Effector profiles distinguish formae speciales of Fusarium oxysporum

van Dam, P.; Fokkens, L.; Schmidt, S.M.; Linmans, J.J.J.; Kistler, H.C.; Ma, L.-J.; Rep, M.

DOI
10.1111/1462-2920.13445

Publication date
2016

Document Version
Other version

Published in
Environmental Microbiology

License
Article 25fa Dutch Copyright Act (https://www.openaccess.nl/en/in-the-netherlands/you-share-we-take-care)

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 426, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Supporting Information

Article title: **Effector profiles distinguish formae speciales of Fusarium oxysporum**
Authors: Peter van Dam¹, Like Fokkens¹, Sarah M. Schmidt¹, Jasper H.J. Linmans¹, H. Corby Kistler², Li-Jun Ma³ and Martijn Rep¹

The following Supporting Information is available for this article:

Fig S1. Phylogenetic relationships of strains considered for whole genome sequencing inferred from the EF-1α sequence.

Fig S2. Repeat-density, RNA and DNA transposon abundance is largely similar both amongst the de novo assemblies as well as compared to the reference assemblies.

Fig S3. Method used for identifying candidate effector genes in a whole genome shotgun assembly.

Fig S4. Clustering of 2.5kb regions downstream of a mimp inverted repeat shows a more fragmented clustering compared to clustering based on effector candidates.

Fig S5. Identical SIX gene sequence types are found in strains belonging to polyphyletic formae speciales, pointing to a combination of vertical and horizontal inheritance of these genes.

Fig S6. RNA sequencing analysis of candidate effectors shows that many candidates show strong expression in planta (i.p.) and little or no expression during in vitro (i.v.) conditions.

Table S1. Statistics of de novo genome assemblies.

Table S2. Effector candidates that were identified in the combined set of genome assemblies.

Table S3. Comparison of candidate genes identified in three genome assemblies of the reference strain Fol4287 shows that each individual assembly yields a few unique candidate genes.

Table S4. Effector identification based on mimp-proximity is applicable only to *F. oxysporum*.

Table S5. RNA sequencing analysis of candidate effectors shows that many candidates show strong expression in planta and little or no expression during in vitro conditions.

Methods S1. Bioassay scoring system

Methods S2. DNA isolation

Methods S3. Read preparation and genome assembly
Fig S1. Phylogenetic relationships of strains considered for whole genome sequencing inferred from the EF-1α sequence.

Part of the EF-1α gene was PCR-amplified using primers FP889 (tcgtcgtcatcggccacgtc) and FP1614 (ggaagtaccagtgatcatgtt) (Van Der Does et al., 2008) and sequenced using an ABI3730XL DNA Analyzer (Macrogen, the Netherlands). A MUSCLE nucleotide sequence alignment was made using 569 nt after trimming. Phylogeny was inferred using PhyML with 100 bootstrap iterations and plotted with F. verticillioides as an outgroup. Branches with most parsimonious bootstrap partitions below 50% were collapsed; values ≥ 50% and < 100% are indicated in red; 100% are not indicated. A coloured circle (wilting), square (root and shoot rot) or triangle (non-pathogenic ["_NP"] / other) representing the strain’s forma specialis was plotted on the leaves of the dendrogram. Strains that are coloured grey were not selected for whole genome sequencing, those that are black and underlined were and those with an asterix (*) have a publically available genome assembly.
Fig S2. Repeat-density, RNA and DNA transposon abundance is largely similar both amongst the de novo assemblies as well as compared to the reference assemblies.

Simple repeats and low complexity comprise c. 0.61% ± 0.03% S.D. of the sequence of all of the Fusarium genome assemblies evaluated, including *F. verticillioides* 7600 (0.70%) and *F. graminearum* PH-1 (0.69%), while both RNA transposons (1.08% ± 0.35% S.D.) and DNA transposons (0.91% ±0.20% S.D.) are much more prevalent in *F. oxysporum* than the other two species.
Fig S3. Method used for identifying candidate effector genes in a whole genome shotgun assembly.

(a) Putative effectors were identified by looking for uninterrupted ORFs in a region of i) 2500 bp or ii) 5000 bp downstream of a miniature Impala (mimp) terminal inverted repeat (TIR). Maximum distance between the mimp terminal inverted repeat (TIR) and the start codon (ATG) of the putative ORF was set in both cases to 2000 bp. (b) Workflow diagram of mimp-associated putative effector discovery pipeline from de novo genome assembly to identification of effector candidates, clustering them and identifying presence/absence patterns in each genome. In blue, the number of records found for the assessed Fo genomes is represented.
Fig S4. Clustering of 2.5kb regions downstream of a mimp inverted repeat shows a more fragmented clustering compared to clustering based on effector candidates.

2.5 kb windows were extracted downstream of each occurrence of a mimp inverted repeats (IRs, ‘TT[TA]TTGCNNCCACTGNN’). Terminal ambiguous bases (‘N’ s) were trimmed from the sequence and redundancy was reduced by self-BLASTN. Presence of 360 regions was detected using BLASTN with the same thresholds as candidate clustering.

