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

Supporting Information

Article title: **Effector profiles distinguish *formae speciales* of *Fusarium oxysporum***
Authors: Peter van Dam[1], Like Fokkens[1], Sarah M. Schmidt[1], Jasper H.J. Linmans[1], H. Corby Kistler[2], Li-Jun Ma[3] and Martijn Rep[1]

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.

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.

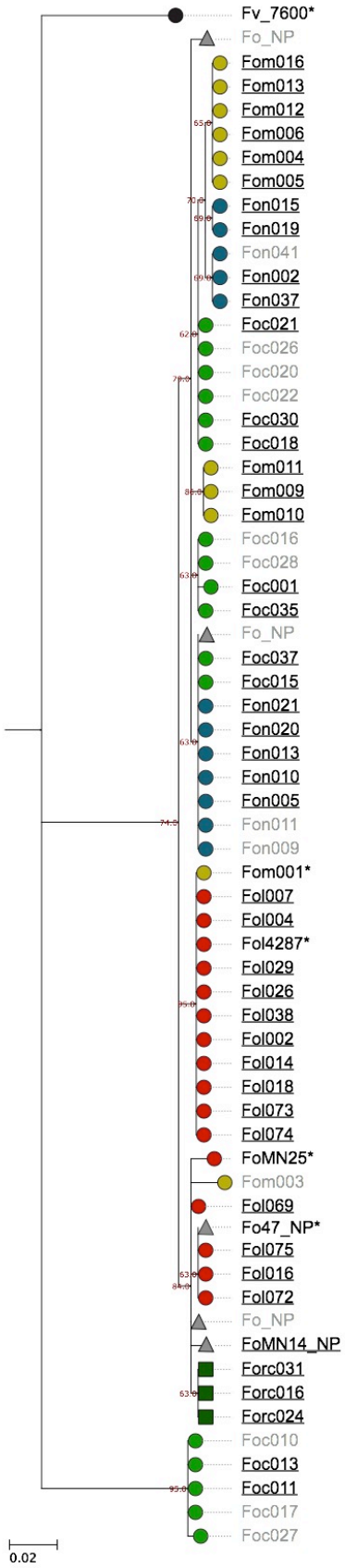


Fig S2. Repeat-density, RNA and DNA transposon abundance is largely similar both amongst the *de novo* assemblies as well as compared to the reference assemblies.
Simple repeats and low complexity comprise *c.* 0.61% ± 0.03% S.D. of the sequence of all of the *Fusarium* genome assemblies evaluated, including *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.

