Segmental duplications drive the evolution of accessory regions in a major crop pathogen

Anouk C. van Westerhoven[1,2] (iD), Carolina Aguilera-Galvez[1] (iD), Giuliana Nakasato-Tagami[1] (iD), Xiaoqian Shi-Kunne[1] (iD), Einar Martinez de la Parte[1] (iD), Edgar Chavarro-Carrero[1] (iD), Harold J. G. Meijer[1,3] (iD), Alice Feurtey[4,5,6] (iD), Nani Maryani[7] (iD), Nadia Ordóñez[1] (iD), Harrie Schneiders[8], Koen Nijbroek[8], Alexander H. J. Wittenberg[8], Rene Hofstede[8], Fernando García-Bastidas[8] (iD), Anker Sørensen[8], Ronny Swennen[9,10] (iD), Andre Drenth[11] (iD), Eva H. Stukenbrock[4,5], Gert H. J. Kema[1]* (iD) and Michael F. Seidl[2]* (iD)

[1]Laboratory of Phytopathology, Wageningen University, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands; [2]Department of Biology, Theoretical Biology & Bioinformatics, Utrecht University, Padualaan 8, 3584 CH, Utrecht, the Netherlands; [3]Department Biointeractions and Plant Health, Wageningen University, Droevendaalsesteeg 1, 6708 PB, Wageningen, the Netherlands; [4]Christian-Albrechts University of Kiel, Christian-Albrechts-Platz 4, 24118, Kiel, Germany; [5]Max Planck Institute for Evolutionary Biology, August-Thienemann-Straße 2, 24306, Plön, Germany; [6]Plant Pathology, Eidgenössische Technische Hochschule Zürich, Rämistrasse 101, 8092, Zürich, Switzerland; [7]Biology Education, Universitas Sultan Ageng Tirtayasa, Jalan Raya Palka No.Km 3, 42163, Banten, Indonesia; [8]KeyGene, Agro Business Park 90, 6708 PW, Wageningen, the Netherlands; [9]Division of Crop Biotechnics, Laboratory of Tropical Crop Improvement, Catholic University of Leuven, Oude Markt 13, 3000, Leuven, Belgium; [10]International Institute of Tropical Agriculture, Plot 15 Naguru E Rd, Kampala, PO Box 7878, Uganda; [11]The University of Queensland, St Lucia, 4072, Brisbane, Queensland, Australia

Summary

- Many pathogens evolved compartmentalized genomes with conserved core and variable accessory regions (ARs) that carry effector genes mediating virulence. The fungal plant pathogen *Fusarium oxysporum* has such ARs, often spanning entire chromosomes. The presence of specific ARs influences the host range, and horizontal transfer of ARs can modify the pathogenicity of the receiving strain. However, how these ARs evolve in strains that infect the same host remains largely unknown.
- We defined the pan-genome of 69 diverse *F. oxysporum* strains that cause Fusarium wilt of banana, a significant constraint to global banana production, and analyzed the diversity and evolution of the ARs.
- Accessory regions in *F. oxysporum* strains infecting the same banana cultivar are highly diverse, and we could not identify any shared genomic regions and *in planta*-induced effectors. We demonstrate that segmental duplications drive the evolution of ARs. Furthermore, we show that recent segmental duplications specifically in accessory chromosomes cause the expansion of ARs in *F. oxysporum*.
- Taken together, we conclude that extensive recent duplications drive the evolution of ARs in *F. oxysporum*, which contribute to the evolution of virulence.

Introduction

The interaction between fungal plant pathogens and their hosts drives rapid evolution (Möller & Stukenbrock, 2017). Plant hosts have evolved sophisticated immune systems to detect pathogens, while adapted pathogens, in turn, secrete effectors to deregulate immune responses and to support host colonization (Rovenich *et al.*, 2014; Cook *et al.*, 2015). To facilitate this co-evolutionary cycle, many filamentous plant pathogens evolved compartmentalized genomes where effector genes are localized in distinct genomic regions (Dong *et al.*, 2015; Torres *et al.*, 2020).

These effector-rich compartments show extensive genetic variation and are enriched for repetitive sequences such as transposable elements (Raffaele & Kamoun, 2012; Dong *et al.*, 2015; Möller & Stukenbrock, 2017; Seidl & Thomma, 2017; Torres *et al.*, 2020). This spatial separation is often referred to as the 'two-speed' genome (Dong *et al.*, 2015) and is thought to facilitate the rapid diversification of effector gene repertoires to enable pathogens to evade host resistances or host jumps (Sánchez-Vallet *et al.*, 2018).

Fusarium oxysporum is a genetically diverse fungal species complex that consists of three clades that can be further separated into different lineages, some of which are recently reclassified as separate species (Maryani *et al.*, 2019). Members of this species

*These authors contributed equally to this work.

complex can collectively infect > 100 economically important crops. By contrast, individual strains typically only infect a single host plant (Edel-Hermann & Lecomte, 2019). *Fusarium oxysporum* evolved a two-speed genome organization where specific genomic regions are characterized by extensive presence-absence polymorphisms between strains (Ma *et al.*, 2010; van Dam *et al.*, 2017; Fokkens *et al.*, 2018; Zhang *et al.*, 2020). These variable accessory regions (ARs) can be embedded in core chromosomes or can encompass entire accessory chromosomes (Ma *et al.*, 2010; Fokkens *et al.*, 2018). Interestingly, shared ARs in otherwise genetically diverse strains have been linked to the adaptation toward the same host (Fokkens *et al.*, 2018; Henry *et al.*, 2021). Moreover, ARs can transfer the capacity to infect a specific host between isolates (Ma *et al.*, 2010; Ayukawa *et al.*, 2021). For example, the transfer of accessory chromosome 14 from a tomato-infecting *F. oxysporum* strain (*Fol*4287) to a nonpathogenic strain converts the nonpathogenic isolate into a tomato-infecting pathogen (Ma *et al.*, 2010). This transfer of pathogenicity is likely linked to effector genes that are localized in ARs such as some of the 14 well-studied *Secreted In Xylem* (*SIX*) effectors (Rep *et al.*, 2002). Consequently, the presence and absence of effectors can group genetically diverse *F. oxysporum* strains based on their host range (van Dam *et al.*, 2016). Whole-genome sequencing of diverse *F. oxysporum* can provide novel insights into the origin of host specificity (Fokkens *et al.*, 2018; Henry *et al.*, 2021). For example, a recent population study of *F. oxysporum* isolates infecting chickpea suggests that, contrary to expectations (Gordon & Martyn, 1997; Henry *et al.*, 2021), *F. oxysporum* might undergo sexual and clonal reproduction (Fayyaz *et al.*, 2023), which could impact the constitution and the dynamics of ARs and, ultimately, influence pathogenicity. However, despite the importance of ARs for host adaption, little is known about the constitution of ARs in *F. oxysporum* strains that cause disease in the same host, and the origin and processes that shape the evolution of ARs.

Fusarium wilt of banana (FWB), which is caused by a suite of *F. oxysporum* species (Maryani *et al.*, 2019), is a major constraint for banana production and threatens food security in tropical and subtropical countries where > 400 million people depend on banana cultivation (FAO, 2020). In *F. oxysporum* causing FWB, three different races can be distinguished based on their pathogenicity toward subsets of banana varieties; race 1 (R1) strains cause FWB in Gros Michel, race 2 (R2) strains infect Bluggoe, and tropical race 4 (TR4) strains infect Gros Michel, Bluggoe, and Cavendish (Ploetz, 2015). In addition to TR4, subtropical race 4 (STR4) strains can cause disease in Cavendish banana under environmental stress conditions, for example caused by lower temperatures in subtropical regions (Ploetz, 2006). All races infect additional locally important banana varieties. TR4 has developed into a pandemic over the last 60 yr (Su *et al.*, 1986; Ordonez *et al.*, 2015; van Westerhoven *et al.*, 2022b) and is of particular concern as it causes FWB in Cavendish, the banana variety that dominates global productions (> 50%) and export trade (> 95%; FAO, 2020). Despite the impact of FWB on banana production, we know little about the ARs in *F. oxysporum* strains that infecting banana and their relation to pathogenicity.

Here, we analyze the dynamics of ARs in a suite of *F. oxysporum* strains causing FWB. To gain insights into the genomic structure, we generate seven chromosome-level reference genome assemblies for R1, R2, and TR4 strains, and use a pan-genomic framework together with a global panel of 69 strains to gain insights into the evolution of ARs in genetically diverse *F. oxysporum* strains.

Materials and Methods

Collection of *Fusarium oxysporum* strains infecting banana

We analyzed a diverse set of 69 *Fusarium oxysporum* strains infecting banana (Supporting Information Table S1). *Fusarium oxysporum* strains were isolated from different banana growing regions world-wide directly from the pseudostem of symptomatic banana plants. Here, 35 *F. oxysporum* strains were newly sequenced and we included 34 publicly available whole-genome sequencing datasets (based on short-read sequencing approaches; García-Bastidas *et al.*, 2014; Guo *et al.*, 2014; Zheng *et al.*, 2018; Warmington *et al.*, 2019; Garcia-Bastidas *et al.*, 2020; Maymon *et al.*, 2020; van Westerhoven *et al.*, 2022a). To obtain high-quality reference assemblies, seven strains were sequenced with Oxford Nanopore Technology (ONT).

Whole-genome sequencing

Fusarium oxysporum isolates were cultured on potato dextrose agar (PDA), and conidia were transferred to liquid media for incubation at 25°C. Fungal material was isolated from liquid culture for genomic DNA extraction. The exact DNA isolation protocol per strain can be found in Methods S1. To obtain long-read, whole-genome sequencing data, seven *F. oxysporum* strains (36102, II5, CR1.1, C058, C135, C081, and M1) were sequenced using Oxford Nanopore Technologies (Oxford, UK). An R9.4.1 flow cell was loaded and run for 24 h. Base calling was performed using GUPPY (v.3.1.5; C058, C135, C081, and M1) and MINIKNOW (v.3.4.6; II5, 36 102, and CR1.1). The other *F. oxysporum* strains were sequenced by Beijing Genomic Institute using Illumina paired-end sequencing, 35 were sequenced during this study, and 34 were obtained from other studies (Guo *et al.*, 2014; Ordonez *et al.*, 2015; Zheng *et al.*, 2018; Warmington *et al.*, 2019; Garcia-Bastidas *et al.*, 2020; Maymon *et al.*, 2020; van Westerhoven *et al.*, 2022a).

Gene expression analysis

Gene expression patterns were compared between *in vitro* growth and 8 d post infection. Cavendish 'Grand Naine' banana plants were grown in a glasshouse compartment (28 ± 2°C, 16 h light, and *c.* 85% relativity humidity) and acclimatized under plastic for 2 wk to maintain high humidity. The roots of *c.* 2.5-month-old plants were dip inoculated with 10^6 spores ml^{-1}. RNA was isolated from the roots of inoculated plants at 8 d post inoculation and from mycelium grown on PDA medium. Samples were ground, and RNA extraction was performed using the Maxwell

Plant RNA kit (Promega, Madison, WI, USA) following the manufacturer's instructions. The quality of the RNA samples was tested by agarose electrophoresis and quantified by Nanodrop (Thermo Fisher, Waltham, MA, USA), and subsequently sent to KeyGene (Wageningen, the Netherlands) for sequencing. Quantifications of gene transcripts from sequenced RNA-Seq reads were performed using Kallisto (Bray *et al.*, 2016).

Genome assembly and annotation

We obtained chromosome-level *de novo* genome assemblies for seven *F. oxysporum* strains sequenced by ONT long-read technology. Adapter sequences were removed with PORECHOP (v.0.2.4, default settings; https://github.com/rrwick/Porechop/tree/master), and reads were self-corrected, trimmed, and assembled using CANU (v.1.8; Koren *et al.*, 2017). Genome assemblies were polished using ONT raw reads with RACON (v.1.5.0; Vaser *et al.*, 2017) and with high-quality Illumina short read using two rounds of PILON (v.1.24; Walker *et al.*, 2014). Contigs were scaffolded, if needed, using LONGSTITCH (v.1.0.2; Coombe *et al.*, 2021).

Fusarium oxysporum strains sequenced solely with short-read sequencing technology were *de novo* assembled with SPADES (v.3.13.0. with default settings; Bankevich *et al.*, 2012). Contigs shorter than 500 bp were subsequently removed. To assess the assembly quality, the genome assemblies were analyzed with QUAST (v.5.0.2; Gurevich *et al.*, 2013), and genome completeness was estimated based on the presence of conserved single-copy genes using BUSCO (v.5.1.2; Simão *et al.*, 2015) with the *hypocreales* odb10 database.

