Plant Disease • 2017 • 101:1515-1521 • http://dx.doi.org/10.1094/PDIS-08-16-1088-RE

e-Xtra*

PCR Markers Derived from Comparative Genomics for Detection and Identification of the Rice Pathogen *Ustilaginoidea virens* in Plant Tissues

Jintian Tang and **Lu Zheng,** The Key Lab of Plant Pathology of Hubei Province, Huazhong Agricultural University, Wuhan 430070, Hubei, PR China; **Qie Jia,** The Key Lab of Plant Pathology of Hubei Province; and College of Horticulture and Landscape Architecture, Yangtze University, Jingzhou 434023, Hubei, PR China; **Hao Liu,** The Key Lab of Plant Pathology of Hubei Province; **Tom Hsiang,** School of Environmental Sciences, University of Guelph, Guelph, ON N1G 2W1, Canada; and **Junbin Huang,** The Key Lab of Plant Pathology of Hubei Province

Abstract

Ustilaginoidea virens is a fungal pathogen that causes the false smut disease in rice and reduces the yield and quantity of the grains. A nested polymerase chain reaction (PCR)-based assay was developed to detect *U. virens* using genes of *U. virens* as specific targets. Ninety-six candidate genes of *U. virens* were found through first-round homology screening against a local database comprising 46 genomes of fungi, bacteria, and plants, with a second-round comparison with the GenBank NR database to further identify genes unique to *U. virens*. Among 96 remaining candidate genes, 20 of them (GenBank accessions KY617806 to KY617825) were randomly selected for further testing and, eventually, six sets of nested PCR primers were developed after further sensitivity, specificity, and detection tests. All six sets could detect DNA of *U. virens* at as little as 1 to 10 fg/μl from field or lab samples. These primers may be used to detect infection by *U. virens* at early stages, for use in research toward mitigating disease spread, as well as for studying the ecology of *U. virens*. This study also serves to illustrate that a comparative genomics method may allow for selection and development of highly specific primers once draft or complete genomes are available.

Rice false smut (RFS) is a serious disease affecting the growth of rice worldwide. It is caused by the fungus, *Ustilaginoidea virens* (Cooke) Takah. (also known as *Villosiclava virens* (M. Sakurai ex Nakata) E. Tanaka & C. Tanaka), which causes the grains to swell as the only external symptom; however, the grains become occupied by hyphae and chlamydospores which results in reduced yield (Fan et al. 2016). In China, the prevalence of RFS has increased and currently affects almost one-third of cultivated rice areas (Zhang et al. 2014). This higher prevalence is due, in part, to increased cultivation of high-yielding cultivars and hybrids that have more spikelets (and, hence, more false smut), and chemical fertilizer use (presumably increasing tissue susceptibility), as well as climate change creating conditions more suitable for disease development (Guo et al. 2012; Zhang et al. 2014). There have been many reports of severe outbreaks of RFS since 2001 in many rice-growing provinces of China such as Liaoning, Hubei, Sichuan, and Anhui, resulting in 20 to 50% yield loss in different areas and varying with different rice varieties. In 2005, RFS occurred across approximately 330,000 ha in Sichuan Province, and a third of the panicles were estimated to be affected (Lu et al. 2009). Outbreaks of the disease caused not only severe grain yield reductions but also contamination of rice grain by fungal toxins (ustiloxins) produced by false smut balls (Koiso et al. 1994; Singh and Pophaly 2010).

Upon infection, rice grains are transformed into spore balls known as false smut balls, which are coated by dark-green, powdery chlamydospores, giving the spore balls a greenish and velvety look (Ou 1985). Stroma, asci, and ascospores develop from fungal sclerotia that form on the spore balls (Fan et al. 2015; Ikegami 1963; Zhang et al. 2014). Conidia are formed after germination of chlamydospores or ascospores and infect the rice spikelets and florets at the booting stage prior to rice plant heading (Ashizawa and Kataoka 2005; Zhou et al. 2003). However, it is not until after heading that spore balls protruding from the glume are apparent. Examination of artificially inoculated rice seed showed that the fungus could invade rice roots and colonize the surface of roots and leaf sheaths (Tanaka et al. 2016). The hyphae of *U. virens* were observed to invade spikelets through their apices via the small gap between the lemma and palea when inoculating at booting stage (Ashizawa et al. 2012). However, whether the hyphae can invade from the roots to the spikelets is still unknown. The widespread occurrence of *U. virens* infections is presumed to be due to the availability of alternative hosts such as rice weed (e.g., *Digitaria marginata*, *Panicum trypheron*, *Echinochloa crusgalli*, and *Imperata cylindrica*) and other grasses that serve as reservoirs of fungal spores during unfavorable seasons (Guo et al. 2012; Shetty and Shetty 1987). Therefore, it is necessary to detect the pathogen before symptom appearance in the rice fields to initiate effective control methods for RFS.

Molecular biology tools such as polymerase chain reaction (PCR) have been reported to effectively detect *U. virens* spores in asymptomatic rice plants and grains (Chen et al. 2014; Zhou et al. 2003). Moreover, real-time PCR (qPCR) has also been used to correlate the amount of *U. virens* in the soil and the prevalence of RFS (Ashizawa et al. 2010). qPCR was also applied as a tool to quantify *U. virens* in infected rice plants in order to find a method for early defense and control of RFS (Li et al. 2013). However, the previous studies found that some small-spored *Alternaria* spp. could not be differentiated from *U. virens* based on internal transcribed spacer (ITS) sequences (Pryor and Gilbertson 2000; Pryor and Michailides 2002). The same situation also exists in the *Mycena pura* morphospecies (Harder et al. 2013) and some *Fusarium* spp. (Yli-Mattila et al. 2004) and, thus, fungal species separation based on ITS regions may not give sufficient resolution. There is a need to develop a highly sensitive assay to accurately determine *U. virens* disease prevalence, which may currently be over- or underestimated.

Basic Local Alignment Search Tool (BLAST), a sequence similarity search program (McGinnis and Madden 2004), can be used to find differences or similarities between query sequences and genomes, and explore novel specific targets by comparative genomics methods. Recent advances in sequencing technology such as Illumina have provided economically accessible opportunities for whole-genome sequencing (DiGuistini et al. 2009; Haridas et al. 2011). The current GenBank inventory contains 45,000 assembled genomes and more than 300,000 nucleotide sequences for rice species (Benson et al. 2015). This large number of sequences may provide the basis to uncover numerous species-specific detection targets and be useful in other studies involving sequence analyses (Haridas et al. 2011).

