Archives of Virology
https://doi.org/10.1007/s00705-018-3712-2

ORIGINAL ARTICLE

Characterization of small RNAs originating from mitoviruses infecting the conifer pathogen *Fusarium circinatum*

E. J. Muñoz-Adalia[1,2] · J. J. Diez[1,2] · M. M. Fernández[1,3] · J. Hantula[4] · E. J. Vainio[4]

Received: 31 August 2017 / Accepted: 27 December 2017

Abstract

Deep sequencing of small RNAs has proved effective in the diagnosis of mycovirus infections. In this study, the presence of mycoviruses in ten isolates of the phytopathogenic fungus *Fusarium circinatum* was investigated by high-throughput sequencing (HTS) of small RNAs. The contigs resulting from *de novo* assembly of the reads were aligned to viral genome sequences. The presence of each mycovirus detected in the isolates was confirmed by RT-PCR analysis with four previously described primer pairs and seven new pairs designed on the basis of sequencing data. The findings demonstrate the potential use of HTS for reconstructing previously identified mitoviruses infecting *F. circinatum*.

Introduction

Fusarium circinatum Nirenberg & O'Donell (teleomorph: *Gibberella circinata* Nirenberg & O'Donell) is the causal agent of pine pitch canker (PPC), one of the most devastating diseases in coniferous forests, plantations and nurseries worldwide. This pathogen severely affects *Pinus* species and also *Pseudotsuga menziesii* (Mirb.) Franco, causing pre- and post-emergence and late damping off on seedlings, with mortality rates of up to 90%. In mature trees, it causes wilting, slow growth and bleeding cankers, reducing the economic yield of the affected timber and increasing the risk of trees breaking during windstorms [1]. The fungus is widespread in several parts of America, Europe, Asia and Africa and is mainly spread via soil- and airborne spores, movement of infected material (seedlings, seeds, pruning tools, etc.), and carrier and vector insects [2, 3]. Although there are no effective means of controlling the disease in the field, biocontrol is currently one of the most promising management options [4].

Fungal viruses (mycoviruses) are frequent in nature, infecting a huge variety of fungi ranging from edible mushrooms to yeasts [5]. Mycoviruses are of great interest in plant pathology, as some are able to reduce the pathogenicity of their hosts (hypovirulence) [6], although many aspects of their biology and ecology remain poorly understood. The best-known case of a mycovirus infection affecting fungal virulence is the chestnut blight pathosystem (causal agent, *Cryphonectria parasitica* (Murrill) M.E. Barr) [7], in which hypovirulence has been used in practice to control *C. parasitica* infections. This has led to increased interest in the use of mycoviruses as biocontrol agents in infections caused by other phytopathogenic fungi [8, 9]. Three different members of the genus *Mitovirus* (family *Narnaviridae*) have been described to be associated with the forest pathogen *F. circinatum*: Fusarium circinatum mitovirus 1, 2-1 and 2-2 (FcMV1, FcMV2-1 and FcMV2-2) [10]. These viruses are rather prevalent among isolates in northern Spain [11]. Despite the interest surrounding this pathogen in forest health, the effects of FcMV1, FcMV2-1 and FcMV2-2 on *F. circinatum* are not yet fully understood. It has been

Handling Editor: Robert H.A. Coutts.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00705-018-3712-2) contains supplementary material, which is available to authorized users.

✉ E. J. Muñoz-Adalia
emigdiojordan.munoz@uva.es; ejordanmunoz@hotmail.com

[1] Sustainable Forest Management Research Institute, University of Valladolid-INIA, Avenida de Madrid 44, 34071 Palencia, Spain

[2] Department of Vegetal Production and Forest Resources, University of Valladolid, Avenida de Madrid 44, 34071 Palencia, Spain

[3] Department of Agroforestry Sciences, University of Valladolid, Avenida de Madrid 44, 34071 Palencia, Spain

[4] Forest Health and Biodiversity, Natural Resources Institute Finland (Luke), Latokartanonkaari 9, 00790 Helsinki, Finland

Published online: 20 January 2018

reported that FcMV1 infection significantly increases fungal virulence in plants [12]; however, the effect of mycoviral infection may differ depending on environmental conditions [13, 14]. Hence, the antiviral response in phytopathogenic fungi deserves further attention.

In recent years, the use of next-generation sequencing technology has improved our knowledge of mycovirus diversity and some aspects of mycovirus-host interactions. Thus, mycoviruses have been successfully detected by deep sequencing of total RNA [15], dsRNA [16] and small RNAs [17–19]. The identification of virus-derived small RNA fragments (vsRNA) by high-throughput sequencing (HTS) has revealed cryptic viruses and suggested that both ascomycete and basidiomycete fungi have an RNA silencing pathway in which vsRNA plays a key role (e.g., [18, 19]).

The RNA silencing process (RNAi) is an important post-transcriptional control pathway that has been described in plants, animals and fungi [20]. This molecular mechanism is involved in some cellular processes such as gene regulation and defence against selfish nucleic acids. Hence, infection by RNA viruses is thought to trigger the RNA silencing response of the host. The RNAi machinery includes a dicer or dicer-like protein (ribonuclease III-like enzyme) that recognizes viral dsRNA molecules and cleaves them to vsRNA of ~21-25 nucleotides in length [21, 22]. The RNA-induced silencing complex (RISC) includes argonaute-like proteins that unwind the paired strands of vsRNA, degrade one of them, and use the other to identify cognate sequences (viral RNA). A ribonuclease-H-like enzyme associated with argonaute then degrades the viral target resulting in an antiviral response [23–25].

In this study, we hypothesized that *F. circinatum* processes viral RNAs into vsRNA in a manner similar to that described for other pathogenic fungi. We investigated whether deep sequencing of small RNAs can reveal new viral strains. The aim of this study was therefore to investigate the practical use of small-RNA deep sequencing for the identification of mycoviruses infecting *F. circinatum*.

Materials and methods

Fungal isolates, total RNA isolation and high-throughput sequencing of small RNAs

A total of ten *F. circinatum* isolates collected in different regions of Spain were selected for study (Table 1). The fungal isolates were cultured in Petri dishes containing PDA medium (3.90% w/v potato dextrose agar, Scharlab S.L., Spain). After culturing for a week in the dark at 25 °C, each isolate was subcultured in cellophane-membrane-covered PDA medium for total RNA extraction.