We performed gene annotation of 69 *F. oxysporum* genome assemblies using FUNANNOTATE (v.1.8.9; Palmer & Stajich, 2019) with the *F. oxysporum* f. sp. *lycopersici* 4287 Uniprot proteins, Uniport database, and previously predicted effectors (van Dam *et al.*, 2016) as external evidence. Genomes were masked using REPEATMASKER (v.4.1.1) following repeat prediction with REPEATMODELER (http://www.repeatmasker.org). For strains II5, CR1.1, and 36 102, RNA-Seq data were used to guide gene model prediction, and *ab initio* training parameters estimated using RNA-Seq data of II5 (TR4, *F. odoratissimum* Maryani, Lombard, Kema & Crous) were re-used to support the annotation in other genome assemblies. Effectors genes were predicted based on their association with Miniature IMPala (MIMP) elements in *F. oxysporum* (van Dam *et al.*, 2016). Briefly, ORFs were predicted 5-kb up- and downstream to MIMPs ('AGT[GA][GA]G[GAT][TGC]GCAA[TAG]AA') and these were refined using AUGUSTUS (v.3.3.3; Stanke & Morgenstern, 2005). In the set of predicted MIMP-genes, secreted proteins were predicted using SIGNALP (v.5) and subsequently, effectors were predicted using EFFECTORP (v.3.0). The predicted effectors were added to the Funannotate annotation using AGAT (v.0.8.0).

To recover potentially missing genes in the gene annotations, we used the comparative annotation toolkit (v.2; Fiddes *et al.*, 2018). CACTUS (v.2.0.1; Armstrong *et al.*, 2020) was used to create whole-genome alignments of all 69 genomes, using the Minigraph-Cactus approach (Hickey *et al.*, 2023). Consecutively, the CAT was used to carry over gene annotations based on Funannotate gene

predictions with the RNA-Seq data (with Augustus settings pre trained on II5, --global-near-best 1, --augustus-utr-off, --augustus-cgp, --maf-chunksize 550 000 –maf-overlap 50 000), which added *c.* 200 genes per genome assembly.

Core gene phylogeny

The phylogenetic relationship between *F. oxysporum* strains infecting banana together with 55 publicly available nonbanana-infecting *F. oxysporum* strains was inferred from 3811 conserved single-copy BUSCO genes. Protein sequences were aligned using MAFFT (v.7.453; Katoh *et al.*, 2002), and a maximum-likelihood phylogeny was determined using IQTREE (v.1.6; -m MFP+MERGE, -B 1000 with 1000 ultra-fast bootstrap; Chernomor *et al.*, 2016); *Fusarium verticillioides* (ASM14955v1; Cuomo *et al.*, 2007) was included as an outgroup.

Pan-genome analysis

We detected conserved and ARs in the set of *F. oxysporum* isolates by performing all-vs-all whole-genome alignments with NUCMER (v.3.1; −−max-match; Kurtz *et al.*, 2004). We removed small alignments with delta-filter (-l 5000) and only retained the best match per alignment (-1). Genomic regions present in all 69 *F. oxysporum* isolates were considered core, regions covered by < 80% were considered accessory, and regions between these values were considered softcore. Pairwise similarity between ARs was calculated using custom python scripts (https://github.com/Anouk-vw/Fusarium_pg), and alignments were visualized with PYGENOMEVIZ (https://github.com/moshi4/pyGenomeViz) and CIRCLIZE (Gu *et al.*, 2014).

To identify orthologous groups (OGs), we used BROCCOLI (v.1.1; Derelle *et al.*, 2020) on the predicted protein-coding genes from the 69 *F. oxysporum* strains infecting banana; per gene, only the longest transcript was included. Orthologous groups present in all isolates were considered core, OGs found in > 55 genomes (80%) were considered softcore, and genes found in < 55 of the genomes were considered accessory.

We determined the number of nonsynonymous and synonymous substitutions (dN/dS) per orthologous group. All genes in an OG were aligned using MAFFT (v.7.490; Katoh *et al.*, 2002) and nonsynonymous and synonymous substitutions per gene pair in the OG were inferred based on a codon-guided nucleotide alignment, created by PAL2NAL (v.14; Suyama *et al.*, 2006), with CODEML (from PAML, v.4.9; Yang, 2007). Differences between groups of genes were further analyzed using custom python scripts and compared using a two-sided Mann–Whitney *U*-test in SCIPY (v.1.10.1; Virtanen *et al.*, 2020).

Estimation of gene age

ORTHOFINDER (v.2.5.4; Emms & Kelly, 2019) was used to detect the presence of homologs in 308 phylogenetically distinct fungal species from the joint genome institute; per fungal family, one fungal genome assembly was randomly selected (Table S2). Homologs that occur at least once outside the phylum

Ascomycota are considered *old* homologs; homologs occurring only within Ascomycota are further subdivided into *Hypocreales* and *Fusarium.*

Detection of duplication events

We detected gene duplication events by self-comparisons of the translated protein sequences (only the longest isoform) of the chromosome-level genome assemblies of *F. oxysporum* infecting banana, as well as *Fusarium oxysporum* f. sp. *lycopersici* (strain *Fol*4287; Ma *et al.*, 2010; GCF_000149955.1), *Fusarium graminearum* (Cuomo *et al.*, 2007; GCF_000240135.3), *Fusarium verticillioides* (Cuomo *et al.*, 2007; GCF_000149555.1), and *Fusarium solani* (Mesny *et al.*, 2021; GCF_020744495.1). Per *Fusarium* genome assembly, the predicted protein sequences were compared with BLAST and the best five hits were retained after filtering for an e-value $1e^{-10}$ and query coverage larger than 50%. Then, BLAST alignments were used to detect collinear duplicated blocks with MCscanX (Wang *et al.*, 2012); blocks of five genes were considered collinear, and matches were filtered with the parameters matchscore 50 and e-value 10^{-5}. Subsequently, we classified duplicated genes into dispersed, proximal, tandem, or segmental duplications with MCscanX's 'duplicate_gene_classifier' script.

Copy number variations were identified by mapping short-read, whole-genome sequencing datasets against the chromosome-level genome assembly of strain II5 using BWA-MEM (v.0.7.17; Li, 2013). We determined the sequencing coverage over the reference genome with BEDTOOLS (v.2.30), which was visualized Wgscoverageplotter from JVARKIT (v.1.1.0; Lindenbaum, 2015), filtering reads with a mapping quality of 0. Single nucleotide polymorphisms were called using GATK (v.4.0.2; −ploidy 1) and filtered according to GATK best practices (Van der Auwera *et al.*, 2013).

Results

Fusarium oxysporum strains infecting banana are genetically diverse

To study the occurrence and evolution of ARs in *Fusarium oxysporum* strains infecting banana, we sequenced and assembled 69 strains (Fig. 1a; Table S1) that were isolated from all major banana growing regions and classified into R1 (39 strains), R2 (two strains), and TR4 (28 strains; Fig. 1a,b; Table S1). *De novo* assembly of seven strains infecting different banana varieties (three TR4, two R1, and two R2) sequenced with long-read sequencing technology achieved chromosome-level contiguity containing 12–15 contigs, most of which represent complete chromosomes flanked by telomeric repeats (Tables S3, S4). While genome assemblies obtained solely from short-read data yielded more fragmented assemblies, all assemblies contained at least 96.9% of the single-copy BUSCO genes. The assemblies range from 43 to 51 Mb in size (Fig. 1b) and encode between 15 664 and 17 865 predicted proteins, of which 551–671 are effector candidates, and consist of 2.6–8.9% of repeats (Fig. 1e, f), similar to previous *F. oxysporum* genome assemblies (van Dam *et al.*, 2016; Warmington *et al.*, 2019; Zhang *et al.*, 2020).

To assess the genetic diversity of *F. oxysporum* strains infecting banana in relation to 55 publicly available *F. oxysporum* strains infecting eight different host plants (van Dam *et al.*, 2016; Table S1), we conducted a phylogenetic analysis based on 3811 single-copy orthologous genes (2226 902 amino acid positions). As expected, *F. oxysporum* strains infecting banana are polyphyletic and associated with *F. oxysporum* clades 1, 2, and 3 (Fig. 1a; O'Donnell *et al.*, 1998; Maryani *et al.*, 2019). We observed low nucleotide diversity between TR4 strains (pi-nucleotide diversity 0.0018) compared with a 100× higher diversity between R1 strains (pi-nucleotide diversity 0.185), corroborating that most TR4 strains evolved from a single recent clonal origin (Ordonez *et al.*, 2015; Maryani *et al.*, 2019). Based on their genetic diversity, *F. oxysporum* strains infecting banana have recently been separated into different species, and the lineage that encompasses TR4 strains (Fig. 1d) is now referred to as *Fusarium odoratissimum* (Ordonez *et al.*, 2015; Maryani *et al.*, 2019).

Fusarium oxysporum strains causing Fusarium wilt of banana have a compartmentalized genome

To identify ARs in *F. oxysporum* strains infecting banana, we performed whole-genome alignments of the seven chromosome-level genome assemblies to the reference genome assembly of the tomato-infecting *F. oxysporum* strain 4287 (*Fol*4287; Ma *et al.*, 2010). We observed 11 homologous chromosomes between the *F. oxysporum* strain II5 (TR4; *F. odoratissimum*) and *Fol*4287, which are considered the core genomes (chromosomes 1–11; Fig. 2a). In addition to the core chromosomes, we observed two large ARs specific for strain II5. The first spans 1.8 Mb, occupying 27% of chromosome 1 in proximity to one telomere and the second region is 1.1 Mb in size and constitutes the entire contig 12 (Fig. 2a). Contig 12 only contains one telomere (Fig. 2b), and therefore, it remains currently unknown whether this AR represents an extra chromosome or is attached to one of the core chromosomes. Importantly, however, we assembled a similar contig in TR4 strain M1, and a similar contig is present in the independently assembled TR4 strain UK0001/Eden (Warmington *et al.*, 2019), which strongly suggests that this contig represents an independent AR.

Compared with the 11 core chromosomes, the two ARs have an increased repeat content (28.30% vs 6.41%) and a decreased gene density (32.15% vs 59.71%) and GC content (46.4% vs 47.5%; Fig. 2b). Importantly, the ARs in strain II5 encode 463 genes, nine are homologous to *SIX* genes (Fig. 2b) and 22 are predicted effectors (out of 629 predicted effectors).