Comparative genomics analyses have been reported as a new approach for mining specific detection targets in microbial species (Malapi-Wight

Corresponding author: J. Huang; E-mail: junbinhuang@mail.hzau.edu.cn

*The ***e*-Xtra** logo stands for "electronic extra" and indicates that two supplementary tables are published online.

Accepted for publication 24 March 2017.

et al. 2016; Yu et al. 2010). Recently, we completed the sequencing of a draft genome of *U. virens* (National Center for Biotechnology Information [NCBI] Short Read Archive: SRR1735304) (Jia et al. 2015). The objectives of this study were to test the development of specific markers using comparative genomics methods to mine for gene targets unique to this pathogen (compared with other sequenced fungal genomes), and to test these markers on DNA extracted from field samples.

Materials and Methods

Strains, culture conditions, and plants. In all, 97 isolates of *U. virens* from different locations in China, 16 other fungal pathogens, two bacterial pathogens, and two widely grown rice cultivars were used in this study (Table 1). Among the 16 other fungal pathogens, *Magnaporthe oryzae* and *Rhizoctonia solani* were selected because they are rice pathogens. *Beauveria bassiana*, *Metarhizium anisopliae*, and *Trichoderma koningiopsis* were chosen because they are closely related to *U. virens*. *Leptosphaeria biglobosa*, *Sclerotinia sclerotiorum*, *S. minor*, *S. nivalis*, *S. trifoliorum*, and *Botrytis cinerea* are common field plant pathogens, especially *L. biglobosa* and *S. sclerotiorum*, the pathogen of oilseed rape (*Brassica napus*). These two were especially targeted because of the ubiquitous rice-rapeseed rotation system in the southern half of China; the oilseed rape pathogens can be easily found in the fields used for rice culture. *Alternaria arborescens*, *A. alternata*, *Colletotrichum higginsianum*, *Coniothyrium minitans*, and *Phoma enqua* were chosen for a greater diversity of test subjects. The bacterial pathogen *Xanthomonas oryzae*, which causes rice bacterial blight, and *Ralstonia solanacearum*, another rice pathogen, are also common in rice fields and were selected for inclusion in the screening process. 'Wanxian98' *indica* rice and 'Huajing924' *japonica* rice were used as controls.

Strains of *U. virens* were cultured from the false smut balls and incubated at 28°C in potato sucrose broth (PSB) with shaking at 150 rpm and 28°C for spore harvesting. Spores were diluted and spread onto potato sucrose agar (PSA) medium for selecting single-spore cultures, which were incubated at 28°C on PSA to produce mycelia. Other fungal cultures capable of spore production were also single spored at least twice. Nonsporulating cultures and single-spored cultures were incubated at 25°C on potato dextrose agar (PDA) medium to obtain mycelia for DNA extraction. The bacterial strains were cultured at 28°C on liquid Luria-Bertani (LB) medium. Rice plants Wanxian98 and Huajing924 were planted in a greenhouse under a regime of 38°C for 12 h of light and 25°C for 12 h of darkness during summer 2014. Seed were first treated for 24 h with 0.15% potassium permanganate solution for sterilization and soaked in water for 2 days. The seed were incubated at 37°C for 48 h before sowing. Wanxian98 and Huajing924 rice seedlings were transplanted into individual plastic buckets (30 by 20 by 30 cm) containing 4 kg of paddy soil (previously autoclaved at 121°C for 1 h) and placed in the greenhouse. Synthetic fertilizer was applied during seedling and tillering stages. RFS-susceptible Wanxian98 rice was used for inoculation in this study.

DNA extraction. After 7 days for growth on PSA or PDA, fungal mycelia were harvested for DNA extraction. DNA of rice plants Wanxian98 and Huajing924 was extracted from leaves at tillering stage. Sampling of other rice tissues for DNA extraction is described later, in the section on detection of *U. virens* in inoculated and naturally infested rice plants. The genomic DNA of fungi and rice plants was extracted using the cetyltrimethyl ammonium bromide method (Sambrook et al. 1989). DNA from bacteria shaken in liquid LB for 1 day was extracted using the MiniBEST Bacterial Genomic DNA Extraction Kit (TaKaRa, Dalian, China).

Sources of genome sequence data. Library preparation, genome sequencing, and genome assembly of *U. virens* strain HWD-2 were done in a previous study (Jia et al. 2015). Scaffolds from the Abyss assembly were processed using AUGUSTUS software (version 2.5.5; http://augustus.gobics.de) for gene prediction (Stanke and Morgenstern 2005; Stanke and Waack 2003). In the Augustus program, the species model *Fusarium graminearum* was selected as the reference species due to the relatively close phylogenetic relationship of these two species (both are in the Sordariomycetous order Hypocreales). Predicted gene sequences from *U. virens* were then used for revealing genes specific to *U. virens*.

To find *U. virens*-specific markers, we set up a 46-genome database for first-round mining, based on the predicted protein sequences of 46 species, with particular attention to the host plant (e.g., *Oryza sativa*), other rice pathogens (e.g., *Magnaporthe oryzae* and *X. oryzae*), and closely related species (e.g., *Metarhizium acridum*, *M. anisopliae*, and *Claviceps purpurea*) (Supplementary Table S1). The genomic data were downloaded from NCBI on 5 July 2014 (ftp://ftp.ncbi.nih.gov/blast/db/FASTA).

***U. virens*-specific target mining and primer design.** Standalone BLAST, version 2.2.28+, was set up on a Microsoft Windows PC system, and BLAST databases were generated for each downloaded genome assembly and its associated predicted gene database (both protein and nucleotide sequences). Then, the 8,231 protein sequences of *U. virens* were searched against the 46-genome amino acid database using Standalone BLASTP for the first round to find matches. For a majority of the analyses, the threshold e-value of 1E-5 was selected as the indicator of homology. To parse the BLAST output files, PERL scripts (adapted from Hsiang and Baillie [2005]) were used (these scripts can be obtained by request to senior author Tom Hsiang). The scripts mined the following data from the BLAST output: query sequence name, query sequence length, top match, length of top match, e-value for top match, number of matches, and record of the next top five matches. The genes with zero matches against the 46-genome database were selected for the second-round testing by searching against the GenBank NR database using Standalone BLASTP. The genes mined through the two rounds which had no matches with *U. virens* genes were considered specific targets (Fig. 1).

After the second round, which identified 96 candidate gene targets, 20 were randomly selected to design primers using the software Oligo 7 (Molecular Biology Insights, Inc., Cascade, CO). All primers used in this study were synthesized by TSINGKE (Wuhan, China). To locate conserved domains within a protein or coding nucleotide sequence, the SMART software (Simple Modular Architecture Research Tool, http://smart.embl-heidelberg.de/) was used.

