FEMS Microbiology Ecology, 95, 2019, fiz084

doi: 10.1093/femsec/fiz084
Advance Access Publication Date: 10 June 2019
Research Article

RESEARCH ARTICLE

A novel metabarcoding approach to investigate *Fusarium* species composition in soil and plant samples

José F. Cobo-Díaz[1,*,†,‡], Riccardo Baroncelli[2,†,¶], Gaétan Le Floch[1,§] and Adeline Picot[1,¥]

[1]Laboratoire Universitaire de Biodiversité et Ecologie Microbienne, IBSAM, ESIAB, Université de Brest, EA 3882, Technopôle Brest-Iroise, 29280 Plouzané, France and [2]Instituto Hispano-Luso de Investigaciones Agrarias (CIALE), University of Salamanca, Calle del Duero 12, 37185 Villamayor (Salamanca), Spain

[*]**Corresponding author:** Laboratoire Universitaire de Biodiversité et Ecologie Microbienne, IBSAM, ESIAB, Université de Brest, EA 3882, Technopôle Brest-Iroise, 29280 Plouzané, France. E-mail: josecobo1983@gmail.com

One sentence summary: New *Fusarium*-specific primers for MiSeq metabarcoding were designed. *F. graminearum* and *F. avenaceum* were the main species found in maize residue samples, and *F. oxysporum* in soil samples.

[†]J.F.C.D. and R.B. contributed equally to this work.

Editor: Ian Anderson

[‡]José F. Cobo-Díaz, http://orcid.org/0000-0002-0898-2358
[¶]Riccardo Baroncelli, https://orcid.org/0000-0002-5878-1159
[§]Gaétan Le Floch, http://orcid.org/0000-0001-7413-8852
[¥]Adeline Picot, https://orcid.org/0000-0001-9740-560X

ABSTRACT

The genus *Fusarium* contains more than 300 species, most of which are plant pathogens. Appropriate molecular tools for accurately and rapidly describing temporal and spatial shifts in *Fusarium* communities would be useful for the development of control strategies. Here, we present a new *Fusarium*-specific primer pair targeting the translation elongation factor 1-α (*EF1*α) gene with amplicons of ~430 bp, suitable for MiSeq metabarcoding sequencing. Mock *Fusarium* communities were used to evaluate its resolution and to optimize read filtering and downstream analyses. The use of the DADA2 pipeline coupled with operational taxonomic unit (OTU) picking at 98% similarity cut-off significantly increased the accuracy of read filtering. Building a phylogenetic tree using a manually curated database as a reference allowed taxonomic assignment at the species or species-complex level. This methodology was tested on soil and maize residue samples collected from crop fields. Up to 18 *Fusarium* OTUs, belonging to 17 species and 8 species complexes, were obtained, with *F. oxysporum* being the most abundant species in soil samples, while *F. graminearum* and *F. avenaceum* were the most abundant in maize residues. We demonstrated the high performance of this workflow which could be further used for profiling *Fusarium* species composition and dynamics during the cultivation cycle.

Keywords: *Fusarium* communities; metabarcoding primer pair; *EF1*α-*Fusarium* database; environmental *Fusarium* species; molecular diversity

Received: 4 March 2019; **Accepted:** 7 June 2019

INTRODUCTION

The genus *Fusarium* (teleomorph *Gibberella*) contains some of the most important plant pathogens worldwide, and can strongly affect crop production and cause important economic losses, notably in small-grain cereals and maize (Smiley *et al.* 2005; Dean *et al.* 2012). Moreover, harvested products can be contaminated with mycotoxins, which represents a major food and feed safety issue (Reddy *et al.* 2010; Sobrova *et al.* 2010). Economically important *Fusarium*-associated plant diseases include Fusarium and Gibberella Ear Rot (FER and GER) in maize (Xu *et al.* 2005; Xu and Nicholson 2009), Fusarium Head Blight (FHB) in small-grain cereals (Dorn *et al.* 2009; Scauflaire *et al.* 2011), Fusarium Crown Rot (FCR) in vegetables and cereals (Chakraborty *et al.* 2006) and several soil-borne diseases caused by *F. oxysporum* and *F. solani* species complexes (Fravel, Olivain and Alabouvette 2003; Zhang *et al.* 2006; Smith 2007). Up to 20 species can be associated with the same *Fusarium*-associated disease. For instance, *F. graminearum*, *F. culmorum*, *F. poae* and *F. avenaceum* are the main causal species for FHB in Europe (Basler 2016; Hellin *et al.* 2016; Karlsson *et al.* 2017), while *F. graminearum* is the main species responsible for GER, *F. verticillioides*, *F. proliferatum* and *F. subglutinans* are the main species responsible for FER (Desjardins 2003; Dorn *et al.* 2009; Summerell *et al.* 2010) and *F. graminearum*, *F. pseudograminearum* and *F. culmorum* are the main causal agents for FCR (Chakraborty *et al.* 2006; Cook, 2010). The control of *Fusarium*-associated diseases therefore represents a major scientific challenge given the complexity of the causal agents, not to mention the equally complex mechanisms that lead to mycotoxin contamination.

Conservation tillage practices, maize-wheat rotation and climate change play important roles in the future spread of FHB and mycotoxin contamination (Dill-Macky and Jones 2000; Cromey *et al.* 2002; Doohan, Brennan and Cooke 2003; Blandino *et al.* 2009). Weather-based models based on climate change augur an increase in FHB severity together with higher levels of mycotoxin contamination in grains in southern England and China by the 2050s (Madgwick *et al.* 2011; West *et al.* 2012; Zhang *et al.* 2014). The spread of FHB is also associated with changes in *Fusarium* species distribution, which are adapted to different climatic conditions. For example, *F. graminearum* is associated with warm and humid conditions (Backhouse 2014), while *F. culmorum* and *F. avenaceum* are associated with cold conditions (Backhouse 2014) and *F. poae* with warm and dry conditions (Xu *et al.* 2008; Yli-Mattila 2010). These differences in climate preferences and global warming may be the reasons for the recent increase of *F. graminearum* at the expense of *F. culmorum* in different cereal crops affected by FHB in several European countries (Waalwijk *et al.* 2003; Jennings *et al.* 2004; Nielsen *et al.* 2011; Scauflaire *et al.* 2011). Moreover, changes in *Fusarium* species composition may cause changes in disease severity levels and mycotoxin production (Nesic, Ivanovic and Nesic 2014). An accurate, fast and wide-ranging description of the *Fusarium* species composition associated with plant tissues is therefore essential for rational control strategies.

Several approaches have been developed to study *Fusarium* species composition in environmental samples, such as culture-dependent methods with a morphological and/or molecular identification of isolates (Arias *et al.* 2013) or culture-independent methods including real-time PCR with species-specific primers (Nicolaisen *et al.* 2009; Lindblad *et al.* 2013) and multiplex real-time PCR for multiple *Fusarium* species (Yli-Mattila *et al.* 2008; Scauflaire *et al.* 2012). These approaches are time-consuming and/or are limited to a low number of species. Alternatively, high-throughput sequencing approaches have recently been developed to describe *Fusarium* species composition including PacBio SMRT (Walder *et al.* 2017), 454 pyrosequencing (Edel-Hermann *et al.* 2015; LeBlanc, Kinkel and Kistler 2015; Karlsson *et al.* 2016) and Illumina Miseq (Boutigny *et al.* 2019) techonologies. Error rates generated by PacBio-CCS (circular consensus sequencing) are slightly reduced compared with Illumina MiSeq technology (Schloss *et al.* 2016; Tedersoo, Tooming-Klunderud and Anslan 2018), but the significantly higher cost per read still remains as an important disadvantage to its use in metarbacoding studies. Moreover, the decline in 454 technologies, due to the commercial decisions of the Roche company, which stopped supporting this platform since mid-2016, encouraged researchers to adapt their methodology to the available platforms, such as Illumina MiSeq or Ion Torrent platforms. In addition, the latter technologies are adapted to sequence shorter regions (up to 550 bp for MiSeq 300 Pair-End sequencing, up to 400 bp for Ion Torrent PGM). Primer pair design is also essential, and ideally they must amplify a gene region with enough resolution to correctly identify *Fusarium* species while being specific to this genus. Translation elongation factor (*EF1α*) is classically used for classification of *Fusarium* strains at species level. It is a single copy gene in *Fusarium*, facilitating quantitative comparisons between species (Geiser *et al.* 2004), and presents enough variability between *Fusarium* species to allow a correct taxonomic classification (O'Donnell *et al.* 2010; Wang *et al.* 2011; Al-Hatmi *et al.* 2016; Karlsson *et al.* 2016). Although a primer pair was recently designed for Illumina sequencing (Boutigny *et al.* 2019), the size of the resulting amplicons (640 bp) does not allow the overlap of the forward and reverse reads obtained by MiSeq Pair-End technology (sequencing size up to 2 × 300 bp), reducing considerably (until 250–300 bp) the sequence size used for taxonomic assignment and, thus, its resolution. The aims of the present study were (i) to design a new specific primer pair to characterize *Fusarium* communities by MiSeq sequencing, (ii) to evaluate whether species abundance is correctly yielded using mock communities, (iii) to optimize data filtering and analysis, including a taxonomic assignment approach and (iv) to test the new primer pair and data analysis workflow in eDNA samples, including soil and plant residues samples.