To further analyze the diversity of ARs, we identified ARs in our global collection using a pan-genomic approach based on all-vs-all whole genome alignments. ARs were defined as regions longer than 5 kb that are absent in > 55 of 69 strains (80%). We identified a varying amount of ARs in *F. oxysporum* strains infecting banana (Fig. 1c), ranging from 6.7 Mb (15% of the genome size) in strain Indo110 (R1) to 15.5 Mb (30%) in strain C135 (R2), in line with the 19 Mb (29%) of ARs in the reference strain *Fol*4287 (Ma *et al.*, 2010). Next to the two ARs in chromosome 1 and contig 12, our pan-genomic approach identified an
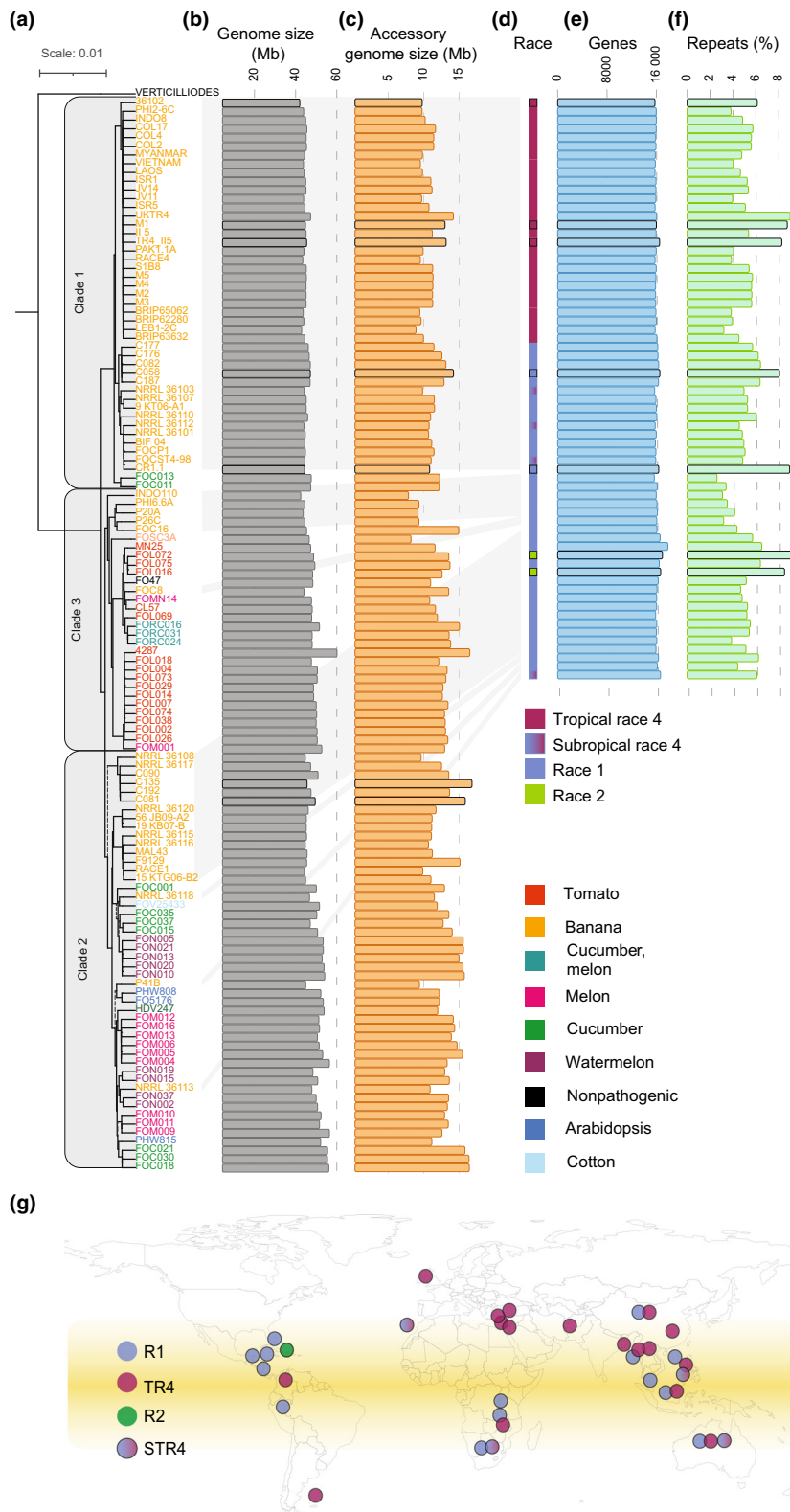
Fig. 1 *Fusarium oxysporum* strains infecting banana are genetically diverse. (a) *Fusarium oxysporum* strains that infect banana are polyphyletic (yellow labels). The relationship between *F. oxysporum* infecting banana and *F. oxysporum* strains that infect other plant hosts (see label color) is estimated using maximum-likelihood phylogenetic analysis based on 3811 concatenated conserved single-copy BUSCO genes, dashed lines indicate branches with a bootstrap value below 80. Black outline indicates high-quality genome assemblies. (b) Genome sizes of *F. oxysporum* strains range between 43 and 61 Mb. (c) The sizes of accessory regions in *F. oxysporum* strains range from 6.7 to 16.5 Mb. (d) TR4 strains (purple squares) occur in clade 1 and are genetically similar, as well as the two race 2 strains (green). By contrast, the race 1 strains (blue squares) are polyphyletic. (e) *F. oxysporum* strains that infect banana encode between 15 664 and 17 865 genes and (f) between 2.9% and 8.9% repeats. (g) The collection of 69 *Fusarium* strains originates from different banana growing regions world-wide; the approximate origin of isolation is indicated by dots, colored according to the race designation: R1, race 1; R2, race 2; STR4 subtropical race 4, and TR4, tropical race 4. Strains sequenced by long-read sequencing technology are highlighted in black.

additional 8 Mb of ARs in strain II5, typically localized at sub-telomeric regions (defined as the first and last 10% of each chromosome; Fig. 2b). These sub-telomeric regions contain 1287 genes, of which 80 are predicted effector genes (6.6%). We also identified that chromosomes 9 (42.6% ARs, 87 single nucleotide polymorphisms (SNPs) kb^{-1}), 10 (27% ARs, 80 SNPs kb^{-1}), and 11 (36% ARs, 92 SNPs kb^{-1}) are less conserved than the other core chromosomes (on average 20% ARs, 59 SNPs kb^{-1};
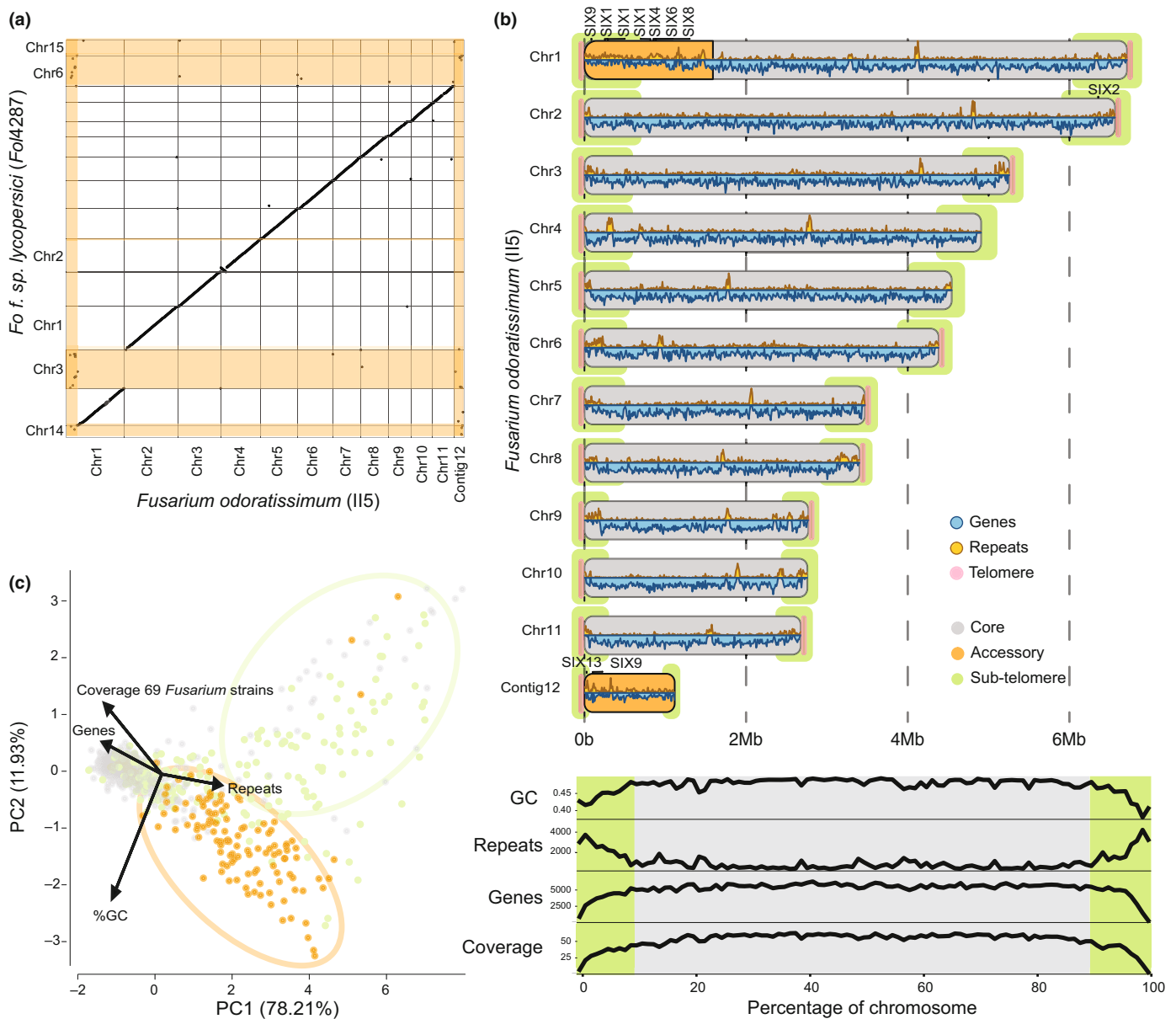
Fig. 2 *Fusarium oxysporum* strains infecting banana carry accessory regions (ARs). (a) The tropical race 4 (TR4) reference strain II5 (*Fusarium odoratissimum*, *TR4*) contains 11 core chromosomes (chromosomes 1–11) as well as two ARs (1.8 Mb on chromosome 1 and entire contig 12). Core chromosomes align to *Fol*4287 in contrast to the ARs that are highlighted in orange. (b) The ARs (orange) in II5 carry most *Secreted In Xylem* (*SIX*) genes and show an increased repeat and lower gene content. A similar pattern is found in the sub-telomeric regions (green, first and last 10% of the chromosome). (c) Principal component analysis on gene-, repeat-, GC content, and coverage of 69 *F. oxysporum* strains separates the genomic regions of II5 into three different clusters. ARs (orange and green) are separated from the core region (grey) and further separated into those ARs localized in sub-telomeric regions (green) and those on chromosome 1 and contig 12 (orange).

Fig. S1), which resembles the situation of the previously identified 'fast-core' chromosomes in *Fol*4287 (Fokkens *et al.*, 2018).

We show that ARs can be defined based on sequence conservation and differ from the core genome in gene, repeat, and GC content. Consequently, principal component analyses distinguished ARs from the core genome based on these four features (Fig. 2c). Unanticipatedly, this analysis also separated the large ARs on chromosome 1 and contig 12 from the smaller ARs localized at the sub-telomeres (Fig. 2c); this separation is driven by

the slightly lower GC content in sub-telomeric regions (44% vs 47%). A reduced GC content can be caused by repeat-induced polymorphisms (RIP) and fungal defense mechanisms that introduces C-to-T mutations in repetitive regions (Galagan & Selker, 2004). In the sub-telomeric region, 25.4% of the repetitive elements are likely subjected to RIP (i.e. composite RIP index > 0), in contrast to only 7.7% in the other ARs, suggesting that repetitive elements at the sub-telomeres are older, or alternatively, that RIP has been more active.
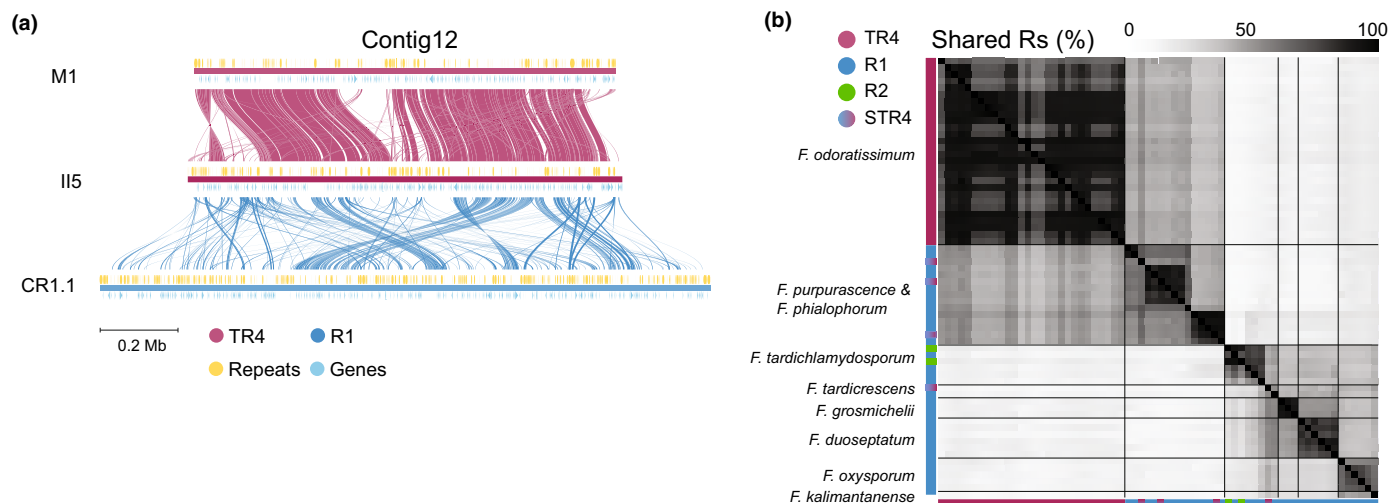
Fig. 3 Accessory regions (ARs) in *Fusarium oxysporum* strains infecting banana differ significantly. (a) Accessory contig 12 of the *F. oxysporum* tropical race 4 (TR4) reference strain II5 (*F. odoratissimum*) is shared with TR4 strain M1; however, this contig shows little similarity to race 1 (R1) strain CR1.1. Lines (red and blue) indicate aligned regions between strains based on nucleotide similarity. Small alignments are present between R1 and TR4 (blue lines); however, these are rearranged and often involve repetitive sequences (yellow blocks). (b) R1 strains (39 strains, highlighted in blue) have a diverse accessory content sharing on average only 20% of the ARs, whereas TR4 strains (28, highlighted in purple) share most of their ARs (91%). Phylogenetically related *F. oxysporum* strains, here indicated by a recently proposed species name (Maryani *et al.*, 2019), share more ARs than distantly related strains.