Traces of the core genome are more clearly visible, for example in the splitting of Fon into two groups, the position of Foc011 and Foc013 (clade 1 isolates, Fig. 2) close to Focub and Fom001 (NRRL26406) in between Fol isolates with a highly similar core genome.
Fig S5. Identical SIX gene sequence types are found in strains belonging to polyphyletic formae speciales, pointing to a combination of vertical and horizontal inheritance of these genes. A MUSCLE alignment was made with the nucleotide sequence of (a) SIX2 (705 nt), (b) SIX4 (793 nt), (c) SIX8 (525 nt), (d) SIX9 (357 nt), (e) SIX11 (335 nt), (f) SIX14 (317 nt). Phylogeny was inferred using PhyML with 100 bootstrap iterations and plotted with mid-point rooting. Branches with most parsimonious bootstrap partitions below 50% were collapsed; values ≥ 50% and < 100% are indicated in red; 100% are not indicated. A coloured circle (wilting), square (root and shoot rot) or triangle (non-pathogenic / other) representing the isolate’s forma specialis was plotted on the leaves of the dendrogram.
Fig S6. RNA sequencing analysis of candidate effectors shows that many candidates show strong expression in planta (i.p.) and little or no expression during in vitro (i.v.) conditions.

Table S1 - Table S4: see “vanDam_etal_SI_tables.xlsx”

Table S5: See “vanDam_etal_SI_table_S5.xlsx”

Methods S1. Bioassay scoring system
In the case of fusarium wilt caused by Fol, Fom, Fon and Foc, the following scoring system was used: 0, no symptoms; 1, plant stunted, slightly swollen and/or bent hypocotyl; 2, one or two brown vascular bundles at height of cotyledons; 3, at least two brown vascular bundles and growth distortion and wilting / clear root rot symptoms; 4, plant either dead or very small and wilted). Root and shoot rot caused by Forc infection was scored as follows: 0, no symptoms; 1, slight root rot symptoms, only at tip of main root; 2, root rot symptoms and stem lesions visible aboveground; 3, very clear root rot symptoms of the entire root system, often with a large lesion extending above the cotyledons; 4, plant either dead or very small and wilted.
Methods S2. DNA isolation

*F. oxysporum* genomic DNA was isolated from freeze-dried mycelium that was harvested from five-day-old NO₃-medium (0.17% yeast nitrogen base, 3% sucrose, 100 mM KNO3) cultures. In the case of all Foc, Forc, Fom, Fon strains as well as Fol007 and Fol4287, DNA was isolated as described by Michielse et al. (Michielse et al., 2009). Briefly, mycelial powder was suspended in 2 ml DNA-extraction buffer (0.2 M TrisHCL pH 8.5, 0.25 M NaCl, 0.05 M EDTA pH 8.0, 48 mg/ml sodium 4-aminosalicylate dihydrate (PAS, Sigma-Aldrich), 8 mg/ml Trisopropylphthalenesulfonic acid sodium (TIPS, Sigma-Aldrich). 2 ml buffer-saturated phenol:chloroform:isoamylalcohol (25:24:1) was added, mixed and centrifuged for 30 minutes at 3,500 rpm and 4°C. DNA was precipitated from the aqueous phase with 0.7 volume isopropanol and 0.1 volume 4 M NaCl and centrifugation for 30 minutes at 3,500 rpm and 4°C. The pellet was resuspended in 500 µl TE buffer (10 mM TrisHCL pH7.5, 1 mM EDTA pH 8.0) and three consecutive rounds of phenol:chloroform:isoamylalcohol isolation were performed. DNA was precipitated from the aqueous phase with 2.5 volume 96% ethanol and 0.1 volume 4 M NaCl. The pellet was resuspended in 100 µl Milli-Q and treated with RNase A (Roche) and proteinase K (Fermentas, pcr grade). 500 µl Milli-Q was added and one more round of phenol:chloroform:isoamylalcohol was executed. DNA was precipitated with 2.5 volume 96% ethanol and 0.1 volume 4 M NaCl, washed with ethanol, air-dried and finally dissolved in 50 µl TE buffer.

DNA isolation of the remaining Fol strains was performed using the Omniprep™ for Fungus (G-Biosciences) kit, followed by an additional purification step by phenol:chloroform extraction and precipitation with 2.5 volume 96% ethanol and 0.1 volume 4 M NaAc.

Methods S3. Read preparation and genome assembly

Sequencing reads were trimmed for quality and adapter sequences with FastqMcf v1.04.676 (http://code.google.com/p/ea-utils, quality threshold=20). A preliminary assembly was made and checked for presence of contaminant contigs with blobology (Kumar et al., 2013) (manual parameters were set to extract these contigs based on best BLAST hit, GC-content and coverage). In the case of Foc030, Fon005, Fon013, Fon015, Fon019 and Fon021, bacterial contamination with *Achromobacter xylosoxidans* was determined. Read pairs that could not be mapped with bowtie2 v2.2.5 against contaminated contigs were de novo assembled, mapped and checked for remaining contaminations with blobology three times, followed by a de novo assembly using CLC-workbench 8.0. Default settings were used, except “minimum contig length=500”. Then, finally, any remaining contaminated contigs were removed manually from the CLC assembly. All other genomes were assembled directly in CLC workbench as described. The completeness of the assemblies was assessed with CEGMA v 2.5 (Parra et al., 2009). We determined the repeat content of all genomes using RepeatMasker (Smit et al., 1996) with ‘-species ascomycota’ and repeat libraries from RepBase (version 20140131) (Jurka et al., 2005).

SI References