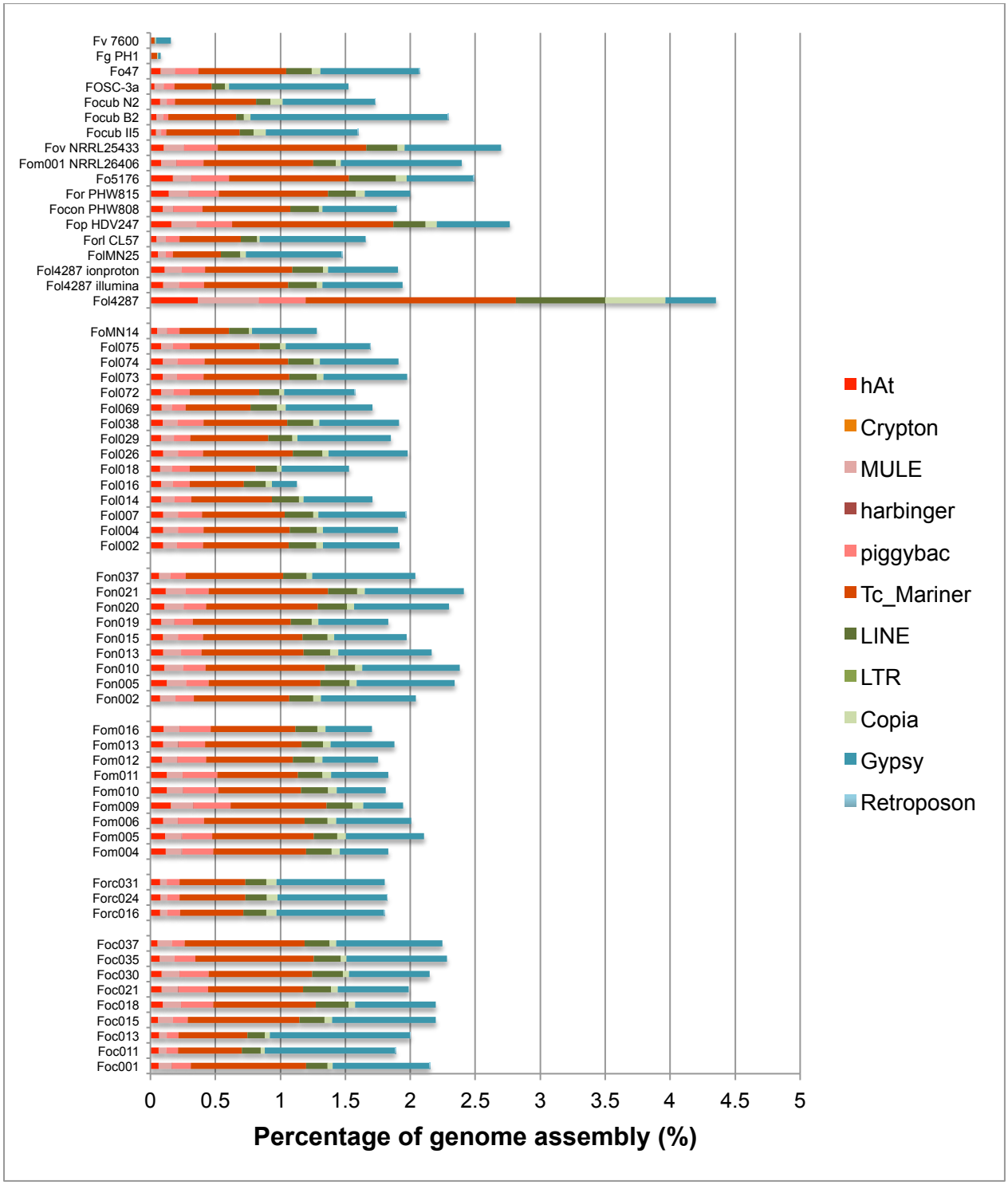


Fig S3. Method used for identifying candidate effector genes in a whole genome shotgun assembly.
(a) Putative effectors were identified by looking for uninterrupted ORFs in a region of i) 2500 bp or ii) 5000 bp downstream of a miniature Impala (mimp) terminal inverted repeat (TIR). Maximum distance between 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 discory 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.

