

# Reference-free identification and pangenome analysis of accessory chromosomes in a major fungal plant pathogen

Anouk C. van Westerhoven<sup>1,2</sup>, Like Fokkens<sup>2</sup>, Kyran Wissink<sup>1</sup>, Gert H.J. Kema<sup>2</sup>, Martijn Rep<sup>3</sup>, Michael F. Seidl<sup>1,\*</sup>

<sup>1</sup>Theoretical Biology and Bioinformatics, Utrecht University, Padualaan 8, 3583CH, Utrecht, the Netherlands

<sup>2</sup>Laboratory of Phytopathology, Wageningen University & Research, Droevendaalsesteeg 1, 6708PB, Wageningen, the Netherlands

<sup>3</sup>Molecular Plant Pathology, Swammerdam Institute of Life Sciences, University of Amsterdam, 1090GE, Amsterdam, the Netherlands

\*To whom correspondence should be addressed. Email: m.f.seidl@uu.nl

## Abstract

Accessory chromosomes, found in some but not all individuals of a species, play an important role in pathogenicity and host specificity in fungal plant pathogens. However, their variability complicates reference-based analysis, especially when these chromosomes are missing in the reference genome. Pangenome variation graphs offer a reference-free alternative for studying these chromosomes. Here, we constructed a pangenome variation graph for 73 diverse *Fusarium oxysporum* genomes, a major fungal plant pathogen with a compartmentalized genome that includes conserved core as well as variable accessory chromosomes. To obtain insights into accessory chromosome dynamics, we first constructed a chromosome similarity network using all-vs-all similarity mapping. We identified eleven core chromosomes conserved across all strains and a substantial number of highly variable accessory chromosomes. Some of these accessory chromosomes are host-specific and likely play a role in determining host range. Using a k-mer based approach, we further identified the presence of these accessory chromosomes in all available (581) *F. oxysporum* assemblies and corroborated the occurrence of host-specific accessory chromosomes. To further analyze the evolution of chromosomes in *F. oxysporum*, we constructed a pangenome variation graph per group of homologous chromosomes. This reveals that accessory chromosomes are composed of different stretches of accessory regions, and possibly rearrangements between accessory regions gave rise to these mosaic accessory chromosomes. Furthermore, we show that accessory chromosomes are likely horizontally transferred in natural populations. Our findings demonstrate that a pangenome variation graph is a powerful approach to elucidate the evolutionary dynamics of accessory chromosomes in *F. oxysporum*, which is not only a useful resource for *Fusarium* but also provides a framework for similar analyses in other species containing accessory chromosomes.

## Introduction

The number of chromosomes of individuals in a single species is generally conserved. However, in various plants, animals, oomycetes, and fungi, a variable number of chromosomes is identified within a species [1]. These accessory chromosomes often encode very few genes, with low transcriptional activity, and little effect on the phenotype [1, 2]. In contrast, accessory chromosomes in various fungal plant pathogens encode genes with important roles in host infection [3–7]. These accessory chromosomes are often highly variable across individuals and show extensive presence/absence variation [6, 8, 9]. The presence of pathogenicity genes on these variable accessory chromosomes separates pathogenicity genes from house-keeping genes and has been thought to facilitate rapid adaptation to changing environments as well as the host immune system [5, 10].

The increasing availability of continuous, chromosome-level genome assemblies enabled the identification of accessory chromosomes in numerous fungi [4, 6, 11–18] and allows for comparative genomic approaches that help to elucidate the evolution of accessory chromosomes [6, 17–19]. However, the analysis of accessory chromosomes is limited by reference-based analyses, where the variation of accessory

chromosomes can only be determined by comparisons to a single reference genome. Because of the variability of accessory regions, the reference genome often lacks specific accessory regions that will not be analyzed. To capture the complete variation in a collection of genome sequences, including the accessory chromosomes, reference-free analysis strategies are needed. Recent advances in pangenome analysis methods can now be used to capture all genetic variation in a species [20, 21]. One approach is the pangenome variation graph [22, 23], where nodes represent sequences and edges connect adjacent nodes; each path through the graph represents a single genome. This graph can be used as a framework to call variants in a population, annotate genes, and analyze chromosome structure [24–26].

Thus far, pangenome variation graphs have been mainly constructed to analyze the human genome as well as some plant genomes [24, 27, 28]. Only recently a pangenome variation graph for the oomycete plant pathogen *Peronospora effusa* has been constructed to study processes driving genome evolution [26]. In this study the construction of the pangenome graph was guided by a reference genome [23, 25], which can offer important insights in organisms where the chromosome structure is largely conserved. However,

Received: December 12, 2024. Revised: February 19, 2025. Editorial Decision: March 11, 2025. Accepted: March 14, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of NAR Genomics and Bioinformatics.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License

(<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact reprints@oup.com for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact journals.permissions@oup.com.

pangenome variation graphs are more challenging to construct for fungal genomes that typically contain more chromosomal variations, especially in accessory chromosomes [8, 10, 29], and fungal pangenome variation graphs have not been published. Due to the lack of co-linearity, the analysis of accessory chromosomes using pangenome variation graphs is not straightforward. Grouping the chromosomes into homologous groups [30] prior to construction of a variation graph is a promising strategy to analyze chromosomes in dynamic fungal species. However, it is unknown how homologous groups should be constructed for species carrying accessory chromosomes.

The fungal plant pathogen *Fusarium oxysporum* can infect various important crops [31–33] and is known to have a compartmentalized genome where accessory regions can span entire chromosomes or can be attached to conserved core chromosomes [4, 6, 18, 29, 34–36]. The core chromosomes are largely co-linear with limited large-scale variation, such as deletions, insertions, inversions, or translocations [4, 6, 37]. In contrast, accessory chromosomes are diverse even between strains infecting the same host [6, 18]. Importantly, horizontal transfer of accessory chromosomes between strains can transfer pathogenicity between strains in the lab [4, 35, 36]. Thus far, the genome structure and diversity of accessory chromosomes in *F. oxysporum* have mainly been analyzed for strains infecting a specific host in comparison to the widely studied reference genome *F. oxysporum* f.sp. *lycopersici* strain 4287 (Fol4287) infecting tomato, which is known to have five accessory chromosomes [4]. Little is known about occurrence, the frequency, and the variation of accessory chromosomes across the *F. oxysporum* species complex.

The presence of a compartmentalized genome together with the availability of various continuous genome assemblies make *F. oxysporum* a suitable model to analyze the dynamics of accessory chromosomes using a pangenome variation graph. Here, we constructed a pangenome variation graph to analyze the evolution of accessory chromosomes in *F. oxysporum*. Some of these accessory chromosomes were specific to strains infecting the same host, suggesting that these play a role in determining host specificity. Furthermore, accessory chromosomes were composed of different combinations of accessory material and likely evolved through recombination of accessory chromosomes. These findings offer insights into the evolution of accessory chromosomes and show how pangenome variation graphs can be applied to analyze chromosome dynamics.

## Materials and methods

### DNA isolation and whole-genome sequencing and assembly

For strains Fol005, Fol007, Fom011, Fom014, Fom021, Fom024, and Fom025, strains were taken from glycerol stock, grown on PDA for 3 days at 25°C, after which 100 ml of NO<sub>3</sub> was inoculated with an overgrown agar plug. Samples were grown in this liquid culture for 5–7 days at 25°C and 150 rpm. Mycelium was harvested by filtering through miracloth, stored in liquid N<sub>2</sub>, and freeze-dried overnight. Per sample, DNA was isolated from ~250 mg of ground mycelium using multiple rounds of phenol-chloroform extractions, two rounds of chloroform extractions, and was pre-

cipitated twice. Quality and quantity of DNA were checked with Nanodrop, Qubit, and agarose gels. DNA samples were sequenced at KeyGene on three PromethION FLO\_PRO002 cells where basecalling was performed with the high-accuracy model in MinKNOW 4.2.6, and reads with quality  $q \geq 7$  were selected. Adapters were trimmed with poreChop, and reads were filtered based on their quality and length with Filt-long (version 0.2.0, `-min_length 1000`, `-min_mean_q 80` `-min_window_q 70`). Per strain, the 80% longest reads were selected and assembled with Flye (v2.8). Strains Fol029 and FolMN25 were sequenced and assembled according to van Dam *et al.*

### Genome selection

*Fusarium oxysporum* genome assemblies were downloaded from NCBI (July 2024), retaining genome assemblies with <100 contigs and an N50 >2 Mb. In addition to the genomes from NCBI, we included 12 previously assembled whole-genome assemblies (Supplementary Table S1). Not all genomes were assembled and processed in the same way. To account for this heterogeneity, we applied a homogeneous filtering strategy and excluded fragmented chromosomes. We filtered contigs <50 kb, removing between 0 and 80 contigs per genome assembly (Supplementary Table S1). We then accessed the genome assembly quality using Quast version 5.0.2 [38] and determined the presence of telomeric repeats using tidk version 0.2.41 [39]. Furthermore, we determined the pairwise average nucleotide identity for all the whole genome assemblies using Pyani (<https://github.com/widdowquinn/pyani>) and constructed a phylogenetic tree of the strains using conserved single-copy BUSCO genes with the BUSCO phylogenomic pipeline, available at [https://github.com/jamiemcg/BUSCO\\_phylogenomics](https://github.com/jamiemcg/BUSCO_phylogenomics).