The total RNA from each isolate was extracted from mycelia ground in liquid nitrogen with the aid of a sterilized pestle. The powdered sample was sequentially treated with TRI Reagent® (Sigma Aldrich Química S.L., Spain) followed by 99% (v/v) chloroform (PanReac Química, Spain), and the total RNA was then precipitated with 70% (v/v) isopropanol (Sigma Aldrich). The pellets were washed with ethanol 75 % v/v (PanReac Química) and resuspended in 20 µl of double-sterilized Milli-Q water. After RNA extraction, samples were transferred to RNase- and DNase-free tubes (Axygen®, USA) and stored at −80 °C. The RNA of each isolate was pooled in two aliquots (20 µl; 2 µl per isolate, final concentration, 1053 ng μl^{-1}). The quality and quantity of the RNA in one pooled extract was measured in a NanoDrop 2000 Spectrophotometer (Thermo Fisher Scientific, USA) and the other aliquot was stored at −80 °C to preserve the integrity of the RNA. The same pooled extract was run in a 1.20% (w/v) agarose D1 low-EEO (Conda Laboratories, Spain) gel, first for 10 min at 70 v and then for 65 min at 50 v. After that, the gel was submerged for 25 min in 10%

Table 1 Isolates of *Fusarium circinatum* used in this study: geographical origin of the strains and the living organisms from which *F. circinatum* was isolated (source) are provided. Viral infection by Fusarium circinatum mitovirus 1, 2-1, 2-2 (FcMV1, FcMV2-1 and FcMV2-2, respectively), contig 50 and contig 571 (model 310) are indicated by the symbol +

Isolate	Origin	Source	FcMV1	FcMV2-1	FcMV2-2	Contig 50	Contig 571
FC5	Asturias	*Pinus canariensis* Sweet ex K. Spreng	-	-	-	-	-
FC122	Cantabria	*Tomicus piniperda* L.	-	+	-	-	+
FC179	Cantabria	*T. piniperda*	+	+	+	+	+
FC213	Cantabria	*Pityophthorus pubescens* Marsham	+	+	-	-	+
FC921	Cantabria	*T. piniperda*	-	+	-	-	-
Va70	Cantabria	*Pinus radiata* D. Don	-	-	-	-	-
FC13	Castile and León	*P. radiata*	-	-	-	-	-
FC14	Castile and León	*P. radiata*	-	-	-	-	-
FC20	Galicia	*Pinus nigra* Arnold	-	+	-	-	+
FC24	Basque Country	*P. radiata*	-	-	-	-	-

(v/v) ethidium bromide (Merck, Germany) staining solution and then visualized under UV light. The sizes of the bands were estimated by comparison with DNA Molecular Marker II (Roche Life Science, Spain).

The remaining pool of RNA was sent to Fasteris SA (Switzerland; https://www.fasteris.com) for siRNA library construction and HTS (Illumina HiSeq 2500, Illumina Inc., USA). The sample processing consisted of acrylamide gel purification of small fragments of RNA, single-stranded ligation (3′ and 5′ adapters) and cDNA library generation by reverse transcription and PCR.

Alignment of Illumina sequence reads and *de novo* assembly

For the analysis, Illumina adapter sequences were removed from reads, using Trimmomatic software v0.32 [26]. Exploratory analysis of reads was performed with Geneious Pro 6.0.6. [27] (Mapping parameters: sensitivity, medium-low; iterations for consensus, two; minimum mapping quality, 99.90%; word length, 18 nt; maximum ambiguity, 4 states), and the complete reads dataset (1-50 nt) was mapped against the following queries: (a) genome of *F. circinatum*, (b) genome of the fungal mitochondria, (c) previously known *F. circinatum* mitoviruses (FcMV1, FcMV2-1 and FcMV2-2), and (d) contigs of interest obtained in this study (see below). The size distribution of reads corresponding to viral genomes (vsRNA) was visualized in a histogram, and the mean size was calculated for each strain of virus. All reads between 19 and 35 nt were mapped against viral genomes using MISIS v2.7 [28] to identify sense and antisense vsRNA. In addition, the hotspots of vsRNA accumulation within viral genomes were analyzed as possible recognition sites for dicer proteins. The peak values were first identified using the "peakPick" package [29] of software R [30] (analysis window, ± 10 positions; limit, 12 standard deviations). All peaks that were not covered by > 100 vsRNAs (sense and antisense) were then removed, and the range of the mean read size of each hotspot was calculated.

De novo assembly was conducted by computing a total of 33 models in the software Velvet 1.2.10 [31]. Each model was defined by three parameters: (a) read size (b) k-mer size, and (c) whether or not AssemblyAssembler 1.4 script (Jacob Crawford, Cornell University) was used for the analysis. The complete dataset of reads (read size between 1 and 50 nt) was used in a total of 12 analyses. In parallel, the complete dataset of reads was mapped against the host genome (i.e., the *F. circinatum* genome, accession ASM49732v2) using Bowtie software 1.2.0 [32]. The resulting non-aligned reads (i.e., reads that did not match with any region of the fungal genome) were selected for a total of 15 analyses in Velvet. The Bowtie output files showed that most informative reads were ≥14 bp long, and therefore the reads between 14 and 31 bp were selected for six additional models. All k-mer values used for analysis were between 9 and 31 bp (Online Resource 1). A total of nine models were run using the post-assembly script AssemblyAssembler 1.4. as reported previously for mycovirus assembly [18].

The models that simultaneously provided the highest values of N50 and number of contigs generated during Velvet computation were selected for further analysis. The contigs were compared against known sequences in the NCBI GenBank Viruses database (taxid: 10239; https://www.ncbi.nlm.nih.gov/) using BLASTx and BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify viral sequences. The resulting contigs of these models were then aligned against the FcMV1, FcMV2-1 and FcMV2-2 genomes (GenBank accession numbers KF803546.1, KF803547.1 and KF803548.1, respectively) in Geneious Pro 6.0.6., in order to estimate the number of effective contigs (contigs that corresponded to the viral genome). Contigs with similarity of less than 90% to the reference sequence were removed from the alignment. The percentage sequence coverage was calculated to enable selection of the best model for each mycovirus.