Nested PCR assays. Preliminary PCR assays for specificity and sensitivity (described later, in the section on PCR specificity and sensitivity) were first developed for all 20 selected target genes. Preliminary screening consisted of testing against 97 isolates of *U. virens*, 16 other species of fungal pathogens, two species of bacterial pathogens, two subspecies of rice, and the diluted DNA of *U. virens* strain HWD-2. After preliminary screening, six sets of primers were selected as the detection set. The PCR mixture consisted of 1 U of Taq DNA polymerase (TaKaRa), 2.5 μl of 10× PCR buffer, 2 μl of dNTP mixture (2.5 mM each), 1 μM each primer and DNA template (20 ng/μl), and sterile distilled water up to 25 μl. The amplification was carried out on a T100 Thermal Cycler (Bio-Rad, Singapore). The first-round thermal cycling conditions of the chosen primers consisted of an initial denaturation at 95°C for 3 min; followed by 30 amplification cycles as 95°C for 30 s, 55 or 50°C (depending on primer set) for 30 s, and 72°C for 30 s; and a final extension step at 72°C for 5 min. For the second-round amplification, 1 μl of the 10-fold diluted first-round PCR production was used as the template, the annealing temperature of the second-round primers ranged from 53 to 55°C, and the other conditions were performed as in the first-round amplification (Table 2). Outcomes from the second round of amplification were examined on 2% agarose gel and the DNA was stained with ethidium bromide.

PCR specificity and sensitivity. Specificity of the nested primer protocol was tested against 97 isolates of *U. virens*, 16 other species of fungal pathogens, two species of bacterial pathogens, and two subspecies of rice (Table 1). To first test the quality of DNA, primers of ITS region (ITS1/ITS4), 16s ribosomal DNA (rDNA) (27f/1492r), and rice gene *Actin1* (OsActin1-F/OsActin1-R) were separately used for positive control testing of the fungal, bacterial, and rice DNA (Fan et al. 2015; Lane 1991; White et al. 1990) (Table 1). For sensitivity testing, purified genomic DNA from *U. virens* strain HWD-2 was serially diluted to 10 ng/μl, 1 ng/μl, 100 pg/μl, 10 pg/μl, 1 pg/μl, 100 fg/μl, 10 fg/μl, and 1 fg/μl. Rice DNA from susceptible Wanxian98 at 100 ng/μl was used as background. Equal volumes of DNA from *U. virens* and rice were added into PCR mixtures and then checked by nested PCR. After sensitivity and specificity testing, six sets of nested primers were selected for *U. virens* detection.

Detection of *U. virens* in inoculated and naturally infested rice plants. Rice panicles (Wanxian98) were inoculated with *U. virens*

strain HWD-2 in the greenhouse following Jia et al. (2014). To produce a sufficient quantity of conidia for artificial inoculation, a 6-mm-diameter mycelial plug incubated on PSA for 7 days was placed in 150 ml of sterilized PSB in a 250-ml flask, which was then put in a shaking incubator at 28°C in the dark at 180 rpm for 7 days. The cultures were filtered through four layers of gauze, and conidia were retrieved from the filtrate. Twenty panicles of rice plants at the booting stage were inoculated with conidial suspensions (1×10^6 conidia/ml) at 5 to 7 days before earing. Twenty rice panicles injected with PSB were used as mock-inoculated controls. The inoculated rice plants were kept at a constant 25°C and 90% relative humidity. Three injected rice panicles and three panicles of controls were randomly collected for detection by nested PCR using the six sets of primers at 2 days after inoculation (DAI). The disease incidence rate of RFS was assessed at 15 DAI.

Five symptomless rice plants at the booting stage were collected randomly from five field locations in Wuhan, China in October 2014 in which RFS was present in the previous year. Each rice plant was divided into four parts, including roots, stems, leaves, and spikelets, giving 20

Table 1. Fungal and bacterial species and plants tested in this study

Species	Isolate	Host	Location	PCR results of each targets[a]						
				G544	G858	G924	G1384	G1581	G5140	ITS[b]
Fungal pathogens										
Ustilaginoidea virens	9 isolates (JC2-3, JC4-2, JC5-3, JC6-3, JC8-1, JC23-2, JC31-1, JC32-3, JC38-1)	Rice	Jilin	+	+	+	+	+	+	+
	10 isolates (DJ-09-1, DT-09-1, LN-LY-1, LN-10-16-1, LZ-09-1, LZ-09-2, LZ-09-3, SY02, TL-09-1, YY8-09-1)	Rice	Liaoning	+	+	+	+	+	+	+
	4 isolates (SX0204, SX0205, SX-10-9-1, SX-10-9-2)	Rice	Shanxi	+	+	+	+	+	+	+
	4 isolates (HN-LS-5-1, HN-LS-6-2, HN-65-1, HN-65-5)	Rice	Henan	+	+	+	+	+	+	+
	16 isolates (HWD-2, S6-3, YH2-3, ZH29-5, JZ2-2, JZ4-4, JZ8-3, XGA-2, XGA-3, XGB-2, XN1-3, XN4-2, XN6-2, HYS-10, HYS-29, HYS-80)	Rice	Hubei	+	+	+	+	+	+	+
	9 isolates (HH4-3, HCA4-3, HCA12-1, HCF2-3, HCF7-2, HCF6-1, HYD5-2, HYD7-1, HYD10-3)	Rice	Hunan	+	+	+	+	+	+	+
	8 isolates (JFC1-1, JFC4-2, JFC5-1, JFD2-1, JFD5-2, JJW1-1, JJW2-1, JJW3-1)	Rice	Jiangxi	+	+	+	+	+	+	+
	12 isolates (ZHD3-1, ZHD4-1, ZHD5-1, ZLL1-3, ZLL6-1, ZLL9-3, ZQK1-2, ZQK4-1, ZQJ1-1, ZWW1-1, ZWW3-3, ZWW5-3)	Rice	Zhejiang	+	+	+	+	+	+	+
	11 isolates (FLW2-2, FLW3-1, FLW6-2, FMB1-1, FMB8-1, FMB10-1, FNY1-3, FNY3-3, FS8-2, FS9-3, FS10-1)	Rice	Fujian	+	+	+	+	+	+	+
	13 isolates (GBL1-1, GBL4-2, GBL5-1, GHT1-2, GHT3-1, GGH7-1, GGH13-1, GGH18-1, GGP1-2, GGP3-1, GY2-1, GY4-1, GY5-2)	Rice	Guangxi	+	+	+	+	+	+	+
	1 isolate (GD-10-3-4)	Rice	Guangdong	+	+	+	+	+	+	+
Alternaria arborescens	THX-1-1	Tobacco	Enshi, Hubei	−	−	−	−	−	−	+
A. alternata	THX-2-1	Tobacco	Enshi, Hubei	−	−	−	−	−	−	+
Botrytis cinerea	GHW-3-2	Strawberry	Wuhan, Hubei	−	−	−	−	−	−	+
Beauveria bassiana	SHW-20	Pine moth	Huanggang, Hubei	−	−	−	−	−	−	+
Colletotrichum higginsianum	C	*Brassica campestris*	Canada	−	−	−	−	−	−	+
Coniothyrium minitans	CHY-9	*Sclerotinia sclerotiorum*	Changyang, Hubei	−	−	−	−	−	−	+
Leptosphaeria biglobosa	Lb-1	Oilseed rape	Wuhan, Hubei	−	−	−	−	−	−	+
Magnaporthe oryzae	189-36-1	Rice	Yichang, Hubei	−	−	−	−	−	−	+
Metarhizium anisopliae	HWC-1	Unknown	Wuhan, Hubei	−	−	−	−	−	−	+
Phoma enqua	EP-5	*Coptis chinensis*	Enshi, Hubei	−	−	−	−	−	−	+
Rhizoctonia solani	199	Rice	Yichang, Hubei	−	−	−	−	−	−	+
Sclerotinia minor	SM-8	*Lactuca sativa*	Huanggang, Hubei	−	−	−	−	−	−	+
S. nivalis	SN-3	Unknown	Wuhan, Hubei	−	−	−	−	−	−	+
S. sclerotiorum	SS-18	Oilseed rape	Wuhan, Hubei	−	−	−	−	−	−	+
S. trifoliorum	ST-36	Unknown	Wuhan, Hubei	−	−	−	−	−	−	+
Trichoderma koningiopsis	F-35	Unknown	Wuhan, Hubei	−	−	−	−	−	−	+
Bacterial pathogens										16s rDNA
Ralstonia solanacearum	QK-45	Tobacco	Enshi, Hubei	−	−	−	−	−	−	+
Xanthomonas oryzae	HW-67	Rice	Wuhan, Hubei	−	−	−	−	−	−	+
Rice species (*Oryza sativa*)										Actin1
Wanxian98	…	…	…	−	−	−	−	−	−	+
Huajing924	…	…	…	−	−	−	−	−	−	+