### Network construction

To detect groups of homologous chromosomes, we first mapped all contigs to each other using wfmash version 0.10.3 [40]. To determine the optimal settings for *F. oxysporum*, we tried different parameters for the length of mapped segments (-s, 2.5, 5, and 10 kb) and the block length of the chained segments (-l, from 5, 10, 25, 50, and 100 kb); increasing segment and block length also increases the length of co-linear blocks between chromosomes. The resulting all-vs-all homology mappings between chromosomes were used to construct a chromosome network. In addition to this network, we also created a network with splitting of chromosomes disabled (wfmash flag -N), removing partial chromosome mappings. These two all-vs-all homology mappings were treated similar for downstream comparisons.

Chromosomes communities were detected in the all-vs-all homology mappings using the Leiden algorithm [41] as described by [24, 42]. The chromosome homology networks were visualized using Gephi version 0.10 [43], and the network layout was calculated using the YiFan Hu algorithm. Community statistics were determined using iGraph Python module version 0.10.8 [44]. To understand the chromosome network in relation to the *F. oxysporum* reference genome assembly Fol4287, we plotted the communities assigned to chromosomes in the Fol4287 using pafR version 0.0.2 (<https://dwinter.github.io/pafR/>).

## Splitting of chromosomes

Before we could construct the pangenome graphs per community, we needed to ensure that each community represent homologous chromosomes. However, in some cases one chromosome maps to two different communities, for example due to chromosome fusion or errors during genome assembly. To include the chromosomes that map to two different communities, we split these chromosomes at the breakpoint and assigned chromosomal regions to the corresponding communities. Chromosomes were split into two different communities when a consecutive stretch of 1 Mb mapped to a different community and when this part occurred at the flank of the chromosome, and thus we only split large co-linear fractions of the chromosome, preventing us to artificially splitting chromosomes into too small sections. Some windows in the chromosome did not map to any community. To maintain the chromosome structure, these unmapped regions were ignored when we determine the flanks of a chromosome and, these unmapped regions were assigned to the new community when they flank the split region.

## Pangenome variation graph construction

We constructed a pangenome variation graph per community detected in the all-vs-all chromosome network. The pangenome graph was constructed with the PanGenomeGraphBuilder version 0.3.0 (PGGB, [30]) with the following parameters: a segment length of 2500 bp (-s), a block length of 50 000 bp (-l), a percent identity threshold of 95% (-p), and 72 haplotypes (n-1). The generated pangenome variation graph was then sorted and visualized using the optimized dynamic genome/graph implementation version 0.8.6.0 (odgi, [24]). The general graph statistics were calculated with odgi stats (-S), the pairwise genome similarity was assessed with odgi similarity, and the percentages of core and accessory nodes were determined using odgi stats (-a). To assess structural variation, the mean inversion rate was calculated using odgi bin -w 50 000 bp, and node presence/absence patterns were visualized with odgi viz. Subgraphs for different communities were combined using odgi squeeze to form the final *F. oxysporum* pangenome variation graph. Finally, this joined *F. oxysporum* variation graph was used to evaluate the pangenome growth curve with Panancus (version 0.2.3, [48]).

## Short read genome assemblies and matching to the pangenome graph

To include a larger variety of *F. oxysporum* strains, we compared the k-mer profile of 588 fragmented *F. oxysporum* assemblies (Supplementary Table S3) to the k-mer profile of the co-linear accessory communities, based on our unsplit mapping approach. The strains were assembled from short read sequencing data downloaded from NCBI using spades version 3.13.0 [45]. To compare the k-mer profiles, we used sourmash gather (-bp-threshold 0), version 4.8.9 [46], and determined the presence of a chromosome community in the *F. oxysporum* strains based on the percentage of k-mer from the chromosome community that matched to the genome assemblies. The k-mer profile of a community is based on all chromosomes in the community, including similar chromosomes, and thus the short read k-mers will always match only a small percentage of the community. We considered that the maximal percentage of k-mers from a genome assembly to match a community to represent the optimal mapping. This maximum mapping per-

centage is used to normalize the matching percentage across all genome assemblies, and a community is considered present in a genome assembly when at least 20% of the maximum percentage is successfully mapped.

## Analysis of accessory communities

To visualize similarity between accessory chromosomes, we used PyGenomeViz version 1.0 (<https://moshi4.github.io/pyGenomeViz/>), in addition to the pangenome based visualizations from odgi viz [47]. To visualize similarities to the Fom0021 reference genome, we used nucmer -maxmatch to align all assemblies to the Fom021 assembly. We filtered the alignments to only retrieve alignments with length >5 kb and a percent identity >90% using custom Python scripts ([https://github.com/LikeFokkens/FOSC\\_multi-speed-genome](https://github.com/LikeFokkens/FOSC_multi-speed-genome)).

## Results

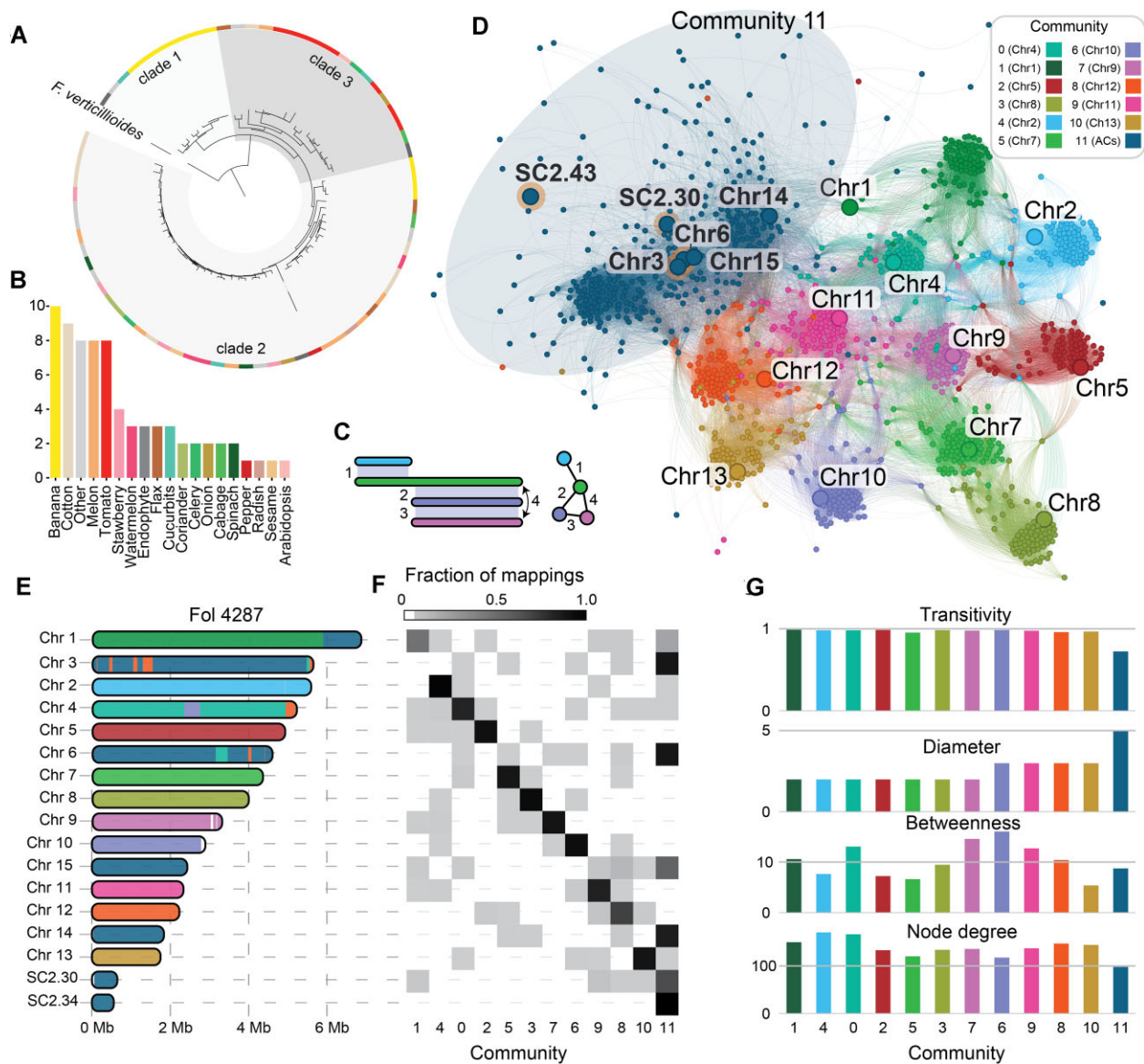
### The *F. oxysporum* species contain eleven core chromosomes and various accessory chromosomes

To study the diversity of core and accessory chromosomes in a large collection of diverse *F. oxysporum* strains, we obtained 73 highly continuous whole-genome assemblies (contig N50 >2 Mb, <100 contigs) (Supplementary Table S1). Together, these strains contain 1261 contigs. Most (1131 out of 1261) contigs encode telomeric repeats on at least one end, and 446 represent complete chromosomes containing telomeric repeats on both ends; we therefore refer to all contigs as chromosomes in this manuscript. The dataset contains *F. oxysporum* strains assigned to 19 different formae speciales based on their capacity to infect specific host plants, three endophytes, and eight strains that cause plant disease but have not been assigned to a forma specialis (Fig. 1A and B). The strains have an average nucleotide similarity of 97.7%, and span all three known phylogenetic clades within the *F. oxysporum* species complex (Fig. 1A; [65, 66]).