Analysis of viral prevalence

RNA samples from the isolates included in pool preparation (see above) were used as templates for RT-PCR with selective primers. cDNA was synthesised using random hexamer primers and PrimeScript™ II Reverse Transcriptase (Takara Bio USA Inc., USA) as described previously [33]. The subsequent PCR reactions were carried out using a KAPA Taq PCR Kit (Kapa Biosystems, USA) and eleven selective primer pairs (Table 2). The PCR reaction volume was 50 µl, and the protocol consisted of denaturation for 10 min at 95 °C followed by 37 cycles of 30 s at 95 °C, 45 s at the different annealing temperatures summarized in Table 2 and 2 min at 72 °C. The final elongation step consisted of 7 min at 72 °C [11]. The PCR products were separated for 50 min at 110 v in a 1.60% (w/v) agarose D1 low-EEO gel stained with 0.004% (v/v) 10,000x GelRed™ (Biotium, USA). The size of the resulting fragments was determined by comparison with a DNA Molecular Marker II (Roche). The products of PCR reactions performed with the primer pairs FcontFOR/FcontREv and FcirCONT/FcirCONTRev were analyzed in a 2% (w/v) agarose D1 low-EEO gel under the above-mentioned staining and electrophoretic conditions. In these cases, the amplicon size was estimated by comparison with a 50 bp DNA ladder (Nippon Genetics Europe, Germany). New primer pairs (Table 2) were designed using Primer3Plus software [34], and their homology to query sequences was confirmed using Geneious Pro. 6.0.6. The amplicons were sent to Macrogen Europe Inc.

Table 2 Primer pairs used for identification of FcMV1, FcMV2-1 and FcMV2-2 by RT-PCR

Primer pair	Forward primer	Reverse primer	Annealing temperature (°C)	Amplicon size (bp)	Virus preference	Reference
FMC1F1/FMC1F1rev1	5'-CGTGGATTAAAACCCACAAA-3'	5'-TGGTAATCTACCATAGCAATTAYTC-3'	49.5	440	FcMV1	[11]
FMC3F1/FMC3Rev1	5'-GAYAGAACTTTTACTCAAGATCC -3'	5'-ATTCATCTYTTGGCAAATTCATA-3'	47.5	461	FCMV2-1* and FCMV2-2*	[11]
FMC2MFor/FMC3Rev1	5'-GCATCAAATAGTGCTCTGAC-3'	5'-ATTCATCTYTTGGCAAATTCATA-3'	49.5	821	FCMV2-1*	[11]
FMC3MidF/FMC3Rev1	5'-TCAACCATAACTGATCCATGT-3'	5'-ATTCATCTYTTGGCAAATTCATA-3'	48.5	996	FCMV2-2*	[11]
FMV1BL/FMV1BLrev	5'-AGGTCAACCTATGGGAGCAT-3'	5'-AGACCACTTATTTCTTTCCCTGA-3'	51	282	FcMV1	This study
FM22for/FM22rev	5'-TGGTTTTGCCAAGGGGTGAC-3'	5'-TACCTCTACTTGGTAGTCCAGTG-3'	53	731	FCMV2-2*	This study
FMNGS1for/FMNGS1rev	5'- CGTGACAGCCCCCTCTTATG-3'	5'-TGACCTTGTAGAATTTCAAGTTGATT-3'	54.5	430	FcMV1	This study
FMNGS21fr/FMNGS21rv	5'-AAAGAAACTTCATGATCAGTGCTT-3'	5'-CGTCTGTCAAGACGAAACTTTA-3'	51	970	FCMV2-1*	This study
FMHTS22/FMHTS22rev	5'-AAACTCTCAAAGAGGTCCAAGG-3'	5'-TCGTATTCAAACATTTCACCAA-3'	51.5	1130	FCMV2-2*	This study
FcontFOR/FcontREV	5'-GGTCTCGTACGTAATGAAATTCAAC-3'	5'-CAATGGTTAACAATGTGGCATA-3'	54.3	163	contig 50* (Model 310)	This study
FcirCONT/FcirCONTRev	5'-CAGTGACAGACTACTTTATAG-3'	5'-ATAGCTCAAGATGAATAACAGCCC ATCGG-3'	42.4	60	contig 571* (Model 310)	This study

*Sequencing of the amplified fragment is required for the identification of each virus strain

(The Netherlands; www.macrogen.com) for DNA purification and sequencing.

Sequences obtained using selective primers and Sanger sequencing were trimmed using Geneious Pro. 6.0.6 and compared to sequences available in the GenBank (NCBI) database using MegaBLAST (nr/nt database; taxid: 10239). All sequences obtained by direct sequencing of amplicons were aligned against FcMV1, FcMV2-1 and FcMV2-2.

Results

Yield and size distribution of the vsRNA reads

The small-RNA deep sequencing analysis included ten *F. circinatum* strains isolated from plant material and bark beetles (Table 1). The quality of the total RNA extracts used for the small-RNA library construction was acceptable, based on spectrophotometry (absorbance ratio 260/280, 1.85) and agarose gel analysis (visible bands of 18S and 28S rRNA). The total number of reads produced by the Illumina sequencer was 24,415,829 (insert size, 1-50 bp), corresponding to 1230 megabases. The proportion of clusters that fulfilled the default Illumina quality criteria was 95.73%, and the proportion of bases with a quality score ≥ 30 (Q30) was 96.12%. The numbers of reads per insert size were distributed as follows: 60.26% inserts of 27-50 bp, 38.56% inserts of 18-26 bp, and 1.17% inserts with a size ≤ 17 bp (Fig. 1).

Regarding the annotation of reads, 86.55% of the obtained reads were assigned to the queries summarized in Figure 2. The sequence coverage was 97.90% for FcMV1 and more than 98.50% for FcMV2-1 and FcMV2-2 (Fig. 3). The mean read sizes that mapped with viral genomes were 25.52 ± 0.18 nt for FcMV1, 24.27 ± 0.10 nt for FcMV2-1, 24.99 ± 0.14 nt for FcMV2-2 (mean and standard error; Fig. 4). A total of 11, 21 and 28 hotspots fulfilled the selection criteria for plausible dicer recognition sites in FcMV1, FcMV2-1 and FcMV2-2, respectively. The mean size of reads accumulated in the hotspots ranged between 21.80 and 31.41 nt in FcMV1, 20.89 and 29.55 nt in FcMV2-1, and 21.27 and 30.80 nt in FcMV2-2.