MATERIALS AND METHODS

Primer design

A set of 144 *EF1α-Fusarium* sequences covering a wide range of the genus' diversity (Al-Hatmi *et al.* 2016) were aligned using MUSCLE (Edgar 2004). Primers were manually designed in Geneious® 10.0.9 (Kearse *et al.* 2012), and primer characteristics were calculated using NetPrimer (http://www.premierbiosoft.com/netprimer/netprlaunch/netprlaunch.html). An additional alignment was done using the sequences included in the new *EF1α* database of *Nectriaceae* strains (hereafter 'FusTEFdb') to check the specificity of primer pairs for *Fusarium* species. FusTEFdb includes sequences belonging to 109 *Fusarium* spp., to species within closely related genera to *Fusarium* genus (*Geejayessia*, *Rectifusarium*, *Neocosmospora* and *Bisifusarium*) and to other associated taxa within the *Nectriaceae* family, such as *Fusicolla* and *Ilyonectria* (Table S1). All reference sequences were manually inspected and taxonomic information was retrieved from published works to ensure correct designation. Primer

pairs with the best characteristics and similar melting temperatures were used to check specificity and correct size of the amplicon by PCR tests on genomic DNA extracted from different isolates from the Culture Collection of the University of Western Brittany (UBOCC, Plouzané, France). This set of isolates includes strains belonging to *Fusarium*, to the *Nectriaceae* family and other associated fungal families (Table S2). Additionally, environmental samples (see section "Field sampling and eDNA extraction") were also used for verifying the correct size of amplification product.

Culture of fungal strains

Mock communities of DNA from 17 *Fusarium* strains representing 16 phylogenetically distinct species (Table 1) predominantly found in wheat and maize crops (Xu *et al.* 2008; Yli-Mattila 2010; Backhouse 2014) were mixed. These strains and those used for checking primers specificity (Table S2) were provided by the UBOCC collection, where they were stored in 30% glycerol at -80°C. Isolates were grown on potato dextrose agar (PDA) during 7 d at 25–30°C, according to strain best growth temperature.

Pure culture genomic DNA extraction, preparation of mock community and DNA amplification

Fungal DNA was extracted from 7-d-old cultures with the FastDNA® SPIN kit (MP Biomedicals, Santa Ana, CA, USA) following the manufacturer's instructions, with an initial homogenization step using the Retsch MM400 instrument (Retsch GmbH) at 30 Hz for 30 s, twice. The DNA was resuspended in 60 μL of sterile nuclease-free water. DNA concentration was measured with a Qubit 3.0 Fluorometer (Invitrogen) and a 17-strain mock community was prepared in triplicate by mixing DNA in equal concentrations (Table 1). DNA aliquots of mock community were diluted in at least 10 ng/μl before submitting samples to the sequencing platform.

PCR amplification of the entire *EF1α* gene of each strain used in mock communities was sequenced to make sure that primers perfectly match with each strain sequence. Primers *EF1f* (5'-ATGGGTAAGGAGGACAAGACTCA-3') and *EF1r* (5'-TGGAGATACCAGCCTCGAAC-3') (Brygoo and Gautier 2007) were used to amplify the *EF1α* gene, obtaining a fragment size of ~700 bp. Amplification reactions were performed in 25 μL volume using 0.025 U/μL of GoTaq Flexi DNA polymerase (Promega, Madison, WI, USA) and 1 × GoTaq Flexi buffer (Promega), 25–50 ng of template DNA, 0.2 μM of each primer, 2 mM of $MgCl_2$ and 0.2 mM of 10 mM dNTP mix (Promega). Amplification conditions were a first denaturation step of 5 min at 95°C; 30 cycles of 60 s at 95°C, 90 s at 62°C and 60 s at 72°C; and a final extension step of 5 min at 72°C.

Amplification products were analyzed by electrophoresis in 1 × TAE buffer with 1% (w/v) agarose gel and detected by UV fluorescence after GelRed (Biotium Inc., CA, USA) staining, according to the manufacturer's instructions. The BenchTop 100-bp DNA ladder (Promega) was used as a molecular size marker. PCR products were sent to Eurofins MWG (Ebersberg, Germany) for purification and for sequencing the forward and reverse strands, which was done on an ABI 3730XL DNA sequencer using the same primers as those used for PCR.

ABI trace files from genomic DNA sequencing were analyzed and consensus sequences were generated using Geneious 10.0.9 (45). A 700 bp amplicon of the *EF1α* gene was obtained and species-level assignment was performed using the targeted sequence amplified by our new primer pair (430 bp).

Field sampling and eDNA extraction

Soils and maize residues were collected from three maize fields in Brittany, France, in November 2016 after the maize harvest. Sampling was done within 3 d after the maize harvest. In each field, 15 points were randomly selected, and the first 5 cm of soil (with a hand auger of 6 cm diameter) as well as the aboveground part of one maize stalk with nodal region and leaves were randomly sampled in each point. Soil samples were stored at 4°C overnight before being sifted with a 2 mm sieve opening mesh followed by DNA extraction the following day. Stalks were stored at 4°C until DNA extraction, performed within a week after sampling.

Maize residue DNA was extracted from 200 mg of crushed maize stalks using FastDNA SPIN kit following the manufacturer's instructions. For soil samples, DNA was extracted from 1 g of soil using NucleoSpin® Kit for Soil (Machery-Nagel, Dueren, Germany) according to the manufacturer's instructions. Quality and concentration of purified DNA were determined using a UV spectrophotometer (NanoDrop1000, Thermo Scientific, USA). Environmental DNA samples were diluted to at least 10 ng/μl before sending them to the sequencing company.

PCR amplification and Illumina sequencing

A total of nine maize residue, nine soil and three mock community samples were selected for PCR amplification and Illumina MiSeq PE300 sequencing, which was performed at the McGill University and Génome Québec Innovation Centre, Montreal, Canada. CS1 universal adapter sequence (underlined before primer sequence) plus primer *Fa_150* (5'-ACACTGACGACATGGTTCTACA CCGGTCACTTGATCTACCAG-3') and CS2 universal adapter sequence (underlined before primer sequence) plus primer *Ra-2* (5'-TACGGTAGCAGAGACTTGGTCT ATGACGGTGACATAGTAGCG-3') were used to amplify *EF1α* genes from *Fusarium* species. Those universal adapter sequences are always used in MiSeq sequencing to add amplicon barcodes during the second round of PCR. Although Hot Start Taq DNA polymerase from QIAGEN is not a High Fidelity Taq, it was used to avoid contamination. PCR mixtures were done in 8 μl of total volume with the following reactive concentrations: each primer 0.4 μM, $MgCl_2$ 1.5 mM, DMSO 5%, dNTP 0.2 mM, Taq polymerase 0.02 U/μl and 1 μl of environmental DNA (~1 ng). Amplification conditions were 15 min at 96°C followed by 35 cycles of 30 s at 96°C, 30 s at 52°C and 60 s at 72°C, with a final extension step of 10 min at 72°C.

After the first round of PCR, the amplification products were run on a 2% agarose gel to determine the appropriate dilution used for the second round of PCR. This second round of amplification was performed with 1 μl of diluted PCR products and primers containing the Illumina adapters and indexes. PCR conditions were 10 min at 95°C, followed by 15 cycles of amplification (15 s at 95°C, 30 s at 60°C and 1 min at 72°C) and a final extension step of 3 min at 72°C. The products of the second PCR were purified with the Agencourt AMPure XP system and quantified with Quant-iT PicoGreen® dsDNA Assay Kit (Life Technologies). The purified amplicons were then pooled in equimolar concentrations, and the final concentration of the library was determined using a quantitative PCR (qPCR) next-generation sequencing (NGS) library quantification kit. Amplicon libraries were mixed with 10% PhiX control according to Illumina's protocols.

Table 1. Strains used for mock communities. Strains UBOCC-A-113 083 and UBOCC-A-101 139 have the same *EF1α* sequence.

