Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium


Published in:
Nature

DOI:
10.1038/nature08850

Link to publication

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: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)

Download date: 12 Sep 2017
Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*

Li-Jun Ma\(^1\), H. Charlotte van der Does\(^2\), Katherine A. Borkovich\(^3\), Jeffrey J. Coleman\(^4\), Marie-Josée Daboussi\(^5\), Antonio Di Pietro\(^6\), Marie Dufresne\(^5\), Michael Freitag\(^7\), Manfred Grabherr\(^1\), Bernard Henrisat\(^8\), Petra M. Houterman\(^9\), Seogchan Kang\(^2\), Won-Bo Shim\(^10\), Charles Woloshuk\(^11\), Xiaohui Xie\(^12\), Jin-Rong Xu\(^11\), John Antoniw\(^13\), Scott E. Baker\(^14\), Burton H. Bluhm\(^11\), Andrew Breakspear\(^15\), Daren W. Brown\(^16\), Robert A. E. Butchko\(^16\), Sinead Chapman\(^1\), Richard Coulson\(^17\), Pedro M. Coutinho\(^8\), Etienne G. J. Danchin\(^8\), Andrew Diener\(^18\), Liane R. Gale\(^19\), Donald M. Gardiner\(^19\), Stephen Goff\(^20\), Kim E. Hammond-Kosack\(^15\), Karen Hilburn\(^15\), Aurélie Hua-Van\(^5\), Wilfried Jonkers\(^2\), Kemal Kazan\(^19\), Chinnappa D. Kodira\(^21\), Michael Koehrens\(^1\), Lokesh Kumar\(^1\), Yong-Hwan Lee\(^22\), Liande Li\(^3\), John M. Manners\(^19\), Diego Miranda-Saavedra\(^22\), Mala Mukherjee\(^10\), Gyungsoon Park\(^3\), Jongsun Park\(^2\), Sook-Young Park\(^9\), Robert H. Proctor\(^16\), Aviv Regev\(^1\), M. Carmen Ruiz-Roldan\(^6\), Divya Sain\(^3\), Sharadha Sakthikumar\(^1\), Sean Sykes\(^4\), David C. Schwartz\(^23\), B. Gillian Turgeon\(^24\), Ilan Wapinski\(^1\), Olen Yoder\(^2\), Sarah Young\(^1\), Qiandong Zeng\(^7\), Shiguo Zhou\(^9\), James Galagan\(^1\), Christina A. Cuomo\(^1\), H. Corby Kistler\(^15\) & Martijn Rep\(^2\)

*Fusarium* species are among the most important phytopathogenic and toxigenic fungi. To understand the molecular underpinnings of pathogenicity in the genus *Fusarium*, we compared the genomes of three phenotypically diverse species: *Fusarium graminearum*, *Fusarium verticillioides* and *Fusarium oxysporum* f. sp. *lycopersici*. Our analysis revealed lineage-specific (LS) genomic regions in *F. oxysporum* that include four entire chromosomes and account for more than one-quarter of the genome. LS regions are rich in transposons and genes with distinct evolutionary profiles but related to pathogenicity, indicative of horizontal acquisition. Experimentally, we demonstrate the transfer of two LS chromosomes between strains of *F. oxysporum*, converting a non-pathogenic strain into a pathogen. Transfer of LS chromosomes between otherwise genetically isolated strains explains the polyphyletic origin of host specificity and the emergence of new pathogenic lineages in *F. oxysporum*. These findings put the evolution of fungal pathogenicity into a new perspective.

*Fusarium* species are among the most diverse and widely dispersed plant-pathogenic fungi, causing economically important blights, root rots or wilts. Some species, such as *F. graminearum* (Fg) and *F. verticillioides* (Fv), have a narrow host range, infecting predominantly the cereals (Fig. 1a). By contrast, *F. oxysporum* (Fo), has a remarkably broad host range, infecting both monocotyledons and dicotyledonous plants and is an emerging pathogen of immuno-compromised humans and other mammals. Aside from their differences in host adaptation and specificity, *Fusarium* species also vary in reproductive strategy. Some, such as Fo, are asexual, whereas others are both asexual and sexual with either self-fertility (homothallism) or obligate out-crossing (heterothallism) (Fig. 1b).