Contig assembly and mapping

The highest values of N50 and the resulting contigs (Online Resource 1) were obtained for models 204, 310 and 360. The viral contigs provided by these models were identified on the basis of comparison of sequences with those included in the NCBI GenBank Viruses database (taxid: 10239) by using BLASTx and BLASTn. Using this approach, we detected 47, 40 and 40 contigs per model (i.e., models 204, 310 and 360, respectively),

corresponding to previously known *F. circinatum* mitoviruses and two contigs that were apparently more similar to mitoviruses infecting other fungal species (i.e., contig 50 and 571 from model 310, Table 3). The possible origin of these contigs was examined on the basis of RT-PCR with specific primers (see below). However, isolate FC179 was simultaneously infected by the three mitovirus strains and showed positive amplification for the two contigs considered (see below). In addition, only one FcMV2-2 infection was identified among the host isolates, whereas FcMV2-1 was represented by multiple strains (N = 5, Table 1).

Based on mapping the contigs against the genomes of known *F. circinatum* mitoviruses, 0.21% of contigs resulting from model 204 corresponded to the FcMV1 genome, and 0.22% corresponded to both FcMV2-1 and FcMV2-2. With model 310, 0.37% of the contigs corresponded to FcMV1, 0.37% to FcMV2-1 and 0.40% to FcMV2-2. With model 360, 0.37% of the contigs corresponded to the genome of FcMV1, 0.37% to FcMV2-1 and 0.40% to FcMV2-2. The three models showed more than 50% coverage of the viral genome (Table 4), with model 310 and model 360 providing the best coverage values.

Detection of viruses by RT-PCR and mitovirus identification

All of the PCR primers used yielded amplification products. The primer pair FMC3F1/FMC3Rev1 amplified sequences of FcMV2-1 and FcMV2-2 indistinctly, while the primer pairs FMC1F1/FMC1F1rev1, FMV1BL/FMV1BLrev, FMNGS1for/FMNGS1rev and FcontFOR/FcontREv produced amplicons only in the isolates infected by FCMV1. Similarly, the primer pairs FMC2MFor/FMC3Rev1 and FMNGS21fr/FMNGS21rv only amplified sequences that corresponded to FcMV2-1, while the primer pairs FMC3MidF/FMC3Rev1, FM22for/FM22rev and FMHTS22/FMHTS22rev amplified fragments of FcMV2-2.

Two new primer pairs were designed based on the contigs 50 and 571 (Table 3). Primer pair FcontFOR/FcontREv amplified a fragment of the expected size (163 nt) in one host

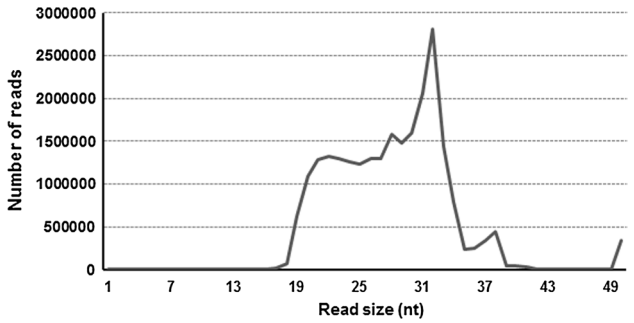

Fig. 1 Size distribution of small RNA reads obtained using Illumina HTS

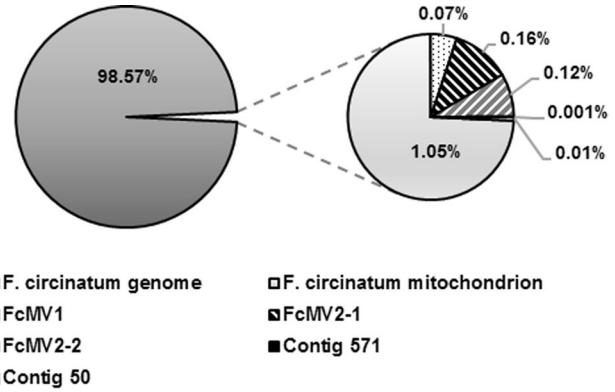

Fig. 2 Distribution of annotated reads from HTS analysis by category

isolate (i.e., FC179). The genome sequence of FcMV2-2 that had been deposited in the GenBank database was incomplete, and we therefore had to determine whether contig 50 represents the 3′- proximal sequence that was lacking for this virus. We found that contig 50 could be aligned with the sequence of a previously unpublished cloned genome fragment from isolate FcCa070 [10], linking it with the end of the FcMV2-2 sequence (Online Resource 2). In addition, the whole dataset was mapped against FcMV2-2, including contig 50 at the 3′ end (Fig. 3d), and the distribution of read sizes was also investigated, revealing a mean size of 24.94 ± 0.13 nt in length (Fig. 4). The primer pair FcirCONT/FcirCONTRev amplified fragments of similar sample size to that expected for contig 571 in four host isolates (Table 1), irrespective of the presence or absence of previously known *F. circinatum* mitovirus strains. However, the sequences obtained were not informative due to the short length of the amplicons (< 50 bp) (Tables 1 and 2). Otherwise, contig 571 shared ~60-65% global nt sequence identity with the three previously known *F. circinatum* mitoviruses.

A total of 50% of the isolates were virus-free, while 10% had a single infection and 40% of the *F. circinatum* strains were co-infected (Table 1). More specifically, FC921 was infected with FcMV2-1, while FC20 and FC122 were co-infected with FcMV2-1 and a possible new mitovirus variant (contig 571). In addition, multiple virus infection was observed in two of the isolates: isolate FC213 was infected with FcMV1, FcMV2-1 and contig 571, and isolate FC179 was infected with FcMV1, FcMV2-1, FcMV2-2 and contig 571.