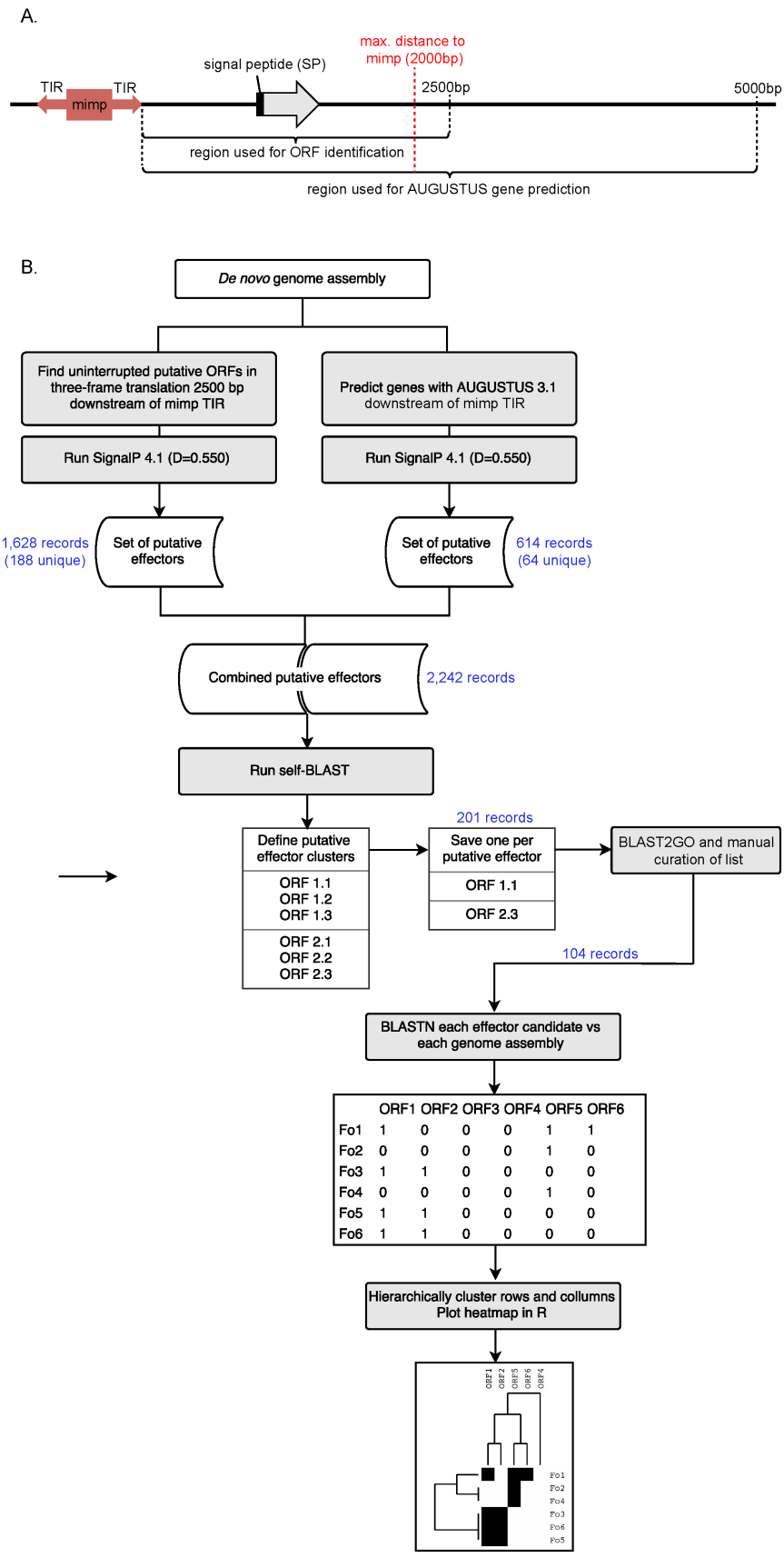


Fig S4. Clustering of 2.5kb regions downstream of a mimp inverted repeat shows a more fragmented clustering compared to clustering based on effector candidates.
2.5 kb windows were extracted downstream of each occurrence of a mimp inverted repeats (IRs, 'TT[TA]TTGCNNCCCACTGNN'). Terminal ambiguous bases ('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.