To study the variation of chromosome content in these different *F. oxysporum* strains and to facilitate pangenome analysis, we first sought to detect groups of co-linear chromosomes based on all-vs-all homology mapping between all 1261 chromosomes (Fig. 1C). We tried various thresholds for mapping, i.e. the percent identity (90%, 95%, or 97%) and the length of the matches (mapped segments length 2.5, 5, or 10 kb and the chained segments length of 5, 10, 25, 50, or 100 kb), to determine the optimal mapping strategy; the mappings need to be specific enough to map homologous chromosomes and to prevent spurious matches, yet sensitive enough to capture similarity between more distantly related chromosomes. These mappings between chromosomes were then translated into a chromosome network, and we used network properties to detect communities within the network that represent groups of homologous chromosomes (Fig. 1D). Since core chromosomes are expected to be conserved and co-linear, we expected to find communities of core chromosomes present in all strains, as well as several accessory chromosomes with a variable presence and absence pattern.

To analyze the distribution of the chromosomes over the communities, we determined the presence of chromosomes of the *F. oxysporum* f. sp. *lycopersici* reference genome Fol4287 in the communities (Fig. 1E and F). The core chromosomes of Fol4287 are grouped into 11 separate communities, whereas





**Figure 1.** Eleven core chromosomes as well as a high number of accessory chromosomes are present in *F. oxysporum*. **(A)** Phylogenetic tree of 73 *F. oxysporum* strains, based on 4443 single-copy BUSCO genes. Different boxes highlight strains belonging to the three taxonomic lineages in the *F. oxysporum* species complex. The outer ring shows the different hosts that can be infected by a strain, colored according to (B). **(B)** Distribution of plant host that can be infected by the 73 *F. oxysporum* strains. Endophytic strains as well as plant pathogens not assigned to a forma specialis (other) are separated. **(C)** The cartoon visualizes how sequence mappings between chromosomes are translated to a chromosome network. **(D)** Chromosome network displays all pairwise mappings between the chromosomes of the 73 *F. oxysporum* strains. Nodes represent chromosomes, and edges represent mappings between chromosomes. Colors indicate the different communities that are detected. Chromosome names correspond to the chromosomes found in the *Fol4287* reference genome, and these chromosomes are depicted by a large node in the graph. Accessory chromosomes of *Fol4287* are indicated by a bold font and an orange circle. All accessory chromosomes found in reference genome *Fol4287* occur in cluster 11. **(E)** Mapping of communities to homologous chromosome in reference *Fol4287*, colored according to (D). **(F)** Communities can be associated to specific chromosomes in *Fol4287*. The squares display the fraction of mappings of the *Fol4287* chromosome (y-axis) with the corresponding community (x-axis). Core chromosomes map largely to a single community, while all accessory chromosomes are assigned to community 11. **(G)** Network statistics of the different communities in the chromosome network. Colors correspond to the communities in the network, see d.

the five accessory chromosomes (chromosomes 3, 6, 14, supercontig 2.30, and supercontig 2.34) all group together into a single large community, which contained 342 chromosomes in total. Interestingly, two out of 73 strains do not contain any accessory chromosome (36 102; infecting banana and Fo5; endophytic, [Supplementary Table S2](#)). The core chromosomes on the other hand are all present in all 73 strains. Of the 11 core-chromosome communities, six communities contain a chromosome of all 73 strains and five communities miss one

or two strains ([Supplementary Table S2](#)). These missing chromosomes are not absent in these strain, but are the result of chromosomal rearrangements, either due to a mis-assembly or due to interchromosomal translocation. These fused chromosomes have similarity to multiple communities but are only assigned to the community with most matches. This also means that none of the strains lacks one of the core chromosomes ([Supplementary Table S2](#)). Irrespective of the mapping threshold that determines the minimal length of mappings between

chromosomes, we identified eleven core chromosomes, with one copy in all strains, as well as one community that contained all accessory chromosomes (Fig. 1D).

### Pangenome graph captures variation of *F. oxysporum* and highlights differences between core and accessory chromosomes

All accessory chromosomes cluster in a single community (community 11; Fig. 1D), suggesting that many accessory chromosomes in *F. oxysporum* share genetic material. However, community 11 has a lower average node degree (89) and a larger diameter [5] compared with the core communities (average node degree 136; average diameter 2.2). Moreover, the nodes in this cluster are generally less well connected (transitivity 0.74 versus average transitivity of 0.97, Fig. 1G), collectively suggesting that some of the chromosomes in the community have limited similarity to each other but are connected through similarity with other chromosomes.

To further analyze the chromosome structure and the similarities between chromosomes, we constructed a pangenome variation graph using the PanGenome Graph Builder (PGGB). This graph consists of 11 subgraphs, one graph per community, and contains 500 Mb of genetic material split over 33 651 077 nodes connected through 48 127 154 edges. 2 559 224 nodes (18.6 Mb) are core (present in all 73 genomes), 13 404 553 nodes (41.1 Mb) are accessory, and 17 687 283 nodes (440.6 Mb) are unique to a single genome (Fig. 2A). The pangenome graph constructed for these 73 *F. oxysporum* genomes is open (Heaps Alpha = 0.327, Fig. 2B), meaning that every addition of a newly sequenced genome is expected to add new genetic material to the graph. Core, accessory, and unique material is found in different quantities in the different communities. The eleven core-chromosome communities contain genetic material present in all genomes. However, in five of these core-chromosome communities, more accessory material is present (Fig. 2C), including communities containing the core chromosomes 11, 12, and 13 of *Fol4287*. Interestingly, these chromosomes have been described as more variable than the other core chromosomes and have been previously referred to as “fast-core” chromosomes [49]. However, a similar pattern is also observed in chromosomes 7 and 10, which are not considered part of the fast-core chromosomes, suggesting that these chromosomes similarly contain high amount of genomic variation. The amount of unique material is consistently low in the core chromosomes, and higher in the accessory community. As may be expected, the accessory community consists solely of accessory material (Fig. 2C), and not a single node is present in all accessory chromosomes.

To further investigate the expected co-linearity and variation between chromosomes in the pangenome graph, we analyzed the orientation and location of the nodes. We observed a higher number of inversions and nodes occur in varying order in the pangenome graph of the accessory community (Fig. 2E and Supplementary Fig. S1). This shows that, while core chromosomes are largely co-linear and conserved, the sequence order between homologous regions in accessory chromosomes is not conserved.

To determine the amount of similarity between different strains, we determined the node-similarity in the pangenome graph. We observed that based on the core communities, most genomes are highly similar (average sharedness 80%).

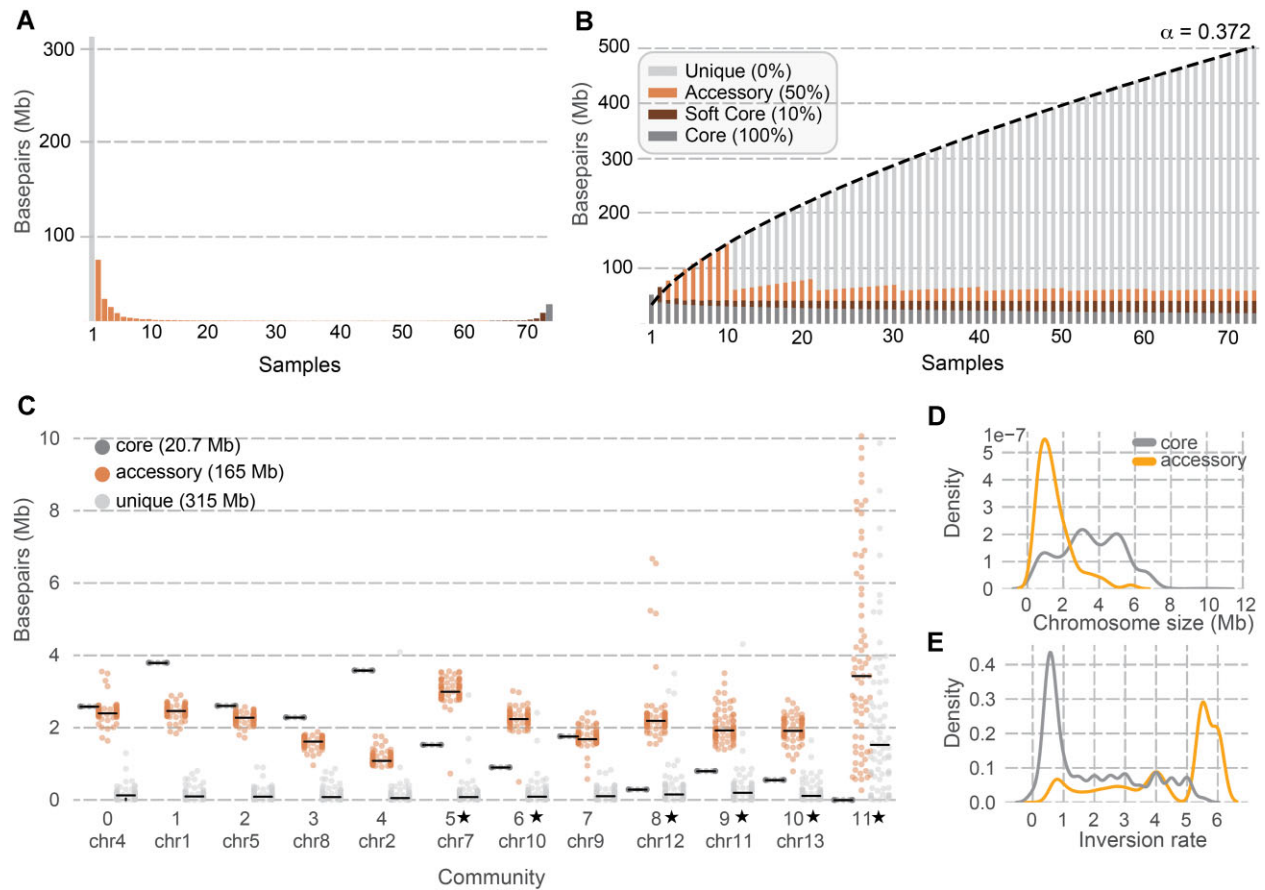
In the accessory community, the average sharedness between genomes is lower (11%), in line with the observed variation and lack of core nodes. The similarity of the accessory regions between strains infecting the same host is higher for some formae speciales, such as the tomato infecting strains (median similarity of 46%, across 64 strains) (Supplementary Fig. S2), but not for all host specific strains. For example, we only observed 14% similarity across 71 banana infecting strains (Supplementary Fig. S1). The high similarity in tomato infecting strains supports earlier reports showing that *F. oxysporum* strains infecting tomato contain a host-specific accessory chromosome that is associated with pathogenicity [4]. The lack of similarity between accessory chromosomes in strains infecting other hosts suggests these might not have host-specific pathogenicity chromosomes, as we have recently proposed for banana-infecting strains [6, 29].

