sigma–Aldrich, Saint Louis, MO, USA) producing a good amount of mycelia in 4 days, with the exception of *Wallemia sebi* isolate

MUT4935 that grew better in malt yeast extracts supplemented with 40% sucrose (MY40) at 24 °C and was harvested after 5 days.

2.2. dsRNA purification

Lyophilized mycelium was first ground with iron beads, then 200 mg were transferred to a 15 ml tube with 4 ml of glass beads (0.5 mm diameter) and homogenized for 30 s in a bead beater (FastPrep24, M.P. Biomedicals, Irvine, CA, USA), added with 5 ml of STE 2X (200 mM NaCl, 10 mM Tris–HCl pH7, EDTA 2 mM pH 8) SDS 1%, and 5 ml of phenol (Phenol BioUltra, Sigma–Aldrich, Saint Louis, MO, USA), mixed again for 30 s in a bead beater and then centrifuged for 15 s at 1000 g. Five ml of the supernatant were transferred to a new tube, added with 5 ml of phenol-chlorofom (1:1), vortexed and then centrifuged for 15 min as above. The supernatant was transferred to a new 50 ml tube with 1 g of Whatman CF11 cellulose previously equilibrated with 15% Et-OH in 1X STE. After discarding the supernatant, the CF11 was re-suspended in 20 ml of 15% ethanol in 1X STE and loaded onto a chromatography column, washed with 20 ml of 15% ethanol in 1X STE, finally eluted with 10 ml of STE 1X. Two volumes of ethanol and 5% of 3 M Na-Acetate pH 5.2 were added to the eluate for precipitation. The dsRNA was kept at −80 °C for 30 min, centrifuged at 12,000 × *g* for 15 min, resuspended in 500 µl of nuclease S1 buffer (sodium acetate 40 mM, sodium chloride 300 mM, zinc sulfate 2 mM) and then added with 2 µl of nuclease S1, incubated at 37 °C for 20 min, added with the same volume of phenol-chloroform, centrifuged for 5 min; the supernatant was precipitated at −80 °C as described above. The pellet was recovered and resuspended in 40 µl of water and traces of DNA were removed by adding 1 µl of DNase I (Thermo Fisher Scientific Inc., Waltham, MA, USA) following the manufacturer's protocol. The dsRNA was then analyzed by electrophoresis in 1% agarose gel in 1X TAE and each distinct single or group of dsRNA fragments was purified from agarose gel using the Zymoclean Gel DNA Recovery kit (Zymoresearch, CA, USA) following manufacturer instructions and eluting the samples in a final volume of 15 µl of water.

2.3. Virus particles purification

Three hundred milligrams of lyophilized mycelia were used for the purification and detection of virus particles using classic differential centrifugation protocol. Each sample was homogenized with a bead beater in 15 ml of extraction buffer (0.25 M potassium–phosphate pH 7, 10 mM EDTA, 0.5% thioglycolic acid), 4 ml of glass beads (0.5 mm diameter) and 1 ml of iron beads (1 mm diameter). After centrifugation at 1000 × *g* for 10 min, the supernatant was transferred into a new 50 ml tube, added with 1% Triton X-100, 0.5% propanesulfonate detergent, 10% PEG8000, and 0.1% sodium chloride, stirred for 60 min and then centrifuged for 10 min at 1000 × *g*. The pellet was resuspended into extraction buffer and then centrifuged on a 20% sucrose layer at 180,000 × *g* (Ciuffo et al., 2008). The resulting pellet was resuspended in 300 µl of extraction buffer and stored at −80 °C.

2.4. Electron microscopy

For TEM, samples were allowed to adsorb for few minutes on carbon and formvar-coated grids. After washing with water, the grids were negatively stained with 0.5% uranyl acetate. Observations and image captures were made using a CM 10 electron microscope (Philips, Eindhoven, The Netherlands).

2.5. Complementary DNA (cDNA) libraries

We performed a reverse transcriptase (RT) protocol using the Invitrogen Thermo-Script RT-PCR kit with the following modifications: (i) 5 µl of the gel purified dsRNA as template; (ii) a RT primer with a defined sequence at the 5' end followed by a random sequence at the 3' end (CCTTCGGATCCTCC-N8) at a concentration of 50 µM to prime the reaction; (iii) a denaturation step at 98 °C followed by a quick chilling on ice.

The polymerization of the second strand was done using a DNA Polymerase I with the following condition: 20 µl cDNA reaction from above, 10 µl of DNA Polymerase I buffer 10X, 5 µl BSA, 0.5 µl RNase H, 2 µl DNA Polymerase I, 62.5 µl H_2O. The mix was incubated at 14 °C for 120 min and then the enzymes were inactivated by heating at 75 °C for 10 min. As a last step we added 2 µl of T4 DNA Polymerase, incubated at 37 °C for 10 min and stopped the reaction on ice for 5 min with 10 µl of 200 mM EDTA.

To clean from residual primer we used DNA clean & concentrator™-5 kit (Zymoresearch, CA, USA) following manufacturer's instructions.

Obtained cDNAs were used as template in a classic PCR protocol using an anchored primer with the specific CCTTCGGATCCTCC sequence at the 3' end and a 4 nucleotide tag at the 5' end (Roossinck et al., 2010). The reaction products were separated on a 1% agarose gel, purified with the Zymoclean Gel DNA Recovery kit (Zymoresearch, CA, USA) and then cloned in the pGEM-T easy vector (Promega, Madison, WI, USA). White colonies were screened by PCR using M13 forward and reverse primers. Plasmids with inserts of the expected sizes were purified from *Escherichia coli* using Zyppy™ Plasmid Miniprep Kit (Zymoresearch, CA, USA) and sequenced using the dideoxy chain termination method (Sanger et al., 1977) at BioFab Research (Rome, Italy).

To confirm the viral origin of the obtained sequences, we searched the NCBI database using BLASTx and BLASTn (Altschul et al., 1990). Once cDNA clones of viral origin were identified, the gaps of the cDNA library were covered through PCR amplification with specific oligonucleotides designed on the known sequences as previously described (Turina et al., 2007). The PCR product in this case was sequenced directly after purification on Zymoresearch purification columns as described above.

To determine the 5′ ends of the different viruses, we used Hirzmann's method (Hirzmann et al., 1993). Gel-purified dsRNA was annealed with a reverse specific primer at the putative 5′ end, and after reverse transcription, RNase H was added to release cDNA from the RNA template and the cDNA was purified using DNA clean & concentrator™-5 kit (Zymoresearch, CA, USA). Dideoxy-GTP was used to synthesize a poly G tail using rTdT terminal deoxynucleotidyl transferase (Promega, Madison, WI, USA) following the manufacturer's instructions. The final PCR step was done using oligo-dC (Eco-bam-dC12) with a specific primer for each dsRNA to be amplified. PCR products were cloned and sequenced as above.

The 3′ ends were obtained by 3′ RACE (rapid amplification of cDNA end), after polyadenylation of each genomic RNA with polyA polymerase and ATP (Ambion, TX USA) following manufacturer's protocols. Reverse transcription and PCR were carried out as mentioned previously using oligo dT-V in the RT step, and oligo dT-V with a specific forward primer upstream for each dsRNA in the PCR step.

A further approach used for the 5′ and 3′ RACE was the ligation of a known blocked adaptor to the 3′ end of the dsRNA viral sequences with the aid of a T4 RNA ligase (Promega, Madison, WI, USA) as described previously in detail (Coutts and Livieratos, 2003).

2.6. Northern blot analysis

Total RNA from fungi was prepared using Total Spectrum RNA Reagent (Sigma–Aldrich, Saint Louis, MO, USA) as suggested by the manufacturer. RNA samples were separated in denaturing conditions (glyoxal method) as detailed, using HEPES-EDTA buffer (Sambrook et al., 1989). Hybridizations were performed using a radio-labeled RNA probe prepared from linearized purified plasmids containing cDNA clones through T7 transcription using the Maxiscript T7 kit reagents (Thermo Fisher Scientific Inc., Waltham, MA, USA), following manufacturer's protocols. A list of the probes used with corresponding viral region is displayed in Supplementary Table 2.

2.7. Total DNA extraction

A classic approach for the detection of DNA viruses was used: DNA was extracted with phenol/chloroform and 5 μg of total DNA, treated with RNAase A, were separated on a 0.8 % agarose gel (Turina et al., 2003).

The same total DNA was also used as template to control the presence/absence of endogenized viral genomes into DNA through PCR. A list of virus-specific primer is present in supplementary material (Supplementary Table 3).

2.8. RCA analysis on fungal isolates

For the rolling circle amplification (RCA) we used the Illustra TempliPhi 100 Amplification Kit (GE Healthcare Bio-Sciences, Uppsala, Sweden) following manufacturer's protocols and reagents. We performed a restriction digestion protocol using *Eco*RI on the amplified fragment. Restriction products, when visible as discrete bands, were cloned in pGEM-T easy vector (Promega, Madison, WI, USA) and white colonies were screened by PCR using M13 forward and reverse primers. Plasmids with positive inserts were purified and sequenced as described above.

2.9. Total RNA extraction and RNAseq

Total RNA was extracted from six isolates positive for dsRNA presence: MUT4330, MUT4358, MUT4359, MUT4379, MUT4917, MUT4935. Four isolates that were negative for dsRNA presence, MUT4366, MUT4370, MUT4405 and MUT4924 were also included. RNA quantification and quality were tested using NanoDrop 2000 Spectrophotometer (Thermoscientific, Waltham, MA, USA). We selected sample with a A260/A280 ratio within 1.8 and 2 and a A260/A230 ratio within 2 and 2.2. Ten micrograms of these RNAs were sent to BMR Genomics (Padova, Italy) for ribosomal RNA (rRNA) depletion (Ribo-Zero™ Gold Kit, Epicentre, Madison, USA), cDNA libraries construction (TrueSeq total RNA sample kit, Illumina) and sequencing by Illumina technologies with a single-end library read length of 150 bp. MUT4358 and MUT4359 were sequenced in the same lane, and also the four negative dsRNA isolates were sequenced in a single mixed sample in the same Illumina lane.

2.10. Assembly of mycovirus sequences from small RNA (sRNA) libraries

For sRNA sequencing we tried to mimic previous work showing that in some instances viruses from endophytic fungi could be detected directly from leaf samples (Al Rwahnih et al., 2011) and we selected two virus positive strains from the dsRNA assay, MUT4330 and MUT4359. Total RNA was isolated from the marine plant *P. oceanica* with Direct-zol RNA MiniPrep kit (Zymoresearch, CA, USA). The quality of total RNA extracted was assessed by testing RIN (RNA Integrity Number) and 28S/18S ratio using the "Agilent RNA 6000 nano Reagents" preparation kit with an Agilent 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). Twenty micrograms of these RNAs were sent to BGI Company (Guangzhou, China) for sRNA gel purification, cDNA synthesis, libraries construction and sequencing by Illumina technologies with a single-end library read length of 50 bp.

2.11. Bioinformatics analysis

Raw data from sRNA sequencing were filtered using an in-house BGI method, and only high quality reads (no more than four bases with quality value lower than 10, and no more than six bases with quality value lower than 13), were considered for further analysis. Adapter sequences (5′ adapter sequence: TTCAGAGTTCTACAGTCCGACGATC, 3′ adapter sequence: TCGTATGCCGTCTTCTGCTTG) were removed using a custom BGI script, allowing four mismatches. The data were used to assemble the viral genomes. Velvet (1.2.10) (Zerbino, 2010) and Oases (0.2.8) (Schulz et al., 2012) with a multi K-mer approach were used for this purpose (Zhao et al., 2011).

De novo assembly from total-RNA sequencing approach was performed using high quality and clean sequences selected using the suite Illumina Real Time Analysis (RTA) v2. For assembly operation we used Trinity (2.0.2) (Haas et al., 2013) and then on the assembled contigs we used the BLAST suite (vers. 2.2.30) to search for homology to virus sequences. Alignments of reads against the viral contigs were performed using BWA 0.5.9 (Li and Durbin, 2009). Read sizes were determined using SAMtools (vers. 0.1.19) (Li et al., 2009).

Coding open reading frames (ORFs) were detected with ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi) and then blasted on NCBI databases.

2.11. Phylogenetic analysis of viral sequences

After characterizing the almost full-length sequences of each virus genome segment, a conserved part of the RNA dependent RNA polymerase (RdRP) coding ORF of each identified virus was used for multiple sequence alignments using ClustalW and MUSCLE (Edgar, 2004) inside the MEGA 6 (Tamura et al., 2012) suite of software. Aligned sequences were used for deriving phylogenetic trees using both the maximum likelihood (Felsenstein, 1981) and the Neighbor-Joining (Saitou and Nei, 1987) methods. Statistical analysis for each clade was carried out through bootstrap analysis with 1000 replicates. Further details of the phylogenetic analysis are included in figure legends.

3. Results

3.1. Mycovirus detected by dsRNA purification

Six isolates out of the 91 tested were positive for dsRNA bands, which varied in numbers and electrophoretic mobility in each of the positive isolate, showing hints of a diverse array of viruses in this collection (Fig. 1). Isolate MUT4330 (*Penicillium aurantiogriseum* var *viridicatum*) shows a complex pattern with at least 5 different bands (2 of about 5 and 6 kbp, one of about 4 kbp, a group of bands around 2 kbp and one of about 800 bp). Isolate MUT4358 (*P. janczewskii)* shows one band of about 3 kbp. Isolate MUT4359 (*P. janczewskii)* shows a different band pattern with one band clearly visible close to the 4 kbp marker band and a group of bands ranging from 3200 to 2500 bp. Isolate MUT4379 (*Pleospora typhicola*) shows only one band of about 6 kbp as well as MUT4935 (*Wallemia sebi*), whereas isolate MUT4917 (*Pleosporales* sp.) shows two bands, one of about 9 kbp and one of about 7 kbp. cDNA cloning using random primers on purified dsRNA resulted in a library of a number of clones corresponding to each dsRNA segment or group of segments. For each fungal isolate we sequenced at least 10 clones of different lengths, and submitted each sequence to BLASTx and BLASTn searches of the GenBank databases. In each of the six fungal isolates positive for dsRNA we identified at least one cDNA clone displaying conserved regions of RdRP of viral origin. In detail, with this technique, in the isolate MUT4330 we identified 4 different putative viruses (based on the number of distinct RdRPs), which, according to the most similar virus present in the database, could be tentatively classified as a fusari-like virus, a virus related to the Aspergillus foetidus slow virus 2, a partitivirus and a totivirus. With the same method we identified: a chrysovirus in *P. janczewskii* isolate MUT4358, another chrysovirus and a virus related to Alternaria longipes virus 1 in another isolate of *P. janczewskii* (MUT4359), a second fusari-like virus species in isolate MUT4379 of *P. tiphycola*, a virus related to the megabirnavirus in a *Pleosporales* sp. (MUT4917), and a virus related to Ustilago maydis virus H1 in the only dsRNA positive basidiomycete *W. sebi* (MUT4935).