Code	Species	Species complex	*EF1α* gb acc. number
UBOCC-A-110 149	*Bisifusarium dimerum*	*dimerum*	MK034339
UBOCC-A-109 122	*F. verticillioides*	*fujikuroi*	MK034344
UBOCC-A-109 151	*F. temperatum*	*fujikuroi*	MK034345
UBOCC-A-112 043	*F. proliferatum*	*fujikuroi*	MK034346
UBOCC-A-109 097	*Fusarium* sp. FIESC	*incarnatum-equiseti*	MK034348
UBOCC-A-113 035	*F. oxysporum*	*oxysporum*	MK034347
UBOCC-A-109 027	*F. sambucinum*	*sambucinum*	MK034354
UBOCC-A-101 142	*F. graminearum*	*sambucinum*	MK034350
UBOCC-A-109 021	*F. poae*	*sambucinum*	MK034352
UBOCC-A-109 103	*F. venenatum*	*sambucinum*	MK034353
UBOCC-A-109 096	*F. avenaceum*	*tricunctum*	MK034343
UBOCC-A-102 014	*Fusarium* sp. FCCSC	*citricola*	MK034341
UBOCC-A-110 136	*F. petroliphilum*	*Neocosmospora* clade (*solani*)	MK034340
UBOCC-A-101 139	*F. culmorum*	*sambucinum*	MK034351
UBOCC-A-113 083	*F. culmorum*	*sambucinum*	MK034351
UBOCC-A-109 148	*F. langsethiae*	*sambucinum*	MK034349
UBOCC-A-109 036	*F. acuminatum*	*tricunctum*	MK034342

Illumina *EF1α* sequences analysis

Paired reads from Illumina MiSeq 300 sequencing were processed with Quantitative Insights Into Microbial Ecology (QIIME) version 1.9.1 (Caporaso *et al.* 2010) using the *multiple_join_paired_ends.py* and *multiple_split_libraries_fastq.py* scripts with default parameters. The UCLUST algorithm (Edgar 2010) was used for operational taxonomic unit (OTU) *denovo* picking at 97% and 98% similarity cut-off, thus generating two different datasets, referred to as OTUs-097 and OTUs-098 approaches, respectively. A third dataset was obtained using the DADA2 library (Callahan *et al.* 2016) in R version 3.5.0 (R Core Team 2017), referred to as DADA2-OTUs-098. Forward and reverse read pairs were trimmed and filtered using DADA2 pipeline, with forward reads truncated at 270 nt and reverse reads at 210 nt. No ambiguous bases were allowed while each read was required to have less than two expected errors based on their quality scores. Amplicon sequence variants (ASV) were independently inferred from the forward and reverse of each sample using the run-specific error rates, and then read pairs were merged, requiring at least 15 bp overlap. ASV were assigned taxonomically using BLASTn versus the BLAST formatted database of FusTEFdb. Non-*Nectriaceae* sequences, characterized by less than 95% of identity with FusTEFdb and/or less than 380 bp alignment length, were removed. The filtered ASV were grouped in OTUs by the *pick_otus.py* QIIME script, with 98% similarity cut-off. A summary of this approach is presented in Fig. 1, and the corresponding scripts are available in supplementary document S1.

An additional species-level assignment method was performed using BLAST (Altschul *et al.* 1990) against the NCBI non-redundant nucleotide database (nt) version updated on 11 January 2019 (ftp://ftp.ncbi.nlm.nih.gov/blast/db/). For the latter species-level assignment method, gDNA sequences from the mock community strains and both environmental and mock OTU representative sequences, along with reference sequences in FusTEFdb (Table S1), were aligned using MAFFT v.7.304 (Katoh, Rozewicki and Yamada 2017). Multiple sequence alignments were exported to MEGA7 (Kumar, Stecher and Tamura 2016) and the best-fit substitution model was calculated for each separate sequence dataset. Using MrBayes 3.2.6 (Ronquist *et al.* 2012), the Markov chain Monte Carlo (MCMC) algorithm was performed to generate phylogenetic trees with Bayesian posterior probabilities for combined sequence datasets using the nucleotide substitution models determined by MEGA7 (Kimura 2-parameter with gamma distributed rate variation among sites [K2-G]). Four MCMC chains were run simultaneously for random trees for 2000 000 generations (the *P*-value reached 0.01). Samples were taken every 2000 generations. The first 25% of trees were discarded as the burn-in phase of each analysis and posterior probabilities were determined from the remaining trees.

To further test the resolution of the proposed pipeline, the 430 bp amplicon-sequence region was extracted from the *Fusarium* and closely related genera sequences in FusTEFdb. These sequences were clustered at 98% similarity cut-off by the *otu_picking.py* QIIME script. Additionally, a linear regression analysis of mock species relative abundance and amplicon length was done using the *lm* function in R (R Core Team 2017) to evaluate length bias. Values for *F. culmorum* were divided by 2 because they were represented twice in the mock communities. All values were normalized by square root transformation before linear regression analysis.

Nucleotide sequences accession numbers

Demultiplexed raw sequence data were deposited in the Sequence Read Archive (http://www.ncbi.nlm.nih.gov/sra) under SRA accession numbers SRX4666001, SRX4665996 and SRX4665997 for mock samples (BioProject accession number PRJNA497205), SRA accession numbers SRX4628227 and SRX4628229–36 for maize samples and SRX4628257–58, SRX4628420–21, SRX4628426–29 and SRX4628544 for soil samples (BioProject PRJNA497210). Representative sequences of OTUs obtained in environmental samples are available in GenBank (MH900113–33) and also *EF1α* sequences of strains used for mock community (MK034339–54).

RESULTS

Primer pair design and specificity

After alignment of the 144 *Fusarium* sequences for primer design, two conserved regions were obtained between nucleotides 85 and 155 and between nucleotides 473 and

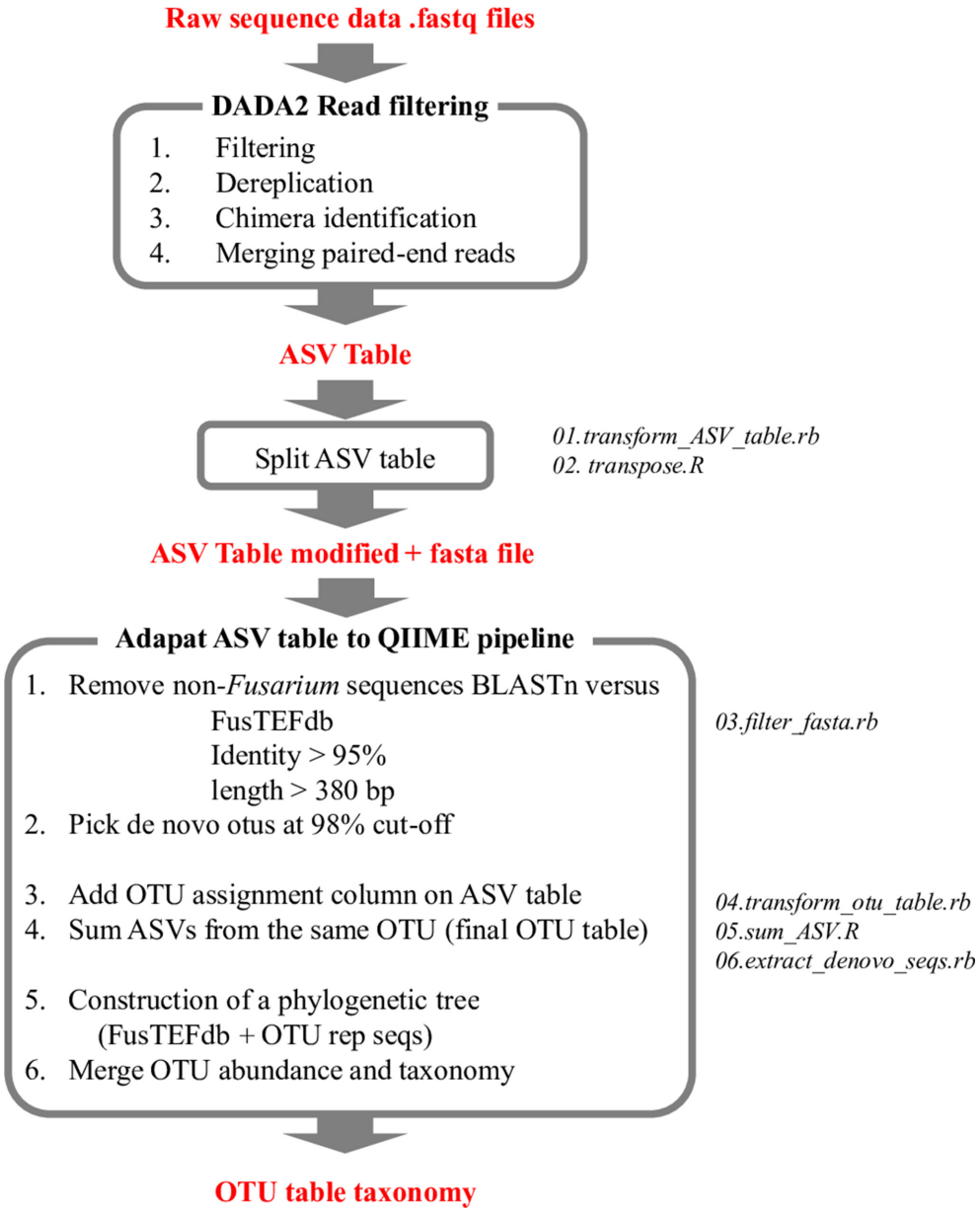

Figure 1. Pipeline of raw read data filtering and analysis by DADA2-OTUs-098 approach for *EF1α-Fusarium* sequences. Scripts used are indicated in italics and presented in the supplementary data text.