### Accessory chromosomes can be grouped into separate co-linear communities

To further delineate subgroups of accessory chromosomes that share more extensive co-linearity, we performed an additional analysis where we only considered mappings between chromosomes that cover most of the corresponding chromosome, i.e. we exclude partial mappings between chromosomes. As a result, chromosomes are now grouped into 36 communities, composed of the same eleven core communities as well as 25 additional accessory communities, demonstrating that accessory chromosomes can be further separated into groups based on co-linearity between chromosomes (Fig. 3A). The retrieval of the same 11 core chromosome communities, even when excluding splitting, further supports that core chromosomes are colinear. Nevertheless, we observed 20 cases where a chromosome maps to two different communities. To construct the final pangenome graph, these chromosomes were manually split and added into the corresponding homologous community. Not all chromosomes present in the 73 *F. oxysporum* strains have a homologous chromosome in another isolate, we find 107 unique chromosomes, with an average size of 1.1 Mb, without any mappings to a homologous chromosome, suggesting that these are strains specific.

The separation of accessory chromosomes into smaller co-linear communities enabled us to limit our analysis to highly similar accessory chromosomes. Interestingly, we observed that some of the co-linear accessory chromosomes are unique to, or at least enriched for one of the formae speciales (Fig. 3C). For example, chromosome 14 of *Fol4287* is well known to play a role in pathogenicity towards tomato [4] and has been found to be present in genetically diverse tomato infecting *F. oxysporum* strains [49]. Similarly, we observed that community 13, containing *Fol4287* chromosome 14, solely consists of tomato infecting *F. oxysporum* strains. The other five accessory chromosomes of *Fol4287* (chromosomes 3, 6, 15, SC2.30, and SC2.34) are found in accessory community 12. The grouping of these five accessory chromosomes supports previous reports that suggested that the accessory chromosomes in *F. oxysporum* strain *Fol4287* are highly similar, due to extensive segmental duplications [4, 6].

In total, we observed 15 clusters that are specific to a formae speciales (Fig. 3C), suggesting that these might also represent pathogenicity chromosomes. However, most for-



**Figure 2.** Pangenome graph of *F. oxysporum* is open and contains extensive accessory material. (A) Histogram shows the amount of genetic material (in base pairs) that is present in an increasing number of samples (x-axis). Gray bar indicates the amount of material present in only one strain (unique material), the orange bars indicate the amount of material present in one to 65 strains, dark orange indicate the amount of material found in 65–72 strains, and dark gray shows the amount of material present in all strains (core). (B) Pangenome growth graph of the complete *F. oxysporum* pangenome, as determined by Panacus [48]. Heaps Alpha < 1 indicates an open pangenome. (C) Proportion (in base pairs) of core, accessory, and unique genetic material found in the different communities. Each dot represents a 500 bp window in the pangenome graph. As expected, community 11, containing all accessory chromosomes, lacks core material. Stars indicate the communities with more accessory than core material. (D) Size of chromosomes in accessory community 11 (orange) compared to the size of chromosomes in the other core communities (gray). (E) Inversion rate per 500 bp window in the pangenome graph in the core communities (gray) and accessory community 11 (orange). The inversion rate is determined by the number of inversions per genome and per node in the pangenome graphs.

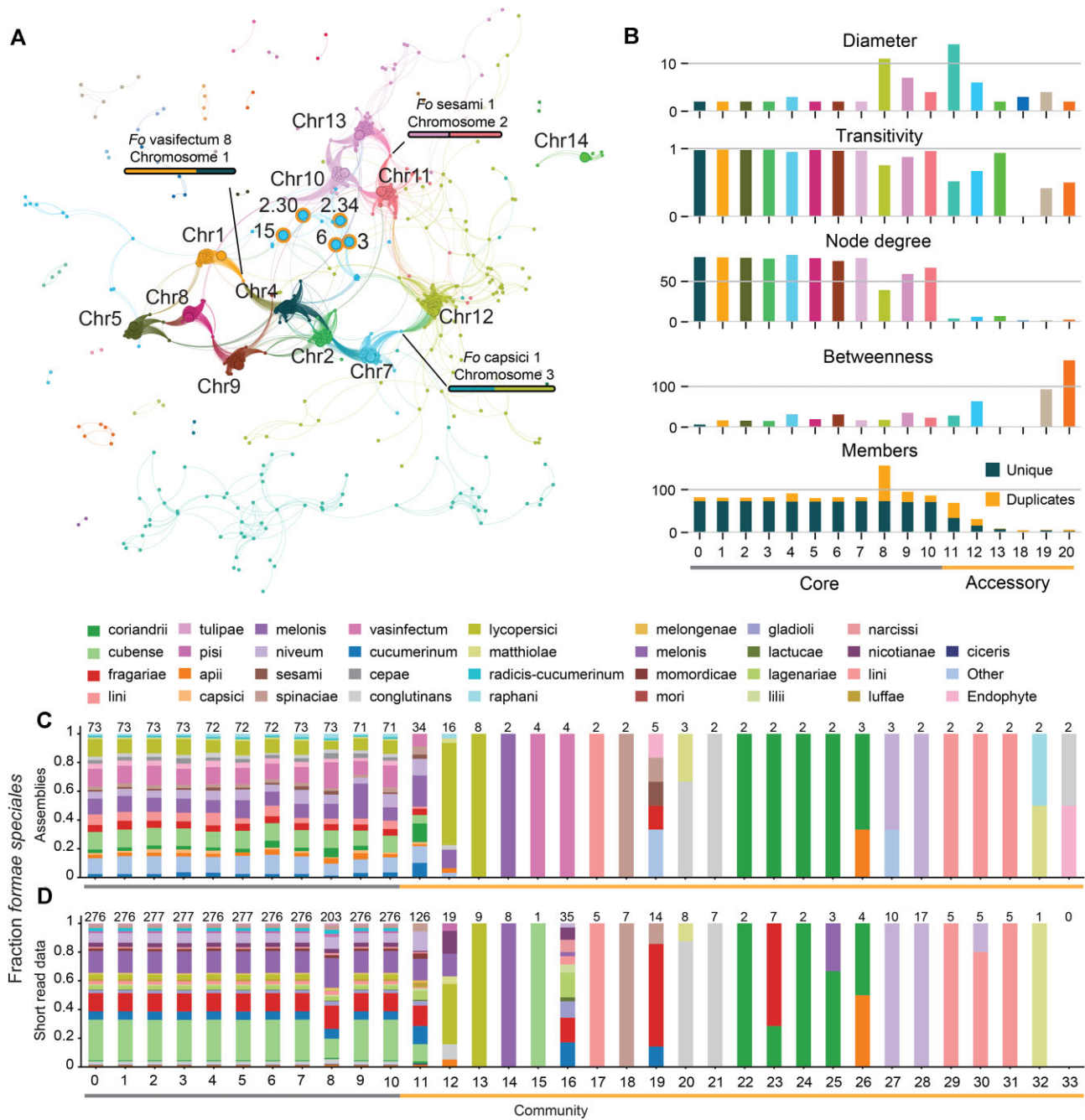
mae speciales were only represented by a limited number of genome assemblies (Fig. 1B and C). To include a broader set of strains, we sought to detect presence of accessory communities in a set of genome assemblies based on short-read sequencing data. We used a collection of 588 genome sequences that had been assembled from public data available at the Sequence Read Archive (Supplementary Table S3). To rapidly determine the chromosomes, present in one of these genome assemblies, we compared the k-mer profiles of these genome assemblies against the k-mer profiles of the chromosome communities. As to be expected, core communities are present in all strains and the accessory communities show a variable presence/absence pattern (Fig. 3D). Using the formae speciales known for 276 strains, we could distinguish 14 host specific chromosomes (present in more than four strains; Fig. 3D), including the tomato-specific community 13 that contains *Fol4287* chromosome 14. These analyses show that some accessory communities are indeed host specific in a wide variety of *F. oxysporum* strains, yet many other accessory communities are not restricted to a single host.