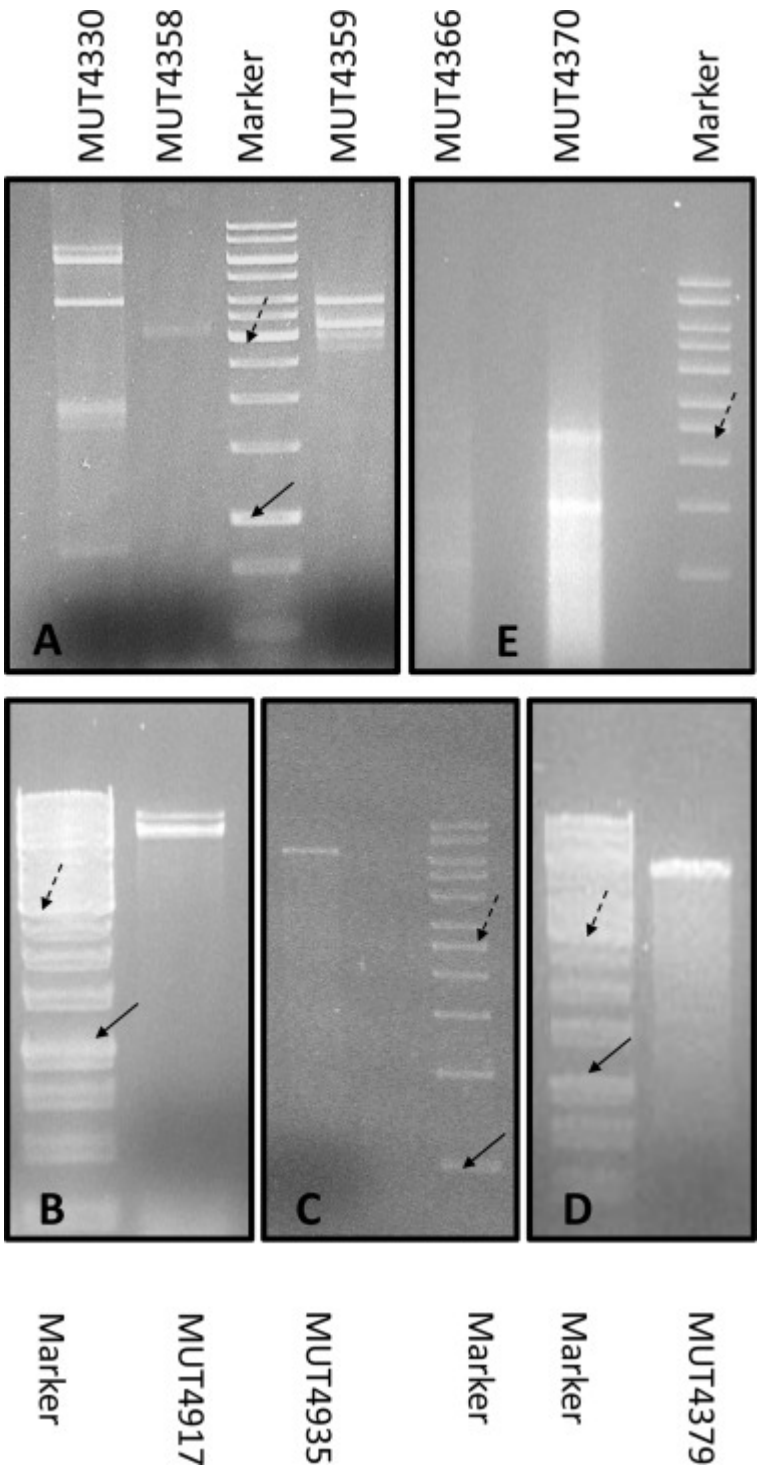


Fig. 1.

Gel electrophoresis of positive dsRNA isolates. (A) The first lane shows the complex pattern of bands isolated from isolate MUT4330 in which we identified 6 different viruses for a total number of eight genome segments. The second lane shows a low accumulation of only 1 of the 4 bands of the Penicillium janczewskii chrysovirus 1. The last lane displays another complex pattern of bands found in isolate MUT4359 in which there are a chrysovirus with 4 genomic bands (Penicillium janczewskii chrysovirus 2) and a dsRNA virus with only 1 genome segment of about 3 kb (Penicillium janczewskii Beauveria bassiana-like virus 1). In (B) two genome segments of the megabirnavirus are visible in the isolate MUT4917. In (C) the dsRNA of Wallemia sebi mycovirus 1 is shown, and in (D) with a similar molecular weight, the dsRNA of Pleospora typhicola fusarivirus 1. In (E) the two isolates (MUT4366

and MUT4370) positive for retrotransposon sequences are shown. Solid arrows point to the 1 kbp band marker, and dashed arrows point to the 3 kbp marker bands of the 1 kb Plus Ladder from Invitrogen-Life Technologies.

To our surprise, during this dsRNA screening we identified also two isolates harboring retrovirus/retrotransposon-like sequences: *P. spinulosum* (MUT4366) and *P. janczewskii* isolate (MUT4370) (Fig. 1E).

3.2. DNA and RCA analysis

The analyses of total DNA on agarose gel did not show any low molecular band other than the high molecular weight band corresponding to genomic DNA (data not shown) suggesting that these isolates do not harbor viruses with circular ssDNA genome: such technique, in facts, was successfully used to discover the presence of a circular ssDNA virus in *Sclerotinia sclerotiorum* (Yu et al., 2010).

RCA is a very powerful and easy technique able to detect the presence of circular DNA, that recently resulted in a wealth of new putative circular ssDNA viruses in metagenomic analysis of insect and marine environmental samples, for which the term “circuloma” was used (Ali et al., 2014 and James et al., 2011). We screened all the 91 isolates with this technique, and surprisingly, all of them amplified a high molecular band, whereas plant DNA healthy controls resulted in no amplification, as expected, or a higher molecular weight band in the case of plants infected with a commonly occurring geminivirus present in Italy (data not shown). We digested each of this amplified bands with *Eco*RI, in order to see if any of this sequence would actually have multimeric sequences containing this restriction enzymes site (resulting therefore in a specific discrete band after digestion): only MUT4379, an isolate of *P. typhicola*, gave a positive result, namely two bands of about 3 kb and 1 kb were obtained after digestion with the *Eco*RI enzyme. We cloned and sequenced such bands, and we obtained the full-length sequence (4439 nucleotides) of a putative previously undescribed fungal plasmid sequence. The nucleotide BLAST matched two supercontigs (7 kb and 4 kb, respectively) from *Leptosphaeria maculans* and *L. biglobosa* whole genome sequences, respectively. Our assembled sequence presents two different ORFs: the first is 1506 nucleotides long, and the encoded protein shows high similarity with DNA polymerases of *Neurospora intermedia* and *Cryphonectria parasitica* encoded by mitochondrial plasmid; ORF2 is 1140 nucleotides long, and shows also, as the ORF1, homology with mitochondrial DNA polymerases.

3.3. Assembly of mycovirus sequences from RNAseq libraries

The main feature and small reads archives (SRA) accession numbers of the RNAseq libraries sequenced are displayed in Table 1. Virus presence can also be detected with deep sequencing of total RNA preparation (most viral RNAs do not have a poly A tail, and therefore transcriptome analyses would not represent exhaustively virus presence) and results obtained with this technique arguably provided the most complete description of the virome present in our fungal collection (see details in discussion).

Table 1.

Summary of Illumina deep sequencing data.

Isolates	Library	SRA accession number	Reads		
			Total clean reads	reads mapping to assembled virus genomes	%
MUT4330 *Penicillium aurantiogriseum* var. *viridicatum*	sRNA	SRR2125723	11.019.948	59.700	0.54%
MUT4359 *Penicillium waksmanii*	sRNA	SRR2125724	9.396.563	9.514	0.10%
Posidonia oceanica	sRNA	SRR2125616	13.464.251		
MUT4330 *Penicillium aurantiogriseum* var. *viridicatum*	RNAseq	SRR2127587	30.234.856	383.220	1.27%
MUT4917 *Pleosporales* sp.	RNAseq	SRR2128152	24.979.455	154.217	0.62%
MUT4935 *Wallemia sebi*	RNAseq	SRR2126280	27.442.254	747.052	2.72%
MUT4379 *Pleospora typhicola*	RNAseq	SRR2131325	27.619.327	564.227	2.04%
MUT4358 *Penicillium waksmanii*, MUT4359 *Penicillium waksmanii*	RNAseq	SRR2131394	32.741.594	95.143	0.29%
MUT4366 *Penicillium spinulosum*, MUT4370 *Penicillium waksmanii*, MUT4405 *Pleosporales* sp., MUT4924 *Cunninghamella bertholletiae*	RNAseq	SRR2131328	27.283.809		

In sample MUT4330 we identified 6 viruses (Table 2 and Fig. 2A). The largest viral genome we identified is 6139 nucleotides long and contains 2 ORFs (Fig. 2A). ORF1 (calculated molecular mass of 174.5 kDa) displays high similarity with the RdRP sequences of already know mycoviruses soon to be classified in a new taxon by the International Committee on the Taxonomy of Viruses (ICTV). ORF2 (53 kDa) is in a different frame and displays two conserved domains: an ATP dependent RNA helicase domain (amino acids 1071–1346) and a myosin tail (amino acids 1633–1952). Similarity to fusari-like viruses prompted us to call this virus Penicillium aurantiogriseum fusarivirus 1 (PaFV1).

Table 2.

List of the new viral species identified with the reference fungal isolate and derived acronym.

Host	Name	Genome segments	Acronym
	Penicillium aurantiogriseum fusarivirus 1	1	PaFV1
	Penicillium aurantiogriseum totivirus 1	1	PaTV1
	Penicillium aurantiogriseum bipartite virus 1	2	PaBV1
MUT4330	Penicillium aurantiogriseum partitivirus 1	2	PaPV1
	Penicillium aurantiogriseum partiti-like virus 1	1	PaPlV1
	Penicillium aurantiogriseum asp-foetidus like virus 1	1	PaFlV1
MUT4358	Penicillium janczewskii chrysovirus 1	4	PjCV1
	Penicillium janczewskii chrysovirus 2	4	PjCV2
MUT4359	Penicillium janczewskii Beauveria bassiana like virus 1	1	PjBlV1
MUT4379	Pleospora typhicola fusarivirus 1	1	PtFV1
MUT4917	Pleosporales megabirnavirus 1	2	PMbV1
MUT4935	Wallemia sebi mycovirus 1	1	WsMV1

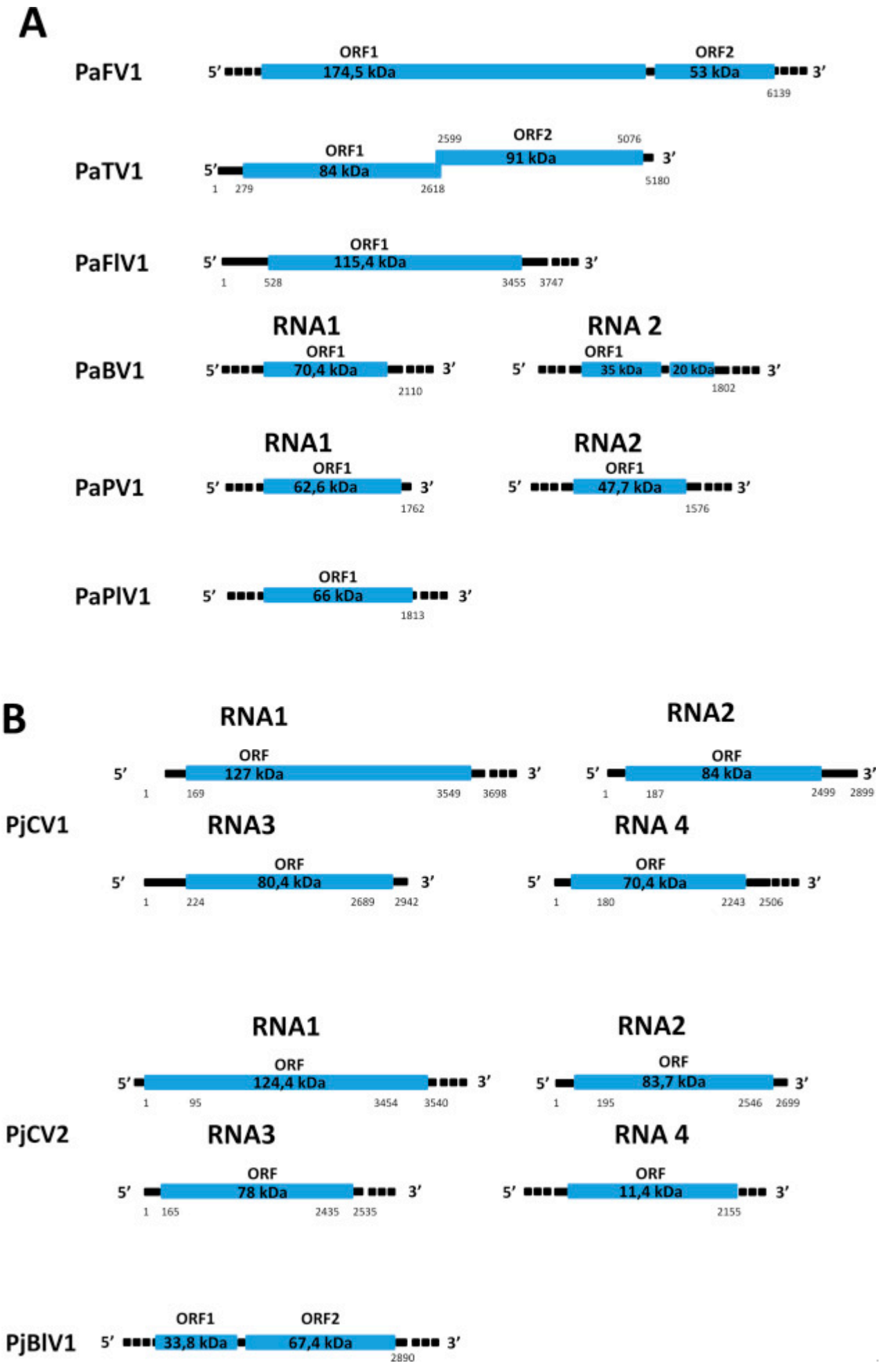

A
PaFV1
ORF1
174,5 kDa
ORF2
53 kDa
5'
3'
6139
PaTV1
ORF1
84 kDa
ORF2
91 kDa
2599
5076
1
279
2618
5180
PaFlV1
ORF1
115,4 kDa
1
528
3455
3747
PaBV1
RNA1
ORF1
70,4 kDa
2110
RNA 2
ORF1
35 kDa
20 kDa
1802
PaPV1
RNA1
ORF1
62,6 kDa
1762
RNA2
ORF1
47,7 kDa
1576
PaPlV1
ORF1
66 kDa
1813
B
PjCV1
RNA1
ORF
127 kDa
1
169
3549
3698
RNA2
ORF
84 kDa
187
2499
2899
RNA3
ORF
80,4 kDa
224
2689
2942
RNA 4
ORF
70,4 kDa
180
2243
2506
PjCV2
RNA1
ORF
124,4 kDa
95
3454
3540
RNA2
ORF
83,7 kDa
195
2546
2699
RNA3
ORF
78 kDa
165
2435
2535
RNA 4
ORF
11,4 kDa
2155
PjBlV1
ORF1
33,8 kDa
ORF2
67,4 kDa
2890

Fig. 2.