623 (data not shown). Six forward and eight reverse primers were designed (data not shown), and only two primer pairs were used in PCR tests (*Fa_150* as forward, *TEF_FUS_R1* and *Ra-2* as reverse) due to their similarity in melting temperatures, the absence of hairpin Tm and the low value of self-dimer Tm. Primer pair *Fa_150* and *Ra-2* showed the highest specificity for *Fusarium* species tested against DNA extracted from pure culture of *Emericellopsis terricola*, *Fusidium terricola*, *Curvularia maculans*, *Cephalosporium acremonium*, *Beauvaria bassiana*, *Acremonium portonii*, *Gliocladium virens*, *Cylindrocarpon heteronema*, *C. radicicola*, *C. suballantoideum* and *Microdochium nivale* (Table S2), and were further used for the metabarcoding approach. The *Ra-2* primer had the same sequence as primer *Ra* (Edel-Hermann *et al.* 2015), but without the first two nucleotides, while *Fa_150* was a newly designed primer located ~150 bp downstream of *Fa* primer (Edel-Hermann *et al.* 2015). It allowed the reduction of amplicon size from 570 to 430 bp and therefore the overlap of the two Miseq reads in paired sequencing.

Only two out of the 144 *Fusarium* sequences used to check the alignment with our primer pair showed a C/T mismatch at position 2 in the forward primer while no mismatch was found in the reverse one. On the additional alignment using all non-*Fusarium* sequences in FusTEFdb (Table S1) along with two *Trichoderma* sequences and the strains used in mock community, we found a strong specificity on *Fusarium* and *Bisifusarium* species, with only one mismatch for *F. plagianthi* and *F. avenaceum* in the forward primer, and one mismatch for *F. plagianthi*, *F. temperatum* and *Bisifusarium domesticum* in the reverse primer. Those mismatches found for *F. avenaceum* and *F. temperatum* did not affect the results presented below.

Evaluation of primer pair resolution in mock communities

A total of 65 644 raw reads were generated by MiSeq sequencing for the three mock community samples, and 51 578 sequences (78.6% of the raw sequences) and 49 222 sequences (75.0% of the raw sequences) were kept for further analysis after read filtering by QIIME and DADA2 pipelines, respectively. Up to 48 353 sequences were clustered in 173 OTUs with more than 10 sequences using the OTUs-097 approach. The 16 most abundant OTUs, representing 88.1% of the total sequences, matched the 16 different species (Fig. 2) present in our mock community (Table 1) and were selected for further analysis. The expected relative abundance values for each species were 5.9%, except for *F. culmorum* (11.8%), given that the DNA of UBOCC-A-101 139 and UBOCC-A-113 083 had a 100% similarity for *EF1α* sequences, both of which belong to *F. culmorum*. The observed average values for each species ranged from 1.7 to 14.5% of relative abundance (Fig. 2). With this OTU-picking method, an important over-representation of *F. culmorum* and *F. poae* was obtained while *F. langsethiae* and *F. avenaceum* were under-represented. Using the OTUs-098 approach, up to 38 762 sequences were clustered in 299 OTUs after removing OTUs with 10 or less reads. The 16 most abundant OTUs, representing 73.0% of the total sequences, matched the 16 different species (Fig. 2) present in our mock community (Table 1). The species distribution was similar to the 97% similarity cut-off, except for the lack of over-representation of *F. culmorum*, and that *F. acuminatum* presented a significant decrease in relative abundance while *F. graminearum* and *F. avenaceum* presented a significant increase (Fig. 2). With the DADA2-OTUs-098 approach, only 16 OTUs were obtained after clustering the 18 ASVs obtained, which matched exactly the number of species used in mock communities. Again, there was an overestimation of *F. poae* and *F. culmorum* and an underestimation of *F. langsethiae* (Fig. 2). Finally, linear regression analysis showed a significant negative correlation (*P*-value = 0.017, R^2 = 0.099) between the amplicon size and species relative abundance obtained in the mock communities (Fig. 3).

Environmental data

A total of 1037 854 raw reads were generated by MiSeq sequencing for the 24 field samples, and 670 498 sequences (64.6% of the raw reads) and 649 204 sequences (62.6% of the raw reads) were kept for further analysis after read filtering by QIIME and DADA2 pipelines, respectively. Up to 32 *Fusarium* OTUs were obtained in environmental samples using the OTUs-097 approach, while 134 OTUs were obtained for OTUs-098. Using only DADA2 pipeline, a total of 117 ASV were obtained and grouped in 18 *Fusarium* OTUs by the DADA2-OTUs-098 approach. ASV associated with *Fusarium* species represented 99.96 and 85.4% of total sequences from maize and soil samples, respectively, although in one soil sample this value decreased to 51.2%. The most abundant non-*Fusarium* ASVs were assigned to *Ilyonectria*, which belonged to *Nectriaceae* (Table S3).

According to the DADA2-OTUs-098 dataset, up to 21 OTUs associated with the family *Nectriaceae* were found in environmental samples, three OTUs assigned to three non-*Fusarium* strains including *Ilyonectria leucospermi*, *Neonectria lugdunensis* and *Fusicolla aquaeductum* (Fig. 4; Table 2). The most abundant OTUs in soil samples were assigned to *F. oxysporum* (60.4%), *F. avenaceum* (10.0%) and *F. flocciferum* (9.0%) followed by 10 additional *Fusarium* OTUs assigned to 10 distinct species, each of which represented less than 0.1 to 4.5% of total sequences (Table 2). Maize residues were dominated by *F. graminearum* (41.9%), *F. avenaceum* (39.0%) and *F. poae* (11.7%) while the remaining sequences included up to eight *Fusarium* OTUs, each of which accounted for 0.5 to 2.4% of total sequences (Table 2).

Taxonomic assignment accuracy

Using 98% similarity cut-off, we found that a few important species within FusTEFdb could not be differentiated from other minor species, including *F. proliferatum* (clustered with *F. globosum*), *F. culmorum* (clustered with *F. cerealis*), *F. tricinctum* (clustered with *F. acuminatum*), *F. oxysporum* (clustered with *F. inflexum*), *F. graminearum* (clustered with *F. austroamericanum* and *F. meridionale*) and *F. temperatum* (clustered with some species within the *fujikuroi* species complex) (Table 3).

DISCUSSION

A new *Fusarium*-specific primer pair designed to characterize *Fusarium* communities in field samples using Illumina high-throughput sequencing was developed and tested in this study. The use of mock communities, recommended as a control community to adapt read filtering and data analysis (Karlsson *et al.* 2016; Walder *et al.* 2017), was employed to test the accuracy of the primer pair. Mock communities were also employed to test and improve the use of the DADA2 R package (Callahan *et al.* 2016), which reduces sequencing errors accumulated in Illumina-sequencing and a subsequent overestimation of OTUs' richness. The use of an OTU-picking approach for the ASVs generated by DADA2 pipeline reduced the species overestimation observed in mock communities, and consequently in environmental samples.

Additionally, the use of a phylogenetic tree with a manually curated reference sequences database allows an accurate species-level designation and the detection of new species or clades; as an example, strain UBOCC-A-102 014 did not cluster with any sequences and this could reflect a putative new species (Fig. 4). The *EF1α* sequence from this isolate, together with an environmental OTU sequence, were clustered into a clade close to *F. tricinctum* species complex, which was recently described as *F. citricola* species complex (Sandoval-Denis *et al.* 2018). These two sequences formed a new branch within this species complex, different from species *F. citricola* and *F. salinense*, and were previously wrongly classified as *F. lateritium* by BLASTn against nt database (data not shown). These findings also highlight the importance of keeping curated databases up to date. For instance, the recent reassignment of *Fusarium dimerum* species complex as *Bisifusarium* was included in our database as well as the assignment of *F. solani* species complex to *Neocosmospora* genus, and *F. ventricosum* clade to *Rectifusarium* genus. These species complexes are now forming three independent well-supported clades closely related to the genus *Fusarium* (Lombard *et al.* 2015). Likewise, *Fusarium merismoides* has been reassigned to *Fusicolla merismoides* and *Fusicolla aquaeductum* (Gräfenhan *et al.* 2011), forming a consistent clade.