Discussion

Next-generation sequencing is a powerful tool for the analysis of viral infections [17, 35]. In this study, HTS of vsRNA was used for the first time to recover members of the family

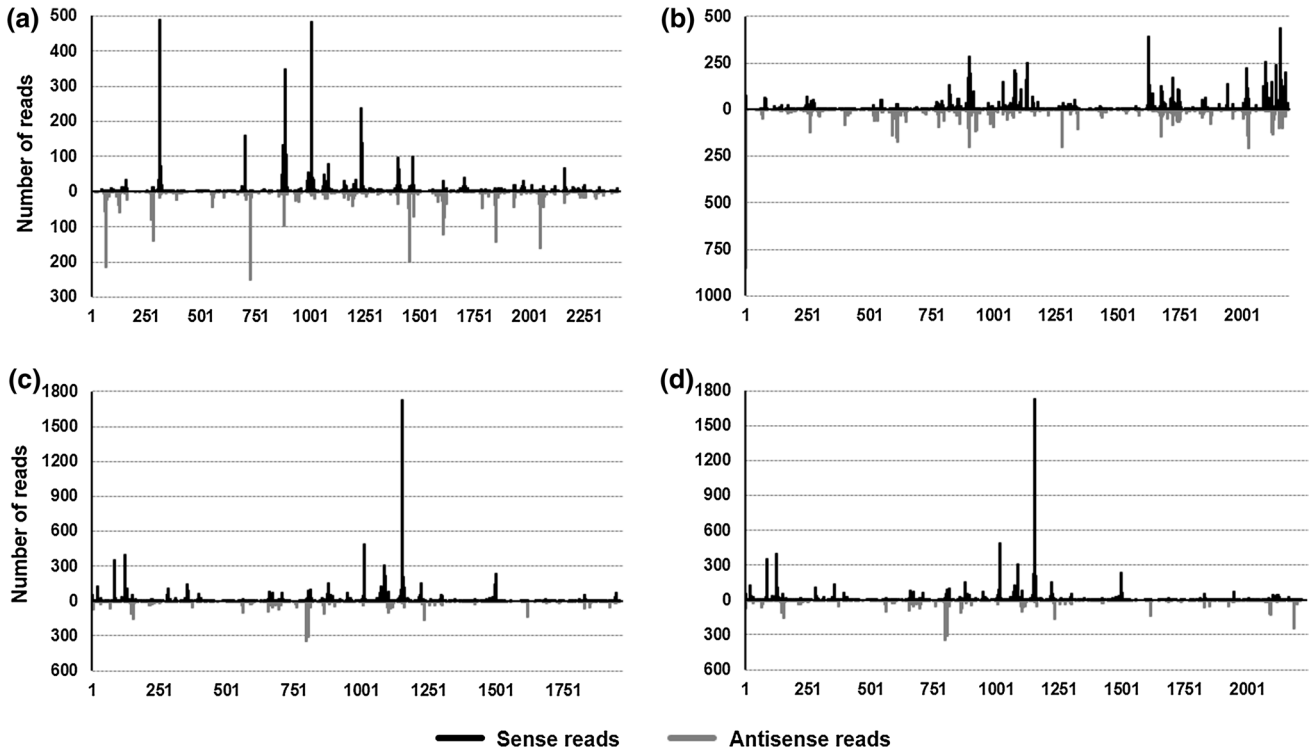

Fig. 3 Distribution of reads (19-35 bp) from small-RNA sequencing along the (**a**) FcMV1 complete genome, (**b**) FcMV2-1 complete genome, (**c**) FcMV2-2 partial genome and (**d**) FcMV2-2 partial genome including contig 50 at the 3′ extreme

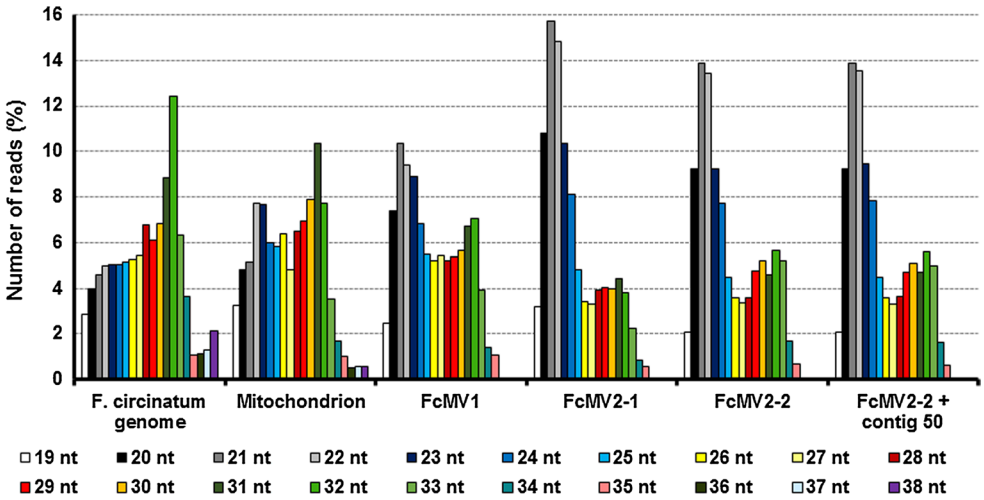

Fig. 4 Size distribution of small RNA reads matching with the genome of *F. circinatum*, the genome of the fungal mitochondrion, and mitoviruses hosted by *F. circinatum* (FcMV1, FcMV2-1, FcMV2-2 and FcMV2-2 including contig 50). Sizes with relative abundance lower than 0.50% are not shown

Narnaviridae infecting *F. circinatum*, demonstrating the suitability of this approach, as shown previously for mycoviruses infecting *Heterobasidion* spp. [18]. Vainio *et al.* [18] used a similar computing method and reported that 3.70% of resulting contigs corresponded to viruses, while less than 1% did so in this study. Similarly, Donaire and Ayllón [19] found a higher proportion of reads corresponding to viral genomes. By contrast, similar yields to those found for *F. circinatum* mitoviruses were obtained in other studies focused on plant viruses and mycoviruses [17, 36]. The low percentage of viral sequences detected might be explained by the low quality of the RNA, although the Q30 values and the low proportion of reads ≤ 17 nt (Fig. 1) did not indicate degradation of the RNA. Possible biological explanations include a low rate of viral replication (latency) at the time when RNA was extracted or reduced activity of the RNAi machinery. In this regard, it is known that mitoviruses replicate inside mitochondria, while RNA silencing mainly occurs in the