Genome organization of the twelve new virus species unveiled by our research: (A) Viruses present in isolate MUT4330 *Penicillium auranteogriseum* var. *viridicatum*. PaFV1 = Penicillium aurantiogriseum fusarivirus 1, PaTV1 = Penicillium aurantiogriseum totivirus 1, PaFlV1 = Penicillium aurantiogriseum asp-foetidus like virus 1, PaBV1 = Penicillium aurantiogriseum bipartite virus 1, PaPV1 = Penicillium aurantiogriseum partitivirus 1 and PaPlV1 = Penicillium aurantiogriseum partiti-like virus 1; (B) Viruses present in two different isolate of *Penicillium janczewskii*, MUT4358 and MUT4359: PjCV1 = Penicillium janczewskii chrysovirus 1, PjCV2 = Penicillium janczewskii chrysovirus 2, PjBlV1 = Penicillium janczewskii Beauveria bassiana-like virus 1; (C) viruses present in *Wallemia sebi* (MUT4935), an undefined Pleosporales sp. (MUT4917) and *Pleospora tiphycola* (MUT4379): PtFV1 = Pleospora typhicola fusarivirus 1, PMbV1 = Pleospora megabirnavirus 1 and WsMV1 = Wallemia sebi mycovirus 1.

For each virus genome, open boxes represent open reading frames. Dash lines indicate lack of 5′ or 3′ RACE characterization. Numbers represent the nucleotide position referred to the viral genome segment, when the complete sequence or the 5′ RACE is available; when 5′ RACE was not successful, the total length of the assembled sequence is reported. Genome segments and ORFs are drawn to proportion.

Sequences from genus *Totivirus* were not detected during dsRNA screening but from the RNAseq analysis and RACE method we could assemble a 5180 nucleotides long sequence with good similarity with totivirus sequences (Fig. 2A). ORF1 encodes a protein of 775 amino acids of about 80.2 kDa molecular mass with a conserved coat protein domain. ORF 2, in frame -1 and overlapping the carboxy terminal segment of the ORF1, encodes for a protein of 826 amino acids with a predicted molecular mass of 90.3 kDa and a conserved RdRP domain. A number of viruses show similar genome organization and many belong to the genus *Totivirus*. For this reason we propose the name Penicillum aurantiogriseum totivirus 1 (PaTV1).

RNAseq and RACE provided the complete genome sequence of a new virus species with similarity to the Aspergillus foetidus slow virus 2 (Kozlakidis et al., 2013) (Fig. 2A): the genome sequence is 3747 nucleotide long and shows a unique ORF spanning 2928 nucleotide. We found this virus during the dsRNA screening but we completed the sequence with data coming from the RNAseq analysis and with 5′ and 3′ RACE. Its single ORF encodes for a protein of 975 amino acids with an approximate molecular mass of 115.4 kDa and contains conserved viral RdRPs domains. This virus has a size and genome organization that is different from the viruses belonging to already described genera in the ICTV report so we propose to call it Penicillium aurantiogriseum asp-foetidus like virus 1 (PaFlV1), since the closest virus described is Aspergillus foetidus slow virus 2.

During the first screening of the dsRNA we found a sequence that encodes for a predicted unknown protein. With the use of RNAseq techniques we identified the unknown sequence as belonging to a bipartite virus, and in particular to the RNA2. The RNA1 is 2110 nucleotides long, it encodes a single protein of 614 amino acids with a predicted molecular mass of 70.4 kDa (Fig. 2A). The deduced protein contains a conserved viral RdRP domain. The RNA2 displays 2 distinct ORFs in the same frame, coding for two proteins of 35 kDa and 20 kDa respectively (Fig. 2A). Both the first and the second ORFs encode for protein of unknown function but while the first show 61% of identity with the homologous protein of Curvularia thermal tolerance virus, the second protein seems to be not conserved. This virus has a genome organization similar to at least 3 other virus: Curvularia thermal tolerance virus, Cryphonectria parasitic bipartite mycovirus 1 and

Heterobasidion RNA virus 6, suggesting the possibility to create new taxa accommodating members of this clade; we propose for this virus the name Penicillium aurantiogriseum bipartite virus 1 (PaBV1).

Sequences coming from the family *Partitiviridae* were also detected from the molecular characterization of the dsRNA screening, but RNAseq helped us to complete the viral genome composed of 2 distinct genomic segments (Fig. 2A). The first segment (RNA1) encodes for a protein with a predicted molecular mass of 62.6 kDa displaying a conserved viral RdRP domain. The RNA2 also contains only 1 ORF encoding for a protein with a predicted molecular mass of 47.7 kDa showing high similarity with RNA2 encoded CP of different partitiviruses. High similarity with members of this genus prompted us to call this virus Penicillium aurantiogriseum partitivirus 1 (PaPV1).

Another sequence that shows low identity with partitivirus (less than 30%) was identified during the RNAseq analysis. This sequence was not detected with the other methods and RNAseq was the only one allowing us to assemble the genome (Fig. 2A). This virus consists only of 1 genomic RNA segment, 1813 nucleotides long, with only 1 ORF encoding for a protein with a predicted molecular weight of 66 kDa with a conserved viral RdRP domain. Considering the low similarity with the partitiviridae RdRP but the absence of the RNA2 with the CP sequence, we decided to call this virus Penicillium aurantiogriseum partiti-like virus 1 (PaPlV1); we cannot rule out the possibility that a second genomic fragment is present, but could not be identified due to the low conservation of the putative CP, although *bona fide* partitivirus coat protein are conserved, and should have been detected if present.

The analysis of the fungus *P. janchzewskii* (isolate MUT4358) revealed the presence of a chrysovirus. Viruses of this group are characterized by the presence of 4 different genomic segments packaged in the same virus particle (Fig. 2B). The RNA1 is about 3.7 kbp and encodes for an RdRP with a predicted molecular mass of 127 kDa. The RNA2 is 2899 nucleotides long and it encodes for the capsid protein (CP) of about 84 kDa; the other two segments, RNA3 and RNA4, are also monocistronic, coding for proteins of a predicted molecular mass of 80.4 and 70.4 kDa respectively of unknown function and display little similarity with other chrysoviruses.

The second isolate of *P. janchzewskii* (MUT4359) revealed the presence of a distinct chrysovirus (Fig. 2B). The RNA1 is 3540 nucleotide long, and its single ORF encodes for a protein with a predicted molecular weight of 124.4 kDa and with conserved viral RdRP domains. The RNA2 is 2699 nucleotides long and it encodes for the CP with a predicted molecular weight of 83.7 kDa. The RNA3 of 2535 bp with a single ORF encodes for a protein of about 78 kDa of unknown function. The RNA4 also contains a single ORF translated in a protein of 70.2 kDa with unknown function. Both identified chrysoviruses from *P. janchzewskii*, named PjCV1 and PjCV2, have genome segment similar in sizes to other chrysoviruses already reported in the databases.

Isolate MUT4359 showed also the presence of a second virus with a genome segment of 2890 nucleotides (Fig. 2B). The genome displays two ORFs encoding for a protein of unknown function of 33.8 kDa in mass and a protein with viral RdRP conserved domains with a putative molecular mass of 67.4 kDa. Due to its highest similarity with Beauveria bassiana RNA virus (Kotta-Loizou et al., 2015), the virus is named Penicillium janczewskii Beauveria bassiana-like virus 1 (PjBlV1) (Fig. 2B).

		RNAseq					sRNA sequencing					
Viral genome	**Accession number**	**Mapping reads**	**Max. depth**	**Av. depth**	**Contig dimension**	**Coverage**	**Mapping reads**	**Max. depth**	**Av. depth**	**Coverage raw reads**	**N° contigs**	**Coverage contigs**
Penicillium aurantiogriseum asp-foetidus like virus	KT601100	47402	5756	1853	3736	99%	9097	230	54	99%	44	86%
Penicillium aurantiogriseum bipartite virus RdRP	KT601101	387	74	27	2110	99%	19	3	0	14%	0	ND
Penicillium aurantiogriseum bipartite virus CP	KT601102	535	103	43	1802	99%	11	2	0	8%	0	ND
Penicillium aurantiogriseum fusarivirus 1	KT601099	60352	3600	1443	6139	99%	17212	326	63	99%	74	94%
Penicillium aurantiogriseum totivirus 1	KT592305	44353	9650	1264	5159	99%	17144	299	75	99%	52	99%
Penicillium aurantiogriseum partiti-like virus	KT601105	103795	19030	8374	1813	99%	9715	364	120	99%	20	98%
Penicillium aurantiogriseum partitivirus RdRP	KT601103	76518	16954	6484	1732	99%	3258	158	45	99%	17	91%
Penicillium aurantiogriseum partitivirus CP	KT601104	49878	11702	4632	1576	99%	3244	159	46	99%	18	79%
Penicillium janczewskii Beauveria bassiana like virus 1	KT601106	20866	2182	1056	2890	99%	2463	238	20	97%	5	10%
Penicillium janczewskii chrysovirus 2 RdRP	KT950836	24128	2015	1002	3529	99%	2588	172	18	94%	12	21%

Penicillium janczewskii chrysovirus 2 RNA2	KT950837	6522	877	352	2709	99%	2124	219	19	92%	3	4%
Penicillium janczewskii chrysovirus 2 RNA3	KT950838	3490	615	202	2524	99%	1450	117	14	97%	1	9%
Penicillium janczewskii chrysovirus 2 RNA4	KT950839	6511	884	446	2143	99%	889	83	10	96%	0	ND
Penicillium janczewskii chrysovirus 1 RdRP	KT601115	3687	218	105	2648	99%						
Penicillium janczewskii chrysovirus 1 RNA2	KT601116	8197	979	436	2756	99%						
Penicillium janczewskii chrysovirus 1 RNA3	KT601117	12190	1152	609	2932	99%						
Penicillium janczewskii chrysovirus 1 RNA4	KT601118	9552	118	552	2532	99%						
Pleospora typhicola fusarivirus 1	KT601107	564227	30386	12410	6733	99%						
Peosporales megabirnavirus L1	KT601119	93306	23059	1550	8822	99%						
Plosporales megabirnavirus L2	KT601120	60911	26649	1742	5136	93%						
Wallemia sebi mycovirus 1	KT601108	747052	41645	18390	5856	99%						

The viral sequence coming from the *P. typhicola* isolate MUT4379, was detected with the dsRNA technique and the sequence was completed with RNAseq analysis (Fig. 2C). We found a 6733 nucleotide long sequence encoding for two ORFs. ORF1 is translated into a protein of 1551 amino acids with a putative molecular mass of 175 kDa which shows two conserved domains: an RdRP domain (nucleotides 1320–2508) and a helicase domain (nucleotides 3243–4215). ORF 2

encodes for a protein of 527 amino acids with a predicted molecular mass of 59 kDa (Fig. 2C). The virus was named Pleospora typhicola fusarivirus 1 (PtFV1).

The only basidiomycete positive to our initial dsRNA screening was the *W. sebi* MUT4935. Its virus (Fig. 2C) display sequences that show high similarity with Ustilago maydis virus H1 and Botrytis porri RNA virus 1. The RNAseq technique gave us a contig that was completed with the 5′ and 3′ RACE into a sequence of 5856 nucleotides (Fig. 2C). The sequence shows a single ORF with a predicted molecular mass of 210.4 kDa encoding for the putative CP-RdRP polyprotein. We suggest for this virus the name Wallemia sebi mycovirus 1 (WsMV1).

Isolate MUT4917, an uncharacterized *Pleosporales* fungus, showed the presence of two dsRNA bands during the first screening. From the cloned fragments we identified the presence of a virus of the family *Megabirnaviridae* (Fig. 2C) and for this reason we named the virus Pleosporales megabirnavirus 1 (PMbV1). The RNA1 is 8845 nucleotides long and presents two distinct ORFs. ORF1 (133 kDa) does not show any conserved domain but found match in database with the CP of another megabirnavirus (Sclerotinia sclerotiorum megabirnavirus 1). ORF2 is translated into a protein with a predicted molecular weight of 122.4 kDa and shows a conserved RdRP domain. We identified also the RNA2 but we were not able to complete the sequence. The obtained contig is 5136 nucleotides long and encodes for 2 distinct ORFs. The first ORF, ORF3, encodes for a protein having a predicted molecular mass of 117.4 kDa and unknown function. ORF4 is only partially characterized since 3′ RACE was not successful for this segment.

In Table 3 we reported a summary of the main statistical results of sRNA and RNAseq techniques, and in Supplementary Table 5 we reported the summary of 5′ and 3′ RACE characterizations. We also indicated for each ORF encoded by each viral genome the best hit in a database search with PSI-BLAST (Altschul et al., 1997) (Supplementary Table 6).

Table 3.

Results of RNAseq and sRNA sequencing: assembled contigs, genome-mapping reads and coverage.