Previously, the genome of the cereal pathogen Fg was sequenced and shown to encode a larger number of proteins in pathogenicity related protein families compared to non-pathogenic fungi, including predicted transcription factors, hydrolytic enzymes, and transmembrane transporters. We sequenced two additional *Fusarium* species, Fv, a maize pathogen that produces fumonisin mycotoxins that can contaminate grain, and *F. oxysporum* f. sp. *lycopersici* (Fol), a tomato pathogen. Here we present the comparative analysis of the genomes of these three species.

**Results**

Genome organization and gene clusters. We sequenced Fv strain 7600 and Fol strain 4287 (Methods, Supplementary Table 1) using a whole-genome shotgun approach and assembled the sequence using Arachne (Table 1, ref. 6). Chromosome level ordering of the scaffolds was achieved by anchoring the assemblies either to a genetic map for Fv (ref. 7), or an optical map for Fol (Supplementary Information A and Supplementary Table 2). We predicted Fol and Fv genes and re-annotated a new assembly of the Fg genome using a combination of manual and automated annotation (Supplementary Information B).

---

1The Broad Institute, Cambridge, Massachusetts 02141, USA. 2University of Amsterdam, Amsterdam 1098XH, The Netherlands. 3University of California Riverside, California 92521, USA. 4University of Arizona, Tucson, Arizona 85721, USA. 5Université Paris-Sud, 91405 Paris, France. 6Universidad de Cordoba, Cordoba 14071, Spain. 7Oregon State University, Corvallis, Oregon 97331, USA. 8CNRS, Universités Aix-Marseille, 13628 Aix-en-Provence, France. 9Penn State University, University Park, Pennsylvania 16802, USA. 10Texas A&M University, College Station, Texas 77843, USA. 11Purdue University, West Lafayette, Indiana 47907, USA. 12University of California, Irvine, California 92697, USA. 13Centre for Sustainable Pest and Disease Management, Rothamsted Research, Harpenden AL5 2JQ, UK. 14Pacific Northwest National Laboratory, Richland, Washington 99352, USA. 15USDA ARS, University of Minnesota, St. Paul, Minnesota 55108, USA. 16USDA-ARS-NCUA, Peoria, Illinois 61604, USA. 17European Bioinformatics Institute, Cambridge CB10 1SD, UK. 18University of California, Los Angeles, California 90095, USA. 19CSIRO Plant Industry, Queensland BioScience Precinct, St Lucia, Brisbane, Queensland, 4067 Australia. 20BRIOD Institute, University of Arizona, Tucson, Arizona 85721, USA. 21Seoul National University, Seoul 151-742, Korea. 22Cambridge Institute for Medical Research, Cambridge CB2 0XY, UK. 23University of Wisconsin-Madison, Madison, Wisconsin 53706 USA. 24Cornell University, Ithaca, New York 14853, USA. 25Camino Del Roca, Ramona, California 92065, USA. 26Present addresses: 454 Life Sciences, Branford, Connecticut 06405, USA (C.D.K.); University of Texas Southwestern Medical Center at Dallas, Dallas, Texas 75390, USA (L.L.); INRA, Institut National de la Recherche Agronomique, 06903 Sophia-Antipolis, France (E.G.J.D.); Seoul National University, Seoul 151-742, Korea (S.-Y.P.).

*These authors contributed equally to this work.*
The Fol genome (60 megabases) is about 44% larger than that of its most closely related species, Fv (42 Mb), and 65% larger than that of Fg (36 Mb), resulting in a greater number of protein-encoding genes in Fol (Table 1).

The relatedness of the three Fusarium genomes enabled the generation of large-scale unambiguous alignments (Supplementary Figs 1–3) and the determination of orthologous gene sets with high confidence (Methods, Supplementary Information C). On average, Fol and Fv orthologues display 91% nucleotide sequence identity, and both have 85% identity with Fg counterparts (Supplementary Fig. 4). Over 9,000 conserved syntenic orthologues were identified among the three genomes. Compared to other ascomycete genomes, these three-species orthologues are enriched for predicted transcription factors (P = 2.6 × 10⁻⁸), lytic enzymes (P = 0.001), and transmembrane transporters (P = 7 × 10⁻⁹) (Supplementary Information C and Supplementary Tables 3–8), in agreement with results reported for the Fg genome.

