

Draft Genome of *Fusarium oxysporum* f. sp. *ubense* Strain Tropical Race-4 Infecting Cavendish (AAA) Group of Banana in India

R. Thangavelu,[†] E. Edwin Raj, P. Pushpakanth, M. Loganathan, and S. Uma

ICAR-National Research Centre for Banana, Tiruchirappalli 620 102, India

Abstract

Fusarium wilt, caused by the fungus *Fusarium oxysporum* f. sp. *ubense*, is the most serious pandemic disease of banana. In this study, we report the draft genome of *F. oxysporum* f. sp. *ubense* vegetative compatibility group (VCG) 01213/16 of strain tropical race 4 (TR4) that infects the Cavendish (AAA) group of banana collected from the subtropical region in India. The genome assembly of *SFoc* TR4 comprises 47,384,463 bp with 4,034 contigs and 15,508 protein-coding regions. Based on VCG analysis, the fungal isolate belongs to *F. oxysporum* f. sp. *ubense* TR4 but the genome sequence of *SFoc* TR4 shows differences in secreted-in-xylem (SIX) protein gene clusters (specifically, SIX8) in comparison with the reference genome of *F. oxysporum* f. sp. *ycopersici* and *F. oxysporum* f. sp. *ubense* TR4.

Fusarium wilt, a serious and destructive fungal disease in banana caused by soilborne *Fusarium oxysporum* f. sp. *ubense*, threatens 15% of the global banana production and approximately 47% of the total banana cultivation area of the world (Ordoñez et al. 2016; Ploetz 2015). The races of *F. oxysporum* f. sp. *ubense* infecting banana have been classified into *F. oxysporum* f. sp. *ubense* race 1 (R1), race 2 (R2), and race 4 (R4). *F. oxysporum* f. sp. *ubense* R4 has been further divided into *F. oxysporum* f. sp. *ubense* tropical race 4 (TR4) and subtropical race 4 (STR4). Among *F. oxysporum* f. sp. *ubense* biotypes, TR4 (VCG 01213/16) is considered to be predominant and most destructive to Cavendish banana in tropical regions and all of the cultivars that are susceptible to other the two races of *F. oxysporum* f. sp. *ubense* (García-Bastidas et al. 2014). The rapid spread of *F. oxysporum* f. sp. *ubense* TR4 into a new Cavendish growing area was facilitated by a limited awareness and lack of quarantine measures along with poor crop protection practices. Dissemination of TR4 has devastated the global banana industry over 30 years, notably in the countries of Asia (Taiwan, Malaysia, Indonesia, India, China, Cambodia, the Philippines, Pakistan, Laos, Vietnam, Myanmar, and Thailand), Africa (Mozambique), the Middle East (Oman, Jordan, Lebanon, Turkey, and Israel), America (Colombia), and Australia (Thangavelu et al. 2020). The senior author recorded the occurrence of *Fusarium* wilt TR4 in Katihar and Purne districts of Bihar, India (Thangavelu et al. 2019). Although the earlier *Fusarium* wilt incidences ranged from 2 to 26.6% (Thangavelu et al. 2019), the recent survey revealed that the incidence was more than 50% in both Bihar and Uttar Pradesh states in India, where Cavendish banana is being cultivated in more than 0.1 million hectares. Because the collected TR4 strain is seriously devastating Cavendish banana in the subtropical regions of India, there could be a chance of variation in virulence-associated genes; specifically, in the *secreted-in-xylem* (SIX) genes (Maldonado Bonilla et al. 2018). In this context, the study has been initiated to understand the genome of a devastating *F. oxysporum* f. sp. *ubense* strain TR4 in detail in comparison with the reference pathogens *F. oxysporum* f. sp. *ycopersici* 4287 (*RFol* 4287;

Funding

Financial support for this study was provided through the extramural project rendered by the ICAR, New Delhi, and the facilities provided by the director.

Keywords

banana wilt disease, *Foc* TR4, secreted in xylem, whole-genome sequencing

[†]Corresponding author: R. Thangavelu; rtbanana@gmail.com

The author(s) declare no conflict of interest.

Accepted for publication 30 July 2020.

Table 1. General genome and predicted proteome statistics for reference and assembled genomes^a

Characteristics	<i>RFol</i> 4287	<i>RFoc</i> TR4	<i>SFoc</i> TR4
Genome size (Mb)	61.5	46.5	47.4
Number of contigs (count)	114	418	4,034
Contig N ₅₀ (Mb)	1.97	1.13	0.03
Maximum contig length	68,54,980	63,07,302	1,90,413
Minimum contig length	900	200	300
Average contig length	6,98,542	63,007	11,746
Median contig length	15,960	579	5,665
GC content	48.4	48.1	49.4
Protein coding genes	27,347	15,519	15,508
Total genes	20,925	16,634	16,129
Gene BUSCO (%)	96	97	98

^a *RFol* 4287 and *RFoc* TR4 are the reference genomes and *SFoc* TR4 is the assembled genome. *RFol* 4287: ASM14995v2; *RFoc* TR4: GCA_000260195.2; and *SFoc* TR4: VYQK00000000 are genome assembly numbers available in the NCBI database.