		RNAseq					sRNA sequencing					
Viral genome	**Accession number**	**Mapping reads**	**Max. depth**	**Av. depth**	**Contig dimension**	**Coverage**	**Mapping reads**	**Max. depth**	**Av. depth**	**Coverage raw reads**	**N° contigs**	**Coverage contigs**
Penicillium aurantiogriseum asp-foetidus like virus	KT601100	47402	5756	1853	3736	99%	9097	230	54	99%	44	86%
Penicillium aurantiogriseum bipartite virus RdRP	KT601101	387	74	27	2110	99%	19	3	0	14%	0	ND

		RNAseq					sRNA sequencing					
Viral genome	**Accession number**	**Mapping reads**	**Max. depth**	**Av. depth**	**Contig dimension**	**Coverage**	**Mapping reads**	**Max. depth**	**Av. depth**	**Coverage raw reads**	**N° contigs**	**Coverage contigs**
Penicillium aurantiogriseum bipartite virus CP	KT601102	535	103	43	1802	99%	11	2	0	8%	0	ND
Penicillium aurantiogriseum fusarivirus 1	KT601099	60352	3600	1443	6139	99%	17212	326	63	99%	74	94%
Penicillium aurantiogriseum totivirus 1	KT592305	44353	9650	1264	5159	99%	17144	299	75	99%	52	99%
Penicillium aurantiogriseum partiti-like virus	KT601105	103795	19030	8374	1813	99%	9715	364	120	99%	20	98%
Penicillium aurantiogriseum partitivirus RdRP	KT601103	76518	16954	6484	1732	99%	3258	158	45	99%	17	91%
Penicillium aurantiogriseum partitivirus CP	KT601104	49878	11702	4632	1576	99%	3244	159	46	99%	18	79%
Penicillium janczewskii Beauveria bassiana like virus 1	KT601106	20866	2182	1056	2890	99%	2463	238	20	97%	5	10%
Penicillium janczewskii chrysovirus 2 RdRP	KT950836	24128	2015	1002	3529	99%	2588	172	18	94%	12	21%
Penicillium janczewskii	KT950837	6522	877	352	2709	99%	2124	219	19	92%	3	4%

		RNAseq					sRNA sequencing					
Viral genome	**Accession number**	**Mapping reads**	**Max. depth**	**Av. depth**	**Contig dimension**	**Coverage**	**Mapping reads**	**Max. depth**	**Av. depth**	**Coverage raw reads**	**N° contigs**	**Coverage contigs**
chrysovirus 2 RNA2												
Penicillium janczewskii chrysovirus 2 RNA3	KT950838	3490	615	202	2524	99%	1450	117	14	97%	1	9%
Penicillium janczewskii chrysovirus 2 RNA4	KT950839	6511	884	446	2143	99%	889	83	10	96%	0	ND
Penicillium janczewskii chrysovirus 1 RdRP	KT601115	3687	218	105	2648	99%						
Penicillium janczewskii chrysovirus 1 RNA2	KT601116	8197	979	436	2756	99%						
Penicillium janczewskii chrysovirus 1 RNA3	KT601117	12190	1152	609	2932	99%						
Penicillium janczewskii chrysovirus 1 RNA4	KT601118	9552	118	552	2532	99%						
Pleospora typhicola fusarivirus 1	KT601107	564227	30386	12410	6733	99%						
Peosporales megabirnavirus L1	KT601119	93306	23059	1550	8822	99%						
Plosporales megabirnavirus L2	KT601120	60911	26649	1742	5136	93%						

		RNAseq					sRNA sequencing					
Viral genome	**Accession number**	**Mapping reads**	**Max. depth**	**Av. depth**	**Contig dimension**	**Coverage**	**Mapping reads**	**Max. depth**	**Av. depth**	**Coverage raw reads**	**N° contigs**	**Coverage contigs**
Wallemia sebi mycovirus 1	KT601108	747052	41645	18390	5856	99%						

Results of RNAseq and sRNA sequencing: assembled contigs, genome-mapping reads and coverage.

Our RNAseq analysis also assembled contigs of two putative viruses for which further evidence of their existence as replicating biological entities could not be provided (Supplementary Table 7). The first group of contigs encodes for various discontinuous fragments with a sequence homologous to a putative negative ssRNA mycovirus related with Sclerotinia sclerotiorum ssRNA virus and was assembled from the RNAseq analysis of the MUT4935 isolate of *W. sebi*; initial RT-PCR analysis to confirm presence of the RNA gave negative results (not shown). The second is a contig obtained from the assembled reads of the mixed sample that contains 4 different fungi: MUT4366, MUT4370, MUT4405 and MUT4924. The contig shows high similarity with the RNA2 of bipartite viruses similar to Curvularia thermal tolerance virus and Cryphonectria parasitica bipartite mycovirus and further analysis with RT-PCR using specific primers and as template cDNA of each of the fungal isolates present in the mix revealed that this transcript is indeed present in MUT4370.

3.4. Assembly of mycovirus sequences from sRNA libraries

NGS of small RNA libraries of plants resulted in a wealth of *de novo* virus genome assembly, and the method demonstrated to be powerful both for detection of known viruses, and new virus genome assembly (Kreuze et al., 2009 and Vainio et al., 2015). A summary of the main features of each of the sequenced libraries is displayed in Table 1 together with the SRA accession number for each library. In order to compare different methods employing NGS to detect mycoviruses we decided to carry out the same protocol on two isolates that were positive for dsRNA presence: MUT4330 *P. aurantiogriseum var. viridicatum* and the MUT4359 *P. janczewskii*. We also included sRNA of the marine plant *P. oceanica* since the same approach in terrestrial plants (grapevine) resulted in the discovery of a number of mycovirus related viruses, likely having as host a fungal endophytes of the plant (Al Rwahnih et al., 2011).

The sRNA size distribution in *P. oceanica* is similar to plant sRNA distribution with a high proportion of 21 and 24 nucleotides reads (Supplementary Fig. 1). In the two fungi we found instead a different distribution, similar to a Gaussian distribution, with a peak for the 19 nucleotide long sRNA in isolate MUT4330 and 25 nucleotide long sRNA in isolate MUT4359 (Supplementary Fig. 2). Distribution of viral sRNA reveals a similar pattern, but with a peak of 21 nucleotides length for isolate MUT4330 and a peak of 22 nucleotides length for isolate MUT4359 (Supplementary Table 4, Supplementary Fig. 2).

Assembly of viral genomes from the sRNA sequences using Velvet and Oases with a multi k-mer approach resulted only in very short sequences and not for all the viruses that we knew being present in the isolate from our preliminary dsRNA approach (Table 2). We also mapped the sRNA sequences on the almost complete genomes obtained compiling the results of all the different approaches and the results are different for the two fungi: the viruses present in the MUT4330 isolate show a high percentage of mapping reads, with an average depth between 45 and 120, with the exception of the virus PaBV1 which showed a very low coverage. In addition to raw reads, we mapped also all the contigs obtained with the different k-mer value in order to understand if each viral genome could be derived mapping them to a reference sequence, since for most of the assembled contigs the viral nature can not be inferred from BLASTx results, particularly when they correspond to not conserved regions of the genomes. Similarly to the results obtained mapping raw sequence all the viruses except PaBV1 show a good percentage of genome coverage with a minimum of 79% for the CP of the PaPV1 up to a maximum of 99% for the PaFV1; in contrast with the result obtained for MUT4330, MUT4359 shows very low accumulation of sRNA sequences of viral origin. None of the 5 viral segments reaches the 99% of coverage with the raw reads, the average depth is low (between 10 and 20) and when we tried to map the assembled contigs we reached a maximum of 21% of coverage on the RNA2 of the PjCV2.

Analysis of sRNA from *P. oceanica* allowed us to assemble some short contigs (between 112 and 270 nucleotides) of viral origin that shared similarity with Stilbocarpa mosaic bacilliform virus and banana streak virus, pararetroviral sequences very likely endogenized in the genome, but no mycovirus-like sequence was detected.

3.5. Virus particle

To determine whether the virus assembled from cDNA from dsRNA and in silico from NGS sequencing are associated with virus particles, we attempted to partially purify (enrich) the viruses present in each fungal isolate. Results are displayed in Fig. 3. Despite the purification of different virus particles, the complexity found in MUT4330, which contains 6 different viruses, did not allow us to distinguish which particle contains which genome: nevertheless we were able to identify at least three different particle morphologies (Fig. 3A).

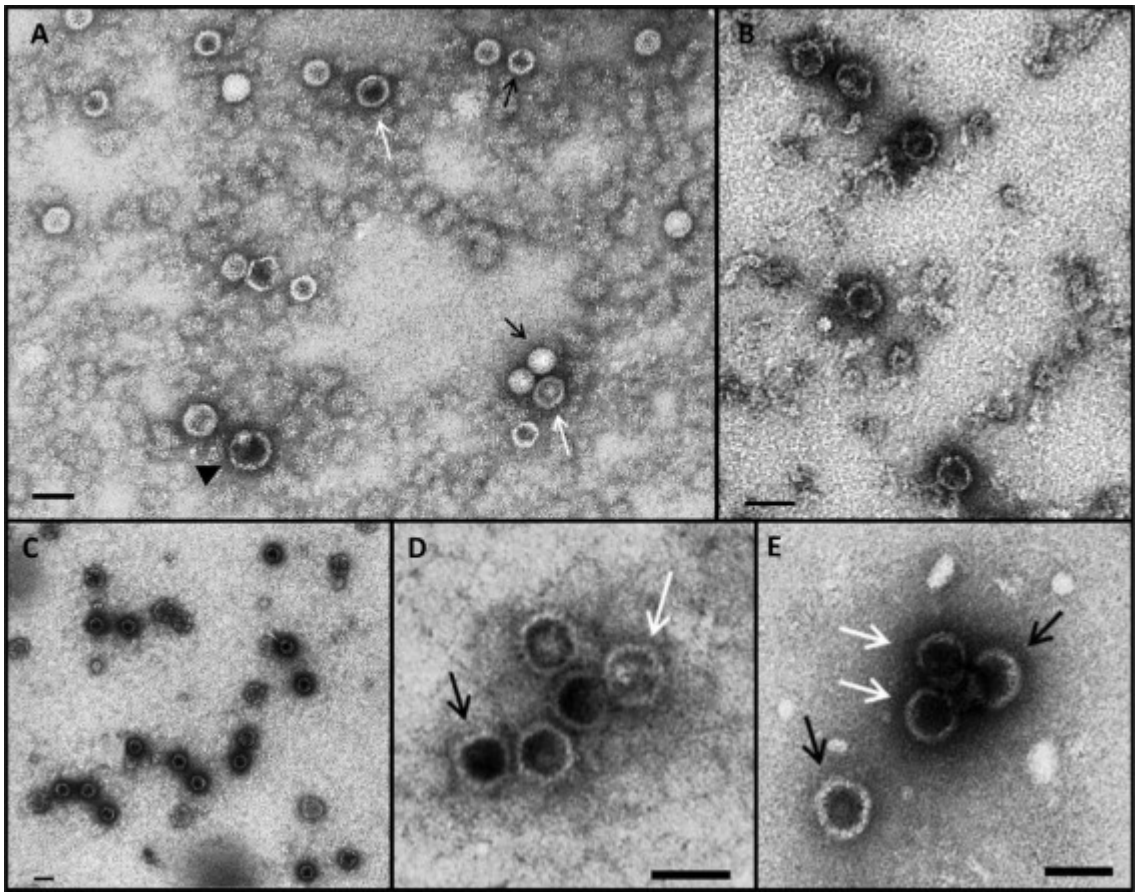


Fig. 3.

Virus particles from partially purified fungal extracts photographed at Transmission Electron Microscope (TEM). (A) In isolate MUT4330 at least three morphologies are visible. Particles of 32–35 nm in diameter, rounded in profile, either penetrated or not penetrated by the stain (black arrows). Particles of 39–42 nm in diameter are hexagonal in profile and always penetrated by the stain (white arrows). Particles of about 50 nm diameter, rounded in profile, always penetrated by the stain (arrowhead). The last two kinds of particles sometimes exhibit a 'core', probably due to an artifact of staining. (B) In isolate MUT4358, particles appear irregularly rounded in profile, about 35 nm in diameter and always penetrated by the stain. (C) Isolate MUT4359 shows particles that appear slightly polygonal in profile, about 35 nm in diameter and always penetrated by the stain. (D) In isolate MUT4917, particles appear polygonal in profile, with diameter of about 35 nm (black arrow) and of about 42 nm (white arrow). These particles are always penetrated by stain and the latter exhibit a 'core', probably due to a staining artifact. (E) Isolate MUT4935 exhibits particles appearing rounded in profile, with two different diameter lengths: about 36 nm (white arrows) and about 42 nm (black arrows) in diameter. They are always penetrated by stain. Bars in each panel represent 50 nm.

The chrysovirus PjCV1 present in isolate MUT4358 was successfully purified (Fig. 3B). The same result was obtained for the MUT4359 which contains 2 different viruses but only one type of particle, similar in shape and size to that found in isolate MUT4358: it is therefore safe to assume that the purified virus particle is that of chrysovirus PjCV2 whereas a second virion shape was absent, suggesting that the second virus genome (PjBlV1) is likely capsid-less (Fig. 3C).

We confirmed the absence of particles in isolate MUT4379, which contains PtFV1, a fusari-like virus with ssRNA genome not associated to virions.

From the MUT4917 isolate, which contains PMbV1, we purified particles of about 45 nm (Fig. 3D) similar to those already described for megabirnavirus (Chiba et al., 2009).

Finally we purified also the 40 nm particles associated with the presence of the WsMV1 (Fig. 3E).

3.6. RT-PCR and Northern blot analysis

For each virus we could amplify a cDNA fragment which was used for detection of the viral RNA in total RNA extracted from each fungal culture, in order to confirm which of the assembled contigs accumulates as RNA and is therefore not simply an assembly artifact, or originated from DNA contamination, or from transcription of a viral fragment endogenized in the genome. For each of the assembled contig we could find corresponding RNA fragments through RT-PCR (Fig. 4). Only the MUT4370, in which we identified a fragment of the RNA2 of a putative bipartite virus, shows the presence of the target sequence in the DNA treated with RNase A. This data seems therefore to confirm the endogenization of a viral fragment into the genome of its fungal host (Fig. 4, panel D and E).

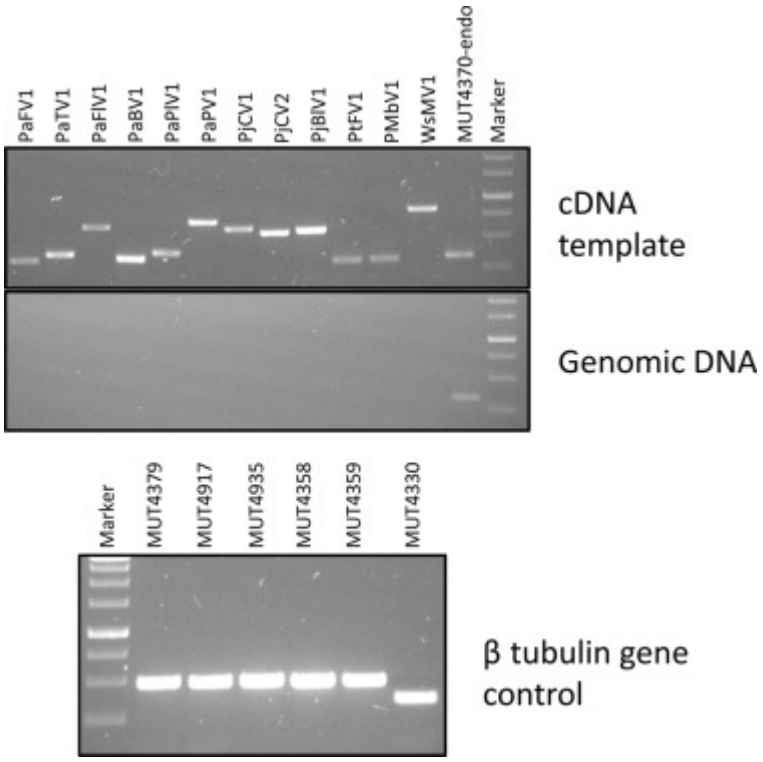


Fig. 4.

Detection of fragments of virus genomes by RT-PCR using DNase digested RNA template (top panel) and by PCR using as template RNase-digested total DNA (lower panel). Each lane is labeled according to the virus genome segment detected by the primer pairs detailed in Supplementary Table 3. In the bottom panel we used universal β-tubulin primers to test the quality of DNA and to exclude the presence of inhibitory substances in the DNA extractions; the assay was carried out for each virus positive isolate (MUT4330, MUT4358, MUT4359, MUT4379, MUT4917 and MUT4935). A specific fragment amplified both from cDNA and DNA template detected the presence of one specific viral sequence assembled *in silico* from RNAseq data from isolate MUT4370 suggesting the endogenization of the viral fragment and its expression as RNA.