Highly variable accessory regions in *Fusarium oxysporum* strains infecting banana

To characterize the similarity of ARs in *F. oxysporum* strains infecting banana, we compared the ARs from the seven chromosome-level genome assemblies. Notably, the ARs are highly diverse with little sequence and structural similarity (Figs 3, S2). In TR4, the ARs from strain II5 on chromosome 1 and contig 12 show extensive similarity to the corresponding chromosomes in strain M1 (Figs 3a, S2). Instead, isolate 36102 (*F. odoratissimum*), causing milder symptoms in Cavendish than TR4 strain II5, carries only the AR on chromosome 1 and does not encode a region similar to contig 12 (Fig. S2). None of the ARs from strain II5 are present in the strains representing race 1 and race 2. Similar to the TR4 strains, the accessory contigs of two race 2 strains correspond to each other (Fig. S2), but no shared ARs are found among R1 strains. Moreover, we observed that no ARs are conserved among different races; for example, only limited genetic material is shared between ARs of TR4 and R2 (Fig. S2). The diversity of ARs was further corroborated by pairwise analyses of all 69-strains, which revealed that, on average, TR4 shared 91% of the ARs, whereas R1 shared only 24% (Fig. 3b). In line with the high number of shared ARs in TR4 (*F. odoratissimum*; Fig. 3b), the amount of shared ARs was the highest among R1 strains belonging to the same *Fusarium* species/lineage (60% shared ARs; FigS 3b, S3); however, R1 strains from different species/lineages share very few ARs (Fig. 3b).

To quantify to what extent *F. oxysporum* strains infecting banana share ARs with *F. oxysporum* strains infecting different plant hosts, we identified and compared ARs in a set of 55 additional *F. oxysporum* strains that are pathogenic to different plant

hosts (van Dam *et al.*, 2016). We observed that genetically related strains share most ARs (Fig. S4); however, we also noticed that *F. oxysporum* strains infecting banana encode more diverse ARs than for instance the tomato-infecting strains; on average, *Fol* strains share 62% ARs, with a minimal 30% of the ARs shared between any two *Fol* strains (Fig. S4). By contrast, banana-infecting strains on average share 44% of the ARs and a pair of strains can share as little as 12% (Fig. S4). Thus, we demonstrate that ARs are highly variable among *F. oxysporum* strains infecting banana, and importantly, strains do not share one host-associated chromosome in contrast to previously reports for *F. oxysporum* infecting tomato (Fokkens *et al.*, 2018).

Evolutionary dynamics of the accessory genome

Fusarium oxysporum strains infecting banana do not share large ARs (Fig. 3b), but we reasoned that ARs nevertheless might share genes that contribute to the pathogenicity of banana. We therefore grouped 1.1 million protein-coding genes predicted in 69 strains into 22 612 OGs and subsequently analyzed the gene-content diversity. The pan-genome consisted of 12 101 core groups (53%) that are present in all 69 strains, 2395 soft-core groups present in at least 80% of the strains, 5595 (25%) accessory groups present in <80% of the strains, and 2521 unique genes (Fig. 4a). Importantly, the pan-genome based on our collection captures the diversity of protein-coding genes in *F. oxysporum* strains infecting banana as we did not observe an increase in the pan-genome size after including >40 strains (Fig. 4a). The conserved core genes are, as expected, enriched with housekeeping genes, while accessory genes are enriched with genes encoding secondary metabolites (Fig. S5). No
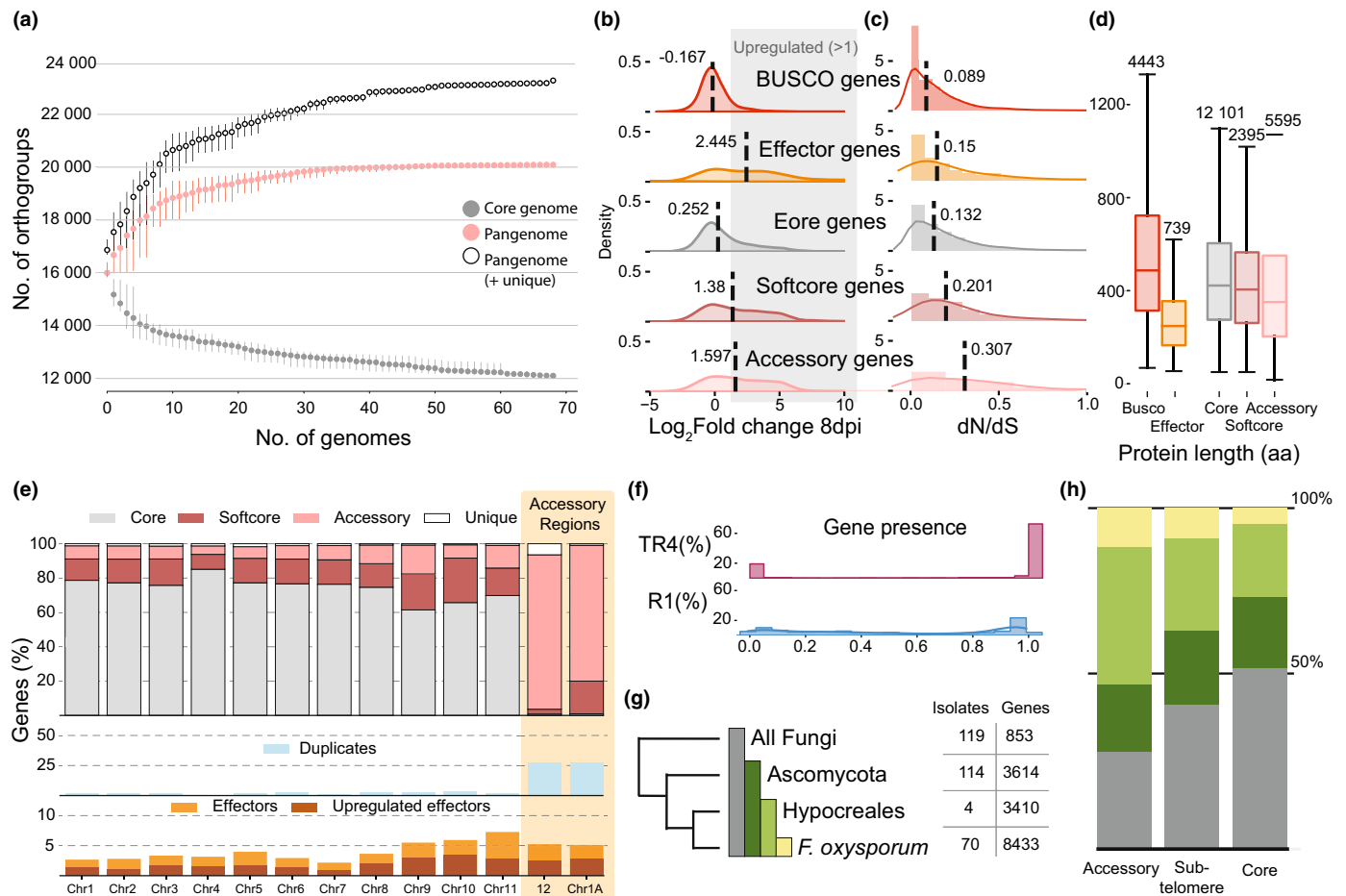
Fig. 4 Evolutionary dynamics of genes in the different genomic compartments in *Fusarium oxysporum* strains infecting banana. (a) Global collection of *F. oxysporum* strains infecting banana has a pan-genome size of 20 091 orthologous groups (OGs, 22 612 when unique genes are included), of which 12 101 are present in all strains (grey). The pan-genome is saturated after adding 40 genomes. The order of genomes was randomly sampled 10 times, dots show the average pan-genome size, and the error bars indicate the SD. (b) Accessory genes are upregulated *in planta* at 8 d post inoculation (dpi) of banana plants. Average $\log_2$ fold change in accessory genes (2416 in II5) and effector genes (629 in II5) shows these genes are typically upregulated upon plant infection, which contrasts the core (11 740 genes, $\log_2$-fold change = 0.252) and BUSCO genes (4364 genes, $\log_2$-fold change = −0.167). The median $\log_2$ fold change between all gene categories is significantly different ($P < 0.05$, two-sided Mann–Whitney U-test). Dotted lines indicate the median value. (c) Sequence conservation estimated by the number of nonsynonymous and synonymous substitutions (dN/dS) values. BUSCO gene families (4443 OGs) and core gene families are most conserved, whereas the accessory genes evolve under more relaxed selective pressure. All gene categories show significant variation in dN/dS value ($P < 0.05$, Mann–Whitney U-test). Dotted lines indicate the median value. (d) Distribution of protein length separated by different gene categories shows that accessory gene families (5595 OGs, pink) are shorter (median length of 315 aa per OG) than core gene families (median length 469 amino acids per OG, 12 101 OGs, grey). All gene categories differed significantly from each other ($P < 0.05$, Mann–Whitney U-test). Boxes indicate the interquartile range of the values, the median values are indicated by horizontal lines, and the whiskers depict the minimum and maximum values. (e) Accessory regions (ARs) in II5 (orange, contig 12, and chromosome 1A) consist mostly of accessory genes (pink, 405 out of 463, 87%), and 28% of these genes are part of expanded orthogroups (129 out of 463). Accessory genes can also be found in core chromosomes. (f) Gene content is variable between race 1 strains, and only 20% of the genes are present in 80–90% of the race 1 strains. Most genes (60%) in tropical race 4 (TR4) are present in all TR4 strains (90–100%). (g) Gene families in strain II5 (*Fusarium odoratissimum*, TR4) differ to which extend they have homologs in 308 fungal phyla; most genes are present in related *F. oxysporum* strains, but only a few gene families (853) are conserved among all considered fungi. (h) ARs contain most recent genes, present in *Fusarium* (51) and Hypocreales (160). Genes in core regions are evolutionarily older with more genes that have homologs in 'all fungi' (7764 out of 11 740 genes).

enrichment of effector genes is found in any of the three categories (core, accessory, or softcore); the pan-genome consists of 739 gene families encoding predicted effectors that are evenly distributed over the different gene categories (3.7% of accessory genes and 3.5% of the core genes). The role of accessory genes in host infection is suggested by gene expression, as accessory genes are upregulated 8 d after banana infection, with