[a] Symbols + and − indicate polymerase chain reaction (PCR)-positive and -negative isolated, respectively. Isolates are available upon request.
[b] Internal transcribed spacer (ITS) positive control.

samples. DNA was isolated from each sample and used in a nested PCR assay. A 1-μl aliquot of each DNA extraction was added to the reaction mixture for nested PCR with the six sets of nested PCR primers.

Results

***U. virens*-specific target mining.** The first-round mining consisted of comparison of the 8,231 predicted genes of *U. virens* against a 46-genome database using Standalone BLASTP. The 46 genomes consisted of 27 Ascomycota (5 Dothideomycetes, 6 Eurotiomycetes, 2 Leiotiomycetes, 2 Saccharomycetes, and 12 Sordariomycetes), 11 Basidiomycota, 1 Chytridiomycota, 2 Mucoromycota, 1 Oomycota, 1 plant, and 3 bacteria. There were no matches for 96 genes; therefore, these were subjected to second-round comparisons with the GenBank NR database, again with no matches found. Thus, the 96 were considered putative *U. virens*-specific targets. From these 96, 20 genes were chosen randomly as candidate targets for primer

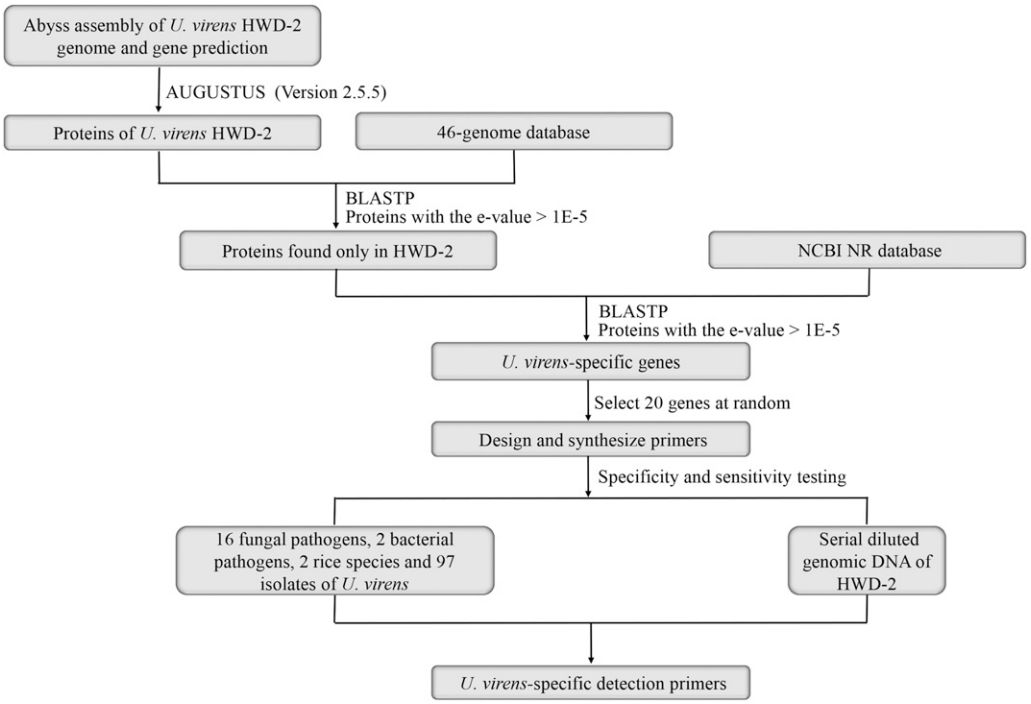

Fig. 1. Pipeline for mining specific genes of *Ustilaginoidea virens* using a complete genome sequence and comparative genomic analysis. NCBI = National Center for Biotechnology Information.