Table 3 Best matches of contig 50 and contig 571 (Model 310) according to BLAST searches against the GenBank database. L, length of each contig expressed as number of nucleotides; Database, nr (non-redundant protein sequences) and nr/nt (nucleotide collection); QC, Query coverage; E, E-value; ID, identity

Model	Contig	L	Algorithm	Database	Limit	Match ID	QC (%)	E	ID (%)	Accession number
310	50	216	BLASTx	nr	Viruses (taxid:10239)	RNA-dependent RNA polymerase [Fusarium poae mitovirus 1]	96 %	< 0.01	80 %	YP_009272898.1
			BLASTn	nr/nt	Viruses (taxid:10239)	Fusarium poae mitovirus 1 genomic RNA, complete genome	97 %	< 0.01	81 %	LC150564.1
310	571	63	BLASTx	nr	Viruses (taxid:10239)	RNA-dependent RNA polymerase [Sclerotinia sclerotiorum mitovirus 12]	93 %	< 0.01	73 %	AHF48628.1
			BLASTn	nr/nt	Viruses (taxid:10239)	Rhizoctonia solani mitovirus 11 isolate 42304-9b RNA-dependent RNA polymerase gene, complete cds	73 %	< 0.01	80 %	KP900906.1

Table 4 Sequence coverage of the FcMV1, FcMV2-1 and FcMV2-2 genomes by the resulting effective contigs of the three best models

Mycovirus	Model 204	Model 310	Model 360
FcMV1	39.7 %	59.9 %	59.9 %
FcMV2-1	52.1 %	52.7 %	52.7 %
FcMV2-2*	35.9 %	53.5 %	53.5 %

*partial genome

cytoplasm. However, effective RNAi has been reported in other fungi infected by members of the family *Narnaviridae* including members of the genus *Mitovirus*. [19]. Thus, the mitochondrion embodies a key element in signalling pathways that trigger the antiviral response in humans and other mammals [37, 38]. Recently, Nibert [39] investigated the relationship between the presence of UGA codons in the genome of mitoviruses and in the genetic code of the host mitochondrion. This triplet is read as a tryptophan codon in the genetic code of mitochondria but as a stop codon in the standard code. The findings reported by Nibert [39] suggest that mitoviruses hosted by *F. circinatum* (i.e., FcMV1 and FcMV2-1), which have a high percentage of UGA codons in their genomes, are unlikely to be translated in the cytosol. This finding might explain the low proportion of reads provided by HTS in this study. On the other hand, a signalling pathway involving mitochondria and possible RNA silencing activity inside mitochondria [19] cannot be ruled out as elements of the antiviral response in this fungus.

To our knowledge, this study represents the first examination of the molecular antiviral response in *F. circinatum*, and hence, vsRNA analysis might provide insights into this process. As shown in Figure 4, the most informative read size ranged between 19 and 35 nt (i.e., >50% of reads between 20-26 nt in each virus strain), which is consistent with the size of vsRNA reported in other filamentous fungi [18, 40, 41]. For *de novo* assembly, either 14 to 31-nt (Bowtie output) or 1 to 50-nt reads were used, resulting in a similar contig yield as when a k-mer of 17 was used. These results show that multiple model computation is required to prevent underestimating reads with biological significance. Chen *et al.* [36] also observed multiple vsRNA sizes among HTS reads and suggested the possible participation of multiple dicer-like proteins in RNAi of maize (*Zea mays* L.). In chestnut blight, two dicer proteins (genes *dcl-1-2*) have been identified, and their participation in the antiviral response has been reported [42]. Similarly, four argonaute-like genes (*agl1-4*) have also been characterized in *C. parasitica*, with the *agl2* gene being the only one required for the induction of RNA silencing of viruses [43]. Otherwise, the results reported here confirmed several hotspots along the viral genome reported by Donaire *et al.* [44], possibly suggesting multiple recognition sites for dicer proteins. Further studies

are therefore required to characterize the pathways involving the RNAi machinery (e.g., cleavage motifs for dicer) in *F. circinatum*.

In this study, the viral prevalence was moderate to high (i.e., 50% of isolates were infected by mycoviruses), while the coinfection rate was moderate despite the small sample size. These findings are consistent with those of Vainio *et al.* [11], who reported that isolates from the Cantabria region showed moderate-high rates of single and double infection (>80% in some locations), while none of the *F. circinatum* strains from the Basque Country showed infection. Regarding other geographical locations, isolates from Castile and León and from Asturias did not show any infection; however, 40% of the isolates from Asturias were infected by FcMV1 in the above-mentioned study. Interestingly, the isolate from Galicia (FC20) was coinfected, although this was not detected in an earlier study [11]. The most plausible explanation for the new detection is the use of novel primer pairs or the possible low rate of viral replication suggested above. Mycoviruses appear to be more abundant in Cantabria than in other locations. However, more samples were colleted there than in other parts of the Iberian Peninsula. Schoebel *et al.* [45] studied the genetic diversity of Hymenoscyphus fraxineus mitovirus 1 (HfMV1; family *Narnaviridae*) throughout the distribution range of its host, the invasive fungus *Hymenoscyphus fraxineus* (T. Kowalski) Baral, Queloz & Hosoya. The researchers concluded that two phylogenetic groups of viruses are present in Europe, which is consistent with the proposed introduction of two individuals of *H. fraxineus* from Asia. Considering all of these data, we conclude that there is no clear pattern in the distribution of mycoviruses hosted by *F. circinatum* in Spain. Nonetheless, further studies are required to clarify whether the prevalence of FcMV1, FcMV2-1 and FcMV2-2 is related to a founder effect, as two introductions of *F. circinatum* in Spain have been suggested [46]. In this study, small-RNA deep sequencing revealed two contigs that showed similarities to mycoviruses not previously reported as being hosted by *F. circinatum*. Regarding the sequencing and read mapping results, we consider it likely that contig 50 represents the 3′ end of the incomplete genome of FcMV2-2. However, the positive RT-PCR amplification using specific primers did not provide enough information to definitively support identification of contig 571 as part of a new viral strain.