Mock communities used in Illumina sequencing confirmed that the targets of our newly designed primer pair were *Fusarium* species and also closely related genera, including *Bisifusarium* and *Neocosmospora*. The over-representation of *F. culmorum* could be due to the high sequence similarity between this strain and strains assigned to *F. graminearum*, with 14 nucleotides of difference in 449 bp amplicon size (3.1% difference between sequences), and these two species sequences could no longer be differentiated with 97% similarity cut-off. Something similar

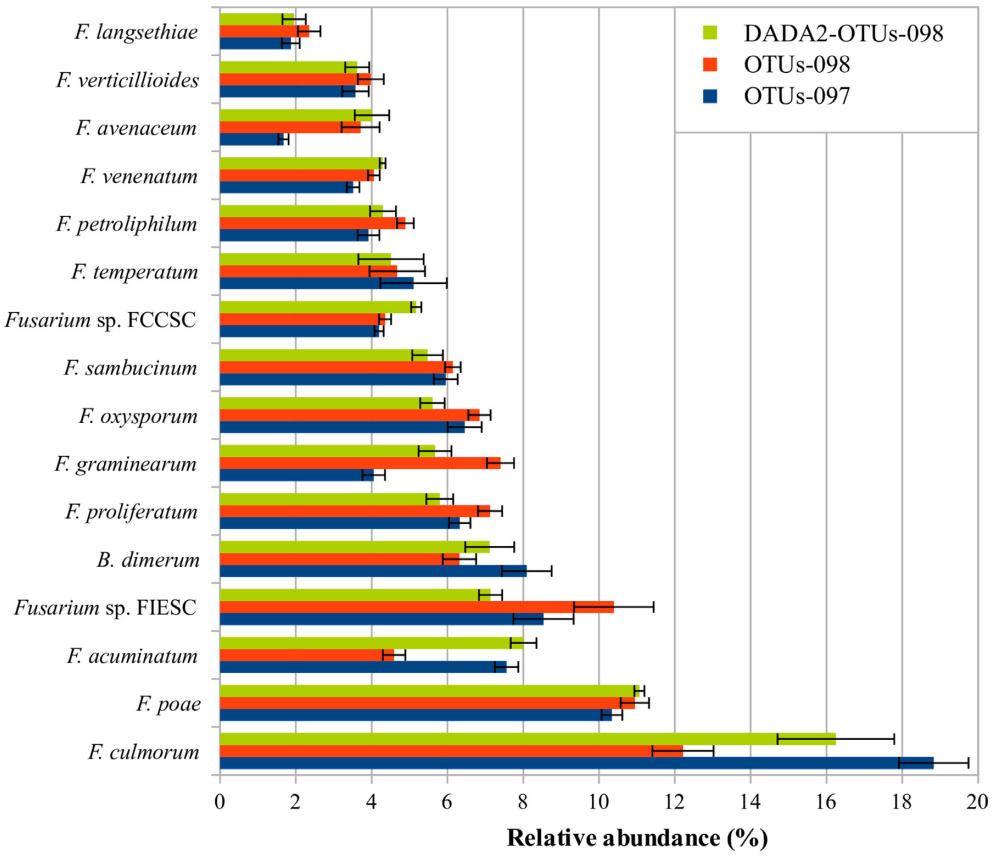

Figure 2. Relative abundance of *Fusarium* species in mock communities. See Fig. 4 for details about the taxonomic assignment of strains. Bars show standard deviation.

was observed for *F. avenaceum* and *F. acuminatum* strains, whose amplicon sequences have only 11 nucleotides of differences in 438 bp amplicon size (2.5%). While the increase of such a cut-off value may eliminate this problem, it could, at the same time, generate an overestimation of *Fusarium*-OTUs' richness in environmental samples. The combined approach based on DADA2 and OTU picking at 98% similarity cut-off fixed these problems since the 16 expected OTUs in mock communities were obtained in our metabarcoding data. Similar issues were found with the *F. culmorum* assignment using *EF1α* metabarcoding (Boutigny *et al.* 2019). In this study, the assignment of species within the *Fusarium graminearum* species complex was also compromised because of the high similarity value (>98%) within the 200 bp *EF1α* sequence for 7 of the 10 species used in the mock community. We also found that ASVs clustered into the environmental OTU assigned to *F. cerealis* were classified as either *F. cerealis* or *F. culmorum* (data not shown), and the corresponding reference sequences from our database clustered together at 98% similarity cut-off. Species assignment is not accurate enough for certain taxa (Table 3) because sequence similarity between those taxa was higher than the selected OTU picking cut-off (98%). ASV sequences within these problematic species-OTUs can be used for phylogenetic analyses to resolve this loss of resolution. Furthermore, proper differentiation of species and subspecies within specific species complexes could be resolved by fungal isolation followed by taxonomic identification using a multilocus approach, as was shown recently for *F. oxysporum* species complex (Lombard *et al.* 2019).

A similar degree of difference between the most and least abundant species were also found in other equally distributed mock communities from past studies aiming at profiling *Fusarium* communities (Edel-Hermann *et al.* 2015; Karlsson *et al.* 2016; Walder *et al.* 2017; Boutigny *et al.* 2019). The use of genomic DNA to equilibrate mock communities carries the bias of different genome sizes between species, which implies differences in the ratio of amplicon to total DNA concentration. The genome of *F. graminearum* has been estimated to be smaller than *F. oxysporum* (Ma *et al.* 2010), which was the main reason put forward to explain the under-representation of *F. oxysporum* (Karlsson *et al.* 2016) and the over-representation of *F. graminearum* (Walder *et al.* 2017) in previous studies with mock communities. Moreover, we found significant negative correlation between amplicon size and relative abundance obtained in mock communities, as has been found also for the *EF1α* primer pair developed by Karlsson *et al.* (2016). This correlation could explain the overestimation of *F. poae*, which has the shortest amplicon size. However, this result has to be treated carefully since genomic DNA was used in mock communities, carrying the above-mentioned bias due to differences in genome size.

The most abundant OTU in soil samples was assigned to *F. oxysporum*, which was found as the second most abundant species in grassland soil samples using RPB2 amplicon sequencing (LeBlanc, Kinkel and Kistler, 2015, 2017) and also the most abundant species through culture-dependent approaches from wheat crop soils (Silvestro *et al.* 2013) and other agricultural crops (Balmas *et al.* 2010; Saremi and Saremi 2013). The OTUs associated with *F. graminearum*, *F. avenaceum* and *F. poae* were dominant in our maize residue samples, as has previously been found in other maize residue samples (Köhl *et al.* 2015), while other studies found that *F. temperatum*, *F. proliferatum* and *F. verticillioides* were the main *Fusarium* species on maize stalks (Dorn

Table 2. Relative abundance and taxonomic assignment of OTUs obtained in environmental samples. See Fig. 4 for details about taxonomic assignment of strains, performed using phylogenetic tree and curated FusTEFdb database.

OTU ID	Species assignment	*Fusarium* species complex	Number of reads	% seqs maize	% seqs soil
denovo5	*F. oxysporum*	*oxysporum*	169 968	0.6	60.4
denovo0	*F. avenaceum* (FTSC 4)	*tricinctum*	152 752	39.0	10.0
denovo9	*F. graminearum*	*sambucinum* (subclade *graminearum*)	146 669	41.9	4.5
denovo15	*F. poae*	*sambucinum* (subclade *sambucinum*)	47 038	11.7	3.4
denovo8	*F. flocciferum*	*tricinctum*	24 999	0	9.0
denovo13	*F. cerealis*	*sambucinum* (subclade *graminearum*)	21 012	2.4	4.8
denovo7	*F. equiseti*	*incarnatum-equiseti*	9386	0	3.4
denovo11	*F. temperatum*	*fujikuroi*	6499	1.6	0.5
denovo2	*Neocosmospora solani* (FSSC 5)	*Neocosmospora* clade 3	4102	0	1.5
denovo4	*F. commune*	*nisikadoi*	3695	0	1.3
denovo3	*F. sporotrichioides*	*sambucinum* (subclade *sporotrichioides*)	3168	1.0	0
denovo6	*Fusarium sp.* (FCCSC)	*citricola*	2459	0.8	0
denovo19	*F. tricinctum*	*tricinctum*	1767	0.6	0
denovo17	*Fusicolla aquaeductuum*	-	1380	0	0.5
denovo1	*F. praegraminearum*	*sambucinum* (subclade *graminearum*)	1024	0	0.4
denovo14	*F. proliferatum*	*fujikuroi*	1006	0.3	0
denovo18	*F. venenatum*	*sambucinum* (subclade *sambucinum*)	682	0.2	0
denovo20	*Ilyonectria leucospermi*	-	534	0	0.2
denovo10	*Neonectria lugdunensis*	-	460	0	0.2
denovo16	*F. oxysporum*	*oxysporum*	20	0	<0.1
denovo12	*Fusarium* f. sp. *cucurbitae* (FSSC 10)	*Neocosmospora* clade 3	5	0	<0.1

Table 3. Reference sequences for *Fusarium* plant pathogens grouped in OTUs at 98% similarity cut-off of the *EF1α* gene region amplified by the proposed primer pair. Those in bold type are important *Fusarium* pathogenic species responsible for FHB, FER, GER or FCRR.