Table 2. Sequence characteristics of the selected nested polymerase chain reaction (PCR) primer sets used in this study

Gene region ID number	GenBank accession	Gene length (bp)	Round	Primer	Sequence(5′–3′)	PCR product length (bp)	Annealing temperature (°C)
G544	KY617817	543	First	Uv544F-O	TCTTGGCTCCTCGGAAGCTC	453	55
	…	…	…	Uv544R-O	CATGTCTTGCCCAAAGATGCG	…	…
	…	…	Second	Uv544F-I	AGGTTCTACATGCGTGTTGTGA	366	53
	…	…	…	Uv544R-I	GCCCTGTGAAACTTGGTGC	…	…
G858	KY617819	573	First	Uv858F-O	ACGAGCAGGCAGCAAACACC	373	55
	…	…	…	Uv858R-O	CCGCCACTTTGCCTGTCGATG	…	…
	…	…	Second	Uv858F-I	TCTTCTTGCGCCTTCTGACA	253	53
	…	…	…	Uv858R-I	GCTCCTCCCATGCGTTTT	…	…
G924	KY617821	651	First	Uv924F-O	TCCTCGTCCCTCCTGTCTT	562	55
	…	…	…	Uv924R-O	CGCCGAGAAGTATCCATCC	…	…
	…	…	Second	Uv924F-I	CGACCTTGTCCACCTCGACCA	432	55
	…	…	…	Uv924R-I	TCACCCTGGCTCTCTTGCGAA	…	…
G1384	KY617823	441	First	Uv1384F-O	AGCCCTTCTTTCTCGTCA	436	50
	…	…	…	Uv1384R-O	TATTTCGGACCCGTGTTT	…	…
	…	…	Second	Uv1384F-I	CGTTTTCAAGCGGGACCCTA	261	54
	…	…	…	Uv1384R-I	TCCGAGTTATTACCAGACATGC	…	…
G1581	KY617824	812	First	Uv1581F-O	CGCCCTGAAATCGCCCTTTA	703	55
	…	…	…	Uv1581R-O	ACCTGACCGTCTTGGCTCCC	…	…
	…	…	Second	Uv1581F-I	TCCTTCTTGTCTCGGCGTCT	341	55
	…	…	…	Uv1581R-I	CCCTTGCCAATGGTCCATGCTT	…	…
G5140	KY617825	578	First	Uv5140F-O	CCAGGCTCAGGTCAACTTCA	484	55
	…	…	…	Uv5140R-O	CGCAGACCTTTCCCTTGC	…	…
	…	…	Second	Uv5140F-I	TCGGCACGCACCAGACACGTT	353	55
	…	…	…	Uv5140R-I	TGGAGCCAGCGGAAACATGCAAA	…	…

design. The primers were generated and used in PCR with DNA from *U. virens* strain HWD-2, and the PCR amplicons were sequenced. The sequencing results were compared with the respective genes, and these showed that all 20 genes were present (Table 2; Supplementary Table S2).

The conserved motifs of the 20 genes were accessed by SMART software and the results showed that no conserved domains were detected in any tested genes, and that only transmembrane segments existed in the genes G702, G859, G924, and G1384. The functions of these 20 genes remain unknown.

Specificity and sensitivity of *U. virens*-specific PCR detection primers. To assess specificity of PCR detection, 97 *U. virens* strains, 16 other fungal pathogens, two bacterial pathogens, and two species of rice were tested by nested PCR using the 20 nested PCR primer sets designed based on the 20 candidate *U. virens*-specific genes. The results of quality testing with the positive control primer sets showed that all of the DNA samples were of sufficient quality for a nested PCR assay (Table 1). Within the 20 nested PCR primer sets, only the gene G192 was not amplified from DNA of any of the 97 *U. virens* strains using primer sets Uv192F-O/Uv192R-O and Uv192F-I/Uv192R-I. For the other 19 nested PCR primer sets, amplification of genomic DNA of 97 *U. virens* strains resulted in the predicted amplicon lengths of 200 to 800 bp, whereas no products were obtained from the 16 other fungal pathogens, two bacterial pathogens, or two species of rice (Table 1). The 19 sets of nested PCR primers were then used for sensitivity tests.

The results of sensitivity PCR assays showed that, among 19 sets of nested PCR primers, amplicons were produced by four sets when template DNA at 1 fg/μl was used: Uv544F-O/Uv544R-O and Uv544F-I/Uv544R-I, Uv858F-O/Uv858R-O and Uv858F-I/Uv858R-I, Uv1384F-O/Uv1384R-O and Uv1384F-I/Uv1384R-I, and Uv1581F-O/Uv1581R-O and Uv1581F-I/Uv1581R-I. Two sets were able to detect template DNA at 10 fg/μl: Uv924F-O/Uv924R-O and Uv924F-I/Uv924R-I, and Uv5140F-O/Uv5140R-O and Uv5140F-I/Uv5140R-I (Fig. 2). The sensitivity of the other 13 sets of nested primers was pegged at 10 or 1 pg/μl; they were not as sensitive as the six sets mentioned above. These results suggested that the six sets of nested PCR primers of the target genes G544, G858, G924, G1384, G1581, and G5140 showed high specificity and sensitivity, and could be used as *U. virens*-specific PCR detection primers.

Detection of *U. virens* in infested rice plants. The detection assay tests for *U. virens* simulated field samples which had contaminating DNA from the rice plants and other microorganisms. Results showed that all target DNA fragments could be detected at 2 dpi. There was no amplification for the PSB control treatment (Fig. 3). Individual grains from the remaining 17 nontested inoculated panicles later developed greenish spore balls (false smut balls), whereas these were not observed in the mock-inoculated controls. The results showed that the primers could detect the pathogen at an early stage of infection with a high background of DNA from other sources.

PCR assays were also performed with field samples. *U. virens* could be detected in the first rice plant (root, stem, and leaves), third rice plant (stem and spikelet), and fourth rice plant (spikelet) but not in the second or fifth rice plants (Fig. 4). All of the detection results of the six sets of nested PCR primers were the same for each sample, showing the consistency and stability of the nested PCR assays based on six unique genes of *U. virens* for pathogen detection in the symptomless rice plants.

Discussion

Molecular biology tools such as PCR are widely used to detect infectious microorganisms, with the success of these assays depending primarily on the sensitivity and specificity of primers to the target sequences. In recent mycological research, many PCR detection targets for fungi and fungal pathogens have been based on the ITS region (rDNA ITS) because of heterogeneity of the noncoding ITS sequences and their high copy number (White et al. 1990). However, some closely related fungal species show extremely high nucleotide