ASM14995v2) and TR4 (*RFoc* TR4; GCA_000260195.2). Therefore, the present study reports the draft genome of the *F. oxysporum* f. sp. *cupense* TR4 (*SFoc* TR4; VCG 01213/16) and discusses a putative virulence-associated unigene, *SIX*, with reference to *RFoc* TR4.

A detailed description of the origin and isolation of *SFoc* TR4, confirmation using nitrate-nonutilizing mutant-M testers of known vegetative compatibility groups (VCGs), characterization of volatile production, and their pathogenicity were reported earlier (Thangavelu et al. 2019). The genomic DNA was isolated from a single-spore culture of late-log-phase growth (grown in potato dextrose broth medium at 25 ± 2°C without shaking) by the cetyltrimethylammonium bromide method followed by proteinase K and RNase digestion. DNA purity was measured by the absorbance at 260/280 nm ratio using NanoDrop and agarose gel electrophoresis. The genomic library was prepared from a sheared DNA fraction of approximately 300 bp using Illumina paired-end sample preparation kits according to the manufacturer's instructions. Fungal libraries were sequenced using the Illumina NextSeq 500 system (M/s. Genotypic, Bangalore, India) for 150 × 2 cycles, generating approximately 2.6 million paired-ends reads amounting to approximately 16.18 million nucleotides. The short-read sequences of *SFoc* TR4 were assembled with *RFol* 4287 using MaSuRCA v.3.2.4 (Zimin et al. 2013), mapped using BWA v0.7.12 (Li and Durbin 2009), and annotated using Fungal Ensemble gene annotation system (<https://fungi.ensembl.org/index.html>). The draft genome sequence of *SFoc* TR4 of VCG 01213/16 was 47.38 mb, with 49.4% GC content. The assembling of sequences resulted in 4,034 contigs, with a length of 300 to 1,90,413 bp and an average contig size of 11,746 bp (Table 1). Functional annotation of AUGUSTUS (v.3.3) predicted (Stanke et al. 2006) protein-coding genes was performed by combining InterProScan (Jones et al. 2014) and BLASTp searches against the UniProtKB Fungi protein database (The UniProt Consortium 2018). In total, 15,508 (96.15%) proteins were annotated from 16,129 using UniProt database with a cutoff E-value of 10⁻⁵, and the remaining 621 (3.85%) were unannotated or uncharacterized proteins. Of the total annotated proteins, 2,197 (14.17%) were associated with biological processes, 6,385 (41.17%) were cellular processes and signaling, and 6,926 (44.66%) were molecular functions. A further comparison with the reference genome of *RFoc* TR4 (GCA_000260195.2) revealed that there were 195 contigs of unaligned sequences, consisting of approximately 3.1 mb in the studied genome. These are considered to be unique regions present in the genome of the study, which differentiates them from the other *F. oxysporum* f. sp. *cupense* TR4 strains based on genome organization. To visualize the locally collinear blocks, MAUVE v20150226 was used with a weight close to 5,000 (Darling et al. 2004) and the result showed that there were 3,202 homologous regions which are conserved in the *SFoc* TR4 of India when compared with the *RFol* 4287 and *RFoc* TR4 genomes.

Plant Host Interaction search (<http://www.phi-base.org>) showed that there were 1,042 genes present in the *SFoc* TR4 genome, of which 19 are unique when compared with *RFoc* TR4. To assess the virulence or genetic relatedness of *SFoc* TR4, the presence of 14 *SIX* gene clusters (*SIX1* to *SIX14*) were examined using BLASTx analyses. Among the *SIX* genes, *SIX1*, *SIX2*, *SIX6*, *SIX8*, and *SIX9* have been found to present in the *SFoc* TR4 genome. Based on the report of Czislowski et al. (2018), STR4 carries *SIX1*, *SIX2*, *SIX7*,