Fusarium species complex	Group 1	Group 2
fujikuroi	AF160297 *F. ananatum* KR071762 *F. ananatum* AF160295 *F. circinatum* AF160291 *F. succisae* AF160294 *F. bulbicola* AF160293 *F. begoniae* KJ541067 *F. sororula* **KF956082 *F. temperatum*** AF160307 *Fusarium* sp.	**KR673923 *F. proliferatum*** **AF160285 *F. globosum***
subclade graminearum	JF740836 *F. austroamericanum* JF740835 *F. meridionale* **JF740867 *F. graminearum***	**AJ543536 *F. cerealis*** **JF740860 *F. culmorum***
oxysporum	**LT746202 *F. oxysporum*** **KU711710 *F. oxysporum*** **HM347117 *F. oxysporum*** AF008479 *F. inflexum*	
tricinctum	**JF740857 *F. acuminatum*** **HM068307 *F. tricinctum***	

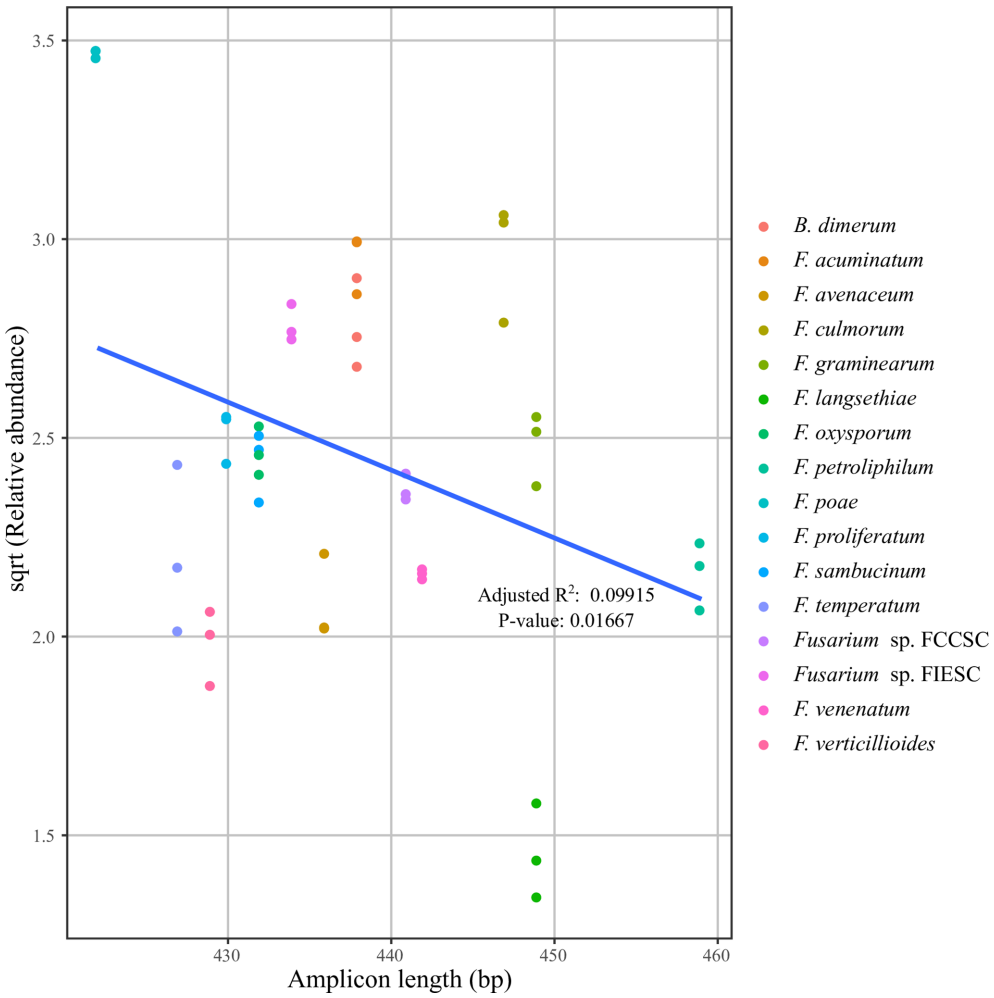

Figure 3. Linear regression of amplicon length and relative abundance of *Fusarium* species used in mock communities.

et al. 2011; Czembor, Stępień and Waśkiewicz 2015). The composition of *Fusarium* communities is strongly influenced by environmental conditions and may thus account for such differences. The species *F. graminearum* and *F. avenaceum* seem to be more adapted to the tempered and humid climatic conditions than species from the *Fusarium fujikuroi* species complex (Xu *et al.* 2008), as was found in our fields located in Brittany, France. Moreover, *F. graminearum*, *F. avenaceum*, *F. culmorum* and *F. poae* were often found to be the most abundant in wheat and barley samples in France (Ioos, Belhadj and Menez 2004; Ioos *et al.* 2005; Boutigny *et al.* 2019). The absence of detection of *F. culmorum* in our samples may also be due to limitations of this approach since, as mentioned before, *F. culmorum* could be assigned to *F. cerealis*. Yet *F. cerealis* was the fourth in abundance in both maize residues and soil samples, and only accounted for 2.4 and 4.8% of the total sequences, respectively. Either way *F. culmorum* remains a minor species in our samples.

Finally, the low percentage of non-*Fusarium* reads obtained on environmental samples, which was almost 0 for maize residues and less than 15% for soil samples, confirmed the *Fusarium* specificity of the designed primer pair. As a basis of comparison, 5 to 95% of non-*Fusarium* species sequences were obtained with other *EF1α Fusarium* specific primer pairs (Boutigny *et al.* 2019) which amplified a wide range or phylogenetically distant species on environmental samples, including the genera *Alternaria*, *Cladosporiun*, *Colletotrichum* and *Trichoderma*.

CONCLUSIONS

The main advantage of this new primer pair, compared with previously developed ones (Edel-Hermann *et al.* 2015; Karlsson *et al.* 2016; Boutigny *et al.* 2019), lies in the fact that the amplicon size (430 bp) is adapted for Illumina technology. Apart from the current disuse of the 454 technology, the use of MiSeq machines avoids bias due to different amplicon sizes (454 Life Sciences Corp. 2011; Ihrmark *et al.* 2012) and allows us to combine this amplicon with others in the same run (Herbold *et al.* 2015). Although the proposed primer pair could be used in single-molecule real-time technologies, such as Nanopore and PacBio, to avoid bias due to PCR steps, the higher cost per read makes it inadvisable. Additionally, the specificity of the presented primer pair to amplify *Fusarium* spp. sequences allows a lower amount of reads per sample to cover the entire species composition, and, therefore, hundreds of samples sequencing in only one run of MiSeq technology. Moreover, we also suggest using a phylogenetic tree for a species-level accurate classification of the OTUs, and DADA2 plus OTU picking at 98% of similarity to avoid

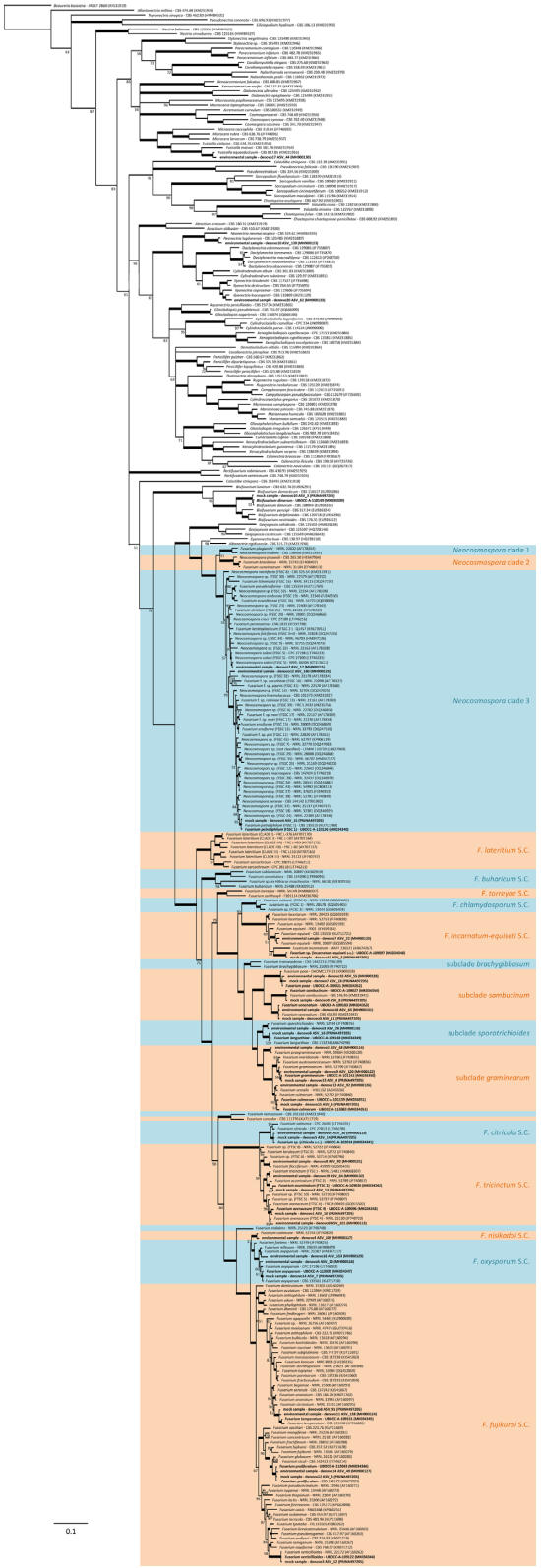

Figure 4. Phylogenetic tree of *EF1α* sequences from gDNA, *denovo* OTUs sequences (DADA2-OTUs-098) of mock and environmental samples and reference sequences of FusTEFdb, whose accession numbers and taxonomic information are indicated. Higher quality figure is available at https://doi.org/10.6084/m9.figshare.8293487.v1