In summary, we used deep sequencing to investigate the infection of *Mitovirus* spp. in the forest pathogen *F. circinatum*. The findings show that HTS followed by *de novo* assembly is suitable for detecting previously identified mitoviruses, even at low viral prevalence. By contrast, this method was of limited value for characterizing new viral strains. Finally, we describe new primer pairs for use in future research focused on viral infection of *F. circinatum*.

Acknowledgements The authors thank Instituto Agroforestal Mediterráneo (Universidad Politécnica de Valencia) for providing the isolates Va70, FC5, FC13, FC14, FC20 and FC24. The authors also thank Diana Bezos for performing the morphological and molecular identification of the isolates FC122, FC179, FC213 and FC921, Milagros de Vallejo and Juan Blanco (Gobierno de Cantabria) for their help in carrying out the study, and Elena Hidalgo (iuFOR, UVa-INIA), María Teresa Pérez-García, Pilar Cidad and Marycarmen Arévalo (IBGM, UVa-CSIC) for their collaboration and valuable advice during laboratory work.

Compliance with ethical standards

Funding This study was supported by two research projects: AGL2015-69370-R "Next Generation Sequencing NGS) technologies for the study of *Fusarium circinatum* mycoviruses" (MINECO/FEDER, UE) and AGL2012-39912 "Biological control of Pine Pitch Canker disease by the use of *Fusarium circinatum* mycoviruses" (Ministerio Economía y Competitividad). This article is based upon work from COST Action FP1406 PINESTRENGTH (Pine pitch canker - strategies for management of *Gibberella circinata* in greenhouses and forests), supported by COST (European Cooperation in Science and Technology). E. J. Muñoz-Adalia is a receipt of grants from the European Social Fund and from the Consejería de Educación de Castilla y León (JCyL) (ORDEN EDU/1083/2013).

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Wingfield MJ, Hammerbacher A, Ganley RJ et al (2008) Pitch canker caused by *Fusarium circinatum*—a growing threat to pine plantations and forests worldwide. Australas Plant Pathol 37:319–334. https://doi.org/10.1071/AP08036
2. Brockerhoff EG, Dick M, Ganley R et al (2016) Role of insect vectors in epidemiology and invasion risk of *Fusarium circinatum*, and risk assessment of biological control of invasive *Pinus contorta*. Biol Invasions 18:1177–1190. https://doi.org/10.1007/s10530-016-1059-8
3. Bezos D, Lomba JM, Martinez-Álvarez P et al (2012) Effects of pruning in monterrey pine plantations affected by *Fusarium circinatum*. For Syst 21:481–488
4. Bezos D, Martínez-Álvarez P, Fernández M, Diez JJ (2017) Epidemiology and management of Pine Pitch Canker Disease in Europe—a review. Balt For 23:279–293
5. Pearson MN, Beever RE, Boine B, Arthur K (2009) Mycoviruses of filamentous fungi and their relevance to plant pathology. Mol Plant Pathol 10:115–128. https://doi.org/10.1111/j.1364-3703.2008.00503.x
6. Xie J, Jiang D (2014) New insights into mycoviruses and exploration for the biological control of crop fungal diseases. Annu Rev Phytopathol 52:45–68. https://doi.org/10.1146/annurev-phyto-102313-050222
7. Zamora P, Martín AB, San Martín R et al (2014) Control of chestnut blight by the use of hypovirulent strains of the fungus *Cryphonectria parasitica* in northwestern Spain. Biol Control 79:58–66. https://doi.org/10.1016/j.biocontrol.2014.08.005

