Gene 585 (2016) 28–34

Contents lists available at ScienceDirect

Gene

journal homepage: www.elsevier.com/locate/gene

Research paper

Survey and analysis of simple sequence repeats in the *Ustilaginoidea virens* genome and the development of microsatellite markers

Mina Yu [a,b], Junjie Yu [a], Huanhuan Li [a], Yahui Wang [a], Xiaole Yin [a], Huiwen Bo [a,c], Hui Ding [a,c], Yuxin Zhou [a], Yongfeng Liu [a,b,*]

[a] Institute of Plant protection, Jiangsu Academy of Agricultural Science, Nanjing 210014, China
[b] Key Lab of Food Quality and Safety of Jiangsu Province—State Key Laboratory Breeding Base, Jiangsu Academy of Agricultural Science, Nanjing 210014, China
[c] Life Sciences College of Nanjing Agricultural University, Nanjing 210095, China

ARTICLE INFO

Article history:
Received 9 December 2015
Received in revised form 8 March 2016
Accepted 9 March 2016
Available online 16 March 2016

Keywords:
Ustilaginoidea virens
Simple sequence repeats
Genome

ABSTRACT

Ustilaginoidea virens is the causal agent of rice false smut, causing quantitative and qualitative losses in rice industry. However, the development and application of simple sequence repeat (SSR) markers for genetic diversity studies in *U. virens* were limited. This study is the first to perform large-scale development of SSR markers of this pathogen at the genome level, to (1) compare these SSR markers with those of other fungi, (2) analyze the pattern of the SSRs, and (3) obtain more informative genetic markers. *U. virens* is rich in SSRs, and 13,778 SSRs were identified with a relative abundance of 349.7 SSRs/Mb. The most common motifs in the genome or in noncoding regions were mononucleotides, whereas trinucleotides in coding sequences. A total of 6 out of 127 primers were randomly selected to be used to analyze 115 isolates, and these 6 primers showed high polymorphism in *U. virens*. This study may serve as an important resource for molecular genetic studies in *U. virens*.
© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Rice false smut (RFS) caused by *Ustilaginoidea virens* (Cooke) Takah (Teleomorph: *Villosiclava virens*), a pathogenic ascomycete fungus, is a plant disease that causes quantitative and qualitative losses in rice industry worldwide. *U. virens* infects individual grains through the small gap at the apex of the grains, and this fungus exclusively attacks the upper parts of the rice filaments during booting stage (Hu et al., 2014; Tang et al., 2013). White hyphae can be observed at the onset of the disease. Subsequently, a layer of powdery chlamydospores forms on the surface of the smut balls, which are yellow in summer and dark -green in autumn (Tang et al., 2013; Fu et al., 2012). This phenomenon is the characteristic symptom of RFS. RFS was once categorized as a minor disease; however, it became prevalent in major rice-growing areas worldwide resulting recently from heavy application of N fertilizers, large-scale expansion of hybrid cultivars with no source of high-level resistance in the existing rice germplasm, and an apparent change in global and regional climates (Zhang et al., 2014; Yu et al., 2015). Furthermore, cyclopeptide mycotoxins produced by *U. virens*, especially chlamydospores, are toxic to both humans and animals (Shan et al., 2012).

Simple sequence repeats (SSRs), also called microsatellites, represent specific DNA sequences that consist of iterations of one to six base pair nucleotide motifs (Sharma et al., 2007; van Belkum et al., 1998). They are randomly distributed in eukaryotic and prokaryotic genomes, and they exhibit a high degree of variations in the number of repeats among individuals, resulting in strand slippage during DNA replication, repair, and recombination (Sharma et al., 2007; Huang et al., 2014). Otherwise, the flanking regions of microsatellites are well conserved, and can be detected using PCR amplification (Li et al., 2014). Thus, SSRs have become one of the most popular markers used in fungal genetic diversity studies because of their high reproducibility, multi-allelic nature, co-dominant mode of inheritability, abundance, and genome-wide coverage (Sharma et al., 2007; Yuan et al., 2015).