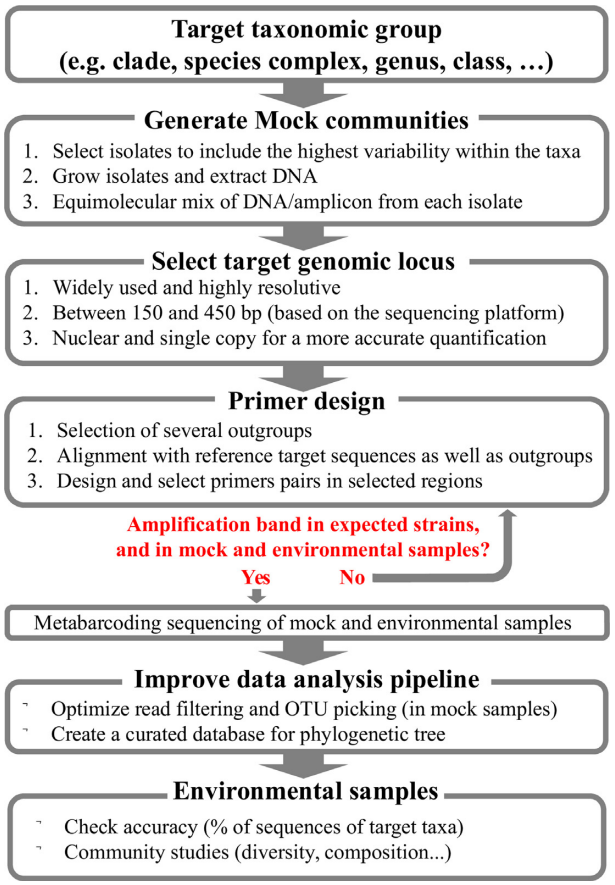

Figure 5. Workflow proposed for specific taxon primer design, improvement of data analysis and validation.

sequencing errors that notably increase both ASV and OTU richness in environmental studies. The limitation of the proposed approach mainly relies on its low discrimination ability for a limited number of very closely related species including *F. culmorum* and *F. cerealis*, or *F. acuminatum* and *F. tricinctum* (Table 3), but it can be solved with a deeper view of ASV sequences belonging to OTUs assigned to these species. Nevertheless, our approach, summarized in Fig. 5, represents an easy, fast and relatively cheap way to study communities of a target taxon and can be adapted to metabarcoding studies focused on other taxonomic groups or functional genes. Other more complex systems, where no species-level marker genes are identified, may need to combine culture-dependent and independent techniques to provide an accurate community profile, as has been performed for *Colletotrichum* species (Da Lio *et al.* 2018). To characterize putative new species or confirm taxonomic assignment of OTUs with relevant importance, an isolation step would ultimately be needed, but limited to the samples where the species or OTU of interest have been detected.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSEC* online.

ACKNOWLEDGEMENTS

We thank the UBO Culture Collection for kindly providing the fungal strains used in this study. We thank the farmers for kindly

giving us access to their fields to take environmental samples. We also thank Martial Briand from the Team Emersys of INRA for providing us with scripts for the DADA2 read filtering pipeline. The authors wish to thank Michael R. Thon for the revision and suggestions.

This work was supported by the Brittany Region [Grant #9097 MycoRes] and the Institut Brestois Santé-Agro-Matière (IBSAM). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Conflict of interest. None declared.

REFERENCES

454 Life Sciences Corp. 2011. *454 sequencing system guidelines for amplicon experimental design GS FLX+/XL+*. Bradford, CT: 454 Life Science Corp.

Al-Hatmi AMS, Van Den Ende AHGG, Stielow JB *et al.* Evaluation of two novel barcodes for species recognition of opportunistic pathogens in Fusarium. *Fungal Biol* 2016;**120**:231–45.

Altschul SF, Gish W, Miller W *et al.* Basic Local Alignment Search Tool. *J Mol Biol* 1990;**215**:403–10.

Arias MMD, Munkvold GP, Ellis ML *et al.* Distribution and frequency of Fusarium species associated with soybean roots in Iowa. *Plant Dis* 2013;**97**:1557–62.

Backhouse D. Global distribution of Fusarium graminearum, F. asiaticum and F. boothii from wheat in relation to climate. *Eur J Plant Pathol* 2014;**139**:161–73.

Balmas V, Migheli Q, Scherm B *et al.* Multilocus phylogenetics show high levels of endemic Fusaria inhabiting Sardinian soils (Tyrrhenian Islands). *Mycologia* 2010;**102**:803–12.

Basler R. Diversity of Fusarium species isolated from UK forage maize and the population structure of F. graminearum from maize and wheat. *PeerJ* 2016;**4**:e2143.

Blandino M, Reyneri A, Vanara F *et al.* Influence of agricultural practices on Fusarium infection, fumonisin and deoxynivalenol contamination of maize kernels. *World Mycotoxin J* 2009;**2**:409–18.

Boutigny A-L, Gautier A, Basler R *et al.* Metabarcoding targeting the EF1 alpha region to assess Fusarium diversity on cereals. *PLoS One* 2019;**14**:e0207988.

Brygoo Y, Gautier A. Polymorphisme moléculaire des souches de Fusarium isolées sur épis de blé et de maïs en France - Récoltes 2003 et 2004. In: *Progrès et perspectives de la recherche sur les mycotoxines de Fusarium dans les céréales*, 2007, 1–21. Presented at Mycotoxines fusariennes des céréales, Arcachon, FRA (2007-09-11 - 2007-09-13). Versailles, France: Editions Quae. https://prodinra.inra.fr/record/196419

Callahan BJ, McMurdie PJ, Rosen MJ *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016;**13**:581–3.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.

Chakraborty S, Liu CJ, Mitter V *et al.* Pathogen population structure and epidemiology are keys to wheat crown rot and Fusarium head blight management. *Australas Plant Pathol* 2006;**35**:643.

Cook RJ. Fusarium root, crown, and foot rots and associated seedling diseases. In: Bockus WW, Bowden R, Hunger R *et al.*. (eds.) *Compendium of wheat diseases and pests*. 3rd edition, PA, USA: The Pennsylvania State University Press, 2010; 37–9.

Cromey MG, Shorter SC, Lauren DR *et al.* Cultivar and crop management influences on Fusarium head blight and mycotoxins in spring wheat (Triticum aestivum) in New Zealand. *New Zeal J Crop Hort* 2002;**30**:235–47.

Czembor E, Stępień L, Waśkiewicz A. Effect of environmental factors on Fusarium species and associated mycotoxins in maize grain grown in Poland. *PLoS One* 2015;**10**:e0133644.

Da Lio D, Cobo-Díaz JF, Masson C *et al.* Combined metabarcoding and multi-locus approach for genetic characterization of Colletotrichum species associated with common walnut (Juglans regia) anthracnose in France. *Sci Rep* 2018;**8**:10765.

Dean R, Van Kan JAL, Pretorius ZA *et al.* The top 10 fungal pathogens in molecular plant pathology. *Mol Plant Pathol* 2012;**13**:414–30.

Desjardins AE. Gibberella from A(venaceae) to Z(eae). *Annu Rev Phytopathol* 2003;**41**:177–98.

Dill-Macky R, Jones RK. The effect of previous crop residues and tillage on Fusarium head blight of wheat. *Plant Dis* 2000;**84**:71–6.

Doohan FM, Brennan J, Cooke BM. Influence of climatic factors on Fusarium species pathogenic to cereals. *Eur J Plant Pathol* 2003;**109**:755–68.

Dorn B, Forrer H-R, Schürch S *et al.* Fusarium species complex on maize in Switzerland: occurrence, prevalence, impact and mycotoxins in commercial hybrids under natural infection. *Eur J Plant Pathol* 2009;**125**:51–61.

Dorn B, Forrer HR, Jenny E *et al.* Fusarium species complex and mycotoxins in grain maize from maize hybrid trials and from grower's fields. *J Appl Microbiol* 2011;**111**:693–706.

Edel-Hermann V, Gautheron N, Mounier A *et al.* Fusarium diversity in soil using a specific molecular approach and a cultural approach. *J Microbiol Methods* 2015;**111**:64–71.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;**26**:2460–1.

Fravel D, Olivain C, Alabouvette C. Fusarium oxysporum and its biocontrol. *New Phytol* 2003;**157**:493–502.

Geiser DM, del Mar Jiménez-Gasco M, Kang S *et al.* FUSARIUM-ID v. 1.0: A DNA sequence database for identifying Fusarium. *Eur J Plant Pathol* 2004;**110**:473–9.

Gräfenhan T, Schroers H-J, Nirenberg HI *et al.* An overview of the taxonomy, phylogeny, and typification of nectriaceous fungi in Cosmospora, Acremonium, Fusarium, Stilbella, and Volutella. *Stud Mycol* 2011;**68**:79–113.