a median $\log_2$ fold change of 1.59 (compared with the *in vitro* control; Fig. 4b), comparable to the upregulation of effector genes (median $\log_2$ fold change of 2.4). The core genes, however, show significantly less increase in gene expression upon plant infection (median $\log_2$ fold change = 0.25, $P < 0.05$, two-sided Mann–Whitney U-test; Fig. 4b). Although this suggests a role of accessory genes in host-pathogenicity, we observed a
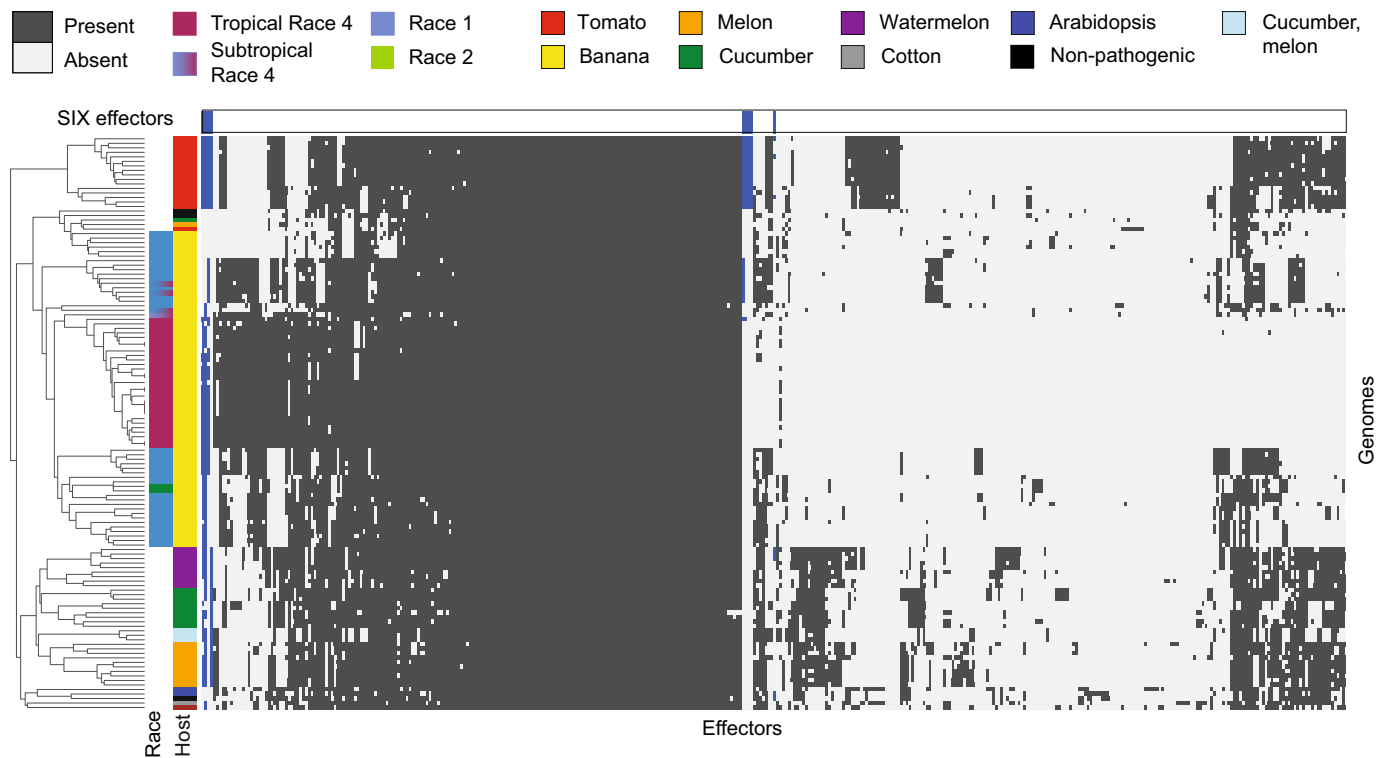
Fig. 5 Predicted effector profile clusters *Fusarium oxysporum* strains infecting the same host. *Fusarium oxysporum* strains infecting banana (yellow) are separated from *F. oxysporum* strains infecting other hosts. In the set of 124 *Fusarium* isolates, 398 Miniature IMPala (MIMP) associated effector families are predicted. The race 1 strains (blue) have diverse effector profiles and do not cluster together, whereas the effector profiles of tropical race 4 strains (purple) are highly similar. Nine Secreted In Xylem (SIX) effectors (blue) are identified in the 398 MIMP-associated effector families and are present in *F. oxysporum* strains infecting tomato, from which they are originally identified, as well as in other *F. oxysporum* genomes.

highly diverse gene content in R1 strains with only 246 out of 7832 accessory genes (3%) being shared among all R1 strains (Fig. 4f). TR4 strains, by contrast, have a more similar gene profile and share 2493 of the 4406 accessory genes (57%).

In TR4 strain II5, ARs have a different gene composition than the core regions, with ARs consisting mostly of accessory genes (405 out of 463; 87%); however, four core and 38 soft-core genes are also found in ARs (Fig. 4e), indicating that some shared genes can be found between otherwise nonconserved regions. ARs seem to contain recently evolved genes, as most AR genes (71%) have homologs in closely related fungi (Fig. 4g,h) and only 29% are considered 'old', that is with homologs in all fungi, which contrasts to the 53% old genes in the core genome (7764, out of 14 774 genes; Fig. 4g,h). Additionally, accessory genes are significantly shorter (315 aa) and evolve under relaxed selective pressure (dN/dS = 0.20), compared with the core genes (469 aa, dN/dS = 0.089; $P < 0.05$, two-sided Mann–Whitney U-test; Fig. 4b,c), both hallmarks of more recent genes (Wolf *et al.*, 2009). This suggests that most genes in the ARs evolved recently and that these regions serve as cradles for the evolution of novel genes. In addition, many genes on the ARs (177 out of 463) belong to expanded gene families based on the defined orthogroups (Fig. 4e), suggesting that gene duplications play a role in the evolution of ARs in *F. oxysporum* strains infecting banana.

Variable effector repertoire in *Fusarium oxysporum* genomes

Effector gene repertoires define host range (van Dam *et al.*, 2016; Brenes Guallar *et al.*, 2022), and we anticipated that effector presence-absence profiles would distinguish *F. oxysporum* strains infecting banana from *F. oxysporum* strains infecting other hosts. We predicted candidate effector genes in all 124 *F. oxysporum* isolates based on their proximity to MIMP elements that are known to co-localize with some effector genes in *F. oxysporum* (van Dam *et al.*, 2016; Brenes Guallar *et al.*, 2022), which yielded 398 MIMP-associated effector gene families (160 of these were predicted in the complete set of 739 candidate effectors). The predicted MIMP-associated effector repertoires clustered *F. oxysporum* isolates in host range (van Dam *et al.*, 2016; Brenes Guallar *et al.*, 2022; Fig. 5). While banana-infecting strains form a clear cluster, only four MIMP-associated effectors were identified that occur in most (i.e. > 80%) of banana-infecting strains but are consistently absent (i.e. present in < 20%) in other *F. oxysporum* strains. Moreover, races within *F. oxysporum* strains infecting banana are not clearly separated; TR4 strains have very similar effector profiles, yet R1 strains differ considerably and do not encode shared effectors. Interestingly, two STR4 strains carry an effector profile remarkably similar to the effector profile of TR4 strains (Fig. 5). This suggests that this effector profile

contributes to the pathogenicity of these STR4 strains to Cavendish banana. However, this effector profile is not STR4 specific as the other two STR4 strains encode effector profiles similar to R1 strains (Fig. 5). This difference can be the result of misclassification of STR4 strains, because STR4 infects Cavendish only under environmental stress conditions and this environmental component makes it difficult to distinguish true STR4 isolates from race 1 isolates (Ploetz, 2006). To further investigate effectors shared between TR4 and STR4, an accurate screening strategy to distinguish true STR4 needs to be conducted. We here reveal a diverse effector repertoire in R1 strains, together with the variable ARs and gene content, this suggests that the species encompassing R1 utilize varying molecular mechanisms to support banana infection.

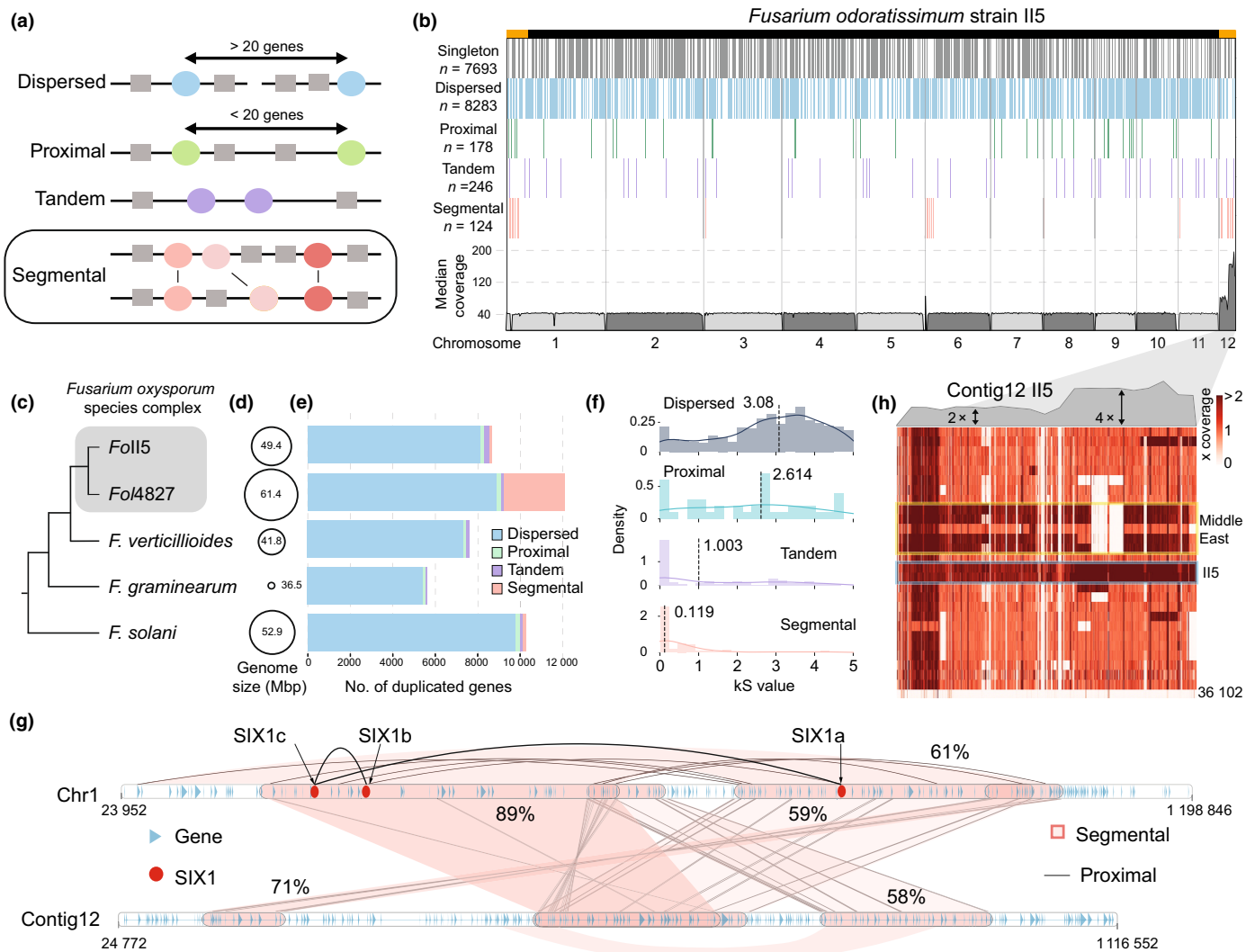## Recent segmental duplications drive accessory genome expansion in *Fusarium oxysporum*

We observed diverse ARs in *F. oxysporum* strains infecting banana with a high abundance of evolutionary young genes and genes that evolved via gene duplications (Fig. 4e). Duplications have been previously reported in *F. oxysporum*, including *F. oxysporum* strains infecting banana (Kistler *et al.*, 1995; Ma *et al.*, 2010; Vlaardingerbroek *et al.*, 2016; Armitage *et al.*, 2018; Li *et al.*, 2020). To better understand the role of gene duplications in the origin and diversification of ARs, we characterized gene duplications in *F. oxysporum*, using MCscanX (Wang *et al.*, 2012) and classified them into four categories (dispersed, proximal, tandem, and segmental; Fig. 6a). Remarkably, most genes in TR4 strain II5 (8831 genes) have been affected by gene duplication during their evolution (Figs 6b, S6). This high number of duplications include old gene duplications, recognized as distant homologs that acquired considerable number of mutations over time. When we apply a more stringent percent identity filtering, the number of duplicated genes is reduced, yet the number of segmental duplicated genes is least affected (Table S5), indicating that the segmental duplicated genes acquired less mutations and thus evolved more recently. We observed similar results for other *F. oxysporum* strains as well as for other *Fusarium* species; *F. verticillioides*, *F. graminearum*, and *F. solani* (syn. *Neocosmospora solani*), suggesting that duplications played a major role in *Fusarium* evolution (Fig. 6b). Interestingly, we observed that segmental duplicated genes occur specifically in ARs (72 out of 124 in II5) and sub-telomeres (48 out of 124 in strain II5) in all chromosome-level *F. oxysporum* assemblies (Figs 6b, S6). In *F. solani*, which contains accessory chromosomes analogous to those observed in *F. oxysporum* (Coleman *et al.*, 2009), 139 segmental duplicated genes were identified. By contrast, in *F. verticillioides* and *F. graminearum*, for which no accessory chromosomes have been described, no segmental duplicated genes are found (Fig. 6c–e), underscoring the link between accessory chromosomes and segmental duplications in the *Fusarium* genus. Moreover, segmental duplicated genes in strain II5 are upregulated during infection (8 d post inoculation; median log$_2$-fold change = 2.69; Fig. S7; Table S6), suggesting a role of these genes in the infection process. These duplications are possibly driven by TE activity (Wicker *et al.*, 2010; Faino *et al.*, 2016), and we observe that segmental duplications in strain II5 occur closer to transposable elements (median distance of 1734 kb, $P < 0.05$, two-sided Mann–Whitney U-test) than other duplication types (dispersed 7721 bp, proximal 7503 bp, and tandem 5941 bp), suggesting that TE activity might influence segmental duplications.