Studies on the diversity of *U. virens* are receiving considerable attention, and the main molecular markers used include random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism, and enterobacterial repetitive intergenic consensus, etc. (Zhou et al., 2004, Zhou et al., 2008; Zhang et al., 2009; Wang et al., 2014). However, the shortcomings common to these markers are their poor repeatability and dominance (Dutech et al., 2007). Sun et al. (2013) identified three single nucleotide polymorphism (SNP)-rich genomic regions as molecular markers to analyze the genetic diversity and population structure of *U. virens* in China (Sun et al., 2013). Wang et al. (2014) also used RAPD

Abbreviations: AFLP, amplified fragment length polymorphism; bp, base pair; CDS, coding sequence; CTAB, cetyltrimethyl ammonium bromide; dNTP, deoxyribonucleotide triphosphate; ERIC, enterobacterial repetitive intergenic consensus; GO, Gene Ontology; NCBI, National Center for Biotechnology Information; PCR, polymerase chain reaction; PDB, potato-dextrose broth medium; PDA, potato-dextrose-agar medium; RAPD, random amplified polymorphic DNA; RFS, rice false smut; SNP, single nucleotide polymorphism; SSR, simple sequence repeat.

* Corresponding author at: Institute of Plant Protection, Jiangsu Academy of Agricultural Science, Zhongling Street 50, Nanjing 210014, China.
E-mail address: liuyf@jaas.ac.cn (Y. Liu).

http://dx.doi.org/10.1016/j.gene.2016.03.016
0378-1119/© 2016 Elsevier B.V. All rights reserved.

and SNP to analyze the genetic diversity of isolates (Wang et al., 2014). However, the high cost of their generation and difficulty in documentation are two of the important relative drawbacks of SNP compared with the use of the equally promising simple sequence repeats (SSRs) (Sharma et al., 2007). Moreover, SNP showed better genetic relatedness with more population number at the population structure level, whereas SSR is more effective in genetic diversity analysis (Singh et al., 2013; Kong et al., 2014).

With the development of sequencing technology, large-scale genome sequencing initiatives on a growing number of fungal species are now providing the opportunity to evaluate the abundance and relative distribution of SSRs in different genera based on whole genomes (Li et al., 2014; Labbé et al., 2011; Murat et al., 2011). Labbé et al. (2011) analyzed the SSRs in *Laccaria bicolor* genome and the sequence polymorphism of 78 SSRs, and the results offer immediate applications in population genetics (Labbé et al., 2011). Murat et al. (2011) investigated the SSRs in perigord black truffle genome and identified new polymorphic SSR markers that can be used in studies on population genetics (Murat et al., 2011). Similar studies were performed on *Gaeumannomyces graminis* var. *tritici*, *Phytophthora* sp., *Candida tropicalis*, *Neurospora crassa* and many other species (Li et al., 2014; Murat et al., 2011; Schena et al., 2008; Wu et al., 2014a; Kim et al., 2008). Large-scale sequencing of RNA and genome in *U. virens* also encouraged the development of database-derived microsatellite markers. Wang et al. (2013) identified 1023 SSR loci in BAC-end sequences (BES) containing approximately 2.7 Mb sequences (Wang et al., 2013).Yu et al. (2014) also identified 12,298 SSRs in a 21 Mb transcriptome (Yu et al., 2014). However, these studies were not conducted at the genome-wide level, limiting the comprehensive assessment of genotype diversity and the subsequent selection of an optimized marker set. Moreover, these studies did not identify primers by using PCR for genetic analysis. In 2014, the whole genome of the *U. virens* isolate Uv_8b was published, providing a great platform for whole genome-level studies (Wang et al., 2014). This genome is rich in repeated sequences, including 18% transposable elements and 7% SSRs. The numerous repeated sequences are possibly a major factor that contributes to the evolution of both genes and genomes.

Therefore, in the present study, we investigated the relative abundance and the motif types of SSRs in the *U. virens* genome. We also compared the patterns of SSR in the *U. virens* genome with those of the sequenced genomes of seven other fungi. Finally, we used long motifs to design primers to find the polymorphic SSR markers for population genetic diversity studies of *U. virens* isolates. Information on the genetic diversity of *U. virens* populations will be useful in studying the genetic variation and population structure of the pathogen and in creating effective prevention and control strategies.

2. Materials and methods

2.1. Fungal isolate and culture conditions

A collection comprising 115 isolates was compiled for this study. All strains were isolated in 2013 from seven fields near the cities of Xuzhou, Yancheng, Yangzhou, Jintan, Nanjing, Huaian, and Ganyu in Jiangsu Province, China, through single-conidia isolation. All isolates were stored as conidial suspensions in potato dextrose broth (PDB) supplemented with 20% glycerol at −80 °C. The isolates were first maintained on potato dextrose agar (PDA) at 28 °C, and then grown in PDB culture prior to DNA isolation.