### Accessory chromosomes are a mosaic of accessory regions

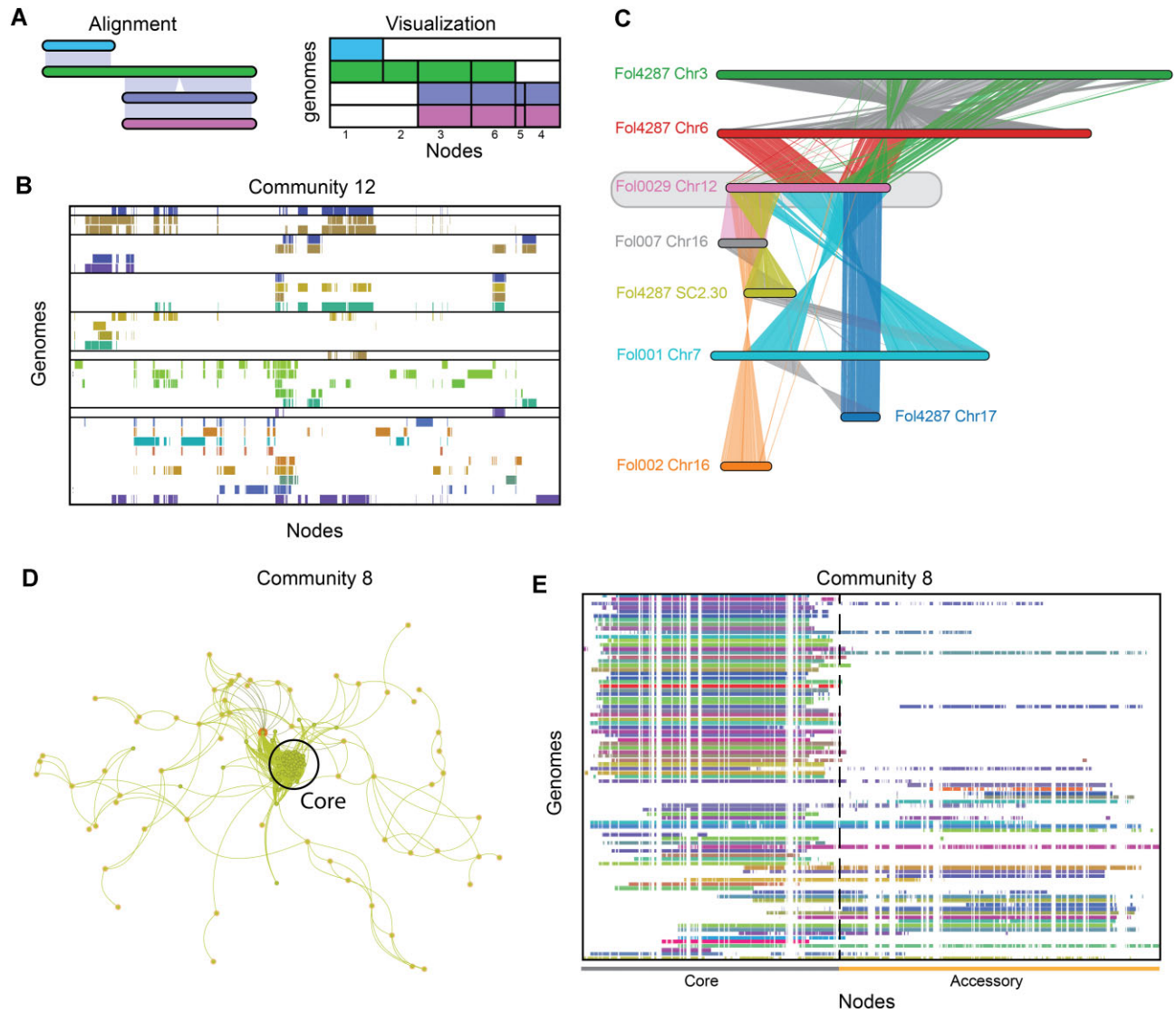
To get a more detailed insight into the variation within homologous chromosomes, we constructed a pangenome variation graph per community. This resulted in a combined pangenome, consisting of 443 Mb of genetic material that is separated over 31 289 751 nodes and connected by 44 436 449 edges. This pangenome graph consists of 18.8 Mb core, 14 Mb soft-core (>90% of strains), 161 Mb accessory, and 250 Mb of unique material, and is open (Heaps Alpha = 0.41 and Supplementary Fig. S3). As to be expected, the fraction of core material is considerably larger than in the previous pangenome graph (Fig. 2), as accessory communities now consist of more similar material (Supplementary Fig. S3).

We observed that the accessory chromosomes contain a high number of inversions (Fig. 2C) and are generally not collinear (Fig. 3A). We distinguished three different types of accessory communities. Most accessory communities [22] are small with only five chromosomes. The chromosomes in these clusters are overall well connected (betweenness: 30.5, diameter: 1.15) and share a large amount of genetic material





**Figure 3.** Accessory chromosomes can be separated into several co-linear communities, revealing host-specific accessory chromosomes. **(A)** The chromosome network displays all pairwise mappings spanning the full length of chromosomes. Nodes represent chromosomes and edges represent mappings between chromosomes. Colors indicate the different communities that are detected in the network. Chromosome names correspond to the chromosomes found in the *Fo4287* reference genome assembly, and these chromosomes are depicted by larger nodes in the graph. Orange circles highlight accessory chromosomes. **(B)** Network statistics of the different communities in the chromosome network. Colors correspond to the communities in the network, see panel (A). **(C)** Distribution of the different strains and their host specificity over the different communities. All strains are found in the core communities (0-10, x-axis). The accessory communities (11-33, x-axis), on the other hand, contain a subset of strains, and 15 accessory communities are host specific; the number on top indicate the total number of genomes in the community. **(D)** Distribution of different *F. oxysporum* strains and their host specificity over the different communities.