The highest number of segmental duplicated genes (2887) is observed in *Fol*4287, carrying the largest accessory genome (19 Mb, 30% of the total genome size; Figs 6d, S6) and 2366 (82%) of the segmental duplicated genes are in ARs. Interestingly, many segmental duplicated genes are shared between accessory chromosomes 3 and 6 (Fig. S8), as observed previously (Ma *et al.*, 2010), indicating that accessory chromosomes evolved by inter-chromosomal segmental duplication that drives the expansion of ARs in *Fusarium*.

Duplications of AR genes significantly affect effector repertoires as 336 out of 669 predicted effectors in II5 evolved via duplications, and nine effectors are part of segmental duplications. Among the segmental duplicated effectors is *SIX1*, an effector that is essential for full virulence to banana (Widinugraheni *et al.*, 2018). Three copies of *SIX1* (*SIX1a,b,c*) are present in the AR on chromosome 1 of the TR4 strains M1, 36 102, and II5. *SIX1a* and *SIX1b* are part of a segmentally duplicated block (Fig. 6g), whereas *SIX1c* is a proximal duplication that shares 82% amino acid sequence identity with *SIX1a* and 73% sequence identity with *SIX1b*. By contrast, the R1 strain CR1.1 contains only one copy of *SIX1*, resembling *SIX1c*. The duplicated blocks with *SIX1* in strain II5 are interspersed with nonduplicated genes (Fig. 6g), indicating that over time these blocks further diverged by gene gains and losses. Interestingly, the segmentally duplicated *SIX1* blocks also share similarity to a region on contig 12 (Fig. 6g), yet no copy of *SIX1* is present, and thus, *SIX1* is either lost or has been gained on chromosome 1 after the segmental duplication between chromosome 1 and contig 12 occurred. Segmental duplications therefore contribute to the evolution of virulence factors such as *SIX1*, which is essential for successful banana infection (Widinugraheni *et al.*, 2018).

To estimate the relative timing of the duplication events, we used the synonymous substitution rate (dS) between duplicated gene pairs as a proxy of time. Dispersed duplicated genes had the highest average dS value (3.08), suggesting that these are the most ancient duplicates, while gene pairs that arose via segmental duplication evolved more recently, with a significantly lower average dS value of 0.61 (Fig. 6f). To assess whether even more recent large-scale duplications are present in our global panel, we determined the read depth of all isolates sequenced with short reads against II5. Based on the genome-wide average read depth (c. 40× coverage), we determined that contig 12 has been entirely duplicated (read depth of c. 80× coverage), while a section (position c. 0.6–1.1 Mb; c. 160× coverage) occurs four times (Fig. 6h), suggesting additional copy number changes in strain II5 next to the previously observed segmental duplications (Fig. 6h). Similar copy number changes were observed when aligning the reads to a repeat masked genome assembly (Fig. S9), indicating that the coverage increase is not due to the expansion

Fig. 6 Segmental duplications are involved in the expansion of accessory regions (ARs) in *Fusarium oxysporum*. (a) Four different types of duplications are distinguished. (b) Segmental duplications (pink) occur mainly in the ARs in strain II5 (tropical race 4 (TR4); *Fusarium odoratissimum*). Sub-telomeric regions have more dispersed duplicated genes (blue) than the chromosomal arms. On average, the read coverage of II5 mapped onto the II5 reference genome is $40\times$, whereas contig 12 has a coverage between $80\times$ and $160\times$, suggesting copy number variations. (c) The schematic tree depicts the relationship of *Fusarium* species with (d) their genome sizes. (e) Extensive segmental duplications are only found in *F. oxysporum* and *Fusarium solani* and are virtually absent in *Fusarium verticillioides* and *Fusarium graminearum*. (f) The age of duplications was estimated from the number of synonymous substitutions (kS) between duplicated gene pairs. Segmental duplications occurred most recently (kS = 0.119). kS distributions differ significantly between all duplication types ($P < 0.05$, two-sided Mann–Whitney U-test). Dotted lines indicate the median values. (g) Genes encoding the virulence factor Secreted In Xylem (SIX) 1 (red) are part of a segmental duplication within chromosome 1. *SIX1a* and *SIX1b* are segmentally duplicated, and *SIX1c* is a proximal duplication sharing 82% amino acid sequence similarity to *SIX1a*. Light red blocks indicate segmental duplications; brown lines between the genes (blue arrows) highlight the genes involved in the segmental duplication. Not all genes in a block are part of the segmental duplication, for example a region on contig 12 is similar to the *SIX1* regions on chromosome 1, yet *SIX1* is absent on contig 12. (h) Sequencing read coverage of *F. oxysporum* strain II5 (TR4; *F. odoratissimum*) and three additional II5 strains from the Middle East (JV11, ISR5, and JV14) show a twofold increase relative to the genome-wide coverage. The other 26 strains do not show an increase in coverage, although smaller duplications are found throughout contig 12.

of repetitive sequences in the re-sequenced genome or to the collapse of repeats in our assembly. Additionally, the alignment of the long-read sequences also supports the observed copy number increased (Fig. S9). Interestingly, no copy number variation of the AR of chromosome 1 is observed. The partial copies of contig 12 possibly originate from additional duplications, yet it currently remains unclear whether these are attached to (core) chromosome(s), or whether they might give rise to an additional

smaller accessory chromosome. To resolve the origin of this duplicated region, we mapped the long reads to the assembly. We observed many clipped reads in this region; however, clipped regions did not align to another chromosome, suggesting that contig 12 or part of it are not attached to another chromosomal arm. Interestingly, the copy number increase in contig 12 can also be observed in three additional TR4 strains from the Middle East, but not in the other 26 TR4 strains (Fig. 6h). Collectively,

we demonstrate that the evolutionary dynamics of ARs in *F. oxysporum* is driven by recent segmental duplications.

Discussion

Many fungal plant pathogens evolved compartmentalized genomes with conserved core regions and variable ARs (Dong *et al.*, 2015; Torres *et al.*, 2020). Some members of the *F. oxysporum* species complex are important plant pathogens and are well-known to have extensive ARs that can encompass entire chromosomes (Sánchez-Vallet *et al.*, 2018), which have been linked to pathogenicity toward specific hosts. Here, we use a pangenomic framework to study the occurrence, composition, and evolution of ARs in a global collection of banana-infecting *F. oxysporum* strains, belonging to the several distinct lineages that have recently recognized as distinct species (Maryani *et al.*, 2019). *Fusarium oxysporum* strains carry up to 15 Mb of ARs that contain *in planta* induced effector genes as well as most *SIX* effectors. ARs are highly divergent among strains and races, and consequently, we cannot identify shared ARs or accessory chromosomes that can be linked to pathogenicity toward banana. We furthermore demonstrate that members of the *Fusarium* genus evolved by extensive duplications and that ARs specifically are shaped by recent segmental duplications in *F. oxysporum* and *F. solani*. Lastly, we uncover that an accessory contig in strain II5 underwent a recent copy number change, suggesting that these processes drive the emergence and evolution of ARs in *F. oxysporum*.

Fusarium wilt of banana is a major threat to food security as banana is a major staple crop in tropical and subtropical regions (Drenth & Kema, 2021; van Westerhoven *et al.*, 2022b). Compared with R1 strains, which caused epidemics in Gros Michel bananas in the 1900s, TR4 strains emerged more recently (1967) and caused the ongoing FWB pandemic in Cavendish plantations (van Westerhoven *et al.*, 2022b). TR4 strains are genetically highly similar, corroborating that a single clone that circumvents Cavendish resistance drives the ongoing pandemic (Ordonez *et al.*, 2015; van Westerhoven *et al.*, 2022b). By contrast, R1 strains are genetically diverse, which is possibly due to prolonged co-evolution of R1 strains with a plethora of genetically diverse banana varieties in the center of origin in Southeast Asia (Perrier *et al.*, 2011). Extensive genetic diversity among *F. oxysporum* strains that infect the same host is not uncommon (Baayen *et al.*, 2000; van Dam *et al.*, 2016; Henry *et al.*, 2021; Fayyaz *et al.*, 2023), yet their ARs and effector profiles are typically remarkably similar (van Dam *et al.*, 2016; Fokkens *et al.*, 2018; Brenes Guallar *et al.*, 2022). We, however, observe that the ARs and effector profile of R1 strains show extensive variation, underscoring their diversity and possibly suggesting that different mechanisms contribute to disease in banana (Henry *et al.*, 2021; Brenes Guallar *et al.*, 2022). *F. oxysporum* strains that infect strawberry have two distinct ARs, one that contributes to yellowing and the other to wilting (Henry *et al.*, 2021). Quantitative differences in banana corm-discoloration have been noted between R1 strains (Garcia-Bastidas, 2019), and our results suggest that this might relate to differences in effector profiles, yet

further functional assays are needed to elucidate the link between specific effectors and phenotypic variation.

We identify abundant and recent segmental duplications in *F. oxysporum* ARs, as well as a copy number change in contig 12 in TR4 strain II5. Previous studies have observed large-scale duplications and deletions in the ARs of *F. oxysporum* during *in vitro* chromosome transfer (Vlaardingerbroek *et al.*, 2016; Li *et al.*, 2020), yet our results show that duplication of ARs also occurs in *F. oxysporum* strains sampled directly from infected banana plants. Although variations in chromosome number or large-scale duplications typically come with a fitness cost (Todd *et al.*, 2017), it can also provide genetic variation necessary for adaptation, for example increasing virulence or fungicide resistance (Sionov *et al.*, 2010; Ropars *et al.*, 2018). We speculate that the observed large-scale duplications play a major role in the origin and evolution of ARs in *Fusarium*. In the wheat pathogen *Zymoseptoria tritici*, chromosomal duplication, together with breakage and fusion, shapes the evolutionary dynamics of ARs (Croll *et al.*, 2013; Möller *et al.*, 2018). Similar dynamics of chromosome duplication, fission, and fusion may underlie the recent segmental duplications observed in *F. oxysporum*. For example, the fusion of a (partially) duplicated contig 12 with the arm of chromosome 1 in strain II5 could explain the observed homology between the ARs, or alternatively, the fission of chromosome 1 could have led to the emergence of contig 12. The process underlying these chromosome dynamics remains unclear. *Fusarium oxysporum* has long been considered to evolve strictly asexually (Gordon & Martyn, 1997); however, the presence of an active sexual cycle in *F. oxysporum* populations has recently been proposed (McTaggart *et al.*, 2022; Fayyaz *et al.*, 2023) and we similarly observe reticulation between strains that caused FWB (Fig. S10), which supports an evolutionary history in which ancient or infrequent sexual recombination is followed by clonal expansion of selected lineages. In this scenario, incomplete chromosome pairing and nondisjunction during meiosis might drive copy number variation of chromosomes, segmental duplications, and ARs diversification. We show that ARs are similar within individual *F. oxysporum* lineages but are genetically distinct when comparing strains from genetically distant lineages. Individual lineages accumulate genomic changes over time, and this process might be sufficient to explain ARs diversity. If new *F. oxysporum* lineages indeed arise from a meiotic cycle followed by clonal expansion, recombination, gain, or loss of ARs during meiosis might further explain the diversity of ARs between lineages. Similarly, the absence of meiotic recombination during subsequent clonal expansion would explain the similarity of ARs within clonal lineages. However, we observe copy number variation within individual clonal lineages, for instance the duplication of contig 12 in some *F. oxysporum* strains, which demonstrates that this variation can occur independent of a meiotic cycle. Thus, chromosome dynamics in *F. oxysporum* require further elucidation, and so far, the mechanisms leading to chromosomal duplications are unclear; they might arise from nondisjunction during mitosis or meiosis (Fayyaz *et al.*, 2023), from incomplete chromosome loss following heterokaryon formation (Harrison *et al.*, 2014; Shahi *et al.*, 2016), or through consecutive horizontal

chromosome transfers when strains acquire the identical chromosome or parts of it several times.

In contrast to the structural variation in the ARs, the core chromosomes in *F. oxysporum* are remarkably stable. Although ARs are likely more tolerant to structural variation as they encode fewer essential genes, additional processes can influence genome stability. First of all, the accessory genome is enriched in transposable elements and the activity of these elements can contribute to structural variation including gene duplications (Faino *et al.*, 2016; Torres *et al.*, 2021; Stalder *et al.*, 2022; Wang *et al.*, 2022). Possibly the activity of transposable elements in the accessory genomic regions gave rise to the observed segmental duplications, for example, in plant genomes, large-scale duplications can arise through double-stranded break repair upon TE-induced double-stranded breaks (Wicker *et al.*, 2010; Wang *et al.*, 2022). Moreover, the presence of histone modifications such as trimethylation of lysine 27 on histone 3 (H3K27me3), a histone modification often enriched in ARs (Fokkens *et al.*, 2018; Möller *et al.*, 2019; Cook *et al.*, 2020; Torres *et al.*, 2020) in fungal plant pathogens, has been implied in genome instability (Seidl *et al.*, 2016; Möller *et al.*, 2019). H3K27me3 also occurs in the ARs of *F. oxysporum* infecting tomato (*Fol*4287), but can also be found in smaller core chromosomes ('fast-core' chromosomes, 9–11; Fokkens *et al.*, 2018) that do not seem to undergo segmental duplications or copy number changes, and thus, additional factors likely influence chromosome stability in *F. oxysporum*.