All the viruses derived from cDNA clones from dsRNA analysis gave positive results in Northern blot analysis (Fig. 5). Size and number of bands are consistent with the predicted genome structure for each virus. As common for mycoviruses – with the exception of Fusarium graminearum virus 1 (Kwon et al., 2007) – no subgenomic RNA could be detected in our total RNA preparations.

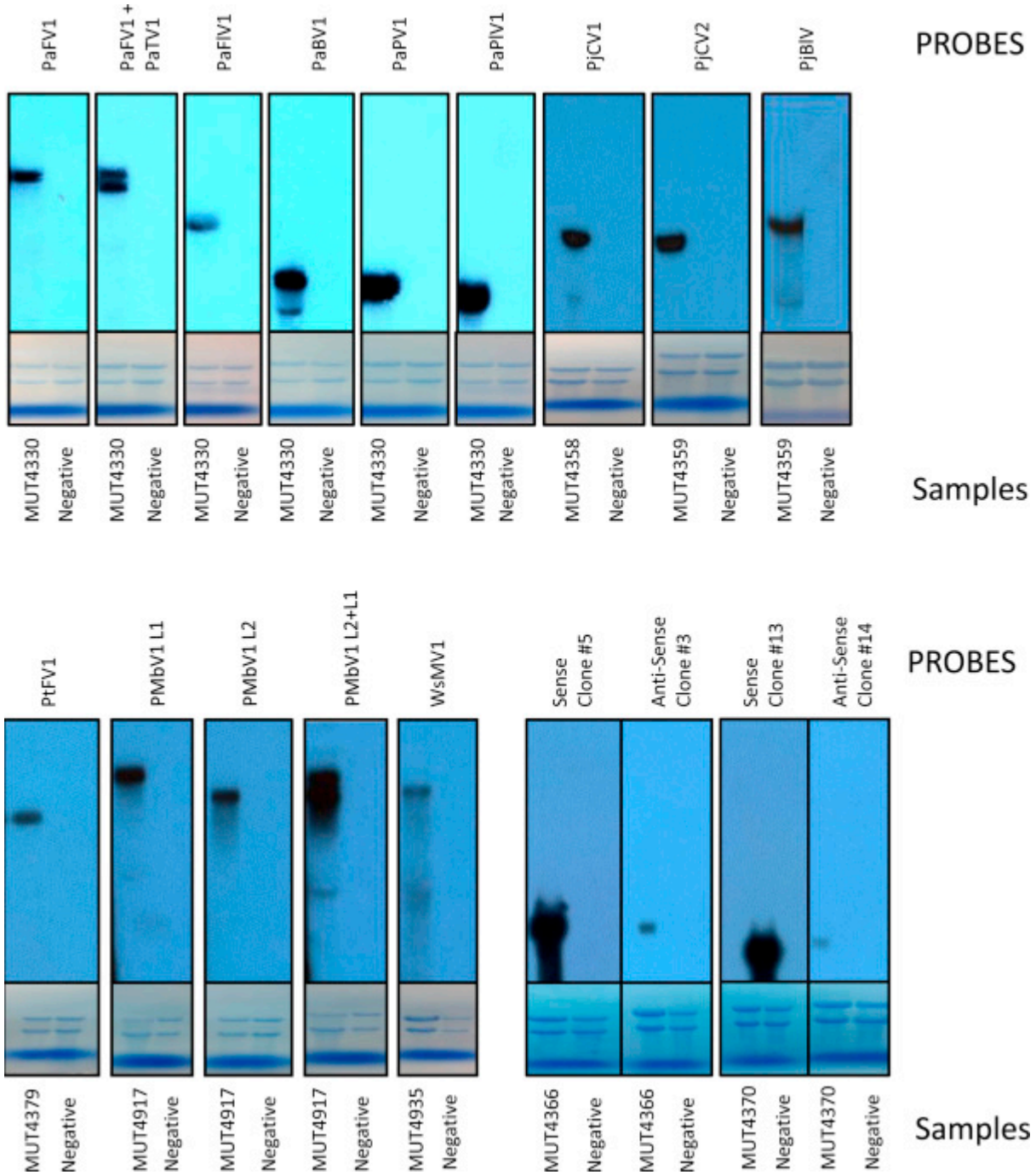


Fig. 5.

Northern blot analysis of total RNA extracted from a number of fungal isolates (labels at the bottom of each panel). Specific riboprobes were used to detect the viral genomes (labels on top of each panel). The Supplementary Table 2 shows the region of each virus genome targeted by each specific probe. Under each panel is a methylene blue stain of the membrane to show ribosomal RNA loadings. Negatives are RNA samples extracted from mycelia of a virus free fungal species from the Mycoteca Universitatis Taurinensis (MUT) collection (same species when available, or same genus).

Northern blot analysis for the putative endogenized retro-like sequences showed great accumulation of positive strand RNA but also some accumulation of minus strand RNA of the same size (Fig. 5).

As mentioned above, we could not find evidence of a negative ssRNA virus genome accumulation in the MUT4935, either by RT-PCR or northern blot, suggesting that the contig we originated could derive from contamination. Regarding the RNA2 contig (Penicillium waksmanii bipartite virus RNA2-like sequence) assembled from isolate MUT4370, northern blot analysis for the negative strand was negative (data not shown) suggesting that the RNA present in this isolate is likely the transcription product of endogenized viral cDNA in the genome (Fig. 4).

3.7. Phylogenetic analysis of viral sequences

In order to provide a proper taxonomical characterization for each of the virus that we found, we aligned each conserved region of the viral polymerase with similar sequences present in the databases (Supplementary Tables 8–10). We then derived a phylogenetic tree for each group of conserved sequences: in order to display alignments with sufficient conservation of amino acids residues among the aligned sequences, we distinguished one phylogenetic tree for dsRNA viruses (Fig. 6) and one for ssRNA viruses (Fig. 7).

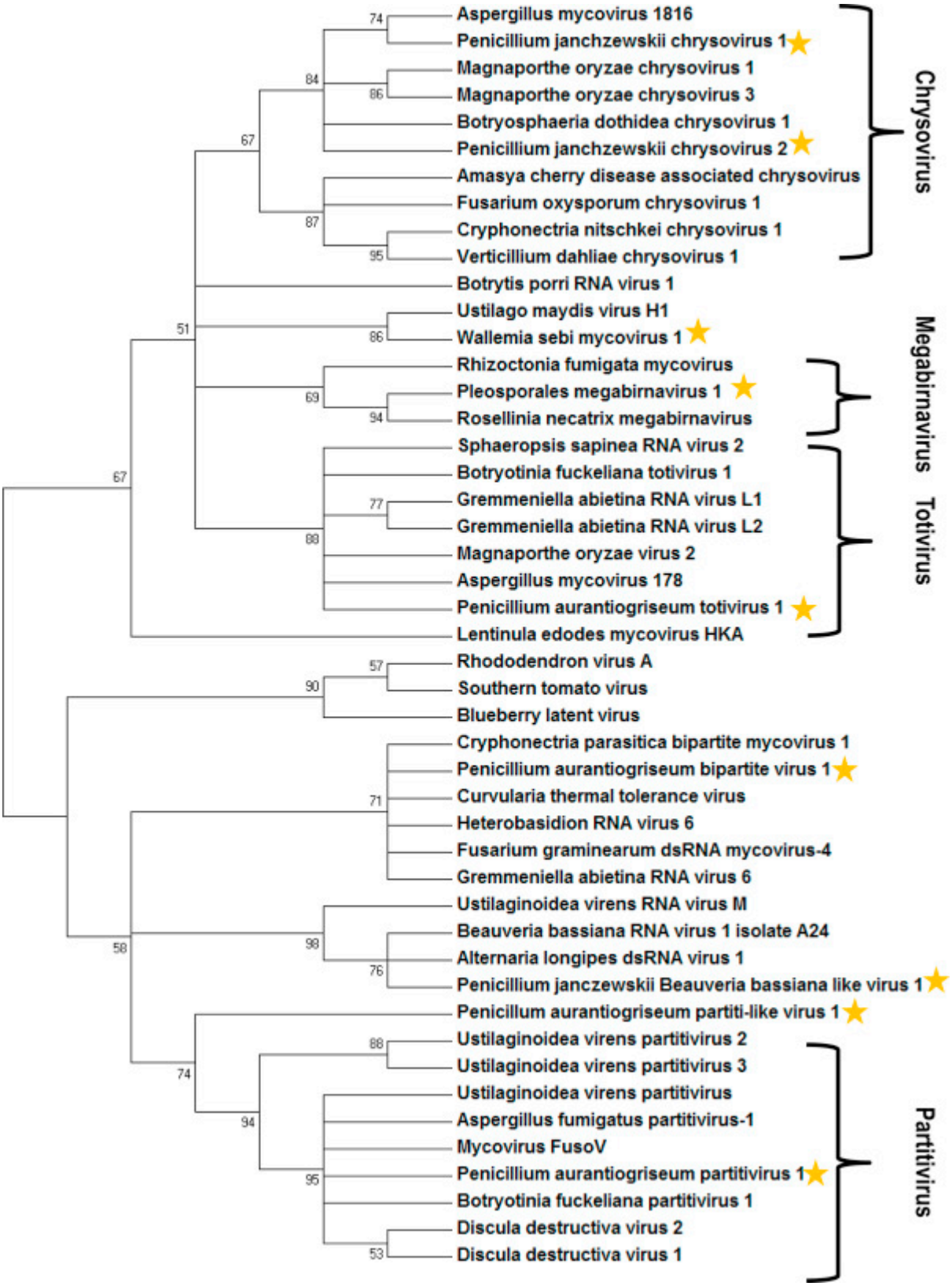


Fig. 6.

Neighbor-Joining phylogenetic analysis of aligned RdRP amino acids sequences from dsRNA viruses. The unrooted trees were generated with MEGA 6.0 software with 1000 bootstrap replicates. The numbers at the nodes indicates the percentage of bootstrap replicates supporting the branch. All nodes with less than 50% bootstrap replicates were collapsed to polytomy. The evolutionary distances were computed using the JTT matrix-

based method and are in the units of the number of amino acid substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). There were a total of 430 positions in the final dataset. Five out of nine viruses belong to already established phylogenetic clades: Penicillium aurantiogriseum totivirus 1 belongs to the *Totiviridae*, PaPV1 belongs to the *Partitiviridae*, Pleosporales megabirnavirus 1 belongs to the *Megabirnaviridae*, Penicillium janczewskii chrysovirus 1 and Penicillium janczewskii chrysovirus 2 belong to the *Chrysoviridae*. The other 4 viruses, Penicillium aurantiogriseum partiti-like virus 1, Wallemia sebi mycovirus 1, Penicillium janczewskii Beauveria bassiana-like virus 1 and Penicillium aurantiogriseum bipartite virus 1 require new taxonomic groups. Accession number of the 18 aligned sequences are reported in Supplementary Table 8.

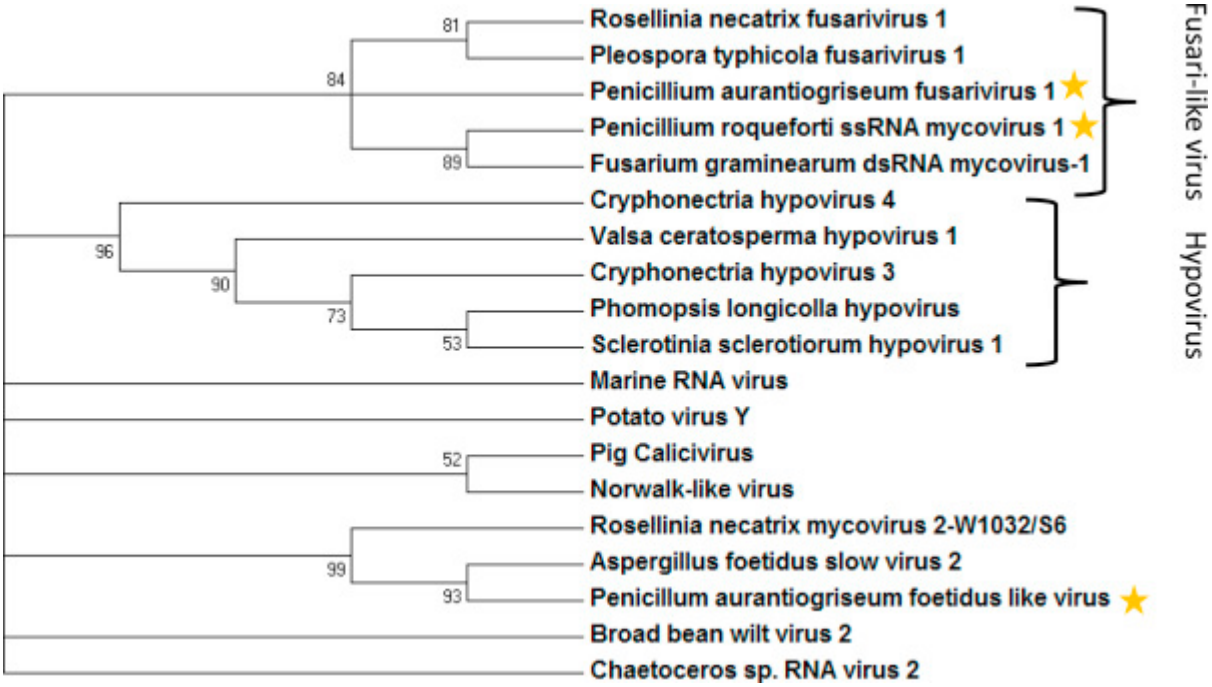


Fig. 7.

Neighbor-Joining phylogenetic analysis of aligned RdRP amino acids sequences from ssRNA viruses. The unrooted phylogenetic trees were generated with MEGA 6.0 software with 1000 bootstrap replicates. The numbers at the nodes indicates the percentage of bootstrap replicates supporting the branch, and all branches with less than 50% support were collapsed to polytomy. The evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. The rate variation among sites was modeled with a gamma distribution (shape parameter = 1). All positions containing gaps and missing data were eliminated resulting in a total of 166 positions in the final dataset. The two fusari-like viruses, Penicillium aurantiogriseum fusarivirus 1 and Pleospora typhicola fusarivirus 1, belong to the putative fusari-like virus group. The Penicillium aurantiogriseum asp-foetidus like virus 1 seems to form a new viral clade with the already described Aspergillus foetidus slow virus 2. Accession numbers of the sequences used for the alignment are reported in Supplementary Table 9.

In Fig. 6, the dsRNA virus Neighbor-Joining tree included 9 dsRNA viruses identified in this work. Most viruses group in clades already established taxonomically by ICTV (chrysoviruses, partitiviruses, totivirus and megabirnavirus), but some viruses belong to well-defined clades that still require a taxonomic definition by ICTV (Fig. 6).