Fusarium species produce diverse secondary metabolites, including mycotoxins that exhibit toxicity to humans and other mammals. In the three genomes, we identified a total of 46 secondary metabolite biosynthesis (SMB) gene clusters. Microarray analyses confirmed the co-expression of genes in 14 of 18 Fg and 10 of 16 Fv SMB gene clusters. Ten out of the 14 Fg and eight out of the 10 Fv co-expressed SMB gene clusters are novel (Supplementary Information D, Supplementary Fig. 5 and Supplementary Table 9, and online materials), emphasizing the potential impact of uncharacterized secondary metabolites on fungal biology.

**Lineage-specific chromosomes and pathogenicity.** The genome assembly of Fol has 15 chromosomes, the Fv assembly 11 and the Fg assembly only four (Table 1). The smaller number of chromosomes in Fg is the result of chromosome fusion relative to Fv and Fol, and fusion sites in Fg match previously described high diversity regions (Supplementary Fig. 3, ref. 5). Global comparison among the three Fusarium genomes shows that the increased genomic territory in Fol is due to additional, unique sequences that reside mostly in extra chromosomes. Syntenic regions in Fol cover approximately 80% of the Fg and more than 90% of the Fv genome (Supplementary Information E and Supplementary Table 10), referred to as the ‘core’ of the genomes. Except for telomere-proximal regions, all 11 mapped chromosomes in the Fv assembly (41.1 Mb) correspond to 11 of the 15 chromosomes in Fol (41.8 Mb). The co-linear order of genes between Fol and Fv has been maintained within these chromosomes, except for one chromosomal translocation event and a few local rearrangements (Fig. 2a).

The unique sequences of Fol are a substantial fraction (40%) of the Fol assembly, designated as Fol lineage-specific (Fol LS) regions, to distinguish them from the conserved core genome. The Fol LS regions include four entire chromosomes (chromosomes 3, 6, 14 and 15), parts of chromosome 1 and 2 (scaffold 27 and scaffold 31, respectively), and most of the small scaffolds not anchored to the optical map (Fig. 2b). In total, the Fol LS regions encompass 19 Mb, accounting for nearly all of the larger genome size of Fol.

Notably, the LS regions contain more than 74% of the identifiable transposable elements (TEs) in the Fol genome, including 95% of all DNA transposons (Fig. 2b, Supplementary Fig. 6 and Supplementary Table 11). In contrast to the low content of repetitive sequence and
translocations, and the red dashed horizontal boxes highlight the chromosome correspondences between the two genomes in the black dashed boxes. The vertical blue lines illustrate the chromosomal translocations, and the red dashed horizontal boxes highlight the Fol LS chromosomes. **Global view of syntenic alignments between F. verticillioides (Fv) and F. oxysporum (Fol)**, showing chromosome correspondences between the two genomes in the black dashed boxes. The vertical blue lines illustrate the chromosomal translocations, and the red dashed horizontal boxes highlight the Fol LS chromosomes. **Global view of syntenic alignments between F. verticillioides (Fv) and F. oxysporum (Fol)**, showing chromosome correspondences between the two genomes in the black dashed boxes. The vertical blue lines illustrate the chromosomal translocations, and the red dashed horizontal boxes highlight the Fol LS chromosomes. **Global view of syntenic alignments between F. verticillioides (Fv) and F. oxysporum (Fol)**, showing chromosome correspondences between the two genomes in the black dashed boxes. The vertical blue lines illustrate the chromosomal translocations, and the red dashed horizontal boxes highlight the Fol LS chromosomes. **Global view of syntenic alignments between F. verticillioides (Fv) and F. oxysporum (Fol)**, showing chromosome correspondences between the two genomes in the black dashed boxes. The vertical blue lines illustrate the chromosomal translocations, and the red dashed horizontal boxes highlight the Fol LS chromosomes. **Global view of syntenic alignments between F. verticillioides (Fv) and F. oxysporum (Fol)**, showing chromosome correspondences between the two genomes in the black dashed boxes. The vertical blue lines illustrate the chromosomal translocations, and the red dashed horizontal boxes highlight the Fol LS chromosomes.
carries a core genome with a high level of synteny whereas Fol and Fs each have LS chromosomes that are distinct with regard to repetitive sequences and genes related to host–pathogen interactions.