Hellin P, Dedeurwaerder G, Duvivier M *et al.* Relationship between Fusarium spp. diversity and mycotoxin contents of mature grains in southern Belgium. *Food Addit Contam Part A* 2016;**33**:1228–40.

Herbold CW, Pelikan C, Kuzyk O *et al.* A flexible and economical barcoding approach for highly multiplexed amplicon sequencing of diverse target genes. *Front Microbiol* 2015;**6**:731.

Ihrmark K, Bödeker ITM, Cruz-Martinez K *et al.* New primers to amplify the fungal ITS2 region - evaluation by 454-sequencing of artificial and natural communities. *FEMS Microbiol Ecol* 2012;**82**:666–77.

Ioos R, Belhadj A, Menez M. Occurrence and distribution of Microdochium nivale and Fusarium species isolated from barley, durum and soft wheat grains in France from 2000 to 2002. *Mycopathologia* 2004;**158**:351–62.

Ioos R, Belhadj A, Menez M *et al.* The effects of fungicides on Fusarium spp. and Microdochium nivale and their associated trichothecene mycotoxins in French naturally-infected cereal grains. *Crop Prot* 2005;**24**:894–902.

Jennings P, Coates ME, Walsh K *et al.* Determination of deoxynivalenol- and nivalenol-producing chemotypes of Fusarium graminearum isolated from wheat crops in England and Wales. *Plant Pathol* 2004;**53**:643–52.

Karlsson I, Edel-Hermann V, Gautheron N *et al.* Genus-specific primers for study of Fusarium communities in field samples. *Appl Environ Microbiol* 2016;**82**:491–501.

Karlsson I, Friberg H, Kolseth A-K *et al.* Agricultural factors affecting Fusarium communities in wheat kernels. *Int J Food Microbiol* 2017;**252**:53–60.

Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform* 2017: bbx108.

Kearse M, Moir R, Wilson A *et al.* Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012;**28**:1647–9.

Köhl J, Lombaers C, Moretti A *et al.* Analysis of microbial taxonomical groups present in maize stalks suppressive to colonization by toxigenic Fusarium spp.: a strategy for the identification of potential antagonists. *Biol Control* 2015;**83**:20–8.

Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol Biol Evol* 2016;**33**:1870–4.

LeBlanc N, Kinkel LL, Kistler HC. Soil fungal communities respond to grassland plant community richness and soil edaphics. *Microb Ecol* 2015;**70**:188–95.

LeBlanc N, Kinkel L, Kistler HC. Plant diversity and plant identity influence Fusarium communities in soil. *Mycologia* 2017;**109**:128–39.

Lindblad M, Gidlund A, Sulyok M *et al.* Deoxynivalenol and other selected Fusarium toxins in Swedish wheat - occurrence and correlation to specific Fusarium species. *Int J Food Microbiol* 2013;**167**:284–91.

Lombard L, van der Merwe NA, Groenewald JZ *et al.* Generic concepts in Nectriaceae. *Stud Mycol* 2015;**80**:189–245.

Lombard L, Sandoval-Denis M, Lamprecht SC *et al.* Epitypification of Fusarium oxysporum - clearing the taxonomic chaos. *Persoonia* 2019;**43**:1–47.

Ma L-J, van der Does HC, Borkovich KA *et al.* Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature* 2010;**464**:367–73.

Madgwick JW, West JS, White RP *et al.* Impacts of climate change on wheat anthesis and Fusarium ear blight in the UK. *Eur J Plant Pathol* 2011;**130**:117–31.

Nesic K, Ivanovic S, Nesic V. Fusarial toxins: secondary metabolites of Fusarium fungi. In: Whitacre DM (ed.). *Reviews of Environmental Contamination and Toxicology*. Vol **228**. Cham: Springer International Publishing, 2014:101–20.

Nicolaisen M, Supronienė S, Nielsen LK *et al.* Real-time PCR for quantification of eleven individual Fusarium species in cereals. *J Microbiol Methods* 2009;**76**:234–40.

Nielsen LK, Jensen JD, Nielsen GC *et al.* Fusarium head blight of cereals in Denmark: species complex and related mycotoxins. *Phytopathology* 2011;**101**:960–9.

O'Donnell K, Sutton DA, Rinaldi MG *et al.* Internet-accessible DNA sequence database for identifying fusaria from human and animal infections. *J Clin Microbiol* 2010;**48**:3708–18.

R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2017.

Reddy K, Salleh B, Saad B *et al.* An overview of mycotoxin contamination in foods and its implications for human health. *Toxin Rev* 2010;**29**:3–26.

Ronquist F, Teslenko M, van der Mark P *et al.* MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 2012;**61**:539–42.

Sandoval-Denis M, Guarnaccia V, Polizzi G *et al.* Symptomatic citrus trees reveal a new pathogenic lineage in Fusarium and two new Neocosmospora species. *Persoonia* 2018;**40**:1–25.

Saremi H, Saremi H. Isolation of the most common Fusarium species and the effect of soil solarisation on main pathogenic species in different climatic zones of Iran. *Eur J Plant Pathol* 2013;**137**:585–96.

Scauflaire J, Mahieu O, Louvieaux J *et al.* Biodiversity of Fusarium species in ears and stalks of maize plants in Belgium. *Eur J Plant Pathol* 2011;**131**:59–66.

Scauflaire J, Godet M, Gourgue M *et al.* A multiplex real-time PCR method using hybridization probes for the detection and the quantification of Fusarium proliferatum, F. subglutinans, F. temperatum, and F. verticillioides. *Fungal Biol* 2012;**116**:1073–80.

Schloss PD, Jenior ML, Koumpouras CC *et al.* Sequencing 16S rRNA gene fragments using the PacBio SMRT DNA sequencing system. *PeerJ* 2016;**4**:e1869.

Silvestro LB, Stenglein SA, Forjan H *et al.* Occurrence and distribution of soil Fusarium species under wheat crop in zero tillage. *Span J Agric Res* 2013;**11**:72.

Smiley RW, Gourlie JA, Easley SA *et al.* Crop damage estimates for crown rot of wheat and barley in the Pacific Northwest. *Plant Dis* 2005;**89**:595–604.

Smith SN. An overview of ecological and habitat aspects in the genus Fusarium with special emphasis on the soil-borne pathogenic forms. *Plant Pathol Bull* 2007;**16**:97–120.

Sobrova P, Adam V, Vasatkova A *et al.* Deoxynivalenol and its toxicity. *Interdiscip Toxicol* 2010;**3**:94–97.

Summerell BA, Laurence MH, Liew ECY *et al.* Biogeography and phylogeography of Fusarium: a review. *Fungal Divers* 2010;**44**:3–13.

Tedersoo L, Tooming-Klunderud A, Anslan S. PacBio metabarcoding of Fungi and other eukaryotes: errors, biases and perspectives. *New Phytol* 2018;**217**:1370–85.

Waalwijk C, Kastekein P, de Vries I *et al.* Major changes in Fusarium spp. in wheat in the Netherlands. *Eur J Plant Pathol* 2003;**109**:743–54.

Walder F, Schlaeppi K, Wittwer R *et al.* Community profiling of Fusarium in combination with other plant-associated fungi in different crop species using SMRT sequencing. *Front Plant Sci* 2017;**8**:2019.

Wang H, Xiao M, Kong F *et al.* Accurate and practical identification of 20 Fusarium species by seven-locus sequence analysis and reverse line blot hybridization, and an in vitro antifungal susceptibility study. *J Clin Microbiol* 2011;**49**:1890–8.

West JS, Holdgate S, Townsend JA *et al.* Impacts of changing climate and agronomic factors on fusarium ear blight of wheat in the UK. *Fungal Ecol* 2012;**5**:53–61.

Xu X-M, Parry DW, Nicholson P *et al.* Predominance and association of pathogenic fungi causing Fusarium ear blightin wheat in four European countries. *Eur J Plant Pathol* 2005;**112**:143–54.

Xu X-M, Nicholson P, Thomsett MA *et al.* Relationship between the fungal complex causing Fusarium head blight of wheat and environmental conditions. *Phytopathology* 2008;**98**:69–78.

Xu X, Nicholson P. Community ecology of fungal pathogens causing wheat head blight. *Annu Rev Phytopathol* 2009;**47**:83–103.

Yli-Mattila T, Paavanen-Huhtala S, Jestoi M *et al*. Real-time PCR detection and quantification of Fusarium poae, F. graminearum, F. sporotrichioides and F. langsethiae in cereal grains in Finland and Russia. *Arch Phytopathology Plant Protect* 2008;**41**:243–60.

Yli-Mattila T. Ecology and evolution of toxigenic Fusarium species in Northern Europe and Asia. *J Plant Pathol* 2010;**92**:7–18.

Zhang N, O'Donnell K, Sutton DA *et al*. Members of the Fusarium solani Species Complex that cause infections in both humans and plants are common in the environment. *J Clin Microbiol* 2006;**44**:2186–90.

Zhang X, Halder J, White RP *et al*. Climate change increases risk of Fusarium ear blight on wheat in central China: Fusarium and climate change in China. *Ann Appl Biol* 2014;**164**:384–95.