The phylogenetic Neighbor-Joining tree displayed for the ssRNA genomes shows that the two new fusari-like virus species we detected are indeed in a well-supported clade with other fusari-like virus, a taxa that is currently under approval from the ICTV (Fig. 7). PaFlV1 instead forms a well-supported clade with Aspergillus foetidus slow virus 2 that stems from a ssRNA tree, therefore showing that this two viruses are indeed phylogenetically ssRNA viruses, and the classification of Aspergillus foetidus slow virus 2 as a dsRNA virus element (Kozlakidis et al., 2013) should, in our opinion, be reviewed.

Supplementary Fig. 3 shows the phylogenetic placement for retrotransposon sequences (currently classified in the family *Pseudoviridae* and *Metaviridae*) identified in two isolates (*P. spinulosum* MUT4366, and *P. janczewskii* MUT4370) and they belong to two different groups: retrotransposon sequences isolated from MUT4370 show similarity to sequences of other retrotransposon of fungal origin whereas the sequence isolated from MUT4366 groups with retrotransposon more similar to those of plants and other fungi (Supplementary Table 7).

Derivation of a maximum likelihood phylogenetic analysis (data not shown) confirmed the results obtained with the Neighbor-Joining trees.

4. Discussion

Our work on the multilayered ecological framework represented by *P. oceanica*, its associated fungal community and the viruses infecting them, is the first reporting the presence of marine mycoviruses: in the past, only two articles described a dsDNA virus in pseudo-fungal marine microorganisms (Dawe and Kuhn, 1983 and Takao et al., 2007), which modern taxonomy has shown to belong to the kingdom Protista. A very recent survey of various oceanic water samples distributed in all the major oceans (the circumglobal Tara Oceans expedition) provided a comprehensive overview of the Eukaryotic plankton diversity (de Vargas et al., 2015) and of the global ocean microbiome (Sunagawa et al., 2015). Dispite the in-depth analysis carried out, this project failed to uncover any mycoviruses because taxa in the eukaryotic plankton fraction (the one containing fungi) were detected through amplification of V9 region of the 18S rRNA gene therefore excluding the possibility of detecting their symbiotic viruses (Brum et al., 2015, de Vargas et al., 2015 and Sunagawa et al., 2015).

4.1. A collection of new mycovirus species isolated from marine fungi and taxonomical implications

In this work we are reporting 12 new viral species resulting from the analysis of 91 isolates representing various fungal taxonomic groups and species (Panno et al., 2013 and Gnavi et al., 2014) from a well characterized collection previously described in its mycological aspects and associated with the seagrass *P. oceanica* (Panno et al., 2013).

Although our work did not have the specific purpose of describing and quantifying the complexity of the three layered biological system under our scrutiny, we did not find any redundancy of virus presence in our collection confirming how difficult it is to fit the rarefaction curve of mycoviral species, as already shown in a previous work on the high prairies in Oklahoma (Feldman et al., 2012).

All the viruses were characterized in depth through dsRNA accumulation, Northern analyses, virus particle purification, and absence/presence of endogenization in the genome. Even though they showed some level of similarity with viruses already described and present in the sequence

databases, they are all considered new viral species. In fact, the maximum level of RdRP identity to the closest species was 85% for the partitivirus from MUT4330, but for the ORF encoded by RNA2 the identity was only 65%, allowing us to safely ascertain the “new species” status also for this virus. No big taxonomical surprise resulted from our inquiry: however, some of the new species we have characterized will help establish or redefine new taxonomic mycovirus groups.

Phylogenetic analyses of some of the newly characterized species prompted us to propose the annotation of some new viral family: in particular the addition of two new members of the fusari-like virus genus contributed to the consolidation of this recently proposed taxon. The report of another bipartite virus, phylogenetically related to Curvularia thermal tolerance virus, contributed to consolidate a group of viruses already large and fairly well characterized, but still not recognized by official taxonomy. In addition to these two groups already under discussion, we propose three new groups of viruses. The first includes the virus PaFlV1 that is phylogenetically a ssRNA virus. The second one includes the PjBlV1, phylogenetically related to Beauveria bassiana RNA virus 1 and Alternaria longipes virus 1. Finally, the association of WsMV1 into a well-defined clade with Ustilago maydis H1 virus shows that these are distinct and very distant from the totiviruses (Ustilago maydis H1 is currently classified as a totivirus) and a new taxon should accommodate these two virus species.

4.2. New tools for searching and assembling mycovirus genomes: suggestion for a multifaceted approach

On some of the samples, different approaches for virus genome detection/assembly were applied. The obtained results are very useful to discuss some of the positive and negative aspects of each methodology. Deriving genome fragments from the cDNA synthesized from dsRNA extracted and purified by CF11 chromatography is a relatively cheap approach, which has the advantage of detecting mostly real and abundant viruses, which are not endogenized; on the downside, completion of the full-length genome through genome walking and RACE, requires some investment in time and reagents, and costs increase. Furthermore this technique will neglect all DNA viruses, and all the ssRNA viruses, which may not accumulate abundant dsRNA as replication intermediate during cell infection. Additionally this technique did not allow to fully describe the virome of samples containing a complex array of viruses, as in the case of the isolate MUT4330, where we were not able to detect some of the viruses. The adaptation of this technique to NGS described by Roossinck and co-authors (Roossinck et al., 2010) retains the limitations of using dsRNA as initial substrate and in facts most of the viruses described with this approach failed to enter in the quoted literature because not sufficiently complete from the sequence point of view or from the biological perspective: ICTV has strict rules for approving a new mycovirus species, and most of the viruses described with ecogenomics approaches did not meet such parameters.

Given the restrictions of using dsRNA as template, we considered other two kinds of RNA substrates, sRNA and total RNA depleted of ribosomal RNA, in both cases with libraries characterized through NGS approaches. Both these approaches rely on the fact that viruses (with DNA or RNA genome) synthesize messenger RNA for translation of the proteins encoded by its genome, and therefore sequencing of RNA (or RNA derived sRNA) will include potentially all possible virus genomes. The downsides of this approach are, on one hand, the high cost per sample, and, on the other hand, the need to confirm the in silico assembly of the virus genomes contigs with some complementary technique, to detect the real presence of the virus entity in the original biological sample, to exclude contamination (Smuts et al., 2014) or artifacts of assembly (as seems the case of our minus strand RNA virus derived contig). Furthermore, such techniques are based on comparison with existing virus genome databases as the key strategy to identify viruses in the RNA

pool that has been sequenced: such approach will neglect all the viruses that do not have any conservation at the protein level with previously characterized viral sequences.

In recent years, works in plants (and to some extent in animals, such as in insects) have shown that the pool of sRNA contains fairly high percentages of virus derived sRNA (vsRNA), with high abundance of 21–24 nucleotides sRNAs generated by the silencing machinery in response to virus infection (Li et al., 2010). The application of sRNA library assembly technique on our samples did not allow to detect all the viruses already known to be present in the sample. Moreover, the sRNA read length profile differed from a characteristic profile produced by Dicer- like RNases from plants with a significant portion of the sRNAs shorter than the typical sRNAs of 21–25 nucleotides. In addition to this, the number of reads for each virus in some cases was very low, and the contigs obtained were difficult to assemble because sRNAs do not easily overlap. This technique has also shown a species-specific ability to detect viruses given that for some virus species the accumulation of sRNA was minimal. In facts, as already shown by the results obtained with different techniques using exactly the same RNA sample as template, in the isolate MUT4330, the majority of viruses are detectable with the sRNA technique but one, the PaBv1, seems to be able to escape almost completely the quelling machinery since only for this virus, the number of sequenced sRNA was minimal (Table 2). Accumulation of vsRNA is characteristic of RNA silencing-based antiviral defense, but in case of dicer-like or argonaute mutants the accumulation of this vsRNA is greatly affected (Segers et al., 2007; Zhang et al., 2008). Moreover some fungi have completely lost the genes for the RNA silencing machinery but do not seem to suffer high incidence of virus-induced symptoms suggesting the possibility of additional and yet uncharacterized mechanism used by fungi to mitigate viral replication (Nakayashiki and Nguyen, 2008). This phenomenon is particularly relevant in some yeasts, where co-evolution with ecologically relevant virus presence (causing the advantageous “killer” phenotype) has brought to the absence of the silencing machinery altogether (Drinnenberg et al., 2011), and therefore sRNA approaches would have failed any detection in these specific fungal species. It is also true that mycoviruses show the ability to suppress the RNA silencing antiviral response (Nuss, 2011). This could also be another explanation for the differences observed in generating sRNA from specific viruses particularly in the case of the isolate MUT4330, which is unique for hosting six different virus specie behaving very differently in respect to generating vsRNA: in this case it is safe to assume that indeed it is a virus encoded factor which determines the different specific outcome for each species.

All these limitations do not arise in using total RNA depleted of ribosomal RNA as substrate for NGS. This technique was the most powerful giving us complete or almost complete genomic sequences even from sample mixes of up to four isolates in a single Illumina sequencing lane. To our knowledge, although RNAseq results have been previously mined to detect the presence of viruses, in our work this approach is used for the first time with the purpose of detecting mycoviruses. The powerfulness of this technique was also demonstrated by the use of already available genome sequence database to describe for the first time a negative ssRNA mycovirus (Kondo et al., 2013).

Our approach for detecting DNA viruses in fungi did not give positive results. Instead, a plasmid was detected in one of the isolate, likely of mitochondrial origin, which is potentially interesting given the precedent for hypovirulence associated to a mitochondrial plasmid in *C. parasitica* (Monteiro-Vitorello et al., 2000). Nevertheless, probably the use of the RCA for the detection of the circuloma is not as effective in fungi as in plants or environmental samples, given that all the fungal samples we tested are positive for the initial amplification product (and that is not the case for plants, insects, or environmental samples). Furthermore, the lack of positive control samples makes it difficult to evaluate whether the approach of running total DNA preparations can be sensitive enough to detect the presence of Gemini-like mycoviruses, although this was exactly the technique

that brought to the discovery of the first ssDNA mycovirus (Yu et al., 2010, Yu et al., 2013 and Du et al., 2014).

4.3. Mycovirus presence should be investigated in fungal isolates of possible biotechnological relevance

In recent years, fungi have been the subject of a great number of works aimed at exploiting their potential for new drug discovery, particularly in their endless array of secondary metabolites (Lim and Keller, 2014), and for enzymes specifically adapted to distinct extreme environmental conditions, in particular in relation to temperature, salinity or pH conditions. For this reasons fungal collections all over the world are becoming a very important potential source of new patents. In this respect, we strongly believe that characterization of their virome is also of utmost importance, particularly for the isolates already identified as potentially interesting for biotechnological exploitation. The virome should be determined for both a general "quality control" pipeline, and also to determine its possible role in conferring the adaptive skills for which some fungal isolates are so interesting for researchers. In the collection of marine fungi we examined, we have already reported some isolates of biotechnological interest (Panno et al., 2013). In this framework, RNAseq associated to bioinformatics pipelines to detect viruses should be carried out for all the most interesting fungal isolates present in our collection.

5. Conclusions

This works resulted in a collection of virus-infected fungal isolates which are all well characterized taxonomically, but for which there is still no biological characterization: multiple approaches will be implemented to associate a specific virus to a phenotype including the comparison with cured or partially cured isolates, transmission (by anastomoses or through protoplast transfection with purified virus preparation) to a specific virus free isolate, or when feasible, through the synthesis of cDNA infectious clones.

One greatly overlooked subject that could stem from our research is what relates to retro-like sequences: (i) to our knowledge our work is the first to present the cloning of retrovirus-like sequences from dsRNA templates from fungi; (ii) the presence of retrovirus or retrotransposon sequences were confirmed by northern blot hybridization; (iii) the probe for anti-sense sequences reacted weakly with the sample indicating a possible low accumulation of the replicative form, that together with the fact that the ORF in the coding region has no stop codon, could suggest that the reverse transcriptase is indeed active. The presence of anti-sense sequences could also be attributed to the specific site of insertion of the retrovirus or retrotransposon. They could be in an area where there are sequences that promote the transcription of anti-sense sequence without the activation of the replicative form. Active retrotransposon have been associated to stress condition during fungal growth (Anaya and Roncero, 1996 and Ikeda et al., 2001).

The absence of DNA viruses in our collection confirms what is already known in literature (Daohong et al., 2013). On the other hand the presence of a plasmid into one of our fungi is an important factor because it could potentially play an important role in the adaptation to new environment (Griffiths, 1995) or give a particular phenotype to the fungus as shown for specific causes of hypovirulence (Monteiro-Virorello et al., 2000).

Another result of our research that requires further attention is the role of the endogenized mycovirus cDNA present in isolate MUT4370: endogenization of virus sequence in fungi was reported previously, and in some cases such event seem to have also a biological significance

(Feschotte and Gilbert, 2012 and Taylor and Bruenn, 2009). Furthermore endogenization during virus infection was recently shown to be a mechanism to temperate virus infection in insects (Goic et al., 2013) and such occurrence/mechanism could be investigated also in a mycovirus-fungal system.

Aknowledgments

We thank Riccardo Lenzi for excellent technical assistance in suggesting and carrying out purification protocols. We also thank Dr. Jeremy Warren, and Dr. Shahideh Nouri for editing the manuscript. This research was conducted in part with Advanced CyberInfrastructure computational resources provided by The Institute for CyberScience at The Pennsylvania State University, University Park Campus (http://ics.psu.edu).

We would like to thank also CRT Fundation for partially supporting this work.