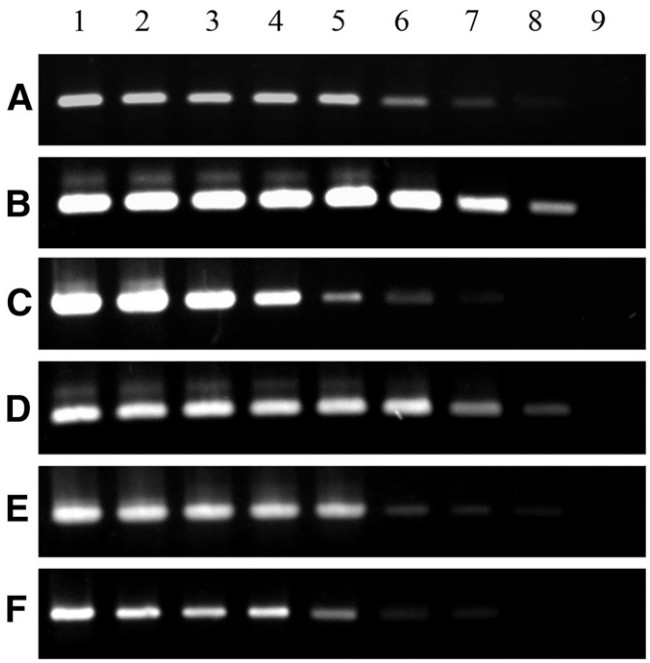

Fig. 2. Sensitivity test of the nested polymerase chain reaction with six primer sets in amplifying extracted DNA of *Ustilaginoidea virens* HWD-2. Amplification results of primer sets **A,** Uv544F-O/Uv544R-O and Uv544F-I/Uv544R-I; **B,** Uv858F-O/Uv858R-O and Uv858F-I/Uv858R-I; **C,** Uv924F-O/Uv924R-O and Uv924F-I/Uv924R-I; **D,** Uv1384F-O/Uv1384R-O and Uv1384F-I/Uv1384R-I; **E,** Uv1581F-O/Uv1581R-O and Uv1581F-I/Uv1581R-I; and **F,** Uv5140F-O/Uv5140R-O and Uv5140F-I/Uv5140R-I. Lanes 1 to 8: *U. virens* DNA at 10 ng/μl, 1 ng/μl, 100 pg/μl, 10 pg/μl, 1 pg/μl, 100 fg/μl, 10 fg/μl, and 1 fg/μl, respectively, plus 100 ng of DNA of rice as background; lane 9: non-DNA negative control.

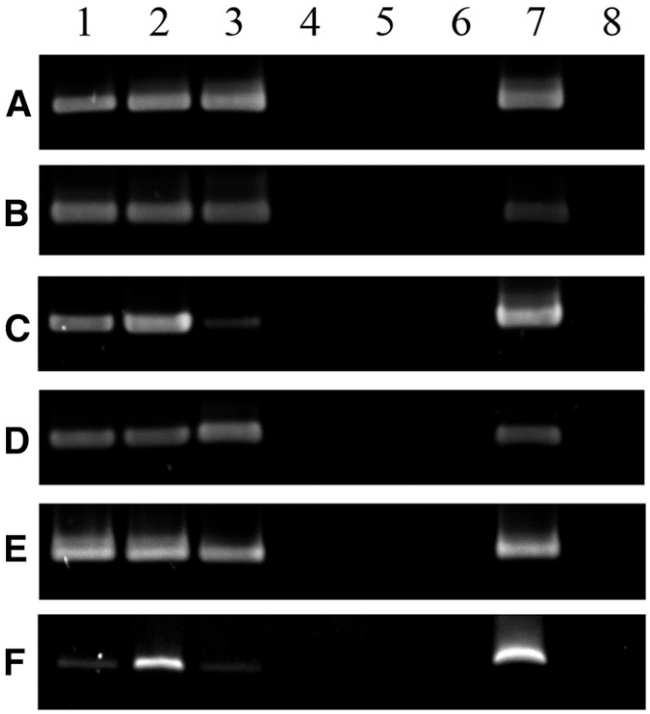

Fig. 3. Detection of *Ustilaginoidea virens* in rice spikelets after inoculation at the booting stage. Amplification results of primer sets **A,** Uv544F-O/Uv544R-O and Uv544F-I/Uv544R-I; **B,** Uv858F-O/Uv858R-O and Uv858F-I/Uv858R-I; **C,** Uv924F-O/Uv924R-O and Uv924F-I/Uv924R-I; **D,** Uv1384F-O/Uv1384R-O and Uv1384F-I/Uv1384R-I; **E,** Uv1581F-O/Uv1581R-O and Uv1581F-I/Uv1581R-I; and **F,** Uv5140F-O/Uv5140R-O and Uv5140F-I/Uv5140R-I. Lane M: 100-bp DNA ladder; lanes 1 to 3: three randomly collected spikelets inoculated with conidial suspensions; lane 4 to 6: three randomly collected spikelets inoculated with potato sucrose broth; lane 7: positive control; lane 8: non-DNA negative control.

sequence identity in the ITS regions, such as the 99% identity observed in the ITS regions between *S. sclerotiorum* and *B. cinerea*, with only eight known polymorphisms (Qin et al. 2011). This poses an issue for specific primer design.

Previously published qPCR and conventional PCR primers for *U. virens* detection have all been based on the ITS region (Ashizawa et al. 2010; Chen et al. 2014; Li et al. 2013; Zhou et al. 2003). In this study, the gene regions targeted for primer design were mined from whole-genome sequences. Specificity of these primers was assessed using 16 other fungal species, plus 97 strains of *U. virens* from different rice-growing regions. To increase the sensitivity of PCR detection, nested PCR was chosen for its higher sensitivity (Zhou et al. 2003). The limit of detection for all six sets of nested PCR primers developed in this study was below 10 fg/μl or even 1 fg/μl, compared with only 20 to 50 fg/μl for published primer sets based on the ITS region. In general, any one of six primer sets obtained in this study should be sufficient for specific detection of *U. virens* but testing with multiple primer sets can provide confirmational results. Use of multiple primer sets for detection may also help to address the issue of potential contamination, especially where nested PCR may generate artifacts by greatly amplifying possible highly similar contaminating DNA sequences.

This study illustrates that the use of comparative genomics for identifying and selecting target genes is highly efficient. From 20 randomly selected genes out of the unique pool of 96 genes, 6 (>30%) were found to provide sensitive and specific detection of *U. virens*. This same approach can be used to generate higher-level taxon-specific primers as well, such as at the genus or family level, if data for the appropriate selection of representative taxa is available. Compared with previous reports on the use of comparative genomics methods for selecting detection targets (Malapi-Wight et al. 2016; Yu et al. 2010), our use of the GenBank NR database in the second screening round of the pipeline provided additional validation that the candidate gene regions were specific to *U. virens*. Because of the huge and ever-growing amount of data in the GenBank NR database, we chose two rounds of mining, accessing the large database only in the reduced second round. Most of the 8,231 predicted genes could be eliminated through screening against a local StandAlone 46-genome database of related and coexisting fungal species using Standalone BLASTP, which saved time and reduced computational requirements compared with screening all 8,231 genes against the GenBank NR database,

In this study, we found 96 unique genes of *U. virens* through genome comparisons, and randomly chose 20 as targets. Only one of these target genes, G192, a hypothetical protein, could not be amplified from most of the 97 test isolates of *U. virens.* This predicted gene or the particular allele may be specific to our sequenced genome and close relatives because it was detected only in the isolates from Hubei Province and Shanxi Province (data not shown). Isolate HWD-2 of *U. virens*, which we used for genome sequencing, was obtained from Hubei Province (Table 1). There may be some relationships between genotypes and geographical locations, because Jia et al. (2015) showed that geographic distance and genetic distance of *U. virens* were highly correlated (Jia et al. 2015). Therefore, multiple introductions or migration of *U. virens* genotypes with particular unique genes could be further explored by comparative genomics once their genomes are available.