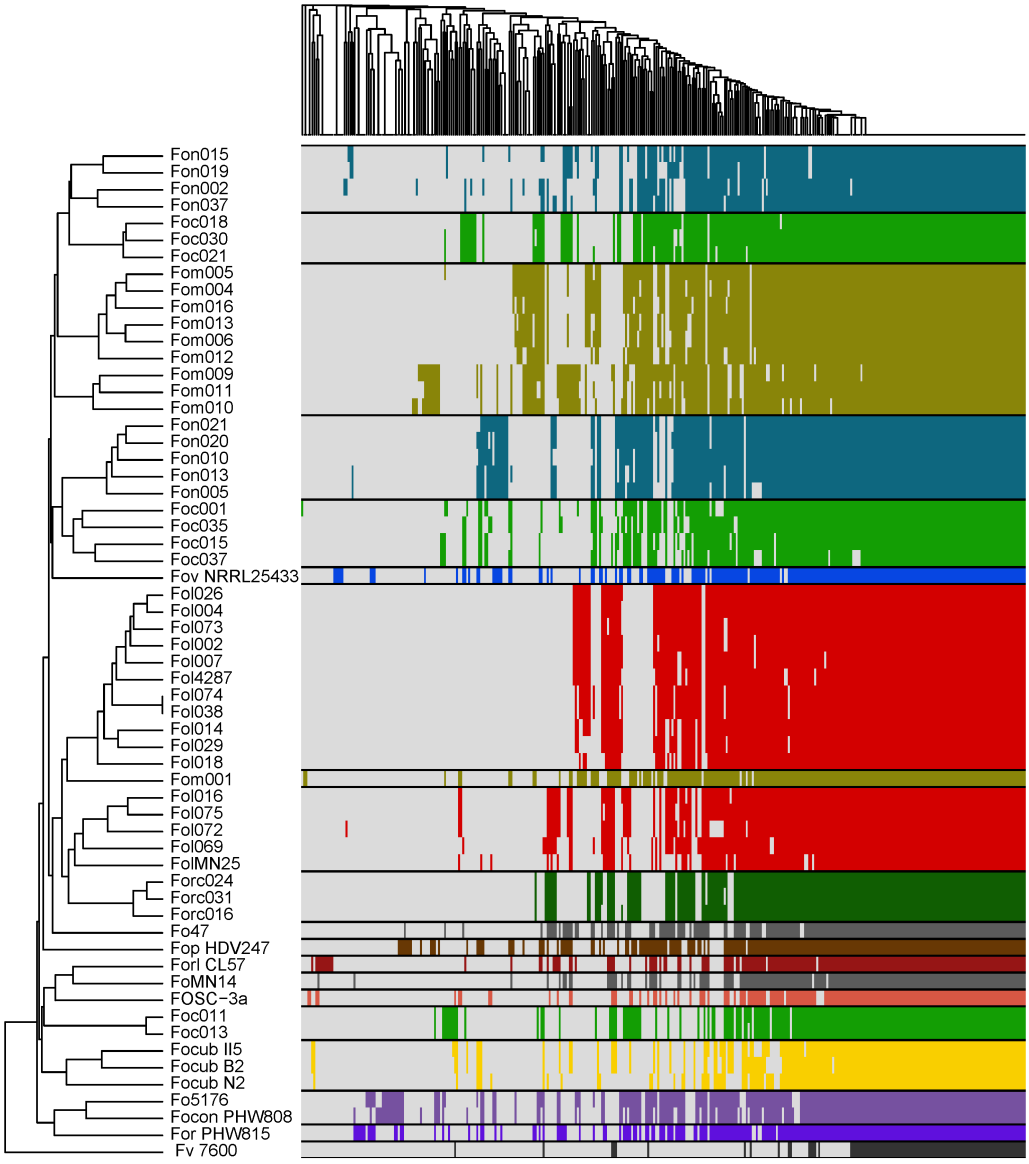