8. Rigling D, Heiniger U, Hohl HR (1989) Reduction of laccase activity in dsRNA-containing hypovirulent strains of *Cryphonectria* (*Endothia*) *parasitica*. Phytopathology 79:219–223. https://doi.org/10.1094/Phyto-79-219
9. Peever TL, Liu YC, Cortesi P, Milgroom MG (2000) Variation in tolerance and virulence in the chestnut blight fungus–hypovirus interaction. Appl Environ Microbiol 66:4863–4869. https://doi.org/10.1128/AEM.66.11.4863-4869.2000
10. Martínez-Álvarez P, Vainio EJ, Botella L et al (2014) Three mitovirus strains infecting a single isolate of *Fusarium circinatum* are the first putative members of the family *Narnaviridae* detected in a fungus of the genus *Fusarium*. Arch Virol 159:2153–2155. https://doi.org/10.1007/s00705-014-2012-8
11. Vainio EJ, Martínez-Álvarez P, Bezos D et al (2015) *Fusarium circinatum* isolates from northern Spain are commonly infected by three distinct mitoviruses. Arch Virol 160:2093–2098. https://doi.org/10.1007/s00705-015-2462-7
12. Muñoz-Adalia EJ, Flores-Pacheco JA, Martínez-Álvarez P et al (2016) Effect of mycoviruses on the virulence of *Fusarium circinatum* and laccase activity. Physiol Mol Plant Pathol. https://doi.org/10.1016/j.pmpp.2016.03.002
13. Vainio EJ, Korhonen K, Tuomivirta TT, Hantula J (2010) A novel putative partitivirus of the saprotrophic fungus *Heterobasidion ecrustosum* infects pathogenic species of the *Heterobasidion annosum* complex. Fungal Biol 114:955–965. https://doi.org/10.1016/j.funbio.2010.09.006
14. Botella L, Dvořák M, Capretti P, Luchi N (2017) Effect of temperature on GaRV6 accumulation and its fungal host, the conifer pathogen *Gremmeniella abietina*. For Pathol 47:1–12. https://doi.org/10.1111/efp.12291
15. Marzano SYL, Domier LL (2016) Novel mycoviruses discovered from metatranscriptomics survey of soybean phyllosphere phytobiomes. Virus Res 213:332–342. https://doi.org/10.1016/j.virusres.2015.11.002
16. Osaki H, Sasaki A, Nomiyama K, Tomioka K (2016) Multiple virus infection in a single strain of *Fusarium poae* shown by deep sequencing. Virus Genes 52:835–847. https://doi.org/10.1007/s11262-016-1379-x
17. Nerva L, Ciuffo M, Vallino M et al (2016) Multiple approaches for the detection and characterization of viral and plasmid symbionts from a collection of marine fungi. Virus Res 219:22–38. https://doi.org/10.1016/j.virusres.2015.10.028
18. Vainio EJ, Jurvansuu J, Streng J et al (2015) Diagnosis and discovery of fungal viruses using deep sequencing of small RNAs. J Gen Virol 96:714–725. https://doi.org/10.1099/jgv.0.000003
19. Donaire L, Ayllón MA (2016) Deep sequencing of mycovirus-derived small RNAs from *Botrytis* species. Mol Plant Pathol. https://doi.org/10.1111/mpp.12466
20. Schumann U, Ayliffe M, Kazan K, Wang M-B (2010) RNA silencing in fungi. Front Biol 5:478–494. https://doi.org/10.1007/s11515-010-0550-3
21. Tinoco MLP, Dias BBA, Dall'Astta RC et al (2010) *In vivo* trans-specific gene silencing in fungal cells by *in planta* expression of a double-stranded RNA. BMC Biol 8:27. https://doi.org/10.1186/1741-7007-8-27
22. Tauati SJ, Pearson MN, Choquer M et al (2014) Investigating the role of dicer 2 (*dcr2*) in gene silencing and the regulation of mycoviruses in *Botrytis cinerea*. Microbiology 83:140–148. https://doi.org/10.1134/S0026261714020180
23. Hammond TM, Andrewski MD, Roossinck MJ, Keller NP (2008) *Aspergillus* mycoviruses are targets and suppressors of RNA silencing. Eukaryot Cell 7:350–357. https://doi.org/10.1128/EC.00356-07
24. Chen Y, Gao Q, Huang M et al (2015) Characterization of RNA silencing components in the plant pathogenic fungus *Fusarium graminearum*. Sci Rep 5:12500. https://doi.org/10.1038/srep12500
25. Zhang DX, Spiering MJ, Nuss DL (2014) Characterizing the roles of *Cryphonectria parasitica* RNA-dependent RNA polymerase-like genes in antiviral defense, viral recombination and transposon transcript accumulation. PLoS One. https://doi.org/10.1371/journal.pone.0108653
26. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170
27. Kearse M, Moir R, Wilson A et al (2012) Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199
28. Seguin J, Otten P, Baerlocher L et al (2014) MISIS: a bioinformatics tool to view and analyze maps of small RNAs derived from viruses and genomic loci generating multiple small RNAs. J Virol Methods 195:120–122. https://doi.org/10.1016/j.jviromet.2013.10.013
29. Weber CM, Ramachandran S, Henikoff S (2014) Nucleosomes are context-specific, H2A.Z-Modulated barriers to RNA polymerase. Mol Cell 53:819–830. https://doi.org/10.1016/j.molcel.2014.02.014
30. R Development Core Team (2015) R: a language and environment for statistical computing. 55:275–286. https://www.r-project.org/
31. Zerbino DR, Birney E (2008) Velvet: algorithms for *de novo* short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107
32. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10:R25. https://doi.org/10.1186/gb-2009-10-3-r25
33. Vainio EJ, Keriö S, Hantula J (2011) Description of a new putative virus infecting the conifer pathogenic fungus *Heterobasidion parviporum* with resemblance to *Heterobasidion annosum* P-type partitivirus. Arch Virol 156:79–86. https://doi.org/10.1007/s00705-010-0823-9
34. Untergasser A, Cutcutache I, Koressaar T et al (2012) Primer3-new capabilities and interfaces. Nucleic Acids Res 40:1–12. https://doi.org/10.1093/nar/gks596
35. Deakin G, Dobbs E, Bennett JM et al (2017) Multiple viral infections in *Agaricus bisporus*—characterisation of 18 unique RNA viruses and 8 ORFans identified by deep sequencing. Sci Rep. https://doi.org/10.1038/s41598-017-01592-9
36. Chen S, Huang Q, Wu L, Qian Y (2015) Identification and characterization of a maize-associated mastrevirus in China by deep sequencing small RNA populations. Virol J 12:156. https://doi.org/10.1186/s12985-015-0384-3
37. McBride HM, Neuspiel M, Wasiak S (2006) Mitochondria: more than just a powerhouse. Curr Biol 16:551–560. https://doi.org/10.1016/j.cub.2006.06.054
38. Moore CB, Ting JPY (2008) Regulation of mitochondrial antiviral signaling pathways. Immunity 28:735–739. https://doi.org/10.1016/j.immuni.2008.05.005
39. Nibert ML (2017) Mitovirus UGA(Trp) codon usage parallels that of host mitochondria. Virology 507:96–100. https://doi.org/10.1016/j.virol.2017.04.010
40. Dang Y, Yang Q, Xue Z, Liu Y (2011) RNA interference in fungi: pathways, functions, and applications. Eukaryot Cell 10:1148–1155. https://doi.org/10.1128/EC.05109-11
41. Zhang X, Segers GC, Sun Q et al (2008) Characterization of hypovirus-derived small RNAs generated in the chestnut blight fungus by an inducible DCL-2-dependent pathway. J Virol 82:2613–2619. https://doi.org/10.1128/JVI.02324-07
42. Segers GC, Zhang X, Deng F et al (2007) Evidence that RNA silencing functions as an antiviral defense mechanism in

fungi. Proc Natl Acad Sci USA 104:12902–12906. https://doi.org/10.1073/pnas.0702500104
43. Sun Q, Choi GH, Nuss DL (2009) A single Argonaute gene is required for induction of RNA silencing antiviral defense and promotes viral RNA recombination. Proc Natl Acad Sci USA 106:17927–17932. https://doi.org/10.1073/pnas.0907552106
44. Donaire L, Wang Y, Gonzalez-Ibeas D et al (2009) Deep-sequencing of plant viral small RNAs reveals effective and widespread targeting of viral genomes. Virology 392:203–214. https://doi.org/10.1016/j.virol.2009.07.005
45. Schoebel CN, Botella L, Lygis V, Rigling D (2017) Population genetic analysis of a parasitic mycovirus to infer the invasion history of its fungal host. Mol Ecol. https://doi.org/10.1111/mec.14048
46. Berbegal M, Pérez-Sierra A, Armengol J, Grünwald NJ (2013) Evidence for multiple introductions and clonality in Spanish populations of *Fusarium circinatum*. Phytopathology 103:851–861. https://doi.org/10.1094/PHYTO-11-12-0281-R