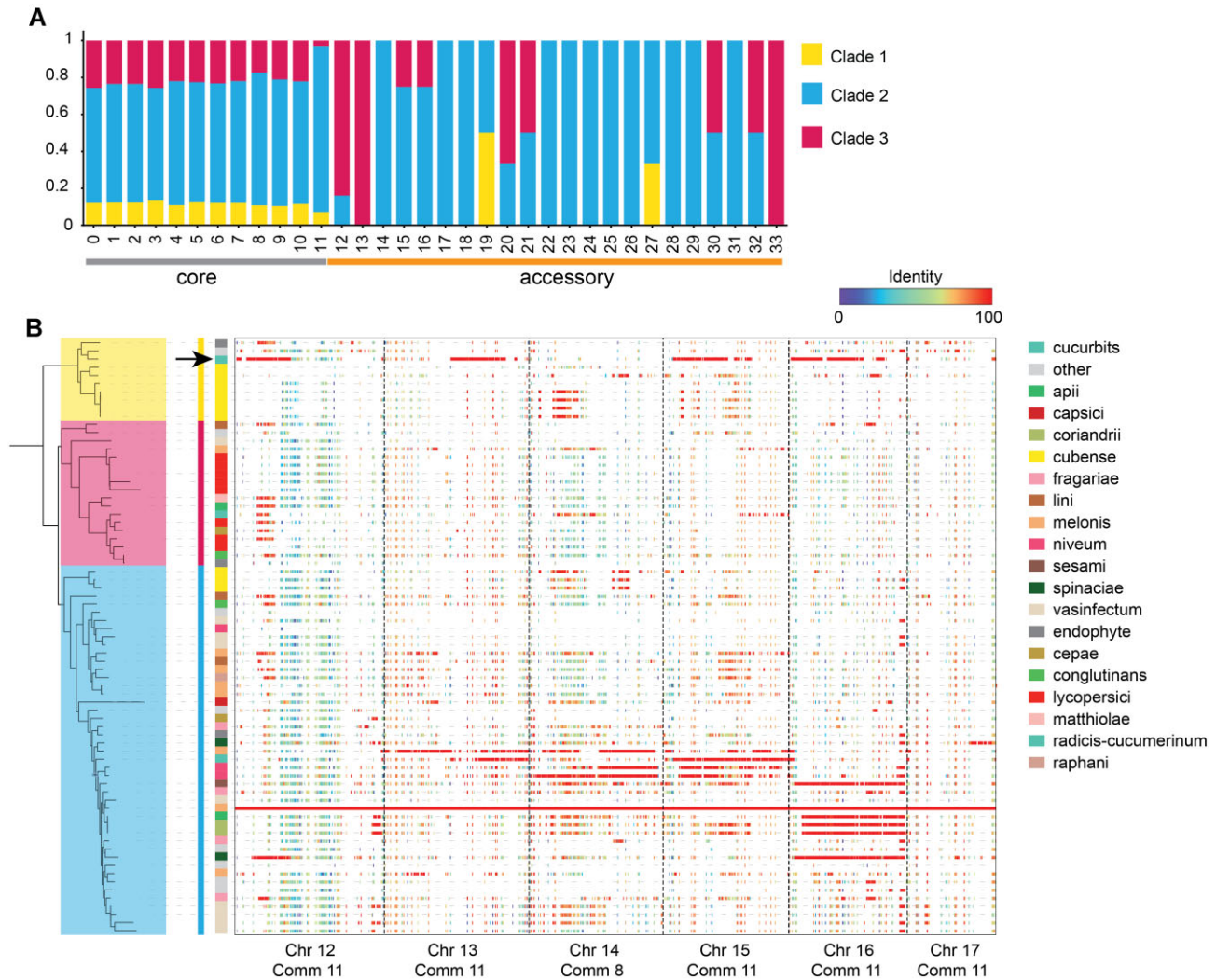


**Figure 4.** Accessory chromosomes are a mosaic of accessory regions from different genomes. **(A)** Translation from a pairwise chromosome alignment to a pangenome graph presence/absence matrix. The visualization places the nodes in the pangenome graph on the x-axis and the genomes on the y-axis. Colored blocks indicate a node that is present in a genome, while white indicates a node that is absent in the respective genome. **(B)** Presence/absence of nodes in the pangenome graph based on the 69 members of community 12. Note, none of the nodes are present in all chromosomes. **(C)** Alignment of eight chromosomes found in community 12. All chromosomes are aligned to *Fol0029* chromosome 12 (pink) and share different but similar genomic regions with the other chromosomes. The connecting lines indicated regions of shared similarity, colored according to query chromosome. Gray lines indicate similarity between two neighboring chromosomes, chromosome 3 and chromosome 6 of *Fol4287* are homologous. **(D)** Community 8 shows a collection of core chromosomes surrounded by various accessory chromosomes (highlighted in orange). **(E)** Node presence/absence matrix of a subset of chromosomes in community 8. The node presence and absence clearly demonstrate that the core part of the chromosomes is similar across all 73 genomes (gray line). Importantly, some core chromosomes contain an accessory region, while in other cases the accessory regions are separate from the core chromosome.

(49 Mb core, 21 Mb accessory, and 27 Mb unique). In addition, we find two large accessory communities (community 11 and community 12 with 69 and 31 members, respectively). These accessory communities have a larger diameter (10.0) and a higher betweenness (45.9) than the small accessory communities, indicating that the nodes in the communities are loosely connected; not all chromosomes share genetic material, analogous to accessory community 11 in our previous analysis (Fig. 1). Moreover, the pangenome graph of communities 11 and 12 does not contain any core nodes (Supplementary Fig. S4), further supporting the lack of shared genetic material between the chromosomes in these communities.

To determine what links the chromosomes in these large and loosely connected accessory communities, we visualized the presence and absence of nodes in the pangenome graph (Fig. 4A). This shows that not all chromosomes share regions with each other but stretches of similar regions are found in different combinations (Fig. 4B). For example, different parts of an accessory chromosome (*Fol029*, chromosome 12) are present in other accessory chromosomes (Fig. 4C), indicating that accessory chromosomes share different accessory region. Together, these findings supports the idea that accessory chromosomes evolve through chromosomal rearrangements, resulting in a mosaic of genetic material from different sources [49].





**Figure 5.** Similar accessory chromosomes can be found in different *F. oxysporum* clades. **(A)** Distribution of *F. oxysporum* clades over the different chromosome communities (x-axis). Ten accessory communities contain strains from different taxonomic clades, indicating that similar chromosomes can be shared by *F. oxysporum* strains from with large phylogenetic distances, even from different taxonomic clades. **(B)** Similarity of six accessory chromosomes found in the melon infecting *F. oxysporum* strain Fom0021 (x-axis) across all 73 *F. oxysporum* strains (y-axis), ordered based on the phylogenetic relationship of the *F. oxysporum* strains (left). Only alignments that span at least 5 kb are shown. The color indicates the percent identity. Accessory regions from Fom0021 are found in different strains in clade 2 as well as in one strain (Foc013) in clade 1, highlighted by an arrow.

Next to these two types of accessory communities, we also identified accessory chromosomes present in two core chromosome communities (community 8 and community 9 with 63 accessory and 14 accessory chromosomes, respectively; Fig. 4D). This pattern can indicate that accessory chromosomes originated from core chromosomes or that accessory chromosomes fused to core chromosomes. We observed that similar core chromosomes are present in community 8, as well as additional separate accessory chromosomes that do not share homology to the core regions (Fig. 4D and E). However, this accessory region is attached to the core chromosome in some strains, whilst in others the accessory region is separate (Fig. 4E). This demonstrates that the link between core and accessory chromosomes in this community is caused by the fusion of an accessory region to a core chromosome, an arrangement that has been previously observed in various *F. oxysporum* genome assemblies [4, 6, 29].

### Accessory chromosomes can be transferred between genetically different *F. oxysporum* strains

Accessory chromosomes can be transferred between *F. oxysporum* strains as has been experimentally shown for a few chromosomes in some strains [4, 35, 36]. Horizontal transfer can also occur in natural populations and has been identified based on the presence of similar accessory chromosomes between distantly related strains [18, 33, 49]. To determine horizontal transfer events in our dataset, we analyzed the similarity of accessory chromosomes between strains from different taxonomic clades. We observed that 10 communities contain members from different taxonomic clades of the *F. oxysporum* species complex (Fig. 5A), suggesting that horizontal transfer might have taken place. For instance, community 11 contains strains from the taxonomic clades 1, 2, and 3. One strain from clade 1 (Foc013) shares 58% similarity with clade 2 genome Fom0021, but only 12% similarity with closest clade 1 neighbor (Fo4, Supplementary Fig. S5), indicating that hori-

zontal transfer has occurred between strains from clade 1 and clade 2.

To get a more detailed insight into the traces of horizontal chromosome transfer between *Fom0021* and *Foc013*, we performed whole-genome alignments between the accessory chromosomes found in *Fom0021* to all 73 genomes. Interestingly, the accessory chromosomes are present in some but not all of the most closely related species, and similar genomic regions could even be found in a strain from different taxonomic clades. For example, chromosome 12, 13, 14, and 16 from *Fom0021* (formae speciales *melonis*) are also found in *Foc013* (formae speciales *cucumerinum*) from clade 1 and a region of accessory chromosome 14 of *Fom0021* can be found in four banana infecting genomes (formae speciales *cubense*) in clade 1. This further corroborates that accessory chromosomes are shared across genomes and further highlights that horizontal transfer occurs between clades of *F. oxysporum*. Interestingly, the shared genomic regions are present in strains infecting different hosts, suggesting that the transfer is not always associated with the same host range.

## Discussion

Pangenome variation graphs offer a reference-free genome representation of different individuals of the same species. This can provide insights into the genome organization and can improve downstream genome-based analysis such as variant calling and gene annotation [24, 26, 28, 50]. Such reference free analysis would be especially useful for analyzing accessory chromosomes in filamentous plant pathogens, as these are not necessarily present in the reference genome assembly. Moreover, accessory chromosomes are often highly diverse, hampering reference-based analyses [17, 51]. By combining a homologous chromosome mapping strategy with the construction of a pangenome variation graph, we captured all accessory chromosomes in a collection of 73 *F. oxysporum* strains and obtained insights into their evolutionary dynamics. We found 11 conserved core chromosomes and a large number of different accessory chromosomes, which evolve through rearrangements that result in a mosaic of different accessory genomic regions. Moreover, we found evidence for horizontal transfer of accessory chromosomes between naturally occurring *F. oxysporum* strains from different phylogenetic clades. These results highlight that pangenome variation graphs can be successfully reconstructed for species carrying accessory chromosomes and can guide the exploration of genetic variation that underlies specific phenotypes and provides insights into their genome organization to obtain insights into genome evolution.

*F. oxysporum* is a highly diverse species complex that infects a wide variety of hosts [31, 32, 37]. We identified an open pangenome, meaning that every new strain adds yet unseen additional genomic regions to the pangenome, which is in line with previous gene-based pangenomes in *F. oxysporum* [6, 37] and is expected for a highly diverse species [20]. The rapid diversification of accessory chromosomes can underlie the observed diversity [12, 52]. Using the pangenome variation graph, we identified a large set of diverse accessory chromosomes with a mosaic composition, encoding various combinations of accessory regions, similar to what has been observed in the accessory genome of other fungi [17, 49, 53]. As a result of these similar accessory regions, all accessory chromosomes can be clustered together in one community, how-

ever, not a single accessory region present in all chromosomes can be identified. Thus, the similarity of accessory chromosomes does not seem to result from a single shared origin. Moreover, similar accessory chromosomes are found between strains from different clades, highlighting that accessory chromosomes can be horizontally transferred between distantly related strains. This supports previous studies suggesting that clades are not genetically isolated in *F. oxysporum* [54] and that genetic material can be exchanged not only within but also between taxonomic clades. We speculate that the horizontal transfer of accessory chromosomes introduces genetic material to strains, through recombination this genetic material can be incorporated in other accessory chromosomes and new combinations of accessory material arise. This means that the horizontal transfer of accessory chromosomes, together with rearrangement of accessory material, can provide genetic diversification that is especially important in asexual fungal species [55, 56].