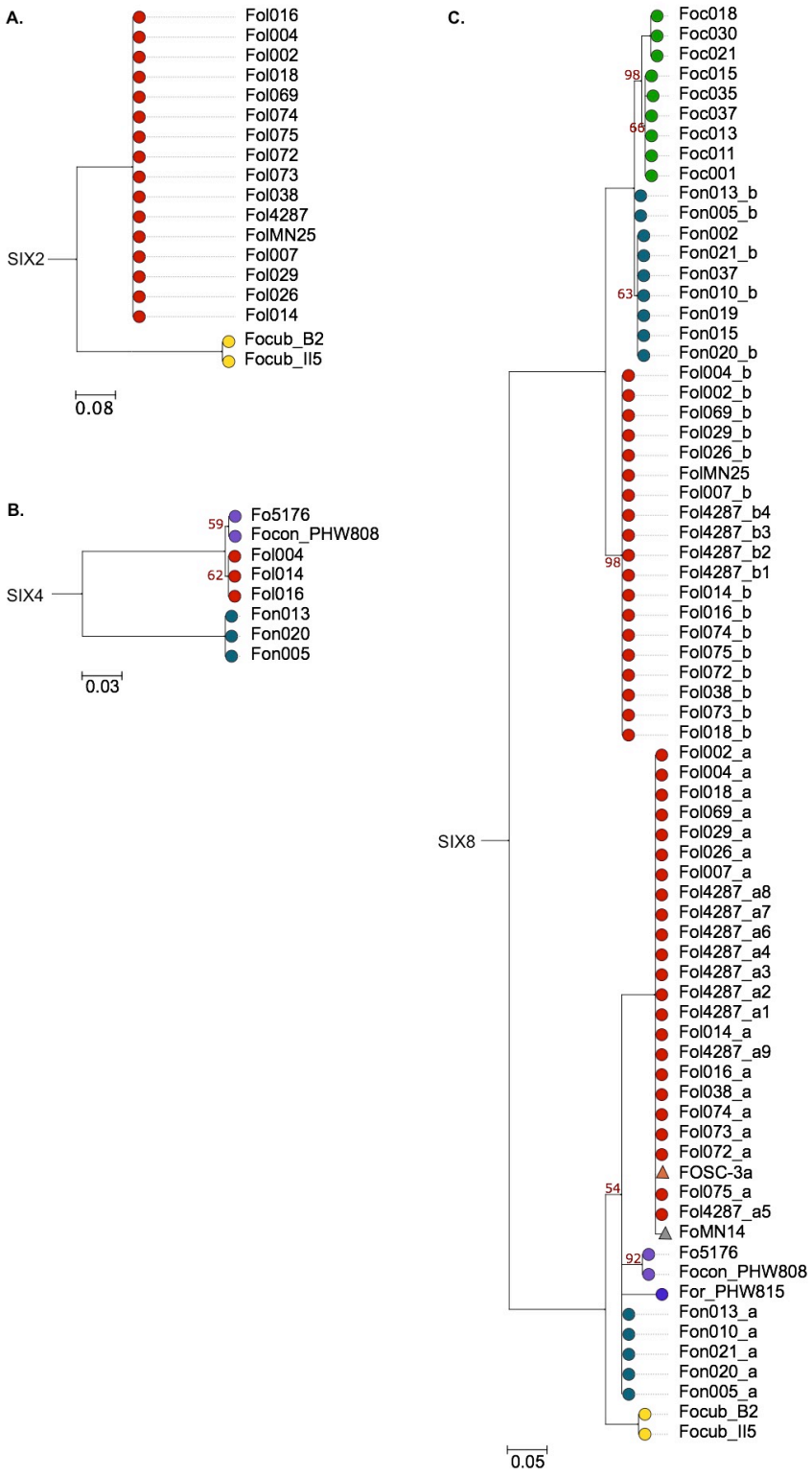