**Origin of LS regions.** Three possible explanations for the origin of LS regions in the Fol genome were considered: (1) Fol LS regions were present in the last common ancestor of the four Fusarium species but were then selectively and independently lost in Fv, Fg and Fs lineages during vertical transmission; (2) LS regions arose from the core genome by duplication and divergence within the Fol lineage; and (3) LS regions were acquired by horizontal transfer. To distinguish among these hypotheses, we compared the sequence characteristics of the genes in the Fol LS regions to those of genes in Fusarium core regions and genes in other filamentous fungi. If Fol LS genes have clear orthologues in the other Fusarium species, or paralogues in the core region of Fol, this would favour the vertical transmission or duplication with divergence hypotheses, respectively. We found that, whereas 90% of the Fol genes in the core regions have homologues in the other two Fusarium genomes, about 50% of the genes on Fol LS regions lack homologues in either Fv or Fg (1 × 10⁻²⁰). Furthermore, there is less sequence divergence between Fol and Fv orthologues in core regions compared to Fol and Fg orthologues (Fig. 3a), consistent with the species phylogeny. In contrast, the LS genes that have homologues in the other Fusarium species are roughly equally distant from both Fv and Fg genes (Fig. 3b), indicating that the phylogenetic history of the LS genes differs from genes in the core region of the genome.

Both codon usage tables and codon adaptation index (CAI) analysis indicate that the LS-encoding genes exhibit distinct codon usage (Supplementary Information E.5, Supplementary Fig. 10 and Supplementary Table 21). Taken together, we conclude that horizontal acquisition from another Fusarium species is the most parsimonious explanation for the origin of Fol LS regions.

**LS regions and host specificity.** F. oxysporum is considered a species complex, composed of many different asexual lineages that can be pathogenic towards different hosts or non-pathogenic. The Fol LS regions differ considerably in sequence among Fol strains with different host specificities, as determined by Illumina sequencing of Fol strain Fol5176, a pathogen of Arabidopsis and EST (expressed sequence tag) sequences from Fol f. sp. vasicinctum, a pathogen of cotton (Supplementary Information E.2). Despite less than 2% overall sequence divergence between shared sequences of Fol and Fol5176 (Supplementary Fig. 13A), for most of the sequences in the Fol LS regions there is no counterpart in Fol5176. (Supplementary Fig. 13B). Also for EST sequences have very high nucleotide sequence identity to the Fol genome (average 99%), but only match the core genes of Fol (Supplementary Information E.2). Large-scale genome polymorphism within Fol is also evident by differences in karyotype between strains (Supplementary Fig. 14). Previous, small, polymorphic and conditionally dispensable chromosomes conferring host-specific virulence have been reported in the fungi Nectria haematococca and Alternaria alternata. Small (<2.3 Mb) and variable chromosomes are absent in non-pathogenic F. oxysporum isolates (Supplementary Fig. 14), indicating that Fol LS chromosomes may also be specifically involved in pathogenic adaptation.

**Transfer of Fol pathogenicity chromosomes.** It is well documented that small proteins are secreted during Fol colonizing the tomato xylem system and at least two of these, Six1 (Avr3) and Six3 (Avr2), are involved in virulence functions. Interestingly, the genes for these proteins, as well as a gene for an in planta-secreted oxidoareductase (ORX1), are located on chromosome 14, one of the Fol LS chromosomes. These genes are all conserved in strains causing tomato wilt, but are generally not present in other strains. The genome data enabled the identification of the genes for three additional small in planta-secreted proteins on chromosome 14, named