Extensive duplications similarly occur in other plant pathogens (Dutheil *et al.*, 2016; Müller *et al.*, 2019; Wyka *et al.*, 2021; Wacker *et al.*, 2023) and are thought to be important drivers in co-evolutionary arms races with their hosts. Understanding the evolution of ARs in *Fusarium oxysporum* can facilitate the discovery of new effector genes and provide insights into effector diversification. Knowledge of effector profiles is crucial for designing effective disease control strategies and supports the identification of durable resistances in crops (Vleeshouwers *et al.*, 2008).

Acknowledgements

Competing interests

None declared.

Author contributions

ACW collected data and performed the analyses. EMP, CAG, ECC, RH, AW, HS, FGB, NM, NO, AS, AD and GNT collected data. EHS, AF, KN and XSK contributed to the analyses. ACW and wrote the manuscript with input from HJGM, GHJK and MFS. All authors contributed to writing and editing the manuscript. GHJK, RS and MFS conceived and supervised the project. GHJK and MFS contributed equally to this work.

ORCID

Carolina Aguilera-Galvez https://orcid.org/0000-0001-7222-0676

Edgar Chavarro-Carrero https://orcid.org/0000-0003-0837-6830

Andre Drenth https://orcid.org/0000-0002-8510-1534

Alice Feurtey https://orcid.org/0000-0001-6315-9631

Fernando García-Bastidas https://orcid.org/0000-0001-6435-1174

Gert H. J. Kema https://orcid.org/0000-0002-2732-6911

Einar Martinez de la Parte https://orcid.org/0000-0002-2322-0888

Nani Maryani https://orcid.org/0000-0003-4599-9295

Harold J. G. Meijer https://orcid.org/0000-0002-0883-219X

Giuliana Nakasato-Tagami https://orcid.org/0000-0003-3240-335X

Nadia Ordóñez https://orcid.org/0000-0001-8440-0770

Michael F. Seidl https://orcid.org/0000-0002-5218-2083

Xiaoqian Shi-Kunne https://orcid.org/0000-0003-3747-3437

Eva H. Stukenbrock https://orcid.org/0000-0001-8590-3345

Ronny Swennen https://orcid.org/0000-0002-5258-9043

Anouk C. van Westerhoven https://orcid.org/0000-0001-7480-2784

Data availability

The data that support the findings of this study are openly available in the Sequence Read Archive at https://www.ncbi.nlm.nih.gov/sra. The data sequenced in this study are available under accession no.: PRJNA979975 the accession nos. for data from additional public projects are listed in Table S1.

References

Armitage AD, Taylor A, Sobczyk MK, Baxter L, Greenfield BPJ, Bates HJ, Wilson F, Jackson AC, Ott S, Harrison RJ *et al.* **2018.** Characterisation of pathogen-specific regions and novel effector candidates in *Fusarium oxysporum* f. sp. *cepae*. *Scientific Reports* **8**: 13530.

Armstrong J, Hickey G, Diekhans M, Fiddes IT, Novak AM, Deran A, Fang Q, Xie D, Feng S, Stiller J *et al.* **2020.** Progressive cactus is a multiple-genome aligner for the thousand-genome era. *Nature* **587**: 246–251.

Ayukawa Y, Asai S, Gan P, Tsushima A, Ichihashi Y, Shibata A, Komatsu K, Houterman PM, Rep M, Shirasu K *et al.* **2021.** A pair of effectors encoded on a conditionally dispensable chromosome of *Fusarium oxysporum* suppress host-specific immunity. *Communications Biology* **4**: 1–12.

Baayen RP, O'Donnell K, Bonants PJM, Cigelnik E, Kroon LPNM, Roebroeck EJA, Waalwijk C. 2000. Gene genealogies and AFLP analyses in the *Fusarium oxysporum* complex identify monophyletic and nonmonophyletic formae speciales causing wilt and rot disease. *Phytopathology* **90**: 891–900.

Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD *et al.* **2012.** SPADES: a new genome

assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* **19**: 455–477.

Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* **34**: 525–527.

Brenes Guallar MA, Fokkens L, Rep M, Berke L, van Dam P. 2022. *Fusarium oxysporum* effector clustering version 2: an updated pipeline to infer host range. *Frontiers in Plant Science* **13**: 1012688.

Chernomor O, von Haeseler A, Minh BQ. 2016. Terrace aware data structure for phylogenomic inference from supermatrices. *Systematic Biology* **65**: 997–1008.

Coleman JJ, Rounsley SD, Rodriguez-Carres M, Kuo A, Wasmann CC, Grimwood J, Schmutz J, Taga M, White GJ, Zhou S et al. 2009. The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genetics* **5**: e1000618.

Cook DE, Kramer HM, Torres DE, Seidl MF, Thomma BPHJ. 2020. A unique chromatin profile defines adaptive genomic regions in a fungal plant pathogen. *eLife* **9**: e62208.

Cook DE, Mesarich CH, Thomma BPHJ. 2015. Understanding plant immunity as a surveillance system to detect invasion. *Annual Review of Phytopathology* **53**: 541–563.

Coombe L, Li JX, Lo T, Wong J, Nikolic V, Warren RL, Birol I. 2021. LONGSTITCH: high-quality genome assembly correction and scaffolding using long reads. *BMC Bioinformatics* **22**: 534.

Croll D, Zala M, McDonald BA. 2013. Breakage-fusion-bridge cycles and large insertions contribute to the rapid evolution of accessory chromosomes in a fungal pathogen. *PLoS Genetics* **9**: e1003567.

Cuomo CA, Güldener U, Xu J-R, Trail F, Turgeon BG, Di Pietro A, Walton JD, Ma L-J, Baker SE, Rep M et al. 2007. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science* **317**: 1400–1402.

van Dam P, Fokkens L, Ayukawa Y, van der Gragt M, ter Horst A, Brankovics B, Houterman PM, Arie T, Rep M. 2017. A mobile pathogenicity chromosome in *Fusarium oxysporum* for infection of multiple cucurbit species. *Scientific Reports* **7**: 9042.

van Dam P, Fokkens L, Schmidt SM, Linmans JHJ, Kistler HC, Ma L-J, Rep M. 2016. Effector profiles distinguish formae speciales of *Fusarium oxysporum*. *Environmental Microbiology* **18**: 4087–4102.

Derelle R, Philippe H, Colbourne JK. 2020. BROCCOLI: combining phylogenetic and network analyses for orthology assignment. *Molecular Biology and Evolution* **37**: 3389–3396.

Dong S, Raffaele S, Kamoun S. 2015. The two-speed genomes of filamentous pathogens: waltz with plants. *Current Opinion in Genetics & Development* **35**: 57–65.

Drenth A, Kema G. 2021. The vulnerability of bananas to globally emerging disease threats. *Phytopathology* **111**: 2146–2161.

Dutheil JY, Mannhaupt G, Schweizer G, Sieber CMK, Münsterkötter M, Güldener U, Schirawski J, Kahmann R. 2016. A tale of genome compartmentalization: the evolution of virulence clusters in smut fungi. *Genome Biology and Evolution* **8**: 681–704.

Edel-Hermann V, Lecomte C. 2019. Current status of *Fusarium oxysporum* formae speciales and races. *Phytopathology* **109**: 512–530.

Emms DM, Kelly S. 2019. ORTHOFINDER: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**: 1–14.

Faino L, Seidl MF, Shi-Kunne X, Pauper M, van den Berg GCM, Wittenberg AHJ, Thomma BPHJ. 2016. Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. *Genome Research* **26**: 1091–1100.

FAO. 2020. *Fruit and vegetables – your dietary essentials: the international year of fruits and vegetables, 2021, background paper*. Rome, Italy: FAO.

Fayyaz A, Robinson G, Chang PL, Bekele D, Yimer S, Carrasquilla-Garcia N, Negash K, Surendrarao A, von Wettberg EJB, Kemal S-A et al. 2023. Hiding in plain sight: genome-wide recombination and a dynamic accessory genome drive diversity in *Fusarium oxysporum* f. sp. *ciceris*. *Proceedings of the National Academy of Sciences, USA* **120**: e2220570120.

Fiddes IT, Armstrong J, Diekhans M, Nachtweide S, Kronenberg ZN, Underwood JG, Gordon D, Earl D, Keane T, Eichler EE et al. 2018. Comparative Annotation Toolkit (CAT)—simultaneous clade and personal genome annotation. *Genome Research* **28**: 1029–1038.

Fokkens L, Shahi S, Connolly LR, Stam R, Schmidt SM, Smith KM, Freitag M, Rep M. 2018. The multi-speed genome of *Fusarium oxysporum* reveals association of histone modifications with sequence divergence and footprints of past horizontal chromosome transfer events. *bioRxiv*. doi: 10.1101/465070.

Galagan JE, Selker EU. 2004. RIP: the evolutionary cost of genome defense. *Trends in Genetics* **20**: 417–423.

Garcia-Bastidas F. 2019. *Panama disease in banana: spread, screens and genes*. PhD thesis, Wageningen University and Research, Wageningen, the Netherlands.

Garcia-Bastidas F, Quintero-Vargas JC, Ayala-Vasquez M, Schermer T, Seidl MF, Santos-Paiva M, Noguera AM, Aguilera-Galvez C, Wittenberg A, Hofstede R et al. 2020. First report of Fusarium wilt tropical race 4 in Cavendish bananas caused by *Fusarium odoratissimum* in Colombia. *Plant Disease* **104**: 994.

García-Bastidas FA, Ordóñez N, Konkol J, Al-Qasim M, Naser Z, Abdelwali M, Salem N, Waalwijk C, Ploetz RC, Kema GHJ. 2014. First report of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 associated with Panama disease of banana outside Southeast Asia. *Plant Disease* **98**: 694.

Gordon TR, Martyn RD. 1997. The evolutionary biology of *Fusarium oxysporum*. *Annual Review of Phytopathology* **35**: 111–128.

Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**: 2811–2812.

Guo L, Han L, Yang L, Zeng H, Fan D, Zhu Y, Feng Y, Wang G, Peng C, Jiang X et al. 2014. Genome and transcriptome analysis of the fungal pathogen *Fusarium oxysporum* f. sp. *cubense* causing banana vascular wilt disease. *PLoS ONE* **9**: e95543.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**: 1072–1075.

Harrison BD, Hashemi J, Bibi M, Pulver R, Bavli D, Nahmias Y, Wellington M, Sapiro G, Berman J. 2014. A tetraploid intermediate precedes aneuploid formation in yeasts exposed to fluconazole. *PLoS Biology* **12**: e1001815.

Henry PM, Pincot DDA, Jenner BN, Borrero C, Aviles M, Nam M-H, Epstein L, Knapp SJ, Gordon TR. 2021. Horizontal chromosome transfer and independent evolution drive diversification in *Fusarium oxysporum* f. sp. *fragariae*. *New Phytologist* **230**: 327–340.

Hickey G, Monlong J, Ebler J, Novak AM, Eizenga JM, Gao Y, Marschall T, Li H, Paten B. 2023. Pangenome graph construction from genome alignments with Minigraph-Cactus. *Nature Biotechnology*. 1–11.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on Fast Fourier transform. *Nucleic Acids Research* **30**: 3059–3066.

Kistler HC, Benny U, Boehm EW, Katan T. 1995. Genetic duplication in *Fusarium oxysporum*. *Current Genetics* **28**: 173–176.

Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. 2017. CANU: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Research* **27**: 722–736.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. *Genome Biology* **5**: 1–9.

Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*: 1303.3997v2.

Li J, Fokkens L, Conneely LJ, Rep M. 2020. Partial pathogenicity chromosomes in *Fusarium oxysporum* are sufficient to cause disease and can be horizontally transferred. *Environmental Microbiology* **22**: 4985–5004.

Lindenbaum P. 2015. JVARKIT: Java-based utilities for bioinformatics. *Figshare*. doi: 10.6084/m9.figshare.1425030.

Ma L-J, van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B et al. 2010. Comparative genomics reveals mobile pathogenicity chromosomes in Fusarium. *Nature* **464**: 367–373.

Maryani N, Lombard L, Poerba YS, Subandiyah S, Crous PW, Kema GHJ. 2019. Phylogeny and genetic diversity of the banana Fusarium wilt pathogen *Fusarium oxysporum* f. sp. *cubense* in the Indonesian centre of origin. *Studies in Mycology* **92**: 155–194.