2.2. Extraction of genomic DNA

U. virens isolates were grown in 30 ml of PDB for 7 days at 28 °C in a 150 rpm orbital shaker. The mycelia were harvested by filtration and immediately ground into a fine powder in liquid nitrogen for DNA isolation using the cetyltrimethyl ammonium bromide method as described in (Yu et al., 2015). The DNA was visualized by 0.8% agarose gel electrophoresis to assess its quality and quantity and then stored at −20 °C for further use.

2.3. Genome sequence data

The genome data used in this analysis were downloaded from the publicly available databases listed in Table 1. The CDSs of these fungi were downloaded from a public website (http://fungi.ensembl.org/info/website/ftp/index.html). All of the complete genome sequences were downloaded in FASTA format for further analysis. The gff files with the localization of the gene CDS (coding sequence), introns and intergenic regions were generated according to the positions in the genome annotations. The positions of the UTR regions were not published, and all of the genomic regions excluding CDSs and introns, were classified as intergenic regions. The predicted protein-coding sequences of the *U. virens* gene models containing SSRs were analyzed using Blast2Go software. Gene Ontology (GO) terms were assigned based on the results for these gene models. A custom perl-based pipeline was used for the above-mentioned annotation.

2.4. SSR identification

The fungal genomes were analyzed for their SSR content by using the MISA software (http://www.pgrc.ipk-gatersleben.de/misa/). Repeat thresholds were adjusted to identify the perfect mono-, di-, tri-, tetra-, penta-, and hexa-nucleotide motifs with a minimum of 10, 6, 5, 5, 5, and 5 repeats, respectively, and an interruption (max difference for 2 SSRs) of 100 bp (Yu et al., 2014). Microsoft Excel was used for further statistical analysis.

2.5. Development of SSR markers and assessment of polymorphisms

In this study, we collected the top 23–28 repeat motifs from each class (di- to hexanucleotide) to design primers; these repeat motifs are the longest in each class, and long SSRs possibly show high frequency of polymorphism (Tang et al., 2008). The SSRs containing mononucleotides were not considered. Primer pairs in the flanking regions of SSRs were designed for PCR using Primer3 (http://pgrc.ipk-gatersleben.de/misa/primer3.html). The primer pairs flanking the SSR motifs were designed using the following parameters: primer product size ranging from 100 nt to 280 nt and stability of 250.

The primer pairs were initially tested and screened on four *U. virens* isolates. PCR was performed in a 20 μl of reaction volume that contained 10 μl of 2× PCR reaction mix buffer (Vazyme Biotech, Nanjing, China), 1 μl of template DNA (approximately 20 ng), and 10 μM of each primer. Amplification was performed in a thermal cycler (Bio-Rad MyCycler) by using the following program: an initial incubation at 94 °C for 3 min, followed by 35 cycles at 94 °C for 30 s, an appropriate annealing temperature (56–60 °C) for 30 s, 72 °C for 30 s, and a final extension reaction at 72 °C for 5 min. The PCR products were resolved on 1% (*w*/*v*) agarose gels, and the primer pairs that showed clear and reproducible amplification patterns were selected for further analysis. To test the applicability of the microsatellite loci in studying genetic diversity of *U. virens*, the operational SSR primers were selected for a second round using another five isolates. The PCR products were resolved on 2.0% (*w*/*v*) gels (1.0% agarose and 1.0% Synergel (GENTAUR, SYN-100)), and the primer pairs that showed polymorphic bands were selected for further analysis. Finally, 6 primers were randomly selected to analyze 115 isolates. PCR was performed following the same protocol described above, and the PCR products were resolved on an 8% (*w*/*v*) acryl-bisacrylamide gel.

The amplified fragments on the acryl-bisacrylamide gel images were scored manually as “1” or “0”, which indicates the presence or absence of equally sized DNA bands, respectively. Genetic diversity, including the number of alleles (Na), the number of effective alleles (Ne), Nei's

Download English Version:

https://daneshyari.com/en/article/2815018

Download Persian Version:

https://daneshyari.com/article/2815018

Daneshyari.com