Fig S5. Identical *SIX* gene sequence types are found in strains belonging to polyphyletic *formae speciales*, pointing to a combination of vertical and horizontal inheritance of these genes.
A MUSCLE alignment was made with the nucleotide sequence of (a) *SIX2* (705 nt), (b) *SIX4* (793 nt), (c) *SIX8* (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.



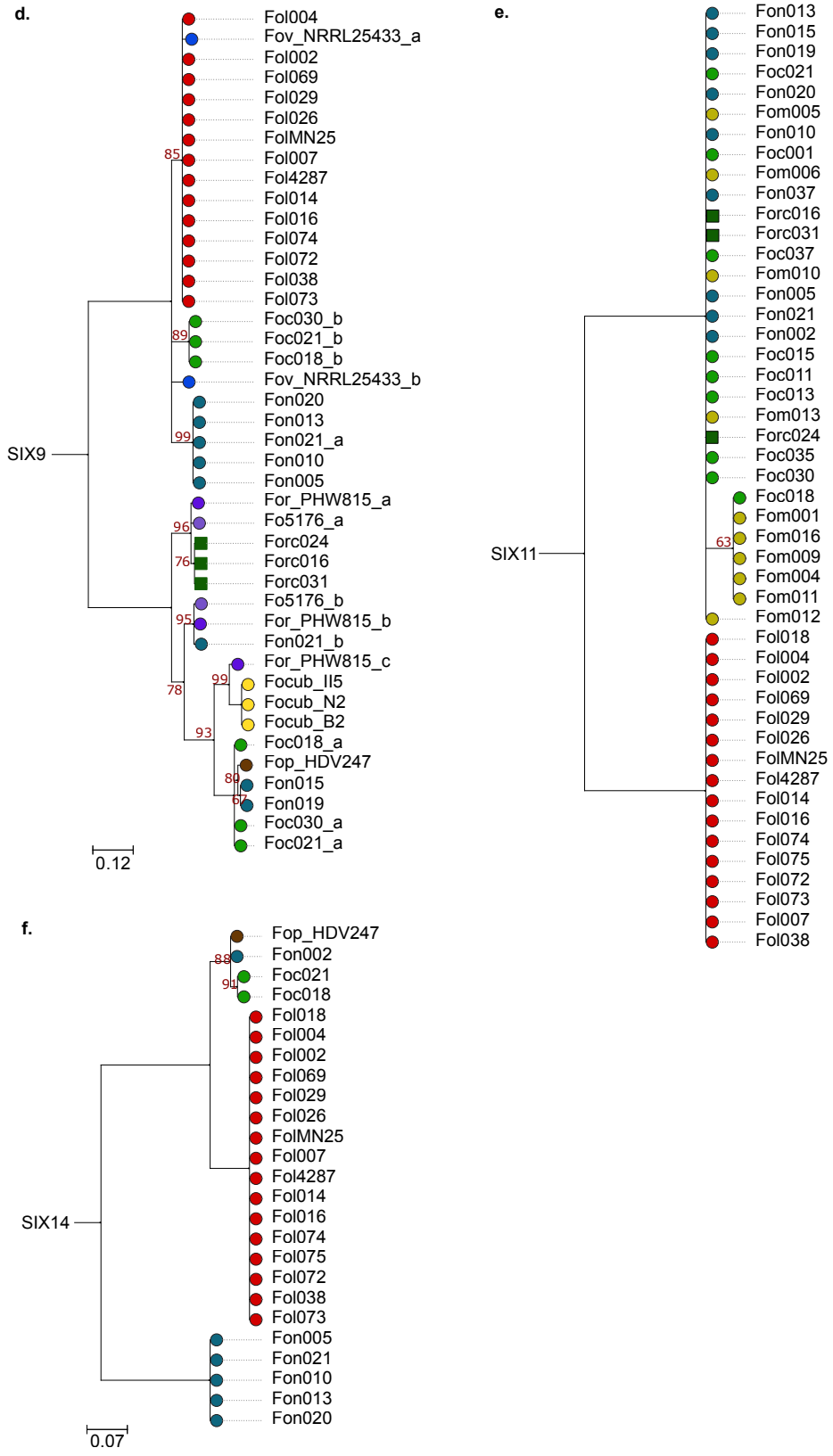

d.
SIX9
Fol004
Fov_NRRL25433_a
Fol002
Fol069
Fol029
Fol026
FolMN25
Fol007
Fol4287
Fol014
Fol016
Fol074
Fol072
Fol038
Fol073
Foc030_b
Foc021_b
Foc018_b
Fov_NRRL25433_b
Fon020
Fon013
Fon021_a
Fon010
Fon005
For_PHW815_a
Fo5176_a
Forc024
Forc016
Forc031
Fo5176_b
For_PHW815_b
Fon021_b
For_PHW815_c
Focub_II5
Focub_N2
Focub_B2
Foc018_a
Fop_HDV247
Fon015
Fon019
Foc030_a
Foc021_a
85
89
99
96
76
95
78
99
93
80
67
0.12
e.
SIX11
Fon013
Fon015
Fon019
Foc021
Fon020
Fom005
Fon010
Foc001
Fom006
Fon037
Forc016
Forc031
Foc037
Fom010
Fon005
Fon021
Fon002
Foc015
Foc011
Foc013
Fom013
Forc024
Foc035
Foc030
Foc018
Fom001
Fom016
Fom009
Fom004
Fom011
Fom012
Fol018
Fol004
Fol002
Fol069
Fol029
Fol026
FolMN25
Fol4287
Fol014
Fol016
Fol074
Fol075
Fol072
Fol073
Fol007
Fol038
63
f.
SIX14
Fop_HDV247
Fon002
Foc021
Foc018
Fol018
Fol004
Fol002
Fol069
Fol029
Fol026
FolMN25
Fol007
Fol4287
Fol014
Fol016
Fol074
Fol075
Fol072
Fol038
Fol073
Fon005
Fon021
Fon010
Fon013
Fon020
88
94
0.07

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.