---

**Figure 3 | Evolutionary origin of genes on the Fol LS chromosomes.** The scatter plots of BLAST score ratio (BSR) based on three-way comparisons of proteins encoded in core regions (a) and the Fol LS chromosomes (b). The numbers indicate the percentage of genes that lack homologous sequences in Fv and Fg (lower left corner), present in Fv but not Fg (x-axis) and present in Fg but not in Fv (y-axis). c, Discordant phylogenetic relationships of proteins encoded in the LS regions. The maximum-likelihood tree was constructed using the concatenated protein sequences of 100 genes randomly selected from 362 genes that share homologues in seven selected ascomycete genomes including the four Fusarium genomes, M. grisea, N. crassa and A. nidulans. The trees were constructed with PHYML (WAG model of evolution). The percentages for the branches represent the value based on a 10,000 bootstrapping data set.
SIX5, SIX6 and SIX7 (Supplementary Table 22) based on mass spectrometry data obtained previously\(^{20}\). Together these seven genes can be used as markers to identify each of the three supercontigs (SC 22, 36 and 51) localized to chromosome 14 (Supplementary Table 23 and Supplementary Fig. 15).

In view of the combined experimental findings and computational evidence, we proposed that LS chromosome 14 could be responsible for pathogenicity of Fol towards tomato, and that its mobility between strains could explain its presence in tomato wilt pathogens, comprising several clonal lineages polyphyletic within the Fo species complex, but absent in other lineages\(^{24}\). To test these hypotheses, we investigated whether chromosome 14 could be transferred and whether the transfer would shift pathogenicity between different strains of Fo, using the genes for in planta-secreted proteins on chromosome 14 as markers. Fol007, a strain that is able to cause tomato wilt, was co-incubated with a non-pathogenic isolate (Fom) or banana (Foc), respectively. A gene conferring resistance against zeocin (BLE) was inserted close to SIX1 as a marker to select for transfer of chromosome 14 from the donor strain into tomato. All strains of chromosome 14 as demonstrated by PCR amplification of the seven gene markers (Fig. 4c, Supplementary Fig. 15 and Supplementary Information F). The parental strains, as well as the sequenced strain Fol4287, each have distinct karyotypes. This enabled us to determine with chromosome electrophoresis whether the entire chromosome 14 of Fol007 was transferred into Fo-47\(^{+}\) strains. All Fo-47\(^{+}\) strains had the same karyotype as Fo-47, except for the presence of one or two additional small chromosomes (Fig. 4d). The chromosome present in all Fo-47\(^{+}\) strains (Fig. 4d, arrow number 1) was confirmed to be chromosome 14 from Fol007 based on its size and a Southern hybridization using a SIX6 probe (Fig. 4e). Interestingly, two double drug-resistant strains (Fo-47\(^{+}\) 1C and Fo-47\(^{+}\) 2A in Fig. 4a), which caused the highest level of disease (Fig. 4a, b), have a second extra chromosome, corresponding in size to the smallest chromosome in the donor strain Fol007 (Fig. 4d, arrow number 2).

To rigorously assess whether additional genetic material other than chromosome 14 may have been transferred from Fol007 into Fo-47\(^{+}\) strains, we developed PCR primers for amplification of 29 chromosome-specific markers from Fol007 but not Fo-47. These markers (on average two for each chromosome) were used to screen Fo-47\(^{+}\) strains for the presence of Fol007-derived genomic regions (Supplementary information F.4 and Supplementary Fig. 16). All Fo-47\(^{+}\) strains were shown to have the chromosome 14 markers (Supplementary Fig. 17), but not Fol007 markers located on any core chromosome, confirming that core chromosomes were not transferred. Interestingly, the two Fo-47\(^{+}\) strains (1C and 2A) that have the second small chromosome and caused more disease symptoms were also positive for an additional Fol007 marker (Supplementary Fig. 17), associated with a large duplicated LS region in Fol0287: scaffold 18 (1.3 Mb on chromosome 3) and scaffold 21 (1.0 Mb on chromosome 6) (Fig. 2c). The presence of most or all of the sequence of scaffold 18/21 in strains 1C and 2A was confirmed with an additional nine primer pairs for genetic markers scattered over this region (data not shown, see Supplementary Tables 25a, b for primer sequences) (Fig. 4d).