In *F. oxysporum*, many accessory chromosomes carry important pathogenicity genes and play a role in determining the host range [4, 34, 36]. Similar to previous studies, we found highly similar and host-specific accessory chromosomes within several formae speciales. However, host-specific accessory chromosomes are not identified for all formae speciales. It remains unclear why some formae speciales have a clear pathogenicity related accessory chromosomes, like tomato [49], whereas others have variable accessory chromosomes [6, 18]. This variability suggests that some host-specificities arose multiple times independently, or the genes required for this host specificity are reshuffled between accessory chromosomes or between accessory chromosomes and the core. Additionally, limited phenotyping data makes it difficult to study host specificity. In general, formae speciales are assigned based on the isolation source, but a strain may infect multiple hosts or cause different symptoms in the same host [18, 57]. The pangenome constructed in this study can serve as a valuable resource for exploring chromosome diversity within and between formae speciales of *F. oxysporum*. This can enable the identification of conserved regions and essential pathogenicity genes, providing insights into the host range of strains. Additionally, the pangenome can place newly sequenced and assembled *F. oxysporum* strains into a larger context, this can reveal the presence of accessory chromosomes, show what accessory regions are shared and thereby provide an important framework to clarify the role of accessory regions in pathogenicity.

Different forms of pangenomes can be constructed, based on genes, whole genomes, or k-mer content [58]. The optimal approach depends on the biology of a species, as well as the exact biological question to be addressed. Research on pangenomes in fungi thus far mainly focused on gene-based pangenomes, representing all genes within a species [59, 60]. This provides important insights into the evolution of gene content, especially of pathogenicity genes, but typically offers limited information on the genome organization and the presence and shape of accessory chromosomes. Insights into accessory chromosome dynamics have been previously obtained by pairwise or reference-based comparisons [6, 17, 61]. The here applied reference-free pangenome variation graph enables large-scale comparisons of accessory chromosomes and can be applied in different fungi to obtain insights into chromosome dynamics. Moreover, a pangenome variation graph is valuable data by improving variant calling [62] as well as

annotation of genes and transposable elements [26, 63, 64], thereby enhancing our capabilities to study a species' biology. Future analyses of pangenome variation graph can thus not only help to reveal insights into the evolution of the genome organization but also can help to reveal new insights into the population structure, pathogenicity genes, and other drivers of phenotypic variations.

## Acknowledgements

A.C.W. and G.H.J.K. were supported by a grant from the Bill and Melinda Gates Foundation to the International Institute of Tropical Agriculture, Agreement No - AG-5797. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Author contributions:** M.F.S. conceived and supervised the project. A.C.W. performed the analysis. L.F. and K.W. contributed to data analysis. G.H.J.K. and M.R. contributed to analyzed material. A.C.W. wrote the manuscript with input from M.F.S. All authors contributed to writing and editing of the manuscript and approved the final manuscript.

## Supplementary data

**Supplementary data** is available at NAR Genomics & Bioinformatics online.

## Conflict of interest

None declared.

## Funding

Bill and Melinda Gates Foundation [AG-5797].

## Data availability

The location of the genome sequencing data and assemblies used in this study can be found in Supplementary Table S1 and S3. Custom code generated for this study can be found in Github, [https://github.com/Anouk-vw/FOSC\\_pangenome](https://github.com/Anouk-vw/FOSC_pangenome) and [https://github.com/LikeFokkens/FOSC\\_multi-speed-genome](https://github.com/LikeFokkens/FOSC_multi-speed-genome), and on Zenodo <https://doi.org/10.5281/zenodo.15006311>.

## References

1. D'Ambrosio U, Alonso-Lifante MP, Barros K *et al.* B-chrom: a database on B-chromosomes of plants, animals and fungi. *Nat Commun* 2017;216:1029–38. <https://doi.org/10.1111/nph.14723>
2. Ma W, Gabriel TS, Martis MM *et al.* Rye B chromosomes encode a functional Argonaute-like protein with *in vitro* slicer activities similar to its A chromosome paralog. *New Phytol* 2017;213:916–28. <https://doi.org/10.1111/nph.14110>
3. Habig M, Quade J, Stukenbrock EH. Forward genetics approach reveals host genotype-dependent importance of accessory chromosomes in the fungal wheat pathogen *Zymoseptoria tritici*. *mBio* 2017;8:e01919–17. <https://doi.org/10.1128/mBio.01919-17>
4. Ma L-J, van der Does HC, Borkovich KA *et al.* Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 2010;464:367–73. <https://doi.org/10.1038/nature08850>
5. Torres DE, Oggenfuss U, Croll D *et al.* Genome evolution in fungal plant pathogens: looking beyond the two-speed genome model. *Fung Biol Rev* 2020;34:136–43. <https://doi.org/10.1016/j.fbr.2020.07.001>
6. van Westerhoven AC, Aguilera-Galvez C, Nakasato-Tagami G *et al.* Segmental duplications drive the evolution of accessory regions in a major crop pathogen. *New Phytol* 2024;242:610–25. <https://doi.org/10.1111/nph.19604>
7. Wei H, Zhong Z, Li Z *et al.* Loss of the accessory chromosome converts a pathogenic tree-root fungus into a mutualistic endophyte. *Plant Commun* 2024;5:100672. <https://doi.org/10.1016/j.xplc.2023.100672>
8. Fouché S, Plissonneau C, McDonald BA *et al.* Meiosis leads to pervasive copy-number variation and distorted inheritance of accessory chromosomes of the wheat pathogen *Zymoseptoria tritici*. *Genome Biol Evol* 2018;10:1416–29. <https://doi.org/10.1093/gbe/evy100>
9. Stalder L, Oggenfuss U, Mohd-Assaad N *et al.* The population genetics of adaptation through copy number variation in a fungal plant pathogen. *Mol Ecol* 2022;32:2443–60. <https://doi.org/10.1111/mec.16435>
10. Croll D, McDonald BA. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS Pathog* 2012;8:e1002608. <https://doi.org/10.1371/journal.ppat.1002608>
11. Coleman JJ, Rounsley SD, Rodriguez-Carres M *et al.* The genome of *Nectria haematococca*: contribution of supernumerary chromosomes to gene expansion. *PLoS Genet* 2009;5:e1000618. <https://doi.org/10.1371/journal.pgen.1000618>
12. Dong S, Raffaele S, Kamoun S. The two-speed genomes of filamentous pathogens: waltz with plants. *Curr Opin Genet Dev* 2015;35:57–65.
13. Goodwin SB, M'Barek SB, Dhillon B *et al.* Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet* 2011;7:e1002070. <https://doi.org/10.1371/journal.pgen.1002070>
14. Haas BJ, Kamoun S, Zody MC *et al.* Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. *Nature* 2009;461:393–8. <https://doi.org/10.1038/nature08358>
15. Rouxel T, Grandaubert J, Hane JK *et al.* Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by repeat-induced point mutations. *Nat Commun* 2011;2:202. <https://doi.org/10.1038/ncomms1189>
16. Wacker T, Helmstetter N, Wilson D *et al.* Two-speed genome evolution drives pathogenicity in fungal pathogens of animals. *Proc Natl Acad Sci USA* 2023;120:e2212633120. <https://doi.org/10.1073/pnas.2212633120>
17. Barragan AC, Latorre SM, Malmgren A *et al.* Multiple horizontal mini-chromosome transfers drive genome evolution of clonal blast fungus lineages. *Mol Biol Evol* 2024;41:msae164. <https://doi.org/10.1093/molbev/msae164>
18. Henry PM, Pincot DDA, Jenner BN *et al.* Horizontal chromosome transfer and independent evolution drive diversification in *Fusarium oxysporum* f. sp. *fragariae*. *New Phytol* 2021;230:327–40. <https://doi.org/10.1111/nph.17141>
19. Plissonneau C, Stürchler A, Croll D. The evolution of orphan regions in genomes of a fungal pathogen of wheat. *mBio* 2016;7:e01231–16. <https://doi.org/10.1128/mBio.01231-16>
20. Tettelin H, Massignani V, Cieslewicz MJ *et al.* Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome”. *Proc Natl Acad Sci USA* 2005;102:13950–5. <https://doi.org/10.1073/pnas.0506758102>
21. The Computational Pan-Genomics Consortium Computational pan-genomics: status, promises and challenges. *Brief Bioinform* 2018;19:118–35.
22. Eizenga JM, Novak AM, Sibbesen JA *et al.* Pangenome graphs. *Annu Rev Genomics Hum Genet* 2020;21:139–62. <https://doi.org/10.1146/annurev-genom-120219-080406>