Appendix A. Supplementary data

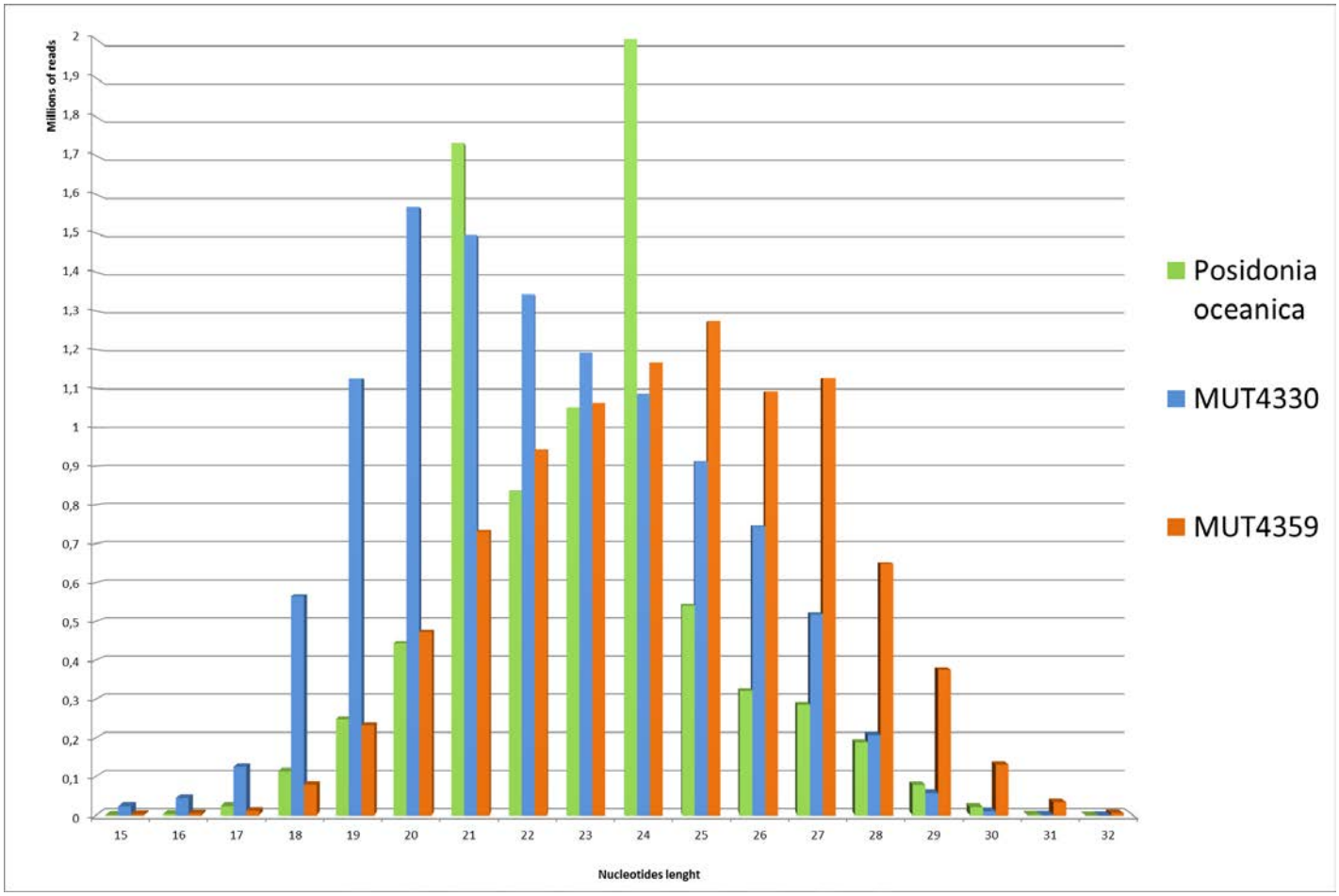
The following are Supplementary data to this article:

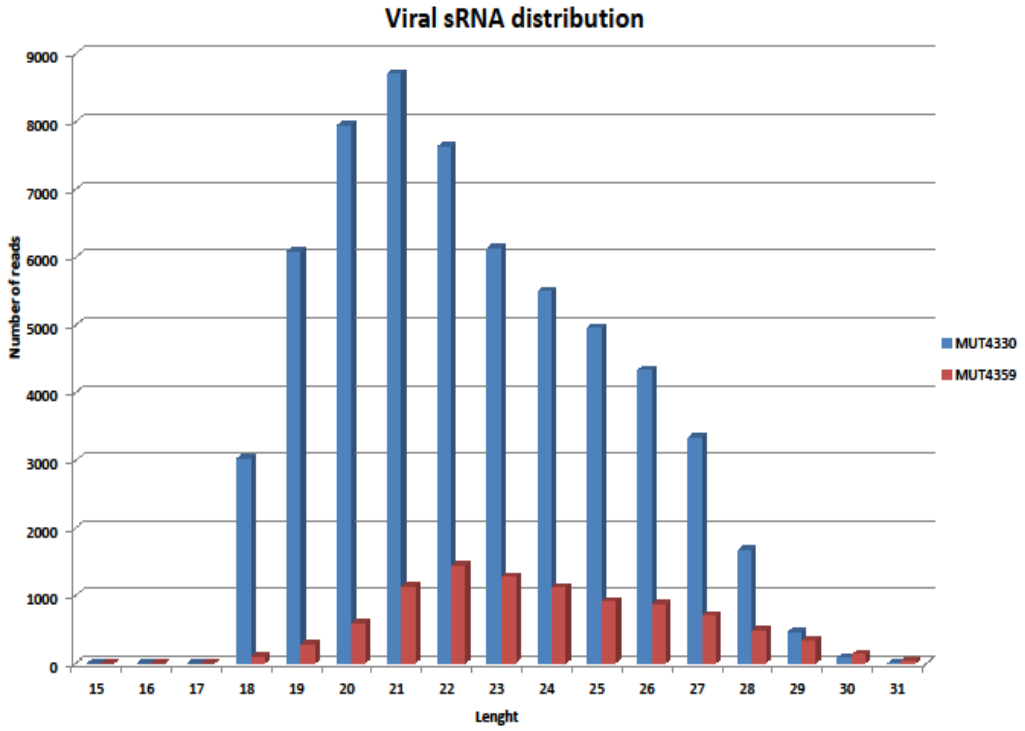
Supplementary material

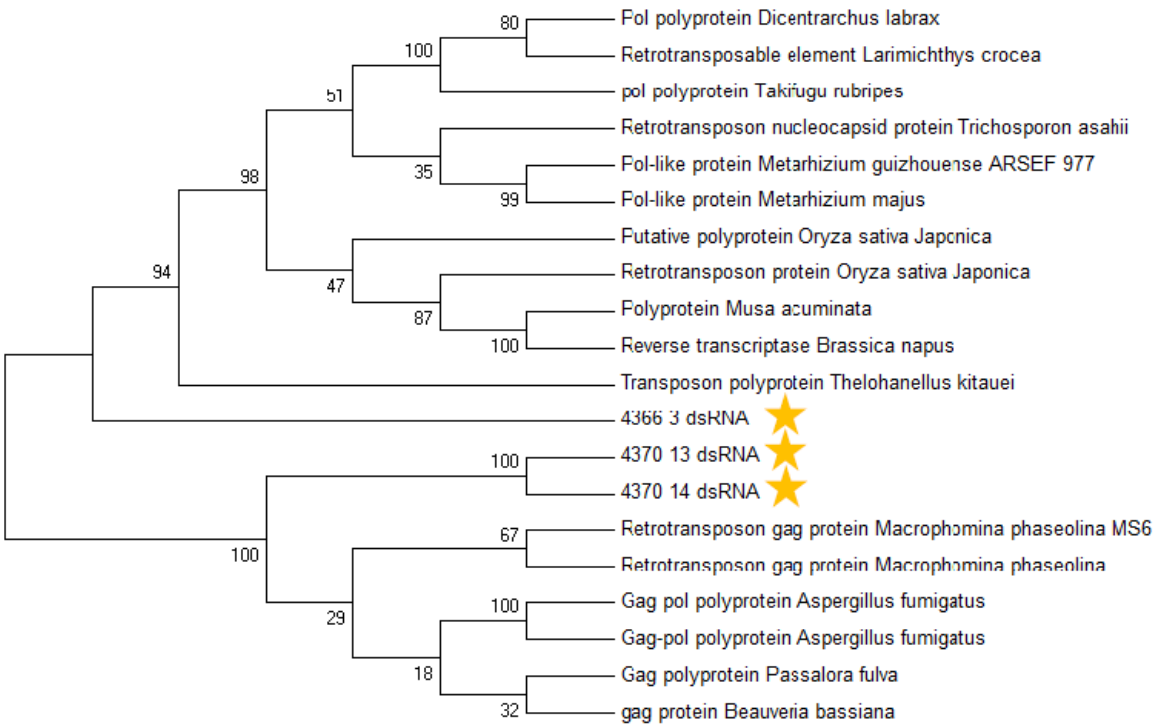
Supplementary Table 1. List of the fungal isolates from *Posidonia oceanica* screened from the MUT collection.

N°	Species	MUT	N°	Species	MUT	N°
1	*Acremonium implicatum*	4262	**31**	*Cladosporium sphaerospermum*	4304	**61**
2	*Acremonium minutisporum*	4263	**32**	*Cladosporium tenuissimum*	4306	**62**
3	*Acremonium* sp.	4264	**33**	*Diaporthe* sp.	4312	**63**
4	*Acremonium* sp.	4265	**34**	Didymellaceae sp.	4313	**64**
5	*Acremonium strictum*	4266	**35**	*Coniothyrium fuckelii*	4320	**65**
6	*Acremonium tubakii*	4267	**36**	*Massarina rubi*	4323	**66**
7	*Acremonium tubakii*	4268	**37**	*Penicillium aurantiogriseum* var. *viridicatum*	4330	**67**
8	*Acremonium tubakii*	4269	**38**	*Penicillium brevicompactum*	4331	**68**
9	*Acremonium verruculosum*	4270	**39**	*Penicillium brevicompactum*	4332	**69**
10	*Alternaria alternata*	4271	**40**	*Penicillium brevicompactum*	4333	**70**
11	*Alternaria alternata*	4272	**41**	*Penicillium brevicompactum*	4334	**71**
12	*Alternaria* sp.	4274	**42**	*Penicillium brevicompactum*	4335	**72**
13	*Aspergillus awamori*	4282	**43**	*Penicillium brevicompactum*	4336	**73**
14	*Aspergillus fumigatus*	4283	**44**	*Penicillium brevicompactum*	4337	**74**
15	*Aspergillus fumigatus*	4284	**45**	*Penicillium brevicompactum*	4338	**75**
16	*Aspergillus versicolor*	4286	**46**	*Penicillium janczewskii*	4356	**76**
17	*Aspergillus versicolor*	4287	**47**	*Penicillium janczewskii*	4357	**77**
18	*Beauveria bassiana*	4288	**48**	*Penicillium janczewskii*	4358	**78**
19	*Beauveria bassiana*	4289	**49**	*Penicillium janczewskii*	4359	**79**
20	*Cladosporium cladosporioides*	4293	**50**	*Penicillium janczewskii*	4360	**80**
21	*Cladosporium cladosporioides*	4294	**51**	*Penicillium janczewskii*	4361	**81**
22	*Cladosporium cladosporioides*	4295	**52**	*Penicillium janczewskii*	4362	**82**
23	*Cladosporium cucumerinum*	4296	**53**	*Penicillium janczewskii*	4363	**83**
24	*Cladosporium cucumerinum*	4297	**54**	*Penicillium janczewskii*	4364	**84**
25	*Cladosporium allicinum*	4298	**55**	*Penicillium spinulosum*	4366	**85**

26	*Cladosporium tenuissimum*	4299	**56**	*Penicillium spinulosum*	4367	**86**
27	*Cladosporium herbarum*	4300	**57**	*Penicillium waksmanii*	4368	**87**
28	*Cladosporium herbarum*	4301	**58**	*Penicillium waksmanii*	4370	**88**
29	*Cladosporium oxysporum*	4302	**59**	*Penicillium waksmanii*	4371	**89**
30	*Cladosporium cladosporioides*	4303	**60**	*Pyrenochaeta* sp.	4378	**90**
						91




Viral sRNA distribution
Number of reads
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
Lenght
MUT4330
MUT4359



References

Al Rwahnih et al., 2011 M. Al Rwahnih, S. Daubert, J.R. Urbez-Torres, F. Cordero, A. Rowhani Deep sequencing evidence from single grapevine plants reveals a virome dominated by mycoviruses Arch. Virol., 156 (2011), pp. 397–403

Ali et al., 2014 M.M. Ali, F. Li, Z. Zhang, K. Zhang, D.-K. Kang, J.A. Ankrum, X.C. Le, W. Zhao Rolling circle amplification: a versatile tool for chemical biology: materials science and medicine Chem. Soc. Rev., 43 (2014), pp. 3324–3341

Altschul et al., 1990 S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman Basic local alignment search tool J. Mol. Biol., 215 (1990), pp. 403–410

Altschul et al., 1997 S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman Gapped BLAST and PSI-BLAST: a new generation of protein database search programs Nucleic Acids Res., 25 (1997), pp. 3389–3402

Anaya and Roncero, 1996 N. Anaya, M.I.G. Roncero Stress-induced rearrangement of *Fusarium* retrotransposon sequences Mol. Gen. Genet, 253 (1996), pp. 89–94

Banks et al., 1968 G.T. Banks, K.W. Buck, E.B. Chain, F. Himmelweit, J.E. Marks, J.M. Tyler, M. Hollings, F.T. Last, O.M. Stone Viruses in fungi and interferon stimulation Nature, 218 (1968), pp. 542–545

Breitbart, 2012 M. Breitbart Marine viruses: truth or dare Ann. Rev. Mar. Sci., 4 (2012), pp. 425–448

Brum et al., 2015 J.R. Brum, J.C. Ignacio-Espinoza, S. Roux, G. Doulcier, S.G. Acinas, A. Alberti, S. Chaffron, C. Cruaud, C. de Vargas, J.M. Gasol, G. Gorsky, A.C. Gregory, L. Guidi, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, B.T. Poulos, S.M. Schwenck, S. Speich, C. Dimier, S. Kandels-Lewis, M. Picheral, S. Searson, P. Bork, C. Bowler, S. Sunagawa, P. Wincker, E. Karsenti, M.B. Sullivan Patterns and ecological drivers of ocean viral communities Science, 348 (2015)

Brussaard et al., 2008 C. Brussaard, S.W. Wilhelm, F. Thingstad, M.G. Weinbauer, G. Bratbak, M. Heldal, S.A. Kimmance, M. Middelboe, K. Nagasaki, J.H. Paul Global-scale processes with a nanoscale drive: the role of marine viruses ISME J., 2 (2008), p. 575

Chiba et al., 2009 S. Chiba, L. Salaipeth, Y.-H. Lin, A. Sasaki, S. Kanematsu, N. Suzuki A novel bipartite double-stranded RNA mycovirus from the white root rot fungus *Rosellinia necatrix*: molecular and biological characterization, taxonomic considerations, and potential for biological control J. Virol., 83 (2009), pp. 12801–12812

Ciuffo et al., 2008 M. Ciuffo, L. Tavella, D. Pacifico, V. Masenga, M. Turina A member of a new Tospovirus species isolated in Italy from wild buckwheat (*Polygonum convolvulus*) Arch. Virol., 153 (2008), pp. 2059–2068

Coutts and Livieratos, 2003 R.H. Coutts, I. Livieratos A rapid method for sequencing the 5′- and 3′-termini of double-stranded RNA viral templates using LM-RACE J. Phytopathol., 151 (2003), pp. 525–527

Daohong et al., 2013 J. Daohong, F. Yanping, G. Li, A.G. Said Viruses of the plant pathogenic fungus *Sclerotinia sclerotiorum* Adv. Virus Res., 86 (2013), pp. 215–248