Considering the gap in our understanding of the RFS cycle and the mechanisms of *U. virens* infection, methods for accurate detection of *U. virens* in the field can be very useful to better explore these processes, which ultimately may lead to improved disease forecasting and disease control. The fungus is able to invade rice spikelets and infect rice florets before the heading stage or at the booting stage (Ashizawa and Kataoka 2005; Zhou et al. 2003). However, how the pathogen penetrates through the sheaths to rice florets is still unknown. In this study, the pathogen in the booting stage was detected using our primers, and there was some correlation between the level of field inoculum and disease severity of disease (data not shown).

In conclusion, we have mined specific markers for detection and identification of *U. virens* within the plant host using a comparative genomics method. The six sets of nested PCR primers had high sensitivity and specificity that could be used to detect *U. virens* in different rice tissues or even in the field, which might help to further elucidate the mechanisms of *U. virens* infection and the rice false smut disease cycle.

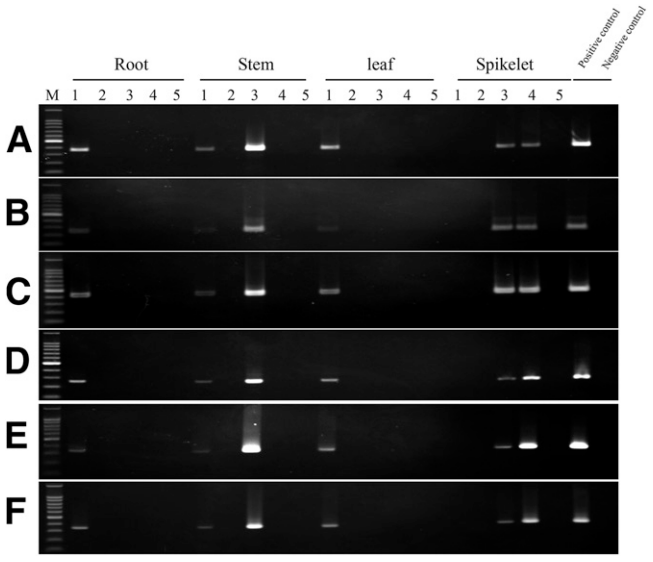

Fig. 4. Detection of *Ustilaginoidea virens* from roots, stems, leaves, and spikelets of rice plants at booting stage in campus experimental fields of Huazhong Agricultural University (Wuhan, China). Amplification results of primer sets **A,** Uv544F-O/Uv544R-O and Uv544F-I/Uv544R-I; **B,** Uv858F-O/Uv858R-O and Uv858F-I/Uv858R-I; **C,** Uv924F-O/Uv924R-O and Uv924F-I/Uv924R-I; **D,** Uv1384F-O/Uv1384R-O and Uv1384F-I/Uv1384R-I; **E,** Uv1581F-O/Uv1581R-O and Uv1581F-I/Uv1581R-I; and **F,** Uv5140F-O/Uv5140R-O and Uv5140F-I/Uv5140R-I. Lane M: 100-bp DNA ladder; lanes 1 to 5: different rice plants collected randomly from the rice fields numbered 1 to 5.

Acknowledgments

This work is supported by the National Key Research and Development Program (2016YFD0300700), the Fundamental Research Funds for the Central Universities of China (2662015QC022, number 2013PY110), and the Program for Changjiang Scholars and Innovative Research Team in University (IRT1247).

Literature Cited

Ashizawa, T., and Kataoka, Y. 2005. Detection of *Ustilaginoidea virens* in rice panicles before and after heading in the field using nested-PCR technique with species-specific primers. Jpn. J. Phytopathol. 71:16-19. (In Japanese)

Ashizawa, T., Takahashi, M., Arai, M., and Arie, T. 2012. Rice false smut pathogen, *Ustilaginoidea virens*, invades through small gap at the apex of a rice spikelet before heading. J. Gen. Plant Pathol. 78:255-259.

Ashizawa, T., Takahashi, M., Moriwaki, J., and Hirayae, K. 2010. Quantification of the rice false smut pathogen *Ustilaginoidea virens* from soil in Japan using real-time PCR. Eur. J. Plant Pathol. 128:221-232.

Benson, D. A., Clark, K., Mizrachi, I. K., Lipman, D. J., Ostell, J., and Sayers, E. W. 2015. GenBank. Nucleic Acids Res. 43:D30-D35.

Chen, Y., Yao, J., Li, Y. F., Wang, W. X., Yang, X., and Zhang, A. F. 2014. Simple and rapid detection of rice false smut pathogen *Ustilaginoidea virens* in rice seeds. Phytoparasitica 42:371-375.

DiGuistini, S., Liao, N. Y., Platt, D., Robertson, G., Seidel, M., Chan, S. K., Docking, T. R., Birol, I., Holt, R. A., Hirst, M., Mardis, E., Marra, M. A., Hamelin, R. C., Bohlmann, J., Breuil, C., and Jones, S. J. M. 2009. De novo genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. Genome Biol. 10:R94.

Fan, J., Guo, X. Y., Li, L., Huang, F., Sun, W. X., Li, Y., Huang, Y. Y., Xu, J. Y., Shi, J., Lei, Y., Zheng, A. P., and Wang, W. M. 2015. Infection of *Ustilaginoidea virens* intercepts rice seed formation but activates grain-filling-related genes. J. Integr. Plant Biol. 57:577-590.

Fan, J., Yang, J., Wang, Y. Q., Li, G. B., Li, Y., Huang, F., and Wang, W. M. 2016. Current understanding on *Villosiclava virens*, a unique flower-infecting fungus causing rice false smut disease. Mol. Plant Pathol. 17:1321-1330.

Guo, X. Y., Li, Y., Fan, J., Li, L., Huang, F., and Wang, W. M. 2012. Progress in the study of false smut disease in rice. J. Agric. Sci. Technol. 2:1211-1217.

Harder, C. B., Læssøe, T., Frøslev, T. G., Ekelund, F., Rosendahl, S., and Kjøller, R. 2013. A three-gene phylogeny of the *Mycena pura* complex reveals 11 phylogenetic species and shows ITS to be unreliable for species identification. Fungal Biol. 117:764-775.