23. Hickey G, Monlong J, Ebler J *et al.* Pangenome graph construction from genome alignments with Minigraph-Cactus. *Nat Biotechnol* 2023;42:663–73. <https://doi.org/10.1038/s41587-023-01793-w>
24. Guarracino A, Buonaio S, de Lima LG *et al.* Recombination between heterologous human acrocentric chromosomes. *Nature* 2023;617:335–43. <https://doi.org/10.1038/s41586-023-05976-y>
25. Hickey G, Heller D, Monlong J *et al.* Genotyping structural variants in pangenome graphs using the vg toolkit. *Genome Biol* 2020;21:35. <https://doi.org/10.1186/s13059-020-1941-7>
26. Skiadas P, Vidal SR, Dommissie J *et al.* Pangenome graph analysis reveals extensive effector copy-number variation in spinach downy mildew. *PLoS Genet* 2024;20:e1011452. <https://doi.org/10.1371/journal.pgen.1011452>
27. Gao L, Gonda I, Sun H *et al.* The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. *Nat Genet* 2019;51:1044–51. <https://doi.org/10.1038/s41588-019-0410-2>
28. Wang T, Antonacci-Fulton L, Howe K *et al.* The human pangenome project: a global resource to map genomic diversity. *Nature* 2022;604:437–46. <https://doi.org/10.1038/s41586-022-04601-8>
29. Dijkstra J, Westerhoven ACv, Gómez-Gil L *et al.* Extensive intrachromosomal duplications in a virulence-associated fungal accessory chromosome. *bioRxiv*, <https://doi.org/10.1101/2024.09.16.611982>, 16 September 2024, preprint: not peer reviewed.
30. Garrison E, Guarracino A, Heumos S *et al.* Building pangenome graphs. *Nat Methods* 2024;21:2008–12. <https://doi.org/10.1038/s41592-024-02430-3>
31. Armer VJ, Kroll E, Darino M *et al.* Navigating the *Fusarium* species complex: host-range plasticity and genome variations. *Fungal Biol* 2024;128:2439–59. <https://doi.org/10.1016/j.funbio.2024.07.004>
32. Edel-Hermann V, Lecomte C. Current status of *Fusarium oxysporum* *formae speciales* and races. *Phytopathology* 2019;109:512–30. <https://doi.org/10.1094/PHYTO-08-18-0320-RVW>
33. van Dam P, Fokkens L, Schmidt SM *et al.* Effector profiles distinguish *formae speciales* of *Fusarium oxysporum*. *Environ Microbiol* 2016;18:4087–102. <https://doi.org/10.1111/1462-2920.13445>
34. Armitage AD, Taylor A, Sobczyk MK *et al.* Characterisation of pathogen-specific regions and novel effector candidates in *Fusarium oxysporum* f. sp. *cepaie*. *Sci Rep* 2018;8:13530. <https://doi.org/10.1038/s41598-018-30335-7>
35. Li J, Fokkens L, Conneely LJ *et al.* Partial pathogenicity chromosomes in *Fusarium oxysporum* are sufficient to cause disease and can be horizontally transferred. *Environ Microbiol* 2020;22:4985–5004. <https://doi.org/10.1111/1462-2920.15095>
36. van Dam P, Fokkens L, Ayukawa Y *et al.* A mobile pathogenicity chromosome in *Fusarium oxysporum* for infection of multiple cucurbit species. *Sci Rep* 2017;7:9042. <https://doi.org/10.1038/s41598-017-07995-y>
37. Fayyaz A, Robinson G, Chang PL *et al.* Hiding in plain sight: genome-wide recombination and a dynamic accessory genome drive diversity in *Fusarium oxysporum* f.sp. *ciceris*. *Proc Natl Acad Sci USA* 2023;120:e2220570120. <https://doi.org/10.1073/pnas.2220570120>
38. Gurevich A, Saveliev V, Vyahhi N *et al.* QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 2013;29:1072–5. <https://doi.org/10.1093/bioinformatics/btt086>
39. Brown M, González De la Rosa PM, Mark B. A telomere identification toolkit. *Zenodo*, 2023. <https://doi.org/10.5281/zenodo.10091385>
40. Marco-Sola S, Moure JC, Moreto M *et al.* Fast gap-affine pairwise alignment using the wavefront algorithm. *Bioinformatics* 2021;37:456–63. <https://doi.org/10.1093/bioinformatics/btaa777>
41. Traag VA, Waltman L, van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities. *Sci Rep* 2019;9:5233. <https://doi.org/10.1038/s41598-019-41695-z>
42. Garrison E, Guarracino A, Heumos S *et al.* Building pangenome graphs. *Nat Methods* 2024. 21 2008–12. <https://doi.org/10.1038/s41592-024-02430-3>
43. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. *Proc Int AAAI Conf Web Social Media* 2009;3:361–2. <https://doi.org/10.1609/icwsm.v3i1.13937>
44. Csárdi G, Nepusz T. The igraph software package for complex network research. 2006. <https://www.semanticscholar.org/paper/The-igraph-software-package-for-complex-network-Cs%C3%A1rdi-Nepusz/1d2744b83519657f5f2610698a8ddd177ced4f5c>. (8 June 2024, date last accessed).
45. Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–77. <https://doi.org/10.1089/cmb.2012.0021>
46. Pierce NT, Irber L, Reiter T *et al.* Large-scale sequence comparisons with *sourmash*. *F1000Research* 2019;8:1006. <https://doi.org/10.12688/f1000research.19675.1>
47. Guarracino A, Heumos S, Nahnsen S *et al.* ODGI: understanding pangenome graphs. *Bioinformatics* 2022;38:3319–26. <https://doi.org/10.1093/bioinformatics/btac308>
48. Parmigiani L, Garrison E, Stoye J *et al.* Panacus: fast and exact pangenome growth and core size estimation. *Bioinformatics* 2024;40:btac720. <https://doi.org/10.1093/bioinformatics/btac720>
49. Fokkens L, Shahi S, Connolly LR *et al.* The multi-speed genome of *Fusarium oxysporum* reveals association of histone modifications with sequence divergence and footprints of past horizontal chromosome transfer events. *bioRxiv*, <https://doi.org/10.1101/465070>, 7 November 2018, preprint: not peer reviewed.
50. Li H, Wang S, Chai S *et al.* Graph-based pan-genome reveals structural and sequence variations related to agronomic traits and domestication in cucumber. *Nat Commun* 2022;13:682. <https://doi.org/10.1038/s41467-022-28362-0>
51. Sherman RM, Salzberg SL. Pan-genomics in the human genome era. *Nat Rev Genet* 2020;21:243–54. <https://doi.org/10.1038/s41576-020-0210-7>
52. Yang H, Yu H, Ma L-J. Accessory chromosomes in *Fusarium oxysporum*. *Phytopathology* 2020;110:1488–96. <https://doi.org/10.1094/PHYTO-03-20-0069-IA>
53. Langner T, Harant A, Gomez-Luciano LB *et al.* Genomic rearrangements generate hypervariable mini-chromosomes in host-specific isolates of the blast fungus. *PLoS Genet* 2021;17:e1009386. <https://doi.org/10.1371/journal.pgen.1009386>
54. van Westerhoven AC, Dijkstra J, Aznar Palop JL *et al.* Frequent genetic exchanges revealed by a pan-mitogenome graph of a fungal plant pathogen. *mBio* 2024;15:e0275824.
55. Fokkens L, Rep M. Population genomics reveals meiotic recombination in *Fusarium oxysporum*. *Proc Natl Acad Sci USA* 2023;120:e2309677120. <https://doi.org/10.1073/pnas.2309677120>
56. Seidl MF, Thomma BPHJ. Sex or no sex: evolutionary adaptation occurs regardless. *Bioessays* 2014;36:335–45. <https://doi.org/10.1002/bies.201300155>
57. Batson AM, Fokkens L, Rep M *et al.* Putative effector genes distinguish two pathogenicity groups of *Fusarium oxysporum* f. sp. *spinaciae*. *Mol Plant Microbe Interact* 2021;34:141–56. <https://doi.org/10.1094/MPMI-06-20-0145-R>
58. Andreade F, Lechat P, Dufresne Y *et al.* Comparing methods for constructing and representing human pangenome graphs. *Genome Biol* 2023;24:274. <https://doi.org/10.1186/s13059-023-03098-2>
59. Badet T, Oggenfuss U, Abraham L *et al.* A 19-isolate reference-quality global pangenome for the fungal wheat pathogen *Zymoseptoria tritici*. *BMC Biol* 2020;18:12. <https://doi.org/10.1186/s12915-020-0744-3>
60. Barber AE, Sae-Ong T, Kang K *et al.* *Aspergillus fumigatus* pan-genome analysis identifies genetic variants associated with

- human infection. *Nat Microbiol* 2021;**6**:1526–36.  
<https://doi.org/10.1038/s41564-021-00993-x>
61. Feurtéy A, Lorrain C, McDonald MC *et al.* A thousand-genome panel retraces the global spread and adaptation of a major fungal crop pathogen. *Nat Commun* 2023;**14**:1059.  
<https://doi.org/10.1038/s41467-023-36674-y>
  62. Sirén J, Monlong J, Chang X *et al.* Pangenomics enables genotyping of known structural variants in 5202 diverse genomes. *Science* 2021;**374**:abg8871.  
<https://doi.org/10.1126/science.abg8871>
  63. Fiddes IT, Armstrong J, Diekhans M *et al.* Comparative annotation toolkit (CAT)—simultaneous clade and personal genome annotation. *Genome Res* 2018;**28**:1029–38.  
<https://doi.org/10.1101/gr.233460.117>
  64. Groza C, Chen X, Wheeler TJ *et al.* A unified framework to analyze transposable element insertion polymorphisms using graph genomes. *Nat Commun* 2024;**15**:8915.  
<https://doi.org/10.1038/s41467-024-53294-2>
  65. Maryani N, Lombard L, Poerba YS *et al.* Phylogeny and genetic diversity of the banana fusarium wilt pathogen *Fusarium oxysporum* f. sp. *cubense* in the Indonesian centre of origin. *Stud Mycol* 2019;**92**:155–94.  
<https://doi.org/10.1016/j.simyco.2018.06.003>
  66. O'Donnell K, Kistler HC, Cigelnik E *et al.* Multiple evolutionary origins of the fungus causing Panama disease of banana: concordant evidence from nuclear and mitochondrial gene genealogies. *Proc Natl Acad Sci USA* 1998;**95**:2044–9.  
<https://doi.org/10.1073/pnas.95.5.2044>