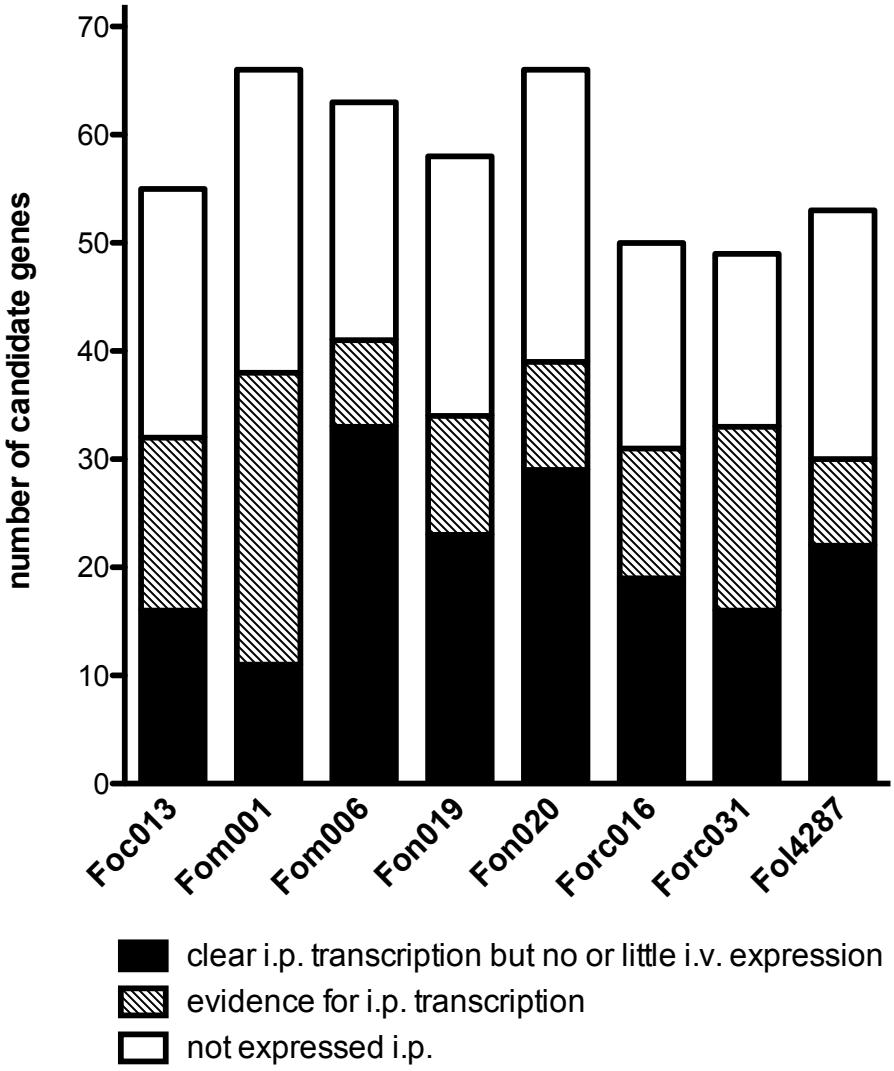


Table S1 - Table S4: see "vanDam_etal_SI_tables.xlsx"

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 NO_3-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 Triisopropylnaphthalenesulfonic 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, airdried 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

Van Der Does HC, Lievens B, Claes L, Houterman PM, Cornelissen BJC, Rep M. **2008**. The presence of a virulence locus discriminates *Fusarium oxysporum* isolates causing tomato wilt from other isolates. *Environmental Microbiology* **10**: 1475–1485.

Jurka J, Kapitonov V V., Pavlicek a., Klonowski P, Kohany O, Walichiewicz J. **2005**. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenetic and Genome Research* **110**: 462–467.

Kumar S, Jones M, Koutsovoulos G, Clarke M, Blaxter M. **2013**. Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Frontiers in genetics* **4**: 237.

Michielse CB, van Wijk R, Reijnen L, Cornelissen BJC, Rep M. **2009**. Insight into the molecular requirements for pathogenicity of *Fusarium oxysporum* f. sp. *lycopersici* through large-scale insertional mutagenesis. *Genome biology* **10**: R4.

Parra G, Bradnam K, Ning Z, Keane T, Korf I. **2009**. Assessing the gene space in draft genomes. *Nucleic Acids Research* **37**: 289–297.

Smit A, Hubley R, Green P. **1996**. RepeatMasker Open-3.0. *RepeatMasker Open-3.0*.