Dawe and Kuhn, 1983 V.H. Dawe, C.W. Kuhn Virus-like particles in the aquatic fungus rhizidiomyces Virology, 130 (1) (1983), pp. 10–20

de Vargas et al., 2015 C. de Vargas, S. Audic, N. Henry, J. Decelle, F. Mahe, R. Logares, E. Lara, C. Berney, N. Le Bescot, I. Probert, M. Carmichael, J. Poulain, S. Romac, S. Colin, J.-M. Aury, L. Bittner, S. Chaffron, M. Dunthorn, S. Engelen, O. Flegontova, L. Guidi, A. Horak, O. Jaillon, G. Lima-Mendez, J. Lukes, S. Malviya, R. Morard, M. Mulot, E. Scalco, R. Siano, F. Vincent, A. Zingone, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, S.G. Acinas, P. Bork, C. Bowler, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, F. Not, H. Ogata, S. Pesant, J. Raes, M.E. Sieracki, S. Speich, L. Stemmann, S. Sunagawa, J. Weissenbach, P. Wincker, E. Karsenti Eukaryotic plankton diversity in the sunlit ocean Science, 348 (2015), p. 1261447

Drinnenberg et al., 2011 I.A. Drinnenberg, G.R. Fink, D.P. Bartel Compatibility with killer explains the rise of RNAi-deficient fungi Science, 333 (2011) 1592

Du et al., 2014 Z. Du, Y. Tang, S. Zhang, X. She, G. Lan, A. Varsani, Z. He Identification and molecular characterization of a single-stranded circular DNA virus with similarities to Sclerotinia sclerotiorum hypovirulence-associated DNA virus 1 Arch. Virol., 159 (2014), pp. 1527–1531

Edgar, 2004 R.C. Edgar MUSCLE: multiple sequence alignment with high accuracy and high throughput Nucleic Acids Res., 32 (2004), pp. 1792–1797

EEC 92/43, 1992 EEC 92/43 On the conservation of natural habitats and of wild fauna and flora European Union (1992)

Espach, 2013 Y. Espach The detection of mycoviral sequences in grapevine using next-generation sequencing Stellenbosch University, Stellenbosch (2013)

Feldman et al., 2012 T.S. Feldman, M.R. Morsy, M.J. Roossinck Are communities of microbial symbionts more diverse than communities of macrobial hosts? Fungal Biol., 116 (2012), pp. 465–477

Felsenstein, 1981 J. Felsenstein Evolutionary trees from DNA sequences: a maximum likelihood approach J. Mol. Evol., 17 (1981), pp. 368–376

Feschotte and Gilbert, 2012 C. Feschotte, C. Gilbert Endogenous viruses: insights into viral evolution and impact on host biology Nature Rev. Genet., 13 (2012), pp. 283–296

Fuhrman, 1999 J.A. Fuhrman Marine viruses and their biogeochemical and ecological effects Nature, 399 (1999), pp. 541–548

Gandy, 1960 D.G. Gandy A transmissible disease of cultivated mushrooms (watery stipe) Ann. Appl. Biol., 48 (1960), pp. 427–430

Gnavi et al., 2014 G. Gnavi, E. Ercole, L. Panno, A. Vizzini, G.C. Varese *Dothideomycetes* and *Leotiomycetes* sterile mycelia isolated from the Italian seagrass *Posidonia oceanica* based on rDNA data SpringerPlus, 3 (2014), p. 508

Goeker et al., 2011 M. Goeker, C. Scheuner, H.-P. Klenk, J.B. Stielow, W. Menzel Codivergence of mycoviruses with their hosts PLoS One, 6 (2011), p. e22252

Goic et al., 2013 B. Goic, N. Vodovar, J.A. Mondotte, C. Monot, L. Frangeul, H. Blanc, V. Gausson, J. Vera-Otarola, G. Cristofari, M.-C. Saleh RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model *Drosophila* Nat. Immunol., 14 (2013), pp. 396–403

Griffiths, 1995 A.J. Griffiths Natural plasmids of filamentous fungi Microbiol. Rev., 59 (1995), pp. 673–685

Haas et al., 2013 B.J. Haas, A. Papanicolaou, M. Yassour, M. Grabherr, P.D. Blood, J. Bowden, M.B. Couger, D. Eccles, B. Li, M. Lieber, M.D. MacManes, M. Ott, J. Orvis, N. Pochet, F. Strozzi, N. Weeks, R. Westerman, T. William, C.N. Dewey, R. Henschel, R.D. Leduc, N. Friedman, A. Regev De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis Nat. Protoc., 8 (2013), pp. 1494–1512

Hirzmann et al., 1993 J. Hirzmann, D. Luo, J. Hahnen, G. Hobom Determination of messenger-rna 5'-ends by reverse transcription of the cap structure Nucleic Acids Res., 21 (1993), pp. 3597–3598

Hollings, 1962 M. Hollings Viruses associated with a die-back disease of cultivated mushroom Nature, 196 (1962), pp. 962–965

Ikeda et al., 2001 K. Ikeda, H. Nakayashiki, M. Takagi, Y. Tosa, S. Mayama Heat shock, copper sulfate and oxidative stress activate the retrotransposon MAGGY resident in the plant pathogenic fungus *Magnaporthe grisea* Mol. Genet. Genom., 266 (2001), pp. 318–325

James et al., 2011 A.P. James, R.J. Geijskes, J.L. Dale, R.M. Harding Development of a novel rolling-circle amplification technique to detect banana streak virus that also discriminates between integrated and episomal virus sequences Plant Dis., 95 (2011), pp. 57–62

Kondo et al., 2013 H. Kondo, S. Chiba, K. Toyoda, N. Suzuki Evidence for negative-strand RNA virus infection in fungi Virology, 435 (2013), pp. 201–209

Kotta-Loizou et al., 2015 I. Kotta-Loizou, J. Sipkova, R.H. Coutts Identification and sequence determination of a novel double-stranded RNA mycovirus from the entomopathogenic fungus *Beauveria bassiana* Arch. Virol., 160 (2015), pp. 873–875

Kozlakidis et al., 2013 Z. Kozlakidis, N. Herrero, S. Ozkan, M.F. Bhatti, R.H.A. Coutts A novel dsRNA element isolated from the *Aspergillus foetidus* mycovirus complex Arch. Virol., 158 (2013), pp. 2625–2628

Kreuze et al., 2009 J.F. Kreuze, A. Perez, M. Untiveros, D. Quispe, S. Fuentes, I. Barker, R. Simon 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 (2009), pp. 1–7

Kwon et al., 2007 S.J. Kwon, W.S. Lim, S.H. Park, M.R. Park, K.H. Kim Molecular characterization of a dsRNA mycovirus, *Fusarium graminearum* virus-DK21, which is phylogenetically related to hypoviruses but has a genome organization and gene expression strategy resembling those of plant potex-like viruses Mol. Cells, 23 (2007), pp. 304–315

Li and Durbin, 2009 H. Li, R. Durbin Fast and accurate short read alignment with Burrows-Wheeler transform Bioinformatics, 25 (2009), pp. 1754–1760

Li et al., 2009 H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan The sequence alignment/map format and SAMtools Bioinformatics, 25 (2009), pp. 2078–2079

Li et al., 2010 L. Li, S.-S. Chang, Y. Liu RNA interference pathways in filamentous fungi Cell. Mol. Life Sci., 67 (2010), pp. 3849–3863

Lim and Keller, 2014 F.Y. Lim, N.P. Keller Spatial and temporal control of fungal natural product synthesis Nat. Prod. Rep., 31 (2014), pp. 1277–1286

Marquez et al., 2007 L.M. Marquez, R.S. Redman, R.J. Rodriguez, M.J. Roossinck A virus in a fungus in a plant: three-way symbiosis required for thermal tolerance Science, 315 (2007), pp. 513–515

Monteiro-Vitorello et al., 2000 C.B. Monteiro-Vitorello, D. Baidyaroy, J.A. Bell, G. Hausner, D.W. Fulbright, H. Bertrand A circular mitochondrial plasmid incites hypovirulence in some strains of *Cryphonectria parasitica* Curr. Genet., 37 (2000), pp. 242–256

Nakayashiki and Nguyen, 2008 H. Nakayashiki, Q.B. Nguyen RNA interference: roles in fungal biology Curr. Opin. Microbiol., 11 (2008), pp. 494–502

Nuss, 2011 D.L. Nuss Mycoviruses, RNA silencing, and viral RNA recombination Adv. Virus Res., 80 (2011), pp. 25–48

Panno et al., 2013 L. Panno, M. Bruno, S. Voyron, A. Anastasi, G. Gnavi, L. Miserere, G.C. Varese Diversity, ecological role and potential biotechnological applications of marine fungi associated to the seagrass *Posidonia oceanica* New Biotechnol., 30 (2013), pp. 685–694

Richards et al., 2012 T.A. Richards, M.D. Jones, G. Leonard, D. Bass Marine fungi: their ecology and molecular diversity Ann. Rev. Mar. Sci., 4 (2012), pp. 495–522

Roossinck et al., 2010 M.J. Roossinck, P. Saha, G.B. Wiley, J. Quan, J.D. White, H. Lai, F. Chavarria, G. Shen, B.A. Roe Ecogenomics: using massively parallel pyrosequencing to understand virus ecology Mol. Ecol., 19 (2010), pp. 81–88

Saitou and Nei, 1987 N. Saitou, M. Nei The neighbor-joining method a new method for reconstructing phylogenetic trees Mol. Biol. Evol., 4 (1987), pp. 406–425

Sambrook et al., 1989 J. Sambrook, E.F. Fritsch, T. Maniatis Molecular Cloning a Laboratory Manual (Second edition)Cold Spring Harbor Laboratory Press (1989)

Sanger et al., 1977 F. Sanger, S. Nicklen, A.R. Coulson DNA sequencing with chain-terminating inhibitors Proc. Natl. Acad. Sci. U. S. A., 74 (1977), pp. 5463–5467

Schulz et al., 2012 M.H. Schulz, D.R. Zerbino, M. Vingron, E. Birney Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels Bioinformatics, 28 (2012), pp. 1086–1092

Segers et al., 2007 G.C. Segers, X. Zhang, F. Deng, Q. Sun, D.L. Nuss Evidence that RNA silencing functions as an antiviral defense mechanism in fungi Proc. Natl. Acad. Sci. U. S. A., 104 (2007), pp. 12902–12906

Smuts et al., 2014 H. Smuts, M. Kew, A. Khan, S. Korsman Novel hybrid parvovirus-like virus, NIH-CQV/PHV, contaminants in silica column-based nucleic acid extraction kits J. Virol., 88 (2014), p. 1398

Stahl et al., 1978 U. Stahl, P.A. Lemke, P. Tudzynski, U. Kueck, K. Esser Evidence for plasmid like dna in a filamentous fungus the ascomycete *Podospora anserina* Mol. Gen. Genet., 162 (1978), pp. 341–343

Sunagawa et al., 2015 S. Sunagawa, L.P. Coelho, S. Chaffron, J.R. Kultima, K. Labadie, G. Salazar, B. Djahanschiri, G. Zeller, D.R. Mende, A. Alberti, F.M. Cornejo-Castillo, P.I. Costea, C. Cruaud, F. d'Ovidio, S. Engelen, I. Ferrera, J.M. Gasol, L. Guidi, F. Hildebrand, F. Kokoszka, C. Lepoivre, G. Lima-Mendez, J. Poulain, B.T. Poulos, M. Royo-Llonch, H. Sarmento, S. Vieira-Silva, C. Dimier, M. Picheral, S. Searson, S. Kandels-Lewis, C. Bowler, C. de Vargas, G. Gorsky, N. Grimsley, P. Hingamp, D. Iudicone, O. Jaillon, F. Not, H. Ogata, S. Pesant, S. Speich, L. Stemmann, M.B. Sullivan, J. Weissenbach, P. Wincker, E. Karsenti, J. Raes, S.G. Acinas, P. Bork Structure and function of the global ocean microbiome Science, 348 (2015)

Suttle, 2007 C.A. Suttle Marine viruses—major players in the global ecosystem Nat. Rev. Microbiol., 5 (2007), pp. 801–812

Tamura et al., 2012 K. Tamura, F.U. Battistuzzi, P. Billing-Ross, O. Murillo, A. Filipski, S. Kumar Estimating divergence times in large molecular phylogenies Proc. Natl. Acad. Sci. U. S. A., 109 (2012), pp. 19333–19338

Takao et al., 2007 Y. Takao, K. Nagasaki, D. Honda Squashed ball-like dsDNA virus infecting a marine fungoid protist *Sicyoidochytrium minutum* (Thraustochytriaceae, Labyrinthulomycetes) Aquat. Microb. Ecol., 49 (2007), pp. 101–108

Taylor and Bruenn, 2009 D.J. Taylor, J. Bruenn The evolution of novel fungal genes from non-retroviral RNA viruses BMC Biol., 7 (2009), p. 88

Turina et al., 2003 M. Turina, A. Prodi, N.K. Van Alfen Role of the Mf1-1 pheromone precursor gene of the filamentous ascomycete *Cryphonectria parasitica* Fungal Genet. Biol., 40 (2003), pp. 242–251

Turina et al., 2007 M. Turina, M. Ricker, R. Lenzi, V. Masenga, M. Ciuffo A severe disease of tomato in the Culiacan area (Sinaloa, Mexico) is caused by a new picorna-like viral species Plant Dis., 91 (2007), pp. 932–941

Vainio et al., 2015 E.J. Vainio, J. Jurvansuu, J. Streng, M.-L. Rajamaki, J. Hantula, J.P.T. Valkonen Diagnosis and discovery of fungal viruses using deep sequencing of small RNAs J. Gen. Virol., 96 (2015), pp. 714–725

Woods and Bevan, 1968 D.R. Woods, E.A. Bevan Studies on the nature of the killer factor produced by *Saccharomyces cerevisiae* fractional precipitation with ammonium sulfate J. Gen. Microbiol., 51 (1968), pp. 115–126

Yu et al., 2010 X. Yu, B. Li, Y. Fu, D. Jiang, S.A. Ghabrial, G. Li, Y. Peng, J. Xie, J. Cheng, J. Huang, X. Yi A geminivirus-related DNA mycovirus that confers hypovirulence to a plant pathogenic fungus Proc. Natl. Acad. Sci. U.S.A., 107 (2010), pp. 8387–8392

Yu et al., 2013 X. Yu, B. Li, Y. Fu, J. Xie, J. Cheng, S.A. Ghabrial, G. Li, X. Yi Extracellular transmission of a DNA mycovirus and its use as a natural fungicide Proc. Natl. Acad. Sci. U.S.A., 110 (2013), pp. 1452–1457

Zerbino, 2010 D.R. Zerbino Using the Velvet de novo assembler for short-read sequencing technologies Curr. Protoc. Bioinf., 31 (2010), pp. 11–15

Zhang et al., 2008 X. Zhang, G.C. Segers, Q. Sun, F. Deng, D.L. Nuss Characterization of hypovirus-derived small RNAs generated in the chestnut blight fungus by an inducible DCL-2-dependent pathway J. Virol., 82 (2008), pp. 2613–2619

Zhao et al., 2011 Q.-Y. Zhao, Y. Wang, Y.-M. Kong, D. Luo, X. Li, P. Hao Optimizing de novo transcriptome assembly from short-read RNA-Seq data: a comparative study BMC Bioinf., 12 (2011), p. S2