SIX8, and *SIX9*, while TR4 carries *SIX1*, *SIX2*, *SIX6*, *SIX8*, *SIX9*, and *SIX13*. In contrast to the above, the present study observed the absence of *SIX13* in *SFoc* TR4 of India. Interestingly, it was also noticed that two homologs of the *SIX1* effector gene in *SFoc* TR4 (in scf7180000021486: nucleotides 1 to 840, g122; and in scf7180000021484: nucleotides 436 to 1,251, g14858) without any sequence variations, which corroborates the report of Czisłowski et al. (2018), especially in VCG 01213/16. In general, *SIX8a* is present in all R4 isolates, whereas *SIX8b* is present in all STR4 isolates and not detected in *F. oxysporum* f. sp. *cubense* TR4 isolates thus far (Fraser-Smith et al. 2014). However, copies of two *SIX8* homologs, *SIX8a* and *SIX8b*, were observed in the genome of *SFoc* TR4 (in scf7180000019319: nucleotides 224 to 3,003, g13480) without sequence variation, which is a major difference noticed in the present study. Because of the *SIX8b* effector gene, *F. oxysporum* f. sp. *cubense* TR4 of India can infect the Cavendish banana grown in the subtropical region as well, and this is confirmed by the fact that TR4 was isolated from the Katihar district of Bihar, India, which comes under subtropical climate in India. A detailed comparative report of the genomes of the *F. oxysporum* f. sp. *cubense* races is being prepared.

Data availability. All of the data generated or analyzed during this study are included in this published article.

Nucleotide sequence accession numbers. The first version of the assembled genome sequences of *SFoc* TR4 (VCG 01213/16) infesting Cavendish banana cultivars have been deposited at DNA Data Bank of Japan/European Nucleotide Archive/GenBank under accession number VYQK00000000.

Acknowledgments

We thank S. Backiyarani, principle scientist, and Sasikala, research scholar, ICAR-NRC for Banana for her suggestions on the manuscript.

Literature Cited

- Czisłowski, E., Fraser-Smith, S., Zander, M., O'Neill, W. T., Meldrum, R. A., Tran-Nguyen, L. T. T., Batley, J., and Aitken, E. A. B. 2018. Investigation of the diversity of effector genes in the banana pathogen, *Fusarium oxysporum* f. sp. *cubense*, reveals evidence of horizontal gene transfer. *Mol. Plant Pathol.* 19: 1155-1171.
- Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. 2004. Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14:1394-1403.
- Fraser-Smith, S., Czisłowski, E., Meldrum, R., Zander, M., O'Neill, W., Balali, G., and Aitken, E. A. B. 2014. Sequence variation in the putative effector gene *SIX8* facilitates molecular differentiation of *Fusarium oxysporum* f. sp. *cubense*. *Plant Pathol.* 63:1044-1052.
- García-Bastidas, F., Ordóñez, N., Konkol, J., Al-Qasim, M., Naser, Z., Abdelwali, M., Salem, N., Waalwijk, C., Ploetz, R. C., and Kema, G. H. J. 2014. First report of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 associated with panama disease of banana outside Southeast Asia. *Plant Dis.* 98:694.
- Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., Pesseat, S., Quinn, A. F., Sangrador-Vegas, A., Scheremetjew, M., Yong, S.-Y., Lopez, R., and Hunter, S. 2014. InterProScan 5: Genome-scale protein function classification. *Bioinformatics* 30:1236-1240. doi:
- Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25:1754-1760.
- Maldonado Bonilla, L. D., Villarruel Ordaz, J. L., Calderón Oropeza, M. A., and Sánchez-Espinosa, A. C. 2018. Secreted in xylem (Six) genes in *Fusarium oxysporum* f. sp. *cubense* and their potential acquisition by horizontal transfer. *Adv. Biotechnol. Microbiol.* 10:555779. <https://juniperpublishers.com/aibm/AIBM.MS.ID.555779.php>
- Ordoñez, N., García-Bastidas, F., Laghari, H. B., Akkary, M. Y., Harfouche, E. N., al Awar, B. N., and Kema, G. H. J. 2016. First report of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 causing panama disease in Cavendish bananas in Pakistan and Lebanon. *Plant Dis.* 100:209.
- Ploetz, R. C. 2015. Management of Fusarium wilt of banana: A review with special reference to tropical race 4. *Crop Prot.* 73:7-15.
- Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. 2006. AUGUSTUS: Ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34: W435-W439.
- Thangavelu, R., Loganathan, M., Arthee, R., Prabakaran, M., and Uma, S. 2020. Fusarium wilt: A threat to banana cultivation and its management. *CAB Review.* CAB International, Wallingford, U.K. <https://www.cabi.org/cabreviews/review/20203073684>
- Thangavelu, R., Mostert, D., Gopi, M., Devi, P. G., Padmanaban, B., Molina, A. B., and Viljoen, A. 2019. First detection of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 (TR4) on Cavendish banana in India. *Eur. J. Plant Pathol.* 154:777-786.
- The UniProt Consortium. 2018. UniProt: The universal protein knowledgebase. *Nucleic Acids Res.* 46:2699.
- Zimin, A. V., Marçais, G., Puiu, D., Roberts, M., Salzberg, S. L., and Yorke, J. A. 2013. The MaSuRCA genome assembler. *Bioinformatics* 29:2669-2677.