Taken together, we conclude that pathogenicity of Fo-47\(^{+}\) strains towards tomato can be specifically attributed to the acquisition of Fol

---

**Figure 4 | Transfer of a pathogenicity chromosome.** a. Tomato plants infected with Fol007, Fo-47 or double drug resistant Fo-47\(^{+}\) strains (1A through 3C) derived from this parental combination, two weeks after inoculation as described for b. b. Eight of nine Fo-47\(^{+}\) strains derived from pairing Fol007 and Fo-47 show pathogenicity towards tomato. Average disease severity in tomato seedlings was measured 3 weeks after inoculation in arbitrary units (a.u.). The overall phenotype and the extent of browning of vessels was scored on a scale of 0–4: 0, no symptoms; 1, slightly swollen and/or bent hypocotyl; 2, one or two brown vascular bundles in hypocotyl; 3, at least two brown vascular bundles and growth distortion (strong bending of the stem and asymmetric development); 4, all vascular bundles are brown, plant either dead or stunted and wilted. c. The presence of SIX genes and ORX1 in Fom, Fo-47 and Fol isolates and in double drug-resistant strains derived from co-incubation of Fol/Fom and Fol/Fo-47, assessed by PCR on genomic DNA. Co-incubations were performed with the isolates shown in bold. Three independent transformants of Fom and Fo-47 with a randomly inserted hygromycin resistance gene (H1, H2, H3) were investigated. d. Fo-47\(^{+}\) strains derived from a Fol007/Fo-47 co-incubation have the same karyotype as Fo-47, plus one or two chromosomes from Fol007. Protoplasts from Fol4287, Fol007 (with BLE on chromosome 14), three independent HYG transformants of Fo-47 (lane Fo-47 H1, H2 and H3) and nine Fo-47\(^{+}\) strains (lane 1A to 3C, the number 1, 2 or 3 referring to the HYG resistant transformant from which they were derived) were loaded on a CHEF (contour-clamped homogeneous electric field) gel. Chromosomes of S. pombe were used as a molecular size marker. Arrows 1 and 2 point to additional chromosomes in the Fo-47\(^{+}\) strains relative to Fo-47. e. Southern blot of the CHEF gel shown in d. hybridized with a SIX6 probe, showing that chromosome 14 (arrow 1 in d) is present in all strains except Fo-47 (H1, H2 and H3).
chromosome 14, which contains all known genes for small in planta-secreted proteins. In addition, genes on other LS chromosomes may further enhance virulence as demonstrated by the two strains containing the additional LS chromosome from Fol007. We did not find a double drug-resistant strain with a tagged chromosome of Fo-47 in the Fol007 background. Also, a randomly tagged transformant of Fol007 did not render any double drug-resistant colonies when co-incubated with Fo-47 (data not shown). This indicates that transfer between strains may be restricted to certain chromosomes, perhaps determined by various factors, including size and TE content of the chromosome. Their propensity for transfer is supported by the fact that the smallest LS chromosome in Fol007 moved to Fo-47 without being selected for drug resistance in two out of nine cases.

Discussion

Comparison of Fusarium genomes revealed a remarkable genome organization and dynamics of the asexual species Fol. This tomato pathogen contains four unique chromosomes making up more than one-quarter of its genome. Sequence characteristics of the genes in the LS regions indicate a distinct evolutionary origin of these regions. Experimentally, we have demonstrated the transfer of entire LS chromosome 14, which contains all known genes for small in planta secreted proteins, from a single, two-locus, genetically isolated members of Fo. The relative ease by which new tomato pathogenic genotypes are generated supports the hypothesis that such transfer between Fo strains may have occurred in nature and has a direct impact on our understanding of the evolving nature of fungal pathogens. Although rare, horizontal gene transfer has been documented in other eukaryotes, including metazoans. However, spontaneous horizontal transfer of such a large portion of a genome and the direct demonstration of associated transfer of host-specific pathogenicity has not been previously reported.

Horizontal transfer of host specificity factors between otherwise distant and genetically isolated lineages of Fo may explain the apparent polyphyletic origins of host specialization and the rapid emergence of new pathogenic lineages in otherwise distinct and incompatible genetic backgrounds. Fo LS regions are enriched for genes related to host–pathogen interactions. The mobilization of these chromosomes could, in a single event, transfer an entire suite of genes required for host compatibility to a new genetic lineage. If the recipient lineage possesses the additional LS chromosome from Fol007, which contains all known genes for small in planta secreted proteins, this may explain the apparent activity of new pathogenic lineages in otherwise distinct and incompatible genetic backgrounds. Fo LS regions are enriched for genes related to host–pathogen interactions. The mobilization of these chromosomes could, in a single event, transfer an entire suite of genes required for host compatibility to a new genetic lineage. If the recipient lineage possesses the additional LS chromosome from Fol007, which contains all known genes for small in planta secreted proteins, this may explain the apparent activity of new pathogenic lineages in otherwise distinct and incompatible genetic backgrounds.