Maymon M, Sela N, Shpatz U, Galpaz N, Freeman S. 2020. The origin and current situation of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 in Israel and the Middle East. *Scientific Reports* **10**: 1590.

McTaggart AR, James TY, Idnurm A, Park RF, Shuey LS, Demers MNK, Aime MC. 2022. Sexual reproduction is the null hypothesis for life cycles of rust fungi. *PLoS Pathogens* 18: e1010439.

Mesny F, Miyauchi S, Thiergart T, Pickel B, Atanasova L, Karlsson M, Hüttel B, Barry KW, Haridas S, Chen C *et al.* 2021. Genetic determinants of endophytism in the Arabidopsis root mycobiome. *Nature Communications* 12: 7227.

Möller M, Habig M, Freitag M, Stukenbrock EH. 2018. Extraordinary genome instability and widespread chromosome rearrangements during vegetative growth. *Genetics* 210: 517–529.

Möller M, Schotanus K, Soyer JL, Haueisen J, Happ K, Stralucke M, Happel P, Smith KM, Connolly LR, Freitag M *et al.* 2019. Destabilization of chromosome structure by histone H3 Lysine 27 methylation. *PLoS Genetics* 15: e1008093.

Möller M, Stukenbrock EH. 2017. Evolution and genome architecture in fungal plant pathogens. *Nature Reviews Microbiology* 15: 756–771.

Müller MC, Praz CR, Sotiropoulos AG, Menardo F, Kunz L, Schudel S, Oberhänsli S, Poretti M, Wehrli A, Bourras S *et al.* 2019. A chromosome-scale genome assembly reveals a highly dynamic effector repertoire of wheat powdery mildew. *New Phytologist* 221: 2176–2189.

O'Donnell K, Kistler HC, Cigelnik E, Ploetz RC. 1998. Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. *Proceedings of the National Academy of Sciences, USA* 95: 2044–2049.

Ordonez N, Seidl MF, Waalwijk C, Drenth A, Kilian A, Thomma BPHJ, Ploetz RC, Kema GHJ. 2015. Worse comes to worst: bananas and Panama disease—when plant and pathogen clones meet. *PLoS Pathogens* 11: e1005197.

Palmer J, Stajich J. 2019. *nextgenusfs/funannotate: funannotate v1.5.3.* doi: 10.5281/zenodo.1134477 [accessed 20 November 2021].

Perrier X, De Langhe E, Donohue M, Lentfer C, Vrydaghs L, Bakry F, Carreel F, Hippolyte I, Horry J-P, Jenny C *et al.* 2011. Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proceedings of the National Academy of Sciences, USA* 108: 11311–11318.

Ploetz RC. 2006. Fusarium wilt of banana is caused by several pathogens referred to as *Fusarium oxysporum* f. sp. *cubense*. *Phytopathology* 96: 653–656.

Ploetz RC. 2015. Management of Fusarium wilt of banana: a review with special reference to tropical race 4. *Crop Protection* 73: 7–15.

Raffaele S, Kamoun S. 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews Microbiology* 10: 417–430.

Rep M, Dekker HL, Vossen JH, de Boer AD, Houterman PM, Speijer D, Back JW, de Koster CG, Cornelissen BJC. 2002. Mass spectrometric identification of isoforms of PR proteins in xylem sap of fungus-infected tomato. *Plant Physiology* 130: 904–917.

Ropars J, Maufrais C, Diogo D, Marcet-Houben M, Perin A, Sertour N, Mosca K, Permal E, Laval G, Bouchier C *et al.* 2018. Gene flow contributes to diversification of the major fungal pathogen *Candida albicans*. *Nature Communications* 9: 2253.

Rovenich H, Boshoven JC, Thomma BPHJ. 2014. Filamentous pathogen effector functions: of pathogens, hosts and microbiomes. *Current Opinion in Plant Biology* 20: 96–103.

Sánchez-Vallet A, Fouché S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D. 2018. The genome biology of effector gene evolution in filamentous plant pathogens. *Annual Review of Phytopathology* 56: 21–40.

Seidl MF, Cook DE, Thomma BPHJ. 2016. Chromatin biology impacts adaptive evolution of filamentous plant pathogens. *PLoS Pathogens* 12: e1005920.

Seidl MF, Thomma BPHJ. 2017. Transposable elements direct the coevolution between plants and microbes. *Trends in Genetics* 33: 842–851.

Shahi S, Beerens B, Bosch M, Linmans J, Rep M. 2016. Nuclear dynamics and genetic rearrangement in heterokaryotic colonies of *Fusarium oxysporum*. *Fungal Genetics and Biology* 91: 20–31.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31: 3210–3212.

Sionov E, Lee H, Chang YC, Kwon-Chung KJ. 2010. *Cryptococcus neoformans* overcomes stress of azole drugs by formation of disomy in specific multiple chromosomes. *PLoS Pathogens* 6: e1000848.

Stalder L, Oggenfuss U, Mohd-Assaad N, Croll D. 2022. The population genetics of adaptation through copy number variation in a fungal plant pathogen. *Molecular Ecology* 32: 2443–2460.

Stanke M, Morgenstern B. 2005. AUGUSTUS: a web server for gene prediction in eukaryotes that allows user-defined constraints. *Nucleic Acids Research* 33: W465–W467.

Su H, Hwang S-C, Ko W-H. 1986. Fusarial wilt of cavendish bananas in Taiwan. *American Phytopathological Society* 70: 814.

Suyama M, Torrents D, Bork P. 2006. PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Research* 34: W609–W612.

Todd RT, Forche A, Selmecki A. 2017. Ploidy variation in fungi: polyploidy, aneuploidy, and genome evolution. *Microbiology Spectrum* 5: 9.

Torres DE, Oggenfuss U, Croll D, Seidl MF. 2020. Genome evolution in fungal plant pathogens: looking beyond the two-speed genome model. *Fungal Biology Reviews* 34: 136–143.

Torres DE, Thomma BPHJ, Seidl MF. 2021. Transposable elements contribute to genome dynamics and gene expression variation in the fungal plant pathogen *Verticillium dahliae*. *Genome Biology and Evolution* 13: evab135.

Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J *et al.* 2013. From FASTQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics* 43: 11.10.1–11.10.33.

Vaser R, Sović I, Nagarajan N, Šikić M. 2017. Fast and accurate *de novo* genome assembly from long uncorrected reads. *Genome Research* 27: 737–746.

Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J *et al.* 2020. SCIPY 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods* 17: 261–272.

Vlaardingerbroek I, Beerens B, Schmidt SM, Cornelissen BJC, Rep M. 2016. Dispensable chromosomes in *Fusarium oxysporum* f. sp. *lycopersici*. *Molecular Plant Pathology* 17: 1455–1466.

Vleeshouwers VGAA, Rietman H, Krenek P, Champouret N, Young C, Oh S-K, Wang M, Bouwmeester K, Vosman B, Visser RGF *et al.* 2008. Effector genomics accelerates discovery and functional profiling of potato disease resistance and *Phytophthora infestans* avirulence genes. *PLoS ONE* 3: e2875.

Wacker T, Helmstetter N, Wilson D, Fisher MC, Studholme DJ, Farrer RA. 2023. Two-speed genome evolution drives pathogenicity in fungal pathogens of animals. *Proceedings of the National Academy of Sciences, USA* 120: e2212633120.

Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK *et al.* 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* 9: e112963.

Wang X, Yan X, Hu Y, Qin L, Wang D, Jia J, Jiao Y. 2022. A recent burst of gene duplications in Triticeae. *Plant Communications* 3: 100268.

Wang Y, Tang H, DeBarry JD, Tan X, Li J, Wang X, Lee T, Jin H, Marler B, Guo H *et al.* 2012. MCSCANX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research* 40: e49.

Warmington RJ, Kay W, Jeffries A, O'Neill P, Farbos A, Moore K, Bebber DP, Studholme DJ. 2019. High-quality draft genome sequence of the causal agent of the current Panama disease epidemic. *Microbiology Resource Announcements* 8: e00904-19.

van Westerhoven AC, Meijer H, Houdijk J, Martínez de la Parte E, Matabuana EL, Seidl M, Kema GHJ. 2022a. Dissemination of Fusarium wilt of banana in Mozambique caused by *Fusarium odoratissimum* tropical race 4. *Plant Disease* 107: 628–632.

van Westerhoven AC, Meijer HJG, Seidl MF, Kema GHJ. 2022b. Uncontained spread of Fusarium wilt of banana threatens African food security. *PLoS Pathogens* 18: e1010769.

Wicker T, Buchmann JP, Keller B. 2010. Patching gaps in plant genomes results in gene movement and erosion of colinearity. *Genome Research* 20: 1229–1237.

Widinugraheni S, Niño-Sánchez J, Van Der Does HC, Van Dam P, García-Bastidas FA, Subandiyah S, Meijer HJG, Kistler HC, Kema GHJ, Rep M. 2018. A SIX1 homolog in *Fusarium oxysporum* f. sp. *cubense* tropical race 4 contributes to virulence towards Cavendish banana. *PLoS ONE* 13: e0205896.

Wolf YI, Novichkov PS, Karev GP, Koonin EV, Lipman DJ. 2009. The universal distribution of evolutionary rates of genes and distinct characteristics of eukaryotic genes of different apparent ages. *Proceedings of the National Academy of Sciences, USA* 106: 7273–7280.

Wyka SA, Mondo SJ, Liu M, Dettman J, Nalam V, Broders KD. 2021. Whole-genome comparisons of ergot fungi reveals the divergence and evolution of species within the genus claviceps are the result of varying mechanisms driving genome evolution and host range expansion. *Genome Biology and Evolution* 13: evaa267.

Yang Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution* 24: 1586–1591.

Zhang Y, Yang H, Turra D, Zhou S, Ayhan DH, DeIulio GA, Guo L, Broz K, Wiederhold N, Coleman JJ et al. 2020. The genome of opportunistic fungal pathogen *Fusarium oxysporum* carries a unique set of lineage-specific chromosomes. *Communications Biology* 3: 1–12.

Zheng S-J, García-Bastidas FA, Li X, Zeng L, Bai T, Xu S, Yin K, Li H, Fu G, Yu Y et al. 2018. New geographical insights of the latest expansion of *Fusarium oxysporum* f. sp. *cubense* tropical race 4 into the greater Mekong subregion. *Frontiers in Plant Science* 9: 457.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Chromosomes 9, 10, and 11 of strain II5 (TR4; *Fusarium odoratissimum*) share few genomic regions with other *Fusarium oxysporum* strains infecting banana and have more polymorphic sites.

Fig. S2 Nucleotide alignments between accessory regions of the chromosome-level *Fusarium oxysporum* genome assemblies identify diverse accessory regions.

Fig. S3 Fraction of shared accessory regions in different sub-groups of *Fusarium oxysporum* strains infecting banana.

Fig. S4 Accessory regions are shared between genetically related *Fusarium oxysporum* strains.

Fig. S5 Functional enrichment of core, softcore, and accessory genes.

Fig. S6 Duplication types found in the core, softcore, and accessory regions in the seven chromosome-level genome assemblies of *Fusarium oxysporum* strains infecting banana and the *Fusarium oxysporum* strain infecting tomato (Fol4287).

Fig. S7 Distribution of $\log_2$-fold gene expression changes per gene in strain II5 (TR4; *Fusarium odoratissimum*), comparing gene expression between *in vitro* growth with 8 d post inoculation of Cavendish banana.

Fig. S8 Segmental duplications between accessory regions.

Fig. S9 Alignment of short reads against the repeat masked strain II5 (TR4; *Fusarium odoratissimum*) and long reads against the unmasked strain II5 (TR4; *Fusarium odoratissimum*).

Fig. S10 Splitstree network shows indications of reticulation that might suggest recombination between some *Fusarium oxysporum* strains infecting banana.

Methods S1 Supporting materials and methods.

Table S1 Genomes of *Fusarium oxysporum* strains infecting banana analyzed in this study.

Table S2 Fungal genomes used in our analyses to estimate the gene age.

Table S3 Statistics of the long-read nanopore sequencing data of seven *Fusarium oxysporum* strains.

Table S4 Genome assembly statistics of seven *Fusarium oxysporum* strains sequenced with nanopore long-read sequencing technology.

Table S5 Number of duplicated genes as reported compared with the number of duplicated genes using a stricter filtering approach.

Table S6 Gene expression statistics per gene duplication category in Fusarium odoratissimum strain II5, both *in vitro* and 8 d post infection on Cavendish banana.

Please note: Wiley is not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.