Haridas, S., Breuill, C., Bohlmann, J., and Hsiang, T. 2011. A biologist's guide to de novo genome assembly using next-generation sequence data: A test with fungal genomes. J. Microbiol. Methods 86:368-375.

Hsiang, T., and Baillie, D. L. 2005. Comparison of the yeast proteome to other fungal genomes to find core fungal genes. J. Mol. Evol. 60:475-483.

Ikegami, H. 1963. Occurrence and development of sclerotia of the rice false smut fungus. Res. Bull. Fac. Agric. Gifu Univ. 18:47-53.

Jia, Q., Gu, Q., Zheng, L., Hsiang, T., Luo, C., and Huang, J. 2015. Genetic analysis of the population structure of the rice false smut fungus, *Villosiclava virens*, in China using microsatellite markers mined from a genome assembly. Plant Pathol. 64:1440-1449.

Jia, Q., Lv, B., Guo, M. Y., Luo, C. X., Zheng, L., Hsiang, T., and Huang, J. B. 2014. Effect of rice growth stage, temperature, relative humidity and wetness duration on infection of rice panicles by *Villosiclava virens*. Eur. J. Plant Pathol. 141:15-25.

Koiso, Y., Li, Y., Iwasaki, S., Hanaoka, K., Kobayashi, T., Sonoda, R., Fujita, Y., Yaejashi, H., and Sato, Z. 1994. Ustiloxins, antimitotic cyclic peptides from false smut balls on rice panicles caused by *Ustilaginoidea virens*. J. Antibiot. 47:765-773.

Lane, D. J. 1991. 16S/23S rRNA sequencing. Pages 115-176 in: Nucleic acid techniques in bacterial systematics. E. Stackebrandt and M. Goodfellow, eds. John Wiley & Sons, Inc., New York.

Li, H., Ni, D. H., Duan, Y. B., Chen, Y., Li, J., Song, F. S., Li, L., Wei, P. C., and Yang, J. B. 2013. Quantitative detection of the rice false smut pathogen *Ustilaginoidea virens* by real-time PCR. Genet. Mol. Res. 12:6433-6441.

Lu, D. H., Yang, X. Q., Mao, J. H., Ye, H. L., Wang, P., Chen, Y. P., He, Z. Q., and Chen, F. 2009. Characterising the pathogenicity diversity of *Ustilaginoidea virens* in hybrid rice in China. J. Plant Pathol. 91:443-451.

Malapi-Wight, M., Demers, J. E., Veltri, D., Marra, R. E., and Crouch, J. A. 2016. LAMP detection assays for boxwood blight pathogens: A comparative genomics approach. Sci. Rep. 6:Article 26140. doi:10.1038/srep26140

McGinnis, S., and Madden, T. L. 2004. BLAST: At the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res. 32:W20-W25.

Ou, S. H. 1985. Rice Diseases, Second Edition. Commonwealth Agricultural Bureaux, Farnham Royal, UK.

Pryor, B. M., and Gilbertson, R. L. 2000. Molecular phylogenetic relationships amongst *Alternaria* species and related fungi based upon analysis of nuclear ITS and mtSSU rDNA sequences. Mycol. Res. 104:1312-1321.

Pryor, B. M., and Michailides, T. J. 2002. Morphological, pathogenic and molecular characterization of *Alternaria* isolates associated with Alternaria late blight of pistachio. Phytopathology 92:406-416.

Qin, L., Fu, Y., Xie, J., Cheng, J., Jiang, D., Li, G., and Huang, J. 2011. A nested-PCR method for rapid detection of *Sclerotinia sclerotiorum* on petals of oilseed rape (*Brassica napus*). Plant Pathol. 60:271-277.

Sambrook, J., Frisch, E. F., and Maniatis, T. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Shetty, S. A., and Shetty, H. S. 1987. Role of *Panicum trypheron* in annual recurrence of false smut of rice. Trans. Br. Mycol. Soc. 88:409-411.

Singh, A. K., and Pophaly, D. J. 2010. An unusual rice false smut epidemic reported in Raigarh District, Chhattisgarh, India. Int. Rice Res. Notes. 35:1-3.

Stanke, M., and Morgenstern, B. 2005. AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. Nucleic Acids Res. 33:W465-W467.

Stanke, M., and Waack, S. 2003. Gene prediction with a hidden Markov model and a new intron submodel. Bioinformatics 19:ii215-ii225.

Tanaka, E., Kumagawa, T., Ito, N., Nakanishi, A., Ohta, Y., Suzuki, E., Adachi, N., Hamada, A., Ashizawa, T., Ohara, T., and Tsuda, M. 2016. Colonization of the vegetative stage of rice plants by the false smut fungus *Villosiclava virens*, as revealed by a combination of species-specific detection methods. Plant Pathol. 66:56-66.

White, T. J., Bruns, T., Lee, S., and Taylor, J. W. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. Pages 315-322 in: PCR Protocols: A Guide to Methods and Applications. M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White, eds. Academic Press, New York.

Yli-Mattila, T. Y., Huhtala, S. P., Parikka, P., Konstantinova, P., and Gagkaeva, T. Y. 2004. Molecular and morphological diversity of *Fusarium* species in Finland and northwestern Russia. Eur. J. Plant Pathol. 110:573-585.

Yu, S. J., Chen, W. Y., Wang, D. P., He, X. H., Zhu, X. N., and Shing, X. M. 2010. Species-specific PCR detection of the food-borne pathogen *Vibrio parahaemolyticus* using the *irgB* gene identified by comparative genomic analysis. FEMS Microbiol. Lett. 307:65-71.

Zhang, Y., Zhang, K., Fang, A. F., Han, Y. Q., Yang, J., Xue, M. F., Bao, J. D., Hu, W. D., Zhou, B., Sun, X. Y., Li, S. J., Wen, M., Yao, N., Ma, L. J., Liu, Y. F., Zhang, M., Huang, F., Luo, C. X., Zhou, L. G., Li, J. Q., Chen, Z. Y., Miao, J. K., Wang, S., Lai, J. S., Xu, J. R., Tom, H., Peng, Y. L., and Sun, W. X. 2014. Specific adaptation of *Ustilaginoidea virens* in occupying host florets revealed by comparative and functional genomics. Nat. Commun. 5:Article 3849. doi: 10.1038/ncomms4849

Zhou, Y. L., Izumitsu, K., Sonoda, R., Nakazaki, T., Tanaka, E., Tsuda, M., and Tanaka, C. 2003. PCR-based specific detection of *Ustilaginoidea virens* and *Ephelis japonica*. J. Phytopathol. 151:513-518.