METHODS SUMMARY

Identification of repetitive sequences. Local-alignment anchors were detected using PatternHunter (1 × 104) (ref. 29). Contiguous sets of anchors with conserved order and orientation were chained together within 10 kb distance and filtered to ensure that no block overlapped another block by more than 90% of its length.

Identification of repetitive sequences. Local-alignment anchors were detected using PatternHunter (1 × 104) (ref. 29). Contiguous sets of anchors with conserved order and orientation were chained together within 10 kb distance and filtered to ensure that no block overlapped another block by more than 90% of its length.

Characterization of promoters. Orthologous genes were determined based on BLASTP and pair-wise syntenic alignments (SI). The blast score ratio tests were used to compare relatedness of proteins among three genomes. The EMBOSs tool ‘cusp’ ([http://emboss.sourceforge.net/) was used to calculate codon usage frequencies. Gene Ontology terms were assigned using Blast2GO software (BLASTP × 1020) and tested for enrichment using Fisher’s exact test, corrected for multiple testing. A combination of homology search and manual inspection was used to characterize gene families. Potentially secreted proteins were identified using SignalP ([http://www.cbs.dtu.dk/services/SignalP/]) after removing trans-membrane/mitochondrial proteins based with TMHMM ([http://www.cbs.dtu.dk/services/TMHMM/]). Phobius (except in the first 50 amino acids), and TargetP (RC score 1 or 2) predictions. Small cysteine-rich secreted proteins were defined as secreted proteins that are less than 200 amino acids in length and contain at least 4% cysteine residues. GPI (glycosyl phosphatidyl inositol)-anchor proteins were identified by the GPI-anchor attachment signal among the predicted secreted proteins using a custom PERL script.

**Supplementary Information** is linked to the online version of the paper at www.nature.com/nature.

**Acknowledgements** The 4× sequence of *F. verticillioides* was provided by Syngenta Biotechnology Inc. Generation of the other 4× sequence of *F. verticillioides* and 6.8× sequence of *F. oxysporum* f. sp. *lycopersici* was funded by the National Research Initiative of USDA’s National Institute of Food and Agriculture through the Microbial Genome Sequencing Program (2005–35600–16405) and conducted by the Broad Institute Sequencing Platform. Wayne Xu and the Minnesota Supercomputing Institute for Advanced Computational Research are also acknowledged for their support. The authors thank Leslie Gaffney at the Broad Institute for graphic design and editing and Tracy E. Anderson of the University of Minnesota, College of Biological Sciences Imaging Center for spore micrographs.

**Author Contributions** L.-J.M., H.C.D., M.R. and H.C.K. coordinated genome annotation, data analyses, experimental validation and manuscript preparation. L.-J.M. and H.C.D. made equivalent contributions and should be considered joint first authors. H.C.K. and M.R. contributed equally as corresponding authors. K.A.B., C.A.C., J.J.C., M.-J.D., A.D.P., M.D., M.F., J.G., M.G., B.H., P.M.H., S.K., W.-B.S., C.W., X.X. and J.-R.X. made major contributions to genome sequencing, assembly, analyses and production of complementary data and resources. All other authors are members of the genome sequencing consortium and contributed annotation, analyses or data throughout the project.

**Author Information** All sequence reads can be downloaded from the NCBI trace repository. The assemblies of Fv and Fol have been deposited at GenBank under the project accessions AA0M02000000 and AA0X01000000. Detailed information can be accessed through the Broad Fusarium comparative website: http://www.broad.mit.edu/annotation/genome/fusarium_group.3/MultiHome.html. Reprints and permissions information is available at www.nature.com/reprints. This paper is distributed under the terms of the Creative Commons Attribution-Non-Commercial-Share Alike licence, and is freely available to all readers at www.nature.com/nature. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to H.C.K. (hckist@umn.edu) or M.R. (m.rep@uva.nl).